Chapter 1

Preliminaries

In this chapter we introduce briefly the most important concepts of the mathematical formalism in quantum information theory. Note that we discuss only those parts of the formalism that are relevant for the lecture. For a more detailed discussion see for example the script in Quantum Information Theory.[1]

1.1 Hilbert spaces and operators on them

An *inner product space* is a vector space (over $\mathbb{R}$ or $\mathbb{C}$) equipped with an inner product $(\cdot, \cdot)$. A *Hilbert space* $\mathcal{H}$ is an inner product space such that the metric defined by the norm $\|\alpha\| \equiv \sqrt{(\alpha, \alpha)}$ is *complete*, i.e., every Cauchy sequence is convergent. We will often deal with finite-dimensional spaces, where the completeness condition always holds, i.e., inner product spaces are equivalent to Hilbert spaces.

We denote the set of *homomorphisms* (i.e., the linear maps) from a Hilbert space $\mathcal{H}$ to a Hilbert space $\mathcal{H}'$ by $\mathrm{Hom}(\mathcal{H}, \mathcal{H}')$. Furthermore, $\mathrm{End}(\mathcal{H})$ is the set of *endomorphism* (i.e., the homomorphisms from a space to itself) on $\mathcal{H}$, that is, $\mathrm{End}(\mathcal{H}) = \mathrm{Hom}(\mathcal{H}, \mathcal{H})$. The identity operator $\alpha \mapsto \alpha$ that maps any vector $\alpha \in \mathcal{H}$ to itself is denoted by id.

The *adjoint* of a homomorphism $S \in \mathrm{Hom}(\mathcal{H}, \mathcal{H}')$, denoted S^*, is the unique operator in $\mathrm{Hom}(\mathcal{H}', \mathcal{H})$ such that

$$(\alpha', S\alpha) = (S^*\alpha', \alpha) \ ,$$

for any $\alpha \in \mathcal{H}$ and $\alpha' \in \mathcal{H}'$. In particular, we have $(S^*)^* = S$. If S is represented as a matrix, then the adjoint operation can be thought of as the conjugate transpose.

In the following, we list some properties of endomorphisms $S \in \mathrm{End}(\mathcal{H})$.

- S is *normal* if $SS^* = S^*S$.
- S is *unitary* if $SS^* = S^*S = \mathrm{id}$. Unitary operators S are always normal.

[1] http://www.itp.phys.ethz.ch/education/fs09/qit/script_05.08.2009.pdf

- S is *Hermitian* if $S^* = S$. Hermitian operators are always normal.
- S is *positive* if $(\alpha, S\alpha) \geq 0$ for all $\alpha \in \mathcal{H}$. Positive operators are always Hermitian. We will sometimes write $S \geq 0$ to express that S is positive.
- S is a *projector* if $SS = S$. Projectors are always positive.

Given an orthonormal basis $\{e_i\}_i$ of $\mathcal{H}$, we also say that S is *diagonal with respect to* $\{e_i\}_i$ if the matrix $(S_{i,j})$ defined by the elements $S_{i,j} = (e_i, Se_j)$ is diagonal.

The bra-ket notation

In this script, we will make extensive use of a variant of Dirac's *bra-ket notation*, where vectors are interpreted as operators. More precisely, we identify any vector $\alpha \in \mathcal{H}$ with an endomorphism $|\alpha\rangle \in \mathrm{Hom}(\mathbb{C}, \mathcal{H})$, called *ket*, and defined as

$$|\alpha\rangle : \quad \gamma \mapsto \alpha\gamma$$

for any $\gamma \in \mathbb{C}$. The adjoint $|\alpha\rangle^*$ of this mapping is called *bra* and denoted by $\langle\alpha|$. It is easy to see that $\langle\alpha|$ is an element of the *dual space* $\mathcal{H}^* := \mathrm{Hom}(\mathcal{H}, \mathbb{C})$, namely the linear functional defined by

$$\langle\alpha| : \quad \beta \mapsto (\alpha, \beta)$$

for any $\beta \in \mathcal{H}$.

Using this notation, the concatenation $\langle\alpha||\beta\rangle$ of a bra $\langle\alpha| \in \mathrm{Hom}(\mathcal{H}, \mathbb{C})$ with a ket $|\beta\rangle \in \mathrm{Hom}(\mathbb{C}, \mathcal{H})$ results in an element of $\mathrm{Hom}(\mathbb{C}, \mathbb{C})$, which can be identified with $\mathbb{C}$. It follows immediately from the above definitions that, for any $\alpha, \beta \in \mathcal{H}$,

$$\langle\alpha||\beta\rangle \equiv (\alpha, \beta) \ .$$

We will thus in the following denote the scalar product by $\langle\alpha|\beta\rangle$.

Conversely, the concatenation $|\beta\rangle\langle\alpha|$ is an element of $\mathrm{End}(\mathcal{H})$ (or, more generally, of $\mathrm{Hom}(\mathcal{H}, \mathcal{H}')$ if $\alpha \in \mathcal{H}$ and $\beta \in \mathcal{H}'$ are defined on different spaces). In fact, any endomorphism $S \in \mathrm{End}(\mathcal{H})$ can be written as a linear combination of such concatenations, i.e.,

$$S = \sum_i |\beta_i\rangle\langle\alpha_i|$$

for some families of vectors $\{\alpha_i\}_i$ and $\{\beta_i\}_i$. For example, the identity $\mathrm{id} \in \mathrm{End}(\mathcal{H})$ can be written as

$$\mathrm{id} = \sum_i |e_i\rangle\langle e_i|$$

for any basis $\{e_i\}$ of $\mathcal{H}$.

Tensor products

Given two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the *tensor product* $\mathcal{H}_A \otimes \mathcal{H}_B$ is defined as the Hilbert space spanned by elements of the form $|\alpha\rangle \otimes |\beta\rangle$, where $\alpha \in \mathcal{H}_A$ and $\beta \in \mathcal{H}_B$, such that the following equivalences hold

- $(\alpha + \alpha') \otimes \beta = \alpha \otimes \beta + \alpha' \otimes \beta$
- $\alpha \otimes (\beta + \beta') = \alpha \otimes \beta + \alpha \otimes \beta'$
- $\mathbf{0} \otimes \beta = \alpha \otimes \mathbf{0} = \mathbf{0}$

for any $\alpha, \alpha' \in \mathcal{H}_A$ and $\beta, \beta' \in \mathcal{H}_B$, where $\mathbf{0}$ denotes the zero vector. Furthermore, the inner product of $\mathcal{H}_A \otimes \mathcal{H}_B$ is defined by the linear extension (and completion) of

$$\langle \alpha \otimes \beta | \alpha' \otimes \beta' \rangle = \langle \alpha | \alpha' \rangle \langle \beta | \beta' \rangle \ .$$

For two homomorphisms $S \in \mathrm{Hom}(\mathcal{H}_A, \mathcal{H}'_A)$ and $T \in \mathrm{Hom}(\mathcal{H}_B, \mathcal{H}'_B)$, the tensor product $S \otimes T$ is defined as

$$(S \otimes T)(\alpha \otimes \beta) \equiv (S\alpha) \otimes (T\beta) \tag{1.1}$$

for any $\alpha \in \mathcal{H}_A$ and $\beta \in \mathcal{H}_B$. The space spanned by the products $S \otimes T$ can be canonically identified[2] with the tensor product of the spaces of the homomorphisms, i.e.,

$$\mathrm{Hom}(\mathcal{H}_A, \mathcal{H}'_A) \otimes \mathrm{Hom}(\mathcal{H}_B, \mathcal{H}'_B) \cong \mathrm{Hom}(\mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{H}'_A \otimes \mathcal{H}'_B) \ . \tag{1.2}$$

This identification allows us to write, for instance,

$$|\alpha\rangle \otimes |\beta\rangle = |\alpha \otimes \beta\rangle \ ,$$

for any $\alpha \in \mathcal{H}_A$ and $\beta \in \mathcal{H}_B$.

1.1.1 Trace and partial trace

The *trace* of an endomorphism $S \in \mathrm{End}(\mathcal{H})$ over a Hilbert space $\mathcal{H}$ is defined by[3]

$$\mathrm{tr}(S) \equiv \sum_i \langle e_i | S | e_i \rangle$$

where $\{e_i\}_i$ is any orthonormal basis of $\mathcal{H}$. The trace is well defined because the above expression is independent of the choice of the basis, as one can easily verify.

The trace operation tr is obviously linear, i.e.,

$$\mathrm{tr}(uS + vT) = u\mathrm{tr}(S) + v\mathrm{tr}(T) \ ,$$

[2] That is, the mapping defined by (1.1) is an isomorphism between these two vector spaces.

[3] More precisely, the trace is only defined for *trace class operators* over a separable Hilbert space. However, all endomorphisms on a finite-dimensional Hilbert space are trace class operators.

for any $S, T \in \mathrm{End}(\mathcal{H})$ and $u, v \in \mathbb{C}$. It also commutes with the operation of taking the adjoint,[4]

$$\mathrm{tr}(S^*) = \mathrm{tr}(S)^* \ .$$

Furthermore, the trace is cyclic, i.e.,

$$\mathrm{tr}(ST) = \mathrm{tr}(TS) \ .$$

Also, it is easy to verify[5] that the trace $\mathrm{tr}(S)$ of a positive operator $S \geq 0$ is positive. More generally

$$(S \geq 0) \wedge (T \geq 0) \implies \mathrm{tr}(ST) \geq 0 \ . \tag{1.3}$$

The *partial trace*[6] tr_B is a mapping from the endomorphisms $\mathrm{End}(\mathcal{H}_A \otimes \mathcal{H}_B)$ on a product space $\mathcal{H}_A \otimes \mathcal{H}_B$ onto the endomorphisms $\mathrm{End}(\mathcal{H}_A)$ on $\mathcal{H}_A$. It is defined as product mapping $\mathcal{I} \otimes \mathrm{tr}$ where $\mathcal{I}$ is the identity operation on $\mathrm{End}(\mathcal{H}_A)$ and tr is the trace mapping elements of $\mathrm{End}(\mathcal{H}_B)$ to $\mathrm{End}(\mathbb{C})$. Because the trace is a completely positive map (see definition below) the same is true for the partial trace.

Similarly to the trace operation, the partial trace tr_B is linear and commutes with the operation of taking the adjoint.

1.2 Postulates of quantum mechanics

Despite more than one century of research, numerous questions related to the foundations of quantum mechanics are still unsolved (and highly disputed). For example, no fully satisfying explanation for the fact that quantum mechanics has its particular mathematical structure has been found so far. As a consequence, some of the aspects to be discussed in the following, e.g., the postulates of quantum mechanics, might appear to lack a clear motivation.

In this section, we pursue one of the standard approaches to quantum mechanics. It is based on a number of postulates about the states of physical systems as well as their evolution. (For more details, we refer to Section 2 of [1], where an equivalent approach is described.) The postulates are as follows:

1. States: The set of states of an isolated physical system is in one-to-one correspondence to the projective space of a Hilbert space $\mathcal{H}$. In particular, any physical state can be represented by a *normalized vector* $\phi \in \mathcal{H}$ which is unique up to a phase factor. In the following, we will call $\mathcal{H}$ the *state space* of the system.[7]

2. Composition: For two physical systems with state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the state space of the product system is isomorphic to $\mathcal{H}_A \otimes \mathcal{H}_B$. Furthermore,

[4] The adjoint of a complex number $\gamma \in \mathbb{C}$ is simply its complex conjugate.

[5] The assertion can, for instance, be proved using the spectral decomposition of S and T (see below for a review of the spectral decomposition).

[6] Here and in the following, we will use subscripts to indicate the space on which an operator acts.

[7] In quantum mechanics, the elements $\phi \in \mathcal{H}$ are also called *wave functions*.

if the individual systems are in states $\phi \in \mathcal{H}_A$ and $\phi' \in \mathcal{H}_B$, then the joint state is

$$\Psi = \phi \otimes \phi' \in \mathcal{H}_A \otimes \mathcal{H}_B \ .$$

3. Evolutions: For any possible evolution of an isolated physical system with state space $\mathcal{H}$ and for any fixed time interval $[t_0, t_1]$ there exists a *unitary* U describing the mapping of states $\phi \in \mathcal{H}$ at time t_0 to states

$$\phi' = U\phi$$

at time t_1. The unitary U is unique up to a phase factor.

4. Measurements: For any measurement on a physical system with state space $\mathcal{H}$ there exists an *observable* O with the following properties. O is a Hermitian operator on $\mathcal{H}$ such that each eigenvalue x of O corresponds to a possible measurement outcome. If the system is in state $\phi \in \mathcal{H}$, then the probability of observing outcome x when applying the measurement is given by

$$P_X(x) = \mathrm{tr}(P_x|\phi\rangle\langle\phi|)$$

where P_x denotes the projector onto the eigenspace belonging to the eigenvalue x, i.e., $O = \sum_x xP_x$. Finally, the state ϕ'_x of the system after the measurement, conditioned on the event that the outcome is x, equals

$$\phi'_x := \sqrt{\frac{1}{P_X(x)}} P_x\phi \ .$$

1.3 Density operators

In quantum information theory, one often considers situations where the state or the evolution of a system is only partially known. For example, we might be interested in a scenario where a system might be in two possible states ϕ_0 or ϕ_1, chosen according to a certain probability distribution. Another simple example is a system consisting of two correlated parts A and B in a state

$$\Psi = \sqrt{\frac{1}{2}}\big(e_0 \otimes e_0 + e_1 \otimes e_1\big) \in \mathcal{H}_A \otimes \mathcal{H}_B \ , \tag{1.4}$$

where $\{e_0, e_1\}$ are orthonormal vectors in $\mathcal{H}_A = \mathcal{H}_B$. From the point of view of an observer who has no access to system B, the state of A does not correspond to a fixed vector $\phi \in \mathcal{H}_A$, but is rather described by a mixture of such states.

Definition. *A* density operator ρ *on a Hilbert space* $\mathcal{H}$ *is a normalized positive operator on* $\mathcal{H}$*, i.e.,* $\rho \geq 0$ *and* $\mathrm{tr}(\rho) = 1$*. The set of density operators on* $\mathcal{H}$ *is denoted by* $\mathcal{S}(\mathcal{H})$*. A density operator is said to be* pure *if it has the form* $\rho = |\phi\rangle\langle\phi|$*. If* $\mathcal{H}$ *is d-dimensional and* ρ *has the form* $\rho = \frac{1}{d} \cdot \mathrm{id}$ *then it is called* fully mixed*.*

For separable Hilbert spaces[8], it follows from the spectral decomposition theorem that any density operator can be written in the form

$$\rho = \sum_x P_X(x)|e_x\rangle\langle e_x|$$

where P_X is the probability mass function defined by the eigenvalues $P_X(x)$ of ρ and $\{e_x\}_x$ are the corresponding eigenvectors. Given this representation, it is easy to see that a density operator is pure if and only if exactly one of the eigenvalues equals 1 whereas the others are 0. In particular, we have the following lemma.

Lemma 1.3.1. *A density operator ρ is pure if and only if* $\mathrm{tr}(\rho^2) = 1$.

Quantum-mechanical postulates in the language of density operators

In a first step, we adapt the postulates of Section 1.2 to the notion of density operators. At the same time, we generalize them to situations where the evolution and measurements only act on parts of a composite system.

1. States: The states of a physical system are represented as density operators on a Hilbert space $\mathcal{H}$. For an isolated system whose state, represented as a vector, is $\phi \in \mathcal{H}$, the corresponding density operator is defined by $\rho := |\phi\rangle\langle\phi|$.[9]

2. Composition: The states of a composite system with state spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ are represented as density operators on $\mathcal{H}_A \otimes \mathcal{H}_B$. Furthermore, if the states of the individual subsystems are independent of each other and represented by density operators ρ_A and ρ_B, respectively, then the state of the joint system is $\rho_A \otimes \rho_B$.

3. Evolution: Any isolated evolution of a subsystem of a composite system over a fixed time interval $[t_0, t_1]$ corresponds to a unitary on the state space $\mathcal{H}$ of the subsystem. For a composite system with Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ and isolated evolutions on both subsystems described by U_A and U_B, respectively, any state ρ_{AB} at time t_0 is transformed into the state[10]

$$\rho'_{AB} = (U_A \otimes U_B)(\rho_{AB})(U_A^* \otimes U_B^*) \tag{1.5}$$

   at time t_1.[11]

4. Measurement: Any isolated measurement on a subsystem of a composite system is is specified by an observable, as defined above. When applying a measurement $O_A = \sum_x xP_x$ on the first subsystem of a composite system $\mathcal{H}_A \otimes \mathcal{H}_B$ whose state is ρ_{AB}, the probability of observing outcome x is

$$P_X(x) = \mathrm{tr}(P_x \otimes \mathrm{id}_B \rho_{AB}) \tag{1.6}$$

[8] It means that the space has a countable orthonormal basis.

[9] Note that this density operator is pure.

[10] In particular, if $\mathcal{H}_B = \mathbb{C}$ is trivial, this expression equals $\rho'_A = U_A \rho_A U_A^*$.

[11] By induction, this postulate can be readily generalized to composite systems with more than two parts.

and the post-measurement state conditioned on this outcome is

$$\rho'_{AB,x} = \frac{1}{P_X(x)}(P_x \otimes \mathrm{id}_B)\rho_{AB}(P_x \otimes \mathrm{id}_B) \ . \tag{1.7}$$

It is straightforward to verify that these postulates are indeed compatible with those of Section 1.2. What is new is merely the fact that the evolution and measurements can be restricted to individual subsystems of a composite system. As we shall see, this extension is, however, very powerful because it allows us to examine parts of a subsystem without the need of keeping track of the state of the entire system.

Partial trace and purification

Let $\mathcal{H}_A \otimes \mathcal{H}_B$ be a composite quantum system which is initially in a state $\rho_{AB} = |\Psi\rangle\langle\Psi|$ for some $\Psi \in \mathcal{H}_A \otimes \mathcal{H}_B$. Consider now an experiment which is restricted to the first subsystem.

It is important to note that the reduced state $\rho_A = \mathrm{tr}(\rho_{AB})$ of a pure joint state ρ_{AB} is not necessarily pure. For instance, if the joint system is in state $\rho_{AB} = |\Psi\rangle\langle\Psi|$ for Ψ defined by (1.4) then

$$\rho_A = \frac{1}{2}|e_0\rangle\langle e_0| + \frac{1}{2}|e_1\rangle\langle e_1| \ , \tag{1.8}$$

i.e., the density operator ρ_A is fully mixed.

Conversely, any mixed density operator can be seen as part of a pure state on a larger system. More precisely, given ρ_A on $\mathcal{H}_A$, there exists a pure density operator ρ_{AB} on a joint system $\mathcal{H}_A \otimes \mathcal{H}_B$ (where the dimension of $\mathcal{H}_B$ is at least as large as the rank of ρ_A) such that

$$\rho_A = \mathrm{tr}_B(\rho_{AB}) \tag{1.9}$$

A pure density operator ρ_{AB} for which (1.9) holds is called a *purification* of ρ_A.

Mixtures of states

We will now give an interpretation of non-pure, or *mixed*, density operators. Consider a quantum system $\mathcal{H}_A$ whose state depends on a classical value Z and let $\rho_A^z \in \mathcal{S}(\mathcal{H}_A)$ be the state of the system conditioned on the event $Z = z$. Furthermore, consider an observer who does not have access to Z, that is, from his point of view, Z can take different values distributed according to a probability mass function P_Z.

Assume now that the system $\mathcal{H}_A$ undergoes an evolution U_A followed by a measurement $O_A = \sum_x xP_x$ as above. Then, according to the postulates of quantum mechanics, the probability mass function of the measurement outcomes x conditioned on the event $Z = z$ is given by

$$P_{X|Z=z}(x) = \mathrm{tr}(P_x U_A \rho_A^z U_A^*) \ .$$

Hence, from the point of view of the observer who is unaware of the value Z, the probability mass function of X is given by

$$P_X(x) = \sum_z P_Z(z) P_{X|Z=z}(x) \ .$$

By linearity, this can be rewritten as

$$P_X(x) = \mathrm{tr}(P_x U_A \rho_A U_A^*) \ . \tag{1.10}$$

where

$$\rho_A := \sum_z P_Z(z) \rho_A^z \ .$$

Alternatively, expression (1.10) can be obtained by applying the postulates of Section 1.3 directly to the density operator ρ_A defined above. In other words, from the point of view of an observer not knowing Z, the situation is consistently characterized by ρ_A.

We thus arrive at a new interpretation of mixed density operators. For example, the density operator

$$\rho_A = \frac{1}{2}|e_0\rangle\langle e_0| + \frac{1}{2}|e_1\rangle\langle e_1| \tag{1.11}$$

defined by (1.8) corresponds to a situation where either state e_0 or e_1 is prepared, each with probability $\frac{1}{2}$. The *decomposition* according to (1.11) is, however, not unique. In fact, the same state could be written as

$$\rho_A = \frac{1}{2}|\tilde{e}_0\rangle\langle\tilde{e}_0| + \frac{1}{2}|\tilde{e}_1\rangle\langle\tilde{e}_1|$$

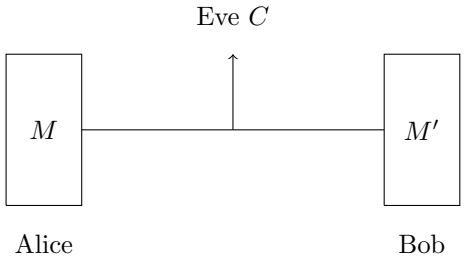
where $\tilde{e}_0 := \frac{1}{\sqrt{2}}(e_0 + e_1)$ and $\tilde{e}_1 := \frac{1}{\sqrt{2}}(e_0 - e_1)$. That is, the system could equivalently be interpreted as being prepared either in state $\tilde{e}_0$ or $\tilde{e}_1$, each with probability $\frac{1}{2}$.

Chapter 2

Quantum Cryptography

2.1 Classical cryptography and the one-time-pad

In a cryptographic scenario Alice wants to send Bob a message M over a public channel that might be eavesdropped by a third party Eve. Let M' be the message Bob receives and C the information Eve obtains about the message from wiretapping.



The requirements are the following.

- $M = M'$
- M is independent of C.

As a consequence of the following observation shown by Shannon (1949), this task cannot be achieved if Alice and Bob have access only to classical systems.

> If Bob can correctly decrypt M then there exists a function d such that $M = d(C)$.

The security of cryptographic schemes used in practise today relies on the assumption that the function d cannot be computed efficiently. For example the security of the well known RSA encryption scheme is based on the assumption that factoring is hard. Here P and Q are two large random prime numbers. Bob first sends Alice a public key $K = P \times Q$ over the public channel. Alice then uses the key to encrypt her message M and sends back the resulting message to Bob. The decryption scheme is such that it requires the knowledge of both

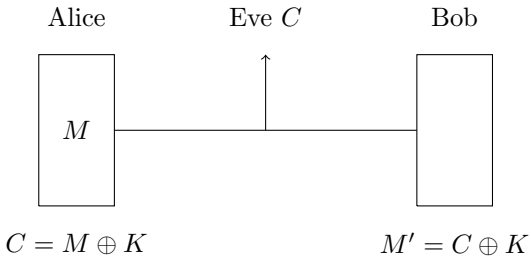
factors Q and P, which is assumed to be hard for Eve to calculate from K. In fact any known algorithm for factoring has a running time which is worse than polynomial in the size of the input. For practical purposes a key length of 1024 bits is assumed to be sufficient.

However, efficient factorisation is possible with a quantum computer. This is achieved with ShorÕs algorithm, which we will see later in this course. Luckily quantum computing is not just a danger for our current communication, but as we will see in the following, it is also the key for schemes which are provably secure.

How do quantum systems escape Shannon's observation? The key point is that it is implicitly assumed that Bob and Eve see the same communication C over the public channel. This is no longer true if quantum information is used for communication. It is a crucial feature of quantum systems that if Eve tries to obtain information about it by measurements she will disturb the state and Bob can therefore find out that she was eavesdropping.

However, this is not yet very useful: Bob only detects eavesdropping after Eve has already gathered secret data. This problem can be solved by using a *one-time-pad* for encryption. Here Alice first sends Bob a secret key over the channel, that does not contain any useful information. More precisely, a key is simply a sequence of bits that are uniformly distributed and independent of any information accessible to an adversary.[1] Let us therefore have a look at how a message can be sent secretly if Alice and Bob share a secret key.

For simplicity we assume that Alice wants to send Bob only one bit $M \in \{0, 1\}$ and they share a random key bit $K \in \{0, 1\}$ which is uniformly distributed and independent of Eve's information C.



The encrypted message sent over the public channel is $C = (M \oplus K)$. For decryption Bob calculates

$$M' = C \oplus K = (M \oplus K) \oplus K = M \oplus (K \oplus K) = M$$

and therefore recovers the original message. However, because K is unknown and distributed uniformly from Eve's point of view she cannot obtain any in-

[1] So-called "direct-communication-quantum-protocols" claim that they are secure by sending a message directly. These claims are however wrong.

formation about M.[2]

Note that one key bit can only be used for the encryption of one message bit (this is why it is called the one-time pad). If the same key was used to encode two bits Eve could obtain information about the correlations.[3]

2.2 Quantum key distribution

We have seen how Alice and Bob can communicate secretly, if they share a key which is independent from Eve. The problem is now: how to get such a key? In the following we will see how Alice and Bob can generate a shared random bit, if they have a maximally entangled state. However, in our proof we will a priori not make any assumptions on the formalism of quantum mechanics. The idea is to make use of the fact that strongly correlated systems are monogamous, i.e., there cannot be any additional correlated system.

Note that if we know that Alice and Bob share a maximally entangled state $|\psi\rangle\langle\psi|_{AB}$, we would directly find that Eve's state ρ_E must be uncorrelated because it holds that

$$\rho_{ABE} \text{ such that } \mathrm{tr}_E(\rho_{ABE}) = |\psi\rangle\langle\psi| \quad \Rightarrow \quad \rho_{ABE} = |\psi\rangle\langle\psi|_{AB} \otimes \rho_E.$$

Key distribution protocols with security proofs that to not make any assumptions about the internal workings of the device (for example that a particular quantum state is realised) are referred to as *device independent*.

We will now illustrate how Alice and Bob can establish a secret key if they can have access to systems that generate values that are correlated in a specific way. The crucial point is to establish the fact that Eve cannot have knowledge about the key (and not that Alice's and Bob's key are correlated – this is true by assumption).

Imagine that Alice and Bob were given two magically linked coins, which always come out the same side up – either two heads or two tails – with equal probabilities. Alice and Bob can then toss such coins at their respective locations, writing '0' for heads and 1' for tails. The resulting binary strings will be random and identical, but will they be secret? Not necessarily. Eve could have manufactured an additional coin, magically linked to the coins held by Alice and Bob. The three coins always tally and Eve knows all the bits in the string.

Clearly, to achieve secrecy we must let Alice and Bob do something that is beyond Eve's control. For example, Alice and Bob may be given a choice between two different coins; Alice can toss either a silver coin S_A or a golden coin G_A and Bob, either S_B or G_B. For each toss they must choose one of the two; tossing both S_A and G_A or both S_B and G_B is forbidden. Suppose, again, that the coins are magically linked; Alice and Bob's coins always come out the

[2]Formally one proves that $P_{MC} = P_M \times P_C$.

[3]See "Venona-Project" for an example where the fact that the same key was used more than once was exploited.

same, except when they toss S_A and G_B, in which case they always come out opposite. The magic can be summarized by the following four conditions

$$S_A = S_B, S_B = G_A, G_A = G_B, G_B \neq S_A. \tag{2.1}$$

These conditions are clearly contradictory; it is impossible to assign values to S_A, G_A, S_B and G_B so that all the four conditions are satisfied. But remember, Alice and Bob can toss only one coin each, and thus they can test only one of the four conditions in equation (2.1) at a time. Unperformed tosses do not have outcomes, and, hence, there is no contradiction here.

What if, say, Alice could break the rule and toss both of her coins, S_A and G_A, in one go? It turns out that she would deprive Bob of his free choice. Suppose that Alice tossed first and that her outcomes are such that $S_A = G_A$. Then Bob has no choice but to toss S_B, because this is the only choice compatible with the conditions in equation (2.1). This simple argument implies that the magic coins cannot be cloned. Having a clone, Z, of, say, S_A, and being able to toss it together with G_A would lead to the same contradictions as tossing both S_A and G_A. The existence of Z deprives Bob of his free choice. The conclusion is that if Alice and Bob have free choice then the magic correlations must be monogamous, that is, nothing else can be correlated to their coins. Therefore, Eve cannot manufacture a coin that will always tally with any of the coins held by Alice or Bob. All ingredients for secure key distribution are in place.

There is only one little problem, which is that the magic correlations do not exist. But all is not lost, because there are physically admissible correlations that are 'magical' enough for our purposes.

Let us assume that we have ε-magical coins satisfying the conditions

$$\Pr[S_A \neq S_B] \leq \varepsilon \tag{2.2}$$

$$\Pr[S_B \neq G_A] \leq \varepsilon \tag{2.3}$$

$$\Pr[G_A \neq G_B] \leq \varepsilon \tag{2.4}$$

$$\Pr[G_B = S_A] \leq \varepsilon. \tag{2.5}$$

We will prove that there cannot be an additional coin $Z = S_A$ provided epsilon is sufficiently small. In order to see this assume by contradiction that there is such a coin, i.e., $\Pr[Z = S_A] = 1$. Condition (2.2) implies

$$\Pr[Z \neq S_B] \leq \varepsilon. \tag{2.6}$$

Combining this with condition (2.3) yields

$$\Pr[Z \neq G_A] \leq 2 \cdot \varepsilon \overset{(2.5)}{\Rightarrow} \Pr[Z \neq G_B] \leq 3 \cdot \varepsilon \tag{2.7}$$

and finally with (2.5)

$$\Pr[Z = S_A] \leq 4 \cdot \varepsilon \tag{2.8}$$

which is a contradiction to our initial assumption that $Z = S_A$ with certainty for $\varepsilon < \frac{1}{4}$.

Even though perfect magic correlations, with $\varepsilon = 0$, do not exist, sufficiently magic coins with $\varepsilon < \frac{1}{4}$ can be constructed using quantum systems as follows.

Let us replace the coin tosses by appropriately chosen polarization measurements on a maximally entangled state

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}\big(\,|00\rangle_{AB} + |11\rangle_{AB}\,\big).$$

Instead of tossing coin S_A, Alice simply measures her photon along $\alpha_1 = 0$ and instead of tossing G_A, she measures the photon along $\alpha_2 = \pi/4$. Similarly, Bob replaces his coin tosses S_B and G_B by measurements along directions $\beta_1 = \pi/8$ and $\beta_2 = 3\pi/8$, respectively. The resulting joint probabilities satisfy conditions (2.2)-(2.5) with $\varepsilon = \sin^2(\pi/8) \leq 0.15$.

Note that Alice in Bob do not have to assume that their system is in the maximally entangled state. In fact, it would even be fine if Eve would manufacture the 'coins'. Once the devices pass a statistical test which ensures that Equations (2.4)-(2.8) hold, they can be used without any knowledge of their internal working. The maximally entangled state simply illustrates that sufficiently magical correlations indeed do exist in nature.

Key distribution protocol

1. Eve distributes n maximally entangled pairs using the quantum communication channel between Alice and Bob.[4]

2. Alice and Bob select some of the pairs at random and carry our measurements to check whether the statistics obey the correlations (2.2)-(2.5).

3. They measure the remaining pairs and keep the resulting bit sequences as raw keys R_A and R_B.

4. They perform error correction and privacy amplification to compute the final keys K_A and K_B (explained in the following).

Note that in the second step Alice and Bob need the existence of *free randomness*, i.e., that Eve cannot predict which pairs are used for the statistical test.

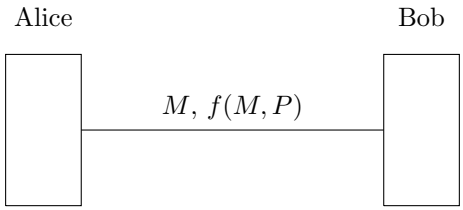
Error correction is needed because some of the bits will not agree as the correlations are not perfect. The purpose of privacy amplification is to reduce any remaining knowledge that Eve may have about the raw key bits. This is achieved by applying a function on the raw key that compresses it to a shorter final key.[5]

In order to exclude the possibility that Eve could tamper with Alice's message in the verification step, we assume that Alice and Bob can communicate authentically, i.e., such that Eve cannot alter messages. This can be achieved in practice, as there exist protocols to obtain authentic communication from an insecure communication link as well as a (short) password P shared by Alice

[4] In a worst case scenario the adversary distributes the state.

[5] This function could for example be the XOR.

and Bob. The rough idea is illustrated in the picture below.



In addition to the message M Alice sends Bob the value of a hash function $f(M,p)$ evaluated on M and a shared secret password p. Therefore Bob would notice if Eve would change the message because without the knowledge of the password she cannot change the value of the hash function.

Note that in order to break the symmetry between Eve and Alice a password is needed. In practise this is not a problem, because it can be short and in contrast to the key in the one-time pad it can be used several times.

In practice QKD systems have been broken because of so-called *loopholes*.[6] The most famous loophole is the *detection loophole*. As a consequence of the limited efficiency for real detectors many of the events will not be detected at all. The straightforeward approach would be to simply discard these events. However, this would bias the statistics (also called *post-selection*) because if Eve would control the detectors she could simply cause them to discard "unwanted" events. Therefore, she could trick Alice and Bob into believing that the system has the desired correlations when in reality it does not. The solution to this problem is to randomly assign "0" and "1" to the the non-detection events. This dilutes the statistics challenging the security proof for real implementations.

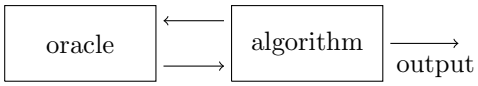
[6] These are deviations of the actual implementation from the theoretical prescription.

Chapter 3

Quantum algorithms

An *algorithm* is a recipe for performing a task (for example adding two numbers). We will look at problems that have classical inputs and outputs, while the algorithm may internally store information in quantum registers. As we will see this allows to solve some problems more efficiently[1] compared to purely classical computation.

If the input is given as an oracle rather than a value we speak of an *oracle based algorithm*.



Reversible computation

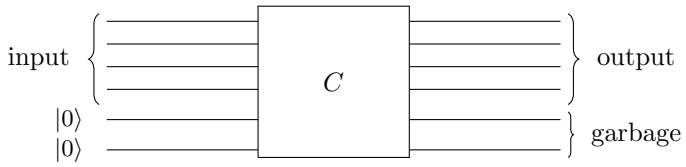
We use the circuit model to represent algorithms. A *circuit* is a sequence of building blocks that carry out elementary computations, called *gates*, connected by wires. In general, these gates may or may not be reversible. However, as quantum theory is unitary, one usually considers gates whose action is unitary, and therefore reversible.

An example for an *irreversible* gate is the AND-gate.

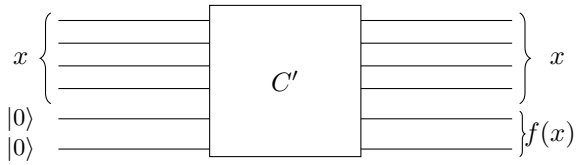
[1] This means that the algorithm requires fewer computational steps or less queries to the oracle as a function of the size of the input.

The CNOT-gate on the other hand is for example reversible.

Let f be a computable function. Then there exists a circuit that evaluates f. Replacing all gates of the circuit by reversible ones (these may take additional inputs and produce additional outputs) one can transform it into a circuit C of the following form.[2]

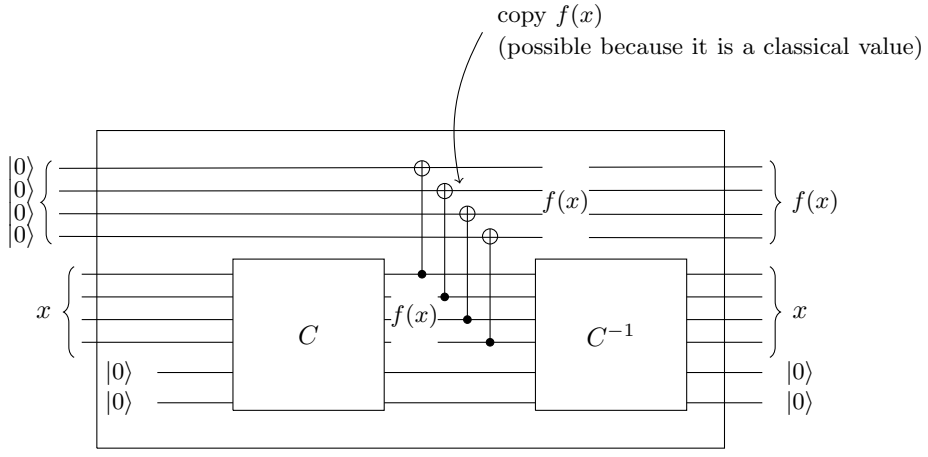


The ultimate goal would be to create a new reversible circuit C' which also computes f but does not produce any garbage. However, this is impossible, because sometimes we lose information in the computation (for example when computing the AND). What we can do is to create no garbage except for a copy of the input. In other words, we want to generate a circuit C' that has the following functionality.



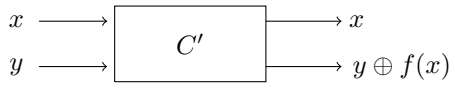
$(\star)$

Such a circuit may be constructed explicitly as follows.

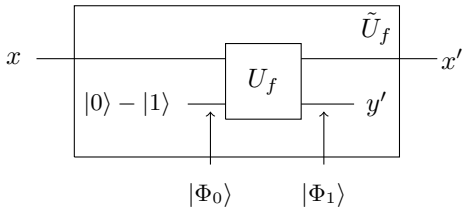


[2] A function is computable if it can be described by an algorithm. If a function can be described by an algorithm, it has a circuit model. An example for a function that is not computable is the Halting problem. Note also that every circuit can be decomposed into a set of universal gates (e.g. the AND and the NOT gate) and that these can be expressed as reversible gates. Thus, every computable function has a reversible circuit.

The above argument implies that for any computable function f there exits a reversible circuit C' that implements the following transformation.



For any unitary U_f that implements a function f we define the unitary $\tilde{U}_f$ by the following circuit. Note that the second input to U_f is a fixed state[3] $|y\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ and the bigger unitary $\tilde{U}_f$ corresponding to the following circuit.



(⋆⋆)

The total state before the application of U_f is given by

$$|\Phi_0\rangle = \frac{1}{\sqrt{2}}|x\rangle(|0\rangle - |1\rangle)$$

and afterwards

$$|\Phi_1\rangle = \begin{cases} |x\rangle|0\rangle - |x\rangle|1\rangle & \text{if } f(x) = 0 \\ |x\rangle|1\rangle - |x\rangle|0\rangle & \text{if } f(x) = 1 \end{cases}$$

$$= (-1)^{f(x)}|x\rangle(|0\rangle - |1\rangle)$$

$$= (-1)^{f(x)}|x\rangle|y\rangle,$$

therefore we can safely ignore $|y\rangle$ as it is unchanged. This proves that $\tilde{U}_f$ implements the following operation

$$|x\rangle \to (-1)^{f(x)}|x\rangle.$$

Deutsch-Josza algorithm

The Deutsch-Jozsa problem is defined as follows. We are fiven a function $f : \{0,1\}^n \to \{0,1\}$ (corresponding to a unitary U_f) with the promise that it is either constant or balanced (i.e. the number of inputs that are mapped to 0 and 1 is equal). The function is realised by an oracle of the form (⋆⋆).

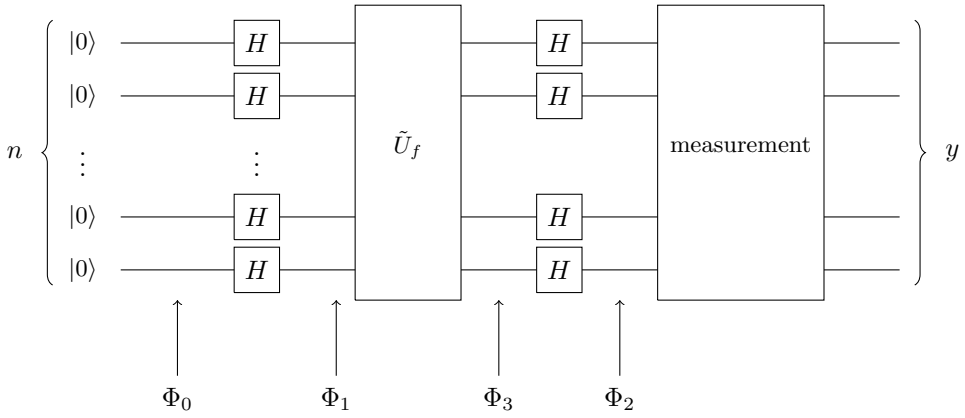
Goal: Determine whether f is constant or balanced.

Classical solution: Requires at least two queries to the oracle.

[3] We omit the normalisation sometimes for simplicity.

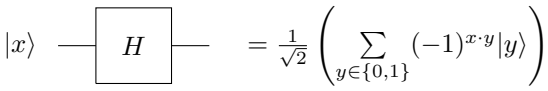
Quantum solution (the Deutsch Jozsa algorithm): Requires only one input (as we will see in the following).

The Deutsch-Jozsa algorithm is specified by the following circuit.

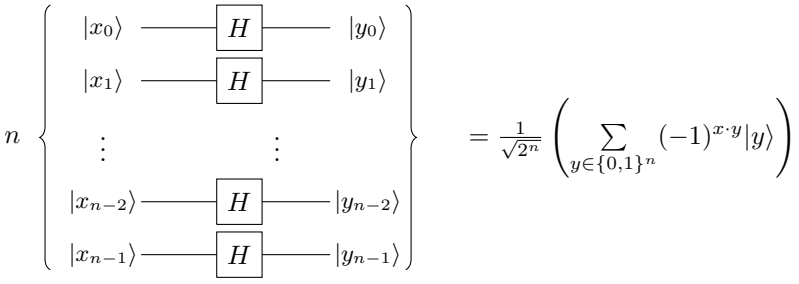


Claim: If the outcome y equals the zero string $\underbrace{00\ldots00}_{n}$ then f was constant, otherwise it was balanced.

Remark 3.0.1. Note that a Hadamard gate performs the following transformation on a state $|x\rangle$ (where $x \in \{0,1\}$)



Therefore we have for $x \in \{0,1\}^n$



where $x \cdot y$ corresponds to the binary scalar product.

Let us now look at the states after each step in the circuit.

$$|\Phi_0\rangle = |0\rangle^{\otimes n}$$

$$|\Phi_1\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle$$

$$|\Phi_2\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle$$

$$|\Phi_3\rangle = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} (-1)^{f(x)} \sum_{y\in\{0,1\}^n} (-1)^{x\cdot y}|y\rangle$$

$$= \frac{1}{2^n} \sum_{y\in\{0,1\}^n} \left[\sum_{x\in\{0,1\}^n} (-1)^{f(x)+x\cdot y} \right] |y\rangle$$

Therefore, the probability to get the zero string as outcome is given by

$$\Pr[y=0] = |\langle 0|^{\otimes n} \cdot |\Phi_3\rangle|^2$$

$$= \left(\frac{1}{2^n}\right)^2 \left(\sum_x (-1)^{f(x)}\right)^2$$

$$= \begin{cases} 1 & \text{if } f \text{ constant} \\ \\ 0 & \text{if } f \text{ balanced.} \end{cases}$$

This completes the proof of the claim.

Grover's algorithm

Grover's algorithm is an algorithm searching an unsorted database with N elements in $O(N^{1/2})$ time.[4] More precisely the problem is the following (where we assume for simplicity that $N = 2^n$):
Given an oracle $\tilde{U}_f$ where $f : \{0,1\}^n \to \{0,1\}$ is such that $f(x) = 1$ for exactly one input w

$$f(x) = \begin{cases} 1 & \text{if } x = w \\ \\ 0 & \text{else} \end{cases}$$

The goal is to find w.
The expected number of oracle calls for the best classical solution is equal to $\frac{2^n}{2}$.

[4]Note that a classical algorithm needs $O(N)$ time.

Grover's algorithm corresponds to the optimal quantum solution and requires $\sqrt{2^n}$ oracle calls.

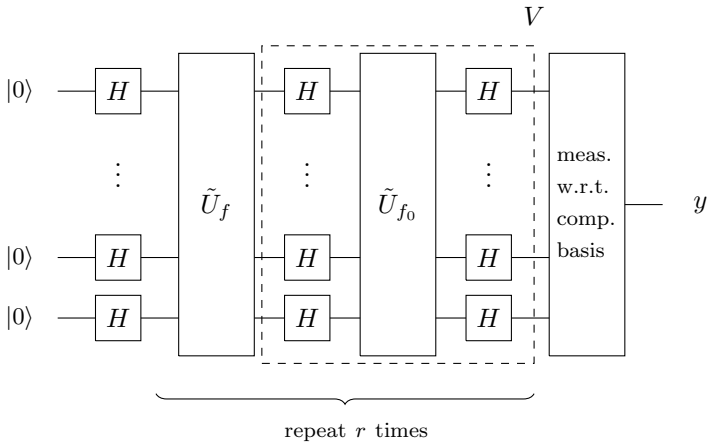
Before we analyse Grover's algorithm we discuss some examples for applications. Note that the algorithm is useful for questions that in the computational class NP, i.e., they are difficult to answer (i.e. there is no known solution in polynomial time w.r.t. the input size), but it is easy to verify that a given answer is correct (i.e., there is a polynomial time algorithm to check).

1. factoring a number m: Here $x = (p, q)$ where $p, q \in \mathbb{N}$ are prime numbers such that $p < q$:
$$f(x) = \begin{cases} 1 & \text{if } m = p \cdot q \\ 0 & \text{else.} \end{cases}$$
Note that this function is a simple multiplication and can therefore be implemented efficiently. Grover's algorithm provides an advantage compared to the basic classical solution (where we essentially have to check all the numbers smaller than $\sqrt{m}$). However, for factoring Shor's algorithm is even even better, namely of order $O(\log m)$.

2. Grover's algorithm is useful for puzzles such a Soduku, where x corresponds to a specific filling of the grid and
$$f(x) = \begin{cases} 1 & \text{if the filling is valid} \\ 0 & \text{else.} \end{cases}$$

3. A more practically useful example would be to find an aerodynamic design of a car. In this case x is the shape of the car and
$$f(x) = \begin{cases} 1 & \text{if air resistance} < \text{some threshold} \\ 0 & \text{else.} \end{cases}$$

Let us now look at the circuit of Grover's algorithm,



where

$$f_0(x) = \begin{cases} 0 & \text{if } x = 00\ldots0 \\ 1 & \text{else.} \end{cases}$$

The claim is that $y = w$ (i.e., the algorithm finds the input y such that $f(y) = 1$).

Let us first look at what the dashed box V does. We have

$$\tilde{U}_{f_0}: \quad |0\rangle^{\otimes n} \longrightarrow |0\rangle^{\otimes n}$$

$$|x\rangle \longrightarrow -|x\rangle \quad \text{if } x \neq 00\ldots0$$

and therefore

$$\tilde{U}_{f_0} = 2|0\rangle\langle 0|^{\otimes n} - \mathrm{id}$$

yielding

$$\begin{aligned} V &= H^{\otimes n}\tilde{U}_{f_0}H^{\otimes n} \\ &= H^{\otimes n}(2|0\rangle\langle 0|^{\otimes n} - \mathrm{id})H^{\otimes n} \\ &= 2 \cdot H^{\otimes n}|0\rangle\langle 0|^{\otimes n}H^{\otimes n} - H^{\otimes n} \cdot H^{\otimes n} \end{aligned}$$

Defining $|s\rangle$ as

$$|s\rangle := H^{\otimes n}|0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle$$

we can write

$$V = 2|s\rangle\langle s| - \mathrm{id}.$$

Furthermore we have

$$\tilde{U}_f = -2|w\rangle\langle w| + \mathrm{id}.$$

Grover's algorithm carries out the following operation $(V \cdot \tilde{U}_f)^r$ on the state

$$|s\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} |x\rangle$$

Let now Σ be the plane spanned by $|s\rangle$ and $|w\rangle$ and let $|s'\rangle$ be the state orthogonal to $|w\rangle$

$$|s'\rangle = \frac{1}{\sqrt{2^n - 1}} \sum_{x\neq w} |x\rangle \quad \in \Sigma.$$

Therefore we have

$$|s\rangle = \sqrt{\frac{2^n - 1}{2^n}}|s'\rangle + \sqrt{\frac{1}{2^n}}|w\rangle.$$

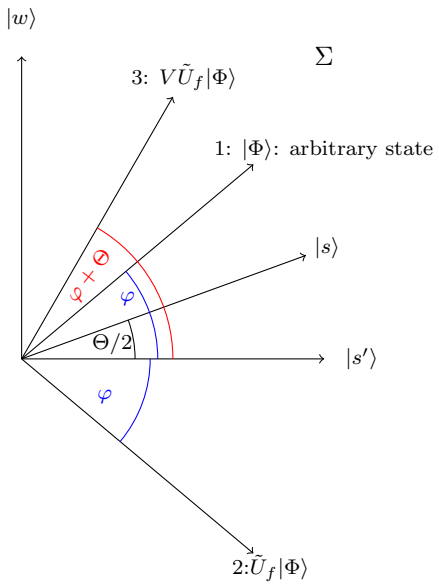
For large n we define the angle Θ such that $\sin\frac{\Theta}{2} = \sqrt{\frac{1}{2^n}}$ and write

$$|s\rangle = \cos\frac{\Theta}{2}|s'\rangle + \sin\frac{\Theta}{2}|w\rangle$$

i.e., $|s'\rangle$ corresponds to a rotation of $|s\rangle$ around $\frac{\Theta}{2}$.

The figure below illustrates the operation implemented by Grover's algorithm. Consider an arbitrary state $|\Phi\rangle$. First we apply $\tilde{U}_f$ corresponding to a reflection at the vector orthogonal to $|w\rangle$ (i.e., at $|s'\rangle$). Then V is applied which corresponds to a reflection at $|s\rangle$. The resulting state is rotated by an angle Θ with respect to the original state

$$|\Phi\rangle \to V\tilde{U}_f|\Phi\rangle = R_\Theta|\Phi\rangle.$$



After r applications of $V\tilde{U}_f$ the state $|s\rangle$ will be rotated by an angle $r \cdot \Theta$. Choose now r such that $r\Theta + \frac{\Theta}{2} \approx \frac{\pi}{2}$. Using that $\Theta \approx 2\sqrt{\frac{1}{2^n}}$ we find that after

$$r \approx \frac{\pi}{4}\sqrt{2^n}$$

calls to the oracle the final measurement will result in w with a probability of almost 1.[5] The number or oracle calls is therefore as promised of order $O(\sqrt{2^n})$ (where 2^n is the size of the input alphabet). It can be shown that the algorithm is optimal (see exercises).

Generalization

If $f(x) = 1$ holds for a set of values x of size k, then the number of oracle calls is of order $O(\sqrt{\frac{2^n}{k}})$ (this is also optimal). The difficulty is that the value of k has to be known. However, it is not a problem if it is not known, because one can simply run the algorithm for all possible values of k and check each time whether the solution is correct.

[5] We can match the state with probability $1 - \frac{1}{2^n}$.

The Quantum Fourier Transform

The quantum Fourier transform with respect to an orthonormal basis $\{|x\rangle\} = \{|0\rangle, \ldots, |N-1\rangle\}$ is defined as the linear operator with the following action on the basis vectors

$$\mathrm{QFT}: \ |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} w^{x\cdot y}|y\rangle,$$

where w is the N-th root of 1

$$w = e^{\frac{2\pi i}{N}}.$$

A general state

$$|\psi\rangle = \sum_{x=0}^{N-1} \alpha_x |x\rangle$$

if therefore mapped to

$$\mathrm{QFT}(|\psi\rangle) = \frac{1}{\sqrt{N}} \sum_{x,y} \alpha_x w^{x\cdot y}|y\rangle = \frac{1}{\sqrt{N}} \sum_{y} \beta_y |y\rangle$$

with

$$\beta_y = \sum_x w^{x\cdot y}\alpha_x = \sum_x e^{i\frac{2\pi x\cdot y}{N}}\alpha_x,$$

i.e., the amplitudes β_y are the discrete Fourier transforms of the amplitudes α_x of the original state.

Let us check that QFT is a unitary transformation:

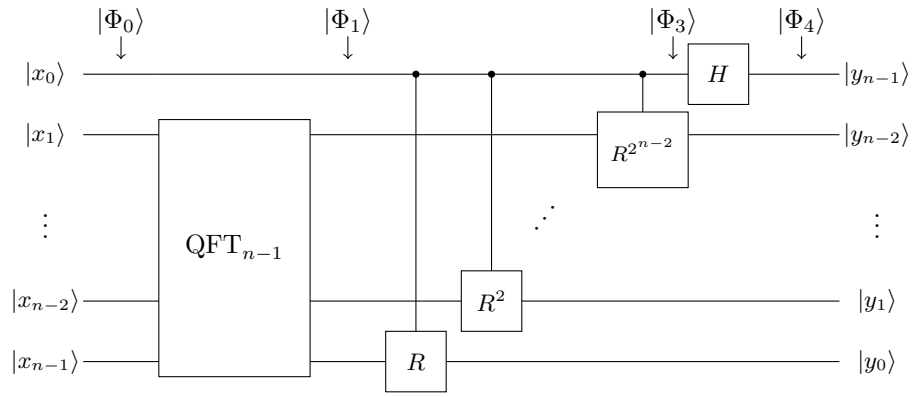
$$\begin{aligned}\mathrm{QFT}^\dagger \circ \mathrm{QFT}(|x\rangle) &= \frac{1}{N}\sum_{x'}\sum_{y} w^{-x'\cdot y} w^{x\cdot y}|x'\rangle \\ &= \frac{1}{N}\sum_{x'} \underbrace{\sum_{y} w^{(x-x')\cdot y}}_{:=c_{x'}} |x'\rangle\end{aligned}$$

Observing that $c_{x'} = w^{x-x'}c_{x'}$ it follows that

$$c_{x'} = \begin{cases} 0 & \text{if } x \neq x' \\ N & \text{if } x = x' \end{cases}$$

and therefore $\mathrm{QFT}^\dagger \circ \mathrm{QFT} = \mathrm{id}$, i.e., QFT is unitary.

Next we will verify that the following recursive circuit implements the QFT (here QFT_{n-1} denotes the QFT on $n-1$ qubits)



with

$$R = \begin{pmatrix} 1 & 0 \\ 0 & w \end{pmatrix}, \quad w = e^{\frac{2\pi i}{N}}.$$

We start with the state

$$|\Phi_0\rangle = |x\rangle = |x_{n-1} \dots x_0\rangle$$

where

$$x = \sum_{i=0}^{n-1} 2^i x_i$$

is the binary representation of x.

Consider the QFT_{n-1} of $|x_{n-1} \dots x_1\rangle$:[6]

$$\mathrm{QFT}(|x_{n-1} \dots x_1\rangle) = \sum_{y_0 \dots y_{n-2}} (\underbrace{e^{\frac{2\pi i}{2^{n-1}}}}_{=w^2})^{(x_{n-1} \dots x_1)(y_{n-2} \dots y_0)} |y_{n-2} \dots y_0\rangle,$$

therefore we have

$$\begin{aligned} |\Phi_1\rangle &= \sum_{y_0 \dots y_{n-2}} w^{2(x_{n-1} \dots x_1)(y_{n-2} \dots y_0)} |y_{n-2} \dots y_0\rangle |x_0\rangle \\ &= \sum_{y_0 \dots y_{n-2}} w^{(x_{n-1} \dots x_1 0)(0 y_{n-2} \dots y_0)} |y_{n-2} \dots y_0\rangle |x_0\rangle. \end{aligned}$$

Now we observe that R is applied if x_0 and y_0 are equal to 1. Therefore, we can simply multiply by $w^{x_0 y_0}$. Analogously the gate R^2 corresponds to a multiplication with $w^{2x_0 y_1}$ and finally R^{2n-2} to a multiplication with $w^{2^{n-2} x_0 y_{n-2}}$. Therefore we get

$$\begin{aligned} |\Phi_3\rangle &= \sum_{y_0 \dots y_{n-2}} w^{x_0 \cdot (y_0 + 2y_1 + \dots + 2^{n-2} y_{n-2})} w^{(x_{n-1} \dots x_1 0)(0 y_{n-2} \dots y_0)} |y_{n-2} \dots y_0\rangle |x_0\rangle \\ &= \sum_{y_0 \dots y_{n-2}} w^{x_0 \cdot (0 y_{n-2} \dots y_0)} w^{(x_{n-1} \dots x_1 0)(0 y_{n-2} \dots y_0)} |y_{n-2} \dots y_0\rangle |x_0\rangle \\ &= \sum_{y_0 \dots y_{n-2}} w^{(x_{n-1} \dots x_1 x_0)(0 y_{n-2} \dots y_0)} |y_{n-2} \dots y_0\rangle |x_0\rangle. \end{aligned}$$

[6] Note that the multiplication in the exponent is with respect to the binary representation and not bitwise.

We can write

$$\begin{aligned} H|x_0\rangle &= \sum_{y_{n-1}} (-1)^{x_0 y_{n-1}} |y_{n-1}\rangle \\ &= \sum_{y_{n-1}} (w^{2^{n-1}})^{x_0 y_{n-1}} |y_{n-1}\rangle \\ &= \sum_{y_{n-1}} w^{x_0 (y_{n-1} 0 \ldots 0)} |y_{n-1}\rangle \\ &= \sum_{y_{n-1}} w^{(x_{n-1} \ldots x_0)(y_{n-1} 0 \ldots 0)} |y_{n-1}\rangle \end{aligned}$$

The last equality holds because

$$w^{(x_{n-1} \ldots x_1 0)(y_{n-1} 0 \ldots 0)} = 1.$$

Therefore we get

$$\begin{aligned} |\Phi_4\rangle &= \sum_{y_0 \ldots y_{n-1}} w^{(x_{n-1} \ldots x_1 x_0)(0 y_{n-2} \ldots y_0)} w^{(x_{n-1} \ldots x_0)(y_{n-1} 0 \ldots 0)} |y_{n-1} \ldots y_0\rangle \\ &= \sum_{y_0 \ldots y_{n-1}} w^{(x_{n-1} \ldots x_1 x_0)(y_{n-1} y_{n-2} \ldots y_0)} |y_{n-1} \ldots y_0\rangle \\ &= \mathrm{QFT}(|x\rangle) \end{aligned}$$

Period finding

The problem is the following. We are given a periodic function

$$f : \{0, \ldots, N-1\} \to \{0, \ldots, N-1\}, \qquad N = 2^n$$

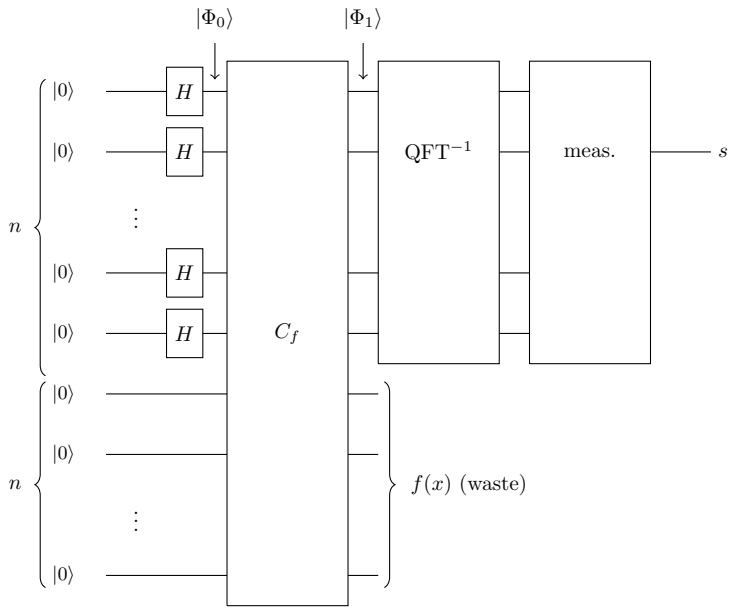
such that

$$f(x) = f(x+r) \quad r \neq 0 \tag{$\triangle$}$$

$$f(x) \neq f(x+s) \quad s < r \tag{$\triangle\triangle$}$$

and the goal is to find the period r. The function is given to us as circuit C_f of the form ($\star$).

The quantum algorithm for period finding (proposed by Shor) is realised by the following circuit.



We will show in the following that the outcome s of this circuit is a multiple of N/r.

We have

$$|\Phi_0\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle |0\rangle^{\otimes n}$$

and

$$|\Phi_1\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle |f(x)\rangle^{\otimes n}.$$

For the analysis we assume that the second output of C_f (which is marked as "waste") is measured and that the outcome is z. Note that at the end we will see that the analysis in independent of that outcome.[7] The post-measurement state of the first n qubits conditioned on that outcome is given by

$$|\Phi_1\rangle_z = \sqrt{\frac{r}{N}} \sum_{x:\ f(x)=z} |x\rangle.$$

Note that $|\Phi_1\rangle_z$ is not obtained by simply taking the partial trace over the second n qubits but corresponds to the state projected onto $\mathrm{id} \otimes |z\rangle$ (see Eq. (1.6)).

Let now x_0 be the smallest value such that $f(x_0) = z$. Because of periodicity ($\triangle$) it follows that

$$f(x_0 + t \cdot r) = z \qquad \forall\ t \in \mathbb{N}$$

and ($\triangle\triangle$) implies

$$f(x) \neq z \qquad \text{if } x \neq x_0 + t \cdot r.$$

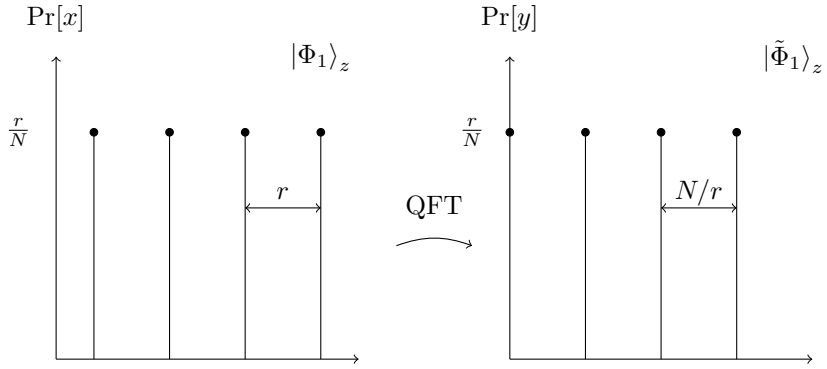
[7] Instead of considering this measurement, we could carry out the analysis for the remaining qubits. But these would be in a mixed state, which makes the analysis more complicated.

Therefore we can write

$$|\Phi_1\rangle_z = \sqrt{\frac{r}{N}} \sum_{t=0}^{N/r-1} |x_0 + t \cdot r\rangle.$$

For simplicity we assume in the following that $r \big| N$, i.e., that r divides N.[8]

The intuition for the circuit is the following. If we would measure $|\Phi_1\rangle_z$ before the application of the QFT the outcomes would be all equally likely and separated by r (this is a direct consequence of the form of $|\Phi_1\rangle_z$ given above), however there would be an unknown offset. The application of the QFT will transform the separation into N/r (which is not a problem, as we will see below) and, crucially, set the offset to zero, allowing us to determine the period by two measurements with a certain probability.



Let $|\tilde{\Phi}_1\rangle_z$ be the QFT^{-1} of $|\Phi_1\rangle_z$

$$|\tilde{\Phi}_1\rangle_z = \sqrt{\frac{r}{N^2}} \sum_{t=0}^{N/r-1} \sum_{y=0}^{N-1} w^{-(x_0+t\cdot r)\cdot y} |y\rangle$$

$$= \sqrt{\frac{r}{N^2}} \sum_{y=0}^{N-1} w^{-x_0\cdot y} \underbrace{\sum_{t=0}^{N/r-1} w^{-t\cdot r\cdot y}}_{c_y} |y\rangle.$$

Note that

$$c_y = \frac{N}{r} \quad \text{if } r \cdot y \text{ is a multiple of } N.$$

As we will see $c_y = 0$ in all other cases. In order to see this note that

$$\Pr[y] = \frac{r}{N^2} |w^{-x_0\cdot y}|^2 c_y^2$$

$$= \frac{1}{r}, \quad \underbrace{\text{if } r \cdot y \text{ is a multiple of } N}_{\Leftrightarrow\ y \text{ is a multiple of } \frac{N}{r}}$$

[8] If this would not be satisfied it it would not be a problem. The analysis would still work, but one would need to take care of the deviations. These deviations can be kept small by making the circuit (i.e., N) larger.

There are r different values y of this type. Therefore, those probabilities sum to one and hence, all other probabilities must be equal to zero.

Conclusion: With probability 1 the circuit outputs

$$s = \mathbb{N}\frac{N}{r}.$$

Important: This conclusion holds independently of the value z. However, we are not there yet – remember that the goal was to find r.

The idea is now to run the algorithm twice and record the output divided by N. We get

$$\frac{k}{r} \text{ and } \frac{k'}{r} \qquad k, k' \in \mathbb{N},$$

where k and k' are unknown. The task is to extract r from $\frac{k}{r}$ and $\frac{k'}{r}$.

Assumption: k and k' are coprime[9], i.e., they share no common prime factor (we will argue that this assumption will be met with high probability).

Note that it follows from number theory that all $x \in \mathbb{N}$ have a unique decomposition into powers of primes

$$x = \prod_i p_i^{k_i},$$

where p_i are distinct prime numbers and k_i their respective multiplicites.

Step 1: Write $\frac{k}{r}$ and $\frac{k'}{r}$ as simplified fractions[10]

$$\frac{k}{r} = \frac{a}{b}, \quad \frac{k'}{r} = \frac{a'}{b'}. \tag{$\circ$}$$

Step 2: Let $\bar{r} = \text{lcm}(b, b')$ be the lowest common multiple of b and b'.[11]

Claim: $\bar{r} = r$.

Proof. Let b be represented by its prime factors

$$b = \prod_i p_i^{k_i}.$$

By assumption ($\circ$) none of the factors p_i divide a (otherwise the fraction could be simplified by dividing by p_i)

$$p_i \nmid a \quad \forall\, p_i.$$

We also have

$$b \cdot k = a \cdot r$$

[9] This is not a restriction on the functionality of the algorithm. The idea is that the algorithm can be run several times until this assumption is satisfied.

[10] This can be done efficiently using Euclid's algorithm.

[11] Also this can be calculated efficiently using Euclid's algorithm.

and because all $p_i^{k_i}$ divide the lhs they must also divide the rhs. Because they cannot divide a they divide r

$$p_i^{k_i} \big| r \quad \forall\, p_i^{k_i}.$$

Therefore r is a multiple of b. By the same reasoning it follows that r is a multiple of b'. Therefore r is a common multiple of b and b'. It remains to show that it is the lowest.

Consider now the decomposition into primes of r

$$r = \prod_i \tilde{p}_i^{m_i},$$

because $a \cdot r = b \cdot k$ and $a' \cdot r = b' \cdot k'$ any $\tilde{p}_i^{m_i}$ must divide $b \cdot k$ and $b' \cdot k'$. Remembering that k and k' are coprime it follows that all $\tilde{p}_i^{m_i}$ must divide either b or b' (or both of them). Therefore r is the lowest common multiple of b and b' (if r would not be the lowest common multiple then we could divide r by a factor that is not contained in neither b nor b').

□

Now we want to discuss that it is not a problem that we assumed that k and k' are coprime.

Claim: In any run of the algorithm the assumption that k and k' are coprime is satisfied with probability of at least 0.35.

Note that if this claim is true then we can simply run the algorithm several times and the success probability converges exponentially.

Proof. Consider the probability that the assumption is wrong, i.e.,

$$\begin{aligned}\Pr[k \text{ and } k' \text{ share at least one prime}] &\leq \sum_{p \in \text{Primes}} \Pr[p \big| k \text{ and } p \big| k'] \\ &\leq \sum_{p \in \text{Primes}} \Pr[p \big| k'] \cdot \Pr[p \big| k']\end{aligned}$$

where the last equality holds because k and k' are obtained from independent runs of the algorithm.

Assume that k and k' are positive (otherwise run the algorithm again). Remember now that the algorithm provides the values $\frac{k}{r}$ each with probability $\frac{1}{r}$. There are r values in total and therefore $\frac{r}{p}$ values that are divided by p (for a fixed prime number p), hence

$$\begin{aligned}\Pr[k \text{ and } k' \text{ share at least one prime}] &\leq \sum_{p : p \text{ is prime}} \frac{1}{p} \cdot \frac{1}{p} \\ &\leq \sum_{p=2}^{\infty} \frac{1}{p^2} = \frac{\pi^2}{6} - 1 \leq 0.65\end{aligned}$$

□

Shor's factoring algorithm

Shor's factoring algorithm [2] is really an application of this quantum algorithm for period finding. The idea is to use the following Lemma from number theory.

Lemma 3.0.2. *Let $M = p \cdot q$, where p and q are primes. Let a be chosen at random and let r be the period of the function*

$$f : \; x \mapsto a^x \mod M,$$

then

$$gcd(M, a^{r/2} - 1) = p \qquad \textit{with probability } \frac{1}{4}.$$

Summary: To factor numbers one uses period finding to find the period of the above function and then uses this Lemma.

Chapter 4

Quantum error correction

We will first introduce some mathematical concepts that will be applied in the following.

Choi-Jamilkowski Isomorphism

Claim: The set of TPCPMs from A to B is isomorphic to the set of density operators of the form

$$\{\rho_{AB} : \mathrm{tr}_B(\rho_{AB}) = \frac{I_A}{d_A}\}$$

where $d_A = \dim(A)$.

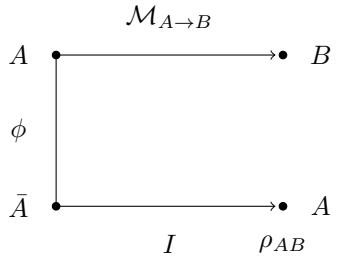
The isomorphism maps any TPCPM $\mathcal{M}_{A\to B}$ to the state

$$\mathcal{M}_{A\to B} \mapsto \rho_{AB} = (\mathcal{M}_{A\to B} \otimes I_{\bar{A}\to A})(|\phi\rangle\langle\phi|_{A\bar{A}}) \qquad (\dagger)$$

where $\mathcal{M}_{A\to B}$ is a TPCPM and

$$|\phi\rangle_{A\bar{A}} = \frac{1}{d_A}\sum |i\rangle_A |i\rangle_{\bar{A}} \qquad (\dagger\dagger)$$

is the maximally entangled state on $A\bar{A}$ for $\bar{A} \cong A$ w.r.t. the basis $\{|i\rangle_A\}$.



ρ_{AB} is called Choi-Jamilkowski representation of $\mathcal{M}_{A\to B}$.

Proof. 1. $\mathrm{tr}_B(\rho_{AB}) = \frac{I_A}{d_A}$:

$$\begin{aligned}\mathrm{tr}_B(\rho_{AB}) &= \mathrm{tr}_B\left((\mathcal{M}\otimes I_{\bar{A}\to A})|\phi\rangle\langle\phi|_{A\bar{A}}\right)\\ &= I_{\bar{A}\to A}\mathrm{tr}_A(|\phi\rangle\langle\phi|_{A\bar{A}})\\ &= I_{\bar{A}\to A}\frac{I_{\bar{A}}}{d_{\bar{A}}}\\ &= \frac{I_A}{d_A}\end{aligned}$$

Where we used that TPCPMs commute with the trace.

2. Invertibility:

   Claim: The inverse of the isomorphism maps

   $$\rho_{AB} \mapsto \mathcal{M}_{A\to B}$$

   with

   $$\mathcal{M}_{A\to B}(\sigma_A) = d_A\cdot \mathrm{tr}_A(\rho_{AB}(I_B\otimes\sigma_A^T))$$

   where the transpose is w.r.t. the basis $\{|i\rangle_A\}$.

The prove that this is indeed the inverse of the isomorphism (†) we will show that

$$d_A\cdot\mathrm{tr}_{\bar{A}}(|\phi\rangle\langle\phi|_{A\bar{A}}(I_A\otimes\sigma_{\bar{A}}^T)) = \sigma_A \qquad \forall \quad \sigma_A \tag{$\star$}$$

holds, which is easy to verify by an explicit calculation as follows:

Because any operator can be written as linear combinations of operators of the the form $\sigma_{\bar{A}} = |i\rangle\langle j|$ we prove it for such operators.

$$\begin{aligned}d_A\cdot\mathrm{tr}_{\bar{A}}(|\phi\rangle\langle\phi|_{A\bar{A}}(I_A\otimes|j\rangle\langle i|_{\bar{A}})) &= d_A\cdot\sum_{k,k'}\frac{1}{d_A}\mathrm{tr}_{\bar{A}}(|k\rangle_A|k\rangle_{\bar{A}}\langle k'|_A\langle k'|_{\bar{A}}(I_A\otimes|j\rangle\langle i|_{\bar{A}}))\\ &= \delta_{k'j}\delta_{ik}|k\rangle\langle k'|_A\\ &= |i\rangle\langle j|_A\end{aligned}$$

Insert now

$$\rho_{AB} = (\mathcal{M}\otimes I)|\phi\rangle\langle\phi|_{A\bar{A}}$$

into

$$\mathcal{M}_{A\to B}(\sigma_A) = d_A\cdot\mathrm{tr}_A(\rho_{AB}(I_B\otimes\sigma_A^T))$$

yielding

$$\begin{aligned}\mathcal{M}_{A\to B}(\sigma_A) &= d_A\cdot\mathrm{tr}_A((\mathcal{M}\otimes I)|\phi\rangle\langle\phi|_{A\bar{A}}(I_B\otimes\sigma_A^T))\\ &= d_A\cdot\mathcal{M}_{A\to B}\circ\mathrm{tr}_{\bar{A}}(|\phi\rangle\langle\phi|_{A\bar{A}}(I_B\otimes\sigma_{\bar{A}}^T))\\ &= \mathcal{M}_{A\to B}(\sigma_A)\end{aligned}$$

where we used Equation ($\star$) in the last step.

□

Stinespring dilation

Define for a TPCPM $\mathcal{E}$ with Kraus operators $\{E_k\}$ the Stinspring-Dilation $V_{A\to BE}$ as

$$V_{A\to BE} = \sum_k E_k \otimes |k\rangle_E$$

where $\{|k\rangle_E\}$ is an orthonormal basis of the environment E. Note that $V_{A\to BE}$ generates a new state on the environment E.

In the following we will simply write V for $V_{A\to BE}$

Claim: V is an isometry ($V^\dagger V = I$) such that

$$\mathrm{tr}_E(V\rho V^\dagger) = \mathcal{E}(\rho).$$

Proof. 1. V is an isometry:

$$\begin{aligned} V^\dagger V &= \sum_{kk'} \left(E_k^\dagger \otimes \langle k|\right)\left(E_{k'} \otimes |k'\rangle\right) \\ &= \sum_k E_k^\dagger E_k \\ &= I \end{aligned}$$

2.

$$\begin{aligned} \mathrm{tr}_E(V\rho V^\dagger) &= \mathrm{tr}_E\left(\sum_{kk'} E_k \otimes |k\rangle\rho E_{k'}^\dagger \otimes \langle k'|\right) \\ &= \sum_k E_k \rho E_k^\dagger \end{aligned}$$

□

The Stinespring representation tells us that it is a matter of perspective whether we say "quantum theory is unitary and reversible" and/or "a quantum measurement is non-unitary". It depends on whether we take both systems B and E into account (first view) or just consider the system B we measure (second view).

Note that the dimension of E can be bounded by $d_A \cdot d_B$ (see exercises).

Purifications

Given a density operator ρ_S on S there exists a pure density operator

$$\bar{\rho}_{S\bar{S}} = |\phi\rangle\langle\phi|_{S\bar{S}}$$

such that

$$\mathrm{tr}(\bar{\rho}_{S\bar{S}}) = \rho_S.$$

Given two purifications $\bar{\rho}_{S\bar{S}}$ and $\bar{\rho}'_{S\bar{S}}$ it can be shown that there exists a unitary

$$U_{S\bar{S}'\to S\bar{S}} = I_S \otimes U_{\bar{S}'\to\bar{S}}$$

acting only on $\bar{S}'$ s.t.

$$\bar{\rho}_{S\bar{S}} = U\bar{\rho}'_{S\bar{S}}U^\dagger.$$

Let us now apply these concepts to quantum error correction.

$$\underset{\substack{\text{logical}\\\text{information}}}{L} \xrightarrow{\text{coding } C} \underset{\substack{\text{physical}\\\text{information}}}{P} \xrightarrow{\text{errors } \mathcal{E}} \underset{\substack{\text{description}\\\text{after}\\\text{errors}}}{P'} \xrightarrow{\text{decoding } D} L'$$

$\mathcal{E}$: mapping that describes error model

For the formal treatment it is useful to think of all these mappings as trace preserving completely positive maps (TPCPMs), which is the most general way to describe valid transformations of quantum states.

The total transformation

$$T = D \circ \mathcal{E} \circ C$$

is a TPCPM and the decoding is successful if T describes the identity transformation

$$T_{L\to L'} = I_{L\to L'}.$$

Let us equip the logical space L with a ONB $\{|i\rangle\}_{i=0,\ldots,d-1}$ called the computational basis. Let $|\varphi_i\rangle$ be the encoded vector corresponding to $|i\rangle$:

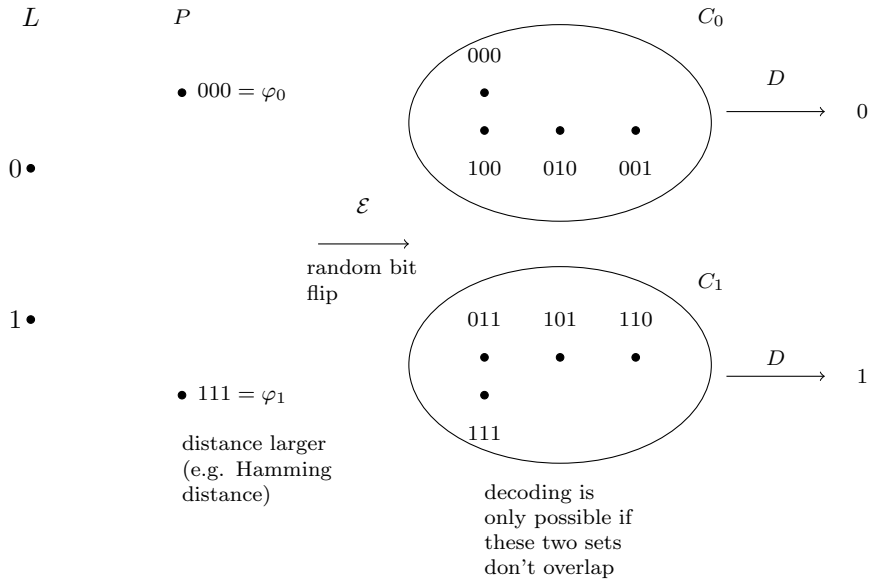
$$C : |i\rangle \to |\varphi_i\rangle \in P.$$

Example. *Repetition code*
This code can correct bit flip but not phase flip errors.

$$|0\rangle \longrightarrow |\varphi_0\rangle = |000\rangle$$

$$|1\rangle \longrightarrow |\varphi_1\rangle = |111\rangle$$

The map $\mathcal{E}$ should model realistic errors as accurately as possible. Because the errors are not under our control the challenge is to find a coding scheme (a combination of the maps C and D) such that $T = I$ for a large class of error models that hopefully include all "physical" errors.



In the diagonal basis

$$|\bar{0}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$$
$$|\bar{1}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right).$$

a phase flip error acts as

$$|\bar{0}\rangle \longrightarrow |\bar{1}\rangle$$
$$|\bar{1}\rangle \longrightarrow |\bar{0}\rangle.$$

Noting that

$$|000\rangle = \frac{1}{2}\left(|\bar{0}\bar{0}\bar{0}\rangle + |\bar{1}\bar{1}\bar{0}\rangle + |\bar{1}\bar{0}\bar{1}\rangle + |\bar{0}\bar{1}\bar{1}\rangle\right)$$
$$|111\rangle = \frac{1}{2}\left(|\bar{1}\bar{0}\bar{0}\rangle + |\bar{0}\bar{1}\bar{0}\rangle + |\bar{0}\bar{0}\bar{1}\rangle + |\bar{1}\bar{1}\bar{1}\rangle\right)$$

we can see that a phase flip error may lead to states that are not orthogonal anymore. For example a phase flip on any of the qubits of the $|000\rangle$ state creates an overlap with the $|111\rangle$ state.

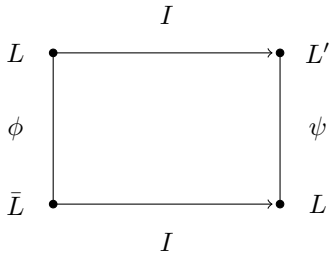
We will see that a necessary condition for a good code is that the spaces C_i on which the errors map the physical states $|\varphi_i\rangle$ are mutually orthogonal.

Remember now that the goal was to find C and D such that

$$T = I_{L\to L'}.$$

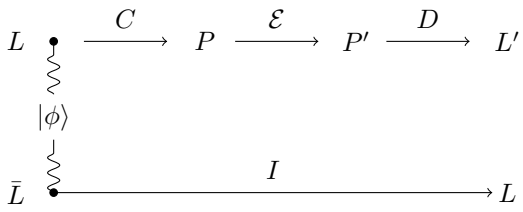
The Choi-Jamilkowski representation of the identity $I_{L\to L'}$ is given by

$$\rho_{LL'} = (I_{L\to L'} \otimes I_{\bar{L}\to L})(|\phi\rangle\langle\phi|_{L\bar{L}})$$
$$= |\phi\rangle\langle\phi|_{L'L}$$



Claim: The coding scheme is successful ($T = I$) if and only if

$$\psi = ((D \circ \mathcal{E} \circ C) \otimes I)|\phi\rangle\langle\phi| = |\phi\rangle\langle\phi|_{L'\bar{L}}.$$

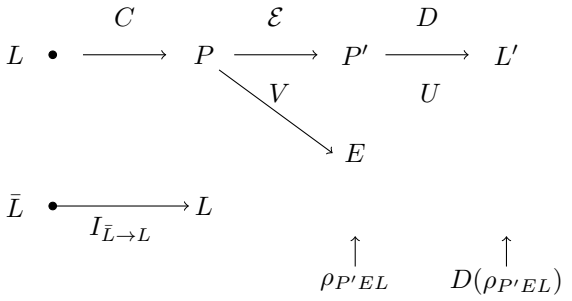


Let now V be Stinespring dilation of $\mathcal{E}$, i.e.,

$$\mathcal{E}(\sigma) = \mathrm{tr}_E(V\sigma V^\dagger).$$

Analogously we define U as the Stinespring dilation of D

$$D(\sigma) = \mathrm{tr}_{\bar{E}}(U\sigma U^\dagger).$$



The goal then rephrases to

$$\underbrace{D(\rho_{P'EL})}_{:=\rho'_{L'EL}} = |\phi\rangle\langle\phi|_{L'L} \otimes \sigma_E. \qquad (\triangle)$$

If we trace out L' we get the following *necessary* condition

$$\rho'_{EL} = \frac{I_L}{d_L} \otimes \sigma_E$$

And because D only acts on P': $\rho'_{EL} = \rho_{EL}$. Hence

$$\rho_{EL} = \frac{I_L}{d_L} \otimes \sigma_E. \qquad (\triangle\triangle)$$

This is surprising: It is a condition that no longer depends on how we encode and decode.

How can we obtain a *sufficient* condition? The idea is to use purifications. Consider a purification of $\rho_{E\bar{L}}$, call it

$$\sigma_{\bar{L}L'E\bar{E}} = |\phi\rangle_{L'\bar{L}} \otimes |\sigma\rangle\langle\sigma|_{E\bar{E}},$$

where $|\sigma\rangle_{E\bar{E}}$ is a purification of σ_E. This is indeed a purification of $\rho_{E\bar{L}}$ because $\mathrm{tr}_{L'\bar{E}}(\sigma_{\bar{L}L'E\bar{E}}) = \frac{I_{\bar{L}}}{d_{\bar{L}}} \otimes \sigma_E$.

It can be verified by a straightforward calculation the the state $\rho_{P'\bar{L}E}$ is pure by construction

$$\rho_{P'\bar{L}E} = |\psi\rangle\langle\psi|_{P'\bar{L}E} \tag{4.1}$$

where

$$|\psi\rangle_{P'\bar{L}E} = \frac{1}{\sqrt{d_{\bar{L}}}} \sum_{i,j} E_j |\varphi_i\rangle_{P'} |i\rangle_{\bar{L}} |j\rangle_E.$$

Hence, if ($\triangle\triangle$) holds, then both $\rho_{P'\bar{L}E}$ and $\sigma_{\bar{L}L'E\bar{E}}$ are purifications of $\rho_{E\bar{L}}$.

Because all purifications are equivalent up to isometries on the purifying system (see before), there is an isometry $U_{P'\to L'\bar{E}}$ such that

$$U\rho_{P'\bar{L}E}U^\dagger = \sigma_{\bar{L}L'E\bar{E}}.$$

$U_{P'\to L'E}$ can be seen as a Stinespring dilation, hence if we trace out $\bar{E}$:

$$D(\rho_{P'\bar{L}E}) = \sigma_{\bar{L}L'E} = |\phi\rangle_{L'\bar{L}} \otimes \sigma_E$$

which implies ($\triangle$), where D is the TPCPM for which U is a Stinespring dilation.

We therefore proved that there is a decoding map D that retrieves all quantum information provided ($\triangle\triangle$). In other words, not only ($\triangle$) $\Rightarrow$ ($\triangle\triangle$) but in fact ($\triangle$) $\Leftrightarrow$ ($\triangle\triangle$) holds.

The condition ($\triangle\triangle$) can be rewritten as follows. Tracing out P' in $\rho_{P'\bar{L}E}$ we get from (4.1)

$$\rho_{\bar{L}E} = \frac{1}{d_{\bar{L}}} \sum_{j,j',i,i'} \langle\varphi_{i'}|E_{j'}^\dagger E_j|\varphi_i\rangle |j\rangle\langle j'| \otimes |i\rangle\langle i'|.$$

Condition ($\triangle\triangle$) can be rewritten as

$$\rho_{\bar{L}E} = \frac{1}{d_{\bar{L}}} \sum_{j,j',i,i'} a_{jj'}\delta_{ii'} |j\rangle\langle j'| \otimes |i\rangle\langle i'|,$$

where $a_{jj'}$ defines the hermitian matrix

$$\sigma_E = \sum_{jj'} a_{jj'} |j\rangle\langle j'|.$$

Comparing coefficients we find

$$a_{jj'}\delta_{ii'} = \langle\varphi_{i'}|E_{j'}^{\dagger}E_j|\varphi_i\rangle \qquad (\square)$$

which is known as the *Knill-Laflamme* condition. In the next Section we will see its application to stabilzer error correcting codes.

Stabilizer quantum error correcting codes

We will now discuss a handy set of tools for describing a class of quantum error correcting codes called *stabilizer formalism*. Here subspaces are not characterised by a set of basis states but instead with a set of operators such that the subspace is an eigenspace. Within coding the relevant subspace of an n-dimensional Hilbert space is the space of the physical qubits. One of the main advantages is that error correcting codes can be described by fewer parameters within the stabilizer formalism (compared to the state vector representation).

Motivation

Suppose we have a set of states $|\psi_i\rangle$ which are $+1$ eigenstates of a hermitian operator S, $S|\psi_i\rangle = |\psi_i\rangle$. Further suppose that T is an operator which anti-commutes with S, $ST = -TS$. Then it is easy to see that

$$S(T|\psi_i\rangle + = -TS|\psi_i\rangle = -(T|\psi_i\rangle).$$

Thus the states $T|\psi_i\rangle$ are -1 eigenstates of S. Since the main idea of quantum error correction is to detect when an error has occurred on a code space, such pairs of operators S and T can be used in such a manner: if we are in the $+1$ eigenvalue subspace of S, then an error of T will move a state to a -1 eigenvalue subspace of S: we can detect that this error has occurred.

An example of this is the bit flip code. Here the code subspace is spanned by $|000\rangle$ and $|111\rangle$, which are eigenstates of $S_1 = Z \otimes Z \otimes I$ and $S_2 = Z \otimes I \otimes Z$. Both S_1 and S_2 anticommute with $X \otimes I \otimes I$.

More generally suppose that we have a set of unitary operators S_i such that our code space is defined by $S_i|\psi\rangle = |\psi\rangle$ for $|\psi\rangle$ in the code subspace. Further we have errors E_j such that the products $E_{j'}^{\dagger}E_j$ always anticommute with at least one S_i. Then the Knill-Laflamme condition ($\square$) can be satisfied, because for the particular S_i we get

$$\begin{aligned}\langle\varphi_{i'}|E_{j'}^{\dagger}E_j|\varphi_i\rangle &= \langle\varphi_{i'}|E_{j'}^{\dagger}E_jS_i|\varphi_i\rangle \\ &= -\langle\varphi_{i'}|S_iE_{j'}^{\dagger}E_j|\varphi_i\rangle \\ &= -\langle\varphi_{i'}|E_{j'}^{\dagger}E_j|\varphi_i\rangle \\ \Rightarrow \langle\varphi_{i'}|E_{j'}^{\dagger}E_j|\varphi_i\rangle &= 0.\end{aligned}$$

The idea is therefore, to define the code states as $+1$ eigenstates of some unitary operators S_i: If the products $E_{j'}^{\dagger}E_j$ anti-commute with at least one S_i then we have a valid code. But what should we use for the S_is?

The Pauli and Stabilizer groups

Recall the definition of a group.

Definition. *A group is a set of objects $\mathcal{G}$ with a binary operation called multiplication such that*

- $g_1 \cdot g_2 \in \mathcal{G}, \quad \forall g_1, g_2 \in \mathcal{G}$
- $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$
- $\exists\ e \in \mathcal{G}$ *such that* $\forall\ g \in \mathcal{G},\ g \cdot e = g$
- $\forall\ g \in \mathcal{G}\ \exists\ g' \in \mathcal{G}$ *such that* $g \cdot g' = e$ *(inverse)*

Recall that the Pauli group operators on a single qubit are $\{I, X, Y, Z\}$.

The representation of the *Pauli group* $\mathcal{P}_n$ acting on n qubits is formed by elements of the form

$$i^k P_1 \otimes P_2 \otimes \dots P_n$$

where $k \in \mathbb{N}$ and $P_i \in \{I, X, Y, Z\}$.

The *stabilizer group* $\mathcal{S}$ is a subgroup of $\mathcal{P}_n$ which has elements which all commute with each other and which does not contain the element $-I$. An example of a stabilizer group on three qubits is the group with the elements $\mathcal{S} = \{I^{\otimes 3}, ZZI, ZIZ, IZZ\}$. We usually don't specify all of the elements of the stabilizer group. Instead we specify a minimal set of *generators*. A set of generators is a set of elements of the group such that multiplication of these generators leads to the full group. A minimal set of such generators is a set of generators of minimal size. In the previous example $\mathcal{S}$ is generated by ZZI and ZIZ. We write this fact as $\mathcal{S} = \langle ZZI, ZIZ \rangle$. For a stabilizer group $\mathcal{S}$ we write a set of minimal generators as $S_1, S_2, \dots S_r$.

Stabilizer Subspace $\mathcal{H}_s$

Given a stabilizer group $\mathcal{S}$ we define the subspace of all states $|\psi\rangle$ which satisfy $S|\psi\rangle = |\psi\rangle$ for all generators S_i. One of the reasons why such a stabilizer subspace is nice is that instead of specifying the states of the subspace we can just specify the generators of the stabilizer group.

Now if $\{E_j\}$ is a set of Pauli group errors, we can satisfy the Knill-Laflamme condition ($\square$) for those errors such that the product $E_{j'}^\dagger E_j$ anti-commute with at least one of the generators S_i. If these elements are themselves elements of the stabilizer: $E_{j'}^\dagger E_j \in \mathcal{S}$, then this is also no problem because in this case

$$\langle \varphi_{i'} | E_{j'}^\dagger E_j | \varphi_i \rangle = \delta_{i'i}$$

holds. Thus, if for an error set $\{E_j\}$ all of the products $E_{j'}^\dagger E_j$ either anticommute with the generators or are elements of the stabilizer, then this set satisfies the Knill-Laflamme condition.

For $\mathcal{S} = \langle ZZI, ZIZ \rangle$ we can consider for example the set of errors $\{III, XII, IXI, IIX\}$.

Dimension of $\mathcal{H}_S$

We will now show that the dimension of the stabilizer subspace with generators $S_1,\ S_2, \ldots, S_r$ is equal to

$$\dim(\mathcal{H}_S) = 2^{n-r}.$$

In order to see this we start with S_1. First note that $S_1^2 = I$ (because $-I \notin \mathcal{S}$) from which it follows that S_1 has eigenvalues ± 1. Further the trace is given by

$$\begin{aligned}\mathrm{tr}(S_1) &= \mathrm{tr}(i^k P_1 \otimes \ldots \otimes P_n)) \\ &= \Pi_i \mathrm{tr}(P_i) \\ &= 0\end{aligned}$$

because at least one of the $P_i \neq I$ if $\mathcal{S}$ is not the trivial group containing only the identity. Because the trace is given by the sum of the eigenvalues, it follows that 2^{n-1} of the eigenvalues must be equal to $+1$ and 2^{n-1} of the eigenvalues must be equal to -1. Therefore, S_1 splits the 2^n dimensional space into half.

As we will see in the following each S_i splits the previous subspace where $S_1, \ldots, S_{i-1}$ have eigenvalues $+1$ into half. In order to see this consider the projector onto the subspace where $S_1, \ldots, S_{i-1}$ have eigenvalues $+1$

$$\frac{1}{2^{i-1}}(I + S_1)\ldots(I + S_{i-1}).$$

The trace of S_i projected onto this subspace vanishes

$$\mathrm{tr}(\frac{1}{2^{i-1}}(I + S_1)\ldots(I + S_{i-1})S_i) = 0.$$

This can be seen be using the linearity of the trace and then using as before that the trace of a tensor product of generators is equal to the product of the individual traces and using that none of the generators is equal to the identity. Therefore the eigenvalues of S_i on this subspace must sum to 0 and therefore, there must be an equal number of $+1$ and -1 eigenvalues. Thus, there are $\frac{1}{2}2^{n-(i-1)} = 2^{n-i}$ $+1$ eigenvalues left as claimed.

Logical operators for $\mathcal{H}_S$

So far we have seen how we can define the coding space given a set of generators $S_1, \ldots, S_r$. The questions is now: What should we use as code words? These are defined as the $+1$ eigenstates of the generators and may look very complicated (see exercises). Instead of specifying the code states $|\bar{0}\rangle$ and $|\bar{1}\rangle$ we therefore specify the logical operators $\bar{X}$ and $\bar{Z}$ that act on the code states like the Pauli operators. It can be shown that $\bar{X}$ and $\bar{Z}$ determine $|\bar{0}\rangle$ and $|\bar{1}\rangle$ uniquely.

Example: 9-qubit Shor code

The Shor code has the following state vector representation

$$|\bar{0}\rangle = (|000\rangle + |111\rangle)^{\otimes 3}$$

$$|\bar{1}\rangle = (|000\rangle - |111\rangle)^{\otimes 3}.$$

Note that in this case the representation is rather simple. However, for a general 9-qubit code the state vector representation would require the specification of $2^9 = 512$ amplitudes.

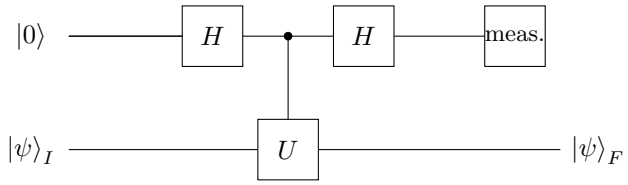
The stabilizer representation of the Shor code is summarized in the following table.

element	operator
S_1	$ZZIIIIIII$
S_2	$ZIZIIIIII$
S_3	$IIIZZIIII$
S_4	$IIIZIZIII$
S_5	$IIIIIIZZI$
S_6	$IIIIIIZIZ$
S_7	$XXXXXXIII$—
S_8	$XXXIIIXXX$
$\bar{X}$	$XXXXXXXXX$
$\bar{Z}$	$ZZZZZZZZZ$

State preparation

As the codewords are the $+1$ eigenstates of each generator S_i we need a method to project qubits onto these eigenstates.

Consider the following circuit that allows to project an arbitrary state $|\psi\rangle_I$ onto the ± 1 eigenstate of an unitary, hermitian operator $U = U^\dagger$.

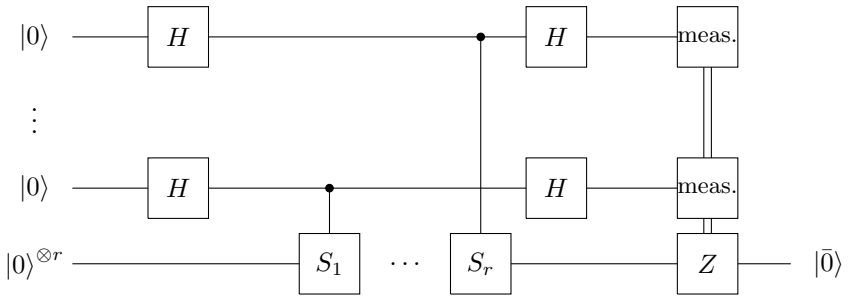


The circuit implements the following transformation

$$\begin{aligned}
|0\rangle|\psi\rangle_I &\to (|0\rangle + |1\rangle)|\psi\rangle_I \\
&\to |0\rangle|\psi\rangle_I + |1\rangle U|\psi\rangle_I \\
&\to (|\psi\rangle_I + U|\psi\rangle_I)|0\rangle + (|\psi\rangle_I - U|\psi\rangle_I)|1\rangle
\end{aligned}$$

Therefore, if we get outcome 0 we have an eigenstate with eigenvalue $+1$ and if we get outcome 1 we have an eigenstate with outcome -1.

This idea is applied for state preparation for stabilizer quantum error correction with generators $S_1, \ldots, S_r$ by the following circuit.

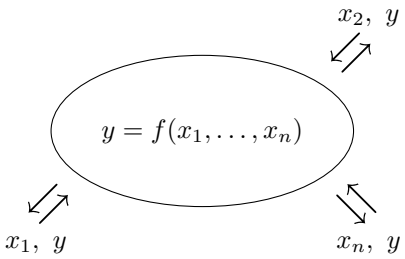


Here Z is a single qubit Z gate that transforms a -1 eigenstate into a $+1$ eigenstate.

Chapter 5

Bit commitment

Secure multiparty computation is one among many other applications of QIP.[1]



Secure function evaluation is the task in which n parties connected by communication channels want to compute a function

$$f:\ (x_1, \ldots, x_n) \mapsto y,$$

where each party provides an input x_i and receives output y, such that the following requirements are satisfied:

Correctness: $y = f(x_1, \ldots, x_n)$

Secrecy: No party learns more about $x_1,\ x_2, \ldots, x_n$ than what is implied by $f(x_1, \ldots, x_n)$ and the partyÕs input x_i.

A specific application of such a protocol is *electronic voting*. Here the input $x_i \in \{0, 1\}$ is the vote of the i-th party and

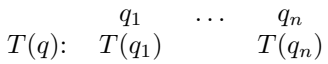
$$f(x_1, \ldots, x_n) = \sum_i x_i.$$

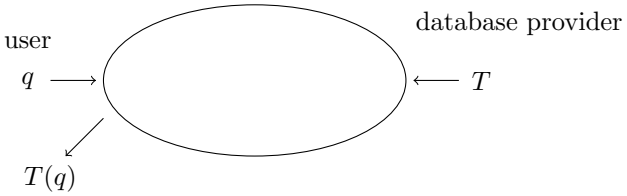
This example also illustrates that it is necessary to include in the secrecy requirement "what is implied by f and x_i. For example for $n = 2$ it is obvious that each party knows the vote of the other party in the end.

Another example is *secure data base search*, where we don't want the database

[1] Other applications include algorithms, metrology, thermodynamics and many more.

provider to know what we ask. The database corresponds to a table

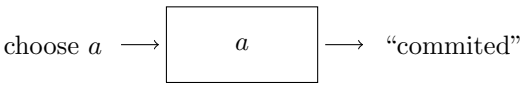




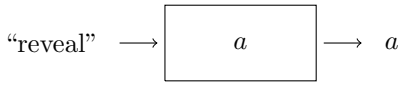
Rather than inventing a new protocol for each application, one develops methods for certain building block functions (*primitives*) from which the evaluation can be built. An example of such a primitive is *bit commitment*. It can be proved that bit commitment is *universal* for quantum protocols, i.e., any secure multiparty computation can be implemented from it. For purely classical protocols bit commitment is not universal but a stronger primitive called *oblivious transfer* is needed.

The protocol has two phases:

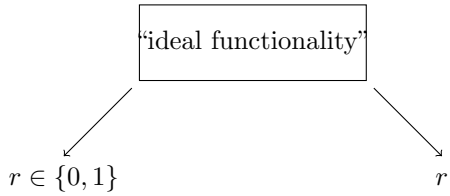
Commit phase:



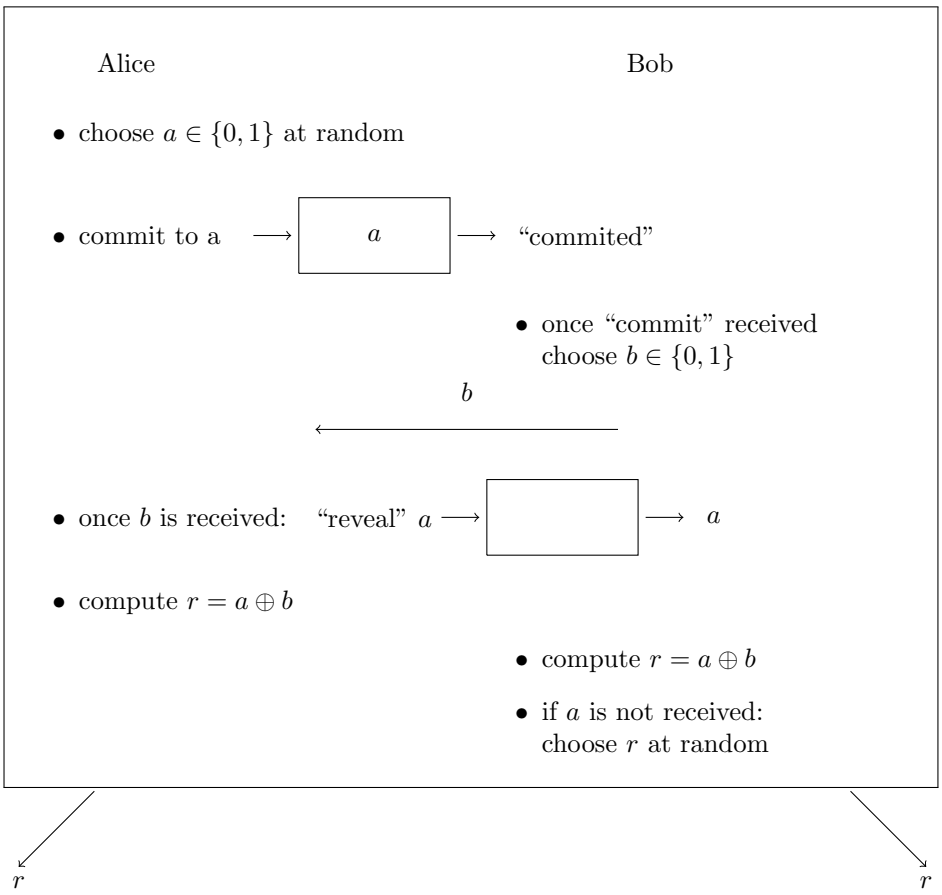
Reveal phase:



An application of bit commitment is *coin tossing*: Here two distrusting parties want to generate a random bit.



The following protocol achieves the ideal functionality.



Classical computationally secure protocol for bit commitment

The idea is to use a *hash function*

$$f:\ \{0,1\}^n \to \{0,1\}^{n'}, \qquad n' \ll n$$

which is *collision resistant*, i.e., it is computationally hard to find $x \neq x'$ with $f(x) = f(x')$. Let us furthermore assume that each output y has approximately equally many pre-images (this property is needed later).

- Commit phase: Alice chooses random bits $r_1,\ r_2, \ldots, r_{n-1}$ computes

  $$m = f(a, r_1,\ r_2, \ldots, r_{n-1})$$

  and sends m to Bob.

- Reveal phase: Alice sends $a, r_1,\ r_2, \ldots, r_{n-1}$ to Bob and he checks if $f(a, r_1,\ r_2, \ldots, r_{n-1}) = m$.

The protocol is hiding because for any inputs $a, r_1,\ r_2, \ldots, r_{n-1}$ there exits a' and suitably chosen $r'_1,\ r_2, \ldots, r'_{n-1}$ such that

$$f(a', r'_1,\ r_2, \ldots, r'_{n-1}) = f(a, r_1,\ r_2, \ldots, r_{n-1})$$

(using the "pre-image property" mentioned above). Therefore the protocol is information theoretically hiding (even if Bob has arbitrary computing power).

That the protocol is also binding follows from the fact that for given $a, r_1, r_2, \ldots, r_{n-1}$ it is hard to find $a', r'_1, r'_2, \ldots, r'_{n-1}$ such that

$$f(a', r'_1, r'_2, \ldots, r'_{n-1}) = f(a, r_1, r_2, \ldots, r_{n-1})$$

(because f is a collision resistant hash function). Therefore the protocol is computationally binding.

Quantum protocol for bit commitment

In the exercises we discussed the BB84 quantum protocol for bit commitment. In the commit phase, Alice chooses her bit $a \in \{0, 1\}$ and generates n random numbers $r_1, .., r_n \in \{0, 1\}$. She then prepares n qubits in states $|(r_1)_a\rangle, .., |(r_n)_a\rangle$, where a determines the basis and r_i the basis element for qubit i such that

$\|(r_i)_a\rangle$	$r_i = 0$	$r_i = 1$
$a = 0$	$\|0\rangle$	$\|1\rangle$
$a = 1$	$\|+\rangle$	$\|-\rangle$

and sends them to Bob.

Upon receiving the n qubits Bob measures each of them randomly in one of the two bases $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ and keeps the outcomes in a table together with the basis of the corresponding measurement.

According to the protocol in the reveal phase Alice sends a together with the random numbers $r_1, .., r_n$. Bob then checks in his table whether the outcomes coincide with r_i for those measurements which he carried out in the basis corresponding to a. If this is the case he accepts, otherwise he rejects.

The protocol is information theoretically hiding because the Bob's local state is identical in both cases $a = 0$ and $a = 1$. However, it is binding only under the assumption that Alice cannot coherently store quantum information.

As it was shown in the exercises, Alice can change her bit after the commit phase by preparing EPR pairs and storing one of the systems. Here we show that secure bit commitment is generally impossible, if Alice has the possibility to store quantum information. Consider the following generic bit commitment protocol:

- Depending on a Alice chooses $\rho_M^{a=0}$ or $\rho_M^{a=1}$ $\xrightarrow{M}$ Bob stores it
- Alice sends M' that depends on a $\xrightarrow{M'}$ Bob checks consistency

Claim: Whatever the protocol is, if it has this structure it is insecure.

Let us denote the joint state of M and M' depending on by a by $\rho_{MM'}^{a=0}$ or $\rho_{MM'}^{a=1}$, corresponding to Bob's information after the protocol is executed but before he measures.

The protocol is hiding if

$$\rho^{a=0}_{MM'} = \rho^{a=1}_{MM'}, \tag{5.1}$$

because otherwise there exists a measurement such that Bob could notice the difference with some probabillity.

As we will see, this implies that the protocol cannot be binding. Assume that Alice keeps a purification of ρ_M, i.e., Alice has a system E such that ρ_{EM} is pure. Remember that all purifications ρ_{EM} or a state ρ_M are equivalent up to local isometries on E. Therefore, for two purifications ρ_{EM} and $\rho_{E'M}$ with identical ρ_M there exists an isometry $U_{E\to E'}$ such that

$$\rho_{E'M} = U_{E\to E'}\rho_{EM}U^\dagger_{E\to E'}.$$

Let now $\rho_{E_0MM'}$ and $\rho_{E_1MM'}$ be purifications of $\rho^{a=0}_{MM'}$ and $\rho^{a=1}_{MM'}$. Because they are both purifications of Bob's local state (5.1) Alice can generate each of them with a suitable chosen isometry $U_{E\to E_{0/1}M'}$ from ρ_{EM}. Therefore, the protocol cannot be binding.

Bibliography

[1] Isaac L. Chuang Michael A. Nielsen. *Quantum Computation and Quantum Information*. Cambridge University Press.

[2] Peter Shor. Algorithms for quantum computation: Discrete logarithms and factoring. *Proc. 35th Annual Symposium on Foundations of Computer Science, , IEEE Computer Society Press, Los Alamitos, CA*, pages 124–134, 1994.