8-2020

Informationally Complete Measurements and Optimal Representations of Quantum Theory

John B. DeBrota

INFORMATIONALLY COMPLETE MEASUREMENTS AND OPTIMAL
REPRESENTATIONS OF QUANTUM THEORY

A Dissertation Presented
by
JOHN B. DEBROTA

Submitted to the Office of Graduate Studies,
University of Massachusetts Boston,
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2020

Applied Physics Program

INFORMATIONALLY COMPLETE MEASUREMENTS AND OPTIMAL REPRESENTATIONS OF QUANTUM THEORY

A Dissertation Presented

by

JOHN B. DEBROTA

Approved as to style and content by:

Christopher A. Fuchs, Professor
Chairperson of Committee

Alioscia Hamma, Associate Professor
Member

Maxim Olchanyi, Professor
Member

Ingemar Bengtsson, Professor
Stockholm University
Member

Jonathan Celli, Program Director
Applied Physics Program

Rahul Kulkarni, Chairperson
Physics Department

ABSTRACT

INFORMATIONALLY COMPLETE MEASUREMENTS AND OPTIMAL REPRESENTATIONS OF QUANTUM THEORY

August 2020

John B. DeBrota, B.S., Indiana University Bloomington
M.Sc., University of Waterloo
Ph.D, University of Massachusetts Boston

Directed by Professor Christopher A. Fuchs

Minimal informationally complete quantum measurements (MICs) furnish probabilistic representations of quantum theory. These representations cleanly present the Born rule as an additional constraint in probabilistic decision theory, a perspective advanced by QBism. Because of this, their structure illuminates important ways in which quantum theory differs from classical physics. MICs have, however, so far received relatively little attention. In this dissertation, we investigate some of their general properties and relations to other topics in quantum information. A special type of MIC called a symmetric informationally complete measurement makes repeated appearances as the optimal or extremal solution in distinct settings, signifying they play a significant foundational role. Once the general structure of MICs is more fully explicated, we speculate that the representation will have unique advantages analogous to the phase space and path integral formulations. On the conceptual

side, the reasons for QBism continue to grow. Most recently, extensions to the Wigner's friend paradox have threatened the consistency of many interpretations. QBism's resolution is uniquely simple and powerful, further strengthening the evidence for this interpretation.

ACKNOWLEDGMENTS

I owe my utmost thanks for the success of this work to my advisor and friend Christopher Fuchs. I have never met anyone with Chris's brand of intuition for the physical and philosophical; it has and continues to fascinate and inspire me. Since day one, five years ago, Chris has tirelessly sought opportunities for me and worked on my behalf. I would not be where I am nor who I am had anyone else taken his place.

Very special thanks go to Blake Stacey, who has also been with me for the duration. Blake is a remarkable person, one of my closest friends, and most regular collaborator. His depth and breadth of knowledge never cease to amaze me.

Very special thanks also go to my family, especially my parents, David and Erin DeBrota. I have had the rare gift of wonderful, supportive parents. At every turn, they made sure that I could follow my passions.

I thank the QBists and QBissels, Marcus Appleby, Gabriela Baretto Lemos, Jacques Pienaar, and Rüdiger Schack for all the discussions, drinks, jokes, hikes, and company. They have each significantly shaped my view of the world.

In the course of graduate school, the following groups and individuals greatly impacted my life. In some way or another, they all partially enabled this dissertation. I thank: The Glütenhäus, Jeffrey La, Kathryn Hausler, and Talena Gandy. My climbing friends, Adrian Collado, Mario Esquivel, Mike Fannon, Sam Lee, Carmelo Locurto, Benjamin Murphy, AJ Pualani, Dan Stuligross, Eric Wong, and each of their families. UMass friends, Benjamin Cruikshank, Vanja Dunjko, Jake Golde, Zai Hwang, Ian Mulligan, and Joanna Ruhl. Perimeter friends, Nick White and Shuyi Zhang. Harald Atmanspacher, Ingemar Bengtsson, Florian Boge, Irina-Mihaela Dumitru, David Glick, Richard Healey, Jan-Åke Larsson, Gustavo Rodrigues Rocha, Thomas Ryckman, and Christopher Timpson.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Ninety-five years ago, physicists got something right. Quantum theory has since enjoyed experimental confirmation of a degree unprecedented in the history of science. Its application in materials allowed us to understand semiconductors well enough to design microchips, enabling the modern technological era. Stimulated emission led to lasers and nuclear spin led to magnetic resonance imaging. In the decades following 1925, physicists produced each part of what is now known as the Standard Model, a "quantum field theory" which has allowed for incredible predictive precision in high energy collision experiments. For example, the National Institute of Standards and Technology reports the value of the fine structure constant to a relative standard uncertainty of 1.5×10^{-10} [1]. Even more excitingly, the very same quantum theory is now pointing the way to the next technological era; fully fault-tolerant quantum computers are still speculative, but the so-called Noisy Intermediate-Scale Quantum era, which already promises advantages, has begun [2].

So quantum theory is "right," but what is it right about? What about nature makes it the right theory? In classical physics, physical postulate and formalism closely play off of one another. One might, roughly speaking, think about the classical laws of physics in much the same way that an engineer looks at the schematics of a machine, with the calculations neatly following from a description of its constituents. Classical mechanics even has this conception in the name—one was to think of it as the physics of the machine that is our reality. Quantum theory frustrates this view. Some try to view quantum theory as a point-for-point substitution of classical notions, with wavefunctions as direct, drop-in replacements for the so-called elements of reality. From this perspective, quantum mechanics is simply a more complicated schematic; one with complex numbers and operators, but conceptually the same kind of entity as classical mechanics before it. However, this style of approach is fraught with difficulties which may be insuperable. More importantly, such an association ignores all evidence that the formalism is a different kind of theory, one more about our knowledge than a representation of physical fact [3].

Discussing the relative merits and weaknesses of different approaches to the understanding of quantum theory is not our purpose in this dissertation.[1] Instead, our primary intent is to pursue technical questions which arise from one particularly flexible and compelling epistemic view: QBism, the "subjective Bayesian" approach to quantum theory. There are several excellent accounts of this interpretation [4–8].[2] Although the technical questions we address were motivated by QBist intuitions, it is not necessary for the reader to fully familiarize themselves with QBism prior to reading this dissertation as the pertinent concepts are introduced as they arise in each chapter.

QBism says that quantum theory is an addition to probabilistic decision theory. What does this mean? The basic quantum scenario goes like this. An experimentalist plans to perform an experiment and would like to forecast their expectations for what they will find upon making a measurement at the end. Quantum theory instructs them to pick an initial quantum state for reasoning about the system, a quantum channel to transform the state in accordance with their understanding of the experiment, and a particular quantum

[1] Although we do indulge in Chapter 8.

[2] For nearly 3000 combined pages of correspondence chronicling the development of QBism, see [9] and [10].

measurement comprised of effect operators associated to the outcomes of the measurement they intend to perform. If they manage to do so, the probabilities they seek can then be computed from the final quantum state and the measurement with the Born rule.[3]

The theory tells you how to use quantum states, channels, and measurements, but it doesn't prescribe a particular way of determining them. Take, for example, quantum states. In textbook exercises, the quantum state one needs is typically in the problem statement, but in the lab it is never so simple. Experimenters need to familiarize themselves with the lab, their equipment, calibration samples, and likely many other things before they start to have a hunch that one quantum state will work better than another. Having finally done enough to confidently express a quantum state is often a momentous occasion, involving quite a bit of statistical and probabilistic legwork. This process may take many forms, but the key point is that one's expectations for a sufficient diversity of potential interactions with a system is what determines the state that an experimenter uses. In other words, probabilities fix the quantum state and the quantum state, in turn, may be used to produce probabilities. A similar story can be told for channels and measurements [4]. Cast in this way, quantum theory is about how such input and output probabilities are constrained and these constraints are *in addition* to the logical constraints probability theory itself already mandates.

If this is what quantum theory is from a conceptual standpoint, one immediately questions the necessity of actually using the intermediate mathematical tools of states, channels, and effects. After all, these objects only figure in the figuring; probabilities are the ultimate deliverable. That quantum theory can, in principle, be expressed purely in terms of probabilities is not altogether surprising. And, of course, the particular packaging rep-

[3]The technical setting for what we have sketched is the general formulation used in quantum information theory. For a concise and complete introduction, see [11].

resented by the standard representation has many advantages—there are reasons it is the standard and little reason to expect that any alternative would ever fully supplant it. But useful alternatives may expose traits of the theory which are obscured in the typical conception. For example, in phase space representations of quantum theory, the principle advantage is that aspects of classical phase space analysis become possible within quantum theory. Another example is the path integral formulation. Although equivalent to the standard representation, this purely representational innovation enabled significant advances. On the face of it, a purely probabilistic representation is likely to be hopelessly messy and unilluminating. Wootters put it this way [12]:

> It is obviously *possible* to devise a formulation of quantum mechanics without probability amplitudes. One is never forced to use any quantities in one's theory other than the raw results of measurements. However, there is no reason to expect such a formulation to be anything other than extremely ugly.

If this intuition holds, then while the considerations above may express quantum theory's conceptual meaning, these insights may have less lasting practical value than the analogies which led to previous representations have had.

The central theme of this dissertation is that this intuition does not hold. Not only are there formulations which are not extremely ugly, there are clean and physically illuminating options. The key to these formulations is a class of quantum measurements called minimal informationally complete quantum measurements (MICs).

In Chapter 2 we will thoroughly introduce this class of measurements. MICs are *quantum reference measurements*. This means that a probability distribution for the outcomes of a single MIC measurement is equivalent to a specification of a quantum state, and the fact that they are "minimal" refers to the fact that the probability distribution has the fewest number of entries possible. MICs illuminate the structure of quantum theory and how it departs from the classical. In this chapter, we establish general properties of MICs, explore constructions of several classes of them, and make some developments to the theory of MIC Gram matrices, that is, the matrix of Hilbert–Schmidt inner products of MIC effect operators. These Gram matrices turn out to be a rich subject of inquiry, relating linear algebra, number theory and probability. Unlike the measurement operators corresponding to quantum observables, a MIC can never be an orthogonal set. In a deep sense, we will see that the ideal measurements of quantum physics are not orthogonal bases. In this chapter we also meet the symmetric informationally complete quantum measurements (SICs), a special class of MICs distinguished by being *rank-1*, meaning all effects are proportional to rank-1 projectors, and *equiangular*, meaning the Hilbert–Schmidt inner product between any two distinct effects is the same. SICs are in many ways optimal among MICs. This chapter provides further context to this view.

In Chapter 3 we describe a general procedure for associating a MIC with a purely probabilistic representation of the Born Rule. Such representations provide a way to understand the Born Rule as a consistency condition between probabilities assigned to the outcomes of one experiment in terms of the probabilities assigned to the outcomes of other experiments. In this setting, the difference between quantum and classical physics is the way their physical assumptions augment bare probability theory: Classical physics corresponds to a trivial augmentation—one just applies the Law of Total Probability (LTP) between the scenarios—while quantum theory makes use of the Born Rule expressed in one or another of the forms of our general procedure. To mark the *irreducible* difference between quantum and classical, one should seek the representations that minimize the disparity between the expressions. We prove that the representation of the Born Rule obtained from a SIC minimizes this distinction in at least two senses—the first to do with unitarily invariant distance measures between the rules, and the second to do with available volume in a reference probability simplex (roughly speaking a new kind of uncertainty principle). Both of these arise from a useful result in majorization theory.

The reference process described in Chapter 3 constitutes an "entanglement breaking" quantum channel, a channel which always produces a probabilistic mixture of a previously agreed upon set of states. In Chapter 4 we consider an alternative channel based on the most familiar quantum state update rule, the Lüders rule. Using this construction, we establish an if-and-only-if condition for the existence of a d-dimensional SIC in terms of a particular depolarizing channel. Moreover, the channel in question satisfies two entropic optimality criteria.

The desire to retain some phase space concepts in quantum mechanics inspired the development of Wigner functions. Although Wigner functions provide a way to do quantum physics with a phase space, they require the use of "probability" distributions that can go negative, that is, quasiprobabilities. This is an extremely unsatisfying move from our perspective, as the decision theoretic character of true probabilities is central to our strategy for developing a new representation. In Chapter 5 we lay the groundwork for relating MIC and Wigner function representations and identify the need to translate the many decades of phase space quasiprobability results to the language of reference probabilities. We observe that the operator bases corresponding to minimal discrete Wigner functions, which we call Wigner bases, are orthogonalizations of MICs. By not imposing a particular discrete phase space structure at the outset, we are able to push Wigner functions to their limits in a suitably quantified sense, revealing a new way in which SICs are significant. Finally, we speculate that astute choices of MICs from the orthogonalization preimages of Wigner bases may in general give quantum measurements conceptually underlying the associated quasiprobability representations.

The appearance of negative terms in quasiprobability representations of quantum theory is known to be inevitable, and, due to its equivalence with the onset of contextuality, of central interest in quantum computation and information. Until recently, however, nothing has been known about how much negativity is necessary in a quasiprobability representation. Zhu [13] proved that the upper and lower bounds with respect to one type of negativity measure among unbiased Wigner bases are in one-to-one correspondence with the SICs. In Chapter 6 we define a family of negativity measures which includes Zhu's as a special case and consider another member of the family which we call "sum negativity." We prove a sufficient condition for local maxima in sum negativity and find exact global maxima in dimensions 3 and 4. Notably, we find that Zhu's result on the SICs does not generally extend to sum negativity, although the analogous result does hold in dimension 4. At the end of the chapter, the Hoggar SIC in dimension 8 makes an appearance in a conjecture on sum negativity. We anticipate further developments along these lines once the relation between MICs and Wigner bases described in Chapter 5 is brought into the fold.

MICs have thus far received very little attention. In fact, the work in this dissertation represents a significant portion of what has been written about them. The situation may be beginning to change [14, 15]. We hope that the subject will continue to grow in popularity. As we will articulate many times in subsequent chapters, we believe that developing a deep understanding of the structure of these measurements will be instrumental in discovering precisely what it is about reality that makes quantum theory our best means for navigating it.

Our attitude towards physics is an intuitive interplay between the technical and conceptual, striving to knead out a coherent point of view. Chapters 7 and 8 are drawn from the conceptual side. The emphasis of each is QBism itself.

The focus of Chapter 7 concerns two very instructive modifications of the Wigner's friend paradox [16]. One concerns a recent no-go theorem by Frauchiger and Renner [17] and the other is a thought experiment by Baumann and Brukner [18]. The resolution of Wigner's original thought experiment was central to the development of QBist thinking. Careful treatment of these modifications has accordingly sharpened our intuitions. We show that the paradoxical features emphasized in these works disappear once both friend and Wigner are understood as agents on an equal footing with regard to their individual uses of quantum theory. Wigner's action on his friend then becomes, from the friend's perspective, an action the friend takes on Wigner. When two agents take actions on each other, each agent has a dual role as a physical system for the other agent. No user of quantum theory is more privileged than any other.

Chapter 8 is a catalog of frequently asked questions about QBism. These remarks (many of them lighthearted) should be considered supplements to more systematic treatments appearing in the literature.

We take stock of the journey so far in Chapter 9, the conclusion. Although there we conclude this dissertation, the journey has just begun.

1.1 Author Contributions

The following chapters comprise monographs coauthored by the author of this dissertation during his time in the QBism Group at the University of Massachusetts Boston. Three have been published, one will soon appear, two others have been submitted for publication, and one stands alone as a whimsical collection of essay answers to frequently asked questions. If read like a book, the reader will encounter repetition of some concepts in this dissertation from chapter to chapter, presented in different ways. These repetitions are partially an artifact of the chapters being separate manuscripts, each of which had to set the scene anew. But preserving them was also a conscious choice because some points bear repeating; the author himself has occasionally had to hear the same thing many times before it stuck.

Each of the following chapters, excluding the conclusion, is a manuscript. All but Chapter 8 have appeared in publications or have been submitted to journals. In our group, author ordering is alphabetical rather than in contribution order. Accordingly, in the following list we make explicit our role in each paper, referring to other authors by their initials. The chapters have been arranged in a pedagogically motivated order, rather than chronologically.

1. Chapter 2 "MIC Facts" is the main text of [19]:

   J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, "Analysis and Synthesis of Minimal Informationally Complete Quantum Measurements," submitted to *Physical Review X Quantum*.

   I was the primary contributor to this work. I produced about two thirds of the proofs and all of the numerical content and figures. The majority of the remaining third was contributed by BCS and a smaller amount is due to CAF, although typeset by BCS and myself. This document has a long history and it underwent many significant revisions. At a conceptual level, all three authors contributed substantially. The writing portion was close to evenly split between BCS and myself. BCS is responsible for much of the text in the introduction. CAF contributed some of the direction of research, asked important questions, suggested analysis methods for a few of the results, and edited the draft.

2. Chapter 3 "LTP analogs" is the main text and appendices of [20]:

J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems,” *Physical Review Research*, vol. 2, pp. 013074, 2020.

I was the primary contributor this work. I produced the final versions of all proofs and was the primary, although not sole, contributor to their development. I was responsible for the initial mathematical framing which led to the questions we explored; subsequently, we three were equally involved in the research directions. I drafted all technical parts and figures. CAF and BCS wrote the introductory paragraphs, asked important questions, produced early versions of some proof arguments, otherwise suggested analysis methods, and edited the draft.

3. Chapter 4 “Lüders MIC Channels” is the main text and appendices of [21]:

   J. B. DeBrota and B. C. Stacey, “Lüders channels and the existence of symmetric informationally-complete measurements,” *Physical Review A*, vol. 100, p. 062327, 2019.

   I was the primary contributor to this work, but not by a large margin. BCS and I closely collaborated at every stage of this work. I produced the final versions of all proofs. BCS asked important questions, originated some proof direction ideas, and advocated for the emphases which appear in the final version. I drafted all or nearly all technical parts and BCS drafted the majority of the conceptual content.

4. Chapter 5 “The Principal Wigner Function” is the main text of [22]:

   J. B. DeBrota and B. C. Stacey, “Discrete Wigner Functions from Informationally Complete Measurements,” submitted to *Physical Review X Quantum*.

   I was the primary contributor to this work. I proposed the direction of research and conjectured and proved every result in the draft. All technical sections were entirely written by me. BCS wrote about half of the introduction, the majority of Section 2, and about half of the penultimate section, all of which are predominately conceptual content. BCS additionally asked important questions during the research phase.

5. Chapter 6, "Sum Negativity" is the main text of [23]:

   J. B. DeBrota and C. A. Fuchs, "Negativity bounds for Weyl–Heisenberg quasiprobability representations," *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

   I was the primary contributor to this work. I proved all of the results, conducted all numerics, and drafted the paper. CAF proposed some of the direction of research, suggested key analysis methods, asked important questions, and edited the draft.

6. Chapter 7 "Wigner's Friends" is the main text of [24]

   J. B. DeBrota, C. A. Fuchs, and R. Schack, "Respecting One's Fellow: QBism's Analysis of Wigner's Friend," to appear in *Foundations of Physics*, 2020.

   I was the least significant contributor to this work. My primary role was participation in the numerous discussions and brainstorming sessions over the course of the last couple of years which eventually led to the positions represented in the paper. A few of the key conclusions appearing in the draft can be traced to my questions and perspectives. RS wrote the bulk of the draft and CAF wrote some parts. I proofread and suggested revisions and emphases at several stages of the writing process.

7. Chapter 8 "FAQBism" is the main text of [25]:

   J. B. DeBrota and B. C. Stacey, "FAQBism," *arXiv:1810.13401*, 2018.

   BCS was the primary author of this work. I was involved in discussions surrounding all of the topics addressed and I primarily drafted a minority of the subsections.

LIST OF REFERENCES

[1] "Fine-structure constant." The NIST Reference on Constants, Units, and Uncertainty: Fundamental Physical Constants, `https://physics.nist.gov/cgi-bin/cuu/Value?alph`.

[2] J. Preskill, "Quantum Computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[3] R. W. Spekkens, "Evidence for the epistemic view of quantum states: A toy theory," *Physical Review A*, vol. 75, no. 3, p. 032110, 2007.

[4] C. A. Fuchs and B. C. Stacey, "QBism: Quantum Theory as a Hero's Handbook," in *Proceedings of the International School of Physics "Enrico Fermi": Course 197, Foundations of Quantum Theory*, pp. 133–202, 2019.

[5] C. A. Fuchs and R. Schack, "Quantum-Bayesian coherence," *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[6] C. A. Fuchs, "QBism, The Perimeter of Quantum Bayesianism," *arXiv:1003.5209*, 2010.

[7] C. A. Fuchs, "Notwithstanding Bohr, the Reasons for QBism," *Mind and Matter*, vol. 15, no. 2, 2017.

[8] R. Healey, "Quantum-Bayesian and Pragmatist Views of Quantum Theory," in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University, spring 2017 ed., 2017.

[9] C. A. Fuchs, *Coming of age with quantum information: notes on a Paulian idea: selected correspondence, 1995-2001*. Cambridge ; New York: Cambridge University Press, 2011.

[10] C. A. Fuchs, M. Schlosshauer, and B. C. Stacey, "My Struggles with the Block Universe," *arXiv:1405.2390*, 2014.

[11] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, tenth anniversary ed., 2010.

[12] W. K. Wootters, "Quantum mechanics without probability amplitudes," *Foundations of Physics*, vol. 16, no. 4, pp. 391–405, 1986.

[13] H. Zhu, "Quasiprobability representations of quantum mechanics with minimal negativity," *Physical Review Letters*, vol. 117, no. 12, p. 120404, 2016.

[14] V. I. Yashin, E. O. Kiktenko, A. S. Mastiukova, and A. K. Fedorov, “Minimal informationally complete measurements for probability representation of quantum dynamics,” *arXiv:2006.13727*, 2020.

[15] E. O. Kiktenko, A. O. Malyshev, A. S. Mastiukova, V. I. Man’ko, A. K. Fedorov, and D. Chruściński, “Probability representation of quantum dynamics using pseudostochastic maps,” *Physical Review A*, vol. 101, no. 5, p. 052320, 2020.

[16] E. P. Wigner, “Remarks on the Mind-Body Question,” in *The Scientist Speculates* (I. J. Good, ed.), pp. 284–302, London: William Heinemann, Ltd., 1961.

[17] D. Frauchiger and R. Renner, “Quantum theory cannot consistently describe the use of itself,” *Nature Communications*, vol. 9, no. 1, p. 3711, 2018.

[18] V. Baumann and Č. Brukner, “Wigner’s Friend as a Rational Agent,” in *Quantum, Probability, Logic* (M. Hemmo and O. Shenker, eds.), pp. 91–99, Cham: Springer International Publishing, 2020.

[19] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Analysis and Synthesis of Minimal Informationally Complete Quantum Measurements,” *arXiv:1812.08762*, 2020.

[20] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems,” *Physical Review Research*, vol. 2, p. 013074, 2020.

[21] J. B. DeBrota and B. C. Stacey, “Lüders channels and the existence of symmetric-informationally-complete measurements,” *Physical Review A*, vol. 100, no. 6, p. 062327, 2019.

[22] J. B. DeBrota and B. C. Stacey, “Discrete Wigner Functions from Informationally Complete Quantum Measurements,” *arXiv:1912.07554*, 2020.

[23] J. B. DeBrota and C. A. Fuchs, “Negativity bounds for Weyl–Heisenberg quasiprobability representations,” *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

[24] J. B. DeBrota, C. A. Fuchs, and R. Schack, “Respecting One’s Fellow: QBism’s Analysis of Wigner’s Friend,” to appear in *Foundations of Physics*, 2020.

[25] J. B. DeBrota and B. C. Stacey, “FAQBism,” *arXiv:1810.13401*, 2018.

CHAPTER 2

MIC FACTS

2.1 Introduction

A significant part of science is the pursuit of measurements that are as informative as possible. Attempts to provide an elementary explanation of “the scientific method” sometimes convey the notion that an ideal measurement is one which is exactly reproducible, always yielding the same answer when applied in succession. But this notion has fairly obvious problems, for example, when the system being measured is dynamical. When the experiment’s sought outcome is the position of Mars at midnight, the numbers will not be the same from one night to the next, and yet Kepler could run a scientific revolution on that data. A more refined standard would be that an ideal measurement is one that provides enough information to project the complete dynamical trajectory of the measured system through phase space. Quantum physics frustrates this ambition by denying the phase space: Quantum uncertainties are not uncertainties about the values of properties that pre-exist the act of measurement. Yet the ideal of a sufficiently informative measurement, the expectations for which fully fix the expectations for any other, can still be translated from classical thought to quantum, and doing so illuminates the nature of quantum theory itself.

Let $\mathcal{H}_d$ be a d-dimensional complex Hilbert space, and let $\{E_i\}$ be a set of positive semidefinite operators on that space which sum to the identity:

$$\sum_{i=1}^{N} E_i = I. \tag{2.1}$$

The set $\{E_i\}$ is a *positive-operator-valued measure* (POVM), which is the mathematical representation of a measurement process in quantum theory. Each element in the set — called an *effect* — stands for a possible outcome of the measurement [1, §2.2.6]. A POVM is said to be *informationally complete* (IC) if the operators $\{E_i\}$ span $\mathcal{L}(\mathcal{H}_d)$, the space of Hermitian operators on $\mathcal{H}_d$, and an IC POVM is said to be *minimal* if it contains exactly d^2 elements. For brevity, we can call a minimal IC POVM a MIC.

A matrix which captures many important properties of a MIC is its Gram matrix, that is, the matrix G whose entries are given by

$$[G]_{ij} := \operatorname{tr} E_i E_j \ . \tag{2.2}$$

Of particular note among MICs are those which enjoy the symmetry property

$$[G]_{ij} = [G_{\text{SIC}}]_{ij} := \frac{1}{d^2}\frac{d\delta_{ij} + 1}{d+1} \ . \tag{2.3}$$

These are known as *symmetric* informationally complete POVMs, or SICs for short [2–5]. In addition to their purely mathematical properties, SICs are of central interest to the technical side of QBism, a research program in the foundations of quantum mechanics [6–9]. Investigations motivated by foundational concerns led to the discovery that SICs are in many ways optimal among MICs [10–12]. In this paper, we elaborate upon some of those results and explore the conceptual context of MICs more broadly.

MICs provide a new way of understanding the Born Rule, a key step in how one uses quantum physics to calculate probabilities. The common way of presenting the Born Rule suggests that it fixes probabilities in terms of more fundamental quantities, namely quantum states and measurement operators. MICs, however, suggest a change of viewpoint. From this new perspective, the Born Rule should be thought of as a *consistency condition between the probabilities assigned in diverse scenarios* — for instance, probabilities assigned to the outcomes of complementary experiments. The bare axioms of probability theory do not themselves impose relations between probabilities given different conditionals: In the abstract, nothing ties together $P(E|C_1)$ and $P(E|C_2)$. Classical intuition suggests one way to fit together probability assignments for different experiments, and quantum physics implies another. The discrepancy between these standards encapsulates how quantum theory departs from classical expectations [13, 14]. MICs provide the key to addressing this discrepancy; any MIC may play the role of a reference measurement through which the quantum consistency condition may be understood. To understand MICs is to understand how quantum probability is like, and differs from, classical.

In the next section, we introduce the fundamentals of quantum information theory and the necessary concepts from linear algebra to prove a few basic results about MICs and comment on their conceptual meaning. Among the results included are a characterization of unbiased MICs, a condition in terms of matrix rank for when a set of vectors in $\mathbb{C}^d$ can be fashioned into a MIC, and an explicit example of an unbiased MIC which is not group covariant. In Section 2.3, we show how to construct several classes of MICs explicitly and note some properties of their Gram matrices. In Section 2.4, we explore several ways in which SICs are optimal among MICs for the project of differentiating the quantum from the classical, a topic complementing one of our recent papers [12]. To conclude, in Section 2.5, we conduct an initial numerical study of the Gram matrix eigenvalue spectra of randomly-chosen MICs of four different types. The empirical eigenvalue distributions we find have intriguing features, not all of which have been explained yet.

2.2 Basic Properties of MICs

We begin by briefly establishing the necessary notions from quantum information theory on which this paper is grounded. In quantum physics, each physical system is associated with a complex Hilbert space. Often, in quantum information theory, the Hilbert space of interest is taken to be finite-dimensional. We will denote the dimension throughout by d. A *quantum state* is a positive semidefinite operator of unit trace. The extreme points in the space of quantum states are the rank-1 projection operators:

$$\rho = |\psi\rangle\langle\psi|. \tag{2.4}$$

These are idempotent operators; that is, they all satisfy $\rho^2 = \rho$. If an experimenter ascribes the quantum state ρ to a system, then she finds her probability for the i^{th} outcome of the measurement modeled by the POVM $\{E_i\}$ via the Hilbert–Schmidt inner product:

$$p(E_i) = \operatorname{tr} \rho E_i. \tag{2.5}$$

This formula is a standard presentation of the Born Rule. The condition that the $\{E_i\}$ sum to the identity ensures that the resulting probabilities are properly normalized.

If the operators $\{E_i\}$ span the space of Hermitian operators, then the operator ρ can be reconstructed from its inner products with them. In other words, the state ρ can be calculated from the probabilities $\{p(E_i)\}$, meaning that the measurement is "informationally complete" and the state ρ can, in principle, be dispensed with. Any MIC can thus be considered a "Bureau of Standards" measurement, that is, a reference measurement in terms of which all states and processes can be understood [15]. Writing a quantum state ρ is often thought of as specifying the "preparation" of a system, though this terminology is overly restrictive, and the theory applies just as well to physical systems that were not processed on a laboratory workbench [16].

Given any POVM $\{E_i\}$, we can always write its elements as unit-trace positive semidefinite operators with appropriate scaling factors we call *weights*:

$$E_i := e_i \rho_i, \text{ where } e_i = \operatorname{tr} E_i. \tag{2.6}$$

If the operators ρ_i are all rank-1 projectors, we will refer to the set $\{E_i\}$ as a *rank-1 POVM*. We will call a POVM *unbiased* when the weights e_i are all equal. Such operator sets represent quantum measurements that have no intrinsic bias: Under the Born Rule they map the "garbage state" $(1/d)I$ to a flat probability distribution. For an unbiased MIC, the condition that the elements sum to the identity then fixes $e_i = 1/d$.

A *column (row) stochastic matrix* is a real matrix with nonnegative entries whose columns (rows) sum to 1. If a matrix is both column and row stochastic we say it is *doubly stochastic*. The following theorem allows us to identify an unbiased MIC from a glance at its Gram matrix or Gram matrix spectrum.

Theorem 1. *Let $\{E_i\}$ be a MIC and $\lambda_{\max}(G)$ be the maximal eigenvalue of its Gram matrix G. The following are equivalent:*

1. *$\{E_i\}$ is unbiased.*

2. *dG is doubly stochastic.*

3. *$\lambda_{\max}(G) = 1/d$.*

Proof. The equivalence of the first two conditions is readily shown. We show (2) $\iff$ (3). Let $|v\rangle := \frac{1}{d}(1, \ldots, 1)^{\mathrm{T}}$ be the normalized d^2 element uniform vector of 1s. If dG is doubly stochastic, $|v\rangle$ is an eigenvector of dG with eigenvalue 1, and the Gershgorin disc theorem [17] ensures $\lambda_{\max}(G) = 1/d$. For any MIC,

$$\lambda_{\max}(G) \geq \langle v|G|v\rangle = \frac{1}{d} \;, \tag{2.7}$$

with equality iff $|v\rangle$ is an eigenvector of G with eigenvalue $1/d$. Since $G|v\rangle = (e_1, \ldots, e_{d^2})^{\mathrm{T}}$, $|v\rangle$ is an eigenvector of G iff $e_i = 1/d$ for all i. □

Given a basis for an inner product space, the *dual basis* is defined by the condition that the inner products of a vector with the elements of the dual basis provide the coefficients in the expansion of that vector in terms of the original basis. In our case, let $\{\widetilde{E}_i\}$ denote the basis dual to $\{E_i\}$ so that, for any vector $A \in \mathcal{L}(\mathcal{H}_d)$,

$$A = \sum_j (\mathrm{tr}\, A\widetilde{E}_j)E_j \;. \tag{2.8}$$

One consequence of this definition is that if we expand the original basis in terms of itself,

$$E_i = \sum_j (\mathrm{tr}\, E_i\widetilde{E}_j)E_j \;, \tag{2.9}$$

linear independence of the $\{E_i\}$ implies that

$$\mathrm{tr}\, E_i\widetilde{E}_j = \delta_{ij} \;, \tag{2.10}$$

from which one may easily see that the original basis is the dual of the dual basis,

$$A = \sum_j (\mathrm{tr}\, AE_j)\widetilde{E}_j \; . \tag{2.11}$$

In the familiar case when the original basis is orthonormal, the dual basis coincides with it: When we write a vector $\mathbf{v}$ as an expansion over the unit vectors $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$, the coefficient of $\hat{\mathbf{x}}$ is simply the inner product of $\hat{\mathbf{x}}$ with $\mathbf{v}$.

A MIC is a positive semidefinite operator basis. For positive semidefinite operators A and B, $\mathrm{tr}\, AB = 0$ iff $AB = 0$. Recall that a Hermitian matrix which is neither positive semidefinite nor negative semidefinite is known as an *indefinite* matrix.

Theorem 2. *The dual basis of a MIC is composed entirely of indefinite matrices.*

Proof. Suppose $\widetilde{E}_1 \geq 0$. The definition of a dual basis tells us $\mathrm{tr}\, \widetilde{E}_1 E_k = 0$ for all $k \neq 1$. Because they are both positive semidefinite, $\widetilde{E}_1 E_k = 0$ for all $k \neq 1$. This means the $d^2 - 1$ MIC elements other than E_1 are operators on a $d - \mathrm{rank}(\widetilde{E}_1)$ dimensional subspace. But

$$\dim \left[\mathcal{L}\left(\mathcal{H}_{d-\mathrm{rank}(\widetilde{E}_1)} \right) \right] \leq (d-1)^2 < d^2 - 1, \tag{2.12}$$

so they cannot be linearly independent. If $\widetilde{E}_1 \leq 0$, $-\widetilde{E}_1$ is positive semidefinite and the same logic holds. □

Corollary. *No element in a MIC can be proportional to an element of the MIC's dual basis.*

Corollary. *No MIC can form an orthogonal basis.*

Proof. Suppose $\{E_i\}$ is a MIC which forms an orthogonal basis, that is, $\mathrm{tr}\, E_i E_j = c_i \delta_{ij}$ for some constants c_i. Summing this over i reveals $c_j = e_j$, the weights of the MIC. Thus the dual basis is given by $\widetilde{E}_j = E_j / e_j = \rho_j$ which is a violation of Corollary 2.2. □

Corollary. *No MIC outcome can ever be assigned probability* 1.

Proof. MIC probabilities provide the expansion coefficients for a state in the dual basis. If $P(E_i) = 1$ for some i, the state would equal the dual basis element, but a state must be positive semidefinite. $\square$

Corollary. *No effect of a MIC can be an unscaled projector.*

Proof. Suppose E_1 were equal to an unscaled projector P. Then any eigenvector of P is a pure state which would imply probability 1 for the MIC outcome E_1, which is impossible. $\square$

Theorem 2.2 and the subsequent corollaries have physical meaning. In classical probability theory, we grow accustomed to orthonormal bases. For example, imagine an object that can be in any one of N distinct configurations. When we write a probability distribution over these N alternatives, we are encoding our expectations about which of these configurations is physically present — about the "physical condition" of the object, as Einstein would say [18], or in more modern terminology, about the object's "ontic state" [19]. We can learn everything there is to know about the object by measuring its "physical condition", and any implementation of such an ideal measurement is represented by conditional probabilities that are 1 in a single entry and 0 elsewhere. In other words, the map from the object's physical configuration to the reading on the measurement device is, at its most complicated, a permutation of labels. Without loss of generality, we can take the vectors that define the ideal measurement to be the vertices of the probability simplex: The measurement basis is identical with its dual, and the dual-basis elements simply label the possible "physical conditions" of the object which the measurement reads off.

In quantum theory, by contrast, no element of a MIC may be proportional to an element in the dual. This stymies the identification of the dual-basis elements as intrinsic "physical conditions" ready for a measurement to read.

Theorem 3. *No elementwise rescaling of a proper subset of a MIC may form a POVM.*

Proof. Since a MIC is a linearly independent set, the identity element is uniquely formed by the defining expression

$$I = \sum_{i=1}^{d^2} E_i. \tag{2.13}$$

If a linear combination of a proper subset Ω of the MIC elements could be made to also sum to the identity,

$$I = \sum_{i\in\Omega} \alpha_i E_i, \tag{2.14}$$

then subtracting (2.14) from (2.13) implies

$$0 = \sum_{i\in\Omega}(1-\alpha_i)E_i + \sum_{i\notin\Omega} E_i \tag{2.15}$$

which is a violation of linear independence. □

Corollary. *No two elements in a $d = 2$ MIC may be orthogonal under the Hilbert–Schmidt inner product.*

Proof. An orthogonal pair of elements in dimension 2 may be rescaled such that they sum to the identity element. Therefore, by Theorem 3, they cannot be elements of a MIC. □

These results also have physics implications. For much of the history of quantum mechanics, one type of POVM had special status: the *von Neumann measurements,* which consist of d elements given by the projectors onto the vectors of an orthonormal basis of $\mathbb{C}^d$. Indeed, in older books, these are the only quantum measurements that are considered (often being defined as the eigenbases of Hermitian operators called "observables"). We can now see that, from the standpoint of informational completeness, the von Neumann measurements are rather pathological: There is no way to build a MIC by augmenting a von Neumann measurement with additional outcomes.

Another holdover from the early days of quantum theory concerns the process of updating a quantum state in response to a measurement outcome. If one restricts attention to von Neumann measurements, one may feel tempted to grant special importance to the post-measurement state being one of the eigenvectors of an "observable". This type of updating is a special case of the more general theory developed as quantum mechanics was understood more fully. The *Lüders Rule* [20, 21] states that the post-measurement state upon obtaining the outcome associated with effect E_i for a POVM $\{E_i\}$ is

$$\rho_i' := \frac{\sqrt{E_i}\rho\sqrt{E_i}}{\operatorname{tr}\rho E_i} \ . \tag{2.16}$$

In the special case of a von Neumann measurement, this reduces to replacing the state for the system with the eigenprojector corresponding to the measurement outcome. A physicist who plans to follow that procedure and then repeat the measurement immediately afterward would expect to obtain the same outcome twice in succession. Some authors regard this possibility as the essential point of contact with classical mechanics and attempt to build an understanding of quantum theory around such "ideal" measurements [22]. But, as we said in the introduction, obtaining the same outcome twice in succession is not a good notion of a "classical ideal". Especially in view of the arbitrariness of von Neumann measurements from our perspective, we regard this possibility as conceptually downstream from the phenomenon of informationally complete measurements.

Corollary 2.2 prompts a question: May any elements of a MIC in arbitrary dimension be orthogonal? In other words, can any entry in a G matrix equal zero? We answer this question in the affirmative with an explicit example of a rank-1 MIC in dimension 3 with 7 orthogonal pairs.

Example. When multiplied by $1/3$, the following is a rank-1 unbiased MIC in dimension 3 with 7 orthogonal pairs.

$$\left\{\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix},\begin{bmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&0&0\\\frac{1}{2}&0&\frac{1}{2}\end{bmatrix},\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{1}{2}\end{bmatrix},\begin{bmatrix}\frac{1}{2}&0&\frac{i}{2}\\0&0&0\\-\frac{i}{2}&0&\frac{1}{2}\end{bmatrix},\begin{bmatrix}0&0&0\\0&\frac{1}{2}&\frac{i}{2}\\0&-\frac{i}{2}&\frac{1}{2}\end{bmatrix},\begin{bmatrix}\frac{1}{3}&\frac{i}{3}&-\frac{i}{3}\\-\frac{i}{3}&\frac{1}{3}&-\frac{1}{3}\\\frac{i}{3}&-\frac{1}{3}&\frac{1}{3}\end{bmatrix},\right.$$
$$\left.\begin{bmatrix}\frac{5}{8}&-\frac{1}{8}-\frac{i}{4}&-\frac{3}{8}-\frac{i}{8}\\-\frac{1}{8}+\frac{i}{4}&\frac{1}{8}&\frac{1}{8}-\frac{i}{8}\\-\frac{3}{8}+\frac{i}{8}&\frac{1}{8}+\frac{i}{8}&\frac{1}{4}\end{bmatrix},\begin{bmatrix}\frac{1}{24}&\frac{1}{8}-\frac{i}{12}&-\frac{1}{8}-\frac{i}{24}\\\frac{1}{8}+\frac{i}{12}&\frac{13}{24}&-\frac{7}{24}-\frac{3i}{8}\\-\frac{1}{8}+\frac{i}{24}&-\frac{7}{24}+\frac{3i}{8}&\frac{5}{12}\end{bmatrix}\right\}. \tag{2.17}$$

These are projectors onto the following vectors in $\mathcal{H}_d$:

$$\left\{(1,0,0),(0,1,0),\frac{1}{\sqrt{2}}(1,0,1),\frac{1}{2}(0,1,1),\frac{1}{\sqrt{2}}(1,0,-i),\frac{1}{\sqrt{2}}(0,1,-i),\frac{1}{\sqrt{3}}(1,-i,i),\right.$$
$$\left.\frac{1}{\sqrt{40}}(5,-1+2i,-3+i),\frac{1}{\sqrt{24}}(1,3+2i,-3+i)\right\}. \tag{2.18}$$

The Gram matrix of the MIC elements is

$$\begin{bmatrix} \frac{1}{9} & 0 & \frac{1}{18} & 0 & \frac{1}{18} & 0 & \frac{1}{27} & \frac{5}{72} & \frac{1}{216} \\ 0 & \frac{1}{9} & 0 & \frac{1}{18} & 0 & \frac{1}{18} & \frac{1}{27} & \frac{1}{72} & \frac{13}{216} \\ \frac{1}{18} & 0 & \frac{1}{9} & \frac{1}{36} & \frac{1}{18} & \frac{1}{36} & \frac{1}{27} & \frac{1}{144} & \frac{5}{432} \\ 0 & \frac{1}{18} & \frac{1}{36} & \frac{1}{9} & \frac{1}{36} & \frac{1}{18} & 0 & \frac{5}{144} & \frac{1}{48} \\ \frac{1}{18} & 0 & \frac{1}{18} & \frac{1}{36} & \frac{1}{9} & \frac{1}{36} & 0 & \frac{5}{144} & \frac{1}{48} \\ 0 & \frac{1}{18} & \frac{1}{36} & \frac{1}{18} & \frac{1}{36} & \frac{1}{9} & \frac{1}{27} & \frac{1}{144} & \frac{5}{432} \\ \frac{1}{27} & \frac{1}{27} & \frac{1}{27} & 0 & 0 & \frac{1}{27} & \frac{1}{9} & \frac{1}{54} & \frac{1}{18} \\ \frac{5}{72} & \frac{1}{72} & \frac{1}{144} & \frac{5}{144} & \frac{5}{144} & \frac{1}{144} & \frac{1}{54} & \frac{1}{9} & \frac{1}{27} \\ \frac{1}{216} & \frac{13}{216} & \frac{5}{432} & \frac{1}{48} & \frac{1}{48} & \frac{5}{432} & \frac{1}{18} & \frac{1}{27} & \frac{1}{9} \end{bmatrix}. \tag{2.19}$$

The process of finding this example led us to formulate the following:

Conjecture 1. A rank-1 MIC in dimension 3 can have no more than 7 pairs of orthogonal elements.

Our next result characterizes when it is possible to build a rank-1 POVM out of a set of vectors and specifies the additional conditions which must be met in order for it to form a MIC. We make use of the *Hadamard product* [23], denoted $\circ$, which is element-wise multiplication of matrices.

Theorem 4. *Consider a set of N normalized vectors $|\phi_i\rangle$ in $\mathcal{H}_d$ and real numbers $0 \leq e_i \leq 1$. The following are equivalent:*

1. $E_i := e_i|\phi_i\rangle\langle\phi_i|$ *forms a rank-1 POVM.*

2. *The Gram matrix g of the rescaled vectors $\sqrt{e_i}|\phi_i\rangle$ is a rank-d projector.*

Furthermore, if $N = d^2$ *and* $\mathrm{rank}(g \circ g^*) = d^2$*,* $\{E_i\}$ *forms a rank-1 MIC.*

Proof. Suppose E_i forms a rank-1 POVM, that is,

$$\sum_i e_i |\phi_i\rangle\langle\phi_i| = I \; . \tag{2.20}$$

It is easy to see that this is only possible if the set $\{\sqrt{e_i}|\phi_i\rangle\}$ spans $\mathcal{H}_d$, and, consequently, $N \geq d$. It now follows that g is a rank-d projector because the left hand side of (2.20) is a matrix that has the same nonzero spectrum as g [24]. On the other hand, if g is a rank-d projector, $N \geq d$ and $\{\sqrt{e_i}|\phi_i\rangle\}$ spans a d dimensional space because the rank of a Gram matrix is equal to the dimension of the space spanned by the vectors. Using again the fact the left hand side of (2.20) has the same nonzero spectrum as the Gram matrix, it must equal the identity and thus the rank-1 POVM condition holds.

To be a MIC, N must equal d^2. The remaining condition on $\{E_i\}$ for it to form a rank-1 MIC is that its elements be linearly independent. This is equivalent to the condition that its Gram matrix G is full rank. The relation between g and G is given by the Hadamard product of g with its conjugate,

$$g \circ g^* = G \; , \tag{2.21}$$

and so, if $N = d^2$ and $\mathrm{rank}(g \circ g^*) = d^2$, $\{E_i\}$ forms a rank-1 MIC. □

For any two matrices A and B, the Hadamard product satisfies the rank inequality

$$\text{rank}(A \circ B) \leq \text{rank}(A)\ \text{rank}(B) \; , \tag{2.22}$$

so a rank-1 MIC is produced when rank$(g \circ g^*)$ achieves its maximal value with the minimal number of effects. Perhaps this criterion will lead to a way to conceptualize rank-1 MICs directly in terms of the vectors in $\mathcal{H}_d$ from which they can be constructed.

As a brief illustration, all rank-d projectors are unitarily equivalent so the specification of g for the rescaled vectors of any rank-1 MIC is obtainable from any rank-d projector by conjugating it with the right unitary. Specifying g specifies the MIC: g is equal to its own square root, so its columns are the vectors up to unitary equivalence which form this Gram matrix. To obtain these vectors as elements of $\mathcal{H}_d$, one can simply write them in the basis provided by the eigenvectors of g with nonzero eigenvalues. From a fixed starting projector, then, finding a rank-1 MIC is equivalent to choosing a unitary in $U(d^2)$ which maximizes rank$(g \circ g^*)$. Numerically this maximization appears to be typical, but we are not aware of an explicit characterization. A further question to ask is whether there are special classes of unitaries which give particular types of MICs.

We finish this section with a very brief discussion of the geometry of MIC space. For this purpose it is not necessary to distinguish between MICs which differ only in permutations of their effects. We further discuss this in the loose sense of not having chosen any particular metric. The full sets of N-outcome POVMs are in general convex manifolds [25], but the requirement of linear independence prevents this from being true for MICs — it is possible for a convex mixture of MICs to introduce a linear dependence and thus step outside of the set. There do, however, exist infinite sequences and curves lying entirely within the set of MICs. In these terms one can see that the space of MICs lacks much of its boundary, that is, one can construct infinite sequences of MICs for which the limit point is not a MIC. The simplest such limit point is the POVM consisting of the identity and $d^2 - 1$ zero matrices. Similarly there are MICs arbitrarily close to any POVM with fewer than d^2 elements which has been padded by zero matrices. Among unbiased MICs, another limit point lying outside of the set is the trivial POVM consisting of d^2 identical matrices $E_i = \frac{1}{d^2}I$. Provided they exist, SICs are limit points, at least among equiangular MICs (see section 2.3.4), which are contained within the set.

2.3 Explicit Constructions of MICs

2.3.1 SICs

The MICs that have attracted the most interest are the SICs, which in many ways are the optimal MICs [10–12, 26, 27]. SICs were studied as mathematical objects (under the name "complex equiangular lines") before their importance for quantum information was recognized [28–31]. Prior to SICs becoming a physics problem, constructions were known for dimensions $d = 2$, 3 and 8 [32]. Exact solutions for SICs are now known in 79 dimensions:

$$\begin{aligned} d = {} & 2\text{–}28, 30, 31, 35, 37\text{–}39, 42, 43, 48, 49, 52, 53, 57, 61\text{–}63, 67, 73, 74, 78, 79, 84, 91, 93, \\ & 95, 97\text{–}99, 103, 109, 111, 120, 124, 127, 129, 134, 143, 146, 147, 168, 172, 195, 199, \\ & 228, 259, 292, 323, 327, 399, 489, 844, 1299. \end{aligned} \tag{2.23}$$

The expressions for these solutions grow complicated quickly, but there is hope that they can be substantially simplified [33]. Numerical solutions have also been extracted, to high precision, in the following dimensions:

$$d = 2\text{–}193, 204, 224, 255, 288, 528, 725, 1155, 2208. \tag{2.24}$$

Both the numerical and the exact solutions have been found in irregular order and by various methods. Many entries in these lists are due to A. J. Scott and M. Grassl [4, 34, 35]; other explorers in this territory include M. Appleby, I. Bengtsson, T.-Y. Chien, S. T. Flammia, G. S. Kopp and S. Waldron.

Together, these results have created the community sentiment that SICs *should* exist for every finite value of d. To date, however, a general proof is lacking. The current frontier of SIC research extends into algebraic number theory [36–40], which among other things has led to a method for uplifting numerical solutions to exact ones [41]. The topic has begun to enter the textbooks for physicists [42] and for mathematicians [24].

The effects of a SIC are given by

$$E_i = \frac{1}{d}\Pi_i, \text{ where } \Pi_i = |\pi_i\rangle\langle\pi_i| \,, \tag{2.25}$$

where we will take the liberty of calling any of the sets $\{E_i\}$, $\{\Pi_i\}$, and $\{|\pi_i\rangle\}$ SICs. It is difficult to find a meaningful visualization of structures in high-dimensional complex vector space. However, for the $d = 2$ case, an image is available. Any quantum state for a 2-dimensional system can be written as an expansion over the Pauli matrices:

$$\rho = \frac{1}{2}\left(I + x\sigma_x + y\sigma_y + z\sigma_z\right). \tag{2.26}$$

The coefficients (x, y, z) are then the coordinates for ρ in the *Bloch ball*. The surface of this ball, the *Bloch sphere,* lives at radius 1 and is the set of pure states. In this picture, the quantum states $\{\Pi_i\}$ comprising a SIC form a regular tetrahedron; for example,

$$\Pi_{s,s'} = \frac{1}{2}\left(I + \frac{1}{\sqrt{3}}\left(s\sigma_x + s'\sigma_y + ss'\sigma_z\right)\right), \tag{2.27}$$

where s and s' take the values ± 1.

The matrix G_{SIC} has the spectrum

$$\lambda(G_{\text{SIC}}) = \left(\frac{1}{d}, \frac{1}{d(d+1)}, \ldots, \frac{1}{d(d+1)}\right). \tag{2.28}$$

The flatness of this spectrum will turn out to be significant; we will investigate this point in depth in the next section.

2.3.2 MICs from Random Bases

It is possible to construct a MIC for any dimension d. Let $\{A_i\}$ be any basis of positive semidefinite operators in $\mathcal{L}(\mathcal{H}_d)$ and define $\Omega := \sum_i A_i$. Then

$$E_i := \Omega^{-1/2} A_i \Omega^{-1/2} \tag{2.29}$$

forms a MIC. If $\{A_i\}$ consists entirely of rank-1 matrices, we obtain a rank-1 MIC.[1] If $\{A_i\}$ is already a MIC, $\Omega = I$ and the transformation is trivial; MICs are the fixed points of this mapping from one positive semidefinite operator basis to another. Thanks to this property, this method can produce *any* MIC if the initial basis is drawn from the full space of positive semidefinite operators.

This procedure was used by Caves, Fuchs and Schack in the course of proving a quantum version of the de Finetti theorem [43]. (For background on this theorem, a key result in probability theory, see [44, §5.3] and [45].) We refer to the particular MICs they constructed as the *orthocross MICs.* As the orthocross MICs are of historical importance, we explicitly detail their construction and provide some first properties and conjectures about it in the remainder of this subsection.

To construct an orthocross MIC in dimension d, first pick an orthonormal basis $\{|j\rangle\}$. This is a set of d objects, and we want a set of d^2, so our first step is to take all possible

[1] In the rank-1 case, this procedure is equivalent to forming what is called the *canonical tight frame* associated with the frame of vectors in $\mathcal{H}_d$ whose outer products form the A_i matrices. For more information on this, see [24].

combinations:

$$\Gamma_{jk} := |j\rangle\langle k|. \tag{2.30}$$

The orthocross MIC will be built from a set of d^2 rank-1 projectors $\{\Pi_\alpha\}$, the first d of which are given by

$$\Pi_\alpha = \Gamma_{\alpha\alpha}. \tag{2.31}$$

Then, for $\alpha = d+1, \ldots, \frac{1}{2}d(d+1)$, we take all the quantities of the form

$$\frac{1}{2}\left(|j\rangle + |k\rangle\right)\left(\langle j| + \langle k|\right) = \frac{1}{2}(\Gamma_{jj} + \Gamma_{kk} + \Gamma_{jk} + \Gamma_{kj}), \tag{2.32}$$

where $j < k$. We construct the rest of the $\{\Pi_\alpha\}$ similarly, by taking all quantities of the form

$$\frac{1}{2}\left(|j\rangle + i|k\rangle\right)\left(\langle j| - i\langle k|\right) = \frac{1}{2}(\Gamma_{jj} + \Gamma_{kk} - i\Gamma_{jk} + i\Gamma_{kj}), \tag{2.33}$$

where again the indices satisfy $j < k$. That is, the set $\{\Pi_\alpha\}$ contains the projectors onto the original orthonormal basis, as well as projectors built from the "cross terms".

The operators $\{\Pi_\alpha\}$ form a positive semidefinite operator basis which can be plugged into the procedure described above. Explicitly,

$$\Omega = \sum_{\alpha=1}^{d^2} \Pi_\alpha \ , \tag{2.34}$$

and the orthocross MIC elements are given by

$$E_\alpha := \Omega^{-1/2}\Pi_\alpha\Omega^{-1/2} \ . \tag{2.35}$$

The operator Ω for the initial set of vectors has a comparatively simple matrix representation: The elements along the diagonal are all equal to d, the elements above the diagonal are all equal to $\frac{1}{2}(1-i)$, and the rest are $\frac{1}{2}(1+i)$, as required by $\Omega = \Omega^\dagger$. The matrix Ω is not quite a circulant matrix, thanks to that change of sign, but it can be turned into one by conjugating with a diagonal unitary matrix. Consequently, the eigenvalues of Ω can be found explicitly via discrete Fourier transformation. The result is that, for $m = 0, \ldots, d-1$,

$$\lambda_m = d + \frac{1}{2}\left(\cot\frac{\pi(4m+1)}{4d} - 1\right). \tag{2.36}$$

This mathematical result has a physical implication [15].

Theorem 5. *The probability of any outcome E_α of an orthocross MIC, given any quantum state ρ, is bounded above by*

$$P(E_\alpha) \leq \left[d - \frac{1}{2}\left(1 + \cot\frac{3\pi}{4d}\right)\right]^{-1} < 1. \tag{2.37}$$

Proof. The maximum of $\mathrm{tr}\,(\rho E_\alpha)$ over all ρ is bounded above by the maximum of $\mathrm{tr}\,(\Pi E_\alpha)$, where Π ranges over the rank-1 projectors. In turn, this is bounded above by the maximum eigenvalue of E_α. We then invoke that

$$\lambda_{\max}(E_\alpha) = \lambda_{\max}(\Omega^{-1/2}\Pi_\alpha\Omega^{-1/2}) = \lambda_{\max}(\Pi_\alpha\Omega^{-1}\Pi_\alpha) \leq \lambda_{\max}(\Omega^{-1}). \tag{2.38}$$

The desired bound then follows. □

Note that all the entries in the matrix 2Ω are Gaussian integers, that is, numbers whose real and imaginary parts are integers. Consequently, all the coefficients in the characteristic polynomial of 2Ω will be Gaussian integers, and so the eigenvalues of 2Ω will be roots of a monic polynomial with Gaussian-integer coefficients. This is an example of how, in the study of MICs, number theory becomes relevant to physically meaningful quantities — in this case, a bound on the maximum probability of a reference-measurement outcome. Number theory has also turned out to be very important for SICs, in a much more sophisticated way [36–40].

The following conjectures about orthocross MICs have been motivated by numerical investigations. We suspect that their proofs will be relatively straightforward, but so far they have eluded us.

Conjecture 2. The entries in G for orthocross MICs can become arbitrarily small with increasing d, but no two elements of an orthocross MIC can be exactly orthogonal.

Conjecture 3. For any orthocross MIC, the entries in G^{-1} are integers or half-integers.

2.3.3 Group Covariant MICs

The method discussed in the previous subsection allows us to make fully arbitrary MICs, but it is also possible to construct MICs with much more built-in structure. The MICs which have received the most attention in the literature to date are the *group covariant* MICs — those whose elements are the orbit of a group of unitary matrices acting by conjugation. For additional discussion of group covariant IC POVMs, see [46].

The Gram matrix of a group covariant MIC is very simple. Suppose $\{E_i\}$ is a group covariant MIC, so $E_i = U_i E_0 U_i^\dagger$ where E_0 is the first element of the MIC and the index i gives the element of the unitary representation of the group sending this element to the ith element. Then all distinct elements of the Gram matrix are present in the first row because

$$[G]_{ij} = \mathrm{tr}\, E_i E_j = \mathrm{tr}\, U_i E_0 U_i^\dagger U_j E_0 U_j^\dagger = \mathrm{tr}\, E_0 U_k E_0 U_k^\dagger\,, \tag{2.39}$$

for some k determined by the group. Another way to say this is that every row of the Gram matrix of a group covariant MIC is some permutation of the first row.

Note that any group covariant MIC is unbiased because conjugation by a unitary cannot change the trace of a matrix, but the converse is not true; the simplest example of an unbiased MIC which is not group covariant which we have encountered is the one given in Example 2.2.

The most important group covariant MICs are the *Weyl–Heisenberg* MICs (WH MICs), which are covariant with respect to the Weyl–Heisenberg group, defined as follows. Let $\{|j\rangle : j = 0, \ldots, d-1\}$ be an orthonormal basis, and define $\omega = e^{2\pi i/d}$. Then the operator

$$X|j\rangle = |j+1\rangle, \tag{2.40}$$

where addition is interpreted modulo d, effects a cyclic shift of the basis vectors. The Fourier transform of the X operator is

$$Z|j\rangle = \omega^j |j\rangle, \tag{2.41}$$

and together these operators satisfy the Weyl commutation relation

$$ZX = \omega XZ. \tag{2.42}$$

The Weyl–Heisenberg displacement operators are

$$D_{k,l} := (-e^{\pi i/d})^{kl} X^k Z^l, \tag{2.43}$$

and together they satisfy the conditions

$$D_{k,l}^\dagger = D_{-k,-l}, \qquad D_{k,l} D_{m,n} = (-e^{\pi i/d})^{lm-kn} D_{k+m,l+n}. \tag{2.44}$$

Each $D_{k,l}$ is unitary and a d^{th} root of the identity. The Weyl–Heisenberg group is the set of all operators $(-e^{\pi i/d})^m D_{k,l}$ for arbitrary integers m, and it is projectively equivalent to $\mathbb{Z}_d \times \mathbb{Z}_d$. Then, for any density matrix ρ such that

$$\mathrm{tr}\Big(D_{k,l}^{\dagger}\rho\Big) \neq 0, \quad \forall (k,l) \in \mathbb{Z}_d \times \mathbb{Z}_d \; , \tag{2.45}$$

the set

$$E_{k,l} := \frac{1}{d} D_{k,l}\rho D_{k,l}^{\dagger} \tag{2.46}$$

forms a WH MIC.

2.3.4 Equiangular MICs

An *equiangular*[2] MIC is one for which the Gram matrix takes the form

$$[G]_{ij} = \alpha\delta_{ij} + \zeta \; . \tag{2.47}$$

Equiangular MICs are unbiased (see Corollary 3 in [11]) and, because $\sum_{ij}[G]_{ij} = d$, it is easy to see that $\alpha = 1/d - d^2\zeta$ and that

$$\frac{1}{d^2(d+1)} \leq \zeta < \frac{1}{d^3} \; . \tag{2.48}$$

SICs are rank-1 equiangular MICs for which ζ achieves the minimum allowed value. The upper bound ζ value is approached by MICs arbitrarily close to $E_i = \frac{1}{d^2} I$ for all i.

[2] Appleby and Graydon introduced the term *SIM* for an equiangular MIC of arbitrary rank; a rank-1 SIM is a SIC [47, 48].

Armed with a SIC in a given dimension, one can construct an equiangular MIC for any allowed ζ value by mixing in some of the identity to each element:

$$E_i = \frac{\beta}{d}\Pi_i + \frac{1-\beta}{d^2}I \ , \quad \frac{-1}{d-1} \leq (\beta \neq 0) \leq 1 \ . \tag{2.49}$$

Even if a SIC is not known, it is generally much easier to construct equiangular MICs when the elements are not required to be rank-1. One way to do this which always works for any $\beta \leq \frac{1}{d+1}$ is by replacing the SIC projector in equation (2.49) with a quasi-SIC.[3] Depending on the quasi-SIC, higher values of β may also work.

Another construction in odd dimensions are the *Appleby MICs* [50]. The Appleby MICs are WH covariant and constructed as follows. Let the operator B be constructed as

$$B := \frac{1}{\sqrt{d+1}} \sum_{\{k,l\}\neq\{0,0\}} D_{k,l}, \tag{2.50}$$

and define $B_{k,l}$ to be its conjugate under a Weyl–Heisenberg displacement operator:

$$B_{k,l} := D_{k,l} B D_{k,l}^{\dagger}. \tag{2.51}$$

The elements of the Appleby MIC have rank $(d+1)/2$, and are defined by

$$E_{k,l} := \frac{1}{d^2}\left(I + \frac{1}{\sqrt{d+1}}B_{k,l}\right). \tag{2.52}$$

For any quantum state ρ, the quantities

$$W_{k,l} := (d+1)\mathrm{tr}\left(E_{k,l}\rho\right) - \frac{1}{d} \tag{2.53}$$

[3]See Appendix A of [49] for the definition and a construction of a quasi-SIC.

are *quasiprobabilities*: They can be negative, but the sum over all of them is unity. The quasiprobability function $\{W_{k,l}\}$ is known as the *Wigner function* of the quantum state ρ. This is an example of a relation we will study much more generally in a companion paper [51].

2.3.5 Tensorhedron MICs

So far, we have not imposed any additional structure upon our Hilbert space. However, in practical applications, one might have additional structure in mind, such as a preferred factorization into a tensor product of smaller Hilbert spaces. For example, a register in a quantum computer might be a set of N physically separate qubits, yielding a joint Hilbert space of dimension $d = 2^N$. In such a case, a natural course of action is to construct a MIC for the joint system by taking the tensor product of multiple copies of a MIC defined on the component system:

$$E_{j_1,j_2,\ldots,j_N} := E_{j_1} \otimes E_{j_2} \otimes \cdots \otimes E_{j_N}. \tag{2.54}$$

Since a collection of N qubits is a natural type of system to consider for quantum computation, we define the N-qubit *tensorhedron MIC* to be the tensor product of N individual qubit SICs.

Theorem 6. *The Gram matrix of an N-qubit tensorhedron MIC is the tensor product of N copies of the Gram matrix for the qubit SIC out of which the tensorhedron is constructed.*

Proof. Consider the two-qubit tensorhedron MIC, whose elements are given by

$$E_{d(j-1)+j'} := \frac{1}{4}\Pi_j \otimes \Pi_{j'}, \tag{2.55}$$

with $\{\Pi_j\}$ being a qubit SIC. The Gram matrix for the tensorhedron MIC has entries

$$[G]_{d(j-1)+j',d(k-1)+k'} = \frac{1}{16}\mathrm{tr}\,[(\Pi_j \otimes \Pi_{j'})(\Pi_k \otimes \Pi_{k'})]. \tag{2.56}$$

We can group together the projectors that act on the same subspace:

$$[G]_{d(j-1)+j',d(k-1)+k'} = \frac{1}{16}\mathrm{tr}\,(\Pi_j\Pi_k \otimes \Pi_{j'}\Pi_{k'}). \tag{2.57}$$

Now, we distribute the trace over the tensor product, obtaining

$$[G]_{d(j-1)+j',d(k-1)+k'} = \frac{1}{16}\frac{2\delta_{jk}+1}{3}\frac{2\delta_{j'k'}+1}{3} = [G_{\mathrm{SIC}}]_{jk}[G_{\mathrm{SIC}}]_{j'k'}, \tag{2.58}$$

which is just the definition of the tensor product:

$$G = G_{\mathrm{SIC}} \otimes G_{\mathrm{SIC}}. \tag{2.59}$$

This extends in the same fashion to more qubits. □

Corollary. *The spectrum of the Gram matrix for an N-qubit tensorhedron MIC contains only the values*

$$\lambda = \frac{1}{2^N}\frac{1}{3^m},\ m = 0, \ldots, N. \tag{2.60}$$

Proof. This follows readily from the linear-algebra fact that the spectrum of a tensor product is the set of products $\{\lambda_i\mu_j\}$, where $\{\lambda_i\}$ and $\{\mu_j\}$ are the spectra of the factors. □

We can also deduce properties of MICs made by taking tensor products of MICs that have orthogonal elements. Let $\{E_j\}$ be a d-dimensional MIC with Gram matrix G, and suppose that exactly N elements of G are equal to zero. The tensor products $\{E_j \otimes E_{j'}\}$ construct a d^2-dimensional MIC, the entries in whose Gram matrix have the form $[G]_{jk}[G]_{j'k'}$, as above. This product will equal zero when either factor does, meaning that the Gram matrix of the tensor-product MIC will contain $2d^4N - N^2$ zero-valued entries. It seems plausible that in prime dimensions, where tensor-product MICs cannot exist, the possible number of zeros is more tightly bounded, but this remains unexplored territory.

2.4 SICs are Minimally Nonclassical Reference Measurements

What might it mean for a MIC to be the best among all MICs? Naturally, it depends on what qualities are valued in light of which one MIC may be superior to another. As mentioned in the introduction, for a large number of metrics, SICs are optimal. The authors of this paper particularly value the capacity of MICs to index probabilistic representations of the Born Rule. For this use, the best MIC is the one which provides the most useful probabilistic representation, adopting some quantitative ideal that a representation should approach. One codification of such an ideal is as follows. In essence, we want to find a MIC that furnishes a probabilistic representation of quantum theory which looks as close to classical probability as is mathematically possible. The residuum that remains — the unavoidable discrepancy that even the most clever choice of MIC cannot eliminate — is a signal of what is truly *quantum* about quantum mechanics.

In a recent paper it was shown that SICs are strongly optimal for this project [12]. To see why, consider the following scenario. An agent has a physical system of interest, and she plans to carry out either one of two different, mutually exclusive procedures on it. In the first procedure, she will drop the system directly into a measuring apparatus and thereby obtain an outcome. In the second procedure, she will cascade her measurements, sending the system through a reference measurement and then, in the next stage, feeding it into the device from the first procedure. Probability theory unadorned by physical assumptions provides no constraints binding her expectations for these two different courses of action. Let P denote her probability assignments for the consequences of following the two-step procedure and Q those for the single-step procedure. Then, writing $\{H_i\}$ for the possible outcomes of the reference measurement and $\{D_j\}$ for those of the other,

$$P(D_j) = \sum_i P(H_i)P(D_j|H_i). \tag{2.61}$$

This equation is a consequence of Dutch-book coherence [7, 45] known as the Law of Total Probability (LTP). But the claim that

$$Q(D_j) = P(D_j) \tag{2.62}$$

is an assertion of *physics,* not entailed by the rules of probability theory alone. This assertion codifies in probabilistic language the classical ideal that a reference measurement simply reads off the system's "physical condition" or "ontic state".

We know this classical ideal is not met in quantum theory, that is, $Q(D_j) \neq P(D_j)$. Instead, as detailed in reference [12], $Q(D_j)$ is related to $P(H_i)$ and $P(D_j|H_i)$ in a different way. To write the necessary equations compactly, we introduce a vector notation where the LTP takes the form

$$P(D) = P(D|H)P(H)\ . \tag{2.63}$$

To set up the quantum version of the above scenario, let $\{H_i\}$ be a MIC and $\{D_j\}$ be an arbitrary POVM. Furthermore, let $\{\sigma_i\}$ denote a set of post-measurement states for the reference measurement; that is, if the agent experiences outcome H_i, her new state for the system will be σ_i. In this notation, the Born Rule becomes

$$Q(D) = P(D|H)\Phi P(H)\ ,\ \text{with}\ [\Phi^{-1}]_{ij} := \operatorname{tr} H_i\sigma_j\ . \tag{2.64}$$

The matrix Φ depends upon the MIC and the post-measurement states, but it is always a column quasistochastic matrix, meaning its columns sum to one but may contain negative elements [12]. In fact, Φ *must* contain negative entries; this follows from basic structural properties of quantum theory [52]. Now, the classical intution we mentioned above would be expressed by $\Phi = I$. However, no choice of MIC and set of post-measurement states can achieve this. The MICs and post-measurement sets which give a Φ matrix closest to the identity therefore supply the ideal representation we seek.

Theorem 1 in reference [12] proves that the distance between Φ and the identity with respect to any unitarily invariant norm is minimized when both the MIC and the post-measurement states are proportional to a SIC. Unitarily invariant norms include the Frobenius norm, the trace norm, the operator norm, and all the other Schatten p-norms, as well as the Ky Fan k-norms. Although this theorem was proven for foundational reasons, a special case of the result turns out to answer in the affirmative a conjecture regarding a practical matter of quantum computation [53, §VII.A].

What ended up being important for the optimality proof in [12] was that both the MIC and the post-measurement states be proportional to SICs, but not necessarily that they be proportional to the *same* SIC. Although the measures considered there were not sensitive to this distinction, the same SIC case has obvious conceptual and mathematical advantages. From a conceptual standpoint, when the post-measurement states are simply the projectors Π_i corresponding to the SIC outcome just obtained, our "throw away and reprepare" process is equivalent to Lüders rule updating, which there are independent reasons for preferring [21]. When the post-measurement states are the same SIC as the reference measurement, Φ takes the uniquely simple form

$$\Phi_{\mathrm{SIC}} = (d+1)I - \frac{1}{d}J \ , \tag{2.65}$$

where J is the Hadamard identity, that is, the matrix of all 1s. Inserted into (5.8) and written in index form, this produces the expression

$$Q(D_j) = \sum_i \left[(d+1)P(H_i) - \frac{1}{d}\right] P(D_j|H_i) \ , \tag{2.66}$$

having the advantage that for each conditional probability given an outcome H_i, only the i^{th} reference probability figures into that term in the sum. This is not so for two arbitrarily chosen SICs, and, as such, that case would result in a messier probabilistic representation.

This path is not the only one from which to arrive at the conclusion that SICs furnish a minimally nonclassical reference measurement. Recall the close association of classicality and orthogonality noted in section 2.2. From this standpoint, one might claim that most "classical" or least "quantum" reference measurement is one that is closest to an orthogonal measurement.

While we know from Corollary 2.2 that a MIC cannot be an orthogonal basis, how close can one get? One way to quantify this closeness is via an operator distance between the Gramians of an orthogonal basis and a MIC. From the proof of Corollary 2.2, we know that if a MIC could be orthogonal its Gram matrix would be $[G]_{ij} = e_i\delta_{ij}$. With no further restrictions, we can get arbitrarily close to this ideal, for instance, with a MIC constructed as follows. Consider a set of d^2 matrices $\{A_i\}$ where the first d of them are the eigenprojectors of a Hermitian matrix and the remaining $d^2 - d$ are the zero matrix.

Then, for an arbitrary[4] MIC $\{B_j\}$, we may form a new MIC, indexed by a real number $0 < t < 1$,

$$E_i^t := tA_i + (1-t)B_i \ . \tag{2.67}$$

One may see that the Gram matrix of $\{E_i^t\}$ approaches the orthogonal Gram matrix in the limit $t \to 1$.

But at such an extreme, the usefulness of a MIC is completely destroyed. In the above scenario when t is close to 1, the informational completeness is all but gone, as one has to reckon with vanishingly small probabilities when dealing with a MIC close to the limit point. Such a MIC fails miserably at being anything like a reasonable reference measurement. Although formally capable of being a reference measurement, a biased MIC deprives us of an even-handed treatment of indifference; the garbage state, which is poised in Hilbert space to capture pure state preparation indifference, would be represented by a non-flat probability distribution. Worse, for any sufficiently biased MIC, i.e., one with any weight less than $1/d^2$, the flat probability distribution is not reached by any density matrix. Consequently, what we're really after is an unbiased reference measurement which is as close to an orthogonal measurement as possible. With this additional constraint, the following theorem demonstrates that SICs are the optimal choice.

Theorem 7. *The closest an unbaised MIC can be to an orthogonal basis, as measured by the Frobenius distance between their Gramians, is when the MIC is a SIC.*[5]

[4] As long as a linear dependence does not develop.

[5] An earlier paper by one of us (CAF) and a collaborator [7] made the claim that the condition of being unbiased could be derived by minimizing the squared Frobenius distance; this is erroneous as the unequally weighted example with $\{E_i^t\}$ shows. For the purposes of that earlier paper, it is sufficient to impose by hand the requirement that the MIC be unbiased, since this is a naturally desirable property for a standard reference measurement. Having made this extra proviso, the conceptual conclusions of that work are unchanged.

Proof. We lower bound the square of the Frobenius distance:

$$\begin{aligned}\sum_{ij}\left(\frac{1}{d}\delta_{ij}-\operatorname{tr}E_iE_j\right)^2 &= \sum_i\left(\frac{1}{d}-\operatorname{tr}E_i^2\right)^2+\sum_{i\neq j}(\operatorname{tr}E_iE_j)^2\\ &\geq \frac{1}{d^2}\left(\sum_i\left(\frac{1}{d}-\operatorname{tr}E_i^2\right)\right)^2+\frac{1}{d^4-d^2}\left(\sum_{i\neq j}\operatorname{tr}E_iE_j\right)^2\\ &= \frac{1}{d^2}\left(d-\sum_i\operatorname{tr}E_i^2\right)^2+\frac{1}{d^4-d^2}\left(d-\sum_i\operatorname{tr}E_i^2\right)^2\\ &= \frac{1}{d^2-1}\left(d-\sum_i\operatorname{tr}E_i^2\right)^2\geq\frac{(d-1)^2}{d^2-1}=\frac{d-1}{d+1}\,.\end{aligned}\tag{2.68}$$

The first inequality follows from two invocations of the Cauchy–Schwarz inequality and achieves equality iff $\operatorname{tr}E_i^2$ and $\operatorname{tr}E_iE_j$, for $i\neq j$, are constants, that is, iff the MIC is an equiangular MIC. The third line is easy to derive from the fact that for any MIC, $\sum_{ij}[G]_{ij}=d$. The final inequality comes from noting that $\sum_i\operatorname{tr}E_i^2=\frac{1}{d^2}\sum_i\operatorname{tr}\rho_i^2\leq 1$ with equality iff the MIC is rank-1. Thus the lower bound is saturated iff the equal weight MIC is rank-1 and equiangular, that is, iff it is a SIC. □

Theorem 7 concerned the Gramian of a MIC. We can, in fact, show a stronger result on the inverse of the Gram matrix.

Theorem 8. *Let G be the Gram matrix of an unbiased MIC, and let $\|\cdot\|$ be any unitarily invariant norm (i.e., any norm where $\|A\|=\|UAV\|$ for arbitrary unitaries U and V). Then*

$$\left\|I-\frac{1}{d}G^{-1}\right\|\geq\left\|I-\frac{1}{d}G_{\mathrm{SIC}}^{-1}\right\|\,,\tag{2.69}$$

with equality if and only if the MIC is a SIC.

Proof. This is a special case of Theorem 1 in [12]. □

As with the theorems we proved above about MICs in general, this mathematical result has physical meaning. Classically speaking, the "ideal of the detached observer" (as Pauli phrased it [13]) is a measurement that reads off the system's point in phase space, call it λ_i, without disturbance. A state of maximal certainty is one where an agent is absolutely certain which λ_i exists. An agent having maximal certainty about each of a pair of identically prepared systems implies that she expects to obtain the same outcome for a reference measurement on each system. In other words, her "collision probability" is unity:

$$\sum_i p(\lambda_i)^2 = 1. \tag{2.70}$$

There is also a quantum condition on states of maximal certainty. As before, we can approach the question, "What is the unavoidable residuum that separates quantum from classical?" by finding the form of this quantum condition that brings it as close as possible to the classical version.

Lemma 1. *Given a MIC $\{E_i\}$ with Gramian G, a quantum state is pure if and only if its probabilistic representation satisfies*

$$\sum_{ij} p(E_i)p(E_j)[G^{-1}]_{ij} = 1. \tag{2.71}$$

Proof. Let $\{E_i\}$ be a MIC. The expansion of any quantum state ρ in the dual basis is

$$\rho = \sum_i (\mathrm{tr}\, E_i \rho) \widetilde{E}_i \ . \tag{2.72}$$

By the Born Rule, the coefficients are probabilities:

$$\rho = \sum_i p(E_i) \widetilde{E}_i \ . \tag{2.73}$$

Now, recall that while $\operatorname{tr}\rho = 1$ holds for any quantum state, $\operatorname{tr}\rho^2 = 1$ holds if and only if that operator is a pure state, i.e., a rank-1 projector. These operators are the extreme points of quantum state space; all other quantum states are convex combinations of them. In terms of the MIC's dual basis, the pure-state condition is

$$\sum_{ij} p(E_i)p(E_j)\operatorname{tr}\widetilde{E}_i\widetilde{E}_j = 1\ , \tag{2.74}$$

and so, because the Gramian of the dual basis is the inverse of the MIC Gram matrix,

$$\sum_{ij} p(E_i)p(E_j)[G^{-1}]_{ij} = 1\ , \tag{2.75}$$

as desired. □

Equation (2.75) closely resembles the collision probability, (2.70). If G^{-1} were the identity, they would be identical. On the face of it, it looks as though we should see how close G^{-1} can get to the identity. One minor wrinkle is that we should actually compare G^{-1} with dI instead of just with I, because an unbiased, orthogonal MIC (if one could exist) would have the Gram matrix $\frac{1}{d}I$. So, how close can we bring G^{-1} to dI, by choosing an appropriate unbiased MIC? We know the answer to this from Theorem 8: The best choice is a SIC.

2.5 Computational Overview of MIC Gramians

In order to explore the realm of MICs more broadly, and to connect them with other areas of mathematical interest, it is worthwhile to generate MICs *randomly* and study the typical properties which result. In this section we focus on the Gram matrix spectra of four MIC varieties whose constructions are described in section 2.3. These types are:

1. Generic MICs: a MIC generated from an arbitrary positive semidefinite basis
2. Generic Rank-1 MICs: a MIC generated from an arbitrary rank-1 positive semidefinite basis
3. WH MICs: a MIC obtained from the WH orbit of an arbitrary density matrix
4. Rank-1 WH MICs: a MIC obtained from the WH orbit of an arbitrary pure state density matrix.

In Hilbert space dimensions 2 through 5 we generated 10^5 MICs with the following methodologies. We constructed the generic MICs as in section 2.3.2 and the WH MICs as in section 2.3.3. Each generic MIC was obtained from a basis of positive semidefinite operators and each WH MIC was obtained from the orbit of an initial density matrix. In the generic rank-1 case, the pure states defining the basis of projectors were sampled uniformly from the Haar measure. Likewise, in the rank-1 WH case, the initial vector was also sampled uniformly from the Haar measure. The positive semidefinite bases for the arbitrary-rank generic MICs and the initial states for the arbitrary-rank WH MICs were constructed as follows. First, Hermitian matrices M were sampled from the Gaussian Unitary distribution, and, for each of these, the positive semidefinite matrix $M^{\dagger}M$ was formed. d^2 of these sufficed to form a positive semidefinite basis without loss of generality and a trace-normalized instance served as the initial state for the WH MICs. For each MIC, we constructed its Gram matrix and computed the eigenvalues. Figures 2.1, 2.2, 2.3, and 2.4 are histograms of the eigenvalue distributions for dimensions 2, 3, 4, and 5, respectively.

We note some expected and unexpected features of these distributions. In accordance with Theorem 1, both group covariant types, being unbiased, always have the maximal eigenvalue $1/d$, while this is the lower bound for the maximal eigenvalue for the other two types. Particularly in the unbiased cases, because the eigenvalues must sum to 1, not all of them can be too large, so it is perhaps not surprising that there are few eigenvalues approaching $1/d$ and that all families show exponential decay until that value. However, the spectra of rank-1 MICs, especially in dimensions 2 and 3 (Figures 2.1 and 2.2), display a richness of features for which we have no explanation.

Most surprising of all is the small eigenvalue plateau in Figure 2.2 for the $d = 3$ rank-1 WH MICs. Further scrutiny has revealed that the plateau ends precisely at $1/12$, the average value for the non-maximal eigenvalues of an unbiased $d = 3$ MIC Gram matrix. The Gram matrix for a $d = 3$ SIC has the spectrum

$$\left(\frac{1}{3}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}\right) , \tag{2.76}$$

which has the maximal amount of degeneracy allowed. As dimension 3 is also exceptional in the study of SICs, we conjecture that the two are related.

Conjecture 4. The plateau in the eigenvalue distribution for $d = 3$, seen in Figure 2.2, is related to the existence of a continuous family of unitarily inequivalent SICs in that dimension [54, 55].

2.6 Conclusions

We have argued that informational completeness provides the right perspective from which to compare the quantum and the classical. The structure of Minimal Informationally Complete quantum measurements and especially how and to what degree this structure requires the abandonment of classical intuitions therefore deserves explicit study. We have surveyed the domain of MICs and derived some initial results regarding their departure from such classical intuitions as orthogonality, repeatability, and the possibility of certainty. Central to understanding MICs are their Gram matrices; it is through properties of these matrices that we were able to derive many of our results. We have only just scratched the surface of this topic, as our conjectures and unexplained numerical features of Gram matrix spectra can attest. In a sequel, we will explore another application of Gram matrices. They hold a central role in the construction of *Wigner functions* from MICs [56–58], and Wigner functions are a topic pertinent to quantum computation [59–63].

Many properties of MIC Gram matrices remain unknown. Numerical investigations have, in some cases, outstripped the proving of theorems, resulting in the conjectures we have enumerated. Another avenue for potential future exploration is the application of Shannon theory to MICs. Importing the notions of information theory into quantum mechanics has proved quite useful over the years at illuminating strange or surprising features of the physics [64–66]. One promising avenue of inquiry is studying the probabilistic representations of quantum states using entropic measures. In the case of SICs, this has already yielded intriguing connections among information theory, group theory and geometry [56, 67–70]. The analogous questions for other classes of MICs remain open for investigation.

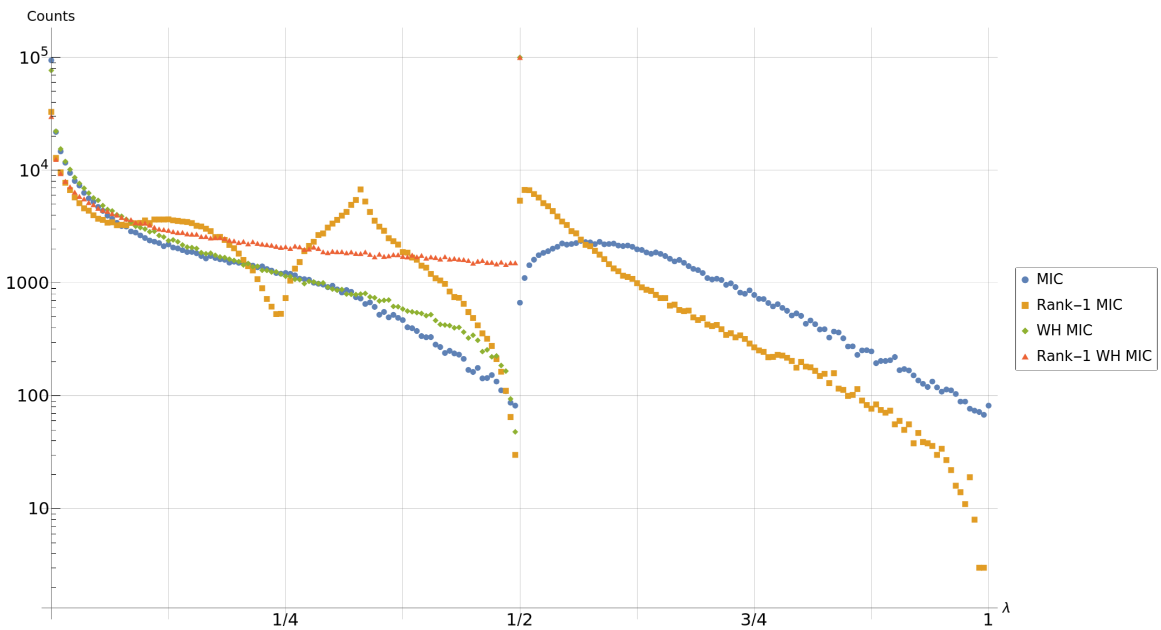

Figure 2.1: $d = 2$ random MIC Gram matrix spectra, $N = 10^5$, bin size $1/200$.

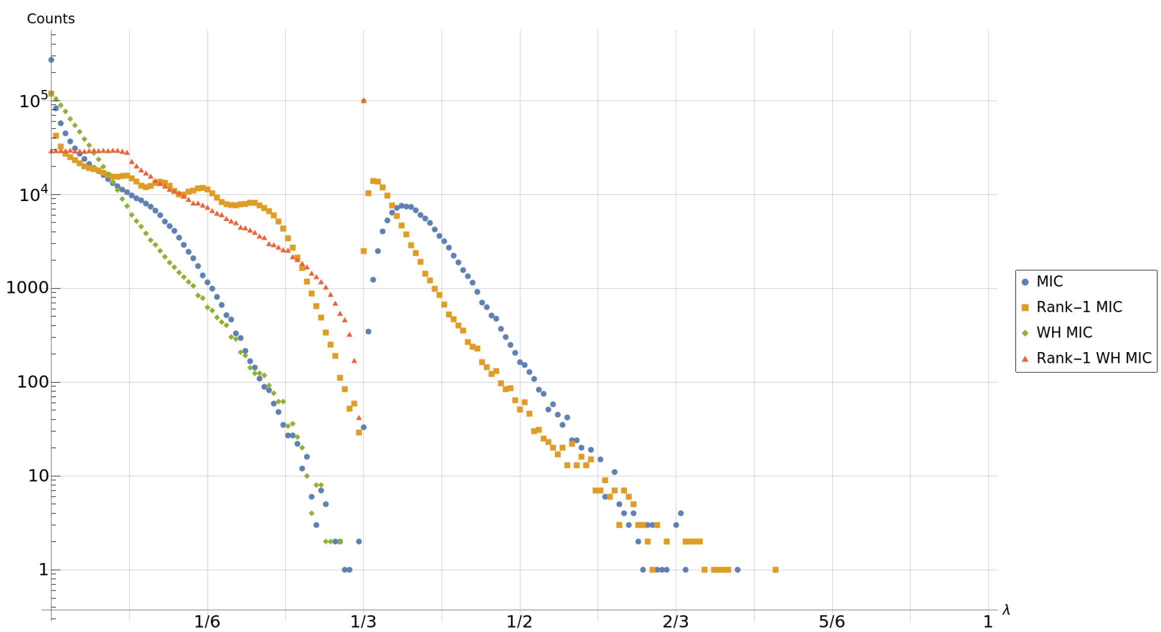

Figure 2.2: $d = 3$ random MIC Gram matrix spectra, $N = 10^5$, bin size $1/198$.

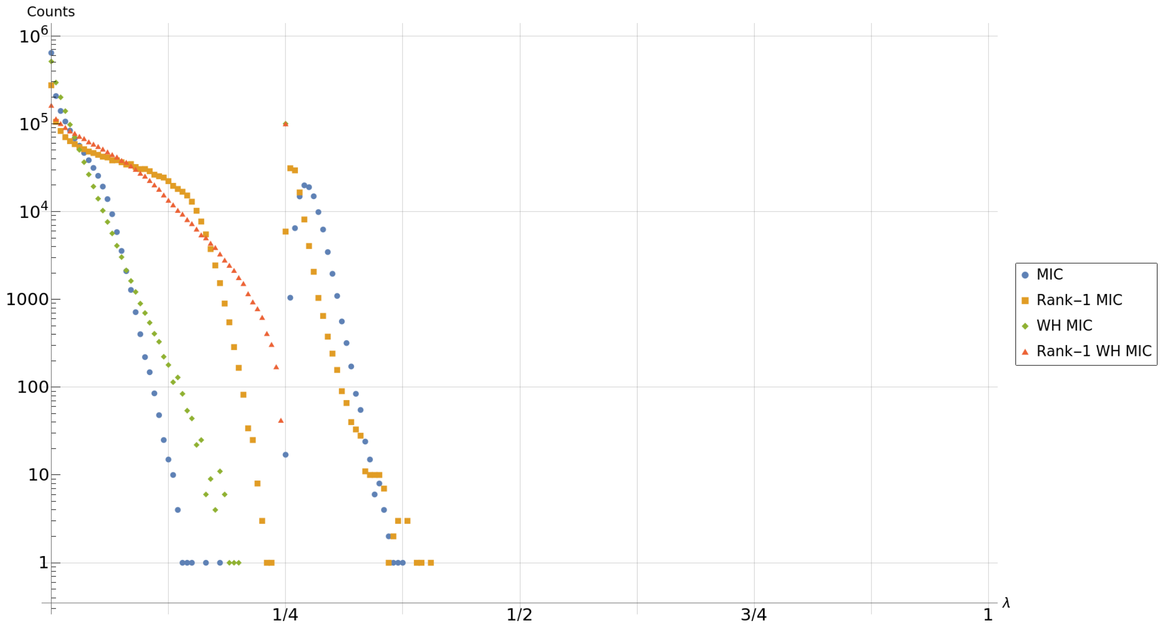

Figure 2.3: $d = 4$ random MIC Gram matrix spectra, $N = 10^5$, bin size $1/200$.

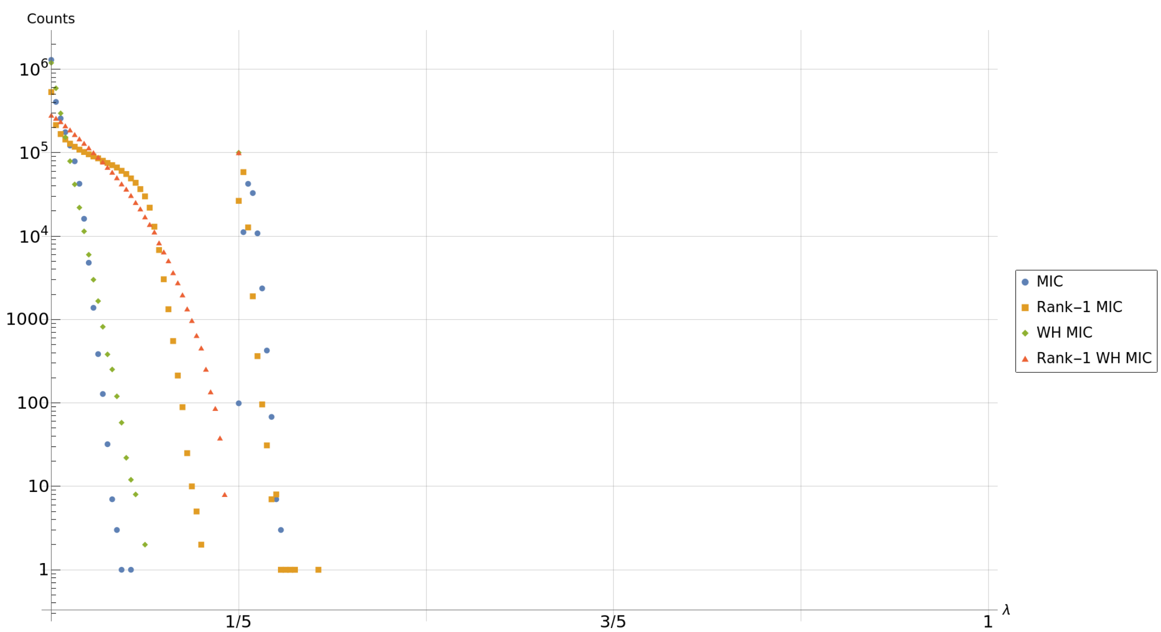

Figure 2.4: $d = 5$ random MIC Gram matrix spectra, $N = 10^5$, bin size $1/200$.

LIST OF REFERENCES

[1] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, tenth anniversary ed., 2010.

[2] G. Zauner, *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*. PhD thesis, University of Vienna, 1999. Published in English translation: G. Zauner, "Quantum designs: foundations of a noncommutative design theory," *International Journal of Quantum Information*, vol. 9, pp. 445–508, 2011.

[3] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, “Symmetric informationally complete quantum measurements,” *Journal of Mathematical Physics*, vol. 45, pp. 2171–2180, 2004.

[4] A. J. Scott and M. Grassl, “Symmetric informationally complete positive-operator-valued measures: A new computer study,” *Journal of Mathematical Physics*, vol. 51, p. 042203, 2010.

[5] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, “The SIC question: History and state of play,” *Axioms*, vol. 6, p. 21, 2017.

[6] C. A. Fuchs, N. D. Mermin, and R. Schack, “An introduction to QBism with an application to the locality of quantum mechanics,” *American Journal of Physics*, vol. 82, no. 8, pp. 749–754, 2014.

[7] C. A. Fuchs and R. Schack, “Quantum-Bayesian coherence,” *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[8] R. Healey, “Quantum-Bayesian and Pragmatist Views of Quantum Theory,” in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University, spring 2017 ed., 2017.

[9] C. A. Fuchs and B. C. Stacey, “QBism: Quantum Theory as a Hero’s Handbook,” in *Proceedings of the International School of Physics "Enrico Fermi": Course 197, Foundations of Quantum Theory*, pp. 133–202, 2019.

[10] D. Appleby, H. Dang, and C. Fuchs, “Symmetric Informationally-Complete Quantum States as Analogues to Orthonormal Bases and Minimum-Uncertainty States,” *Entropy*, vol. 16, no. 3, pp. 1484–1492, 2014.

[11] D. M. Appleby, C. A. Fuchs, and H. Zhu, “Group theoretic, Lie algebraic and Jordan algebraic formulations of the SIC existence problem,” *Quantum Information & Computation*, vol. 15, pp. 61–94, 2015.

[12] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems,” *Physical Review Research*, vol. 2, p. 013074, 2020.

[13] C. A. Fuchs, “Notwithstanding Bohr, the Reasons for QBism,” *Mind and Matter*, vol. 15, no. 2, 2017.

[14] B. C. Stacey, “Is the SIC Outcome There When Nobody Looks?,” *arXiv:1807.07194*, 2018.

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

[16] C. A. Fuchs, *Coming of age with quantum information: notes on a Paulian idea: selected correspondence, 1995-2001*. Cambridge ; New York: Cambridge University Press, 2011.

[17] T. Leinster, “In Praise of the Gershgorin Disc Theorem.” The n-Category Café, 2016.

[18] B. C. Stacey, “Misreading EPR: Variations on an Incorrect Theme,” *arXiv:1809.01751*, 2018.

[19] R. W. Spekkens, “Evidence for the epistemic view of quantum states: A toy theory,” *Physical Review A*, vol. 75, no. 3, p. 032110, 2007.

[20] P. Busch and P. Lahti, “Lüders Rule,” in *Compendium of Quantum Physics* (D. Greenberger, K. Hentschel, and F. Weinert, eds.), pp. 356–358, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[21] H. Barnum, “Information-disturbance tradeoff in quantum measurement on the uniform ensemble and on the mutually unbiased bases,” *arXiv:quant-ph/0205155*, 2002.

[22] A. Cabello, “Quantum correlations from simple assumptions,” *Physical Review A*, vol. 100, no. 3, p. 032120, 2019.

[23] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*. Cambridge University Press, 2008.

[24] S. Waldron, *An Introduction to Finite Tight Frames*. Springer, 2018.

[25] G. M. D’Ariano, P. L. Presti, and P. Perinotti, “Classical randomness in quantum measurements,” *Journal of Physics A: Mathematical and General*, vol. 38, no. 26, pp. 5979–5991, 2005.

[26] C. A. Fuchs and M. Sasaki, “Squeezing Quantum Information through a Classical Channel: Measuring the "Quantumness" of a Set of Quantum States,” *Quantum Information & Computation*, vol. 3, pp. 377–404, 2003.

[27] A. J. Scott, “Tight informationally complete quantum measurements,” *Journal of Physics A: Mathematical and General*, vol. 39, no. 43, pp. 13507–13530, 2006.

[28] P. Delsarte, J. M. Goethels, and J. J. Seidel, “Bounds for systems of lines and Jacobi polynomials,” *Philips Research Reports*, vol. 30, pp. 91–105, 1975.

[29] S. G. Hoggar, “Two quaternionic 4-polytopes,” in *The Geometric Vein: The Coxeter Festschrift* (C. Davis, B. Grünbaum, and F. A. Sherk, eds.), Springer, 1981.

[30] H. S. M. Coxeter, *Regular Complex Polytopes*. Cambridge University Press, second ed., 1991.

[31] S. G. Hoggar, “64 lines from a quaternionic polytope,” *Geometriae Dedicata*, vol. 69, pp. 287–289, 1998.

[32] A. Koldobsky and H. König, “Aspects of the isometric theory of Banach spaces,” in *Handbook of the geometry of Banach spaces* (W. Johnson and J. Lindenstrauss, eds.), vol. 1, pp. 899–939, Amsterdam: North-Holland, 2001.

[33] M. Appleby and I. Bengtsson, “Simplified exact SICS,” *Journal of Mathematical Physics*, vol. 60, no. 6, p. 062203, 2019.

[34] A. J. Scott, “SICs: Extending the list of solutions,” *arXiv:1703.03993*, 2017.

[35] M. Grassl and A. J. Scott, “Fibonacci–Lucas SIC-POVMs,” *Journal of Mathematical Physics*, vol. 58, p. 122201, 2017.

[36] M. Appleby, H. Yadsan-Appleby, and G. Zauner, “Galois automorphisms of a symmetric measurement,” *Quantum Information & Computation*, vol. 13, pp. 672–720, 2013.

[37] M. Appleby, S. Flammia, G. McConnell, and J. Yard, “Generating ray class fields of real quadratic fields via complex equiangular lines,” *Acta Arithmetica*, vol. 192, no. 3, pp. 211–233, 2020.

[38] I. Bengtsson, “The number behind the simplest SIC-POVM,” *Foundations of Physics*, vol. 47, pp. 1031–41, 2017.

[39] M. Appleby, S. Flammia, G. McConnell, and J. Yard, “SICs and algebraic number theory,” *Foundations of Physics*, vol. 47, pp. 1042–59, 2017.

[40] G. S. Kopp, "SIC-POVMs and the Stark conjectures," *International Mathematics Research Notices*, p. rnz153, 2019.

[41] M. Appleby, T.-Y. Chien, S. Flammia, and S. Waldron, "Constructing exact symmetric informationally complete measurements from numerical solutions," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 16, p. 165302, 2018.

[42] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, second ed., 2017.

[43] C. M. Caves, C. A. Fuchs, and R. Schack, "Unknown quantum states: The quantum de Finetti representation," *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4537–4559, 2002.

[44] B. C. Stacey, *Multiscale Structure in Eco-Evolutionary Dynamics*. PhD thesis, Brandeis University, 2015.

[45] P. Diaconis and B. Skyrms, *Ten great ideas about chance*. Princeton: Princeton University Press, 2018.

[46] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, "Informationally complete measurements and group representation," *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 6, no. 6, pp. S487–S491, 2004.

[47] M. A. Graydon and D. M. Appleby, "Quantum conical designs," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 8, p. 085301, 2016.

[48] M. A. Graydon and D. M. Appleby, "Entanglement and designs," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 33, p. 33LT02, 2016.

[49] J. B. DeBrota and B. C. Stacey, "Lüders channels and the existence of symmetric-informationally-complete measurements," *Physical Review A*, vol. 100, no. 6, p. 062327, 2019.

[50] D. M. Appleby, "Symmetric informationally complete measurements of arbitrary rank," *Optics and Spectroscopy*, vol. 103, no. 3, pp. 416–428, 2007.

[51] J. B. DeBrota and B. C. Stacey, "Discrete Wigner Functions from Informationally Complete Quantum Measurements," *arXiv:1912.07554*, 2020.

[52] C. Ferrie, "Quasi-probability representations of quantum theory with applications to quantum information science," *Reports on Progress in Physics*, vol. 74, no. 11, p. 116001, 2011.

[53] A. Veitia and S. J. van Enk, "Testing the context-independence of quantum gates," *arXiv:1810.05945*, 2018.

[54] G. N. M. Tabia and D. M. Appleby, “Exploring the geometry of qutrit state space using symmetric informationally complete probabilities,” *Physical Review A*, vol. 88, no. 1, p. 012131, 2013.

[55] L. P. Hughston and S. M. Salamon, “Surveying points in the complex projective plane,” *Advances in Mathematics*, vol. 286, pp. 1017–1052, 2016.

[56] B. Stacey, “SIC-POVMs and Compatibility among Quantum States,” *Mathematics*, vol. 4, no. 2, p. 36, 2016.

[57] H. Zhu, “Quasiprobability representations of quantum mechanics with minimal negativity,” *Physical Review Letters*, vol. 117, no. 12, p. 120404, 2016.

[58] J. B. DeBrota and C. A. Fuchs, “Negativity bounds for Weyl–Heisenberg quasiprobability representations,” *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

[59] W. K. Wootters, “A Wigner-function formulation of finite-state quantum mechanics,” *Annals of Physics*, vol. 176, no. 1, pp. 1–21, 1987.

[60] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, “Discrete phase space based on finite fields,” *Physical Review A*, vol. 70, no. 6, p. 062101, 2004.

[61] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, “Negative quasi-probability as a resource for quantum computation,” *New Journal of Physics*, vol. 14, no. 11, p. 113011, 2012.

[62] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, “The resource theory of stabilizer computation,” *New Journal of Physics*, vol. 16, p. 013009, 2014.

[63] M. Howard and E. T. Campbell, “Application of a resource theory for magic states to fault-tolerant quantum computing,” *Physical Review Letters*, vol. 118, no. 9, p. 090501, 2017.

[64] S. L. Braunstein and C. M. Caves, “Wringing out better Bell inequalities,” *Annals of Physics*, vol. 202, no. 1, pp. 22–56, 1990.

[65] C. A. Fuchs, “Just Two Nonorthogonal Quantum States,” *arXiv:quant-ph/9810032*, 1998.

[66] J. R. Buck, S. J. van Enk, and C. A. Fuchs, “Experimental proposal for achieving superadditive communication capacities with a binary quantum alphabet,” *Physical Review A*, vol. 61, no. 3, p. 032309, 2000.

[67] B. C. Stacey, “Geometric and Information-Theoretic Properties of the Hoggar Lines,” *arXiv:1609.03075*, 2016.

[68] W. Słomczyński and A. Szymusiak, “Highly symmetric POVMs and their informational power,” *Quantum Information Processing*, vol. 15, no. 1, pp. 565–606, 2016.

[69] A. Szymusiak, “Maximally informative ensembles for SIC-POVMs in dimension 3,” *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 44, p. 445301, 2014.

[70] A. Szymusiak and W. Słomczyński, “Informational power of the Hoggar symmetric informationally complete positive operator-valued measure,” *Physical Review A*, vol. 94, p. 012122, 2015.

CHAPTER 3

LTP ANALOGS

Quantum information theory represents a change of perspective. Rather than regarding quantum physics as a limitation on our abilities—the typical sentiment of older texts—we have learned that it can augment them. In frustrating some ambitions, it enables more subtle ones. Deviation from classicality is a *resource,* and the idea that this resource can be quantified as a modification of the classical probability calculus dates to the beginning of the field [1]. More recent inquiries have developed this notion precisely: The "negativity" in a Wigner-function representation of quantum states is now understood to be valuable in its own right [2–11]. But what does this line of thinking say about quantum mechanics itself? Can one, following the lead of Carnot, take what might seem a statement of "mere" engineering and find a physical principle? In this paper, we prove some strong results in this regard in the context of finite dimensional Hilbert spaces. In particular, we find the unique form of the quantum mechanical Born Rule that makes it resemble the classical Law of Total Probability (LTP) as closely as possible in at least two senses. Both come from a significant majorization result which may be of general interest for resource theory. This way of tackling the distinction between quantum and classical arises naturally in the quantum interpretive project of QBism [12, 13], where the Born Rule is seen as an empirically motivated constraint that one adds to probability theory when using it in the context of alternative (complementary) quantum experiments. We expect the techniques developed here to give an alternative way to explore the paradigm of negativity and to be of use for a range of practical problems.

The standard procedure in quantum theory for generating probabilities starts with an observer, or agent, assigning a quantum state ρ to a system. When the agent plans to measure the system, she represents the outcomes of her measurement with a positive operator-valued measure (POVM) $\{D_j\}$. Assigning ρ implies that she assigns the Born Rule probabilities $Q(D_j) = \mathrm{tr}\,\rho D_j$ for the outcomes of her measurement. In this way, any quantum state ρ may be regarded as a compilation of probability distributions for all possible measurements. However, one does not have to consider all possible measurements to completely specify ρ. In fact, there exist measurements which are informationally complete (IC) in the sense that ρ is uniquely specified by the agent's expectations for the outcomes of that single measurement [14]. With respect to an IC measurement, any quantum state, pure or mixed, is equivalent to a single probability distribution. In this paper, we consider minimal informationally-complete POVMs (MICs) for finite dimensional quantum systems. These sets of operators form bases for the vector space of Hermitian operators and lead to probability distributions with the fewest number of entries necessary for reconstructing the quantum state. MICs furnish a convenient way to bypass the language of quantum states, making quantum theory analogous to classical stochastic process theory, in which one puts probabilities in and gets probabilities out.

One can eliminate the need to use the operators ρ and D_j in the Born Rule by reexpressing it as a relation between an agent's expectations for different experiments. Suppose our agent has a preferred reference process consisting of a measurement to which she ascribes the MIC $\{H_i\}$, and, upon obtaining outcome i, the preparation of a state σ_i, drawn from a linearly independent set of post-measurement states $\{\sigma_i\}$. (See Fig. 3.1.) In her choice of this reference process, she requires linearly independent post-measurement states so that the inner products $\mathrm{tr}\, D_j\sigma_i$ will uniquely characterize the operators D_j. Let $P(H_i)$ be her probabilities for the measurement $\{H_i\}$ and $P(D_j|H_i)$ be her conditional probabilities for a subsequent measurement of $\{D_j\}$. What consistency requirement among $Q(D_j)$, $P(H_i)$, and $P(D_j|H_i)$ does quantum physics entail?

Using the fact that $\{\sigma_i\}$ is a basis, we may write

$$\rho = \sum_j \alpha_j \sigma_j \ , \tag{3.1}$$

for some set of real coefficients α_j. The probability of outcome H_i is then

$$P(H_i) = \sum_j \alpha_j \,\mathrm{tr}\, H_i \sigma_j = \sum_j \left[\Phi^{-1}\right]_{ij} \alpha_j \ , \tag{3.2}$$

where we have defined the matrix Φ via its inverse,

$$\left[\Phi^{-1}\right]_{ij} := \mathrm{tr}\, H_i \sigma_j = h_i \mathrm{tr}\, \rho_i \sigma_j \ , \tag{3.3}$$

for $\rho_i := H_i/h_i$ and $h_i := \mathrm{tr}\, H_i$. The invertibility of Φ is assured by the linear independence of the MIC and post-measurement sets. This implies that the coefficients of ρ in the σ_i basis may be written as an application of the Φ matrix on the vector of probabilities,

$$\rho = \sum_i \left[\sum_k [\Phi]_{ik} P(H_k)\right] \sigma_i \ . \tag{3.4}$$

The probability of D_j is given by another application of the Born Rule, which becomes

$$Q(D_j) = \sum_{i=1}^{d^2} \left[\sum_{k=1}^{d^2} [\Phi]_{ik} P(H_k)\right] P(D_j|H_i) \ , \tag{3.5}$$

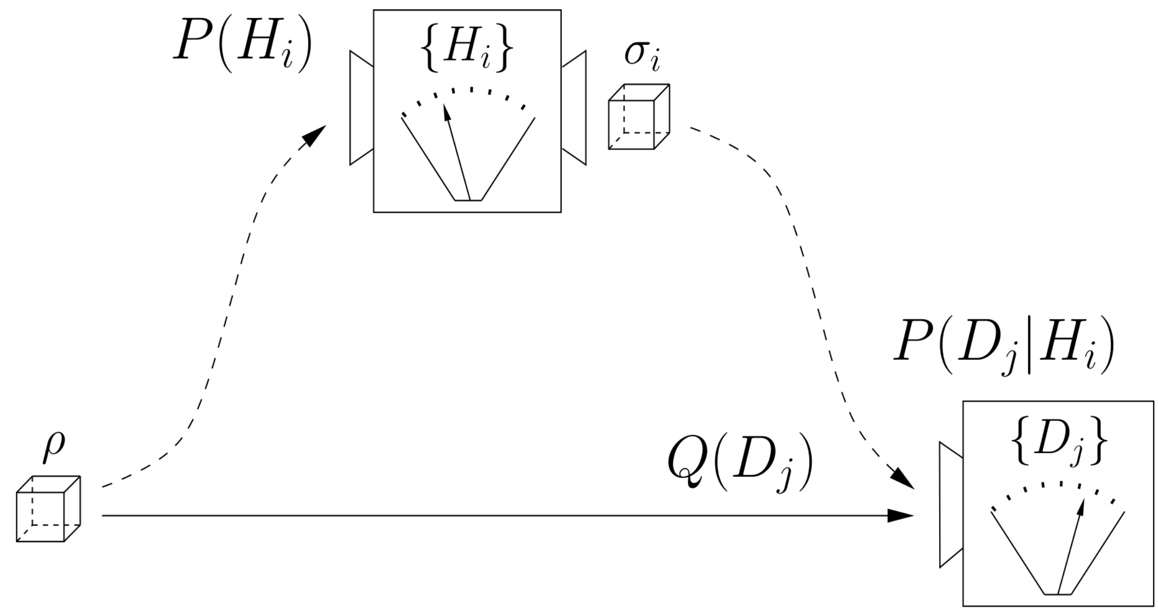

Figure 3.1: The solid and dashed lines represent two hypothetical procedures an agent contemplates for a system assigned state ρ. The solid line represents making a direct measurement of a POVM $\{D_j\}$. The dotted line represents making the MIC measurement $\{H_i\}$ first, preparing a post-measurement state σ_i, and then finally making the $\{D_j\}$ measurement. For the solid path, the agent assigns one set of probabilities $Q(D_j)$. For the dotted path, she assigns two sets of probabilities: $P(H_i)$ and $P(D_j|H_i)$. Unadorned by physical assumptions, probability theory does not suggest a relation between these paths. The Born Rule in the form of Eq. (5.8) is such a relation.

where $P(D_j|H_i) = \operatorname{tr} D_j\sigma_i$ is the probability for outcome D_j conditioned on obtaining H_i in the reference measurement. In more compact matrix notation, we can write

$$Q(D) = P(D|H)\,\Phi\,P(H)\ , \tag{3.6}$$

where $P(D|H)$ is a matrix of conditional probabilities.

A SIC [15–24] is a MIC for which all the H_i are rank-1 and

$$\operatorname{tr} H_iH_j = \frac{1}{d^2}\frac{d\delta_{ij}+1}{d+1}\ . \tag{3.7}$$

SICs have yet to be proven to exist in all finite dimensions d, but they are widely believed to [23] and have even been experimentally demonstrated in some low dimensions [25–28]. The *SIC projectors* associated with a SIC are the pure states $\rho_i = dH_i$. In dimension 2, a SIC can be represented as a regular tetrahedron inscribed in the Bloch sphere. (States defining a qubit SIC can be extracted from Feynman's 1987 essay "Negative probabilities" [29].) In higher dimensions, they are, of course, harder to visualize. When there is no chance of confusion, we will refer to the set of projectors as SICs as well. Prior work has given special attention to the reference procedure where the measurement and post-measurement states are the same SIC [12, 30, 31]. In this case we denote Φ by Φ_{SIC} and Eq. (3.5) takes the particularly simple form

$$Q(D_j) = \sum_{i=1}^{d^2} \left[(d+1)P(H_i) - \frac{1}{d}\right] P(D_j|H_i) \; . \tag{3.8}$$

Recall that the LTP expresses the simple consistency relation between the probabilities one assigns to the second of a sequence of measurements, the probabilities one assigns to the first, and the conditional probabilities for the second given the outcome of the first. Written in vector notation, this is

$$P(D) = P(D|H)P(H) \; . \tag{3.9}$$

We write $P(D)$ as opposed to $Q(D)$ to indicate that it is the probability vector for the *second* of two measurements. $Q(D)$, on the other hand, is the vector of probabilities associated with a single measurement. Aside from the presence of Φ matrix, Eq. (5.8) is functionally equivalent to the LTP.

Although $P(H)$, $P(D|H)$, and $Q(D)$ are probabilities, $\Phi P(H)$ often is not. One may see by summing both sides of Eq. (3.5) over j that the vector is normalized, but in general it may contain negative numbers and values greater than 1. Such a vector is known as a quasiprobability, and matrices like Φ—real-valued matrices with columns summing to 1—which take probabilities to quasiprobabilites are called column-quasistochastic matrices [32]. The subset of column-quasistochastic matrices with nonnegative entries are the column-stochastic matrices. The inverse of a column-stochastic matrix is generally a column-quasistochastic matrix; in our case, inspection of Eq. (3.3) reveals that Φ^{-1} is column-stochastic.

What would it mean if Φ could equal I? In this case we would have $Q(D) = P(D)$. Then, conceptually, it wouldn't matter if the intermediate measurement were performed or not. Put another way, we could behave as though measurements simply revealed a preexisting property of the system, as in classical physics where measurements provide information about a system's coordinates in phase space.

Some amount of what makes quantum theory nonclassical resides in the fact that Φ cannot equal I. How close, then, can we make Φ to I by wisely choosing our MIC and post-measurement states? It turns out that $\Phi_{\rm SIC}$ is closest to the identity with respect to the distance measure induced by any member of a large family of operator norms called unitarily invariant norms (see section 3.5 in [33]). A unitarily invariant norm is one such that $\|A\| = \|UAV\|$ for all unitary matrices U and V. These norms include the Schatten p-norms (among which are the trace norm, the Frobenius norm, and the operator norm when $p = 1, 2$, and ∞ respectively) and the Ky Fan k-norms. This result codifies the intuition that Eq. (3.8) represents the "simplest modification one can imagine to the LTP" [34, p. 1971].

To prove this, we will make use of the theory of majorization [33, 35]. Suppose x and y are vectors of N real numbers and that $x^{\downarrow}$ and $y^{\downarrow}$ are x and y sorted in nonincreasing order. Then we say that x weakly majorizes y from below, denoted $x \succ_w y$, if

$$\sum_{i=1}^{k} x_i^{\downarrow} \geq \sum_{i=1}^{k} y_i^{\downarrow} \ , \quad \text{for } k = 1, \ldots, N \ . \tag{3.10}$$

If the last inequality is an equality, we say x majorizes y, denoted $x \succ y$.

Another variant of majorization, called log majorization or multiplicative majorization, is also studied [35]. We say that x weakly log majorizes y from below, denoted $x \succ_{w \log} y$, if

$$\prod_{i=1}^{k} x_i^{\downarrow} \geq \prod_{i=1}^{k} y_i^{\downarrow} \ , \quad \text{for } k = 1, \ldots, N \ . \tag{3.11}$$

If the last inequality is an equality, we say x log majorizes y, denoted $x \succ_{\log} y$. Taking the log of both sides of Eq. (3.11) demonstrates that log majorization is majorization between the vectors after an element-wise application of the log map. Log majorization is strictly stronger than regular majorization; $x \succ_{w \log} y \implies x \succ_w y$, but the reverse implication is not true. Majorization is a partial order on vectors of real numbers sorted in nonincreasing order.

Throughout this paper we will make use of the standard inequalities between the arithmetic, geometric, and harmonic means for vectors of n positive numbers x_i:

$$\frac{1}{n} \sum_{i=1}^{n} x_i \geq \left(\prod_{i=1}^{n} x_i \right)^{1/n} \geq \left(\frac{1}{n} \sum_{i=1}^{n} \frac{1}{x_i} \right)^{-1} \ . \tag{3.12}$$

with equality in all cases if and only if $x_i = c$ for all i. We now turn to two lemmas.

Lemma 2. *Let* Φ_{p} *denote the column-quasistochastic matrix associated with a MIC and a proportional post-measurement set. Then* $\det \Phi_{\mathrm{p}} \geq \det \Phi_{\mathrm{SIC}}$ *with equality iff the MIC is a SIC.*

Proof. We may write $\Phi_{\mathrm{p}}^{-1} = GA^{-1}$ where $G_{ij} := \mathrm{tr}\, H_i H_j$ is the Gram matrix of the MIC elements and $A_{ij} := h_i \delta_{ij}$. Note that Φ_{p}^{-1} has real, positive eigenvalues because it has the same spectrum as the positive definite matrix $A^{-1/2} G A^{-1/2}$. Also note that

$$\sum_i \frac{1}{\lambda_i(\Phi_{\mathrm{p}})} = \mathrm{tr}\, \Phi_{\mathrm{p}}^{-1} = \sum_i h_i \mathrm{tr}\, \rho_i \sigma_i \leq \sum_i h_i = d\ . \tag{3.13}$$

One of the eigenvalues of Φ_{p}, which we denote $\lambda_{d^2}(\Phi_{\mathrm{p}})$, must equal 1 because an equal-entry row vector is always a left eigenvector with eigenvalue 1 of a matrix with columns summing to unity. Therefore, we may write

$$\sum_{i<d^2} \frac{1}{\lambda_i(\Phi_{\mathrm{p}})} \leq d - 1. \tag{3.14}$$

The reciprocal of this expression is proportional to the harmonic mean of the first $d^2 - 1$ eigenvalues of Φ_{p}. Thus, because the geometric mean is always greater than or equal to the harmonic mean,

$$\left(\prod_{i=1}^{d^2-1} \lambda_i(\Phi_{\mathrm{p}}) \right)^{\frac{1}{d^2-1}} \geq \left(\frac{1}{d^2-1} \sum_{i=1}^{d^2-1} \frac{1}{\lambda_i(\Phi_{\mathrm{p}})} \right)^{-1} \geq d + 1\ , \tag{3.15}$$

which, noting that $\lambda_{d^2}(\Phi_{\mathrm{p}}) = 1$, implies

$$\det \Phi_{\mathrm{p}} \geq (d+1)^{d^2-1} = \det \Phi_{\mathrm{SIC}}\ . \tag{3.16}$$

Equality is achieved in this iff all the $\lambda_i(\Phi_{\mathrm{p}})$ are equal, so Eq. (3.16) is saturated iff $\lambda(\Phi_{\mathrm{p}}) = \lambda(\Phi_{\mathrm{SIC}})$. We next show this implies that in fact the MIC is a SIC.

For any Φ_{p}^{-1}, we may write $\Phi_{\mathrm{p}}^{-1} = P^{-1} D P$ where the rows of P are the left-eigenvectors of Φ_{p}^{-1} and D is the diagonal matrix of eigenvalues of Φ_{p}^{-1}. Since Φ_{p}^{-1} is column-stochastic, the row vector $(1/d, \ldots, 1/d)$ is the (scaled) left-eigenvector of Φ_{p}^{-1} with eigenvalue 1, and so it is the first row of P when the eigenvalues are in descending order. Left-eigenvectors of a matrix are right-eigenvectors of the transpose of the matrix, so we have

$$\begin{aligned}(\Phi_{\mathrm{p}}^{-1})^T|v\rangle &= A^{-1}G|v\rangle = A^{-1}GA^{-1}A|v\rangle \\ &= A^{-1}\Phi_{\mathrm{p}}^{-1}A|v\rangle = \lambda|v\rangle\ , \\ \implies \Phi_{\mathrm{p}}^{-1}A|v\rangle &= \lambda A|v\rangle\ ,\end{aligned} \tag{3.17}$$

where $\langle v|$ is an arbitrary left-eigenvector of Φ_{p}^{-1}. Combined with our choice of scale for the first row of P, we conclude that the first column of P^{-1} is $(h_1, h_2, \ldots, h_{d^2})^T$.

Now suppose Φ_{p} is such that $\lambda(\Phi_{\mathrm{p}}) = \lambda(\Phi_{\mathrm{SIC}})$. Then $G = P^{-1}DPA$ where $[D]_{ij} = \frac{1}{d+1}(\delta_{ij} + d\delta_{i1}\delta_{j1})$, and

$$\begin{aligned}[G]_{ij} &= \sum_{klm}[P^{-1}]_{ik}[D]_{kl}[P]_{lm}[A]_{mj} \\ &= \sum_{klm}[P^{-1}]_{ik}\left[\frac{1}{d+1}(\delta_{kl} + d\delta_{k1}\delta_{l1})\right][P]_{lm}\delta_{mj}h_m \\ &= \frac{1}{d+1}\sum_{kl}[P^{-1}]_{ik}(\delta_{kl} + d\delta_{k1}\delta_{l1})[P]_{lj}h_j \\ &= \frac{1}{d+1}(h_j\delta_{ij} + dh_j[P^{-1}]_{i1}[P]_{1j}) \\ &= \frac{1}{d+1}(h_j\delta_{ij} + h_ih_j)\ .\end{aligned} \tag{3.18}$$

In the last step we used that $[P]_{1j} = 1/d$ and $[P^{-1}]_{i1} = h_i$. If this Gram matrix comes from a MIC, one may use

$$[G]_{ii} = h_i^2 \mathrm{tr}\,\rho_i^2 = \frac{1}{d+1}(h_i + h_i^2)\ , \tag{3.19}$$

and the fact that $\operatorname{tr}\rho_i \leq 1$ to show that $h_i \geq 1/d$. As the average h_i value must be $1/d$, this implies that $h_i = 1/d$ for all i and furthermore that each ρ_i is rank-1. Substituting this into Eq. (3.18) gives

$$[G]_{ij} = \frac{d\delta_{ij} + 1}{d^2(d+1)} \,, \tag{3.20}$$

that is, the MIC is a SIC and $\Phi_{\rm p} = \Phi_{\rm SIC}$. $\square$

Let $s(A)$ denote the vector of singular values of the matrix A in nonincreasing order. The proof of the following lemma may be found in Appendix A.

Lemma 3. *Let Φ be the column-quasistochastic matrix associated with an arbitrary reference process. Then*

$$s(\Phi) \succ_{w\log} s(\Phi_{\rm SIC}) \,, \tag{3.21}$$

with equality iff the MIC and post-measurement states are SICs.

We are now poised to prove:

Theorem 9. *Let Φ be the column-quasistochastic matrix associated with an arbitrary reference process. Then for any unitarily invariant norm $\|\cdot\|$,*

$$\|I - \Phi\| \geq \|I - \Phi_{\rm SIC}\| \,, \tag{3.22}$$

with equality iff the MIC and post-measurement states are SICs.

Proof. By Corollary 3.5.9 in [33], every unitarily invariant norm is monotone with respect to the partial order on matrices induced by weak majorization of the vector of singular values. $I - \Phi$ is singular with exactly one eigenvalue equal to zero, so one of its singular values is zero as well. Then

$$\begin{aligned} s(I-\Phi) &\succ \left\{\frac{\sum_i s_i(I-\Phi)}{d^2-1},\ldots,\frac{\sum_i s_i(I-\Phi)}{d^2-1}\right\} \\ &\succ_w \{d,\ldots,d\} = s(I-\Phi_{\rm SIC}) \end{aligned} \tag{3.23}$$

if

$$\sum_i s_i(I-\Phi) \geq d(d^2-1)\ . \tag{3.24}$$

We have

$$\begin{aligned} \sum_i s_i(I-\Phi) &\geq \sum_i |\lambda_i(I-\Phi)| = \sum_i |\lambda_i(\Phi)-1| \\ &\geq \sum_i (|\lambda_i(\Phi)|-1) \geq \sum_i \lambda_i(\Phi_{\rm SIC}) - d^2 \\ &= d(d^2-1)\ , \end{aligned} \tag{3.25}$$

where the first inequality follows Eq. 3.3.13a in [33], the second follows from the triangle inequality, and the last follows from Lemma 3. □

It is known that no quasiprobability representation of quantum theory can be entirely nonnegative [36]. What does this mean in our formalism?

Let $\mathcal{N}$ be the normalized hyperplane of d^2-element quasiprobability vectors. Within this is the (d^2-1)-simplex of probability vectors, Δ. For any MIC, d-dimensional quantum state space $\mathcal{Q}_d$ is mapped by the Born Rule to a convex subset of Δ, denoted $\mathcal{P}$. Note that $\Phi^{-1}(\Delta)$ is equal to the convex hull of the d^2 probability vectors $\mathrm{tr}\, H_j\sigma_i$, that is, the probabilities for the MIC measurement for each post-measurement state. Consequently, $\Phi^{-1}(\Delta) \subset \mathcal{P}$, which implies $\Delta \subset \Phi(\mathcal{P})$. These inclusions must be strict, i.e., $\Phi \neq I$: When the MIC and post-measurement states are rank-1, the vertices of the simplex will be among the pure-state probability vectors, but $\mathcal{P}$ contains more pure states than there are

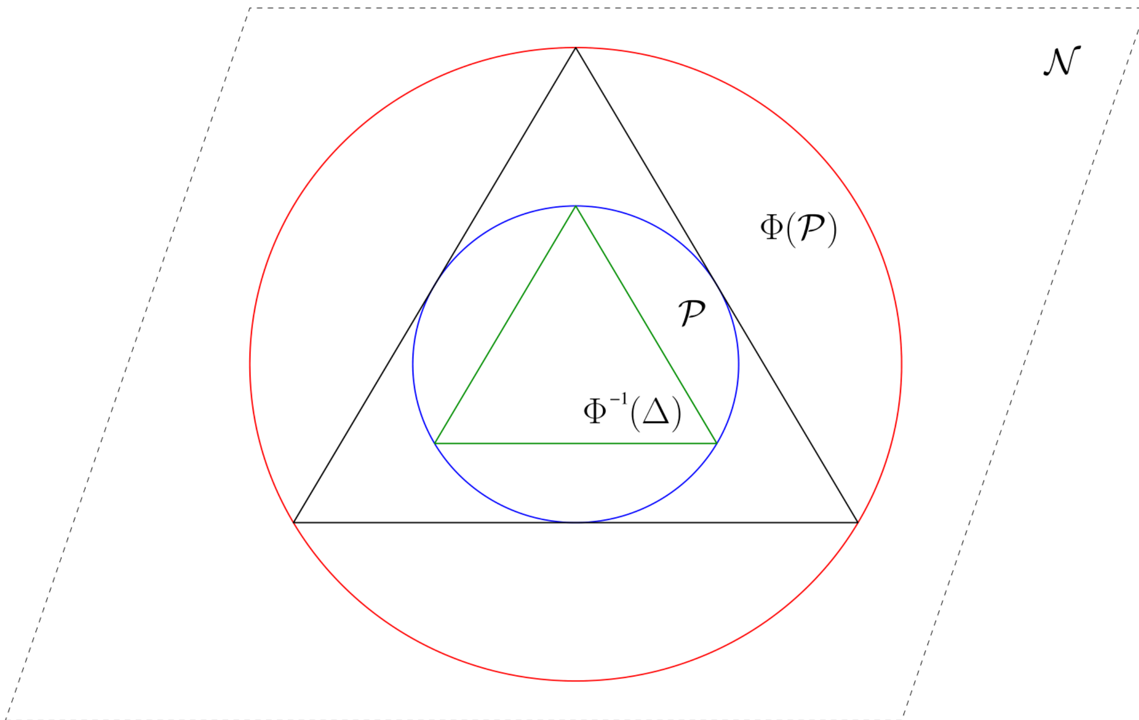

Figure 3.2: $\mathcal{N}$ is the normalized hyperplane of d^2-element quasiprobability vectors and the outer, black triangle represents the (d^2-1)-simplex Δ of probabilities. For a given MIC, the inner, green triangle is the simplex $\Phi^{-1}(\Delta)$, the blue circle is the image of $\mathcal{Q}_d$ under the Born Rule, denoted $\mathcal{P}$, and the red circle is $\Phi(\mathcal{P})$. $\mathcal{P}$ and $\Phi(\mathcal{P})$ are portrayed with circles to capture convexity and inclusion relationships only; they need not bear any resemblance to spheres.

vertices of $\Phi^{-1}(\Delta)$. Since the image of some probability vectors consistent with quantum theory must leave the probability simplex under the application of Φ, we have demonstrated that the appearance of negativity is unavoidable in our framework and is in fact characterized by the fact that Φ cannot equal the identity. Figure 3.2 illustrates the situation.

The weak log majorization result of Lemma 3 has at least one more important implication for quantifying the quantum deviation from classicality. Instead of looking at the functional form of Eq. (5.8) and considering how much of a deviation from the LTP it represents, one may approach the problem from a geometric perspective.

Classically one can always imagine assigning probability 1 to an outcome of a putative "maximally informative measurement"—for instance when one knows the system's exact phase space point. However, in an interpretation of quantum theory without hidden variables, whatever one might mean by "maximally informative," one cannot mean that the reference measurement's full probability simplex is available. Indeed, quantum mechanics does not allow probability 1 for the outcome of any MIC measurement [37]. Thus deviation from classicality can also be captured by the fact that the region of probabilities compatible with quantum states is strictly smaller than the full (d^2-1)-simplex. In this setting, the irreducible deviation from classicality is defined by the largest possible region for a reference measurement's probability simplex. The following theorem establishes that a SIC measurement uniquely maximizes the Euclidean volume of this region, thereby answering a question raised by one of us in 2002 [34, pp. 475, 571].

Theorem 10. *For any MIC in dimension d, let $\mathcal{P}$ denote the image of $\mathcal{Q}_d$ under the Born Rule and let $\mathrm{vol}_{\mathrm{E}}(\mathcal{P})$ denote its Euclidean volume. Then*

$$\mathrm{vol}_{\mathrm{E}}(\mathcal{P}) \leq \mathrm{vol}_{\mathrm{E}}(\mathcal{P}_{\mathrm{SIC}}) , \tag{3.26}$$

with equality iff the MIC is a SIC. Furthermore,

$$\mathrm{vol}_{\mathrm{E}}(\mathcal{P}_{\mathrm{SIC}}) = \sqrt{\frac{(2\pi)^{d(d-1)}}{d^{d^2-2}(d+1)^{d^2-1}}\frac{\Gamma(1)\cdots\Gamma(d)}{\Gamma(d^2)}}\ . \tag{3.27}$$

The proof of Theorem 10 involves methods of differential geometry which would be distracting here. We direct the interested reader to Appendix B for details.

The (d^2-1)-simplex Δ has Euclidean volume [38]

$$\mathrm{vol}_{\mathrm{E}}(\Delta) = \frac{d}{\Gamma(d^2)}\ , \tag{3.28}$$

so we can calculate the ratio of the Euclidean volumes of $\mathcal{P}_{\mathrm{SIC}}$ and the simplex it lies within,

$$\frac{\mathrm{vol}_E(\mathcal{P}_{\mathrm{SIC}})}{\mathrm{vol}_{\mathrm{E}}(\Delta)} = \sqrt{\frac{(2\pi)^{d(d-1)}}{d^{d^2}(d+1)^{d^2-1}}\Gamma(1)\cdots\Gamma(d)}\ . \tag{3.29}$$

When $d = 2$, quantum state space is the Bloch ball and $\mathcal{P}_{\mathrm{SIC}}$ is the largest ball which can be inscribed in the regular tetrahedron Δ_3,

$$\frac{\mathrm{vol}_{\mathrm{E}}(\mathcal{P}_{\mathrm{SIC}})}{\mathrm{vol}_{\mathrm{E}}(\Delta_3)} = \frac{\pi}{6\sqrt{3}} \approx 0.3023\ . \tag{3.30}$$

When $d = 3$,

$$\frac{\mathrm{vol}_{\mathrm{E}}(\mathcal{P}_{\mathrm{SIC}})}{\mathrm{vol}_{\mathrm{E}}(\Delta_8)} = \frac{\pi^3}{1296\sqrt{3}} \approx 0.0138\ . \tag{3.31}$$

In general, the ratio is very rapidly decreasing, signifying a greater and greater deviation from classicality with each Hilbert space dimension.

Theorems 9 and 10 show that the SICs provide a way of casting the Born Rule in wholly probabilistic terms, which by two different standards make the difference between classical and quantum as small as possible. Of all the representations deriving from our general procedure, the representation given by Eq. (3.8) is the essential one for specifying how quantum is quantum.

3.1 Appendix A: Proof of Lemma 3

For a MIC $\{E_i\}$ and a post-measurement set $\{\sigma_j\}$,

$$\left[\Phi^{-1}\right]_{ij} = \mathrm{tr}\, E_i \sigma_j \; . \tag{3.32}$$

The elements of the MIC may be expanded in the SIC basis

$$E_i = \sum_k [\alpha]_{ik}\, H_k \; , \tag{3.33}$$

so we may write

$$\left[\Phi^{-1}\right]_{ij} = \sum_k [\alpha]_{ik}\, \mathrm{tr}\, H_k \sigma_j = \sum_k [\alpha]_{ik}\, p(k|j) \; , \tag{3.34}$$

where $p(k|j)$ is the probabilistic representation of the state σ_j with respect to the SIC $\{H_k\}$. The α matrix must be invertible because it is a transformation between two bases, so the probability vectors can be written

$$p(i|j) = \sum_k \left[\alpha^{-1}\right]_{ik} \left[\Phi^{-1}\right]_{kj} \; . \tag{3.35}$$

We know that SIC probability vectors satisfy [12]

$$\sum_i p(i|j)^2 \le \frac{2}{d(d+1)} \quad \forall j \; , \tag{3.36}$$

so we have

$$\sum_i \left(\sum_k \left[\alpha^{-1}\right]_{ik} \left[\Phi^{-1}\right]_{kj} \right)^2 \leq \frac{2}{d(d+1)} \quad \forall j \ . \tag{3.37}$$

Summing over j, we then have

$$\sum_{ij} \left(\sum_k \left[\alpha^{-1}\right]_{ik} \left[\Phi^{-1}\right]_{kj} \right)^2 \leq \frac{2d}{d+1} \ . \tag{3.38}$$

This expression is the sum of the absolute square entries of a matrix, which is equivalent to the square of the Frobenius norm of the matrix:

$$\|\alpha^{-1}\Phi^{-1}\|_2^2 = \sum_i s^2(\alpha^{-1}\Phi^{-1}) \leq \frac{2d}{d+1} \ . \tag{3.39}$$

From [33] 3.1.11, for any square matrix A,

$$\sum_i |\lambda_i(A)|^2 \leq \sum_i \varsigma^2(A) \ , \tag{3.40}$$

so we have a general bound on the absolute squared spectrum:

$$\sum_i |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \leq \frac{2d}{d+1} \ . \tag{3.41}$$

Eq. (3.35) shows that $\alpha^{-1}\Phi^{-1}$ is column-stochastic and thus that one of its eigenvalues is 1, so we may write:

$$\sum_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \leq \frac{2d}{d+1} - 1 = \frac{d-1}{d+1} \ . \tag{3.42}$$

Now, using the arithmetic-geometric mean inequality,

$$
\begin{aligned}
\frac{d-1}{d+1} &\geq \sum_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2 \\
&\geq (d^2-1)\left(\prod_{i>1} |\lambda_i(\alpha^{-1}\Phi^{-1})|^2\right)^{\frac{1}{d^2-1}} \\
&= (d^2-1)|\det\alpha^{-1}\Phi^{-1}|^{\frac{2}{d^2-1}} \ ,
\end{aligned}
\tag{3.43}
$$

which implies

$$
\begin{aligned}
|\det\alpha^{-1}\Phi^{-1}| &\leq \left(\frac{d-1}{(d+1)(d^2-1)}\right)^{\frac{d^2-1}{2}} \\
&= \left(\frac{1}{d+1}\right)^{d^2-1} = \det\Phi_{\text{SIC}}^{-1} \ .
\end{aligned}
\tag{3.44}
$$

From Eq. (3.33), we can write

$$
\begin{aligned}
\operatorname{tr} E_i E_j = \sum_{kl} \alpha_{ik}\alpha_{jl} \operatorname{tr} H_k H_l \iff G = \alpha G_{\text{SIC}} \alpha^T \\
\iff \det G = (\det\alpha)^2 \det G_{\text{SIC}} \ ,
\end{aligned}
\tag{3.45}
$$

where G is the MIC Gram matrix and G_{SIC} is the SIC Gram matrix. Recall the definition of the A matrix from the proof of Lemma 2. The arithmetic-geometric mean inequality shows $\det A \leq (1/d)^{d^2}$ with equality iff $h_i = 1/d$. Then, since $G = \Phi_{\text{p}}^{-1} A$, Lemma 2 shows

$$
\begin{aligned}
\det G = (\det\Phi_{\text{p}}^{-1})(\det A) &\leq (\det\Phi_{\text{SIC}}^{-1})(1/d)^{d^2} \\
&= \det G_{\text{SIC}} \ ,
\end{aligned}
\tag{3.46}
$$

with equality iff the MIC is a SIC. This implies $(\det \alpha)^2 \leq 1$, and so $|\det \alpha| \leq 1$. Since $|\det \alpha^{-1}\Phi^{-1}| = |\det \alpha^{-1}||\det \Phi^{-1}|$, we conclude that

$$|\det \Phi^{-1}| \leq \det \Phi_{\text{SIC}}^{-1} \; . \tag{3.47}$$

Equivalently, $\det \Phi_{\text{SIC}} \leq |\det \Phi|$. Theorem 3.3.2 in [33] shows $s(A) \succ_{\log} |\lambda(A)|$ for an arbitrary matrix A. To show the desired weak log majorization result, we wish to prove $|\lambda(\Phi)| \succ_{w\log} \lambda(\Phi_{\text{SIC}})$. For this we show weak majorization of the log of the entries.

$$\begin{aligned} \log|\lambda(\Phi)| &\succ \left(\frac{\sum_{i=1}^{d^2}\log|\lambda_i(\Phi)|}{d^2-1},\ldots,\frac{\sum_{i=1}^{d^2}\log|\lambda_i(\Phi)|}{d^2-1},0\right) \\ &= \left(\frac{\log|\det\Phi|}{d^2-1},\ldots,\frac{\log|\det\Phi|}{d^2-1},0\right) \\ &\succ_w \left(\frac{\log\det\Phi_{\text{SIC}}}{d^2-1},\ldots,\frac{\log\det\Phi_{\text{SIC}}}{d^2-1},0\right) \\ &= (\log(d+1),\ldots,\log(d+1),0) = \lambda(\log\Phi_{\text{SIC}}) \; . \end{aligned} \tag{3.48}$$

Thus,

$$s(\Phi) \succ_{\log} |\lambda(\Phi)| \succ_{w\log} \lambda(\Phi_{\text{SIC}}) = s(\Phi_{\text{SIC}}) \; . \tag{3.49}$$

If $\{H_i\}$ and $\{\sigma_j\}$ are SICs, $\Phi^{-1}_{ij} = \frac{1}{d}\mathrm{tr}\,\Pi_i\Pi'_j$, where $\{\Pi_i\}$ and $\{\Pi'_j\}$ are SIC projectors in dimension d. Then

$$
\begin{aligned}
\left[\Phi^{-1}\Phi^{-1\dagger}\right]_{ij} &= \frac{1}{d^2}\sum_k (\mathrm{tr}\,\Pi_i\Pi'_k)(\mathrm{tr}\,\Pi_j\Pi'_k) \\
&= \frac{1}{d^2}\mathrm{tr}\left[(\Pi_i\otimes\Pi_j)\left(\sum_k \Pi'_k\otimes\Pi'_k\right)\right] \\
&= \frac{1}{d^2}\mathrm{tr}\left[(\Pi_i\otimes\Pi_j)\left(\frac{2d}{d+1}P_{\text{sym}}\right)\right] \\
&= \frac{1}{d(d+1)}\mathrm{tr}\left[(\Pi_i\otimes\Pi_j)\left(I\otimes I+\sum_{kl}^{d}|k\rangle\langle l|\otimes|l\rangle\langle k|\right)\right] \\
&= \frac{1+\mathrm{tr}\,\Pi_i\Pi_j}{d(d+1)} = \frac{d\delta_{ij}+d+2}{d(d+1)^2} = \left[\Phi_{\text{SIC}}^{-2}\right]_{ij}\ ,
\end{aligned}
\tag{3.50}
$$

where P_{sym} is the projector onto the symmetric subspace of $\mathcal{H}_d^{\otimes 2}$ and in the third step we employed the fact that the SICs form a minimal 2-design [16]. This shows that the modulus of Φ is equal to Φ_{SIC} and thus the singular values of Φ and Φ_{SIC} coincide.

On the other hand, suppose $s(\Phi) = s(\Phi_{\text{SIC}})$. The product of all the singular values is the absolute value of the determinant [33], so $|\det\Phi^{-1}| = \det\Phi_{\text{SIC}}^{-1} \implies |\det\alpha| = 1 \implies \det G = \det G_{\text{SIC}} \iff \{E_i\}$ is a SIC. Carrying through the consequences of the MIC being a SIC allows us to see from Eq. (3.37) that σ_j is rank-1 because the upper bound is saturated for SIC probability vectors. We may expand the $\{\sigma_j\}$ in the SIC projector basis,

$$
\sigma_j = \sum_k [\beta]_{jk}\,\Pi_k\ . \tag{3.51}
$$

Acting on both sides by a SIC POVM element and computing the trace of both sides, we see

$$
\left[\Phi^{-1}\right]_{ij} = \mathrm{tr}\,E_i\sigma_j = \sum_k [\beta]_{jk}\,\mathrm{tr}\,E_i\Pi_k = [\Phi_{\text{SIC}}^{-1}\beta^T]_{ij}\ , \tag{3.52}
$$

so $|\det\Phi^{-1}| = |\det\Phi_{\text{SIC}}^{-1}||\det\beta^T| = \det\Phi_{\text{SIC}}^{-1}$ implies $|\det\beta| = 1$. Denoting the Gram matrix of states by g, we have, in the same way as before,

$$\det g = (\det\beta)^2 \det g_{\text{SIC}} = \det g_{\text{SIC}} \; . \tag{3.53}$$

We now prove that $\det g = \det g_{\text{SIC}}$ implies that the basis of projectors forms a SIC. The following lemma is due to Huangjun Zhu [39]. We only use part of Zhu's conclusion, but the lemma is of enough interest to present in full.

Lemma 4 (Zhu)**.** *Let λ be the spectrum of the Gram matrix g of a normalized basis of positive semidefinite operators Π_j sorted in nonincreasing order. Then $\lambda \succ \lambda_{\text{SIC}}$ with equality iff Π_j forms a SIC.*

Proof. By assumption $\text{tr}\Pi_j^2 = 1$ for all j. Since the eigenvalues of Π_j are nonnegative,

$$1 = \text{tr}\Pi_j^2 = \sum_i \lambda_i^2(\Pi_j) \leq \sum_i \lambda_i(\Pi_j) = \text{tr}\Pi_j. \tag{3.54}$$

Define the frame superoperator

$$\mathcal{F} = \sum_j |\Pi_j\rangle\rangle\langle\langle\Pi_j| \; , \tag{3.55}$$

where $|A\rangle\rangle := \sum_{ij}[A]_{ij}|i\rangle|j\rangle$. $\mathcal{F}$ has the same spectrum as the Gram matrix $[g]_{ij} = \langle\langle\Pi_i|\Pi_j\rangle\rangle = \text{tr}\,\Pi_i\Pi_j$. To see this, form a projector out of the state $\sum_i|\Pi_i\rangle\rangle|i\rangle$ where $|i\rangle$ is an orthonormal basis in $\mathcal{H}_{d^2}$ and perform partial traces over each subsystem. The results are g^{T} and $\mathcal{F}$, and so, by the Schmidt theorem, the spectra of $\mathcal{F}$ and g are equal: $\lambda(g) = \lambda(\mathcal{F}) = \lambda$.

The expectation value of any operator with respect to an arbitrary normalized state is less than or equal to its maximal eigenvalue. Thus, a lower bound on the maximal eigenvalue λ_1 of $\mathcal{F}$ is given by

$$\lambda_1 \geq \frac{1}{d}\langle\langle I|\mathcal{F}|I\rangle\rangle = \frac{1}{d}\sum_j (\operatorname{tr}\Pi_j)^2 \geq d. \tag{3.56}$$

As our basis is normalized, $\operatorname{tr} g = d^2$, so $\sum_i \lambda_i = d^2$. With this constraint and our bound on the maximal eigenvalue, we have

$$\begin{aligned}\lambda &\succ \left(\lambda_1, \frac{d^2-\lambda_1}{d^2-1}, \ldots, \frac{d^2-\lambda_1}{d^2-1}\right) \\ &\succ \left(d, \frac{d}{d+1}, \ldots, \frac{d}{d+1}\right) = \lambda_{\text{SIC}}.\end{aligned} \tag{3.57}$$

The second majorization becomes an equality when $\lambda_1 = d$. From Eq. (3.56), we can see that all Π_j must be rank-1 for this condition to be satistfied. Furthermore, we see that in this case $\frac{1}{\sqrt{d}}|I\rangle\rangle$ is an eigenvector of $\mathcal{F}$ which achieves the maximal eigenvalue d. When both majorizations are equalities the spectrum λ_{SIC} tells us that $\mathcal{F}$ takes the form of a weighted sum of a projector and the identity superoperator $\mathbf{I}$, specifically

$$\mathcal{F} = \frac{d}{d+1}\left(\mathbf{I} + |I\rangle\rangle\langle\langle I|\right). \tag{3.58}$$

By Cor. 1 in [40], this implies the Π_j form a SIC. □

As in the Lemma, denote by λ the spectrum of g sorted in nonincreasing order. $\operatorname{tr} g = d^2$, so

$$\sum_{i>1} \lambda_i = d^2 - \lambda_1 \ . \tag{3.59}$$

Then because the arithmetic mean is greater than or equal to the geometric mean with equality iff the elements are all equal, we have

$$\frac{1}{d^2-1}\sum_{i>1}\lambda_i = \frac{d^2-\lambda_1}{d^2-1} \geq \left(\prod_{i>1}\lambda_i\right)^{\frac{1}{d^2-1}} \ , \tag{3.60}$$

which implies

$$\det g \leq \lambda_1 \left(\frac{d^2-\lambda_1}{d^2-1}\right)^{d^2-1} \ , \tag{3.61}$$

with equality iff $\lambda_2 = \cdots = \lambda_d^2 = \frac{d^2-\lambda_1}{d^2-1}$. When $\lambda_1 = d$, we then have

$$\det g = \frac{d^{d^2}}{(d+1)^{d^2-1}} = \det g_{\mathrm{SIC}} \ , \tag{3.62}$$

with equality iff $\lambda = \lambda_{\mathrm{SIC}}$. By Lemma 18, we have equality iff the post-measurement states form a SIC.

3.2 Appendix B: Proof of Theorem 10

Equation (3.4) expanded instead in the ρ_i basis allows us to relate the differential elements of operator space and probability space for any MIC basis:

$$\mathrm{d}\sigma = \sum_{i,j}[\Phi]_{ij}\rho_i \mathrm{d}p^j \ . \tag{3.63}$$

The Hilbert–Schmidt line element is then

$$\mathrm{d}s_{\mathrm{HS}}^2 = \mathrm{tr}\,(\mathrm{d}\sigma)^2 = \sum_{ijkl}[\Phi]_{ij}[\Phi]_{kl}(\mathrm{tr}\,\rho_i\rho_k)\mathrm{d}p^j\mathrm{d}p^l \ . \tag{3.64}$$

As in the proof of Lemma 2, we write $\Phi = AG^{-1}$ where $[G]_{ij} = \operatorname{tr} H_i H_j$ is the Gram matrix for the MIC and $[A]_{ij} = h_i \delta_{ij}$. Note further that $\operatorname{tr} \rho_i \rho_j = [A^{-1} G A^{-1}]_{ij}$. Then Eq. (3.64) simplifies to

$$\mathrm{d}s^2_{\mathrm{HS}} = \sum_{ij} \left[G^{-1}\right]_{ij} \mathrm{d}p^i \mathrm{d}p^j \ . \tag{3.65}$$

The Hilbert–Schmidt volume element on the space of Hermitian operators in $\mathcal{L}(\mathcal{H}_d)$ may now be related to the Euclidean volume element in $\mathbb{R}^{d^2}$,

$$\mathrm{d}\Omega_{\mathrm{HS}} = \sqrt{|\det G^{-1}|}\mathrm{d}V_{\mathrm{E}} \ , \tag{3.66}$$

or, equivalently,

$$\mathrm{d}V_{\mathrm{E}} = \sqrt{|\det G|}\mathrm{d}\Omega_{\mathrm{HS}} \ . \tag{3.67}$$

The larger $\det G$, the larger the corresponding Euclidean volume. Recall Eq. (3.46) which says

$$\det G \leq \det G_{\mathrm{SIC}} \ , \tag{3.68}$$

with equality iff the MIC is a SIC. Thus, for any region in operator space, the Euclidean volume is maximal with respect to the SIC basis. In particular, the SIC basis gives the largest volume among positive semidefinite operators A satisfying $1 - \epsilon \leq \operatorname{tr} A \leq 1 + \epsilon$ for any $\epsilon > 0$. As $\epsilon \to 0$, we obtain quantum state space $\mathcal{Q}_d$ and the corresponding region in $\mathbb{R}^{d^2}$ will have the largest hyperarea within Δ when computed with the SIC basis.

To calculate this hyperarea, we need to find the metric on Δ induced by the Hilbert–Schmidt metric in the SIC basis. We may parameterize Δ by

$$X = \left(p^1, \ldots, p^{d^2-1}, 1 - \sum_{i=1}^{d^2-1} p^i\right) \ , \tag{3.69}$$

which has partial derivatives $\partial_i X^\mu = \delta_i^\mu - \delta_{d^2}^\mu$ where the Latin index runs from 1 to $d^2 - 1$ and the Greek index runs from 1 to d^2. For any MIC, the induced metric g is given by

$$[g]_{ij} = \sum_{\mu,\nu=1}^{d^2} \partial_i X^\mu \partial_j X^\nu \left[G^{-1}\right]_{\mu\nu} \ . \tag{3.70}$$

It is easily seen that $G_{\text{SIC}}^{-1} = d(d+1)I - J$ where J is the Hadamard identity. One may then calculate $g_{\text{SIC}} = d(d+1)(I+J)$ and $\det g_{\text{SIC}} = d^2(d^2+d)^{d^2-1}$. The induced volume element on Δ is then

$$\mathrm{d}\omega_{\text{HS}} = d\sqrt{(d^2+d)^{d^2-1}}\mathrm{d}p^1 \cdots \mathrm{d}p^{d^2-1} \ . \tag{3.71}$$

In a similar way, it may be checked that the Euclidean metric in $\mathbb{R}^{d^2}$ induces a volume element $\mathrm{d}\mathcal{A}_{\text{E}}$ on Δ satisfying

$$\frac{1}{d}\mathrm{d}\mathcal{A}_{\text{E}} = \mathrm{d}p^1 \cdots \mathrm{d}p^{d^2-1} \ , \tag{3.72}$$

and so

$$\mathrm{d}\omega_{\text{HS}} = \sqrt{(d^2+d)^{d^2-1}}\mathrm{d}\mathcal{A}_{\text{E}} \ . \tag{3.73}$$

We may now integrate over quantum state space to obtain

$$\text{vol}_{\text{HS}}(\mathcal{Q}_d) = \sqrt{(d^2+d)^{d^2-1}}\text{vol}_{\text{E}}(\mathcal{P}_{\text{SIC}}) \ . \tag{3.74}$$

Życzkowski and Sommers [41] calculate the Hilbert–Schmidt volume of finite-dimensional quantum state space to be

$$
\mathrm{vol}_{\mathrm{HS}}(\mathcal{Q}_d) = \sqrt{d}(2\pi)^{d(d-1)/2} \frac{\Gamma(1)\cdots\Gamma(d)}{\Gamma(d^2)} , \tag{3.75}
$$

from which Eq. (3.27) follows.

LIST OF REFERENCES

[1] R. P. Feynman, “Simulating physics with computers,” *International Journal of Theoretical Physics*, vol. 21, no. 6-7, pp. 467–488, 1982.

[2] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, “The resource theory of stabilizer computation,” *New Journal of Physics*, vol. 16, p. 013009, 2014.

[3] M. Howard, J. Wallman, V. Veitch, and J. Emerson, “Contextuality supplies the ‘magic’ for quantum computation,” *Nature*, vol. 510, pp. 351–355, 2014.

[4] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, “Wigner Function Negativity and Contextuality in Quantum Computation on Rebits,” *Physical Review X*, vol. 5, no. 2, p. 021003, 2015.

[5] H. Pashayan, J. J. Wallman, and S. D. Bartlett, “Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities,” *Physical Review Letters*, vol. 115, no. 7, p. 070501, 2015.

[6] B. Stacey, “SIC-POVMs and Compatibility among Quantum States,” *Mathematics*, vol. 4, no. 2, p. 36, 2016.

[7] H. Zhu, “Permutation Symmetry Determines the Discrete Wigner Function,” *Physical Review Letters*, vol. 116, no. 4, p. 040501, 2016.

[8] J. B. DeBrota and C. A. Fuchs, “Negativity bounds for Weyl–Heisenberg quasiprobability representations,” *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

[9] M. Howard and E. T. Campbell, “Application of a resource theory for magic states to fault-tolerant quantum computing,” *Physical Review Letters*, vol. 118, no. 9, p. 090501, 2017.

[10] R. Raussendorf, D. E. Browne, N. Delfosse, C. Okay, and J. Bermejo-Vega, “Contextuality and Wigner-function negativity in qubit quantum computation,” *Physical Review A*, vol. 95, no. 5, p. 052334, 2017.

[11] L. Kocia and P. Love, “Discrete Wigner formalism for qubits and noncontextuality of Clifford gates on qubit stabilizer states,” *Physical Review A*, vol. 96, no. 6, p. 062134, 2017.

[12] C. A. Fuchs and R. Schack, “Quantum-Bayesian coherence,” *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[13] C. A. Fuchs and B. C. Stacey, “QBism: Quantum Theory as a Hero’s Handbook,” in *Proceedings of the International School of Physics "Enrico Fermi": Course 197, Foundations of Quantum Theory*, pp. 133–202, 2019.

[14] E. Prugovečki, “Information-theoretical aspects of quantum measurement,” *International Journal of Theoretical Physics*, vol. 16, no. 5, pp. 321–331, 1977.

[15] G. Zauner, *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*. PhD thesis, University of Vienna, 1999. Published in English translation: G. Zauner, "Quantum designs: foundations of a noncommutative design theory," *International Journal of Quantum Information*, vol. 9, pp. 445–508, 2011.

[16] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, “Symmetric informationally complete quantum measurements,” *Journal of Mathematical Physics*, vol. 45, pp. 2171–2180, 2004.

[17] A. J. Scott and M. Grassl, “Symmetric informationally complete positive-operator-valued measures: A new computer study,” *Journal of Mathematical Physics*, vol. 51, p. 042203, 2010.

[18] D. Appleby, H. Dang, and C. Fuchs, “Symmetric Informationally-Complete Quantum States as Analogues to Orthonormal Bases and Minimum-Uncertainty States,” *Entropy*, vol. 16, no. 3, pp. 1484–1492, 2014.

[19] M. Appleby, S. Flammia, G. McConnell, and J. Yard, “Generating ray class fields of real quadratic fields via complex equiangular lines,” *Acta Arithmetica*, vol. 192, no. 3, pp. 211–233, 2020.

[20] M. Appleby, T.-Y. Chien, S. Flammia, and S. Waldron, “Constructing exact symmetric informationally complete measurements from numerical solutions,” *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 16, p. 165302, 2018.

[21] A. J. Scott, “SICs: Extending the list of solutions,” *arXiv:1703.03993*, 2017.

[22] M. Grassl and A. J. Scott, “Fibonacci–Lucas SIC-POVMs,” *Journal of Mathematical Physics*, vol. 58, p. 122201, 2017.

[23] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, “The SIC question: History and state of play,” *Axioms*, vol. 6, p. 21, 2017.

[24] B. C. Stacey, “Sporadic SICs and the normed division algebras,” *Foundations of Physics*, vol. 47, pp. 1060–64, 2017.

[25] T. Durt, C. Kurtsiefer, A. Lamas-Linares, and A. Ling, “Wigner tomography of two-qubit states and quantum cryptography,” *Physical Review A*, vol. 78, no. 4, p. 042338, 2008.

[26] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, "Experimental characterization of qutrits using symmetric informationally complete positive operator-valued measurements," *Physical Review A*, vol. 83, no. 5, p. 051801, 2011.

[27] N. Bent, H. Qassim, A. Tahir, D. Sych, G. Leuchs, L. Sánchez-Soto, E. Karimi, and R. Boyd, "Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures," *Physical Review X*, vol. 5, no. 4, p. 041006, 2015.

[28] Z. Hou, J.-F. Tang, J. Shang, H. Zhu, J. Li, Y. Yuan, K.-D. Wu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, "Deterministic realization of collective measurements via photonic quantum walks," *Nature Communications*, vol. 9, no. 1, p. 1414, 2018.

[29] R. P. Feynman, "Negative probability," in *Quantum Implications: Essays in Honour of David Bohm* (B. Hiley and F. D. Peat, eds.), pp. 235–48, Routledge, 1991.

[30] C. A. Fuchs, "Notwithstanding Bohr, the Reasons for QBism," *Mind and Matter*, vol. 15, no. 2, 2017.

[31] M. Appleby, S. Flammia, G. McConnell, and J. Yard, "SICs and algebraic number theory," *Foundations of Physics*, vol. 47, pp. 1042–59, 2017.

[32] J. van de Wetering, "Quantum Theory is a Quasi-stochastic Process Theory," *Electronic Proceedings in Theoretical Computer Science*, vol. 266, pp. 179–196, 2018.

[33] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*. Cambridge University Press, 2008.

[34] C. A. Fuchs, M. Schlosshauer, and B. C. Stacey, "My Struggles with the Block Universe," *arXiv:1405.2390*, 2014.

[35] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*. Springer Series in Statistics, Springer, 2 ed., 2011.

[36] C. Ferrie and J. Emerson, "Framed Hilbert space: hanging the quasi-probability pictures of quantum theory," *New Journal of Physics*, vol. 11, p. 063040, 2009.

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

[38] C. M. Caves, "Measures and volumes for spheres, the probability simplex, projective Hilbert space, and density operators." `http://info.phys.unm.edu/~caves/reports/measures.pdf`, 2001.

[39] H. Zhu, “Geometry of Quantum States,” 2017. Unpublished notes.

[40] D. M. Appleby, C. A. Fuchs, and H. Zhu, “Group theoretic, Lie algebraic and Jordan algebraic formulations of the SIC existence problem,” *Quantum Information & Computation*, vol. 15, pp. 61–94, 2015.

[41] K. Życzkowski and H.-J. Sommers, “Hilbert–Schmidt volume of the set of mixed quantum states,” *Journal of Physics A: Mathematical and General*, vol. 36, no. 39, pp. 10115–10130, 2003.

CHAPTER 4

LÜDERS MIC CHANNELS

4.1 Introduction

A *minimal informationally complete* (MIC) quantum measurement for a d dimensional Hilbert space $\mathcal{H}_d$ is a set of linearly independent positive semidefinite operators $\{E_i\}$, $i = 1, \ldots, d^2$, which sum to the identity [1, 2]. If every element in a MIC is proportional to a rank-n projector, we say the MIC itself is *rank-n*. If the Hilbert–Schmidt inner products $\operatorname{tr} E_i E_j$ equal one constant for all $i \neq j$ and another constant when $i = j$, we say the MIC is *equiangular*. A *symmetric informationally complete* (SIC) quantum measurement is a rank-1 equiangular MIC [3–6]. When a SIC $\{H_i\}$ exists, one can show that $H_i = \frac{1}{d}\Pi_i$ where Π_i are rank-1 projectors and that

$$\operatorname{tr} H_i H_j = \frac{d\delta_{ij} + 1}{d^2(d+1)} \ . \tag{4.1}$$

The theory of quantum channels provides a means to discuss the fully general way in which quantum states may be transformed. A standard result [7] has it that a quantum channel $\mathcal{E}$ may always be specified by a set of operators $\{A_i\}$, called *Kraus operators*, such that for a quantum state ρ,

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger \ . \tag{4.2}$$

Consider a physicist Alice who is preparing to send a quantum system through a channel that she models by $\mathcal{E}$. Alice initially describes her quantum system by assigning to it a density matrix ρ. The state $\mathcal{E}(\rho)$ encodes Alice's expectations for measurements that can potentially be performed after the system is sent through the channel. More specifically, let Alice's channel be a *Lüders MIC channel* (LMC) associated with the MIC $\{E_i\}$, which may be understood in the following way. Alice plans to apply the MIC $\{E_i\}$, and upon obtaining the result of that measurement, invoke the Lüders rule [8, 9] to obtain a new state for her system,

$$\rho_i' := \frac{1}{\operatorname{tr} \rho E_i} \sqrt{E_i} \rho \sqrt{E_i} \; , \tag{4.3}$$

where we have introduced the *principal Kraus operators* $\{\sqrt{E_i}\}$, the unique positive semidefinite square roots of the MIC elements. Before applying her MIC, Alice can write the *post-channel state*

$$\mathcal{E}(\rho) := \sum_i p(E_i)\rho_i' = \sum_i \sqrt{E_i}\rho\sqrt{E_i} \; , \tag{4.4}$$

which is a weighted average of the states from which Alice plans to select the actual state she will ascribe to the system after making the measurement. (For more on the broader conceptual context of this operation, see [10, 11].)

LMCs are a proper subset of all quantum channels as many valid channels are unrelated to a MIC and do not admit a representation in terms of principal Kraus operators. For example, a unitary channel is not an LMC.

Throughout this paper, we will frequently use the fact that any MIC element E_i is proportional to a density matrix $E_i := e_i \rho_i$, where we call the proportionality constants $\{e_i\}$ the *weights* of the MIC. Because the $\{E_i\}$ sum to the identity, the weights $\{e_i\}$ sum to the trace of the identity, which is just the dimension d.

We refer to the LMC obtained from a SIC as the *SIC channel* $\mathcal{E}_{\rm SIC}$. We may characterize the SIC channel in any dimension in which a SIC exists using the convenient notion of a *dual basis*. Given a basis for a vector space, any vector in that space is uniquely identified by its inner products with the basis elements. These inner products are the coefficients in the expansion of the vector over the elements of the dual basis. Likewise, the inner products with the elements of the dual basis are the coefficients in the expansion over the original basis. A consequence of this is that, if $\{H_i\}$ denotes the original basis and $\{\widetilde{H}_j\}$ denotes its dual basis, then

$$\mathrm{tr}\, H_i \widetilde{H}_j = \delta_{ij}. \tag{4.5}$$

It follows that if $\{H_i\}$ is a SIC, then the dual basis is given by

$$\widetilde{H}_j = (d+1)\Pi_j - I \ , \tag{4.6}$$

so we may write an operator $X \in \mathcal{L}(\mathcal{H}_d)$ in the SIC basis as

$$\begin{aligned} X &= \sum_j (\mathrm{tr}\, X\widetilde{H}_j) H_j \\ &= (d+1)\sum_j (\mathrm{tr}\, X\Pi_j) H_j - (\mathrm{tr}\, X) I \ . \end{aligned} \tag{4.7}$$

Noting that $(\mathrm{tr}\, X\Pi_j)H_j = \frac{1}{d}\Pi_j X \Pi_j$ and that $\sqrt{H_j} = \frac{1}{\sqrt{d}}\Pi_j$, we obtain

$$\mathcal{E}_{\rm SIC}(X) = \frac{1}{d}\sum_j \Pi_j X \Pi_j = \frac{(\mathrm{tr}\, X)I + X}{d+1} \ , \tag{4.8}$$

and so for any quantum state ρ,

$$\mathcal{E}_{\mathrm{SIC}}(\rho) = \frac{I+\rho}{d+1} \ . \tag{4.9}$$

Going forward, given an LMC $\mathcal{E}$ and input state ρ, let λ denote the eigenvalue spectrum of the post-channel state $\mathcal{E}(\rho)$ and $\lambda_{\max}$ denote the maximum eigenvalue of this state. We use the notation $\overline{f(|\psi\rangle\langle\psi|)}$ to denote the average value of the function $f(|\psi\rangle\langle\psi|)$ over all pure state inputs $|\psi\rangle\langle\psi|$ with respect to the Haar measure. We now prove a lemma applicable to arbitrary LMCs upon which our later results rely.

Lemma 5. *Let $\mathcal{E}$ be an LMC. Then*

$$\overline{\lambda}_{\max} \geq \frac{2}{d+1} \ . \tag{4.10}$$

Proof. For an arbitrary pure state $\rho = |\psi\rangle\langle\psi|$,

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \sum_i \sqrt{E_i}|\psi\rangle\langle\psi|\sqrt{E_i} \ . \tag{4.11}$$

We may lower bound $\lambda_{\max}$ given such an input as follows:

$$\lambda_{\max} \geq \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle = \sum_i |\langle\psi|\sqrt{E_i}|\psi\rangle|^2 \ . \tag{4.12}$$

If we now average over all pure states with the Haar measure, we will produce a generic lower bound on the average maximal eigenvalue of the post-channel state:

$$\overline{\lambda}_{\max} \geq \sum_i \int_{\mathcal{H}} |\langle\psi|\sqrt{E_i}|\psi\rangle|^2 d\Omega_\psi \ . \tag{4.13}$$

We can evaluate this integral using a known property of the Haar measure [4]. Integrating a tensor power over pure state projectors gives a result proportional to the projector $P_{\rm sym}$ onto the symmetric subspace of $\mathcal{H}_d \otimes \mathcal{H}_d$. Consequently,

$$\overline{\lambda}_{\max} \geq \frac{2}{d(d+1)} \sum_i \mathrm{tr}\left[\left(\sqrt{E_i} \otimes \sqrt{E_i}\right) P_{\rm sym}\right] . \tag{4.14}$$

Using the fact that

$$P_{\rm sym} = \frac{1}{2}\left(I \otimes I + \sum_{kl} |l\rangle\langle k| \otimes |k\rangle\langle l|\right) , \tag{4.15}$$

one may verify that

$$\mathrm{tr}\left[\left(\sqrt{E_i} \otimes \sqrt{E_i}\right) P_{\rm sym}\right] = \frac{1}{2}\left(\mathrm{tr}\sqrt{E_i}\right)^2 + \frac{e_i}{2} . \tag{4.16}$$

Thus equation (4.14) becomes

$$\overline{\lambda}_{\max} \geq \frac{1}{d(d+1)}\left(\sum_i \left(\mathrm{tr}\sqrt{E_i}\right)^2 + \sum_i e_i\right) . \tag{4.17}$$

Now since $\mathrm{tr}\sqrt{E_i} = \sqrt{e_i}\mathrm{tr}\sqrt{\rho_i}$ and $\mathrm{tr}\sqrt{\rho_i} \geq 1$,

$$\overline{\lambda}_{\max} \geq \frac{2}{d(d+1)} \sum_i e_i = \frac{2}{d+1} . \tag{4.18}$$

□

The following theorem reveals that the SIC channel's action is unique to SICs among LMCs.

Theorem 11. *A SIC exists in dimension d iff there is an LMC with action $\mathcal{E}(\rho) = \frac{I+\rho}{d+1}$ for all ρ.*

Proof. If a SIC exists, take $\mathcal{E} = \mathcal{E}_{\mathrm{SIC}}$. For the other direction, we will first demonstrate that the MIC which gives rise to this LMC must be rank-1. Having established this, we will be able to see that the unitaries relating different Kraus operators for this LMC are directly related to the MIC weights. This will allow us to show that the principal Kraus operators have the Gram matrix of a SIC and must form a SIC themselves.

For any pure state input,

$$\lambda = \left(\frac{2}{d+1}, \frac{1}{d+1}, \ldots, \frac{1}{d+1}\right) \ . \tag{4.19}$$

Thus the lower bound in Lemma 5 is saturated. This can only occur when $\mathrm{tr}\sqrt{\rho_i} = 1$ for all i which implies that the MIC is rank-1.

In Appendix A we define and construct the *quasi-SICs*, that is, sets of Hermitian, but not necessarily postive semidefinite, matrices $\{Q_i\}$ which have the same Hilbert–Schmidt inner products as SIC projectors, and we demonstrate that they furnish a Hermitian basis of constant-trace Kraus operators A_i which give the same action as $\mathcal{E}$. Any other set of Kraus operators with the same effect will be related to this set by a unitary remixing, and since $\mathcal{E}$ is an LMC, we must have

$$\sqrt{E_i} = \sum_j [U]_{ij} A_j \ . \tag{4.20}$$

Since we know the MIC is rank-1, we can trace both sides of this expression to obtain the identity $\sqrt{de_i} = \sum_j [U]_{ij}$. Furthermore, since the A_j form a Hermitian basis, one may see that every element of U must be real, rendering U an orthogonal matrix. Then

$$\begin{aligned}\operatorname{tr}\left(\sqrt{E_i}\sqrt{E_j}\right) &= \frac{1}{d}\sum_{k,l}[U]_{ik}[U]_{jl}\operatorname{tr} Q_k Q_l \\ &= \frac{1}{d}\sum_{k,l}[U]_{ik}[U]_{jl}\frac{d\delta_{kl}+1}{d+1} \\ &= \frac{d\delta_{ij}+d\sqrt{e_i e_j}}{d(d+1)}\;.\end{aligned} \tag{4.21}$$

When $i = j$, we have

$$e_i = \frac{1+e_i}{d+1} \implies e_i = \frac{1}{d}\;, \tag{4.22}$$

and so

$$\operatorname{tr}\left(\sqrt{E_i}\sqrt{E_j}\right) = \frac{d\delta_{ij}+1}{d(d+1)}\;, \tag{4.23}$$

and because E_i is rank-1,

$$\operatorname{tr}(E_i E_j) = \frac{d\delta_{ij}+1}{d^2(d+1)}\;, \tag{4.24}$$

that is, the MIC is a SIC. □

4.2 Depolarizing Lüders MIC Channels

The SIC channel falls within a class of channels called *depolarizing channels* [12]. A depolarizing channel is a channel

$$\mathcal{E}_\alpha(\rho) = \alpha\rho + \frac{1-\alpha}{d}I\;, \qquad \frac{-1}{d^2-1} \le \alpha \le 1\;. \tag{4.25}$$

The SIC channel corresponds to $\alpha = \frac{1}{d+1}$. One might wish to know when an LMC is a depolarizing channel. From Theorem 11, we know the only LMC with $\alpha = \frac{1}{d+1}$ is the SIC channel. What range of α are achievable by LMCs?

The answer to this question is any $\frac{1}{d+1} \leq \alpha < 1$. To see this, note that the eigenvalue spectrum for a depolarizing channel given a pure state input is

$$\lambda\left(\mathcal{E}_\alpha(|\psi\rangle\langle\psi|)\right) = \left(\alpha + \frac{1-\alpha}{d}, \frac{1-\alpha}{d}, \ldots, \frac{1-\alpha}{d}\right) . \tag{4.26}$$

Recall the lower bound on the average maximal eigenvalue for any LMC given a pure state input from Lemma 5 is $\frac{2}{d+1}$. As the spectrum for a depolarizing channel is constant for pure state inputs, the lower bound on the average is the lower bound for any pure state input. If $\lambda_{\max} = \frac{1-\alpha}{d}$, then $\frac{1-\alpha}{d} \geq \alpha + \frac{1-\alpha}{d} \implies \alpha \leq 0$. The more negative α is, the larger the maximal eigenvalue would be, so the largest it can get is when $\alpha = \frac{-1}{d^2-1}$, in which case $\lambda_{\max} = \frac{d}{d^2-1} < \frac{2}{d+1}$. So, $\lambda_{\max} = \alpha + \frac{1-\alpha}{d} \geq \frac{2}{d+1} \implies \alpha \geq \frac{1}{d+1}$. When $\alpha = 1$, the channel is the identity channel, in other words, not depolarizing at all. It is easy to prove that were this to be implemented by an LMC, it would require $\sqrt{E_i} = \frac{1}{d}I$, but this does not lead to a linearly independent set and is not a MIC. If a SIC exists, however, a depolarizing LMC exists for any $\frac{1}{d+1} \leq \alpha < 1$, as the next proposition shows.

Proposition 1. *Suppose a SIC exists in dimension d. For a nonzero $\beta \in \left[\frac{-1}{d-1}, 1\right]$ satisfying*

$$\alpha = 1 - \frac{\left(\sqrt{1-\beta+d\beta} - \sqrt{1-\beta}\right)^2}{d+1} , \tag{4.27}$$

or, equivalently,

$$\begin{aligned} \beta = \frac{1}{d^2}\Big(&(d-2)(d+1)(1-\alpha) \\ &+ 2\sqrt{(d+1)(1-\alpha)(1-\alpha+d^2\alpha)}\Big) , \end{aligned} \tag{4.28}$$

the MIC

$$E_i = \frac{\beta}{d}\Pi_i + \frac{1-\beta}{d^2} I , \tag{4.29}$$

where Π_i is a SIC projector, gives rise to the LMC

$$\mathcal{E}_\alpha(\rho) = \alpha\rho + \frac{1-\alpha}{d} I \;, \qquad \frac{1}{d+1} \leq \alpha < 1 \;. \tag{4.30}$$

One may check that the principal Kraus operators associated with the MIC elements (4.29) are given by

$$A_i = \frac{\sqrt{1-\beta+d\beta} - \sqrt{1-\beta}}{d} \Pi_i + \frac{\sqrt{1-\beta}}{d} I \;, \tag{4.31}$$

and then a routine calculation and the characterization of the SIC channel from Theorem 11 confirms the claim of the proposition.

Remark. When $\beta = 1$, the MIC $\{E_i\}$ is the original SIC, whereas when β equals its minimum allowed value $-1/(d-1)$, it is the rank-$(d-1)$ equiangular MIC

$$E_i = \frac{1}{d(d-1)} (-\Pi_i + I) \;, \tag{4.32}$$

indirectly noted in prior work [2, 13, 14] for extremizing a nonclassicality measure based on negativity of quasi-probability.

Do any LMCs give rise to depolarizing channels in dimensions where one does not have access to a SIC? If we replace the SIC projector in equation (4.31) with a quasi-SIC, we may form Kraus operators effecting the same depolarizing channel,

$$K_i = \frac{\sqrt{1-\beta+d\beta} - \sqrt{1-\beta}}{d} Q_i + \frac{\sqrt{1-\beta}}{d} I \;. \tag{4.33}$$

From (4.59), one can check that these will square to a valid MIC. For arbitrary β, however, K_i may fail to be positive semidefinite and would therefore not be a principal Kraus operator. From the definition of a quasi-SIC, one sees that the eigenvalues of Q_i are bounded below by -1. Even in this worst case, one can easily derive that K_i will be positive semidefinite for any nonzero $\beta \leq \frac{3}{d+3}$. This range of β entitles any $\alpha \geq \frac{d^2-d-1}{d^2-1}$. (The minimal α is obtained from the most negative β.) When $d = 2$, this minimal α matches the lower bound achieved by the SIC channel because every quasi-SIC is a SIC in this dimension, but for all $d > 2$ the inequality is strict and monotonically increases with dimension. In practice, the minimal eigenvalue among all of the quasi-SIC operators one constructs will be significantly larger than -1, and so, depending on how close to a SIC one can make their quasi-SIC, one should be able to get significantly closer to the SIC bound than the α we have derived.

Fully classifying the MICs giving depolarizing LMCs for particular values of $\alpha > \frac{1}{d+1}$ appears to be a difficult problem; it is not clear what properties these MICs must satisfy. For example, squaring the K_i operators from equation (4.33) results in MICs which are dependent on one's quasi-SIC implementation and need not be equiangular as the family in equation (4.29) was. All principal Kraus operators which give rise to a depolarizing channel with a given β (and corresponding α) will be related to the operators (4.33) by way of a unitary remixing satisfying

$$\sqrt{E_i} = \sum_j [U]_{ij} K_j \tag{4.34}$$

for some unitary U. As in the proof of Theorem 11, all the elements of the unitary must be real and so it is actually an orthogonal matrix. We have not been able to identify any further necessary characteristics of the U in the completely general case, but the following notable restriction yielded further structure. A MIC is *unbiased* if the traces of all the elements are equal, that is, if $e_i = \frac{1}{d}$ for all i. MICs in this class have the property that their measurement outcome probabilities for the “garbage state” $\frac{1}{d}I$ input is the flat probability distribution over d^2 outcomes. From the standpoint of [1], this means they preserve the intution that the state $\frac{1}{d}I$ should correspond to a prior with complete outcome indifference in a reference process scenario and accordingly warrant special attention. If we demand that $\{E_i\}$ be unbiased, then it is necessary, but not sufficient, that the orthogonal matrix remixing (4.33) be doubly quasistochastic (see Appendix B).

4.3 Entropic Optimality

One way to evaluate the performance of a quantum channel is by using measures based on von Neumann entropy,

$$S(\rho) = -\mathrm{tr}\,\rho \log \rho\ . \tag{4.35}$$

In this section, we consider two such, proving in each case an optimality result for LMCs constructed from SICs. To understand the conceptual significance of the bounds we will derive, consider again Alice who is preparing to send a quantum system through an LMC. Alice initially ascribes the quantum state ρ to her system, and before sending the system through the channel, she computes $\mathcal{E}(\rho)$. After eliciting a measurement outcome, Alice will update her quantum-state assignment, not to $\mathcal{E}(\rho)$ but rather to whichever ρ_i' corresponds to the outcome E_i that actually transpires. The state $\mathcal{E}(\rho)$ will generally be mixed, while ρ_i' will be a pure state in the case of a rank-1 MIC. This change from mixed to pure represents a sharpening of Alice's expectations about her quantum system. We can quantify this in entropic terms, even for MICs that are not rank-1. In fact, for pure state inputs we can calculate Alice's *typical sharpening of expectations* by averaging the post-channel von Neumann entropy over the possible input states using the Haar measure, denoted

$\overline{S(\mathcal{E}(|\psi\rangle\langle\psi|))}$. We will see that SIC channels give the largest possible typical sharpening of expectations.

In the following we make use of a partial ordering on real vectors arranged in non-increasing order called *majorization* [15]. A real vector x rearranged into nonincreasing order is written as $x^{\downarrow}$. Then we say a vector x majorizes a vector y, denoted $x \succ y$, if all of the leading partial sums of $x^{\downarrow}$ are greater than or equal to the leading partial sums of $y^{\downarrow}$ and if the sum of all the elements of each is equal. Explicitly, $x \succ y$ if

$$\sum_{i=1}^{k} x_i^{\downarrow} \geq \sum_{i=1}^{k} y_i^{\downarrow} \ , \tag{4.36}$$

for $k = 1 \ldots N-1$ and $\sum_i x_i = \sum_i y_i$. Speaking heuristically, if $x \succ y$, then y is a flatter vector than x. A *Schur convex* function is a function f satisfying the implication $x \succ y \implies f(x) \geq f(y)$. A function is *strictly* Schur convex if the inequality is strict when $x^{\downarrow} \neq y^{\downarrow}$. When the inequality is reversed the function is called *Schur concave*.

Theorem 12. *Let $\mathcal{E}$ be an LMC. $\overline{S(\mathcal{E}(|\psi\rangle\langle\psi|))} \leq \log(d+1) - \frac{2}{d+1}\log 2$ with equality achievable if a SIC exists in dimension d.*

Proof. From Lemma 5, we know that the average maximal eigenvalue for the output of an arbitrary LMC given a pure state input is lower bounded by $\frac{2}{d+1}$. This implies

$$\begin{aligned} \overline{\lambda} &\succ \left(\overline{\lambda}_{\max}, \frac{1-\overline{\lambda}_{\max}}{d-1}, \ldots, \frac{1-\overline{\lambda}_{\max}}{d-1}\right) \\ &\succ \left(\frac{2}{d+1}, \frac{1}{d+1}, \ldots, \frac{1}{d+1}\right) \ . \end{aligned} \tag{4.37}$$

The Shannon entropy $H(P) = -\sum_i P_i \log P_i$ is a concave and Schur concave function of probability distributions. Furthermore, the von Neumann entropy of a density matrix is equal to the Shannon entropy of its eigenvalue spectrum. Using these facts we have

$$\begin{aligned}\overline{S(\mathcal{E}(|\psi\rangle\langle\psi|))} &= \overline{H\Big(\lambda(\mathcal{E}(|\psi\rangle\langle\psi|))\Big)} \\ &\le H\Big(\overline{\lambda(\mathcal{E}(|\psi\rangle\langle\psi|))}\Big) \\ &\le \log(d+1) - \frac{2}{d+1}\log 2 \ .\end{aligned} \tag{4.38}$$

If a SIC exists, taking $\mathcal{E} = \mathcal{E}_{\mathrm{SIC}}$ achieves this upper bound. □

Theorem 12 would have been more forceful if the upper bound were saturated "only if" a SIC exists, but we were unable to demonstrate this property, and so we leave it as a conjecture:

Conjecture 5. Equality is achievable in the statement of Theorem 12 only if a SIC exists in dimension d.

We were, however, able to prove a strong SIC optimality result in the setting of bipartite systems, applicable for example to Bell-test scenarios. The *entropy exchange* for a channel $\mathcal{E}$ upon input by state ρ is defined [7] to be the von Neumann entropy of the result of sending one half of a purification of ρ, $|\Psi_\rho\rangle$, through the channel:

$$S(\rho, \mathcal{E}) := S\Big(I \otimes \mathcal{E}(|\Psi_\rho\rangle\langle\Psi_\rho|)\Big) \ . \tag{4.39}$$

Theorem 13. *Let $\mathcal{E}$ be an LMC. Then $S\Big(\frac{1}{d}I, \mathcal{E}\Big) \le \log d + \frac{d-1}{d}\log(d+1)$ with equality achievable iff a SIC exists in dimension d.*

Proof. The purification of the state $\frac{1}{d}I$ is the maximally entangled state $|ME\rangle := \frac{1}{\sqrt{d}}\sum_i |ii\rangle$. Let λ be the eigenvalues of $I \otimes \mathcal{E}(|ME\rangle\langle ME|)$ arranged in nonincreasing order. We may lower bound the maximal eigenvalue as follows:

$$\begin{aligned}
\lambda_{\max} &\geq \langle ME|I\otimes\mathcal{E}(|ME\rangle\langle ME|)|ME\rangle \\
&= \frac{1}{d^2}\sum_{ijkl}\langle ii|I\otimes\mathcal{E}(|jj\rangle\langle kk|)|ll\rangle \\
&= \frac{1}{d^2}\sum_{ijkl}\langle ii|(|j\rangle\langle k|\otimes\mathcal{E}(|j\rangle\langle k|))|ll\rangle \\
&= \frac{1}{d^2}\sum_{il}\langle i|\mathcal{E}(|i\rangle\langle l|)|l\rangle \\
&= \frac{1}{d^2}\sum_{ijl}\langle i|\sqrt{E_j}|i\rangle\langle l|\sqrt{E_j}|l\rangle \\
&= \frac{1}{d^2}\sum_{j}\left(\mathrm{tr}\sqrt{E_j}\right)^2 \geq \frac{1}{d^2}\sum_j e_j = \frac{1}{d}\,.
\end{aligned}\tag{4.40}$$

Thus,

$$\begin{aligned}
\lambda &\succ \left(\lambda_{\max}, \frac{1-\lambda_{\max}}{d^2-1}, \ldots, \frac{1-\lambda_{\max}}{d^2-1}\right) \\
&\succ \left(\frac{1}{d}, \frac{1}{d(d+1)}, \ldots, \frac{1}{d(d+1)}\right)\,.
\end{aligned}\tag{4.41}$$

The upper bound now follows from the Schur concavity of von Neumann entropy.

If a SIC exists, it is easy to verify that

$$I\otimes\mathcal{E}_{\mathrm{SIC}}(|ME\rangle\langle ME|) = \frac{\frac{1}{d}I\otimes I + |ME\rangle\langle ME|}{d+1} \tag{4.42}$$

which saturates the upper bound. Von Neumann entropy is strictly Schur concave [16], so the upper bound is saturated iff $\lambda = \left(\frac{1}{d}, \frac{1}{d(d+1)}, \ldots, \frac{1}{d(d+1)}\right)$. Equation (4.40) shows that $|ME\rangle$ is the maximal eigenstate and that $\{E_j\}$ is a rank-1 MIC in the same way as in

Theorem 11. By the spectral decomposition, we may write

$$I \otimes \mathcal{E}(|ME\rangle\langle ME|) = \frac{1}{d}|ME\rangle\langle ME| + \frac{1}{d(d+1)}\sum_{i=2}^{d^2} P_i \tag{4.43}$$

where P_i are projectors into the other d^2-1 eigenstates. As the full set of projectors forms a resolution of the identity, we have

$$\sum_{i=2}^{d^2} P_i = I \otimes I - |ME\rangle\langle ME| \ , \tag{4.44}$$

so

$$I \otimes \mathcal{E}(|ME\rangle\langle ME|) = \frac{\frac{1}{d}I \otimes I + |ME\rangle\langle ME|}{d+1} \ . \tag{4.45}$$

It follows from (4.57) in Appendix A that

$$|ME\rangle\langle ME| = \frac{d+1}{d^2}\sum_{i=1}^{d^2} Q_i^T \otimes Q_i - \frac{1}{d} I \otimes I \ , \tag{4.46}$$

where the Q_i are elements of a quasi-SIC. From the previous expression we now have

$$I \otimes \mathcal{E}(|ME\rangle\langle ME|) = \frac{1}{d^2}\sum_i Q_i^T \otimes Q_i \ . \tag{4.47}$$

Applying $I \otimes \mathcal{E}$ directly to equation (4.46) gives us

$$\begin{aligned} I \otimes \mathcal{E}(|ME\rangle\langle ME|) &= \frac{d+1}{d^2}\sum_i Q_i^T \otimes \mathcal{E}(Q_i) - \frac{1}{d} I \otimes I \\ &= \frac{1}{d^2}\sum_i Q_i^T \otimes [(d+1)\mathcal{E}(Q_i) - I] \ , \end{aligned} \tag{4.48}$$

where we used that every LMC is unital and that $\frac{1}{d}\sum_i Q_i^T = I$. Comparing equations (4.47) and (4.48), we may see that

$$Q_i = (d+1)\mathcal{E}(Q_i) - I \tag{4.49}$$

by multiplying both sides by $\widetilde{Q}_j^T \otimes I$ and tracing over the first subsystem. The quasi-SICs form a basis for operator space, so it follows by linearity that

$$\mathcal{E}(\rho) = \frac{I+\rho}{d+1}\;, \tag{4.50}$$

and so by Theorem 11 we are done. □

4.4 Conclusions

In prior works we have emphasized the importance of MICs as a special class of measurements. The considerations of this paper developed from the idea that MICs may naturally furnish important classes of quantum channels as well. We affirmed this intuition with the introduction of LMCs which enabled us to discover several new ways in which SICs occupy a position of optimality among all MICs, supposing they exist. The appearance of additional equivalences with SIC existence plays two important roles. First, it should aid those trying to prove the SIC existence conjecture in all finite dimensions, and second, to our minds, it suggests that LMCs are a more important family of quantum channels than has been realized. We hope this work will inspire more study of LMCs and other types of channels derived from MICs not investigated here.

One example of such an alternative is a procedure where, when the agent implementing the channel applies the MIC, they *reprepare* the measured system in such a way that they ascribe a fixed quantum state to it, the choice of new state being made based on the measurement outcome. The action of such a channel is

$$\mathcal{E}(\rho) = \sum_i (\operatorname{tr} \rho E_i)\sigma_i, \tag{4.51}$$

where the states $\{\sigma_i\}$ are the new preparations applied in consequence to the measurement outcomes. Channels defined by a POVM and a set of repreparations are known as *entanglement-breaking channels* [17]. When the POVM is a MIC, we can speak of an *entanglement-breaking MIC channel* (EBMC). EBMCs coincide with LMCs for rank-1 MICs and repreparations proportional to the MIC, but not in general. While earlier work already gives some indication that SIC channels are significant among EBMCs [1], we suspect that there is much more to be discovered about EBMCs as a class.

Postscript

Due to a breakdown of our university email system, it was not until after *Physical Review A* published this article that we became aware of the preprint "Entanglement Breaking Rank" by Pandey *et al.* [18]. Their Corollary 3.3 is equivalent to our Theorem 1, albeit proved from a different starting point. They consider all channels having the same action as $\mathcal{E}_{\text{SIC}}$ and ask when those channels can be achieved using only d^2 rank-1 Kraus operators. We consider channels defined by d^2 Kraus operators (of arbitrary rank) and ask when they can have the action of the SIC channel. We regret this oversight, and we commend their paper to the reader's attention. The silver lining is that we can now say the SIC problem has attracted sufficient interest that the literature is not easy to keep up with. Moreover, our attention having been called back to this paper after an interlude thinking about other aspects of SICs, we now believe that Conjecture 1 can be proven for the special case of unbiased LMCs in $d = 2$. We now sketch the argument here.

Consider a MIC whose elements are constructed by taking the orbit of an operator under the action of a discrete group of unitaries. Such a MIC is known as *group covariant* and is necessarily unbiased. If $\{|\psi_j\rangle\}$ is a set covariant with respect to the same group as the MIC, then the post-channel states $\{\mathcal{E}(|\psi_j\rangle\langle\psi_j|)\}$ are unitarily equivalent and thus isospectral. We have the eigenvalue bound

$$\lambda_{\max}(\mathcal{E}(|\psi_0\rangle\langle\psi_0|)) \geq \langle\psi_0|\mathcal{E}(|\psi_0\rangle\langle\psi_0|)|\psi_0\rangle \ . \tag{4.52}$$

In the case that the states $\{|\psi_j\rangle\}$ comprise a SIC and the LMC is unbiased and rank-1, that is, the MIC elements take the form $E_i = \frac{1}{d}|\phi_i\rangle\langle\phi_i|$, we can evaluate this bound by transferring the unitary transformations from the LMC elements to the SIC states:

$$\lambda_{\max}(\mathcal{E}(|\psi_0\rangle\langle\psi_0|)) \geq \frac{1}{d}\sum_i |\langle\phi_0|\psi_i\rangle|^4 = \frac{2}{d+1} \ . \tag{4.53}$$

Supposing the entropic bound in Theorem 2 is saturated, then the MIC must be rank-1, and the average maximum eigenvalue is equal to $2/(d+1)$. If a SIC exists, then the discrete average of the maximum eigenvalue over the SIC-state inputs is equal to the continuous average over all pure states, because a SIC is a 2-design. Therefore, if a SIC exists and the entropic bound is saturated, then the maximum eigenvalue of each post-channel state for any SIC-state input is exactly $2/(d+1)$. In addition, the eigenvector of the post-channel state corresponding to this eigenvalue is the SIC-state input itself.

Eigenvalue information is most helpful in $d = 2$. Knowing the maximum eigenvalue fixes the only other eigenvalue, and from the above, we have the complete eigendecomposition of the post-channel state for any SIC-state input. From here, we can essentially do quantum channel tomography, fixing by linearity the action of the channel.

A careful study of Bloch-sphere geometry shows that an unbiased rank-1 MIC in $d = 2$ is necessarily group covariant, and in fact is unitarily equivalent to a MIC covariant under the Pauli group. Therefore, knowing that an LMC in $d = 2$ is unbiased and that the entropic bound is saturated, we know the MIC is group covariant, and the above argument applies.

4.5 Appendix A

Here we define and construct the quasi-SICs which furnish the Kraus operators needed in Theorem 11 and which were referenced in Theorem 13. Although SIC existence is not assured, one may always form a *quasi-SIC* in any finite dimension d. A quasi-SIC is a set of Hermitian operators obeying the same Hilbert–Schmidt inner product condition as the SIC projectors. As positivity is not demanded, it is relatively easy to construct a quasi-SIC as follows [19]. Start with an orthonormal basis for the Lie algebra $\mathfrak{su}(d)$ of traceless Hermitian operators. With the Hilbert–Schmidt inner product this space is a $(d^2 - 1)$-dimensional Euclidean space, so it is possible to construct a regular simplex $\{B_i\}$ consisting of d^2 normalized traceless Hermitian operators. In this case $\mathrm{tr}\, B_i B_j = \frac{-1}{d^2-1}$ when $i \neq j$. Then the operators

$$Q_i = \sqrt{\frac{d-1}{d}} B_i + \frac{1}{d} I \tag{4.54}$$

form a quasi-SIC.

It turns out that $A_i = \frac{1}{\sqrt{d}} Q_i$ give a set of Kraus operators such that

$$\mathcal{E}(\rho) = \sum_j A_j \rho A_j^\dagger = \frac{I + \rho}{d+1} , \tag{4.55}$$

or, more generally, for an arbitrary operator X,

$$\mathcal{E}(X) = \frac{(\operatorname{tr} X)I + X}{d+1} \ , \tag{4.56}$$

that is, equivalent to the action of $\mathcal{E}_{\text{SIC}}$. To see this, first observe from Corollary 1 in [20] that

$$\begin{aligned} \frac{1}{d}\sum_i Q_i \otimes Q_i^T &= \frac{2}{d+1} P_{\text{sym}}^{T_B} \\ &= \frac{1}{d+1}\left(I^{\otimes 2} + \sum_{kl} |kk\rangle\langle ll|\right) , \end{aligned} \tag{4.57}$$

where T_B indicates the partial transpose over the second subsystem. Then, with the help of the vectorized notation for an operator $|A\rangle\!\rangle := \sum_i A \otimes I |i\rangle|i\rangle$ and the identity $|BAB\rangle\!\rangle = B \otimes B^T |A\rangle\!\rangle$, we have

$$\begin{aligned} |\mathcal{E}(X)\rangle\!\rangle &= \frac{1}{d}\sum_i Q_i \otimes Q_i^T |X\rangle\!\rangle \\ &= \frac{1}{d+1}\left(I^{\otimes 2} + \sum_{kl} |kk\rangle\langle ll|\right) |X\rangle\!\rangle \\ &= \frac{1}{d+1}\left(|X\rangle\!\rangle + \sum_{klm} |k\rangle\langle l|X|m\rangle \otimes |k\rangle\langle l|I|m\rangle\right) \\ &= \frac{1}{d+1}\left(|X\rangle\!\rangle + \sum_{km} \langle m|X|m\rangle |k\rangle|k\rangle\right) \\ &= \frac{1}{d+1}\left(|X\rangle\!\rangle + |(\operatorname{tr} X)I\rangle\!\rangle\right) \\ &= \left|\frac{(\operatorname{tr} X)I + X}{d+1}\right\rangle\!\!\!\rangle , \end{aligned} \tag{4.58}$$

which is equivalent to (4.56). Sending $X = I$ through equation (4.56) reveals the identity

$$\frac{1}{d}\sum_i Q_i^2 = I \ , \tag{4.59}$$

which, since the quasi-SICs are Hermitian, is equivalent to the requirement that Kraus operators satisfy $\sum_i A_i^\dagger A_i = I$.

4.6 Appendix B

A doubly quasistochastic matrix is a matrix of real numbers whose rows and columns sum to 1. If we assume that $\{E_j\}$ is an unbiased MIC, $E_i = \frac{1}{d}\rho_i$, we will now show that U is furthermore doubly quasistochastic.

The Gram matrix for the K_i operators (4.33) is

$$\mathrm{tr}\, K_i K_j = (1/d - \gamma)\delta_{ij} + \gamma \tag{4.60}$$

where

$$\gamma = \frac{d - 1 - (d-2)\beta + 2\sqrt{(1-\beta)(1-\beta+d\beta)}}{d(d+1)} \,. \tag{4.61}$$

Since $e_i = 1/d = \mathrm{tr}\, E_i = \mathrm{tr}\, \sqrt{E_i}\sqrt{E_i}$, we have

$$\begin{aligned} \frac{1}{d} &= \sum_{jk} [U]_{ij}[U]_{ik} \mathrm{tr}\, K_j K_k \\ &= \sum_{jk} [U]_{ij}[U]_{ik} \Big((1/d - \gamma)\delta_{jk} + \gamma \Big) \\ &= (1/d - \gamma) \sum_{jk} [U]_{ij}[U]_{ik}\delta_{jk} + \gamma \left(\sum_j [U]_{ij} \right)^2 \\ &= (1/d - \gamma) + \gamma \left(\sum_j [U]_{ij} \right)^2 \,, \end{aligned} \tag{4.62}$$

from which we obtain

$$\sum_j [U]_{ij} = 1 \,. \tag{4.63}$$

Now note that

$$\operatorname{tr} K_i = \frac{\sqrt{1-\beta+d\beta} + (d-1)\sqrt{1-\beta}}{d} \ . \tag{4.64}$$

Tracing both sides of (4.34) reveals that $\operatorname{tr} \sqrt{E_i} = \operatorname{tr} K_i$ is a constant. Corollary 3 from [20] then asserts that

$$\begin{aligned} \sum_i \sqrt{E_i} &= \sum_j \left(\sum_i [U]_{ij} \right) K_j \\ &= \sqrt{d(1/d - \gamma) + d^3\gamma} I \ . \end{aligned} \tag{4.65}$$

Summing equation (4.33) gives

$$\sum_i K_i = \left(\sqrt{1-\beta+d\beta} + (d-1)\sqrt{1-\beta} \right) I \ . \tag{4.66}$$

As the K_j form a linearly independent set, combining the previous two equations requires

$$\sum_i [U]_{ij} = \frac{\sqrt{1-d\gamma+d^3\gamma}}{\sqrt{1-\beta+d\beta} + (d-1)\sqrt{1-\beta}} = 1 \ . \tag{4.67}$$

Thus U is doubly quasistochastic, as claimed.

LIST OF REFERENCES

[1] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems,” *Physical Review Research*, vol. 2, p. 013074, 2020.

[2] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Analysis and Synthesis of Minimal Informationally Complete Quantum Measurements,” *arXiv:1812.08762*, 2020.

[3] G. Zauner, *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*. PhD thesis, University of Vienna, 1999. Published in English translation: G. Zauner, "Quantum designs: foundations of a noncommutative design theory," *International Journal of Quantum Information*, vol. 9, pp. 445–508, 2011.

[4] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, “Symmetric informationally complete quantum measurements,” *Journal of Mathematical Physics*, vol. 45, pp. 2171–2180, 2004.

[5] A. J. Scott and M. Grassl, “Symmetric informationally complete positive-operator-valued measures: A new computer study,” *Journal of Mathematical Physics*, vol. 51, p. 042203, 2010.

[6] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, “The SIC question: History and state of play,” *Axioms*, vol. 6, p. 21, 2017.

[7] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, tenth anniversary ed., 2010.

[8] H. Barnum, “Information-disturbance tradeoff in quantum measurement on the uniform ensemble and on the mutually unbiased bases,” *arXiv:quant-ph/0205155*, 2002.

[9] P. Busch and P. Lahti, “Lüders Rule,” in *Compendium of Quantum Physics* (D. Greenberger, K. Hentschel, and F. Weinert, eds.), pp. 356–358, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[10] C. A. Fuchs and R. Schack, “Bayesian Conditioning, the Reflection Principle, and Quantum Decoherence,” in *Probability in Physics* (Y. Ben-Menahem and M. Hemmo, eds.), pp. 233–247, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[11] B. C. Stacey, “Quantum Theory as Symmetry Broken by Vitality,” *arXiv:1907.02432*, 2019.

[12] M. Wilde, *Quantum information theory*. Cambridge ; New York: Cambridge University Press, 2013.

[13] H. Zhu, "Permutation Symmetry Determines the Discrete Wigner Function," *Physical Review Letters*, vol. 116, no. 4, p. 040501, 2016.

[14] J. B. DeBrota and C. A. Fuchs, "Negativity bounds for Weyl–Heisenberg quasiprobability representations," *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

[15] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*. Springer Series in Statistics, Springer, 2 ed., 2011.

[16] G. M. Bosyk, S. Zozor, F. Holik, M. Portesi, and P. W. Lamberti, "A family of generalized quantum entropies: definition and properties," *Quantum Information Processing*, vol. 15, no. 8, pp. 3393–3420, 2016.

[17] M. B. Ruskai, "Qubit Entanglement Breaking Channels," *Reviews in Mathematical Physics*, vol. 15, no. 06, pp. 643–662, 2003.

[18] S. K. Pandey, V. I. Paulsen, J. Prakash, and M. Rahaman, "Entanglement breaking rank and the existence of SIC POVMs," *Journal of Mathematical Physics*, vol. 61, no. 4, p. 042203, 2020.

[19] M. Appleby, T.-Y. Chien, S. Flammia, and S. Waldron, "Constructing exact symmetric informationally complete measurements from numerical solutions," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 16, p. 165302, 2018.

[20] D. M. Appleby, C. A. Fuchs, and H. Zhu, "Group theoretic, Lie algebraic and Jordan algebraic formulations of the SIC existence problem," *Quantum Information & Computation*, vol. 15, pp. 61–94, 2015.

CHAPTER 5

THE PRINCIPAL WIGNER FUNCTION

5.1 Introduction

In the practical course of doing physics, managing coordinate systems is an important skill. It is helpful to know how to choose a basis that makes a problem as simple as possible, and it is beneficial to understand what can and cannot be eliminated by a clever choice of reference frame. “One good coordinate system may be worth more than a hundred blue-in-the-face arguments,” a colleague advises us [1]. To that end, this article will explore two particular classes of normalized operator bases and the relations between them.

We will work in the quantum theory of finite-dimensional systems familiar from the study of quantum information and computation [2]. In this theory, each physical system is associated with a Hilbert space $\mathcal{H}_d \simeq \mathbb{C}^d$, where the dimension d can be taken as a physical characteristic of the system. For example, in a quantum computer containing N *qubits,* the dimension is $d = 2^N$. A *quantum state,* the means of expressing the preparation of a system, is an operator on this Hilbert space that is positive semidefinite and has a trace of unity. Measurements that one can perform upon a system are represented mathematically as *positive-operator-valued measures* (POVMs), which are resolutions of the identity into positive semidefinite operators:

$$\sum_{i=1}^{n} E_i = I. \tag{5.1}$$

Each *effect* E_i corresponds to a possible outcome of the measurement $E = \{E_i\}$, and the probability of that outcome is calculated by the *Born Rule*:

$$p(E_i) = \mathrm{tr}\,(\rho E_i). \tag{5.2}$$

That is, probabilities are Hilbert–Schmidt inner products between the operators that stand for outcomes and for preparations (or *priors,* in more probabilistic language). If the set of effects $\{E_i\}$ spans $\mathcal{L}(\mathcal{H}_d)$, the space of linear operators on $\mathcal{H}_d$, then any operator in this space can be expressed as a list of inner products. Such a POVM is *informationally complete,* since any quantum-mechanical calculation about the system may be done in terms of these inner products. In order to be informationally complete, a POVM must span $\mathcal{L}(\mathcal{H}_d)$, so it must contain at least d^2 effects. An informationally complete POVM with exactly d^2 effects is a *minimal informationally complete* measurement, or MIC.

Because a MIC is informationally complete, it can serve as a *reference measurement.* This means that any MIC has the property that if an agent has written a probability distribution over its d^2 possible outcomes, she can then compute the probabilities that she should assign to the outcomes of any other measurement; in other words, a MIC measurement allows us to think of the Born Rule as furnishing a fully probabilistic representation of quantum theory. As we and collaborators have emphasized in the past, and will revisit below, proofs of the impossibility of probabilistic representations of quantum theory should actually be understood as proofs of the impossibility of probabilistic representations *where the probabilities combine in a particular way inspired by classicality* [3].

When using a MIC E as a reference measurement, one must be mindful of its *bias* $\{e_i\}$, composed of the *weights* $e_i := \operatorname{tr} E_i$; if the weights are equal, then $e_i = 1/d$ and we call the MIC *unbiased*. The bias of a MIC is an informational characteristic of the measurement — it is equal to d times the Born Rule probabilities for the MIC measurement given the quantum state of maximal indifference, the *garbage state*, $\rho = \frac{1}{d}I$.

Much of the community's interest in MICs has focused upon the fascinating special case of the *symmetric informationally complete* measurements, the SICs [4–7]. A SIC is an unbiased MIC where each effect is proportional to a rank-1 projector — so, each outcome of the measurement is specified by a ray in the Hilbert space — and the inner products between any two effects are constant. This latter condition is captured by the Gram matrix, that is, the matrix of inner products $[G]_{ij} := \operatorname{tr} E_i E_j$, of a SIC taking the particularly simple form

$$[G_{\rm SIC}]_{ij} = \frac{1}{d^2}\frac{d\delta_{ij}+1}{d+1}. \tag{5.3}$$

SICs have proved in many ways optimal among MICs [3, 8–12]. SICs will once again occupy a privileged position from the perspective taken in this paper.

Projective measurements (i.e., those corresponding to projections onto the eigenspaces of quantum observables) are not informationally complete; probabilities corresponding to the outcomes of such measurements only provide a partial picture of one's expectations for all possible measurements. It seems the feeling that projective measurements are nevertheless the most conceptually significant variety combined with the development of quasiprobabilistic representations with powerful phase space analogies produced in some the intuition that probability theory itself was insufficient or at least inconvenient to handle the oddities of quantum mechanics. The quantum foundations, information and computation communities have, accordingly, often looked for deviations from classical behavior in the peculiarities of quasiprobability distributions (e.g., the appearance of negativity) [13–16]. Discrete Wigner function approaches in particular have received substantial theoretical and applied interest. Our purpose in this paper is to connect reference measurement based probabilistic representations with a suitably generalized notion of discrete Wigner function representations. Our hope is that a deeper understanding of this association will be valuable to both probabilistic and quasiprobabilistic points of view and ultimately help us better understand what is possible in a quantum world.

Our plan for this paper is as follows. In §5.2, we will use the MIC concept to introduce probabilistic and quasiprobabilistic representations of quantum theory and identify the intuition that there should be a natural relation between them. In §5.3 we consider a generalization of a MIC we call a measure basis and use it to define discrete minimal Wigner bases, setting the scene for a formal discussion of the intuitive relation noted earlier. §5.4 is a brief interlude providing background and motivation from frame theory; in particular we introduce the frame operator and the canonical tight frame. With the background in place, in §5.5 we define the principal Wigner basis and derive some of its first properties, such as an induced equivalence class on the set of MICs. In §5.6 we begin to apply the perspective gained in the previous section to MICs and Wigner bases appearing in the literature. Here, we will prove Theorem 16, which bounds the distance between an unbiased MIC and an unbiased Wigner basis. Theorem 17 then captures the special role that SICs play in these considerations. In the broader picture of our research program, Theorem 17 is the key result of the paper, for it demonstrates a new way in which SICs are extremal among MICs. Finally, in §5.7 we discuss what has been learned, record a few open questions, and anticipate a few fruitful directions for further research.

5.2 Probability and Quasiprobability

As we noted above, a MIC may serve as a reference measurement. We can illustrate the meaning of this by comparing and contrasting it with classical particle mechanics. There, a "reference measurement" would just be an experiment that reads off the system's phase-space coordinates, i.e., the positions and momenta of all the particles making up the system. Any other experiment, such as observing the total kinetic energy, is in principle a coarse-graining of the information that the reference measurement itself provides.

To develop the analogy, consider the following scenario [17]. An agent Alice has a physical system of interest, and she plans to carry out either one of two different, mutually exclusive laboratory procedures upon it. In the first protocol, she will drop the system directly into a measuring apparatus and thereby obtain an outcome. In the second protocol, she will cascade her measurements, sending the system through a reference measurement and then, in the next stage, feeding it into the device from the first protocol. Probability theory in the abstract *provides no consistency conditions* between Alice's expectations for these two protocols. Different circumstances, different probabilities! Let P denote her probability assignments for the consequences of following the two-step procedure and Q those for the single-step protocol. Then, writing $\{H_i\}$ for the possible outcomes of the reference measurement and $\{D_j\}$ for those of the other,

$$P(D_j) = \sum_i P(H_i)P(D_j|H_i). \tag{5.4}$$

This much is just logic, or more specifically speaking, a consequence of Dutch-book coherence [10, 18]. It is known as the Law of Total Probability (LTP). *However,* the claim that

$$Q(D_j) = P(D_j) \tag{5.5}$$

is an assertion of *physics,* above and beyond probabilistic self-consistency. It codifies in probabilistic language the idea that the classical ideal of a reference measurement simply reads off the system's pre-existing phase-space coordinates, or data equivalent thereto.

In quantum physics, life is very different. Instead of taking a weighted average of the $\{P(D_j|H_i)\}$ as in the LTP, Alice instead uses a mapping

$$Q(D_j) = \mu\left(\{P(H_i)\}, \{P(D_j|H_i)\}\right), \tag{5.6}$$

where the exact form of the function μ depends upon her choice of MIC. Conveniently, quantum theory is only so nonclassical that μ is a bilinear form, rather than a more convoluted function.

We can write our equations more compactly by introducing a vector notation, in which omitted subscripts imply that an entire vector or matrix is being treated as an entity. Then the LTP has the expression

$$P(D) = P(D|H)P(H)\ , \tag{5.7}$$

while the quantum relation, the Born Rule, is[1]

$$Q(D) = P(D|H)\Phi P(H)\ ,\ \text{with } [\Phi^{-1}]_{ij} := \frac{1}{\operatorname{tr} H_j} \operatorname{tr} H_i H_j\ . \tag{5.8}$$

[1]To obtain (5.8), expand an arbitrary state ρ in the basis $\{\rho_i\}$, $\rho_i = H_i/h_i$, of quantum states proportional to the MIC basis: $\rho = \sum_i \alpha_i \rho_i$. Then, computing the MIC probabilities for ρ with the Born Rule, it follows that $\alpha = \Phi P(H)$. Thus, since $P(D_j|H_i) = \operatorname{tr} D_j \rho_i$, (5.8) follows from another application of the Born Rule with the D measurement.

The Born matrix Φ depends upon the MIC[2] , but it is always a column quasistochastic matrix, meaning its columns sum to one but may contain negative elements [3]. In fact, Φ *must* contain negative entries; this follows from basic structural properties of quantum theory [13]. As a consequence, $\Phi P(H)$ is a quasiprobability. Considering the operation of Φ on $P(H)$ as a single term results in an equation algebraically equivalent to the LTP aside from the appearance of negativity in the last term. The same thing happens if we regard Φ as acting to the left on $P(H|D)$, but now the negativity has been relegated to the first term. For an unbiased MIC, Φ is a symmetric matrix and we can do the same thing in yet another way by "splitting the map down the middle" and acting with one factor in either direction:

$$Q(D) = \left(P(D|H)\Phi^{1/2}\right)\left(\Phi^{1/2}P(H)\right) \ , \tag{5.9}$$

where the principal square root $\Phi^{1/2}$, which also turns out to be quasistochastic, provides a similar quasiprobabilistic interpretation to $\Phi^{1/2}P(H)$ and the rows of $P(D|H)\Phi^{1/2}$. One might then say that, for a given MIC, there is a certain "gauge freedom" about where the negativity can occur if we wish to massage (5.8) into a form which fits together in exactly the same way as the LTP [19]. As we will see later on, the even-handed gauge choice (5.9) amounts to a Wigner function representation naturally associated with the reference measurement MIC.

5.3 Discrete Minimal Wigner Functions

We now begin a more formal discussion of quasiprobability representations of quantum theory constructed on orthogonal operator bases [20, 21], which were motivated in

[2] And, in general, a set of post-measurement states; in this case the post-measurement states ρ_i are proportional to the MIC itself and so do not constitute another dependency. We considered the more general update procedure extensively in [3].

the first place by the demand that we cast the Born Rule structurally analogous to the LTP. We first define the following generalization of a MIC:

Definition 1. A *measure basis* is a Hermitian basis L for $\mathcal{L}(\mathcal{H}_d)$ for which

$$\sum_i L_i = I \quad \text{and} \quad \operatorname{tr} L_i \geq 0 \; . \tag{5.10}$$

The sum condition ensures that its inner products with a quantum state give a quasiprobability in the same way the Born Rule gives probabilities for a POVM. As with a MIC, the bias $\{l_i\}$ of a measure basis consists of the weights $l_i := \operatorname{tr} L_i$. The condition that the weights are nonnegative permits us to carry over the probabilistic significance of a MIC's bias to a measure basis in general. It will also sometimes be useful to define the diagonal bias matrix $[A]_{ij} := l_i\delta_{ij}$ of a measure basis. We extend the Born matrix definition to any measure basis in the natural way: In terms of its Gram matrix and bias matrix, $\Phi = AG^{-1}$. Unless otherwise specified, a measure basis has a bias with weights denoted by the lowercase letter equivalent of the basis elements, e.g. the bias of B is $\{b_i\}$. In this language, a MIC is a measure basis consisting of positive semidefinite operators.

A MIC cannot be orthogonal [22], but a measure basis can — this is the distinguishing property of the operator bases corresponding to discrete Wigner functions which presently concerns us. Accordingly, we define

Definition 2. A *discrete minimal Wigner basis* is an orthogonal measure basis.

We will generally omit the additional qualifiers "discrete" and "minimal" since we are only considering finite dimensions and we will not be discussing quasiprobability representations of quantum theory that use overcomplete operator bases.[3] Up to proportionality, what we call a Wigner basis is the set of "phase point operators" for most Wigner

[3]For a generalization of Wigner functions which does use overcomplete bases, see [23].

function approaches in the literature. We prefer to use phase space agnostic terminology as we have intentionally avoided building any particular conception of a discrete phase space into our generalization. This move will hopefully allow us to eventually understand precisely when, and to what extent, such conceptions are most useful. For the remainder of the paper, when we speak of a Wigner function, we mean the quasiprobability distribution one may obtain from a Wigner basis and a density matrix. There is not much to say about Wigner bases at this level of generality, so for now we note two basic properties before proceeding.

Proposition 2. *The Gram matrix of a Wigner basis is equal to its bias matrix,* $[G]_{ij} = f_i\delta_{ij}$.

Proof. Let $\{F_i\}$ be a Wigner basis. As it is orthogonal, the Gram matrix is diagonal, that is, $\operatorname{tr} F_iF_j = c_i\delta_{ij}$ for some constants $\{c_i\}$. The dual basis is the unique basis $\{\widetilde{F}_i\}$ such that $\operatorname{tr} \widetilde{F}_iF_j = \delta_{ij}$, so we must have $F_i = c_i\widetilde{F}_i$. Then the sum constraint enforces $\operatorname{tr} \widetilde{F}_i = 1$ and $c_i = f_i$ follows. □

Definition 3. Let F be a Wigner basis. The *shifted Wigner basis* of F is the set F^{S},

$$F_i^{\mathrm{S}} := -F_i + \frac{2f_i}{d}I \; . \tag{5.11}$$

Proposition 3. *The shifted Wigner basis is a Wigner basis with the same bias.*

Proof. $\operatorname{tr} F_i^{\mathrm{S}} = f_i$ and $\sum_i F_i^{\mathrm{S}} = I$ are obvious from inspection. Then

$$\operatorname{tr}\left(-F_i + \frac{2f_i}{d}I\right)\left(-F_j + \frac{2f_j}{d}I\right) = \operatorname{tr} F_iF_j - \frac{2f_j}{d}\operatorname{tr} F_i - \frac{2f_i}{d}\operatorname{tr} F_j + \frac{4f_if_j}{d^2}\operatorname{tr} I = f_i\delta_{ij} \tag{5.12}$$

demonstrates orthogonality. □

5.4 The Frame Operator and the Canonical Tight Frame

In addition to the Gram matrix, another important operator for classifying sets of vectors is the *frame operator* which appears in the theory of frames. For an introduction to frame theory, we recommend the reference [24]. In the finite dimensional setting, a frame for a Hilbert space is a spanning set of vectors; a basis is a frame with the minimal number of vectors to be a spanning set. Specializing to $\mathcal{L}(\mathcal{H}_d)$, we define

Definition 4. The *frame operator* of a frame L is the superoperator $S : \mathcal{L}(\mathcal{H}_d) \to \mathcal{L}(\mathcal{H}_d)$ defined by

$$X \mapsto \sum_i (\operatorname{tr} XL_i)L_i \ , \qquad \forall X \in \mathcal{L}(\mathcal{H}_d) \ . \tag{5.13}$$

The frame operator is self-adjoint with respect to the Hilbert–Schmidt inner product and has the same nonzero spectrum as the Gram matrix; for a basis, such as a MIC, they are isospectral. By inspection of the definition, one sees that the frame operator allows us to construct and interconvert between the frame and the *dual*, $\widetilde{L}_i := S^{-1}(L_i)$; for a basis, the dual frame is exactly the dual basis. It is not hard to show that the inverse of the frame operator is the frame operator of the dual. Likewise, for bases, the Gram matrix is also invertible and its inverse is given by the Gram matrix of the dual basis

$$[G^{-1}]_{ij} = \operatorname{tr} \widetilde{L}_i \widetilde{L}_j \ . \tag{5.14}$$

We will also make use of the fact that the inner products of a vector with a frame give the expansion coefficients for expressing the vector in the dual frame:

$$X = \sum_i (\operatorname{tr} X\widetilde{L}_i)L_i = \sum_i (\operatorname{tr} XL_i)\widetilde{L}_i \ , \qquad \forall X \in \mathcal{L}(\mathcal{H}_d) \ . \tag{5.15}$$

When the frame operator for a set of vectors is proportional to the identity, a frame is called *tight*; and when the constant of proportionality is unity, a tight frame is called *normalized*. For all frames there is a naturally associated normalized tight frame known as the *canonical tight frame* which is "halfway" to the dual, obtained by applying the inverse square root of the frame operator to each vector. Intuitively, halfway between a basis and its dual is an orthonormal basis, that is, a self-dual basis. Indeed, for bases, the canonical tight frame construction is an orthogonalization procedure corresponding to a symmetric version of the Gram–Schmidt algorithm [24]. The canonical tight frame is distinguished by being the *closest* tight frame to the original frame in the sense of minimizing the least squares error (Theorem 3.2 in [24]).

A MIC cannot be a tight frame[4] — if its frame operator were proportional to the identity, its Gram matrix would have to be as well and a MIC cannot be an orthogonal basis. For the same reason, an unbiased Wigner basis *is* a tight frame. The canonical tight frame for a MIC $\{E_i\}$ is the orthonormal operator basis $\{S^{-1/2}(E_i)\}$. This cannot be a measure basis because the sum condition fails for a normalized operator basis, but if the MIC is unbiased, $1/\sqrt{d}$ times the canonical tight frame is an unbiased Wigner basis. In what follows, we will extend this observation to MICs and Wigner bases of arbitrary bias and begin to explore the consequences.

[4]This is a good place to clear up a confusion that we have encountered a few times during conferences, when people from different subfields try to communicate. A MIC is a basis for the d^2-dimensional space $\mathcal{L}(\mathcal{H}_d)$. It is not overcomplete, but exactly complete, having just the right number of elements to span the operator space while keeping itself a linearly-independent set. A rank-1 MIC, for which $E_i = e_i|\psi_i\rangle\langle\psi_i|$, is specified by a set of weights $\{e_i\}$ and a set of vectors $\{|\psi_i\rangle\}$. These vectors are d^2 in number and live within $\mathcal{H}_d$, so for *that* space, they would be overcomplete. The vectors $\{\sqrt{e_i}|\psi_i\rangle\}$ may be considered a frame for the d-dimensional space $\mathcal{H}_d$, and, in fact, this is a normalized tight frame because the frame operator and POVM sum condition coincide: $S = \sum_i e_i|\psi_i\rangle\langle\psi_i| = I$. As an operator basis, however, a MIC is never a tight frame.

5.5 The Principal Wigner Basis

In this section we define the principal Wigner basis which is the most significant association of a Wigner basis to a given MIC. We will see how the principal Wigner basis induces an equivalence class among measure bases and MICs more specifically. We then begin to study the structure of that class.

An arbitrary Wigner basis is not a tight frame; for bases this is because the tight frame concept effectively demands orthogonality *and* that the vectors all have the same norm. From Proposition 2, we see that the Hilbert–Schmidt norm of the Wigner basis elements are the square roots of the weights. So for any Wigner basis, if we divide each element by the square root of its weight, the result is a normalized tight frame. We would like to think of Wigner bases playing the analogous role among measure bases that the canonical tight frame plays among frames. This motivates the following modification of the frame operator:

Definition 5. The *rescaled frame operator* $\mathcal{S}_L$ of a measure basis L is the frame operator of the *rescaled basis* $\left\{\frac{1}{\sqrt{l_i}}L_i\right\}$. For $X \in \mathcal{L}(\mathcal{H}_d)$,

$$\mathcal{S}_L(X) = \sum_j \left(\operatorname{tr} X \frac{L_j}{\sqrt{l_j}}\right) \frac{L_j}{\sqrt{l_j}} = \sum_j \left(\frac{\operatorname{tr} X L_j}{l_j}\right) L_j \ . \tag{5.16}$$

For all measure bases, $\mathcal{S}_L(I) = I$ and, consequently, $\operatorname{tr} \mathcal{S}_L(X) = \operatorname{tr} \mathcal{S}_L(X) I = \operatorname{tr} X \mathcal{S}_L(I) = \operatorname{tr} X$. In accordance with the motivation above, $\mathcal{S}_F$ is the identity superoperator for a Wigner basis F, which follows from $F_i = f_i \widetilde{F}_i$. For a MIC E, $\mathcal{S}_E$ is a trace-preserving quantum channel we called the entanglement breaking MIC channel (EBMC) in a previous paper; an EBMC for a rank-1 MIC is also the Lüders MIC channel for that MIC [25]. It also turns out that the rescaled frame operator for a MIC has independently arisen in a different context pertaining to quantum state tomography, suggesting a deeper significance [26]. When acting on one of the MIC elements we have

$$\mathcal{S}_E(E_i) = \sum_j \left(\frac{\operatorname{tr} E_i E_j}{e_j}\right) E_j = \sum_j [\Phi^{-1}]_{ij} E_j \ , \tag{5.17}$$

from which it follows that

$$\mathcal{S}_E^{-1}(E_i) = \sum_j [\Phi]_{ij} E_j \ . \tag{5.18}$$

The action of the Born matrix on the reference probability vector is thus equivalent to the inverse rescaled frame operator on the effects. It is easy to see that $\mathcal{S}_E^{-1}(E_i)$ is the dual basis element to the state ρ_i, so the expression (5.8) simply follows from inserting the identity into the Born Rule:

$$Q(D_j) = \operatorname{tr} D_j \rho = \sum_i (\operatorname{tr} D_j \rho_i) \left(\operatorname{tr} \mathcal{S}_E^{-1}(E_i)\rho\right) \ . \tag{5.19}$$

This lets us understand the first type of quasiprobability grouping we discussed; $\operatorname{tr} \mathcal{S}_E^{-1}(E_i)\rho$ is the ith element of the quasiprobability vector $\Phi P(E)$. We are now in place to see that the even-handed gauge choice discussed above corresponds to a Wigner function.

Definition 6. The *principal Wigner basis* of a measure basis L is the set $PW(L) = \{F_i\}$,

$$F_i := \mathcal{S}_L^{-1/2}(L_i) \ . \tag{5.20}$$

Proposition 4. *For a measure basis L, $PW(L)$ is a Wigner basis with the same bias as L.*

Proof. As the identity is an eigenvector with eigenvalue 1 of $\mathcal{S}_L$, this will also be true of $\mathcal{S}_L^{-1/2}$, so

$$\sum_i F_i = \sum_i \mathcal{S}_L^{-1/2}(L_i) = \mathcal{S}_L^{-1/2}(I) = I \ . \tag{5.21}$$

The bias is preserved because $\mathcal{S}_L^{-1/2}$ is trace-preserving. Orthogonality follows from the self-adjointness of the rescaled frame operator and the relation between a basis and its dual:

$$\operatorname{tr} F_i F_j = \operatorname{tr}\left(\mathcal{S}_L^{-1/2}(L_i)\mathcal{S}_L^{-1/2}(L_j)\right) = \operatorname{tr} L_i \mathcal{S}_L^{-1}(L_j) = \sqrt{l_i l_j}\operatorname{tr}\left(\frac{L_i}{\sqrt{l_i}}\mathcal{S}_L^{-1}\left(\frac{L_j}{\sqrt{l_j}}\right)\right) = l_i\delta_{ij} \ . \tag{5.22}$$

□

Proposition 5. *Let L be a measure basis with Gram matrix G and bias matrix A. The square root of the Born matrix with all positive eigenvalues,*

$$\sqrt{\Phi} := A^{1/2}\left(A^{1/2}G^{-1}A^{1/2}\right)^{1/2} A^{-1/2} \ , \tag{5.23}$$

gives the expansion coefficients of the principal Wigner basis in the measure basis:

$$F_i = \sum_j [\sqrt{\Phi}]_{ij} L_j \ . \tag{5.24}$$

Proof. Let D be the rescaled basis of L. Since $F_i = \mathcal{S}_L^{-1/2}(L_i) = \sum_j (\operatorname{tr}\mathcal{S}_L^{-1/2}(L_i)\widetilde{L}_j)L_j$, we must show $[\sqrt{\Phi}]_{ij} = \operatorname{tr}\mathcal{S}_L^{-1/2}(L_i)\widetilde{L}_j$. Note that

$$[A^{1/2}G^{-1}A^{1/2}]_{ij} = \operatorname{tr}\widetilde{D}_i\widetilde{D}_j = \operatorname{tr}\mathcal{S}_L^{-1}(D_i)\widetilde{D}_j \implies \left[\left(A^{1/2}G^{-1}A^{1/2}\right)^{1/2}\right]_{ij} = \operatorname{tr}\mathcal{S}_L^{-1/2}(D_i)\widetilde{D}_j \ . \tag{5.25}$$

Thus

$$\begin{aligned} [\sqrt{\Phi}]_{ij} &= \left[A^{1/2}\left(A^{1/2}G^{-1}A^{1/2}\right)^{1/2} A^{-1/2}\right]_{ij} = \sqrt{\frac{l_i}{l_j}}\left[\left(A^{1/2}G^{-1}A^{1/2}\right)^{1/2}\right]_{ij} \\ &= \sqrt{\frac{l_i}{l_j}}\operatorname{tr}\mathcal{S}_L^{-1/2}(D_i)\widetilde{D}_j = \operatorname{tr}\mathcal{S}_L^{-1/2}(L_i)\widetilde{L}_j \ , \end{aligned} \tag{5.26}$$

where we used the fact that $\widetilde{D}_j = \sqrt{l_j}\widetilde{L}_j$. $\square$

Applied to MICs, Proposition 4 demonstrates that the principal Wigner basis is in fact a Wigner basis and Proposition 5 connects it to the probabilistic representations discussed in §5.2; given any MIC reference measurement, there is an associated Wigner function formed by dividing the effect of the Born matrix equally among the probabilities and quasiprobabilities in (5.8).

The principal Wigner basis map satisfies several of the properties one would desire for a "principal" definition. Global unitary conjugation acts equally on a MIC and its principal Wigner basis; for any unitary U, if $\{F_i\}$ is the principal Wigner basis of the MIC $\{E_i\}$, $\{UF_iU^\dagger\}$ is the principal Wigner basis of the MIC $\{UE_iU^\dagger\}$. This follows from the invariance of the Gram matrix under a global unitary conjugation of the MIC. As a consequence, if a MIC is *group covariant*, i.e., if it can be produced by taking the orbit of an initial element under the action of a group, thereby making d^2 elements out of one, its principal Wigner basis is group covariant with respect to the same group. The principal Wigner basis also respects tensor products in the following way. Let D and E be MICs, not necessarily for the same dimensional Hilbert space. The elementwise tensor products of their elements, $\{D_i \otimes E_j\}$, forms a MIC $D \otimes E$ for the product dimension. The principal Wigner basis for this MIC is equal to the tensor product of the principal Wigner bases for the constituent MICs, that is, $PW(D \otimes E) = PW(D) \otimes PW(E)$. This follows from the observation that the tensor product of the Born matrices for the constituent MICs is equal to the Born matrix of the tensor product MIC.

Definition 6 furnishes a map from any MIC to a particular Wigner basis. But which MICs are in the preimage of a particular Wigner basis? As one might expect from an orthogonalization procedure, there are infinitely many MICs which share a principal Wigner basis. We group those that do into an equivalence class.

Definition 7. If $PW(L) = PW(M)$ for two distinct measure bases L and M, we say L and M are *Wigner equivalent*, denoted $L \sim_{\mathrm{W}} M$.

What can we say about the Wigner equivalence classes? The first thing to realize is that the effects of a MIC of a given bias can have different norms depending on the purity of ρ_i while, by contrast, for a Wigner basis, the norm of the elements is fixed by the bias. This inspires us to guess that MICs which are "in the same direction" will be Wigner equivalent. The following few results will make this intuition precise.

Proposition 6. *Given a measure basis L and a real parameter $t \neq 0$, the set $L^t = \{L_i^t\}$,*

$$L_i^t := tL_i + (1-t)\frac{l_i}{d}I \tag{5.27}$$

is a measure basis with the same bias as L.

Proof. $\sum_i L_i^t = I$ and $\operatorname{tr} L_i^t = l_i$ are obvious. What remains is to prove linear independence. Suppose there is a set $\{\beta_i\}$ not all zero for which $\sum_i \beta_i L_i^t = 0$. Then

$$\sum_i \beta_i L_i^t = t\sum_i \beta_i L_i + \frac{(1-t)}{d}\left(\sum_i \beta_i l_i\right) I = 0 \implies \sum_i \beta_i L_i = \sum_i \left(\frac{t-1}{td}\sum_j \beta_j l_j\right) L_i \ , \tag{5.28}$$

where we used the fact that $\sum_i L_i = I$. Because $\{L_i\}$ is a basis, the expansion coefficients are unique, so $\beta_i = \frac{t-1}{td}\sum_j \beta_j l_j$ for all i, that is, the β_i are constant, say $\beta_i = \beta$. Then $\sum_i \beta_i L_i^t = \beta \sum_i L_i^t = \beta I = 0 \implies \beta = 0$, but this is a contradiction, so L^t is a measure basis. $\square$

Definition 8. Let L be a measure basis. Any measure basis L^t, defined by (5.27), is *collinear* with L. If $t > 0$, it is *parallel* to L and if $t < 0$, it is *antiparallel* to L.

Now we can prove that, in fact, all parallel measure bases are Wigner equivalent. Let $SPW(L)$ denote the shifted principal Wigner basis of a measure basis L.

Theorem 14. *For any measure basis L,*

$$PW(L^t) = \begin{cases} PW(L), & \textit{if}\, t > 0 \\ SPW(L), & \textit{if}\, t < 0 \end{cases} . \tag{5.29}$$

Proof. Because the relation between L and L^t is relatively simple, their Born matrices $\Phi = AG^{-1}$ also end up simply related:

$$\Phi^t = \frac{1}{t^2}\Phi + \left(1 - \frac{1}{t^2}\right)\frac{1}{d}AJ \; . \tag{5.30}$$

In particular note that Φ^t for two t values with the same magnitude but opposite signs are equal. Φ is column quasistochastic, so the left eigenvector with eigenvalue 1 is the all 1s vector; it is easy to show that the corresponding right eigenvector is the vector of weights. $\sqrt{\Phi}$ shares these properties and because of them one may deduce that

$$\sqrt{\Phi}AJ = AJ\sqrt{\Phi} = AJ \; . \tag{5.31}$$

From this, it is easy to verify that

$$\sqrt{\Phi^t} = \frac{1}{|t|}\sqrt{\Phi} + \left(1 - \frac{1}{|t|}\right)\frac{1}{d}AJ \; , \tag{5.32}$$

where $|t|$ is the absolute value of t. Now we may use Proposition 5 to compute the principal Wigner basis for L^t:

$$\begin{aligned}
F_i^t &= \sum_j [\sqrt{\Phi^t}]_{ij} L_j^t = \sum_j \left[\frac{1}{|t|}\sqrt{\Phi} + \left(1 - \frac{1}{|t|}\right)\frac{1}{d}AJ\right]_{ij} \left(tL_j + (1-t)\frac{l_j}{d}I\right) \\
&= \frac{t}{|t|}\sum_j [\sqrt{\Phi}]_{ij} L_j + \frac{1-t}{|t|d}\sum_j [\sqrt{\Phi}]_{ij} l_j I + \frac{t}{d}\left(1 - \frac{1}{|t|}\right)\sum_j [AJ]_{ij} L_j + \left(1 - \frac{1}{|t|}\right)\frac{1-t}{d^2}\sum_j [AJ]_{ij} l_j I \\
&= \frac{t}{|t|}F_i + \left[\frac{1-t}{|t|} + t\left(1 - \frac{1}{|t|}\right) + (1-t)\left(1 - \frac{1}{|t|}\right)\right]\frac{l_i}{d}I = \frac{t}{|t|}F_i + \left(1 - \frac{t}{|t|}\right)\frac{l_i}{d}I\ ,
\end{aligned} \tag{5.33}$$

from which the claim follows. □

Since all parallel measure bases are Wigner equivalent, all parallel MICs are as well. Conveniently, we have also learned that the shifted principal Wigner basis is the principal Wigner basis of the antiparallel MICs. As the next lemma shows, there is always an interval of MICs collinear with any measure basis. This interval corresponds to the intersection of collinear measure bases with the cone of positive semidefinite operators.

Lemma 6. *Let L be a measure basis. Define $\sigma_i := L_i/l_i$. Let $\lambda_i^{\max}$ and $\lambda_i^{\min}$ be the maximum and minimum eigenvalues of σ_i. Then L^t is a MIC iff $t \neq 0$ and*

$$\max_i \left\{\frac{1}{1 - d\lambda_i^{\max}}\right\} \leq t \leq \min_i \left\{\frac{1}{1 - d\lambda_i^{\min}}\right\}\ . \tag{5.34}$$

Proof. As L^t is a measure basis, for it to be a MIC the elements must all be positive semidefinite. Thus we need to show $\lambda_{\min}(L_i^t) \geq 0$ for all t in the range (5.34). Since $\operatorname{tr} L_i = l_i$, the average eigenvalue of L_i is l_i/d, and so $\lambda_{\min}(L_i) \leq l_i/d$. Suppose $t > 0$. Then

$$\lambda_j(L_i^t) = t\lambda_j(L_i) + (1-t)\frac{l_i}{d}\ , \tag{5.35}$$

so

$$\lambda_{\min}(L_i^t) = t\lambda_{\min}(L_i) + (1-t)\frac{l_i}{d} \geq 0 \iff t \leq \frac{1}{1 - d\lambda_i^{\min}} \,. \tag{5.36}$$

For this to be true for all i, t must be less than the minimum value of the right hand side. The lower bound is similarly obtained by assuming $t < 0$ and making the appropriate adjustments. $\square$

Starting with a particular MIC, we were able to see that all those parallel to it were Wigner equivalent. But there are Wigner equivalent MICs which are not collinear; in fact, the following theorem opens the way for the construction of arbitrarily many Wigner equivalent MICs which are not collinear.

Theorem 15. *Let F be a Wigner basis. For any measure basis L, $\mathcal{S}_L^{1/2}(F)$ is a measure basis collinear with MICs in the Wigner equivance class of F.*

Proof. Note $\operatorname{tr}\mathcal{S}_L^{1/2}(F_i) = \operatorname{tr}\mathcal{S}_L^{1/2}(F_i)I = \operatorname{tr} F_i\mathcal{S}_L^{1/2}(I) = \operatorname{tr} F_i = f_i$. Now from Definition 5 and the proportionality of a Wigner basis to its dual it follows that

$$\mathcal{S}_{\mathcal{S}_L^{1/2}(F)}(X) = \sum_j \left(\frac{\operatorname{tr}\mathcal{S}_L^{1/2}(F_j)X}{\operatorname{tr}\mathcal{S}_L^{1/2}(F_j)}\right)\mathcal{S}_L^{1/2}(F_j) = \mathcal{S}_L^{1/2}\left(\sum_j \left(\frac{\operatorname{tr} F_j\mathcal{S}_L^{1/2}(X)}{f_j}\right)F_j\right) = \mathcal{S}_L(X) \,. \tag{5.37}$$

From this we see that $PW(\mathcal{S}_L^{1/2}(F)) = F$. Then from Lemma 6 we may construct MICs collinear with $\mathcal{S}_L^{1/2}(F)$ in the Wigner equivalence class of F. $\square$

We have likely only just begun to scratch the surface of this topic, but we have seen enough to begin applying what we've learned. Before doing so in the next section we remark on the significance of our observations. Wigner equivalence is a very broad grouping. Every MIC furnishes a distinct probabilistic representation of the Born Rule, but the probabilistic representations obtained from any two Wigner equivalent MICs are related to the same quasiprobabilistic representation in the way we have seen. In other words, a Wigner function representation of quantum theory is insensitive to differences among the informationally complete measurements consistent with it. This is at least in part an artifact of our level of treatment. For our purposes here, any MIC sufficed to furnish a probabilistic representation, but this is not to say that any MIC would be practically useful in that capacity. Wigner equivalence alone preserves bias, but, it seems, little else. A good reference measurement should unmask features of quantum state space that were otherwise hard to detect. It seems likely that the reference measurements standing a chance of being useful in this regard would possess mathematical properties beyond the basic definition of a MIC, perhaps being rank-1 or symmetric with respect to a particular group. Indeed, as the phase point operators of most Wigner function approaches are covariant with respect to the Weyl–Heisenberg group, furnishing discrete analogs of position and momentum operators, we speculate that Weyl–Heisenberg covariant MICs are a class of reference measurements which provide something reminiscent of a phase space without sacrificing operational significance. We hope the general association we've identified will allow for a dramatic sharpening of the correspondence between probabilistic and quasiprobabilistic representations once refinements of this variety and others are adopted.

5.6 Behind Every Great Wigner Basis is a Great MIC

The previous section introduced the principal Wigner basis and the induced relation of Wigner equivalence. We can carry this study further by specializing to *unbiased* Wigner bases, the case of most practical and theoretical interest as, to our knowledge, it encompasses all of the Wigner functions derived from operator bases to date in the literature. Pursuing this line of inquiry will lead us to a new way in which SICs are extremal among MICs. We begin by showing that the principal Wigner basis of an unbiased MIC inherits the closeness property enjoyed by the canonical tight frame. Because of the sum normalization condition, there is also a *farthest* Wigner basis from a given unbiased MIC, namely the shifted principal Wigner basis.

Theorem 16. *Let E be an unbiased MIC and F be an unbiased Wigner basis. Let λ_k be the kth eigenvalue of the MIC's frame operator S. Then*

$$\sum_k \left(\sqrt{\lambda_k} - \sqrt{1/d}\right)^2 \leq \sum_i \|E_i - F_i\|^2 \leq \sum_k \left(\sqrt{\lambda_k} + \sqrt{1/d}\right)^2 - \frac{4}{d}\,, \tag{5.38}$$

where the lower bound is saturated iff $F = PW(E)$ and the upper bound is saturated iff $F = SPW(E)$.

Proof. The lower bound follows from Theorem 3.2 in [24]. The key step is the demonstration that for any MIC E and Wigner basis F,

$$\sum_j \operatorname{tr} E_j F_j \leq \frac{1}{\sqrt{d}} \sum_k \sqrt{\lambda_k}\,, \tag{5.39}$$

with equality iff

$$F_j = \frac{1}{\sqrt{d}} S^{-1/2}(E_j)\,, \tag{5.40}$$

that is, $F = PW(E)$. For the upper bound,

$$\begin{aligned}\sum_i \|E_i - F_i\|^2 &= \sum_i \left\| E_i + F_i^{\mathrm{S}} - \frac{2}{d^2} I \right\|^2 \\ &= \sum_i \operatorname{tr}\Big(E_i + F_i^{\mathrm{S}} - \frac{2}{d^2} I \Big)\Big(E_i + F_i^{\mathrm{S}} - \frac{2}{d^2} I \Big) \\ &= \sum_i \left(\operatorname{tr} E_i^2 + \operatorname{tr} F_i^{\mathrm{S}2} + 2\operatorname{tr} E_i F_i^{\mathrm{S}2} - \frac{4}{d^3} \right) \\ &\le \sum_k \left(\lambda_k + \frac{1}{d} + \frac{2}{\sqrt{d}} \sqrt{\lambda_k} \right) - \frac{4}{d} \\ &= \sum_k \left(\sqrt{\lambda_k} + \sqrt{1/d} \right)^2 - \frac{4}{d} \,, \end{aligned} \tag{5.41}$$

with equality iff $F^{\mathrm{S}} = PW(E)$, that is, iff $F = SPW(E)$. □

Theorem 16 does not seem to exactly generalize to the biased case; perhaps a different condition is more appropriate for biased measure bases. However, this theorem inspires us to ask the reverse question: What is the closest MIC to a given Wigner basis? This is probably quite hard to answer in general, but it turns out we can answer it in an important special case to which we now turn.

Let E be a SIC, that is,

$$E_i := \frac{1}{d}\Pi_i, \text{ with } \operatorname{tr} \Pi_i \Pi_j = \frac{d\delta_{ij} + 1}{d+1} \,. \tag{5.42}$$

One may calculate

$$[\sqrt{\Phi}_{\mathrm{SIC}}]_{ij} = \frac{1}{\sqrt{d}} [G_{\mathrm{SIC}}^{-1/2}]_{ij} = \sqrt{d+1}\delta_{ij} + \frac{1}{d^2}(1 - \sqrt{d+1}) \,, \tag{5.43}$$

which gives the following principal and shifted principal Wigner basis:

$$F_j = \frac{1}{d}\left(\sqrt{d+1}\right)\Pi_j + \frac{1}{d^2}\left(1-\sqrt{d+1}\right)I \quad \text{and } F_j^{\mathrm{S}} = -\frac{1}{d}\left(\sqrt{d+1}\right)\Pi_j + \frac{1}{d^2}\left(1+\sqrt{d+1}\right)I\ . \tag{5.44}$$

As the next theorem demonstrates, the smallest *and* largest that the bounds in Theorem 16 can be both occur when the unbiased MIC is a SIC. The minimal lower bound result complements the prior discovery that SICs are the closest MICs can come to being orthogonal bases [11]. Consequently, the next theorem supports the intuition that SICs are the natural analogues of orthogonal operator bases contained within the cone of positive semidefinite operators.

Theorem 17. *Let $\{E_i\}$ be an unbiased MIC and $\{F_j\}$ be an unbiased Wigner basis. Then*

$$\frac{d-1}{d}\left(d+2-2\sqrt{d+1}\right) \leq \sum_i \|E_i - F_i\|^2 \leq \frac{d-1}{d}\left(d+2+2\sqrt{d+1}\right)\ , \tag{5.45}$$

where the lower bound is saturated iff $\{E_i\}$ is a SIC and $\{F_j\}$ is its principal Wigner basis and the upper bound is saturated iff $\{E_i\}$ is a SIC and $\{F_j\}$ is its shifted principal Wigner basis.

Proof. The matrix dG is doubly stochastic for an unbiased MIC, so the maximal eigenvalue of S is always $1/d$. Furthermore, because the diagonal entries of an unbiased MIC's Gram matrix are bounded above by $1/d^2$, we also know that $\operatorname{tr} S \leq 1$. It is then straightforward to perform a constrained optimization to see that the bounds in (5.38) achieve their extreme values when

$$\lambda = \left(\frac{1}{d}, \frac{1}{d(d+1)}, \ldots, \frac{1}{d(d+1)}\right)\ . \tag{5.46}$$

Plugging this spectrum in to (5.38) gives the upper and lower bounds in (5.45). Such a spectrum occurs iff the MIC is a SIC, a fact that is easy to derive from Lemma 1 in [3].

□

The Wigner bases (6.13) were identified by Zhu [20] for a different reason.[5] Given a Wigner basis, the *ceiling negativity* of a quantum state ρ is the magnitude of the most negative entry in the quasiprobability vector that represents ρ. Maximizing the ceiling negativity over all quantum states yields the ceiling negativity of the Wigner basis. Zhu proved that the principal and shifted principal Wigner bases associated with a SIC provide, respectively, the *lower and upper bounds* on the ceiling negativity over all unbiased Wigner bases in dimension d. Our orthogonalization procedure sets Zhu's result in a broader conceptual context: Zhu's Wigner bases are the output of applying to a SIC a procedure that works for any MIC. Our quite general definition of a Wigner basis was partly inspired by Zhu's approach. His relaxation of the requirement of a discrete phase space interpretation for his Wigner bases allowed him to propose quasiprobability representations extremizing the computational resource of negativity beyond what would have been possible within a narrower scope. We have similarly aimed to impose very few constraints at the outset to see to what extent quantum theory, thus unrestrained, might offer replacements for our presuppositions.

SICs are exceptional among MICs, so finding them in a Wigner equivalence class as the closest member to the Wigner basis prompts us to postulate in general that the closest MIC or MICs in an equivalence class to their principal Wigner basis may be a quantum measurement of particular conceptual similarity to the Wigner basis. As we alluded earlier, finding representationally significant refinements to the set of principal Wigner basis

[5]Zhu calls unbiased Wigner bases "NQPRs" and prefers to report the dual basis elements. In his notation, $Q_j^- = dF_j$ and $Q_j^+ = dF_j^{\mathrm{S}}$.

preimages stands a chance of enriching our understanding of both Wigner function representations and informationally complete measurements. With this program in mind, we present some initial observations about a few Wigner bases and some candidate MICs "behind" them.

The discrete Wigner functions most familiar from the literature are those introduced by Wootters [27]. He constructs Wigner bases for prime dimensions and uses tensor products of these to form a Wigner basis for any composite dimension. As we noted in the previous section, the same can be done with Wigner equivalent MICs in the component dimensions to form Wigner equivalent MICs in any composite dimension. For $d = 2$, all unbiased Wigner bases are equivalent up to an overall unitary transformation and permutation; in particular, we may view any of them as the principal Wigner basis for some qubit SIC. Thus, as Zhu notes, reproducing the Wootters–Wigner basis is a matter of choosing the proper SIC — that is, picking a regular tetrahedron with the correct orientation in the Bloch sphere [20]. Similarly, Wootters' qutrit Wigner basis is exactly the shifted principal Wigner basis for a special SIC in dimension 3, the Hesse SIC [20, 28]. More generally, MICs parallel to the remaining necessary Wootters–Wigner bases, those in odd prime dimensions, were first constructed by Appleby [29]. The Appleby MIC can be constructed in any odd dimension d. The elements $\{E_{k,l}\}$ of this MIC are labeled by ordered pairs of integers $k, l \in \{0, \ldots, d-1\}$, and each element has rank $(d+1)/2$. Together, the elements of the Appleby MIC comprise an orbit under the action of the Weyl–Heisenberg group. Like a SIC, the Appleby MIC is equiangular. In dimension 3, the Appleby MIC is apparently the MIC antiparallel to the Hesse SIC with the minimal t value in (5.34). Nice properties like equiangularity, relatively low rank elements, and the group covariance suggest these MICs may be the most significant MICs Wigner equivalent with Wootters–Wigner bases.

Another example may be found in a discrete extension of the Cahill–Glauber formalism [30]. The authors there define an orthogonal operator basis in odd dimensions which takes operators to functions on a discrete phase space $\{\mu, \nu\}$. One can form an unbiased Wigner basis from their notation via $F_{\mu,\nu} = \frac{1}{d}\mathbf{T}^{(0)}(\mu, \nu)$, where the unitary operators $\mathbf{U}$ and $\mathbf{V}$ in their definition are the Weyl–Heisenberg "shift" and "phase" operators, respectively. The parallel, equiangular MIC with largest t value may again be a good choice of associated reference measurement, but this choice is not as compelling as in the Appleby case because the rank of its effects is $d - 1$. Equiangularity may not be worth such a large rank tradeoff. We speculate that there is a better association among the Wigner equivalent MICs to this family of Wigner bases.

A case of particular interest for quantum computation is N-qubit systems. The Wootters–Wigner basis for such a system is the N-fold tensor product of the qubit Wootters–Wigner basis. Probably the most significant Wigner equivalent MIC to this Wigner basis would be the N-fold tensor product of the appropriately oriented qubit SIC. Let $\{E_i : i = 1, \ldots, 4\}$ be a qubit SIC. Up to the weighting factor, each effect is a rank-1 projector, and the set of four such projectors can be portrayed as a regular tetrahedron inscribed in the Bloch sphere [5]. A *tensorhedron MIC* [22] is a POVM whose elements are tensor products of operators chosen from the qubit SIC $\{E_i\}$:

$$E_{i_1,\ldots,i_N} = E_{i_1} \otimes \cdots \otimes E_{i_N}. \tag{5.47}$$

Up to an overall unitary conjugation, every qubit SIC is covariant under the Pauli group, and so every tensorhedron MIC has an N-qubit Pauli symmetry. Perhaps thinking about tensorhedron measurements in N-qubit computation scenarios will inspire insights that didn't come from their quasiprobability counterparts.

The study of MICs may also suggest directions of research on the Wigner function side. One of the mysteries of the SICs is that, in all known cases, the SICs are group covariant. The *definition* of a SIC does not mention group covariance anywhere — the only symmetry in it is the equality of the inner products — and so the fact that the known SICs are all group covariant might be a subtle consequence we do not yet understand, or it might be an accident of convenience. We do know, thanks to Zhu, that in prime dimensions, if a SIC is group covariant then it must be covariant under the Weyl–Heisenberg group specifically [31, 32]. This leaves open the cases of dimensions that are higher prime powers or products of distinct primes. And in dimension $d = 8$, there exists in addition to the Weyl–Heisenberg SICs the class of *Hoggar-type SICs,* which are related to the octonions and are covariant under the three-qubit Pauli group [33–37]. All of these SICs can be converted to unbiased Wigner bases in the manner described above, and the resulting Wigner bases will inherit the group-covariance properties of the original SICs. Therefore, for a three-qubit system, the construction of principal Wigner bases from unbiased MICs furnishes three inequivalent Wigner bases of interest: Wootters, Weyl–Heisenberg, and Hoggar. The Wootters version is distinguished by particularly nice permutation symmetry properties [16].

We conclude this section by noting an example of a measure basis property which may be studied for both MICs and Wigner bases in light of the principal Wigner basis concept. In order to study time evolution, Wootters [27] explores the *triple products* of his Wigner basis, which in our notation are

$$\Gamma_{jkl} = d^2 \mathrm{tr}\,(F_j F_k F_l). \tag{5.48}$$

These can of course be defined for any Wigner basis. Of particular note is the case where the Wigner basis is the principal Wigner basis of a SIC, because the SIC triple products $\mathrm{tr}\,\Pi_j\Pi_k\Pi_l$ are remarkable numbers [12, 38–41]. We have that

$$
\begin{aligned}
d^3 \mathrm{tr}\,(F_j F_k F_l) = \pm(d+1)^{3/2}\,\mathrm{tr}\,(\Pi_j\Pi_k\Pi_l) \\
+ (1-\sqrt{d+1})(\delta_{jk}+\delta_{kl}+\delta_{jl}) \\
- \frac{1}{d^2\sqrt{d+1}}\Big(2\sqrt{d+1}+(d+1)(d-2)\Big)\,.
\end{aligned}
\tag{5.49}
$$

For Wootters' definition of the discrete Wigner basis, the triple products can be found using the geometry of the finite affine plane on d^2 points. Essentially, one takes the triangle formed by three points in that phase space, and the triple product depends upon the "area" of it. Specifically, the triple product (for the case of odd prime dimension) is given by the complex exponential

$$
\Gamma_{jkl} = \frac{1}{d}\exp\left(\frac{4\pi i}{d}A_{jkl}\right)\,. \tag{5.50}
$$

This leads naturally to an interpretation of the triple products in terms of *geometric phases.* The larger the enclosed area, the greater the geometric phase. The SIC triple products, and thus by extension those of their associated Wigner bases, have rich number- and group-theoretic properties [36–41]. What these properties imply for the Wigner bases derived from SICs is largely an open question. For early results in this vein, see Theorems 6 and 13 of [12] and also [42].

5.7 Discussion

A MIC is a basis for the operator space $\mathcal{L}(\mathcal{H}_d)$, or in other words, a coordinate system for doing quantum mechanics. Because MIC elements are required to be positive semidefinite, no MIC can ever be an *orthogonal* basis; the closest that a MIC can come to orthogonality is by being a SIC [11]. Prior work has shown that this expresses how much of the oddity of quantum theory is an artifact of coordinates, versus what is the unavoidable residuum of nonclassicality [3]. Other reference measurements may be optimally suited for other purposes, say, potentially, for algorithm design or to account for experimental specifics. If one abandons direct operational meaning in terms of probabilities, one can push basis elements outside of the positive semidefinite cone and achieve orthogonality. In this paper, we have shown well-defined procedures for doing so, and we have quantified how far an orthogonalized basis — a Wigner basis — can deviate from the original MIC in the unbiased case.

In exploring the consequences of the principal Wigner basis definition, we have found ourselves with a number of thus far unresolved questions. Several pertain to what we believe will be a fruitful direction for further research, namely the disambiguation of Wigner equivalent MICs: When is there a rank-1 MIC in an equivalence class? How is the rank of the MIC related to its distance to the principal Wigner basis? Does the principal Wigner basis suggest anything about the operational significance of a Weyl–Heisenberg covariant reference measurement? While negative quasiprobabilities do not have direct operational meaning as probabilities do, they can be made meaningful in combination with additional data. Of particular relevance is the discovery that negativity can be a resource for quantum computation [14, 15]. With a suitable Wigner equivalent MIC, the analog of negativity may be studied in reference probabilities. In the other direction, perhaps one could explore how useful statistical properties which are easily displayed by probabilistic representations are reflected in the principal Wigner function. In pursuing this inverse problem, one potential place to turn is to resource theory, especially in light of a majorization lemma concerning Born matrices we proved in a previous paper [3]. Grasping the variety of Wigner functions, and how they relate to the most economical of probabilistic representations of quantum theory, may prove helpful in advancing our understanding of this intriguing subject.

LIST OF REFERENCES

[1] C. A. Fuchs, “Charting the Shape of Quantum-State Space,” in *Quantum Communication, Measurement and Computing (QCMC): The Tenth International Conference* (T. Ralph and P. K. Lam, eds.), vol. 1363 of *AIP Conference Proceedings*, (American Institute of Physics, Melville, NY), pp. 305–314, 2011.

[2] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, tenth anniversary ed., 2010.

[3] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems,” *Physical Review Research*, vol. 2, p. 013074, 2020.

[4] G. Zauner, *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*. PhD thesis, University of Vienna, 1999. Published in English translation: G. Zauner, "Quantum designs: foundations of a noncommutative design theory," *International Journal of Quantum Information*, vol. 9, pp. 445–508, 2011.

[5] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, “Symmetric informationally complete quantum measurements,” *Journal of Mathematical Physics*, vol. 45, pp. 2171–2180, 2004.

[6] A. J. Scott and M. Grassl, “Symmetric informationally complete positive-operator-valued measures: A new computer study,” *Journal of Mathematical Physics*, vol. 51, p. 042203, 2010.

[7] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, “The SIC question: History and state of play,” *Axioms*, vol. 6, p. 21, 2017.

[8] C. A. Fuchs and M. Sasaki, “Squeezing Quantum Information through a Classical Channel: Measuring the "Quantumness" of a Set of Quantum States,” *Quantum Information & Computation*, vol. 3, pp. 377–404, 2003.

[9] A. J. Scott, “Tight informationally complete quantum measurements,” *Journal of Physics A: Mathematical and General*, vol. 39, no. 43, pp. 13507–13530, 2006.

[10] C. A. Fuchs and R. Schack, “Quantum-Bayesian coherence,” *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[11] D. Appleby, H. Dang, and C. Fuchs, “Symmetric Informationally-Complete Quantum States as Analogues to Orthonormal Bases and Minimum-Uncertainty States,” *Entropy*, vol. 16, no. 3, pp. 1484–1492, 2014.

[12] D. M. Appleby, C. A. Fuchs, and H. Zhu, “Group theoretic, Lie algebraic and Jordan algebraic formulations of the SIC existence problem,” *Quantum Information & Computation*, vol. 15, pp. 61–94, 2015.

[13] C. Ferrie, “Quasi-probability representations of quantum theory with applications to quantum information science,” *Reports on Progress in Physics*, vol. 74, no. 11, p. 116001, 2011.

[14] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, “Negative quasi-probability as a resource for quantum computation,” *New Journal of Physics*, vol. 14, no. 11, p. 113011, 2012.

[15] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, “The resource theory of stabilizer computation,” *New Journal of Physics*, vol. 16, p. 013009, 2014.

[16] H. Zhu, “Permutation Symmetry Determines the Discrete Wigner Function,” *Physical Review Letters*, vol. 116, no. 4, p. 040501, 2016.

[17] M. Appleby, C. A. Fuchs, B. C. Stacey, and H. Zhu, “Introducing the Qplex: a novel arena for quantum theory,” *The European Physical Journal D*, vol. 71, no. 7, p. 197, 2017.

[18] P. Diaconis and B. Skyrms, *Ten great ideas about chance*. Princeton: Princeton University Press, 2018.

[19] B. C. Stacey, “Is the SIC Outcome There When Nobody Looks?,” *arXiv:1807.07194*, 2018.

[20] H. Zhu, “Quasiprobability representations of quantum mechanics with minimal negativity,” *Physical Review Letters*, vol. 117, no. 12, p. 120404, 2016.

[21] J. B. DeBrota and C. A. Fuchs, “Negativity bounds for Weyl–Heisenberg quasiprobability representations,” *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

[22] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Analysis and Synthesis of Minimal Informationally Complete Quantum Measurements,” *arXiv:1812.08762*, 2020.

[23] M. Zurel, C. Okay, and R. Raussendorf, “A hidden variable model for universal quantum computation with magic states on qubits,” *arXiv:2004.01992*, 2020.

[24] S. Waldron, *An Introduction to Finite Tight Frames*. Springer, 2018.

[25] J. B. DeBrota and B. C. Stacey, “Lüders channels and the existence of symmetric-informationally-complete measurements,” *Physical Review A*, vol. 100, no. 6, p. 062327, 2019.

[26] H. Zhu, “Quantum state estimation with informationally overcomplete measurements,” *Physical Review A*, vol. 90, no. 1, p. 012115, 2014.

[27] W. K. Wootters, “A Wigner-function formulation of finite-state quantum mechanics,” *Annals of Physics*, vol. 176, no. 1, pp. 1–21, 1987.

[28] B. Stacey, “SIC-POVMs and Compatibility among Quantum States,” *Mathematics*, vol. 4, no. 2, p. 36, 2016.

[29] D. M. Appleby, “Symmetric informationally complete measurements of arbitrary rank,” *Optics and Spectroscopy*, vol. 103, no. 3, pp. 416–428, 2007.

[30] M. Ruzzi, M. A. Marchiolli, and D. Galetti, “Extended Cahill–Glauber formalism for finite-dimensional spaces: I. Fundamentals,” *Journal of Physics A: Mathematical and General*, vol. 38, no. 27, pp. 6239–6251, 2005.

[31] H. Zhu, “SIC POVMs and Clifford groups in prime dimensions,” *Journal of Physics A: Mathematical and Theoretical*, vol. 43, no. 30, p. 305305, 2010.

[32] H. Zhu, *Quantum state estimation and symmetric informationally complete POMs*. PhD thesis, National University of Singapore, 2012.

[33] S. G. Hoggar, “Two quaternionic 4-polytopes,” in *The Geometric Vein: The Coxeter Festschrift* (C. Davis, B. Grünbaum, and F. A. Sherk, eds.), Springer, 1981.

[34] S. G. Hoggar, “64 lines from a quaternionic polytope,” *Geometriae Dedicata*, vol. 69, pp. 287–289, 1998.

[35] A. Szymusiak and W. Słomczyński, “Informational power of the Hoggar symmetric informationally complete positive operator-valued measure,” *Physical Review A*, vol. 94, p. 012122, 2015.

[36] B. C. Stacey, “Sporadic SICs and the normed division algebras,” *Foundations of Physics*, vol. 47, pp. 1060–64, 2017.

[37] B. C. Stacey, “Geometric and Information-Theoretic Properties of the Hoggar Lines,” *arXiv:1609.03075*, 2016.

[38] M. Appleby, S. Flammia, G. McConnell, and J. Yard, “SICs and algebraic number theory,” *Foundations of Physics*, vol. 47, pp. 1042–59, 2017.

[39] M. Appleby, I. Bengtsson, I. Dumitru, and S. Flammia, “Dimension towers of SICs. I. Aligned SICs and embedded tight frames,” *Journal of Mathematical Physics*, vol. 58, no. 11, p. 112201, 2017.

[40] G. S. Kopp, "SIC-POVMs and the Stark conjectures," *International Mathematics Research Notices*, p. rnz153, 2019.

[41] O. Andersson and I. Dumitru, "Aligned SICs and embedded tight frames in even dimensions," *Journal of Physics A: Mathematical and Theoretical*, vol. 52, no. 42, p. 425302, 2019.

[42] D. M. Appleby, S. T. Flammia, and C. A. Fuchs, "The Lie algebraic significance of symmetric informationally complete measurements," *Journal of Mathematical Physics*, vol. 52, no. 2, p. 022202, 2011.

CHAPTER 6

SUM NEGATIVITY

6.1 Introduction

The mathematical machinery of quantum theory has persisted without substantial modification for nearly a century, but we are still waiting for a compelling set of physical principles upon which to hang the theory's predictions. It is the hope of some quantum foundations researchers that looking at standard quantum theory re-represented in an appropriate fashion will help suggest these principles [1].

Probabilities are central objects in canonical quantum theory; at the end of a quantum mechanical calculation, we are left with a probability distribution or a simple consequence of one. It is tempting, therefore, to behave as though quantum theory *gives* us a probability distribution—*the* probabilities for a set of outcomes. Careful consideration reveals, however, that *given a quantum state*, quantum theory allows us to calculate a probability distribution. But where does the quantum state come from in the first place? The apparatus of quantum theory is unable to say. In practice, an experimenter eventually settles on a quantum state for her preparation procedure after a suite of tests and calibrations, and, ultimately, statistical methods go into the state determination itself. That is, we have probabilities at the beginning and probabilities at the end with the formal apparatus of quantum theory gluing it all together. It is possible that what goes on in between these ends stands alone, but such a circumstance is far from guaranteed. The proper understanding of quantum theory may depend upon conceiving of probability theory in the proper way.

In fact, revisiting and deciding on the proper understanding of probability theory is the starting point of QBism [2–4]. QBists take a strict personalist Bayesian [5–7] stance on probability theory. A probability is a valuation an agent places on his or her degree of belief in a possible outcome, nothing more and nothing less. As a consequence, probabilities are not empirically determinable quantities because they do not independently exist outside of an agent's mind.

As an example, consider repeated flips of a coin. A frequentist conception of probability asserts that the probability of heads for the coin is the long-run ratio of number of heads to number of flips. The Bayesian first points out that to regard the coin flipped at different times as an "equivalent" or "exchangeable" process amounts to a belief the experimenter has about the situation—perhaps nothing is wrong with this belief, but she should be cognizant of its influence on the conclusions of the experiment. Secondly, the Bayesian asks just how many times the experimenter plans to actually flip the coin before she decides that the relative frequency is the probability of heads for that coin. If she flips it a finite number of times, by her own admission, any frequency is technically possible (although she believes some are unlikely). The usual answer to such a question is that she plans to flip it until the deviations in the ratio with further experimentation are small enough so as to be negligible. In other words, in order to *define* the likelihood of an outcome, she asserts that the ratio obtained by a finite series of experiments is *likely* to be close to the "true probability". The circularity of such an argument should be evident and worrying to anyone espousing the frequentist paradigm. Properly understood, then, probabilities are single-case; no probability is meaningfully right or wrong by any *external* criterion. The knee-jerk reaction to this statement is to exasperatedly throw up one's hands and exclaim that Bayesianism is just the claim that probability theory is useless! Nothing could be further from the truth: Although a probability is not subject to objective external validation or invalidation, it commands the same sway over our lives that it would if it were; if an agent wishes to avoid sure loss—to avoid being demonstrably stupid—she must take steps to ensure that she never assigns probabilities which mutually contradict each other. When a theory tells us what we “should” do or “strive for”, it is a *normative* theory. In other words, the personalist Bayesian view is that probability theory is a normative theory. In this community, compatible probabilities are called *coherent*. Remarkably, nearly all of the standard rules of probability theory are consequences of coherence [6].

One consequence that we will reference soon is the Law of Total Probability (LTP):

$$q(j) = \sum_i p(i)r(j|i). \tag{6.1}$$

The LTP describes a scenario where two actions are taken, one after another. $p(i)$ represents the probability we ascribe to getting outcome i from the first action, $r(j|i)$ is the probability we ascribe to getting outcome j from the second action conditioned on outcome i for the first action, and $q(j)$ is the probability we ascribe to outcome j for the second action, not conditioned on anything other than the operational procedure we have laid out. The commitment to coherence alone (and independent from any possible nature of reality) requires that our probabilities assigned at any given moment should hold together in accordance with (6.1).

As we explain below, it is possible to represent any quantum state as a single probability distribution over the possible outcomes of an appropriately chosen measurement. If we take this fact seriously, a quantum state is conceptually nothing more than a probability distribution. In QBism, all of the personalist Bayesian properties of probability theory carry over to quantum states; that is, quantum states, like probabilities, are valuations of belief for future experiences. However, application of the rules of quantum theory reveal that not all of the probability distributions in the probability simplex correspond to a valid quantum state. Just as in probability theory where an agent strives to be consistent with herself in her probability assignments, an agent should not ascribe a probability distribution she knows to be in conflict with the quantum mechanical formalism. In this way we arrive at an understanding that quantum theory is an empirically-motivated normative addition to probability theory.

If the functional form of the additions to probability calculus are cumbersome, then there may be no reason to adopt it for everyday use—furthermore, it may not shed any light on the "nature of reality." What would constitute a nice looking addition to probability theory? One possibility would be if the normative rules of quantum theory could be made to mirror those of probability theory in a suggestive way. It turns out that just this sort of situation can be made to occur.

An *informationally complete quantum measurement (IC-POVM)* for a Hilbert space $\mathcal{H}_d$ is a set of at least d^2 positive semi-definite operators E_i which span $\mathcal{L}(\mathcal{H}_d)$, the vector space of linear operators on $\mathcal{H}_d$, and satisfy $\sum_i E_i = \mathbb{I}$. When such a measurement consists of exactly d^2 elements, density matrices ρ and the Born rule probabilities $p(i) = \text{Tr}(\rho E_i)$ are in bijective correspondence because the E_i form a basis for $\mathcal{L}(\mathcal{H}_d)$. Such minimal IC-POVMs are known to exist in all dimensions [8]. Which one we choose for a representation, however, stands a chance of revealing or obscuring the properties which probability distributions equivalent to quantum states must have. Very often in mathematics and physics, a hard problem becomes easy when we choose the right basis. For example, the Eddington–Finkelstein coordinates revealed that the event horizon of a nonrotating black hole is not a physical singularity. Is there a best or particularly nice IC-POVM which will reveal hidden properties of quantum theory?

It is not possible for an IC-POVM to be an orthonormal basis [9–11], so perhaps our first hope is ruled out. If the elements of an IC-POVM cannot be orthogonal, can they at least be equiangular? It turns out that they can [12]. That is, we can find a set of E_i which satisfy $\mathrm{Tr}(E_i E_j) = c$ for $i \neq j$. Can such a POVM consist of only rank-one matrices? Remarkably for such a simply-stated question, it is not generally known. A set of d^2 rank-one matrices Π_i such that

$$\mathrm{Tr}\left(\Pi_i \Pi_j\right) = \frac{d\delta_{ij} + 1}{d+1} \tag{6.2}$$

defines an IC-POVM $E_i = \frac{1}{d}\Pi_i$ called a *Symmetric IC-POVM (SIC)* [13–15]. High-precision numerical SICs have been found in all dimensions 2–151 [16, 17] and in a few sporadic higher dimensions. In many cases, exact SICs have been constructed among these dimensions as well [18]. All indications are that SICs exist in all dimensions, but the proof continues to evade us.

If a SIC exists in dimension d, it is possible to rewrite the Born rule in a uniquely simple form analogous to the LTP, an equation called the *urgleichung* [3, 4], German for “primal equation” in reference [4]:

$$q(j) = \sum_i \left[(d+1)p(i) - \frac{1}{d}\right] r(j|i), \tag{6.3}$$

where $q(j)$ is the probability for obtaining outcome j of a general quantum measurement, $p(i)$ is the probability an agent ascribes to obtaining outcome i in the imagined scenario where a SIC measurement is performed on the system instead, and $r(j|i)$ is the probability for obtaining the equivalent outcome j conditional on obtaining outcome i that the agent ascribes in the imagined scenario (See [2] for a detailed exposition). It is essential to recognize the operational difference between the urgleichung and the LTP: the urgleichung describes a scenario where the first measurement is not *actually* made—just imagined. If we actually planned to implement the first measurement, our probabilities must hold together according to the familiar LTP. We take equation (6.3) very seriously. In fact, it motivated the most recent development in QBism—a reconstruction of quantum theory featuring a generalization of the urgleichung as the key assumption [19]. See references [20] and [21] for critical review and discussion of the urgleichung in QBism as well as comparison to other contexts.

Probability theory itself has no tether to physical reality—rather it is a tool that anyone, *anywhere*, can use to manage their expectations for further experiences. Those expectations will certainly be influenced by deeply-held convictions that the agent has about the nature of reality around them, but the way those probabilities must hang together if the agent is to be coherent is unaffected. Quantum mechanics, on the other hand, although evidently a normative theory like probability theory, *is* tethered to physical reality. A reformulation of quantum theory which brings to the forefront this normative structure allows us to examine the threads of this tether without confusing the subjective and the objective. This way, we may hopefully more readily determine the aspects of reality which forced quantum theory to be the way that it is. The LTP is a direct consequence of coherence in one's probability assignments. Is the urgleichung, which presents as an almost trivial modification of the LTP, on the right track for expressing the conditions for a kind of quantum coherence that an agent should strive for by virtue of being in our universe? We would like to accumulate as much evidence as possible that it is. Often one can gather more evidence for a sentiment simply by looking where it's least expected. In this case, Huangjun Zhu recently demonstrated additional evidence for this line of reasoning by instead departing from probability theory [22].

If we formally relax the positivity condition for minimal IC-POVMs (keeping the fact that they sum to the identity), we are dealing with the larger space of Hermitian operator bases. Denoting such an operator F_i, if we also keep the form of the Born rule for a quantum state ρ, we obtain a set of real numbers $\mathfrak{p}(i) = \text{Tr}(\rho F_i)$, some of which may be negative, such that $\sum_i \mathfrak{p}(i) = 1$. This set of numbers is referred to as a *quasiprobability vector* (we will always denote quasiprobability vectors with $\mathfrak{fraktur}$ script) and a quasiprobability vector obtained in this way from a quantum state is called a *quasiprobability representation* of the state.

Quantum opticians have benefited from the ease of plotting quasiprobability distributions over phase space [23], so there is some utility in their use, but what is a quasiprobability? If a probability is a valuation of belief, what meaning can we attach to a quasiprobability? There does not appear to be a simple meaning—some attempts at attaching operational substance to quasiprobabilities have been made, for example, see references [24–28], but if these solutions get us no closer to understanding *why* quantum theory is the normative probability calculus an agent of our universe should use to productively navigate, then they amount to duct tape over a structural weakness. However unsatisfying a quasiprobability may be on principled grounds, it turns out that permitting them for the time being gets us something desirable in return: the Born rule obtains an even closer functional analogy to the LTP.

Consider a Hermitian operator basis $\{F_i\}$ such that $\sum_i F_i = \mathbb{I}$ and a dual basis[1] $\{Q_j\}$ constrained to satisfy $\text{Tr}(Q_j) = 1$. From a state ρ and an arbitrary POVM $\{G_j\}$ we form the quasiprobabilities $\mathfrak{p}(i) = \text{Tr}(\rho F_i)$ and the conditional quasiprobabilities $\mathfrak{r}(j|i) =$

[1] A dual basis is one for which the bases considered together are biorthogonal, $\text{Tr}F_iQ_j = \delta_{ij}$.

$\mathrm{Tr}(Q_iG_j)$. Using the identity

$$\sum_j \mathrm{Tr}(BF_j)\mathrm{Tr}(Q_jC) = \mathrm{Tr}(BC), \tag{6.4}$$

we can rewrite the Born rule as

$$q(j) = \mathrm{Tr}(\rho G_j) = \sum_i \mathfrak{p}(i)\mathfrak{r}(j|i). \tag{6.5}$$

Like the urgleichung, (6.5) looks very much like the LTP. In fact, (6.5) is functionally *identical* to the LTP. The difference is that this equation is written in terms of quasiprobabilities instead of probabilities. Negativity must pop up somewhere, for it is known that negativity *must*[2] appear in quasiprobability representations of quantum theory [10, 29].

Another advantage when we are not burdened with positivity is that we may choose the F_i to form an orthogonal basis for operator space. If a basis is orthogonal, it is proportional to its dual basis and called *self-dual*. In this case, the sum constraint on the basis automatically fixes $\mathrm{Tr}(Q_j) = 1$ and the constant of proportionality $F_j = \frac{1}{d}Q_j$. Here we depart from Zhu's terminology and refer to a quasiprobability representation obtained from this sort of self-dual basis as a *Q-rep*. Q-reps account for most of the quasiprobability representations considered in the literature [30–33]. Importantly, however, Q-reps do not account for all finite dimensional quasiprobability representations; any nonorthogonal basis provides an example outside of this set. We can identify a Q-rep with the dual basis, $\{Q_i\}$, which defines it. We will reserve $\mathfrak{q}$ for quasiprobability representations of states with respect to a Q-rep $\{Q_j\}$ (when there is no index, we are referring to the full vector).

[2] That is, it is impossible to represent quantum theory in a way which eliminates the appearance of negativity in *both* $\mathfrak{p}$ and $\mathfrak{r}$ in (6.5) for all quantum states and POVMs. It is possible, in general, to eliminate the negativity appearing in one or the other.

Mathematically speaking, a probability is a type of quasiprobability; that is, the kind without any negative entries. If we believe that the urgleichung differing from the LTP captures some aspect of the essential difference between quantum and classical, then some of this essential difference is also contained in the difference between the LTP and (6.5), that is, in the appearance of negativity in quasiprobability representations of quantum theory. Indeed, Spekkens showed that the presence of negative elements in quasiprobability representations of quantum theory and the impossibility of noncontextual hidden variable models are equivalent notions of nonclassicality [9]. Therefore, insofar as we think contextuality is an important ingredient in the quantum–classical distinction, we should be interested in negativity as well. Investigating the negativity in Q-reps also seems to be a promising approach to identifying exactly what advantages quantum computation affords us over classical computation, for example, Veitch *et al.* showed that negativity is a resource for quantum computation [34] and Howard *et al.* recently showed that contextuality enables universal quantum computation via 'magic state' distillation [35]. Additionally, and, as we will see, of particular note for this paper, Pashayan *et al.* have shown that a value related to the sum of the negative entries in a quasiprobability representation may be thought of as a measure that bounds the efficiency of a classical estimation of probabilities [36]. In light of these facts, Zhu's recent result is especially exciting. In his paper, Zhu developed a natural measure of negativity for Q-reps and established strict upper and lower bounds for this measure in one-to-one correspondence with SICs in each dimension. SICs are related to the bounds of this measure of negativity and the appearance of negativity in quasiprobability representations of quantum theory seems to contain some hints toward what "quantum" really means. Are there more hints to be found? Motivated to answer this question, we investigate how robust Zhu's result is to modifications in the negativity measure.

In Section 6.2 we motivate and define a general negativity measure for quasiprobability representations which includes Zhu's measure and the measure of principle interest in this paper, called sum negativity, as special cases and state Zhu's theorem that the bounds on his measure of negativity are achieved by Q-reps in one-to-one correspondence with SICs. In Section 6.3 we address the sum negativity for the SIC Q-reps in the first few dimensions. In Section 6.4 we argue that Weyl-Heisenberg covariant Q-reps are a natural subset to consider while looking for counterexamples to Zhu's theorem and establish the explicit conditions for a Weyl-Heisenberg covariant Q-rep in dimension 3. Section 6.5 contains the main results: we explicitly demonstrate that Zhu's theorem does not generally extend to sum negativity in either bound, we prove a general sufficiency theorem for a Q-rep being a local maximum for sum negativity, use this theorem to prove the exact upper bound for sum negativity among Q-reps in dimension 3, and state a conjecture regarding the lower bound among Weyl-Heisenberg Q-reps. In Section 6.6 we again apply our theorem to prove that, although not generally the case, one of the SIC Q-reps achieves the exact upper bound for sum negativity in dimension 4. We also briefly discuss the sum negativity for one of the Hoggar SIC Q-reps in dimension 8. In Section 6.7 we discuss further questions and directions.

6.2 Negativity and Sum Negativity

Setting aside Q-rep vectors for a moment, we start by proposing a family of negativity measures for general quasiprobability vectors. Qualitatively speaking, we want a measure of the "amount" of negativity that appears in a vector with entries which sum to 1. A few candidates immediately stand out as especially natural measures: Perhaps a measure proportional to the sum of the negative elements or to the most negative element appearing in the quasiprobability vector—indeed, as we will see, the latter choice is taken by Zhu in [22]. We want a family of measures of negativity to meaningfully capture the deviation or "distance" from quasiprobabilities which have no negative elements. So we are faced with the task of measuring something we might understand as a distance for elements in a finite dimensional vector space (constrained by the normalization condition, of course). From this vantage point, the L^p-norms offer a very compelling family of generalized distance measures which we might utilize. Recall the definition of the L^p-norm of a vector $\mathbf{x}$,

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x(i)|^p \right)^{1/p}, \tag{6.6}$$

and the limiting expression for $p \to \infty$,

$$\|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}. \tag{6.7}$$

We might hope that the L^p-norm of a quasiprobability vector will itself be a useful quantity which is immediately related to the negativity present in the vector. If $p = 1$, for example, we can see that if there are any negative elements in the quasiprobability vector $\mathfrak{p}$ then $\|\mathfrak{p}\|_1 > 1$, so perhaps we could associate the amount by which the L^p-norm is larger than 1 with the negativity of $\mathfrak{p}$. However, $\|\mathfrak{p}\|_p \leq \|\mathfrak{p}\|_q$ when $p > q$ and, in fact, $\|\mathfrak{p}\|_p$ need not be greater than 1 even when negativity is present. So what we really want is to measure the deviation of only the negative part of the quasiprobability vector from the zero vector, disregarding all positive entries entirely. With this in mind, define the *negative part* of a quasiprobability vector $\mathfrak{p}$ to be the vector

$$\mathfrak{p}^{(-)}(j) := \frac{|\mathfrak{p}(j)| - \mathfrak{p}(j)}{2} \tag{6.8}$$

which replaces the positive elements of $\mathfrak{p}$ with zero and the negative elements with their absolute value. This definition makes it easy to isolate properties of the negative elements of a quasiprobability vector. Now that we have done away with the positive entries in our quasiprobability vector, we define the N^p *negativity* of a quasiprobability vector $\mathfrak{p}$ to be the L^p-norm of the negative part of a quasiprobability vector:

$$N^p(\mathfrak{p}) := \left\| \mathfrak{p}^{(-)} \right\|_p. \tag{6.9}$$

We will refer to the special cases N^1 and N^∞, which we see are equivalent to the two natural candidates proposed above, as the *sum negativity*[3] and the *ceiling negativity* respectively.

Zhu defines the negativity of a quantum state ρ with respect to a Q-rep $\{Q_j\}$ to be d times the magnitude of most negative element appearing in the quasiprobability representation $\mathfrak{q}(j) = \mathrm{Tr}(\rho F_j)$. In our framework, this corresponds to d times $N^\infty(\mathfrak{q})$. He then defines the negativity of a Q-rep itself to be the maximum of this value over all of quantum state space. In our framework, this corresponds to d times $\max_\rho N^\infty(\mathfrak{q})$. Thus, in the general case, we define the N^p *negativity*[4] *of a quantum state ρ with respect to a Q-rep* $\{Q_j\}$ to be

$$N^p(\rho, \{Q_j\}) := N^p(\mathfrak{q}), \tag{6.10}$$

[3]Veitch *et al.* use the term sum negativity specifically for the sum of the negative elements of the discrete Wigner function for a quantum state [37] whereas we will be considering the equivalent notion with respect to any Q-rep.

[4]We have chosen to omit multiplication by d in our negativity definitions so that the negativity can more immediately be associated with the negative values in a quasiprobability vector. As the dual basis is calculationally easier to work with, a downside of our convention is that factors of $1/d$ crop up more frequently.

and the N^p *negativity of a Q-rep* $\{Q_j\}$ to be

$$N^p(\{Q_j\}) := \max_\rho N^p(\rho, \{Q_j\}). \tag{6.11}$$

Since L^p-norms are convex and nondecreasing for $p \geq 1$ on the positive reals [38] and the negative part map (6.8) is a convex function, the composition (6.10) is also convex [39]. This means the maximum occurs on the boundary of the domain so we may take the maximum in (6.11) to be over pure states in these cases.

As it will be useful later, note that the L^2 norm of a Q-rep vector corresponding to a pure state always equals $\sqrt{1/d}$. To see this, let $B = C = \rho$ and $\mathrm{Tr}(\rho^2) = 1$ in (6.4). This is another reason why the L^p norms are not a good choice for a family of negativity measures.

Zhu notes that the ceiling negativity of a Q-rep $\{Q_j\}$ takes the simple form

$$N^\infty(\{Q_j\}) = \frac{1}{d}\left|\min_j \lambda_{\text{min}}(Q_j)\right|, \tag{6.12}$$

where $\lambda_{\text{min}}(Q_j)$ is the minimal eigenvalue of Q_j.

If a SIC, denoted $\{\Pi_j\}$, exists in dimension d, we may construct two Q-reps $\{Q_j^+\}$ and $\{Q_j^-\}$ called the *SIC Q-reps*:

$$Q_j^\pm = \mp\left(\sqrt{d+1}\right)\Pi_j + \frac{1}{d}\left(1 \pm \sqrt{d+1}\right)\mathbb{I} \tag{6.13}$$

which have ceiling negativities

$$N^+ \equiv N^\infty(\{Q_j^+\}) = \frac{(d-1)\sqrt{d+1}-1}{d^2}, \quad N^- \equiv N^\infty(\{Q_j^-\}) = \frac{\sqrt{d+1}-1}{d^2}. \tag{6.14}$$

We introduce these Q-reps because of the following theorem.

Theorem 18 (Zhu)**.** *Every Q-rep $\{Q_j\}$ in dimension d satisfies $N^{-} \leq N^{\infty}(\{Q_j\}) \leq N^{+}$. The lower bound is saturated if and only if $\{Q_j\}$ has the form $\{Q_j^{-}\}$ where Π_j is a SIC. If $\{Q_j\}$ is group covariant, then the upper bound is saturated if and only if $\{Q_j\}$ has the form $\{Q_j^{+}\}$.*

Zhu's theorem identifies SICs as centrally important to the study of Q-reps and more broadly for quantum theory because the Q-reps which achieve both bounds on ceiling negativity in any dimension are related to SICs by a simple affine transformation. Is the ceiling negativity unique in this way? If so, it would be interesting to understand why. If not, where does it fail?

We will address this question for the sum negativity. Like ceiling negativity, there is a more manageable expression for the sum negativity of a Q-rep. The following argument is due to Appleby and Zhu.

Lemma 7. *An equivalent form of the sum negativity of a Q-rep $\{Q_j\}$ is*

$$N^1(\{Q_j\}) = -\frac{1}{d}\min\left\{\lambda_{\{1\}}, \lambda_{\{2\}}, \ldots, \lambda_{\{2^{d^2}-1\}}\right\}, \tag{6.15}$$

where $\lambda_{\{i\}}$ is the minimal eigenvalue of the ith partial sum matrix of the $\{Q_j\}$ matrices.

Proof. For a quasiprobability representation $\mathfrak{q}$ of a state ρ with respect to Q-rep $\{Q_j\}$, if $\mathcal{S} := \{i|\mathrm{Tr}(\rho Q_i) < 0\}$, then

$$N^1(\mathfrak{q}) = \left|\frac{1}{d}\sum_{i\in\mathcal{S}}\mathrm{Tr}(\rho Q_i)\right| = \left|\frac{1}{d}\mathrm{Tr}\left(\rho\sum_{i\in\mathcal{S}}Q_i\right)\right|. \tag{6.16}$$

As in (6.12), the minimal value of the expression $\mathrm{Tr}(\rho F)$ over quantum state space is the minimal eigenvalue of F and the state ρ which minimizes the expression is the corresponding eigenvector. Thus, for a fixed subset $\mathcal{G}$ of the $\{Q_j\}$ matrices, the state which minimizes $\mathrm{Tr}(\rho \sum_{i \in \mathcal{G}} Q_i)$ is the minimal eigenvector of the matrix $\sum_{i \in \mathcal{G}} Q_i$. Via the definition above, for each ρ there is a subset $\mathcal{S}$; in particular, there is a subset $\mathcal{S}'$ for a state whose quasiprobability representation has the sum negativity value $N^1(\{Q_j\})$ and, furthermore, the magnitude of the minimal eigenvalue of the matrix $\sum_{i \in \mathcal{S}'} Q_i$ is equal to d times $N^1(\{Q_j\})$. Thus determining the sum negativity is equivalent to looking for the minimal eigenvalue over all partial sum matrices of $\{Q_j\}$. There are $2^{d^2} - 1$ entries to minimize over because we ignore the partial sum corresponding to the empty subset. □

Does Theorem 18 extend to sum negativity? In Section 6.5, we will demonstrate that it generally does not with explicit counterexamples in the first nontrivial dimension, $d = 3$.

6.3 Sum Negativities of $\{Q_j^+\}$ and $\{Q_j^-\}$

Sum negativity is notably harder to work with than ceiling negativity, both analytically and numerically. As such, using the method described in Lemma 7, analytic results for the sum negativity of $\{Q_j^+\}$ and $\{Q_j^-\}$ have only so far been obtained for dimensions 2, 3, and 4. Numerically exact results have also been obtained for $d = 5$.

For $d = 2$, sum negativity and ceiling negativity are equivalent measures because a $d = 2$ Q-rep vector cannot contain more than one negative element. This property can be proven easily with equation (6.4) and the fact quasiprobabilities are normalized. In addition, the ceiling negativities for $\{Q^+\}$ and $\{Q^-\}$ are equivalent. Thus,

$$N^1(\{Q_j^\pm\}) = N^\infty(\{Q_j^\pm\}) = \frac{\sqrt{3}-1}{4}. \tag{6.17}$$

In fact, all N^p negativities for the SIC Q-reps are equal in dimension 2. This is a reflection of the fact that all Q-reps are equivalent to the Wootters discrete Wigner function in this dimension [22].

There is a continuous one-parameter family of SICs in dimension 3 [40]. For any of them, we may construct the SIC Q-reps. In Zhu's paper we easily see that ceiling negativity is insensitive to the SIC chosen—all that matters is that the defining property of a SIC is satisfied. This turns out to also be true for sum negativity[5]. What's more, the sum negativities of $\{Q_j^+\}$ and $\{Q_j^-\}$ are equal:

$$N^1(\{Q_j^\pm\}) = \frac{1}{3}. \tag{6.18}$$

Recall that Theorem 18 establishes that the upper and lower bounds for ceiling negativity over all Q-reps are achieved by the SIC Q-reps. We see now that in dimension 3, the sum negativities of $\{Q_j^+\}$ and $\{Q_j^-\}$ are equal. This tells us that if any $d = 3$ Q-rep has a sum negativity other than $1/3$, the analog of Theorem 18 does not hold for sum negativity in dimension 3.

For $d = 4$,

$$N^1(\{Q_j^+\}) = \frac{1}{2}, \tag{6.19}$$

[5]We won't explore it further here, but the situation is more interesting. Although the sum negativity is insensitive to the value of the parameter t which defines the inequivalent SIC Q-reps, the quantum states whose quasiprobability representations achieve these sum negativity values *do* depend on the parameter. It turns out that the sum negativity for $\{Q_j^-\}$ constructed from the Hesse SIC ($t = 0$ in [40]) is achieved by a complete set of mutually unbiased bases [41], that is, 12 $(= d(d+1))$ vectors which form four orthogonal bases such that any vector from one basis has an equal overlap with any vector from another basis. For all the other SICs in dimension 3, the states which achieve the sum negativity of $\{Q^-\}$ form a single basis instead. The complete set of mutually unbiased bases also turns out to be the set of states which minimize the Shannon entropy in the Hesse SIC representation [42, 43].

and

$$N^1(\{Q_j^-\}) = -\frac{1}{16}\left(5+\sqrt{5}-2\sqrt{2\left(1+\sqrt{5}\right)}-2\sqrt{23-2\sqrt{5}+2\sqrt{-22+10\sqrt{5}}}\right)$$
$$\approx 0.420967. \tag{6.20}$$

Value (6.19) is surprisingly nice. We will comment briefly on this in Section 6.6. Value (6.20) is shocking. However, in light of recent results relating the SIC problem to algebraic number theory ([44], see the contributions of Appleby *et. al* [45] and Bengtsson [46] to this volume for a review), it is worth mentioning a few possibly relevant facts about this number and how it arose.

The sum negativity for the $d = 4$ SIC Q-rep is $(1/4$ times) the minimal eigenvalue of certain 7-element partial sums of the SIC Q-rep matrices. The characteristic polynomial which has this eigenvalue as a root is:

$$\begin{aligned} x^4 - 7x^3 + \frac{21}{2}x^2 + \left(\frac{129}{8} - \frac{35\sqrt{5}}{8} + \sqrt{2\left(31+17\sqrt{5}\right)}\right)x \\ - \frac{1293}{32} + \frac{293\sqrt{5}}{32} - \sqrt{5\left(22+29\sqrt{5}\right)} \end{aligned} \tag{6.21}$$

The coefficients in equation (6.21) are members of the field $\mathbb{Q}\left(\sqrt{5}+\sqrt{362+313\sqrt{5}}\right)$. The factor multiplying $-1/16$ in (6.20) is an algebraic integer, but not an algebraic unit. Finally, the minimal polynomial for (6.20) is degree 8.

For $d = 5$, $N^1(\{Q^+\}) \approx 0.584277$ and $N^1(\{Q^-\}) \approx 0.501957$. These answers are numerically correct, but do not lend themselves readily to conversion to exact values.

6.4 Weyl-Heisenberg Q-reps in $d = 3$

We say that a set of vectors is *group covariant* if it is the orbit of some group action on an initial vector, which we call the *fiducial*. All known SICs are group covariant and all but one of those are covariant with respect to the *Weyl-Heisenberg (WH) group*. In dimension d, let $\omega_d = e^{2\pi i/d}$ be a dth root of unity, and define the shift and phase operators

$$X|j\rangle = |j+1\rangle, \, Z|j\rangle = \omega_d^j |j\rangle, \tag{6.22}$$

where the shift is modulo d. Products of powers of X and Z and powers of ω_d define the WH group. The order of the WH group in dimension d is d^3, but for the purposes of constructing a measurement operator or Q-rep, we can neglect the phase factors. In other words, for some fiducial matrix Q_0, the other matrices in the orbit would take the form

$$Q_{ij} = X^i Z^j Q_0 (X^i Z^j)^\dagger. \tag{6.23}$$

The only known exception to WH covariance for SICs is the Hoggar SIC in dimension 8, but even this outlier is group covariant with respect to the tensor product of three $d = 2$ WH groups. What about Q-reps? By their construction, the non-Hoggar SIC Q-reps are WH covariant and these Q-reps achieve the bounds for ceiling negativity in all dimensions (provided a SIC exists in that dimension). Thus, if we were to consider any subset of the full space of Q-reps for computational study, the set of WH covariant Q-reps (WH Q-reps) is likely the best starting point. In any case, due to the ubiquity of the WH group in quantum information theory, the bounds of sum negativity within WH Q-reps may be of independent interest.

A Q-rep is associated with an orthogonal basis of operators $\{Q_j\}$ with norm 3, that is,

$$\mathrm{Tr}(Q_iQ_j) = 3\delta_{ij}. \tag{6.24}$$

Therefore, the general conditions for a WH Q-rep may be obtained by requiring that the elements of the WH orbit of a general unit-trace Hermitian matrix

$$\begin{bmatrix} z & y & x \\ y^* & w & v \\ x^* & v^* & 1-z-w \end{bmatrix} \tag{6.25}$$

satisfy equation (6.24). Imposing this condition results in a number of equations which can be algebraically simplified to the following three:

$$\begin{aligned} z^2 + zw + w^2 &= z + w, \\ |y|^2 + |x|^2 + |v|^2 &= 1, \\ xy + y^*v + v^*x^* &= 0. \end{aligned} \tag{6.26}$$

In terms of real variables, an arbitrary unit-trace Hermitian matrix

$$\begin{bmatrix} a & b+ic & d+ie \\ b-ic & f & g+ih \\ d-ie & g-ih & 1-a-f \end{bmatrix} \tag{6.27}$$

is a WH Q-rep fiducial if

$$\begin{aligned} a^2 + af + f^2 &= a + f, \\ b^2 + c^2 + d^2 + e^2 + g^2 + h^2 &= 1, \\ dg + bd + bg + ch &= ec + eh, \\ cd + be + bh &= cg + dh + eg. \end{aligned} \tag{6.28}$$

From (6.26) or (6.28) we can see that $d = 3$ WH Q-rep fiducials define a 4 dimensional subspace of the 8 dimensional space of 3×3 unit-trace Hermitian matrices. The main diagonal elements are independent of the off-diagonal elements and satisfy the equation of an ellipse (the first equation in (6.26) or (6.28)). The magnitude of the off-diagonal elements in (6.25) lie on the unit 2-sphere, but their exact values only lie at points where the expression $xy + y^*v + v^*x^*$ vanishes.

The following two matrices are explicit examples of $d = 3$ *Non-SIC* WH Q-rep fiducials:

$$Q^{\text{min}} = \begin{bmatrix} 0 & -\frac{1}{3} + \frac{i}{3} & \frac{2}{3} - \frac{i}{3} \\ -\frac{1}{3} - \frac{i}{3} & 1 & -\frac{1}{3} + \frac{i}{3} \\ \frac{2}{3} + \frac{i}{3} & -\frac{1}{3} - \frac{i}{3} & 0 \end{bmatrix}, \tag{6.29}$$

$$Q^{\text{max}} = \begin{bmatrix} 0 & \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) + \frac{i}{4} & \left(\frac{2}{3} + \frac{\sqrt{7}}{12}\right) - \frac{i}{4} \\ \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) - \frac{i}{4} & 1 & \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) + \frac{i}{4} \\ \left(\frac{2}{3} + \frac{\sqrt{7}}{12}\right) + \frac{i}{4} & \left(-\frac{1}{3} + \frac{\sqrt{7}}{12}\right) - \frac{i}{4} & 0 \end{bmatrix}. \tag{6.30}$$

In the following section we will see why they are presented with the designations “max” and “min”.

6.5 Sum Negativity Bounds in $d = 3$

The WH Q-reps generated by (6.29) and (6.30) will be denoted $\{Q_j^{\min}\}$ and $\{Q_j^{\max}\}$ respectively. Their sum negativities are

$$N^1(\{Q_j^{\min}\}) = \frac{1}{3}\left(2\cos\frac{\pi}{9} - 1\right) \approx 0.293128, \tag{6.31}$$

and

$$N^1(\{Q_j^{\max}\}) = \frac{2}{9}\left(\sqrt{7} - 1\right) \approx 0.365723. \tag{6.32}$$

Recall from Section 6.3 that the sum negativities of both SIC Q-reps is $1/3$, so $\{Q_j^{\min}\}$ and $\{Q_j^{\max}\}$ are explicit counterexamples to Zhu's theorem for sum negativity. It turns out that (6.32) is a strict upper bound on the sum negativity, not only of WH Q-reps, but of all Q-reps in dimension 3. We start with two lemmas.

Lemma 8. *For quasiprobability vectors with d^2 elements lying in the sphere of radius $\sqrt{1/d}$, the stationary points for sum negativity are:*

1. *Those vectors whose entries consist only of two distinct values.*

2. *Those vectors whose entries consist only of three distinct values including zero.*

Proof. Quasiprobability vectors in the sphere of radius $\sqrt{1/d}$ satisfy the following constraints:

$$\sum_{j=1}^{d^2} \mathfrak{p}(j) = 1 \quad \text{and} \quad \sum_{j=1}^{d^2} \mathfrak{p}(j)^2 = \frac{1}{d}. \tag{6.33}$$

The definition of the sum negativity of a quasiprobability vector (6.34) gives us

$$N^1(\mathfrak{p}) = \|\mathfrak{p}^{(-)}\|_1 = \sum_{j=1}^{d^2} \frac{|\mathfrak{p}(j)| - \mathfrak{p}(j)}{2} = \frac{1}{2}\sum_{j=1}^{d^2} |\mathfrak{p}(j)| - \frac{1}{2}. \tag{6.34}$$

From this it is clear that the stationary points of the sum negativity are exactly the stationary points of the sum of the absolute values. Absolute values are often difficult to deal with in optimization problems, but it turns out that the function $|x|$ may be approximated efficiently by $\sqrt{x^2+c}$ where c is taken to zero after any differentiation [47]. Thus, we want to extremize the function

$$\sum_{j=1}^{d^2} \sqrt{\mathfrak{p}(j)^2 + c} \tag{6.35}$$

subject to constraints (6.33) in the small c limit. To do this we construct a Lagrangian

$$\mathcal{L}(\mathfrak{p}, \lambda, \mu) = \sum_{j=1}^{d^2} \left(\sqrt{\mathfrak{p}(j)^2 + c}\right) - \lambda \left(\sum_{j=1}^{d^2} \mathfrak{p}(j) - 1\right) - \mu \left(\sum_{j=1}^{d^2} \mathfrak{p}(j)^2 - \frac{1}{d}\right) \tag{6.36}$$

where λ and μ are Lagrange multipliers. Varying this Lagrangian, we see that the stationary points must satisfy

$$\frac{\mathfrak{p}(j)}{\sqrt{\mathfrak{p}(j)^2 + c}} - \lambda - 2\mu\mathfrak{p}(j) = 0 \tag{6.37}$$

for all j. This expression has 4 solutions, two of which become zero in the $c = 0$ limit:

$$\mathfrak{p}(j) = \frac{\pm 1 - \lambda}{2\mu} \quad \text{or} \quad \mathfrak{p}(j) = 0. \tag{6.38}$$

Consider the case where $\mathfrak{p}(j) \neq 0$ for all j. Now, in order for the constraints (6.33) to hold, some number n of the entries are $\frac{-1-\lambda}{2\mu}$ and the other $d^2 - n$ are $\frac{1-\lambda}{2\mu}$. That is,

$$n\left(\frac{-1-\lambda}{2\mu}\right) + (d^2 - n)\left(\frac{1-\lambda}{2\mu}\right) = 1 \quad \text{and} \quad n\left(\frac{-1-\lambda}{2\mu}\right)^2 + (d^2 - n)\left(\frac{1-\lambda}{2\mu}\right)^2 = \frac{1}{d}. \tag{6.39}$$

It is easy to verify that $0 < n < d^2$. In this case the quasiprobability vector $\mathfrak{p}$ consists of two distinct values. If $\mathfrak{p}(j) = 0$ for m of the indices then $d^2 - n - m$ of the entries are $\frac{1-\lambda}{2\mu}$

and the appropriately modified form of (6.39) holds in which case the quasiprobability vector $\mathfrak{p}$ consists of three distinct values including zero. $\square$

A stationary point can be a local maximum, a local minimum, or a saddle point. Sufficient conditions for maxima and minima in Lagrangian systems with equality constraints are known. We will need the following tool (which can be found in chapter 2 of [48]):

Theorem 19 (Sufficient Conditions for Constrained Maxima)**.** *Consider a constrained maximization problem for a twice-differentiable function of n variables $y(\mathbf{x})$ with m twice-differentiable equality constraints $f_i(\mathbf{x}) = 0$, $i = 1, \ldots, m$. The Lagrangian is*

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = y(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}), \tag{6.40}$$

where λ_i are Lagrange multipliers. If there exist vectors $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$ such that $\partial_i \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$, $i = 1, \ldots, n$ and $f_i(\mathbf{x}^*) = 0$, $i = 1, \ldots, m$ and if

$$(-1)^s \begin{vmatrix} \partial_{11}\mathcal{L} & \cdots & \partial_{1s}\mathcal{L} & \partial_1 f_1 & \cdots & \partial_1 f_m \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_{1s}\mathcal{L} & \cdots & \partial_{ss}\mathcal{L} & \partial_s f_1 & \cdots & \partial_s f_m \\ \partial_1 f_1 & \cdots & \partial_s f_1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_s f_m & 0 & \cdots & 0 \end{vmatrix} > 0 \tag{6.41}$$

for $s = m+1, \ldots, n$ ($\partial_{ij}\mathcal{L}$ indicates a second partial derivative of $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ with respect to x_i and x_j evaluated at $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$ and $\partial_i f_j$ indicates the first partial derivative of constraint function $f_j(\mathbf{x})$ with respect to x_i evaluated at $\mathbf{x}^*$), then $y(\mathbf{x})$ has a strict local maximum at $\mathbf{x}^*$.

A consequence of this is the following:

Lemma 9. *The stationary points for sum negativity of quasiprobability vectors with d^2 elements lying in the sphere of radius $\sqrt{1/d}$ are all local maxima or global minima.*

Proof. From Lemma 8 we know that the stationary points for sum negativity of quasiprobabilities lying on the sphere of radius $\sqrt{1/d}$ are those consisting of exactly two distinct elements or those consisting of three if one of them is zero. If there are two distinct elements, they can either both be nonnegative or of opposite signs. If they are both nonnegative, the sum negativity is zero which is the global minimum sum negativity value. If there are three distinct elements and they are all nonnegative, then it is also a global minimum for sum negativity. We will now show that when the nonzero elements of a stationary point are of opposite signs, it is a local maximum.

Consider first the stationary points with two distinct elements. In terms of the Lagrange multipliers from the proof of Lemma 8, the stationary quasiprobability vectors are comprised of opposite signed values when $\mu \neq 0$ and $-1 < \lambda < 1$. Without loss of generality we choose $\frac{-1-\lambda}{2\mu}$ to be the negative value so that we may enumerate the number of negative entries with index n as in (6.39). This amounts to the further restriction $\mu > 0$. Note that by substituting $\frac{-1-\lambda}{2\mu} = a$ and $\frac{1-\lambda}{2\mu} = b$ in (6.39), we can solve for the positive and negative entries in the quasiprobability vector in terms of the dimension d and number of negative entries $1 \leq n < d(d-1)$:

$$a = \frac{1}{d^2} + \frac{1}{d^2}\sqrt{\frac{n(d-1)}{d^2-n}} - \sqrt{\frac{d-1}{n(d^2-n)}} < 0, \tag{6.42}$$

$$b = \frac{1}{d^2} + \frac{1}{d^2}\sqrt{\frac{n(d-1)}{d^2-n}} > 0. \tag{6.43}$$

Sum negativity is invariant to the ordering of the entries in a quasiprobability vector so we need only prove that one ordering is a local maximum for each n. Without loss of generality, we fix $\mathfrak{p}(1) = a$ and demand that $\mathfrak{p}(2)$ and $\mathfrak{p}(3)$ are not both also equal to a. Consider again the Lagrangian (6.40). Taking the relevant derivatives and c to zero, we may construct the bordered Hessian matrix from (6.41) for our problem. Therefore, if $\vec{\mathfrak{p}}$ is a stationary point and if

$$(-1)^s \begin{vmatrix} -2\mu & \cdots & 0 & 1 & 2\mathfrak{p}(1) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & -2\mu & 1 & 2\mathfrak{p}(s) \\ 1 & \cdots & 1 & 0 & 0 \\ 2\mathfrak{p}(1) & \cdots & 2\mathfrak{p}(s) & 0 & 0 \end{vmatrix} > 0 \tag{6.44}$$

for $s = 3, \dots, d^2$, then $\vec{\mathfrak{p}}$ is a local maximum in sum negativity.

Recall the following identity for matrix blocks A, B, C, and D which holds when A is invertible:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D - CA^{-1}B|. \tag{6.45}$$

Using (6.45), the determinant of the bounded Hessian matrix in (6.44) is

$$(-1)^s (2\mu)^{s-1} \left(4s \sum_{i=1}^{s} \mathfrak{p}(i)^2 - 4 \sum_{i,j=1}^{s} \mathfrak{p}(i)\mathfrak{p}(j) \right). \tag{6.46}$$

With the additional $(-1)^s$ term in equation (6.44), we see that the inequality is satisfied if

$$s \sum_{i=1}^{s} \mathfrak{p}(i)^2 - \sum_{i,j=1}^{s} \mathfrak{p}(i)\mathfrak{p}(j) > 0 \tag{6.47}$$

for all s. Let l denote the number of negative elements in the truncated quasiprobability vectors appearing in (6.44). Note that $1 \leq l < s$ by our earlier assumption about the first three elements. Then

$$s\sum_{i=1}^{s}\mathfrak{p}(i)^2-\sum_{i,j=1}^{s}\mathfrak{p}(i)\mathfrak{p}(j) = s(la^2+(s-l)b^2)-(la+(s-l)b)^2 = (a-b)^2(s-l)l > 0. \quad (6.48)$$

Thus the stationary points consisting of two distinct values, one negative and one positive, are local maxima.

The values of a and b in the case with three distinct values are more complicated, but may still be obtained. Otherwise the proof in this case carries through in the same way as above with a slight modification in the last step. Let m denote the number of elements equal to zero in the truncated quasiprobability vectors appearing in (6.44). Note that $l + m \leq s$. Then it can be verified that

$$s\sum_{i=1}^{s}\mathfrak{p}(i)^2 - \sum_{i,j=1}^{s}\mathfrak{p}(i)\mathfrak{p}(j) = s(la^2 + (s-l-m)b^2) - (la + (s-l-m)b)^2 > 0, \quad (6.49)$$

which completes the proof. □

We may now return to the question of extremality among Q-reps. We say that a pure state ρ *achieves the sum negativity* if it is an eigenvector with eigenvalue magnitude equal to d times $N^1(\{Q_j\})$ of one of the partial sum matrices of the Q-rep $\{Q_j\}$.

Theorem 20. *If the quasiprobability representation of a state which achieves the sum negativity of a Q-rep $\{Q_j\}$ consists of two distinct elements or three including zero, then $N^1(\{Q_j\})$ is a local maximum among all Q-reps.*

Proof. Recall that the quasiprobability representation of a pure state with respect to a Q-rep lies in the sphere of radius $\sqrt{1/d}$. Also recall that the convexity of the sum negativity function implies that any state which achieves the sum negativity for a Q-rep $\{Q_j\}$ is a pure state. Therefore, if, as we vary $\{Q_j\}$ over the space of Q-reps, the quasiprobability representation of a pure state which achieves the sum negativity of $\{Q_j\}$ consists of two distinct values or three distinct values including zero for some Q-rep $\{Q'_j\}$, then by Lemmas 8 and 9, there is a local maximum or global minimum of the sum negativity function at $\{Q'_j\}$. As we know the appearance of negativity is inevitable, the global minimum possibility is avoided. $\square$

Theorem 21. *The exact upper bound for sum negativity among Q-reps in dimension* 3 *is* $\frac{2}{9}(\sqrt{7}-1)$.

Proof. When $d = 3$, explicit calculation reveals that the sum negativity of a quasiprobability vector lying in the sphere of radius $\sqrt{1/d}$ constructed with n values equal to $a < 0$, m values equal to 0, and $9 - n - m$ values equal to $b > 0$ is maximized when $n = 2$ and $m = 0$. Those values are, from (6.42) and (6.43),

$$a = \frac{1}{9}(1-\sqrt{7}) \text{ and } b = \frac{1}{63}(7+2\sqrt{7}). \tag{6.50}$$

Lemmas 8 and 9 and the fact that our domain has no boundary (lying on the sphere of radius $\sqrt{1/d}$ and quasiprobability normalization together define a (d^2-2)–sphere) imply that the sum negativity of this quasiprobability vector is the global maximum value over this domain. The quasiprobability representations of the states which achieve (6.32) with respect to the Q-rep $\{Q_j^{\max}\}$ consist of these values, and so, by Theorem 20, $\frac{2}{9}(\sqrt{7}-1)$ is the maximum value for sum negativity over all Q-reps in dimension 3. $\square$

It is important to note that while Theorem 21 shows that (6.32) is the strict upper bound for sum negativity among all Q-reps, it does not imply that $\{Q_j^{\max}\}$ is the unique

Q-rep which achieves this bound. Numerical searching suggests that, for the fiducial main diagonal $\{0, 1, 0\}$, the WH orbit of (6.30) is the unique Q-rep which achieves the sum negativity (6.32), but that there is at least one other fiducial main diagonal which achieves this bound, namely the main diagonal corresponding to the major axis vertices of the ellipse defined by the first equation in (6.26): $\{\frac{1}{3}, \frac{1}{3} + \frac{1}{\sqrt{3}}, \frac{1}{3} - \frac{1}{\sqrt{3}}\}$. Unfortunately, for this fiducial, we were unable to convert the numerical result to exact values.

What can be said about the lower bound for sum negativity in dimension 3? So far, less is known, but we present the following numerically motivated conjecture:

Conjecture 6. The exact lower bound for sum negativity among WH Q-reps in dimension 3 is $\frac{2}{3}(\cos\frac{\pi}{9} - \frac{1}{2})$.

This statement resisted our attempts to prove it in a fashion similar to Theorem 21 because for every Q-rep there exist states with zero negative elements in their quasiprobability representation (for example, the maximally mixed state). The Q-rep vector corresponding to the minimal eigenstate over all partial sums of $\{Q_j^{\min}\}$ consists of three distinct values which each appear three times:

$$\frac{1}{9}\left(1 - 2\cos\frac{\pi}{9}\right), \frac{1}{9}\left(1 + 2\cos\frac{2\pi}{9}\right), \text{ and } \frac{1}{9}\left(1 + 2\cos\frac{\pi}{9} - 2\cos\frac{2\pi}{9}\right). \tag{6.51}$$

The values in (6.51) may not have significance as deep as those in the upper bound quasiprobability vector, however, as numerical searching has revealed that (6.31) is achieved by a WH Q-rep for *every* valid main diagonal. The off-diagonal elements in other cases were too difficult to convert to exact values. The fact that there seems to be a WH Q-rep which achieves (6.31) for any main diagonal satisfying the first equation in (6.26) suggests the lower bound among Q-reps may not be saturated by WH Q-reps. Additionally, the extremal properties of general quasiprobability vectors lying on the sphere of radius $\sqrt{1/d}$

do not come to our aid when we try to find the lower bound because there is no obvious reason to hope that the process of maximizing over quantum state space (in the definition of a Q-rep sum negativity) followed by minimizing over all Q-reps (to find the lower bound for sum negativity) should result in one of the local maxima for general quasiprobability vectors on the sphere of radius $\sqrt{1/d}$ (recall, of course, that it cannot result in the global minimum of zero negativity on this sphere because we know that the appearance of negativity is inevitable). Before, we were maximizing over both; in some sense we got lucky that the global maximum sum negativity for a quasiprobability vector on this sphere was achieved by a Q-rep vector.

6.6 Further Observations about SIC Q-reps

In Section 6.3, we noted the appearance of a rational value for the sum negativity of $\{Q_j^+\}$ in dimension 4. The quasiprobability vectors which achieve the sum negativity of $\{Q_j^+\}$ are of a special and familiar form; they consist of only the values $-1/8$ and $1/8$. We know from Theorem 20, therefore, that this Q-rep is a local maximum for sum negativity among Q-reps in dimension 4. In fact, following the exact same procedure as in Theorem 21, we see that $N^1(\{Q_j^+\}) = 1/2$ is the exact upper bound over all Q-reps!

In footnote 5, we mentioned that the states which achieve the sum negativity of $\{Q_j^-\}$ for the Hesse SIC in dimension 3 form a complete set of mutually unbiased bases. The states which achieve the sum negativity for $\{Q_j^-\}$ in dimension 4 also form a structure of possible interest. They consist of a set of 16 vectors which have two nontrivial squared overlaps. In the terminology of reference [13], these states form a quantum design of degree 2.

Although dimension 5 was the last in which we were able to explicitly calculate the sum negativity for the SIC Q-reps by exhaustive combinatorial searching, we suspect that we have found the correct sum negativity for $\{Q_j^-\}$ constructed with the Hoggar SIC in dimension 8. Rather than calculating the eigenvalues of every partial sum matrix (since this is infeasible for 2^{64}, 8×8 matrices), we used a numerical local maximization procedure and around 10^6 random pure state seeds. The overall maximum value we found, $7/8$, occurred frequently in our data and is significantly larger than all of the smaller local maxima. Of course, we could still be falling short of the global maximum value if it occurs at very hard to access positions. The states whose quasiprobability representations achieve the sum negativity of $7/8$ consist of 28 copies of value $-1/32$ and 36 copies of value $5/96$. Therefore, by Theorem 20, if these states achieve the actual sum negativity, then $\{Q_j^-\}$ constructed with the Hoggar SIC is a local maxima among all Q-reps in dimension 8. Surprisingly, these quantum states also minimize the Shannon entropy of their Hoggar SIC representations and thus compose the "twin" Hoggar SIC [43, 49]. This result and the one for $\{Q_j^-\}$ in dimension 4 in the previous paragraph parallel the one mentioned in footnote 5.

6.7 Discussion

In QBism, quantum states are probability distributions over a set of possible outcomes for an appropriately chosen measurement. With the understanding that quantum theory is an addition to coherence which rational agents should use to help inform their expectations for future experiences in terms of their past ones, QBists hope that the structure of quantum theory, and in particular its boundaries, can be made to suggest nature's motives. A century of quantum foundational debate should by now have convinced us that these motives will not conform to our prejudices about reality. Although the weirdness of quantum theory has convinced some that we can no longer pretend physics is more than an exercise in instrumentalism, QBists are optimistic that there are ubiquitous and recognizably physical statements about nature yet to be made.

Prior to Zhu's paper, we had focused on characterizing the bounds of quantum theory only from within probability theory. His results reveal this approach was nearsighted. In the initial stages of this project, we had hoped to find further evidence for the centrality of SICs in quantum theory by showing that Zhu's theorem extends to another natural measure of negativity for quasiprobability representations. Indeed, it is interesting to find that the SICs do not generally play the same role in this alternate context. Why do they not? Do they still always play *some* role which is not immediately apparent? Maybe there are other families of Q-reps constructed from SICs or another structure which naturally play the same role for sum negativity. $\{Q_j^+\}$ in dimension 4 did achieve the upper bound for sum negativity. Does this happen again? Although we cannot calculate it exactly, we have some numerical evidence that the sum negativities for the Hoggar SIC Q-reps differ from the non-Hoggar SIC Q-reps in dimension 8. This suggests that there might be an essential relation between sum negativity and group covariance. Furthermore, the appearance of the complete set of mutually unbiased bases in dimension 3 (mentioned in footnote 5) and the "twin" Hoggar SIC in dimension 8 suggest a deep connection between sum negativity and minimizing Shannon entropy which warrants further exploration.

Due to the connection to contextuality, the bounds on sum negativity and N^p negativity in general are likely to be of interest to the quantum computation community. A natural further direction for this research is the consideration of negativities other than ceiling and sum negativity. Perhaps the next to consider is the only non-convex integral negativity, N^0, which tells us the maximum number of negative elements which can appear in a Q-rep vector. On the other hand, N^2 negativity makes use of the most familiar distance function, and, as such, may warrant special attention. In Section 6.4 we established the general conditions for a $d = 3$ WH Q-rep. This result and any analogous results in higher dimensions[6] may be of independent interest. Likewise, in order to pursue the exact lower bound for sum negativity in dimension 3, we need strategies to construct non-WH Q-reps. Towards this, the general structure and symmetries inherent in Q-reps warrants exploration. It may further be interesting to consider what can be said about quasiprobability representations obtained from non-orthogonal bases or even redundant operator frames in the negativity context.

[6] And in dimension 8, it may be interesting to look at general WH $\otimes$ WH $\otimes$ WH covariant Q-reps.

LIST OF REFERENCES

[1] S. Weinberg, “The Trouble with Quantum Mechanics,” *The New York Review of Books*, 2017.

[2] C. A. Fuchs and B. C. Stacey, “QBism: Quantum Theory as a Hero’s Handbook,” in *Proceedings of the International School of Physics "Enrico Fermi": Course 197, Foundations of Quantum Theory*, pp. 133–202, 2019.

[3] C. A. Fuchs, “QBism, The Perimeter of Quantum Bayesianism,” *arXiv:1003.5209*, 2010.

[4] C. A. Fuchs and R. Schack, “Quantum-Bayesian coherence,” *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[5] B. de Finetti, *Theory of Probability*. Wiley, 1990.

[6] F. Lad, *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*. Wiley-Interscience, New York, 1996.

[7] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*. Wiley, 1994.

[8] C. M. Caves, C. A. Fuchs, and R. Schack, “Unknown quantum states: The quantum de Finetti representation,” *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4537–4559, 2002.

[9] R. W. Spekkens, “Negativity and Contextuality are Equivalent Notions of Nonclassicality,” *Physical Review Letters*, vol. 101, p. 020401, 2008.

[10] C. Ferrie and J. Emerson, “Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations,” *Journal of Physics A: Mathematical and Theoretical*, vol. 41, p. 352001, 2008.

[11] D. Appleby, H. Dang, and C. Fuchs, “Symmetric Informationally-Complete Quantum States as Analogues to Orthonormal Bases and Minimum-Uncertainty States,” *Entropy*, vol. 16, no. 3, pp. 1484–1492, 2014.

[12] D. M. Appleby, “Symmetric informationally complete measurements of arbitrary rank,” *Optics and Spectroscopy*, vol. 103, no. 3, pp. 416–428, 2007.

[13] G. Zauner, *Quantendesigns. Grundzüge einer nichtkommutativen Designtheorie*. PhD thesis, University of Vienna, 1999. Published in English translation: G. Zauner, "Quantum designs: foundations of a noncommutative design theory," *International Journal of Quantum Information*, vol. 9, pp. 445–508, 2011.

[14] C. M. Caves, "Symmetric informationally complete POVMs." `http://info.phys.unm.edu/~caves/reports/infopovm.pdf`, 2002.

[15] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Symmetric informationally complete quantum measurements," *Journal of Mathematical Physics*, vol. 45, pp. 2171–2180, 2004.

[16] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, "The SIC question: History and state of play," *Axioms*, vol. 6, p. 21, 2017.

[17] A. J. Scott, "SICs: Extending the list of solutions," *arXiv:1703.03993*, 2017.

[18] M. Appleby, T.-Y. Chien, S. Flammia, and S. Waldron, "Constructing exact symmetric informationally complete measurements from numerical solutions," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 16, p. 165302, 2018.

[19] M. Appleby, C. A. Fuchs, B. C. Stacey, and H. Zhu, "Introducing the Qplex: a novel arena for quantum theory," *The European Physical Journal D*, vol. 71, no. 7, p. 197, 2017.

[20] A. Khrennikov, "External Observer Reflections on QBism, Its Possible Modifications, and Novel Applications," in *Quantum Foundations, Probability and Information* (A. Khrennikov and B. Toni, eds.), pp. 93–118, Cham: Springer International Publishing, 2018.

[21] A. Khrennikov, "Towards Better Understanding QBism," *Foundations of Science*, vol. 18, pp. 1–15, 2017.

[22] H. Zhu, "Quasiprobability representations of quantum mechanics with minimal negativity," *Physical Review Letters*, vol. 117, no. 12, p. 120404, 2016.

[23] W. P. Schleich, *Quantum Optics in Phase Space*. Wiley-VCH, 2001.

[24] E. Prugovečki, "A Postulational Framework for Theories of Simultaneous Measurement of Several Observables," *Foundations of Physics*, vol. 3, pp. 3–18, 1973.

[25] R. P. Feynman, "Negative probability," in *Quantum Implications: Essays in Honour of David Bohm* (B. Hiley and F. D. Peat, eds.), pp. 235–48, Routledge, 1991.

[26] A. Khrennikov, "On the physical interpretation of negative probabilities in Prugovečki's empirical theory of measurement," *Canadian Journal of Physics*, vol. 75, pp. 291–298, 2011.

[27] A. Khrennikov, *Interpretations of Probability*. Berlin: De Gruyter, second ed., 2009.

[28] M. Burgin, "Interpretations of Negative Probabilities," *arXiv:1008.1287*, 2010.

[29] C. Ferrie, R. Morris, and J. Emerson, "Necessity of negativity in quantum theory," *Physical Review A*, vol. 82, p. 044103, 2010.

[30] W. K. Wootters, "A Wigner-function formulation of finite-state quantum mechanics," *Annals of Physics*, vol. 176, no. 1, pp. 1–21, 1987.

[31] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, "Discrete phase space based on finite fields," *Physical Review A*, vol. 70, no. 6, p. 062101, 2004.

[32] C. Ferrie and J. Emerson, "Framed Hilbert space: hanging the quasi-probability pictures of quantum theory," *New Journal of Physics*, vol. 11, p. 063040, 2009.

[33] C. Ferrie, "Quasi-probability representations of quantum theory with applications to quantum information science," *Reports on Progress in Physics*, vol. 74, no. 11, p. 116001, 2011.

[34] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, "Negative quasi-probability as a resource for quantum computation," *New Journal of Physics*, vol. 14, no. 11, p. 113011, 2012.

[35] M. Howard, J. Wallman, V. Veitch, and J. Emerson, "Contextuality supplies the 'magic' for quantum computation," *Nature*, vol. 510, pp. 351–355, 2014.

[36] H. Pashayan, J. J. Wallman, and S. D. Bartlett, "Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities," *Physical Review Letters*, vol. 115, no. 7, p. 070501, 2015.

[37] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer computation," *New Journal of Physics*, vol. 16, p. 013009, 2014.

[38] R. Bhatia, *Matrix Analysis*. Graduate Texts in Mathematics, Springer, New York, 1997.

[39] D. Bertsekas, *Convex Analysis and Optimization*. Athena Scientific, 2003.

[40] G. N. M. Tabia and D. M. Appleby, "Exploring the geometry of qutrit state space using symmetric informationally complete probabilities," *Physical Review A*, vol. 88, no. 1, p. 012131, 2013.

[41] W. K. Wootters and B. D. Fields, "Optimal state-determination by mutually unbiased measurements," *Annals of Physics*, vol. 191, pp. 363–381, 1989.

[42] A. Szymusiak, "Maximally informative ensembles for SIC-POVMs in dimension 3," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 44, p. 445301, 2014.

[43] B. C. Stacey, “Geometric and Information-Theoretic Properties of the Hoggar Lines,” *arXiv:1609.03075*, 2016.

[44] M. Appleby, S. Flammia, G. McConnell, and J. Yard, “Generating ray class fields of real quadratic fields via complex equiangular lines,” *Acta Arithmetica*, vol. 192, no. 3, pp. 211–233, 2020.

[45] M. Appleby, S. Flammia, G. McConnell, and J. Yard, “SICs and algebraic number theory,” *Foundations of Physics*, vol. 47, pp. 1042–59, 2017.

[46] I. Bengtsson, “The number behind the simplest SIC-POVM,” *Foundations of Physics*, vol. 47, pp. 1031–41, 2017.

[47] C. Ramirez, R. Sanchez, V. Kreinovich, and M. Argaez, “$\sqrt{x^2+c}$ is the Most Computationally Efficient Smooth Approximation to $|x|$,” *Journal of Uncertain Systems*, vol. 8, pp. 205–210, 2014.

[48] R. Pike, *Optimization for Engineering Systems*. Van Nostrand Reinhold, 1986.

[49] A. Szymusiak and W. Słomczyński, “Informational power of the Hoggar symmetric informationally complete positive operator-valued measure,” *Physical Review A*, vol. 94, p. 012122, 2015.

CHAPTER 7

WIGNER'S FRIENDS

7.1 Introduction

Wigner's famous thought experiment [1] is a tale of two agents. One agent, Wigner's friend, performs a quantum measurement in a lab and obtains an outcome. The other agent, Wigner, treats the lab containing his friend and the friend's experimental setup as one large quantum system and writes down a joint quantum state which evolves continuously in time. Thus for Wigner there is no measurement outcome. Who is right, Wigner or his friend? There is a difficulty here if one thinks of a measurement outcome as something objective in the sense that it can be verified in principle by anybody.

The lesson QBism [2–4] draws from Wigner's thought experiment is that, for consistency's sake, measurement outcomes must be regarded as personal to the agent who makes the measurement. This idea first appeared in 2006 in Ref. [5], where the phrase "facts for the agent" was coined. Its connection with Wigner's friend was eventually spelled out fully in 2010 with Ref. [2], and the personal nature of an agent's measurement outcomes was further emphasized in Ref. [4], where outcomes were identified with the "experiences" of the agent doing the measurement. As emphasized, e.g., by Pusey [6], Wigner's friend thought experiments thus pose no problem for QBism. In fact, Wigner's friend was central to the development of QBist thinking [7].

Recently, variations of Wigner's original thought experiment were introduced by Brukner [8, 9], Frauchiger and Renner (FR) [10], and Baumann and Brukner (BB) [11].[1] In Brukner's thought experiment, first described in 2015 [8] and thoroughly analyzed in Ref. [9], the assumption that measurement outcomes are objective leads to a Bell inequality and thus to a conflict with the predictions of quantum mechanics. Brukner concludes from this that measurement outcomes should be regarded as "facts relative to the observer," the same conclusion QBism reached by considering the original Wigner's friend thought experiment.

The main innovation in the very interesting FR scenario is that both Wigner and a friend make predictions about the outcome of one and the same measurement, which is performed by Wigner. A seemingly straightforward application of the quantum formalism then appears to show that the predictions of Wigner and friend are mutually contradictory. Frauchiger and Renner turn this into a formal contradiction by postulating a small number of what they believe to be intuitive assumptions. They conclude that any interpretation of quantum mechanics has to abandon at least one of these assumptions.

From a QBist perspective, however, there is a fundamental problem with Frauchiger and Renner's analysis. In their thought experiment, both Wigner and the friend are agents applying the quantum formalism, but Frauchiger and Renner treat them in an asymmetric way. Because Wigner is outside the lab that contains the friend, this asymmetry seems to be inherent in the very setup of the experiment. We will show that this is not so. QBism both requires and makes possible a fully symmetric treatment of Wigner and his friend. Wigner's action on his friend then becomes, from the friend's perspective, an action the

[1] There are too many responses to these papers to cite here, but a sampling of those which attempt to analyse QBism's relation to the thought experiments can be found in Refs. [12–20]. Though Refs. [19, 20] are both very relevant to QBist interests, neither of these get at the heart of the argument made here.

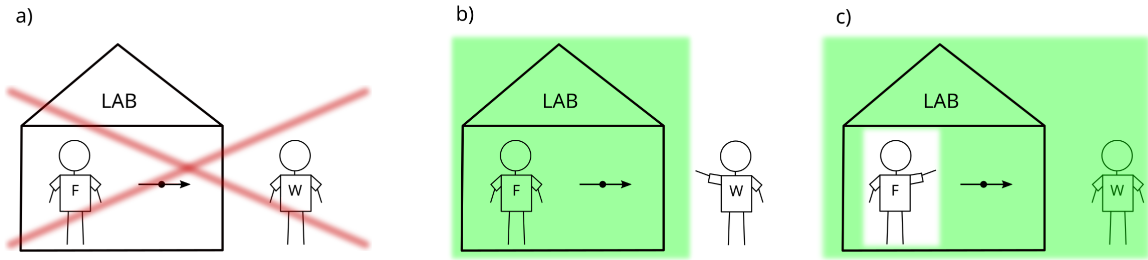

Figure 7.1: (a) In usual descriptions of various Wigner's-friend thought experiments, there is an urge to portray everything from a God's eye view. Here, we depict the BB thought experiment, where Wigner, his friend, and a spin-1/2 particle interact, and we symbolize QBism's disapproval of such portrayals with a big red X. In QBism, the quantum formalism is only used by agents who stand *within* the world; there is no God's-eye view. (b) Instead, in QBism, to make predictions, Wigner treats his friend, the particle, and the laboratory surrounding her (all shaded in green) as a physical system external to himself. While (c) to make her own predictions, the friend must reciprocally treat Wigner, the particle, and her surrounding laboratory (all shaded in green) as a physical system external to herself. It matters not that the laboratory spatially surrounds the friend; it, like the rest of the universe, is external to her agency, and that is what counts.

friend takes on Wigner. Once this is taken into account, the paradoxical features of the FR thought experiment disappear.

The more recent BB scenario is similar to the FR thought experiment in that, again, both Wigner and friend make predictions about the outcome of Wigner's measurement. Baumann and Brukner appear to show that applying the standard quantum formalism leads the friend to make a bad prediction. As in the FR case, the problem with Baumann and Brukner's analysis is that they fail to treat the friend as an agent on the same footing as Wigner. If, instead, Wigner and his friend are treated symmetrically, the BB scenario loses its seemingly paradoxical character. Because the BB scenario is much simpler than the FR scenario, we will discuss it first. Indeed it was through thinking about the BB thought experiment that we finally arrived at our present understanding of the more intricate FR thought experiment.

Our paper is organized as follows. In Section 7.2 we summarize the main QBist principles. We spell out what we mean by a user of the quantum formalism, and how quantum states and quantum measurements are thought of as personal judgments in our framework. We further explain how Asher Peres's dictum that unperformed experiments have no results remains true even when an agent is certain of what he will find. In Section 7.3 we review Wigner's original thought experiment and explain what it means for one agent to be a physical system for another agent, distinguishing our notion from Wigner's original where he argued that the friend must be described as in a "suspended animation" unless the laws of physics are changed. Section 7.4 contains the main argument of the paper. It shows that the BB thought experiment can be understood fully within the standard quantum formalism if Wigner and his friend are treated in a fully symmetric fashion. We turn to the FR thought experiment in Section 7.5. We show that, exactly as in the BB case, the apparent contradiction derived by Frauchiger and Renner is due to a failure to treat one of the participants in the thought experiment as an agent in the full sense of the word. Finally we clarify the circumstances in which one agent may adopt another agent's quantum state assignments and thereby address a challenge posed by Frauchiger and Renner [10] in their subsection titled "Analysis within QBism."

7.2 Agents and QBism

As there exist several authoritative accounts of QBism [21–23], this section focuses on those aspects of QBism that are important for our argument about agents and Wigner's friend. We start by defining the terms "agent" and "user of quantum mechanics" and discuss some key tenets of QBism. We then give an account of the elementary double-slit experiment in QBist terms, in order to set the scene for the discussion of Wigner's friend in the next section.

7.2.1 Agents and users of quantum mechanics

According to QBism, the quantum formalism is a tool decision-making agents can adopt to better guide their decisions when faced with the inevitable uncertainties of the quantum world. Particularly, the theory guides its users in how to gamble on the personal consequences of their measurement actions. Thus for QBism, the quantum formalism plays a *normative* role for its users, not a descriptive role for exactly how the world is: It suggests how a user *should* gamble.

Users of the theory are thus at the center of the QBist approach. It is therefore important to spell out what we mean by the term. In the following, we will make a distinction between agents and users of quantum mechanics:

- *Agents* are entities that can take actions freely on parts of the world external to themselves, so that the consequences of their actions matter for them.
- A *user of quantum mechanics* is an agent that is capable of applying the quantum formalism normatively.

While our definition of a user is narrow, our definition of an agent is broad: It does not rule out attributing agency to dogs, euglenas, or artificial life. However, it does exclude a computer program that deterministically "chooses" an action from a look-up table. On the other hand, as Khrennikov emphasizes in Ref. [24], "The idea is that QM is something used only by a privileged class of people. Those educated in the methods of QM are able to make better decisions (because of certain basic features of nature) than those not educated in the methods of QM." This notion of a user of the theory is sufficiently open to allow for additional details in the future, but it is also precise enough for the purposes of this paper.

There exists a range of definitions of agency in the philosophical literature that overlap with our definition to different degrees [25, 26]. According to the above definitions, a team of scientists sharing notebooks, calculations, observations, etc., can act as a single agent and even a user of quantum mechanics [27].

7.2.2 Some tenets of QBism

The previous subsection started with a one-paragraph summary of QBism. The following five tenets provide more detail. Taken together they ensure QBism's consistency.

What is a measurement?

A measurement is an action of an agent on its external world, where the consequences of the action, or its *outcomes*, matter to the agent.

Like our definition of agent before, this definition of measurement is very broad. Basically anything an agent can do to its external world—from opening a box of cookies, to crossing a street, to performing a sophisticated quantum optics experiment—counts as a measurement in our sense. The only thing that sets quantum measurement as normally construed apart from the more mundane examples given above is whether it is fruitful or worth one's while to apply the quantum formalism to guide one's actions. But, in principle a user of quantum mechanics could use the formalism to make decisions in any measurement situation, including measurements on living systems as in the Wigner's friend thought experiment that are of concern here.

By applying the term measurement only to actions on the agent's external world, we exclude the case where an agent, directly or indirectly, acts on him or herself. We thus require a strict separation between the agent performing the measurement and the measured system.

Measurement outcomes are personal

When an agent performs a measurement—that is, takes an action on its external world—the "outcome" of the measurement is the consequence of this action for the agent. A measurement outcome is personal to the agent doing the measurement. Thus two agents cannot experience the same outcome. Different agents may inform each other of their outcomes and thus agree upon the consequences of a measurement, but a measurement outcome should not be viewed as an agent-independent fact which is available for anyone to see [5].

This tenet has led some commentators to claim mistakenly that QBism is a form of solipsism. This claim has been thoroughly refuted (see, e.g., Ref. [2, 28–31]). That QBism is not solipsism follows immediately from the premise that a measurement is an action on the *world external to the agent*. A QBist assumes the existence of an external world from the outset. Furthermore, the consequences of measurement actions are beyond the agent's control—the world can surprise the agent. The world is thus capable of genuine novelty complementary to the agent's actions—the world and the agent cannot be identified with each other. (See Refs. [30, pp. 6–10] and [22, pp. 19–20], arXiv versions.)

A quantum state is an agent's personal judgment

In QBism, the only purpose of the quantum formalism is to help an agent make better decisions. Rigorous use of the formalism enables an agent to make more successful gambles. The term "gamble" evokes games of luck, but here it is meant to encompass any action by an agent where the consequences matter to the agent. Any physics experiment is thus a gamble in this sense.

As we will explain in more detail in the next subsection, the quantum formalism can be viewed as an addition to classical decision theory [3, 22]. Following the approach to decision theory pioneered by Savage [32, 33], QBism takes all probabilities, including those equal to zero and one, to be an agent's personal degrees of belief concerning future measurement outcomes. Personalist probabilities [34, 35] acquire an operational meaning by their use in decision making. A key consequence of this theory is that, to avoid sure loss, an agent's gambles must be constrained by the rules of probability theory.

In QBism, a quantum state is also an agent's personal judgment, reflecting the agent's degrees of belief in the outcomes of all possible measurements he or she might perform. A quantum state, rather than being a property of a quantum system, thus encodes an agent's expectations regarding the outcomes of future measurements.

The quantum formalism is normative rather than descriptive

We will see below that the quantum-mechanical Born rule can be viewed as placing additional constraints on an agent's probability assignments to the outcomes of different measurements [5] in situations where pure probability theory is simply silent. In line with the central place that QBism gives to measurement, QBism treats the Born rule as fundamental: To understand the quantum formalism, one has to understand the Born rule first.

For a measurement with outcomes labeled $j = 1, \ldots, n$, the Born rule is usually given in the form $p_j = \mathrm{tr}(\rho E_j)$, where E_j is a measurement operator or *effect* corresponding to outcome j (for a von Neumann measurement this will be a projection operator), ρ is the density operator for the measured system, and p_j is the probability for outcome j. In QBism, ρ represents the agent's belief about the system, and the list of effects (or POVM) $\{E_1, \ldots, E_n\}$ represents the agent's belief about the measurement.

In contrast to the usual reading of the Born rule as a formula for computing p_j given ρ and E_j, in QBism the Born rule functions as a consistency requirement [2, 3]. If an agent has beliefs p_j, ρ and E_j that do not satisfy the Born rule, he or she should modify at least one of these beliefs. The formalism does not prescribe which one to modify or how to modify it.[2]

In QBism the Born rule is thus a consistency criterion that an agent should strive to satisfy in its probability and quantum state assignments. It is a single-agent criterion; it says nothing about consistency between the probability and quantum state assignments of different agents. It is entirely about internal consistency of an agent's expectations. This is what is meant by saying that the Born rule, and thus the quantum formalism, is normative rather than descriptive.

This tenet has led some commentators to claim mistakenly that QBism is a form of instrumentalism. This, as with the claim of solipsism, is also easily refuted; see, e.g., Refs. [30, 37]. Indeed from its earliest days, the very goal of QBist research has been to distill a statement about the character of the world from the fact that gambling agents should use the quantum formalism [38]. Even though this remains an ongoing project, it has already led to a number of strong ontological claims on the part of QBism—from the world being capable of genuine novelty and being in constant creation, to the Born rule expressing a novel form of structural realism [22, 30].

[2] It is easier to see how p_j, ρ, and E_j are on an equal footing if ρ and E_j are expressed as probabilities. That this can be done is well known: with respect to an appropriately chosen informationally complete measurement, any density operator is equivalent to a vector of probabilities [36], and any POVM $\{E_1, \ldots, E_n\}$ is characterized by a stochastic matrix of conditional probabilities [3].

Probability-1 assignments are judgments

QBism regards even probability-1 (and probability-0) assignments as an agent's personal judgments. Assigning probability-1 to an outcome expresses the agent's supreme confidence that the outcome will occur, but does not imply that anything in nature guarantees that the outcome *will* occur.

Similarly, QBism regards both pure and mixed quantum states as an agent's personal judgments. This implies in particular that even a statement such as "this outcome is certain to occur" reflects an agent's judgment rather than a fact of nature. In other words, nothing in nature guarantees that an outcome to which an agent has assigned proobability-1 will in fact occur.

7.2.3 Unperformed measurements have no outcomes, even when an agent is certain what the outcome will be

In his 1964 *Lectures on Physics*, Richard Feynman famously stated that the double-slit experiment exhibits "the basic peculiarities of all quantum mechanics". "In reality, it contains the *only* mystery," he said. This idea seems to overlook the importance of quantum-foundational results such as Bell inequalities, Kochen-Specker style noncolorability theorems, or contextuality inequalities. Yet, in the end QBism believes Feynman was on the right track. Not only does the double-slit experiment exhibit the basic peculiarities of quantum mechanics, it points to the solution of the Wigner's friend conundrum as well.

The double-slit experiment consists of a particle source, a screen with two slits, and a second screen farther away from the source where the particle position is recorded. Assume an agent has made assignments $P(H_0)$ and $P(H_1)$ for the probability that the particle passes through the left or right slit, respectively, and $P(D_j|H_0)$ and $P(D_j|H_1)$ that the particle is detected at position D_j given that it passes through the left or right slit, respectively. We assume $P(H_0) + P(H_1) = 1$ and $\sum_j P(D_j|H_0) = \sum_j P(D_j|H_1) = 1$, as must be the case for probabilities and conditional probabilities. As explained above, for these probabilities to have operational, decision-theoretic, meaning they have to refer to actual outcomes for the agent. Spelled out, this means $P(H_0)$ is the probability that the agent sees the particle pass through the left slit, allowing the agent to gamble on this outcome, and $P(D_j|H_0)$ is the probability that the agent detects the particle at D_j given that he or she has seen it pass through the left slit, allowing the agent to make the corresponding conditional gamble. (The same applies, of course, to the right slit.)

In the uncontroversial case where the agent actually intends to check which slit the particle passes through before it hits the second screen, the agent's probability $P(D_j)$ for finding the particle at D_j is given by

$$P(D_j) = P(H_0)P(D_j|H_0) + P(H_1)P(D_j|H_1) \; . \tag{7.1}$$

This follows from probability theory alone.

But what if the agent does not intend to check which slit the particle passes through? In this case we are dealing with a *different experiment* for which probability theory alone does not constrain the agent's probabilities. The classical intuition in this situation is to continue to use Eq. (7.1) for the probability of detecting the particle at D_j, which amounts to the physical postulate that as far as this probability is concerned, it does not matter whether the agent does or does not check which slit the particle goes through.

This classical intuition has to be abandoned in quantum mechanics. Whether a measurement is performed or not matters profoundly. Asher Peres expressed this in his famous slogan "Unperformed experiments have no results." [39]

Crucially, in a quantum analysis of the double-slit experiment, there is no change in either the values or the meaning of the probabilities $P(H_0)$, $P(H_1)$, $P(D_j|H_0)$, and $P(D_j|H_1)$. For instance, $P(D_j|H_0)$ is still the agent's probability to detect the particle at D_j given that the agent saw it pass through the left slit. In his paper "The Concept of Probability in Quantum Mechanics" [40], Feynman makes a similar point and then goes on to write that "[w]hat is changed, and changed radically, is the method of calculating probabilities."

Here QBism departs from Feynman's view in one significant way. In the uncontroversial case that the agent intends to do an intermediate measurement to check which slit the particle goes through, the quantum rules do not lead to a change of the probabilities $P(D_j)$ as given in Eq. (7.1). If the agent does not intend to do the intermediate measurement, however, Eq. (7.1) no longer applies because the probabilities $P(H_0)$, $P(H_1)$, $P(D_j|H_0)$, and $P(D_j|H_1)$ now refer to a *hypothetical* intermediate measurement. Thus probability theory alone no longer gives a formula for the probability of finding the particle at D_j. The "radically" new quantum method of calculating probabilities in this case should therefore not be viewed as a change of, but as an addition to, existing methods. Probability theory remains fully valid in the quantum realm.

As we stated in Section 7.2.2, QBism takes the stand that even when an agent assigns probability-1 to one of the possible outcomes of a measurement, there is nothing in the agent's external world that metaphysically ensures it will necessarily come about. For "unperformed measurements have no outcomes" is a statement about the assumed character of the world, whereas a probability-1 assignment is only a belief (supremely strong, but nonetheless a belief) that someone happens to have in the moment. That unperformed measurements have no results is, for QBism, the great lesson of all the Bell-inequality and Kochen-Specker-theorem results of the last half century, more recently reinforced by the "no-go theorems" of Pusey, Barrett, and Rudolph [41] and Colbeck and Renner [42]. It plays a central role in the QBist approach to Wigner's friend.

7.3 Wigner's Friend

Wigner described his thought experiment in a 1961 paper entitled "Remarks on the Mind-Body Question" [1]. Below we use a slightly modernized version of Wigner's notation. The friend (who prefers the pronouns "she" and "her") performs a two-outcome measurement on a quantum system, where the outcomes correspond to the states $|\psi_1\rangle$ and $|\psi_2\rangle$, respectively. In order to be consistent with the BB scenario discussed in Section 7.4 below, we assume that $|\psi_1\rangle$ and $|\psi_2\rangle$ are states of a spin-1/2 particle corresponding to "spin up" and "spin down", respectively. After the friend's measurement, Wigner contemplates a simple measurement on her, consisting of the question: what was the result of your spin measurement?

The assumption is now that Wigner assigns a quantum state to the joint system consisting of particle and his friend, and treats it as a closed quantum system. After the friend has measured the spin, Wigner's joint state becomes

$$|\Phi\rangle = \alpha|\psi_1\rangle|\chi_1\rangle + \beta|\psi_2\rangle|\chi_2\rangle \ , \tag{7.2}$$

where $|\chi_1\rangle$ and $|\chi_2\rangle$ are Wigner's states for his friend and correspond to her responding to the question "what was the result of your spin measurement" with "up" and "down", respectively.

In his 1961 paper, Wigner argued that the question of whether the friend saw "up" or "down" was already decided in her mind, before Wigner asked her. Here is what Wigner concludes from this, in his own words (excepting our modernized notation):

If we accept this, we are driven to the conclusion that the proper wave function immediately after the interaction of friend and object was already either $|\psi_1\rangle|\chi_1\rangle$ or $|\psi_2\rangle|\chi_2\rangle$ and not the linear combination $\alpha|\psi_1\rangle|\chi_1\rangle+\beta|\psi_2\rangle|\chi_2\rangle$. This is a contradiction, because the state described by the wave function $\alpha|\psi_1\rangle|\chi_1\rangle + \beta|\psi_2\rangle|\chi_2\rangle$ describes a state that has properties which neither $|\psi_1\rangle|\chi_1\rangle$ nor $|\psi_2\rangle|\chi_2\rangle$ has. If we substitute for "friend" some simple physical apparatus, such as an atom [...], this difference has observable effects and *there is no doubt that* $\alpha|\psi_1\rangle|\chi_1\rangle + \beta|\psi_2\rangle|\chi_2\rangle$ *describes the properties of the joint system correctly, the assumption that the wave function is either* $|\psi_1\rangle|\chi_1\rangle$ *or* $|\psi_2\rangle|\chi_2\rangle$ *does not* [Wigner's italics]. If the atom is replaced by a conscious being, the wave function $\alpha|\psi_1\rangle|\chi_1\rangle + \beta|\psi_2\rangle|\chi_2\rangle$ (which also follows from the linearity of the equations) appears absurd because it implies that my friend was in a state of suspended animation before [she] answered my question.

It follows that the being with a consciousness must have a different role in quantum mechanics than the inanimate measuring device: the atom considered above. In particular, the quantum mechanical equations of motion cannot be linear if the preceding argument is accepted. This argument implies that "my friend" has the same types of impressions and sensations as I—in particular, that, after interacting with the object, [she] is not in that state of suspended animation which corresponds to the wave function $\alpha|\psi_1\rangle|\chi_1\rangle + \beta|\psi_2\rangle|\chi_2\rangle$.

QBism, along with most other current interpretations of quantum mechanics, does not restrict the applicability of quantum mechanics to inanimate devices. An agent can apply the normative quantum calculus to any part of the world external to him or herself, including conscious beings and other agents. So how does QBism escape the conclusions that Wigner draws from his thought experiment?

Part of the answer is straightforward and follows directly from the QBist tenets. Wigner's quantum state $|\Phi\rangle$ is not descriptive: it does not describe properties of the joint system to which it refers. The exclusive role of $|\Phi\rangle$ is to help Wigner quantify his expectations regarding the consequences of his future actions on friend and particle. In the same way, the friend's state assignments refer to her expectations regarding the consequences of her future actions. There is simply no conflict between these two perspectives. In particular, Wigner's state assignment has no bearing on whether the friend is or is not in a "state of suspended animation".

This straightforward resolution of Wigner's paradox has profound implications for the QBist worldview. A few lines below the quoted passage, Wigner points out that insisting on the superposition state $|\Phi\rangle$ for particle and friend, though not necessarily a contradiction, amounts to denying "the existence of the consciousness of a friend" to an intolerable extent. In QBism, a quantum state assignment has no bearing on the existence of the consciousness of a friend. It follows that a QBist can simultaneously assign the state $|\Phi\rangle$ and grant his friend a conscious experience of having seen either "up" or "down".

This claim requires some elaboration. The scenario of Wigner's friend can be understood as a version of the double-slit experiment, in line with Feynman's dictum that the latter contains the basic peculiarities of all quantum mechanics. As in our analysis of the double-slit experiment, none of the probabilities considered in Wigner's paper change their value or their meaning when Wigner writes down his quantum state $|\Phi\rangle$. The only implications of Wigner assigning a quantum state to the friend are that, (i) as far his probabilities for the outcomes of some future quantum measurement on the friend are concerned, it matters whether or not he first asks her whether she saw up or down, and (ii) that he should use the Born rule to compute these probabilities. There is no reason why Wigner cannot assign a quantum state that respects all of his beliefs about the friend's inner life, conscious experiences, or agenthood.

The parallel with the double-slit experiment is somewhat hidden in Wigner's original argument, because he only considers measurements on the friend that, in the double-slit experiment, correspond to determining which slit the particle went through. But, as Wigner makes clear when he writes that "there is no doubt that [$|\Phi\rangle$] describes the properties of the joint system correctly", assigning the state $|\Phi\rangle$ amounts to committing to predictions for the outcomes of a wide range of quantum measurements on the friend, including those for which the predictions depend on whether or not the friend is first asked what she saw. In the context of the Wigner's friend scenario, such measurements were first considered by David Deutsch [43]. They are crucial for the BB thought experiment, to which we will now turn.

7.4 The Friend's Perspective: Response to Baumann & Brukner

Baumann and Brukner's thought experiment [11] is a simple modification of Wigner's original scenario. After the friend has made her measurement, Wigner's joint state of particle and friend is again given by Eq. (7.2), where it is now assumed that $\alpha = \beta = 1/\sqrt{2}$, so that we have

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|\psi_1\rangle|\chi_1\rangle + |\psi_2\rangle|\chi_2\rangle\Big) \, . \tag{7.3}$$

Whereas in Wigner's paper, Wigner contemplates a simple measurement on the friend consisting of asking her about the result of her spin measurement, Baumann and Brukner let Wigner do the measurement

$$M_W : \{|\Phi\rangle\langle\Phi|, 1 - |\Phi\rangle\langle\Phi|\} . \tag{7.4}$$

Such a measurement is far beyond any current and probably future experimental possibilities, but if we allow Wigner to write down the state $|\Phi\rangle$, we must also allow him to contemplate the measurement M_W. Clearly, Wigner has a probability $p = 1$ of obtaining the outcome corresponding to $|\Phi\rangle\langle\Phi|$, which is labeled "+" in Ref. [11].

Baumann and Brukner's main claim concerns the friend's prediction for the "+" outcome. They argue that, if in her spin measurement the friend obtains "up", her probability of "+" is given by applying M_W to the state $|\psi_1\rangle|\chi_1\rangle$, and if she obtains "down", she should apply M_W to the state $|\psi_2\rangle|\chi_2\rangle$. In both cases, her probability for "+" is 1/2. Since this probability is the same for "up" and "down", she can communicate her prediction to Wigner without affecting the rest of the experiment. Baumann and Brukner's claim thus leads to the troubling conclusion that two different ways of applying the rules of quantum mechanics give contradictory numbers for the probability of "+". [3]

In the above account it might appear problematic that, in QBism, the outcome of the measurement M_W is personal to Wigner. But Baumann and Brukner show a valid way around this problem by stipulating that Wigner record the outcome of his measurement on a piece of paper. The friend's probability assignment can then be regarded as referring to her finding "+" upon checking the piece of paper.

The real problem with the BB analysis is that for the friend to base her prediction on the state $|\psi_1\rangle|\chi_1\rangle$ (or $|\psi_2\rangle|\chi_2\rangle$) amounts to assigning a quantum state to herself, which vi-

[3] Baumann and Brukner assume the experiment is repeated many times, so that the alleged contradiction can be phrased in frequentist terms. From a QBist perspective, the full force of the contradiction arises already in the single-case analysis given here.

olates the QBist tenet that there must be a clear separation between agent and measured system. It is easy to see why this leads to a serious difficulty. Assume for the moment that the friend writes down $|\psi_1\rangle|\chi_1\rangle$ for the joint system of particle and herself and uses it to compute her probability for what she will see on the piece of paper. This state assignment would not just commit her to a probability for outcome "+". It would commit her to probabilities for any measurement that she could perform on the particle and herself. For instance, according to our discussion in Section 7.3, the state $|\chi_1\rangle$ corresponds to her responding with "up" to the question "what was the result of your spin measurement". But since she is a free agent, she has control over the answer to this question. It is up to her whether she replies "up", "down", or by sticking her tongue out. Since she has at least partial control over these measurement outcomes, the above quantum-state assignment cannot form a reliable basis for guiding her actions.

So what should the friend do instead? The answer has already been given in the Introduction and Figure 1. Rather than adopting Wigner's viewpoint, she needs to analyze the experiment as an action that she takes on the particle, the lab, Wigner, and the piece of paper on which Wigner records his outcome. In particular, this requires her reversing roles and treating Wigner as a physical system. That this would be an enormously complex and practically infeasible task is hardly a valid objection given the assumption that Wigner is able to write down a quantum state for her. Indeed, an even-handed analysis of the thought experiment clearly requires the assumption that the friend is as skillful a user of quantum mechanics as Wigner himself.

What probability should the friend assign to her finding "+" on the piece of paper? This depends on her prior states, unitaries, and POVMs regarding the lab and Wigner. The only constraint on her probability for "+" is that it should be consistent with her prior quantum assignments in the sense given in Section 7.2.2, i.e., it should be consistent with the Born rule. This implies that the friend's probability for "+" cannot be derived from the details provided in the BB thought experiment. Furthermore, if the experiment is repeated many times as envisaged by Baumann and Brukner, the formalism will typically lead her to update her assignments after each repetition. Her probabilities will thus reflect what she learned in previous runs of the experiment.

Here is a summary of our argument. In the same way that Wigner does not take the friend's viewpoint into account when he computes his probability of "+", the friend need not take Wigner's viewpoint into account when she computes her probability of "+". This puts Wigner and the friend on an equal footing. In particular, the friend's quantum state assignments is not a function of Wigner's quantum states $|\chi_1\rangle$ and $|\chi_2\rangle$. We thus explicitly reject Baumann and Brukner's claim that standard quantum theory requires the friend to base her probability assignments on $|\chi_1\rangle$ and $|\chi_2\rangle$. In contrast to Baumann and Brukner, who propose that the friend uses a modified Born rule incorporating Wigner's perspective, our symmetric QBist treatment of Wigner and his friend requires no such modification.

7.5 Reasoning about Other Agents: Response to Frauchiger & Renner

Frauchiger and Renner's thought experiment, which considers four agents, is somewhat intricate, but here only the following broad outline is needed. Two agents, F (the friend) and $\bar{\mathsf{F}}$, are located in separate labs. The other two agents, W (Wigner) and $\bar{\mathsf{W}}$, are on the outside and perform measurements on the labs. At time $t = 0$, agent $\bar{\mathsf{F}}$ prepares a qubit in a given state, measures it, prepares a spin-1/2 particle in a state that depends on the measurement outcome, and sends the particle to agent F's (the friend's) lab. At $t = 1$, the friend measures the particle. At $t = 2$, agent $\bar{\mathsf{W}}$ measures $\bar{\mathsf{F}}$'s lab in a given basis. Finally, at $t = 3$, Wigner measures the friend's lab in a given basis.

Wigner now uses two different methods to make predictions for the outcome of his measurement. He is interested in the probability of one of the outcomes, labeled w=fail. His first prediction uses the quantum formalism. For his second prediction, he reasons about what predictions the other agents would have made at earlier stages of the experiment, assuming all agents start from the same initial quantum state assignment. Frauchiger and Renner argue that the two methods lead to mutually contradictory predictions.

When reasoning about other agents, Wigner, agent $\bar{\mathsf{W}}$, and the friend apply FR's

> **Assumption (C)**: Suppose that agent A has established that "I am certain that agent A′, upon reasoning within the same theory as the one I am using, is certain that $x = \xi$ at time t." Then agent A can conclude that "I am certain that $x = \xi$ at time t."

The FR argument starts with agent $\bar{\mathsf{F}}$ making, immediately after time $t = 0$, a prediction about Wigner's measurement outcome. Since the measured system is the friend's lab, agent $\bar{\mathsf{F}}$'s prediction is about a part of her external world. The next step is that the friend applies Assumption (C) to make $\bar{\mathsf{F}}$'s prediction her own. Then $\bar{\mathsf{W}}$ applies Assumption (C) to make the friend's and thus $\bar{\mathsf{F}}$'s prediction his own. Finally Wigner applies Assumption (C) to make $\bar{\mathsf{W}}$'s and thus also the friend's and $\bar{\mathsf{F}}$'s prediction his own. These steps are cleverly arranged in time so that they don't clash with the different measurements.

Notice that all four agents' predictions concern the same outcome, namely w=fail in Wigner's measurement of the lab containing the friend. This means in particular that the FR argument depends on the friend making a prediction about Wigner's measurement on herself. In their table 3, Frauchiger and Renner make this explicit by stating the friend's conclusion as "I am certain that [Wigner] will observe w=fail at time [$t = 3$]."

But this means we are now in the same situation as in the previous section when we analyzed the BB thought experiment. The friend *can* use the quantum formalism to make a prediction for Wigner's outcome (more precisely, for what she will find when she checks a record of Wigner's outcome). But she is not required to base her prediction on agent $\bar{\mathsf{F}}$'s or Wigner's state assignments. She will have to analyze the experiment as an action that she takes on the other lab and the other agents. Similar to the discussion of the BB thought experiment, the contradiction derived by FR is resolved if the symmetry between Wigner and his friend is recognized.

Frauchiger and Renner state correctly that Assumption (C) is rejected by QBism. This does not mean that there is a prohibition in QBism for one agent to adopt another agent's probability or state assignments. A QBist agent will have to decide on a case by case basis whether or not to do so. A straightforward way of making use of other agents' probabilities in one's decision making is simply to ask them what their probabilities are and to treat the answers as data which they may or may not take into account in their own probabilities.

The most common scenario in which scientists adopt each others' probability and state assignments is that of a team working jointly on a quantum experiment and acting as a single agent and user of quantum mechanics. It follows from the definitions given in Section 7.2.1 that scientists in such a team must have common probability and state assignments. The requirement of a strict separation between agent and measured system now translates into a strict separation of team and measured system. For the FR thought experiment this means that its four players cannot be thought of as acting as a single agent, because they perform measurements on each other.

In a subsection of their paper, titled "Analysis within QBism", Frauchiger and Renner write that "Nevertheless, there should be ways for agents to consistently reason about each other." In this paper we have provided such a way. For two users of quantum mechanics who interact, it requires each of them to treat their interaction as an action he or she takes freely on the other.

7.6 Conclusion

We have seen that the thought experiments described by Frauchiger and Renner and by Baumann and Brukner have a key aspect in common. In each of them, an agent (the friend) is using quantum mechanics to predict the consequences of an action performed on her by another agent (Wigner). We have shown that the paradoxa found by these authors disappear if the friend analyses the experiment as an action she performs on the world outside herself, which includes Wigner. These thought experiments thus illustrate what we have called a quantum Copernican principle: when two agents take actions on each other, each agent has a dual role as a physical system for the other agent.

LIST OF REFERENCES

[1] E. P. Wigner, "Remarks on the Mind-Body Question," in *The Scientist Speculates* (I. J. Good, ed.), pp. 284–302, London: William Heinemann, Ltd., 1961.

[2] C. A. Fuchs, "QBism, The Perimeter of Quantum Bayesianism," *arXiv:1003.5209*, 2010.

[3] C. A. Fuchs and R. Schack, "Quantum-Bayesian coherence," *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[4] C. A. Fuchs, N. D. Mermin, and R. Schack, "An introduction to QBism with an application to the locality of quantum mechanics," *American Journal of Physics*, vol. 82, no. 8, pp. 749–754, 2014.

[5] C. M. Caves, C. A. Fuchs, and R. Schack, "Subjective probability and quantum certainty," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, vol. 38, no. 2, pp. 255–274, 2007.

[6] M. F. Pusey, "An inconsistent friend," *Nature Physics*, vol. 14, no. 10, pp. 977–978, 2018.

[7] C. A. Fuchs, M. Schlosshauer, and B. C. Stacey, "My Struggles with the Block Universe," *arXiv:1405.2390*, 2014.

[8] Č. Brukner, "On the Quantum Measurement Problem," in *Quantum [Un]Speakables II* (R. Bertlmann and A. Zeilinger, eds.), pp. 95–117, Cham: Springer International Publishing, 2017.

[9] Č. Brukner, "A No-Go Theorem for Observer-Independent Facts," *Entropy*, vol. 20, no. 5, p. 350, 2018.

[10] D. Frauchiger and R. Renner, "Quantum theory cannot consistently describe the use of itself," *Nature Communications*, vol. 9, no. 1, p. 3711, 2018.

[11] V. Baumann and Č. Brukner, "Wigner's Friend as a Rational Agent," in *Quantum, Probability, Logic* (M. Hemmo and O. Shenker, eds.), pp. 91–99, Cham: Springer International Publishing, 2020.

[12] R. Healey, "Quantum Theory and the Limits of Objectivity," *Foundations of Physics*, vol. 48, no. 11, pp. 1568–1589, 2018.

[13] N. Nurgalieva and L. del Rio, "Inadequacy of Modal Logic in Quantum Settings," *Electronic Proceedings in Theoretical Computer Science*, vol. 287, pp. 267–297, 2019.

[14] M. Proietti, A. Pickston, F. Graffitti, P. Barrow, D. Kundys, C. Branciard, M. Ringbauer, and A. Fedrizzi, "Experimental test of local observer independence," *Science Advances*, vol. 5, no. 9, p. eaaw9832, 2019.

[15] R. Krismer, "Representation Lost: The Case for a Relational Interpretation of Quantum Mechanics," *Entropy*, vol. 20, no. 12, p. 975, 2018.

[16] F. J. Boge, "Quantum Information Versus Epistemic Logic: An Analysis of the Frauchiger–Renner Theorem," *Foundations of Physics*, vol. 49, no. 10, pp. 1143–1165, 2019.

[17] A. Sudbery, "The hidden assumptions of Frauchiger and Renner," *arXiv:1905.13248*, 2019.

[18] P. W. Evans, "Perspectival objectivity: Or: how I learned to stop worrying and love observer-dependent reality," *European Journal for Philosophy of Science*, vol. 10, no. 2, p. 19, 2020.

[19] A. Suarez, "The limits of quantum superposition: Should "Schrödinger's cat" and "Wigner's friend" be considered "miracle" narratives?," *arXiv:1906.10524*, 2019.

[20] B. C. Stacey, "On QBism and Assumption (Q)," *arXiv:1907.03805*, 2019.

[21] C. A. Fuchs and B. C. Stacey, "QBism: Quantum Theory as a Hero's Handbook," in *Proceedings of the International School of Physics "Enrico Fermi": Course 197, Foundations of Quantum Theory*, pp. 133–202, 2019.

[22] C. A. Fuchs, "Notwithstanding Bohr, the Reasons for QBism," *Mind and Matter*, vol. 15, no. 2, 2017.

[23] B. C. Stacey, "Ideas Abandoned en Route to QBism," *arXiv:1911.07386*, 2019.

[24] A. Khrennikov, "Towards Better Understanding QBism," *Foundations of Science*, vol. 23, no. 1, pp. 181–195, 2018.

[25] T. Müller and H. J. Briegel, "A Stochastic Process Model for Free Agency under Indeterminism: Free Agency under Indeterminism," *Dialectica*, vol. 72, no. 2, pp. 219–252, 2018.

[26] H. J. Briegel, "On creative machines and the physical origins of freedom," *Scientific Reports*, vol. 2, no. 1, p. 522, 2012.

[27] J. B. DeBrota and B. C. Stacey, "FAQBism," *arXiv:1810.13401*, 2018.

[28] C. G. Timpson, “Quantum Bayesianism: A study,” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, vol. 39, no. 3, pp. 579–609, 2008.

[29] N. D. Mermin, “Why QBism Is Not the Copenhagen Interpretation and What John Bell Might Have Thought of It,” in *Quantum [Un]Speakables II* (R. Bertlmann and A. Zeilinger, eds.), pp. 83–93, Cham: Springer International Publishing, 2017.

[30] C. A. Fuchs, “On Participatory Realism,” in *Information and Interaction* (I. T. Durham and D. Rickles, eds.), pp. 113–134, Cham: Springer International Publishing, 2017.

[31] R. Healey, “Quantum-Bayesian and Pragmatist Views of Quantum Theory,” in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University, spring 2017 ed., 2017.

[32] L. J. Savage, *The foundations of statistics*. New York: Dover Publications, 2d rev. ed ed., 1972.

[33] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*. Wiley, 1994.

[34] B. De Finetti, A. Machì, A. F. M. Smith, and B. De Finetti, *Theory of probability: a critical introductory treatment*. Chichester, UK ; Hoboken, NJ: John Wiley & Sons, 2017.

[35] J. Berkovitz, “On de Finetti’s instrumentalist philosophy of probability,” *European Journal for Philosophy of Science*, vol. 9, no. 2, p. 25, 2019.

[36] C. M. Caves, C. A. Fuchs, and R. Schack, “Unknown quantum states: The quantum de Finetti representation,” *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4537–4559, 2002.

[37] D. Wallace, “A Case for QBism.” Presentation at the XII International Ontology Conference, San Sebastian, Spain, 5 October 2016.

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

[39] A. Peres, “Unperformed experiments have no results,” *American Journal of Physics*, vol. 46, no. 7, pp. 745–747, 1978.

[40] R. P. Feynman, “The concept of probability in quantum mechanics,” in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, (Berkeley, Calif.), pp. 533–541, University of California Press, 1951.

[41] M. F. Pusey, J. Barrett, and T. Rudolph, "On the reality of the quantum state," *Nature Physics*, vol. 8, no. 6, pp. 475–478, 2012.

[42] R. Colbeck and R. Renner, "Is a System's Wave Function in One-to-One Correspondence with Its Elements of Reality?," *Physical Review Letters*, vol. 108, no. 15, p. 150402, 2012.

[43] D. Deutsch, "Quantum theory as a universal physical theory," *International Journal of Theoretical Physics*, vol. 24, no. 1, pp. 1–41, 1985.

CHAPTER 8

FAQBISM

As proponents of an interpretation of quantum mechanics, we are accustomed to encountering puzzlement. We know that this is an unusual, perhaps even disreputable activity for professional physicists to engage in. One of the authors came to QBism from quantum chemistry, and the other from nonequilibrium statistical physics. We are here largely because we sought to apply the same habits to quantum foundations that we would find virtuous in any other field of physics: separating principles from convenient conventions, reformulating old mathematics in new ways to make a different set of questions easy, maintaining a healthy disregard for philosophers' judgments about what is impossible. Yet even a mostly ordinary upbringing can lead to a surprising place.

What follows is our attempt to take the questions we have encountered, select those that have been posed in good faith, and provide responses that summarize QBist thinking on those topics. We have tried to be serious without being dour. Finding quantum foundations important at all may be controversial among physicists, potentially more so than any specific choice of interpretation. More shocking still is the suggestion that the material be approached with a sense of fun and adventure.

8.1 Why did you change from Quantum Bayesianism to QBism?

Originally, the *Q* stood for *Quantum* and the *B* for *Bayesian.* The former is still true.

Back in the 1990s and early 2000s, the term "Quantum Bayesianism" was serviceable. However, it had its issues. For one thing, nobody was consistent on whether to capitalize the *Q*: Those who called themselves Quantum Bayesians preferred it uppercase, so that neither half of the term had undue emphasis over the other, but try to convince a copy editor of that simple point! More importantly, there are many varieties of Bayesianism, and plenty of self-declared Bayesians disagreed in fundamental ways with the particular variety that our school found necessary for quantum physics. For a while, N. David Mermin joked that the *B* should stand for *Bruno* de Finetti [1], and Fuchs suggested that the *QB* was like *KFC,* which once stood for "Kentucky Fried Chicken" but is now a stand-alone trademark [2].

More recently, we found a way to expand the *B* that we had never anticipated — a rolling, Lewis-Carroll-esque word: *bettabilitarianism* [3]. This word comes, of all places, from the jurist Oliver Wendell Holmes, Jr. To quote Louis Menand's history of American pragmatism, *The Metaphysical Club: A Story of Ideas in America*:

> 'The loss of certainty' is a phrase many intellectual historians have used to characterize the period in which Holmes lived. But the phrase has it backward. It was not the loss of certainty that stimulated the late-nineteenth-century thinkers with whom Holmes associated; it was the discovery of uncertainty. Holmes was, in many respects, a materialist. He believed, as he put it, that "the law of the grub … is also the law for man." But he was not entirely a determinist, because he did not think that the course of human events was fixed …. Complete certainty was an illusion; of that he was certain. There were only greater and lesser degrees of certainty, and that was enough. It was, in fact, better than enough; for although we always want to reduce the degree of uncertainty in our lives, we never want it to disappear entirely, since un-

> certainty is what puts the play in the joints. Imprecision, the sportiveness, as it were, of the quantum, is what makes life interesting and change possible. Holmes liked to call himself a "bettabilitarian": we cannot know what consequences the universe will attach to our choices, but we can bet on them, and we do it every day.

A QBist declares, "I strive to be the very model of a Quantum Bettabilitarian!"

We have occasionally seen manglings like "QBian" and even "Qubian". These spoil the pun of *QBism* and are thus strongly deprecated.

8.2 Why do you call QBism "local"?

A journey of a thousand perspective shifts begins with a single step. In this case, the first step is to realize why the scenarios trotted out to imply "quantum nonlocality" actually don't [4]. The standard argument for "nonlocality" rests upon entanglement. Conjure up a pair of qubits, for example, assign to them a maximally entangled state and ship one of the pair off to Mars. Upon measuring the qubit left on Earth, "the state of the qubit on Mars changes instantaneously". But this is not a *physical* change of any property of a material object! Compare this with classical electromagnetism: In that subject, if we could toggle a quantity at a distance but only in ways that could not effect a transmission of information, we'd have no hesitation in calling that quantity unphysical — an artifact, we'd say, of choosing a gauge that does not respect relativistic causality. QBism says that the right way to interpret quantum theory is to take this helpful and uncontroversial move seriously. When Alice measures her Earth-bound qubit, what changes for its partner on the red planet? Only *Alice's expectations* for what might happen *to Alice* herself, *if* she were to make the journey and intervene upon that qubit.

The fact that nature violates Bell inequalities is reason to reject the hypothesis used to derive those inequalities, "local realism". But when we decide to adopt *non-(local realism)*, we have a choice of how to clear those parentheses. We can put the *non-* on either half, and when we consider the highly specialized character of what is actually meant by *realism* in this context, keeping the *local* turns out to be the natural move. Another way of saying this is that adopting quantum theory does not force us to revise the notions of "causal structure" developed in classical physics, such as the conceptual tool of Minkowski spacetime.

8.3 What does it feel like to be in a quantum superposition?

A QBist affirms, "My quantum states are mine, your quantum states are yours. If someone else considers a quantum system containing me and ascribes to that system a quantum superposition state, so be it. That is their quantum state assignment. It doesn't make sense for me to assign myself a quantum state if a quantum state is an encoding of my own beliefs for the outcomes of my freely chosen actions." If something feels off about this answer, consider whether you are assuming that there is a "correct" — i.e., purely physically mandated — quantum state in this scenario. For a QBist, there is never such a quantum state just as in personalist probability theory there is never an ontologically "correct" probability distribution. The answer to the question "What does it feel like to be in a quantum superposition?" is the same as the answer to the question "What does it feel like to be in someone else's probability distribution about me?".

Incidentally, this is one problem with the no-go argument made by Frauchiger and Renner, originally intended to rule out what they called "single-world interpretations" of quantum theory [5]. This argument, as well as others closely related to it [6], all at some point make an assumption that amounts to an agent putting herself into a quantum superposition. Ultimately, this makes no more sense than trying to crawl inside a probability distribution and live there.

8.4 Isn't it just solipsism?

On the issue of solipsism, QBism stands with Martin Gardner [7]:

> The hypothesis that there is an external world, not dependent on human minds, made of *something,* is so obviously useful and so strongly confirmed by experience down through the ages that we can say without exaggerating that it is better confirmed than any other empirical hypothesis. So useful is the posit that it is almost impossible for anyone except a madman or a professional metaphysician to comprehend a reason for doubting it.

For a QBist, the basic subject matter of quantum theory is *an agent's interactions with the outside world*; the formalism of quantum theory makes no sense otherwise. Were there no systems outside the QBist's mind, there would be no interface between agent and world, and quantum theory would have no subject matter. QBism is a full-throated rejection of metaphysical solipsism.

Someone once asked us, "Where is the real world in such a view?" The real world is exactly where it always has been. It is the world in which our species evolved. It is the world in which we grow and strive and protest, where we learn by individual experience — including our encounters with the words of others — the pain of heartbreak and the utility of the Lorentz transform. It has conditioned our calculus of expectations, even as those expectations themselves remain intensely personal.

To a QBist, "measurement" is that variety of interaction which physics understands best, precisely because experiments are actions whose potential outcomes we can catalogue. Measurements are to quantum physics what "model organisms" are to biology. Why do developmental biologists know so much about zebrafish? Because their embryos are transparent! But life flourishes on, unanalysed, beyond our microscopes. Far from being solipsistic, QBism recognises just how little of nature we have managed to touch.

It is true that QBists refuse to make an upfront definite claim about what the stuff of the world is. How then, can they have a consistent doxastic interpretation?[1] This is accomplished by being clear on what quantum information is actually information about: A quantum state encodes a user's beliefs about the experience they will have as a result of taking an action on an external part of the world. Among several reasons that such a position is defensible is the fact that any quantum state, pure or mixed, is equivalent to a probability distribution over the outcomes of an informationally complete measurement [8]. Accordingly, QBists say that a quantum state is conceptually no more than a probability distribution. Okay, fine, but what is the stuff of the world? QBism is so far mostly silent on this issue, but *not* because there is no stuff of the world. The character of the stuff is simply not yet understood well enough. Answering this question is the goal, rather than the premise.

Is this an unacceptable weakness of the interpretation? Well, that's a matter of opinion, but ours is that it is not. Must we demand that a complete ontology be laid out before one's ramblings graduate to the status of an "interpretation"? If taken to the extreme, this is clearly unfair: One might claim that no one has a qualifying interpretation because we don't have a successful theory for quantum gravity and so every proposed ontology *necessarily* fails. More practically, feeling pressured to commit to an ontology prematurely may leave physicists unable to imagine one which departs sufficiently from classical intu-

[1] A brief note on useful terminology: *epistemic* refers to knowledge and information, *doxastic* to belief, and *ontic* to brute elements of physical reality.

itions. Why not see if the right ontology can be teased out from the formalism itself and a principled stance on the meaning of its more familiar components (such as probability distributions)?

In fact, QBism has had ontological aspirations ever since the beginning. (It's hard to have ontological aspirations for a theory if you think you'd have to be a solipsist to hold it.) There are *structural realist* and *neutral pluralist* elements in QBism, and there seems to be a process or event ontology underlying it all, somewhere in a spectrum of things suggested by William James, Henri Bergson, Alfred North Whitehead, and John Archibald Wheeler [9]. *The stuff of the world is the becomings of the world.* However, we really don't believe we'll be able to say anything in proper detail until we get the quantum formalism into a better shape. (That's what all the SIC research described in §8.10 is about.) So, from this perspective, QBism is a project.

Mermin [10] argues,

> QBists are often charged with solipsism: a belief that the world exists only in the mind of a single agent. This is wrong. Although I cannot enter your mind to experience your own private perceptions, you can affect my perceptions through language. When I converse with you or read your books and articles in *Nature*, I plausibly conclude that you are a perceiving being rather like myself, and infer features of your experience. This is how we can arrive at a common understanding of our external worlds, in spite of the privacy of our individual experiences.

This leads us to an important topic: communication between agents. What does the idea of agents comparing notes mean when we interpret quantum mechanics as a single-user theory? Consider an agent Alice. She can use quantum mechanics as a "manual for good living", a way to organize her expectations while navigating an irreducibly unpredictable world. Alice encounters a system which she designates as "Bob". Alice can ask Bob about his experiences and use quantum mechanics to predict his answer. If Alice so chooses, she can incorporate Bob's responses into her expectations. Nothing in the formalism of quantum theory forces her to do so, however.

8.5 Isn't it just the Copenhagen intepretation?

This has been said a lot through the years, and we continue to hear it today. Sometimes, it's said that QBism is trying to be more Copenhagen than the Copenhagen interpretation itself. As if QBism had a fever, and the only prescription were more Copenhagen! But the idea that there ever was a unified "Copenhagen interpretation" — i.e., that the definite article is remotely applicable — was a myth of the 1950s. Trying to exceed "the" Copenhagen interpretation in any respect is to race against a phantom.

QBism does not have, for example, Bohr's emphasis on "ordinary language" [11], whatever that might mean. Nor does it have the quantum-classical cut of Heisenberg, the classical laboratory equipment of Landau and Lifshitz [12], the public experimental records of Pauli [13], the essentially ontic state vectors of early Bohm [14], or the frequentism of early von Neumann [15]. Unlike van Kampen, QBism does not presume that the vanishing of interference terms will solve all riddles [16]. Unlike Wheeler, QBism does not posit that all observers should ideally have the same information about a system and thus the same quantum state for it [13, footnote 9]. There simply is not a way to summarize this overflow of differences by claiming that QBism is "more Copenhagen".

At one point, the Wikipedia article on QBism claimed that it "is very similar to the Copenhagen interpretation that is commonly taught in textbooks". What does this even mean? First, as we noted, there's no such thing as "the" Copenhagen interpretation. In addition, claiming that "the Copenhagen interpretation" is "commonly taught in textbooks" conflates the early developers of quantum theory *and* the varied modern expositions of it into a vague mishmash. Asher Peres' textbook is more instrumentalist than the undergraduate standards; the *Feynman Lectures* handle probability in a less frequentist way than Peres. Are all common textbooks Copenhagen, or is Copenhagen that which is commonly taught in all textbooks? Better to strike the term "Copenhagen interpretation" from our lexicon going forward and instead be precise about what views we mean!

8.6 Why does the interpretation of probability theory matter?

E. T. Jaynes put the basic point rather well [17]:

> [O]ur present QM formalism is not purely epistemological; it is a peculiar mixture describing in part realities of Nature, in part incomplete human information about Nature — all scrambled up by Heisenberg and Bohr into an omelette that nobody has seen how to unscramble. Yet we think that the unscrambling is a prerequisite for any further advance in basic physical theory. For, if we cannot separate the subjective and objective aspects of the formalism, we cannot know what we are talking about; it is just that simple.

According to Jaynes, the way to unscramble that Heisenberg–Bohr omelette will be "to find a different formalism, isomorphic in some sense but based on different variables" [18].

The *results* of quantum-mechanical calculations are generally probabilities, or stand-ins for them like rates and effective cross-sections, and so the question "what means probability?" must be addressed sooner or later. This question gains urgency when we realize that the *inputs* to those calculations are just as probabilistic as the outputs. We can see this concretely by focusing on the simplest possible quantum system, a single qubit. A quantum state for a qubit can be written as a linear combination of the Pauli matrices, where the coefficients are expectation values for the outcomes of the three Pauli measurements:

$$\rho = \frac{1}{2}\left(I + \langle x \rangle \sigma_x + \langle y \rangle \sigma_y + \langle z \rangle \sigma_z\right). \tag{8.1}$$

Because each Pauli measurement has only two possible outcomes, $+$ and $-$, we can write the expectation value $\langle x \rangle$ as

$$\langle x \rangle = p(+|x) - p(-|x) = 2p(+|x) - 1, \tag{8.2}$$

and similarly for $\langle y \rangle$ and $\langle z \rangle$. Mathematically, any qubit state ρ, whether pure or mixed, is nothing more than a convenient packaging of the three probabilities $p(+|x)$, $p(+|y)$ and $p(+|z)$. Thus, whatever status one grants to quantum states, one must grant that same status to at least some probabilities. Conversely, if a particular interpretation of probability theory turns out to be logically untenable, then that rules out a possible way of interpreting quantum states, too.

Note that the fact that specifying a qubit quantum state requires *three* probability values is more fundamental than any choice of those values. To put it another way, when we take our empirically successful theory of physics and find the simplest case where it applies, we see that the theory has three knobs, not one or two or five. This is a more primitive, more basal statement about our theory than a choice of ρ is!

Modern quantum information theory provides an even deeper take: Any quantum state can in fact be specified, not just as a compendium of probabilities for different experiments, but as a probability distribution over the outcomes of a *single* experiment. For more on the concept of a *reference measurement*, see §8.10.

8.7 What is the meaning of the double-slit experiment in QBism?

For a QBist, the double-slit experiment is about the peculiarities that happen when an agent tries to relate their expectations for one hypothetical scenario to their expectations for another. Per tradition, we can call this agent Alice. She might compute her probability for a detector click given that she will place the detector at position x and open slit $\#1$ — call it $P_1(x)$. Likewise, she can compute the corresponding quantities for the configuration with only slit $\#2$ open, $P_2(x)$; and for when both slits are open, $P_{12}(x)$. All of these quantities are, by themselves, rather ordinary probabilities: None of them end up being negative, let alone complex. Nor is it surprising that $P_1(x)$ might be discrepant from $P_2(x)$ or from $P_{12}(x)$. Different conditions, different probabilities! The puzzle is that

$$P_{12}(x) \neq P_1(x) + P_2(x). \tag{8.3}$$

The strangeness lies not in the curve for any particular scenario, but in how the scenarios fit together.

Rob Spekkens likes to point out that the *mere fact* of interference is not a very deep probe of quantum theory, because interference can arise in models based on local hidden variables [19, 20]. You just have to be careful and consistent when constructing your model. In his toy theory, where states are probability distributions over discrete local hidden variables, we can build a test for double-slit-type oddities (a toy Mach–Zehnder interferometer), and indeed, interference occurs.

In order to test quantum theory more stringently, we have to find probes of nonclassical expectation-meshing that resist easy emulation. This, from the QBist perspective, is what Bell inequality violations are all about: Given any particular choice of detector settings, the outcome probabilities are just probabilities. The power and the mystery of quantum theory reside in the relation between probabilities for different choices of detector settings.

Our research on SICs is also in this vein (see §8.10). Using a SIC as a reference measurement is like considering a generalized interference experiment, where the outcomes for the "which-way" measurement correspond to *nonorthogonal* quantum states. This generalization takes us out of the realm of easy classical emulation, letting us investigate the quantum formalism more deeply.

8.8 Doesn't the PBR theorem prove QBism wrong?

The Pusey–Barrett–Rudolph (PBR) no-go theorem demonstrates, as the authors put it, "that any model in which a quantum state represents mere information about an underlying physical state of the system, and in which systems that are prepared independently have independent physical states, must make predictions which contradict those of quantum theory" [21]. In the years since its appearance, many have claimed that the PBR theorem proves quantum states are ontic — that it rules out all epistemic and doxastic interpretations. One often hears that QBism, having itself a doxastic conception of quantum states, should therefore be ruled out by the lack of any experimental violations of quantum theory.

But one should not believe these rumors. The PBR theorem does no damage to QBism. PBR say so themselves at the end of their paper. This is because what they demonstrate is the inconsistency of the idea of holding epistemic quantum states at the same time as holding that they are epistemic *about* ontic states. In QBism, quantum states represent one's beliefs, not about some ontic variable, but about one's future *personal* experiences which come in consequence of taking an action on the external world. I.e., they are epistemic (or better, doxastic) *about* personal experiences. Technically, this means there are no compelling reasons in QBism to adopt the very starting point of PBR — namely, trying to use an integral over ontic states λ to get probabilities. The PBR theorem is a no-go result for a direction in which we never wanted to go.

The foundational assumption of the PBR theorem is a rule for computing some quantities $p(k|\Psi(x_1,\ldots,x_n))$, probabilities for a measurement outcome k given preparation of a product state $\Psi(x_1,\ldots,x_n)$. This rule is a statement about conditional probabilities:

$$p(k|\Psi(x_1,\ldots,x_n)) = \int_\Lambda \cdots \int_\Lambda p(k|\lambda_1,\ldots,\lambda_n)\mu_{x_1}(\lambda_1)\cdots\mu_{x_n}(\lambda_n)d\lambda_1\cdots d\lambda_n\,. \quad (8.4)$$

Here, λ_i in a measure space Λ is a possible physical state that a system can be in, $\mu_{x_i}(\lambda_i)$ is a probability distribution over Λ for the ith system, and $p(k|\lambda_1,\ldots,\lambda_n)$ is a probability for obtaining outcome k given a set of physical states for each system. In other words, the whole approach of PBR is trying to identify the Born Rule with an application of the Law of Total Probability (LTP). It can't be done, and they have rediscovered that in their own way.[2]

The LTP is familiar and what one would use *if there were* underlying hidden variables. One avenue of QBist technical research currently ongoing is to explore an *alternative* to the LTP which expresses the fact that such hidden variables do not exist. The crucial idea is a *reference measurement,* a procedure with the property that a probability distribution over its outcomes can be used to compute the probabilities for all the outcomes of any other measurement. Let $P(H_i)$ be Alice's probability for obtaining outcome H_i in an optimal reference measurement (many criteria for optimality turn out to be equivalent for

[2]We note that one philosopher of physics has declared, speaking of an assumption equivalent to Eq. (8.4), "If you don't believe that, you don't believe in physics at all." As best as we can tell, there is no reason to accept such a claim, other than an underdeveloped imagination.

this problem). Classical intuition suggests that the best possible reference measurement would just be to read off the ontic state, and so by the LTP,

$$P(D_j) = \sum_i P(H_i)P(D_j|H_i)\,, \tag{8.5}$$

for any other measurement $\{D_j\}$. But in the quantum world, this does not apply, and the closest we can get to it, by cannily choosing our reference measurement, is

$$Q(D_j) = \sum_{i=1}^{d^2} \left[(d+1)P(H_i) - \frac{1}{d}\right] P(D_j|H_i)\,. \tag{8.6}$$

$Q(D_j)$ now represents an agent's probability for obtaining the experience D_j from a measurement she represents with the POVM $\{D_j\}$, $P(H_i)$ is her probability for obtaining outcome H_i in a hypothetical reference measurement, and $P(D_j|H_i)$ is her probability, asserted now, for obtaining the experience D_j supposing she had previously made the reference measurement and obtained experience H_i. Note that the only difference from the LTP is a constant shift and rescaling of $P(H_i)$ for each i. In fact, this is the *closest* [8] the two expressions can come, suggesting that this expression may provide insight into what it is about the universe that makes it "quantum".

So, in all, QBists say this about the PBR theorem (and similarly about Bell's theorem): Rather than denigrate the QBist conception of quantum theory, they actually help compel it. There are so many arguments of analogy for epistemic quantum states (Rob Spekkens' toy model nails about 25 of them [19]), but what the PBR and Bell theorems compel and the toy theories can't is that, if quantum states are epistemic, they cannot be epistemic about some ontic variables. The most the PBR theorem can do is rule out a middle ground that we are not sure anyone actually occupied in the first place.

8.9 Is QBism about the Bayes rule?

Adopting a personalist Bayesian interpretation of probability does *not* mean treating all changes of belief as applications of the Bayes rule. This is shocking to some people! And distancing ourselves from the dogmatists who claim to follow that creed is one reason why we prefer *QBism* over “Quantum Bayesianism”.

In the tradition of Ramsey, Savage and de Finetti, there are consistency conditions that an agent’s probability assignments should meet at any given time, *and then* there are guidelines for *updating* probability assignments in response to new experiences. Going from the former to the latter requires making extra assumptions — the two are not as strongly coupled as many people think. The Bayes rule is not a condition on how an agent *must* change her probabilities, but rather a condition for how she should *expect* that she will modify her beliefs in the light of possible new experiences. For this observation, we credit Hacking, Jeffrey and van Fraassen.

Fuchs and Schack go into more detail on this point in an article [22], and BCS wrote a pedagogical treatment in a book [23, §5.1].

There is a common misconception afoot that being “Bayesian” fundamentally means using the Bayes rule to update probabilities. For example, the Wikipedia page that lists things named after Thomas Bayes says that “Bayesian” refers to “concepts and approaches that are ultimately based on Bayes’ theorem”. This may be historically correct, but it is not logically correct. In the personalist Bayesian school, we first start with the idea of quantifying beliefs and expectations as gambling commitments. Then, we impose a consistency condition, from which the familiar rules of probability theory follow. The idea of *updating probabilities over time* in accord with the Bayes rule arrives rather late in this development. One must first establish the standards for probabilities being consistent with each other at a particular time, before invoking further considerations to establish a scheme for *changing* probabilities in response to new experiences. Bayes' theorem *is* a theorem, not an axiom.

The "collapse of the wavefunction" is analogous to, and an algebraic variant of, Bayesian conditionalization [24]. Having recognized this, we can appreciate that it clears up a mystery (or, perhaps better put, allows us to identify a pseudo-mystery for what it is). But the recognition of the "quantum Bayes rule" was an early step on the path to QBism, and its relevance in more recent years has if anything been rather peripheral.

8.10 What technical questions have been motivated by QBism?

The development of Quantum Bayesianism, and its progressive evolution into QBism, is a story of feedback loops between technical and philosophical questions.

The quantum de Finetti theorem was sought and proved in order to show there could be a meaning to the phrase "unknown quantum state" even from a subjectivist perspective [25]. The Quantum Bayesians thought that without such a theorem, a subjectivist reading of probability in quantum theory wouldn't be possible after all. This theorem then outgrew its foundational origins, becoming a powerful tool for the practical problem of analyzing the security of quantum key distribution. A quantum de Finetti theorem for "unknown processes" followed from the same motivation as that for "unknown states" [26].

Asher Peres pointed out that quantum states are more analogous to probability distributions over phase space — that is, to Liouville density functions — than to points in phase space. In 1995, Fuchs followed this lead and searched for examples within Liouville mechanics that echoed quantum theory, including the aspects of quantum theory that had been declared uniquely nonclassical. He found that the quantum no-cloning theorem was just one such feature: A no-cloning theorem holds in Liouville mechanics, exactly as in the quantum case. Trying to further refine the enquiry led to the quantum no-broadcasting theorem [27].

In 2002, Caves, Fuchs and Schack took on the question of whether or not quantum theory implied any kinds of compatibility conditions for disparate agents' quantum state assignments [28]. This is a natural question to ask, if quantum states are to be interpreted doxastically. The work resulted in solid theorems — and, in a twist whose irony has gone underappreciated, Pusey, Barrett, and Rudolph [21] used one of these notions to prove the PBR theorem. (For QBism's response to the PBR theorem, see §8.8.)

Another example came from trying to understand what it could mean for quantum states to be "disturbed by measurement" if they are not ontic. Answering this led to [29] and [30], which Fuchs later turned to the purpose of defining a threshold for successful quantum teleportation in Jeff Kimble's lab [31]. Discussion of this point can be found in Fuchs and Jacobs [32].

More recently, at the creative interface between conceptual and technical matters, Fuchs and Schack have made the case that the right way to think about decoherence is with van Fraassen's *reflection principle* [22]. We suspect that there are new theorems to be proved in this area, in addition to the conceptual implications (such as putting a sharper point on an old argument of Asher Peres about when black-hole evaporation should *not* be modeled with a unitary evolution [33]).

The most active technical topic in contemporary QBism research is the project of reconstructing quantum theory from physical principles. Central to this is our ongoing research into symmetric informationally complete quantum measurements (SICs). A SIC for a d dimensional Hilbert space is a set of d^2 pure quantum states with equal pairwise overlaps:

$$|\langle\psi_i|\psi_j\rangle|^2 = \frac{d\delta_{ij}+1}{d+1}\ . \tag{8.7}$$

A uniform rescaling of these states defines a POVM which is uniquely suited to be a "standard quantum measurement".

Not everyone who works on SICs is devoted to QBism. Indeed, we gather from conversations in hotel bars that one of the prime movers in SIC-hunting doesn't particularly care about quantum mechanics; their appeal as geometrical objects is enough. (Historically speaking, one of the most closely studied SIC constructions originally flowed from the pen of Coxeter, who just really liked polytopes [34]. But in a surprise twist, this SIC arises in the study of quantum-state compatibility that Caves, Fuchs and Schack initiated [35]!) Another SIC researcher is not a QBist, but came to the problem through Fuchs's advocacy and over the years has displayed many sympathies.

Going in the other direction, being a QBist doesn't mean you have to live and breathe SICs. Fuchs and BCS put it the following way [3]:

> If all that you desire is a story that you can tell about the current quantum formalism, then all this business about SICs and probabilistic representations might be of little moment. Of our fellow QBists, we know of one who likely doesn't care one way or the other about whether SICs exist. Another would like to see a general proof come to pass, but is willing to believe that QBism can just as well be developed without them — i.e., they are not part of the essential philosophical ideas — and is always quick to make this point. On the other hand, we two are inclined to believe that QBism will become stagnant in the way of *all other* quantum foundations programs without a deliberate effort to rebuild the formalism.

We find that SICs cut to the heart of quantum theory in a way that other ideas for rebuilding the formalism do not. This is a point we discuss elsewhere in this collection (§8.7), and in earlier papers [3, 11, 36]. The representation of quantum theory that SICs furnish has natural connections with the study of Wigner-function negativity, which is important for quantum computation [37]. In addition, the discovery of a connection between SICs and algebraic number theory reshapes the boundary between physics and pure mathematics in a remarkable way [38, 39].

8.11 Isn't quantum probability just classical probability but noncommutative?

There's a *Far Side* cartoon that shows a man waking up in bed and staring at a giant note he wrote for himself on the wall: "First pants, *then* shoes!" The lesson is that order of operations matters in daily life, long before it matters in quantum physics. So, we have to be careful what we mean by "noncommuting", if we want it to have any meaningful content. And when we do get appropriately mathematical about it, we find that it is *not* the signature of the quantum. The Spekkens toy model, which has a simple statement in terms of underlying local hidden variables, has observables that do not commute [19].

There is a common sentiment about that quantum mechanics is "a noncommutative generalization of probability theory": Instead of using vectors that sum to 1, one has matrices whose trace is 1, and so forth. This is a fine approach for many applications, but in physics, there is never a guarantee that a method which works for one set of problems will do equally well with another. Taking one representation of the theory as defining its essence can cloud your physical insight. In this case, the "we must generalize probability to make it noncommutative" impulse obscures the fact that *given a specific experimental scenario,* the probabilities of quantum physics are just probabilities — numbers that play together in accord with Kolmogorov's rules. As we noted in §8.7, it is the meshing of expectations for one scenario with those of another which reveals the fundamental enigma of quantum theory. Noncommutativity is a secondary property, and as the Spekkens toy model teaches us, not a quintessentially quantum one at that.

BCS, who came to QBism from statistical physics, likes to point out that the Doi–Peliti formalism for nonequilibrium stochastic dynamics has noncommuting operators, and also complex numbers, Feynman diagrams, renormalization, Glauber states, the Heisenberg equation of motion, and even the Schwinger representation of $\mathfrak{su}(N)$. Yet it is all a fully classical theory [23, 40]. It borrows calculational devices from quantum mechanics, but the stochasticity it considers is, at root, ignorance about pedestrian hidden variables.

8.12 Doesn't decoherence solve quantum foundations?

The theory of decoherence is a set of calculations which enable one to write a density matrix that is nearly diagonal in some basis of interest. This does not tell you what a density matrix *means.*

Max Schlosshauer, who wrote the canonical textbook on decoherence, recently summarized the situation as follows [41]:

> Decoherence, at its heart, is a technical result concerning the dynamics and measurement statistics of open quantum systems. From this view, decoherence merely addresses a *consistency problem,* by explaining how and when the quantum probability distributions approach the classically expected distributions. Since decoherence follows directly from an application of the quantum formalism to interacting quantum systems, it is not tied to any particular interpretation of quantum mechanics, nor does it supply such an interpreta-

> tion, nor does it amount to a theory that could make predictions beyond those of standard quantum mechanics.
>
> The predictively relevant part of decoherence theory relies on reduced density matrices, whose formalism and interpretation presume the collapse posultate and Born's rule. If we understand the "quantum measurement problem" as the question of how to reconcile the linear, deterministic evolution described by the Schrödinger equation with the occurrence of random measurement outcomes, then decoherence has not solved this problem.

For a deeper dive into the QBist take on decoherence, see [22].

8.13 Is QBism like Rovelli's "Relational Quantum Mechanics"?

Several people have made the comparison between QBism and Rovelli's "Relational Quantum Mechanics" [42], and it is not unjust. Some slogans of RQM can be carried over to QBism with only a little modification, and the *motivation* for the research program that Rovelli suggested in his original paper has certain affinities with our own. However, there are important differences between QBism and RQM, and moreover, we find the statements of RQM imprecise on key points.

Both QBism and Rovellian RQM reject the notion of a single quantum state for the entire universe. In QBism, measurement outcomes are personal experiences for the agent who elicits them, while in RQM, physical properties exist "relationally" between systems. As the *Stanford Encyclopedia of Philosophy* says, in RQM, "Quantum events only happen in interactions between systems, and the fact that a quantum event has happened is only true with respect to the systems involved in the interaction" [43]. This motto is not unlike what we have written about QBism. For example,

> Certainly QBism has creation going on all the time and everywhere; quantum measurement is just about an agent hitching a ride and partaking in that ubiquitous process.

But we can already start to see a divergence. Rovellian RQM downplays the idea of agency: In RQM, a grain of sand can be an "observer" of another quantum system. Given any two systems S_1 and S_2, there is a quantum state of S_2 relative to S_1, just as in Newtonian physics, S_2 always has a velocity relative to S_1.

Likewise, QBism and RQM differ on how to interpret probability. While we find the foundational papers of RQM somewhat vague on this point, our overall impression is that RQM leans more to a Jaynesian kind of Bayesianism, more objective and less personalist than the Ramseyian/de Finettian school to which QBism adheres. This is tied to a point emphasized in the technical side of QBism (§8.10). Mathematically speaking, a quantum state *is* a probability distribution. Pick any informationally complete POVM, and you can replace density operators with probability distributions over the outcomes of that POVM (even when the density operators are rank-1 projectors, i.e., pure states). As best we can tell from reading Rovelli *et al.,* whenever an "observer" S_1 coexists with another system S_2, there exists a unique, physically correct quantum state for S_2 relative to the observer S_1. Therefore, there exists a unique, physically mandated set of probabilities concerning S_2, which happen to be relative to S_1. We find this philosophy of probability ultimately untenable [44, 45].

We must also admit, we're not great fans of the word *relational.* This adjective naturally carries the connotation of "just like in relativity theory". But in relativity, we can readily transform between reference frames. A statement like "the clocks C_1 and C_2 are synchronized" is relational: Its truth or falsity depends on whether it is evaluated by Alice or by Bob. Yet if Alice knows Bob's trajectory relative to herself, she can take what she sees and Lorentz-transform her figures to compute what Bob must see.

In quantum theory, there is no analogue of this. (Emphasizing this point of *dis*-analogy is another way QBism distinguishes itself from Bohr [11].) RQM tries to invent one, but the attempt flounders. We can see exactly how this happens if we examine Smerlak and Rovelli's paper "Relational EPR" [46]. The authors take a certain notion of consistency among multiple observers over from Rovelli's original paper:

> It is one of the most remarkable features of quantum mechanics that indeed it automatically guarantees precisely the kind of consistency that we see in nature [Rovelli 1996]. Let us illustrate this assuming that both A and B measure the spin in the same direction, say z, that is $n = n' = z$.

But on the very next page, they describe the following scenario:

> A observes the spin in a given direction to be $\uparrow$ and B observes the spin in the same direction to be also $\uparrow$.

And they say that *this* is an ill-posed statement, because

> it does not happen either with respect to A or with respect to B. The two sequences of events (the one with respect to A and the one with respect to B) are distinct accounts of the same reality that cannot and should not be juxtaposed.

But if the second statement is an invalid proposition, *then the first must be as well.* The description "both A and B measure the spin in the same direction" cannot apply "either with respect to A or with respect to B"; it presumes a view from nowhere. (One could try to evade this by interpreting the story of what both A and B measure as told relative to a third party, the superobserver C. This might look like it could ameliorate the problem, at least if the difficulties we saw above could be resolved. But presuming that a superobserver is *always* available, and that the expectations of the superobserver override those of any other participant, just de-relationalizes the theory all over again. And why *should* physics guarantee on a fundamental level that a superobserver is always available? When children or politicians quarrel, life does not always provide a responsible adult who can restore the peace.) In short, the description of the gedankenexperiment that Smerlak and Rovelli use to put forth their notion of "consistency" is exactly the kind of language which they elsewhere insist is meaningless.

One philosophy paper that compared QBism and RQM [47] must be mentioned in particular.[3] We reproduce the relevant passage with its absence of citations preserved intact:

> QBism is the view that quantum mechanics is not a theory about the world, but about our degrees of credence concerning predictions. The theory provides universal, objective rules for updating these degrees from the information one gets on the world through events. All this is shared by RQM. One difference is that QBism is human-centered, while RQM is not: any physical object qualifies as a potential observer. But what remains of it if all talk of external observers boils down to talk of events relative to us? If anything, RQM is more radically instrumentalist than QBism: after all, the latter assumes that events are objective and publicly accessible…

Most of this is at least a little wrong, so we will go through it in detail.

[3] We have the sense that, like Bohmian mechanics, RQM has been of interest to philosophers more than it has been to physicists. The question of what biases the philosophy community perpetuates by always turning to its familiar authorities for opinions is an interesting one.

> QBism is the view that quantum mechanics is not a theory about the world, but about our degrees of credence concerning predictions.

In QBism, quantum mechanics is not a theory *directly* about the world, but rather, a theory that any of us can use to manage our "degrees of credence" *in light of the fact that* the world has a specific character.

> The theory provides universal, objective rules for updating these degrees from the information one gets on the world through events.

Yes, the rules that quantum theory provides are "universal" (anyone can pick up the hero's handbook [3]) and "objective" (or as objective as anyone could want of a physical theory). The emphasis on "updating" echoes a misconception we have seen elsewhere, that Bayesian probability is fundamentally about the Bayes update rule (see §8.9). And in the QBist understanding of personalist probability, the rules allow more loose play in updating expectations than this formulation grants.

> All this is shared by RQM.

To us, it seems a better fit for RQM than for QBism. As we wrote above, a preference for objective probability runs through RQM, holding it back.

> One difference is that QBism is human-centered, while RQM is not: any physical object qualifies as a potential observer.

Human-centered, no, but *agent*-centered, yes. An agent does not have to be human (see sections §8.17 and §8.19).

> If anything, RQM is more radically instrumentalist than QBism: after all, the latter assumes that events are objective and publicly accessible…

No, it doesn't. Fuchs put it this way in 2010 [48]:

> *Whose information?* "Mine!" *Information about what?* "The consequences (for *me*) of *my* actions upon the physical system!" It's all "I-I-me-me mine," as the Beatles sang.

That article goes on to draw an explicit contrast between QBism and Pauli's claim that measurement outcomes "are objectively available for anyone's inspection".

The introductory paper by Fuchs, Mermin and Schack [4] expresses the point as follows:

> The personal internal awareness of agents other than Alice of their own private experience is, by its very nature, inaccessible to Alice, and therefore not something she can apply quantum mechanics to. But verbal or written reports to Alice by other agents that attempt to represent their private experiences are indeed part of Alice's external world, and therefore suitable for her applications of quantum mechanics. Having always stressed the crucial importance of stating the results of experiments in ordinary language, Bohr would probably have been comfortable with Alice's indirect access to Bob's experience, through language.
>
> But Bohr would not have approved of Alice superposing reports from Bob about his own experience, as QBism requires her to do if she wants to subject those reports to analysis before they enter her own experience. We believe that Bohr would have viewed Bob's reports — formulations in ordinary language — as beyond the scope of quantum mechanics. But because Alice can treat Bob as an external physical system, according to QBism she can assign him a quantum state that encodes her probabilities for the possible answers to

> any question she puts to him. When Alice elicits an answer from Bob, she treats this as she treats any other quantum measurement. Bob's answer is created for Alice only when it enters her experience. A QBist does not treat Alice's interaction with Bob any differently from, say, her interaction with a Stern–Gerlach apparatus, or with an atom entering that apparatus.

Or, later and more compactly:

> What the usual story [of Wigner's Friend] overlooks is that the coming into existence of a particular measurement outcome is valid only for the agent experiencing that outcome.

8.14 Why do QBists prefer de Finetti over Cox?

The Cox approach is too psychologically loaded in the direction of hidden variables and inferences about them. This sentiment dates back to the 1990s, when Fuchs and colleagues were hashing out the basics of being Bayesian in a quantum world. During 1993 and 1994, Fuchs and Schack became disenchanted with Cox's development of probability theory and attracted instead more to the development of de Finetti and Savage and others. The essence of the latter school is the Dutch-book notion and/or the simultaneous development of probabilities with utilities (i.e., decision theory). Looking back on it, the attraction to the one over the other cuts to a rather fundamental point:

QBism regards physics, and science in general, in Darwinian terms. The mathematics we develop is practical because, at root, it helps agents to survive. From this point of view, the idea of a probability as a gambling commitment, a belief made quantitative and ready to be acted upon, is an attractive notion. On the other hand, the idea of probability being used for a "theory of inference" in the usual sense — i.e., a measure of plausibility for something that is "out there" but unknown — is a bit off-putting.

(This also seems to be a fundamental distinction between our program and that of Rob Spekkens. The general tenor of the Spekkensian program has been to interpret quantum states as states of information about some type of hidden variable as yet unspecified, perhaps degrees of freedom that are "relational" in some way. The Coxian attitude is a natural fit for this view, but it is not so for QBism.)

All the way back in July 1996, Fuchs wrote the following, in a note to Sam Braunstein:

> While in Torino, you really got me interested in the old [Cox derivation] question again. I noticed in this version of the book that Jaynes makes some points about how there are still quite a few questions about how to set priors when you don't even know how many outcomes there are to a given experiment, i.e., you don't even know the cardinality of your sample space. That, it seems to me, has something of the flavor of quantum mechanics … where you have an extra freedom not even imagined in classical probability. The states of knowledge are now quantum states instead of probability distributions; and one reason for this is that the sample space is not fixed — any POVM corresponds to a valid question of the system. The number of outcomes of the experiment can be as small as two or, instead, as large as you want.
>
> However I don't think there's anything interesting to be gained from *simply* trying to redo the Coxian "plausibility" argument but with complex numbers. It seems to me that it'll more necessarily be something along the lines of: "When you ask me, "Where do all the quantum mechanical outcomes

> come from?" I must reply, "There is no where there." [. . .] That is to say, my favorite "happy" thought is that when we know how to properly take into account the piece of prior information that "there is no where there" concerning the origin of quantum mechanical measurement outcomes, then we will be left with "plausibility spaces" that are so restricted as to be isomorphic to Hilbert spaces. But that's just thinking my fantasies out loud.

More recently, we have made steps in this direction, as documented in our earlier papers [3, 4, 36] and outlined in §8.10.

8.15 Why so much emphasis on finite-dimensional Hilbert spaces?

Quantum theory can be formulated for finite- and infinite-dimensional systems. By any standard, genuinely nonclassical effects are present in finite-dimensional systems, suggesting that these may be all that is strictly necessary for capturing the conceptual core of the theory. Indeed, it might even be distracting to let infinite dimensions complicate foundational considerations. In some ways the infinite-dimensional situation is the limit of large dimensions, but in other ways it isn't.

Infinite dimensions are subtle and complicated, but it seems they are not so for "quantum" reasons.

The goal of our research is to bring clarity to the quantum mysteries. When one looks up what the "quantum mysteries" are, one finds that either they are expressed in finite-dimensional terms from the get-go [49], or, if the presentation includes continuous degrees of freedom, all the interesting stuff happens in the finite-dimensional part. For example, Asher Peres' book explains a Bell–EPR scenario using both position and spin degrees of freedom, but the essence of the problem lies in the spins, while the position coordinates just provide conceptual scaffolding. To "go for the jugular" of the quantum enigmas, we have chosen to focus on finite dimensions — and the results have been so pretty that we can't help but wonder if they offer a guide for where physics should go *next,* as it pushes beyond the continuum theories we all know so well.

The authors of this FAQ spend our weekdays reformulating finite-dimensional quantum theory (see §8.10). However, we would have nothing personal against anyone who tried to find a new representation for, say, algebraic quantum field theory. We do offer a cautionary note: Even the most successful and most "fundamental" physical theories are provisional, their applicability contingent on physicists' limited abilities as agents to intervene into the affairs of other natural systems. Indeed, the way we extract any empirical utility from a QFT is, in practice, to remind ourselves that it *cannot* be valid to arbitrarily high energies, and then deal with that limitation in a mature way (a process technically known as regularization and renormalization). When one cannot trust *any* physical theory to provide ultimate, metaphysical bedrock; when *all* the theories one might wish to reformulate and reconstruct are inextricably provisional — then, unavoidably, picking the theory to focus upon becomes a judgment call.

It is intriguing that the possibility that physically accessible Hilbert-space dimension is always finite — possibly quite large, but still finite — is a recurring theme in quantum-gravity research. For various flavors of this idea, see, e.g., [50–53]. Fuchs and BCS gave a QBist spin on this speculation in 2016 [3], following a lead that Fuchs set out in 2010 [48] and 2004 [54].[4]

[4]And, in correspondence with Bill Unruh and others, even before that [55, pp. 659–52].

8.16 Aren't the probabilities in quantum physics objective?

The intuition that the probabilities in quantum physics are objective properties of a system is deeply ingrained. For many, the suggestion that it might be otherwise is so outlandish as to obviate the need for rebuttal. Thus the starting point of QBism, adopting a strict, de Finettian/Ramseyan interpretation of *all* probabilities, turns out to be a big pill to swallow once the full seriousness of its consequences are realized. However, QBists do not deny the objective probability intution. What we claim is that the advantages that subjectivity brings (which may be found in any exposition of QBism) outweigh the draw of untutored impulses. In fact, the appeal of this intuition may be understood from and thereby absorbed into a purely personalist point of view.

There is nothing about the intuition which demands the invocation of quantum theory. For instance, we might just as well consider a coin or a die. One often hears that the symmetries of the matter distribution making up a "fair" coin or a die *determine* the probability of a flip landing "heads" or of rolling a "3". But what does it mean for a coin to be "fair"? It means that one assigns equal probability to the heads and tails outcomes. How does one certify that a coin is fair? If the answer involves checking that the coin's mass distribution closely matches that of a thin cylinder, claiming that the probability distribution comes from the mass distribution is circular. We bring many expectations and a lifetime of experience to the table when asserting a probability. Among these is experience with the effects of gravity on differently shaped objects. The reason that it feels our probabilities are properties of objects is just that we feel the force of our priors so strongly that we feel they were given to us by nature.

More generally, if we wanted the probability to be physically determined, a little reflection reveals it couldn't be a property *only* of the coin itself. It must also depend on the flipping process. A coin can have a very even mass distribution while it sits forgotten on the bedside table. For that matter, it is quite possible to engineer a machine which precisely flips a coin to land heads up every time [56]. Furthermore, couldn't a high-speed camera and a sufficiently advanced computer program predict the result of any particular coin toss with amazingly few errors given the first few fractions of a second of the flip? With such a setup, what should we say is the probability of heads after the machine announces its prediction?

Supposing the force of these arguments is felt and the conclusion that probability is about personal expectations is accepted, there remains one refuge for the objective probabilists — essentially, that quantum theory legitimizes them. Classically, one might argue, complete information is in principle possible, but quantum mechanically, maximal information is incomplete. What's left over is the objective chance. If one knew the objective chance, they would be best served by setting their personal expectation equal to it.

First, we note that maximal information being incomplete doesn't require the *nature* of probability to change. Supposing there is a correct probability in a given circumstance remains a big leap. But there is a more critical issue, namely, if there were a correct probability, there's no way to be sure you've got it. Here's how Fuchs and BCS put it in a previous paper [3].

> Previous to Bayesianism, probability was often thought to be a physical property—something objective and having nothing to do with decision-making or agents at all. But when thought so, it could be thought only inconsistently so. And hell hath no fury like an inconsistency scorned. The trouble is always the same in all its varied and complicated forms: If probability is to be a physical property, it had better be a rather ghostly one—one that can be told of in

campfire stories, but never quite prodded out of the shadows. Here's a sample dialogue:

Pre-Bayesian: Ridiculous, probabilities are without doubt objective. They can be seen in the relative frequencies they cause.

Bayesian: So if $p = 0.75$ for some event, after 1000 trials we'll see exactly 750 such events?

Pre-Bayesian: You might, but most likely you won't see that exactly. You're just likely to see something close to it.

Bayesian: "Likely"? "Close"? How do you define or quantify these things without making reference to your degrees of belief for what will happen?

Pre-Bayesian: Well, in any case, in the infinite limit the correct frequency will definitely occur.

Bayesian: How would I know? Are you saying that in one billion trials I could not possibly see an "incorrect" frequency? In one trillion?

Pre-Bayesian: OK, you can in principle see an *incorrect* frequency, but it'd be ever less *likely*!

Bayesian: Tell me once again, what does "likely" mean?

This is a cartoon of course, but it captures the essence and the futility of every such debate. It is better to admit at the outset that probability is a degree of belief, and deal with the world on its own terms as it coughs up its objects and events. What do we gain for our theoretical conceptions by saying that along with each actual event there is a ghostly spirit (its "objective probability," its "propensity," its "objective chance") gently nudging it to happen just as it did? Objects and events are enough by themselves.

To see how quantum physics does not make probabilities somehow more objective, consider the following [4]. Take a two-qubit system for which an agents could make either of the two quantum state assignments ρ_+ and ρ_-, defined by

$$\rho_\pm = \frac{1}{2}\left(|0\rangle\langle 0|^{\otimes 2} + |\pm\rangle\langle\pm|^{\otimes 2}\right) \tag{8.8}$$

where we have used the common notation

$$|\pm\rangle = \sqrt{\frac{1}{2}}(|0\rangle \pm |1\rangle)\,. \tag{8.9}$$

These state assignments are "compatible" in that they have overlapping supports on the two-qubit state space. Yet suppose the first qubit is measured in the "computational basis" $\{|0\rangle, |1\rangle\}$ and outcome 1 is found. The agent updates her state accordingly, using the standard Lüders rule, and her postmeasurement state for the second qubit is then $|+\rangle$. However, if she had begun with the joint state ρ_-, then experiencing outcome 1 would have led her to update her state for the second qubit to $|-\rangle$ instead. The two possibilities for the initial state were compatible, but the two possible final states, updated in response to exactly the same data, are orthogonal! This is an illustrative extreme case of a phenomenon that is much more general: Priors do not inevitably wash out, even in the limit of infinite data [57].

8.17 Don't you have to define "agent"?

Fuchs and BCS wrote the following in an earlier paper:

> Thinking of probability theory in the personalist Bayesian way, as an extension of formal logic, would one ever imagine that the notion of an agent, the user of the theory, could be derived out of its conceptual apparatus? Clearly not. How could you possibly get flesh and bones out of a calculus for making wise decisions? [...] Look as one might in a probability textbook for the ingredients to reconstruct the reader herself, one will never find them. So too, the QBist says of quantum theory.

This perspective is essentially that of L. J. Savage, who developed rational decision theory in terms of "consequences", "acts" and "decisions" [58], though where Savage says "person" we say *agent* instead.

An analogy may be helpful. In the Peano axioms for arithmetic [59], the terms *number*, *zero* and *successor* are undefined primitives. They gain meaning by how they play together. Seeking a more elementary meaning of those terms within the same theory is not helpful. Instead of trying an analysis — in the literal sense, a "breaking down" — one develops an understanding by synthesis, by a bringing-together. The same can be said of Hilbert's axiomitazation of geometry, in which *point* and *line* are undefined primitives [60].

The situation in personalist Bayesian probability is somewhat similar. There is no way of carving up the terms *gambler* or *expectation* into smaller conceptual atoms, at least not within probability theory itself. Personalist Bayesianism is a *synthetic theory of quantified expectations,* and there is nothing troublesome about this. QBism simply inherits this situation, applying that synthetic understanding to quantum phenomena.

Just like *point* and *line*, or *zero* and *number* and *successor*, the terms *agent* and *experience* gain meaning through their interplay. Using them in physics brings some baggage from their use in everyday speech, though their meaning is altered — refined, honed —

by deployment in the more quantified setting. This is nothing remarkable: Think of *force*, *potential*, *field* and so forth.

8.18 Does QBism lose the "explanatory power" of other interpretations?

In the philosophy of science, explanations can be causal, unificationist, deductive-nomological, statistical relevantist, inducto-statistical, asymptotic and probably other types besides [61]. Sometimes, epochal progress is made by declaring that an entire genre of attempted explanations is unnecessary, misguided and counterproductive. We've been doing that ever since some clever ancient Greek decided that they could contemplate thunder without drawing the family tree of the Thunderer. While Descartes pictured the planets as being dragged about in a material whirlpool, Newton declared, "I feign no hypotheses" and gave us classical mechanics. The manifold complexities of living beings did not require central planning — only, as Darwin taught us, heredity and luck. Einstein postulated the constancy of the speed of light, without worrying about how moving through the ether might elastically deform the electron, and that is why we learn Lorentz's equations but with Einstein's motivation.

In a sense, Newton explained *less* than Kepler did, because Kepler had a reason why there were six and only six planets: After six planets, we run out of Platonic solids. We can rightly reject Kepler's explanation, even in the absence of a complete story about how the solar system happened — and even though Newton's explanation was, by the standards of his time, frankly un-"physical".[5] Quantum physics leads us to go further than

[5] Kepler's image of nested spheres and regular solids seems absurdly numerological today, though anyone who has wanted E_8 or the Monster group to appear in fundamental physics, just for the æsthetics of it, should feel the tug of the Platonic solids! (We strongly doubt that there is any "theory of everything" inside E_8 [62], although the corresponding lattice does turn out to involve a peculiarly nice quantum measurement [63, 64].) Kepler's geometrical model was wrong, but it was specific, quantitative, directly inspirational and, unlike many bits of our scientific heritage [65], not breathtakingly racist, which maybe counts for something.

Newton. Instead of merely saying “I feign no hypothesis”, we can declare that *the character of the natural world is such that “feigning a hypothesis” — erasing agency and telling a story from a God’s-eye perspective — is a bad idea.* This is an affirmative statement about ontology, and the furthest thing possible from asserting that the world vanishes when I close my eyes (see §8.4).

To ask quantum theory for a story about what happens at the slits of a double-slit experiment “when nobody is looking” is like taking thermodynamics and saying, “OK, but where is the phlogiston?”, or seeing the inverse-square law of gravity and demanding to be shown the dodecahedron that makes it go.

One motivation for the technical side of QBism (see §8.10), particularly the project of reconstructing quantum theory from physical principles, is to elevate the quality of explanations of which quantum physics is capable. The quantum formalism can be applied to any physical system, minuscule or vast, and so any lesson gleaned from the formalism itself must be a very general one — a *why* that pertains, in some measure, anywhere. We physicists tend to like explanations that cut to the fundamental principles of a subject, particularly with a dramatic twist that makes the argument more obvious in retrospect. The opaque nature of the textbook quantum formalism doesn’t just make teaching the subject difficult. (“Master these fifty pages of differential equations and operator theory. Just trust us. Yours not to question why.”) It also buries the enigmatic features of the theory, like the violation of Bell inequalities, and limits physicists’ abilities to devise *good* explanations. We aim to fix this — but *that* is a whole project (§8.10).

When critics have challenged us on the issue of QBism’s “explanatory power”, the type of explanation they’ve often had in mind is something like what solid-state physics has to say about matter being solid. *Pauli exclusion keeps you from falling through the floor; checkmate, QBists!* And in fairness, this does sound rather removed from the scenarios that the QBist literature has mostly dwelled upon — an example of QBism showing its ancestry in quantum information science. Where are agents and interventions in the topics preferred in solid-state society?

In physics, an explanation is not a statement made in isolation. We do not just say, "That rock will sit there without collapsing in on itself." We naturally go a step further: "That rock will resist being squeezed." Squeezing a rock in one's hands *is a quantum measurement* — merely a very imprecise one, for which the textbooks don't say much about representing by a POVM. When we invest meaning in words like *solid* and *rigid* and *incompressible,* we are, at least tacitly, making claims about how a physical system will react against interventions. And thus, even in solid-state mechanics, agenthood was there all along. The fact that we do not make single predictions in isolation is ultimately baked into the formalism, because asserting a quantum state assignment ρ for a system implies quantitative expectations about the outcomes of *any* experiment that one can represent in the theory. No expectation value stands alone.[6]

8.19 Where does the agent end?

At a conference in 2016, Wayne Myrvold asked this:

> Okay, help me understand this restriction of [the] scope of quantum mechanics you're proposing, because you're telling me I should only use quantum mechanics to calculate probabilities for outcomes of my future experiences, and that, compared to what most people think is the scope of the theory, is a

[6] When our measurements are sloppy, we can typically get by without the full apparatus of quantum theory to guide our actions. We can use dodges like average densities of energy levels. We can cheat and model a phenomenon as a classical stochastic process with mundane parameters like average reaction rates. The more closely we interrogate the world, the more we need quantum theory in order to prosper in it. Freedom to intervene, and precision of intervention, are *resources.* When an agent is limited in these regards, the full vitality of quantum phenomena is denied them.

> really serious restriction of scope. So imagine that yesterday someone came to me and said, "Wayne I want your advice on how to construct a nuclear waste storage facility." To do this I need to know about calculating probabilities of decays. So should I not care about any decays that might happen after I'm gone? Would it be a mistake to use quantum mechanics to calculate probabilities of radioactive decays hundreds of years after I'm dead?

The quantum formalism, understood as a normative criterion for an agent's behavior, is rather agnostic about the character of the agent. It says nothing about the agent's memory capacity, their rate of energy consumption, how long they maintain conscious thought at a stretch, or how quickly the molecules of their body are replaced by food. Looking for this kind of information in the quantum formalism confuses the roles of agent and object. If one is dully reductionist and tries to specify the properties of an agent in more and more physical detail, one will eventually be writing a many-body wavefunction. But any wavefunction is only meaningful as a mental tool an agent carries to manage their expectations about something else.

Likewise, the quantum formalism itself does not tell Alice how to attach POVM elements to her experiences. Instead, it is a handbook that she can use to help herself be consistent, howsoever she sets about mathematizing her life. The formalism does not care whether she believes that she will die tomorrow, whether she thinks she can cryogenically freeze herself and wake up on Mars a thousand years from now still essentially Alice, whether she regards potential genetic descendants of herself as sharing in her good or ill fortune — nothing of the sort. Instead, the formalism helps her gamble consistently, using whatever beliefs she currently has about such matters.

In the case of a gamble with consequences beyond an individual's expectations for their own longevity, the "agent" making the bet may be a community, rather than a single human being. Perhaps it is a collaboration of a number of scientists which grows or shrinks as years go by. The situation is similar to that of an individual buying life insurance. Why would anyone ever do this? Life insurance pays out only if the individual making the purchase dies — it's impossible for anyone to reap the benefits of their own life insurance policy. The answer is quite intuitive: Because they consider their family to be an extension of themselves. Even though they, personally, will be gone, a conceptual part of themselves remains which can cash the check. It is like the couple who shares a bank account and makes purchasing decisions on the basis of "us" rather than either of them alone. The concept of an agent is extremely flexible.

Quantum theory tells us that an agent can express her expectations in terms of probabilities for a hypothetical "Bureau of Standards" experiment (see §8.8 and §8.10). The BoS experiment might be exceedingly difficult to carry out: Perhaps it costs a hundred million dollars in optical equipment. But, even though Alice does not physically perform it, it is mentally useful for her in her cogitations.

What about an experiment that requires a forbiddingly large investment of another resource — not money, but *time?* The same binding of expectations between different hypothetical scenarios should still apply. Mathematically, all the prolongation implies is an orthogonal transformation of her probability vector.

To push it a step further: What if Alice contemplates the hypothetical experiment of extending her own life radically? She sees no ready path to doing so, but she lets her imagination wander. Could she replace her neurons one by one with nanomachines? Does her overall mesh of beliefs about her own agenthood permit the idea that any meaningful aspect of her could persist? Even if Alice finds the whole notion exceedingly implausible, can she treat it simply as another experiment that would require a large resource investment to realize?

The QBist answer is “Yes” — or, more carefully, that nothing in the quantum formalism itself forbids it.

We are reminded of a lesson from a colleague.

> This is a good example of the primary point of Dirac notation: it has many built in ambiguities, but it is designed so that any way you chose to resolve those ambiguities is correct. In this way elementary little theorems become consequences of the notation. Mathematicians tend to loathe Dirac notation, because it prevents them from making distinctions they consider important. Physicists love Dirac notation, because they are always forgetting that such distinctions exist and the notation liberates them from having to remember.
>
> — N. David Mermin, “Lecture Notes on Quantum Computation” (2003)

The philosophy of personal identity is brimming with ambiguities, but living in accord with the normative principles of the quantum formalism means that any way I choose to resolve them is correct.

8.20 Aren’t probabilities an insufficient representation of beliefs?

We don’t claim that personalist Bayesian probability theory is the end of the story. We only hold that it is adequate where it is needed: It is a tool applicable when experiments can be defined quantitatively and the sample spaces of their potential outcomes tabulated in advance. BCS notes, “This is one reason why I say *expectation* instead of *belief* sometimes. It carries a bit of a connotation of belief quantified and rigorized, rather than left raw. Plus, the *X* makes it sound cool.”

A great amount of confusion has been stirred up by the misconception that personalist Bayesianism presumes that living human beings actually do act as perfectly rational expectation-balancing agents. In this regard, we share a wry observation of Diaconis and Skyrms [56]:

> In a large and growing experimental literature in psychology and behavioral economics, it appears that *almost all theories are systematically violated by some significant proportion of the population.* It also appears that there are different types in the population. Some violate one principle; some violate another. And there are even some expected utility maximizers.

In other words, the theory of personalist Bayesian probability is normative, not descriptive.

8.21 What does unitary time evolution mean in QBism?

This is a point that we addressed a bit tersely in earlier publications [3, 4, 15, 24], and which we approached from multiple directions amid a large samizdat of miscellany [55]. In this section, we will attempt a balance between these two levels of verbosity.

Fuchs pointed out some time ago that the arguments for the subjectivity of quantum states also apply to unitary time evolutions [24]. Unitaries can be toggled from a distance; they can be teleported. More recent work on the probabilistic representation of quantum theory makes the point even more directly: Quantum states are probability distributions (§8.8), and unitaries are *conditional probabilities*, used in a way that respects the nonexistence of hidden variables.

Consider an agent Alice, who uses quantum theory to help herself navigate the world. Accordingly, she carries a probability distribution for an informationally complete measurement, which she uses to summarize her expectations. Alice can calculate other probability distributions from it, including distributions for other informationally complete measurements which she might carry out in the distant future. The textbook way of writing a unitary evolution is to say

$$\rho' = U(t)\rho U(t)^\dagger, \tag{8.10}$$

where the operator $U(t)$ models the passage of an amount of time t. Both density operators ρ and ρ' express beliefs that Alice holds *now.* The former encodes her present beliefs about a reference measurement performed immediately, while the latter encodes her beliefs about what she might experience were she to instead perform that reference measurement at a later time. All of these beliefs, which she expresses quantitatively as gambling commitments, are commitments she makes at the present time. If time 0 is Monday at noon, and time t is noon on Tuesday, then ρ is Alice's gambling commitment about a measurement on Monday, and ρ' is the commitment she holds simultaneously about a measurement to potentially be done on Tuesday. The unitary operator $U(t)$ is, likewise, a belief that she holds as part of the same mesh of expectations that includes ρ and ρ'. It is, in this sense, a statement *synchronic* with ρ and ρ'. It does *not* express how Alice's beliefs must necessarily change as time passes, though if Tuesday rolls around and Alice has not yet performed a measurement, she can adopt her old numbers ρ' as her new expectations for a reference measurement.

For simplicity, let's assume that the system Alice is contemplating experiments upon is one for which she knows a SIC. She represents her quantum state ρ using the Born Rule as

$$P(H_i) = \frac{1}{d}\mathrm{tr}\left(\rho\Pi_i\right), \tag{8.11}$$

where the projectors $\{\Pi_i\}$ satisfy

$$\mathrm{tr}\left(\Pi_i\Pi_j\right) = \frac{d\delta_{ij}+1}{d+1}. \tag{8.12}$$

Then, she can write the Born Rule for any other POVM $\{D_j\}$ as a simple modification of the Law of Total Probability:

$$Q(D_j) = \sum_{i=1}^{d^2} \left[(d+1)P(H_i) - \frac{1}{d} \right] P(D_j|H_i) \,. \tag{8.13}$$

If Alice uses something other than a SIC as her reference measurement $\{H_i\}$, the formula will be more complicated, but the concepts are the same. A unitary transformation of ρ can be shifted onto the elements of the reference measurement, since

$$P'(H_j) = \mathrm{tr}\,(\rho' H_j) = \mathrm{tr}\,[U(t)\rho U^\dagger(t) H_j] = \mathrm{tr}\,[\rho(U^\dagger(t) H_j U(t))] \,. \tag{8.14}$$

Expressing ρ in terms of $P(H_i)$, this becomes

$$P'(H_j) = \sum_{i=1}^{d^2} \left[(d+1)P(H_i) - \frac{1}{d} \right] R(H_j|H_i, t) \,, \tag{8.15}$$

where

$$R(H_j|H_i, t) = \frac{1}{d} \mathrm{tr}\,(U(t)\Pi_i U^\dagger \Pi_j) \tag{8.16}$$

is a doubly stochastic matrix. Note that classically, we would express the relation between Alice's current probabilities for a measurement now and her current probabilities for a measurement later as

$$P'(H_j) = \sum_{i=1}^{d^2} P(H_i) R(H_j|H_i, t) \,. \tag{8.17}$$

In quantum mechanics, we cannot think of time evolution as shifts in the values taken by hidden variables, so we do not use this expression, but rather its quantum replacement, which simplifies to

$$P'(H_j) = (d+1) \sum_{i=1}^{d^2} P(H_i) R(H_j|H_i,t) - \frac{1}{d}\,. \tag{8.18}$$

All of the quantities $P(H_i)$, $P'(H_j)$ and $R(H_j|H_i,t)$ are beliefs that Alice holds simultaneously. They all have the same status, in that they are personalist Bayesian probabilities, every one of them.

Over the years, we have noticed that some people who are on board with quantum states being subjective still balk at the prospect of regarding quantum *operations* that way. Imbuing unitaries with subjectivity, they fear, risks the whole Standard Model going up in a puff of arthouse smoke. This concern is understandable, but misplaced. We find that a personalist Bayesian take on unitaries, and the "all QFTs are effective QFTs" ethos of weekday field theory [66], meet quite nicely if we only let them. What follows is speculation for the future development of physics, guided by the lessons of practical applications.

Let's suppose that Alice is a physicist who is preparing to do an experiment, say on a spin system. Following ordinary procedure, she writes down a unitary time-evolution operator generated by a Hamiltonian. What does this Hamiltonian encode? Well, it expresses what Alice is doing with her laboratory equipment: the $\vec{E}$ and $\vec{B}$ fields established by charged capacitor plates and current-carrying wires, for example. An old book might have called this information "a complete description of the apparatus in everyday language, suitably augmented with the concepts of classical physics". But Alice knows that she can treat any item of her laboratory apparatus as a quantum system in its own right. For instance, she can use the quantum theory of solids to explain why she can force a current through her coiled wire. So, that which she expresses as a unitary operator, she also recognizes from a broader perspective as a mathematical consequence, in principle, of a quantum state assignment. The "effective unitary" she implements naturally has, therefore, the same physical status as her quantum-state ascriptions: They are all, at root, personalist Bayesian expectations.

What, then, of the most "fundamental" time evolutions of all? Let us go all the way, or at least as far as modern physics can take us. What is the status of the Standard Model Lagrangian in QBism? Apart from the last two words, this is a question already generations old; the project of grand unification, seen as quite respectable, has with game persistence tried to understand the Standard Model as the low-energy limit of a new theory, not too dissimilar to it in basic conceptions.

We do not want to prejudge the matter and ennoble some part of a theory too rashly. After all, human mathematicians have yet to express a nontrivial QFT in a way that meets even their own standards of rigor, let alone a way that would be suitable for the "eyes of God". Among the manifold interesting complications is the fact that not all QFTs are written in terms of a Lagrangian, and it is conceivable that not all of them can be [67].

In order to wring practical numbers out of a QFT, one admits that the theory only applies up to some high-energy or short-distance cutoff, and then one deals with this limitation in an emotionally mature manner. This discipline is known as regularization and renormalization [68]. A scattering amplitude is computed as a function, not just of particle momenta and coupling strengths, but also of the ultraviolet cutoff. Changes in some of these parameters can be absorbed by changes in others, leaving the scattering amplitude numerically unchanged. The theory is not just a single choice of terms and coefficients, but the entire renormalization-group flow.

Seen in this light, the core of a QFT begins to take on a role akin to the Born Rule: a normative constraint relating expectations for different experiments. The story of integrating over UV degrees of freedom, beta functions, the running of coupling "constants" — it brings the message that gambles at one energy ought to be tied with gambles at another.[7]

It is conceivable that when the foundational unscrambling (§8.6) is complete, unitary operators will join quantum states on the doxastic side of the line, while the fundamental core of a "grand unified" theory will, like the Born Rule, reveal itself as an empirically motivated, normative addition to probability theory.

8.22 What do the recent Extended Wigner's Friend thought-experiments imply for QBism?

The past few years have seen the birth of a mini-field, where the thought-experiment called Wigner's Friend is wrapped around a no-hidden-variables argument [5, 6, 71, 72].

We've read the papers, we've been to a workshop [73], and we're still not convinced that the introduction of additional friends, robots or Wigners goes beyond the original Wigner's Friend "paradox" that QBism already answered on its own terms [3]. In order to deserve attention, a "paradox" should reveal an actual inconsistency following from the premises of some interpretation. From our perspective, every new variant just does other things that we know are fallacious: treating unitaries as ontic, acting like systems have a quantum state when there is no agent to assign one, pretending that probabilities follow from frequencies, etc. The extra complexity introduces additional opportunities for confusion, without making the argument more forceful. We would *like* this situation to change; for example, it would be interesting to derive a new quantitative criterion of classicality from these considerations. But the mini-field that studies Wigner's friends, cousins and former roommates is not quite there yet. So far, the conditions deduced from

[7] A related hint comes from lattice gauge theory, where the gauge group is specified at a quite primitive level of setting up the problem [69, §VII.1], much like the selection of Hilbert-space dimensionality in quantum computation, and what follows is rather like a complexified MaxEnt [70].

these thought experiments have been Bell-type inequalities given slightly rephrased justifications, and thus they have not pointed to fundamentally novel issues.

8.23 Is QBism compatible with the Many Worlds Interpretation?

Sometimes, when ideas are presented as going off in two opposite directions, the reason is that they really are, and there isn't any secret centrist wisdom in trying to yoke them back together.

There is no one single Everettian faith, any more than there is truly a unified "Copenhagen Interpretation" (see §8.5). Instead, the genus has many species, frequently incompatible with one another [74]. On rare occasions, an apostle of one of these creeds might make a statement that, in isolation, has a vague affinity to a QBist position. That much is to be expected, since we are all talking about quantum physics, and we are not trying to hang a bag of hidden variables on the side of it (as, say, the Bohmians are wont to do). But we QBists have no physical state vector for the entire universe, no All-Function evolving unitarily in the eye of God.

Imagine, if you can, a physical state vector for the entire cosmos $|\Psi\rangle$, and a factorization of the cosmic Hilbert space into distinguished subsystems. (An Everettian creed will either presume this or attempt to derive it, generally by way of an argument that turns out to be circular.) Now, pick one of those subsystems and take the trace of $|\Psi\rangle$ over all the others. The marginal state of the focal subsystem is then the unique, physically mandated density operator for that subsystem, fixed by ontology. But in QBism, there is no such thing.

The same holds true if one tries to decompose the All-Function into "relative states" of observers and observed. When the carving is all done, the pieces are each physically mandated, ontologically fixed — and that's simply not the role that any quantum state plays in QBism.

A typical move for modern Everettians is to take the quantum-mechanical formalism, chop off the Born rule and then claim to re-derive it. Generally, the algebra can be made to cough up a set of numerical weights, but the identification of those weights as probabilities in any meaningful sense turns out rather unwarranted.

Take another look at the infrastructure underlying the Everettian story: complex Hilbert space, time evolution as unitary operator, etc. To us, *all* of those cry out for explanation. Indeed, the Born Rule, the very part of the theory that Everettians wish to excise — the part to be re-derived as a technicality, delegated to the afterthoughts — may be the most important part of all. Properly formulated, it might well bring the essential enigma of the quantum into the spotlight with a clarity never before achieved [36].

By contrast, we see nothing in the Everettian picture that is uniquely compelled by quantum theory specifically. For instance, you *could* invent a Many-Worlds Interpretation of Spekkens' toy model (as John Smolin once admitted [55, p. 1407]). The result would be baroque and contrived, revealing nothing about the model itself.

We suspect that the appeal of multiverse imagery has more to do with psychology than with physics. Quoting a letter Fuchs wrote in 2002 [55, p. 347]:

> What I find egocentric about the Everett point of view is the way it purports to be a means for us little finite beings to get outside the universe and imagine what it is doing as a whole. And what is it doing as a whole? Something fantastic? Something almost undreamable?! Something inexpressible in the words of man?!?! Nope. It's conforming to a scheme some guy dreamed up in the 1950s.

> This whole fantastic universe can be boiled down to something representable within one of its most insignificant components — the brain of man. Even toying with that idea, strikes me as an egocentrism beyond belief. The universe makes use of no principle that cannot already be stuffed into the head of an average PhD in physics? The chain of logic that leads to the truth of the four-color theorem (apparently) can't be stuffed into our heads, but the ultimate operating principle for all that "is" and "can be" can?

Other varieties of multiversitarianism also leave us unmoved. To adapt a line of Martin Gardner, observable universes are not even as common as two blackberries. Proclamations about "the multiverse" appear to us like failures of imagination, wrapped up in extravagances that provide a certain unsubtle, bulk-rate imitation of it. Our cynical view of these proclamations may be due to our preference for the philosophy of pragmatism.[8]

Most likely, we are doing ourselves few favors in the pop-science media by taking this position, but we are willing to be cast as the stodgy ones.

As for the high-flying speculations of the "all mathematical structures are physically real" variety, we find that an observation by the philosopher William James rather encapsulates our sentiments. The quote that follows is from a 1906 lecture. While a modern multiversitarian would use newer terminology, it boils down to nothing essentially differ-

[8]It also seems to us that arguments in this area tend to disconnect from actual scientific progress. For example, it is a genre convention to quote Weinberg's "prediction" of a small, nonzero cosmological constant from anthropic reasoning [75]. Varying one parameter in isolation — a parameter that we have no good reason to consider fundamental [76], at that — while holding all others fixed strikes us as having dubious physical relevance. Moreover, Weinberg's calculation requires as input the maximal observed redshift of a galaxy [77]. His formula coughed up a decent answer when this was $z = 4.4$, but it fares dramatically worse now that we have seen a galaxy at $z = 11.1$ [78]. Weinberg's argument now gives a bound on the vacuum energy density of about 5800 times the present cosmic mass density. This is three orders of magnitude larger than the observed value, a ratio well into the regime where Weinberg himself says the cosmological constant would be "so small that even the anthropic principle could not explain its smallness" [79].

ent from the “Absolute” and the “mind of God” that had taken hold of the “rationalists” at the time.

> The more absolutistic philosophers dwell on so high a level of abstraction that they never even try to come down. The absolute mind which they offer us, the mind that makes our universe by thinking it, might, for aught they show us to the contrary, have made any one of a million other universes just as well as this. You can deduce no single actual particular from the notion of it. It is compatible with any state of things whatever being true here below. [. . .] Absolutism has a certain sweep and dash about it, while the usual theism is more insipid, but both are equally remote and vacuous.

8.24 What are good things to read about QBism?

While we’re quoting William James, it’s a good time to share a remark from his *Pragmatism* (1907), which by itself is enough to elevate him to the first rank of intellectuals:

> Whatever universe a professor believes in must at any rate be a universe that lends itself to lengthy discourse.

Accordingly, there is no shortage of primary sources about QBism. The essay by Fuchs, Mermin and Schack in the *American Journal of Physics* introduces the interpretation with an emphasis on how it gives meaning to the standard mathematical formulation of quantum theory [4]. Mermin [12, 80] and Fuchs [13, 81] have both written pieces that go more in depth on the historical setting of QBism. Of these essays, Fuchs’s explains more of the technical side of current research. Additional details of that technical work are presented in [3, 4]. Fuchs also discusses the genesis of QBism in the introduction to the samizdat compilation [55].

As for secondary sources, the *Stanford Encyclopedia of Philosophy* has a pretty good article on QBism and related interpretations:

`https://plato.stanford.edu/entries/quantum-bayesian/`

This was written by Richard Healey, who is not a QBist but has an interpretational attitude that is in many ways QBism-adjacent. Being written for an *SEoP* audience, it is heavier on the philosophical matters and gives less time to the technical research that those matters have motivated.

If you want a whole book that you can carry around, Hans von Baeyer's *QBism: The Future of Quantum Physics* (Harvard University Press, 2016) is an accurate portrayal, pitched to the interested-layperson audience.

(And incidentally, on the topic of books, Persi Diaconis and Brian Skyrms recently released *Ten Great Ideas about Chance*, which lays out a school of thought about probability that is pretty much aligned with the one QBism adopts. Diaconis and Skyrms confine the quantum stuff to a single chapter, but they do recommend a David Mermin essay on QBism as good reading [56].)

QBism has been written up both in *New Scientist* [82] and in *Scientific American* [83], though not terribly accurately in either case, thanks to the editorial process [80, 84, 85]. A better treatment, albeit in German, appeared in the *Frankfurter Allgemeine Sonntagszeitung* [86]. *Nature* addressed it briefly in the context of information-oriented reconstructions of quantum theory [87].

In June 2015, the pop-science website *Quanta Magazine* ran an interview with Fuchs [2]. The accompanying profile is largely accurate, except for a figure caption that implies QBism is a hidden-variable theory:

> A quantum particle can be in a range of possible states. When an observer makes a measurement, she instantaneously "collapses" the wave function into one possible state. QBism argues that this collapse isn't mysterious. It just reflects the updated knowledge of the observer. She didn't know where the particle was before the measurement. Now she does.

A better caption would go more like the following:

> In the textbook way of doing quantum physics, a quantum particle has a "wave function" that changes smoothly when no one is looking, but which makes a sharp jump or "collapse" when the particle is observed. QBism argues that this collapse isn't mysterious. It just reflects the altered expectations of the observer. Before the measurement, she didn't know what would happen to her when she interacted with the particle. After the measurement, she can update her expectations for her future experiences accordingly.

Originally, the subhead was also misleading; soon after the interview appeared, *Quanta* fixed the subhead, but not the figure caption. So it goes.

Later, Fuchs was interviewed for the Australian Broadcasting Company's program, *The Philosopher's Zone* [88].

LIST OF REFERENCES

[1] N. D. Mermin, “QBism as CBism: Solving the Problem of "the Now",” *arXiv:1312.7825*, 2013.

[2] A. Gefter, “Quantum Bayesianism Explained by Its Founder,” *Quanta Magazine*, 2015.

[3] C. A. Fuchs and B. C. Stacey, “QBism: Quantum Theory as a Hero’s Handbook,” in *Proceedings of the International School of Physics "Enrico Fermi": Course 197, Foundations of Quantum Theory*, pp. 133–202, 2019.

[4] C. A. Fuchs and R. Schack, “Quantum-Bayesian coherence,” *Reviews of Modern Physics*, vol. 85, no. 4, pp. 1693–1715, 2013.

[5] D. Frauchiger and R. Renner, “Quantum theory cannot consistently describe the use of itself,” *Nature Communications*, vol. 9, no. 1, p. 3711, 2018.

[6] Č. Brukner, “On the Quantum Measurement Problem,” in *Quantum [Un]Speakables II* (R. Bertlmann and A. Zeilinger, eds.), pp. 95–117, Cham: Springer International Publishing, 2017.

[7] M. Gardner, “Why I Am Not a Solipsist,” in *The Whys of a Philosophical Scrivener*, Quill, 1983.

[8] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, “Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems,” *Physical Review Research*, vol. 2, p. 013074, 2020.

[9] H. Atmanspacher, “20th century variants of dual-aspect thinking,” *Mind and Matter*, vol. 12, 2014.

[10] N. D. Mermin, “Physics: QBism puts the scientist back into science,” *Nature*, vol. 507, no. 7493, pp. 421–423, 2014.

[11] B. C. Stacey, “Is the SIC Outcome There When Nobody Looks?,” *arXiv:1807.07194*, 2018.

[12] N. D. Mermin, “Making better sense of quantum mechanics,” *Reports on Progress in Physics*, vol. 82, no. 1, p. 012002, 2019.

[13] C. A. Fuchs, “Notwithstanding Bohr, the Reasons for QBism,” *Mind and Matter*, vol. 15, no. 2, 2017.

[14] D. Bohm, *Quantum theory*. New York: Dover Publications, 1989.

[15] B. C. Stacey, “Von Neumann was not a Quantum Bayesian,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2068, p. 20150235, 2016.

[16] N. Van Kampen, “Ten theorems about quantum mechanical measurements,” *Physica A: Statistical Mechanics and its Applications*, vol. 153, no. 1, pp. 97–113, 1988.

[17] E. T. Jaynes, “Probability in quantum theory,” in *Complexity, Entropy and the Physics of Information* (W. H. Zurek, ed.), Addison–Wesley, 1990.

[18] E. T. Jaynes, “Cleaning up the mysteries – the original goal,” in *Maximum-Entropy and Bayesian Methods* (J. Skilling, ed.), Springer, 1989.

[19] R. W. Spekkens, “Evidence for the epistemic view of quantum states: A toy theory,” *Physical Review A*, vol. 75, no. 3, p. 032110, 2007.

[20] R. W. Spekkens, “Reassessing claims of nonclassicality for quantum interference phenomena.” Presentation at Perimeter Institute: `http://pirsa.org/16060102/`, 2016.

[21] M. F. Pusey, J. Barrett, and T. Rudolph, “On the reality of the quantum state,” *Nature Physics*, vol. 8, no. 6, pp. 475–478, 2012.

[22] C. A. Fuchs and R. Schack, “Bayesian Conditioning, the Reflection Principle, and Quantum Decoherence,” in *Probability in Physics* (Y. Ben-Menahem and M. Hemmo, eds.), pp. 233–247, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[23] B. C. Stacey, *Multiscale Structure in Eco-Evolutionary Dynamics*. PhD thesis, Brandeis University, 2015.

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

[25] C. M. Caves, C. A. Fuchs, and R. Schack, “Unknown quantum states: The quantum de Finetti representation,” *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4537–4559, 2002.

[26] C. A. Fuchs, R. Schack, and P. F. Scudo, “De Finetti representation theorem for quantum-process tomography,” *Physical Review A*, vol. 69, no. 6, p. 062305, 2004.

[27] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, “Noncommuting Mixed States Cannot Be Broadcast,” *Physical Review Letters*, vol. 76, no. 15, pp. 2818–2821, 1996.

[28] C. M. Caves, C. A. Fuchs, and R. Schack, "Conditions for compatibility of quantum state assignments," *Physical Review A*, vol. 66, no. 6, 20002.

[29] C. A. Fuchs and A. Peres, "Quantum-state disturbance versus information gain: Uncertainty relations for quantum information," *Physical Review A*, vol. 53, no. 4, pp. 2038–2045, 1996.

[30] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, "Optimal Eavesdropping in Quantum Cryptography. I," *arXiv:quant-ph/9701039*, 1997.

[31] A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, "Unconditional Quantum Teleportation," *Science*, vol. 282, no. 5389, pp. 706–709, 1998.

[32] C. A. Fuchs and K. Jacobs, "Information-tradeoff relations for finite-strength quantum measurements," *Physical Review A*, vol. 63, no. 6, p. 062305, 2001.

[33] A. Peres and D. R. Terno, "Quantum information and relativity theory," *Reviews of Modern Physics*, vol. 76, no. 1, pp. 93–123, 2004.

[34] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, "The SIC question: History and state of play," *Axioms*, vol. 6, p. 21, 2017.

[35] B. Stacey, "SIC-POVMs and Compatibility among Quantum States," *Mathematics*, vol. 4, no. 2, p. 36, 2016.

[36] M. Appleby, C. A. Fuchs, B. C. Stacey, and H. Zhu, "Introducing the Qplex: a novel arena for quantum theory," *The European Physical Journal D*, vol. 71, no. 7, p. 197, 2017.

[37] J. B. DeBrota and C. A. Fuchs, "Negativity bounds for Weyl–Heisenberg quasiprobability representations," *Foundations of Physics*, vol. 47, pp. 1009–30, 2017.

[38] M. Appleby, T.-Y. Chien, S. Flammia, and S. Waldron, "Constructing exact symmetric informationally complete measurements from numerical solutions," *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 16, p. 165302, 2018.

[39] G. S. Kopp, "SIC-POVMs and the Stark conjectures," *International Mathematics Research Notices*, p. rnz153, 2019.

[40] J. C. Baez and J. Biamonte, "Quantum Techniques for Stochastic Mechanics," *arXiv:1209.3632*, 2019.

[41] M. Schlosshauer, "The quantum-to-classical transition and decoherence," *arXiv:1404.2635*, 2019.

[42] C. Rovelli, “Relational quantum mechanics,” *International Journal of Theoretical Physics*, vol. 35, no. 8, pp. 1637–1678, 1996.

[43] F. Laudisa and C. Rovelli, “Relational Quantum Mechanics,” in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), Metaphysics Research Lab, Stanford University, winter 2019 ed., 2019.

[44] D. M. Appleby, “Facts, Values and Quanta,” *Foundations of Physics*, vol. 35, no. 4, pp. 627–668, 2005.

[45] D. M. Appleby, “Probabilities Are Single-Case or Nothing,” *Optics and Spectroscopy*, vol. 99, no. 3, p. 447, 2005.

[46] M. Smerlak and C. Rovelli, “Relational EPR,” *Foundations of Physics*, vol. 37, no. 3, pp. 427–445, 2007.

[47] Q. Ruyant, “Can We Make Sense of Relational Quantum Mechanics?,” *Foundations of Physics*, vol. 48, no. 4, pp. 440–455, 2018.

[48] C. A. Fuchs, “QBism, The Perimeter of Quantum Bayesianism,” *arXiv:1003.5209*, 2010.

[49] N. D. Mermin, “Quantum Mysteries for Anyone,” *The Journal of Philosophy*, vol. 78, no. 7, pp. 397–408, 1981.

[50] G. ’t Hooft, “Dimensional reduction in quantum gravity,” *Conf. Proc. C*, vol. 930308, pp. 284–296, 1993.

[51] E. Witten, “Quantum gravity in de Sitter space,” in *Strings 2001: International Conference*, 2001.

[52] M. Parikh and E. Verlinde, “De Sitter Holography with a Finite Number of States,” *Journal of High Energy Physics*, vol. 2005, no. 01, pp. 054–054, 2005.

[53] T. Banks and W. Fischler, “The holographic spacetime model of cosmology,” *International Journal of Modern Physics D*, vol. 27, no. 14, p. 1846005, 2018.

[54] C. A. Fuchs, “On the Quantumness of a Hilbert Space,” *Quantum Information & Computation*, pp. 467–78, 2004.

[55] C. A. Fuchs, M. Schlosshauer, and B. C. Stacey, “My Struggles with the Block Universe,” *arXiv:1405.2390*, 2014.

[56] P. Diaconis and B. Skyrms, *Ten great ideas about chance*. Princeton: Princeton University Press, 2018.

[57] C. A. Fuchs, R. Schack, L. Accardi, G. Adenier, C. Fuchs, G. Jaeger, A. Y. Khrennikov, J. Å. Larsson, and S. Stenholm, "Priors in Quantum Bayesian Inference," in *AIP Conference Proceedings*, (Växjö, Sweden), pp. 255–259, 2009.

[58] L. J. Savage, *The foundations of statistics*. New York: Dover Publications, 2d rev. ed ed., 1972.

[59] J. C. Baez, "Surprises in Logic." `http://math.ucr.edu/home/baez/surprises.html`, 2016.

[60] M. Shulman, "Homotopy type theory: A synthetic approach to higher equalities," in *Categories for the Working Philosopher* (E. Landry, ed.), Oxford University Press, 2017.

[61] J. O. Weatherall, "On (Some) Explanations in Physics," *Philosophy of Science*, vol. 78, no. 3, pp. 421–447, 2011.

[62] J. Distler and S. Garibaldi, "There is No "Theory of Everything" Inside E8," *Communications in Mathematical Physics*, vol. 298, no. 2, pp. 419–436, 2010.

[63] B. C. Stacey, "Geometric and Information-Theoretic Properties of the Hoggar Lines," *arXiv:1609.03075*, 2016.

[64] B. C. Stacey, "Sporadic SICs and the normed division algebras," *Foundations of Physics*, vol. 47, pp. 1060–64, 2017.

[65] J. Bouie, "How the Enlightenment Created Modern Race Thinking and Why We Should Confront It," *Slate Magazine*, 2018.

[66] J. Zinn-Justin, *Phase transitions and renormalization group*. Oxford graduate texts, Oxford, United Kingdom: Oxford University Press, 2013.

[67] "Lagrangian quantum field theory in nLab." `https://ncatlab.org/nlab/show/Lagrangian+quantum+field+theory`, 2017.

[68] J. C. Baez, "Struggles with the Continuum," *arXiv:1609.01421*, 2020.

[69] A. Zee, *Quantum field theory in a nutshell*. Princeton, N.J: Princeton University Press, 2nd ed., 2010.

[70] J. Baez and B. Pollard, "Quantropy," *Entropy*, vol. 17, no. 2, pp. 772–789, 2015.

[71] M. F. Pusey, "An inconsistent friend," *Nature Physics*, vol. 14, no. 10, pp. 977–978, 2018.

[72] J. Bub, "In defense of a "single-world" interpretation of quantum mechanics," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, p. S1355219817301788, 2018.

[73] "Encapsulated Agents in Quantum Theory: Re-examining Wigner's Friend," (Boston, MA), 2019.

[74] A. Kent, "Does it Make Sense to Speak of Self-Locating Uncertainty in the Universal Wave Function? Remarks on Sebens and Carroll," *Foundations of Physics*, vol. 45, no. 2, pp. 211–217, 2015.

[75] G. F. R. Ellis, "Physics on Edge," *Inference*, vol. 3, no. 2, 2017.

[76] Q. Wang, Z. Zhu, and W. G. Unruh, "How the huge energy of quantum vacuum gravitates to drive the slow accelerating expansion of the Universe," *Physical Review D*, vol. 95, no. 10, p. 103504, 2017.

[77] S. Nadathur, "Blank On The Map: Does the multiverse explain the cosmological constant?." `https://blankonthemap.blogspot.com/2014/02/does-multiverse-explain-cosmological.html`, 2014.

[78] P. A. Oesch, G. Brammer, P. G. v. Dokkum, G. D. Illingworth, R. J. Bouwens, I. Labbé, M. Franx, I. Momcheva, M. L. N. Ashby, G. G. Fazio, V. Gonzalez, B. Holden, D. Magee, R. E. Skelton, R. Smit, L. R. Spitler, M. Trenti, and S. P. Willner, "A Remarkably Luminous Galaxy at z=11.1 Measured with *Hubble Space Telescope* Grism Spectroscopy," *The Astrophysical Journal*, vol. 819, no. 2, p. 129, 2016.

[79] S. Weinberg, "Anthropic Bound on the Cosmological Constant," *Physical Review Letters*, vol. 59, no. 22, pp. 2607–2610, 1987.

[80] N. D. Mermin, "Why QBism Is Not the Copenhagen Interpretation and What John Bell Might Have Thought of It," in *Quantum [Un]Speakables II* (R. Bertlmann and A. Zeilinger, eds.), pp. 83–93, Cham: Springer International Publishing, 2017.

[81] C. A. Fuchs, "Interview with a Quantum Bayesian," *arXiv:1207.2141*, 2012.

[82] M. Chalmers, "QBism: Is quantum uncertainty all in the mind?," *New Scientist*, no. 2968, pp. 32–25, 2014.

[83] H. C. von Baeyer, "Can Quantum Bayesianism Fix the Paradoxes of Quantum Mechanics?," *Scientific American*, vol. 308, pp. 46–51, 2013.

[84] N. D. Mermin, "QBism in the New Scientist," *arXiv:1406.1573*, 2014.

[85] N. D. Mermin, “Putting the Scientist into the Science.” Presentation at *Quantum [Un]Speakables II: 50 Years of Bell’s Theorem*, University of Vienna, 2014.

[86] U. von Rauchhaupt, “Philosophische Quantenphysik : Ganz im Auge des Betrachters [Philosophical Quantum Physics: Only in the Eye of the Beholder],” *FAZ.NET*, 2014.

[87] P. Ball, “Physics: Quantum Quest,” *Nature*, vol. 501, no. 7466, pp. 154–156, 2013.

[88] J. Gelonasi, “Why the multiverse is all about you,” *ABC Radio National Program: The Philosopher’s Zone*, 2015.

CHAPTER 9

CONCLUSION

It is time for some concluding remarks. In the introduction, we noted that our technical investigations were motivated by the QBist foundational perspective. Walking just a few steps along that path revealed unknowns we needed to tackle. Each one we dealt with produced at least a few more. And so it goes and grows. Now, at the end of the adventure that is graduate school, the author finds himself more excited by this general direction of research than ever before. Perhaps most exciting of all, we are beginning to see possibilities for inroads and bridges to other subfields. The next step is to clearly articulate these connections. The payoffs could be remarkable; some of the most important advances happen when diverse backgrounds meet.

Further study of minimal informationally complete measurements and their associations to other structures is likely to produce considerable advances in our understanding of quantum theory. It is a widely unappreciated fact that a quantum state may be thought of as a *single* probability distribution. The vast majority of work on quantum theory restricts attention almost exclusively to orthogonal projective measurements. Obviously these have their place, uses, and practicalities, but if one only ever thinks of measurements derived from "observables," one will never notice when a different kind of measurement is better suited for their purposes. And for the purpose of understanding what a quantum state actually is, a MIC is clearly superior because it allows us to see that a quantum state is conceptually no more than a probability distribution; as a von Neumann measurement is not informationally complete, its outcome probabilities are only a partial picture. The MIC probabilistic approach we have described in several of the chapters of this dissertation similarly clarifies the meaning of a general measurement — with respect to a fixed reference process, any POVM takes the form of a conditional probability matrix.

But this way of thinking stands a chance of doing much more than clarifying. We have every reason to believe this approach could turn out to be as productive as more mature reformulations of quantum theory. Phase space quantum mechanics and the path integral formulation are useful because they preserve key concepts from pre-quantum physics. Our reference measurement approach does the same thing. This time the most basic concept preserved is that of informational completeness. Classically, an informationally complete measurement is one which reads off the "physical condition" of the system. In mechanics, this would be the (potentially only hypothetical) measurement which reads off the system's precise phase space coordinates. If one has a probability distribution for the outcomes of such a measurement, i.e. a Liouville distribution, then one may compute the probability distribution for any other measurement, say, of the system's total energy, by coarse graining appropriately. Bell and Kochen–Specker type results tell us something about the Liouville picture cannot be made to work. If one insists on leaving the phase space intact, one has to abandon probabilities, in effect abandoning the intuition of informationally complete measurements — because of its negativity, a quasiprobability distribution loses the operational significance that probabilities have. However, if one is ready to learn a different lesson about quantum theory, one can forgo directly imposing a phase space and keep informational completeness instead. In the absence of an obvious notion of an ontic state, hidden variable, physical condition, phase space point, element of reality, or otherwise, it is surprising to classical intuitions that we can nonetheless keep the concept of informational completeness. If not the probability for something that is "out there", what does it mean that such probability distributions contain all of the information about a system necessary to consistently assign probabilities to any other measurement? Our approach provides a framework to attack these types of questions head on.

In Chapter 2, we explored some immediate consequences of the definition of a MIC, primarily through the Gram matrix. Even this close to the ground, it is clear that the native quantum replacements for phase space are extremely diverse and fundamentally different in each Hilbert space dimension. In Chapter 3, we provided for the first time (chronologically) a mathematical justification for the intuition that SICs occupy a privileged position from which to assess the differences between the quantum and the classical. In addition to seeking significant bases for the understanding of quantum theory itself, we extended our knowledge of MICs through their connections to other areas. In the other three technical chapters, this is done in two ways: connection to channels and connection to quasiprobability representations. In Chapter 4, we defined a new class channels called Lüders MIC channels and found that their properties allow for another characterization of SIC existence. We also saw that our definition suggests new information-related questions, some of which we solved with entropic bounds. In Chapter 5, we took MICs in another direction when we noticed that the MIC-specific way of expressing the Born rule naturally suggested an associated Wigner basis as kind of orthogonalization. This observation suggests a strategy for further development and classification of MIC properties and has the potential to reciprocally influence the study of discrete Wigner functions. Although we didn't know it at the time, through this association our study of Wigner bases in the manuscript which became Chapter 6 should have downstream effects for understanding what sum negativity of a Wigner function means in the language of MICs.

There, in current terminology, we revealed that while the principal and shifted principal Wigner bases of a SIC extremize ceiling negativity among all unbiased Wigner bases, this fact does not extend in complete generality to the sum negativity setting, which is also of conceptual and practical interest.

On the conceptual side, we believe that evidence continues to mount that the QBists are on the right track. We hope that Chapter 7, QBism's response to recent extensions of the Wigner's friend paradox, will make this intuition more widespread. For many other interpretations, the original Wigner's friend thought experiment was mildly alarming, although perhaps not catastrophic. The appearance of two recent, more elaborate, variations of the experiment, however, have raised concerns as to the foundational strength of many of the same interpretations. The QBists have always thought the Wigner's friend experiment exemplified a powerful and beautiful fact of reality — it was a feature, not a bug. Although slow to digest, the extensions of the thought experiment have actually strengthened our convictions. The resolution is simply a matter of truly treating all agents on equal footing, an attitude long present in descriptions of QBism, but never quite so forcefully brought to the fore.

Speaking of convictions, we QBists have long said that the best judge of an interpretation of quantum mechanics is in what it spurs. A good interpretation would inspire solutions to outstanding problems, suggest particular questions, and provide intuitions for technologies. A bad one tries to close off the story as quickly as possible, leaving everyone wondering whether it was worth the bother to think more about the meaning of quantum theory at all. Of course, this criterion applies more broadly. Good things respond and grow. It feels like we have started a good thing, that we are on the cusp of important breakthroughs. Well, that will only be true if we make it so. It's time to get to work!

BIOGRAPHICAL SKETCH

John Bernard DeBrota was born in Indianapolis, Indiana on September 27, 1991. He earned B.S. degrees in Physics and Mathematics from Indiana University Bloomington in 2014 and an M.Sc. in Physics from the University of Waterloo in 2015. In August 2020, he will graduate with a Ph.D. in Applied Physics from the University of Massachusetts Boston under the supervision of Christopher A. Fuchs.

In addition to physics, mathematics, and philosophy, he enjoys rock climbing, baking bread, playing the guitar and ukulele, spicy food, seeing new parts of the world, and spending time with family and friends.