

Chapter 8

Recent Developments in Quantum Zero-Error Information Theory

In the previous Chaps. 6–7 some recent developments and applications of the quantum zero-error information theory were introduced. In this chapter we introduce some contributions from other authors to the field.

This chapter is organized as follows. We revisit some nonlocal phenomena by Bell's inequalities and their consequences in Sect. 8.1. After that, we introduce some definitions that did not appear so far. We revisit relevant contrasts of classical and quantum correlations and discuss a proof of the Bell's inequality. Also, due to their importance, we revisit Gleason's and Kochen-Specker's theorems. We observe that this section has a historical flavor, so the reader, based on his own background, can skip this first section and go straight to the next section.

The classical zero-error capacity of a quantum channel, introduced in Chap. 5, was defined in terms of the clique number of the characteristic graph of a quantum channel. Now, in Sect. 8.2, we comment on the results introduced more recently: the literature by Scarpa, Severini, and Mancinska [26, 32]. Their contribution, mainly the second one, clearly binds Kochen-Specker (also known as Bell-Kochen-Specker) theorems to the quantum zero-error information theory.

A quantum version of the Wielandt's inequality [31] is described in Sect. 8.3. This inequality states an upper bound to the number of uses of a quantum channel in order to map an arbitrary density operator to a full rank operator. In this interesting paper, the authors state a remarkable relation with the quantum zero-error information theory intermediated by dichotomy theorems.

A variant of the zero-error capacity which considers entanglement assistance is presented in Sect. 8.4. Results of Winter et al. [14, 15, 17] in which non-commutative graphs are used for the quantum zero-error information theory are presented in Sect. 8.5. A quantum version of the Lovász theta function and some alternate definitions for the zero-error capacity of a quantum channel are presented as well.

A non-trivial application of zero-error quantum channels to help in determining the complexity class of a well-known problem was proposed by Beigi and Shor [5] which is now depicted in Sect. 8.5.

8.1 Bell's Inequalities

Entanglement motivated the famous article “Can Quantum Mechanical Description of Physical Reality Be Considered Complete?” written by Einstein, Poldosky, and Rosen, EPR for short [19]. After an important and long discussion, concepts like *principle of locality*, *elements of reality*, and *hidden variables* were defeated. The principle of locality, for example, claims that the events occurring in place are independent of parameters, eventually controlled at another “distant place” in the same time, but it was not confirmed.

The main assumption in EPR argument is the a priori concept of element of reality that could be obeyed by the Nature. The EPR paper aimed to show that quantum mechanics was an incomplete theory based on a *sufficient condition* for a physical property to be an element of reality:

“If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unit) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity” [19].

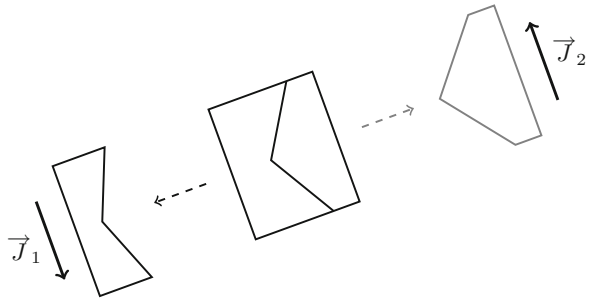
Example 8.1 (Quantum Correlations are Stronger than the Classical Ones [29]). This example sets forth that quantum correlations are in general stronger than the classical ones. Consider a block of explosive material, in rest at $t = 0$, so at this time with angular momentum $\vec{J} = 0$ exploding in two asymmetric parts, as shown in Fig. 8.1. Due to conservation laws [39, p. 323], the two parts carry angular moments, $\vec{J}_1 = -\vec{J}_2$, respectively.

Suppose observers detecting the fragments and measuring the *classical dynamical variables* $a = \text{sign}(\vec{\alpha} \cdot \vec{J}_1)$ and $b = \text{sign}(\vec{\beta} \cdot \vec{J}_2)$, respectively, where $|\alpha\rangle$ and $|\beta\rangle$ are arbitrary unit vectors chosen by the observers. Obviously, $a, b = \pm 1$.

For N repetitions of the experiment, with directions of \vec{J}_1 and \vec{J}_2 randomly distributed, the averages are near to zero, that is,

$$\langle a \rangle = \frac{1}{N} \sum_{j=1}^N a_j \approx 0, \quad \langle b \rangle = \frac{1}{N} \sum_{j=1}^N b_j \approx 0. \quad (8.1)$$

Fig. 8.1 Classical setup with zero angular momentum



In order to compare their results, the observers calculate the *correlation*, defined by

$$\langle ab \rangle = \frac{1}{N} \sum_{j=1}^N a_j b_j. \tag{8.2}$$

The correlation is not zero in general. For concreteness, taking $\vec{\alpha} = \vec{\beta}$, the observers get $a_j = -b_j$, and in this case correlation yields $\langle ab \rangle = -1$.

For arbitrary $\vec{\alpha}$ and $\vec{\beta}$ the solution is [29]:

$$\langle ab \rangle = -1 + \frac{2\theta}{\pi}, \quad 0 \leq \theta \leq \pi, \tag{8.3}$$

where θ stands for the angle between directions $\vec{\alpha}$ and $\vec{\beta}$. Notice that the correlation increases linearly from -1 to $+1$ as θ varies from 0 to π . Such correlation is shown in the plot of Fig. 8.2

Now let's consider the quantum turn. Consider the quantum analogy taking into consideration the singlet (entangled state)

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{8.4}$$

Assume that observers measure the observables $\vec{\alpha} \cdot \sigma_a$ along the axis $\vec{\alpha}$ and $\vec{\beta} \cdot \sigma_b$, along the axis $\vec{\beta}$, respectively. As before, in the classical analog, unit vectors are arbitrarily chosen by the observers and the possible values of a and b measurements are ± 1 . The average values are both zero, that is

$$\langle a \rangle = \langle b \rangle = 0. \tag{8.5}$$

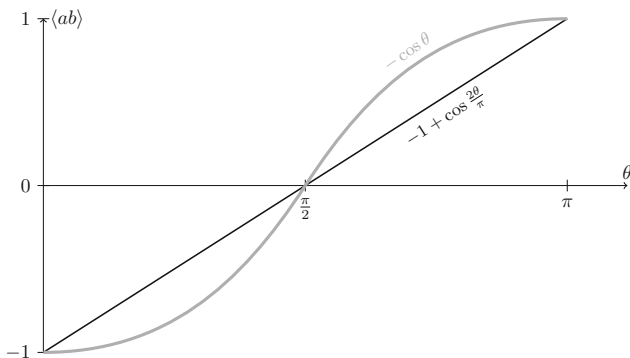


Fig. 8.2 Classical and quantum correlations

Furthermore, correlation can be calculated, according to quantum mechanics rules, as

$$\langle ab \rangle = \langle \psi | (\vec{\alpha} \cdot \sigma_a)(\vec{\beta} \cdot \sigma_b) | \psi \rangle, \quad (8.6)$$

where σ_a and σ_b are the Pauli matrices of the systems a and b , respectively.

For the singlet one has

$$\sigma_a | \psi \rangle = -\sigma_b | \psi \rangle. \quad (8.7)$$

Therefore, using the identity $(\vec{\alpha} \cdot \sigma)(\vec{\beta} \cdot \sigma) = \vec{\alpha} \cdot \vec{\beta} + i(\vec{\alpha} \times \vec{\beta}) \cdot \sigma$, the correlation is then obtained

$$\langle ab \rangle = -\cos \theta. \quad (8.8)$$

The main remark here is that quantum correlation, which is also shown in Fig. 8.2, is stronger than the classical correlation for all values of θ , except $\theta = 0$, $\frac{\pi}{2}$ and π .

The EPR paradox was solved by the *Bell's inequality*. It is interesting to remark that the inequality is not about quantum mechanics, rather its proof is general and independent of Physics. The central statement is that if one assumes validity of principle of locality, then there is an upper bound to the correlation between distant events. What Bell's inequality states is that local realism is incompatible with quantum mechanics.

In order to see an explanation why this happens, one must consider the thought experiment outlined in Fig. 8.3. Assuming the EPR principle, that is assuming the truth of local realism or the existence of hidden variables, we shall perform calculations to obtain the Bell's CHSH inequalities¹ [11]. After, real measurements demonstrate the violation of those inequalities.

In the thought experiment, a physicist, say, Charlie, repeats a large number, N , of preparations of two particles (say “left” and “right,” respectively) and send, one by one, to his colleagues, Alice and Bob. The left particle is sent to Alice and the right one is sent to Bob. Later, Alice and Bob perform simultaneously measurements on their respective particles. Alice's lab is too far from Bob's lab in such a way their respective actions are *concurrent* [30], that is, their actions are *relativistically disconnected*.

Additionally, Alice is free to choose directions $\vec{\alpha}$ or $\vec{\beta}$ to perform her measurements, which results in a random variable denoted by $A = \pm 1$, if she chooses direction $\vec{\alpha}$, or a random variable denoted by $B = \pm 1$ if she chooses direction $\vec{\beta}$. Similarly, Bob is free to choose directions $\vec{\gamma}$ or $\vec{\delta}$, and from his measurement obtains random variables $C = \pm 1$ and $D = \pm 1$, respectively.

¹The letters CHSH are a mention to the authors of this form of Bell's inequality [11].

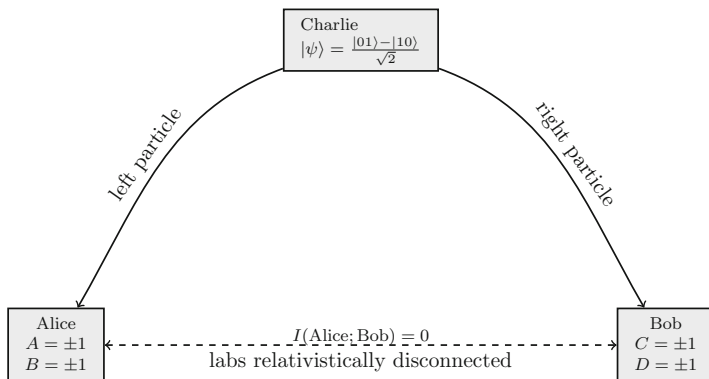


Fig. 8.3 Bell's inequality

Table 8.1 Values assumed by the random variables A and B

	$A = -1$	$A = 1$		$A = -1$	$A = 1$
(a) Values for $A + B$			(b) Values for $A - B$		
$B = -1$	-2	0	$B = -1$	0	2
$B = 1$	0	2	$B = 1$	-2	0

Now consider the random variable V defined by the following sum

$$V = AC + BC + BD - AD \tag{8.9}$$

$$= (A + B)C + (B - A)D, \tag{8.10}$$

where (8.10) follows from a simple rearrangement. From (8.10), it is clear that, as $A = \pm 1$ and $B = \pm 1$, either $(A + B)C = 0$ or $(B - A)D = 0$. From this, according to Table 8.1, it is easy to check that

$$V = AC + BC + BD - AD = \pm 2. \tag{8.11}$$

Now consider the expected value $\mathbb{E}[V]$:

$$\mathbb{E}[V] = \mathbb{E}[AC + BC + BD - AD] \tag{8.12}$$

$$= \sum_{a,b,c,d} \Pr[a, b, c, d](ac + bc + bd - ad) \tag{8.13}$$

$$\leq 2 \sum_{a,b,c,d} \Pr[a, b, c, d] \tag{8.14}$$

$$\leq 2, \tag{8.15}$$

where (8.12) and (8.13) are definitions, (8.14) is justified by (8.11), and (8.15) is because probabilities sum to one. On the other hand, from linearity of the

expectation

$$\mathbb{E}[V] = \mathbb{E}[AC] + \mathbb{E}[BC] + \mathbb{E}[BD] - \mathbb{E}[AD] \quad (8.16)$$

Comparing (8.15) with (8.16), we get the Bell's CHSH inequality:

$$\mathbb{E}[AC] + \mathbb{E}[BC] + \mathbb{E}[BD] - \mathbb{E}[AD] \leq 2. \quad (8.17)$$

Recall that the last inequality, shown in (8.17), was obtained based on the principle of local realism, and there is nothing wrong with this formula under that assumption. However, here, the authors objective was to check its validity for quantum mechanics. A question that arises is: how to perform this task?

Fortunately, the expectations on the left side of (8.17) can be estimated, with accuracy $\frac{1}{\sqrt{N}}$, through repeating the experiments N times. For example, let $\widehat{\mathbb{E}[AC]}$ denote the estimate of $\mathbb{E}[AC]$, then

$$\widehat{\mathbb{E}[AC]} = \frac{\sum_{j=1}^N a_j c_j}{N} \longrightarrow \mathbb{E}[AC], \quad (8.18)$$

with high probability, as the number of repetitions, N , increases [28].

Now, consider quantum mechanics into account. As it is suggested in Fig. 8.3, Charlie sends the left qubit to Alice and the right qubit to Bob. The random variables A, B, C , and D are defined by the results of measurements of the following observables:

$$A = Z_1 \quad C = \frac{-Z_2 - X_2}{\sqrt{2}} \quad (8.19)$$

$$B = X_1 \quad D = \frac{Z_2 - X_2}{\sqrt{2}} \quad (8.20)$$

where subscripts 1, 2 stand for the left and right qubits sent by Charlie, respectively. The calculation of the expected values is straightforward, for example,

$$\mathbb{E}[AC] = \langle \psi | \left(Z_1 \otimes \frac{-Z_2 - X_2}{\sqrt{2}} \right) | \psi \rangle, \quad (8.21)$$

and similar to BC, BD , and AD . But these calculations turn out to:

$$\mathbb{E}[AC] = \mathbb{E}[BC] = \mathbb{E}[BD] = -\mathbb{E}[AD] = \frac{1}{\sqrt{2}}. \quad (8.22)$$

But, with these values, the sum adds up to:

$$\mathbb{E}[AC] + \mathbb{E}[BC] + \mathbb{E}[BD] - \mathbb{E}[AD] = 2\sqrt{2}. \quad (8.23)$$

The result obtained in (8.23) shows a clear violation to the upper bound obtained in (8.17).

The conflicting results of Bell's CSHS inequality and the last result obtained from quantum mechanics vide (8.23) only can be solved via experimental procedures. Several such procedures were performed starting in the decades of 1960 and 1970. One of the most important was the work of Aspect et al. [3] that used two-photon atomic transitions in the setup. The results corroborated the predictions of quantum mechanics.

What the prior description shows is that for entangled states it is viable to find a pair of observables correlated in such a way their correlations violate the Bell's inequality. The meaning is that quantum mechanics produces statistical predictions that cannot be explained if one assumes the *Einstein locality*, that is, assuming that the results of experiments performed in a location are independent of another, discretionary one performed in another distant location, simultaneously.

8.1.1 Functional Consistency

Due to the complexity in demonstrating mentioned violations of Bell's CHSH inequality, new ways for demonstrated nonlocality were proposed [27]. Given a set of commuting observables A, B, C, \dots and a set of quantum states $|\psi\rangle, |\phi\rangle, \dots$, then it is viable measuring the observables simultaneously and to obtain the joint distribution of the values of the observables chosen from that set. Consider an ensemble of identically prepared systems, in the state, say, $|\phi\rangle$, and suppose these states are described by observables A, B, C, \dots . Each measure shall assign numerical values for each observable, $v(A), v(B), v(C), \dots$. Quantum rules require that $v(A)$ is dependent only on the operator A , not on the state $|\phi\rangle$, and require also that in a commuting set of observables the only allowed results of simultaneous measurement are in the set of simultaneous eigenvalues.

From requirements considered [27], it is possible to notice that for any particular functional identity

$$f(A, B, C, \dots) = 0, \quad (8.24)$$

fulfilled by the commuting observables, should be fulfilled by the set of eigenvalues, that is

$$f(v(A), v(B), v(C), \dots) = 0. \quad (8.25)$$

For example, if A and B commute, then

$$C = A + B \xrightarrow{([A,B]=0)} v(C) = v(A) + v(B), \quad (8.26)$$

or, equivalently,

$$C - A - B = 0 \stackrel{([A,B]=0)}{\implies} v(C) - v(A) - v(B) = 0. \quad (8.27)$$

But, implications (8.26) or (8.27) are valid only if A and B commute! Because if it is not this way, A 's and B 's eigenvalues are different and they cannot be simultaneously measured. There is no evidence supporting those identities. However, in the sense of mean, (8.27) holds ever, that is, for any quantum state $|\phi\rangle$, it is true that

$$\langle \phi | A + B | \phi \rangle = \langle \phi | A | \phi \rangle + \langle \phi | B | \phi \rangle, \quad (8.28)$$

even despite the fact that A and B do not commute. Historically, famous mistakes happened probably motivated by this caveat [36].

We have seen that one important meaning of the Bell's CHSH inequalities is that for an entangled quantum state it is viable to find pairs of observables such that quantum mechanics statistics predictions are incompatible with the requirement of *locality* (also referred to as *Einstein locality*). Equivalently, this means that the results of measurements made at a given place are not all independent of those obtained at a remote lab.

An $n \times n$ matrix M is said to be diagonalizable if and only if the sum of the dimensions of the its eigenspaces equals n or, equivalently, if and only if M is similar to a diagonal matrix, that is, there exist an invertible matrix P such that

$$P^{-1}MP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad (8.29)$$

where the scalars λ_i ($i = 1, 2, \dots, n$) are the eigenvalues of M . From

$$MP = P \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = (\lambda_1 |\alpha_1\rangle \lambda_2 |\alpha_2\rangle \dots \lambda_n |\alpha_n\rangle), \quad (8.30)$$

where P can be written as

$$P = (|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle). \quad (8.31)$$

In this representation $|\alpha_i\rangle$ ($i = 1, 2, \dots, n$) and α_i ($i = 1, 2, \dots, n$) are the eigenvectors and eigenvalues of M , respectively.

If for a matrix M there is only one basis in which it is diagonal, that basis corresponds to a *maximal quantum test* which is equivalent to a measurement of

the observable represented by M . If $[M, N] = 0$ (M and N commute), it is possible to find a basis in which both matrices are diagonal. This basis corresponds to a maximal test, which provides a measurement of both M and N . Therefore two commuting operators can be simultaneously measured, otherwise they are said to be *incompatible*.

Generalization is straightforward. A set of matrices are said to be simultaneously diagonalizable if there exists a single invertible matrix P such that $P^{-1}MP$ is a diagonal matrix for every M in the set. A set of diagonalizable matrices commutes if and only if the set is simultaneously diagonalizable. A set of commuting operators is said to be *complete* if there exists a single basis in which all these operators are diagonal.

8.1.2 Context

If, regardless of the previously mentioned ambiguity, we insist in assuming that the measurement of an operator M depends uniquely on the objective properties of the measured quantum system, then we are assuming validity of contextualization of the setup to determine the measurement results completely. For example, if $[M, N] = 0$ and also, $[M, V] = 0$, we can jointly measure M and N , or jointly M and V , then we wait that the result of the measurement of M does not depend on its *context*, specifically, whether we measure only M , or M and N , etc. Notice that the assumption made is clearly counterfactual, that is, it cannot be put under an experimental setup.

Considered jointly, *contextuality* and *functional consistency* are indeed incompatible with the predictions of quantum theory, in spite of their “reasonability.” The following example is particularly illustrative of this point.

Example 8.2 (Peres [29]). Consider a pair of qubits, not necessarily entangled (singlets), and the operators displayed in the following matrix

$$\mathcal{A} = \begin{pmatrix} \mathbb{1} \otimes Z & Z \otimes \mathbb{1} & Z \otimes Z \\ X \otimes \mathbb{1} & \mathbb{1} \otimes X & X \otimes X \\ X \otimes Z & Z \otimes X & Y \otimes Y \end{pmatrix}. \quad (8.32)$$

It is possible to verify that:

- Each operator has eigenvalue ± 1 ;
- In each row the three operators commute;
- In each column the three operators commute;
- Each operator is product of the two others, with exception of the third column, that requires a minus sign.

On the other hand, consider the sequences

$$(X \otimes Z)(Z \otimes X) = (XZ) \otimes (ZX) \quad (8.33)$$

$$= (-iY) \otimes (iY) \quad (8.34)$$

$$= Y \otimes Y, \quad (8.35)$$

and

$$(Z \otimes Z)(X \otimes X) = (ZX) \otimes (XZ) \quad (8.36)$$

$$= (-iY) \otimes (-iY) \quad (8.37)$$

$$= -Y \otimes Y. \quad (8.38)$$

Due to the opposite signs in (8.35) and (8.38), we cannot assign ± 1 values to entries of matrix \mathcal{A} in such a way that those came out of measuring the operators form \mathcal{A} .

8.1.3 Gleason's Theorem

This important theorem states that for Hilbert spaces of dimension at least three, the only possible probability measures are that of the form

$$\langle A \rangle = \text{Tr}(\rho A), \quad (8.39)$$

where ρ stands for a prepared quantum state and A is an observable. This means that there is not observable other than (8.39). The proof of that theorem is recognized as difficult.²

One of the issues raised by Gleason is that assuming only the primitives:

- Decision tests (only yes/no answers allowed) are represented by projectors in a Hilbert space;
- Compatible tests simultaneously correspond to commuting projectors;
- If P and Q are orthogonal projectors, then their sum, $S = P + Q$, which is itself a projector, obeys

$$\langle S \rangle = \langle P \rangle + \langle Q \rangle, \quad (8.40)$$

then, for Hilbert spaces with dimension larger than 2, (8.39) is the only that gives the corrected statistics for the measurements. The main remark is that the projector S , $\text{Tr} S = 2$, can be decomposed in unlimited number of manners. For instance, take

$$P_1 = |\alpha\rangle \langle \alpha|, \quad (8.41)$$

$$P_2 = |\beta\rangle \langle \beta|, \quad (8.42)$$

²See an interesting “geometry oriented” discussion in [6, p. 151].

projectors onto orthonormal vectors $|\alpha\rangle$ and $|\beta\rangle$, respectively. Consider the next “rotations”:

$$|a\rangle = \frac{|\alpha\rangle + |\beta\rangle}{\sqrt{2}}, \quad (8.43)$$

$$|b\rangle = \frac{|\alpha\rangle - |\beta\rangle}{\sqrt{2}}. \quad (8.44)$$

Clearly $|a\rangle$ and $|b\rangle$ are also orthonormal. The projectors onto these last vectors are

$$Q_1 = |a\rangle \langle a|, \quad (8.45)$$

$$Q_2 = |b\rangle \langle b|. \quad (8.46)$$

The last two decompositions satisfy

$$Q_1 + Q_2 = |a\rangle \langle a| + |b\rangle \langle b| \quad (8.47)$$

$$= \frac{(|\alpha\rangle + |\beta\rangle)(\langle\alpha| + \langle\beta|)}{2} + \frac{(|\alpha\rangle - |\beta\rangle)(\langle\alpha| - \langle\beta|)}{2} \quad (8.48)$$

$$= P_1 + P_2. \quad (8.49)$$

The identity $Q_1 + Q_2 = P_1 + P_2$ is considered as trivial, but, in contrast, the similar statement about the averages is not. Such statement regarding the averages can be formally written as:

$$\langle P_1 \rangle + \langle P_1 \rangle = \langle Q_1 \rangle + \langle Q_2 \rangle. \quad (8.50)$$

Considering such non-trivial nature, it deserves to be experimentally verified [29].

8.1.4 The Kochen-Specker Theorem

Mermin [27] introduced the following reasoning. Take observables with eigenvalues 1 or 0 with corresponding spin components $0, \pm 1$. The sums of the squared spin components along any three orthogonal axis, x, y, z obey

$$S_x^2 + S_y^2 + S_z^2 = s(s+1) = 2. \quad (8.51)$$

This is valid for particles with spin $s = 1$. Additionally, the squared components S_x^2, S_y^2 , and S_z^2 form a mutually commuting set. The results of measurement are 0 or 1 for each direction, x, y , or z , additionally, that results must fulfill condition (8.51).

Assume that a *set of directions* with many orthogonal trials is given, in conjunction with the set of observables (squared spins components) alongside that

directions. As the observables along orthogonal axes, the squared components are mutually commuting and due to constraint (8.51), the measured values equals one of the three triads $(1, 1, 0)$, $(1, 0, 1)$ or $(0, 1, 1)$.

The contextuality, that is, the impossibility of description supported on *local hidden variables*, can be proved revealing quantum states for which the statistics of their respective observables (S_x^2, S_y^2, S_z^2) , connected with the orthogonal axes, cannot be obtained by any assignments of 1s and 0s to every direction in the set, such that condition (8.51) holds.

The Kochen-Specker theorem exhibits a set of vectors, called *Kochen-Specker (KS) sets* such that it is impossible to assign 1's (associated with a color red) and 0's (associated with a color blue) and condition (8.51) is kept.

Notice that no statistics relative of the states is necessary for justification. This exclusion of statistics from the problem is similar to the change from the classical information theory, where asymptotically small probability of error is admitted to the zero-error information theory where no error is admitted. In the first, the ordinary information theory, probability measures are essential. For the last, zero-error information theory, graph theory and combinatorics are the main tools for analysis.

From the last discussion, it is possible to give a concise statement of the KS theorem in terms of the following problem.

Problem 8.1 (Mermin [27]). Determine a set of directions (vectors) in a 3-dimensional space such that it is impossible to assign a color (red or blue) to each direction under the condition that every subset of three mutually orthogonal vectors contains exactly one blue and two red vectors.

For the sake of completeness, the solution (proof) given by Mermin [27] is sketched here. Firstly, notice that only directions are essential, one is free to modify the size of the vectors. Without loss of generality assume the unit vector \mathbf{z} , blue, defining this axis. Take the red vector \mathbf{a} living in the y - z plane:

$$\mathbf{a} = \mathbf{z} + \alpha\mathbf{y}, \quad 0 < \alpha < 0.5. \quad (8.52)$$

Then consider the following remarks:

- As \mathbf{z} is blue, \mathbf{x} and \mathbf{y} are both red. Indeed, any vector in the x - y must be red, due to the condition that one cannot have two orthogonal blue vectors, that is

$$\mathbf{c} = \beta\mathbf{x} + \mathbf{y}, \quad \text{for any } \beta. \quad (8.53)$$

must be red.

- Additionally, since \mathbf{a} and \mathbf{x} are red, any vector in their plane must be red. To the proof, we shall soon verify that an interesting red vector in this plane is

$$\mathbf{d} = \frac{1}{\beta}\mathbf{x} - \frac{1}{\alpha}\mathbf{a}. \quad (8.54)$$

- Notice that because $\mathbf{a} = \mathbf{z} + \alpha\mathbf{y}$, then d is orthogonal to $\mathbf{c} = \beta\mathbf{x} + \mathbf{y}$. To see this, it is enough to perform the scalar product (denoted by “ \cdot ”):

$$\mathbf{d} \cdot \mathbf{c} = \left(\frac{1}{\beta}\mathbf{x} - \frac{1}{\alpha}\mathbf{a} \right) \cdot (\beta\mathbf{x} + \mathbf{y}) \quad (8.55)$$

$$= \|\mathbf{x}\|^2 - \frac{\beta}{\alpha}(\mathbf{z} + \alpha\mathbf{y}) \cdot \mathbf{x} - \frac{1}{\alpha}(\mathbf{z} + \alpha\mathbf{y}) \cdot \mathbf{y} \quad (8.56)$$

$$= \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 \quad (8.57)$$

$$= 0. \quad (8.58)$$

- Recall that \mathbf{c} and \mathbf{d} are both red, so the normal to their plane must be blue. Therefore, any vector in their plane must be red. So, the following sum is red:

$$\mathbf{e} = \mathbf{d} + \mathbf{c}. \quad (8.59)$$

- Since $\alpha \in (0, 0.5)$, then $\frac{1}{\alpha} > 2$, and, for $\beta \in \mathbb{R}$,

$$\left| \beta + \frac{1}{\beta} \right| \in (2, \infty), \quad (8.60)$$

it is viable determined a value of β such that \mathbf{e} is parallel to

$$\mathbf{f} = \mathbf{x} - \mathbf{z}. \quad (8.61)$$

Also, changing the signal of β results in another \mathbf{e} parallel to

$$\mathbf{g} = -\mathbf{x} - \mathbf{z}. \quad (8.62)$$

- Since \mathbf{e} is red independently of the β value, both \mathbf{f} and \mathbf{g} must be red.
- But, $\mathbf{f} \cdot \mathbf{g} = 0$, they are orthogonal; so the normal to their plane is blue and any vector in their plane is surely red.
- Notice that

$$\mathbf{z} = -\frac{1}{2}\mathbf{f} - \frac{1}{2}\mathbf{g} \quad (8.63)$$

lives in the \mathbf{f} , \mathbf{g} plane, but \mathbf{z} is blue. This is the contradiction sought.

A chain of simpler and simpler proofs of quantum contextuality has been introduced since the KS theorem has appeared in the literature. One of these recent simplifications was introduced by Cabello et al. [9]. Experimental apparatus has also been explored in this same kind of sequence of simplifications [16].

An interesting approach concerning the plethora of nonlocality proofs is given in van Dam et al. [35], where the authors, rooted on Kullback-Lieber distance, propose a measure for the amount of evidence provided by the experimental setup.

In the following sections, recent results related with quantum zero-error information theory are examined.

8.2 Quantum Chromatic Number and Kochen-Specker Sets

Throughout this section we shall use definitions from graph theory introduced earlier in Chap. 4. Eventually, for the sake of easiness, some of those may be redundant.

Scarpa and Severini [32] and Mancinska et al. [26] introduced conditions for equality and strict inequality between three quantities associated with a graph G :

- The minimum dimension of orthogonal representation, denoted by $\xi(G)$;
- The quantum chromatic number, denoted by $\chi_q(G)$;
- The *ordinary* quantum chromatic number, denoted by $\chi(G)$.

One remarkable contribution introduced in the mentioned works is the outstanding role performed by the KS sets in the proofs. These sets are collections of vectors with applications to investigations about the calculation of quantum zero-error capacities of quantum channels.

The *quantum chromatic number* is a remarkable parameter for at least one reason: it is a tool for differentiating aspects of quantum and classical behavior, in particular, for entanglement-assisted communications. Also the quantum chromatic eases comprehension of combinatorial parameters as, e.g., the Lovász theta function and the minimum dimension of an orthogonal representation of a simple graph.

In this section only simple graphs (unweighted, undirected graphs without self-loops) are considered and, as before, for a graph G , $V(G)$ and $E(G)$ denote its vertex and edge sets, respectively.

Before introducing the relationship between the concepts of quantum chromatic number, KS sets, and their consequences for the zero-error information theory, however, due to its importance, we shall review main concepts related to the Kochen-Specker theorem, following mainly the reading given in Peres [29].

A proper k -coloring of a graph is an assignment of k colors such that every two adjacent vertices have different colors. The chromatic number of a graph G , $\chi(G)$, is the minimum number of colors k needed to build a proper k -coloring map of G .

We now introduce a coloring game for a graph $G = (V, E)$. We consider that Alice and Bob claim that they have a proper k -coloring for G and a referee tests this claim with a one-round game. The rules of the game do not allow communication between the players. The referee asks Alice the color, say a , for the vertex v and Bob for the color, say b , for the vertex w . Alice and Bob win the game if, for $a, b \in \{1, \dots, k\}$:

- IF $v = w$, THEN $a = b$;
- IF $(v, w) \in E$, THEN, $a \neq b$.

A classical strategy is formed by two deterministic functions:

$$g_A : V \rightarrow \{1, \dots, k\}, \quad (8.64)$$

$$g_B : V \rightarrow \{1, \dots, k\}. \quad (8.65)$$

It is clear that, independently of the strategy chosen by the players, including even probabilistic strategies, they cannot win the game with probability 1 if $k < \chi(G)$, that is, using less than the chromatic number of colors in their assumed proper k -coloring procedure.

A quantum strategy the players can take advantage of is a convenient entangled state $|\psi\rangle$ living in a Hilbert space of dimension d and two collections of POVMs in the following way:

- For all $v \in V$, Alice owns $\{\mathbf{E}_{va}\}_{a=1,\dots,k}$ and similarly Bob owns $\{\mathbf{F}_{vb}\}_{b=1,\dots,k}$;
- Alice applies her POVM $\{\mathbf{E}_{va}\}_{a=1,\dots,k}$ to her part of the entangled state and get the value a ;
- Bob applies his POVM $\{\mathbf{F}_{vb}\}_{b=1,\dots,k}$ to his part of the entangled state and get the value b .

In order to have consistent conditions, for a quantum strategy, the rules are adapted in such a way that Alice and Bob win the game, if only if

$$\forall v \in V, \forall a \neq b, \langle \psi | \mathbf{E}_{va} \otimes \mathbf{F}_{vb} | \psi \rangle = 0, \quad (8.66)$$

$$\forall (v, w) \in E, \forall a, \langle \psi | \mathbf{E}_{va} \otimes \mathbf{F}_{wa} | \psi \rangle = 0. \quad (8.67)$$

If these conditions are attained, the strategy is said to be a *winning strategy*. Notice that only the number of measurement operator is fixed, neither the dimension of the entangled stated nor the rank of the measurement operator is taken into consideration. This motivates the following definition.

Definition 8.1 (Quantum Chromatic Number [32]). For all graphs G , the quantum chromatic number $\chi_q(G)$ is the minimum number k such that there exists a quantum k -coloring of G .

In the following \overline{W} stands for the complex conjugate of the complex matrix W , that is, the (i, j) entry of \overline{W} is obtained from the corresponding (i, j) entry of W taking its complex conjugate. The *Hilbert-Schmidt product* of two complex matrices W and V is given, as usual, by

$$\langle W | V \rangle = \text{Tr } W V^\dagger. \quad (8.68)$$

The normal form of a k -coloring emphasizes the simplicity of its structure.

Proposition 8.1 (Normality [32]). *If G has a quantum k -coloring, then there exists a quantum k -coloring of G in normal form, with the following properties:*

1. All POVMs are projective measurements with k projectors of rank r ;
2. The state $|\psi\rangle$ is the maximally entangled of local dimension rk ;
3. For all pairs v, a , the projectors of Alice and Bob are conjugate, that is, $\mathbf{E}_{v,a} = \overline{\mathbf{F}}_{v,a}$;
4. The consistency conditions can be represented as

$$\forall (v, w) \in E, \quad \forall a \in \{1, \dots, k\}, \quad \langle \mathbf{E}_{va}, \mathbf{E}_{wa} \rangle = 0. \quad (8.69)$$

The proof for this proposition was introduced by Scarpa and Severini [32]. The authors emphasize that the quantum chromatic number depends on the rank of the POVM elements adopted by Alice and Bob. This remark motivates the following definition.

Definition 8.2 (Rank- r Quantum Chromatic Number [32]). The rank- r quantum chromatic number $\chi_q^{(r)}(G)$ of G is the minimum number of colors k such that G has quantum k -coloring formed by projectors of rank r and a maximally entangled state of local dimension rk .

It can be observed that $\chi_q^{(r)}(G) \leq \chi_q^{(s)}(G)$ if $r \geq s$ [32]. For rank-1 quantum coloring, the dimension of the maximally entangled state equals k and, the rank-1 projectors for each vertice v , can be represented as the outer product

$$|e_{va}\rangle \langle e_{va}|, \quad a \in \{1, \dots, k\}, \quad (8.70)$$

for an orthonormal basis $\{|e_{va}\rangle_{a \in \{1, \dots, k\}}\}$. Therefore (8.69) can be rewritten as:

$$\forall (v, w) \in E(G), \quad \forall a \in \{1, \dots, k\}, \quad \langle e_{va}| e_{wa} \rangle = 0. \quad (8.71)$$

If a quantum k -coloring of a graph $G(V, E)$ is given, then a *matrix representation* of G can be constructed with the map:

$$\Phi : V \rightarrow \mathbb{C}^{k \times k}, \quad (8.72)$$

such that for all $(v, w) \in E$ it is required

$$\text{diag}(\Phi(v)^\dagger \Phi(w)) = 0. \quad (8.73)$$

Here, $\text{diag}(A)$ stands for the vector formed with the A diagonal entries. The map Φ is built taking for all $v \in V$ a unitary matrix U_v mapping the computational basis $\{|i\rangle_{i \in \{1, \dots, k\}}\}$ to $\{|e_{va}\rangle, a \in \{1, \dots, k\}\}$. Notice that U_v is a $k \times k$ matrix and (8.71) condition assures that if (v, w) is an edge then the diagonal entries of $U_v^\dagger U_w$ are zero.

8.2.1 Relationship Between $\xi(G)$, $\chi_q^{(1)}(G)$ and $\chi(G)$

A k -dimensional *orthogonal representation* of $G = (V, E)$ is a map

$$\phi : V \rightarrow \mathbb{C}^k, \quad (8.74)$$

such that for all $(v, w) \in E$, the inner product $\langle \phi(v) | \phi(w) \rangle = 0$. The *orthogonal rank* of a graph G , denoted by $\xi(G)$, is defined as the minimum k such that an orthogonal representation of G in \mathbb{C}^k exists.

For all graphs G the following inequalities hold [10]:

$$\xi(G) \leq \chi_q^{(1)}(G) \leq \chi(G). \quad (8.75)$$

If we have two graphs G and H , we can now define the *Cartesian product* GH .

- The vertex set $V(GH) = V(G) \times V(H)$ is the Cartesian product of the vertex sets of G and H ;
- An edge between vertices $(a_1, b_1), (a_2, b_2) \in V(GH)$ is either
 - $a_1 = a_2$ and $(b_1, b_2) \in E(H)$, or
 - $(a_1, a_2) \in E(G)$ and $b_1 = b_2$.

Notice that a vertex in $V(GH)$ corresponds to a pair (a, b) of vertices where a is from G and b is from H .

Let K_k be a complete graph with k vertices. The next proposition clarifies the relation between the minimum dimension of orthogonal representation and the quantum chromatic number.

Proposition 8.2 (Scarpa and Severini [32]). *For all graphs G :*

$$\chi_q^{(1)} = \min\{k : \xi(GK_c) = k\}. \quad (8.76)$$

From this proposition follows a condition for equality between the rank-1 quantum chromatic number and orthogonal rank of a graph G .

Theorem 8.1 (Scarpa and Severini [32]). *For all graphs G :*

$$\chi_q^{(1)}(G) = \xi(G) \text{ if only if } \xi(GK_{\xi(G)}) = \chi(G). \quad (8.77)$$

8.3 Wielandt's Inequality

We have defined a quantum channel, denoted \mathcal{E} , as a *trace preserving completely positive linear map* (TPCP), that is,

$$\mathcal{E} : \mathcal{M}_{D \times D} \rightarrow \mathcal{M}_{D \times D}, \quad (8.78)$$

where $\mathcal{M}_{D \times D}$ is the space of the complex $D \times D$ matrices. The Kraus operators $W_k \in \{\mathcal{M}_{D \times D}\}_{k=1}^d$ are a versatile representation of quantum channel \mathcal{E}_W :

$$\mathcal{E}_W(\rho) = \sum_{k=1}^d W_k \rho W_k^\dagger. \quad (8.79)$$

Unitary operations, some kind of measurements, addition of uncorrelated quantum systems, substituting of a input state by other state are some operations in a quantum channel that can be well represented by Kraus operators.

Another useful representation, that allows modeling any operation by a unitary operation on a larger Hilbert space, is the Stinespring theorem:

Theorem 8.2 (Stinespring Theorem). *Let \mathcal{E} be a trace-preserving quantum operation on a Hilbert space \mathcal{H} . Then there is an ancilla space \mathcal{K} of dimension $\dim \mathcal{K} \leq (\dim \mathcal{H})^2$ so that for any fixed $|\chi\rangle \in \mathcal{K}$ there is a unitary transformation \hat{U} on $\mathcal{H} \otimes \mathcal{K}$ with*

$$\mathcal{E}(\rho) = \text{Tr}_{\mathcal{K}}\{\hat{U}(\rho \otimes |\chi\rangle\langle\chi|)\hat{U}^\dagger\}. \quad (8.80)$$

However, for quantum channels the representation by means of Kraus operators is actually more usual, at least for applications where discrete classical information are to be transmitted. In this section we discuss some relevant development of the notion of zero-error communications through quantum channels.

As usual we begin with a classical concept to later introduce extensions of that concept into the quantum framework. Recall the definition of a classical discrete memoryless channel (DMC) $(\mathcal{X} \times \mathcal{Y}, W(Y|X))$ for which $|\mathcal{X}| = |\mathcal{Y}| = D$. The main elements of a DMC are shown in Fig. 8.4. The matrix $W(Y|X)$ is defined as a stochastic matrix whose rows are indexed by the elements of \mathcal{X} while the columns are indexed by those of the finite set \mathcal{Y} . The (x, y) entry of $W(Y|X)$ is the probability that $Y = y$ is received when $X = x$ is transmitted.

Source messages V are picked from a finite set (alphabet of the source) and are mapped by means of the encoder in codewords $X^n = (X_1, X_2, \dots, X_n)$. Each X_i , $i = 1, 2, \dots, n$, is transmitted through the memoryless channel that produces an output Y_i in such a way that

$$\Pr(Y^n = y^n | X^n = x^n) = \prod_{i=1}^n \Pr(Y_i = y_i | X_i = x_i). \quad (8.81)$$

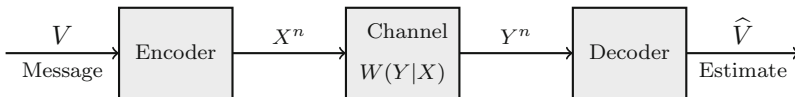


Fig. 8.4 Classical communication systems

The right-hand product means the i -th output y_i of a DMC depends by a stochastic map only on the i -th input x_i . This is the motivation to the term *memoryless* in the definition.

If we take into account the input and output blocks, X^n and Y^n , respectively, it is natural to define a stochastic matrix for these blocks, say,

$$W^{(n)}(Y^n|X^n) \triangleq \Pr(Y^n|X^n) = \underbrace{W \otimes W \otimes \dots \otimes W}_{n \text{ times}}. \quad (8.82)$$

Notice that the probability distributions $W(\cdot|x)$ and $W^{(n)}(\cdot|x^n)$ correspond to the x -th and x^n -th lines of the product matrix, respectively. If two input blocks X^n and \tilde{X}^n can lead to the same output block Y^n , decoding cannot be performed without error. We say that X^n and \tilde{X}^n are *confusable* or *indistinguishable*.

Sanz et al. [31] proposed an extension of the classical Wielandt's inequality to quantum channels. The concept concerns the number of applications of the channel to any source (probability distribution) for which any output will be reached.

Before introducing the formal definition for the quantum case, we will recall some important notions. A matrix is said to be *positive* if its entries are all strictly positive.³ This means that for a positive DMC matrix W , any output $y \in \mathcal{Y}$ can be reached from any input $x \in \mathcal{X}$ at the input.

Definition 8.3 (Primitive Matrix). A square stochastic matrix W is said to be primitive if there is an $m \in \mathbb{N}$ such that $(W^m)_{ij} > 0$ for all (i, j) , that is if W^m is positive. The minimum m for which this occurs, denoted $p(W)$, is said to be the classical index of primitivity of W .

It means that if a DMC probability transition matrix W is primitive, then using $p(W)$ times a DMC $(\mathcal{X} \times \mathcal{Y}, W(Y|X))$, any $y \in \mathcal{Y}$ can be reached from each input $x \in \mathcal{X}$ transmitted or, equivalently, for the product channel displayed in Fig. 8.4, all input blocks X^n are confusable.

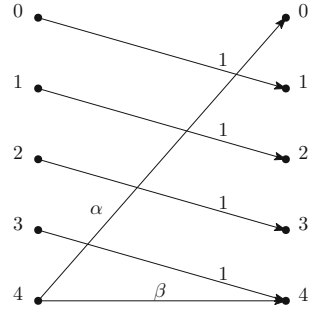
The *Wielandt's inequality* [22, p. 520] states that, for every primitive matrix W , then:

$$p(W) \leq D^2 - 2D + 2. \quad (8.83)$$

Observe that the Wielandt's inequality does not depend on the matrix elements; only its primitivity is required. There are applications of the Wielandt's inequality for several fields, e.g., to graph theory, number theory, numerical analysis, etc. An extension of the concept of index of primitivity to the quantum framework was introduced by Sanz et al. [31] which is defined by the number of times a channel must be used, so that it maps any density operator to one with full rank.

³We call the attention that there is no connection of definition of *positive matrix* with the definition of *positive definite matrix*.

Fig. 8.5 Graph of the DMC induced by the stochastic matrix of (8.84) for $D = 5$. Labels on the edge (x, y) stand for conditional probabilities, for instance, $\alpha = \Pr[Y = 0|X = 4]$



Example 8.3. The following $D \times D$ matrix

$$W = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ \alpha & \beta & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \tag{8.84}$$

is primitive, for $\alpha > 0$ and $\beta > 0, \alpha + \beta = 1$. The primitivity is due to the fact that W^m is positive for $m = D^2 - 2D + 2$. This is just the Wielandt bound [28, p. 730]. For instance, fixing $D = 5$, we have $p(W) = m = 17$, which means that 17 uses of that DMC, every output block $Y^{17} \in \mathcal{Y}^{17}$ can be reached from any $X^{17} \in \mathcal{X}^{17}$ input block of symbols.

The graph of transitions for the DMC induced by this matrix ($D = 5$) is displayed in Fig. 8.5.

Consider the probability row vector \mathbf{p} of the input X as:

$$\mathbf{p} = (p_0, p_1, p_2, p_3, p_4). \tag{8.85}$$

That is, $p_x = \Pr[X = x], x \in \mathcal{X}$. Similarly denote

$$\mathbf{q} = (q_0, q_1, q_2, q_3, q_4), \tag{8.86}$$

the probability row vector of the output Y , that is, $p_y = \Pr[Y = y], y \in \mathcal{Y}$. The transition channel matrix W defines the relationship between probability vectors \mathbf{p} and \mathbf{q} , as follows:

$$\mathbf{q} = \mathbf{p}W. \tag{8.87}$$

8.3.1 Quantum Index of Primitivity

Let \mathcal{E}_W be the quantum channel defined by Kraus operators $\{W_k \in \mathcal{M}_{D \times D}\}_{k=1}^d$, that is

$$\mathcal{E}_W(\rho) = \sum_{k=1}^d W_k \rho W_k^\dagger. \quad (8.88)$$

The quantum index of primitivity, denoted by q , is defined to a quantum channel (TPCP map) by the least $m \in \mathbb{N}$ such that m uses of the channel assures that every positive semidefinite operator is mapped onto a positive definite operator, where D is the dimension of the Hilbert space, and d the number of linearly independent Kraus operators.

It is possible to show that

$$q \leq (D^2 - d + 1)D^2, \quad (8.89)$$

where D is the dimension of the Hilbert space, and d the number of linearly independent Kraus operators.

Our proposal now is to focus on the issues of the inequality (8.89) connected with the notion of quantum channels with positive zero-error capacity. In this way, we need to recall some preliminaries given in [31]. Firstly, the authors define $S_n(W) \in \mathcal{M}_{D \times D}$ as the linear space spanned by all possible products of *exactly* n Kraus operators, $W_{k_1} W_{k_2} \dots W_{k_n}$ and denote $W_k^{(n)}$ the elements of $S_n(W)$, with this, they define

$$H_n(W, \varphi) \triangleq S_n(W) |\varphi\rangle \subseteq \mathbb{C}^D, \quad (8.90)$$

as the space spanned by all vectors $W_{k_1} W_{k_2} \dots W_{k_n} |\varphi\rangle$, where $|\varphi\rangle \in \mathbb{C}$. Secondly, Sanz et al. [31] recall the one-to-one correspondence between a quantum channel \mathcal{E} and its *Choi matrix*

$$\omega(\mathcal{E}) \triangleq (\mathbb{1} \otimes \mathcal{E})(\Omega), \quad (8.91)$$

where $\Omega = \sum_{i,j}^D |ii\rangle \langle jj|$. Then, the observed rank $[\mathcal{E}_W^n(|\varphi\rangle \langle \varphi|)] = \dim[H_n(W, \varphi)]$.

Equipped with prior discussion, three properties are introduced.

1. **Primitive Quantum Channel.** A quantum channel \mathcal{E}_W is said to be *primitive* if there exists some $n \in \mathbb{N}$ such that for all $|\phi\rangle \in \mathbb{C}^D$, $H_n(W, \varphi) = \mathbb{C}^D$. The number $q(\mathcal{E}_W)$ stands for the minimum n for which the condition is reached. This means that for every input density operator ρ the output $\mathcal{E}_W^n(\rho)$, obtained after n applications of the channel is full-rank. It is observed that if \mathcal{E}_W is primitive, then for every $m \in \mathbb{N}$, \mathcal{E}_W^m is also primitive and we have

$$H_n(W, \varphi) = \mathbb{C}^D \text{ for all } n \geq q(\mathcal{E}_W). \quad (8.92)$$

2. **Eventually Full Kraus Rank Quantum Channel.** A quantum channel \mathcal{E}_W is called *eventually full Kraus rank* if there exists some $n \in \mathbb{N}$ such that $S_n(W) = \mathcal{M}_{D \times D}$. This means that $\text{rank} [{}_W(\mathcal{E}_W^n)] = D^2$. The number $i(W)$ stands for the minimum n for which that condition is satisfied. Notice that if \mathcal{E}_W fulfills this property, then $S_n(W) = \mathcal{M}_{D \times D}$ for all $n \geq i(W)$.
3. **Strongly Irreducible Quantum Channel.** A quantum channel \mathcal{E}_W is said to be *strongly irreducible* if the following two conditions are fulfilled:
- \mathcal{E}_W has a unique eigenvalue, λ , with $|\lambda| = 1$;
 - The corresponding eigenvector ρ is a positive definite operator ($\rho > 0$).

An important question now is how the classical Wielandt bound relates with the quantum one. The main tool is to make an embedding of the classical channel in the quantum framework, as it is shown in the next example.

Example 8.4. Consider again the classical DMC illustrated in Fig. 8.5 (for $D = 5$) and respective stochastic matrix W given by (8.84). It is easy to see that embedding is obtained by intermediate the following map \mathcal{E}_W defined by the Kraus operators

$$W_{x,y} = \sqrt{w_{x,y}} |y\rangle \langle x|, \quad x \in \mathcal{X}, y \in \mathcal{Y}. \quad (8.93)$$

For an input (diagonal) operator

$$\rho = \delta_{x,y} p_x = \begin{pmatrix} p_0 & 0 & 0 & 0 & 0 \\ 0 & p_1 & 0 & 0 & 0 \\ 0 & 0 & p_2 & 0 & 0 \\ 0 & 0 & 0 & p_3 & 0 \\ 0 & 0 & 0 & 0 & p_4 \end{pmatrix}. \quad (8.94)$$

Here, $\delta_{x,y}$ stands for the Kronecker function and p_x , $x \in \{0, 1, 2, 3, 4\}$, are the entries of input probability vector \mathbf{p} (recall Example 8.3). The output is given by

$$\sigma = \mathcal{E}_W(\rho) \delta_{x,y} q_y = \begin{pmatrix} q_0 & 0 & 0 & 0 & 0 \\ 0 & q_1 & 0 & 0 & 0 \\ 0 & 0 & q_2 & 0 & 0 \\ 0 & 0 & 0 & q_3 & 0 \\ 0 & 0 & 0 & 0 & q_4 \end{pmatrix}, \quad (8.95)$$

where q_y , $y \in \{0, 1, 2, 3, 4\}$, are the entries of the output vector \mathbf{q} .

The set of Kraus operators are

$$\{ |1\rangle \langle 0|, |2\rangle \langle 1|, |3\rangle \langle 2|, |4\rangle \langle 3|, \sqrt{\alpha} |0\rangle \langle 4|, \sqrt{\beta} |4\rangle \langle 4| \}. \quad (8.96)$$

Assume W is a primitive DMC stochastic probability transition matrix with primitivity index $p(W)$. The following is proved in [31].

Proposition 8.3. *Let W be a primitive stochastic map and let \mathcal{E}_W be the corresponding TPCP. The channel \mathcal{E}_W is also primitive and*

$$q(W) = p(W) = i(W). \quad (8.97)$$

Notice that equality of (8.97) holds for quantum channels resulting of the above DMC embedding. This fact is illustrated next.

Example 8.5. Let us consider a “genuine” quantum channel, with $D = 2$ and $d = 3$ given by the Kraus operators

$$\left\{ W_1 = \frac{1}{\sqrt{3}}X, W_2 = \frac{1}{\sqrt{3}}Y, W_3 = \frac{1}{\sqrt{3}}Z \right\} \quad (8.98)$$

where X , Y , and Z are the Pauli matrices. Explicitly, the map is given by

$$\mathcal{E}_W(\rho) = \frac{1}{3} \left(X\rho X^\dagger + Y\rho Y^\dagger + Z\rho Z^\dagger \right), \quad (8.99)$$

for an input state ρ .

It is straightforward to check that in this case $q(\mathcal{E}_W) = 1$ and $i(W) = 2$. The quantum Wielandt’s bound is $8 = (D^2 - d + 1)D^2 \geq q(\mathcal{E}_W)$.

There is an open question if the quantum Wielandt’s bound is sharp, however the following theorem is very important because establishing a *universal dichotomy* behavior of the zero-error capacity of the important class of unital quantum channels. This dichotomy result is universal in the sense that it depends only on the dimension of the Hilbert space, D , and not on the channel itself.

Theorem 8.3 (Dichotomy Behavior [31]). *Let $C^{(0)}(\mathcal{E})$ be the zero-error classical capacity of the quantum channel \mathcal{E} . If \mathcal{E} is a quantum channel with full-rank fixed point, then either $C^{(0)}(\mathcal{E}^n) \geq 1$ for all n or $C^{(0)}(\mathcal{E}^{q(\mathcal{E})}) = 0$.*

Notice that if \mathcal{E}^n stands for the input–output relationship after n units of time or space then the theorem reveals the existence of a *universal critical length* $n = q(\mathcal{E})$ such that once a transmission is successfully viable then a successful transmission $m \geq n$ is possible.

8.4 Entanglement-Assisted Zero-Error Capacity

In Sect. 3.2.5 we saw different capacities of quantum channels. The entanglement-assisted capacity in particular considers that the parties share an unrestricted amount of previously shared entanglement which they can use in order to maximize the information changed through the quantum channel. In this section, we describe how entanglement can be used in a zero-error scenario for exchanging classical information.

Before introducing the entanglement-assisted zero-error communication, we need some background concepts. We start with the hypergraph of a quantum channel.

Definition 8.4 (Hypergraph of a Quantum Channel [13]). Let \mathcal{E} be a quantum channel. The hypergraph of \mathcal{E} , denoted by $\mathbf{H}(\mathcal{E})$, is a set of vertices, denoted by S and a set of subsets of S . The set of vertices S is composed by the channel input. There is one hyperedge for each of the outputs, which contains all the inputs that have a nonzero probability of causing that output.

We also use the concept of clique in hypergraphs. A clique of $\mathbf{H}(\mathcal{E})$ is a set κ_i of possible inputs of a given output in a confusability graph. In other words, the clique κ_i contains all the inputs that can cause the same output.

In this current scenario, in particular, prior to information exchange through the quantum channel, Alice and Bob share a d -dimensional entangled state $\rho_{AB} = |\Phi_{AB}\rangle\langle\Phi_{AB}|$ given by

$$|\Phi_{AB}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i_A\rangle |i_B\rangle \quad (8.100)$$

Considering that such pre-shared entanglement is available, Cubitt et al. [13] proposed a protocol for entanglement-assisted zero-error communication that is described as follows:

1. Alice chooses a message $m \in \{1, \dots, K\}$ from a set of messages, where K is the number of messages. Alice wants to send the message m to Bob;
2. Alice measures her half of the entangled system using a complete orthogonal basis, say $B_m = \{|\psi_{\hat{x}}\rangle\}$, where \hat{x} is a vertice in a clique κ_m from the hypergraph $\mathbf{H}(\mathcal{E})$;
3. Alice sends the result of her measurement to Bob.

Some clarifications are needed before proceeding. In the hypergraph $\mathbf{H}(\mathcal{E})$ the vertice x represents the unit vector $|\psi_x\rangle$ such that if x and \hat{x} are adjacent then $\langle\psi_x|\psi_{\hat{x}}\rangle = 0$. Recalling that K is the size of the messages set, the hypergraph has K cliques of size d , say $\{\kappa_1, \dots, \kappa_K\}$. It is analogous to say that each message m has a d -size clique κ_m in the hypergraph $\mathbf{H}(\mathcal{E})$.

4. After Alice's measurement, Bob's state will collapse to $|\psi_x\rangle^*$;
5. Bob will measure his state in $B_m = \{|\psi_x\rangle\}$ in order to get the final state $|\psi_{\hat{x}}\rangle^*$;
6. Bob output is denoted by y . His possible states are determined by those vertices x , for which $p(y|\hat{x}) > 0$ and these adjacent states are mutually orthogonal, i.e., for any \hat{x}_1 and \hat{x}_2 , then $\langle\hat{x}_1|\hat{x}_2\rangle = 0$ [23].

For short, to send K messages using entanglement, Alice and Bob can use a maximally entangled state of rank d : to send m , Alice measures her side of the state in the bases B_m and obtains the outcome j (at random). She inputs $(m;j)$ to the channel. Bob's output tells him that Alice's input was in some particular mutually confusable subset, but by construction, these inputs correspond to mutually orthogonal residual

states of his subsystem, so he can perform a projective measurement to determine precisely which input Alice made to the classical channel, and hence which of the K messages she chose to send, with certainty [13].

Using the previously defined elements and protocol characterized, we can now characterize the entanglement-assisted zero-error classical capacity of quantum channels.

Theorem 8.4 (Entanglement-Assisted Zero-Error Capacity [13]). *Let \mathcal{E} be a quantum channel. The entanglement-assisted zero-error capacity of \mathcal{E} , denoted by $C_E^{(0)}(\mathcal{E})$, is given by*

$$C_E^{(0)}(\mathcal{E}) = \lim_{n \rightarrow \infty} \frac{1}{n} \log K^E(\mathcal{E}^{\otimes n}) \geq \log(K^E) \quad (8.101)$$

where K^E is the number of mutually non-adjacent input messages with entanglement assistance.

Theorem 8.5 (Cubitt-Leung-Matthews-Winter Theorem [13] via [23]). *For a quantum channel \mathcal{E} with hypergraph $\mathbf{H}(\mathcal{E})$, there exists an entanglement-assisted quantum communication protocol that can send one of K messages with zero error; hence for entanglement-assisted asymptotic classical zero error capacity*

$$\begin{aligned} \log(K) \leq C^{(0)}(\mathcal{E}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log(K(\mathcal{E}^{\otimes n})) < C_E^{(0)}(\mathcal{E}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log K^E(\mathcal{E}^{\otimes n}) \geq \log(K^E). \end{aligned} \quad (8.102)$$

This theorem shows us that entanglement can sometimes be used to increase the number of classical messages which can be sent perfectly over quantum channels [13].

Some results in the literature have interesting connections with the entanglement-assisted zero-error capacity. Leung et al. [25], using certain input codewords (based on a Pauli graph), show that entanglement can help to increase the classical zero-error capacity to the maximum achievable HSW capacity.

In general, it is possible to observe the following relation between the classical zero-error quantities:

$$C^{(0)}(\mathcal{E}) < C_E^{(0)}(\mathcal{E}) \leq C_{1,\infty}(\mathcal{E}) \quad (8.103)$$

Recalling that the zero-error capacity is a quantity hard to compute even for small characteristic graphs, upper bounds play an important role, in particular, the Lovász ϑ function is commonly considered. Beigi [4] verified that the Lovász ϑ function is an upper bound on the zero-error capacity even in the presence of entanglement between the sender and receiver.

8.5 Non-Commutative Graphs and Quantum Lovász ϑ Function

The works of Duan, Cubitt, Severini, Winter, and other collaborators [14, 15, 17] introduce the theory of non-commutative graphs in the study of quantum zero-error capacity problem. Starting with the Kraus form of representation of a quantum channel, the authors define a generalization of the classical adjacency graph called *non-commutative (confusability) graph* by the operator space:

$$S \triangleq \text{span}\{W_j^\dagger W_k : j, k\} < \mathcal{L}(\mathcal{H}) \quad (8.104)$$

where $\mathcal{L}(\mathcal{H})$ stands for the set of observable on Hilbert space \mathcal{H} .

According to this graph definition, a zero-error code consists in the anti-clique of the corresponding graph. The biggest anti-clique, called *independence number* and denoted by $\alpha(G)$, corresponds to the maximum number of messages that can be transmitted through the channel with probability of error equal to zero. This way, the classical zero-error capacity of a channel with graph G is given by

$$C^{(0)}(G) = \lim_{n \rightarrow \infty} \log \alpha(G^n) = \sup_n \log \alpha(G^n). \quad (8.105)$$

Translating this concept to the quantum scenario, we have a quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_X) \rightarrow \mathcal{B}(\mathcal{H}_Y)$, where $\mathcal{B}(\cdot)$ is the space of linear operators from a given Hilbert space. This way, the event $E_{x,y} : \mathcal{H}_X \rightarrow \mathcal{H}_Y$, that corresponds to the input of a quantum state $x \in \mathcal{X}$ and the output of a quantum state $y \in \mathcal{Y}$ in this quantum channel, is given by

$$E_{x,y} = \sqrt{p(y|x)} |y\rangle \langle x|. \quad (8.106)$$

This way, we can define the *confusability graph of a quantum channel* (or non-commutative graph) as being a subspace

$$S = \text{span} \left\{ E_{x',y'}^\dagger \cdot E_{x,y} \neq 0; x, x' \in \mathcal{X}, y, y' \in \mathcal{Y} \right\}. \quad (8.107)$$

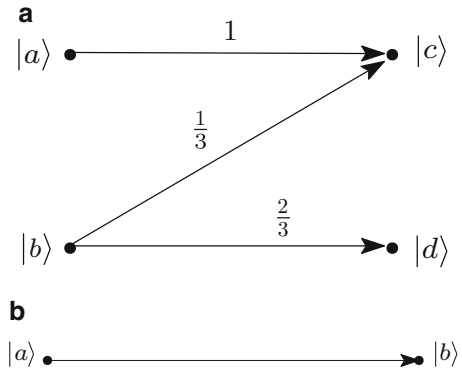
It is interesting to notice that such definition emphasizes the channel's input that can be confused, while in Chap. 5 the approach was to emphasize inputs that are not adjacent. Despite this difference, both definitions are equivalent.

Example 8.6 (Confusability Graph of a Quantum Channel). Let \mathcal{E} be a quantum channel shown in Fig. 8.6a. The input alphabet contains the symbols $\mathcal{X} = \{a, b\}$ and the output alphabet contains the symbols $\mathcal{Y} = \{c, d\}$.

According to (8.106), we have the following events:

$$E_{a,c} = 1 \cdot |c\rangle \langle a| = |c\rangle \langle a|, \quad (8.108)$$

Fig. 8.6 Example of a graph of a quantum channel (a) transitions probabilities at the channel's end (b) confusability graph



$$E_{a,d} = 0 \cdot |d\rangle \langle a| = 0, \tag{8.109}$$

$$E_{b,c} = \sqrt{\frac{1}{3}} |c\rangle \langle b|, \tag{8.110}$$

$$E_{b,d} = \sqrt{\frac{2}{3}} |d\rangle \langle b|. \tag{8.111}$$

It is important to emphasize that these events have a straight correspondence with the quantum channel it is related. From these events, we can consider the following elements that will compose the confusability graph (see (8.107))

$$E_{a,c}^\dagger \cdot E_{a,c} = |a\rangle \langle a|, \tag{8.112}$$

$$E_{a,c}^\dagger \cdot E_{b,c} = \sqrt{\frac{1}{3}} |a\rangle \langle b|, \tag{8.113}$$

$$E_{b,c}^\dagger \cdot E_{a,c} = \sqrt{\frac{1}{3}} |b\rangle \langle a|, \tag{8.114}$$

$$E_{b,c}^\dagger \cdot E_{b,c} = \frac{1}{3} |b\rangle \langle b|, \tag{8.115}$$

$$E_{b,d}^\dagger \cdot E_{b,d} = \frac{2}{3} |b\rangle \langle b|. \tag{8.116}$$

Thus,

$$S = \text{span} \left\{ E_{a,c}^\dagger \cdot E_{a,c}, E_{a,c}^\dagger \cdot E_{b,c}, E_{b,c}^\dagger \cdot E_{a,c}, E_{b,c}^\dagger \cdot E_{b,c}, E_{b,d}^\dagger \cdot E_{b,d} \right\}. \tag{8.117}$$

Considering the subspace S we can denote the zero-error capacity of a quantum channel. This capacity is given by the following expression, considering that the biggest set of self-orthogonal states is given by $\{|\phi_m\rangle : m = 1, \dots, N\}$

$$\forall m \neq m' : |\phi_m\rangle \langle \phi_{m'}| \in S^\perp, \tag{8.118}$$

Table 8.2 Alternative definitions for zero-error capacities considering independence numbers

Capacity	Expression	Observations
Classical zero-error capacity	$C^{(0)}(S) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \alpha(S^{\otimes n})$	Shown previously in (8.105)
Quantum zero-error capacity	$Q^{(0)}(S) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \alpha_q(S^{\otimes n})$	α_q denotes the quantum independence number whose value depends on the existence of a Stienespriug dilatation in the channel.
Entanglement assisted zero-error capacity	$C_E^{(0)}(S) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\alpha}(S^{\otimes n})$	$\tilde{\alpha}$ denotes the higher integer N for which there are (i) Hilbert spaces \mathcal{H}_{X_0} and \mathcal{H}_{Y_0} ; (ii) $\omega \in S(\mathcal{H}_{X_0} \otimes \mathcal{H}_{Y_0})$; (iii) a map $\mathcal{E}_m : \mathcal{B}(\mathcal{H}_{X_0}) \rightarrow \mathcal{B}(\mathcal{H}_X)$, such that there are N states $\rho_m = (\mathcal{E} \circ \mathcal{E}_m \otimes \mathbb{1}_{Y_0})\omega$ which are mutually adjacent.
Generalized entanglement assisted zero-error capacity	$\hat{C}_E^{(0)}(S) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \hat{\alpha}(S^{\otimes n})$	$\hat{\alpha}$ denotes the higher independence number assisted by generalized entanglement which demands that $\mathcal{E}_m(\sigma) = \sum_j E_{jm} \sigma E_{jm}^\dagger$ and that $\sum_j E_{jm}^\dagger E_{jm} \in GL(\mathcal{H}_{X_0})$ to be invertible.

where S^\perp is an orthogonal subspace to S given in (8.107). The expression in (8.118) has some relations with the corresponding independence number of the graph. In Example 8.6, for instance, we have that $S^\perp = \emptyset$. It implies that the classical zero-error capacity of the quantum channel \mathcal{E} is equal to zero.

For every confusability graph $S \leq \mathcal{B}(\mathcal{H}_X)$ we have the following relation

$$\alpha_q(S) \leq \alpha(S) \leq \tilde{\alpha}_U(S) \leq \tilde{\alpha}(S) \leq \hat{\alpha}(S) \quad (8.119)$$

where each of these α s, called *independence numbers*, has relation with a different kind of zero-error capacity, as shown in Table 8.2. Detailed information about how to obtain such numbers can be found in the work of Duan et al. [17, 18].

The classical zero-error capacity $C^{(0)}$ and the quantum zero-error capacity $Q^{(0)}$ of a quantum channel were deeply discussed in Chap. 5. The zero-error capacity assisted by entanglement was introduced in the previous section.

Among the independence numbers showed, the only which is not directly related to a zero-error capacity is $\tilde{\alpha}_U(S)$ because it considers unitary restrictions in its definitions. The numbers $\alpha_q(S)$, $\alpha(S)$, $\tilde{\alpha}(S)$ are $\hat{\alpha}(S)$ computable. However, finding a computable expression to the associated zero-error capacity cannot be a trivial task.

Regarding the Lovász theta function, presented in Sect. 4.3 for the classical scenario, it works as an upper bound for the zero-error capacity of a DMC. It is natural to pursue a quantum version of such definition. It was presented by the authors and is formally defined as follows.

Definition 8.5 (Quantum Lovász Theta Number [17]). Let S be the non-commutative graph of a quantum channel \mathcal{E} . The quantum Lovász theta number is given by

$$\tilde{\vartheta}(S) = \sup_n \vartheta(S \otimes \mathcal{L}(\mathbf{C}^n)) \quad (8.120)$$

$$= \sup_n \max\{\|\mathbb{1} + T\| : T \in S^\perp \otimes \mathcal{L}(\mathbf{C}^n), \mathbb{1} + T \geq 0\}, \quad (8.121)$$

where the supremum is over all integers n , and the maximum is taken over Hermitian operators T .

The authors show more results regarding characteristics of this quantum version of the Lovász theta number, such as its monotonicity and supermultiplicativity.

Considering the independence numbers and their corresponding zero-error capacities as well as the Lovász theta number, both characterized using the same approach of non-commutative graphs, the authors believe that these results suggest that there might be a much more systematic way in which operator systems generalize Graph Theory to the non-commutative domain. They are pursuing new results in such direction.

8.6 QMA-Completeness of Quantum Clique

In this section we are going to explore the results of Beigi and Shor regarding the computational complexity of the quantum clique problem which was found out to be QMA-complete [5]. According to the authors, the original problem of finding the quantum clique can be written in terms of finding the zero-error capacity of a quantum channel. Exploring the zero-error behavior in this scenario brought relevant contributions to the theory of complexity, enlightening the classification of an important problem according to quantum complexity classes.

Theory of Complexity is a subarea of computer science whose goal is to prove for important problems that their solutions require certain minimum resources [38]. When considering each solution, it takes into account a *model of computation* (classical Turing machines, probabilistic Turing machines, quantum Turing machines, for instance) and a *certain resource* (for example, memory or time) [34].

Problems are grouped into *complexity classes*, according to the model used and to a function of the amount of resources their best solutions demand over a certain size of input in a worst-case scenario. For classical models of computation and when time resources are considered, the complexity classes \mathcal{P} and \mathcal{NP} are widely studied.

An algorithmic problem belongs to the complexity class \mathcal{P} of *polynomially solvable problems* if it can be solved by an algorithm with polynomial worst-case runtime [38]. The class \mathcal{P} is felt to capture the notion of problems with efficient time solutions considering classical (deterministic) Turing machines [2].

The complexity class \mathcal{NP} contains problems having efficiently verifiable solutions. In other words, if x is a solution to the problem (certificate), it is possible to verify that in deterministic polynomial time [2]. For example, the subset sum is in the \mathcal{NP} class. In this problem, given a list of n numbers A_1, A_2, \dots, A_n and a number T , one must decide if there is a subset of numbers that sums up to T . The certificate for this problem is the list of members in this subset. For a practical example in the problem considered, if the list of numbers is $\{4, -8, 0, 22, -17, 3, 2\}$ and $T = 1$, then the certificate $\{4, -8, 3, 2\}$ can be verified as a solution to the problem in polynomial time.

Any problem in \mathcal{P} is also in the \mathcal{NP} class because we can solve it in polynomial time even without the need of a certificate [12]. However, the question if the classes \mathcal{P} and \mathcal{NP} are equal or different remains as one of the most important challenges for computer science. Certain problems in the \mathcal{NP} class, in particular, have a special classification. A problem in \mathcal{NP} is called an \mathcal{NP} -complete problem if any efficient algorithm for it can be converted into an efficient algorithm for any other problem in \mathcal{NP} [20].

Considering this brief introduction of computational complexity, from now on we will examine more closely the contribution and the results of Beigi and Shor [5]. Some definitions and argumentation presented below in this section are from their original work and the reader is referred to it for more details.

The *clique of a graph* is a widely known \mathcal{NP} -complete problem. Given a graph G a clique is a subset of vertices, every two of which are adjacent, and the size of a clique is the number of its vertices. The clique problem is that given a graph G and an integer number k , decide whether G contains a clique of size k or not.

Let G^c be the complement of the graph G . In the complement of G a clique is changed to an *independent set*. A subset of a graph G where no two vertices are adjacent characterize an independent set. The maximum size of an independence set is the *independence number* of the graph G , denoted by $\alpha(G)$. So the clique problem in the complement graph reduces to decide whether $\alpha(G) \geq k$, and then it is \mathcal{NP} -complete. This reduction is important because the problem of computing $\alpha(G)$ is related to the problem of computing the zero-error capacity of a classical discrete memoryless channel, as shown extensively on Chap. 4.

The quantum version of the clique problem, known as *quantum clique problem*, is also to decide whether $\alpha(\mathcal{E}) \geq k$ for a given quantum channel \mathcal{E} . It is equivalent to decide whether there exists quantum states ρ_1, \dots, ρ_k such that $\mathcal{E}(\rho_1), \dots, \mathcal{E}(\rho_k)$ have orthogonal supports or not. Note that, for any two states σ_1, σ_2 , then $\text{Tr}(\sigma_1\sigma_2) \geq 0$ and equality holds iff σ_1 and σ_2 have orthogonal supports.

Let $\sigma_{1,2} = \sigma_1 \otimes \sigma_2$ then $\text{Tr}(\sigma_1\sigma_2) = \text{Tr}(S\sigma_{1,2})$, where S is the swap gate ($S|\psi\rangle|\varphi\rangle = |\varphi\rangle|\psi\rangle$). We can estimate $\text{Tr}(\sigma_1\sigma_2)$ by applying the swap gate. We must notice that if $\sigma_{1,2}$ is not separable then the equality does not hold and the orthogonality is not implied by $\text{Tr}(S\sigma_{1,2}) = 0$. To avoid this problem we must restrict ourselves to entanglement breaking channels.

Definition 8.6 (Entanglement Breaking Quantum Channel [5]). A quantum channel \mathcal{E} is called an entanglement breaking quantum channel if there are POVM

$\{M_i\}$ and states σ_i such that

$$\mathcal{E}(\rho) = \sum_i \text{Tr}(M_i \rho) \sigma_i \quad (8.122)$$

for any ρ . In this case, $\mathcal{E}^{\otimes 2}(\rho_{1,2})$ is always separable, $\text{Tr}(S\mathcal{E}^{\otimes 2}(\rho_{1,2})) \geq 0$ and equality implies $\mathcal{E}(\rho_1)$ and $\mathcal{E}(\rho_2)$ are orthogonal.

Putting all the concepts together, we can now formally define the quantum clique problem.

Definition 8.7 (Quantum Clique Problem [5]). The quantum clique problem (\mathcal{E}, k, a, b) is defined as follows:

- **Input.** Integer numbers n and k ; non-negative real numbers a and b with an inverse polynomial gap $b - a > n^{-c}$; and \mathcal{E} an entanglement breaking quantum channel that acts on n -qubit states;
- **Promise.** Either exists $\rho_1 \otimes \dots \otimes \rho_k$ such that $\sum_{i,j} \text{Tr}(S\mathcal{E}(\rho_i) \otimes \mathcal{E}(\rho_j)) \leq a$ or for any states $\rho_{1,2,\dots,k}$ we have $\sum_{i,j} \text{Tr}(S\mathcal{E}^{\otimes 2}(\rho_{i,j})) \geq b$;
- **Output.** Decide which one is the case.

Despite the deep understanding of the clique problem as an \mathcal{NP} -complete problem in the classical case, the same does not happen to the quantum clique problem prior to the work discussed here. Even nowadays, our knowledge regarding quantum complexity theory is still not rich as its classical analogue.

Now we are going to characterize the QMA quantum complexity class that, loosely speaking, is the quantum version of the \mathcal{NP} class [37]. The acronym of this complexity class stands for *Quantum-Merlin-Arthur* where Merlin is an oracle with infinite computational power and Arthur is a quantum polynomial time verifier. Merlin answers decision problems of the type “*Is x in L ?*” and accompany the answer with a polynomial certificate y which Arthur can verify in polynomial time using a quantum machine [1]. We associate two probabilities with the QMA class which are related to the completeness and soundness. The formal definition of such complexity class is presented as follows.

Definition 8.8 (QMA Complexity Class [37]). A language L is said to be in $\text{QMA}(2/3, 1/3)$ if there exists a quantum polynomial time verifier V such that

- **Completeness.** $\forall x \in L, \exists |\xi\rangle \in \mathcal{H}^{p(|x|)}, \Pr(V(|x\rangle |\xi\rangle) = 1) \geq 2/3$;
- **Soundness.** $\forall x \notin L, \forall |\xi\rangle \in \mathcal{H}^{p(|x|)}, \Pr(V(|x\rangle |\xi\rangle) = 1) \leq 1/3$;

After the QMA complexity class was characterized, Kitaev introduced a problem from Physics called the “*local Hamiltonian problem*” and showed that it is QMA-complete. This problem is the quantum analogue of the classical SAT problem and these results are the analogue of the Cook-Levin theorem [5]. The formal definition of this problem is described below.

Definition 8.9 (Local Hamiltonian Problem [24]). The k -local Hamiltonian problem (H_1, \dots, H_s, a, b) is defined as follows:

- **Input.** An integer n , real numbers a, b such that $b - a > n^{-c}$, and polynomially many Hermitian non-negative semidefinite matrices H_1, \dots, H_s with bounded norm $\|H_i\| \leq 1$, such that each of them acts just on k of n qubits;
- **Promise.** The smallest eigenvalue of $H_1 + \dots + H_s$ is either less than a or greater than b ;
- **Output.** Decide which one is the case.

Intuitively, a k -local matrix assigns a real number to any quantum state on n qubits. This number depends only on the reduced state of the k qubits where a quantum operator M acts non-trivially, and can be thought of as a locally defined penalty on a given quantum state. Loosely speaking, the k -local Hamiltonian problem asks whether there exists a quantum state that can significantly avoid a collection of such penalties [37].

Considering the quantum clique problem, the following theorem states its complexity.

Theorem 8.6 (Quantum Clique is \mathcal{QMA} -Complete [5]). *The quantum clique problem (\mathcal{E}, k, a, b) where \mathcal{E} is an entanglement breaking channel on n -qubit states and has the operator-sum representation*

$$\mathcal{E}(\rho) = \sum_{i=1}^r E_i \rho E_i^\dagger, \quad (8.123)$$

where $\sum_i E_i^\dagger E_i = \mathbb{1}$ and $r = \text{poly}(n)$, is \mathcal{QMA} -complete.

The proof of this theorem consists in showing that (\mathcal{E}, k, a, b) is \mathcal{QMA} . To prove the hardness, the authors establish a polynomial time reduction from the local Hamiltonian problem to quantum clique. In this result, a is a positive number that means that some probability of error is allowed.

If we consider the case $a = 0$, we will try to find a protocol with no error. In this case, $(\mathcal{E}, k, a = 0, b)$ exactly says that whether $\alpha(\mathcal{E}) \geq k$ or not. We can achieve this by using a zero-error quantum channel \mathcal{E} where

$$\mathcal{E}(\rho) = \sum_{i=1}^r \text{Tr}(M_i \rho) |i\rangle \langle i|, \quad (8.124)$$

where $\{M_1, \dots, M_r\}$ is a POVM and $|1\rangle, \dots, |r\rangle$ are orthogonal states. Checking orthogonality of two outcomes states is accomplished in the following way: given two states $\mathcal{E}(\rho)$ and $\mathcal{E}(\rho')$ we measure them in the basis $|1\rangle, \dots, |r\rangle$. If the outcome of the measurements are the same, then their supports are not orthogonal.

So, the quantum clique problem $(\mathcal{E}, k, a = 0, b)$ where \mathcal{E} is a zero-error quantum channel that can be implemented exactly by a polynomial time verifier is \mathcal{QMA}_1 -complete. The article of Beigi and Shor contains the entire version of the proofs briefly discussed here [5].

Some promise problems in the literature are known to be \mathcal{QMA} -complete, such as variants of the local Hamiltonian problem, the density matrix consistency problem, and also other problems about quantum circuits [37]. However, the quantum clique problem is the only so far in this complexity class whose complexity was described using zero-error quantum channels.

This section described the complexity classification of the quantum clique problem. We presented the results of Beigi and Shor that showed a non-trivial result where zero-error quantum channels helped in determining the quantum complexity of a problem. The contributions of these authors enrich the knowledge regarding quantum complexity classes and the classification of an important quantum problem, which may have implications in algorithms and protocols for practical applications.

8.7 Further Reading

In this section we could see some developments in the literature that provide new results and insights into the quantum zero-error information theory. We saw the background which relates quantum zero-error information theory with Kochen-Specker sets and Bell's inequality. The quantum version of the Wielandt's inequality which states an upper bound to the number of uses of a quantum channel in order to map an arbitrary density operator to a full rank operator was also discussed. An alternative version of the zero-error capacity of quantum channels considering entanglement assistance was introduced. The approach for zero-error capacity considering non-commutative graphs and the quantum counterpart of the Lovász theta function was also considered. Lastly, an application of quantum zero-error channels to find the complexity class of a problem showed a non-trivial application of the concepts discussed along the book.

Other recent results besides those discussed here can also be found in the literature. Blume-Kohout et al. [7] developed a framework to handle quantum information that can be perfectly preserved (i.e, with zero-error) by the system dynamics. According to the authors, the system dynamics affects the kind of information that can be carried or store (classical, quantum or neither, for instance). Taking that into account, the main purpose of their operational framework is to describe how to perfectly preserve information despite the system dynamics. This framework considers not only quantum channels with positive zero-error capacity, but also quantum error-correcting codes, decoherence-free subspaces and subsystems and even other methods proposed by the own authors, such as the unconditionally preserved codes. This work provides an exhaustive classification of ways that information can be preserved.

Regarding practical implementations, Gyongyosi and Imre [21] considered the use of multiple optical channels to send information. Each of these individual channels has no positive zero-error capacity, but when used jointly the zero-error capacity is superactivated. Their idea is to adopt such strategy as part of the implementation of quantum repeaters, devices that can extend the range of quantum communication between sender and receiver.

Besides the already described results on superactivation of zero-error capacity, Shirokov [33] showed a special kind of superactivation of quantum channels under block coding.

Brït et al. considered the use of quantum entanglement in the zero-error source-channel coding problem [8]. In their scenario, Alice and Bob are each given an input from a random source and get access to a noisy channel through which Alice can send messages to Bob. Their goal is to minimize the average number of channel uses per source input such that Bob can learn Alice's inputs with zero probability of error. Their results show lower bound and optimum rate of entanglement-assisted source codes and the advantage that entanglement can give in the source-channel coding problem.

We hope that much more results on quantum zero-error information theory are yet to come.

References

1. Aharonov D, Naveh T (2002) Quantum NP – a survey. <http://arxiv.org/abs/quant-ph/0210077>. Accessed 02 Feb 2016
2. Arora S, Barak B (2009) Computational complexity: a modern approach. Cambridge University Press, Cambridge
3. Aspect A, Grangier P, Roger G (1982) Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: a new violation of Bell's inequalities. *Phys Rev Lett* 49(2):91–94
4. Beigi S (2010) Entanglement-assisted zero-error capacity is upper bounded by the Lovász theta function. *Phys Rev A* 82:010303
5. Beigi S, Shor PW (2008) On the complexity of computing zero-error and Holevo capacity of quantum channels. <http://arxiv.org/abs/0709.2090>. Accessed 25 Oct 2014
6. Bengtsson I, Życzkowski K (2006) Geometry of quantum states. Cambridge University Press, Cambridge
7. Blume-Kohout R, Ng HK, Poulin D, Viola L (2010) Information-preserving structures: A general framework for quantum zero-error information. *Phys Rev A* 82:062306
8. Briet JB, Buhman H, Laurent M, Piovesan T, Scarpa G (2015) Entanglement-assisted zero-error source-channel coding. *IEEE Trans Inf Theory* 61(2):1124–1138
9. Cabello A, Badziąg P, Cunha MT, Bourennane M (2013) Simple Hardy-like proof of quantum contextuality. *Phys Rev Lett* 111:180404
10. Cameron PJ, Montanaro A, Newman MW, Severini S, Winter A (2007) On the quantum chromatic number of a graph. *Electron J Comb* 14(1):1–15
11. Clauser J, Horne MA, Shimony A, Holt RA (1969) Proposed experiment to test local hidden-variable theories. *Phys Rev Lett* 41:880
12. Cormen TH, Leiserson CE, Rivest RL, Stein C (2009) Introduction to algorithms, 3rd edn. MIT, Cambridge
13. Cubitt T, Leung D, Matthews W, Winter A (2010) Improving zero-error classical communication with entanglement. *Phys Rev Lett* 104:230503
14. Cubitt TS, Leung D, Matthews W, Winter A (2010) Improving zero-error classical communication with entanglement. *Phys Rev Lett* 104
15. Cubitt TS, Leung D, Matthews W, Winter A (2011) Zero-error channel capacity and simulation assisted by non-local correlations. *IEEE Trans Inf Theory* 57(8):5509–5523
16. D'Ambrosio V, Herbauts I, Anselm E, Nagah E (2013) Experimental implementation of a Kochen-Sepecker set of quantum tests. *Phys Rev X* 3:011012

17. Duan R, Severini S, Winter A (2011) Zero-error communication via quantum channels, non-commutative graphs and a quantum Lovasz ϑ function. In: IEEE international symposium on information theory, Russia, pp 64–68
18. Duan R, Severini S, Winter A (2013) Zero-error communication via quantum channels, non-commutative graphs and a quantum Lovasz ϑ function. *IEEE Trans Inf Theory* 59(2):1164–1174
19. Einstein A, Podolsky B, Rosen N (1935) Can quantum-mechanical description of physical reality be considered complete? *Phys Rev* 47:777–780
20. Goldreich O (2010) P, NP, and NP completeness. Cambridge University Press, Cambridge
21. Gyongyosi L, Imre S (2012) Long-distance quantum communications with superactivated gaussian optical quantum channels. *Opt Eng* 51(1):1–16
22. Horn RA, Johnson CR (1985) Matrix analysis. Cambridge University Press, Cambridge
23. Imre S, Gyongyosi L (2012) Advanced quantum communications: an engineering approach, 1st edn. Wiley, New York
24. Kitaev A, Shen A, Vyalyi MN (2002) Classical and quantum computation. American Mathematical Society, Rhode Island
25. Leung D, Mancinska L, Matthews W, Ozols M, Roy A (2012) Entanglement can increase asymptotic rates of zero-error classical communication over classical channels. *Commun Math Phys* 311:97–111
26. Mancinska L, Scarpa G, Severini S (2013) A generalization of Kochen-Specker sets relates quantum coloring to entanglement-assisted channel capacity. *IEEE Trans Inf Theory* 59(6):4025–4032
27. Mermin ND (1993) Hidden variables and the two theorems of John Bell. *Rev Mod Phys* 65(3):803–815
28. Papoulis A, Pillai SU (2002) Probability, random variables and stochastic processes. McGraw-Hill, New York
29. Peres A (1995) Quantum theory: concepts and methods. Kluwer Academic Publishers, Dordrecht
30. Petri CA (1996) Nets, time and space. *Theor Comput Sci* 3:3–48
31. Sanz M, Pérez-García D, Woff MM, Cirac JI (2010) A quantum version of Wielandt’s inequality. *IEEE Trans Inf Theory* 56(9):4668–4673
32. Scarpa G, Severini S (2012) Kochen-Specker sets and the rank-1 quantum chromatic number. *IEEE Trans Inf Theory* 58(4):2524–2529
33. Shirokov ME (2015) On quantum zero-error capacity. *Commun Mosc Math Soc* 70(1):176–178
34. Sipser M (2012) Introduction to the theory of computation, 3rd edn. Course Technology, Boston
35. van Dam W, Gill RD, Grünwald PD (2005) The statistical strength of nonlocality proofs. *IEEE Trans Inf Theory* 51(8):2812–2835
36. von Neumann J (1955) Mathematical foundations of quantum mechanics. Princeton University Press, Princeton
37. Watrous J (2008) Quantum computational complexity. <http://arxiv.org/abs/0804.3401>. Accessed 02 Feb 2016
38. Wegener I (2005) Complexity theory – exploring the limits of efficient algorithms. Springer, Dortmund
39. Young HD (1964) Fundamentals of mechanics and heat, 1st edn. McGraw-Hill Book Company, Inc., New York