Quantum Communication & Networks

Joseph M. Renes
Institute for Theoretical Physics, ETH Zürich, 8093 Zürich, Switzerland

Notes for the ICTP Advanced School on Ubiquitous Quantum Physics: The New Quantum Revolution.

1 Overview/Syllabus

The goal of these notes is to provide an overview of quantum error-correction and quantum cryptography. For the former we will address

- What is a quantum error-correcting code (QEC)?
- How do they work?
- What communication rates can they achieve?

We will also touch on how quantum error-correcting codes can be used for reliable quantum computation (the question of fault-tolerance). For the latter we will focus on the task of quantum key distribution (QKD) and address

- What is a QKD protocol?
- How does it work?
- How do we know a QKD protocol is secure?

These two topics are in fact quite closely related by the uncertainty principle, and examining two useful uncertainty relations will be our route to understanding QEC and QKD.

These topics are in the intersection of information theory and quantum physics, so first let us get more acquainted with the setting of information theory by examining the famous teleportation protocol.

2 The setting of information theory: Teleportation

Actually, the setting of information theory is not so foreign to physics; it is very similar to thermodynamics. In both the goal is to understand the possibilities and limitations of machines we would like to build, though in information theory we usually talk about "protocols" instead of "machines". For thermodynamics the machines have to do with performing useful work or regulating temperature and so forth, whereas in information theory we want to transmit, store, or manipulate data of some kind. But the broad point is the same:

1. We have some ideal behavior in mind, say keeping the inside of the refrigerator at 4 degrees Celsius,
2. We have some available resources for doing so, say electricity from the grid and an ambient heat bath,
3. Then we try to construct a machine that will enable us to simulate the ideal behavior by using the resources.

Note that the ideal behavior is thought of a little un-physically in the example: It's just the idealized situation in which the inside of the refrigerator is at the given temperature. Also note that the available resources actually also include the material we use to build the machine itself.

We can write this symbolically as a *resource inequality*,

$$\text{given resources} \geq \text{ideal behavior} \tag{1}$$

by which we mean there is a specific means of using the given resources, a protocol, to simulate the ideal behavior.

Teleportation can be thought of in a similar way. The goal is to transmit a quantum system, a qubit say, from one laboratory to another. Usually we call the transmitter Alice and the receiver Bob. There are two available resources in the above sense:

1. classical communication between the labs, and
2. shared entanglement.

That is, the labs can communicate using normal non-quantum means, such as the internet, and they also share maximally entangled biparte quantum systems. Defining the following state of two qubits

$$|\Phi\rangle_{AB} = \tfrac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B), \tag{2}$$

Alice has qubit A in her lab and Bob has qubit B in his lab. These qubits could be realized in a myriad of different ways, but perhaps the simplest example is to think of them as spin-$1/2$ particles such as electrons, where $|0\rangle$ corresponds to spin up along the $\hat{z}$ axis and $|1\rangle$ to spin down.

The teleportation protocol is then a means by which Alice and Bob can use the entanglement and classical communication to transfer one qubit from Alice to Bob. Denoting the shared entanglement symbolically by $[qq]$, the ability to transmit a single classical bit by $[c \to c]$, and the ability to transmit a single qubit by $[q \to q]$, teleportation is a protocol which achieves the following resource inequality:

$$[qq] + 2[c \to c] \geq [q \to q]. \tag{3}$$

That is, teleportation will consume a single entangled pair $|\Phi\rangle_{AB}$ and require transmitting two classical bits in order to transmit a single qubit.

Before stating the precise protocol, we need a few more definitions. First, call the Pauli x and z matrices X and Z, respectively. In the $|0\rangle/|1\rangle$ basis we've already used, in Dirac notation we have

$$X = |1\rangle\langle 0| + |0\rangle\langle 1|, \qquad Z = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{4}$$

In more familiar matrix notation, with $|0\rangle = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $|1\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, we have

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5}$$

Now define the *Bell states*

$$|\Phi_{jk}\rangle_{AB} = \mathbb{1}_A \otimes X_B^j Z_B^k |\Phi\rangle_{AB}, \tag{6}$$

for j and k in $\{0, 1\}$. Here the subscripts are just labels to indicate which qubit the operator acts on or which qubits are part of a given quantum state, while the superscripts are usual powers. Thus, the operator XZ is proportional to the Pauli y operator, while $X^2 = Z^2 = \mathbb{1}$. It's easy to work out the explicit form of the Bell states, which shows that they form an orthonormal basis:

$$|\Phi_{00}\rangle_{AB} = \tfrac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B), \tag{7}$$

$$|\Phi_{01}\rangle_{AB} = \tfrac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B - |1\rangle_A \otimes |1\rangle_B), \tag{8}$$

$$|\Phi_{10}\rangle_{AB} = \tfrac{1}{\sqrt{2}}(|0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |0\rangle_B), \tag{9}$$

$$|\Phi_{11}\rangle_{AB} = \tfrac{1}{\sqrt{2}}(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B). \tag{10}$$

Since the Bell states form an orthonormal basis, according to the rules of quantum mechanics it is possible to make a measurement that leaves any input quantum state in a particular Bell state indicated by the measurment outcome (the projection postulate). In the spin language, the measurement is actually detecting whether the two spins are in a joint state of total angular momentum zero ($|\Phi_{11}\rangle$, the spin-singlet) or a magnetic angular momentum zero state of total angular momentum 1, aligned along the $\hat{x}$, $\hat{y}$, or $\hat{z}$ directions ($|\Phi_{01}\rangle$, $|\Phi_{00}\rangle$, and $|\Phi_{10}\rangle$, respectively). More usefully, the Bell states are the joint eigenstates of two simple commuting operators, $X_A \otimes X_B$ and $Z_A \otimes Z_B$. These describe rotations of both spin-1/2 particles by π around the x or z axes, respectively. In particular, $|\Phi_{jk}\rangle$ has eigenvalue $(-1)^k$ of $X_A \otimes X_B$ and $(-1)^j$ of $Z_A \otimes Z_B$.

Now we can define the teleportation protocol. The input to the protocol is a qubit that Alice wishes to transfer to Bob. Call this qubit A'. Then the protocol consists of three steps:

1. Alice jointly measures qubits A and A' in the Bell basis, obtaining outcomes $j \in \{0,1\}$ and $k \in \{0,1\}$.
2. Alice communicates j and k to Bob using the classical channel.
3. Bob applies the Pauli operator product $X^j Z^k$ to qubit B.

Note that in the protocol Alice does not do anything which depends on the particular input state: No matter what it is, she just measures in the Bell basis. Thus, the protocol is designed to accept an arbitrary input.

Here's a brief sketch of why it works. Calling the input qubit state $|\psi\rangle_{A'}$, the joint state of all qubits before step 1 is $|\Psi\rangle_{A'AB} = |\psi\rangle_{A'} \otimes |\Phi\rangle_{AB}$. According to the projection postulate, outcome j,k will result in the unnormalized state $\Pi^{(j,k)}_{A'A}|\Psi\rangle_{A'AB}$, where $\Pi^{(j,k)}_{A'A} = |\Phi_{jk}\rangle\langle\Phi_{jk}|_{A'A}$, and the normalization will give the probability of outcome j,k. (Note that the Bell states appearing in the initial state and in the projector pertain to different systems! Qubits A and B are involved in the former, qubits A' and A in the latter.) Bob will apply the corresponding product of Pauli operators, so the final, still-unnormalized state will be $\left(\Pi^{(j,k)}_{A'A} \otimes (X^j Z^k)_B\right)|\Psi\rangle_{A'AB}$. Then a direct calculation of all cases shows that, for all j,k,

$$\left(\Pi^{(j,k)}_{A'A} \otimes (X^j Z^k)_B\right)|\Psi\rangle_{A'AB} = \frac{1}{2}|\Phi_{jk}\rangle_{A'A} \otimes |\psi\rangle_B \,. \tag{11}$$

Thus, the probability of any particular outcome is uniformly 1/4.

This gives us a look at the way we think about quantum information processing and the types of calculations that we commonly do. The major difference to the setting encountered in research is that this case was "exact" and we are ultimately more interested in the "approximate" case. For instance, here the classical communication is noiseless and the initial entangled state is precisely $|\Phi\rangle$. Moreover, the protocol *perfectly* transfers the qubit from Alice to Bob. But of course we can ask what happens when the entanglement isn't quite maximally entangled. In this case we can only hope to approximately simulate the ideal behavior of ideally transmitting the qubit. To treat this setting we require suitable mathematical measures of approximation, such as trace distance or fidelity, but we will not go into those here.

3 Uncertainty relations

3.1 Setup: Guessing games

One of the nice aspects of the study of quantum information is that in trying to obtain tight bounds on various kinds of protocols, we are forced to sharpen existing results from quantum mechanics, so we learn a bit more about quantum mechanics itself in the process. One example of this is improved versions of the adiabatic theorem that have come out of studying adiabatic quantum computation. Another example is given by various uncertainty relations.

Recall the uncertainty *principle*, which is a vague statement along the lines of

> Complementary physical properties, like position and momentum, cannot be simultaneously known precisely.

The job of uncertainty *relations* is make a precise statement in this direction. (And there can be many such useful statements; I don't think there's necessarily one uncertainty relation to rule them all.) Observe that, since they don't commute, X and Z are complementary in the same sense as position and momentum.

Uncertainty relations are mathematical statements, and in quantum mechanics we have to be especially careful to make sure that mathematical quantities we use are meaningful in some way. One way to do this is to try to formulate a statement that has a direct operational meaning in that it says a particular process is constrained in some way. This distinction is a bit like the various forms of the second law. One formulation, due to Carathéodory, pertains more to the mathematical formulation of the theory: "In every neighborhood of any state S of an adiabatically enclosed system there are states inaccessible from S." The notion of adiabatic accessibility is not so immediate. But the Kelvin-Planck formulation is more direct: "It is impossible to devise a cyclically operating device, the sole effect of which is to absorb energy in the form of heat from a single thermal reservoir and to deliver an equivalent amount of work."

We can make two concrete uncertainty relations that are more in the Kelvin-Planck vein and which will illuminate why quantum error correction and key distribution are possible. Instead of the task of delivering some amount of work, consider the following two guessing games played by Alice and Bob. The overarching goal of both games is for Bob to prepare a quantum system in a state such that he can predict the outcome of a measurement made by Alice on that system. Here is version 1:

1. Bob prepares a qubit A in any manner of his choosing and delivers it to Alice.
2. Bob announces his guess of an X measurement on A and his guess of a Z measurement on A.
3. Alice randomly chooses X or Z and performs the corresponding measurement on A.
4. They compare the outcome with the guess.

According to the uncertainty principle, it should be impossible to always win the game. Mathematically, we anticipate this because there is no joint eigenvector of X and Z. One can prove that Bob's probability of guessing X, $P_{\text{guess}}(X)_\rho$ and his probability of guessing Z, $P_{\text{guess}}(Z)_\rho$ cannot both be 1. In fact, we can compute the precise region of feasible guessing probabilities.

Version 2 is only slightly different:

1. Bob prepares a qubit A in any manner of his choosing and delivers it to Alice.
2. Alice randomly chooses X or Z and tells Bob her choice.
3. Bob announces his guess of the measurement.
4. Alice performs the measurement on A.
5. They compare the outcome with the guess.

Now, however, it is possible to win the game with certainty! In step 1, Bob should prepare $|\Phi\rangle_{AB}$ and keep B for himself. In step 3, he makes the same measurement on B as Alice has announced in step 2. Clearly, due to the form of the state, if Alice chooses Z, then Bob's Z meausurement on B will be the same as Alice's. But note that, for the X eigenstates $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle$, it holds that

$$|\Phi\rangle_{AB} = \tfrac{1}{\sqrt{2}}(|+\rangle_A \otimes |+\rangle_B + |-\rangle_A \otimes |-\rangle_B). \tag{12}$$

Therefore, if Alice chooses X, then his measurement result will *also* be identical to hers.

In this way he can circumvent the straightforward reading of the uncertainty principle; the difference between version 1 and version 2 hinges on what we mean by "simultaneously". Version 1 corresponds to the more straightforward reading, since we directly demand both pieces of information from Bob. In version 2, though, he has to be ready to guess either, which makes it seem that he would just need to have concrete guesses for both. But not quite. What he has instead is a "quantum guess" in the form of qubit B, since it will tell him either result (but not both at the same time). And in fact it is only possible to certainly win the game by using a "quantum guess": If both guessing probabilities are 1, then the initial state Bob starts with must be maximally entangled. We might wonder if Bob can win a variant of version 1 in which he is allowed to produce a quantum guess for Alice, but alas the answer is no.

3.2 Entropic uncertainty relations

We can try to capture the limitations and possibilities of both versions of the game in an uncertainty relation. We could hope to show that there is no strategy to always win version 1 by appealing to the Heisenberg-Robertson relation relating the variances of the observables in a state $|\psi\rangle$ to the expectation of their commutator.

$$(\Delta X)_\psi (\Delta Z)_\psi \geq \tfrac{1}{2} \left| \langle [X, Z] \rangle_\psi \right| . \tag{13}$$

In this case, however, the bound is trivial. Since the operators X and Z anticommute ($XZ + ZX = 0$), the righthand side reduces to $|\langle XZ \rangle_\psi|$. Choosing $|\psi\rangle = |0\rangle$ immediately yields zero.

Fortunately, there exist uncertainty relations for which the bound on the righthand side is state-independent. In particular, the *entropic uncertainty relation* of Maassen and Uffink. To state it, we first define the entropy. The measurement of an observable, say Z, generates outcomes distributed according to a probability distribution $P_Z(z)$ specified by the quantum state $|\psi\rangle$ via the Born rule: $P_Z(z) = |\langle z|\psi\rangle|^2$, with $z \in \{0, 1\}$. More generally, for a quantum state described by a density operator ρ, the Born rule is $P_Z(z) = \mathrm{Tr}[|z\rangle\langle z|\rho]$. The Shannon entropy of the distribution is defined by $H(Z)_P = -\sum_{z \in \{0,1\}} P_Z(z) \log_2 P_Z(z)$ and is a measure of the uncertainty in the value of Z. Since P_Z comes from ρ, we'll denote the entropy by $H(Z)_\rho$. Then we have the following entropic uncertainty relation, which states that for all qubit states ρ,

$$H(X)_\rho + H(Z)_\rho \geq 1 . \tag{14}$$

We can extend the uncertainty relation to the conditional entropy, which measures the uncertainty of a random variable when we have access to some other piece of information. The result even holds when conditioning on quantum information! Let us set aside the question of what it means to condition on quantum information for a moment and state the definition.

In order to do so, we need to define "classical-quantum states", which allow us to treat classical and quantum information on the same footing. Note that any probability distribution like P_Z above can be turned into a valid density operator ρ by setting $\rho = \sum_z P_Z(z)|z\rangle\langle z|$. In this way, (classical) probability theory is a special case of quantum theory, the case where all density operators commute. Correlations between random variables and quantum states can then be captured by bipartite density operators in which one part is diagonal. Thus, if we are interested in a situation in which quantum state $|\psi_x\rangle$ occurs with probability p_x, we can describe both the quantum state and the random variable by the density operator

$$\rho_{XB} = \sum_x p_x |x\rangle\langle x|_X \otimes |\psi_x\rangle\langle\psi_x|_B . \tag{15}$$

Here we've arbitrarily called the random variable X and the quantum system B, just to give them concrete names. Note that the conditional quantum states of B given X need not be pure states; they could be arbitrary mixed density operators.

The conditional entropy of A given B can be defined as the total entropy of A and B minus the entropy of B alone. For this definition we need the *von Neumann* entropy of a quantum system, which is $H(B)_\rho = -\mathrm{Tr}[\rho \log_2 \rho]$. Note that this is just the Shannon entropy of the eigenvalues of ρ. Now we can define the conditional entropy of A given B:

$$H(A|B)_\rho = H(AB)_\rho - H(B)_\rho \,. \tag{16}$$

We can apply it to the case of classical quantum states, replacing A by X. This quantity is meant to capture the uncertainty of the random variable X, conditioned on having access to quantum system B. Crucially, the conditional entropy satsifies the *data-processing inequality* which states that, among other things, if we perform a measurement on B and obtain random variable Y, then $H(X|Y) \geq H(X|B)$. So $H(X|B)$ puts a lower limit on uncertainty we have about X even if we get to use B in order to guess X.

The extension of the Maassen-Uffink relation states, for any tripartite density operator ρ_{ABC},

$$H(X_A|B)_\rho + H(Z_A|C)_\rho \geq 1 \,. \tag{17}$$

The notation $H(X_A|B)_\rho$ means that the X measurement is performed on A in state ρ_{ABC}, system C is ignored, and the conditional entropy of the resulting classical-quantum state on XB is computed. Similarly for $H(Z_A|C)$. In version 1 of the game Bob must report guesses for the Alice's measurement result up front; we can think of his X guess as being stored in system B and his Z guess as stored in system C. Importantly, they have to be different systems since the game requires both guesses. Since perfect guessing would correspond to $H(X_A|B)_\rho = H(Z_A|C)_\rho = 0$, this relation implies that there is no strategy enabling Bob to certainly win version 1 of the guessing game, even if his guesses can be "quantum".

In the winning strategy for version 2 of the game, Bob stores his guess in a single quantum system, B, and it turns out there is a bipartite version of the above uncertainty relation which covers this case. It states that for all bipartite density operators ρ_{AB},

$$H(X_A|B)_\rho + H(Z_A|B)_\rho \geq 1 + H(A|B)_\rho \,. \tag{18}$$

At first glance it doesn't look like this will be useful, since we've added a term to the righthand side. However, for bipartite quantum states, the conditional entropy can actually be negative! In fact, negative conditional entropy implies that the quantum state is entangled. And that is precisely the case for the winning strategy, where $\rho_{AB} = |\Phi\rangle\langle\Phi|_{AB}$. It is easy to see that $H(AB)_\rho = 0$ since the state has only one non-zero eigenvalue, but $H(B) = 1$ since the state $\mathrm{Tr}_A[|\Phi\rangle\langle\Phi|_{AB}] = \frac{1}{2}\mathbb{1}_B$. Thus, the righthand side is zero, as intended.

3.3 Relation to QEC and QKD

Version 1 of the game and the tripartite uncertainty relation (17) are relevant to QKD while version 2 of the game and the bipartite relation (18) are related to QEC. Later we will gain a more precise understanding of this relationship, but here let us just sketch the idea broadly so we can see where we are going.

3.3.1 QKD

The goal of QKD is for Alice and Bob to produce random numbers that are unknown to a would-be eavesdropper, who is usually called Eve. To do this, they will use a quantum channel—a means of transmitting quantum systems—under the control of Eve. So in the course of any protocol that attempts to produce these random numbers, we may anticipate that Eve will gain some information about what Alice and Bob are doing. Whatever it is, it will have to be stored in some possibly-large, possibly-quantum system C. Suppose that the random numbers are to be produced by Alice

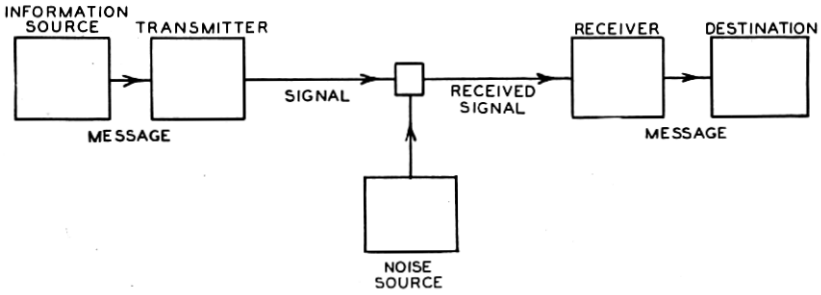

Figure 1: Shannon's depiction of the general communication task.

measuring her qubits in the Z basis. Then, what we require from QKD is for the uncertainty of Z_A to be large, even when conditioned on C. That is, given the information that Eve has collected during the course of the protocol and stored in C, it should nonetheless not be possible for her to predict the value of Z_A. One way to say this in a more quantitative way is that for the protocol to be secure, we need a lower bound on $H(Z_A|C)$.

The above holds for any kind of cryptographic protocol, really, and the main difficulty is to get such a lower bound, since by definition the honest parties Alice and Bob do not directly know what Eve is doing, so they don't directly know anything about C. But the uncertainty relation (17) gives us a way to bound $H(Z_A|C)$ in terms of quantities that the honest parties can in principle measure. For we may imagine that B is held by Bob and that Alice and Bob can determine $H(X_A|B)$ (even just by an upper bound), then they can infer how much information the eavesdropper could in principle have. Reliance on the uncertainty principle is what makes quantum key distribution different from any protocol relying on only classical information. In the latter case, it is impossible to bound Eve's knowledge in principle, because she sees all the communication between Alice and Bob and could just copy it without being detected.

3.3.2 QEC

Turning to quantum error correction, here the goal is to reliably transmit information over a noisy channel. The channel need not describe communication from one lab to another, it could also represent the action of noise on a quantum memory or in a quantum computer. The setup is the same as Shannon sketched in his seminal 1948 paper on classical communication, shown in Figure 1. The job of the transmitter is to "encode" the message such that the receiver will be able to reconstruct the input after obtaining the output of the channel (i.e. "decode" the message). A protocol for reliable communication consists of an encoder and a decoder. In general, we cannot expect the transmission to be absolutely perfect; we will have to tolerate some amount of error. Without going into the details of how we should mathematically specify the error (especially in the quantum case), we can nevertheless be a little more concrete and say that an (M, ϵ) protocol for reliable transmission over a noisy channel $\mathcal{N}$ consists of an encoder and a decoder such that arbitrary messages of size M are transmitted with error no larger than ϵ. In the quantum case, M refers to the dimension of the quantum state we would like to send, and is usually a power of 2, i.e. $M = 2^m$, so that we speak of transmitting m qubits.

Initially it was thought that error correction is not really possible for quantum systems, since they appear to be more like analog information than digital information. Analog in this context means the information can take a continuous range of values, rather than a finite and discrete set, and correction of analog information is fraught because it would appear to ultimately require being able to perform operations to arbitrary precision. Otherwise, errors due to finite-precision operations will build up and eventually make correction (to arbitrary precision) impossible. Since describing even just a qubit requires three real numbers, this was thought to rule out the possibility of quantum error-correction.

The breakthrough in Shor's and Steane's construction of quantum error correcting codes in 1995 and 1996 was to realize that quantum information is digital for the purposes of correction—it is only necessary to correct a small set of discrete errors. Indeed, it is sufficient to correct just two: X and Z errors!

We can appreciate this from the form of (18). If we manage to find an (M, ϵ) protocol, or code, one thing this should be able to do is transmit entanglement. That is, the sender Alice could create an entangled state $|\Phi\rangle_{AB}$ and then use the code to transmit B to Bob over the noisy channel. Let us denote the output by ρ_{AB}. Ideally ρ_{AB} is nearly equal to $|\Phi\rangle\langle\Phi|_{AB}$. Now, we are not yet sure this is possible. But imagine Alice measured A in the X basis or Z basis prior to transmission of B. In either case Bob will obtain some output at his end that he could use to guess Alice's measurement result. We see from (18) that if his guesses of X and Z would both be reliable, then the conditional entropies $H(X_A|B)_\rho$ and $H(Z_A|B)_\rho$ on the lefthand side must be small. And then the conditional entropy $H(A|B)_\rho$ must be nearly -1, which means that ρ_{AB} is indeed close to $|\Phi\rangle\langle\Phi|_{AB}$. Therefore, *it is enough to ensure that classical X and Z information are reliably transmitted in order to transmit entanglement.* This simplifies the construction of QECs, because we can break down the task into reliable transmission of classical information, which is a lot easier.

4 Error correction

4.1 The classical case in quantum language

The action of the X and Z operators, as unwanted modulation of the quantum state, are usually referred to as bit flips and phase flips, for the following reason. One commonly fixes a basis and calls it the amplitude basis, and then for an arbitrary qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ an amplitude error resulting from an unwanted X operator just flips the states $|0\rangle$ and $|1\rangle$, hence the name bit flip. Similarly, phase flips interchange the states $|+\rangle$ and $|-\rangle$, or equivalently, flips the phase of $|1\rangle$, taking (α, β) to $(\alpha, -\beta)$.

Either type of error by itself could be corrected in exactly the way a classical error would be corrected, through repetition. To correct a single bit flip error classically, we can *encode* it into three bits as follows,

$$0 \to \overline{0} = 000 \qquad 1 \to \overline{1} = 111. \tag{19}$$

These two bitstrings are called *codewords*, and the overline denotes a logical value of the encoded bit, as opposed to the values of the individual physical bits. Then, if one error occurs, we can correct it by examining each string and flipping the one bit which is different from the other two. Equivalently, the error may be diagnosed by computing the two parities, generally called *syndromes*, $s_1 = b_1 \oplus b_3$ and $s_2 = b_2 \oplus b_3$, where b_1, b_2, and b_3 are the three bit values. The syndromes associated to each error position are shown in Table 1. Note that the bit is encoded in the value of $\overline{b} = b_1 \oplus b_2 \oplus b_3$.

Bitstring pair $(\overline{0}, \overline{1})$	Error Position	Syndrome (s_1, s_2)
$(000, 111)$	$\emptyset$	$(0,0)$
$(100, 011)$	1	$(1,0)$
$(010, 101)$	2	$(0,1)$
$(001, 110)$	3	$(1,1)$

Table 1: The three-bit repetition code. The first column gives the bitstrings corresponding to the encoded logical zero $\overline{0}$ and logical one $\overline{1}$ after a bitflip error whose position is given in the second column. The third column lists the syndrome information which allows the error position to be diagnosed.

Seen from a different perspective, the reason this works is that the eight possible three-bit strings are grouped into four pairs, as in Table 1. One pair is given by the codewords themselves, and the

other pairs are the images of the codewords under the three single-bit errors. In each pair one string corresponds to $\overline{0}$ and the other to $\overline{1}$ as defined by this mapping. The syndromes reveal precisely which pair is present, but importantly they do not reveal anything about the logical bit value. Error-correction corresponds to mapping the noisy pair of strings back to the original pair.

To correct qubit bit flip errors we may simply use the same repetition code in the computational basis. Since the syndrome and correction procedure for a given error are independent of the encoded information, superpositions are also maintained by the error-correcting code. Thus, the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is encoded as $|\overline{\psi}\rangle = \alpha|000\rangle + \beta|111\rangle$, a process which can be implemented as a unitary transformation on the input and two auxiliary systems, each in some given state we can take to initially be prepared in the state $|0\rangle$. The necessary syndrome information can be generated by measuring the two *stabilizer* operators $Z\mathbb{1}Z = Z \otimes \mathbb{1} \otimes Z$ and $\mathbb{1}ZZ$, which we can write as Z_1Z_3 and Z_2Z_3. Each of these has the same action on the two logical states in each subspace, returning the values $(-1)^{s_1}$ and $(-1)^{s_2}$, respectively.

The name stabilizer reflects the fact that the code subspace is stabilized by the two operators, as it is the simultaneous $+1$ eigensubspace of both operators. The encoded subspace supports a single qubit, and so it must be possible to represent its amplitude and phase operators. One possibility is given by $\overline{Z} = Z_1Z_2Z_3$ and $\overline{X} = X_1X_2X_3$. These each commute with the stabilizers, but anticommute with each other as intended. Note that $\overline{Z}$ gives the encoded bit, just as in the classical case.

We can also think of the stabilizers and encoded amplitude operator as defining a new complete set of commuting observables for the set of physical qubits. Such a set fixes a basis in the state space of the three qubits, and each of the operators is the amplitude operator for a corresponding "virtual" qubit. Labeling the virtual qubit operators with primes, we can write $Z_1' = Z_1Z_2$, $Z_2' = Z_2Z_3$, and $Z_3' = \overline{Z} = Z_1Z_2Z_3$. Conjugate to the new amplitude observables are phase observables $X_1' = X_2X_3$, $X_2' = X_1X_3$, and $X_3' = \overline{X} = X_1X_2X_3$, which are found by ensuring that they anticommute with the amplitude operators of the same qubit but commute with all other operators. The entire collection is shown in Table 2. The code subspace is then defined by the first two virtual qubits being in the $+1$ amplitude state. Bit flip errors change the amplitude of the encoded qubit and at least one of the virtual qubits, and the stabilizer measurement determining the location of the error translates into an amplitude measurement of the first two virtual qubits.

Virtual qubit	Amplitude	Phase
1	$ZZ\mathbb{1}$	$\mathbb{1}XX$
2	$\mathbb{1}ZZ$	$XX\mathbb{1}$
3	ZZZ	XXX

Table 2: Virtual qubits associated with the three-qubit amplitude repetition code. Note that amplitude and phase anticommute for each qubit, but commute for different qubits.

Discretization is automatically provided by the measurement of the stabilizer operators. Consider an error operator of the form $E = e_0I + e_1X_1$, with $e_0, e_1 \in \mathbb{C}$, which is a sort of combination bit flip error and no error on the first qubit. It produces a superposition between two code subspaces,

$$|\overline{\psi}'\rangle = E|\overline{\psi}\rangle = e_0|\overline{\psi}\rangle + e_1X_1|\overline{\psi}\rangle = e_0\left(\alpha|000\rangle + \beta|111\rangle\right) + e_1\left(\alpha|100\rangle + \beta|011\rangle\right). \tag{20}$$

Measurement of the stabilizer operators destroys this superposition, forcing the system to the state of either one error or no error, but leaves the logical qubit superposition intact. Here the measurement has two possible syndrome outcomes, either $(0,0)$ or $(1,0)$, with probabilities $|e_0|^2/(|e_0|^2+|e_1|^2)$ and $|e_1|^2/(|e_0|^2+|e_1|^2)$, respectively. Conditioned on these outcomes, the state becomes $|\overline{\psi}\rangle$ or $X_1|\overline{\psi}\rangle$, respectively, and can therefore be corrected using the syndrome information.

4.2 Correcting Both Kinds of Errors

Since phase flips are just bit flips in the basis $|\pm\rangle$, the above analysis immediately applies to this case upon changing $X \leftrightarrow Z$ and working in the new basis. The insight of Shor and Steane was to realize that a single error of either type can be corrected by appropriately combining these procedures. Shor's scheme is conceptually somewhat simpler, and is based on concatenating the two error-correcting codes. That is, we take the codewords of the phase flip repetition code and replace each of the three qubits with qubits appropriately encoded in the bit flip repetition code. This produces codewords of nine qubits, as follows (here ignoring normalization),

$$|+\rangle \longrightarrow |\overline{+}\rangle = |+{+}{+}\rangle = (|0\rangle + |1\rangle)(|0\rangle + |1\rangle)(|0\rangle + |1\rangle) \tag{21}$$

$$|\overline{+}\rangle \longrightarrow |\widetilde{+}\rangle = (|\overline{0}\rangle + |\overline{1}\rangle)(|\overline{0}\rangle + |\overline{1}\rangle)(|\overline{0}\rangle + |\overline{1}\rangle), \tag{22}$$

$$|-\rangle \longrightarrow |\overline{-}\rangle = |-{-}{-}\rangle = (|0\rangle - |1\rangle)(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) \tag{23}$$

$$|\overline{-}\rangle \longrightarrow |\widetilde{-}\rangle = (|\overline{0}\rangle - |\overline{1}\rangle)(|\overline{0}\rangle - |\overline{1}\rangle)(|\overline{0}\rangle - |\overline{1}\rangle). \tag{24}$$

The repetition in the amplitude basis in the second step protects the encoded qubit from bit flip errors, since a single bit flip can always be detected and corrected by applying the 3-qubit repetition procedure to each block of three qubits. This corresponds to measuring the Z-parity observables Z_1Z_3, Z_2Z_3, Z_4Z_6, Z_5Z_6, Z_7Z_9, and Z_8Z_9. Phase flips are slightly more involved, but consider what happens when a single phase flip error plagues, say, the fourth qubit. This is the first qubit of the second block, so we can zoom in on this block to determine the effect on the encoded states. Applying the error operator Z_1 to the encoded states we find $Z_1|\overline{0}\rangle = Z_1|000\rangle = |000\rangle = |\overline{0}\rangle$, while $Z_1|\overline{1}\rangle = Z_1|111\rangle = -|111\rangle = -|\overline{1}\rangle$. Thus, the error causes the action

$$|\widetilde{+}\rangle \to (|\overline{0}\rangle + |\overline{1}\rangle)(|\overline{0}\rangle - |\overline{1}\rangle)(|\overline{0}\rangle + |\overline{1}\rangle) \tag{25}$$

$$|\widetilde{-}\rangle \to (|\overline{0}\rangle - |\overline{1}\rangle)(|\overline{0}\rangle + |\overline{1}\rangle)(|\overline{0}\rangle - |\overline{1}\rangle), \tag{26}$$

which is precisely a phase flip at the "inner" level. We could detect and correct this at the inner level by measuring the X-parities X_1X_3 and X_2X_3. Translating to the outer level of actual qubits, we replace each of the constituent X operators on the inner level by its encoded $\overline{X}$ operator on the outer level and instead measure $X_1X_2X_3X_7X_8X_9$ and $X_4X_5X_6X_7X_8X_9$. The outcomes for the damaged states are $+1$ and -1 respectively, for both encoded states, implying that to correct the error we merely need apply Z_4.[1]

The six amplitude parities and two phase parities commute pairwise and stabilize the code subspace. As with the repetition code, the error analysis is made simpler by thinking in terms of virtual qubits, in this case nine, as shown in Table 3. Observe that the concatenated structure is reflected in the operators: three copies of the repetition code in virtual qubits one through six, followed by the same repetition code on the three blocks. The code subspace is fixed by requiring virtual qubits one through six to be in the $+1$ amplitude eigenstate and virtual qubits seven and eight in the $+1$ phase eigenstate, but this structure makes it clear that we could have defined the code the other way around.

Using this framework it is easy to see that the Shor code also enables detection and correction of joint bit and phase errors. A joint bit and phase flip of the fourth qubit, for instance, would reveal itself by the fourth virtual qubit having the wrong amplitude and the seventh having the wrong phase, corresponding to -1 eigenvalues of the stabilizers Z_4Z_6 and $X_4X_5X_6X_7X_8X_9$. From the structure of the virtual amplitude and phase operators it is clear that the code can actually detect and correct one bit and one phase error, irrespective of their locations.

Again error discretization is provided by the stabilizer measurement, and fortunately, being able to correct just these two types of error is sufficient to correct any conceivable single-site error. Just as with the repetition code, we can consider the effect of arbitrary errors which are linear combinations

[1] Z_5 or Z_6 would also work just as well.

Virtual qubit #	Amplitude	Phase
1	$ZZ\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}$	$\mathbb{1}XX\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}$
2	$\mathbb{1}ZZ\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}$	$XX\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}$
3	$\mathbb{1}\mathbb{1}\mathbb{1}ZZ\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}$	$\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}XX\mathbb{1}\mathbb{1}\mathbb{1}$
4	$\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}ZZ\mathbb{1}\mathbb{1}\mathbb{1}$	$\mathbb{1}\mathbb{1}\mathbb{1}XX\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}$
5	$\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}ZZ\mathbb{1}$	$\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}XX$
6	$\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}ZZ$	$\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1}XX\mathbb{1}$
7	$ZZZZZZ\mathbb{1}\mathbb{1}\mathbb{1}$	$\mathbb{1}\mathbb{1}\mathbb{1}XXXXXX$
8	$\mathbb{1}\mathbb{1}\mathbb{1}ZZZZZZ$	$XXXXXX\mathbb{1}\mathbb{1}\mathbb{1}$
9	$ZZZZZZZZZ$	$XXXXXXXXX$

Table 3: Virtual qubits associated with the nine-qubit Shor code. Note that amplitude and phase anticommute for each qubit, but commute for different qubits.

of all the correctable errors. Since the Shor code can correct any single flip of bit and/or phase, errors of the form $E = e_{00}I + e_{10}X_1 + e_{01}Z_1 + e_{11}X_1Z_1$ with $e_{jk} \in \mathbb{C}$ can also be corrected. But, as can be readily verified, any operator can be expressed in this way as a complex combination of these four operators, meaning arbitrary single-site errors can be digitized to amplitude and/or phase errors and corrected. Despite initial appearances to the contrary, quantum information is therefore in a critical sense digital.

5 Quantum key distribution

Now let us examine the use of (17) for quantum cryptography, in particular quantum key distribution. First a little background on the uses and need for such a protocol. Imagine that our two parties Alice and Bob, who are sitting in separated labs, would like to communicate privately. However, they do not have any private means of communication between them, only public classical channels. Any eavesdropper, usually named Eve, could simply listen to their conversation.

In resource terms, we would like to have private communication by somehow making use of untrusted means of communication. With only classical resources, there is no solution to this problem such that the private communication is information-theoretically secure. This is the ultimate cryptographic goal in which the eavesdropper's collected information, call it E gives no information about the message, call it M, beyond what could be guessed in the absence of E. In terms of entropy, a scheme is information-theoretically secure if, for all possible E that Eve could obtain, we have

$$H(M) = H(M|E)\,. \tag{27}$$

The classical solution is *encryption*. Alice and Bob share a secret key K, which should be a uniformly-distributed random variable. Given a message M, the sender Alice computes the ciphertext $C = f(K, M)$ according to some function f and transmits C to the receiver Bob via the untrusted channel. The function f is known to all parties, including Eve; if a secret function should be used, then the choice of precisely which function is just part of the secret key K. The requirements for perfect security are that $H(M) = H(M|C)$, so that the ciphertext by itself offers no benefit in determining the message, but $H(M|C,K) = 0$, so that the key allows the message to be recovered from the ciphertext.

It turns out, as was shown by Shannon in 1949, that to meet these requirements we require $H(K) \geq H(M)$. If we model the message random variable M as uniform, so that all possible messages are equally-likely (this can be achieved by data compression), then the length of the key must be as long as the message.

The proof is relatively straightforward and makes repeated use of chain rules. Expand $H(MCK)$ in two ways. First,

$$H(MCK) = H(C) + H(M|C) + H(K|CM) = H(C) + H(M) + H(K|CM)\,, \tag{28}$$

where we have used the first security requirement in the second equality. Second,

$$H(MCK) = H(K) + H(M|C) + H(M|CK) + H(C|K) = H(K) + H(C|K), \tag{29}$$

where we use the second security requirement. The two equalities together imply

$$H(K) - H(M) = H(C) - H(C|K) + H(K|CM) = I(C:K) + H(K|CM) \geq 0. \tag{30}$$

Equality holds here if C and K are independent and K is a deterministic function of C and M. This can be achieved by the "one-time pad" encryption protocol. Supposing the random variables are all in $\{0,1\}^m$, take K to be uniformly random and set $C = M \oplus K$, where $\oplus$ denotes bitwise addition modulo 2. Then $M = C \oplus K$.

The security requirement $H(M) = H(M|C)$ is extremely stringent, in the sense that it means that the ciphertext is useless for determining the message, no matter what means are used. In practice, this requirement is usually relaxed to a computational security assumption, which roughly states that the ciphertext is useless for determining the message by reasonable computational means. More specifically, the claim could be that determining M from C is equivalent to solving a certain computational task, which is thought to be infeasible even in a relatively large amount of time.

The second difficulty of this security requirement is that it leads to the need for very long keys. And the above argument also rules out key expansion using an untrusted classical device, for the following reason. Suppose Alice and Bob already share a key K and they would like to generate a larger K' by exchanging additional messages S over the untrusted classical channel. The messages might go back and forth between the two of them, but here we collect everything together in S. At the end of the exchange, they each compute K' from K and S. Then they would like to use K' to encrypt a longer message M'. However, Eve has access to S, so now the relevant security statement is $H(M') = H(M'|C', S)$, where C' is the ciphertext. And C' depends on S, since K' does. Therefore, the only secret protecting the message M' is K itself, and so by the above argument we will need $|K| \geq |M'|$. The only solution to this problem is for Alice and Bob to generate a long enough key in the past to encrypt all the messages they need encrypting until they meet again.

If the untrusted channel is quantum, however, the situtation is different. Using the fact that Alice and Bob can detect Eve's interference, they can expand the size of their key. Very roughly speaking, Alice attempts to send Bob a randomly-generated string, as well as some additional information that enables them to detect Eve's interference. If they can determine how much information Eve has about the transmitted string, they can extract a shorter, secret version from it. Using entropic uncertainty relations, they can estimate the amount of information Eve has about the string from the correlations present in the additional information.

The simplest protocol for quantum key distribution (QKD) is the earliest, proposed by Bennett and Brassard in 1984 and known as BB84. It requires an untrusted quantum channel as well as a noise-free, authenticated classical channel. The latter is a channel in which Eve obtains all inputs to the channel, but cannot modify them. The protocol then works as follows. Alice and Bob repeat the following two steps n times:

1. Alice randomly generates one of the qubit states $|0\rangle$, $|1\rangle$, $|+\rangle$, or $|-\rangle$ and transmits it to Bob.

2. Bob randomly chooses the σ_z or σ_x basis and measures the incoming qubit in this basis.

These are the only quantum parts of the protocol. The next steps are

3. Bob announces which bases he measured (but not the measurement result).

4. Alice announces in which rounds the basis choices match. This results in strings $X_A^{r'}$, $X_B^{r'}$ of their X basis measurement results or state preparation choices and Z_A^r, Z_B^r of the same for the Z basis. Here r' is the number which match in X and r the number which match in Z.

5. Alice randomly selects $r-k$ numbers in $1,\ldots,r$ and transmits the Z_A values in these positions to Bob. She also announces the entire string $X_A^{r'}$. Bob computes and announces the fractions ϵ_z and ϵ_x of mismatches between the correspoinding portion of Z_B^r and $X_B^{r'}$.

6. Alice and Bob perform an error-correction procedure designed to remove the mismatches between Z_A^k and Z_B^k, where these are the remaining parts of Z_A^r and Z_B^r. This involves public communication of information S between them, the size of which depends on ϵ_x.

7. Alice computes the possible size m of the output key from ϵ_x, chooses a random function $f : \{0,1\}^k \to \{0,1\}^m$ with output size m, and announces her choice to Bob. They compute $K_A = f(Z_A^k)$ and $K_B = f(Z_A^k)$ use this as the key. If m is not positive, she announces that they should abort the protocol.

Before embarking on a sketch of why the protocol is secure, first note that we can alter the initial, quantum phase of the protocol to use entanglement. Instead of step 1, Alice can do the following

0. Alice generates a maximally entangled qubit pair $|\Phi\rangle_{AB}$ and transmit system B to Bob.

1. Alice randomly chooses the σ_z or σ_x basis and measures A in this basis.

The rest of the protocol proceeds as before. The two protocols look identical from Eve's point of view, and therefore if the latter is secure, then so is the former.

The advantage of the entanglement-based protocol is that we can appeal to entropic uncertainty relations to establish security. In step 7, Alice needs to compute the size m of available key that can be obtained from Z_A^k. In the limit of large k, it turns out that $m \approx H(Z_A^k|ES)$, where E is whatever information Eve has obtained during the execution of the protocol, e.g. from spying on the initial transmission. It could be a quantum system. This bound is the asymptotic achievability of randomness extraction; it is a general bound and not tied to the protocol in any way.

Using the chain rule, we can simplify the entropy quantity:

$$H(Z_A^k|ES) = H(Z_A^k S|E) - H(S|E) \geq H(Z_A^k|E) - \log|S|\,. \tag{31}$$

In the second equality we use the fact that the error-correction information S is a deterministic function of Z_A^k and the general bound $H(S|E) \leq \log|S|$. Using the entropic uncertaintly relation, $H(Z_A^k|E) + H(X_A^k|B) \geq k$, where X_A^k refers to the results of hypothetical σ_x measurements that Alice could have made on the qubits that she actually measured in the σ_z basis. Meanwhile, B is just Bob's part of the entangled state, and we can use montonicity of the conditional entropy to obtain $H(Z_A^k|E) + H(X_A^k|X_B^k) \geq k$. Even though the σ_x measurements were not performed, assuming that the choice of basis in each round is random, the estimate ϵ_x indicates how many mismatches there would have been between X_A^k and X_B^k. Hence, ϵ_x can be used to bound $H(X_A^k|X_B^k)$. Then the amount m of key is at least (roughly) $k - H(X_A^k|X_B^k) - \log|S|$. Crucially, the second two quantities can be computed, or at least bounded, using ϵ_x and ϵ_z, respectively.