Chapter 11

Accessible information about quantum states: An open optimization problem

Jun Suzuki, Syed M. Assad, and Berthold-Georg Englert

Abstract We give a brief summary of the current status of the problem of extracting the accessible information when a quantum system is received in one of a finite number of pre-known quantum states. We review analytical methods as well as a numerical strategy. In particular, the group-covariant positive-operator-valued measures are discussed, and several explicit examples are worked out in detail. These examples include some that occur in the security analysis of schemes for quantum cryptography.

11.1 Introduction

A sender, traditionally called Alice, sends quantum states, one by one, to a receiver, Bob. Bob then wishes to perform measurements on the quantum states he receives to find out, the best he can, what Alice has sent. Generally speaking, owing to the nature of quantum mechanics, it is impossible for Bob to obtain full knowledge about the states which he is receiving. Instead, he has to choose his measurements judiciously from all measurements permitted by

quantum mechanics. A natural question one might ask is then:

What is the best strategy for Bob to maximize his knowledge about the states he is receiving from Alice? (11.1)

The answer to this question is not only of importance for our understanding of the implications of quantum mechanics, it also has great practical significance for most areas in quantum information, in particular for the capacity of quantum channels and the security analysis of schemes for quantum cryptography under powerful eavesdropping attacks. Indeed, our own interest in the matter originates in its relevance to the security of "tomographic quantum cryptography," a class of protocols for quantum key distribution developed in Singapore [1, 2, 3, 4].

The main objective of this chapter is to provide a concise introduction to this problem with a summary of ongoing research in this field. For this purpose we will not give a rigorous mathematical exposition, and we will be content with stating most of theorems without proof. We suggest that readers who are interested in the technical mathematical details consult the pertinent literature referred to in the text.

Here is a brief preview of coming attractions. In Section 11.2 we remind the reader of a few basic concepts and, at the same time, establish the terminology and the notational conventions we are using. Then, in Section 11.3, we state question (11.1) as an optimization problem, for which the mutual information between Alice and Bob is the figure of merit. Section 11.4 reports essential properties of this mutual information and important theorems about known properties of the solution. A numerical procedure for searching the optimum by a steepest-ascent method is described in Section 11.5. Examples are presented in Section 11.6, where we limit the choice to cases with a structure as one meets it in the security analysis of quantum cryptography schemes. We close with a summary and outlook.

11.2 Preliminaries

11.2.1 States and measurements

We set the stage by first providing a brief mathematical description of the physical situation that (11.1) refers to, that is: Alice sends certain physical states to Bob who measures them to find out which states she sent. For simplicity and for concreteness, we consider only finite-dimensional systems.

The quantum states prepared by Alice are denoted by $\rho_1, \rho_2, \ldots, \rho_J$ whereby $J \geq 1$ is finite, and the set $\mathscr{E} = \{\rho_j | j = 1, 2, \ldots, J\}$ is the *ensemble* of quantum states sent by Alice. Each of the ρ_js is a *density matrix*: a semi-definite positive, and therefore hermitian, matrix with finite trace.[1] One calls the jth state a *pure state* when the density matrix ρ_j is essentially a projector, otherwise it is a *mixed state*,

$$\text{state } \rho_j \text{ is } \left\{ \begin{array}{c} \text{pure} \\ \\ \text{mixed} \end{array} \right\} \text{ if } \left\{ \begin{array}{l} \mathrm{Tr}(\rho_j^2) = (\mathrm{Tr}\,\rho_j)^2 , \\ \\ \mathrm{Tr}(\rho_j^2) < (\mathrm{Tr}\,\rho_j)^2 . \end{array} \right. \tag{11.2}$$

By convention we normalize the ρ_js such that their traces are the probabilities a_j with which Alice is sending them. Thus, Bob knows that the probability of receiving ρ_j as the next state is $a_j = \mathrm{Tr}\,\rho_j$. Since the next state is surely one of the ρ_js, these probabilities have unit sum,

$$1 = \sum_{j=1}^{J} a_j = \sum_{j=1}^{J} \mathrm{Tr}\,\rho_j . \tag{11.3}$$

It follows that the total density matrix $\rho = \sum_{j=1}^{J} \rho_j$ has unit trace, $\mathrm{Tr}\,\rho = 1$. The rank of ρ is the dimension d of the space under consideration, which is to say that we represent all ρ_js, and all other linear operators, by $d \times d$ matrices.[2]

It is often convenient to represent a pure-state matrix ρ_j as a product of a d-component column $|j\rangle$ and its adjoint d-component row $\langle j| = |j\rangle^\dagger$, that is $\rho_j = |j\rangle\langle j|$. In the standard terminology of quantum physics, one speaks of *kets* and *bras* when referring to the columns $|j\rangle$ and the rows $\langle j|$, respectively. The numerical row-times-column product of $\langle j_1|$ with ket $|j_2\rangle$ is denoted by $\langle j_1|j_2\rangle$ and is called their *bracket*; it is equal to the trace of their column-times-row product $|j_2\rangle\langle j_1|$,

$$\mathrm{Tr}\big(|j_2\rangle\langle j_1|\big) = \langle j_1|j_2\rangle . \tag{11.4}$$

Bob's measurement is specified by a decomposition of the $d \times d$ identity matrix $\mathbf{1}_d$ into a set of semi-definite positive, hermitian matrices,

$$\mathbf{1}_d = \sum_{k=1}^{K} \Pi_k \quad \text{with} \quad K \geq 1 \quad \text{and} \quad \Pi_k \geq 0 , \tag{11.5}$$

[1] More generally, a quantum state is specified by a semi-definite positive linear operator with finite trace and each of its equivalent matrix representations is a corresponding density matrix. By choosing one particular orthonormal basis in the Hilbert space, we specify one set of density matrices for the set of states under consideration.

[2] More generally, d is the dimension of the relevant subspace of a possibly much larger Hilbert space.

which is the general[3] form of a so-called *positive operator valued measure* (POVM) [5], here with K *outcomes* Π_k. Bob's *a priori* probability of getting the kth outcome is

$$b_k = \mathrm{Tr}(\rho\Pi_k), \tag{11.6}$$

which is properly normalized to unit sum as a consequence of the unit trace of ρ. Two special cases are worth mentioning: the von Neumann measurements, and the tomographically complete measurements.

We have a *von Neumann measurement* when the outcomes of the POVM are pairwise orthogonal projectors, $\Pi_k\Pi_l = \Pi_k\delta_{kl}$. When all Π_ks are rank-1 projectors, one speaks of a *maximal* von Neumann measurement, for which $K = d$, of course.

The POVM is *tomographically complete* if ρ can be inferred from the knowledge of all of Bob's probabilities b_k, which is to say that the map $\rho \mapsto \{b_k | k = 1, \ldots, K\}$ is injective. A tomographically complete POVM, has at least d^2 outcomes; in the case of $K = d^2$, one speaks of a *minimal* tomographically complete POVM.

Every outcome of a POVM can be written as a square, $\Pi_k = A_k^\dagger A_k$, but this factorization is not unique.[4] Typically, there is one such factorization for each physical implementation of the POVM. Then, given an ideal—that is, noise-free and nondestructive—implementation, the final state of the physical system after the measurement is

$$\rho^{(k)} = \frac{A_k \rho A_k^\dagger}{\mathrm{Tr}(\rho\Pi_k)} \tag{11.7}$$

if ρ is the state before the measurement and the kth outcome is obtained. Therefore, in general, the possible final states are mixed states when POVMs are performed on mixed states.

When Bob performs the POVM (11.5) on the states $\rho_1, \rho_2, \ldots, \rho_J$ sent by Alice, the joint probability that Alice sends the jth state and Bob gets the kth outcome is

$$p_{jk} = \mathrm{Tr}(\rho_j \Pi_k). \tag{11.8}$$

The respective marginal probabilities

$$a_j = \sum_{k=1}^{K} p_{jk} = \mathrm{Tr}\,\rho_j, \quad b_k = \sum_{j=1}^{J} p_{jk} = \mathrm{Tr}(\rho\Pi_k) \tag{11.9}$$

[3] Somewhat more generally, the label k could be continuous and the summation replaced by an integration. We do not need to consider such general cases.

[4] More generally, Π_k could be a sum of squares, $\Pi_k = \sum_l A_{kl}^\dagger A_{kl}$, even in the case of a von Neumann measurement, as is illustrated by $A_{kl} = V_{kl}\Pi_k^{1/2}$ with $\sum_l V_{kl}^\dagger V_{kl} = \mathbf{1}_d$ for all k. The case of $A_k = \Pi_k^{1/2}$ is sometimes referred to as an *ideal* POVM.

are the probabilities that Alice sends the jth state and the *a priori* probabilities that Bob gets the kth outcome.

The *conditional probabilities* $p(k|j) = p_{jk}/a_j$ and $p(j|k) = p_{jk}/b_k$ have the following significance, respectively: *If* Alice sends the jth state, she can predict that Bob will get the kth outcome with probability $p(k|j)$; *if* Bob receives the kth outcome, he can infer that Alice sent the jth state with probability $p(j|k)$.

It is worth noting that there is a reciprocal situation with exactly the same joint probabilities. It is specified by Alice measuring the POVM

$$\mathbf{1}_d = \sum_{j=1}^{J} \widetilde{\Pi}_j \quad \text{with} \quad \widetilde{\Pi}_j = \rho^{-1/2} \rho_j \rho^{-1/2} \tag{11.10}$$

and Bob sending her the states $\widetilde{\rho}_k = \rho^{1/2} \Pi_k \rho^{1/2}$.

11.2.2 Entropy and information

Next, we define several quantities that will be used for the quantification of information in the sequel [6, 7]: the von Neumann entropy, the Shannon entropy, the Kullback–Leibler relative entropy, the mutual information, and the accessible information.

von Neumann entropy: The von Neumann entropy $S(\rho)$ of a density matrix ρ is[5]

$$S(\rho) = -\operatorname{Tr}\left(\frac{\rho}{\operatorname{Tr}\rho} \log \frac{\rho}{\operatorname{Tr}\rho}\right) = -\frac{\operatorname{Tr}(\rho \log \rho)}{\operatorname{Tr}\rho} + \log \operatorname{Tr}\rho\,, \tag{11.11}$$

which has the more familiar appearance

$$S(\rho) = -\operatorname{Tr}(\rho \log \rho) \quad \text{if} \quad \operatorname{Tr}\rho = 1\,. \tag{11.12}$$

By construction, we have $S(x\rho) = S(\rho)$ for all $x > 0$. Further we note that the mapping $\rho \mapsto \operatorname{Tr}(\rho) S(\rho)$ is concave:

$$\operatorname{Tr}(\rho_1 + \rho_2) S(\rho_1 + \rho_2) \geq \operatorname{Tr}(\rho_1) S(\rho_1) + \operatorname{Tr}(\rho_2) S(\rho_2) \tag{11.13}$$

for any two density matrices ρ_1 and ρ_2.

[5] Historically, the von Neumann entropy involves the natural logarithm and also the Boltzmann constant to establish contact with the thermodynamical entropy, whereas the Shannon entropy uses the logarithm to base 2 and the value is usually stated in units of bits. We use the logarithm to base 2 throughout.

With the convention $\lambda \log \lambda = 0$ for $\lambda = 0$, the von Neumann entropy (11.12) is expressed in terms of the eigenvalues λ_i ($i = 1, 2, \ldots, d$) of ρ as

$$S(\rho) = -\sum_{i=1}^{d} \lambda_i \log \lambda_i \quad \text{if} \quad \sum_{i=1}^{d} \lambda_i = 1\,. \tag{11.14}$$

We remark that the von Neumann entropy is zero for pure states and only for pure states, for which a single eigenvalue is positive and all others are zero.

Shannon entropy: Given Alice's ensemble $\mathscr{E} = \{\rho_j | j = 1, 2, \ldots, J\}$, we have the set $P = \{a_j = \mathrm{Tr}\rho_j | j = 1, 2, \ldots, J\}$ that is composed of the probabilities of occurrence, which have unit sum, $\sum_j a_j = 1$. The Shannon entropy $H(P)$ of such a normalized set of probabilities P is defined by[5]

$$H(P) = -\sum_{j=1}^{J} a_j \log a_j\,. \tag{11.15}$$

For any two sets of normalized probabilities $P^{(1)} = \{a_j^{(1)} | j = 1, 2, \ldots, J\}$ and $P^{(2)} = \{a_j^{(2)} | j = 1, 2, \ldots, J\}$, we can consider their convex sums $xP^{(1)} + (1 - x)P^{(2)} = \{xa_j^{(1)} + (1 - x)a_j^{(2)} | j = 1, 2, \ldots, J\}$ with $0 \le x \le 1$, for which the concavity

$$H(xP^{(1)} + (1-x)P^{(2)}) \ge xH(P^{(1)}) + (1-x)H(P^{(2)}) \tag{11.16}$$

holds.

As a consequence of the concavity of the von Neumann entropy in (11.13), we have the inequalities (see, e.g., Subsection 11.3.6 in [7])

$$H(P) + \sum_{j=1}^{J} a_j S(\rho_j) \ge S(\rho) \ge \sum_{j=1}^{J} a_j S(\rho_j)\,, \tag{11.17}$$

where $\rho = \sum_{j=1}^{J} \rho_j$ is the total density matrix. On the left, the equal sign applies if and only if all ρ_js are pairwise orthogonal pure states. On the right, the equal sign applies if the ρ_js are essentially equal to each other in the sense that $a_j \rho_k = \rho_j a_k$ for all j and k.

Kullback–Leibler relative entropy: For any two sets of normalized probabilities $P = \{p_j | j = 1, 2, \ldots, J\}$ and $\widetilde{P} = \{\widetilde{p}_j | j = 1, 2, \ldots, J\}$, the Kullback–Leibler relative entropy $D(P||\widetilde{P})$ is defined by

$$D(P||\widetilde{P}) = \sum_{j=1}^{J} p_j \log \frac{p_j}{\widetilde{p}_j} \ge 0\,, \tag{11.18}$$

whereby the equal sign applies only if $p_j = \widetilde{p}_j$ for all j. The Kullback–Leibler relative entropy may serve as a rough measure of difference between two probability distributions P and $\widetilde{P}$. But, since it is not symmetric, $D(P||\widetilde{P}) \neq D(\widetilde{P}||P)$ as a rule, and does not satisfy the triangle inequality, it is not a distance or metric in the mathematical sense.

Mutual information: For any normalized set of joint probabilities $A\&B = \{p_{jk}|j = 1,2,\ldots,J; k = 1,2,\ldots,K\}$ with $\sum_{jk} p_{jk} = 1$, and its two sets of marginals $A = \{a_j = \sum_{k=1}^{K} p_{jk}|j = 1,2,\ldots,J\}$ and $B = \{b_k = \sum_{j=1}^{J} p_{jk}|k = 1,2,\ldots, K\}$, the mutual information $I(A;B)$ is the relative entropy between the joint probabilities $A\&B$ and the set $AB = \{a_j b_k|j = 1,2,\ldots,J; k = 1,2,\ldots,K\}$ of product probabilities,

$$I(A;B) = D(A\&B||AB) = \sum_{j=1}^{J}\sum_{k=1}^{K} p_{jk} \log \frac{p_{jk}}{a_j b_k}$$
$$= H(A) + H(B) - H(A\&B)\,, \qquad (11.19)$$

where the last version expresses the mutual information in terms of the various Shannon entropies.

The mutual information is a measure of the strength of the statistical correlations in joint probabilities. If there are no correlations at all, that is, if $p_{jk} = a_j b_k$ for all j and all k, the mutual information vanishes; otherwise it is positive.

In the physical situation to which the question (11.1) refers, we have the joint probabilities of (11.8) and the marginals of (11.9). Therefore, the mutual information $I(\mathscr{E};\Pi)$ between $\mathscr{E}$, the ensemble of Alice's states, and Π, Bob's POVM, quantifies his knowledge about the quantum states she is sending. This brings us, finally, to the accessible information for Bob about Alice's quantum states.

Accessible information: The accessible information $I_{\rm acc}$ is the maximum of the mutual informations for all possible POVMs that Bob can perform, that is

$$I_{\rm acc}(\mathscr{E}) = \max_{\text{all } \Pi} I(\mathscr{E};\Pi)\,. \qquad (11.20)$$

This poses the challenge of determining the value of $I_{\rm acc}(\mathscr{E})$ for the given set $\mathscr{E}$ of quantum states.

In addition to the accessible information, there are other numerical measures [8] that can be used for the quantification of Bob's knowledge about Alice's states, such as the Bayes cost (see, e.g., [9, 5]), which is essentially the probability for guessing wrong, or the probability that Bob can unambiguously identify the state he just received (see, e.g., chapter 11 in [10]). In the

context of studying the security of quantum cryptography schemes, however, the figure of merit is the accessible information. Also, the history of the subject seems to indicate that it is substantially more difficult to determine the accessible information than the Bayes cost or the probability of unambiguous discrimination.

11.3 The optimization problem

We now state the main problem (11.1) in technical terms as a double question:

Given an ensemble of quantum states $\mathscr{E} = \{\rho_j | j = 1, 2, \ldots, J\}$,
(a) what is the value of the accessible information $I_{\rm acc}(\mathscr{E})$, and
(b) what is the optimal POVM $\Pi = \{\Pi_k | k = 1, 2, \ldots, K\}$ for which the mutual information is the accessible information, $I(\mathscr{E}; \Pi) = I_{\rm acc}(\mathscr{E})$? (11.21)

Part of the answer to query (b) is to establish the number K of outcomes in the optimal POVM.

This problem was first formulated by Holevo in 1973 [9]. After more than three decades, it remains unsolved. The major difficulty is a lack of sufficient conditions that ensure the optimality of POVMs in general. Sufficiency is known only when the ensemble of quantum states possesses certain symmetry properties; see Subsection 11.4.4 below. The obvious nonlinearity that originates in the logarithms is another hurdle.

The current situation is still rather unsatisfactory even for seemingly simple ensembles $\mathscr{E}$. For instance, we do not have analytical expressions for the optimal POVMs in the case where $\mathscr{E}$ consists of only two full-rank mixed quantum states for $d = 2$;[6] see Subsection 11.6.1 below for details.

There are, of course, very special cases for which the answer is immediate. One extreme situation is

(i) all states commute with each other, $\rho_j \rho_{j'} = \rho_{j'} \rho_j$; then the optimal POVM is a von Neumann measurement composed of the projectors to the joint eigenstates. A special case thereof is

[6] Two mixed single-qubit states in the jargon of quantum information.

(ii) all states are pairwise orthogonal, $\rho_j \rho_{j'} = \delta_{jj'} \rho_j^2$, so that they can be distinguished without effort and we have essentially the situation of Bob receiving a classical signal.

A related, yet different problem is the determination of the so-called *quantum channel capacity* [7, 11]. A quantum channel turns any input quantum states into an output quantum state, always preserving the positivity and usually also the trace of the input. The ensemble $\mathscr{E}$ received by Bob, for which he has to find the optimal POVM, then comes about by processing Alice's input ensemble $\mathscr{E}_{\text{in}}$ through the quantum channel. There is then a two-fold optimization problem: one needs to find both Alice's optimal input ensemble as well as Bob's optimal POVM. The quantum channel capacity problem is also an *open problem*. It is clear that any progress with the accessible-information problem (11.21) means corresponding progress with the channel-capacity problem.

11.4 Theorems

Before going to the actual computation of the accessible information, we give a brief summary of established properties of the mutual information and the accessible information [6, 7].

11.4.1 Concavity and convexity

Let us regard the joint probabilities $p_{jk} = a_j p(k|j)$ as the product of Alice's probabilities a_j and the conditional probabilities $p(k|j)$. Then, the mutual information $I(\mathscr{E};\Pi)$ is a concave function of the a_js for given $p(k|j)$s, and a convex function of the $p(k|j)$s for given a_js. In other words, the mutual information is a convex functional on the set of all possible POVMs. Therefore, all optimal POVMs are located on the boundary of the POVM space.

Since this convexity of the mutual information is of some importance in our discussion, we give more details. Suppose we have two POVMs $\Pi^{(i)} = \{\Pi_k^{(i)} | k = 1,2,\ldots,K\} (i = 1,2)$, then the combined new POVM $\Pi(\lambda) = \{\lambda \Pi_k^{(1)} + (1-\lambda)\Pi_k^{(2)} | k = 1,2,\ldots,K\}$ with $0 < \lambda < 1$ obeys the following inequality for the mutual information:

$$I(\mathscr{E};\Pi(\lambda)) \leq \lambda I(\mathscr{E};\Pi^{(1)}) + (1-\lambda) I(\mathscr{E};\Pi^{(2)}) . \tag{11.22}$$

The equality is satisfied if and only if

$$p^{(1)}_{jk}/b^{(1)}_k = p^{(2)}_{jk}/b^{(2)}_k \quad \text{or} \quad p^{(1)}(j|k) = p^{(2)}(j|k) \tag{11.23}$$

holds for all j and k, wherein we meet the joint probabilities, $p^{(i)}_{jk} = \mathrm{Tr}(\rho_j \Pi^{(i)}_k)$ and the marginals $b^{(i)}_k = \sum_{j=1}^{J} p^{(i)}_{jk}$, as well as the resulting conditional probabilities $p^{(i)}(j|k)$.

A particular situation in which the equal sign applies in (11.22) is as follows. Let $\Pi^{(1)}$ and $\Pi^{(2)}$ be two K-outcome POVMs with null outcomes such that $\Pi^{(1)}_k = 0$ for $\bar{k} < k \leq K$ and $\Pi^{(2)}_k = 0$ for $1 \leq k \leq \bar{k}$ with $1 \leq \bar{k} < K$. Then the outcomes of $\Pi(\lambda)$ are given by $\Pi_k(\lambda) = \lambda \Pi^{(1)}_k$ for $1 \leq k \leq \bar{k}$ and $\Pi_k(\lambda) = (1-\lambda)\Pi^{(2)}_k$ for $\bar{k} < k \leq K$, and it is clear that

$$I(\mathscr{E};\Pi(\lambda)) = \lambda I(\mathscr{E};\Pi^{(1)}) + (1-\lambda) I(\mathscr{E};\Pi^{(2)}) \tag{11.24}$$

holds in this situation.

11.4.2 Necessary condition

For a POVM Π to be optimal, it is necessary that the accessible information $I(\mathscr{E};\Pi)$ is stationary with respect to infinitesimal variations of Π. These variations are, however, constrained by both the positive nature of each outcome Π_k and the unit sum of all outcomes.

The first constraint is accounted for by writing $\Pi_k = A_k^\dagger A_k$, whereby the factor A_k is rather arbitrary and may differ from the physical A_k in (11.7) by a unitary matrix multiplying A_k on the left. The second constraint, that is $\sum_{k=1}^{K} \delta\Pi_k = 0$, then requires the infinitesimal variations of the A_ks to be of the form

$$\delta A_k = \mathrm{i} \sum_{k'=1}^{K} \varepsilon_{kk'} A_{k'} \quad \text{with} \quad {\varepsilon_{kk'}}^\dagger = \varepsilon_{k'k}\,, \tag{11.25}$$

where the $\varepsilon_{kk'}$s are otherwise arbitrary infinitesimal matrices.

We note that the mutual information is expressed as

$$I(\mathscr{E};\Pi) = \sum_{k=1}^{K} \mathrm{Tr}(R_k \Pi_k) \tag{11.26}$$

with the hermitian matrices R_k given by

$$R_k = \sum_{j=1}^{J} \rho_j \log \frac{p_{jk}}{a_j b_k}\,. \tag{11.27}$$

It turns out that there is no contribution from the variation of the R_ks to

$$\delta I(\mathscr{E};\Pi) = -\mathrm{i}\sum_{k,k'=1}^{K} \mathrm{Tr}\big(\varepsilon_{kk'}A_{k'}(R_{k'}-R_k)A_k^\dagger\big)\,. \tag{11.28}$$

Therefore, a necessary condition for Π to be an optimal POVM is

$$A_{k'}(R_{k'}-R_k)A_k^\dagger = 0 \quad \text{for all } \ k,k'\,, \tag{11.29}$$

or

$$\Pi_{k'}(R_{k'}-R_k)\Pi_k = 0 \quad \text{for all } \ k,k'\,. \tag{11.30}$$

Upon summing over k or k' we arrive at an equivalent set of equations,

$$R_k\Pi_k = \Lambda\Pi_k \quad \text{and} \quad \Pi_k\Lambda = \Pi_k R_k \quad \text{for all } \ k\,, \tag{11.31}$$

which are adjoint statements of each other because

$$\Lambda = \sum_{k=1}^{K} R_k\Pi_k = \sum_{k=1}^{K} \Pi_k R_k \tag{11.32}$$

is hermitian. Mathematically speaking, Λ is the Lagrange multiplier of the unit-sum constraint in (11.5), and its significance is revealed by noting that $I_{\mathrm{acc}}(\mathscr{E}) = \mathrm{Tr}\,\Lambda$ for an optimal POVM.

Equations (11.30)–(11.32) have been investigated by Holevo [9]. These equations are nonlinear and there does not seem to be any efficient method for finding their solutions. Indeed, the $\frac{1}{2}K(K-1)$ equations (11.30) are not solved directly in the numerical approach described in Section 11.5. Rather, we exploit the observation that (11.28) identifies the gradient in the POVM space.

We remark that a POVM obeying (11.30) is not guaranteed to be an optimal POVM. Strictly speaking, $I(\mathscr{E};\Pi)$ is only ensured to be extremal, but it could be a local maximum rather than a global maximum, or a local minimum, or even a saddle point. Whereas local minima and saddle points tend to be unstable extrema for the numerical procedure of Section 11.5, local maxima are just as stable as global maxima.

11.4.3 Some basic theorems

We state four basic theorems about $I(\mathscr{E};\Pi)$ and $I_{\mathrm{acc}}(\mathscr{E})$ without proof. The reader is invited to consult the respective references for proofs and further details.

Theorem 11.1: Number of outcomes

The accessible information is always achievable by an optimal POVM whose outcomes are rank-1 operators, so that $\Pi_k^2 = \Pi_k \operatorname{Tr}(\Pi_k)$ for $1 \le k \le K$. The number of outcomes needed in such an optimal POVM is bounded by the rank d of the total density matrix ρ, which is also the dimension of the relevant Hilbert space,[7] in accordance with [12]

$$d \le K \le d^2 . \tag{11.33}$$

When all quantum states ρ_j can be represented as matrices with real numbers, then the upper bound is reduced to $K \le d(d+1)/2$ [13].

(Davies [12]; Sasaki *et al.* [13])

For the following theorems we introduce two quantities that are defined by

$$\chi(\mathscr{E}) = S(\rho) - \sum_{j=1}^{J} a_j S(\rho_j) \ge 0 \tag{11.34}$$

and

$$\chi(\mathscr{E};\Pi) = \sum_{k=1}^{K} \Big(b_k S\big(\rho^{(k)}\big) - \sum_{j=1}^{J} p_{jk} S\big(\rho_j^{(k)}\big) \Big) \ge 0 , \tag{11.35}$$

where $\rho^{(k)}$ is the final total state conditioned on Bob's kth outcome, as in (11.7), and $\rho_j^{(k)}$ is the corresponding conditional final state when ρ_j is the initial state. That is

$$\rho^{(k)} = A_k \rho A_k^\dagger \quad \text{and} \quad \rho_j^{(k)} = A_k \rho_j A_k^\dagger , \tag{11.36}$$

where the normalizing denominators of (11.7)—respectively $\operatorname{Tr}\big(\rho^{(k)}\big) = b_k$ and $\operatorname{Tr}\big(\rho_j^{(k)}\big) = p_{jk}$—are irrelevant here because these conditional density matrices appear only as arguments of the von Neumann entropy function of (11.11).

Theorem 11.2: Upper bound on $I(\mathscr{E};\Pi)$

The mutual information is bounded by the difference of $\chi(\mathscr{E})$ and $\chi(\mathscr{E};\Pi)$,

$$I(\mathscr{E};\Pi) \le \chi(\mathscr{E}) - \chi(\mathscr{E};\Pi) . \tag{11.37}$$

(Schumacher, Westmoreland, and Wootters [14])

[7] If ρ is embedded in a larger Hilbert space, there is one more outcome in the POVM, namely, the projector on the orthogonal complement of the range of ρ.

Since the term $\chi(\mathscr{E};\Pi)$ that is subtracted on the right-hand side of (11.37) is nonnegative and vanishes if and only if all outcomes Π_k are of rank 1, we have $I(\mathscr{E};\Pi) \leq \chi(\mathscr{E})$ for all POVMs, in particular for all optimal POVMs. This implies the following theorem.

Theorem 11.3: Upper bound on the accessible information
An upper bound on the accessible information is given by

$$I_{\rm acc}(\mathscr{E}) \leq \chi(\mathscr{E}), \tag{11.38}$$

the so-called *Holevo bound*. **(Holevo [15])**

We remark that the equal sign holds in (11.38) if and only if all quantum states ρ_j commute with each other, and hence the Holevo bound is *not tight* in general.

Theorem 11.4: Lower bound on the accessible information
A lower bound of the accessible information is given by

$$I_{\rm acc}(\mathscr{E}) \geq \mathscr{Q}(\rho) - \sum_{j=1}^{J} a_j \mathscr{Q}(\rho_j/a_j), \tag{11.39}$$

wherein the so-called *subentropy* $\mathscr{Q}(\rho)$ of a unit-trace density matrix ρ with eigenvalues λ_i $(i = 1,2,\ldots,d)$ is defined by

$$\mathscr{Q}(\rho) = -\sum_{i=1}^{d} \left(\prod_{i'(\neq i)} \frac{\lambda_i}{\lambda_i - \lambda_{i'}} \right) \lambda_i \log \lambda_i . \tag{11.40}$$

If there are degenerate eigenvalues, one treats them as the limit of nondegenerate ones. **(Jozsa, Robb, and Wootters [16])**

We should also mention that one can establish substantially tighter upper and lower bounds for the accessible information by taking more specific properties of the ρ_js into account than the rather global entropies and subentropies that enter the right-hand sides of (11.38) and (11.39); see, in particular, the work of Fuchs and Caves [17, 8].

11.4.4 Group-covariant case

Following Holevo [9], an ensemble $\mathscr{E} = \{\rho_j | j = 1,2,\ldots,J\}$ of quantum states ρ_j is said to be covariant with respect to a group G if there exists a faithful projective unitary representation $\{U_g | g \in G\}$ of G such that

$$U_g \rho_j U_g^\dagger \in \mathscr{E} \quad \text{for all } \rho_j \in \mathscr{E} \text{ and all } g \in G. \tag{11.41}$$

A projective unitary representation of a group G means that for any pair g_1, g_2 of group elements $U_{g_1}U_{g_2} = U_{g_1 g_2}\,\mathrm{e}^{\mathrm{i}\phi(g_1,g_2)}$ holds with a real phase function $\phi(g_1,g_2)$. Several remarks are in order.

1. If an ensemble $\mathscr{E}$ is covariant with respect to a group G, then $\mathscr{E}$ is also covariant with respect to any subgroup of G.
2. When a group G acts transitively on an ensemble $\mathscr{E}$, then $\mathscr{E}$ constitutes a single orbit of G. In this case the order of the group $|G|$ is equal to the number of elements of the ensemble, i.e., $|G| = J$, and the group parameterizes the input states ρ_j. Furthermore, Alice's probabilities of occurrence are all equal, i.e., $a_j = 1/J$.
3. It is always possible to construct a nonprojective unitary representation of the group by a central extension of the original group. In other words, a projective unitary representation is not essential in our discussion.

In this chapter we will only consider nonprojective unitary representations.

In general, a group has a direct sum of irreducible unitary representations of the form

$$U_g = \bigoplus_{\ell=1}^{L} \mathbf{1}_{m_\ell} \otimes u_g^\ell\,, \tag{11.42}$$

where m_ℓ is the multiplicity of inequivalent unitary irreducible representation of u_g^ℓ in d_ℓ dimensions, and L is the number of inequivalent irreducible representations. By construction one has $\sum_{\ell=1}^{L} m_\ell d_\ell = d$. The following theorem [18] is crucial for the discussion below.

Theorem 11.5: Optimal POVM for group-covariant ensemble

Let the ensemble of quantum states $\mathscr{E}$ be covariant with respect to the group G, which has a representation (11.42). Then there exists rank-1 projectors S_m $(m = 1,2,\ldots,M)$, the so-called *seeds*, whose orbits

$$\mathscr{C}_m = \left\{ \frac{d}{|G|} U_g S_m U_g^\dagger \,\middle|\, g \in G \right\} \tag{11.43}$$

constitute an optimal POVM with $K = M|G|$ outcomes. The count M of the seeds is bounded by

$$M \le \sum_{\ell=1}^{L} m_\ell^2\,, \tag{11.44}$$

and the POVM is given by the weighted union of the orbits,

$$\Pi = \bigcup_{m=1}^{M} \lambda_m \mathscr{C}_m = \left\{ \frac{\lambda_m d}{|G|} U_g S_m U_g^\dagger \,\middle|\, 1 \le m \le M\,,\ g \in G \right\}, \tag{11.45}$$

where the values of the nonnegative weights λ_m are determined by the identity decomposition requirement of (11.5). **(Davies [12], Decker [18])**

We remark the following:

1. The labels k of the outcomes Π_k are here identified with the pairs (m,g) with $m = 1,2,\ldots,M$ and $g \in G$.
2. The construction implies $\sum_{m=1}^{M} \lambda_m = 1$, which is the reason for the normalizing factor $d/|G|$ in (11.43).
3. When the group G is irreducible, we have $m_1 = d$ and $L = 1$, and theorem 11.5 reduces to the case studied by Davies and Sasaki *et al.* [12, 13].
4. Although the group-covariant POVM is an optimal POVM, it may not be the only one which maximizes the mutual information. In other words, also for group-covariant ensembles $\mathscr{E}$, the optimal POVM is not unique as a rule; there can be other POVMs that are as good as the optimal group-covariant POVM. This situation occurs typically for $|G| > d$. We will illustrate this point in several examples in Section 11.6.
5. Since $\mathscr{C}_m$ is an orbit, $U_g S_m {U_g}^\dagger$ and S_m are equivalent seeds. Whereas the orbits of the optimal group-covariant POVM may be unique, the seeds are not.
6. When one orbit is enough to attain the accessible information, Schur's lemma provides the following restriction on the structure of the seed:

$$S_m = \bigoplus_{\ell=1}^{L} \frac{d_\ell}{d} \mathbf{1}_{m_\ell} \otimes s^\ell \,, \tag{11.46}$$

where the s^ℓs are rank-1 projectors in the d_ℓ-dimensional subspaces identified by the decomposition (11.42).

7. If the group G acts transitively on the ensemble $\mathscr{E}$, we have $J = |G|$ and $U_g \rho {U_g}^\dagger = \rho$ for all $g \in G$, and the marginals are

$$a_j = \sum_{k=1}^{K} p_{jk} = \frac{1}{|G|} \,, \quad b_k = \sum_{j=1}^{J} p_{jk} = \frac{\lambda_m d}{|G|} \operatorname{Tr}(\rho S_m) \,. \tag{11.47}$$

Bob's *a priori* probabilities b_k, with $k \equiv (m,g)$, are the same for all outcomes within one orbit $\mathscr{C}_m$; their unit sum gives

$$\sum_{m=1}^{M} \lambda_m \operatorname{Tr}(\rho S_m) = \frac{1}{d} \,. \tag{11.48}$$

11.5 Numerical search

Any numerical procedure that is capable of finding maxima of a function could be used in the numerical search for the optimal POVM. In particular, the method of simulated annealing performed well in practice [19]. Such general procedures, however, are unspecific; they do not take full advantage of the structural properties of the mapping $\Pi \to I(\mathscr{E};\Pi)$ and are, therefore, not tailored to the problem at hand.

One algorithm that exploits the structure of $I(\mathscr{E};\Pi)$ is the iterative procedure of Ref. [20]. It implements a steepest-ascent approach to the extremal points in the POVM space, locally proceeding into the direction of the gradient of $I(\mathscr{E};\Pi)$ with respect to Π.

The gradient in steepest ascent is essentially composed of the operators that multiply the infinitesimal increments $\varepsilon_{kk'}$ in (11.28). Accordingly, if we choose the $\varepsilon_{kk'}$s proportional to the respective components of the gradient, the altered POVM will yield a larger value for $I(\mathscr{E};\Pi)$ than the original POVM.

More specifically, we put

$$\varepsilon_{kk'} = \mathrm{i}\alpha\big[A_{k'}(R_{k'} - R_k)A_k^\dagger\big]^\dagger, \tag{11.49}$$

where the value chosen for the "small" parameter α determines the step size. For $\alpha > 0$, the right-hand side of (11.28) is assuredly nonnegative,

$$\begin{aligned}\Delta I(\mathscr{E};\Pi) &= \alpha \sum_{k,k'=1}^{K} \mathrm{Tr}\big(\big[A_{k'}(R_{k'} - R_k)A_k^\dagger\big]^\dagger \big[A_{k'}(R_{k'} - R_k)A_k^\dagger\big]\big) \\ &= \alpha \sum_{k,k'=1}^{K} \mathrm{Tr}\big((R_{k'} - R_k)\Pi_{k'}(R_{k'} - R_k)\Pi_k\big) \geq 0, \end{aligned} \tag{11.50}$$

whereby the equal sign applies only if the POVM obeys the necessary condition (11.30) of an extremal point.

The increment (11.49), which is first-order in α for A_k, gives rise to a term $\propto \alpha^2$ in Π_k, so that we must ensure proper normalization of the improved POVM. This is the purpose of the $T^\dagger \cdots T$ sandwich in

$$\Pi_k \to \Pi_k^{(\mathrm{new})} = T^\dagger\big(\mathbf{1}_d + \alpha G_k^\dagger\big)\Pi_k\big(\mathbf{1}_d + \alpha G_k\big)T \tag{11.51}$$

$$\text{with} \quad G_k = R_k - \sum_{k'=1}^{K} R_{k'}\Pi_{k'} \tag{11.52}$$

$$\text{and} \quad TT^\dagger = \Big(\mathbf{1}_d + \alpha^2 \sum_{k=1}^{K} G_k^\dagger \Pi_k G_k\Big)^{-1}. \tag{11.53}$$

So, given the ensemble $\mathscr{E}$ of Alice's quantum states with its marginals a_j, the numerical procedure of one round of iteration is as follows. For the present nonoptimal POVM Π, we evaluate the joint probabilities p_{jk} of (11.8), the marginals b_k, and the R_ks of (11.27). Then we choose the step size $\alpha > 0$, compute the G_ks of (11.52) as well as T of (11.53), and finally determine the outcomes $\Pi_k^{(\mathrm{new})}$ of the improved POVM in accordance with (11.51). In view of the first-order increase of (11.50), we will have $I(\mathscr{E};\Pi^{(\mathrm{new})}) > I(\mathscr{E};\Pi)$ unless α is too large.

The procedure (11.51)–(11.53) is repeated until no further improvement can be achieved, which happens when the POVM obeys (11.30). Since local minima and saddle points are numerically unstable, the iteration terminates when a local maximum is reached.

Several remarks are in order.

1. If the POVM obeys (11.30), the right-hand side of (11.53) is $\mathbf{1}_d$, and then we have to choose $T = \mathbf{1}_d$ to ensure that the iteration halts. Otherwise, as long as the POVM does not obey (11.30), we have $0 < TT^\dagger < \mathbf{1}_d$ and $T = \left(\mathbf{1}_d + \alpha^2 \sum_{k=1}^K G_k^\dagger \Pi_k G_k\right)^{-1/2} U$ with U unitary and such that $U \to \mathbf{1}_d$ when $TT^\dagger \to \mathbf{1}_d$.
2. Here is an iteration that yields T in a few rounds without the need of calculating the reciprocal square root of a possibly large matrix: Starting with $T_0 = \mathbf{1}_d$ compute T_1, T_2, ... successively with the aid of the recurrence relation

$$T_{n+1} = T_n - \mathrm{e}^{\mathrm{i}\pi/3} T_n \left[T_n^\dagger \left(TT^\dagger\right)^{-1} T_n - \mathbf{1}_d\right], \tag{11.54}$$

wherein $\left(TT^\dagger\right)^{-1}$ is the given inverse of the right-hand side in (11.53). As long as the step size α is so small that all eigenvalues of $\left(TT^\dagger\right)^{-1}$ are less than 2, which is typically the case in practice without particular precautions, we have $T_n \to T$ with a cubic convergence because

$$\begin{aligned} &T_{n+1}^\dagger \left(TT^\dagger\right)^{-1} T_{n+1} = \mathbf{1}_d + \left[T_n^\dagger \left(TT^\dagger\right)^{-1} T_n - \mathbf{1}_d\right]^3, \\ &\text{implying} \quad T_n^\dagger \left(TT^\dagger\right)^{-1} T_n = \mathbf{1}_d + \left(\alpha^2 \sum_{k=1}^K G_k^\dagger \Pi_k G_k\right)^{3^n}. \end{aligned} \tag{11.55}$$

3. A quadratically convergent iteration is obtained by the replacement $\mathrm{e}^{\mathrm{i}\pi/3} \to \frac{1}{2}$ in (11.54); this may be preferable if $\left(TT^\dagger\right)^{-1}$ is a real matrix and one wishes to have a real matrix for T as well.

4. As mentioned earlier, the POVM resulting from the iteration procedure (11.51)–(11.53) could be a local maximum rather than a global one. Since there are no known sufficiency conditions for a global maximum, one cannot prevent convergence toward a local maximum. All numerical schemes face this generic problem. As a remedy, we run the iteration many times with different initial POVMs, and so reduce the risk of mistaking a local maximum for a global one.
5. Theorem 11.1 states that we can restrict the numerical search to POVMs with rank-1 outcomes that are no more than $K = d^2$ (or $K = \frac{1}{2}d(d+1)$ if all ρ_js are real) in number. To determine the actual value of K, we begin with optimizing for $K = d$, then for $K = d+1$, then for $K = d+2$, until an increase of K no longer gives an increase of the maximal mutual information.—Alternatively, we start with optimizing for $K = d^2$ or $K = \frac{1}{2}d(d+1)$, and then reduce the number of outcomes by identifying equivalent ones. Outcomes Π_k and $\Pi_{k'}$ are equivalent if $p_{jk}p_{j'k'} = p_{j'k}p_{jk'}$ for all j and j', for then $R_k = R_{k'}$, and the pair of outcomes $\left(\Pi_k + \Pi_{k'}, 0\right)$ is as good as the pair $\left(\Pi_k, \Pi_{k'}\right)$. Incidentally, numerical experience seems to indicate [21] that by choosing the initial K value substantially larger than d^2, so that there will surely be superfluous outcomes in the POVM, one reduces substantially the risk of ending up in a local maximum.
6. The choice (11.49) is the basic steepest-ascent strategy where one proceeds in the direction of the gradient. As usual, convergence is improved markedly when one employs *conjugated gradients* instead; see Section 10.6 in [22] or Shewchuk's tutorial [23] and the references therein.

11.6 Examples

11.6.1 Two quantum states in two dimensions

We first consider the simplest example: the situation of two states, $\mathscr{E} = \{\rho_1, \rho_2\}$, in two dimensions, $d = \mathrm{rank}(\rho_1 + \rho_2) = 2$. Since any 2×2 matrix is a linear combination of the identity matrix $\mathbf{1}_2$ and the three familiar matrices of Pauli's matrix vector $\vec{\sigma}$, we write

$$\rho_j = \frac{a_j}{2}(\mathbf{1}_2 + \vec{r}_j \cdot \vec{\sigma}), \quad j = 1,2, \tag{11.56}$$

for the two quantum states. The *Pauli vector* $\vec{r}_j$ is of unit length if ρ_j is a pure state, and shorter if ρ_j is a mixed state. The probabilities of occurrence are both nonzero, $0 < a_1 = 1 - a_2 < 1$.

Numerical studies by ourselves and others, such as work by Fuchs and Peres as reported by Shor [24], strongly suggest the conjecture that there is always a von Neumann measurement among the optimal POVMs if $\mathscr{E}$ is a two-state ensemble. This observation is very important in practice but, unfortunately, no proofs seem to exist in the published literature.

Bearing in mind this conjecture, we restrict the search to POVMs of the form

$$\Pi_1 = \frac{1}{2}(\mathbf{1}_2 + \vec{n}\cdot\vec{\sigma}), \quad \Pi_2 = \frac{1}{2}(\mathbf{1}_2 - \vec{n}\cdot\vec{\sigma}), \tag{11.57}$$

where the unit vector $\vec{n}$ specifies the POVM. Therefore the optimization of the POVM amounts to determining the direction of $\vec{n}$, which is an optimization over two angle parameters.

Then, the joint probabilities $p_{jk} = \mathrm{Tr}(\rho_j \Pi_k)$ and their marginals are

$$\begin{aligned} p_{11} &= \frac{a_1}{2}(1+x_1), & p_{12} &= \frac{a_1}{2}(1-x_1), \\ p_{21} &= \frac{a_2}{2}(1+x_2), & p_{22} &= \frac{a_2}{2}(1-x_2), \\ b_1 &= p_{11} + p_{21} = \frac{1}{2}(1+X), & b_2 &= p_{12} + p_{22} = \frac{1}{2}(1-X), \end{aligned} \tag{11.58}$$

wherein

$$x_1 = \vec{n}\cdot\vec{r}_1, \quad x_2 = \vec{n}\cdot\vec{r}_2, \quad X = a_1 x_1 + a_2 x_2. \tag{11.59}$$

They give

$$I(\mathscr{E};\Pi) = a_1\Phi(x_1) + a_2\Phi(x_2) - \Phi(X) \tag{11.60}$$

with

$$\Phi(x) = \frac{1}{2}\big[(1+x)\log_2(1+x) + (1-x)\log_2(1-x)\big] \tag{11.61}$$

for the information accessed by the POVM (11.57).

When the two quantum states commute with each other, the two Pauli vectors are parallel, $\vec{r}_1 \parallel \vec{r}_2$, and then the optimal POVM is given by $\vec{n} \parallel \vec{r}_1 \parallel \vec{r}_2$. This covers as well the case that one, or both, of the Pauli vectors vanishes. Therefore, in the following we take for granted that $r_1 = |\vec{r}_1| > 0$ and $r_2 = |\vec{r}_2| > 0$, and denote by θ the angle between the two Pauli vectors, $\vec{r}_1\cdot\vec{r}_2 = r_1 r_2 \cos\theta$ with $0 < \theta < \pi$.

An infinitesimal variation of the unit vector $\vec{n}$ is an infinitesimal rotation, $\delta\vec{n} = \vec{\varepsilon} \times \vec{n}$, where $\vec{\varepsilon}$ is an arbitrary infinitesimal vector. The resulting variation

of $I(\mathscr{E};\Pi)$ is of the form $\delta I = \vec{\varepsilon}\cdot\big[\vec{n}\times(\cdots)\big]$, so that $\vec{n}\parallel(\cdots)$ if the POVM (11.57) is optimal.

Since the vector $(\cdots)$ is a linear combination of $\vec{r}_1$ and $\vec{r}_2$, the POVM vector $\vec{n}$ is such a linear combination as well. In fact, then, the optimization of $\vec{n}$ is reduced to finding its orientation in the plane spanned by $\vec{r}_1$ and $\vec{r}_2$, which constitutes a one-parameter problem. Expressed in terms of the angles ϑ_1 and ϑ_2 between $\vec{n}$ and the Pauli vectors,

$$x_1 = \vec{n}\cdot\vec{r}_1 = r_1\cos\vartheta_1\,, \qquad x_2 = \vec{n}\cdot\vec{r}_2 = r_2\cos\vartheta_2 \tag{11.62}$$

with $0\le\vartheta_1,\vartheta_2\le\pi$, we have

$$(\sin\theta)^2\vec{n} = (\cos\vartheta_1 - \cos\vartheta_2\cos\theta)\frac{\vec{r}_1}{r_1} + (\cos\vartheta_2 - \cos\vartheta_1\cos\theta)\frac{\vec{r}_2}{r_2}\,. \tag{11.63}$$

The unit length of $\vec{n}$ implies

$$\big[\cos(\vartheta_1+\vartheta_2)-\cos\theta\big]\big[\cos(\vartheta_1-\vartheta_2)-\cos\theta\big] = 0\,. \tag{11.64}$$

It turns out that the second, not the first, factor vanishes when $I(\mathscr{E};\Pi)$ is maximal, so that the actual constraint is $\cos(\vartheta_1-\vartheta_2)=\cos\theta$, and since the POVM to $-\vec{n}$ is equivalent to the one to $\vec{n}$, we can insist on $\vartheta_2-\vartheta_1=\theta$. The optimization of $\vec{n}$ thus amounts to determining ϑ_1, say.

With $\theta=\vartheta_2-\vartheta_1$ in (11.63), we have

$$\vec{n} = \frac{\sin\vartheta_2}{\sin\theta}\frac{\vec{r}_1}{r_1} - \frac{\sin\vartheta_1}{\sin\theta}\frac{\vec{r}_2}{r_2} \tag{11.65}$$

and the requirement $\vec{n}\parallel(\cdots)$ reads

$$a_1 r_1 \sin\vartheta_1 \log\frac{(1+x_1)(1-X)}{(1-x_1)(1+X)} + a_2 r_2 \sin\vartheta_2 \log\frac{(1+x_2)(1-X)}{(1-x_2)(1+X)} = 0\,, \tag{11.66}$$

which we regard as the equation for ϑ_1 as the basic unknown, with $\vartheta_2=\vartheta_1+\theta$ and x_1,x_2,X as given in (11.62) and (11.59). The variables a_1,a_2,r_1,r_2,θ specify Alice's states, and once the value of ϑ_1 is determined, Bob's optimal POVM is known.

For arbitrary values of a_1,a_2,r_1,r_2,θ, there is no known analytical solution of (11.66). But, as noted by Levitin [25] as well as Fuchs and Caves [17, 8], there is a notable special situation, for which the solution is known and simple: the case of $\det\rho_1=\det\rho_2$ or

$$a_1^2(1-r_1^2) = a_2^2(1-r_2^2)\,. \tag{11.67}$$

When this equation is obeyed, the optimal POVM coincides with the measurement for error minimization [5], that is,

$$\vec{n} = \frac{a_1\vec{r}_1 - a_2\vec{r}_2}{\left|a_1\vec{r}_1 - a_2\vec{r}_2\right|}, \tag{11.68}$$

so that

$$\begin{aligned} x_1 &= \frac{\left(a_1 r_1 - a_2 r_2\cos\theta\right) r_1}{\left|a_1\vec{r}_1 - a_2\vec{r}_2\right|}, \\ x_2 &= \frac{\left(a_1 r_1\cos\theta - a_2 r_2\right) r_2}{\left|a_1\vec{r}_1 - a_2\vec{r}_2\right|}, \\ \text{and} \quad X &= \frac{\left(a_1 r_1\right)^2 - \left(a_2 r_2\right)^2}{\left|a_1\vec{r}_1 - a_2\vec{r}_2\right|}. \end{aligned} \tag{11.69}$$

To justify these remarks, we first note that, if $\vec{n}$ is of the form (11.68), (11.65) implies

$$a_1 r_1 \sin\vartheta_1 = a_2 r_2 \sin\vartheta_2, \tag{11.70}$$

and then (11.66) requires

$$\frac{(1+x_1)(1-X)}{(1-x_1)(1+X)} - 1 = \frac{(1-x_2)(1+X)}{(1+x_2)(1-X)} - 1. \tag{11.71}$$

The subtraction of 1 serves the purpose of making both sides vanish for $x_1 = x_2 = X$, which solution results in $I(\mathscr{E};\Pi) = 0$ and is, therefore, of no further interest. Upon dividing by $x_1 - x_2$, (11.71) turns into

$$a_1(1 - x_1 X) = a_2(1 - x_2 X) \quad \text{or} \quad a_1^2(1 - x_1^2) = a_2^2(1 - x_2^2). \tag{11.72}$$

The identity $(a_1x_1)^2 - (a_2x_2)^2 = (a_1r_1)^2 - (a_2r_2)^2$, which follows from (11.70), now establishes (11.67) as the condition that, indeed, must be met by Alice's states if Bob's optimal POVM is given by the unit vector in (11.68).

Two details of (11.67) are worth pointing out: It does not involve the angle θ between the two Pauli vectors; and, irrespective of the probabilities of occurrence a_1 and a_2, (11.67) is always obeyed if both states are pure ($r_1 = r_2 = 1$).

11.6.2 Trine: Z_3 symmetry in two dimensions

We next discuss the celebrated example of the "trine," where no von Neumann measurement can achieve the accessible information. This example was

proposed and solved partially by Holevo in 1973 [26]. The complete solution was obtained by Sasaki *et al.* in their discussion of Z_N symmetry in the two-dimensional Hilbert space [13].

Three pure states $\rho_j = |j\rangle\langle j|$ $(j = 1,2,3)$ with equal probabilities of occurrence, $a_1 = a_2 = a_3 = \frac{1}{3}$, are given in $d = 2$ dimensions by their kets

$$|1\rangle = \frac{1}{2\sqrt{3}}\begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix},\ |2\rangle = \frac{1}{2\sqrt{3}}\begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix},\ |3\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{11.73}$$

or equivalently by their Pauli vectors,

$$\vec{r}_1 = \frac{1}{2}(-\sqrt{3},0,-1)\,,\ \vec{r}_2 = \frac{1}{2}(\sqrt{3},0,-1)\,,\ \vec{r}_3 = (0,0,1)\,. \tag{11.74}$$

These three vectors are coplanar and point to the corners of an equilateral triangle in the *xz*-plane: they form a *trine*.

The cyclic symmetry of the trine is made explicit by noting that

$$|j\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} \cos(j\theta_0) \\ \sin(j\theta_0) \end{pmatrix} \quad \text{for} \quad j = 1,2,3 \quad \text{with} \quad \theta_0 = \frac{2\pi}{3} \tag{11.75}$$

and

$$U|1\rangle = |2\rangle,\ U|2\rangle = |3\rangle,\ U|3\rangle = |1\rangle \quad \text{with} \quad U = \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{pmatrix}. \tag{11.76}$$

Since $U^3 = \mathbf{1}_2$, the 2×2 matrices $\mathbf{1}_2$, U, U^2 are an irreducible unitary representation of Z_3 on a *real* field, the cyclic group of period 3, and the group acts transitively on the ensemble $\mathscr{E} = \{\rho_1, \rho_2, \rho_3\}$.

According to Subsection 11.4.4, the outcomes Π_k of the optimal POVM can be generated by these unitary matrices from a seed S:

$$\Pi_k = \frac{2}{3}U^k S U^{-k} \quad \text{for} \quad k = 1,2,3 \quad \text{with} \quad S = |v\rangle\langle v|\,. \tag{11.77}$$

The seed ket $|v\rangle$ has to be normalized to unit length, $\langle v|v\rangle = 1$, so we write

$$|v\rangle = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \tag{11.78}$$

where the angle parameter θ specifies the POVM.

Therefore, the problem is to maximize the mutual information $I[\theta] = I(\mathscr{E};\Pi_\theta)$ as a function of θ, with

$$I[\theta] = \frac{1}{3}\sum_{j=1}^{3}\bigl(1+\cos(2\theta+j\theta_0)\bigr)\log\bigl(1+\cos(2\theta+j\theta_0)\bigr)\,. \tag{11.79}$$

This function is $\frac{1}{2}\theta_0$-periodic in θ, $I[\theta+\frac{1}{2}\theta_0] = I[\theta]$, because the POVM with the outcomes of (11.77) and (11.78) does not change as a whole when θ is replaced by $\theta+\frac{1}{2}\theta_0$. It is, therefore, sufficient to consider the range $0 \le \theta < \frac{1}{2}\theta_0$, and one verifies easily that the global maximum of $I[\theta]$ is obtained for $\theta = \frac{1}{6}\pi = \frac{1}{4}\theta_0$. Accordingly, the accessible information is

$$I_{\mathrm{acc}}(\mathscr{E};\Pi) = \log\frac{3}{2} \tag{11.80}$$

in the case of the trine.

The optimal POVM of (11.77) with $\theta = \frac{1}{6}\pi$ consists of three rank-1 operators, $\Pi_k = \frac{1}{3}\bigl(1-\vec{r}_k\cdot\vec{\sigma}\bigr)$, with the vectors $\vec{r}_k$ of (11.74). Thus, whereas the state ensemble $\mathscr{E}$ makes up the trine of $\vec{r}_1$, $\vec{r}_2$, and $\vec{r}_3$, the POVM makes up the "anti-trine" composed of $-\vec{r}_1$, $-\vec{r}_2$, and $-\vec{r}_3$. Since $\rho = \frac{1}{2}\mathbf{1}_2$ here, the roles of the trine and the anti-trine are simply interchanged in the reciprocal situation of (11.10).

When we regard the three two-dimensional kets of (11.73) as spanning a plane in a three-dimensional space, we can lift them jointly out of this plane by giving each the same third component. The cyclic symmetry is maintained thereby. Such a *lifted trine* actually consists of the edges of an obtuse pyramid. As Shor established [24], one needs two seeds for the optimal six-outcome POVM of the lifted trine.

If one lifts the trine by so much that the edges of the pyramid are perpendicular to each other, then clearly a three-outcome POVM of von Neumann type is optimal. In fact, there is a large range of angles between the edges, around the perpendicular-edges geometry, for which the optimal POVM has three outcomes. But for acute pyramids with a rather small angle between the edges, one needs a four-outcome POVM [3, 27].

Instead of lifting the trine, one can distort it in the original two-dimensional space, so that the cyclic symmetry is lost. The optimal POVMs for distorted trines have been found quite recently [28].

11.6.3 Six-states protocol: symmetric group S_3

As a practical example, we now turn to an application that occurs in the security analysis in quantum cryptography. In the raw-data attack on the six-states version [29] of the BB84 protocol [30], eavesdropper Eve gains knowledge by discriminating six rank-2 states in $d = 4$ dimensions [20].

11.6.3.1 States received by Eve

We denote these states by ρ_{js} whereby $j = 1,2,3$ is a ternary index and $s = \pm$ is a binary index, so that we are dealing with three pairs of states. It is expedient to use the following 4×4 matrices for the six states:

$$\rho_{1\pm} = \frac{\varepsilon}{24}\begin{pmatrix} z^2 & \pm z & 0 & 0 \\ \pm z & 1 & 0 & 0 \\ 0 & 0 & 1 & \pm \mathrm{i} \\ 0 & 0 & \mp \mathrm{i} & 1 \end{pmatrix},$$

$$\rho_{2\pm} = \frac{\varepsilon}{24}\begin{pmatrix} z^2 & 0 & \pm z & 0 \\ 0 & 1 & 0 & \mp \mathrm{i} \\ \pm z & 0 & 1 & 0 \\ 0 & \pm \mathrm{i} & 0 & 1 \end{pmatrix},$$

$$\rho_{3\pm} = \frac{\varepsilon}{24}\begin{pmatrix} z^2 & 0 & 0 & \pm z \\ 0 & 1 & \pm \mathrm{i} & 0 \\ 0 & \mp \mathrm{i} & 1 & 0 \\ \pm z & 0 & 0 & 1 \end{pmatrix}, \tag{11.81}$$

where the parameter ε measures the level of noise between the communicating parties that results from the eavesdropping, and $z = \sqrt{4/\varepsilon - 3}$ is a convenient abbreviation. The physically reasonable range of the noise parameter is $0 \le \varepsilon \le 1$ but only communications with $\varepsilon < \frac{2}{3}$ are potentially useful for the purpose of quantum cryptography. Indeed, we will see below that the optimal POVMs are structurally different for $\varepsilon < \frac{2}{3}$ and $\varepsilon \ge \frac{2}{3}$.

The two nonzero eigenvalues of each ρ_{js} are $(2-\varepsilon)/12$ and $\varepsilon/12$, so that all six probabilities are $\frac{1}{6}$ and the six matrices of (11.81) are unitarily equivalent,

$$\rho_{js} = U_{js}\,\rho_{1+}\,U_{js}{}^{\dagger}. \tag{11.82}$$

Here,

$$\rho_{1+} = |1\rangle\langle 1| + |2\rangle\langle 2| \quad \text{with} \quad \langle 1| = \sqrt{\frac{\varepsilon}{24}}(z,1,0,0)$$
$$\text{and} \quad \langle 2| = \sqrt{\frac{\varepsilon}{24}}(0,0,1,\mathrm{i}) \tag{11.83}$$

state the spectral decomposition of ρ_{1+} and so makes its rank-2 nature explicit, and the unitary matrices U_{js} are given by

$$U_{1+} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{1}_4, \qquad U_{1-} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$
$$U_{2+} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad U_{2-} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$
$$U_{3+} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad U_{3-} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{11.84}$$

They form a multiplicative group of order 6 with this group table:

	U_{1+}	U_{1-}	U_{2+}	U_{2-}	U_{3+}	U_{3-}
U_{1+}	U_{1+}	U_{1-}	U_{2+}	U_{2-}	U_{3+}	U_{3-}
U_{1-}	U_{1-}	U_{1+}	U_{3-}	U_{3+}	U_{2-}	U_{2+}
U_{2+}	U_{2+}	U_{2-}	U_{3+}	U_{3-}	U_{1+}	U_{1-}
U_{2-}	U_{2-}	U_{2+}	U_{1-}	U_{1+}	U_{3-}	U_{3+}
U_{3+}	U_{3+}	U_{3-}	U_{1+}	U_{1-}	U_{2+}	U_{2-}
U_{3-}	U_{3-}	U_{3+}	U_{2-}	U_{2+}	U_{1-}	U_{1+}

(11.85)

which shows that it is a nonabelian group that is isomorphic to the symmetric group S_3. It is well known that the representation (11.84) is not irreducible. To get an irreducible representation, we need to carry out the similarity transformation

$$U_{js} \to \widetilde{U}_{js} = T^{-1} U_{js} T \tag{11.86}$$

with the transformation matrix

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 0 & 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 0 & 1/\sqrt{3} & 1/\sqrt{6} & 1\sqrt{2} \end{pmatrix}. \tag{11.87}$$

The transformed unitary matrices give us a direct sum of irreducible representations for the group,

$$\widetilde{U}_{1\pm} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\widetilde{U}_{2\pm} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \mp 1/2 & -\sqrt{3}/2 \\ 0 & 0 & \pm\sqrt{3}/2 & -1/2 \end{pmatrix},$$

$$\widetilde{U}_{3\pm} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \mp 1/2 & \sqrt{3}/2 \\ 0 & 0 & \mp\sqrt{3}/2 & -1/2 \end{pmatrix}. \tag{11.88}$$

They combine a $\phi_0 = 2\pi/3$ rotation and a reflection,

$$\widetilde{U}_{js} = U_{(j-1)\phi_0} \Sigma_s \quad \text{for} \quad j = 1,2,3 \quad \text{and} \quad s = \pm, \tag{11.89}$$

where

$$U_\vartheta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\vartheta & -\sin\vartheta \\ 0 & 0 & \sin\vartheta & \cos\vartheta \end{pmatrix}, \tag{11.90}$$

with ϑ taking on the values 0, ϕ_0, $2\phi_0$ for $j = 1, 2$, and 3, respectively, and

$$\Sigma_+ = \mathbf{1}_4 = \widetilde{U}_{1+}, \quad \Sigma_- = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \widetilde{U}_{1-}. \tag{11.91}$$

Eve's states ρ_{js} are transformed correspondingly, resulting in

$$\widetilde{\rho}_{1\pm} = \frac{\varepsilon}{24} \begin{pmatrix} z^2 & \pm z/\sqrt{3} & \mp z\sqrt{2/3} & 0 \\ \pm z/\sqrt{3} & 1 & 0 & \pm \mathrm{i}\sqrt{2/3} \\ \mp z\sqrt{2/3} & 0 & 1 & \pm \mathrm{i}/\sqrt{3} \\ 0 & \mp \mathrm{i}\sqrt{2/3} & \mp \mathrm{i}/\sqrt{3} & 1 \end{pmatrix},$$

$$\widetilde{\rho}_{2\pm} = \frac{\varepsilon}{24} \begin{pmatrix} z^2 & \pm z/\sqrt{3} & \pm z/\sqrt{6} & \mp z/\sqrt{2} \\ \pm z/\sqrt{3} & 1 & \mp \mathrm{i}/\sqrt{2} & \mp \mathrm{i}/\sqrt{6} \\ \pm z/\sqrt{6} & \pm \mathrm{i}/\sqrt{2} & 1 & \pm \mathrm{i}/\sqrt{3} \\ \mp z/\sqrt{2} & \pm \mathrm{i}/\sqrt{6} & \mp \mathrm{i}/\sqrt{3} & 1 \end{pmatrix},$$

$$\widetilde{\rho}_{3\pm} = \frac{\varepsilon}{24} \begin{pmatrix} z^2 & \pm z/\sqrt{3} & \pm z/\sqrt{6} & \pm z/\sqrt{2} \\ \pm z/\sqrt{3} & 1 & \pm \mathrm{i}/\sqrt{2} & \mp \mathrm{i}/\sqrt{6} \\ \pm z/\sqrt{6} & \mp \mathrm{i}/\sqrt{2} & 1 & \pm \mathrm{i}/\sqrt{3} \\ \pm z/\sqrt{2} & \pm \mathrm{i}/\sqrt{6} & \mp \mathrm{i}/\sqrt{3} & 1 \end{pmatrix}. \tag{11.92}$$

In summary then, the inputs are generated by the group $\widetilde{U}_{js}$ ($j = 1, 2, 3$; $s = \pm$) as

$$\widetilde{\rho}_{js} = \widetilde{U}_{js} \widetilde{\rho}_{1+} \widetilde{U}_{js}^{\dagger}, \tag{11.93}$$

where

$$\widetilde{\rho}_{1+} = |\tilde{1}\rangle\langle\tilde{1}| + |\tilde{2}\rangle\langle\tilde{2}| \quad \text{with} \quad \langle\tilde{1}| = \sqrt{\frac{\varepsilon}{24}}\Big(z, \sqrt{\frac{1}{3}}, -\sqrt{\frac{2}{3}}, 0\Big)$$
$$\text{and} \quad \langle\tilde{2}| = \sqrt{\frac{\varepsilon}{24}}\Big(0, \sqrt{\frac{2}{3}}, \sqrt{\frac{1}{3}}, \mathrm{i}\Big)\,. \tag{11.94}$$

11.6.3.2 Eve's POVM

We find the optimal POVM for Eve by an application of Theorem 11.5. The group structure for the six-states protocol is given by

$$\widetilde{U}_g = \bigoplus_{\ell=1}^{3} (\mathbf{1}_{m_\ell} \otimes u_g^\ell)\,, \tag{11.95}$$

with unit multiplicity for all ℓ values,

$$m_\ell = 1 \quad \text{for} \quad \ell = 1,2,3\,, \tag{11.96}$$

and the inequivalent irreducible representations are

$$\ell = 1 \text{ or } \ell = 2\text{:}\;\; u^1_{g\pm} = 1\,, \quad u^2_{g\pm} = \pm 1 \quad \text{for all } g\,;$$
$$\ell = 3\text{:}\;\; u^3_{1\pm} = \begin{pmatrix} \pm 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$u^3_{2\pm} = \frac{1}{2}\begin{pmatrix} \mp 1 & -\sqrt{3} \\ \pm\sqrt{3} & -1 \end{pmatrix},$$
$$u^3_{3\pm} = \frac{1}{2}\begin{pmatrix} \mp 1 & \sqrt{3} \\ \mp\sqrt{3} & -1 \end{pmatrix}. \tag{11.97}$$

These representations exhaust all inequivalent irreducible representations, since the sum of the squares of the dimensions of the irreducible representations is equal to the order of the group. Indeed, $1^2 + 1^2 + 2^2 = 6$ holds here.

According to Theorem 11.5, an optimal POVM can be generated by the same group by means of

$$\widetilde{\Pi}_g = \frac{4}{6}\widetilde{U}_g \widetilde{S} \widetilde{U}_g^{\,\dagger}\,, \tag{11.98}$$

with the seed $\widetilde{S}$ of the form

$$\widetilde{S} = \bigoplus_{\ell=1}^{3} \frac{d_\ell}{4} |\tilde{v}_\ell\rangle\langle\tilde{v}_\ell| , \tag{11.99}$$

where d_ℓ is the dimension of the respective irreducible representation, and $\langle\tilde{v}_\ell|\tilde{v}_\ell\rangle = 1$ is required for each ℓ. In general, we may need more than one rank-1 state $\widetilde{S}$, and the upper bound is $\sum_\ell m_\ell^2 = 3$. A single seed is, however, enough to reach the accessible information for the specific example under consideration.

Hence we write $\widetilde{S} = |\tilde{v}\rangle\langle\tilde{v}|$ where

$$|\tilde{v}\rangle = \begin{pmatrix} e^{i\phi_1}/2 \\ e^{i\phi_2}/2 \\ e^{i\phi_3}\cos\theta/\sqrt{2} \\ e^{i\phi_4}\sin\theta/\sqrt{2} \end{pmatrix}, \tag{11.100}$$

with real angle parameters $\phi_1, \ldots, \phi_4, \theta$. Since the global phase is irrelevant, the value of one of the ϕ_js can be chosen by a convenient convention, and we set $\phi_1 = 0$ from now on.

Upon defining f_i by

$$\langle\tilde{v}|\tilde{\rho}_{i\pm}|\tilde{v}\rangle = \frac{1}{24}(1 \pm f_i), \tag{11.101}$$

we find

$$\begin{aligned} f_i &= \eta g_i - \frac{\varepsilon}{\sqrt{3}} h_i, \\ g_i &= \frac{1}{2}\cos\phi_2 - \cos\phi_3 \cos\varphi_i \cos\theta - \cos\phi_4 \sin\varphi_i \sin\theta, \\ h_i &= \sin\phi_{23} \sin\varphi_i \cos\theta - \sin\phi_{24} \cos\varphi_i \sin\theta - \sin\phi_{34} \cos\theta \sin\theta. \end{aligned} \tag{11.102}$$

Here $\eta = z\varepsilon/\sqrt{3} = \sqrt{4\varepsilon/3 - \varepsilon^2}$, $\varphi_i = 2\pi(i-1)/3$, and ϕ_{ij} denotes $\phi_{ij} = \phi_i - \phi_j$. The mutual information $I(\rho, \Pi)$ is then given by

$$I(\mathscr{E}; \Pi) = \frac{1}{3} \sum_{i=1}^{3} \Phi(f_i), \tag{11.103}$$

where $\Phi(\ )$ is the function introduced in (11.61).

The accessible information is now obtained by maximizing this mutual information $I(\mathscr{E};\Pi)$ over the four parameters $\phi_2, \phi_3, \phi_4, \theta$. With the help of numerical analysis, we observe that increasing the number of seeds does not provide a larger mutual information than what we get for a single seed.

As we noted above, the cases $\varepsilon < \frac{2}{3}$ and $\varepsilon \geq \frac{2}{3}$ are physically different. This is reflected in the structural difference between the optimal POVMs in these two parameter ranges.

Case $0 \leq \varepsilon < \frac{2}{3}$**:** The optimal POVM is given by

$$\phi_2 = \phi_3 = \phi_4 = 0 \quad \text{and} \quad \theta = \pi. \tag{11.104}$$

In the original representation of (11.81), this is expressed as

$$|v\rangle = T|\tilde{v}\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{3} \\ 0 \\ 0 \end{pmatrix}. \tag{11.105}$$

The accessible information is

$$I_{\text{acc}}(\mathscr{E}) = \frac{1}{3}\Phi(3\eta/2) \tag{11.106}$$

with $\Phi(\,)$ of (11.61) and η as in (11.102). We remark that this optimal POVM is independent of the noise parameter ε, and all its outcomes are real. These findings agree with those obtained in [20], which were obtained with the aid of a numerical search by the method of Section 11.5. This demonstrates the optimality of this POVM.

Case $\frac{2}{3} \leq \varepsilon \leq 1$**:** The optimal POVM has a more complicated structure here, namely it is specified by

$$\begin{aligned} \phi_2 &= -\tan^{-1}\sqrt{\frac{2(3\varepsilon-2)}{4-3\varepsilon}}, \\ \phi_3 &= \tan^{-1}\sqrt{\frac{3\varepsilon-2}{2(4-3\varepsilon)}}, \\ \phi_4 &= 0, \\ \theta &= \pi + \tan^{-1}\sqrt{\frac{3\varepsilon-2}{2-\varepsilon}}, \end{aligned} \tag{11.107}$$

where $-\frac{1}{2}\pi < \phi_2, \phi_3, \theta - \pi < \frac{1}{2}\pi$. This POVM amounts to $f_1 = 1$ and $f_2 = f_3 = 0$ in (11.103), so that the accessible information is

$$I_{\rm acc}(\mathscr{E}) = \frac{1}{3}\Phi(1) = \frac{1}{3}. \tag{11.108}$$

We note in passing that there are other POVMs that also give $I_{\rm acc} = \frac{1}{3}$ for the whole range $\frac{2}{3} \le \varepsilon \le 1$.

11.6.4 Four-group in four dimensions

As a simplest nontrivial group, we study the four-group—r *Klein group*, or *vierergruppe*—which is the noncyclic group of order four. One meets this group structure in the eavesdropping analysis for the BB84 protocol [31]. Here we give a discussion based on a toy model for the four-group in a 4-dimensional Hilbert space.

Each of the four quantum states $\rho_1, \ldots, \rho_4$ is a rank-2 state, and the total state $\rho = \rho_1 + \cdots + \rho_4$ has rank 4, and we have equal probabilities of occurrence:

$$\rho_j = |\psi_j\rangle\langle\psi_j| + |\phi_j\rangle\langle\phi_j| \quad \text{with} \quad \mathrm{Tr}\,\rho_j = \frac{1}{4} \tag{11.109}$$

and

$$\left.\begin{matrix}|\psi_1\rangle\\ |\psi_2\rangle\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} a \\ \pm b \\ 0 \\ 0 \end{pmatrix}, \qquad \left.\begin{matrix}|\psi_3\rangle\\ |\psi_4\rangle\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} a \\ 0 \\ \pm b \\ 0 \end{pmatrix},$$

$$\left.\begin{matrix}|\phi_1\rangle\\ |\phi_2\rangle\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ \pm c \\ d \end{pmatrix}, \qquad \left.\begin{matrix}|\phi_3\rangle\\ |\phi_4\rangle\end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} 0 \\ \pm c \\ 0 \\ -d \end{pmatrix}. \tag{11.110}$$

Here a, b, c, d are real constants satisfying $a^2 + b^2 + c^2 + d^2 = 1$. We express these states using unitary matrices U_j as $\rho_j = U_j \rho_1 U_j{}^\dagger$, whereby

$$\left.\begin{matrix}U_1\\ U_2\end{matrix}\right\} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \left.\begin{matrix}U_3\\ U_4\end{matrix}\right\} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{11.111}$$

They constitute the four-group with the familiar group table

$$\begin{array}{c|cccc} & U_1 & U_2 & U_3 & U_4 \\ \hline U_1 & U_1 & U_2 & U_3 & U_4 \\ U_2 & U_2 & U_1 & U_4 & U_3 \\ U_3 & U_3 & U_4 & U_1 & U_2 \\ U_4 & U_4 & U_3 & U_2 & U_1 \end{array} \tag{11.112}$$

where we note that the four-group is abelian and has order-2 subgroups consisting of $U_1 = \mathbf{1}_4$ and either U_2 or U_3 or U_4.

The representation (11.111) is not irreducible. In order to obtain a direct sum of inequivalent irreducible representations $\widetilde{U}_j$, we introduce the following transformation T:

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{11.113}$$

As is fitting for an abelian group, the transformed unitary matrices $\widetilde{U}_j = T^{-1} U_j T$ have diagonal components only:

$$\left.\begin{matrix} \widetilde{U}_1 \\ \widetilde{U}_2 \end{matrix}\right\} = \mathrm{diag}(1, \pm 1, \pm 1, 1), \qquad \left.\begin{matrix} \widetilde{U}_3 \\ \widetilde{U}_4 \end{matrix}\right\} = \mathrm{diag}(1, \pm 1, \mp 1, -1). \tag{11.114}$$

They are indeed the direct sum of irreducible four-dimensional representations of the four-group. These representations consist of a direct sum of four different inequivalent representations. Each of inequivalent representations is one-dimensional. We also note the unit multiplicity for all four representations.

According to Theorem 11.5, we could need as many as 4 seeds. It is important to know that if we restrict ourself to the single-orbital case, then the optimal POVM generated by this group cannot have real outcomes. This is so because the seed has to have a unit length for each component by Schur's lemma. Therefore, we encounter the perhaps unexpected situation where we need a complex seed even though all input states and group representations are expressed as real quantities. As we will see later, there also exist real seeds

which provide the accessible information, but then we need more than a single orbit.

We parameterize the seed ket $|\tilde{v}_1\rangle$ by three angle parameters θ_1, θ_2 and θ_3,

$$|\tilde{v}_1\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ e^{i\theta_1} \\ e^{i\theta_2} \\ e^{i\theta_3} \end{pmatrix}. \tag{11.115}$$

The group generated outcomes of the POVM are then given by

$$\widetilde{\Pi}_k = \widetilde{U}_k|\tilde{v}_1\rangle\langle\tilde{v}_1|\widetilde{U}_k^\dagger = |\tilde{v}_k\rangle\langle\tilde{v}_k|, \tag{11.116}$$

and the optimization requires the determination of the three θ_ks.

Corresponding to the transformation on the unitary matrices, the quantum states ρ_j are transformed into $\tilde{\rho}_j = T^{-1}\rho_j T$, or $|\widetilde{\psi}_j\rangle = T^{-1}|\psi_j\rangle$ and $|\widetilde{\phi}_j\rangle = T^{-1}|\phi_j\rangle$. Explicitly we have

$$\left.\begin{matrix} |\widetilde{\psi}_1\rangle \\ |\widetilde{\psi}_2\rangle \end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} a \\ \pm b/\sqrt{2} \\ \mp b/\sqrt{2} \\ 0 \end{pmatrix}, \qquad \left.\begin{matrix} |\widetilde{\psi}_3\rangle \\ |\widetilde{\psi}_4\rangle \end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} a \\ \pm b/\sqrt{2} \\ \pm b/\sqrt{2} \\ 0 \end{pmatrix},$$

$$\left.\begin{matrix} |\widetilde{\phi}_1\rangle \\ |\widetilde{\phi}_2\rangle \end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} 0 \\ \pm c/\sqrt{2} \\ \pm c/\sqrt{2} \\ d \end{pmatrix}, \qquad \left.\begin{matrix} |\widetilde{\phi}_3\rangle \\ |\widetilde{\phi}_4\rangle \end{matrix}\right\} = \frac{1}{2}\begin{pmatrix} 0 \\ \pm c/\sqrt{2} \\ \mp c/\sqrt{2} \\ -d \end{pmatrix}. \tag{11.117}$$

For $\bar{\rho}_j$ defined by

$$\bar{\rho}_j = \langle\tilde{v}_1|\tilde{\rho}_j|\tilde{v}_1\rangle = |\langle\tilde{v}_1|\tilde{\psi}_j\rangle|^2 + |\langle\tilde{v}_1|\tilde{\phi}_j\rangle|^2, \tag{11.118}$$

we find

$$\begin{aligned} \bar{\rho}_{1,2} = \frac{1}{16}\big[1 - (b^2 - c^2)\cos(2\theta_-) \mp 2\sqrt{2}ab\sin\theta_+\sin\theta_- \\ \pm 2\sqrt{2}cd\cos(\theta_+ - \theta_3)\cos\theta_-\big], \\ \bar{\rho}_{3,4} = \frac{1}{16}\big[1 + (b^2 - c^2)\cos(2\theta_-) \pm 2\sqrt{2}ab\cos\theta_+\cos\theta_- \\ \pm 2\sqrt{2}cd\sin(\theta_+ - \theta_3)\sin\theta_-\big], \end{aligned} \tag{11.119}$$

where $\theta_\pm = (\theta_1 \pm \theta_2)/2$. Finally, the mutual information is expressed as

$$I(\mathscr{E};\Pi)[\theta_k] = \frac{1}{4}\sum_{j=1}^{4}(16\bar{p}_j)\log(16\bar{p}_j), \tag{11.120}$$

which is to be regarded as a function of the three θ_ks.

The general solution to this optimization problem is not known as yet. But if the parameters b and c are equal, we have the analytical solution at hand.

Upon setting $b = c$, the ensemble of states is characterized by two independent parameters because a, b, and d must obey $a^2 + d^2 + 2b^2 = 1$. We define A and θ_0 by

$$\begin{aligned} A &= 2b\sqrt{2(a^2+d^2)} = 2b\sqrt{2(1-2b^2)}, \\ \theta_0 &= \tan^{-1}\frac{d}{a}. \end{aligned} \tag{11.121}$$

The expression for the mutual information then simplifies to

$$I(\mathscr{E};\Pi)[\theta_k] = \frac{1}{2}\sum_{j=1}^{2}\Phi(f_j), \tag{11.122}$$

where f_1 and f_2 are

$$\begin{aligned} f_1 &= A\left[\cos\theta_0\sin\theta_+\sin\theta_- - \sin\theta_0\cos(\theta_+-\theta_3)\cos\theta_-\right], \\ f_2 &= A\left[\cos\theta_0\cos\theta_+\cos\theta_- + \sin\theta_0\sin(\theta_+-\theta_3)\sin\theta_-\right]. \end{aligned} \tag{11.123}$$

The partial derivatives with respect to the θ_ks are

$$\begin{aligned} \frac{\partial}{\partial\theta_1}I[\theta_k] &= \frac{A}{8}\left[\cos\theta_0\sin\theta_1\log\frac{R_1}{R_2} + \sin\theta_0\sin(\theta_1-\theta_3)\log(R_1R_2)\right], \\ \frac{\partial}{\partial\theta_2}I[\theta_k] &= \frac{A}{8}\left[-\cos\theta_0\sin\theta_2\log(R_1R_2) + \sin\theta_0\sin(\theta_2-\theta_3)\log\frac{R_1}{R_2}\right], \\ \frac{\partial}{\partial\theta_3}I[\theta_k] &= -\frac{\sin\theta_0 A}{8}\left[\sin(\theta_1-\theta_3)\log(R_1R_2) + \sin(\theta_2-\theta_3)\log\frac{R_1}{R_2}\right], \\ \text{with } R_j &= \frac{1+f_j}{1-f_j}. \end{aligned} \tag{11.124}$$

The right-hand sides are of the form $X_i\log R_1 + Y_i\log R_2$, and the necessary conditions for stationary points, that is: $\frac{\partial}{\partial\theta_i}I[\theta_k] = 0$ for $i = 1,2,3$, are then equivalent to

$$(X_iY_j - X_jY_i)\log R_l = 0 \quad \text{for} \quad l = 1,2 \quad \text{and} \quad i,j = 1,2,3. \tag{11.125}$$

Since $\log R_1 = \log R_2 = 0$ gives zero mutual information, the coefficients must be zero, i.e.,

$$X_i Y_j - X_j Y_i = 0 \quad \text{for} \quad (i,j) = (1,2),(2,3),(3,1)\,. \tag{11.126}$$

Explicitly, they are

$$\begin{aligned}
&\cos^2\theta_0 \sin\theta_1 \sin\theta_2 + \sin^2\theta_0 \sin(\theta_1-\theta_3)\sin(\theta_2-\theta_3) = 0\,,\\
&\big[\cos\theta_0 \sin\theta_2 + \sin\theta_0 \sin(\theta_1-\theta_3)\big]\sin(\theta_2-\theta_3) = 0\,,\\
&\big[\cos\theta_0 \sin\theta_1 - \sin\theta_0 \sin(\theta_2-\theta_3)\big]\sin(\theta_1-\theta_3) = 0\,,
\end{aligned} \tag{11.127}$$

two of which imply the third. One verifies immediately that

$$\theta_1 = \theta_0\,, \quad \theta_2 = -\theta_0\,, \quad \theta_3 = -\frac{\pi}{2} \tag{11.128}$$

solve these equations, and this solution gives the accessible information.

The optimal POVM thus found consists of a single orbit with the seed ket given by

$$|v_1\rangle = T|\tilde{v}_1\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2}\mathrm{i}\sin\theta_0 \\ \sqrt{2}\cos\theta_0 \\ -\mathrm{i} \end{pmatrix} \tag{11.129}$$

in the original representation of (11.109). This corresponds to $f_1 = A$ and $f_2 = 0$ in (11.123). The resulting accessible information is

$$I_{\mathrm{acc}}(\mathscr{E}) = \frac{1}{2}\Phi(A) = \frac{1}{2}\Phi\Big(2b\sqrt{2(1-2b^2)}\Big)\,. \tag{11.130}$$

Rather intriguingly, the accessible information depends only on one of the parameters.

We next show how to construct a real optimal POVM out of this complex solution. Split the optimal seed $|\tilde{v}_1\rangle$ into real and imaginary parts,

$$|\tilde{v}_1\rangle = |\tilde{v}_{1\mathrm{r}}\rangle + \mathrm{i}|\tilde{v}_{1\mathrm{i}}\rangle\,, \tag{11.131}$$

and consider another set of outcomes generated by the complex conjugate seed

$$|\tilde{v}_1^*\rangle = |\tilde{v}_{1\mathrm{r}}\rangle - \mathrm{i}|\tilde{v}_{1\mathrm{i}}\rangle\,. \tag{11.132}$$

These two POVMs give the same joint probabilities and, therefore, the same amount of mutual information. It then follows from the convexity of the mutual information (11.22) that a real rank-2 seed

$$\widetilde{S}_{\text{real}} = \frac{1}{2}\left(|\tilde{v}_1\rangle\langle\tilde{v}_1| + |\tilde{v}_1^*\rangle\langle\tilde{v}_1^*|\right) = |\tilde{v}_{1\text{r}}\rangle\langle\tilde{v}_{1\text{r}}| + |\tilde{v}_{1\text{i}}\rangle\langle\tilde{v}_{1\text{i}}| \tag{11.133}$$

gives the accessible information as well.

As we mentioned before, the optimal POVM is not unique as a rule. Here we have already a choice between a POVM with four complex rank-1 outcomes, its complex conjugate POVM, or a POVM with four real rank-2 outcomes. These three POVMs can be regarded as equivalent in the sense that they give rise to the same joint probabilities.

In addition, there is a one-parameter family of inequivalent POVMs, each having four real rank-2 outcomes. In the original representation of (11.109) the outcomes are of the form

$$\Pi_k = |u_k\rangle\langle u_k| + |v_k\rangle\langle v_k| \quad \text{for} \quad k = 1,\ldots,4 \tag{11.134}$$

with the kets $|u_k\rangle$ and $|v_k\rangle$ depending on the real parameter r in the following way:

$$\left.\begin{matrix}|u_1\rangle\\ |u_2\rangle\end{matrix}\right\} = \frac{\sqrt{\cos^2\theta_0 + r}}{2}\begin{pmatrix}1/\cos\theta_0\\ \pm\sqrt{2}\\ 0\\ 0\end{pmatrix},$$

$$\left.\begin{matrix}|u_3\rangle\\ |u_4\rangle\end{matrix}\right\} = \frac{\sqrt{\cos^2\theta_0 - r}}{2}\begin{pmatrix}1/\cos\theta_0\\ 0\\ \pm\sqrt{2}\\ 0\end{pmatrix},$$

$$\left.\begin{matrix}|v_1\rangle\\ |v_2\rangle\end{matrix}\right\} = \frac{\sqrt{\sin^2\theta_0 + r}}{2}\begin{pmatrix}0\\ 0\\ \pm\sqrt{2}\\ 1/\sin\theta_0\end{pmatrix},$$

$$\left.\begin{matrix}|v_3\rangle \\ |v_4\rangle\end{matrix}\right\} = \frac{\sqrt{\sin^2\theta_0 - r}}{2} \begin{pmatrix} 0 \\ \mp\sqrt{2} \\ 0 \\ 1/\sin\theta_0 \end{pmatrix}. \tag{11.135}$$

The parameter r is restricted to the range

$$|r| \leq \min\bigl(\cos^2\theta_0, \sin^2\theta_0\bigr) = \frac{\min(a^2, d^2)}{a^2 + d^2} \tag{11.136}$$

but is otherwise arbitrary. We note that, when $|r|$ is maximal, the corresponding optimal POVM has two rank-2 outcomes and two rank-1 outcomes, rather than four rank-2 outcomes. We note further that the POVM is not group-covariant when $r \neq 0$.

11.7 Summary and outlook

We have given a brief introduction to, and summary of, the problem of determining the accessible information about a given set of quantum states. At present, the problem (11.21) remains open because there is no generally applicable method by which we can determine the optimal POVM and the accessible information. We note in particular the lack of sufficient conditions by which one could judge whether a candidate POVM is optimal. Until such conditions are established, the strategy of choice is a combination of a numerical search—possibly by the method described in Section 11.5—with an analytical check of the necessary conditions (11.30).

We recall further that the seemingly simple conjecture mentioned after (11.56) has not been proven as yet. A proof would surely constitute a major step forward because in practice one often encounters the situation of the conjecture, namely the task of distinguishing optimally between two quantum states.

We also emphasize that obtaining analytical expressions for the optimal POVM usually requires solving a set of nonlinear equations, and we would not expect that they can be solved routinely, with closed-form solutions. This point is well illustrated by the example in Subsection 11.6.1, arguably the simplest nontrivial situation.

In practice, however, we are rarely looking for the accessible information about random quantum states. Rather, the quantum states of interest tend

to possess certain symmetries among them. We can then apply the group-covariant POVM method of Subsection 11.4.4 for solving the problem as demonstrated by the examples of Subsections 11.6.2–11.6.4. Nevertheless, the numerical strategy explained in detail in Section 11.5 lends us significant help in the search for optimal POVMs. A major problem thereby is, of course, to discriminate between local and global maxima. Further studies are clearly necessary.

We remark that in general the optimal POVM is not unique for a given set of quantum states. We have demonstrated this nonuniqueness by the example of Subsection 11.6.4, where we report an optimal group-covariant von Neumann measurement, an optimal group-covariant POVM that is not of von Neumann type, and a family of inequivalent optimal POVMs that are not group-covariant. From the purely theoretical point of view, these POVMs are equally good in the sense of providing the accessible information. On the other hand, however, there are great differences between them when a physical implementation of the POVM is required. Generally speaking, von Neumann projection measurements and nonprojection measurements belong to different classes of measurement schemes.

This suggests that one should examine thoroughly under which conditions a von Neumann measurement can extract the accessible information about the given quantum states. The conjecture mentioned above is particularly relevant in this context.

Acknowledgments

We are grateful for numerous discussions with Janet Anders, Wee Kang Chua, Thomas Decker, Dagomir Kaszlikowski, Shang Yong Looi, and Jaroslav Řeháček. J. S. and B.-G. E. wish to thank Hans Briegel for the generous hospitality extended to them at the Institute for Quantum Optics and Quantum Information in Innsbruck, where part of this work was done. This work is supported by A*STAR Temasek Grant 012-104-0040 and NUS Grant WBS R144-000-116-101.

References

[1] D. Bruß, M. Christandl, A. Ekert, B.-G. Englert, D. Kaszlikowski, and C. Macchiavello, Phys. Rev. Lett. **91** (2003) art. 097901 (4 pages).

[2] Y. C. Liang, D. Kaszlikowski, B.-G. Englert, L. C. Kwek, and C. H. Oh, Phys. Rev. A **68** (2003) art. 022324 (9 pages).

[3] D. Kaszlikowski, A. Gopinathan, Y. C. Liang, L. C. Kwek, and B.-G. Englert, Phys. Rev. A **70** (2004) art. 032306 (5 pages).

[4] B.-G. Englert, D. Kaszlikowski, H. K. Ng, W. K. Chua, J. Řeháček, and J. Anders, *Highly Efficient Quantum Key Distribution With Minimal State Tomography*, eprint arXiv:quant-ph/0412089 (2004).

[5] C. W. Helstrom, *Quantum Detection and Estimation Theory*, (Academic Press, New York 1976).

[6] T. Cover and J. Thomas, *Elements of Information Theory*, (John Wiley & Sons, New York 1991).

[7] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, Cambridge 2000).

[8] C. A. Fuchs, *Distinguishability and Accessible Information in Quantum Theory*, Ph.D. Thesis, University of New Mexico (1995), eprint arXiv:quant-ph/9601020 (1996).

[9] A. S. Holevo, J. Multivariate Anal. **3** (1973) 337–394.

[10] *Quantum State Estimation*, edited by M. Paris and J. Řeháček, Lecture Notes in Physics, Vol. 649 (Springer, Berlin 2004).

[11] A. S. Holevo, *Coding Theorems for Quantum Channels*, eprint arXiv:quant-ph/9809023 (1998).

[12] E. B. Davies, IEEE Trans. Inf. Theory **24** (1978) 596–599.

[13] M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, and O. Hirota, Phys. Rev. A **59** (1999) 3325–3335.

[14] B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. Lett. **76** (1996) 3452–3455.

[15] A. S. Holevo, Probl. Peredachi Inf. **9** (1973) 3–11; English translation: Probl. Inf. Transm. (USSR) **9** (1973) 177–183.

[16] R. Jozsa, D. Robb, and W. K. Wootters, Phys. Rev. A **49** (1994) 668–677.

[17] C. A. Fuchs and C. M. Caves, Phys. Rev. Lett. **73** (1994) 3047–3050.

[18] T. Decker, *Symmetric measurements attaining the accessible information*, eprint arXiv:quant-ph/0509122 (2005).

[19] F. H. Willeboordse, A. Gopinathan, and D. Kaszlikowski, Phys. Rev. A **71** (2005) art. 042310 (4 pages).

[20] J. Řeháček, B. -G. Englert, and D. Kaszlikowski, Phys. Rev. A **71** (2005) art. 054303 (4 pages).

[21] J. Řeháček, private communication (2005).

[22] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (2nd edition, Cambridge University Press 1992).

[23] J. R. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain (Edition* $1\frac{1}{4}$*)*, available at www.cs.cmu.edu/∼quake-papers/painless-conjugate-gradient.pdf.

[24] P. W. Shor, *On the Number of Elements in a POVM Attaining the Accessible Information*, eprint arXiv:quant-ph/0009077 (2000).

[25] L. B. Levitin, *Optimal quantum measurements for two pure and mixed states*, in: *Quantum Communications and Measurement*, edited by V. P. Belavkin, O. Hirota, and R. L. Hudson, (Plenum Press, New York 1995) 439–448.

[26] A. S. Holevo, Probl. Peredachi Inf. **9** (1973) 31–42; English translation: Probl. Inf. Transm. (USSR) **9** (1973) 110–118.

[27] D. Kaszlikowski, A. Gopinathan, Y. C. Liang, L. C. Kwek, and B.-G. Englert, *How well can you know the edge of a quantum pyramid?* eprint arXiv:quant-ph/0307086 (2003).

[28] M. R. Frey, Phys. Rev. A **73** (2006) art. 032309 (7 pages).

[29] D. Bruß, Phys. Rev. Lett. **81** (1998) 3018–3021.

[30] C. H. Bennett and G. Brassard, in: *Proceedings of the IEEE Conference on Computers, Systems, and Signal Processing Bangalore, India, December 1984* (IEEE, New York 1984) 175–179.

[31] S. M. Assad, J. Suzuki, and B.-G. Englert, Int. J. Quant. Inf. **4** (2006) 1003–1012.