Frames, Designs, and Spherical Codes in Quantum Information Theory

by

Joseph M. Renes

Bachelor of Science, California Institute of Technology, 1999
Master of Science, University of New Mexico, 2002

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physics

The University of New Mexico

Albuquerque, New Mexico

May, 2004

With knowledge comes power; from the pursuit, humanity.

Acknowledgments

Thanks go first and foremost to my advisor, Carl Caves, for guidance not only in the subject matter of my research, but also the style in which to carry it out. Carl holds firmly to the view that the job of a theoretical physicist is to make mistakes as fast as possible, and for a graduate student this is perhaps doubly true. He backs this view with thoughtful, reasoned consideration of the inevitable products of this mode of research, unlikely as they may initially appear.

I thank the past and present members of the information physics group for a varied and exciting research atmosphere, as well as friendship. Ivan Deutsch and his research group, particularly Gavin Brennen, Shohini Ghose, René Stock, and Trace Tessier have helped by looking at the field from a different point of view and sharing their outlook. I'm grateful to the "home team" of Kiran Manne, Mark Tracy, and Andrew Scott and my colleagues and coworkers at other insitutions, especially Dave Bacon, Robin Blume-Kohout, and Josh Bienfang, for their patience and willingness to help with problems large and small, often at the spur of the moment.

To Norbert Lütkenhaus I owe a great deal for his incredible ability to frame the core issues of quantum information theory and offer an elegant and coherent overall perspective. This has helped me immensely in setting my own course of research. I am also indebted to Chris Fuchs and Howard Barnum for their eagerness to suggest new research problems, not a few of which form the basis for this dissertation.

I thank the remaining members of my committee, Daniel Finley and Richard Hughes, for their scrutiny of this work and the valuable time it takes. I also appreciate the efforts of Tamara Baca, Mary DeWitt, and Betty Fry, as well as the entire office staff.

To my parents I owe thanks for their unfailing encouragement and support. Finally, Andrea Ostermeyer deserves special thanks, for without her this dissertation would not have been possible.

Abstract

Frame theory offers a lens through which to view a large portion of quantum information theory, providing an organizational principle to those topics in its purview. In this thesis, I cut a trail from foundational questions to practical applications, from the origin of the quantum probability rule to quantum cryptography, by way of a standard quantum measurement helpful in quantum tomography and representation of quantum theory. Before embarking, preparations are undertaken by outlining the relevant aspects of frame theory, particularly the characterization of generalized orthonormal bases in terms of physical quantum measurements, as well as several aesthetically appealing families of measurements, each possessing a high degree of symmetry.

Much more than just elegant, though, these quantum measurements are found to be useful in many aspects of quantum information theory. I first consider the foundational question of justifying the quantum probability rule, showing that putting a probability valuation on generalized quantum measurements leads directly to the Born rule. Moreover, for qubits, the case neglected in the traditional formulation of Gleason's theorem, a symmetric three-outcome measurement called the trine is sufficient to impel the desired form. Keeping with foundational questions, I then turn to the problem of establishing a symmetric measurement capable of effortlessly rendering quantum theory in terms of classical probability theory. Numerical results provide an almost utterly convincing amount of evidence for this, justifying the subsequent study of its use in quantum tomography and detailed account of the properties of the reduction to probabilistic terms.

Saving perhaps the most exciting topic for last, I make use of these aesthetic ensembles in the applied field of quantum cryptography. A large class of streamlined key distribution protocols may be cut from the cloth of these ensembles, and their symmetry affords them improved tolerance to eavesdropping over the traditionally-studied schemes. Because the ability to put quantum key distribution protocols into practice is essentially right around the corner, I conclude by examining the prospects for implementing the new protocols in free space and their ability to boost the operating signal intensity, currently a major obstacle in the development of practical schemes.

Contents

List of Figures

List of Tables

Part I

Prolegomenon

Chapter 1

Introduction

Ranging from the categorization of new complexity classes in theoretical computer science to fine-tuning scattering potentials of neutral atoms trapped in optical lattices, quantum information theory is by now a vast subject. By steadily amalgamating pieces of any field it might find relevant, it has come to count computer science, physics, mathematics, engineering, as well as philosophy toward its corpus. Nominally a marriage of information theory and quantum mechanics, quantum information theory finds that each informs the other. Not only can quantum mechanics solve problems suggested by information theory, but concepts in information theory also aid the understanding of quantum mechanics. Though perhaps originating in physics, its goals are not strictly those of understanding the nature of physical law in its own right. Rather, the field is a good deal more applied, typically oriented around developing useful tools to solve particular problems.

These, then, are the two complementary research thrusts of the field: finding novel solutions to practical problems while developing concepts and methods to better characterize modern physical theory. This approach is evident in the early works of the field, appearing in what would now be regarded as its prehistory, the 1970s. Among the first problems considered was the paradigmatic one, the transmission of classical signals by quantum systems. In principle, for a given physical system the

quantum description includes a wealth of states which a classical description does not, so that one may hope for an improved ability to communicate signals from point to point. This enlarged space turns out to be mostly a mirage, generally worthless for this purpose, as demonstrated by Holevo in 1973 [78]. These additional states are not all mutually-distinguishable, meaning that information encoded into them cannot be reliably recovered. Evidently, whatever is meant by "quantum state" has quite different attributes than "classical state." Not only does this result tell us something useful, namely straightforward use of quantum states for communication won't help, but also something fundamental, quantum systems don't always give straightforward answers to questions the way classical systems do. Improperly-encoded information is trapped by the quantum system.

These dual lines of investigation persist, despite the unclear border between them. In the subsequent years, the field's pioneers discovered more things quantum systems can and cannot do, how they do them, and why. Related to the aforementioned problem of distinguishing among a set of quantum states is the fact that quantum states cannot be cloned, or copied. Of course, copying a classical state is no trouble; whole industries are devoted to high-quality reproductions of printed material, after all. More importantly, in any communication scheme the transmitted signals are invariably amplified at some stage, but this amounts to copying, and so is forbidden in quantum mechanics. Were quantum states cloneable, they could be distinguished: even if a given state is hard to distinguish from others in the sense of having low probability of being done correctly, by creating a huge number of copies eventually it becomes tenable. Hence indistinguishability and no-cloning come as a pair. The mathematical reasons forbidding cloning are simple enough, at least in the restricted case of pure states. Suppose U is a quantum cloning machine, so that feeding in a given state $|\psi\rangle$ and a "blank page" $|0\rangle$, i.e. a physical state prepared in a standard way, results in the output of two copies $|\psi\rangle|\psi\rangle$. Now let $|\phi\rangle$ be another state for which the machine U again outputs two copies. What happens if we try to copy the

superposition of the two $|\psi\rangle + |\phi\rangle$?[1] Surely this is a legitimate quantum state, but now linearity implies that we obtain

$$U(|\psi\rangle + |\phi\rangle)|0\rangle = |\psi\rangle|\psi\rangle + |\phi\rangle|\phi\rangle \neq (|\psi\rangle + |\phi\rangle)(|\psi\rangle + |\phi\rangle), \tag{1.1}$$

which isn't at all what we wanted. Since quantum mechanics is a linear theory, copying of arbitrary states cannot be performed [167]. Poor-quality copies may be obtained, but this would be akin to placing a sheet of text into a copying machine and obtaining two versions each of which is somewhat different than the original. Copying machines today may not be perfect—or even reliable—but at a minimum we expect them not to alter the original, at least not much. Perhaps more than their inherent indistinguishability, the fact that quantum states cannot even be copied or amplified signifies a gigantic departure from classical physics.

Bizarrely, despite the injunction against amplifying quantum states, the information contained can be blindly transmitted. Without knowing the input state it would seem to be impossible to reliably read the information and then send it, and indeed it is. However, with the help of auxiliary systems having nothing to with the input, the state can be measured at one point, the results transmitted in the standard classical fashion, and the state faithfully reconstructed at the other end. This process is termed teleportation [14] since the quantum state is not bodily transmitted—the quantum state at the transmitter's end is destroyed otherwise teleportation would imply cloning. Moreover, the transmitted measurement results are essentially random, making this an almost ridiculous-sounding scheme. The secret, of course, lies in the non-classical properties of the auxiliary state, termed *entangled* by Schrödinger [136]. The output state of the cloning machine in equation 1.1 is entangled; the quantum state is shared between the two systems in a way which cannot be described by considering each individually. Due to the usefulness of entangled states, they are the subject of much current research, which in turn sheds light on how entanglement signals the departure from classical physics.

[1] Here we are avoiding issues of normalization.

Beyond this, quantum systems also exhibit an optimality in the way they convey information through measurement [165, 166]. Considering the restricted case of linearly-polarized photons, the probability distribution proscribed by quantum mechanics conveys more information about the polarization than any other distribution. That is to say, considering the mutual information between the state of the photon polarization and the measurement of this quantity as a function of the probability distribution for the measurement outcome given the polarization, the distribution given by quantum mechanics maximizes this mutual information. This signals a feature of fundamental importance, like the minimization of action in classical mechanics, but now the important quantity is a concept borrowed from information theory, not a strictly physical quantity. Perhaps like the reduction of classical mechanics to the principle of least action a similar reduction can be made for quantum mechanics in terms of information-theoretic quantities.

Quantum systems can also be used for cryptographic purposes, as chapter five examines in detail, a phenomenon which relies entirely on the odd fact that information encoded in quantum states generally cannot even be read without causing some disturbance to it. This would be akin to picking up a book and finding yourself reading Don Quixote the first time around, The Theory of the Leisure Class the second, and the entire collection of Physical Review articles for 1928 the third. Such peculiar behavior was exploited in the first application of quantum mechanics to real world problems, the creation of unforgeable bank notes [163].

The most propulsive idea in quantum information theory arose in the 1980s but did not catch fire until 1994: the quantum computer. An exciting combination of words if ever there was one, the quantum computer was first imagined by Feynman at the same time Wheeler contemplated the merits of viewing the universe—and by extension, physical law—as a computer. Feynman hoped to simulate physical systems more efficiently [62], while Wheeler's ideas provided an impetus for viewing physical problems in information-theoretic terms[2] [162]. Happily, a problem which

[2] "It from bit" as the saying goes.

the quantum computer could solve faster than a classical computer was found in 1985 [51], but it wasn't until the development in rapid succession of an algorithm to very rapidly factor large numbers [142] and an error-correction scheme to ensure the quantum computer's signal would not drown in noise of its own making [143, 146] did researchers at large take notice. A veritable explosion of work followed, from new algorithmic software fodder such as faster searching of databases [72], speedier solution of algebraic equations [74], and quicker transversing of graphs using quantum random walks [35], to fanciful hardware on which to run them, ranging from electromagnetically trapped single neutral atoms [85] or ions [73, 164] to superconducting Josephson-junction circuits.

And this isn't even the most immediately realizable application of quantum information theory. A quantum computer capable of performing tasks faster than available classical computers is likely decades away. For the impatient, quantum mechanics also offers the ability to improve measurement of classical parameters [36] (probably a requirement for fruitful functioning of gravitational-wave detectors [31]), as well as frequency standards [80], lithographic techniques [17], and clock synchronization [38], in addition to the aforementioned cryptographic applications.

It has been remarked that to do research in quantum information theory, one ought to pick a favorite text on classical information theory [110], open to a chapter, and translate the contents into quantum-mechanical language [138]. Naturally, if everyone followed this advice the field would quickly be overrun with books entitled "Elements of Quantum Information Theory" [42]. But the statement succinctly captures the fundamental point of research in quantum information—quantum mechanics offers a fresh look at existing problems, and in venturing to adopt this perspective and solve these problems we are likely to learn a lot besides.

Instead of looking to information theory for guidance, this dissertation draws heavily on the subject of *frame theory* for inspiration and structure. Frame theory doesn't offer a comprehensive framework for quantum information, but does provide a detailed tour through a wide-ranging section of it. Several good introductions to

the field exist, particularly by Casazza [29]. Topics from frame theory are used both to develop useful tools in quantum information theory, as described in part one, and to explore new and improved methods of quantum cryptography, the subject of part two.

First we take a brief look at frame theory in chapter two, amounting to a tiny fraction of the theory, but just enough for our purposes. At its most general, frame theory generalizes the notion of an orthogonal basis in a vector space and proceeds to consider the questions of functional analysis in these terms. For instance, conditions on frames ensuring their completeness as well as those establishing equivalence to ordinary bases are well-known. The usefulness of such sets of vectors was elucidated by their creator, Dennis Gabor, who in 1946 sought a signal decomposition localized both in time *and* frequency [66], much like the way a conductor's score describes a symphony.

Often two research fields may proceed in isolation, each one developing many of the same concepts and results without realizing it. Such is the case for frame theory and quantum information, since upon close inspection, frames are found to be equivalent to generalized quantum measurements, POVMs. Any vector may be expressed in terms of the elements of a frame, this being the very definition. That is to say, for any $|\psi\rangle$ we may write $|\psi\rangle = \sum_k c_k(\psi)|\phi_k\rangle$ for some generally non-unique choice of constants c_k. Any frame may be transformed in a canonical sort of way into an isomorphic frame $\{|\widetilde{\phi}_k\rangle\}$ for which

$$\sum_k |\widetilde{\phi}_k\rangle\langle\widetilde{\phi}_k| = aI, \tag{1.2}$$

where $a > 0$ and I represents the identity operator. But now we have a set of positive operators $E_k = |\widetilde{\phi}_k\rangle\langle\widetilde{\phi}_k|/a$ which sum to the identity, i.e. a set of operators forming a POVM. Too, the condition expressing the equivalence of a frame to an orthonormal basis in a larger vector space is also well-known in quantum information theory as Neumark's theorem. Several authors have noticed this and other connections and attempted to close the gap between the two fields [58, 59]. Hopefully this dissertation

will also serve to establish the frame-theoretic point of view as useful in quantum information theory.

Two specialized types of frames are also examined in chapter 2: spherical codes and spherical t-designs. These are simply sets of vectors arranged in an appealing fashion, and have typically found application in classical coding theory (the former) and numerical estimation theory (the latter). We shall import them into a quantum-mechanical setting in chapter four, and explicitly employ the spherical codes for cryptographic ends in chapters five and six.

But first, the frame theory perspective is put to the test in chapter three, wherein the quantum probability rule is shown to follow quite simply from the structure of generalized measurements. Ostensibly this issue pertains to the foundations of quantum mechanics, not to applied matters. However, such foundational questions do bear directly on practical questions, for the very notion of possible measurements is at stake. Gleason's theorem [68] establishes that the probabilistic structure of quantum theory, the Born probability rule, follows from the structure of ordinary *projective* measurements, which were already imagined by von Neumann to be the fundamental elements of the theory [153]. In this formulation, the standard for quantum mechanics textbooks, a possible quantum measurement corresponds to a set of orthogonal projection operators Π_k such that $\Pi_j \Pi_k = 0$ for $j \neq k$ and $\sum_k \Pi_k = I$. Gleason's theorem then determines the probabilistic nature of quantum mechanics by showing that if a measurement is made, the outcomes must occur with probability given by the rule $p_k = \mathrm{Tr}[\Pi_k \rho]$, where the *density operator* ρ is a positive, trace-one operator on the same space as the Π_k. Generally there aren't any dispersion-free density operators causing all probabilities to be either one or zero, so given the nature of measurement, quantum theory is inescapably probabilistic.

Generalized measurements, POVMs, drop the orthogonality requirement and relax the restriction to projection operators. These were seen as an afterthought, a trick that could be performed by writ from the church of the larger Hilbert space. However, by necessity Gleason left out our favorite quantum system, the two-level qubit,

and the focus on projective measurements obscured the very *useful* role POVMs have to play. Statements like "position and momentum cannot be simultaneously measured" are patently absurd; it's just they cannot be *projectively* measured together.[3] By considering generalized measurements from the outset, it becomes very simple to follow the trail of implications to the Born rule, as well as net the outstanding qubit case. This chapter represents original work reported in [33], which was later discovered to have been proven in [24].

This approach cements the importance of POVMs, which in addition to being useful, provide an elegant, streamlined, and coherent foundation for quantum theory. In some sense this only serves to invite the question of why measurements should be considered fundamental in the first place, and how their structure ought to follow from more basic principles. This question has been with us since von Neumann, of course, but with the rest of the theory more or less in place, has taken on ever-more urgency. The quantum logic program, started by von Neumann and Birkhoff [16] and continued by Mackey [111, 112], Jauch [84], Piron [124], and many others, sets itself the task of solving this problem. Today quantum logic has evolved to the study of *effect algebras*, the generalization of POVMs (whose elements are termed effects) in hopes of providing a satisfactory answer to the second half of the aforementioned question. The first half has, in one form or another, raged since antiquity. A flavor of the argument is given in the beginning of chapter three, but suffice it to say here that positing measurements as fundamental seems the best way to reconcile the varying, if vague, notions of probability, measurement, physical properties, and the observer's role in the whole affair.

Chapter three turns to a more concrete application of frame theory: the search for a suitable frame with which to effortlessly reduce finite-dimensional quantum mechanics to more familiar, probabilistic terms. Known by the unwieldy name "symmetric, informationally-complete positive operator-valued measure", SICPOVM for

[3]Naturally the limiting accuracy of such a joint measurement is set by the uncertainty principle.

short, this aesthetically appealing set of operators offers elegant and efficient distributions and quasidistributions akin to the Glauber-Sudarshan P and Husimi Q functions of infinite dimensions. The SICPOVM appears in other contexts under various names, because more than an ordinary frame, it is simultaneously an equiangular spherical code and a spherical 2-design. Thus one of the earliest investigations gave it title "equiangular lines" [104], work continued by many others under various names [50, 98, 77, 97, 147, 99, 60, 61, 169].

Because these POVMs induce probability distributions which are faithful representations of any density operator or indeed any operator they earn the title "informationally-complete" [125, 22, 137, 46], but the elegance begins with "symmetric". This makes them simple to define, too: a set of $n = d^2$ normalized vectors $|\phi_k\rangle$ in $\mathbb{C}^d$ is equivalent to a SICPOVM when it satisfies

$$|\langle\phi_j|\phi_k\rangle|^2 = \frac{1}{d+1}\,, \qquad \forall\, j \neq k\,. \tag{1.3}$$

The actual SICPOVM comes about by subnormalizing these vectors so that their outer products sum to the identity.

Sadly, elegance does not directly translate into provably existent, and the first half of chapter 4 is devoted to examining the structure in detail before providing analytical examples in dimensions two, three, and four, and numerical solutions up to dimension 45. This work represents collaboration with Andrew Scott, Kiran Manne, and Robin Blume-Kohout [129]. Convinced of their existence, the second half of the chapter elucidates what they can do. Just like their infinite-dimension cousin, there exists a "P" and a "Q" function based on the SICPOVM, with the major advantage that transforming from one to the other is trivial. Given any informationally-complete measurement, P and Q functions can always be found; the Q function is just the probability distribution of the measurement, while the P function arises when writing the density operator in terms of the measurement elements, like $\rho = \sum_k P_k(\rho)E_k$.

Typically the connection between the two representations will involve matrix inversion, and may be singular in places. However, for the SICPOVM Ps and Qs, we'll

find that $P_k(\rho) = d(d+1)Q_k(\rho) - 1$, so effectively the P and Q are the same up to an overall scale and translation. This makes the SICPOVM appealing as a sort of "standard quantum measurement", a future measurement to which any given state could be subjected. With this, the whole apparatus of quantum theory may be transcribed into ordinary probability theory by defining measurements and dynamical operators in terms of their effect on the to be performed standard measurement. What were measurement operators will now be linear functionals of distributions, and dynamical operators stochastic maps from distributions to distributions. In this new guise the content of quantum theory remains, limiting the set of allowed distributions, functionals, and stochastic maps.

Such a prescription applies immediately to tomography, the reconstruction of quantum states from a large number of measurements. Suppose a physical system reliably produces a particular quantum state whose identity we would like to verify. Perhaps we have constructed a device to reliably prepare a spin-1/2 system along the z axis for use in a quantum magnetometer. Before putting it into production we should first verify that this is indeed the state prepared. In principle we have access to an unlimited supply of states, and by using the SICPOVM to measure each one, we build up a statistical estimate of the Q function. This can be translated back into a density operator via the P function if we like, but this step is no longer necessary if we are content to stay in this "SICPOVM representation."

Tomography applies to the other pieces of quantum mechanics as well: measurements and dynamics. Faced instead with an unknown measurement apparatus, one simply turns the preceding procedure inside out, feeding in different states corresponding to elements of the SICPOVM one at a time until enough statistics have been collected to determine the form of the measurement operators. In the service of determining the dynamics of a quantum system, the procedure is termed quantum process tomography. To accomplish this task, we simply sandwich the two preceding procedures together, feeding in SICPOVM quantum states and using the SICPOVM measurement at the other end.

Using the SICPOVM for tomography is simple, but not particularly efficient, and in practice would not be attempted unless no background information about the item in question were available. But it does provide a conceptually clear way to think about tomography and an elegant means of representing the objects in quantum theory.

Chapters two and three of part two represent applications of frame theory to finding mathematical tools useful in quantum information theory. Part three considers frame theory in the context of quantum cryptography, the most practical application of quantum information theory thus far. Specifically, chapter four adapts equiangular spherical codes for use in key distribution protocols, and chapter five examines the experimental prospects for actual implementation.

Cryptography is a huge subject in and of itself, but a brief self-contained introduction is given in the beginning of chapter four before considering what quantum mechanics has to offer: security guaranteed by (currently-understood) physical law not computational difficulty. The goal of cryptography is to encode data in such a way that only trusted parties can read it. This is achieved by using a shared key which "locks" and "unlocks" the message, preventing prying eyes from learning it. The sender uses the key to encode the message at which point it may be securely broadcast, since only the intended recipient, i.e. another person with the correct key, can distinguish it from noise. Closely-related, then, is the problem of secure key distribution. If the sender and receiver do not share a key and are separated by a large distance preventing them from meeting in private, they cannot even begin the secure protocol without first establishing the key. This would seem to imply a catch-22, since because the key needs to be secret, how will they agree on it without resorting to a secure means of communication? The trick is to realize that establishing keys securely is not like transmitting messages securely; the key is not data possessed by the sender which the receiver would like to have. Rather, it is simply a random string shared between the two parties and unknown to anyone else. Hence it can be *created* at each end, rather than being *transported* from sender to receiver.

With a working key distribution protocol, the sender and receiver can first establish the key and then use it to safely encrypt and transmit the actual message. Typical current protocols for both data encryption and secure key exchange between remote parties are based on the current computational difficulty of certain mathematical problems, such as factoring large numbers. These schemes would all be broken should efficient algorithms be found, and in principle the new quantum factoring algorithm and others do exactly this.

But what quantum mechanics takes by destroying the security of current cryptographic protocols, it gives in the form of new, *unconditionally* secure methods. Using systems described by quantum mechanics, protocols can be created which are secure in as much as quantum mechanics is correct. That is to say, the security of such protocols issues from physical law, not the current computational state of the art. As already mentioned, this behavior relies on the curious feature of quantum systems that they cannot be measured without being disturbed. In a different guise, this is the trusty old quantum measurement problem that from one point of view the state of the system instantaneously "collapses" when it is measured. Such is life in the quantum world, and by holding our philosophical objections just long enough to consider the consequences, we can make use of this discomforting state of affairs.

Formally, this feature manifests itself when considering the post-measurement state. Consider a simple, rank-one projective measurement for the moment, whose elements are $\Pi_k = |e_k\rangle\langle e_k|$. Students of introductory quantum courses know that if a quantum state $|\psi\rangle$ is measured by $\{\Pi_k\}$, then the kth outcome will obtain with probability $p_k = \langle\psi|\Pi_k|\psi\rangle$. Moreover, the quantum state of the system must now by given by $|e_k\rangle$, for an immediate subsequent measurement must certainly yield the same result.

In the context of communication, this translates into a tradeoff between the amount of information which can be gathered by a measurement and the disturbance caused in the process. Suppose the sender encodes a message from a given set into a quantum system, and transmits it to the receiver. Ideally, the receiver would like to

distinguish all the possible messages and so arranges a measurement in which each outcome corresponds to a potential message. Calling the message set $\{\pi_j, |\phi_j\rangle\}$, in which the jth message is sent with probability π_j, and the measurement $\{E_k\}$, the overall probability of error when using a noiseless channel becomes

$$p_{\text{err}} = \sum_{j \neq k} \pi_j \langle \phi_j | E_k | \phi_j \rangle = 1 - \sum_k \pi_k \langle \phi_k | E_k | \phi_k \rangle \tag{1.4}$$

Should the encoded messages be non-orthogonal, as for instance would be the case if the number of possible messages is larger than the dimension of the quantum system, then the receiver cannot read the message without on average causing some disturbance to the system. Sometimes an incorrect outcome will obtain, and the state will collapse to this incorrect message state. Conversely, the only sure way to extract the message without disturbing the state is for the messages to be encoded into a set of orthogonal quantum states, for then the corresponding measurement yields the message sent. The cross-terms for which $j \neq k$ in equation 1.4 can be made zero by choosing $E_k = |\phi_k\rangle\langle\phi_k|$. But this is effectively the classical case of perfectly distinguishable messages, so any quantum departure from it invites the information/disturbance tradeoff.

Now suppose an eavesdropper intercepts the message before the message signal reaches the receiver. If in the attempt to decode the message disturbance is introduced, the probability of correctly decoding the message will decrease. For instance, if the eavesdropper also uses the measurement $\{E_k\}$ and forwards the state $|\phi_k\rangle$ upon obtaining outcome k, the probability of error becomes

$$p_{\text{err}} = \sum_{j \neq k,l} \pi_j \langle \phi_j | E_l | \phi_j \rangle \langle \phi_l | E_k | \phi_l \rangle = 1 - \sum_{k,l} \pi_k \langle \phi_k | E_l | \phi_k \rangle \langle \phi_l | E_k | \phi_l \rangle \tag{1.5}$$

The legitimate parties may utilize this feature to check for the presence of an eavesdropper. By publicly comparing a subset of the messages sent with those received, using a classical broadcast channel, an increased error rate over that indicated by equation 1.4 will alert them to interference with the quantum channel.

This ability of quantum states to record attempts at copying fits perfectly with the goal of key distribution. By using quantum states, the sender and receiver can

create a shared key from strings of single codewords, and then check the error rate to determine if it is truly secret. Should the error rate be too high to ensure security, the key creation attempt may be abandoned. When an error rate exists below which the sender and receiver can be guaranteed that illegitimate parties *cannot* know enough of the key to be in any way useful, the protocol exhibits unconditional security.

Equiangular spherical codes (ESCs) are found to be particularly well-suited to the task of secure key creation. These are sets of n states in $\mathbb{C}^d$ with equal overlap arranged to be as widely-spaced in the vector space as possible, conforming to the requirement

$$|\langle\phi_j|\phi_k\rangle|^2 = \frac{n-d}{d(n-1)} \tag{1.6}$$

Such equiangular sets with this overlap only exist when $d \leq n \leq d^2$, so the SICPOVM is the largest possible ESC. In general, key distribution protocols exist with arbitrary n between these limits, and we shall see that the wide spacing translates well into eavesdropper sensitivity.

Establishing the unconditional security of a key distribution protocol is a subtle art, doubly so due to the use of quantum mechanics. In chapter five the beginning of such an analysis is presented, gradually approached by first studying the relevant issues and complications in the context of the original proposal for quantum cryptography, put forth by Bennett and Brassard in 1984 [12]. In this scheme, two sets of polarization states, horizontal/vertical (+) and 45/135 degree diagonal ($\times$) are used to encode key bits. Both horizontal and 45 degree states may be taken as `0`. for instance, and the others `1`. The receiver randomly chooses one of the two sets in which to measure the incoming signal, and when that choice matches the encoding a key bit is created. The other cases are simply discarded. The protocol is secure since through the use of non-orthogonal signal states the legitimate users can infer the eavesdropper's disturbance via the observed error rate. From this they may infer the eavesdropper's knowledge and discard the key if it is too large. In general, however, the eavesdropper may perform quantum manipulations to the signals and also store

whatever results in quantum form. The real difficulty of establishing security lay in quantifying the eavesdropper's information when encoded in quantum states.

Instead of attempting a full analysis straight away, the equiangular spherical code states are assumed to be subjected to the more limited eavesdropping method described above by equation 1.6. This "intercept/resend" attack has both the advantage and drawback of being the simplest to consider. It immediately provides a comparison between ESC protocols and the original scheme, as well as its variants, but cannot make concrete statements in the most general setting.

In this setting spherical codes are found to offer several advantages over the more standard protocols which use collections of orthogonal bases. Like the BB84 protocol which uses the two bases + and $\times$, in arbitrary dimensions many different bases can be used, giving rise to a plethora of protocols. This is similar to the choice of n for spherical code protocols, but there are simply more spherical code sets, so this allows ESC protocols to offer a wider range of possibilities in terms of key generation speed and noise tolerance. Faster protocols can be used when noise is low; safer protocols when noise is high. Second, for a given number of signal states in a fixed dimension, arranging them into a spherical code provides more noise tolerance, but a lower key generation rate than using bases. Finally, ESC protocols automatically estimate the channel noise, removing the need for the legitimate parties to do this manually. The most practical version utilizes two-level systems, like polarization, for which the details are examined separately. Here there are two possible equiangular spherical codes, called the trine and tetrahedron after their shapes in the Bloch-sphere representation. The trine consists of three coplanar states spaced by 120 degrees, and the tetrahedron the four vertices of that regular Platonic solid. For these two protocols slightly stronger eavesdropping methods can be formulated and analyzed, again demonstrating the increased security of equally-spaced message states.

Qubit-based protocols enjoy the advantage of easy implementation using polarization states of light. With polarizers and waveplates arbitrary superpositions of the

two orthogonal states are simple to prepare at the transmission end, while polarizing beamsplitters in conjunction with phase shifters and ordinary beamsplitters suffice to implement any potential measurement at the receiving end. However, faster protocols tolerating more eavesdropper interference are to be found among those having more states in higher dimensions, so chapter six examines the prospects for using transverse spatial modes of light along with passive linear-optical elements to encode and decode states of higher-dimensional quantum systems.

Experiments creating, controlling, and measuring transverse modes, particularly those carrying so-called "orbital" angular momentum[4] have recently been performed, suggesting that cryptographic protocols based on them are not far off. These modes carry angular momentum in their spiral-shaped wavefront (surfaces of constant phase resembling fusili) and may be created from ordinary Gaussian-profile laser beams with simple blazed phase holograms. Superpositions of such states, too, result from the use of holograms, or may be created by interferometry. At the decoding end, the angular momentum analog of a polarizing beamsplitter may be realized as a sequence of interferometers employing phase shifters and Dove prisms to effect transverse mode rotations. Like the polarizing beamsplitter, this optical network essentially transcribes the quantum state from a superposition of spatial mode states to one involving distinct propagation channels. These modes may be manipulated again using linear optical elements to implement any desired quantum measurement.

Orbital angular momentum modes are not the only available which are suitably influenced by linear optics, but the use of linear optics is a must for quantum cryptography. Although quantum states may not be blindly copied, the sender should take care not to send many photons per pulse, since each is effectively a copy of the signal. When communicating over a lossy channel, an eavesdropper could in principle hijack the channel, replacing the loss mechanism with a means of splitting off some of the photons from each signal, with the legitimate parties none the wiser. Though not very realistic, it is possible, so to ensure unconditional security, this case must

[4]as opposed to "spin" angular momentum carried by polarization.

be avoided or dealt with. This requirement mandates a sparing number of photons per pulse and requisite measurement devices capable of handling such faint signals. Though daunting, this is just feasible. Already qubit-based quantum cryptography has been demonstrated experimentally, both in free-space and in optical fibers, and commercial applications are even available.

Chapter seven gives a summary of the work presented herein before turning to the consideration of what future work should be undertaken. For the SICPOVM, clearly one would like to analytically establish its existence. The *prima facie* simplicity of doing so, contrasted with the actual difficulty points to the possibility of a fruitful and enlightening proof, work on which is indeed underway [156]. Other applications of the SICPOVM abound, from the efficiency of state reconstruction to the characterization of ensembles of quantum states as the most "non-classical" [65]. The unconditional security of ESC-based protocols also presents itself as an obvious choice for completion as soon as possible, but along with this comes the bigger and perhaps more interesting question of the *optimal* ensemble for cryptography, in the sense of providing the most security. Heuristic arguments can be given as to why equiangular spherical codes ought to be the optimal ensembles, and numerical results would immediately indicate if this is likely the be the case. Spherical code protocols may also help in increasing the number of photons which can be safely sent per pulse, and the resulting intensity increase has useful practical implications.

These topics and their order reflect my own progression through the field of quantum information theory, shaped both by interest and coincidence. Being one part practical and one part foundational, the dichotomy inherent to the field at once offers a wealth of diverse and interesting research questions and provides a certain "intellectual-tension" conducive to solving them. When progress on one of these questions slows, the opposite perspective may be adopted for a fresh look. New lines of research can be developed in conjunction with scientists in seemingly-distant fields. That the thesis delivers this material organized around the topic of frame theory reflects the fact that much of the research carried out could also justifiably be

considered as part of that field. All of this contributes to the sense that intellectual curiosity can perhaps never be satisfied, but is certainly never bored.

A thesis can never contain solely original work; the work must always be placed in the appropriate context and important background information given. To be clear about what is original, the following table spells it out by section.

Section	2.3	3.1–3	4.1–3	5.4–5	6.4
Collaboration	Ref. [129]	Ref. [33]	Ref. [129]	—	—

Chapter 2

Frame Theory Basics

From a purely mathematical point of view, frames, spherical designs, and spherical codes are simply appealing sets of vectors in a linear space. Each rightly commands its own field of study, often closely related to the main application, classical coding theory. Perhaps unsurprisingly, these structures can be applied to quantum information theory as well, and they underlie many of the results presented in this work. In this chapter, the aim is to develop only the main features of each mathematical object, paying special attention to the aspects that will be relevant for use later in the setting of quantum information theory.

Frames are the most general of the three, encompassing the other two, so we begin with them. Simply put, frames are generalizations of the familiar orthonormal bases of linear spaces, with the requirements of orthogonality, normalization, and even linear independence relaxed. Absent these restrictions a frame is an arbitrary basis that represents arbitrary vectors with a (typically) redundant and non-unique set of coefficients. The redundancy of the frame representation is advantageous in questions of signal processing, where frames were first introduced by Gabor in 1946 [66] as a means of signal decomposition into elementary signals localized in time and frequency. Drawing on this work, in 1952 Duffin and Schaeffer [55] introduced frames in a Hilbert space for use in the study of nonharmonic Fourier series. The

field lay dormant, however, until the work of Daubechies, Grossmann, and Meyer in 1986 [47]; currently frame theory is a quite active and rapidly growing field [29].

2.1 Simple Frames

For a finite-dimensional vector space $\mathcal{H}$ of dimension d, a collection of n vectors $|\phi_k\rangle \in \mathcal{H}$ is a frame if there exist constants $0 < a \leq b < \infty$ such that

$$a\langle\xi|\xi\rangle \leq \sum_k |\langle\phi_k|\xi\rangle|^2 \leq b\,\langle\xi|\xi\rangle \tag{2.1}$$

for all $|\xi\rangle \in \mathcal{H}$. The constants a and b are called the *frame bounds*. The lower bound ensures that a frame spans the space, whence $n \geq d$. The upper bound is satisfied for $n < \infty$ by using the Cauchy-Schwartz inequality and setting $b = \sum_{k=1}^{n}\langle\phi_k|\phi_k\rangle$. Thus any finite collection of vectors is a frame for its span. If $a{=}b$, the frame is said to be *tight*, if $a = b = 1$ *normalized*, and if upon deletion of an element the collection ceases to be a frame, it is said to be *exact*.

We may work with frames more easily by formulating them in terms of operators. The *analysis operator*, or *frame transform*, $T : \mathcal{H} \to \ell_2$ decomposes any vector into the sequence of overlaps with frame elements: $T|\psi\rangle = \{\langle\phi_k|\psi\rangle\}$. The adjoint is the *synthesis operator*, or *preframe operator*, $T^\dagger : \ell_2 \to \mathcal{H}$ which creates a vector from a sequence: $T^\dagger\{a_k\} = \sum_k a_k|\phi_k\rangle$. [1] The *frame operator* is then the positive operator $S : \mathcal{H} \to \mathcal{H}$ such that $S = T^\dagger T$, i.e.

$$S{=}\sum_k |\phi_k\rangle\langle\phi_k| \ . \tag{2.2}$$

Putting the analysis and synthesis operators in the other order $TT^\dagger$ yields the Gram matrix $G_{j,k} = \langle\phi_j|\phi_k\rangle$ of the frame elements. In terms of the frame operator, equation 2.1 now reads $aI \leq S \leq bI$ where I is the identity matrix.

[1] Note that in mathematical literature the analysis and synthesis operators are defined in the opposite sense, stemming from the use of the inner product in the conjugate sense.

In order to express a vector in terms of a frame, one must determine the coefficients, and in general this expansion is not unique. However, the frame operator immediately yields an expansion in terms of the *canonical dual*. The elements of the canonical dual are given by $|\widetilde{\phi}_k\rangle = S^{-1}|\phi_k\rangle$; its frame operator is obviously S^{-1}, and frame bounds $1/b, 1/a$. The reconstruction formula is given by

$$|\psi\rangle = SS^{-1}|\psi\rangle = \sum_k |\phi_k\rangle\langle\widetilde{\phi}_k|\psi\rangle. \tag{2.3}$$

This method can be difficult because it requires a matrix inversion, though of all the possible expansion coefficients, those found by using the canonical dual have the smallest possible ℓ_2 norm, as the following theorem shows.

Theorem 1 (Duffin and Schaeffer [55]) *Let $\{|\phi_k\rangle\}$ be a frame for a Hilbert space $\mathcal{H}$ and $|\psi\rangle \in \mathcal{H}$. If $\{b_n\}$ is any sequence of scalars satisfying $|\psi\rangle = \sum_k b_k|\phi_k\rangle$ then*

$$\sum_k |b_k|^2 = \sum_k |\langle\widetilde{\phi}_k|\psi\rangle|^2 + \sum_k |\langle\widetilde{\phi}_k|\psi\rangle - b_k|^2. \tag{2.4}$$

The proof is quite simple. Starting from

$$|\psi\rangle = \sum_k b_k|\phi_k\rangle = \sum_k |\phi_k\rangle\langle\widetilde{\phi}_k|\psi\rangle, \tag{2.5}$$

take the inner product with $S^{-1}|\psi\rangle$ to obtain

$$\sum_k b_k\langle\psi|\widetilde{\phi}_k\rangle = \sum_k |\langle\widetilde{\phi}_k|\psi\rangle|^2, \tag{2.6}$$

from which equation 2.4 follows. □

Tight frames, meanwhile, make such expansions simple since $S = aI$. Associated to any frame is the *canonical tight frame*, generated by using the square root of the inverse of the frame operator: $|\overline{\phi}_k\rangle = S^{-1/2}|\phi_k\rangle$. The canonical tight frame is the normalized tight frame closest to the original in the sense of minimizing the squared error [58, 59]. We can establish this by a judicious use of the singular value decomposition. First note that one may think of the columns of the matrix $T^\dagger$ as the expansion coefficients of the $|\phi_k\rangle$ in an orthonormal basis $|e_k\rangle$. Suppose that $|\overline{\phi}_k\rangle$

is any normalized tight frame with analysis operator F. Then the error quantity we seek to minimize is given by

$$\sum_k \| |\phi_k\rangle - |\overline{\phi}_k\rangle \|^2 = \mathrm{Tr}[(T-F)(T-F)^\dagger], \tag{2.7}$$

subject to the condition that $F^\dagger F = I$. Consider the singular value decomposition $T^\dagger = U\Sigma V^\dagger$ where U, V are unitary operators and Σ is diagonal. We may expand the trace in the orthonormal basis of U, denoted $\{|u_k\rangle\}$, and define $|a_k\rangle = F|u_k\rangle$, an orthonormal basis since $\langle a_j|a_k\rangle = \langle u_j|F^\dagger F|u_k\rangle = \delta_{jk}$. Note that the singular value decomposition implies $T|u_k\rangle = \sigma_k|v_k\rangle$, where $\{|v_k\rangle\}$ is the orthonormal basis associated with V and σ_k the associated singular value. Thus the error may be written as $\sum_k \|\sigma_k|v_k\rangle - |a_k\rangle\|^2$ with the minimization over the orthogonal $|a_k\rangle$. The solution, of course, is $|a_k\rangle = |v_k\rangle$, whence $F|u_k\rangle = |v_k\rangle$. Hence $F^\dagger = (T^\dagger T)^{-1/2}T^\dagger$, or $|\overline{\phi}_k\rangle = S^{-1/2}|\phi_k\rangle$.

In the special case that the frame elements are equal in number to the dimension, the canonical tight frame will be an orthonormal basis, simply because in order for d rank-one projectors to sum to the identity operator on d dimensions, they must be mutually orthogonal. This in turn implies that the frame and its canonical dual form a biorthogonal system:

$$\delta_{jk} = \langle\overline{\phi}_j|\overline{\phi}_k\rangle = \langle\phi_j|S^{-1}|\phi_k\rangle = \langle\phi_j|\widetilde{\phi}_k\rangle. \tag{2.8}$$

Note that the synthesis operator takes an orthonormal basis of ℓ_2 to the frame in $\mathcal{H}$, though for a finite-element frame in finite dimensions, one needs only a finite subspace of ℓ_2 [11]. We may go in reverse and *dilate* a frame to an orthonormal basis in a higher dimensional space as well. A result of Casazza, Han, and Larson shows that this is equivalent to the notion of a frame.

Theorem 2 (Casazza, Han, and Larson [30]) *A set $\{|\phi_k\rangle\}$ is a frame for a Hilbert space $\mathcal{H}$ if and only if there exists a Hilbert space $\mathcal{K} \supset \mathcal{H}$ with an orthonormal basis $\{|e_k\rangle\}$ and a (not necessarily orthogonal) projection $P : \mathcal{K} \to \mathcal{H}$ such that*

$P|e_k\rangle = |\phi_k\rangle$ for all k. Further, $\{|\phi_k\rangle\}$ is a normalized tight frame if and only if P is an orthogonal projection.

The portion of the theorem pertaining to tight frames is simple to prove. Beginning with the reverse implication, let $\{|e_k\rangle\}$ be an orthogonal basis for $\mathcal{K}$ and P an orthogonal projection. Then for any $|\psi\rangle \in P(\mathcal{K}) = \mathcal{H}$,

$$\sum_k |\langle\psi|P|e_k\rangle|^2 = \sum_k |\langle e_k|P|\psi\rangle|^2 = \sum_k |\langle e_k|\psi\rangle|^2 = \langle\psi|\psi\rangle, \tag{2.9}$$

so $\{P|e_k\rangle = |\phi_k\rangle\}$ is a normalized tight frame. Conversely, whenever $\{|\phi_k\rangle\}$ is a normalized tight frame, $\||T|\psi\rangle\|^2 = \langle\psi|T^\dagger T|\psi\rangle = \langle\psi|\psi\rangle$, so that the analysis operator is an into isometry. Thus we may associate $T(\mathcal{H}) \subset \ell_2$ with $\mathcal{H}$ itself and set P as the projection from ℓ_2 to $T(\mathcal{H})$. Let $\{|e_k\rangle \in \ell_2\}$ be an orthonormal basis such that $T^\dagger|e_k\rangle = |\phi_k\rangle \in \mathcal{H}$. Then

$$\langle T\psi|Pe_k\rangle = \langle PT\psi|e_k\rangle = \langle T\psi|e_k\rangle = \langle\psi|T^\dagger e_k\rangle = \langle\psi|\phi_k\rangle = \langle T\psi|T\phi_k\rangle, \tag{2.10}$$

for all k and $|\psi\rangle \in \mathcal{H}$, and therefore any normalized tight frame $\{|\phi_k\rangle\} \approx \{P|e_k\rangle\}$. Of course, all other tight frames can be obtained by an appropriate scaling of the projection operator P. □

In the language of quantum mechanics, normalized tight frames are POVMs consisting of rank-one elements, as they are positive operators which decompose the unit operator. Conversely, since any POVM with arbitrary rank elements can be decomposed into rank-one elements, all POVMs can be thought of as normalized tight frames. In this context, theorem 2 is equivalent to Neumark's theorem which asserts that any POVM can be realized as an orthogonal projection-valued measure in a higher dimensional space [122].

Now specialize to the case $\mathcal{H} = \mathbb{C}^d$ and let $\mathbb{S}^d \subset \mathbb{C}^d$ be the subset consisting of vectors that have unit norm. Any frame can be rewritten in terms of the corresponding normalized vectors, but tightness is not preserved under this transformation. For a frame $\{|\phi_k\rangle \in \mathbb{S}^d\}_{k=1}^n$ made up of normalized vectors, the quantity

$$\mathrm{Tr}[S^2] = \sum_{j,k} |\langle\phi_j|\phi_k\rangle|^2 \tag{2.11}$$

is called the *frame potential*. Throughout the following, only frames made up of normalized vectors are considered, although suitable frame potentials can be defined for frames consisting of vectors of arbitrary norm [155]. A useful theorem due to Benedetto and Fickus states the following.

Theorem 3 (Benedetto and Fickus [11]) *Given any d and n, let $\{|\phi_k\rangle \in \mathbb{S}^d\}_{k=1}^n$ be a set of normalized vectors with frame operator S. Then*

$$\mathrm{Tr}[S^2] \geq \max(n, n^2/d) \; . \tag{2.12}$$

Furthermore, the bound is achieved if and only if $\{|\phi_k\rangle\}$ consists of orthonormal vectors, when $n \leq d$, or is a tight frame, when $n \geq d$.

Proof Denoting the ordered eigenvalues of S by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d$, we first note that the number of nonzero eigenvalues is at most $q = \min\{n, d\}$. Thus we have

$$\mathrm{Tr}[S] = n = \sum_{k=1}^{q} \lambda_k \qquad \text{and} \qquad \mathrm{Tr}[S^2] = \sum_{k=1}^{q} \lambda_k^2 \; . \tag{2.13}$$

Minimizing $\mathrm{Tr}[S^2]$ subject to the constraint $\mathrm{Tr}[S] = n$ gives the inequality. Equality holds if and only if $\lambda_k = n/q$, $k = 1, \ldots, q$, whence $S = (n/q)\Pi_q$ where Π_q is the projector onto the subspace corresponding to nonzero eigenvalues. Thus for $n \leq d$, S is a projector onto an n-dimensional subspace, implying that the vectors $|\phi_k\rangle$ are orthogonal, and if $n \geq d$, $S = (n/d)I$, implying that the set $\{|\phi_k\rangle\}$ is a tight frame. □

2.2 Spherical Codes

Equation 2.12 is called the Welch bound from classical signal processing [160], and frames meeting this bound are also called Welch bound equality sequences. These sequences find heavy application in code-division multiple-access (CDMA) systems,

such as personal wireless communication [147]. Welch established equation 2.12 by proving that

$$\max_{i\neq j} |\langle\phi_i|\phi_j\rangle|^2 \geq \min\{0, \frac{n-d}{d(n-1)}\}. \tag{2.14}$$

Working backwards, this follows from theorem 3 by replacing "off-diagonal" terms in the frame potential by the maximum value.

This bound on the maximum overlap leads directly to the concept of spherical codes, sets of unit vectors spread widely throughout a given vector space. For current purposes, a spherical code $\mathcal{S}(d, n, s)$ is a set of n unit vectors in $\mathbb{C}^d$ such that

$$\max_{j\neq k} |\langle\phi_j|\phi_k\rangle|^2 \leq s. \tag{2.15}$$

Finding the *smallest* s for a given number of vectors in a given dimension is called Tammes's problem, after the Dutch botanist who studied the distribution of pores on pollen grains [148]. The complementary task of finding the largest n for a given dimension and maximal overlap s is called the kissing problem [41]. Clearly the smallest s (i.e. the shortest distance on the sphere) is that given by equation 2.14. Spherical codes meeting this bound are called, variously, *equiangular spherical codes* (ESCs) or *optimal Grassmann frames*.

The existence of ESCs isn't known for arbitrary n and d, though some general statements can be made [129]. They always exist for $n = d + 1$ (a regular simplex), but never when $n > d^2$. Equiangular spherical codes with $n = 2d$ are known to exist for dimensions $d = (p^\alpha + 1)/2$ or $d = 2^\alpha$ where p is an odd prime and α is any integer; those having $n = d^2$ elements are believed to exist in any dimension, as will be seen in chapter four. For $n \leq d^2$, when a Grassman frame exists, it is a spherical code, but for $n > d^2$, spherical codes aren't equiangular.

2.3 Spherical Designs

Instead of looking for interesting arrangements of vectors by minimizing the maximal overlap of a set of vectors, one might simply minimize the frame potential instead.

The form of the frame potential itself suggests a generalization to higher orders, so that we can define the tth-order frame potential as

$$V_t = \sum_{j,k} |\langle \phi_j | \phi_k \rangle|^{2t}. \tag{2.16}$$

A *spherical t-design*, a set of unit vectors useful in numerical integration of polynomial functions on the sphere, always results from minimization of the t-th-order frame potential. Specifically, a spherical t-design is a set of n normalized vectors $\{|\phi_k\rangle \in \mathbb{S}^d\}$ such that the average value of any t-th order polynomial $f_t(\psi)$ over the set $\{|\phi_k\rangle\}$ is equal to the average of $f_t(\psi)$ over *all* normalized vectors $|\psi\rangle$. Note that if a set is a t-design, it is also an s-design for all $s \leq t$, since an s-th order polynomial is also a t-th order polynomial. Spherical t-designs were originally developed as subsets of the real sphere S^d; here the concept is applied to the set $\mathbb{S}^d$. To make the connection between the averaging property and the frame potential minimization, we begin with the definition of polynomial functions on $\mathbb{S}^d$.

Let $\mathcal{H} = \mathbb{C}^d$, $\mathcal{H}_t$ be the t-fold tensor product of such spaces, and $\mathcal{S}_t$ be the symmetric subspace of $\mathcal{H}_t$, and consider a function $f_t : \mathcal{H} \to \mathbb{C}$ defined as

$$f_t(\psi) = \langle \Psi^t | F_t | \Psi^t \rangle \ , \qquad |\Psi^t\rangle = |\psi\rangle^{\otimes t} \ , \qquad |\psi\rangle \in \mathcal{H} \ , \tag{2.17}$$

where the choice of f_t is equivalent to a choice of a symmetric operator $F_t \in \mathcal{B}(\mathcal{S}_t)$. Such a function is a t-th order polynomial function on $\mathcal{H}$. We can decompose F_t into a sum of product operators, i.e., $F_t = \sum_k \bigotimes_{j=1}^t A_{j;k}$; thus any such function can be decomposed into monomial terms like

$$\left\langle \Psi^t \middle| \bigotimes_{j=1}^t A_j \middle| \Psi^t \right\rangle = \prod_{j=1}^t \langle \psi | A_j | \psi \rangle \ . \tag{2.18}$$

Without loss of generality, we can restrict our attention to such monomial functions and rewrite them as

$$f_t(\psi) = \prod_{j=1}^t \mathrm{Tr}\Big[A_j |\psi\rangle\langle\psi| \Big] = \mathrm{Tr}\left[\left(\bigotimes_{j=1}^t A_j \right) \Pi_\psi^{\otimes t} \right] , \qquad \Pi_\psi = |\psi\rangle\langle\psi| \ . \tag{2.19}$$

Since the set $\{|\phi_k\rangle\}$ is a t-design if and only if the average of any f_t over $\{|\phi_k\rangle\}$ is equal to its average over all $|\psi\rangle \in \mathbb{S}^d$, we are led to compute the average of an arbitrary monomial term:

$$\langle f_t \rangle = \int \mathrm{d}\psi \, \mathrm{Tr}\left[\left(\bigotimes_{j=1}^{t} A_j\right)\Pi_\psi^{\otimes t}\right] = \mathrm{Tr}\left[\left(\bigotimes_{j=1}^{t} A_j\right)\int \mathrm{d}\psi \, \Pi_\psi^{\otimes t}\right] = \mathrm{Tr}\left[\left(\bigotimes_{j=1}^{t} A_j\right)K_t\right] . \tag{2.20}$$

Hence we focus on finding K_t, since it effectively takes the average of f_t. A spherical t-design is then a set of vectors for which

$$S_t = \sum_{k=1}^{n} |\Phi_k^t\rangle\langle\Phi_k^t| = nK_t \ , \qquad |\Phi_k^t\rangle = |\phi_k\rangle^{\otimes t} \ . \tag{2.21}$$

Note that S_t is the t-fold tensor-product analog of the frame operator S.

To find the operator K_t, note that K_t has support only on the symmetric subspace $\mathcal{S}_t$. Further, because K_t is invariant under any $U^{\otimes t}$ for $U \in SU(d)$, we conclude that $K_t \propto \Pi_{\mathrm{sym}}$, the projector onto $\mathcal{S}_t$. (Recall that $\mathcal{S}_t$ is an irreducible invariant subspace of the group consisting of the operators $U^{\otimes t}$.) Finally, to determine the constant of proportionality, we consider the average of the trivial function $f_t(\psi)=1$. Equation 2.20 then becomes $\mathrm{Tr}[K_t]=1$, and since $\mathcal{S}_t$ has dimension $\binom{t+d-1}{d-1}$, we have

$$K_t = \frac{t!(d-1)!}{(t+d-1)!}\Pi_{\mathrm{sym}} \ . \tag{2.22}$$

Equation 2.21 now says that the set $\{|\phi_k\rangle\}$ is a t-design if and only if the set $\{|\Phi_k^t\rangle\}$ is a tight frame on $\mathcal{S}_t$, whence we can apply Theorem 3 to obtain the following result.

Theorem 4 *A set of normalized vectors* $\{|\phi_k\rangle \in \mathbb{S}^d\}_{k=1}^n$ *with* $n \geq \binom{t+d-1}{d-1}$ *forms a spherical* t*-design if and only if*

$$\mathrm{Tr}[S_t^2] = \sum_{j,k} |\langle\phi_j|\phi_k\rangle|^{2t} = \frac{n^2 t!\,(d-1)!}{(t+d-1)!} \ . \tag{2.23}$$

Furthermore, this value is the global minimum of $\mathrm{Tr}[S_t^2]$.

This theorem yields a very nice characterization of spherical t-designs; however, it does not establish the minimum number n_t of elements required nor guarantee

existence. The bound on n in the theorem is very loose, arising only from the dimension of $\mathcal{S}_t$, but a general method of determining n_t can be outlined here, taking the case $t = 2$ for which $n_2 = d^2$ (see also [97]). Consider again the steps leading to the definition of the operator K_t. In carrying out the average of the function f_2, we could have written

$$\langle f_2 \rangle = \mathrm{Tr}\Big[A_2 \int \mathrm{d}\psi \, |\psi\rangle\langle\psi|A_1|\psi\rangle\langle\psi|\Big] = \mathrm{Tr}[A_2 \mathcal{G}(A_1)], \tag{2.24}$$

and thus considered the superoperator $\mathcal{G} : \mathcal{B}(\mathbb{C}^d) \rightarrow \mathcal{B}(\mathbb{C}^d)$. Here $\mathcal{G}(UAU^\dagger) = U\mathcal{G}(A)U^\dagger$ for any $U \in SU(d)$, so by Schur's lemma, $\mathcal{G}$ is some linear combination of projectors onto the invariant subspaces of U acting on $\mathcal{B}(\mathbb{C}^d)$. These invariant subspaces are (i) the (d^2-1)-dimensional subspace of traceless operators and (ii) the one-dimensional subspace spanned by the identity operator I. Thus we can write $\mathcal{G} = a\mathcal{I} + b\mathbf{I}$ where $\mathcal{I}(A) = A$ and $\mathbf{I}(A) = \mathrm{Tr}[A]I$ (i.e., $\mathcal{I}$ is the identity superoperator, and $\mathbf{I}$ projects onto the identity operator). To find a and b, we first let $A_1 = A_2 = I$, which gives the function $f_2(\psi) = 1$, so that equation 2.24 yields $d(a + bd) = 1$. Next we consider $A_1 = A_2 = |\phi\rangle\langle\phi|$, for which $\langle f_2 \rangle = \int \mathrm{d}\psi \, |\langle\phi|\psi\rangle|^4 = a + b$. We can use equations 2.20 and 2.22 to show that $\langle f_2 \rangle = 2/d(d+1)$; combined with the previous result, this implies $a = b = 1/d(d+1)$. Therefore, $\mathcal{G}$ has no null subspace, must be rank-d^2, and cannot be constructed from less than d^2 linearly independent rank-one superoperators. Similar arguments can be applied to all spherical t-designs. By similar rearrangements, we can make several different types of operators K'_t, and the rank of each serves as a lower bound on the number of vectors required to comprise a t-design.

Equiangular spherical codes can also be thought of as arising from the minima of the frame potential in the following way. Restricting attention to normalized tight frames (i.e. frames which minimize V_1), consider minimizing V_2. Let $\lambda_{jk} = |\langle\phi_j|\phi_k\rangle|^2$, so that from the minimum of V_1 we have $\sum_{j\neq k} \lambda_{jk} = V_1 - n = n(n-d)/d$. Now $\sum_{j\neq k} \lambda_{jk}^2 = V_2 - n$, whence the minimum of V_2 over all sets minimizing V_1 is bounded below by making all the λ_{jk} the same and given by equation 2.14. When this lower bound is achieved, i.e $V_2 = n^2(n-2d+d^2)/d^2(n-1)$, the result is a Grassman frame.

Note that for n up to d^2 the result of the minimization is an equiangular spherical code, while for n beyond d^2 the result is a spherical 2-design. For $n = d^2$ the result is both; this will be the subject of inquiry in chapter four.

2.4 Weyl-Heisenberg Frames

One particular frame stands out as especially useful in quantum mechanics as well as classical coding theory and signal processing theory: the Weyl-Heisenberg or Gabor frame. So-named because it was the the frame originally introduced by Gabor, it is based on a projective representation of the Weyl-Heisenberg group. This group has elements in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{T}$ which are multiplied according to the rule

$$(q, p, t) \cdot (q', p', t') = (q + q', p + p', t + t' + \frac{1}{2}(q \cdot p' - q' \cdot p)). \tag{2.25}$$

Now we can let the projective representation $\pi(\cdot)$ of this group on $L^2(\mathbb{R})$ be

$$\pi(q, p, t)\phi(x) = E_p T_q \phi(x) = e^{2\pi i p(x-q)}\phi(x - q), \tag{2.26}$$

where T_q is the translation or shift operator and E_p the phase or modulation operator. A Weyl-Heisenberg frame is constructed from these elements by simply picking a suitable "window" function $\phi(x)$, shift/phase parameters q and p and then forming the set $\{\phi_{j,k}(x) = E_p^k T_q^j \phi(x)\}$ for $j, k \in \mathbb{Z}$. Not all functions nor shift/phase parameters can be used to define legitimate frames (though note that in infinite dimensions the requirement that the upper frame bound be finite is dropped) and much research in frame theory is concerned with fully characterizing the set of Gabor frames. In particular, in order that the set be complete it must be that $qp \leq 1$ [48]. For combinations of (ϕ, q, p) which do generate frames, the set is overcomplete when $qp < 1$ but forms a Riesz basis (i.e. a set isomorphic to an orthonormal basis) when $qp = 1$. However, in this latter case a famous result severely restricts the functions which can be used for WH frames.

Theorem 5 (Balian-Low [8, 107, 48]) *If $qp = 1$ and $\phi \in L^2(\mathbb{R})$ generates a WH-frame, then either $x\phi(x) \notin L^2(\mathbb{R})$ or $\phi'(x) \notin L^2(\mathbb{R})$.*

This theorem immediately implies that should $qp = 1$ then functions suitable for making WH frames are either not smooth or not rapidly decaying. Gabor himself used the window function $\phi_0(x) = \pi^{-1/4}e^{-x^2/2}$ so, although choosing $qp = 1$ yields a complete set, it doesn't yield a frame. The difficulty stems from the fact that when $qp = 1$

$$\inf_{\psi \in L^2(\mathbb{R})} \sum_{jk} |\langle \phi_{jk} | \psi \rangle|^2 = 0, \tag{2.27}$$

meaning that the lower frame bound is zero [7]. Strictly speaking, no function is orthogonal to the set, though there are functions arbitrarily close to being so. Hence the eigenvalues of the frame operator are not bounded away from zero, making its inversion generally impossible. This implies that there is no stable way to reconstruct arbitrary functions ψ from the sequence $\{\langle \phi_{jk} | \psi \rangle\}$.

In dealing with simple harmonic oscillator in quantum mechanics the construction of the Weyl-Heisenberg frame as a discrete set of functions is abandoned in favor of allowing *all* possible displacements of the window function g, the ground state wave-function. This makes the set vastly overcomplete, but removes the difficulty associated with the use of Gaussian window functions, and simplifies many expressions by allowing integration instead of summation. More importantly, however, the continuous set of elements provides a phase space description of quantum dynamics. This is the result of geometric quantization.

Generally, this construction begins with a dynamical symmetry group G of the quantum system (a simply connected Lie group), and creates the phase space by considering a unitary irreducible representation U_g acting on $\mathcal{H}$. Fixing a reference state $|\phi_0\rangle$ implicitly defines an isotropy subgroup $H \subset G$, consisting of all U_h such that $U_h|\phi_0\rangle = e^{i\theta(h)}|\phi_0\rangle$ for some $\theta \in \mathbb{T}$. Dividing H out of G yields the coset space $G/H = X$ which is a coarse-graining of the group such that any element g may be written $g = xh$. Since the phase of the state is irrelevant, elements of X define the coherent states via the action of G: $|\phi_x\rangle = U_{xh}|\phi_0\rangle$ for any $h \in H$. By Schur's lemma the coherent states satisfy $\int_X \mathrm{d}\mu |\phi_x\rangle\langle\phi_x| = I$, where μ is the invariant measure on the group G restricted to X, so that arbitrary states may be represented in terms

of the coherent states. If coset space forms a homogeneous Kählerian manifold, i.e. it is locally isomorphic to $\mathbb{C}^n$, then X may be thought of as a phase space. Thus geometric quantization is the map $x \to |\phi_x\rangle\langle\phi_x|$ in this situation. Typically one takes the highest-weight state of the representation U_g for the coherent state $|\phi_0\rangle$, and the ground state of the harmonic oscillator is indeed such a state. In this case we recover from the Weyl-Heisenberg group H_3 the "classical" phase space $H_3/U(1) = \mathbb{C}$. Given the annihilation and creation operators a and $a^\dagger$ (which together with I make up the Lie algebra of H_3), the coherent states are simply $|\alpha\rangle = \exp(\alpha a^\dagger - \alpha^* a)|0\rangle$.

In finite dimensions the situation is much less complicated, since any set of vectors complete in a finite dimensional vector space forms a frame. Here, the relevant representations of the Weyl-Heisenberg group are also representations of $\mathbb{Z}_d \times \mathbb{Z}_d$, generated by the operators

$$V = \sum_k |k \oplus 1\rangle\langle k| \qquad\qquad U = \sum_k \omega^k |k\rangle\langle k|, \tag{2.28}$$

where $|k\rangle$ is the kth element of a fixed basis, arithmetic inside the ket symbol is taken modulo d and ω is the d-th primitive root of unity. Thinking of the basis vectors $|k\rangle$ as the "position" basis, the operator V advances the position by one unit. Eigenstates of this translation operator are clearly Fourier transforms of the position basis states, i.e. the momentum basis states. In this basis U advances the momentum by one unit. Combining position and momentum translations leads to the full set of group elements, called the displacement operators, and for reasons having to do with the projective nature of the representation, they are best defined as

$$D_{jk} = \omega^{jk/2} \sum_{m=0}^{d-1} \omega^{jm} |k \oplus m\rangle\langle m| \,, \tag{2.29}$$

Sets of vectors based on the discrete analog of the Weyl-Heisenberg group are complete for any "window vector"; that is, for any normalized $|\psi\rangle \in \mathbb{S}^d$,

$$S_\psi = \sum_{jk} D_{jk} |\psi\rangle\langle\psi| D_{jk}^\dagger = dI \,, \tag{2.30}$$

a fact readily checked by direct calculation. Hence the displacement operators applied to any vector immediately yield a tight frame.

This definition of the displacement operators implies several useful relationships. First, checking the group representation property, note that

$$D_{jk}D_{lm} = \omega^{(jm-kl)/2}D_{j+l,k+m}. \tag{2.31}$$

Due to the minus sign in the difference of the two exponential terms, these operators and representations derived from them inherit the symplectic structure of phase space. Some other nice properties which can be derived by direct calculation:

$$\begin{aligned}
D_{jk}^{\dagger} &= D_{-j,-k} && (2.32)\\
D_{jk}D_{lm}D_{jk}^{\dagger} &= \mathcal{D}_{jk}(D_{lm}) = \omega^{jm-kl}D_{lm} && (2.33)\\
FD_{jk}F^{\dagger} &= D_{k,-j} && (2.34)\\
[D_{jk}, D_{lm}] &= 2i\sin\left(\frac{jm-kl}{d}\right)D_{j+l,k+m} && (2.35)\\
\mathrm{Tr}\Big[D_{jk}^{\dagger}D_{lm}\Big] &= d\,\delta_{jl}\delta_{km}, && (2.36)
\end{aligned}$$

where F is the Fourier transform operator.

Part II

Foundations: Probability and Representation

Chapter 3

The Quantum Probability Rule

Frame theory is immediately relevant to quantum mechanics through the identification of normalized tight frames with quantum measurements, POVMs. Borrowing methods from frame theory to find optimal measurements for one problem or another will be a topic of later chapters, but now the focus turns to more foundational questions: what characterizes the nature of quantum mechanics as a probabilistic theory? Is the theory to be understood as is classical mechanics, with underlying physical properties and probability used to encode an observer's knowledge of these? Or perhaps as promoting probability *itself* to an ontological status, in the guise of the state vector?

Naturally, this is a false dilemma; neither of these choices is satisfactory. Indeed, one should fall back to a third, more operational, position to avoid the logical abyss that marks interpretations of quantum mechanics. In this context we'll find that from the structure of measurements themselves, the POVMs, follows the probabilistic structure of quantum theory, the density operator. Sections 3.1 and 3.2 consider the general case, and section 3.3 specializes to qubits. With this structure in place, it becomes clear what form dynamics may take, and the rest of the theory falls into place. In this sense we are reworking the axioms of quantum mechanics to start from a different point of view. From this vantage point we are allowed to glimpse beyond the

usual intellectual knot springing from the collision of the poorly-demarcated concepts *probability* and *measurement*. Traditional interpretations of quantum mechanics sail well at a distance from these two, but here they run aground.

In classical mechanics these two concepts arouse no difficulty: measurements reveal certain pre-existing properties, and probabilities are a way to keep track of what is known and how well. One may usefully think of probabilities as simply a gambler's aid; asking the gambler for the odds at which he's willing to bet is akin to asking how certain he is of a particular outcome. Such is the view when considering probabilities subjectively or epistemically, as did Bayes, Laplace, and a great many of the physicists of the 19th century.

Zeal for objective physical theories attempted to place probability itself in the more objective or ontological realm in the 20th century. Though clearly appealing, this attempt fails because in the final analysis, no sensible statements can be made about objective probability. The most famous attempt to do so is to consider probability of an outcome as identical to its frequency of occurrence in a large ensemble. This is indicative of how probability is used, and would explain where it comes from. However, such ensembles don't actually exist. Imagine the weatherman at his work examining the set of all tomorrows and counting the number in which it rains. One might like to instead appeal to the frequency of a hypothetical ensemble. But this has two problems. First, since the ensemble is hypothetical, on what grounds should we regard the resultant probability as objective? Second, the statement that probability is frequency comes from the law of large numbers, which is itself couched in probabilistic terms. So instead of defining probability by frequency, the opposite is accomplished.

Having been deprived of this bit of objectivity, quantum mechanics invites us to cast even more aside, for one cannot make sense of what it predicts by appealing to "certain pre-existing properties", or at least not local ones. This is Bell's famous result. Considering two separated physical systems and the joint probability distributions of measurements that observers might make on them, he demonstrated that

quantum mechanics offers some probability distributions which cannot arise from distributions of locally realistic quantities. Thus we are forced to choose. If we desire a local theory, then we must abandon the idea that there is a real state of affairs of each of these two subsystems. On the other hand, if we wish to hold on to reality, we must accept that the states of affairs are correlated in a manner that can't be explained locally. In some sense the two particles aren't distinct at all. Though the nonlocal path may indeed lead out of this dilemma, it appears from this vantage point rather to lead deeper into the thicket as it questions the very understanding of distinct physical systems. Not that the other choice is any more palatable, for no longer can measurements be said to reveal properties of systems.

Perhaps we should look at quantum mechanics with fresh eyes and consider that perhaps probability *itself* is the realistic part of the theory. The idea of using frequencies is still a nonstarter for the same reasons mentioned above. However, quantum mechanics gives more credence to the notion of propensity, or objective chance. We might simply think of the state vector of a system as objective, and since it encodes probabilities for any measurements done, the probabilities are therefore also objective. Viewing pure states as akin to classical properties naturally leads to considering convex combinations of them, where the coefficients represent probabilities of the system to be in one state or another. This step, however, causes the whole edifice to unravel. Though convex, the set of all mixed states is not a simplex: there's no unique way to consider a given mixed state as a convex combination of pure states. Thus the propensities and the probabilities are thoroughly mixed, and there's no sense in differentiating between them. Probabilities are either entirely objective or entirely subjective.

We could eschew subjective probabilities by omitting mixed states entirely, basing the theory and its interpretation solely on pure states. But we must still contend with the measurement problem. Measurements are repeatable if nothing else, so if a pure state yields a certain outcome on the first round—but was not destined to do so with probability one—it must somehow change in order to certainly yield

that outcome on the second round. One remedy is to consider that measurement instantly changes the state vector, a process termed "wavefunction collapse". Now there are two rules for dynamics, the normal unitary evolution and instantaneous measurement dynamics. As if this dichotomy weren't bad enough for a universal theory, the exact conditions for what constitutes a measurement are vague, so it's problematic at best to say which one is happening when.

Collapse interpretations aren't necessary if we are willing to include the observer, which we should do anyway for an objective theory. The many-worlds interpretation provides a scheme for doing this, whereby to say that superposition states are measured is to say that the observer becomes part of the superposition, with one term describing each possible outcome. Since the state vector is objective, all these "branches" are somehow objective, too; we just never directly contact them.[1] But due to the linear structure of the theory, there's no unique way to express the whole tree as a sum of branches; the branching process could be written using different sets of pure states. Hence the act of measuring does not pin down exactly what was measured, or what outcomes occurred in each branch. This is the analog of the nonuniqueness of mixed-state decompositions just encountered.

Like so much ether, new concepts which have no purpose other than to explain the theory are invoked in order to interpret quantum mechanics in the familiar language of classical mechanics. Neither of the alternatives offered in the opening paragraph is truly satisfying. By tossing them out and concentrating on a more operational approach, we may hope to build up the theory in a conceptually clean manner, so that the appropriate concepts may later be discovered. Such a scheme begins with measurement directly, positing simply that measurements happen and outcomes occur, whatever the reason.

To adopt this perspective, imagine that for any experiment, there exist mathematical objects which represent the possible outcomes. Likewise, there exists a

[1] Though if quantum mechanics were even weakly nonlinear, this could be accomplished, a phenomenon called the "Everett phone."

mathematical object the experimenter uses to describe the system under investigation. The minimal task of any physical theory is to determine what these objects are and to use them to furnish the probabilities for the various outcomes. Though the aforementioned problems with measurement and probability do not arise in this context, a new conceptual question emerges: Why the Hilbert-space description of quantum measurement outcomes [64, 71, 108]?

Leaving this question aside as too ambitious for the present, start by assuming the form of measurements as given by classical or quantum theory. In classical physics, the measurement objects are the points in a phase space, while in quantum physics, they are traditionally one-dimensional projectors on a Hilbert space. Classically, only one measurement exists—a full accounting of the phase space. In quantum mechanics, on the other hand, any complete set of one-dimensional orthogonal projectors suffices.

The description of the system can be given, in classical mechanics, by a phase-space point. This point is the "true" point—others are "false"—so the outcome of a measurement can be predicted with certainty. Attempting such a concrete description in quantum mechanics is ruled out by the Kochen-Specker theorem: There is no way to assign truth and falsity to all the one-dimensional projectors in such a way that in any measurement there is only one true outcome [96]. At this point in the development, quantum mechanics becomes an irreducibly probabilistic theory; the possibility of underlying certainties has been ruled out.

With certainties ruled out, Gleason's theorem delineates the allowable descriptions of the system, i.e., the form the probabilities can take [68]. Keeping with the linear structure, every outcome probability is an inner product of the corresponding measurement projector and a *density operator* for the system. The density operator—any convex combination of one-dimensional projection operators—represents the description or "state" of the system. Thus Gleason's theorem gives a means to go from the structure of measurements to the structure of states. It immediately implies the Kochen-Specker result, as there are no density operators that

yield probability distributions for all measurements that are valued only on zero and one.

Neither theorem holds for two-dimensional quantum systems, so-called qubits, as long as the measurement objects are restricted to being one-dimensional projection operators. We can include this outstanding case, however, by widening the class of allowed measurements to include positive-operator-valued measures (POVMs), comprised of measurement operators called *effects*, and in so doing, we also considerably simplify the derivation. To coherently present these results, the remainder of the chapter is organized as follows. Section 3.1 describes how measurement outcomes in quantum mechanics are associated with effects. Section 3.2 shows that given the structure of effects, the usual quantum-mechanical probability rule follows simply, even for qubits. Section 3.3 investigates several specific restricted classes of measurements for qubits and what kinds of probability distributions these classes permit.

It should be emphasized from the outset that this is an inherently *noncontextual* approach, meaning that a shared outcome of two distinct measurements has the same probability in both contexts. For instance, if an experimenter has the choice between two measurement devices, both of which contain the same outcome, for the same input state this same outcome occurs with the identical probability in either setup. This fact is not a consequence of this approach, but rather an assumption used in its construction. In this construction a description of the measurement device and a description of the physical system are composed into a probability for each possible outcome. In defining measurements by their physical set-ups, this approach initially pertains to each measurement situation individually. The Hilbert-space formalism we are using to describe measurements takes a further step by associating the same measurement object with outcomes in different measurements. For this reason, we *assume* that these outcomes have the same probabilities, this being the noncontextual assumption. To abandon this assumption at the level of finding allowed probability assignments via Gleason's theorem would be to ignore, wholly or partially, the framework provided by the vector-space structure of measurements.

3.1 Effect Operators

In the typical von Neumann formulation of quantum measurement theory, measurements are described by complete sets of orthogonal projection operators. Here we consider quantum measurements in their full generality, the so-called POVMs [101, 23]. A POVM is also a complete set of operators resolving the identity operator, but comprised of positive operators less than the identity. These operators, called *effects*, can also be characterized as Hermitian operators having eigenvalues in the unit interval. The set of effects in d dimensions is denoted by $\mathcal{E}_d$.

That $\mathcal{E}_d$ is a convex set is clear. The projection operators of all ranks, including the zero operator $\mathbf{0}$ and the identity operator I, form the extreme points of $\mathcal{E}_d$, a fact demonstrated via the following construction. For a given effect E, order its d eigenvalues (including zero eigenvalues) from smallest to largest, $\{\lambda_1, \ldots, \lambda_d\}$. Associated with each eigenvalue λ_j is a one-dimensional projector π_j onto the corresponding eigenvector; these projectors can be chosen to make up a complete, orthonormal set. Now form the projection operators $\Pi_k = \sum_{j=k}^{d} \pi_j$. Clearly $\Pi_1 = I$ and $\Pi_d = \pi_d$. These projectors get smaller in rank as the index gets bigger. The effect E can be written as the following convex combination:

$$E = \lambda_1 \Pi_1 + \sum_{m=2}^{d} (\lambda_m - \lambda_{m-1}) \Pi_m + (1 - \lambda_d)\mathbf{0} \; . \tag{3.1}$$

The zero operator is included to make the sum of the coefficients unity without affecting E.

Since every effect can be expanded as a convex combination of projectors, only projectors can be extreme points of $\mathcal{E}_d$. To show that all the projectors are extreme points, we need to show that a projector Π cannot be written as a *proper* convex combination of other projectors, i.e., cannot be written as $\Pi = a\Pi_1 + (1 - a)\Pi_2$, where $0 < a < 1$ and $\Pi_1 \neq \Pi_2$. To show this, suppose Π could be so written. For any normalized vector $|\psi\rangle$, we have $\langle\psi|\Pi|\psi\rangle = a\langle\psi|\Pi_1|\psi\rangle + (1 - a)\langle\psi|\Pi_2|\psi\rangle$. If $|\psi\rangle$ is in the null subspace of Π, i.e., is orthogonal to the support of Π, we have

$\langle\psi|\Pi|\psi\rangle = 0$, which implies that $\langle\psi|\Pi_1|\psi\rangle = \langle\psi|\Pi_2|\psi\rangle = 0$; this shows that the supports of Π_1 and Π_2 are contained in the support of Π. If $|\psi\rangle$ is in the support of Π, we have $\langle\psi|\Pi|\psi\rangle = 1$, which implies that $\langle\psi|\Pi_1|\psi\rangle = \langle\psi|\Pi_2|\psi\rangle = 1$; this shows that the support of Π is contained in the supports of Π_1 and Π_2. Together these conclusions imply that $\Pi_1 = \Pi_2 = \Pi$, contradicting our assumption of a proper convex combination for Π. Thus the extreme points of $\mathcal{E}_d$ are the projectors of all ranks, including $\mathbf{0}$.

Projection operators are a limiting case of effect operators, the latter being a "fuzzy" or "unsharp" version of the former by convex combination. Similarly, the von Neumann projective measurements are a limiting case of POVMs.

In two dimensions the set of effects, $\mathcal{E}_2$, has an appealing geometric picture. Beginning with the parameterization of Hermitian operators by the Pauli matrices, write the general two-dimensional effect as

$$E = rI + \mathbf{s}\cdot\boldsymbol{\sigma} = rI + s\mathbf{n}\cdot\boldsymbol{\sigma}\ , \tag{3.2}$$

where $\mathbf{n}$ is a unit vector. The eigenvalues $r \pm ||\mathbf{s}|| = r \pm s$ must lie in the unit interval, which is equivalent to the conditions $0 \leq r \leq 1$ and $0 \leq s \leq \min(r, 1-r)$. Since s characterizes the radius of a sphere, the full set can be pictured in the following way: Starting with the unit interval for r, associate with each point a ball of radius r for $r \leq 1/2$ and of radius $1-r$ for $r > 1/2$. The set of two-dimensional effects is thus the intersection of two three-dimensional cones (i.e., two cones in four real dimensions) both having an opening angle of 45°, one extending up from a vertex at $r = s = 0$ ($E = \mathbf{0}$) and the other extending down from a vertex at $r = 1$, $s = 0$ ($E = I$). The intersection of the boundaries of the two cones, $r = s = 1/2$, is the surface of the Bloch sphere, where the effects are one-dimensional projectors. The effects on the boundary of the lower cone, $r = s \leq 1/2$, are multiples of one-dimensional projectors.

3.2 The Quantum Probability Rule

The task of quantum theory is, minimally, to associate with every measurement a probability distribution for its outcomes. This is done noncontextually for the reasons given above. The probability rule is thus a function from the set of effects to the unit interval, which is normalized on any subset that makes up a POVM. Such a function is known as a *frame function*. More precisely, a frame function is a function $f : \mathcal{E}_d \to [0, 1]$ that satisfies

$$\sum_{E_j \in X} f(E_j) = 1 \tag{3.3}$$

on any subset $X = \{E_j \in \mathcal{E}_d | \sum_j E_j = I\}$. This section is devoted to proving the following Gleason-type theorem, which was first proved by Busch [24].

Theorem 6 *For every frame function $f : \mathcal{E}_d \to [0, 1]$, there is a unique unit-trace positive operator W such that $f(E) = (W, E) = \mathrm{Tr}(WE)$.*

The operator W is the density operator that gives rise to the frame function probabilities $f(E)$. The proof of this theorem is divided into several parts, each of which occupies a subsection.

3.2.1 Linearity with Respect to the Nonnegative Rationals

Every frame function is trivially additive, for consider two POVMs, $\{E_1, E_2, E_3\}$ and $\{E_1 + E_2, E_3\}$. Clearly both are POVMs if either is, and the frame-function requirement immediately yields

$$f(E_1) + f(E_2) = f(E_1 + E_2) \, . \tag{3.4}$$

From this we obtain a homogeneity property for multiplication by rational numbers. We can break an effect nE into m pieces to form the effect $(n/m)E$. Using the additivity property twice, we obtain

$$mf\Big(\frac{n}{m}E\Big) = f(nE) = nf(E) \qquad \Longrightarrow \qquad f\Big(\frac{n}{m}E\Big) = \frac{n}{m}f(E) \, . \tag{3.5}$$

The function f is thus established to be linear in the nonnegative rationals. We can extend to full linearity by proving continuity. Alternately, adopting the strategy of Busch [24], we can demonstrate the homogeneity of f, from which linearity follows immediately. These two arguments are taken up in turn in the next two subsections.

3.2.2 Continuity

Continuity of the frame function can be established via *reductio ad absurdum*: A contradiction with the definition of a frame function arises if f is discontinuous. Recall the definition of continuity for metric spaces: f is continuous at x_0 if for all $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ for all x satisfying $|x - x_0| < \delta$. On the space of operators, we use the Hilbert-Schmidt inner product $(A, B) = \mathrm{Tr}(A^\dagger B)$ and the associated norm $|A| \equiv \sqrt{(A, A)}$.

Consider first continuity at the zero operator(since $f(E) = f(E + \mathbf{0}) = f(E) + f(\mathbf{0})$, we know that $f(\mathbf{0}) = 0$). If we assume f is discontinuous at the zero operator, then there exists an $\epsilon > 0$ such that for all $\delta > 0$, there exists an effect E satisfying $|E| < \delta$ and $f(E) \geq \epsilon$. Choose $\delta = 1/N < \epsilon$, where N is an integer, and let E be an effect satisfying $|E| < 1/N$ and $f(E) \geq \epsilon$. Now $F = NE$ satisfies $|F| = N|E| < 1$, which implies that F is an effect, since the sum of the squares of its eigenvalues is less than 1. But we also have from additivity that $f(F) = Nf(E) \geq N\epsilon > 1$, contradicting the definition of a frame function. Hence f is continuous at the zero operator.

We can easily translate the continuity at $\mathbf{0}$ to the entire set of effects. To prove continuity at an arbitrary effect E_0, we need to consider neighboring operators E and the difference $E - E_0$. Diagonalizing the difference, we can write $E - E_0 = A - B$, where A is the nonnegative-eigenvalue part of the eigendecomposition and $-B$ is the negative-eigenvalue part. It is clear that A and B are positive operators satisfying $|A|, |B| \leq |A - B| = |E - E_0|$, which implies that A and B are effects (provided $|E - E_0| \leq 1$). Applying the frame function to the equation $E + B = E_0 + A$

yields $f(E) - f(E_0) = f(A) - f(B)$ by additivity. Invoking continuity at zero establishes that for every $\epsilon = \epsilon'/2 > 0$, there exists a $\delta > 0$ such that $|A|, |B| < \delta \Rightarrow f(A), f(B) < \epsilon'$. Thus if $|E - E_0| = |A - B| < \delta$, we have $|A|, |B| < \delta$ and $|f(E) - f(E_0)| = |f(A) - f(B)| \leq |f(A)| + |f(B)| < 2\epsilon' = \epsilon$. This establishes the continuity of f on all of $\mathcal{E}_d$, which in turn shows that f is a linear function on $\mathcal{E}_d$.

3.2.3 Homogeneity

An alternative route to linearity is to prove the homogeneity of the frame function. Following Busch's proof, first note that the frame function preserves order; i.e., if $E_1 < E_2$ for any pair of measurement operators, then $f(E_1) \leq f(E_2)$. This follows immediately from the definition, for $E_1 < E_2 \Leftrightarrow E_2 - E_1 \equiv E_3 > 0$, so E_3 is an effect. Writing $E_2 = E_1 + E_3$, which implies $f(E_2) = f(E_1) + f(E_3)$ by additivity, we find that $f(E_1) \leq f(E_2)$ since $f(E_3) \geq 0$.

Now the pinching theorem can be used to establish the homogeneity of f. Consider two sequences of rational numbers sharing the same irrational limit α: $\{q_i\}$ is an increasing sequence and $\{p_i\}$ a decreasing sequence. By order preservation and linearity in the nonnegative rationals, we have

$$q_i f(E) = f(q_i E) \leq f(\alpha E) \leq f(p_i E) = p_i f(E) \tag{3.6}$$

for all i. The pinching theorem shows that $f(\alpha E) = \alpha f(E)$, establishing that f is a homogeneous and, hence, linear function.

3.2.4 Linearity and the Inner Product

Since the frame function is linear, it arises from an inner product. To show this, the definition of f is extended to the entire vector space of operators; then it is a trivial theorem of linear algebra to recast a linear function on a vector space as an inner product.

The frame function is extended in the most straightforward fashion. Let H be an arbitrary Hermitian operator. Every such operator can be written as the difference of two positive operators G_1 and G_2; one way to do so is simply to diagonalize H and to let G_1 be the positive-eigenvalue part and $-G_2$ the negative-eigenvalue part. Further, for any positive operator G there exists a positive number α such that $\alpha E = G$ for some effect E. Now define $f(H) = f(G_1) - f(G_2) = \alpha_1 f(E_1) - \alpha_2 f(E_2)$. Though the unraveling of H is not unique, the extension is. Suppose $H = \alpha_1 E_1 - \alpha_2 E_2 = \alpha_3 E_3 - \alpha_4 E_4$, which implies $\alpha_1 E_1 + \alpha_4 E_4 = \alpha_2 E_2 + \alpha_3 E_3$. Choose β such that $\beta \geq \max\{\alpha_j\}$ so that we have

$$\frac{\alpha_1}{\beta} E_1 + \frac{\alpha_4}{\beta} E_4 = \frac{\alpha_2}{\beta} E_2 + \frac{\alpha_3}{\beta} E_3 \; . \tag{3.7}$$

Since every operator is now in the original domain of f, we can apply the frame function to find

$$\alpha_1 f(E_1) + \alpha_4 f(E_4) = \alpha_2 f(E_2) + \alpha_3 f(E_3) \; . \tag{3.8}$$

The extension being manifestly linear, we now have a linear function f on the entire space of Hermitian operators. It can be extended to all operators by complexification.

To rewrite this linear function as an inner product, choose an orthonormal operator basis $\{\tau_j\}$, and write an arbitrary operator as $A = \sum_j \tau_j (\tau_j, A)$. Clearly, then $f(A) = \sum_j f(\tau_j)(\tau_j, A)$. Now define W as the unique solution of the d^2 equations $f(\tau_j) = (W, \tau_j)$, so that the frame function is $f(A) = \sum_j (W, \tau_j)(\tau_j, A) = (W, A)$. The d^2 equations are, of course, nonsingular since $\{\tau_j\}$ is an orthonormal basis.

The nonnegativity and normalization of the frame function induce the density operator properties of W, i.e., positivity and unit trace. Given an arbitrary normalized vector $|\psi\rangle$, $0 \leq f(|\psi\rangle\langle\psi|) = \langle\psi|W|\psi\rangle$, showing that W is positive. The condition of unit trace follows from normalization:

$$\operatorname{Tr} W = (W, I) = \Big(W, \sum_j E_j\Big) = \sum_j (W, E_j) = \sum_j f(E_j) = 1 \; . \tag{3.9}$$

The reader should note that if the extension of linearity to real numbers is omitted (sections 3.2.2 and 3.2.3), the arguments in sections 3.2.1 and 3.2.4 demonstrate

the quantum probability rule for vector spaces over rational fields. This result has implications for recent discussions in the literature about the possibility of describing the finite precision of real-world measurements via vector spaces over complex numbers with rational parts (see, in particular, references [116, 40, 4, 27]). Without dwelling on these points, note that this result shows that rational POVMs cannot be assigned truth values, the only frame functions being those derived from density operators. The most straightforward approach is to think of POVMs as the preferred description of finite-precision measurements, and thus that is all that needs to be said about finite-precision quantum measurements.

3.3 Frame Functions for Qubits

The POVM version of Gleason's theorem works even for qubits, unlike the original Gleason theorem, which was based on measurements described by one-dimensional orthogonal projectors. We now turn our attention specifically to qubits and investigate whether several restricted sets of POVMs enforce the quantum probability rule. In particular, the quantum probability rule is necessitated by a particular subset of POVM measurements called *trines*. Other measurements are studied to shed light on what kinds of measurements yield the quantum probability rule and why.

3.3.1 General Description of Restricted Sets of POVMs

To start, recall that the two-dimensional effects are parameterized by four real parameters as in equation 3.2. In any of the restricted sets of measurements considered in this section, the allowed POVMS are made up of effects that are multiples of one-dimensional projectors, i.e., $r = s \leq 1/2$ and $E = r(I + \mathbf{n} \cdot \boldsymbol{\sigma})$, and all the effects have the same value of r. An allowed POVM is thus specified by a set of unit vectors that sum to the zero vector; if there are N outcomes in the POVM, then $r = 1/N$.

Finally, we assume that the allowed POVMs are rotationally invariant; i.e., they are obtained by applying all possible 3-dimensional rotations to any particular POVM in the allowed subset.

With these assumptions, we have the following structure. For POVMs having N outcomes, the allowed effects have the form

$$E = \frac{1}{N}(I + \mathbf{n}\cdot\boldsymbol{\sigma}) \; , \tag{3.10}$$

where $\mathbf{n}$ can be any unit vector. Frame functions defined on this set, i.e.,

$$f\Big(\frac{1}{N}(I + \mathbf{n}\cdot\boldsymbol{\sigma})\Big) \equiv F(\mathbf{n}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} c_{lm}Y_{lm}(\mathbf{n}) \; , \tag{3.11}$$

are functions on the unit sphere and thus can be expanded in terms of spherical harmonics. In writing the spherical-harmonic expansion, we are assuming that F is continuous on the unit sphere. (Please be careful to note that this extra continuity assumption was not made in the full-fledged Gleason-type theorem of the previous section.) Properties of the spherical harmonics that we need in the following are the following: (i) the separation into θ and ϕ (or n_z and $n_x + in_y$) dependencies,

$$Y_{lm}(\mathbf{n}) = Y_{lm}(\theta,\phi) = \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi} = h_{lm}(n_z)(n_x + in_y)^m \; , \tag{3.12}$$

where $P_l^m(x)$ is an associated Legendre function and h_lm is defined implicitly, and (ii) the changes under conjugation, reflection, and parity,

$$\begin{aligned} Y_{lm}(\mathbf{n}) &= (-1)^m Y^*_{l,-m}(\mathbf{n}) = (-1)^{l+m}Y_{lm}(\pi-\theta,\phi) \\ &= (-1)^m Y_{lm}(\theta,\phi+\pi) = (-1)^l Y_{lm}(-\mathbf{n}) \; . \end{aligned} \tag{3.13}$$

Particularly useful is the form of equation 3.12 for $m = l$: $Y_{ll}(\mathbf{n}) \propto (n_x + in_y)^l$.

The sought-after quantum rule is

$$F(\mathbf{n}) = \mathrm{Tr}(WE) = \frac{1}{N}(1 + \mathbf{n}\cdot\mathbf{P}) \; , \tag{3.14}$$

where $\mathbf{P} = \mathrm{Tr}(W\boldsymbol{\sigma})$ is any 3-vector such that $\|\mathbf{P}\| \leq 1$. The quantum rule evidently contains only $l = 0, 1$ spherical harmonics, with $c_{00} = \sqrt{4\pi}/N$, $c_{10} = \sqrt{4\pi/3}\,P_z/N$, and $c_{1,\pm1} = \sqrt{2\pi/3}\,(\mp P_x + iP_y)/N$.

Now let the set of unit vectors $\{\mathbf{n}_1, \ldots, \mathbf{n}_N\}$ specify a "fiducial" POVM; the completeness property of a POVM implies that these vectors specify a POVM if and only if

$$0 = \sum_{j=1}^{n} \mathbf{n}_j \ . \tag{3.15}$$

Any other set, $\{R\mathbf{n}_j\}$, where R is a three-dimensional rotation, is also a POVM. The frame condition is that

$$1 = \sum_{j=1}^{N} F(R\mathbf{n}_j) = \sum_{l,m} c_{lm} \sum_{j=1}^{N} Y_{lm}(R\mathbf{n}_j) = \sum_{l=0}^{\infty} \sum_{m,r=-l}^{l} c_{lm} \mathcal{D}^{(l)*}_{mr}(R) \sum_{j=1}^{N} Y_{lr}(\mathbf{n}_j) \tag{3.16}$$

for all rotations R. In writing the last equality, we use

$$Y_{lm}(R\mathbf{n}) = \sum_{r=-l}^{l} Y_{lr}(\mathbf{n}) \mathcal{D}^{(l)}_{rm}(R^{-1}) = \sum_{r=-l}^{l} Y_{lr}(\mathbf{n}) \mathcal{D}^{(l)*}_{mr}(R) \ , \tag{3.17}$$

where $\mathcal{D}^{(l)}_{mr}(R)$ is the irreducible (unitary) matrix representation of the rotation R in the angular-momentum subspace with angular momentum l.

We can now use the fundamental orthogonality property of the irreducible representations of the rotation group [149]:

$$\int d\mu_R \, \mathcal{D}^{(l)*}_{mr}(R) \mathcal{D}^{(l')}_{m'r'}(R) = \frac{1}{2l+1} \delta_{ll'} \delta_{mm'} \delta_{rr'} \ . \tag{3.18}$$

Here the integration is over the invariant measure $d\mu_R$ of the rotation group. Noting that $\mathcal{D}^{(0)}_{00}(R) = 1$, we can use this orthogonality relation to invert equation 3.16, obtaining the condition $c_{lm} \sum_{j=1}^{N} Y_{lr}(\mathbf{n}_j) = \delta_{l0}$ for all l, m, and r. For $l = 0$ this is a trivial normalization constraint, satisfied by choosing $c_{00} = \sqrt{4\pi}/N$. For $l \geq 1$, we can write the condition in a more illuminating, equivalent form,

$$c_{lm} = 0 \text{ for } m = -l, \ldots, l, \qquad \text{or} \tag{3.19}$$

$$\sum_{j=1}^{N} Y_{lr}(\mathbf{n}_j) = 0 \text{ for } r = -l, \ldots, l. \tag{3.20}$$

These are necessary and sufficient conditions for a frame function $F(\mathbf{n})$.

The frame conditions 3.19 and 3.20 are a potent restriction. They say that if the lth harmonic is allowed in $F(\mathbf{n})$, then the unit vectors for the fiducial POVM must satisfy the sum condition 3.20. The choice of fiducial POVM being arbitrary, the sum condition must be satisfied by the unit vectors for all POVMs in the restricted set under consideration. This extension from a fiducial POVM to all POVMs is, however, automatic: If a fiducial POVM satisfies the sum condition 3.20, then the rotation property 3.17 of spherical harmonics guarantees that the condition is satisfied by all POVMs in the restricted set.

The sum condition 3.20 is automatically satisfied for $l = 1$ by virtue of the completeness condition 3.15. For higher l, if one finds a nonzero value of $\sum_{j=1}^{N} Y_{lr}(\mathbf{n}_j)$ for just one value of r and just one set of POVM vectors $\{\mathbf{n}_j\}$, then the lth harmonic must be absent from frame functions. On the other hand, if the sum condition 3.20 is satisfied for a fiducial POVM, then $F(\mathbf{n})$ can contain the lth harmonic. The expansion coefficients c_{lm} cannot be chosen arbitrarily, of course, since $F(\mathbf{n})$ must be real and nonnegative. Making $F(\mathbf{n})$ real is trivial—simply choose $c_{l,-m} = (-1)^m c_{lm}^*$—but delineating the region of coefficients that gives rise to nonnegative functions is generally quite a difficult task. Nonetheless, we can conclude that the lth harmonic is allowed whenever the sum condition is met, for the following reason. Because the spherical harmonics are bounded functions, sufficiently small expansion coefficients c_{lm} can be combined with $c_{00} = \sqrt{4\pi}/N$ without making F negative.

The spherical harmonic $Y_{lr}(\mathbf{n})$ can be regarded as the rth component of a rank-l spherical tensor formed from $\mathbf{n}$; it is a linear combination of the components of the rank-l symmetric trace-free Cartesian tensor formed from $\mathbf{n}$. The vanishing of $\sum_{j=1}^{N} Y_{lr}(\mathbf{n}_j)$ simply says that the sum of these tensors over all the unit vectors in a POVM vanishes. Harmonics $l = 1$ and $l = 2$ illustrate what is going on. The $l = 1$ spherical-tensor components $Y_{1r}(\mathbf{n})$ are linear combinations of the Cartesian components of $\mathbf{n}$ and thus always sum to zero over the unit vectors in a POVM as a consequence of the constraint 3.15. The $l = 2$ spherical-tensor components $Y_{2r}(\mathbf{n})$ are

linear combinations of Cartesian components of the symmetric trace-free two-tensor $n_k n_l - \frac{1}{3}\delta_{kl}$. Thus $\sum_{j=1}^{N} Y_{2r}(\mathbf{n}_j) = 0$ if and only if

$$\sum_{j=1}^{N} (n_j)_k (n_j)_l = \frac{N}{3}\delta_{kl} \; ; \tag{3.21}$$

i.e., the sum of the projectors onto the 3-vectors $\mathbf{n}_j$ is proportional to the three-dimensional identity operator.

We can make another general statement. If the allowed POVMs are made up of pairs of (subnormalized) orthogonal projectors—i.e., $-\mathbf{n}$ is in the set of unit vectors if $\mathbf{n}$ is—then the parity property of the spherical harmonics implies that all odd harmonics are allowed in the frame function. This, of course, is the reason that Gleason's theorem does not hold for qubits if the allowed measurements are restricted to orthogonal projectors.

We turn now to applying the sum condition 3.20 to particular sets of POVMs.

3.3.2 Restricted Sets of POVMs

Trine measurements

Consider first the trine measurements, three-outcome measurements described by three unit vectors equally spaced in a plane, which thereby sum to zero. We consider the sum $\sum_{j=1}^{3} Y_{ll}(\mathbf{n}_j)$ for the particular trine

$$\mathbf{n}_1 = \mathbf{e}_x \; , \qquad \mathbf{n}_{2,3} = -\frac{1}{2}\,\mathbf{e}_x \pm \frac{\sqrt{3}}{2}\,\mathbf{e}_z \; . \tag{3.22}$$

Using $Y_{ll}(\mathbf{n}) \propto (n_x + i n_y)^l$, we find that the sum is proportional to $1 + (-1)^l/2^{l-1}$. This being nonzero for all l except $l = 1$, we conclude that a frame function has no harmonics higher than $l = 1$. To establish the precise form of the quantum rule, we must require that the frame function be real and nonnegative. Having ruled out all but harmonics $l = 0, 1$, we can write the general form as $F(\mathbf{n}) = (1 + \mathbf{n}\cdot\mathbf{P})/N$,

where $\mathbf{P}$ is a 3-vector required to be real by the reality of $F(\mathbf{n})$. The nonnegativity of $F(\mathbf{n})$ then requires that $\|\mathbf{P}\| \leq 1$, leaving us with the standard quantum rule.

Tetrahedral Measurements

Now consider the tetrahedral measurements, four-outcome measurements whose unit vectors point to the vertices of a tetrahedron. Tetrahedral measurements are important as the two-dimensional example of a *symmetric informationally complete* POVM, as will be seen in the following chapter.

Since the vertices of a tetrahedron satisfy

$$\sum_{j=1}^{4}(n_j)_k(n_j)_l = \frac{4}{3}\delta_{kl} \;, \tag{3.23}$$

we can conclude immediately, as discussed above, that $\sum_{j=1}^{4} Y_{2r}(\mathbf{n}_j) = 0$ for all r and all tetrahedra. This means that a frame function for a tetrahedral measurement can contain $l = 2$ harmonics and thus does not necessarily follow from the quantum probability rule.

To investigate the possibility of higher harmonics, consider the particular tetrahedron

$$\mathbf{n}_1 = \mathbf{e}_x \;, \qquad \mathbf{n}_2 = -\frac{1}{3}\,\mathbf{e}_x - \frac{2\sqrt{2}}{3}\,\mathbf{e}_z \;, \qquad \mathbf{n}_{3,4} = -\frac{1}{3}\,\mathbf{e}_x + \frac{\sqrt{2}}{3}\,\mathbf{e}_z \pm \sqrt{\frac{2}{3}}\,\mathbf{e}_y \;. \tag{3.24}$$

For this tetrahedron the sum $\sum_{j=1}^{4} Y_{ll}(\mathbf{n}_j)$ is proportional to

$$1 + \left(-\frac{1}{3}\right)^l + \left(-\frac{1}{3} + i\sqrt{\frac{2}{3}}\right)^l + \left(-\frac{1}{3} - i\sqrt{\frac{2}{3}}\right)^l = 1 + \frac{(-1)^l + 2(\sqrt{7})^l \cos l\alpha}{3^l} \;, \tag{3.25}$$

where $e^{i\alpha} = -1/\sqrt{7} + i\sqrt{6/7}$. It is not hard to verify that this quantity is nonzero for all $l \neq 1, 2, 5$ ($\cos 2\alpha = -5/7$ and $\cos 5\alpha = -121/49\sqrt{7}$), implying that a frame function cannot have any harmonics other than $l = 0, 1, 2, 5$.

To show that a frame function can have $l = 5$ harmonics, rotate the tetrahedron 3.24 by $-90°$ about the y axis, obtaining

$$\mathbf{n}_1 = \mathbf{e}_z \;, \qquad \mathbf{n}_2 = -\frac{1}{3}\,\mathbf{e}_z + \frac{2\sqrt{2}}{3}\,\mathbf{e}_x \;, \qquad \mathbf{n}_{3,4} = -\frac{1}{3}\,\mathbf{e}_z - \frac{\sqrt{2}}{3}\,\mathbf{e}_x \pm \sqrt{\frac{2}{3}}\,\mathbf{e}_y \;. \tag{3.26}$$

It is easy to see that for this tetrahedron, $\sum_{j=1}^{4} Y_{lr}(\mathbf{n}_j)$ vanishes when r is not a multiple of 3. For $r = 0$ this sum is proportional to $\sum_{j=1}^{4} P_l((n_j)_z) = 1 + 3P_l(-1/3)$, and for $r = 3n \neq 0$, it is proportional to $\sum_{j=2}^{4} P_l^{3n}((n_j)_z) = 3P_l^{3n}(-1/3)$. For $l = 5$, it is easy to check that $P_5(-1/3) = -1/3$ and $P_5^3(-1/3) = 0$, so the sum condition 3.20 holds for $l = 5$, and a frame function can contain $l = 5$ harmonics.

Other Measurements

We now apply our technique to other measurements with nice symmetry properties: the five platonic solids, planar polygons of any degree, and the uniform POVM.

Easiest is the case of the uniform POVM, i.e., a measurement whose outcomes include every direction on the Bloch sphere. There being only one POVM in the restricted set, the frame conditions reduce to the requirement that $F(\mathbf{n})$ be a real, nonnegative function that integrates to unity on the sphere. In terms of a spherical harmonic expansion, this fixes $c_{00} = 1/\sqrt{4\pi}$, while the other coefficients are only restricted so as to provide real, nonnegative function values.

Continuing in reverse order, consider the case when the allowed unit vectors lie in a plane, forming a regular polygon with N vertices. For a POVM with unit vectors lying in the x-y plane, it is easy to see that $\sum_{j=1}^{N} Y_{lr}(\mathbf{n}_j)$ is zero unless r is a multiple of N; when $r = nN$, the sum is proportional to $P_l^{nN}(0)$, which is zero (nonzero) if $l + nN$ is odd (even). As a consequence, the allowed harmonics for N even are $l = 0$ and all odd harmonics, and the allowed harmonics for N odd are $l = 0$ and the odd harmonics such that $l \leq N - 2$. Notice that the trine is the only regular polygon that gives the quantum probability rule.

The vertices of each of the five platonic solids yield a symmetric set of unit vectors for constructing restricted sets of POVMs. For all five, the projectors onto the unit vectors sum to a multiple of the 3-dimensional identity operator, implying that $l = 2$ harmonics are allowed; moreover, for all but the tetrahedron, the unit vectors come in antipodal pairs, meaning that all odd harmonics are allowed. For

Platonic solid	*Allowed harmonics*
tetrahedron	0, 1, 2, 5
octahedron	0, 2, & odds
cube	0, 2, & odds
dodecahedron	0, 2, 4, 8, 14, & odds
icosahedron	0, 2, 4, 8, 14, & odds

Table 3.1: Allowed spherical harmonics in the frame functions for POVMs based on the platonic solids.

the octahedron, by considering the sum $\sum_j Y_{ll}(\mathbf{n_j})$ over the six unit vectors that point along the x and $-x$ axes and into the four quadrants in the y-z plane, one finds that all even harmonics except $l = 0, 2$ are ruled out. Similarly, for the cube, this same sum over the tetrahedron 3.24 and its antipodal points rules out all even harmonics except $l = 0, 2$. For the dodecahedron and icosahedron, additional even harmonics are allowed, and numerical investigation using *Mathematica* shows these to be $l = 4, 8, 14$. The results for the platonic solids are summarized in Table 3.3.2.

Gleason's theorem reveals that from this "measurement first" point of view, the quantum probability rule is implicit in the structure of measurements themselves. Use of generalized measurements greatly simplifies the proof of Gleason's theorem, extends its applicability, and reduces the conceptual overhead of quantum theory. Attention is thereby refocused to the tougher task of justifying the form of quantum measurements [64, 71, 108].

Investigating particular measurements and their allowable probability distributions complements the investigation of Kochen-Specker colorable sets [28]. Considering particular sets of POVMs allows us to explore the range of probability distributions that lie between the quantum probability rule and Kochen-Specker truth assignments.

This work naturally leads to the question of whether POVMs ought to be considered, in some sense, more fundamental than standard projection-valued measures.

The reason for thinking so is not just the simplicity of proofs. Foremost is the notion of "fuzziness" that effects capture, a notion essential for practical purposes. Secondly, in thinking of quantum mechanics operationally, nothing singles out projection measurements for *fundamental* status. Physically implementing them is no more or less difficult in most circumstances than projective-valued measurements. What's more, while it's true that effects themselves are convex combinations of projection operators, POVMs needn't be convex combinations of projection-valued measures [45]. Finally, POVMs are *useful* measurements as we'll see in the following chapters, as can be seen in the wider field of quantum information as a whole [120]. All these reasons point toward a fundamental role.

Chapter 4

Standard Quantum Measurements

In the previous chapter, frames provided a method of describing the primitive concept of measurement in quantum mechanics and led to a simple method of determining the form of probability distributions quantum mechanics provides. In addition to describing measurements, frames are also suited to describing the quantum states which follow from them. Using tight frames to represent quantum states means representing states by measurements, which in turn casts quantum states as probability distributions over, or quasiprobability distributions in terms of, such measurements.

For systems described by an infinite-dimensional Hilbert space, such as a harmonic oscillator or a quantum field, such a construction, based on the standard coherent states, is quite well known. Summarizing the main features of representations based on this construction will allow us to compare with the new results that follow. Following chapter two, let $|0\rangle$ be the vacuum state and $D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$ the displacement operator, so that the coherent states are simply $|\alpha\rangle = D(\alpha)|0\rangle$. They form an overcomplete set for the underlying vector space $\mathcal{H}$, and their outer products form a complete set for the vector space of bounded operators on it, $\mathcal{B}(\mathcal{H})$. Since the coherent states are operator-complete, they are in principle suitable for representing arbitrary operators. However, in practice, such representations are highly singular for almost all operators except mixtures of coherent states.

Following chapter two, the Weyl-Heisenberg system offers four interrelated representations of any operator: the characteristic function, the Wigner function, and the P and Q functions. The first two are based on the group itself, while the latter two issue from the coherent states. The displacement operators can be shown to form an overcomplete basis, and the characteristic function of $A \in \mathcal{B}(\mathcal{H})$ is simply the representation of A in this basis. The Wigner function, meanwhile, is its Fourier transform:

$$\chi(\alpha) = \mathrm{Tr}[D^\dagger(\alpha)A] \qquad A = \frac{1}{\pi^2}\int \mathrm{d}^2\alpha\ \chi(\alpha)D(\alpha) \tag{4.1}$$

$$W(\alpha) = \frac{1}{\pi^2}\int \mathrm{d}^2\beta\ e^{\beta^*\alpha - \beta\alpha^*}\chi(\beta) \qquad A = \frac{1}{\pi^2}\int \mathrm{d}^2\alpha\ W(\alpha)\Delta(\alpha)\ . \tag{4.2}$$

The operator $\Delta(\alpha)$ is defined by displacing the parity operator Π:

$$\Delta(\alpha) = D(\alpha)\Pi D^\dagger(\alpha)\ . \tag{4.3}$$

This reconstruction formula from the Wigner function can be verified by first reverting to the characteristic function and then using the definition of the displacement operators.

The dual P and Q representations involve the coherent states themselves. While the P function expresses an operator in terms of the basis of coherent states, the Q function uses the coherent states to determine the expansion coefficients

$$A = \int \mathrm{d}\mu(\alpha)P_A(\alpha)|\alpha\rangle\langle\alpha| \tag{4.4}$$

$$Q_A(\alpha) = \langle\alpha|A|\alpha\rangle\ . \tag{4.5}$$

The Q and Wigner functions can be obtained from the P function by simple Gaussian convolutions

$$W(\alpha) = \frac{2}{\pi}\int \mathrm{d}^2\beta P(\beta)\exp[-2|\beta-\alpha|^2]\ , \tag{4.6}$$

$$Q(\alpha) = \frac{1}{\pi}\int \mathrm{d}^2\beta P(\beta)\exp[-|\beta-\alpha|^2]\ , \tag{4.7}$$

the only difference being the width of the Gaussian used. Reversing direction and determining P in terms of Q, for instance, is generally more difficult, because the convolution is with a divergent Gaussian [157].

The coherent states in infinite dimensions have a wealth of structure and symmetry which has only been touched on here. Much of this can be replicated in finite dimensions, and this has become a well-traveled path of inquiry. One fork in this path considers coherent states arising from various Lie groups, and is certainly a study of its own. Another fork considers the discrete Weyl-Heisenberg frames, and we shall stick to this path. Here one particular, as yet unstudied WH frame stands above the rest for its symmetry and simplicity: a symmetric, informationally-complete POVM, or SICPOVM for short. Informationally-complete measurements are those whose statistics uniquely determine the state being measured [125, 22, 137, 46]. However, the SICPOVM is much more besides. Simultaneously a WH frame, an equiangular spherical code, and a spherical 2-design, this ensemble is a collision of the elegant properties of the various frames, designs, and codes introduced in chapter 2. Such a measurement is not only elegant from a representation-theory point of view, but also useful. In quantum information theory the SICPOVM is relevant to quantum state tomography [32], quantum cryptography [65], and to foundational studies [64] where it would make for a particularly interesting "standard quantum measurement". The main question is whether or not it exists in any given dimension. We shall take up this problem for the next three sections, delving into the symmetry of the SICPOVM in section 4.1, then finding anlaytic and numerical solutions in sections 4.2 and 4.3. Then we can turn to a more detailed examination of the P and Q representation possibilities presented by the SICPOVM (section 4.4) and tomographic uses (section 4.5). The SICPOVM also provides a method to render quantum mechanics in the language of classical probabiliy theory, as shown in section 4.6.

The SICPOVM problem has been studied in a different context, first by Lemmes and Seidel [104] looking for a set of equiangular lines in $\mathbb{C}^d$, and subsequently by many others [50, 98, 77, 97, 147, 99, 60, 61, 169], but no general result is known. However, extensive numerical analysis, performed by the author and collaborators and including all dimensions up to 45, provides strong evidence that these coherent states exist in all dimensions.

Before tackling the existential problem, let us consider its precise definition. We may specify the SICPOVM solely by its symmetry condition, as its remaining properties follow. A SICPOVM $\mathcal{P}$ is a set of $n = d^2$ normalized vectors $|\phi_k\rangle$ in $\mathbb{C}^d$ satisfying

$$|\langle\phi_j|\phi_k\rangle|^2 = \frac{1}{d+1}\,, \qquad \forall\, j \neq k\,. \tag{4.8}$$

More precisely, the elements of $\mathcal{P}$ are the subnormalized projectors $|\phi_k\rangle\langle\phi_k|/d = \Pi_k/d = E_k$, which have pairwise Hilbert-Schmidt inner product $(E_j, E_k) = \mathrm{Tr}[E_j^\dagger E_k] = 1/d^2(d+1)$ for $j \neq k$.

The frame potential V_1 of such a collection of vectors is manifestly equal to $n/d = d$, so $\mathcal{P}$ is indeed a POVM. For $\mathcal{P}$ to be informationally complete, the d^2 operators $\Pi_k = |\phi_k\rangle\langle\phi_k|$ must be linearly independent so that they span the space of operators. The linear independence follows from considering the rank of their Gram matrix $(\Pi_j, \Pi_k) = \mathrm{Tr}[\Pi_j^\dagger \Pi_k] = (d\delta_{jk} + 1)/(d+1)$, which being circulant (each row is a cyclic shift of the previous row), has eigenvalues given by the Fourier transform of one of the rows. A simple calculation reveals that due to the combination of constant term and Kronecker delta, the eigenvalues are exactly the same as the values in any row. Since no eigenvalues are zero, the Gram matrix has full rank, the projection operators Π_k are linearly independent, and $\mathcal{P}$ is informationally complete.

The SICPOVM is also a 2-design, since it has enough elements and the global minimum of the V_2 potential is achieved. This fact also suffices to establish that $\mathcal{P}$ is informationally-complete, since the 2-design condition is essentially an operator-completeness condition.

We can now state our conjecture.

Conjecture 1 *For any dimension $d \in \mathbb{N}$, let $\{|k\rangle\}_{k=0}^{d-1}$ be an orthonormal basis for $\mathbb{C}^d$, and define*

$$\omega = \exp(2\pi i/d)\,, \qquad D_{jk} = \omega^{jk/2} \sum_{m=0}^{d-1} \omega^{jm} |k \oplus m\rangle\langle m|\,, \tag{4.9}$$

where $\oplus$ denotes addition modulo d. Then there exists a normalized $|\phi\rangle \in \mathbb{C}^d$ such that the set $\{D_{jk}|\phi\rangle\}_{j,k=1}^{d}$ is a SICPOVM $\mathcal{P}$.

Analytic solutions are known for $d = 2, 3, 8$ [97], and to this list $d = 4$ is added. Additionally, computer calculations reveal numerical solutions (with an accuracy better than 1 part in 10^8) in dimensions up to 45. Not all known SICPOVM examples have the precise group covariance described in the conjecture, so at this point it is appropriate to loosen the restraints slightly from the original goal of finding a Weyl-Heisenberg SICPOVM and consider the group covariant case generally. Thereafter the analytic and numerical results may be more easily formulated.

4.1 Group Covariance

To say that the SICPOVM $\mathcal{P}$ is group covariant is to say that there exists a group G with a d-dimensional projective unitary representation $\{U_g\}$ such that (i) $\mathcal{P}$ is invariant under any U_g, i.e., for any $|\phi_j\rangle \in \mathcal{P}$ and any U_g, $U_g|\phi_j\rangle \in \mathcal{P}$ (up to a phase), and (ii) $\{U_g\}$ acts transitively on $\mathcal{P}$, i.e., for any $|\phi_j\rangle, |\phi_k\rangle \in \mathcal{P}$, there exists U_g such that $U_g|\phi_j\rangle = |\phi_k\rangle$ (also up to a phase). Assuming group covariance simplifies the search for SICPOVMs. We simply search for a fiducial vector such that $\mathcal{P} = \{U_g|\phi\rangle\}$ is a SICPOVM (note that the transitivity property implies that the order of G must be at least d^2). To do this, we use groups such that $\{U_g|\phi\rangle\}$ is a tight frame for any normalized vector $|\phi\rangle$, and then we search for a particular vector $|\phi\rangle$ such that $|\langle\phi|U_g|\phi\rangle|^2 = 1/(d+1)$ for all $g \neq e$. All other inner products are then guaranteed to have this value due to the group action.

We suspect the case of group covariance to be general for the following reason. Consider the map $\alpha : \mathbb{S}^d \to \mathcal{B}(\mathbb{C}^d)$ that takes a normalized vector to the corresponding projector, i.e., $\alpha(|\phi_j\rangle) = |\phi_j\rangle\langle\phi_j|$. Now consider the operators

$$\sigma_j = \sqrt{\frac{d}{d-1}} \left(|\phi_j\rangle\langle\phi_j| - \frac{I}{d} \right) . \tag{4.10}$$

Being both traceless and Hermitian, these operators lie in a subspace of $\mathcal{B}(\mathbb{C}^d)$ that is isomorphic to $\mathbb{R}^{d^2-1}$; indeed, since $(\sigma_j, \sigma_j) = 1$, they all lie on the unit sphere in $\mathbb{R}^{d^2-1}$. This sphere is a generalization of the Bloch sphere for two-dimensional systems, the difference being that for $d > 2$, not all operators on the (d^2-2)-sphere are images of vectors in $\mathbb{S}^d$ under the map α. From the SICPOVM condition 4.8, one finds immediately that

$$(\sigma_j, \sigma_k) = -\frac{1}{d^2-1} \qquad \forall\, j \neq k \;. \tag{4.11}$$

This is the condition for the d^2 operators $\{\sigma_j\}$ to form a regular simplex in $\mathbb{R}^{d^2-1}$, whose automorphism group is the permutation group S_{d^2}. Given this result, some group covariance seems natural. One is tempted to think that from here it is a simple matter to establish the existence of the set $\mathcal{P}$. This is not the case, however, as working in the operator space obscures the very difficult task of determining when a given operator is the image of some element of $\mathbb{S}^d$ under the map α. In the same vein, most of the elements of the permutation group cannot be represented in this framework as unitary transformations of $\mathbb{C}^d$; thus, while we know that any G satisfying the conditions above must be a subgroup of S_{d^2}, it is not obvious which subgroups are candidates.

The outstanding choice for G is, of course, the group $\mathbb{Z}_d \times \mathbb{Z}_d$. The group's immediate usefulness here stems from the fact that, for any normalized $|\psi\rangle \in \mathbb{S}^d$,

$$S_\psi = \sum_{jk} D_{jk}|\psi\rangle\langle\psi|D_{jk}^\dagger = dI \;, \tag{4.12}$$

which we met in equation 2.30.

This property of producing a tight frame for any input state is quite general. The following argument is adapted from Proposition 3 of [161]. Any set of d^2 orthogonal unitary operators T_j, thus satisfying $\mathrm{Tr}[T_j^\dagger T_k] = d\delta_{jk}$, is a complete set for expanding operators in $\mathcal{B}(\mathbb{C}^d)$; the unitary operators $\{D_{jk}\}$ are one example of operators that satisfy this orthogonality condition. It is a simple matter to turn the completeness relation into $\sum_k T_k C T_k^\dagger = d\,\mathrm{Tr}[C]I$ for any operator C. Simply consider the inner

product of two arbitrary operators A and B. The completeness relation means that

$$(A, B) = \frac{1}{d} \sum_k (A, T_k)(T_k, B) \ . \tag{4.13}$$

Setting $A = |\phi_1\rangle\langle\phi_2|$ and $B = |\psi_1\rangle\langle\psi_2|$ (which we can do without loss of generality because such outer products span $\mathcal{B}(\mathbb{C}^d)$), we find

$$\langle\phi_1|\psi_1\rangle\langle\psi_2|\phi_2\rangle = \frac{1}{d} \sum_k \langle\phi_1|T_k|\phi_2\rangle\langle\psi_2|T_k^\dagger|\psi_1\rangle \ , \tag{4.14}$$

from which it follows that $\sum_k T_k|\phi_2\rangle\langle\psi_2|T_k^\dagger = d\,\langle\psi_2|\phi_2\rangle$, whence the result follows.

Thus the property of producing a tight frame regardless of the fiducial $|\phi_0\rangle$ is common to all groups of size d^2 whose representation operators are a complete, orthogonal set. Such groups were introduced by Knill in connection with quantum error-correcting codes and are called "nice error bases" or unitary error bases [94, 95]. Klappenecker and Rötteler have kindly detailed all such nice error bases up to dimension 10, so we can apply them to the problem at hand [91, 92, 93]. Only the nice error bases associated with the group $\mathbb{Z}_d \times \mathbb{Z}_d$ exist in every dimension, thus accounting for our focus on this group.

By considering the generalized problem of groups generating higher spherical t-designs, our problem makes contact with the deep subject of the relationship between frames, designs, and groups [75].

4.2 Analytic SICPOVMs

Now we concentrate specifically on using the group $\mathbb{Z}_d \times \mathbb{Z}_d$. Fixing the representation operators D_{jk} of this group, we can determine the set of fiducial vectors that under the group action make a SICPOVM. From this we can determine also the number of distinct SICPOVMs generated by our fixed representation. In three dimensions there are an uncountably infinite number of such covariant SICPOVMs, but in two

dimensions there are just two, and in four, 16. We write the fiducial state as

$$|\phi\rangle = \sum_k r_k e^{i\theta_k} |k\rangle \ , \tag{4.15}$$

where we can, of course, immediately choose $\theta_0 = 0$.

4.2.1 $d = 2$

The two solutions, represented as column vectors in the standard basis, are

$$\left\{ \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3+\sqrt{3}} \\ e^{i\pi/4}\sqrt{3-\sqrt{3}} \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -\sqrt{3-\sqrt{3}} \\ e^{i\pi/4}\sqrt{3+\sqrt{3}} \end{pmatrix} \right\} . \tag{4.16}$$

These have a simple interpretation on the Bloch sphere, where the nontrivial group operators are simply rotations by π about the x, y, and z axes, respectively. Then the Bloch vectors of the two fiducial states are $\pm(1,1,1)/\sqrt{3}$, and the two SICPOVM states thus formed are regular tetrahedra, each one related to the other by inversion of the Bloch vectors.

4.2.2 $d=3$

For r_0 satisfying $1/\sqrt{2} < r_0 < \sqrt{2/3}$, define

$$r_\pm(r_0) = \frac{1}{2} r_0 \pm \frac{1}{2}\sqrt{2 - 3r_0^2} \ . \tag{4.17}$$

Hence $0 < r_- \le 1/\sqrt{6} \le r_+ < 1/\sqrt{2} < r_0 \le \sqrt{2/3}$. The complete set of fiducial states, represented as column vectors in the standard basis, is then

$$\left\{ \begin{pmatrix} r_0 \\ r_+ e^{i\theta_1} \\ r_- e^{i\theta_2} \end{pmatrix}, \begin{pmatrix} \text{plus all vectors formed} \\ \text{by permuting of elements} \end{pmatrix} \middle| \, \theta_1, \theta_2 \in \left\{ \frac{\pi}{3}, \pi, \frac{5\pi}{3} \right\} \right\}$$
$$\bigcup \left\{ \begin{pmatrix} 1/\sqrt{2} \\ e^{i\theta_1}/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} \text{plus all vectors formed} \\ \text{by permuting of elements} \end{pmatrix} \middle| \, 0 \le \theta_1 < 2\pi \right\} . \tag{4.18}$$

4.2.3 $d = 4$

Now let

$$r_0 = \frac{1 - 1/\sqrt{5}}{2\sqrt{2 - \sqrt{2}}}\,, \qquad r_1 = (\sqrt{2} - 1)r_0\,, \qquad r_\pm = \frac{1}{2}\sqrt{1 + 1/\sqrt{5} \pm \sqrt{1/5 + 1/\sqrt{5}}}\,, \tag{4.19}$$

along with

$$a = \arccos\frac{2}{\sqrt{5 + \sqrt{5}}}\,, \qquad b = \arcsin\frac{2}{\sqrt{5}}\,, \tag{4.20}$$

and define the set

$$\begin{aligned}\Omega \;\equiv\; \Big\{\Big((-1)^m(a/2 + b/4) + \pi(m + 2n + 7j + 1)/4\,,\, \pi(2k + 1)/2\,,\\ (-1)^m(-a/2 + b/4) + \pi(m + 2n + 3j + 4k + 1)/4\Big)\\ \Big|\; j, k, m = 0, 1 \text{ and } n = 0, \ldots, 3\Big\}\,.\end{aligned} \tag{4.21}$$

The complete set of fiducial states, represented as column vectors in the standard basis, is now

$$\left\{\begin{pmatrix} r_0 \\ r_+ e^{i\theta_+} \\ r_1 e^{i\theta_1} \\ r_- e^{i\theta_-} \end{pmatrix}, \begin{pmatrix} r_0 \\ r_- e^{i\theta_-} \\ r_1 e^{i\theta_1} \\ r_+ e^{i\theta_+} \end{pmatrix}, \begin{pmatrix} \text{plus all vectors formed} \\ \text{by cycling of elements} \end{pmatrix} \;\Big|\; (\theta_+, \theta_1, \theta_-) \in \Omega\right\} \tag{4.22}$$

4.3 Numerical SICPOVMs

Because analytic solutions to the SICPOVM condition equation 4.8 are so few, our conjectures are based almost entirely on numerical evidence (even the $d = 4$ solution was originally inspired by close examination of numerical solutions). To find numerical instances of P, we simply minimize the second frame potential $\mathrm{Tr}[S_2^2]$ over sets of d^2 normalized vectors generated by a representation of $\mathbb{Z}_d \times \mathbb{Z}_d$ from a vector $|\phi\rangle$. It is also possible to vary independently the d^2 elements of $\mathcal{P}$, but this is much less

efficient; taking advantage of the group-covariance conjecture permits us to search a space of $O(d)$ complex parameters instead of $O(d^3)$ complex parameters.

The quantity that we minimize, $\sum_{j,k} |\langle\phi|D_{jk}|\phi\rangle|^4$, is proportional to the frame potential because of the group covariance. Since it is a quartic function of $|\phi\rangle$, we have to use numerical methods to minimize it, using either Mathematica (simpler) or C++ (much faster). The method used is an adaptive conjugate gradient method; this has the advantage of converging with exponential rapidity to a local minimum, but the disadvantage of being insensitive to global conditions. As a result, the most time-intensive portion of the computation by far is identifying one of the global minima among the many local minima.

Once the correct minimum is located, we quickly obtain $\mathcal{P}$ such that equation 4.8 is satisfied to an accuracy of 10^{-8}. The sole exception to this rule is $d = 3$ (where an exact analytic solution is known): in $d = 3$ there exists a continuously infinite family of solutions, and this degeneracy makes numerical solution difficult. For every dimension between $d = 5$ and $d = 45$, however, we have found $\mathbb{Z}_d \times \mathbb{Z}_d$-covariant solutions to within machine precision [130].

Additionally, in small dimensions, one can attempt an exhaustive search for *all* possible $\mathbb{Z}_d \times \mathbb{Z}_d$-covariant SICPOVMs, by simply running the minimization many times with differing presumptive fiducial states, tabulating all the while the distinct SICPOVM fiducial states found. Table 4.3 lists the results for the number of distinct SICPOVMs.

Finally, we have tested some of the other nice error bases tabulated by Klappenecker and Rötteler. These are also easy handled, and although not all groups were tested, at least four groups were found to generate SICPOVM sets. In the notation of the library of small groups used by GAP3, GAP4, and MAGMA, these groups are G(36,11), G(36,14), G(64,8), and G(81,9). Each of these solutions has an accuracy of 10^{-15} in the individual vector inner products. Perhaps surprisingly, many of the tabulated groups do not seem to yield group-covariant SICPOVMs.

d	#(SICPOVMs)
2	2
3	∞
4	16
5	80
6	96
7	336

Table 4.1: Number of SICPOVM sets generated by a fixed representation of the group $\mathbb{Z}_d \times \mathbb{Z}_d$ in dimensions two through seven. The infinity in dimension three is uncountable.

A rigorous proof of existence of SICPOVMs in all finite dimensions seems tantalizingly close, yet remains somehow distant. Although the numerical evidence makes very clear the relevance of the group $\mathbb{Z}_d \times \mathbb{Z}_d$, this is not definitively established. Given the apparent importance of $\mathbb{Z}_d \times \mathbb{Z}_d$, it would seem to be just a short step to some general form for an operator whose eigenvectors could be a fiducial state, but a proof by this method has not been forthcoming. For instance, in three dimensions the Fourier transform operator has an eigenvector that is a fiducial state (the one associated with the eigenvalue i), but this does not hold in general. In five dimensions a fiducial vector can be found among the degenerate eigenvectors of a particular $\mathbb{Z}_3$ subgroup of the normalizer of $\mathbb{Z}_d \times \mathbb{Z}_d$ in $SU(d)$, but there is no such subgroup at all in the normalizer for dimension seven. The group-theoretic structure of SICPOVMs is exceedingly rich, however, and ongoing efforts to understand the full automorphism group of a SICPOVM might yield insights into operators that yield fiducial states. Perhaps by here establishing the framework and providing motivating numerical results, a proof might yet be found.

4.4 SICPOVM Representations

Regardless of whether their existence can be proved rigorously, SICPOVMs appear to exist in many dimensions, so we are justified in examining their properties. The

next chapter deals with quantum cryptography in detail so here we confine ourselves to questions of representation theory and quantum state tomography.

Using the SICPOVMs we can create an analog of the coherent states and the P and Q representations that follow. For the P representation we expand a given state in terms of the SICPOVM states, whereas the Q function is the set of coefficients formed using the states, like so:

$$\rho = \sum_{jk} P_{jk}(\rho)|\phi_{jk}\rangle\langle\phi_{jk}| \,, \tag{4.23}$$

$$Q_{jk}(\rho) = \langle\phi_{jk}|\rho\,|\phi_{jk}\rangle/d \,. \tag{4.24}$$

The Q function is normalized such that it forms a probability distribution. As in the continuous, infinite-dimensional version, the P and Q are dual representations as we may find a set of operators R_{jk} such that

$$\rho = \sum_{jk} Q_{jk}(\rho)R_{jk} \,, \tag{4.25}$$

$$P_{jk}(\rho) = \mathrm{Tr}[R_{jk}\rho] \,. \tag{4.26}$$

Such a dual system of representations can, in principle, be constructed using any set of operator-complete quantum states, pure or mixed. However, the SICPOVM offers an exceedingly simple connection between the two halves of the dual representations. To see this, we will first determine R_{jk}, by using equation 4.25 in equation 4.24 or similarly for the P function. Then we find

$$\langle\phi_{jk}|R_{lm}|\phi_{jk}\rangle = d\,\delta_{jl}\delta_{km} \,. \tag{4.27}$$

Since the $|\phi_{jk}\rangle$ are operator-complete, these equations determine the R_{jk}. Simply guessing that $R_{jk} = a|\phi_{jk}\rangle\langle\phi_{jk}| + bI$, the constants a and b are determined by the previous equation to be such that

$$R_{jk} = (d{+}1)|\phi_{jk}\rangle\langle\phi_{jk}| - I \,. \tag{4.28}$$

This immediately implies an essentially trivial relationship between the Q and P functions. Using this form of R_{jk} in equation 4.26 we have

$$P_{jk}(\rho) = (d{+}1)\langle\phi_{jk}|\rho|\phi_{jk}\rangle - 1 = d(d{+}1)Q_{jk}(\rho) - 1. \tag{4.29}$$

Thus the P function is simply a shifted and rescaled version of the Q function, a much simplified version of the Gaussian convolution relation in infinite dimensions.

Practically any operator can be used in conjunction with the displacement operators to create a set of "coherent states", to stretch the term. For some operator H define the set of displaced versions as $H_{jk} = D_{jk} H D_{jk}^\dagger$. For H to generate a linearly-independent set of operators, we must ensure that the Gram matrix has full rank, as we did with the SICPOVM. Owing to the structure of the displacement operators the Gram matrix is block-circulant with circulant blocks, each of size d. Two Fourier transforms suffice to bring it into diagonal form, so that the eigenvalues may be expressed as

$$\lambda_{jk} \propto \frac{1}{d} \sum_{lm} \omega^{jm-kl} \mathrm{Tr}[D_{lm} H D_{lm}^\dagger H], \tag{4.30}$$

where again $\omega = \exp[2\pi i/d]$. From the explicit form of the D_{jk} we can calculate that

$$\frac{1}{d} \sum_{s,t=0}^{d-1} \omega^{jt-ks} D_{st} D_{lm} D_{st}^\dagger = d\, \delta_{jl} \delta_{km} D_{lm}. \tag{4.31}$$

Now writing H in terms of the characteristic function the eigenvalue equation becomes

$$\begin{aligned} \lambda_{jk} &\propto \frac{1}{d} \sum_{lmst} \omega^{jm-kl} \mathrm{Tr}[D_{lm} D_{st}^\dagger D_{lm}^\dagger H] \mathrm{Tr}[H^\dagger D_{st}] && (4.32) \\ &\propto |\mathrm{Tr}[D_{jk}^\dagger H]|^2. && (4.33) \end{aligned}$$

Thus to ensure full rank it is enough to demand that H not be orthogonal to any of the displacement operators, i.e., that when expanded in terms of the displacement operators, there be no zero terms.

If we pick H to be a positive-semidefinite operator with non-zero trace, we can emulate the Q and P functions by defining

$$\begin{aligned} Q_{jk}^H(\rho) &= \mathrm{Tr}[H_{jk}\rho]/d\mathrm{Tr}[H] && (4.34) \\ \rho &= \sum_{jk} P_{jk}^H(\rho) H_{jk} && (4.35) \end{aligned}$$

This definition ensures that the set $\{D_{jk}HD_{jk}^{\dagger}/\mathrm{Tr}[H]\}$ forms a POVM, so that the Q function is again a probability distribution and P a quasi-distribution. The Gram matrix and its inverse provide the means to go between the two, since we may simply form the Q function according to equation 4.34 from the form of the state given in equation 4.35. In this way we obtain

$$Q_{jk}^{H}(\rho) = \sum_{lm} \frac{\mathrm{Tr}[H_{jk}H_{lm}]}{d\,\mathrm{Tr}[H]} P_{lm}^{H}(\rho). \tag{4.36}$$

Defining a matrix $\mathcal{M}_{jk;lm} = \mathrm{Tr}[H_{jk}H_{lm}]/d\,\mathrm{Tr}[H]$ we may express the relationship as

$$Q(\rho) = \mathcal{M}P(\rho) \qquad P(\rho) = \mathcal{M}^{-1}Q(\rho). \tag{4.37}$$

The point of this derivation is to show that in general the connection between Q and P can be complicated to compute due to the inverse of $\mathcal{M}$. On the other hand, this connection is trivial for the SICPOVM coherent states, a fact which will be important when considering quantum state tomography.

4.4.1 Wigner Functions

It is also possible to link the P and Q functions to the displacement operators and a Wigner function, though not with the elegance found in infinite dimensions. Fourier transforms of the displacement operators create Hermitian operators suitable for use in the Wigner function, but only in odd dimensions are these operators complete. For even dimensions it has been long known that the finite-dimensional Wigner function must be extended to four times as many values to ensure completeness. While using the Fourier transform in odd dimensions results in a complete set of operators, in fact an orthonormal operator basis, the parity operator is not among them. Generally it is impossible to import all of the nice features from the infinite-dimensional case due to the nature of the projective representation. In what follows, only odd dimensions will be considered, for only there do such issues not cloud the connection between the various representations.

For convenience, when working in odd dimensions, we may reset the range of the indices (j,k) of the displacement operator to run from $-n=(1-d)/2$ to n, so that they are located symmetrically about zero. All the properties of the displacement operator developed in chapter 1 remain intact. Then we can define the Wigner function by starting with the parity operator and displacing it about

$$\Pi = \sum_{k=-n}^{n} |-k\rangle\langle k|, \tag{4.38}$$

$$\Delta_{jk} = D_{jk}\Pi D_{jk}^{\dagger} = \sum_{m} \omega^{-2jm}|k-m\rangle\langle k+m|. \tag{4.39}$$

Though Hermitian, these operators aren't positive, having eigenvalues ± 1, $(d+1)/2$ positive and the remainder negative. They are, however, orthogonal:

$$\mathrm{Tr}[\Delta_{jk}\Delta_{lm}] = \mathrm{Tr}[D_{jk}\Pi D_{jk}^{\dagger} D_{lm}\Pi D_{lm}^{\dagger}] = \mathrm{Tr}[\Delta_{l-j,m-k}\Pi] = d\,\delta_{jl}\delta_{km}. \tag{4.40}$$

Arbitrary operators may be easily expressed in terms of the "point operators" Δ_{jk}, yielding the Wigner function

$$W_{jk}(\rho) = \mathrm{Tr}[\Delta_{jk}\rho]/d \qquad \rho = \sum_{jk} W_{jk}\Delta_{jk}. \tag{4.41}$$

Note that unlike the P and Q representations, the Wigner function is self-dual since the operators Δ_{jk} both generate the coefficients and form the expansion basis. Apart from this, the salient feature of the Wigner function is that although it is a quasi-distribution, its marginals are proper probability distributions. Summing over the first index in Δ_{jk}, we obtain

$$\sum_{j} \Delta_{jk} = \sum_{jm} \omega^{-2jm}|k-m\rangle\langle k+m| = d|k\rangle\langle k|, \tag{4.42}$$

proportional to the projection onto the position state k. To sum on the second index, first note that the parity operator is invariant under Fourier transform: $\Pi = F\Pi F^{\dagger}$. Then we may proceed, finding

$$\sum_{k} \Delta_{jk} = \sum_{k} D_{jk}\Pi D_{jk}^{\dagger} = \sum_{k} FF^{\dagger}D_{jk}FF^{\dagger}\Pi FF^{\dagger}D_{jk}^{\dagger}FF^{\dagger} \tag{4.43}$$

$$= \sum_{k} FD_{-k,j}\Pi D_{-k,j}^{\dagger}F^{\dagger} = \sum_{k} F\Delta_{-k,j}F^{\dagger} = d\,F|j\rangle\langle j|F^{\dagger}, \tag{4.44}$$

the rescaled projector onto the momentum state j. In terms of the Wigner function, then,

$$\sum_j W_{jk}(\rho) = \langle q_k|\rho|q_k\rangle \qquad \sum_k W_{jk}(\rho) = \langle p_j|\rho|p_j\rangle. \tag{4.45}$$

Here $|q_k\rangle$ is the state $|k\rangle$ defined earlier. Since confusion can arise between the position and momentum bases, q serves to label position and p momentum.

In odd dimensions the Wigner function point operators exhibit an interesting connection to a version of SICPOVM. As the parity operator has trace equal to unity, only its non-positivity stops it from being a POVM. (Note that in even dimensions the parity operator is traceless, so the following construction would not work at all.) This is easy to fix, however, since it has only two distinct eigenvalues. The point operators can be "boosted" to a POVM, simply by combining the parity operator or its negative with enough identity operator to make the negative eigenvalues zero. Clearly this can be done in two different ways:

$$E_{jk} = \frac{I + \Delta_{jk}}{d(d+1)} \qquad F_{jk} = \frac{I - \Delta_{jk}}{d(d-1)}. \tag{4.46}$$

The first is rank $(d+1)/2$ and the second $(d-1)/2$. The interesting bit now is that both E_{jk} and F_{jk} form higher-rank SICPOVMs, for the inner products of elements obey

$$\mathrm{Tr}[E_{jk}E_{lm}] = \frac{d+2}{d^2(d+1)^2} \qquad \mathrm{Tr}[F_{jk}F_{lm}] = \frac{d-2}{d^2(d-1)^2} \qquad (j,k) \neq (l,m). \tag{4.47}$$

These two higher-rank SICPOVMs can be used to construct representations in just the same way as the standard SICPOVM.

4.5 Tomography

Tomography with the SICPOVM is quite simple, and works for quantum states as well as measurements and processes. This ability follows from the completeness of the SICPOVM states, so there is no doubt of its efficacy in principle. Here, however,

we elucidate the precise method by which tomography is performed, starting with quantum state tomography.

Imagine an experimenter can repeatedly perform quantum state preparation and wishes to check that the quantum state prepared is indeed the state desired. In this case, quantum state tomography is called for. Quantum states are not observable themselves, so there is no measurement the experimenter can perform on a single system whose outcomes label the possible states; state tomography is inherently statistical. Instead, the experimenter can perform a *tomographically-complete* or informationally-complete measurement (sets of measurements can be amalgamated into a single POVM) on many iterations of the same state from whose statistics the state can be inferred.

An informationally-complete measurement was said to be one comprised of elements spanning the space of operators. In the previous section we examined how to represent any given operator in terms of such a set, but that the set made up a measurement was incidental. Here it is the salient feature, for by repeated measurement a good guess can be made of the underlying probability distribution of outcomes, and this in turn uniquely specifies the quantum state. In terms of the various representations, the statistics of the measurement realize the Q function.

The great advantage of the SICPOVM should now be apparent: converting the measurement statistics into the form of the state itself is trivial. After much data has been collected, this can be arranged into the Q function which can then simply be converted into the P function, yielding the state itself in the next step. Thus we may write the inferred state in terms of the frequencies of the various outcomes as

$$\rho_*(f_{jk}) = \sum_{jk} [d(d+1)f_{jk} - 1]|\phi_{jk}\rangle\langle\phi_{jk}|, \tag{4.48}$$

where the asterisk denotes the state obtained from measurement results. For more general "coherent states" as described in the previous section, two difficulties arise in converting the measured data into quantum state form. First, if an eigenvalue of the Gram matrix is near zero, its inverse will obviously be very large. If the data approx-

imate the corresponding eigenvector, the variance will be amplified relative to other possible sets of data, meaning that the measurement will be far more sensitive to fluctuations in the corresponding regions of the operator space. State reconstruction becomes unstable in such regions. The SICPOVM avoids this problem as there are no such strong singularities. In contrast, the SICPOVM is afflicted with the second problem, as are all measurements: varying sensitivity of the data itself. Even with a stable state reconstruction algorithm, every measurement has an inbuilt variable sensitivity *of the data* due to the arrangement of the measurement operators. For instance if the measured state is near one of the SICPOVM states then the variance will be low, as this outcome will be more likely. Similarly for states which are orthogonal to a particular SICPOVM state. However, in between these two regions the variance of the data increases and the measurement is less sensitive in these regions. The high symmetry of the SICPOVM prevents wild swings in this variance, but in general one could crowd all the measurement operators near one "corner" of the operator space, spreading them just widely enough to span the space and thereby lowering the average sensitivity.

The SICPOVM states can also be enlisted to verify the correct functioning of a quantum measurement. In this problem the experimenter constructs a device corresponding to a particular POVM and wishes to check that it is indeed the desired measurement. Essentially this is the dual of the previous problem. By preparing and measuring the SICPOVM states repeatedly the experimenter collects statistics on each outcome given the various inputs. Let f^i_{jk} be the frequency of outcome E_i given the input ρ_{jk}. By collecting enough data it will be true that

$$f^i_{jk} \approx p_{i|jk} = \mathrm{Tr}[E_i \rho_{jk}] = d\, Q_{jk}(E_i). \tag{4.49}$$

Then it is a simple matter to convert this to the P function and obtain for the form of the ith outcome operator

$$E_i = \sum_{jk} [(d+1) f^i_{jk} - 1] \rho_{jk}. \tag{4.50}$$

Note that the form has changed slightly from the case of state tomography because the sum over j, k is no longer the sum over all values in a probability distribution.

Finally, quantum dynamics can be investigated, a method called quantum process tomography. As is the case with all tomographic methods, this scheme of process tomography is only warranted when little is known about the process. Otherwise other, more efficient, processes may be used. In general, this inefficiency precludes a quantum computer from being very effective at simulating other physical systems and providing a method to determine the dynamical rules. It is another matter altogether as to whether the quantum computer will be widely useful if we wish to simulate a physical system for the purposes of extracting a *particular* dynamical parameter.

Let us name the dynamical process $\mathcal{G}$ such that $\rho' = \mathcal{G}(\rho)$. If we think of $\mathcal{G}$ as a sort of black box, we may input SICPOVM states and perform the SICPOVM measurement at the output. Thus we would collect frequency statistics of the form $f_{jk;lm} = \mathrm{Tr}[\rho_{jk}\mathcal{G}(\rho_{lm})]/d$. We can also use the SICPOVM as a basis in which to represent the action of the process. Using the R_{jk} is simpler, yielding

$$\mathcal{G}(\rho) = \sum_{jklm} \mathcal{G}_{jk;lm} R_{jk} \mathrm{Tr}[\rho R_{lm}]. \tag{4.51}$$

Now the statistics may be used directly to find that

$$\mathcal{G}_{jk;lm} = f_{jk;lm}/d. \tag{4.52}$$

4.6 Quantum Theory as a Probability Theory

Putting all three pieces together, we may place quantum mechanics in a purely classical probabilistic framework, eschewing the operator representation. The most convenient method begins by promoting the quantum state to the Q function. Then, just as measurement probabilities may be obtained as a linear function on the quantum state, we may convert the measurement operators into linear functions of the Q

function. We'll need the P function so that

$$p_i = \mathrm{Tr}[\rho E_i] = \sum_{jklm} \mathrm{Tr}[P_{jk}(E_i) Q_{lm}(\rho) \rho_{jk} R_{lm}] = d \sum_{jk} P_{jk}(E_i) Q_{jk}(\rho). \tag{4.53}$$

Now quantum states and measurement are each elements of a real-valued vector space of dimension d^2. The analog of quantum states are probability distributions, though not all distributions correspond to quantum states.

In this framework quantum dynamical maps become stochastic maps on the set of allowed distributions. Given a particular map $\mathcal{G}$, the associated stochastic map has the form

$$\mathcal{G} \to G_{jk;lm} = \mathrm{Tr}[\rho_{jk} \mathcal{G}(R_{lm})]/d, \tag{4.54}$$

which ensures that

$$\rho' = \mathcal{G}(\rho) \;\Rightarrow\; Q_{jk}(\rho') = \sum_{lm} G_{jk;lm} Q_{lm}(\rho). \tag{4.55}$$

These three pieces — quantum states, measurements, and dynamics — are the essential features of the theory, and with the aid of the SICPOVM we can work in a real space in which quantum states correspond to probability distributions if we desire. The probabilities are distributions over a future measurement using the SICPOVM, rather than distributions over physical variables as is the case in classical mechanics. Such a presentation in a purely probabilistic framework blurs the line between quantum mechanics and probability theory, as intended, for quantum mechanics is quite rightly a modified version of classical probability theory in which the underlying quantities are non-commutative.

4.7 The de Finetti Theorem

There is a somewhat thorny issue of principle lurking in the background of the tomographic program. Strictly speaking, if the quantum state is thought to be akin to a probability distribution, it is nonsensical to consider an "unknown" quantum

state. Probability distributions and quantum states alike are assigned by observers on the basis of facts already known and are thus never themselves unknown. For example, if nothing or virtually nothing is known about a given physical system, an observer will want to describe it by a highly mixed state. Little will be known about various properties of the system—its energy, its spin or momentum, etc.—yet the state itself is perfectly well-known.

However, probabilities of probabilities are used all the time, and fortunatly the SICPOVM can help put this concept on a firm foundational footing in the quantum case. For now, take the canonical example, determining if a coin is unbiased or not. Initially nothing is known about the bias of the coin, so the sensible distribution of the bias is uniform. As the coin is tossed and results accumulated, our probability distribution changes according to Bayes' rule. When n heads have been observed in N trials, the probability distribution of heads on the next toss will have mean value $(n+1)/(N+2)$. This is Laplace's rule of succession. All along we've used the probability of the bias of the coin, but the bias of the coin is itself a probability. This is nonsense logically even if it works well in practice.

Fortunately there is a solution to this dilemma, first outlined by de Finetti. Certainly the experimenter would not object if we point out that the many measurements performed are invariant under permutation. That is, the order of the measurements doesn't matter. If we imagine that the experimenter could prepare arbitrarily many instances of the system, the combined quantum state of any number of instances is called *exchangeable*. We should consider this global system and its quantum state rather than the individual systems since the exchangeability is one fact that we know about the setup. However, as de Finetti first showed for classical probability distributions, an exchangeable distribution can be *thought of* as a distribution over independent and identically-distributed single-trial probability distributions: a probability of probabilities. That is, suppose $p_N(x_1, x_2, \ldots, x_N)$ is the joint distribution of N runs of the experiment, and $p_1(x_j)$ is a single-trial distribution. Then de Finetti's theorem asserts that whenever p_N is exchangeable, there exists a generating function

$P(p_1) > 0$ such that

$$p_N(x_1, x_2, \ldots, x_N) = \int \mathrm{d}p_1 P(p_1) \prod_{j=1}^{N} p_1(x_j) \tag{4.56}$$

$$1 = \int \mathrm{d}p_1 P(p_1) \tag{4.57}$$

Since the generating function is normalized and positive, we may think of it as the probability of the single-trial probability p_1, a convenient fiction. Thus this concept is rehabilitated in an operational sense.

For quantum systems the theorem is entirely analogous. Let ρ_N be an exchangeable quantum state of N systems, a set of systems invariant under permutations and which could have involved in principle any number N of systems. Then there exists a positive, normalized generating function $P(\rho)$ on single-system density operators ρ such that

$$\rho_N = \int \mathrm{d}\rho \, P(\rho) \, \rho \otimes \rho \otimes \cdots \otimes \rho = \int \mathrm{d}\rho \, P(\rho) \, \rho^{\otimes N}. \tag{4.58}$$

Again we may interpret $P(\rho)$ as the probability of the state ρ. Now it is again clear how tomography works logically. Each run of the experiment yields an outcome which updates the generating function via Bayes' rule. With enough data the generating function will become sufficiently peaked as to indicate a particular quantum state.

If we take the classical de Finetti theorem as given, the SICPOVM immediately allows us to prove the quantum version by simply reducing it to the classical case [32]. First, "measure" ρ_N with N copies of the SICPOVM. This produces a probability distribution p_N, which is exchangeable if ρ_N was. Next, apply the de Finetti theorem to convert this distribution into a convex combination of i.i.d. single-system probabilities. The last step is to back-track to the operator setting, achieved by noting that p_N is essentially a multipartite Q function which can be converted into a density operator via the P function. This example shows the power the SICPOVM has in rendering quantum mechanics in probabilistic terms: theorems from probability theory can often be directed imported to the quantum setting.

In this chapter we have explored the structure of SICPOVMs to find numerical

examples and seen that they are not only elegant, but useful. Such an ensemble bridges the gap between quantum mechanics and probability theory, showing just how similar the two are. More importantly, however, doing so exposes the differences, and by making them more plain, perhaps a deeper understanding of why they are different will come to light.

Part III

Applications: Quantum Cryptography

Chapter 5

Cryptography Old and New

Several tools based on frames have been added to the quantum information theory toolbox in the previous two chapters. Now we shall focus attention on using these tools to solve practical problems which originate outside the field, specifically those stemming from cryptography. Frames, in particular equiangular spherical codes, are well-suited to this task, though this use differs somewhat from their original use as classical error-correcting codes. In the setting of classical coding, all ESC codewords are perfectly distinguishable in principle and their wide spacing ensures that they are more impervious to noise and less likely to be incorrectly decoded. As quantum states, however, the codewords are not so readily distinguishable. This difficulty can be made an asset, as their indistinguishability translates into an ability to detect tampering through the inevitable noise that tampering generates. Such tamper-evidence is naturally useful in transmitting messages which can't be easily read by anyone but the intended recipient.

This chapter examines spherical code quantum key distribution according to the following trajectory. First, a background of classical cryptography is given, ranging from ancient uses to modern developments. The possibility of using quantum mechanics for new, unbeatable schemes is introduced by examining the first quantum key distribution scheme, called BB84 after its inventors Bennett and Brassard, in

section 5.2. General schemes are considered in section 5.3, before taking up the case of spherical codes in section 5.4. Spherical code schemes offer several advantages over protocols using collections of orthogonal bases, in particular a wider range of protocols, higher eavesdropping tolerance, and a streamlined key creation process. Finally, with an eye toward practical implementation, section 5.5 considers qubit-based protocols.

5.1 Background: Classical Cryptography

Cryptography is an ancient subject: the oldest known encrypted text is an inscription carved on the tomb of Khumhotep II in 1900 BCE, using unusual hieroglyphic symbols as a partial substitution code. By the sixteenth century BCE Assyrian merchants used flat stones carved with symbols for identification, a use of digital signatures. Steganography, the practice of hiding data, was also in widespread use, for instance by carving a message in a piece of wood used as a wax plate. Wax melted over the message obscured it until later removed. In the fifth century BCE Spartans used a staff of wood called the *scytale* (rhymes with Italy) to encrypt military messages. A strip of parchment was wrapped around the staff and the message written down the length. The circumference of the staff was, in effect, a secret key, for without knowing the correct circumference the message was difficult to read [89].

Along with methods of creating cryptographic systems came methods of breaking them, cryptanalysis. As early as 300 BCE the *arthasastra*, an Indian manual on the methods of statecraft, discussed the importance and methods of cryptanalysis. Abu Yahmadi discussed the plaintext attack method of breaking codes in the first book devoted to the subject, written around 800 CE. Yahmadi solved a Byzantine cryptographic puzzle written in Greek by guessing at the form of the beginning of the text. Shortly thereafter in the 9th century the 'philosopher of the Arabs' al-Kindi wrote a treatise entitled *A Manuscript on Deciphering Cryptographic Messages*, detailing the use of frequency analysis to decipher messages [6]. By examining the

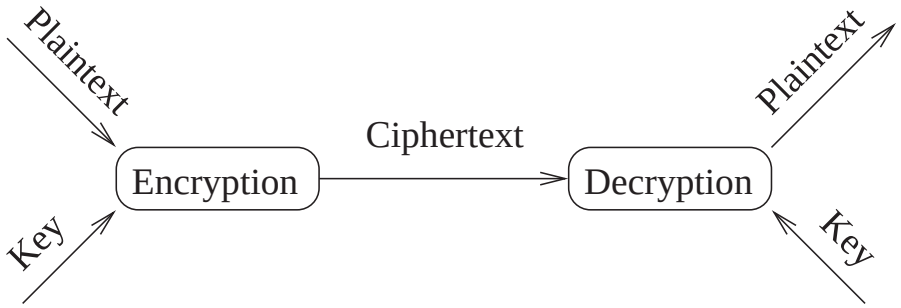

Figure 5.1: Schematic diagram showing encryption and decryption with a shared private key.

average frequency of letters used in a language one may look for correspondingly frequent symbols in the encrypted message. The British automated these and other plaintext attack methods and used them over a millennium later to break the German Enigma code during World War II.

With the modern advent of the field of information theory, cryptographic methods could be rigorously analyzed. Formally, the goal is to take a plaintext message m and produce an encrypted version, called the ciphertext c, so that only chosen parties may invert the process and read the message. This is accomplished using a secret key k to produce c from m by way of the encryption algorithm E_k: $c = E_k(m)$ and to decrypt by way of the decryption algorithm D_k. Whoever has the key may decrypt the ciphertext, but without it the message isn't easily readable. The strategy is to give each of the chosen parties the key and keep it from everyone else. Figure 5.1 depicts this setup. To be certain of security against attack, cryptographic protocols follow Kerckhoffs' principle [90] that security reside entirely in the key.[1] This way, should the encryption and decryption algorithms be known to illegitimate parties, security will survive.

In 1949 Shannon proved that the only unconditionally secure cryptographic code uses a secret key which is as long as the message [141]. A secure cryptosystem creates

[1] In modern parlance this is the argument against "security through obscurity."

a ciphertext that reveals nothing of the plaintext, so that the probability of the plaintext equals the probability of the plaintext given the ciphertext, $p(m) = p(m|c)$. Similarly, since possession of c and k leads deterministically to m, the uncertainty of the distribution of messages m cannot be less than that of the possible keys, given a particular ciphertext. Shannon formalized this statement, proving

$$H(K) \geq H(M), \tag{5.1}$$

where K is the random variable corresponding to the possible keys, M the possible messages. Since we can compress all the messages m down to a fraction $H(M)$ by Shannon's noiseless channel coding theorem, and similarly for the keys k, this implies the key cannot be shorter than the message.

Invented in 1917, the eponymous Vernam cipher is by the Shannon criterion the optimal, unconditionally secure method of data encryption [154]. Vernam created a machine capable of generating a pseudo-random string called a one-time pad for use as a secret key. The message is first converted to binary, say, and added to the secret key. To decode, one simply needs to subtract the key from the encrypted message, recovering immediately the actual, plaintext message. As long as the key is only used once, the encrypted message is essentially unbreakable, as it only indicates the length of the message. Hence the name "one-time pad".

The one-time pad is unwieldy, however, as it requires long keys. This presents a difficulty to separated parties who wish to exchange many messages. Either they must meet at some point and establish new keys, carry all the keys they will ever need from their only encounter, or have some method of creating new keys at a distance. Absent a private means of communicating, creating keys at a distance seems to imply a catch-22. In order to agree on a secret key over a public channel, encryption should be used, but this requires the key needed in the first place.

To get around this problem of key distribution, in 1976 Diffie, Hellman, and Merkle developed a method of key distribution based on "one-way" functions, functions easy to implement but difficult to invert [54]. Before examining their protocol,

consider the following insecure scheme, which demonstrates the principle of public key distribution. Two parties wishing to establish a key—traditionally named Alice and Bob—may do so by each starting with a copy of a public random string γ. Alice adds a privately generated random string α to her copy, as does Bob for his copy. Then they exchange their strings $\alpha+\gamma$ and $\beta+\gamma$ publicly and again add their private strings. Now they are in possession of the same string $\alpha+\beta+\gamma$ which may be used as a key. The protocol is obviously not secure since an eavesdropper, Eve, may simply subtract the known random string γ from the publicly exchanged copies to obtain α and β, and thus the key. But if the addition operations were instead commuting *one-way* functions, it would be difficult for Eve to do this, leaving Alice and Bob with a secure key.

Such one-way functions abound in number theory. For instance, suppose that Alice and Bob agree on numbers g and p such that p is prime and g generates all integers up to $p-1$ under exponentiation modulo p (g is primitive with respect to p). Then the pair is analogous to the public random string γ, for Alice and Bob may each choose a random number, Alice a and Bob b, compute $\alpha = g^a \bmod p$ and $\beta = g^b \bmod p$, and exchange α and β. Now if Alice computes $\beta^a \bmod p$ and Bob $\alpha^b \bmod p$, they're each guaranteed to have $g^{ab} \bmod p$. Eve, however, can't simply compute this value from α, β, g, and p since determining the discrete logarithms a and b is difficult. This is the Diffie-Hellman key exchange protocol.

Public key exchange protocols solve one problem but create another: authentication. Alice and Bob need to be certain they are communicating with each other, lest one inadvertently share a secret key with Eve and then broadcast the ciphertext. This would seem to destroy all hopes for key exchange, but the problem of message authentication is not as daunting as key exchange, since secrecy is no longer the goal. Still, it should be clear that each party must know something unique and unforgeable about the other party for identification purposes. A shared secret key accomplishes this task, and, importantly, for authentication purposes the key needn't be terribly long. Using a key-dependent one-way hash function Alice and Bob can

"sign" their messages with the shared key. A hash function of a long string is a shorter "fingerprint" of the string such that small changes to the input string generate largely different hashes [135]. This feature, combined with the one-way and the key-dependence aspects, prohibits Eve from forging the signature.

Note that inverting the one-way function in Diffie-Hellman key exchange is referred to as "difficult", but not "impossible". Precisely for this reason the protocol is not unconditionally secure, as there is no absolute guarantee that Eve does not know how to invert the function. In practice, though, no efficient method of finding discrete logarithms is known. The best known method, the number field sieve, requires a running time superpolynomial in the prime p, roughly $O(\exp[n_p^{1/3}(\ln n_p)^{2/3}])$ steps where n_p is the number of bits in the binary expansion of p.

Computational security concerns aside, Diffie-Hellman suffers one other drawback: the requirement of two-way communication between the parties. Such interaction may not always be feasible. To solve this remaining problem, Diffie, Hellman, and Merkle invented public-key cryptography which uses different keys for encryption and decryption, obviating the problem of key exchange. The encryption key is made public, while the decryption key is kept secret. Anyone wishing to encrypt data such that only the holder of the decryption key may decrypt it is free to do so at his or her leisure. The simultaneous ease of decryption for legitimate users and difficulty presented to the eavesdropper relies on a "trapdoor one-way" function which is one-way without knowing the secret trapdoor, but two-way otherwise. Here, encryption using the public key is simple, but the decryption phase is difficult without knowing the private key, the secret trapdoor.

The premier example of a public key cryptosystem is the RSA system, invented in 1977 by Rivest, Shamir, and Adleman [131]. RSA again uses modular exponentiation as the one-way function, but instead of performing operations modulo a prime number, n is the product of two large primes. Knowledge of the prime factors is the secret trapdoor. To see how this works, let Alice choose large primes p and q. She then makes public $n = pq$ and some number e, the public key, which is rel-

atively prime to $(p-1)(q-1)$. Bob uses the function $E(m) = c = m^e \bmod n$ to encipher the message m. To decipher, Alice computes d, the private key, such that $d = e^{-1} \bmod (p-1)(q-1)$, in other words $ed = 1 \bmod (p-1)(q-1)$. Then Alice recovers m with the function $D(c) = c^d \bmod n$, since

$$D(E(m)) = m^{ed} \bmod n = m^{k(p-1)(q-1)+1} \bmod n = m. \tag{5.2}$$

The last step follows from Euler's totient theorem, which states that for a relatively prime to n, $a^{\phi(n)} \bmod n = 1$. Here $\phi(n)$ is the totient function, the number of numbers less than n relatively prime to n; for prime p, $\phi(p) = p-1$.

RSA is as difficult to crack as the Diffie-Hellman key exchange, for the best known attack requires factoring n, and the algorithm for doing so is again the number field sieve. However, since the public key is likely to be used many times, it is vulnerable to attack, akin to reusing a one-time pad. In practice one may simply combine RSA with a one-time pad by generating a random key, encrypting it with the public key of the intended recipient, and sending it on its way. Now reusing the public key is less insecure, as it is only used to encrypt random strings. This is the basis for many internet security protocols, including the secure sockets layer (SSL) used in secure hypertext transfer protocol (https) transactions [139] and the secure shell (SSH) remote terminal [83].

Quantum mechanics has the possibility to completely change this security landscape. No known efficient *classical* algorithm exists for factoring or finding discrete logarithms, but the quantum information theory toolbox does contain an efficient factoring tool for use on a quantum computer. Invented by Shor in 1994 [142], this algorithm requires only $\mathrm{O}(n^3)$ steps to factor an n-bit number, rendering public-key cryptosystems based on the difficulty of factoring and finding discrete logarithms useless. Not all such cryptosystems are based on these problems, though the majority and most widely-used are; the McEliece protocol for instance uses the ease of encoding but difficulty of decoding a general linear error-correcting code to provide a one-way function.

The Shor algorithm does require a quantum computer, which may be more difficult to realize than an efficient classical factoring algorithm, so for the moment public-key cryptography is practically secure. However, the damage is done in principle. Fortunately, just as the tool for breaking widely-used cryptographic protocols exists in the toolbox, so does the fix: *quantum cryptography*. Quantum key distribution is the main application of quantum cryptography, a marriage of the Vernam cipher with quantum communication, uniting the unconditional security of the former with a key-exchange ability provided by the latter. In this sense, quantum key distribution is a key exchange protocol, or more properly a key expansion protocol, since it too requires authentication.

The advantage gained by using quantum systems is the ability to detect tampering and prevent copying, as first realized by Wiesner in the early 1970s [163]. His ideas for unforgeable bank notes, called quantum money, were unfortunately not followed up until 1984 when Charles Bennett and Gilles Brassard adapted them to key distribution [12]. Their scheme, known as BB84, is worth examining in detail before pressing forward with the main focus of this chapter, quantum key distribution based on equiangular spherical codes.

5.2 The BB84 Protocol

The BB84 protocol is designed to solve exactly the same problems that Diffie-Hellman key exchange solves, only in an unconditionally-secure fashion. Our two separated parties, Alice and Bob, wish to agree on a secret key through public communication channels. We assume that the eavesdropper Eve, meanwhile, can monitor all classical and quantum communication, as well as perform unlimited (quantum) computation in order to break the scheme. Alice and Bob already share a short key which provides the means to authenticate messages from each other, preventing Eve from "spoofing" messages from either party. However, this key is too short to encrypt data they wish

to share but keep secret. If they can successfully use the public channels to create a longer key, then they can securely send the actual message.

Key creation is accomplished in the following manner. The public communication channels they each have access to are photon-transmitting channels, perhaps free-space or a fiber optic cable. For now we'll think of the channels as noiseless. Focusing on the polarization modes of the photons, Alice and Bob agree publicly on two bases to use and an encoding scheme within each one. Traditionally they choose the horizontal/vertical basis (+) and the 45/135 degree linear polarization basis (×), agreeing that vertical or 45-degree corresponds to `0` and the others to `1`. In the spin-1/2 language we may describe this as using the σ_x or σ_z bases, with spin up in either basis corresponding to `0` and down to `1`, which we will do from now on. To make matters concrete, we are letting vertical polarization correspond to spin up along z and 45 degrees to spin up along x. Alice randomly chooses a binary string α and an encoding string ε, each of the same length N. She then encodes α into a sequence of photon polarization states, using the encoding string to determine the basis used, `0` corresponding to z and `1` to x, say. For instance if the string α were `010011` and the encoding string `100110`, Alice would prepare the state $|\uparrow_x\rangle|\downarrow_z\rangle|\uparrow_z\rangle|\uparrow_x\rangle|\downarrow_x\rangle|\downarrow_z\rangle$.

Upon receipt of the state, Bob randomly picks a decoding string ξ and measures the polarizations in the corresponding bases. Naturally, ξ doesn't perfectly coincide with ε, but they agree on average half the time. When Bob measures in the same basis Alice used to prepare the signal, he obtains the bit she encoded. If not, he obtains either outcome with equal chance, since the measurement of a spin-z eigenstate along x is random and vice versa. In all, the measurement process leaves Bob with a string β. Following the example above, if the decoding string were `101000`, then the string β would be any of the strings `01???1`, where `?` denotes a random bit. This is depicted in figure 5.2. To remove the random outcomes, Bob then announces the bases he used to measure, i.e. the string ξ, and Alice replies with the encoding string ε. Whenever these strings disagree, they discard the corresponding bit from their strings α and β, a process called *sifting*. A simple way to perform sifting in this case

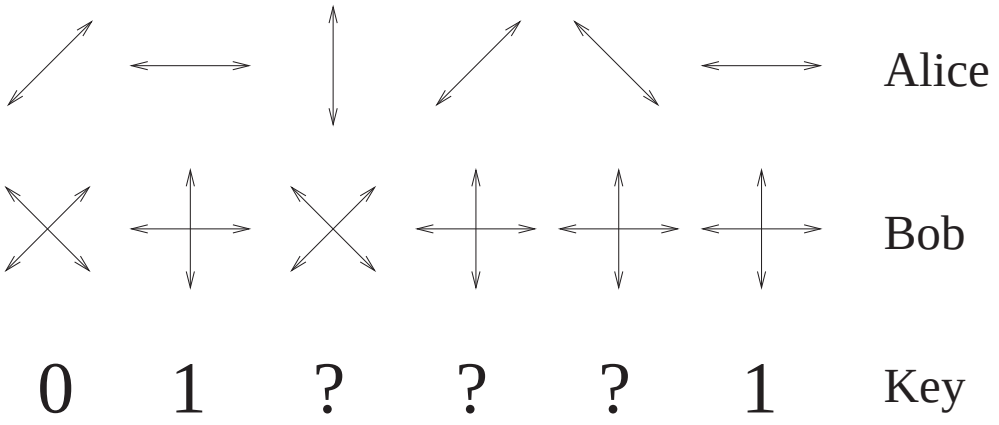

Figure 5.2: The BB84 protocol example given in the text. The first line corresponds to Alice's encoding into polarization states, and the second to Bob's choice of measurement. The third line gives the resulting key string, in which half of the values are discarded due different choices of encoding and decoding bases.

is to add the strings ξ and ε modulo 2 and keep only the bits corresponding to 0. For a noiseless channel, sifting leaves them each with a copy of the key κ, a random subset of the randomly-generated string α.

If Eve listens to their basis announcement messages, she knows only the encoding and decoding strings. Since these have nothing to do with the key κ, she learns nothing from the public communication, and Alice and Bob may safely encrypt their data. The quantum channel Alice uses to transmit the photons to Bob is also public, so Eve may try her luck listening to the quantum channel as well. However, in listening to the quantum channel, she will introduce noise, which Alice and Bob will be able to detect by comparing a portion of their (now slightly different) key strings κ_a and κ_b. The more information Eve wishes to gain from the photons, the more disturbance she will incur, all due to the fact that Alice encoded her string α into two different bases. The error rate enables Alice and Bob to infer how much Eve knows about their key, at which point they may use error-correction and a procedure called *privacy amplification* to create a shortened key, which has no errors and about which Eve is ignorant. Both procedures will be considered in more detail later in

the chapter; the important point is that they are made possible by this information-disturbance tradeoff, the *sine qua non* of quantum key distribution.

To see this tradeoff in action, we may consider a specific model of eavesdropping which Eve might use: the intercept/resend attack. Here Eve simply intercepts the photons individually, measures them in some fashion, and prepares a new photon depending on the measurement result, which she forwards to Bob. One obvious choice for Eve's measurement is to randomly select either the x or z basis, note the result, and resend the corresponding state to Bob. If she happens to guess the basis correctly, she'll be able to copy the result without any disturbance, but if she guesses incorrectly, she'll learn nothing and cause maximal disturbance to the polarization state in the process.

The joint probability of Alice's encoding, Eve's attack, and Bob's measurement may be written

$$\begin{aligned} p(j,k;l,m;s,t) &= \frac{1}{16}|\langle\phi_{jk}|\phi_{lm}\rangle|^2|\langle\phi_{lm}|\phi_{st}\rangle|^2 \\ &= \frac{1}{64}[\delta_{jl}(2\delta_{km}-1)+1][\delta_{sl}(2\delta_{tm}-1)+1], \end{aligned} \tag{5.3}$$

where the (j,k) indices correspond to Alice, (l,m) to Eve, and (s,t) to Bob, while the state $|\phi_{lm}\rangle$ stands for an encoded bit m in basis l. The basis announcements allow Alice and Bob to select the elements of the distribution for which $j=s$, so the (renormalized) distribution for the sifted strings is simply

$$p(s,k;l,m;s,t) = \frac{1}{32}[\delta_{sl}(4\delta_{km}\delta_{tm}-1)+1]. \tag{5.4}$$

The indices k and t now refer directly to Alice and Bob's key bits, a and b, so we may use these labels, reserving j,k,l,m,n,s,t to state labels. Eve's information includes both basis and element information, so in keeping with this naming scheme, formally $e=(l,m)$. This shall always be the naming convention, since Alice and Bob generate a raw string which requires processing into a putative key string. Note that due to the symmetric nature of the attack, the resulting distribution is also symmetric, and

we can express everything in terms of the error probability

$$\epsilon = p_{a\neq b} = \sum_{s,l,m,a\neq b} p(s,a;l,m;s,b) = 1/4. \tag{5.5}$$

Suppose that Eve only performs the intercept/resend attack a fraction η of the time. Were η zero, in each sifted step the probability for $a \neq b$ would be zero. At the other extreme, $\eta = 1$ implies an error rate of 1/4, in accordance with the above calculation, as half the interceptions cause no error, and of the half that do alter the polarization state, Bob still has an even chance of obtaining the correct result. In between, the error rate increases linearly with η, so it is given generally by $p(a \neq b) = \eta/4$. By sacrificing some of the created key, Alice and Bob can determine the error rate and thus the value of η.

The intercept rate also determines how much information Eve has about the key; if she intercepts every signal, she learns half of Alice's string, since half the time she guesses correctly. This is equal to her knowledge of Bob's string by the symmetry of the distribution. Writing expressions in terms of the error rate $\epsilon = \eta/4$, we end up with the following relations on the mutual information shared by the various parties:

$$I(A{:}B) = 1 + \epsilon \log \epsilon + (1-\epsilon)\log(1-\epsilon) \;=\; 1 - H_2(\epsilon) \tag{5.6}$$

$$I(A{:}E) = I(B{:}E) \;=\; 2\epsilon, \tag{5.7}$$

where $0 \leq \epsilon \leq 1/4$. In the first equation we've implicitly defined the binary entropy function H_2.

The error rate determines the amount of information Eve has about the key, so sacrificing a few key bits to determine the error rate allows Alice and Bob to get on with correcting these errors and performing privacy amplification to create a truly secret key. Error-correction is made possible with a classical error-correcting code Γ. Given the error rate, Alice and Bob choose a suitable code which can correct these errors. Alice then selects a random codeword γ from Γ and sends $\kappa_a + \gamma$ to Bob. His key string κ_b may be written $\kappa_b = \kappa_a + \delta$, where δ is a string having a fraction ϵ of 1s. Bob subtracts Alice's message from his key string, leaving him with $\gamma + \delta$, which

he corrects to obtain the codeword γ. By Shannon's noisy channel coding theorem, there are (asymptotically) reliably $2^{NI(A\not B)}$ choices for the codeword γ, where N is the length of the key string, so this procedure yields Alice and Bob the same string κ' of length $I(A:B)N$.

Next privacy amplification steps in to create a shortened string κ shared by Alice and Bob of which Eve is completely ignorant. One method is based on computing the parity of a large enough number of bits so that Eve won't know the result. If Eve knew roughly half the bits, for instance, Alice and Bob could agree simply to divide κ into two random substrings and use the parity, or exclusive-or, of the two as a key. Basically, Eve will only know one bit of each pair used in the XOR, so she won't have any information about the result.

Eve's information does not always come in the form of knowledge of a certain subset of bits. Perhaps instead she learns the parity of $N/2$ pairs of bits. Then parity checks aren't a good method of privacy amplification. Generally, Eve will have a joint probability $p(\kappa', e)$ of the key κ' and measurement results denoted here by e. She can crack the cryptosystem if the conditional distribution of key given measurement, $p(\kappa'|e)$, is peaked enough such that she can try all the highly probable keys in a short amount of time. Now consider the average probability of picking the same key twice, $p_c = \sum_{\kappa',e} p(e)p(\kappa'|e)^2$, called the average collision probability. Given the amount of information $t \approx I(A:E)N$ Eve has about κ', Bennett, Brassard, Crépeau, and Maurer [15] show that Alice and Bob can pick a security parameter s such that whenever $-\log p_c \geq N - t$, they can *distill* a key κ with length $N - t - s$ such that Eve's information about κ is less than $2^{-s}/\ln 2$.

The analysis of error-correction and privacy amplification can become quite complicated. Rather than delving into these details to determine the secret key generation rate and the associated secure error rates, we may instead appeal to the following bounds on the optimal key rate R:

$$I(A:B) - \min\{I(A:E), I(B:E)\} \leq R \leq \min_{E \to \overline{E}} I(A:B|\overline{E}). \tag{5.8}$$

The quantity in the upper bound, called the intrinsic information, is the mutual information shared by Alice and Bob given Eve, where Eve is allowed to process her measurement outcomes in any way so as to minimize this quantity [115]. More intuitive is the lower bound, which holds in the case of one-way communication [43]. If Alice and Bob share more information than either does with Eve, the party with the least in common with Eve may initiate an error-correction and privacy amplification procedure as outlined above. To progress beyond this lower bound, a procedure known as *advantage distillation* must be used, though this is of limited efficiency [114]. Note that in the absence of eavesdropping, both bounds equate to $I(A:B)$, the classical capacity of the channel.

We may immediately apply the lower bound, having already calculated the relevant quantities. From this we obtain that for zero error, the key rate is 1/2, as expected, while the maximum tolerable bit error rate *assuming the intercept/resend attack*, obtained as the solution of $2\epsilon = 1 - H_2(\epsilon)$, is roughly 17.1%. Stronger attacks glean more information for the same level of disturbance, so this can only be an upper bound on the maximum secure error rate given the one-way protocol used here.

5.3 Generalized Key Distribution Protocols

The preceding discussion of the BB84 protocol only included the most basic attack, and so can only be considered as a first step toward a proof of unconditional security. But it does establish the basic elements of many quantum key distribution protocols, as follows. Alice and Bob publicly agree on a set of quantum states to be used for encoding and a measurement to be used for decoding. By repeated use of this scheme over the quantum channel, they each generate raw strings, requiring processing into a secret key. In general the generated strings may be only weakly correlated because the original classical signal was encoded into a set of noncommuting quantum states, from which not all the input information may be reliably extracted. In the second step, the

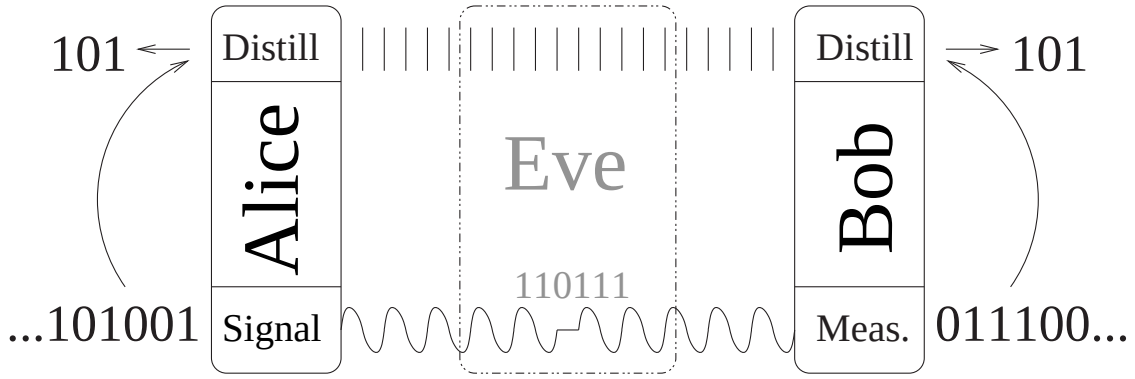

Figure 5.3: Schematic depiction of generalized quantum key distribution. Alice wishes to establish a key with Bob using the insecure quantum channel (bottom) and authenticated classical broadcast channel (top). First she sends quantum signals to Bob, who measures them in a predetermined fashion. Eve can tamper with the signals, shown as the phase delay of the sinusoidal (quantum) signal. After establishing a putative key, shown at bottom for each party, Alice and Bob try to distill a shorter sequence of which Eve is ignorant by communicating on the classical channel. The main difficulty of proving the security of such protocols is determining what Eve knows about the putative key from the noise in the quantum channel.

classical broadcast channel is used to increase the mutual information between the strings, through exchange of information about the quantum states sent and received. In BB84 this stage corresponds to sifting the strings with the basis information. Next Alice and Bob must estimate the error-rate by making use of the statistics they have accumulated thus far in the protocol. In BB84 they announce a portion of the sifted key, though key sacrifice is not necessary in equiangular spherical code protocols, as we shall see later. Finally come the error-correction and privacy amplification steps as described in the previous section. Error-correction in this step differs from the second step, which may also be thought of as error-correction, because in this step there is no explicit reference to the quantum signals. Thus, the public communication at this phase gives Eve only information about their classical strings. The generalized scheme is shown in figure 5.3.

This crucial difference between the error-correction in steps two and four is the reason that proving the security of a quantum key distribution protocol is difficult. In

the course of eavesdropping, Eve acquires *quantum* information about the eventual key, and it is not immediately obvious how much more Eve can do with this and what forms of privacy amplification will be required to combat it. When faced with cracking the BB84 protocol, Eve could for instance attempt to clone the signal, and then wait for the basis information before measuring her copy. Naturally her copy will be of low quality, due to the no-cloning theorem, and introduce noise to the signal, due to the information-disturbance tradeoff. Such features make it difficult to reduce the problem to a situation in which the rate bounds in equation 5.8 apply. To establish unconditional security of any given protocol, we must look for the strongest attack, i.e. the one with the lowest key rate. In practice, however, this is quite complicated at the outset, so the analysis builds up to this case through a series of ever-stronger attacks. Several such attacks have been considered for the BB84 protocol, confined to the requirement that Eve tackle each signal individually [57, 109, 63].

Other promising avenues of handling Eve's quantum information have been recently developed. The first method begins by recasting the protocol in a fully coherent fashion, and using the properties of entanglement to create a key. In this version, described by Ekert in 1991 [56], Alice prepares the bipartite Bell state $(|\uparrow\uparrow\rangle+|\downarrow\downarrow\rangle)/\sqrt{2}$ in the z basis and sends one half to Bob. The key is created when each measures in this basis as this will generate an identical random string for each party. Such a bipartite state is entangled, as it cannot be thought of as the convex combination of locally-defined, or product, states such as $|\uparrow\uparrow\rangle$. For pure states we may measure the entanglement by the von Neumann entropy of either subsystem. Thus for a pure product state, each subsystem state is pure and has entropy zero, meaning the entanglement is also zero. For the Bell state prepared by Alice, each subsystem state is maximally-mixed, an incoherent superposition of spin up and down. Such a state has entropy unity, meaning the entanglement is also unity, the maximum value it can take on in two two-dimensional systems.

The important property of entanglement here is that it cannot be shared: the

more entangled systems A and B are, the less either could be with a third system C. In particular, should systems A and B be maximally entangled, then system C is completely uncorrelated with either. To see this, let A and B be described by the Bell state above. To have an overall pure state of ABC in which the AB subsystem is also pure, it must be that $\rho_{ABC} = \rho_{AB} \otimes \rho_C$. Hence if Alice and Bob generate a key from a shared entangled state, they can be certain that Eve knows nothing about it. Conversely, entanglement is in some sense required for key distribution, for the condition that Alice and Bob end up with shared information that can be made into a secret key is akin to sharing an entangled state [44].

The BB84 protocol may be adapted to this method; after preparing the requisite state, Alice randomly rotates the second system to the x basis with probability $1/2$ [13]. Bob proceeds as usual, randomly choosing a basis and making the corresponding measurement. In this fashion we may convert "prepare & measure" schemes into coherent form. If the quantum channel is noisy, Alice and Bob may use quantum error correction to faithfully transmit the states. Now the problem of Eve's optimal attack is dealt with automatically: the overall coherent protocol can be seen as implementing quantum error-correction and *quantum* privacy amplification [52], ensuring that Eve will not know anything about the final key. Understanding the precise relationship between the coherent and incoherent versions is a subject of ongoing research; see for instance [67, 1, 53, 21]

Both of these steps require a greater level of coherent control than does preparation and measurement of isolated systems, and so are more demanding—really this scheme is an entirely different protocol. Remarkably, however, the coherent protocol complete with quantum error-correction can be converted back into a prepare and measure scheme, at least for BB84 and similar protocols [144]. Such a construction allows the maximum secure error rate to be estimated by considering "virtual quantum privacy amplification", i.e. considering the stronger coherent version without actually altering the physical implementation of the protocol. For the BB84 protocol

using one-way communication, the maximum secure error rate is proven to be 11% using these methods.

Very recent advances in classical privacy amplification allow us to return to the prepare and measure formalism and confront the problem head-on. Ostensibly privacy amplification combats Eve's classical information about Alice and Bob's key, and the contortions of considering coherent schemes, using quantum privacy amplification, and returning to the incoherent protocol work around this limitation. However, classical privacy amplification may be modified to include the case that Eve gains quantum information, and it turns out that doing so yields her only a negligible advantage over classical information [100]. This immediately implies a simplification of security proofs; for one-way communication the secret key rate is provably

$$R = I(A:B) - \max_{\rho_{AB}} S(\rho_{AB}), \tag{5.9}$$

where the maximization of the von Neumann entropy is taken over all bipartite density operators ρ_{AB} consistent with the measurement outcomes resulting in the strings described by the random variables A and B [37]. Although simple looking, this method hides a complication that the rate expression is also valid when conditioning A and B and the density operator ρ_{AB} on further random variables—anything we care to define. Thus, in principle, one must maximize over all possible conditional expressions.

For generic quantum key distribution protocols, we thus have our choice of roughly three methods of analyzing privacy amplification: brute force, quantum error-correction methods, and the recently improved direct quantum privacy amplification. Other methods exist, but they are much more complicated and beyond the scope of this work. Surely the last of these is the most straightforward and simplest, yet because it is so recent, its full implications have not been studied as of this writing. We'll mainly focus on brute force methods, on the assumption that the analysis of the intercept/resend attack and its variants is indicative of the general trend.

5.4 Equiangular Spherical Code Protocols

The use of equiangular spherical codes offers several advantages over "traditional" protocols, which, like BB84, use sets of bases as the signal and measurement ensembles. First, ESCs offer a modest improvement in security on the one hand (again for the intercept/resend attack), and an increased key generation rate on the other, though typically not both simultaneously. Second, we can do away with sacrificing key bits to estimate the error rate and instead rely on the success probability of the protocol itself to furnish this information. This simplifies the protocol considerably and allows for that much more key generation.

To analyze the use of spherical codes, we begin by recounting the security situation for the more oft-studied protocols, which encode using *mutually-unbiased* bases. Using this point of reference clarifies the analysis for ESC protocols and more plainly highlights their advantages. When characterizing any *one* protocol, the key error rate is customarily employed as the relevant parameter and the key generation rate expressed as a function of it. This is sensible because Alice and Bob can easily determine the key rate in practice. To properly *compare* the protocols, though, we should convert the key error rate of each to an actual noise rate of the quantum channel, which we may take as the depolarizing channel. This standardizes Eve's interference across different protocols, each of which has a different sensitivity of the key error rate to channel noise rate. The actual, physical noise rate of the channel is the universal quantity, so it makes for an appropriate comparison.

The discussion thus far has been somewhat framed from Eve's point of view by concentrating on which eavesdropping methods she could use and the security implications for Alice and Bob. Thinking in terms of the noise rate shifts the focus to Alice and Bob's point of view. The noise rate is the only relevant quantity to them, for it is what they confront in practice. For a given noise rate they would like to know if their protocol is secure; considering a specific eavesdropping method is a means to attribute a certain amount of information to Eve. Hence, again, unconditional

security at a given noise rate is established by demonstrating the security of the protocol to *any* attack. Here we make the further assumption that the channel is a depolarization channel, which determines the noise model facing Alice and Bob. Then the intercept/resend attack steps in to establish how much information corresponds to a given depolarization rate. The depolarizing channel is a simple yet realistic channel, respecting the symmetry of the eavesdropping attacks typically considered. It takes an input state ρ to output state $\rho' = (1-q)\rho + q\, I/d$, depolarizing (mixing) it completely with some probability q. In the following both perspectives are used, though they are complementary sides of the same issue.

5.4.1 Mutually-Unbiased Bases

The original protocol of BB84 exploits an appealing feature of the x and z bases of a spin-1/2 particle: each is *unbiased* to the other. This notion may be extended to higher dimensions d by defining two bases to be unbiased whenever the squared overlap of an element from one basis with an element from another is uniformly $1/d$. Again employing the notation $|\phi_{jk}\rangle$ for states in such a collection of bases, where j labels the basis and k the state within that basis, we have

$$|\langle\phi_{jk}|\phi_{lm}\rangle|^2 = \begin{cases} 1 & j = l, k = m \\ 0 & j = l, k \neq m \\ 1/d & j \neq l \end{cases} \tag{5.10}$$

A collection of bases such that all pairs are unbiased is called a set of *mutually unbiased bases*, and it is known that no more that $n_{\text{mub}} = d + 1$ such bases can exist in d dimensions, though this full set is only known to exist in prime or prime-power dimensions. In other dimensions it seems likely that only three unbiased bases exist [169]. Two being prime, in $\mathbb{C}^2$ three unbiased bases exist; quite obviously we may append the y basis to the x and z bases we have been using. Key distribution using all three bases is termed the six-state protocol. In terms of photon polarization this corresponds to using the bases of horizontal/vertical (+) linear polarization,

right/left circular polarization (↻↺), and 45/135 degree (×) linear polarization. Key distribution protocols may be simply constructed using any number of unbiased bases, in a manner completely analogous to the BB84 protocol. Both more bases and higher dimensions lead to improved security [18, 34].

The one-way key generation rate based on the lower bound is simple to determine. We have the ingredients necessary in the analog of equations 5.3 and 5.4. First, assume that Alice chooses signals with equal probability,[2] and she and Bob employ the sifting postselection. When Eve intercepts a signal, she measures in the wrong basis with the probability $(n_{\text{mub}} - 1)/n_{\text{mub}}$. In this case she forwards Bob a state which generates a completely random outcome since it is an element of an unbiased basis. Should she measure in the correct basis, she forwards Alice's signal unchanged to Bob. Thus the total probability of error in Bob's key string is the product of the probability for Eve to intercept, to measure in the wrong basis, and for Bob to get the wrong outcome. Symbolically,

$$p_{a\neq b} = \eta \left(\frac{n_{\text{mub}} - 1}{n_{\text{mub}}} \right) \left(\frac{d-1}{d} \right) = \epsilon. \tag{5.11}$$

Recall that subscripts a, b, e are used to refer to the key letters of the corresponding parties, while the j, k, l, m, s, t to signal states and measurements. The information Alice and Bob share in their key strings is then simply

$$I(A{:}B) = \log d + (1-\epsilon)\log[1-\epsilon] + \epsilon \log \left[\frac{\epsilon}{d-1} \right]. \tag{5.12}$$

In the case of a noiseless channel, a moment's thought reveals that this expression reduces to $\log d$.

Eve, for her part, gets no information when she doesn't intercept the signal, nor if she measures in the wrong basis. Otherwise, she gains the maximum amount, $\log d$. In all, this yields her an amount of information given by

$$I(A{:}E) = I(B{:}E) = \frac{\eta}{n_{\text{mub}}} \log d = \frac{\epsilon}{n_{\text{mub}} - 1} \left(\frac{d}{d-1} \right) \log d. \tag{5.13}$$

[2]Note that this need not be the case, and in fact faster key rates can be achieved at the same level of security by using asymmetric signal probabilities [105].

The value of ϵ for which these two information expressions are equal is the error rate at which the protocol becomes insecure, which must then be converted into an equivalent depolarizing rate. For the unbiased bases this conversion is quite simple: a depolarized signal causes an error in a fraction $(d-1)/d$ of samples, so the depolarizing rate corresponding to error rate ϵ is $q = \epsilon\, d/(d-1)$.

5.4.2 Equiangular Spherical Codes

Turning attention, finally, to the equiangular spherical code protocols, suppose Alice and Bob publicly agree on a ESC with n elements in d dimensions ($d \leq n \leq d^2$). Alice then selects states randomly and transmits them to Bob, who uses the same ESC as a measurement. The symmetric nature of the spherical codes translates into the following probability distribution when transmitting on a noiseless channel:

$$p(a,b) = \frac{d}{n^2}|\langle\phi_a|\phi_b\rangle|^2 = \begin{cases} d/n^2 & a = b \\ (n-d)/n^2(n-1) & a \neq b \end{cases} \tag{5.14}$$

In this case we can dispense with the labels $j, k \ldots$ and proceed to a, b, e since the meaning is clear. Alice and Bob's key letters will agree with probability d/n. Then classical error correction may be used to yield a shorter string, identical for the the two parties with high probability. Labeling the spherical code states 0 to $n-1$, Bob's string β is simply Alice's string α plus a string δ having a fraction $(n-d)/n$ of non-zero elements. From this point they proceed exactly as in section 5.2, usingclassical error correction to yield with high probability a shorter string, identical for the two parties.

This protocol is not particularly robust, however, and Alice and Bob can do better by announcing some of the signals *not* received. Upon receipt of each signal, Bob publicly broadcasts m randomly-chosen outcomes he did not obtain. If Alice's signal is among these, she publicly announces this fact and they throw the signal away and proceed to the next. This occurs with probability $m/(n-1)$ as Bob could send any of $\binom{n-1}{m}$ outcomes and $\binom{n-2}{m-1}$ of these contain Alice's signal. For those which pass

the test, Alice and Bob relabel the remaining states in order from 0 to $n-m-1$ and follow the error-correcting procedure. The protocol itself succeeds with probability

$$p_{\text{succeed}} = \frac{n(n-1) - m(n-d)}{n(n-1)} = \frac{s}{n(n-1)}, \tag{5.15}$$

where we have implicitly defined the constant s. The joint distribution of Alice and Bob's key letters, given that the protocol succeeds, becomes

$$p(a,b) = \frac{1}{s} \times \begin{cases} d(n-1) & a = b \\ (n-d)(n-m-1) & a \neq b \end{cases}. \tag{5.16}$$

In the bargain they gain improved security. The whole protocol will tolerate more noise since much of it will be discarded. This will become more apparent when considering the intercept/resend attack; for now, we find the following expression for the key rate using an n-word equiangular spherical code in d dimensions excluding m outcomes:

$$R = \log[n-m] + \frac{d(n-1)}{s}\log[d(n-1)] + \left(1 - \frac{d(n-1)}{s}\right)\log[n-d] - \log s. \tag{5.17}$$

Removing errors in this fashion improves the key rate, up to a point. By sacrificing all but one outcome not obtained—leaving two possibilities for the key letter—Bob reduces the possibility of error, but also reduces the number out possible key letters, which also affects the key rate. For a small number of omitted non-outcomes, the change in number of letters isn't drastic, and the error cleanup helps matters. The value of m which yields the highest key rate is found by locally maximizing $I(A:B)$ by setting the derivative to zero. To simplify the expression, let $\ell = \log[d(n-1)/(n-d)]$, whence we obtain

$$m_{\max} = \max\left\{0, \frac{n(n-1)\left(n(d-1) - d(n-d)\ell\right)}{(n-d)\left(n(d-1) - d(n-1)\ell\right)}\right\}. \tag{5.18}$$

The maximum is included to take care of those values of n and d for which $m = 0$ yields the largest key value, but which doesn't have a zero derivative.

Now for the intercept/resend eavesdropping analysis. As in the case of unbiased bases, suppose Eve intercepts a fraction η of signals. She measures these using

the same equiangular spherical code as does Bob, resending him the output of this process. Eve simply assumes that her outcome corresponds to Alice's signal, unless it is excluded by Bob's announcement. In this case she may still guess, retaining the information that she was forced to do so. The increased security of the protocol stems from Bob forcing Eve to abandon her outcome in this manner.

By delineating the various cases, as shown in figure 5.4, it is uncomplicated to arrive at the relevant probabilities for the case of full interception, $\eta = 1$. However, the breakdown in the figure gives us only the raw case, which we must renormalize to omit the cases in which the protocol fails. This complicates matters as we cannot simply mix this distribution with the $\eta = 0$ case because this renormalization itself varies with η. Otherwise, the situation is linear. Therefore with an explicit expression for the probability of protocol success, we can fix the normalization and then linearly interpolate between the two cases. The success rate of the protocol itself can be found simply by mixing the cases, and depends on η as in the following expression. Letting $t = s(n-1) - \eta m(n-d)(d-1)$, the success rate is

$$p_{\text{succeed}} = t/n(n-1)^2. \tag{5.19}$$

Before proceeding to the key letter probabilities, note that this is already a major departure from the unbiased bases protocols. They don't share this feature of a success rate varying with the noise rate: should Bob measure in the same basis Alice prepared the state, the protocol succeeds, regardless of the noise. The probability of the protocol succeding is clearly $1/n_{\text{mub}}$. In contrast, for equiangular spherical codes the protocol succeeds with higher probability when Bob receives a state which is close to Alice's original signal. Coupled with Alice's random signal selection and the wide spacing of signal states, no matter what form the noise takes, its rate affects the protocol success rate. The noisier the channel, the likelier the protocol is to fail.

This feature makes the overall protocol much simpler as Alice and Bob no longer need to sacrifice key bits to estimate the error rate. Together with the ability to vary the number m of outcomes to exclude, Alice and Bob can adapt the protocol to the channel noise on the fly without varying any of the hardware setup. This works on a

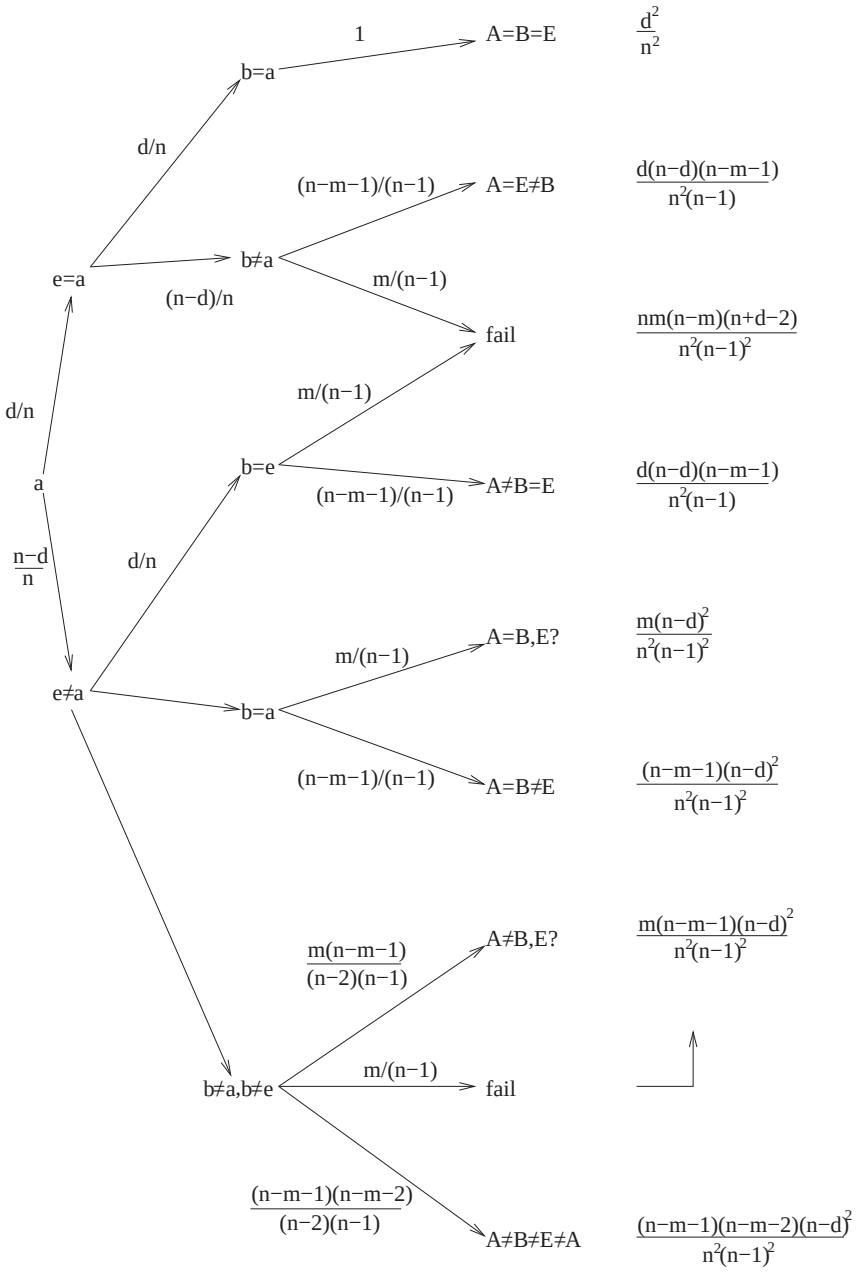

Figure 5.4: Tallying the possibilities in the ESC protocol using n states in d dimensions and discarding m outcomes. Beginning on the left with Alice's state, probabilities label the arrows to the various cases. Lower-case letters stand for signal states or outcomes, while upper case stand for the key letters. The total probability of each case is shown to the right.

block by block basis, where the blocks of signals are chosen to be long enough such that the error rate can be reliably estimated from the success rate. For the initial block, Bob conservatively chooses m based on what he thinks the channel error is likely to be. After running the protocol through the block, he estimates the error rate and alters m to suit. He need not confer with Alice to do so, since they don't need to compute the putative key and compare parts of it to determine the error rate. Because altering m does not involve the hardware, it makes sense to think of ESC protocols as defined by n and d, each one including all the possible choices of m. The protocol runs fastest when $m = 0$ and most securely when $m = n - 2$.

Now we return to finding the joint key letter probability distributions, employing the modified mixing procedure described above. Alice and Bob's joint distribution is characterized solely by the agreement probability

$$p_{a=b} = (n-1)[d(n-1)-q(n-d)(d-1)]/t = 1 - \epsilon. \tag{5.20}$$

The mutual information shared by Alice and Bob then follows immediately:

$$I(A:B) = \log[n-m] - \epsilon \log[n-m-1] - H_2(\epsilon). \tag{5.21}$$

Since Alice and Bob use the same ESC ensemble, Eve's joint probability with Alice is the same as with Bob. In order to account for the cases in which Eve measures the signal but this outcome is later excluded by the protocol, we may append an event to her probability distribution, denoted by `?`. Now she has $n - m + 1$ total outcomes, and the `?` outcome functions as a guess as to the key letter in the cases it occurs. We didn't need to deal with this directly for unbiased bases since there it was a simpler matter to go directly to the mutual information, counting only the cases in which she need not guess.

The joint probability of such an exclusion and the particular signal j is plainly the same for all j, and together with the probability of agreement between Alice and Eve, these quantities fully describe the overall distribution:

$$p_{a=e} \;=\; qd(n-1)s/nt, \tag{5.22}$$

$$p_{?} = 1 - qs^2/nt. \tag{5.23}$$

From these we find immediately the entropies

$$H(A) = \log[n-m], \tag{5.24}$$

$$H(E) = -p_{?}\log[p_{?}] - (1-p_{?})\log\left[\frac{1-p_{?}}{n-m}\right], \tag{5.25}$$

$$\begin{aligned} H(AE) = {} & -p_{a=e}\log\left[\frac{p_{a=e}}{n-m}\right] - p_{?}\log[p_{?}] \\ & -(1-p_{?}-p_{a=e})\log\left[\frac{1-p_{?}-p_{a=e}}{(n-m)(n-m-1)}\right]. \end{aligned} \tag{5.26}$$

which together make the mutual information

$$I(A{:}E) = p_{a=e}\log\left[p_{a=e}\right] - (1-p_{?})\log\left[\frac{1-p_{?}}{n-m}\right] + (1-p_{a=e}-p_{?})\log\left[\frac{1-p_{a=e}-p_{?}}{n-m-1}\right]. \tag{5.27}$$

Thus far all quantities are expressed in terms of the key error rate; the final step is to convert this error rate into the depolarization rate. This task is best accomplished by finding an expression for the key error rate in terms of the depolarization rate and then solving for the latter in terms of the former. Supposing Bob receives the state ρ_a when Alice sends the signal $|\phi_a\rangle\langle\phi_a|$, first note that the raw probability of agreement is given by

$$p_{\mathrm{a=b}} = \frac{d}{n^2}\sum_a \langle\phi_a|\rho_a|\phi_a\rangle = \frac{d}{n}\left(1 - q + \frac{q}{d}\right). \tag{5.28}$$

Meanwhile, the protocol fails with probability

$$p_{\mathrm{fail}} = \frac{m}{n-1}\left(1 - p_{\mathrm{a=b}}\right), \tag{5.29}$$

so the error probability in the processed key string is simply $1-p_{\mathrm{a=b}}/(1-p_{\mathrm{fail}})$. With a little algebraic manipulation we can massage this expression into

$$q = \frac{s}{m(d-1)} - \frac{n(n-1)(n-m-1)}{m(d-1)(n-1-m(1-\epsilon))}. \tag{5.30}$$

Finally, we are in possession of all the ingredients needed to find the maximum tolerable noise rate. First we find the intercept rate η for which $I(A:B) = I(A:E)$, then convert it to the error rate $\epsilon = 1 - p_{a=b}$, and lastly the equivalent noise rate using this expression.

5.4.3 Comparison

Though easily adaptable and more efficient at estimating the noise rate, ESC protocols will not be of much use if they are only secure in a narrow range of parameters. Generically, the key generation rate decreases monotonically with noise rate, as shown in figure 5.5. Each protocol can be characterized by a pair of quantities: the noiseless key generation rate and the maximum tolerable noise rate. As we standardized the security analysis by referring everything to the depolarization channel, we must also standardize the absolute key rate by including the success rate of the protocol. The former is $\log[d]/n_{\text{mub}}$ for unbiased bases and given by the product of equations 5.17 and 5.15 for spherical code protocols. For the latter we may again appeal to the one-way communication lower bound.

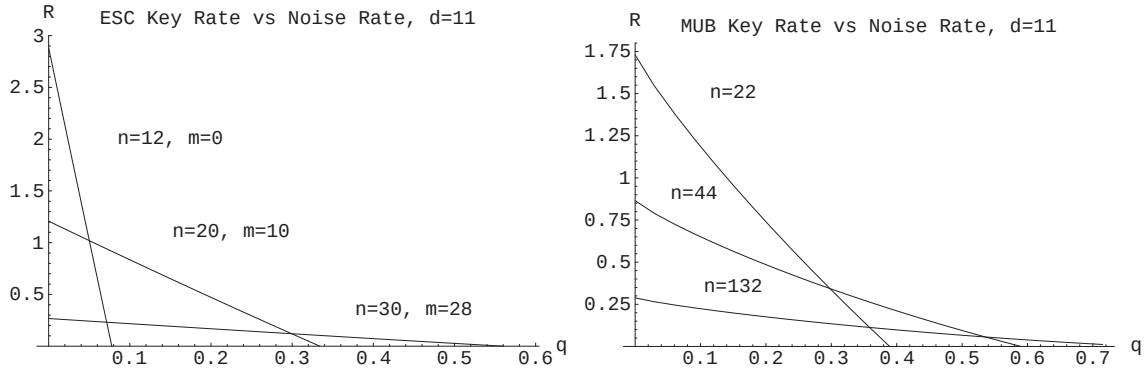

Figure 5.5: Key rate R versus depolarization noise rate q for various ESC and MUB protocols in $d = 11$ dimensions. Key rate generically decreases monotonically with noise rate, so each protocol may be characterized by the horizontal and vertical intercepts.

Neither of the maximum secure noise rate quantities for MUB nor ESC protocols lend themselves to analytic expressions, so we resort to numerical investigation using *Mathematica*. We've already established how the spherical codes are able to estimate the error rate from the success rate and that this translates into a simple scheme for adapting to varying channel noise. The numerical results paint the rest of the picture: how the key generation rates and maximum secure noise rates compare. Two main

conclusions follow. First, spherical codes provide both increased noise tolerance and key generation rates for fixed dimension d. Second, for fixed dimension and number of signals, spherical codes provide more security but lower key rates than their basic cousins.

Broadly speaking, one may demand security or speed, but not both simultaneously. Examining the key generation rate in a fixed dimension d, there are more possible ESC ensembles to choose from, and in particular protocols for which $n < 2d$, the minimum number for unbiased bases. Since each protocol employs a number of signals greater than the dimension of the quantum states used to encode them, the more signals used, the lower the key generation rate. Therefore, consider the versions of each protocol which employ the fewest number of signals. A minimum of two unbiased bases must be used, while the spherical codes employ a minimum of $d+1$ signals. In the latter case, this translates into a key generation rate of $\log[d]/2$, while the latter amounts to $(d-1)\log[d]/(d+1)$. Already in three dimensions the two are equal and for any higher dimension the spherical codes enjoy a speed advantage. Figure 5.6 shows this maximum key rate for each protocol, normalized to the capacity of the channel, $\log[d]$. Asymptotically the key generation rate for $n = \alpha d$ spherical code signals behaves roughly like $\log[d]/\alpha$, so for any number $n < 2d$ the key generation rate is similar or higher to that of two unbiased bases.

The speed/security tradeoff is already apparent, however, as $d + 1$ signals are not too different from an orthogonal basis which itself provides no defense against copying by Eve. In a given dimension, however, the ESC protocols may employ up to d^2 signals, and by announcing all but two of the outcomes not obtained, Bob aims for maximum security. Similarly, MUB protocols are most secure when using the full complement of $d + 1$ bases. In this case, numerical results reveal the spherical code protocols already tolerate more noise in just four dimensions. Figure 5.7 shows the maximum tolerable noise rate as a function of dimension.

Thus in a particular *dimension* ESC protocols may be found which are faster or more secure than their MUB counterparts. Alice and Bob simply have more

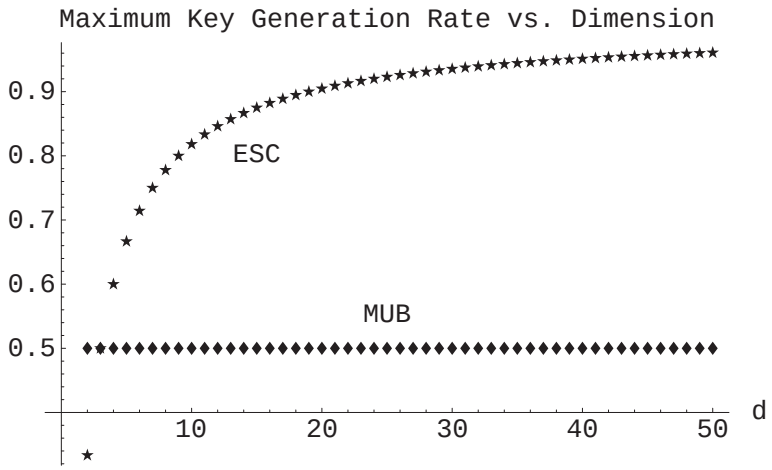

Figure 5.6: Maximum possible key generation rate versus dimension for the two ensembles, normalized to the classical capacity of the channel. Two unbiased bases is the minimum, optimal number where speed is concerned, corresponding to half the capacity no matter the dimension. The spherical codes may employ $d+1$ states, which offer little security, but asymptotically approach the maximum capacity.

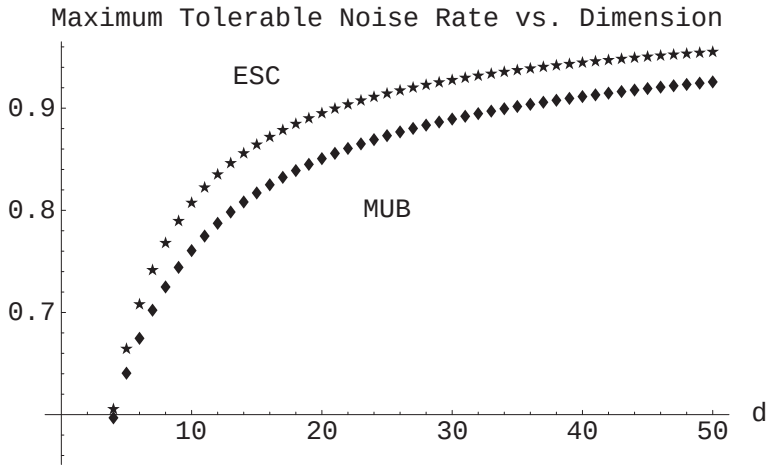

Figure 5.7: The maximum tolerable depolarizing rate versus dimension for the two protocols. Using the full complement of $d+1$ unbiased bases and d^2 spherical code states achieves the maximum security, with the ESC protocol excluding all but two possibilities. Both asymptotically tolerate total depolarization, but the spherical codes offer more security for any finite dimension four or greater.

choices when using spherical codes, but they cannot have it both ways in a particular

protocol, i.e. for a particular choice of n and d. Comparing the two protocols at fixed dimension and fixed number of signal states, we'll find the spherical codes offer more security but lower key rates. Figure 5.8 shows the tradeoff between speed (horizontal axis) and security (vertical) when using $n = 2d$ signal states in each protocol, for dimensions two through 100. The unbiased bases are faster, but tolerate less noise than the spherical codes.

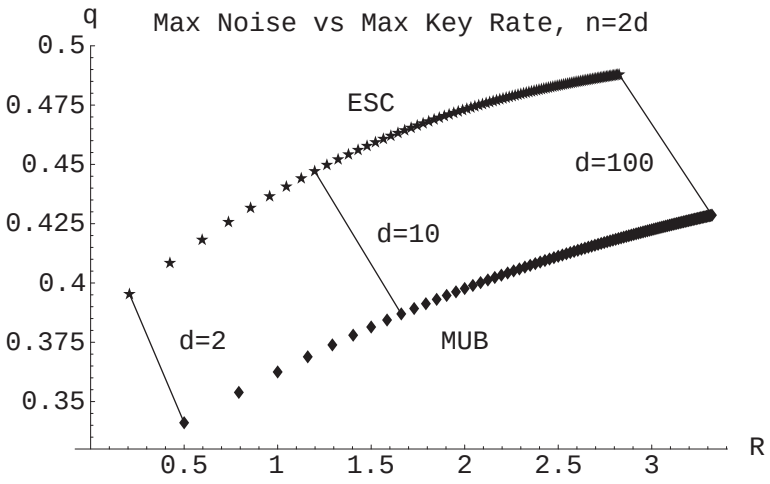

Figure 5.8: Maximum depolarizing rate versus maximum key generation rate when using $n = 2d$ signal states in either protocol. Unbiased bases are faster, but tolerate less noise; the three connecting lines for dimensions two, 10, and 100 link corresponding protocols.

Suppose, though, that the noise rate of the channel is low. In this case Alice and Bob would prefer to use as few signal states as possible. Note that using two unbiased bases in two dimensions provides security up to noise rates of roughly $q = 0.34$, so should Alice and Bob independently estimate the channel noise to be, say, $q = 0.25$, even the BB84 protocol would be overkill. Instead, by using a number of spherical codes roughly 4/3 of dimension, the security of the resulting protocol can be ensured and higher key generation rates achieved. Comparing the key generation rates on the noiseless channel, the two unbiased bases again yield $\log[d]/2$ bits per signal, while the $n = 4d/3$ ESC protocol achieves a rate of $\log[d] - \log[4d-3]/4$, asymptotically $3\log[d]/4$, an improvement of 50%. Figure 5.9 shows the exact behavior.

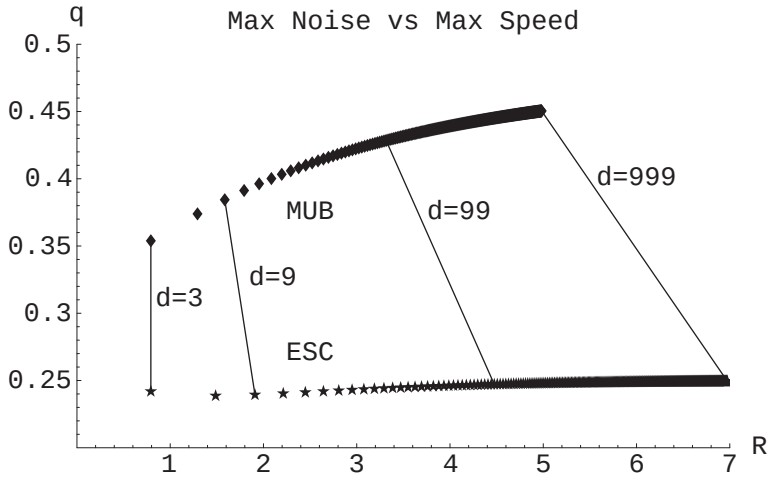

Figure 5.9: For low noise, say $q = 0.25$, ESC protocols with roughly $4d/3$ elements provide sufficient security and offer higher key generation rates. In contrast, to maintain security, at least two unbiased bases must be used, though these are better suited to higher noise rates.

These findings do not indicate a preference of one type of quantum key distribution protocol over another in all cases. Ultimately equiangular spherical codes complement the use of unbiased bases. In the abstract they are both faster and more robust for fixed dimension, due primarily to the wider range of options. In practice, however, a protocol by protocol comparison fixing both n and d is more appropriate. In cases of fixed n and d where more noise tolerance is required, or in cases of low noise, equiangular spherical codes are to be preferred.

Returning to the concept of the frame potential, we can find a general reason why we should have expected ESC protocols would offer the most security. In the intercept/resend cryptographic context, the suitably-rescaled V_2 frame potential can be interpreted as the marginal probability of agreement between Alice and Bob, given that Eve intercepts each signal:

$$p_{a=b} = \frac{d}{n^2} \sum |\langle \phi_j | \phi_k \rangle|^4 = \frac{d^2}{n^3} V_2. \tag{5.31}$$

This holds no matter what the ensemble used by Alice, provided it forms a POVM which Bob may use at his end. Thus the prefactor comes from the product of the prior

probability of the signal, $1/n$, and the normalization associated with the POVM, d/n, once each for Bob and Eve. Meanwhile, by the same token, the agreement probability over a noiseless channel is simply d/n. The difference in the agreement probabilities may be taken as a measure of the disturbance to the signals caused by the eavesdropping. Since $V_2 \leq V_1$ by inspection, the difference

$$\mathcal{D}_s = \frac{d}{n}\left(1 - \frac{d}{n^2}V_2\right) \tag{5.32}$$

is positive and the maximum obtains whenever V_2 is minimal. Given the assumption that V_1 is already minimized, the greatest disturbance therefore occurs in the ensembles which are spherical codes (for $n \leq d^2$) or spherical 2-designs (for $n \geq d^2$). This sensitivity to eavesdropping translates into increased noise tolerance, since Alice and Bob may be more certain what knowledge Eve has of the key.

One caveat should be mentioned. This analysis is well-suited to a *symmetric* protocol, one for which the resulting joint probability between Alice and Bob can be characterized solely by the probability of agreement. This need not hold generally; instead the pair may end up with an essentially arbitrary distribution. In this case they might use a generalized version of the disturbance measure

$$\mathcal{D} = \left(\sum_{a,b}(p_{ab} - p^e_{ab})^2\right)^{1/2} \tag{5.33}$$

in which discrepancies in any signal/outcome pair are taken into account. Here p_{ab} is the joint probability absent noise, and p^e_{ab} given full signal interception. The earlier measure now provides a lower bound on the generalized disturbance measure. We may write this in a simpler form as the Frobenius norm of the matrix M_{ab} having elements $p_{ab} - p^e_{ab}$, that is $\mathcal{D}^2 = \mathrm{Tr}[MM^T]$. Then the desired result follows according to

$$\begin{aligned} \mathcal{D}^2 &= \sum_k \sigma_k(MM^T) \\ &\geq \sum_k \sigma_k(M)\sigma_k(M^T) \end{aligned}$$

$$\begin{aligned} &\geq \frac{1}{r(M)}\left(\sum_k \sigma_k(M)\right)^2 \\ &\geq \frac{1}{r(M)}\mathrm{Tr}[M]^2 \\ &= \frac{1}{r(M)}\mathcal{D}_s^2, \end{aligned} \tag{5.34}$$

where $\sigma_k(M)$ are the singular values and $r(M)$ the rank. The first line follows by definition, the second from a well-known inequality [79], the third from the fact that $\sum_k^n x_k^2 \leq (\sum_k^n x_k)^2/n$ for all real x_k, and the fourth from the triangle inequality of the trace norm. Intuitively we might expect the spherical codes to have high noise tolerance because by having the greatest lower bound on the disturbance $\mathcal{D}$ they are very sensitive to the eavesdropper's presence.

Of course, this isn't the end of the story, as most of the preceding analysis involved the simplest attack, intercept/resend. The speed advantage enjoyed at low noise rates is certain to be valid, as this does not depend much on the attack, but we must be concerned with the maximum noise rate estimates. Clearly these must be revised downward upon considering stronger attacks. The crucial question is the following: will the two protocols require the same revision, or will the spherical codes become comparatively better or worse than the intercept/resend attack indicates?

To give a flavor of the analysis which awaits us in solving this problem, we may consider the next two rungs on the ladder toward unconditional security. First, keeping within the framework in which Eve handles signals one at a time, instead of simply intercepting them, she may choose to weakly clone them. This unitary operation attempts to make two copies of an arbitrary input state, each having a different fidelity with the input state. Naturally, both copies cannot be of perfect fidelity, and Eve may select the quality of each copy within this overall requirement. The unitary quantum cloning machine (UQCM) is known to be the optimal attack against $n_{\mathrm{mub}} = d+1$ bases when monitoring each signal separately [18], and a related cloning attack is thought optimal for two bases [34].

Beyond this attack a more coherent approach in the manner of Preskill and Shor

may be appropriate. The first step is to translate the bare protocol into coherent form. For BB84, this is easily done by preparing the maximally-entangled state and randomly applying a rotation to switch bases, a procedure that works for general protocols using mutually-unbiased bases. The target bipartite state for Alice to begin with is simply $|\Psi\rangle = (1/\sqrt{d}) \sum_k |k\rangle|k\rangle$. Spherical code protocols, too, can be easily made coherent in this fashion. For a ESC set with states $|\phi_k\rangle$, consider the conjugate set consisting of elements $|\phi_k^*\rangle$, obtained from the original by conjugation in the standard basis. A maximally-entangled state may be written $|\Psi\rangle = (\sqrt{d}/n) \sum_k |\phi_k\rangle|\phi_k^*\rangle$, ensuring that the two protocols do start on the same footing. This fact is easily established by forming the density operator from the state vector and taking the partial trace over the second system to obtain the description of the first system alone, as follows.

$$\begin{aligned} \rho_{AB} &= \frac{d}{n^2} \sum_{jk} |\phi_j\rangle\langle\phi_k| \otimes |\phi_j^*\rangle\langle\phi_k^*| \\ &\Downarrow \\ \rho_A &= \frac{d}{n^2} \sum_{jk} \langle\phi_j|\phi_k\rangle \, |\phi_j\rangle\langle\phi_k|. \end{aligned} \tag{5.35}$$

The expression for ρ_A shows that it has no kernel since the ESC states span the vector space. If we can show that $\rho_A^2 = \rho_1/d$, then we know that $\rho_A = I/d$. Squaring the operator is simple enough, so

$$\rho_A^2 = \frac{d^2}{n^4} \sum_{jklm} \langle\phi_j|\phi_k\rangle\langle\phi_k|\phi_l\rangle\langle\phi_l|\phi_m\rangle|\phi_j\rangle\langle\phi_m|. \tag{5.36}$$

We can sum over the k and then l indices, using the fact that $\sum_k |\phi_k\rangle\langle\phi_k| = (n/d)I$, to obtain the desired result. The complex conjugate is simply used to interchange the order of the inner product: $\langle\phi_j^*|\phi_k^*\rangle \to \langle\phi_k|\phi_j\rangle$. Note that with ESC protocols, though, Alice need not apply any further action to the signal subsystem, since a simple transmission of it and measurement of her own half suffices to realize the prepare and measure setup.

These two steps are beyond the scope of this work, but by focusing on a particular

case, we can make progress beyond the simple intercept/resend attack toward the goal of provable unconditional security.

5.5 Two Qubit Protocols

The outstanding choice for quantum key would certainly be the qubit case, for it lends itself more readily to practical implementation by encoding into photon polarization states as with the original BB84 protocol. Reviewing figures 5.6 and 5.7 one finds that spherical code protocols for qubits perform quite poorly, being both slower and less resistant to noise. This arises from the poor choice of measurement ensemble on Bob's part. By choosing the *same* set of codewords as Alice he simplifies the process of key creation since their goal is simply to agree on the identity of the transmitted and received states, but decreases the amount of classical information that can be sent through the quantum channel. Intuition leads us to believe that if more information can be transmitted, one can find a way to ensure the secrecy of at least a part of that information.

Such is indeed the case for qubits, and may be so in higher dimensions as well. Two equiangular spherical codes exist for qubits, the trine and tetrahedron ensembles. For the trine it is known that if Bob makes each state in his measurement ensemble orthogonal to the corresponding state of Alice's signal ensemble, the classical capacity of the resulting channel is maximized [133], and a similar result is conjectured for the tetrahedron [49]. In higher dimensions, to take a numerically-generated example, six equiangular states in three dimensions have a nominal capacity of 0.424 bits when both parties use the same ensemble. This increases by roughly 50% to about 0.638 bits if Bob picks a suitable version of Alice's signal ensemble, in which each element is unitarily transformed with the same unitary operator. This example was generated by locally-minimizing the frame potential until an ESC signal set was found. Bob's measurement ensemble was then found by optimizing over all en-

sembles unitarily-equivalent to the original. Both operations were performed using *Mathematica*™.

Increased capacity is no doubt useful, but the difficulty lies in converting the signal record and measurement outcomes into highly-correlated key strings. Take the trine for instance. Since Bob's ensemble is the inverse of Alice's in the sense of the Bloch sphere, upon obtaining a particular measurement result, Bob can only be certain that Alice did not send the corresponding antipodal signal. He remains completely uninformed as to the other two possibilities. The same applies when using the tetrahedron, except Bob is completely uninformed about the three possibilites not ruled out by his measurement outcome, and the numerical example of six states in three dimensions is still more complicated.

Again the trick is for Bob to reveal outcomes not obtained, proceeding by elimination. Such a method was applied to the trine ensemble by Phoenix, *et al.* [123]. For a given outcome, Bob already knows one signal Alice didn't send, and conversely, given the transmitted signal, Alice may already eliminate one outcome Bob might receive. Starting from this little bit of knowledge about what the other doesn't have, they may proceed by publicly announcing further results not obtained or signals not sent and eventually exclude enough possibilities to agree on a bit. The shared (anti-) correlation between signal and outcome allows them to remain one step ahead of an eavesdropper Eve, ensuring that unless she tampers with the quantum signal, she knows nothing of the created key.

Unlike the previously-studied protocols, in which Alice's choice of signal or Bob's outcome determined the key letter, for the trine and tetrahedron it is only the relation between Alice's signal and Bob's outcome that determines the key bit. In the trine protocol Alice's choice of signal narrows Bob's possible outcomes to the two lying 60 degrees on either side. Each is equally likely, and they publicly agree beforehand that the one clockwise from Alice's signal corresponds to `1` and the other `0`. Alice hopes to determine which is the case when Bob announces one outcome that he *didn't* receive. For any given outcome, he chooses randomly between the other two

and publicly announces it. Half the time he announces he did not receive an outcome which Alice already knows to be impossible. This tells Alice nothing new, and she announces that the protocol failed. In the other half of cases, Alice learns Bob's outcome and announces success.

Upon hearing his message was a success, Bob can determine the signal Alice sent. For any outcome Bob receives, he immediately knows one signal Alice couldn't have sent, and the message that his announcement was successful indicates to him that she also didn't send the signal orthogonal to his message. Had she sent that signal, she would have announced failure; thus between his measurement outcome and the success message, Bob learns the identity of Alice's signal. Each knowing the relative position of signal and outcome, they can each generate the same requisite bit.

Mathematically, we might consider the protocol as follows. Suppose we label the signal states and measurement outcomes clockwise from one to three such that Bob's j outcome is orthogonal to Alice's jth signal. When she sends signal j, Bob necessarily obtains $k = j+1$ or $k = j+2$. Then he announces that he didn't receive some $l \neq k$. If $l = j$, Alice announces failure. Otherwise each party knows the identity of j, k, and l, and they compute the key bit as $(1-\epsilon_{jkl})/2$. Figure 5.10 shows the case that they agree on a `1`.

Though Eve may listen to the messages on the classical channel, she won't have any knowledge of the bit value, for all she knows is one outcome Bob didn't receive and the corresponding antipodal state that Alice didn't send. Of the two remaining equally-likely alternatives, one corresponds to a `0` and the other a `1`. Hence the protocol establishes one fully secret bit half the time, analogous to the BB84 protocol.

The strategy for the tetrahedron is entirely similar, except that Bob must now reveal two outcomes he didn't receive. As depicted in figure 5.11, Alice uses four tetrahedral states in the Bloch-sphere picture, and as before Bob uses the inverse of Alice's tetrahedron for measurement. Alice sends signal j and Bob receives $k \neq j$. He then randomly chooses two outcomes l and m he didn't receive and announces

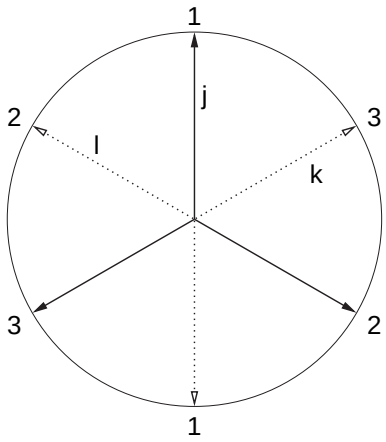

Figure 5.10: Bloch-sphere representation of the trine-based protocol by which Alice and Bob create a secret key bit, shown here creating a `1`. Alice's three possible signal states are shown in black and Bob's measurement outcomes in dotted lines; antipodal points are orthogonal. Without loss of generality we may assume that Alice sends the state $j = 1$. The antipodal point is the impossible outcome for Bob; here he obtains the outcome $k = 3$. Of the two outcomes he did not get, he picks one at random and announces this to Alice. Here he announces the outcome $l = 2$, and Alice infers the value of k. Had Bob announced the other outcome, the protocol would fail, as this doesn't tell Alice anything she doesn't already know. Here she announces that she is satisfied with Bob's message, and Bob infers the value of j, since Alice's signal could not have been l. Now they compute the bit $(1-\epsilon_{jkl})/2 = 1$. The announcement only reveals l, so the bit is completely secret.

them. One-third of the time this is successful, in that $l \neq j$ and $m \neq j$. This allows Alice to infer k, and her message of satisfaction allows Bob to infer j, just as for the trine. They then each compute the bit $(1 + \epsilon_{jklm})/2$. Note that in this case the order of Bob's messages effectively determines the key bit, so he should take care to ensure that they are sent in a random order each time.

Again they stay one step ahead of Eve as she listens to the messages, as she can only narrow Alice's signal down to two possibilities. Given the order of Bob's messages, one of these corresponds to `0` and the other to `1`, so Eve is ignorant of the bit's identity. Using the tetrahedron allows Alice and Bob to establish one fully secret bit one third of the time, analogous to the six-state protocol.

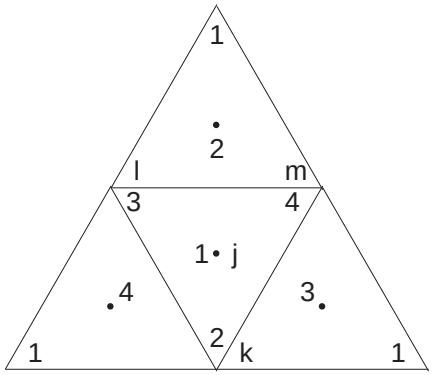

Figure 5.11: Unfolded view of the Bloch-sphere tetrahedron states. Vertices of triangles correspond to Bob's outcomes, their centers Alice's signals; all three vertices of the large triangle represent the same point antipodal to its center. Suppose Alice sends signal j; Bob necessarily receives $k \neq j$. Here we suppose $j = 1$ and $k = 2$. Bob then announces two outcomes he didn't obtain, here shown as $l = 3$ and $m = 4$. Had either message equaled j, which happens 2/3 of the time, Alice announces failure. Otherwise, as here, she accepts. Thus Alice determines k, and Bob finds out j. They compute the bit $(1 + \epsilon_{jklm})/2 = 1$. The announcement reveals only l and m, so the bit is secret.

In the two protocols, the dual arrangement of signals and measurements allows Alice and Bob to proceed by elimination to establish a putative key. To establish security of the protocols we again start with the standard intercept/resend attack and then proceed on to a "gentler" version of the same. Just as before, we're assuming that Eve simply intercepts a fraction η of the signals, measures them, and sends a new state on to Bob. The first task is then to determine R as a function of η and then to relate η to the statistics compiled in the course of the protocol. As always, we should compare the protocols in terms of the maximum secure noise rate they would actually witness experimentally, but we shall see that for the four protocols under consideration here the noise rate is simply twice the bit error rate.

For either ESC protocol, Eve's best attack is to use *both* Alice's and Bob's trines for measurement, half the time pretending to be Alice and the other half Bob. This holds for the tetrahedron as well and is due to the minimum in the lower bound of equation 5.8. By pretending to be Alice, Eve can guess Bob's result fairly well,

but Alice's less so: $I(B:E)$ is large but $I(A:E)$ small. By pretending to be Bob, the situation is reversed, with the mutual information quantities changing roles. By mixing the two strategies, Eve increases the minimum knowledge she has about either party's bit string. In [123] the scheme in which Eve pretends to be Bob is noted to be the measurement maximizing her mutual information with Alice; however as the analysis stops there and doesn't proceed to consider the rate bounds, it's insufficient as a cryptographic analysis.

To determine the mutual information quantities as functions of η, it suffices to consider first the case in which Eve intercepts every signal and uses Alice's ensemble for measurement. With these quantities in hand, we can mix Eve's strategies appropriately and then include her probability of interception. We begin with the trine. Given a signal state from Alice, there are two cases to consider. Either Eve measures and gets the same state, which happens with probability 2/3, or she obtains one of the other two results, with probability 1/6 for each. Whatever her outcome, she passes the corresponding state along to Bob and guesses that it was the state sent by Alice, *unless the subsequent exchange of classical messages eliminates this possibility*, at which point she reserves judgment.

Suppose her outcome corresponds to Alice's signal, and thus no disturbance is caused. Naturally, Alice and Bob go on to establish a bit half the time, a bit whose value Eve now knows. On the other hand, should her outcome not coincide with Alice's signal, there are two further possibilities. Half the time Bob obtains a result consistent with Alice's signal, and a further half the time the protocol succeeds. However, for this round to succeed, the required messages will eliminate Eve's outcome as Alice's signal, thus forcing Eve to abandon her guess. In the other case, Bob's result is orthogonal to Alice's signal, and from here the pair are guaranteed to think the procedure was a success, but also guaranteed to compute different bit values. Eve's guess at the bit value corresponds to Bob's in this case. Figure 5.12 lays out this train of thought.

Putting all this together, one obtains that the protocol succeeds with probability

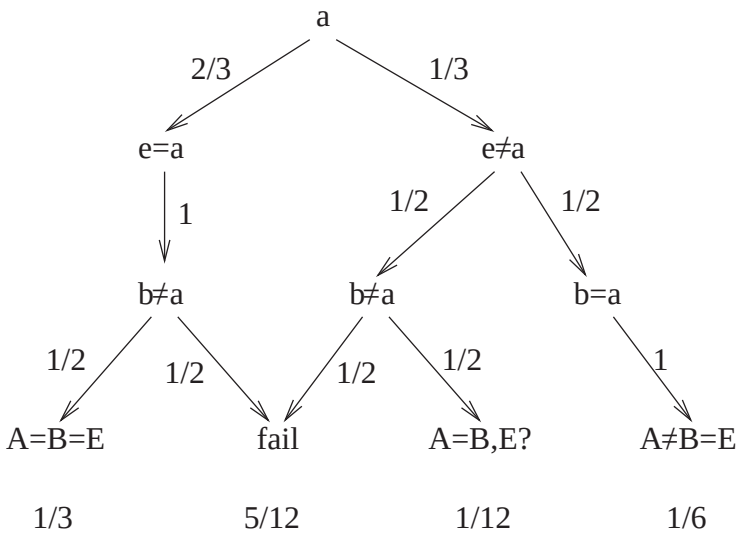

Figure 5.12: Thinking through the various cases when Eve measures the intercepted signals using Alice's ensemble. The variables a, b, and e correspond to signals or measurement outcomes for Alice, Bob, and Eve, while A, B, and E refer to key bit values.

7/12. Of the key bits created, Bob agrees with Alice with probability 5/7, while Eve agrees with overall probability 4/7, only guessing at all with probability 5/7. Eve and Bob agree on their bit values with probability 5/7. These numbers are obtained by considering the raw probabilities of agreement and renormalizing by 12/7. Should Eve instead measure the signals using Bob's trine ensemble, her agreement probabilities with Alice and Bob are swapped. Mixing the two eavesdropping strategies yields her an agreement probability of 9/14, an error probability of 1/14, and a no-guess probability 2/7. To interpolate between the endpoints of no action and full interception, note that to condition on the cases of successful bit creation, the probability of bit agreement must be renormalized by the probability of success. This probability depends linearly on η: $p_{\text{success}} = (6+\eta)/12$. All probabilities must therefore contain $6+\eta$ in the denominator, whence we may derive the form of the probabilities that Bob's and Eve's bit values correspond to Alice's:

$$p_{a=b} = \frac{6-\eta}{6+\eta} \qquad p_{a=e} = \frac{9\eta}{2(6+\eta)}. \tag{5.37}$$

Eve's probability to not guess at all is $2(3-2\eta)/(6+\eta)$.

By determining the probability of error in Alice's and Bob's bit strings as a function of η, we may compare to other protocols. For the trine, errors occur in the key string with probability $2\eta/(6+\eta)$. Using the expressions in equation 5.37 in the rate bound, one obtains that $R = 0$ corresponds to a maximum tolerable error rate of 20.4%. This compares favorably with the BB84 protocol's maximum tolerable error of 17.1% [57].

Analysis of the tetrahedron protocol proceeds similarly by examining the various cases. In this case, when $\eta = 1$ the failure rate of the protocol drops to 5/9, while Alice and Bob agree with probability 5/8, Eve has probability 7/16 of knowing Alice's or Bob's bit value, and she reserves judgment half the time. As the success rate of the protocol increases with η as $(3+\eta)/9$, we may determine the form of the probabilities using the same method to be

$$p_{a=b} = \frac{6-\eta}{2(3+\eta)} \qquad p_{a=e} = \frac{7\eta}{4(3+\eta)}, \tag{5.38}$$

while the error rate in the key string is $3\eta/2(3+\eta)$ and Eve's probability of not guessing is $(3-\eta)/(3+\eta)$. Again using these probabilities in the rate bound yields a maximum error rate of 26.7%. Like before, this compares favorably to the maximum tolerable error rate in the six-state protocol of 22.7%.

Key error rates in all protocols under consideration translate directly into noise rates by the same factor, so the bit error rate itself provides an appropriate means of comparison, as we now show. Again considering the quantum channel to be a depolarizing channel instead of that arising from Eve's interactions, first take the case of unbiased bases. If the state Bob receives is maximally-mixed, every outcome has the same probability to occur. In the cases the protocol succeeds, Bob therefore has a 50% chance of error. As much is true for the spherical code protocols, too. If Bob has a uniform probability of any outcome, then the one orthogonal to Alice's signal generates an error, and always causes the protocol to succeed. The remaining $n-1$ outcomes lead to identical bits for Alice and Bob, but only succeed with probability $1/(n-1)$. Hence in total, Bob again has a 50% chance of error. It should not be surprising that the same holds for both protocols, since if the signal

state is totally depolarized then it carries no information and the two key strings will have zero mutual information. To obtain noise rates from bit error rates for these protocols, we simply multiply by two.

Eve's attack could be gentler, however. In the version of the attack already considered, the measurement POVM consists of subnormalized projectors onto the code states in addition to an element proportional to the identity operator, corresponding to the case in which Eve doesn't intercept the signal. Instead of this incoherent mixture of measurement and no measurement, Eve might "coherently" mix the two by forming a set in which each effect operator is a linear combination of the identity operators and the same projectors; that is, we would change the POVM in the following way.

$$\left\{\eta\frac{d}{n}\Pi_0,\ldots,\eta\frac{d}{n}\Pi_{n-1},(1-\eta)I\right\}\rightarrow\left\{\eta\frac{d}{n}\Pi_k+\frac{1-\eta}{n}I\right\}. \tag{5.39}$$

Both measurements have the same outcome statistics; the crucial difference is that the *dynamics* associated with the new POVM can be different than the original. Using the square root of each measurement operator for the dynamics, the new measurement yields Eve more information for the same amount of disturbance. Mathematically, what once was a straight-forward wavefunction collapse onto the measured state now becomes a slightly more complicated transition

$$\rho\rightarrow\frac{\sqrt{E_k}\rho\sqrt{E_k}}{\mathrm{Tr}[E_k\rho]}\qquad E_k=\eta\frac{d}{n}\Pi_k+\frac{1-\eta}{n}I. \tag{5.40}$$

For the BB84 protocol, this attack was determined to be optimal when Eve doesn't wait to hear in which basis the signal was prepared [109].

As with the general spherical code protocols, we must take care in this gentler attack to properly translate Eve's measurement into her protocol for guessing the key bit. In particular, she should again reserve judgment in case the classical messages exclude her measurement outcome. This is easily accomplished by letting the computer do the bookkeeping, and to this end *Mathematica* was again enlisted. Since Eve causes less disturbance for the same information gain, the maximum tolerable

	BB84	Trine	Six-state	Tetrahedron
Normal IR	17.1	20.4	22.7	26.7
Gentle IR	15.3	16.6	21.0	22.6

Table 5.1: Maximum tolerable bit error rates for the four qubit-based protocols under consideration for the two versions of the intercept/resend attack. Doubling the figures yields the maximum tolerable noise rate, defined as probability of total depolarization in a uniform channel.

error decreases: the trine tolerates 16.6%, as opposed to 15.3% for its cousin BB84. The tetrahedron remains the most robust, tolerating a maximum error rate of 22.6%, as compared to 21.0% for the six-state protocol. These figures along with those corresponding to the normal intercept/resend attack are summarized in table 5.5.

Though the gap has narrowed between the ESC and MUB protocols, there's good reason to think that the spherical codes will remain more robust in the face of increasing eavesdropping. Suppose for a moment that Eve is granted the "super-quantum" power to make a perfect copy of an arbitrary quantum state, but she is still bound by the structure of quantum measurements. Now she can crack any protocol using unbiased bases by simply copying each signal state, and measuring it in the basis announced on the classical channel. When faced with the trine or tetrahedron protocol, Eve cannot obtain perfect information about the key in this manner. Now the classical messages narrow the set among which Eve must discriminate, but unfortunately for her, these remaining possibilities are not orthogonal. Hence returning to the situation in which Eve can only make low-quality copies, the protocol offers a tiny amount of "residual secrecy" which Alice and Bob may exploit. Apart from the noise introduced in trying to copy the state, Eve must also deal with the nondeterministic state discrimination difficulty. This may play an important role in more worrisome eavesdropping attacks on channels which contain loss, as we'll see in the next chapter.

The qubit protocols inherit the ability to determine the error rate from the success rate, though now as the channel becomes noisier and Bob's outcome becomes less

correlated to Alice's signal, the success rate *increases*. Of course, not all of this increase provides useful key: most of it leads to errors. But Eve cannot substitute signals solely for the purpose of modifying the success rate, as her signals won't be correlated with Alice's and will therefore also lead to an increase in the success rate. Hence Alice and Bob are safe in using the success rate to estimate the error rate.

In this chapter we've seen how quantum mechanics has the potential to both ruin and rescue cryptography as well as the role spherical codes can play. Unconditional security remains to be established, but by studying the intercept/resend attack we have outlined the main features: a streamlined protocol which automatically estimates channel noise, a wider range of speed and robustness, as well as a complementary security/speed tradeoff for fixed resources. No mention has yet been made of the practicalities of implementing such schemes, a topic taken up in the next chapter.

Chapter 6

Experimental Realizations

The previous chapter outlined how equiangular spherical codes may be used for quantum cryptography, as well as their strengths and weaknesses, at least in theory. Now we turn to the question of putting these ensembles into practice. The contents of this chapter also bear on all parts of the thesis, for it serves to examine both the details of what it means operationally to "prepare" and "measure" a quantum state and how one determines if it is done properly.

Spherical codes, designs, and frames can be realized in many physical systems, but here we focus on electromagnetic field modes since Alice and Bob invariably resort to them in any quantum communication setting. Easily produced and manipulated, electromagnetic modes retain quantum coherence over long times and distances, making them ideal for key distribution in which the parties are assumed to be separated by a large distance. Further, passive linear optics, together with photodetection, provides all the required tools for manipulating *isolated* quantum systems of a fixed dimension: state preparation, unitary transformations, and measurement. Though working with two-level systems based on polarization states is by far the easiest and already enjoys widespread use, systems with higher dimension may be created using the transverse spatial modes of the laser field, techniques which are still in their infancy. Quantum states so encoded may be transferred or "written"

into a superposition of various directional modes by using polarizing beam splitters and their spatial-mode analog for higher-dimensional systems. Passive linear optical elements such as beamsplitters and phase shifters then take over, allowing the implementation of any desired unitary transformation of the state. Ancillary states may be introduced by adding vacuum modes. Finally, measurement may be performed by ordinary photodetection.

In this context, working with spherical codes turns out to be no more difficult than with unbiased bases; indeed really *any* fixed-size ensemble in a fixed dimension requires roughly the same resources: real estate for the required modes. This resource demand is the reason why one should compare key distribution protocols in these terms, as was done in the previous chapter.

Beginning first by detailing how linear optics implements arbitrary unitary operations, we then consider the qubit case and the setups necessary for state preparation and measurement in section 6.2. From there we proceed to analogously to examine higher dimensional systems in section 6.3 before finally remarking on practical limitations of such realizations, especially the need for single-photon pulses and associated high-efficiency optical networks in section 6.4.

6.1 Linear Optics

Directional modes of the electromagnetic field provide a means of expressing an arbitrary quantum state. Each mode is described by a creation operator $a^\dagger_{\mathbf{k}}$, which acts to create an excitation in the mode propagating in direction $\mathbf{k}$. Ostensibly these modes are plane waves extending over all space, though we intend them here to be states emitted from a conventional laser. The spatial dependence of these beam modes are correctly described by making use of the paraxial approximation, whose details we examine later when considering transverse spatial mode encoding. For the moment it is sufficient to describe the mode by both the direction $\mathbf{k}$ and a point $\mathbf{x}$

through which it passes. In doing so, we are implicitly assuming that the spatial description of each mode is irrelevant and is confined to a small region transverse to the direction of propagation, i.e. a "laser beam". Even this is usually more than sufficient, and it will be simplest to label the modes just by a number, $1 \ldots d$ as it will be clear from the context what each mode describes.

By restricting the excitations in all modes to a contain just a single photon, each individual mode becomes a basis state in the space $\mathbb{C}^d$. Now passive linear optical elements may be used to implement arbitrary quantum operations in the one-photon sector. Only beamsplitters and phase shifters are required, though in principle elements which alter the polarization or the transverse spatial mode are also allowed. Since the linear elements do not alter the number of photons, we may describe their effects by their action on the creation and annihilation operators. Consider a single photon in the mode k, described by $a_k^\dagger|0\rangle$. An element involving d modes may then be described by a unitary matrix U_{jk}, which acts according to the following rule:

$$Ua_k^\dagger|0\rangle = Ua_k^\dagger U^\dagger|0\rangle = \sum_{j=1}^{d} U_{jk}a_k^\dagger|0\rangle, \tag{6.1}$$

where U is the associated *operator* in the space of quantum electromagnetic fields. A phase shifter simply imparts a phase to the propagation of the mode, and may be represented by $P_k(\phi) = \exp[i\phi]$ when acting on mode k. A beamsplitter simply mixes two modes with a particular strength, parameterized by an angle θ. When acting on modes j and k, it may be represented as

$$B_{jk}(\theta) = \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix}, \tag{6.2}$$

where the output modes are ordered such that the first input mode is reflected to the right and the second to the left, relative to their propagation directions. The intensity transmissivity and reflectivity are $\cos^2\theta$ and $\sin^2\theta$. Between the two elements we may implement any $d \times d$ unitary we desire, simply by breaking it into a sequence of operations acting on only one or two modes at a time. Such an array is termed an

optical multiport by Reck *et al.*, who first developed the method [127]. To break down an arbitrary operator, we proceed by a diagonalization process similar to Gaussian elimination. Step one is to find a sequence of beamsplitters $B_{d,d-1}B_{d,d-2}\ldots B_{d,1}$ (generally interspersed with phase shifters) such that $UB_{d,d-1}B_{d,d-2}\ldots B_{d,1}$ is block-diagonal, consisting of a $(d-1)\times(d-1)$ unitary matrix in modes $1,\ldots,d-1$ and a diagonal entry $\exp[i\alpha]$ in the dth mode. Proceeding recursively through each dimension builds up a sequence of beamsplitter unitaries (interspersed with phase shifters) which is the inverse of the matrix U. For d modes, then, $\binom{d}{2}$ beamsplitters and phase shifters each will generally be required. Figure 6.1 shows the case of three modes.

As an aside, although in principle such a multiport could be used for quantum computation, scaling limitations prevent this from being practical. Again, d modes are required for simulating a d-dimensional system with linear optical elements, and for useful computations, d would be huge. One main advantage of quantum computation is that two-level *physical* systems, e.g. spin states of ions or neutral atoms, can be used to build up a huge Hilbert space, and the number of atoms, say, would only be roughly the logarithm of the dimension.

The optical multiport realizes unitary operations on a set of modes, but these modes are not practical for transporting quantum states from point to point. Continuing to encode information spatially, one would either require a set of d optical fibers or need to establish the modes as separate channels in free space. Both are impractical, and as we shall see, wasteful. By using short-duration pulses, *temporal* modes can be hewn from a single spatial channel, but decoding the resulting (coherent!) pulse train back to separate spatial channels requires fast, efficient, and reliable switching, which is not currently available. Instead, the quantum state can be encoded into a single mode, using degrees of freedom beyond simply the direction of propagation. In particular, the transverse field mode can be so utilized for arbitrary dimensions, but for two-level systems the method is more straightforward, since the two orthogonal polarization states of the field can be used.

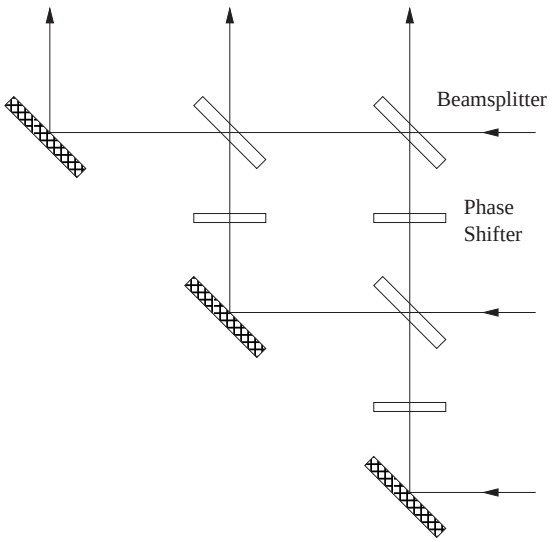

Figure 6.1: A generic optical multiport for three modes, input on the right, and output on top. Three beamsplitters and three phase shifters are required to implement a generic 3×3 unitary operator.

6.2 Polarization Qubits

Polarization-state preparation is trivial for Alice: ordinary polarizers and waveplates suffice to create states as desired. To make one of a set of states on demand, Alice could simply use a set of polarizers, as well as shutters to activate the desired one in any given instance. Figure 6.2 Measurement at Bob's end utilizes the methods just developed in the previous section. First the polarization state is converted to a mode state by means of a polarizing beam splitter and waveplate, after which point the polarization need not be referred to again. The polarizing beamsplitter first entangles the polarization and mode states, taking the input state $|\psi_{\text{in}}\rangle = (\alpha|\updownarrow\rangle + \beta|\leftrightarrow\rangle)|0\rangle$ to the output state $|\psi_{\text{out}}\rangle = \alpha|\updownarrow\rangle|0\rangle + \beta|\leftrightarrow\rangle|1\rangle$. With a suitable waveplate in one of the output beams we may rotate that polarization to match the other, thereby completely erasing the polarization information. In what follows we'll assume that this is the case and absorb into the polarizing beamsplitter element this further mode-dependent polarization rotation. Then in the diagrams to follow, the polarizing beamsplitter's effect on the state can be written $(\alpha|\updownarrow\rangle + \beta|\leftrightarrow\rangle)|0\rangle \rightarrow (\alpha|0\rangle + \beta|1\rangle)|\updownarrow\rangle$. Generic

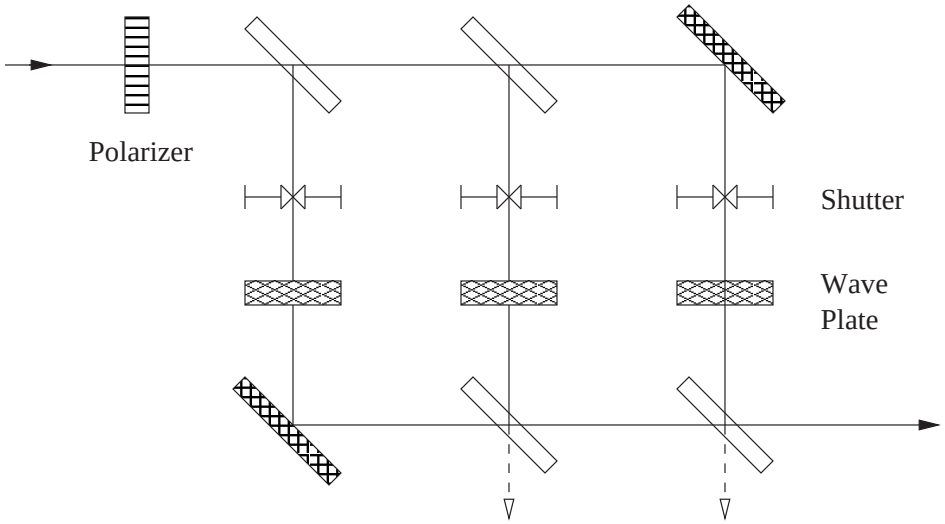

Figure 6.2: An optical network Alice might use to make the trine states. A polarized beam is input at the upper left and distributed by the beamsplitters to the different paths. By adjusting the wave plates appropriately, Alice may create any trine state she desires simply by opening the corresponding shutter. Note that no matter which state is created, it travels through the same number of beamsplitters, so that the input beam is attenuated the same for each state. To ensure a single-photon output, Alice adjusts the input intensity appropriately. This scheme doesn't require changing any optical elements except the shutter to prepare the state.

quantum measurements may be implemented by finding a Neumark extension and rotating the initial mode basis into the appropriate measurement basis. The process is completed by using photodetectors on the output modes to finally register an outcome. In chapter two, the Neumark extension was shown to be another name for frame dilation, and generall the situation is as follows.

Suppose $\{E_j = r_j|\phi_j\rangle\langle\phi_j|\}_{j=1}^n$ is a POVM to be implemented, where r_j takes care of the normalization. Including the unpopulated auxiliary modes, write the input state as

$$|\psi_{\text{in}}\rangle = (\alpha|\updownarrow\rangle + \beta|\leftrightarrow\rangle)|0\rangle + \sum_{k=1}^{n-1} \gamma_k|\updownarrow\rangle|k\rangle \ ,$$

where all the $\gamma_k = 0$ initially. After transcribing the polarization state into channel

mode states, the state becomes

$$|\psi'_{\rm in}\rangle = (\alpha|0\rangle + \beta|1\rangle + \sum_{k=2}^{n-1} \gamma_k |k\rangle)|\updownarrow\rangle \ ,$$

so that we can now omit reference to the polarization. The entire linear optical transformation in the optical network is given by a unitary U, transforming the input state to the output

$$U|\psi'_{\rm in}\rangle = \sum_j |j\rangle\langle j|U|\psi'_{\rm in}\rangle \ .$$

To measure the desired POVM, we must find U such that $|\langle j|UP|\psi'_{\rm in}\rangle|^2 = r_j|\langle\phi_j|\psi\rangle|^2$, where $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is the initial polarization state (now in the mode basis) and P projects onto the $|0\rangle, |1\rangle$ subspace. This leads immediately to

$$\langle j|UP|\psi'_{\rm in}\rangle = \sqrt{r_j}e^{i\mu_j}\langle\phi_j|\psi'_{\rm in}\rangle \ ,$$

where μ_a is an arbitrary phase that could be absorbed into the POVM states. This works just in case

$$U_{j0} = \langle j|U|0\rangle = \sqrt{r_j}e^{i\mu_j}\langle\phi_j|0\rangle \qquad \text{and} \qquad U_{j1} = \langle j|U|1\rangle = \sqrt{r_j}e^{i\mu_j}\langle\phi_j|1\rangle \ .$$

Apart from the normalizing factor and the phase, the first two elements in row j of the unitary transformation—i.e., the amplitudes to go from $|0\rangle$ and $1\rangle$ to $|j\rangle$—are the complex conjugates of the amplitudes of the states $|\phi_j\rangle$ comprising the POVM.

To see this procedure in action, consider the trine measurement, which may be achieved using only two beamsplitters. If the trine states are given by

$$|\phi_1\rangle = |0\rangle \tag{6.3}$$

$$|\phi_2\rangle = -\frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|1\rangle \tag{6.4}$$

$$|\phi_3\rangle = -\frac{i}{2}|0\rangle - i\frac{\sqrt{3}}{2}|1\rangle, \tag{6.5}$$

then a 2:1 (transmission/reflection) and 1:1 pair of beamsplitters suffice to realize the Neumark extension, as shown in figure 6.3. The associated unitary matrix describing

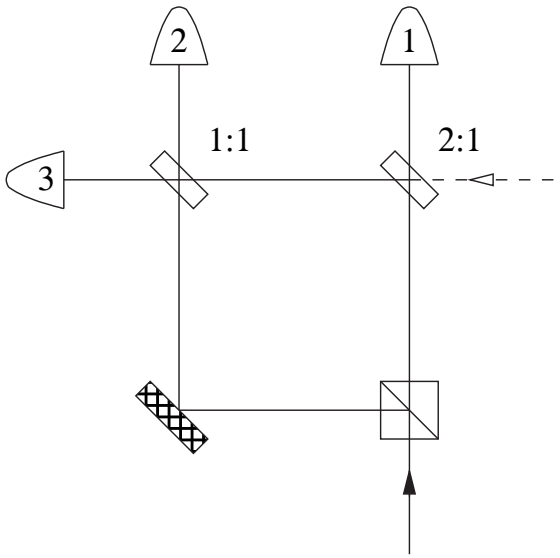

Figure 6.3: Passive linear optical implementation of the trine measurement. The polarizing beam splitter "writes" the quantum state onto modes one and two. Mode three, in the vacuum state, enters the first beamsplitter at the upper right. The two beamsplitters, 2:1 and 1:1 in transmission to reflection intensity, transform the mode basis to the Neumark-extended trine basis, and qubit polarization states input on the lower right are thus measured by the trine via the photodetectors.

the two beamsplitters may be written

$$U_{\text{trine}} = \sqrt{\frac{2}{3}} \begin{pmatrix} 1 & 0 & i/\sqrt{2} \\ -1/2 & \sqrt{3}/2 & i/\sqrt{2} \\ i/2 & i\sqrt{3}/2 & 1/\sqrt{2} \end{pmatrix}. \tag{6.6}$$

Note that the first two elements of each row are the trine states themselves, renormalized to $\sqrt{2/3}$, so that the input mode states are indeed suitably transformed.

The trine measurement is less involved than making the measurement of two conjugate bases as in the BB84 protocol. Figure 6.4 shows how this measurement can be realized with three 50:50 beamsplitters.

Fortunately the tetrahedron measurement is also simple to perform, requiring only four beamsplitters, as shown in figure 6.5. This corresponds to the measurement

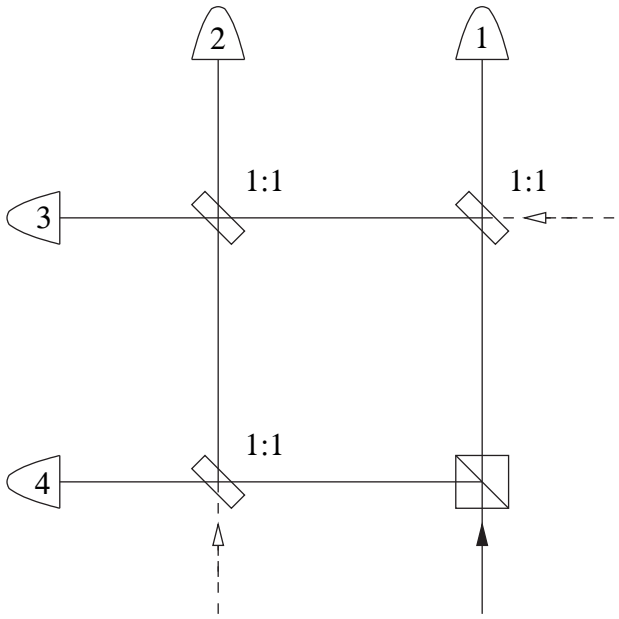

Figure 6.4: The BB84 measurement. The polarizing beam splitter "writes" the quantum state onto modes one and four. Modes two and three enter in the vacuum state. All beamsplitters are 50:50, resulting in a measurement of the two conjugate bases.

of the following tetrahedron states:

$$|\phi_1\rangle = |0\rangle, \tag{6.7}$$

$$|\phi_2\rangle = -\frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle, \tag{6.8}$$

$$|\phi_3\rangle = -\frac{1}{\sqrt{3}}|0\rangle + e^{-i\frac{2\pi}{3}}\sqrt{\frac{2}{3}}|1\rangle, \tag{6.9}$$

$$|\phi_4\rangle = -\frac{i}{\sqrt{3}}|0\rangle + e^{-i\frac{5\pi}{6}}\sqrt{\frac{2}{3}}|1\rangle. \tag{6.10}$$

Such inelegant-looking states are the result of simplifying the number of beamsplitters and phase shifters required.

In the measurement schemes presented up till now, we eschew the use of elements which reference the polarization directly after the initial polarizing beamsplitter and waveplate, but sometimes retaining polarization information in the optical network makes the measurement more compact. In this picture BB84 becomes slightly sim-

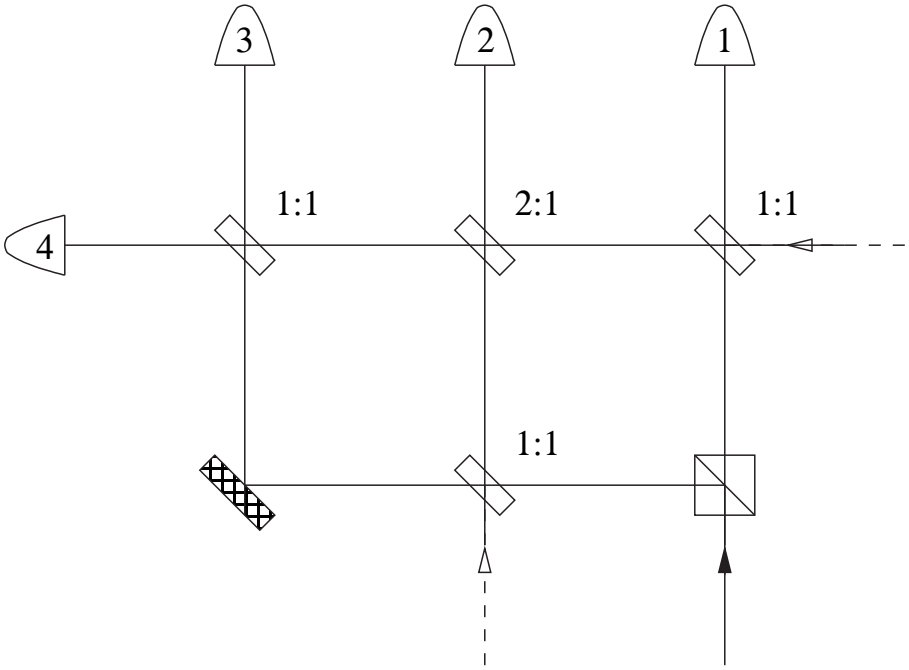

Figure 6.5: The tetrahedron measurement. The polarizing beam splitter writes the quantum state onto modes one and three. Modes two and four, both in the vacuum state, enter along the dotted lines.

pler, and the measurement of the three unbiased bases vastly more so. To measure the combination of horizontal/vertical (+) and diagonal (×) polarization bases, it suffices to use an ordinary 50:50 beamsplitter, two polarizing beamsplitters, and one waveplate. A polarizing beamsplitter by itself effectively measures in the basis corresponding to its orientation. Thus to make the measurement consisting of both bases, halve the input among the two possibilities with an ordinary beamsplitter. The waveplate is inserted in one output to transform the resulting measurement from + to ×, and the polarizing beamsplitters finish the task. With one more ordinary beamsplitter, polarizing beamsplitter, and waveplate we may realize the measurement of three unbiased bases, and so on. Now the first beamsplitter is 2:1 and the second 1:1, but otherwise the scheme is perfectly analogous. Without making use of the polarization states in the various modes, i.e. without using waveplates, this measurement could require up to 15 beamsplitters and phase shifters each, instead

of the seven optical elements used here. The trine and tetrahedron measurements have also been performed in this context, with rms errors in the observed statistical distributions of a few percent for each [39].

6.3 Higher-Dimensional Systems

To encode quantum states of higher dimension into a field mode, something beyond simple polarization must be used. For this purpose we may turn to the spatial field mode transverse to the direction of propagation. In this context we're thinking mainly in terms of free-space cryptography, rather than via optical fibers, which tend to offer only very noisy multi-mode channels. A proper description of the available modes is given by the paraxial approximation. Consider a monochromatic electric field propagating in the z direction described by the expression $\mathbf{E}(\mathbf{x},t) = \mathbf{E}(\mathbf{x})\exp[i(kz-\omega t)]$. Let the spatial envelope be given by the function $\mathbf{E}(\mathbf{x}) = E_T(\mathbf{x})\hat{\mathbf{e}} + E_L(\mathbf{x})\hat{\mathbf{z}}$, where $\hat{\mathbf{e}}$ is the polarization vector, and E_T and E_L are transverse and longitudinal components, respectively. If the transverse portion of the function varies slowly with respect to the wavenumber k in the z direction, i.e. $\partial E_T/\partial z \ll k$, then the wave equation may be simplified to the paraxial wave equation

$$\nabla_T^2 E_T + 2ik\frac{\partial E_T}{\partial z} = 0. \tag{6.11}$$

Here ∇_T^2 is the two-dimensional Laplace operator pertaining to the transverse directions x and y. This equation looks strikingly similar to the Schrödinger equation, and in fact in a proper treatment of field quantization, one obtains transverse mode wavefunctions obeying this equation.

As such, there are a variety of useful solutions of the paraxial wave equation, but two are more often-used than the rest: Hermite-Gaussian and Laguerre-Gaussian modes. The former are simply products of eigenstates of the one-dimensional harmonic oscillator, i.e. solutions found by separating variables in the Cartesian coordinate system, while the latter are solutions in cylindrical coordinates. The Hermite-

Gaussian beam HG_{mn} is given by

$$\begin{aligned} E_{mn} &= \frac{C^H_{mn}}{w(z)} \exp\left[-ik\frac{x^2+y^2}{2R(z)}\right] \exp[-i(kz-\psi^G_{mn})] \\ &\quad \times H_m\left(\frac{x\sqrt{2}}{w(z)}\right) H_n\left(\frac{y\sqrt{2}}{w(z)}\right) \exp\left[-\frac{x^2+y^2}{w(z)^2}\right]. \end{aligned} \tag{6.12}$$

Here k is the wavenumber $2\pi/\lambda$, one of a host of constants, functions, and beam parameters contained in this expression. First, H_m is a Hermite polynomial of order m, defined by $H_m(u) = (-1)^m e^{u^2} \frac{\mathrm{d}^m}{\mathrm{d}u^m} e^{-u^2}$. Out front is the normalization constant C^H_{mn}, defined by $C^H_{mn} = (2/2^{n+m} n! m! \pi)^{1/2}$. The remaining functions are beam parameters, starting with the width $w(z)$, defined in terms of the nominal width, or *beam waist*, w_0 as $w(z) = w_0\sqrt{1+(\lambda z/\pi w_0^2)^2}$ for wavelength λ. The function $R(z)$ is the wavefront's radius of curvature at position z, determined by $R(z) = z(1+(\pi w_0^2/\lambda z)^2)$. Finally, ψ^G_{mn} is the so-called Guoy phase, given by $\psi^G_{mn} = (m+n+1)\arctan(\lambda z/\pi w_0^2)$. The numbers n and m determine the number of nodes in the x and y directions, as shown in the intensity profiles in Figure 6.6.

Meanwhile, the Laguerre-Gaussian beam LG_{pl} is written in cylindrical coordinates (r, ϕ) as

$$\begin{aligned} E_{pl} &= (-1)^p \frac{C^L_{pl}}{w(z)} \exp\left[-\frac{ikr^2}{2R(z)}\right] \exp[-i(kz-\psi^G_{pl})] \\ &\quad \times \exp[-il\phi] \left[\frac{r\sqrt{2}}{w(z)}\right]^{|l|} L_p^{|l|}\left(\frac{2r^2}{w(z)^2}\right) \exp\left[-\frac{r^2}{w(z)^2}\right]. \end{aligned} \tag{6.13}$$

Now the normalization constant becomes $C^L_{lp} = \sqrt{2(p)!}\Big/\sqrt{\pi(|l|+p)!}$ while the Guoy phase is $\psi^G_{pl} = (2p+|l|+1)\arctan(\lambda z/\pi w_0^2)$. The $L^p_|l|(u)$ are the associated Laguerre polynomials. The radial index $p \in \mathbb{Z}$ determines the number of radial nodes $p+1$ and the angular index $l \in \mathbb{N}$ the angular nodes.

The phase of the field does not depend on the transverse spatial coordinates for Hermite-Gaussian beams, except so as to show a spherical wavefront of curvature $R(z)$ at position z. Laguerre-Gaussian modes, on the other hand, exhibit a screw-like wavefront due to the "winding" phase dependence on ϕ and the associated singularity at the origin. Such a wavefront is shown in figure 6.7. Interestingly, this spiraling

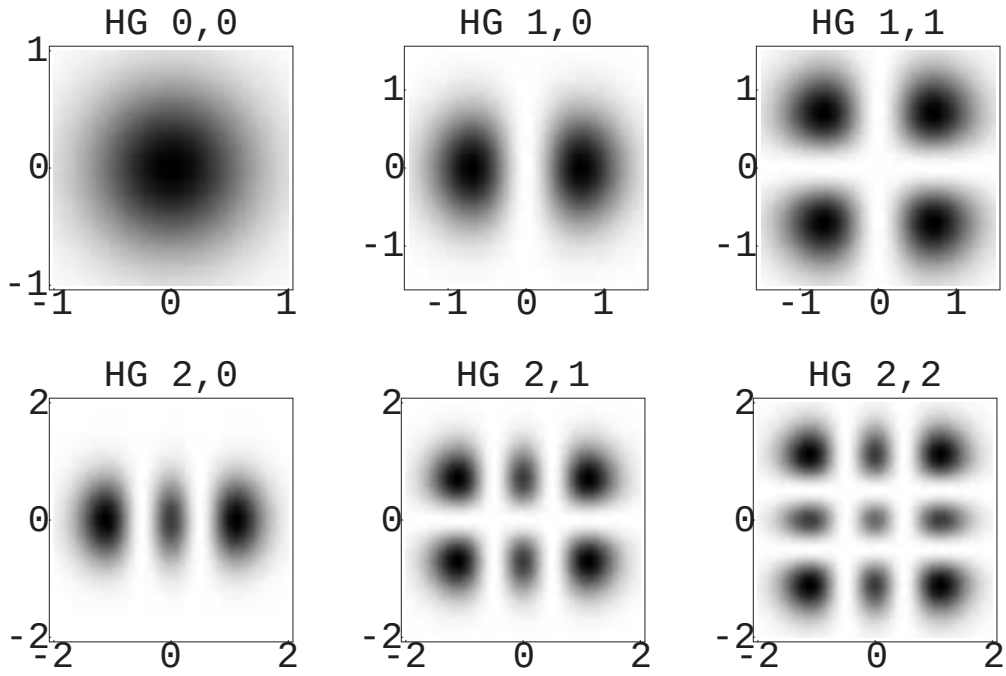

Figure 6.6: Intensity plots of the first few Hermite-Gaussian modes. Note the increasing area footprint with increasing mode number.

wavefront carries angular momentum, playing the role of orbital angular momentum complementing the intrinsic "spin" of the angular momentum carried by the polarization [3]. A photon in an LG_{pl} mode carries angular momentum $l\hbar$, as verified in experiment [76].

Before describing how to prepare and measure such states, we return to the point alluded to at the beginning of section 6.2, the wastefulness of using many distinct spatial channels compared to using distinct spatial modes of a single beam [158]. To communicate a d-level quantum state using distinct beams, d beam are required. The fundamental LG_{00} =HG_{00} Gaussian beam is best for this purpose, since it has the most compact profile for a given waist w_0. By packing the beams in a hexagonal or square fashion we can fit them all into an area of roughly $4w_0^2 d$ while avoiding crosstalk between channels. In contrast, the effective size of an LG or HG mode is proportional to the square-root of its order [145]. Hence the first d modes fit into

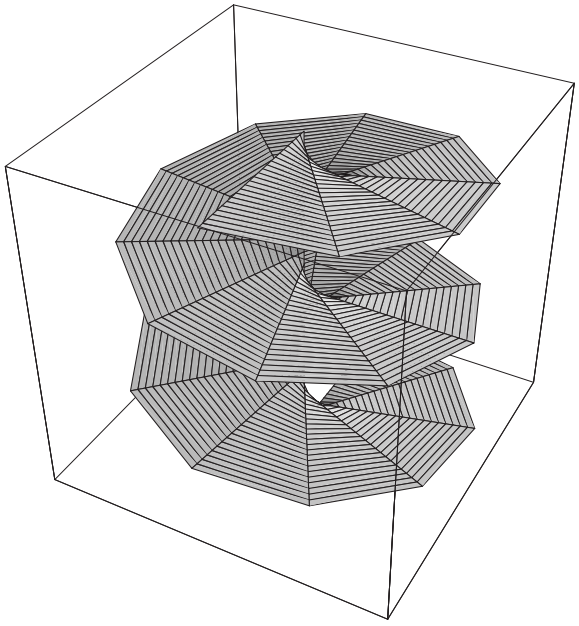

Figure 6.7: Screw-like wavefront of Laguerre-Gaussian modes. The phase singularity at the origin implies zero intensity on the beam axis.

a square of area approximately $4w_0^2\sqrt{d}$, a major improvement. Put differently, if the area A supports d distinct channels, then it could instead support d^2 transverse spatial modes. Both methods rely on an increased spatial area, but the transverse modes utilize this area more efficiently [117].

6.3.1 State Preparation

Now we turn to state preparation. In contrast to encoding two-level quantum states in polarization states, creation of arbitrary spatial mode states is not so easy in practice, though simple-enough in principle. Physically creating *specific* states can be accomplished in a variety of ways. For instance, Hermite-Gaussian modes may be created by introducing thin wires into the laser cavity corresponding to node positions, and astigmatic optics can convert a HG_{mn} mode into an LG_{pl} Laguerre-Gaussian mode with $l = |m-n|$ and $p = \min(l, m)$ [10]. These basis states could

then be combined into arbitrary superpositions with an interferometer, though one would be required for each state Alice wished to make on demand.

In principle, all that is required to change one mode state into another is the ability to affect the phase of the transverse field at every point. By rephasing an incoming beam in just the right fashion it can be transformed into any other mode state. If we label the input beam by $E_r \exp[i\phi(x,y)]$, then a phase transformation takes it to $E_r \exp[i\phi(x,y)] \exp[if(x,y)]$, where $f(x,y)$ is the phase modulation function. Trivially, and immediately, this scheme affects the *phase* $\phi(x,y)$ of the beam, but to affect the distribution of *intensity* E_r, the appropriate rephasing $f(x,y)$ must be chosen such that the beam subsequently propogates with the desired intensity pattern.

For a one-step solution suitable for state-prepratation we may turn to computer-generated holograms [9, 5, 152, 128]. In principle, an amplitude transmission hologram can transform a given input state into an arbitrary desired output state by recording the intensity interference pattern $I(x,y)$ between the two. Suppose the input beam (the reference beam) is described by $E_r \exp[i\mathbf{k}_r \cdot \mathbf{x}]$ and the desired output (the object) by $E_o \exp[i\mathbf{k}_o \cdot \mathbf{x}]$, where the functions E_r and E_o contain the transverse phase and amplitude profile of the corresponding beams. Without loss of generality we take $\mathbf{k}_r = k\hat{\mathbf{z}}$, and superimposing the beams at the hologram ($z = 0$) then gives the intensity/transmission function $I(x,y) = |E_r + E_o \exp[i\mathbf{k}_o \cdot \mathbf{x}_T]|^2$. Here $\mathbf{x}_T$ is perpendicular to the direction of propagation, $\mathbf{z}$. Illuminating the hologram with the reference beam alone transforms it according to the rule

$$I(x,y)E_r e^{ikz} = \left(|E_r|^2 + |E_o|^2\right) E_r e^{ikz} + E_r^2 E_o^* e^{i(2k\hat{\mathbf{z}} - \mathbf{k}_o)\cdot\mathbf{x}} + |E_r|^2 E_o e^{i\mathbf{k}_o\cdot\mathbf{x}}. \quad (6.14)$$

If the input beam is taken to be a plane wave such that E_r is a real constant, then the third term is the the quantity of interest, representing the desired output beam. The first term is simply a continuation of the reference beam, amplitude-modulated by the intensity profile of the object, while the second is the conjugate image. The desired phase interference pattern may be determined by computer and from this a hologram generated simply by recording on holographic film. Figure 6.8 shows the

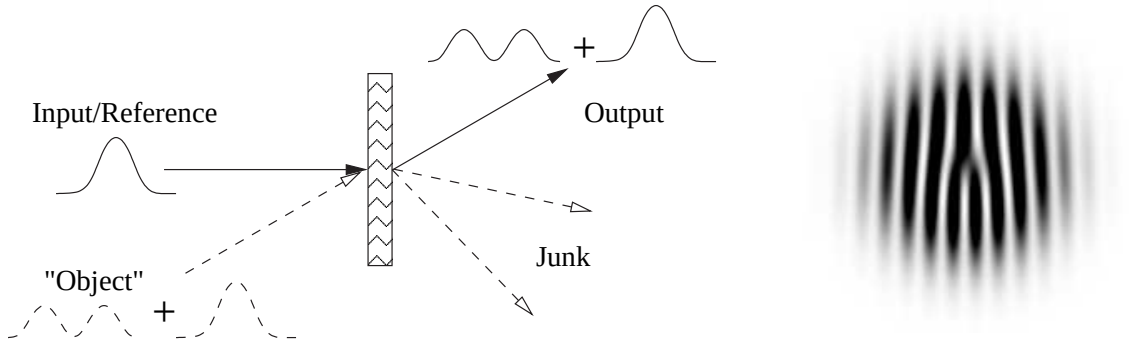

Figure 6.8: Setup for producing desired transverse states with a transmission hologram. Given the input (reference) beam and the "object" beam, the desired output, the phase pattern may be computed and transferred to holographic film. When illuminated with the input beam, the desired output emerges from the corresponding location on the other side of the hologram. Due to diffraction effects, other states are also created in various diffraction orders, but Alice may simply ignore these and select the desired mode. On the right is shown the (negative of the) interference pattern of an on-axis Gaussian beam and an $LG_{0,1}$ mode 30 degrees off-axis. This pattern may be recorded to holographic film in order to transform a Gaussian beam into a $LG_{0,1}$ beam, as well as superpositions of such states. The diffractive nature of the hologram is readily apparent.

interference pattern of an on-axis Gaussian beam and an $LG_{0,1}$ mode at 30 degrees off-axis.

Amplitude transmission holograms suffer from unavoidable diffraction effects inherent in their nature. For instance, the hologram shown is more or less a diffraction grating, plus a fork dislocation at the center. In contrast, a blazed grating is theoretically capable of 100% transmission efficiency in the desired diffraction order. To improve efficiency, a blazed phase hologram may be used instead of the amplitude hologram, returning to the abstract method outlined at the beginning of the section. For the case depicted in figure 6.8, we may simplify matters by considering the interference pattern of a plane wave and one with winding phase as in figure 6.7. The former is described by $E_r \exp[i(k_x x + k_z z)]$, propagating off the hologram axis at angle $\arctan(k_z/k_x)$, and the latter by $E_o \exp[il\phi] \exp[ikz]$. The resulting interference pattern becomes $I = E_r^2 + E_o^2 + 2E_o E_r \cos(k_x x - l\phi)$. Neglecting the constant terms, the

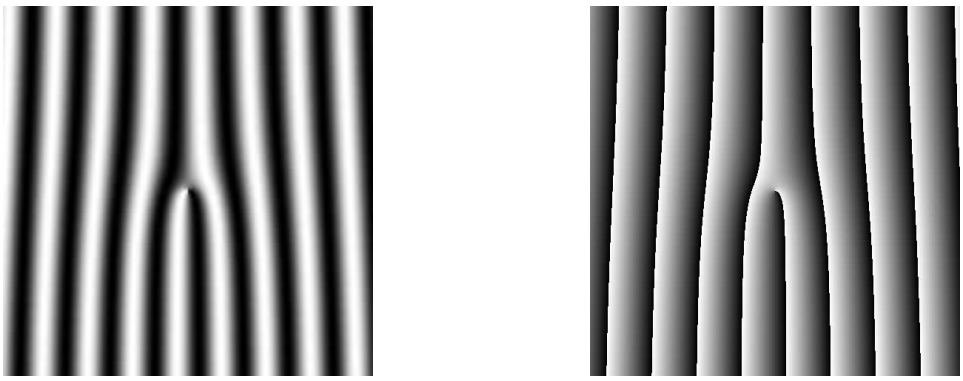

Figure 6.9: Hologram patterns for converting Gaussian beams into $p = 0, l = \pm 1$ LG modes. The intensity pattern for an amplitude transmission hologram is shown on the left and that for a blazed phase transmission hologram on the right. In the latter the grayscale level indicates the effective thickness of the hologram, so that it can be seen to correspond to a blazed diffraction grating, plus a fork dislocation.

bare (amplitude) hologram pattern is simply $I = \cos(k_x x - l\phi)$, shown in figure 6.9 with $l = 1$. To realize this as a phase transmission hologram, i.e. one with transmission function $T(r, \phi) = \exp[if(r, \phi)]$, we simply use the argument $k_x x - l\phi \bmod 2\pi$ to generate the amplitude hologram as before, and then bleach the film. This transfers the amplitude information into phase information by altering the index of refraction at each point. The hologram modulates the incident plane wave by this phase factor, producing $E_r \exp[i(k_x x + k_z z) \exp[-i\delta(k_x x - l\phi)]$, where δ is the magnitude of the phase modulation. Should $\delta = 1$, 100% of the incident beam would be modulated into the desired output. The blazed pattern is shown in figure 6.9.

Blazed phase holograms concentrate output intensity into the desired order, but the absorption and diffraction effects cannot be completely eliminated. This doesn't bother Alice, however, as she may adjust the input intensity and select the appropriate outgoing mode as required. In practice such computer-generated holograms have been used to create Laguerre-Gaussian states as well as superpositions of them [113], achieving roughly 14% efficiency in the desired output mode.

6.3.2 Measurement

In analogy with the qubit polarization-state case, Bob requires the equivalent of a polarizing beamsplitter and waveplate for the transverse modes. Once he has split the incoming state into various spatial channels by basis state, the "waveplate" is used to erase the transverse mode state. Then he may use an optical multiport to coax the quantum state into being measured by his desired POVM. In principle, the analog of a waveplate may be accomplished using a phase mask as described at the beginning of section 6.3.1. Bob could also make use of holograms, as does Alice, but high-efficiency devices must be fabricated in some other fashion. He doesn't have the luxury of increasing the input intensity; every signal absorbed by the hologram slows down the protocol. In any case, we'll simply consider whatever method is used in practice as part of the mode sorter. Fortunately, high-efficiency methods to accomplish the sorting have recently been developed. One method, used to sort states based on their angular momentum, uses Dove prisms and phase shifters in the arms of a Mach-Zender interferometer [102, 158, 103]. Another method sorts HG modes by using fractional Fourier transformers in the interferometer arms, a scheme which may also be used to sort LG modes based on the radial index [168, 159]. Since both the production and sorting of $p = 0$ LG modes is simplest, we'll consider the problem mainly from this perspective. Moreover, sorting of angular momentum states in this fashion has been carried out experimentally. Afterwards we shall comment on the similarities between this and the HG mode sorting method.

A Dove prism, which has a trapezoidal shape in the propagation direction, simply reflects the input image along the axis perpendicular to the top and bottom faces. Two prisms oriented relative to one another by an angle $\theta/2$ effect a rotation by θ, and by placing these together in one arm of an interferometer, the rotated and unrotated images can be made to interfere with each other. Due to the phase profile of LG modes, rotation by θ shifts the phase by an amount $l\theta$, so the LG_{pl} modes are eigenstates of the rotation operator $\mathcal{R}_\theta$ satisfying $\mathcal{R}_\theta[e^{il\phi}] = e^{il(\phi+\theta)}$. Thus by adjusting the path lengths of the two arms, modes with particular l values can be made

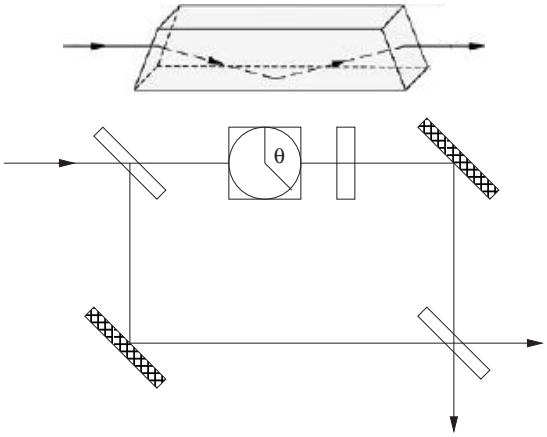

Figure 6.10: Illustration of a Dove prism and its use in a Mach-Zender interferometer setup to sort angular momentum modes. In the upper arm the beam acquires an l-dependent phase shift $l\theta$ from the two Dove prisms, while a phase shifter imparts a fixed phase irrespective of mode number. For $\theta = \pi$, modes with even l can be made to exit to the right and modes with odd l to the bottom.

to exit in certain directions after the beams recombine at the final beamsplitter, as shown in figure 6.10. Suppose for instance that the induced relative phase difference in the two arms is $l\pi$, accomplished with the two prisms oriented 90 degrees relative to one another. In this case modes with even l are unchanged but those having odd values of l are 180 degrees out of phase. Then recombining the beams at a beamsplitter causes even values to exit in one direction and odd values the other.

By cascading several interferometers one can immediately sort superpositions of modes in which l is a power of two, with the nth stage employing a rotation angle $\theta = \pi/2^n$. After the first stage $l = 1$ is output while the remainder go to the second stage. There l for which $l \bmod 4 = 0$ continue to the next stage as $l = 2$ is output, and so on, i.e. sorting in the manner of the Fourier transform. To construct a general sorting scheme, an ordinary phase shifter may be introduced into the interferometer. Using an overall phase shift of $-k\pi/2^n$ with the same rotations as before, modes with $l \bmod 2^n = k$ can be sorted into groups having $l \bmod 2^{n+1} = k$ and $l \bmod 2^{n+1} = k+2^n$. Suppose we want to sort the lowest seven angular momentum states, that is $|l| \leq 3$. In the first step we separate odd from even using $\theta = \pi$. We then direct the even

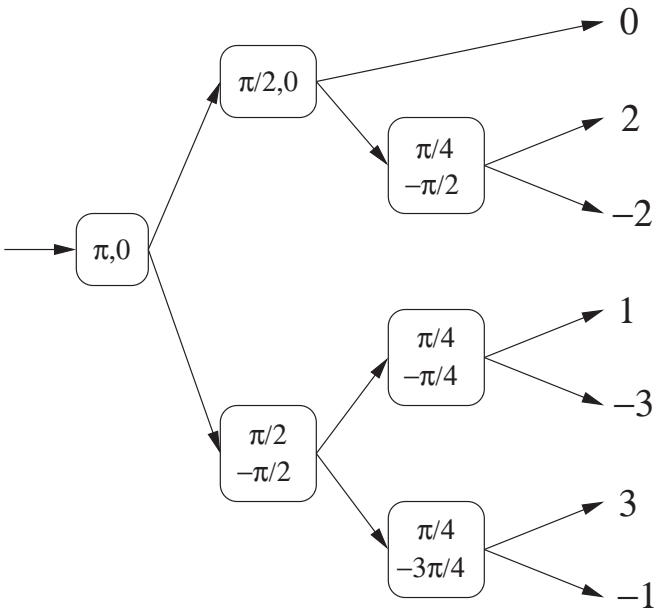

Figure 6.11: Mode sorter for the lowest seven angular momentum states, i.e. $|l| \leq 3$. The first phase refers to the l-dependent rotation and the second the global phase.

states $0, \pm 2$ into a second interferometer set to $\theta = \pi/2$. One output will be $l = 0$, while the remaining two go into a third interferometer set to $\theta = \pi/4$ plus a phase shift of $-\pi/2$, which separates $+2$ from -2. Meanwhile the odd l modes are sorted as follows. First, they are input into a system with $\theta = \pi/2$ and phase shift $-\pi/2$, which sends $-1, 3$ out one port and $-3, 1$ out the other. To sort the former pair, set $\theta = \pi/4$ with a global phase shift of $-3\pi/4$, and to sort the latter change the phase shift to $-\pi/4$. This is shown schematically in figure 6.11. Since all optical elements in the mode sorter are in principle high efficiency, Bob can combine the mode sorter and an optical multiport to implement any quantum measurement with efficiency limited primarily by the photodetectors.

Hermite-Gaussian beams may be sorted in essentially the same fashion, for the fractional Fourier transform affects HG modes the way the image rotation affects LG modes, though one may in addition act on x and y separately. Hermite functions are eigenstates of fractional Fourier transform operator $\mathcal{F}_\theta$ (which in quantum-mechanical language, is just the time-development operator for a harmonic oscilla-

tor), as it acts in accordance with the rule $\mathcal{F}_\theta[H_m(x)] = \exp[im\theta]H_m(x)$. Xue *et al.* [168] suggest an immediate method of sorting HG modes using a sequence of graded-index (GRIN) rods with quadratic profile $n(x) = n_0 - n_2x^2$ [106]. The length of these rods controls the parameter θ, which is analogous to the rotation angle when using Dove prisms to sort angular momentum states. We can then employ the same method to sort the spatial modes as for angular momentum states.

6.4 Practical Limitations

Alice and Bob would like to use high-efficiency measurement schemes, because if Alice sends more than one photon per signal, she is effectively handing Eve that many copies with which to break the protocol. This *photon number splitting attack* would not pose a problem but for omnipresent losses in any electromagnetic channel. To be on the safe side, Alice and Bob must assume that losses are due to Eve redirecting some of the sent photons, so sending more than one photon per signal is hazardous. In the standard protocols, unbiased bases offer zero protection against this attack, since knowledge of the basis allows Eve to determine the key letter with certainty.

Modifications have been proposed to combat this problem, which of course lower the attainable key rate [2, 134]. Use of spherical codes may offer a slight edge in combating this problem, as none of the states are orthogonal to begin with, and thus the classical messages never provide a deterministic procedure for learning the key letter. Returning to the comment at the end of subsection 5.4.3, ESC protocols may permit a slightly higher signal intensity to be safely used due to this "residual" security.

To see how this works, consider the trine protocol implemented over a lossy but otherwise noiseless channel. Eve learns one signal that Alice did not send from the classical channel in every successful round. Having narrowed the choices, she must

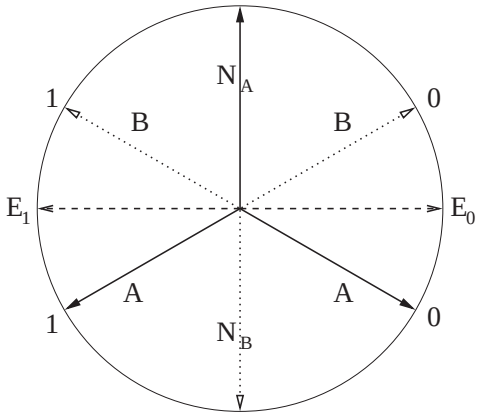

Figure 6.12: A diagram of the discrimination problem facing Eve. Communication over the classical channel has eliminated the vertical state from the possible signals sent by Alice, leaving the two black arrows labeled 'A.' Consequently, the optimal measurement Eve can make to determine which of the remaining two was actually sent is shown by the dashed arrows pointing to E_0 and E_1. Since the signal states are not orthogonal, there is a probability of $(2-\sqrt{3})/4 \approx 0.067$ of incorrectly identifying the state.

decide among the remaining two, which have overlap 1/4. The optimal measurement for doing this is simply a projection measurement along the perpendicular to the bisector of the two states, i.e. in the Bloch-sphere representation each outcome is 30 degrees from the corresponding trine state so as to be back-to-back, as shown in figure 6.12. Because the channel is otherwise noiseless, Bob's and Alice's key bits agree every time the protocol is successful, but Eve has a roughly 6.7% chance of error when measuring one copy of her own. This seemingly-small probability of error actually translates into only ≈ 0.65 bits of information, meaning the other ≈ 0.35 remain secret. In this manner Alice and Bob can give one copy to Eve and still expect to create a truly secret bit from about every three successful rounds of the protocol. For the tetrahedron the numbers are slightly better since the states have greater overlap, i.e. they are less orthogonal. Now one copy of the signal allows Eve to guess the key bit with error probability $\approx 9.2\%$. Again she makes the same orthogonal measurement perpendicular to the bisector of the states. The secret key rate given that Eve has one copy of her own then becomes ≈ 0.44. If Eve has three

copies and takes the majority of three independent measurements, she reduces the error rate to $\approx 2.4\%$. But this still leaves Alice and Bob with a key rate of ≈ 0.16.

In practice, Alice is likely to use strongly attenuated coherent states, which have a Poisson distribution in photon number. To suppress the probability of a multiphoton signal to one percent, for instance, Alice should attenuate the coherent state to have an average of just under 0.15 photons per pulse. The use of such faint pulses is responsible for most of the practical difficulties in implementing a truly quantum key distribution cryptographic system, since only very high-efficiency elements may be used in order to obtain a useful key generation rate.

By using the spherical code protocols, one copy of the signal may be granted to Eve in order to trade a lower key generation rate per successfully-received signal for a higher actual transmission rate. In order to restrict the probability of a three-or-more photon signal to one percent, Alice may safely send an average of just under 0.45 photons per pulse, a three-fold increase in intensity for a 56% cutback in key generation rate.

Other means of obtaining higher intensities are the subject of much current research, such as reliably sending only one photon per pulse to begin with. Such a single-photon source may be one half of an entangled two-photon source, such as those resulting from parametric down-conversion. This phenomenon occasionally provides a pair of entangled photons when a laser is incident on a type-I nonlinear crystal. Normally one imagines creating polarization-entangled states in this fashion, though recent experiments have demonstrated that the orbital angular momentum states of the photons are also entangled [113, 151]. Conventional state preparation may now be dispensed with in favor of measurement by exploiting the simple formulation of a maximally-entangled state in terms of either unbiased bases or spherical codes found at the end of section 5.4.3. However, in any scenario, low yields associated with small photon number coupled inevitable channel loss in the channel severely limit the speed of quantum cryptographic protocols and the distance over which they can be securely carried out.

Despite these difficulties, several experiments have made use of the relative simplicity of creating, manipulating, transmitting, and measuring electromagnetic modes and demonstrated the feasibility of quantum cryptography, both in fiber and free-space [121, 82, 25, 81]. Two commercial applications are even available for the security-conscious user with a hefty bankroll, both based on fiber-optic implementations. For free-space implementations, such as rekeying of satellites, the straightforward use of higher-dimensional protocols via encoding into transverse spatial modes make higher efficiency and security protocols within reach. The ability to rekey satellites on demand from the ground would prevent unauthorized use, as has been witnessed on commercial satellites in the past [26], as well as costly space shuttle missions to manually rekey sensitive military satellites [119].

Part IV

Epilogue

Chapter 7

Conclusion

7.1 Summary

From the foundations of quantum mechanics to the practicalities of ground-based rekeying of satellites, frame theory has helped see us through a wide-ranging terrain of topics. Though perhaps appearing as a gallimaufry of various research questions, the selection of topics here tends from foundational, theoretical questions toward practical, useful applications, without nearly running the gamut of different types of theoretical research a person can do in quantum information theory.

Fundamentally, frame theory and quantum mechanics are tightly interwoven as they make use of the same mathematical structures, so it is sensible to begin with an overview of frame theory as it applies to quantum information theory. In this thesis we've mainly considered questions involving properties of collections of measurement operators or quantum states—What probability distributions are possible? How can we represent quantum states? How can suitable collections of states enable cryptography? Frame theory, then, is perfectly suited to this purpose, as it too asks questions of sets of vectors in a very general setting. Chapter two provided the necessary background for our purposes, detailing among other things, how to transform an arbitrary collection of vectors into a measurement by the canonical

dual, and the relationship between various elegant sets of vectors to minima of simple functions of the vectors, the frame potentials.

With the background firmly in hand, part one turns to foundational questions of importance to physics. Chapter three starts by laying out a coherent way to interpret quantum mechanics before moving to investigating how the quantum probability rule follows from the structure of measurements. The most sensible way to think of quantum mechanics may in the end be to posit that measurements simply happen and work from there. Otherwise we invite the measurement problem and a host of other paradoxes based on the conceptual collision between probability and measurement in a physical theory without underlying properties, like quantum mechanics. This point of view shares much with the Copenhagen interpretation[1], but is primarily an operational approach, eschewing concepts which make no difference in the laboratory.

Taking measurements as the fundamental building blocks of the theory only works if the rest follows without too much further input, which is indeed the case. By utilizing quantum measurements in their most general form, POVMs, the probability rule follows immediately. From there dynamical rules are only a short step away, and the foundations of the theory are set. Generally, all possible POVMs should be considered, but for qubits even just the trine measurement suffices to imply the Born rule. Such a successful application of generalized measurements should serve to cement their foundational status. Additionally, chapter six demonstrates that actually implementing POVMs is no more or less daunting than projective measurements, which should overcome any lingering doubts stemming from adherence to the axioms as originally formulated by von Neumann.

Having determined that probabilities in quantum mechanics come encoded in the form of a density operator, chapter four sets out to build up a representation of finite-dimensional quantum theory purely in terms of measurement probabilities. The ideal measurement candidate on which to base the construction is the most

[1]Summed up neatly by Lawrence Bragg as "Everything in the future is a wave, everything in the past is a particle."

symmetric measurement available, the SICPOVM. In the terms of frame theory, the set of vectors associated with the SICPOVM is simultaneously an equiangular spherical code, a spherical 2-design, and a discrete version of a Weyl-Heisenberg frame. This rich structure accounts for its simple properties and usefulness to this problem, but unfortunately does not lead so easily to an existence proof. Instead, the first half of chapter four details the relevant structure before presenting numerical results showing existence in dimensions up to 45.

Satisfied, at least provisionally, that the SICPOVM does always exist, the second half of the chapter develops an elegant formulation of quantum mechanics—density operators, measurements, and dynamics—explicitly in terms of it. Quantum states are replaced with the analog of the Husimi Q function, i.e. the probability distribution when measuring the SICPOVM. Measurements are represented by the analog of the Glauber-Sudarshan P function, while quantum operations become stochastic maps on the Q function. This formalism provides a means to work with quantum mechanics in terms familiar from classical probability theory. However, the mathematical difficulties inherent to the theory carry over to the new representation. In particular, the condition demarcating the density operators from all operators is quite complicated and difficult to deal with. In the new representation things are no different: just as not all operators are density operators, not all probability distributions are Q functions. The new representation does not offer any immediate help in clearly describing the boundary.

Study of the SICPOVM bridges the gap between the foundational topics of part two and the practical applications of part three, because not only does it offer appealing representations for quantum mechanics, but it is also useful for quantum cryptography. Chapter five provides a concise background of cryptography and cryptographic methods before explaining what quantum mechanics has to offer. Cryptography itself is charged the nominal task of encoding messages so that they cannot be read by unintended parties. This involves using another string, the key, to transform the message into something unintelligible. Using the key again recovers the message,

so in practice the real difficulty lay in ensuring each intended party has a copy of the key. This is not an easy task when the parties are separated by a large distance. Now the peculiar (even offensive) features of quantum mechanics come to our aid to provide a means of key distribution whose security is grounded in physical law, not on the apparent computational difficulty of certain tasks as it is currently.

If quantum mechanics offers security in key distribution protocols, spherical codes offer more of it, as the second half of chapter five examined in detail. The workhorse of these schemes is the phenomenon that measuring a quantum state to determine its identity invariably perturbs it. This allows Alice and Bob to establish a putative key in almost any way they like using quantum signals, and then simply check if it has been compromised. If it has, they simply throw the key away and begin again when circumstances change. To investigate how much eavesdropper interference can be sustained before Alice and Bob must abandon their efforts, analysis of spherical code protocols was performed in the context of the intercept/resend attack, and direct comparison with more traditionally studied schemes was made. This simple attack lays out the general features of the spherical code protocols, which are important to understand before mounting a full-scale attempt to prove unconditional security. At the outset, it's not apparent that spherical code protocols will be at all useful, so going straight for the ultimate answer of unconditional security against all attacks is not immediately warranted. For the intercept/resend method of eavesdropping, the ability of a set of quantum states (which can be assembled into a POVM for use by Bob) to record tampering was found to be related to the minimum of the $t = 2$ frame potential. This explains the numerical results showing that the available ESC protocols in a given dimension, and the SICPOVM in particular, offer more security, more resistance to eavesdropping. These protocols trade speed of key creation to achieve higher security, as was evident when comparing protocols having a fixed number of elements, again in a given dimension. Besides these two direct comparisons, two further observations should be made of the versatility of ESC protocols. First, spherical codes offer a greater variety of protocols, so that although they trade speed for se-

curity in any one instance, there may be faster ESC protocols providing just enough security for the particular implementation in question. Should the apparent noise level in the communication channel used by Alice and Bob be low, there's really no reason to use a protocol which resists eavesdropping associated with very much larger error rates; in a particular use the goal is of course to create the required keys as fast as possible. Second, the ESC protocols are capable of automatically estimating the error rate of the channel, saving a step for Alice and Bob and making the possibility of autocompensating protocols easier to realize. If, unlike in the previous point, noise levels in the channel fluctuate over time, Alice and Bob might instead desire a protocol which automatically compensates for this, particularly one which does not involve changes to the physical hardware used.

To take a step toward establishing their unconditional security, the end of chapter five examines the concrete case of the two qubit-based protocols, the trine and tetrahedron. By a slight modification to the general ESC protocols, these two were demonstrated to be just as fast as their two cousin protocols, BB84 and six-state, but offer the same advantage of automatic noise estimation as before. Stronger eavesdropping attacks could be considered in this case, and the results were the same, if less dramatic.

The use of spherical codes represents a conceptual step away from the traditionally studied protocols which provide Alice and Bob the opportunity to deterministically create key letters in each step. When Bob measures in the basis Alice used to prepare the state, he is certain to obtain the identical value if the channel is noiseless. In contrast, even using a noiseless channel, the raw strings of signals and outcomes generated by ESC protocols contain discrepancies which are probabilistically removed using classical error-correction. The exceptions are the trine and tetrahedron protocols, for which suitable measurements "dual" to Alice's signal ensemble can be be used by Bob. Placing the reliability of the protocol in the hands of the data processing instead of the signal set reflects the same spirit Shannon introduced into the problem of reliable communication: establishing *some*, but not total,

correlation is typically sufficient if not optimal for classical communication. The trick in applying this idea to cryptography is to find a signal set which offers something in return for the decrease in key rate. Spherical codes, due to their appealing structure, were in retrospect perhaps an obvious candidate.

Finally, chapter six tackles the question of how the higher-dimensional protocols can be implemented, be they ESC or MUB based. Since Alice and Bob must communicate over in principle large distances, methods of encoding quantum states into electromagnetic modes are investigated. To lay the groundwork, polarization-based encoding is first examined, as this method is straightforward and already in widespread use. Each photon may be thought of as a qubit using this scheme, and arbitrary quantum states are simple to prepare with polarizers and waveplates. Arbitrary measurements are also easy to realize by first transcribing the polarization state into various spatial channels with a polarizing beamsplitter and waveplate. From there, ordinary beamsplitters and phase shifters suffice along with photodetection suffice to realize any measurement Bob wishes to perform. These three building blocks, state preparation, transformation into spatial modes, and subsequent processing and detection, are then applied to the question of encoding higher-dimensional systems into transverse spatial field modes. For state preparation Alice may resort to holography, even though it is of low efficiency. Processing and detection by beamsplitters, phase shifters, and photodetectors again applies, once the superposition states Alice has created can be broken down into components traveling in different spatial channels. The tricky part, then, is the analog of the polarizing beamsplitter/waveplate combination. Due to the phase profiles of the transverse mode states under consideration, mode sorting is possible using interferometric techniques. An analog of the polarizing beamsplitter can be realized which sorts the incoming beam into it various constituents, suitable for further processing and measurement. A transverse-mode waveplate can be realized using holograms, though this suffers from low efficiency.

In its last section, chapter six examines the practical issues surrounding quantum

cryptography. Foremost is the interdiction against using high-intensity pulses, lest Alice simply give away many copies of the state she intends for Bob. The methods developed for higher-dimensional encoding fit into this scheme, as Bob's measurement device consists of high-efficiency elements, at least in principle. On the state preparation side, Alice need not worry about using high-efficiency elements since she is required to strongly attenuate the beam anyway. Physical means may be able to boost the signal intensity, such as reliably sending one photon per pulse, but here again spherical codes offer a "software" advantage. Considering the qubit case, the nonorthogonality of all the signal states makes it impossible for Eve to unambiguously determine the key bit even when in possession of several copies.

7.2 Topics for Future Work

Having concluded our tour through only a tiny section of quantum information, the question naturally turns to future applications of this work. Good science is that which not only provides answers but asks new questions, and this work will hopefully prove no different. Topics deserving further inquiry abound; some are rather obvious extensions of results reported here, and some are applications of the tools developed herein to other problems.

Clearly one would like to place the existence of SICPOVMs on firm ground. An existence proof would likely reveal something mathematically interesting, for it has not been established by straightforward means, as detailed in chapter four. Because the SICPOVM is apparently strongly related to the displacement operators, and these operators find wide application in quantum information theory, such a proof may also have direct relevance to the field. Finally, in the face of considerable numerical evidence, it is mathematically unjust not to establish analytically what is clearly true.

Studying the SICPOVM is also worthwhile because it is useful, and not just for

quantum cryptography. Along with the question of what measurements are suitable for state tomography comes the problem of finding an optimal one, a measurement which determines the state with as efficiently as possible. It is known that for qubits, the uniform POVM extracts the most information about the state's identity per measurement, so it is the fastest possible means to determine an unknown quantum state [86, 87]. But in addition to being wildly impractical, it is also incredibly wasteful. For each outcome we would need to store the associated spherical coordinate to high precision. The entropy of the generated data will be gigantic in comparison with the minuscule amount of information extracted. Very well, only speed was demanded. If instead we wish to determine the state without requiring a massive data storage capacity, the SICPOVM seems to be the logical choice. Having the fewest number of outcomes possible, we simply need to keep track of how often each outcome occurs, vastly reducing our data storage requirements. But is this indeed the case? This question is likely not too difficult to answer. Much more difficult, but more rewarding would be to look for an "information-maximization" principle akin to that mentioned in the introduction for linear polarization. Using the SICPOVM in a particular dimension to again determine an unknown quantum state, can it be established that the quantum probability rule maximizes the information gain in this process?

Naturally, quantum key distribution provides ample opportunities for future work. Foremost is proving the unconditional security of the spherical code protocols, starting with the qubits as a concrete case. Now that the intercept/resend analysis has shown the usefulness of ESC protocols, the newly-developed methods of establishing privacy amplification should be of help. If such a direct appeal is not forthcoming with results, a more limited step can be taken by examining the various cloning protocols which have been studied in the context of unbiased basis protocols. These were determined to be optimal attacks when only eavesdropping on signals one at a time. Such a result is of great practical utility, because demanding unconditional security is an example of over-engineering for two reasons. First,

in practice quantum key distribution only forms a part of a larger cryptosystem, which itself is not unconditionally secure. Second, assuming that Eve can perform arbitrary interactions on arbitrary numbers of signal states is hopelessly unrealistic. Ultimately it is of theoretical interest that quantum key distribution systems exhibit security against all possible eavesdropping attacks consistent with physical law, but these are not all likely to be the attacks in practice. Restricting attention to the cases in which Eve only interrogates signals one at a time is far more realistic, because nothing like reliable quantum memory currently exists. Coupled with the first point it's unrealistic to assume that an eavesdropper with a fully-functional quantum computer would choose to attack the cryptosystem at its *strongest* point. Returning to the theoretical perspective, such a gradual building-up of eavesdropping methods is how the analysis of BB84 proceeded historically, so it's not a bad idea to travel this proven path.

Within the context of the intercept/resend attack, the security optimality of spherical code protocols could be examined. The argument given in chapter five relating to the frame potential is only heuristic as it does not establish the maximum tolerable noise rate, but rather shows that eavesdropping on the spherical codes is easily detected. Numerical investigation should suffice to establish quickly how the ESC protocols stand in relation to random protocols. Additionally, with the tools of frame theory, the restrictions on investigating signal sets which also form POVMs can be relaxed. Arbitrary vectors can be collected for the signal set, and Bob assigned the associated canonical tight frame for measurement. Finding the optimal ensemble, be it an equiangular spherical code or not, is certainly of interest, since although quantum mechanics offers the possibility of secure key distribution, it isn't limitless.

Related to the optimal ensemble problem is determining how these prepare and measure schemes stand in relation to protocols making full use of quantum coherence and entanglement. Normally one considers key distribution protocols which have as little to do with quantum mechanics as possible. This makes them easy

to implement, but one wonders how much is sacrificed. In principle, quantum key distribution is quite straightforward—distribute halves of entangled states, concentrate the entanglement into a fewer number of systems if it was corrupted, and then simply measure each part to establish the key. Since entanglement cannot be shared, truly secret keys follow from measurements of entangled particles. Concentrating the entanglement, a process known as entanglement distillation, is the difficult part. In general this is a massively multi-particle process, and is therefore extremely difficult to implement. Prepare and measure schemes, in contrast, are as simple to implement as possible, focusing on one system at a time and doing the key distillation classically. But how much is really lost in the conversion of this process into something easier to implement? We've restricted attention to the lower bound of the key rate bounds, the one achieved using one-way communication. With two-way communication between parties longer keys can be cut from the same raw strings, but so far a large gap exists between this rate and that which could be achieved using entanglement distillation [70]. Will this gap necessarily persist in the face of improved key distribution protocols, or can everything that can be achieved using fully quantum protocols also be done with simpler, stripped-down protocols? Though the answer to this question is of practical value in determining how much effort should be put into realizing more coherent protocols, it also bears on the fundamental issues of how we conceive of the information processing tasks. If the prepare and measure schemes are sufficient, then quantum information processing is conceptually similar to classical information processing, albeit in a context with more than a few peculiar features. On the other hand, if such schemes are insufficient to realize the full secrecy capacity of quantum mechanics, then something inescapably different is taking place. Though performing similar tasks, the quantum information processing cannot be accounted for by adding the quantum features to classical information processing tasks. This question mirrors the debate on the computer science side of quantum information theory about what gives a quantum computer its apparent power over classical computation, another practical question with foundational implications.

Returning to the remarks at the end of chapter six, determining the ability of spherical code protocols and others like them to resist number splitting attacks deserves attention. In this instance a purely practical consideration motivates this problem, but as with most research topics in quantum information theory, again a foundational connection can be made. Now the issue is *resistance* to copying, not just the ability to bear witness to it. The qubit spherical code protocols accomplish this by combining their inherent indistinguishability with a structure that allows Alice and Bob to create key bits deterministically with some probability. Since Eve has less than unit probability of correctly copying the state and guessing the key letter, Alice and Bob can simply give her a copy and proceed anyway. This phenomenon is subtler than the use of the information/disturbance tradeoff to provide security against single signal interference, since now it is the combination of the ensemble and classical communication protocol which makes the scheme possible. Practically, resistance to number splitting would be of enormous utility in extending the range over which quantum key distribution can be performed. Current commercial applications quote a maximum distance of 100km in ideal circumstances, a distance which could conceivably be doubled by suitable encoding.

Finally, note that key distribution protocols make up only one aspect of quantum cryptography. The information/disturbance tradeoff can also be used to improve actual encryption schemes by making the strings hard for Eve to even read, let alone crack [69]. The prohibition against recycling one-time pads stems from the assumption that Eve has access to the ciphertext once it is transmitted, but if she cannot even read it, Alice and Bob can confidently reuse the key for the next message. By combining a classical one-time pad encryption system with an "eavesdropper-detecting" quantum code, Alice and Bob can establish how much of the message Eve was able to read, and thus how much of the key may safely be recycled. Spherical codes may again find application in such schemes.

References

[1] A. Acín, L. Masanes, and N. Gisin, "Equivalence Between Two-Qubit Entanglement and Secure Key Distribution," *Physical Review Letters*, vol. 91(16), no. 167901, 2003.

[2] A. Acín, N. Gisin, and V. Scarani, "Coherent Pulse Implementations of Quantum Cryptography Protocols Resistant to Photon Number Splitting Attacks," *Physical Review A*, vol. 69(1), no. 012309, 2004.

[3] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, "Orbital Angular Momentum of Light and the Transformation of Laguerre-Gaussian Laser Modes," *Physical Review A*, vol. 45(11), pp. 8185–9, 1992.

[4] D. M. Appleby, "Existential Contextuality and the Models of Meyer, Kent, and Clifton," *Physical Review A*, vol. 65(2), no. 022105, 2002.

[5] J. Arlt, K. Dholakia, L. Allen, and M. J. Padgett, "The Production of Multiringed Laguerre-Gaussian Modes by Computer-Generated Holograms," *Journal of Modern Optics*, vol. 45(6), pp. 1231–7, 1998.

[6] I. A. al-Kadi, "The Origins of Cryptology: The Arab Contributions", *Cryptologia*, vol. 16(2), pp. 97–126, 1992.

[7] H. Bacry, A. Grossmann, and J. Zak, "Proof of the Completeness of Lattice States in kq Representation," *Physical Review B*, vol. 12(4), pp. 1118–20, 1975.

[8] R. Balian, "Un principe d'incertitude fort en théorie du signal en mécanique quantique (A Strong Uncertainty Principle in Signal Theory or in Quantum Mechanics)," *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique*, vol. 292(20), pp. 1357–62, 1981.

[9] V. Y. Bazhenov, M. V. Vasnetsov, and M. S. Soskin, "Laser-Beams with Screw Dislocations in Their Wave-Fronts," *JETP Letters*, vol. 52(8), pp. 429–31, 1990.

[10] M. W. Beijersbergen, L. Allen, H. E. L. O. van der Veen, and J. P. Woerdman, "Astigmatic Laser Mode Converters and Transfer of Orbital Angular-Momentum," *Optics Communications*, vol. 96(1), pp. 123–32, 1993.

[11] J. J. Benedetto and M. Fickus, "Finite Normalized Tight Frames," *Advances in Computational Mathematics*, vol. 18(2), pp. 357–85, 2003.

[12] C. H. Bennett and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," *Proceedings of the IEEE International Conference on Computers Systems and Signal Processing, Bangalore, India*, pp. 175–9, 1984.

[13] C. H. Bennett, G. Brassard, and N. D. Mermin, "Quantum cryptography without Bell's theorem," *Physical Review Letters*, vol. 68(5), pp. 557–9, 1992.

[14] C. H. Bennett, G. Brassard, C. Crépaeu, R. Jozsa, A. Peres, and W. Wootters, "Teleporting an Unknown Quantum State Via Dual Classical and EPR Channels," *Physical Review Letters*, vol. 70(13), pp. 1895–9, 1993.

[15] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized Privacy Amplification," *IEEE Transactions on Information Theory*, vol. 41(6), pp. 1915–23, 1995.

[16] G. Birkhoff and J. von Neumann, "The Logic of Quantum Mechanics", *Annals of Mathematics*, vol. 37(4), pp. 823–43, 1936.

[17] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, "Quantum Interferometric Optical Lithography: Exploiting Entanglement to Beat the Diffraction Limit," *Physical Review Letters*, vol. 85 (13), pp. 2733–6, 2000.

[18] M. Bourennane, A. Karlsson, G. Björk, N. Gisin, and N. Cerf, "Quantum Key Distribution Using Multilevel Encoding: Security Analysis," *Journal of Physics A*, vol. 35(47), pp. 10065–76, 2002.

[19] G. Brassard, N. Lütkenhaus, T. Mor, and B. Sanders, "Limitations on Practical Quantum Cryptography," *Physical Review Letters*, vol. 85(6), pp. 1330–3, 2000.

[20] D. Bruß, "Optimal Eavesdropping in Quantum Cryptography with Six States," *Physical Review Letters*, vol. 81(14), pp. 3018–21, 1998.

[21] D. Bruß, M. Christandl, A. Ekert, B.-G. Englert, D. Kaszlikowski, and C. Macchiavello, "Tomographic Quantum Cryptography: Equivalence of Quantum and Classical Key Distribution," *Physical Review Letters*, vol 91(9), no. 097901, 2003.

[22] P. Busch, "Informationally Complete-Sets of Physical Quantities", *International Journal of Theoretical Physics*, vol. 30(9), pp. 1217–27, 1991.

[23] P. Busch, M. Grabowski, and P. Lahti, *Operational Quantum Physics* Berlin: Springer, 1995, 2nd corrected printing, 1997.

[24] P. Busch, "Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem," *Physical Review Letters*, vol. 91(12), no. 120403, 2003. First appeared as `quant-ph/9909073`.

[25] W. T. Buttler, R. J. Hughes, S. K. Lamoreaux, G. L. Morgan, J. E. Nordholt, and C. G. Peterson, "Daylight Quantum Key Distribution Over 1.6 km," *Physical Review Letters*, vol. 84(24), pp. 5652–55, 2000.

[26] J. Byers and J. Cramer, "A Daring Intruder Airs the Beefs of Dish Owners," *Time*, 12 May 1986.

[27] A. Cabello, "Finite-Precision Measurement Does Not Nullify the Kochen-Specker Theorem," *Physical Review A*, vol. 65(5), no. 052101, 2002.

[28] A. Cabello, "Kochen-Specker Theorem for a Single Qubit using Positive Operator-Valued Measures," *Physical Review Letters*, vol. 90(19), no. 190401, 2003.

[29] P. G. Casazza, "The Art of Frame Theory," *Taiwanese Journal of Mathematics*, vol. 4(2), pp. 129–201, 2000.

[30] P. Casazza, D. Han, and D. R. Larson, "Frames for Banach Spaces," preprint. `http://www.math.missouri.edu/~pete/pdf/57.FramesBSpaces.pdf`

[31] C. M. Caves, "Quantum-Mechanical Noise in an Interferometer," *Physical Review D*, vol. 23(8), pp. 1693–708, 1981.

[32] C. M. Caves, C. A. Fuchs, and R. Schack, "Unknown Quantum States: The Quantum de Finetti Representation," *Journal of Mathematical Physics*, vol. 43(9), pp. 4537–59, 2002.

[33] C. M. Caves, C. A. Fuchs, K. K. Manne, and J. M. Renes, "Gleason-Type Derivations of the Quantum Probability Rule for Generalized Measurements," *Foundations of Physics*, vol. 34(2), pp. 193–209, 2004.

[34] N. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, "Security of Quantum Key Distribution Using d-Level Systems," *Physical Review Letters* vol. 88(12), no. 127902, 2002.

[35] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, "Exponential algorithmic speedup by quantum walk," *Proceedings of the 35th ACM Symposium on Theory of Computing (STOC)*, pp. 59–68, 2003.

[36] A. M. Childs, J. Preskill, and J. Renes, "Quantum Information and Precision Measurement," *Journal of Modern Optics*, vol. 47(2), pp. 155–76, 2000.

[37] M. Christandl, A. Ekert, and R. Renner, "A Generic Security Proof for Quantum Key Distribution," `quant-ph/0402131`.

[38] I. L. Chuang, "Quantum Algorithm for Distributed Clock Synchronization", *Physical Review Letters*, vol. 85(9), pp. 2006–9, 2000.

[39] R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, E. Riis, and M. Sasaki, "Experimental Realization of Optimal Detection Strategies for Overcomplete States," *Physical Review A*, vol. 64(1), no. 012303, 2001.

[40] R. Clifton and A. Kent, "Simulating Quantum Mechanics by Non-Contextual Hidden Variables," *Proceedings of the Royal Society of London A*, vol. 456(2001), pp. 2101–14, 2000.

[41] J. H. Conway and N. J. A. Sloane, editors, *Sphere Packings, Lattices, and Groups*. Springer, Berlin, 1993.

[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, New York, 1991.

[43] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages", *IEEE Transactions on Information Theory*, vol. 24(3), pp. 339–48, 1978.

[44] M. Curty, M. Lewenstein, and N. Lütkenhaus, `quant-ph/0307151`.

[45] G. M. D'Ariano and P. Lo Presti, "Classical and quantum noise in measurements and transformations," `quant-ph/0301110`.

[46] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, "Informationally Complete Measurements and Groups Representation," `quant-ph/0310013`.

[47] I. Daubechies, A. Grossmann, and Y. Meyer, "Painless Nonorthogonal Expansions", *Journal of Mathematical Physics*, vol. 27(5), pp. 1271–1283, 1986.

[48] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Transactions on Information Theory*, vol. 36(5), pp. 961–1005, 1990.

[49] E. B. Davies, "Information and quantum measurement," *IEEE Transactions on Information Theory*, vol. 24(5), pp. 596–9, 1978.

[50] 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.

[51] D. Deutsch, "Quantum-Theory, the Church-Turing Principle and the Universal Quantum Computer," *Proceedings of the Royal Society of London, Series A*, vol. 400(1818), pp. 97–117, 1985.

[52] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, "Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels," *Physical Review Letters*, vol. 77(13), pp. 2818–21, 1996.

[53] I. Devetak and A. Winter, "Distillation of Secret Key and Entanglement From Quantum States," `quant-ph/0306078`.

[54] W. Diffie and M. Hellman, "New Directions in Cryptography", *IEEE Transactions on Information Theory*, vol. 22(6), pp. 644–54, 1976.

[55] R. J. Duffin and A. C. Schaeffer, "A Class of Nonharmonic Fourier Series," *Transactions of the American Mathematical Society*, vol. 72, pp. 341–366, 1952.

[56] A. K. Ekert, "Quantum Cryptography Based on Bell's Theorem," *Physical Review Letters*, vol. 67(6), pp. 661–3, 1991.

[57] A. Ekert and B. Huttner, "Information Gain in Quantum Eavesdropping," *Journal of Modern Optics*, vol. 41(12), pp. 2455–66, 1994.

[58] Y. C. Eldar and G. D. Forney, "On Quantum Detection and the Square-Root Measurement," *IEEE Transactions on Information Theory*, vol. 47(3), pp. 858–72, 2001.

[59] Y. C. Eldar and G. D. Forney, "Optimal Tight Frames and Quantum Measurement," *IEEE Transactions on Information Theory*, vol. 48(3), pp. 599–610, 2002.

[60] B. Et-Taoui, "Equiangular Lines in C^r," *Indagationes Mathematicae (New Series)*, vol. 11(2), pp. 201–207, 2000.

[61] B. Et-Taoui, "Equiangular Lines in C^r (Part II)," *Indagationes Mathematicae (New Series)*, vol. 13(4), pp. 483–486, 2002.

[62] R. P. Feynman, "Simulating Physics with Computers," *International Journal of Theoretical Physics*, vol. 21(6), pp. 467–88, 1982.

[63] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, "Optimal Eavesdropping in Quantum Cryptography. I. Information Bound and Optimal Strategy," *Physical Review A*, vol. 56(2), pp. 1163–72, 1997.

[64] C. A. Fuchs, "Quantum Mechanics as Quantum Information (and only a little more)," `quant-ph/0205039`, and private communication.

[65] C. A. Fuchs and M. Sasaki, "Squeezing Quantum Information through a Classical Channel: Measuring the 'Quantumness' of a Set of Quantum States," *Quantum Information and Computation*, vol. 3(5), pp. 377–404, 2003.

[66] D. Gabor, "Theory of Communications," *Journal of the Institute of Electrical Engineering*, vol. 93(21), pp. 429–457, 1946.

[67] N. Gisin and S. Wolf, "Quantum Cryptography on Noisy Channels: Quantum versus Classical Key-Agreement Protocols", *Physical Review Letters*, vol. 83(20), pp. 4200–3, 1999.

[68] A. M. Gleason, "Measures on the Closed Subpaces of a Hilbert Space", *Journal of Mathematics and Mechanics*, vol. 6, pp. 885-89, 1957.

[69] D. Gottesman, "Uncloneable Encryption," *Quantum Information & Computation*, vol. 3(6), pp.581–602, 2003.

[70] D. Gottesman and H.-K. Lo, "Proof of Security of Quantum Key Distribution with Two-Way Classical Communications," *IEEE Transactions on Information Theory*, vol. 49(2), pp. 457–75, 2003.

[71] R. J. Greechie and D. J. Foulis, "Transition to Effect Algebras," *International Journal of Theoretical Physics*, vol. 34(8), pp. 1369–82, 1995.

[72] L. K. Grover, "Quantum Mechanics Helps in Searching For a Needle in a Haystack," *Physical Review Letters*, vol. 79(2), pp. 325–8, 1997.

[73] S. Gulde, M. Riebe, G. P. T. Lancaster, C. Becher, J. Eschner, H. Häffner, F. Schmidt-Kaler, I. L. Chuang, and R. Blatt, "Implementing the Deutsch-Jozsa Algorithm on an Ion-Trap Quantum Computer," *Nature*, vol. 421(6918), pp. 48–50, 2003.

[74] S. Hallgren, "Polynomial-Time Quantum Algorithms for Pell's Equation and the Principal Ideal Problem," *Proceedings of the 34th ACM Symposium on Theory of Computing (STOC)*, pp. 653–8, 2002.

[75] D. Han and D. R. Larson, "Frames, Bases, and Group Representations," *Memoirs of the American Mathematical Society*, vol. 147, 2000.

[76] H. He, M. E. J. Friese, N. R. Heckenberg, and H. Rubinsztein-Dunlop, "Direct Observation of Transfer of Angular Momentum to Absorptive Particles from a Laser Beam with a Phase Singularity," *Physical Review Letters*, vol. 75(5), pp. 826–9, 1995.

[77] S. G. Hoggar, "64 Lines from a Quaternionic Polytope," *Geometriae Dedicata*, vol. 69(3), pp. 287–289, 1998.

[78] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problemy Peredachi Informatsii*, vol. 9(3), pp. 3–11, 1973. [A. S. Kholevo, *Problems of Information Transmission*, vol. 9(3), pp. 177–83 (1973)].

[79] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[80] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, "Improvement of Frequency Standards with Quantum Entanglement," *Physical Review Letters*, vol. 79(20), pp. 3865–8, 1997.

[81] R. J. Hughes, W. T. Buttler, P. G. Kwiat, S. K. Lamoreaux, G. L. Morgan, J. E. Nordholt, and C. G. Peterson, "Free-Space Quantum Key Distribution in Daylight," *Journal of Modern Optics*, vol. 47(2), pp. 549–62, 2000.

[82] R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson, "Practical Free-Space Quantum Key Distribution Over 10 km in Daylight and at Night," *New Journal of Physics*, vol. 4, no. 43, 2002.

[83] Internet Engineering Task Force (IETF) secure shell working group draft standard, `http://www.ietf.org/html.charters/secsh-charter.html`

[84] J. M. Jauch, *Foundations of Quantum Mechanics*. Addison Wesley, Reading, MA, 1968.

[85] P. S. Jessen, D. L. Haycock, G. Klose, G. A. Smith, I. H. Deutsch, and G. K. Brennen, "Quantum Control and Information Processing in Optical Lattices," *Quantum Information and Computation*, vol. 1, pp. 20–32, 2001.

[86] K. R. W. Jones, "Principles of Quantum Inference," *Annals of Physics*, vol. 207(1), pp. 140–170, 1991.

[87] K. R. W. Jones, "Fundamental Limits Upon the Measurement of State Vectors," *Physical Review A*, vol. 50(5), pp. 3682–99, 1994.

[88] R. Jozsa, D. S. Abrams, J. P. Dowling, and C. P. Williams, "Quantum Clock Synchronization Based on Shared Prior Entanglement," *Physical Review Letters*, vol. 85(9), pp. 2010–3, 2000.

[89] D. Kahn, *The Codebreakers: The Story of Secret Writing*. MacMillan Publishing, New York, 1967.

[90] A. Kerckhoffs, "La Cryptographie Militaire (Military Cryptography)," *Journal des Sciences Militaires*, vol. 9, pp. 5–83; 161–91, 1883.

[91] A. Klappenecker and M. Rötteler. "Beyond Stabilizer Codes I: Nice Error Bases," *IEEE Transactions on Information Theory*, vol. 48(8), pp. 2392–5, 2002.

[92] A. Klappenecker and M. Rötteler. " Unitary error bases: Constructions, equivalence, and applications," *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Proceedings*, vol. 2643, pp. 139–149, 2003.

[93] A. Klappenecker and M. Rötteler, "Catalogue of Nice Error Bases," `http://faculty.cs.tamu.edu/klappi/ueb/ueb.html`.

[94] E. Knill, "Non-binary Unitary Error Bases and Quantum Codes," LANL report LAUR-96-2717; `quant-ph/9608048`.

[95] E. Knill, "Group Representations, Error Bases and Quantum Codes," LANL report LAUR-96-2807; `quant-ph/9608049`.

[96] S. Kochen and E. P. Specker, *Journal of Mathematics and Mechanics*, vol. 17, pp. 59, 1967.

[97] A. Koldobsky and H. König, "Aspects of the Isometric Theory of Banach Spaces," in *Handbook of the Geometry of Banach Spaces*, Vol. 1, edited by W. B. Johnson and J. Lindenstrauss, pp. 899–939. North Holland, Dordrecht, 2001.

[98] H. König and N. Tomczak-Jaegermann, "Norms of Minimal Projections," `math.fa/9211211`.

[99] H. König, "Cubature Formulas on Spheres," 2003. Available online at `http://analysis.math.uni-kiel.de/koenig/preprints.html`.

[100] R. König, U. Maurer, and R. Renner, "On the Power of Quantum Memory," `quant-ph/0305154`.

[101] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*. Springer-Verlag, Berlin, 1983.

[102] J. Leach, M. J. Padgett, S. M. Barnett, S. Franke-Arnold, and J. Courtial, "Measuring the Orbital Angular Momentum of a Single Photon," *Physical Review Letters*, vol. 88(25), no. 257901, 2002.

[103] J. Leach, J. Courtial, K. Skeldon, S. M. Barnett, S. Franke-Arnold, and M. J. Padgett, "Interferometric Methods to Measure Orbital and Spin, or the Total Angular Momentum of a Single Photon," *Physical Review Letters*, vol. 92(1), no. 013601, 2004.

[104] P. W. H. Lemmens and J. J. Seidel, "Equiangular Lines," *Journal of Algebra*, vol. 24, pp. 494–512, 1973.

[105] H.-K. Lo, H. F. Chau, and M. Ardehali, "Quantum Key Distribution Scheme and Proof of its Unconditional Security", `quant/ph-0011056`.

[106] A. W. Lohmann, D. Mendlovic, and Z. Zalevsky, "Fractional Transformations in Optics", in *Progess in Optics*, Vol. 38, E. Wolf, editor, pp. 263. Pergamon, London, 1998.

[107] F. Low, "Complete Sets of Wave Packets", in: *A Passion for Physics—Essays in Honor of Geoffrey Chew*, pp. 17–22. World Scientific, Singapore, 1985.

[108] G. Ludwig, *Foundations of Quantum Mechanics* Springer-Verlag, Berlin, 1983.

[109] N. Lütkenhaus, "Security Against Eavesdropping in Quantum Cryptography," *Physical Review A*, vol. 54(1), pp. 97–111, 1996.

[110] D. J. C. Mackay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, 2003.

[111] G. Mackey, "Quantum Mechanics and Hilbert Space", *American Mathematical Monthly*, vol. 64, pp. 45–57, 1957.

[112] G. Mackey, *The Mathematical Foundations of Quantum Mechanics*. Benjamin, New York, 1963.

[113] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of the Orbital Angular Momentum States of Photons," *Nature*, vol. 412(6844), pp. 313–6, 2001.

[114] U. Maurer, "Secret Key Agreement by Public Discussion From Common Information," *IEEE Transactions on Information Theory*, vol. 39(3), pp. 733–42, 1993.

[115] U. Maurer and S. Wolf, "Unconditionally Secure Key Agreement and the Intrinsic Conditional Information," *IEEE Transactions on Information Theory*, vol. 45(2), pp. 499–514, 1999.

[116] D. A. Meyer, "Finite Precision Measurement Nullifies the Kochen-Specker Theorem," *Physical Review Letters*, vol. 83(19), pp. 3751–4, 1999.

[117] D. A. Miller, "Spatial Channels for Communicating with Waves Between Volumes," *Optics Letters*, vol. 23(21), pp. 1645–7, 1998.

[118] G. Molina-Terriza, J. P. Torres, and L. Torner, "Management of the Angular Momentum of Light: Preparation of Photons in Multidimensional Vector States of Angular Momentum," *Physical Review Letters*, vol. 88(1), no. 13601, 2002.

[119] W. Munro, private communication, 2004.

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

[121] J. E. Nordholt, R. J. Hughes, G. L. Morgan *et al.*, Present and future free-space quantum key distribution, in *Free-Space Laser Communication Technologies XIV*, Proceedings of SPIE Vol. 4635, pp. 116–26, SPIE, Bellingham, WA, 2002.

[122] A. Peres, *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, Dordrecht, 1993.

[123] S. J. D. Phoenix, S. M. Barnett, and A. Chefles, "Three-State Quantum Cryptography," *Journal of Modern Optics*, vol. 47(2), pp. 507–16, 2000.

[124] C. Piron, *Foundations of Quantum Physics*. W. A. Benjamin, Reading, MA, 1976.

[125] E. Prugovečki, "Information-Theoretical Aspects of Quantum Measurement", *International Journal of Theoretical Physics*, vol. 16, pp. 321–31, 1977.

[126] R. Raussendorf, D. E. Browne, and H.-J. Briegel, "The One-Way Quantum Computer—A Non-Network Model of Quantum Computation," *Journal of Modern Optics*, vol. 49(8), pp. 1299–1306, 2002.

[127] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental Realization of Any Discrete Unitary Operator," *Physical Review Letters*, vol. 73(1), pp. 58–61, 1994.

[128] X.-F. Ren, G.-P. Guo, B. Yu, J. Li, G.-C. Guo, "Orbital Angular Momentum of the Down Converted Photons," `quant-ph/0309044`

[129] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Symmetric Informationally Complete Quantum Measurements," *Journal of Mathematical Physics*, vol. 45(6), pp. 1–10, 2004.

[130] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Numerical SICPOVM Solutions," `http://info.phys.unm.edu/papers/reports/sicpovm.html`

[131] R. L. Rivest, A. Shamir, and L. M. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," *Communications of the ACM*, vol. 21(2), pp. 120–6, 1978.

[132] H. Sasada and M. Okamoto, "Transverse-Mode Beam Splitter of a Light Beam and its Application to Quantum Cryptography," *Physical Review A*, vol. 68(1), no. 012323, 2003.

[133] M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, and O. Hirota, "Accessible Information and Optimal Strategies for Real Symmetrical Quantum Sources," *Physical Review A*, vol. 59(5), pp. 3325–35, 1999.

[134] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, "Quantum Cryptography Protocols Robust Against Photon Number Splitting Attacks for Weak Laser Pulses Implementations," *Physical Review Letters*, vol. 92(5), no. 057901, 2004.

[135] B. Schneier, *Applied Cryptography Second Edition: Protocols, Algorithms, and Source Code in C*. John Wiley & Sons, New York, 1996.

[136] E. Schrödinger, "Die Gegenwärtige Situation in der Quantenmechanik", *Naturwissenschaften*, vol. 23, pp. 807, 1935.
Translated and reprinted in *Quantum Theory and Measurement* Princeton University Press, Princeton, 1983.

[137] F. E. Schroeck, *Quantum Mechanics on Phase Space*. Kluwer Academic Press, Dordrecht, 1996.

[138] B. Schumacher, "Multiparticle Quantum Entanglement," lecture in the Complexity, Entropy, and the Physics of Information (CEPI) seminar series, held at the Santa Fe Institute, March 2000.

[139] Secure hypertext transfer protocol RFC standard,
`ftp://ftp.rfc-editor.org/in-notes/rfc2660.txt`

[140] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948. Reprinted in book form, with postscript by Warren Weaver: C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL, 1949.

[141] C. E. Shannon, "Communication Theory of Secrecy Systems," *Bell System Technical Journal*, vol. 28, pp. 656, 1949.

[142] P. W. Shor, "Algorithms for Quantum Computation: Discrete Logarithms and Factoring," *Proceedings of the 35th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 124–34, 1994.

[143] P. W. Shor, "Scheme For Reducing Decoherence in Quantum Computer Memory." *Physical Review A*, vol. 52(4), pp. 2493–6, 1995.

[144] P. Shor and J. Preskill, "Simple Proof of Security of the BB84 Quantum Key Distribution Protocol," *Physical Review Letters*, vol. 85(2), pp. 441–4, 2000.

[145] A. E. Siegman, *Lasers*. University Science Books, Mill Valley, CA, 1986.

[146] A. M. Steane, "Error Correcting Codes in Quantum Theory," *Physical Review Letters*, vol. 77(5), pp. 793–7, 1996.

[147] T. Strohmer and R. Heath, "Grassmanian Frames with Applications to Coding and Communication," *Applied and Computational Harmonic Analysis*, vol. 14(3), pp. 257–75, 2003.

[148] P. M. L. Tammes, "On the Origin of Number and Arrangements of the Places of Exit on the Surface of Pollen-Grains", *Recl. Trav. Bot. Neerl.*, vol. 27, pp. 1–84, 1930.

[149] W.-K. Tung, *Group Theory in Physics*, World Scientific, Philadelphia, 1993.

[150] A. Uhlmann, "The 'transition probability' in the state space of a *-algebra," *Reports on Mathematical Physics*, vol. 9, pp. 273–9, 1976.

[151] A. Vaziri, G. Weihs, and A. Zeilinger, "Experimental Two-Photon, Three-Dimensional Entanglement for Quantum Communication," *Physical Review Letters*, vol. 89(24), no. 240401, 2002.

[152] A. Vaziri, G. Weihs, and A. Zeilinger, "Superpositions of the Orbital Angular Momentum for Applications in Quantum Experiments," *Journal of Optics B*, vol. 4(2), pp. S47–S51, 2002.

[153] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, Berlin, 1932. Translated by E. T. Beyer, *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, 1955.

[154] G. S. J. Vernam and J. Mauborgne, U. S. Patent 1310719. Available online at `http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=1310719`

[155] S. Waldron, "Generalised Welch Bound Equality Sequences Are Tight Frames," *IEEE Transactions on Informtion Theory*, vol. 49(9), pp. 2307–9, 2003.

[156] S. Waldron, private communication, 2004.

[157] D. Walls and G. J. Milburn, *Quantum Optics*. Springer-Verlag, Berlin, 1995.

[158] H. Wei, X. Xue, J. Leach, M. J. Padgett, S. M. Barnett, S. Franke-Arnold, E. Yao, and J. Courtial, "Simplified Measurement of the Orbital Angular Momentum of Single Photons," *Optics Communications*, vol. 223(1), pp. 117–22, 2003.

[159] H. Wei and X. Xue, "Comment on 'Measuring the Orbital Angular Momentum of a Single Photon,' " `quant-ph/0208146`, 2002.

[160] L. Welch, "Lower Bounds on the Maximum Cross-Correlation of Signals," *IEEE Transactions on Information Theory*, vol. 20(3), pp. 397–9, 1974.

[161] R. F. Werner, "All Teleportation and Dense Coding Schemes," *Journal of Physics A*, vol. 34(35), pp. 7081–94, 2001.

[162] J. A. Wheeler, "The Computer and the Universe," *International Journal of Theoretical Physics*, vol. 21(6), pp. 557–72, 1982.

[163] S. Wiesner, "Conjugate Coding," *Sigact News*, vol. 15(1), pp. 78–8, 1983. Originally written c. 1970 but unpublished.

[164] D. J. Wineland, M. Barrett, J. Britton, J. Chiaverini, B. L. DeMarco, W. M. Itano, B. M. Jelenkovic, C. Langer, D. Leibfried, V. Meyer, T. Rosenband, and T. Schaetz, "Quantum Information Processing with Trapped Ions,"

Philosophical Transactions of the Royal Society of London A, vol. 361(1808), pp. 1349–61, 2003.

[165] W. K. Wootters, "Information is Maximixed in Photon Polarization Measurements," in *Quantum Theory and Gravitation*, pp. 13–36. Academic Press, Boston, 1980.

[166] W. K. Wootters, *The Acquisition of Information from Quantum Measurements*. PhD thesis, University of Texas at Austin, 1980.

[167] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299(5886), pp. 802–3, 1982.

[168] X. Xue, H. Wei, and A. Kirk, "Beam Analysis by Fractional Fourier Transform," *Optics Letters*, vol. 26(22), pp. 1746–8, 2001.

[169] G. Zauner, *Quantendesigns - Grundzüge einer nichtkommutativen Designtheorie (Quantum Designs — Foundations of a Non-Commutative Theory of Designs)*. PhD thesis, University of Vienna, 1999. Available online at `http://www.mat.univie.ac.at/~neum/papers.html`