University of Tartu Prof. Dr. Dominique Unruh

Quantum Cryptography

Short notes, spring 2021

Important note: These notes are not supposed to be self-contained. Instead, they are intended as a reminder about which topics were discussed in the lecture. If you find mistakes in these notes, please send them to unruh@ut.ee.

Knowlets: The identifiers in boxes in the right margin are names of "knowlets". A knowlets is a self-contained piece of knowledge. Those labels will be use to refer to knowlets, e.g., from homework sheets, etc. (Knowlets are an experiment that I try for the first time this semester.)

Contents

1 Lecture timeline

- **Lecture 2021-02-10:** Quantum systems, quantum states, unitary operations.
  Video, Whiteboard. Covered:
- **Practice 2021-02-16:** Small exercises with single qubits. Polarization invariant under rotation.
  Whiteboard. Covered:
- **Lecture 2021-02-17:** Measurements in computational basis. Elitzur-Vaidman bomb tester. Complete measurements.
  Video, Whiteboard. Covered:
- **Practice 2021-02-23:** Light filters as measurements. Improved bomb tester. Quantum Zero effect.
  Whiteboard. Covered:
- **Practice 2021-03-02:** Initializing using measurements. States from prob. distributions. Equivalent definitions of unitary.
  Whiteboard. Covered:
- **Lecture 2021-03-03:** Projective measurements. Tensor product. Composition of quantum systems / quantum states / unitaries / measurements.
  Video, Whiteboard. Covered:
- **Practice 2021-03-09:** Using tensor product and proj. measurements. Quantum teleportation.
  Whiteboard. Covered:
- **Lecture 2021-03-10:** Deutsch's algorithm. Quantum state probability distributions (ensembles). Operations on ensembles. Density operators.
  Video, Whiteboard. Covered:
- **Practice 2021-03-16:** Distinguishing incorrect toy crypto. Density operators. Creating unitaries for boolean functions.
  Whiteboard. Covered:

- **Lecture 2021-03-17:** Operations on density operators. Theorem: Physically indistinsuishable iff same density operator. Toy crypto protocol is secure. Quantum one-time pad.
  Video, Whiteboard. Covered:
- **Practice 2021-03-23:** Unitaries and measurements on density ops. Observables.
  Whiteboard. Covered:
- **Lecture 2021-03-24:** Partial trace. Quantum operations.
  Video, Whiteboard. Covered:
- **Practice 2021-03-30:** Tracing out buffer qubits in U_f. Trace as a quantum op. Replace operation.
  Whiteboard. Covered:
- **Lecture 2021-03-31:** Statistical distance. Trace distance. Short mentions not in notes: Fidelity. Optimal distinguisher
  Video, Whiteboard. Covered:
- **Practice 2021-04-06:** Trace distance of biased distributions. QOTP without 0-keys. TD between any two states.
  Whiteboard. Covered:
- **Lecture 2021-04-07:** Quantum key distribution: Intro. Security definition. Protocol overview. First step (distributing Bell pairs).
  Video, Whiteboard. Covered:
- **Practice 2021-04-13:** Prob. of measuring key after QKD. Alternate sec def of QKD. SMT from QKD.
  Whiteboard. Covered:
- **Lecture 2021-04-14:** Quantum key distribution: Bell test. Measuring the raw key.
  Video, Whiteboard. Covered:
- **Practice 2021-04-20:** Alternate sec def for QKD (continued). Measuring key with t errors.
  Whiteboard. Covered:
- **Lecture 2021-04-21:** Error correcting codes. Error correction step in QKD. Strong randomness extractors. Universal hash functions. Privacy amplification in QKD. Finished QKD security proof.
  Video, Whiteboard. Covered:
- **Practice 2021-04-27:** Last bit of key deleted or set 0. Error correction after randomness extraction. Extracting too much from a key. Problems with deterministic randomness extractors.
  Whiteboard. Covered:
- **Lecture 2021-04-28:** Shor's algorithm. Period finding. Factoring. Discrete logarithm.
  Video, Whiteboard. Covered:
- **Practice 2021-05-04:** Implementing the Quantum Fourier Transform. The von Neumann extractor.
  Whiteboard. Covered:

- **Lecture 2021-05-05:** LWE problem (computational and decisional). Regev's cryptosystem. IND-CPA security of Regev's cryptosystem.
  Video, Whiteboard. Covered:
- **Practice 2021-05-11:** Example of Regev's cryptosystem. The Short Integer Solutions problem. Collision-Resistant hash functions from SIS.
  Whiteboard. Covered:
- **Lecture 2021-05-12:** Classical/quantum zero knowledge. Difficulty with rewinding in the quantum case.
  Video, Whiteboard. Covered:
- **Practice 2021-05-18:** Aborting simulators - classical and quantum.
  Whiteboard. Covered:
- **Lecture 2021-05-19:** Quantum rewinding. Constructing a quantum ZK simulator.
  Video, Whiteboard. Covered:
- **Lecture 2021-05-26:** SchrÃ¶dinger equation. Particle in an infinite potential well.
  Video, Whiteboard. Covered:

2 Quantum systems

QState

Definition 1 (Quantum states) *An n-dimensional* quantum state *is represented by a vector* $|\Psi\rangle \in \mathbb{C}^n$ *with* $\||\Psi\rangle\| = 1$ *(here* $\mathbb{C}^n$ *is a Hilbert space).*

CompBasis

In most cases, we assume some canonical orthonormal basis of $\mathbb{C}^n$ (representing the classical possibilities of the system) which we call the *computational basis*. We then use the following convention: If $|b_1\rangle, \ldots, |b_n\rangle$ are the basis vectors, and $b_1, \ldots, b_n$ are some labels we assign to these vectors sorted according to some natural ordering (e.g., for an m-qubit system (i.e., $n = 2^m$) b_i is the bitstring $b_i \in \{0,1\}^m$ which is the binary representation of $i-1$), then $|b_i\rangle = (0, \ldots, 0, 1, 0, \ldots, 0)^t$ where the 1 is at the i-th position.

2.1 Unitary transformations

There are two kinds of operations on quantum states, unitary transformations and measurements.

UniTrafo

Definition 2 (Unitary transformation) *A* unitary transformation *on a quantum state* $|\Psi\rangle \in \mathbb{C}^n$ *is represented by a unitary matrix* $U \in \mathbb{C}^{n \times n}$*. The state after the transformation is* $U|\Psi\rangle$*.*

Important simple examples of unitary transformations are:

PauliX

Definition 3 (Bit flip) *The* bit flip *(also called* not-gate *or* X-gate *or* Pauli-X*) is defined by*

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

or equivalently

$$X|0\rangle = |1\rangle$$
$$X|1\rangle = |0\rangle$$

The bit flip corresponds to a negation. It can, however, be applied in superposition.

PauliY

Definition 4 *The* Pauli-Y matrix *(or* gate*) is defined by*

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

or equivalently

$$Y|0\rangle = i|1\rangle$$
$$Y|1\rangle = -i|0\rangle$$

PauliZ

Definition 5 *The* Pauli-Z matrix *(or* gate*) is defined by*

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

or equivalently

$$Z|0\rangle = |0\rangle$$
$$Z|1\rangle = -|1\rangle$$

Hada

Definition 6 (Hadamard) *The* Hadamard gate *(usually denoted H) is defined by*

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

or equivalently

$$H|0\rangle = \tfrac{1}{\sqrt{2}}(|1\rangle + |0\rangle)$$
$$H|1\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

The Hadamard gate is useful for introducing superpositions as it takes a classical bit ($|0\rangle$ or $|1\rangle$ and transforms it into a superposition).

Rota

Definition 7 (Rotation) *The rotation by angle θ is defined by*

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

or equivalently

$$R_\theta|0\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$$
$$R_\theta|1\rangle = -\sin\theta|0\rangle + \cos\theta|1\rangle$$

PhaseShift

Definition 8 (Phase shift) *The phase shift S is defined by*

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

or equivalently

$$S|0\rangle = |0\rangle$$
$$S|1\rangle = i|1\rangle$$

More generally, we can parametrise the phase shift by an angle θ:

$$S_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$$

or equivalently

$$S_\theta|0\rangle = |0\rangle$$
$$S_\theta|1\rangle = e^{i\theta}|1\rangle$$

Note that $S = S_{\frac{\pi}{2}}$.

Further reading: [NC00, Section 1.2.1, 1.3.1], and [NC00, Section 4.2] for the single qubit gates.

2.2 Measurements

There are many possible ways how to formalize how a quantum system is measured. (I.e., how information is extracted about a quantum state.) We will show three here, of different generality. They all have in common that a measurement is a process that operates on a quantum system in some quantum state, returns an outcome (which is a classical value that is determined according to a probability distribution that depends on the state that is measured), and leaves the quantum system in a possibly changed quantum state (because measurements affect the measured system.

CBMeas

The simplest form of measurement is a measurement in the computational basis:

Definition 9 (Measurement in the computational basis) *When* measuring *a state $|\Psi\rangle = (\alpha_1 \ \ldots \ \alpha_n)^T \in \mathbb{C}^n$ (i.e., a state in a quantum system with n classical possibilities)* in the computational basis*, the probability of getting the outcome $i \in \{1, \ldots, n\}$ is $|\alpha_i|^2$. And the post-measurement state (the state after getting outcome i) is $|i\rangle = (0 \ \ldots \ 0 \ 1 \ 0 \ \ldots \ 0)^T$. (The 1 is in the i-th position.)*

Note that the indices in the above definition do not necessarily have to be $1, \ldots, n$. E.g., if $|\Psi\rangle = (\alpha_{up} \ \alpha_{middle} \ \alpha_{down})^T$, then we get analogously a measurement with outcomes *up*, *middle*, *down*.

ComplMeas

Somewhat more general that the above is the following kind of measurement:

Definition 10 (Complete measurement) *A* complete measurement *on $\mathcal{H}$ is specified by an orthonormal basis $B = \{|i\rangle\}_{i\in I}$ of $\mathcal{H}$ labelled with the possible measurement outcomes $i \in I$.*

When measuring a state $|\Psi\rangle \in \mathcal{H}$, the outcome i occurs with probability

$$|\langle i|\Psi\rangle|^2.$$

and the corresponding post-measurement state is

$$|i\rangle$$

Or alternatively, the post-measurement state is

$$\frac{\langle i|\Psi\rangle}{|\langle i|\Psi\rangle|} \cdot |i\rangle$$

(The two alternatives are physically equivalent because there is no experiment that can notice a scalar factor $\frac{\langle i|\Psi\rangle}{|\langle i|\Psi\rangle|}$ of absolute value 1*, called a* global phase*. The latter form is more complicated but has the advantage of being compatible with the notion of a projective measurement introduced below.)*

Note that the measurement in the computational basis (Definition 9) is a special case of the complete measurement. If we set $|i\rangle := (0 \ \ldots \ 0 \ 1 \ 0 \ \ldots \ 0)^T$ (computational basis) in Definition 10, we get Definition 9.

ProjMeas

Measurements as described above have the limitation that they require us to measure the whole quantum system (i.e., get out as much information as possible). Since measurements change the quantum state, it can be important to measure less than the whole system. The following formalism models that:

Definition 11 (Projective measurement) *A* (projective) measurement *on a Hilbert space $\mathcal{H}$ is specified by a family $\{P_i\}_{i\in I}$ of orthogonal projections on $\mathcal{H}$ labelled with the possible measurement outcomes $i \in I$. The projections have to be pairwise orthogonal, i.e., $P_iP_j = 0$ for $i \neq j$. And the projections sum to* 1*, i.e., $\sum_i P_i = 1_{\mathcal{H}}$ where $1_{\mathcal{H}}$ is the identity on $\mathcal{H}$.*

When measuring a state $|\Psi\rangle \in \mathcal{H}$, the outcome i occurs with probability

$$\|P_i|\Psi\rangle\|^2.$$

If the outcome i occurs, the state after the measurement (post-measurement state) is

$$\frac{P_i|\Psi\rangle}{\|P_i|\Psi\rangle\|}.$$

Note that the complete measurement is a special case of this with $P_i := |i\rangle\langle i|$ (projector onto $|i\rangle$).

We can also formalize projective measurements differently:

ProjMeasVS

Definition 12 (Projective Measurement (alternative definition)) *A* (projective) measurement *on a Hilbert space $\mathcal{H}$ is specified by a family $\{V_i\}_{i\in I}$ of subspaces of $\mathcal{H}$ labelled with the possible measurement outcomes $i \in I$. The spaces V_i have to be pairwise orthogonal, i.e., $\langle\psi|\phi\rangle = 0$ for all $|\psi\rangle \in V_i, |\phi\rangle \in V_j, i \neq j$. And the $\sum_i V_i = \mathcal{H}$, i.e., every $|\psi\rangle \in \mathcal{H}$ is a linear combination of vectors from $\bigcup V_i$.*

For a state $|\Psi\rangle \in \mathcal{H}$, let $|\Psi_i\rangle$ be vectors with $\sum_{|\Psi_i\rangle} = |\Psi\rangle$ and $|\Psi_i\rangle \in V_i$. (The conditions on the V_i guarantee that those $|\Psi_i\rangle$ exist and are uniquely determined.)

When measuring a state $|\Psi\rangle \in \mathcal{H}$, the outcome i occurs with probability

$$\||\Psi_i\rangle\|^2.$$

If the outcome i occurs, the state after the measurement (post-measurement state) is

$$\frac{|\Psi_i\rangle}{\||\Psi_i\rangle\|}.$$

Definitions 11 and 12 are equivalent: We can convert a measurement according to Definition 11 into one according to Definition 12 by setting $V_i := \operatorname{im} P_i$. And we can convert a measurement according to Definition 12 into one according to Definition 11 by letting P_i be the orthogonal projector onto V_i.

Note that the complete measurement is a special case of the projective measurement with $V_i := \operatorname{span}\{|i\rangle\}$.

Further reading: [NC00], Section 2.2.1, 2.2.2, and 2.2.5 for states, unitary evolution, and projective measurements, respectively. Section 2.2.7 for information in the global phase.

3 Elitzur-Vaidman Bomb Tester

Bomb

A *beam splitter* is a device into which a photon can enter in two positions (call them *up* and *down*), and exit in two positions (call them *up* and *down*, too). The input to the beam splitter is a qubit that is represented as a superposition between $|\text{up}\rangle := \begin{pmatrix}1\\0\end{pmatrix}$ and $|\text{down}\rangle := \begin{pmatrix}0\\1\end{pmatrix}$. Then the beam splitter performs the following linear transformation $B_{\frac{\pi}{4}}$:

$$B_{\frac{\pi}{4}}|\text{up}\rangle = \tfrac{1}{\sqrt{2}}(|\text{up}\rangle + |\text{down}\rangle) = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}$$

$$B_{\frac{\pi}{4}}|\text{down}\rangle = \tfrac{1}{\sqrt{2}}(-|\text{up}\rangle + |\text{down}\rangle) = \begin{pmatrix}-\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}$$

Another variant of the beam splitter is given by the linear transformation

$$B_{-\frac{\pi}{4}}|\text{up}\rangle = \tfrac{1}{\sqrt{2}}(|\text{up}\rangle - |\text{down}\rangle) = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$B_{-\frac{\pi}{4}}|\text{down}\rangle = \tfrac{1}{\sqrt{2}}(|\text{up}\rangle + |\text{down}\rangle) = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Note that $B_{\frac{\pi}{4}}$ and $B_{-\frac{\pi}{4}}$ are unitary, and that $B_{\frac{\pi}{4}}B_{-\frac{\pi}{4}} = B_{-\frac{\pi}{4}}B_{\frac{\pi}{4}} = 1$.

The *Elitzur-Vaidman bomb tester* is the following construction. We are given a box that may or may not contain a bomb. The bomb explodes if a single photon falls onto it. We want to find out whether the box contains a bomb. To do so, we take a $B_{\frac{\pi}{4}}$ beam splitter and send an $|\text{up}\rangle$ photon through it. The state that comes out of the beam splitter is $\frac{1}{\sqrt{2}}(|\text{up}\rangle + |\text{down}\rangle)$. Now we put the box in the path of the $|\text{down}\rangle$ photon. Assume for the moment that a bomb is in that box. Then the box constitutes a measurement whether the photon takes the up- or the down-path. Since the state of the photon is $\frac{1}{\sqrt{2}}(|\text{up}\rangle + |\text{down}\rangle)$, the measurement outcome will be up or down, each with probability $\frac{1}{2}$. In the case of a down-outcome, the bomb explodes. In the case of an up-outcome, the resulting state is $|\text{up}\rangle$ (i.e., the photon takes the upper path). Then the photon passes the $B_{-\frac{\pi}{4}}$ beam splitter and is transformed into $\frac{1}{\sqrt{2}}(|\text{up}\rangle - |\text{down}\rangle)$. Now we measure whether the photon is in the up state or the down state (by simply putting a photon detector in at the end of both paths). With probability $\frac{1}{2}$ the photon will be up (conditioned on the fact that the bomb did not explode), with probability $\frac{1}{2}$ it will be down. Altogether we get the following predictions for this experiment.

Event	**Probability**
Bomb explodes	$\frac{1}{2}$
Photon is in up-path	$\frac{1}{4}$
Photon is in down-path	$\frac{1}{4}$

On the other hand, if no bomb is in the box, the box has no effect on the photon. In this case, the experiment consists of two beam splitters $B_{\frac{\pi}{4}}$ and $B_{-\frac{\pi}{4}}$ in a row. Because these beam splitters are inverses of each other, they cancel each other out, and the photon coming out of the second beam splitter will be in state $|\text{up}\rangle$. Thus in this case we get the following probabilities:

Event	**Probability**
Bomb explodes	0
Photon is in up-path	1
Photon is in down-path	0

In other words, if the outcome of the experiment is "down", we know for sure that there is a bomb in the box without having caused it to explode. Unfortunately, with

probability $\frac{1}{2}$ the bomb still explodes. The experiment can, however, be improved to make the probability of the bomb exploding arbitrarily small (homework).

Further reading: For the modelling of the beam splitter: [NC00, Section 7.4] (uses some physics we have not discussed yet). For the bomb tester: [Wik, Elitzur-Vaidman bomb-tester].

4 Composite Systems

ComposQSys

Definition 13 (Composite systems) *Given n quantum systems $\mathcal{H}_i$, the* composite system *is $\mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n$.*

ComposQState

Definition 14 (Composite states) *Given n quantum states $|\Psi_i\rangle \in \mathcal{H}_i$, the* composite state *consisting of n independent subsystems in states $|\Psi_i\rangle$ is*

$$|\Psi_1\rangle \otimes \ldots \otimes |\Psi_n\rangle \in \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n.$$

ComposUni

Definition 15 (Composite unitary operations) *Given a composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$, performing the unitary operation U_1 on $\mathcal{H}_1$ and U_2 on $\mathcal{H}_2$ independently is equivalent to performing the unitary operation $U_1 \otimes U_2$ on $\mathcal{H}_1 \otimes \mathcal{H}_2$.*

A special case is performing an operation U only on $\mathcal{H}_1$ and not touching $\mathcal{H}_2$. This is represented by $U \otimes I$ where I is the identity.

ComposMeas

Definition 16 (Composite measurements) *Given a measurement M_1 specified by projections $P_1, \ldots, P_n$ on $\mathcal{H}_1$ and a measurement M_2 specified by projections $P'_1, \ldots, P'_m$ on $\mathcal{H}_2$, performing each of the measurements independently is equivalent to performing the measurement M specified by the projections $P_{ij} := P_i \otimes P_j$ with $i = 1, \ldots, n$ and $j = 1, \ldots, m$. (I.e., the possible outcomes of M are pairs i, j with $i = 1, \ldots, n$ and $j = 1, \ldots, m$.)*

Note that the measurement that does nothing and has no effect on the state is given by the single projector I (the identity). Thus a measurement M on $\mathcal{H}_1$ only extends to a measurement M' on $\mathcal{H}_1 \otimes \mathcal{H}_2$ as follows: If M consists of $P_1, \ldots, P_n$, then M' consists of $P_1 \otimes I, \ldots, P_n \otimes I$.

Further reading: [NC00], Section 2.2.8.

5 Multi-qubit gates

CNOT

Definition 17 (Controlled NOT) *The CNOT gate on $\mathbb{C}^4$ is defined to be the linear operation defined by*

$$\begin{aligned}\text{CNOT}|00\rangle &= |00\rangle\\ \text{CNOT}|01\rangle &= |01\rangle\\ \text{CNOT}|10\rangle &= |11\rangle\\ \text{CNOT}|11\rangle &= |10\rangle\end{aligned}$$

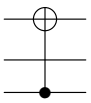
or equivalently

$$\text{CNOT}|a,b\rangle = |a, a\oplus b\rangle \qquad (a,b\in\{0,1\})$$

where $\oplus$ denotes XOR.

In circuits, we write CNOT as follows:

The dot represents the controlling qubit, and the $\oplus$ represents the qubit that is conditionally flipped. The dot does not have to be one the qubit above the $\oplus$. For example,



represents the operation defined by

$$|a,b,c\rangle \mapsto |a\oplus c, b, c\rangle \qquad (a,b,c\in\{0,1\})$$

Swap

Definition 18 (SWAP) *The SWAP gate on $\mathbb{C}^4$ is defined to be the linear operation defined by*

$$\text{SWAP}|a,b\rangle = |b,a\rangle.$$

The swap gate is represented by

Again, the two $\times$ do not have to be on adjacent lines.

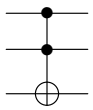
Toff

Definition 19 (Toffoli) *The Toffoli gate on $\mathbb{C}^8$ is defined to be the linear operation defined by*

$$\text{Toffoli}|a,b,c\rangle = |a,b,(a\cdot b)\oplus c\rangle$$

where $\cdot$ is the multiplication modulo 2, or equivalently, the and-operation.

The Toffoli gate is usually represented as follows:



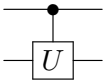
As with the CNOT, the two dots can be on arbitrary lines, not only those adjacent to the $\oplus$. Furthermore, the symbol generalises to more than two controlling qubits in the obvious way.

CU

Definition 20 (Controlled-U) *Given a unitary transformation $U \in \mathbb{C}^n$, the controlled-U gate $C(U)$ is defined to be the linear operation on $\mathbb{C}^{2n}$ defined by*

$$C(U)|0,j\rangle = |0,j\rangle$$
$$C(U)|1,j\rangle = |1\rangle \otimes U|j\rangle.$$

The controlled-U is depicted as follows:



Again, the dot can be on an arbitrary qubit.

Further reading: [NC00], Section 4.3

6 The Deutsch-Jozsa Algorithm

Deutsch's algorithm. Assume we are given a function $f : \{0,1\} \to \{0,1\}$. We ask the question which of the following two cases applies: Deutsch

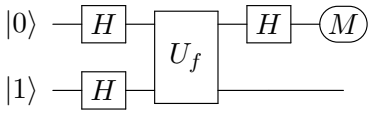
- f is constant ($f(0) = f(1)$), or
- f is balanced ($f(0) \neq f(1)$).

We further assume that f is implemented as a unitary transformation U_f on two qubits that performs the following operation:

$$U_f|x,y\rangle = |x, y \oplus f(x)\rangle \qquad (x,y \in \{0,1\})$$

(Such a unitary can be efficiently implemented if f has a poly-size classical circuit.)

Deutsch's algorithm performs the following operations:



(Here —(M) denotes a complete measurement of the first qubit in the computational basis, i.e., we look whether it is $|0\rangle$ or $|1\rangle$.)

Computing the output of this circuit, we get the following:

- If f is constant, then with probability 1 the measurement M has outcome 0.
- If f is balanced, then with probability 1 the measurement M has outcome 1.

Thus with one evaluation of f we have determined whether f is constant or balanced. Classically, we would have needed two evaluations.

An extension of this algorihm, the Deutsch-Jozsa algorithm, can even handle functions $f : \{0,1\}^n \to \{0,1\}$ and decide whether they are constant or balanced (same number of 0 and 1 outputs). It needs only one evaluation of f. (There is no guarantee if f is neither constant nor balanced.)

Further reading: [NC00, Section 1.4.3].

7 Density Operators

Intuitively, a quantum state probability distribution is a probability distribution on quantum states.

QDistr

Definition 21 (Quantum state probability distribution) *A* quantum state probability distribution *(a.k.a. ensemble) E over a Hilbert space $\mathcal{H}$ is a (possibly infinite) set of pairs $E = \{|\Psi_i\rangle @ p_i\}_i$ satisfying:*

- *For all i we have $|\Psi_i\rangle \in \mathcal{H}$.*
- *The vectors $|\Psi_i\rangle$ are normalized ($\||\Psi_i\rangle\| = 1$).*
- *We have $p_i \geq 0$ for all i and $\sum_i p_i = 1$.*

The interpretation is that a system is in state $|\Psi_i\rangle$ with probability p_i.

Operations performed on quantum states generalise to quantum state probability distributions.

QDistrU

Definition 22 (Unitary transformation) *Let U be a unitary matrix on $\mathcal{H}$. Let $E = \{|\Psi_i\rangle @ p_i\}_i$ be an quantum state probability distribution over $\mathcal{H}$.*

Then applying U to the quantum state probability distribution E leads to the quantum state probability distribution

$$UE = \{U|\Psi_i\rangle \,@\, p_i\}_i.$$

QDistrM

Definition 23 (Measurement) *Let $M = \{Q_1, \ldots, Q_n\}$ be a projective measurement over $\mathcal{H}$ consisting of projectors Q_i. Let $E = \{|\Psi_i\rangle @ p_i\}_i$ be an quantum state probability distribution over $\mathcal{H}$.*

If we measure the state described by E with M, the outcome j has probability

$$\Pr[\textit{Outcome } j] = \sum_i p_i \|Q_j|\Psi_i\rangle\|^2.$$

After measuring the outcome j, the system state is described by the following quantum state probability distribution:

$$\left\{ \frac{Q_j|\Psi_i\rangle}{\|Q_j|\Psi_i\rangle\|} \ @ \ \frac{p_i\|Q_j|\Psi_i\rangle\|^2}{\Pr[Outcome\ j]} \right\}_i.$$

QDistrX

Definition 24 (Extending the state space) *Let $E = \{|\Psi_i\rangle @ p_i\}_i$ be an quantum state probability distribution over $\mathcal{H}$. Let $|\Gamma\rangle \in \mathcal{H}'$, $\||\Gamma\rangle\| = 1$.*

Then extending the state described by E by adding another quantum system described by $|\Gamma\rangle$ results in the following quantum state probability distribution over $\mathcal{H} \otimes \mathcal{H}'$:

$$E \otimes |\Gamma\rangle = \{|\Psi_i\rangle \otimes |\Gamma\rangle \ @ \ p_i\}_i.$$

PhysInd

Definition 25 (Physical indistinguishability) *We call two quantum state probability distributions physically indistinguishable if all sequences of operations according to Definitions 22, 23, and 24 lead to the same probabilities of measurement outcomes.*

Density

A density operator is a compact representation of a quantum quantum state probability distribution. This representation looses some information contained in the description of an quantum state probability distribution,[1] but it still contains enough information to predict the outcome of physical experiments.

Definition 26 (Density operator) *Let $E = \{|\Psi_i\rangle @ p_i\}_i$ be a quantum state probability distribution over $\mathcal{H}$. The* density operator (density matrix, mixed state) *corresponding to E is the linear transformation ρ_E on $\mathcal{H}$ defined as follows:*

$$\rho_E = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|.$$

We call ρ a density operator over $\mathcal{H}$ if it is a density operator for some quantum state probability distribution E over $\mathcal{H}$. By $S(\mathcal{H})$ we denote the set of all density operators over $\mathcal{H}$.

Note: The usage of the words *mixed state* and *pure state* is ambiguous. There are two usages:

- A mixed state is a density operator $\rho \in S(\mathcal{H})$ and a pure state is a state described by a vector $|\Psi\rangle \in \mathcal{H}$.
- A pure state is a density operator of the form $|\Psi\rangle\langle\Psi|$ (i.e., a density operator corresponding to an quantum state probability distribution with only one entry), and a mixed state is a density operator that cannot be written as $|\Psi\rangle\langle\Psi|$.

DensityAlt

Lemma 1 *The set $S(\mathcal{H})$ consists of all positive Hermitian matrices with trace* 1.

Due to its mathematical simplicity, one usually takes Lemma 1 as the definition of density operators.

DensityU

Definition 27 (Unitary transformation) *Let U be a unitary matrix on $\mathcal{H}$. Let $\rho \in S(\mathcal{H})$ be a density operator over $\mathcal{H}$.*

Then applying U to the state ρ leads to the state $U\rho U^\dagger$.

DensityM

Definition 28 (Measurement) *Let $M = \{Q_1, \ldots, Q_n\}$ be a projective measurement over $\mathcal{H}$ consisting of projectors Q_i. Let $\rho \in S(\mathcal{H})$ be a density operator over $\mathcal{H}$.*

If we measure the state ρ with M, the outcome j has probability

$$\Pr[\textit{Outcome } j] = \operatorname{tr} Q_j \rho Q_j^\dagger = \operatorname{tr} Q_j \rho.$$

After measuring the outcome j, the system state is $\frac{Q_j \rho Q_j^\dagger}{\operatorname{tr} Q_j \rho Q_j^\dagger}$.

DensityX

Definition 29 (Extending the state space) *Let $\rho \in S(\mathcal{H})$ be a density operator over $\mathcal{H}$.*

Then extending the state ρ by adding another quantum system described by $\sigma \in S(\mathcal{H}')$ results in the density operator $\rho \otimes \sigma$ over $\mathcal{H} \otimes \mathcal{H}'$

DensityPhysInd

The following theorem states that density operators characterise physical indistinguishability of quantum state probability distributions.

Theorem 1 *Let E, E' be quantum state probability distributions over $\mathcal{H}$ and ρ, ρ' the corresponding density operators. Then E and E' are physically indistinguishable if and only if $\rho = \rho'$.*

Since in physics, there is no reason to assume that some distinction exists if it is principally impossible to measure it, one usually directly says that the physical system is in the state ρ and does not assume that there is some hidden quantum state probability distribution behind this state that contains more information than the density operator ρ.

Further reading: [NC00, Section 2.4.1 and 2.4.2]. Note that they define a density operator as being positive Hermitian (and omit the condition $\operatorname{tr} \rho = 1$).

8 Quantum One-Time Pad

OTP

The classical one-time pad is an encryption scheme that works as follows: Let $k \in \{0,1\}^n$ be a uniformly random key and let $m \in \{0,1\}^n$ be a message. Then the encryption

[1] E.g., the following two quantum state probability distributions both have the same representation as a density operator: $\{(|0\rangle, \frac{1}{2}), (|1\rangle, \frac{1}{2})\}$ and $\{(|+\rangle, \frac{1}{2}), (|-\rangle, \frac{1}{2})$ with $|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ and $|-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$.

of m using k is simply $c := m \oplus k$ (XOR). If we formulate this for the special case $n = 1$ (encrypting a single bit message), this means that if $k = 0$, the message bit m is unchanged, and if $k = 1$, the message bit is flipped.

The one-time pad has what is called "perfect secrecy", meaning that if k is uniformly random (and used only once), then the distribution of c does not depend on m. (In fact, c is a uniformly random bit for any m.)

QOTP

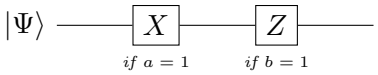
In the situation of the quantum one-time pad, instead of encrypting a classical message, we want to encrypt a quantum state. For simplicity, we only describe how to encrypt a single qubit (analogous to the $n = 1$ case in the classical one-time pad).

The obvious analogy to the classical one-time pad would be to have a one bit key $k \in \{0, 1\}$, and to apply the X-gate (quantum bit flip, Definition 3) if $k = 1$.

Unfortunately, the resulting encryption scheme is insecure. For example, when encrypting $|+\rangle$ or $|-\rangle$, respectively, we get the ciphertext $|+\rangle$ and $|-\rangle$, respectively, no matter what k is. (More precisely, we get something physically indistinguishable from $|+\rangle$, $|-\rangle$.) So one can distinguish the plaintexts $|+\rangle$ and $|-\rangle$ perfectly.

However, with two-bit keys, the QOTP is secure:

Definition 30 (Quantum one-time pad) *Let a, b be uniformly random independent bits. To encrypt a qubit $|\Psi\rangle \in \mathbb{C}^2$, we apply the following circuit:*



Security of this scheme is shown by the following lemma:

Lemma 2 *For any quantum state $|\Psi\rangle \in \mathbb{C}^2$, the density operator at the end of the QOTP circuit is $\rho = \frac{1}{2}I = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.*

Note that this lemma only states the security of the quantum one-time-pad in the case that the plaintext is a qubit that is not entangled with anything else. If we encrypt a qubit that is part of a larger quantum system, the above lemma does not guarantee anything. But this is only a problem of the formulation of the lemma, the QOTP is secure also in that case. (We do not have the tools yet to express the security in that case, we would need to partial trace introduced later to do so.)

9 Partial Trace and Purification

ParTr

Definition 31 (Partial trace) *Let a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B$ be given.*

The partial trace $\mathrm{tr}_B : S(\mathcal{H}_A \otimes \mathcal{H}_B) \to S(\mathcal{H}_A)$ *is the linear transformation defined by*

$$\mathrm{tr}_B\, \sigma \otimes \tau = \sigma \cdot \mathrm{tr}\, \tau \qquad \sigma \in S(\mathcal{H}_A),\ \tau \in S(\mathcal{H}_B).$$

We say that $\mathcal{H}_B$ (or just B) is traced out. *Analogously we can also trace out $\mathcal{H}_A$ or consider multipartite systems.*

Given a state $\rho \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$, the state $\rho^A := \mathrm{tr}_B\, \rho$ describes the state resulting from destroying (or locking away) the B-part of the system. Or equivalently, ρ^A represents all information that can be extracted about the state ρ from the A-part of the system alone.

Purif

Theorem 2 (Purification) *Let a state $\rho \in S(\mathcal{H}_A)$ be given. Then for any space $\mathcal{H}_B$ such that $\dim \mathcal{H}_B \geq \dim \mathcal{H}_A$, there is a quantum state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ such that*

$$\mathrm{tr}_B |\Psi\rangle\langle\Psi| = \rho.$$

We call Ψ a *purification* of ρ. Note that the purification is not unique.

This theorem means that any mixed state can be considered as a part of some larger pure state (we usually call the added subsystem $\mathcal{H}_B$ the *environment*).

In many cases, analysing a pure system may be simpler than analysing a mixed one. In these cases Theorem 2 allows to simplify the analysis.

Further reading: [NC00, Section 2.4.3] for the partial trace and [NC00, Section 2.5] for purification.

10 Quantum Operations

QOper

Definition 32 (Quantum Operations) *A* quantum operation $\mathcal{E}$ *is a map $\mathcal{E} : S(\mathcal{H}) \to S(\mathcal{H}')$ of the form*

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger \tag{1}$$

where $E_k : \mathcal{H} \to \mathcal{H}'$ are linear operators satisfying $\sum_k E_k^\dagger E_k = I$ (where I is the identity on $\mathcal{H}$).

We sometimes write $\mathcal{E} = \{E_k\}_k$ to denote the fact that $\mathcal{E}$ is the operation defined by (1)*. The operator E_k are called the* Kraus operators *of $\mathcal{E}$.*

Quantum operations describe all operations that can be applied to a mixed state ρ, including unitary transformations, measurements (when the outcomes are erased). Also the partial trace is an example of a quantum operation.

Quantum operations are also called *superoperators*.

ComposQOper

Definition 33 (Composing operations) *Let $\mathcal{E}$ and $\mathcal{F}$ be two quantum operations (over $\mathcal{H}_E$ and $\mathcal{H}_F$, respectively). Then $\mathcal{E} \otimes \mathcal{F}$ is the linear operation defined by*

$$(\mathcal{E} \otimes \mathcal{F})(\sigma \otimes \tau) = \mathcal{E}(\sigma) \otimes \mathcal{F}(\tau).$$

Note that $\mathcal{E} \otimes \mathcal{F}$ is a quantum operation over $\mathcal{H}_E \otimes \mathcal{H}_F$.

QOperAlt

Theorem 3 *$\mathcal{E} : S(\mathcal{H}) \to S(\mathcal{H}')$ is a quantum operation if and only if it satisfies the following three conditions:*

- *It is linear.*

- *It is trace-preserving (i.e.,* $\operatorname{tr}\mathcal{E}(\rho) = \operatorname{tr}\rho$*).*
- *It is* completely positive. *That is, for any vector space* $\tilde{\mathcal{H}}$ *and any positive* $\rho \in S(\mathcal{H} \otimes \tilde{\mathcal{H}})$*, we have that* $(\mathcal{E} \otimes I)(\rho)$ *is positive, too. (Here* I *is the identity on* $\tilde{\mathcal{H}}$*.)*

Further reading: [NC00, Section 8.2].

11 Trace distance

Note: In the following, we will use random variables and probability distributions interchangeably. That is, if we say "X is a probability distribution over A", we may then use X as a random variable taking values in A and write $\Pr[X = a]$ for the probability assigned by the distribution X to a.

SD

Definition 34 (Statistical distance) *Let* X *and* Y *be probability distributions over some (countable) set* A*. Then the* statistical distance $\mathrm{SD}(X, Y)$ *between* X *and* Y *is defined as*

$$\mathrm{SD}(X, Y) := \max_{T \subseteq A} \big|\Pr[X \in T] - \Pr[Y \in T]\big|.$$

Intuitively, the statistical distance tells us how good a sample chosen according to the distribution X and a sample chosen according to Y can be distinguished by an optimal statistical test T.

SDSumDef

Lemma 3 (Alternative definition of statistical distance) *Let* X *and* Y *be probability distributions over some (countable) set* A*. Then*

$$\mathrm{SD}(X, Y) = \tfrac{1}{2} \sum_{a \in A} \big|\Pr[X = a] - \Pr[Y = a]\big|.$$

This lemma is often taken as the definition of statistical distance. However, it does not have an operational meaning like Definition 34 and it does not generalise to uncountable sets A.

The statistical distance is often used in cryptography in definitions of security against computationally unlimited adversaries: If we have some random variable I that describes what the output/communication of the protocol should ideally look like (e.g., it should be stochastically independent of the secrets used in the protocol), and the random variable R describes the actual output/communication, then one would require that $\mathrm{SD}(R, I)$ is sufficiently small.

SDProps

Lemma 4
- *The statistical distance* SD *is a metric (on the set of probability distributions over a given set* A*).*
- *For any (possibly randomized) function* F *we have that*

$$\mathrm{SD}(F(X), F(Y)) \leq \mathrm{SD}(X, Y)$$

If F is injective, equality holds.

(This means that applying a function to some data may not make it more distinguishable, it may only loose information.)

- *Let X, Y, Z be stochastically independent. Then*

$$\mathrm{SD}((X,Z),(Y,Z)) \leq \mathrm{SD}(X,Y)$$

where (X,Z) is the random variable describing pairs chosen according to X and Z.

(Adding independent information does not help in distinguishing.)

TD

Definition 35 (Trace distance) *Given density operators $\sigma, \rho \in S(\mathcal{H})$, we define the* trace distance $\mathrm{TD}(\sigma,\rho)$ *as*

$$\mathrm{TD}(\sigma,\rho) := \tfrac{1}{2}\,\mathrm{tr}|\sigma-\rho|.$$

Here $|M|$ denotes the absolute value of the matrix M, see Definition 67.

TDMaxDef

Lemma 5 (Alternative definition of the trace distance) *Given density operators $\sigma, \rho \in S(\mathcal{H})$ we have that*

$$\mathrm{TD}(\sigma,\rho) = \max_P \big|\mathrm{tr}\,P\sigma - \mathrm{tr}\,P\rho\big|.$$

Here P ranges over all orthogonal projectors on $\mathcal{H}$.

In other words, the trace distance tells us how good we can distinguish the states σ and ρ by a measurement $\{P, 1-P\}$. This is analogous to Definition 34 since a quantum measurement is the analogue of a statistical test in the classical world.

TDSD

This analogy is made even stronger by the following lemma:

Lemma 6 *Let X and Y be probability distributions over A. Let*

$$\rho_X := \sum_{a\in A} \Pr[X=a]\,|a\rangle\langle a| \in S(\mathbb{C}^A)$$

(in other words, ρ_X describes the distribution X over classical states $|a\rangle$) and ρ_Y analogous.

Then $\mathrm{SD}(X,Y) = \mathrm{TD}(\rho_X,\rho_Y)$.

TDProps

Lemma 7

- *The trace distance* TD *is a metric (on $S(\mathcal{H})$).*
- *For any quantum operation $\mathcal{E}$ and any $\sigma, \rho \in S(\mathcal{H})$ we have that*

$$\mathrm{TD}(\mathcal{E}(\sigma),\mathcal{E}(\rho)) \leq \mathrm{TD}(\sigma,\rho).$$

If $\mathcal{E}$ applies a unitary (i.e., $\mathcal{E}(\rho) := U\rho U^\dagger$), then equality holds.

- *Let $\sigma, \rho \in S(\mathcal{H})$ and $\tau \in S(\mathcal{H}')$. Then*

$$\mathrm{TD}(\sigma \otimes \tau, \rho \otimes \tau) = \mathrm{TD}(\sigma, \rho).$$

Note the one-to-one correspondence with the properties in Lemma 4.

TDMeasLemma

Lemma 8 *Let P be an orthogonal projector on $\mathcal{H}$, let $\rho \in S(\mathcal{H})$, let $\varepsilon \geq 0$. Assume that $\operatorname{tr} P\rho \geq 1 - \varepsilon$ (i.e., the measurement $\{P_{yes} := P, P_{no} := 1 - P\}$ returns yes with high probability).*

Then there is a state $\rho' \in S(\mathcal{H})$ such that

(a) $\mathrm{TD}(\rho, \rho') \leq \sqrt{\varepsilon}$.

(b) There are states $|\Psi_i\rangle \in \operatorname{im} P$ and values p_i with $\sum_i p_i = 1$, $p_i \geq 0$ such that $\rho' = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$. (In other words, when measuring ρ', the measurement would always return yes, i.e., ρ' satisfies the property specified by P.)

This lemma gives a criterion to show that the trace distance between some state ρ and some set of states S is small: Find a projector P such that S consists of all states satisfying (b). Then show that with high probability, measuring P would succeed.

TDConvex

Lemma 9 (Convexity of the trace distance) *Let $\rho = \sum_i p_i \rho_i$ and $\sigma = \sum_i p_i \sigma_i$ with $\sum_i p_i = 1$, $p_i \geq 0$. Then*

$$\mathrm{TD}(\rho, \sigma) \leq \sum_i p_i \mathrm{TD}(\rho_i, \sigma_i).$$

This lemma is sometimes useful because it allows to remove some initial random choices from the analysis

A generalisation of this lemma that does not require the probabilities p_i to be the same in ρ and σ also exists.

Further reading: [NC00, Section 9.2.1].

12 Quantum key distribution

QKDIntro

The goal of quantum key distribution (QKD, a.k.a. quantum key exchange) is the following. Two parties Alice and Bob communicate over two kinds of channels. The first channel allows to send classical information and is authenticated (but not secret). The second channel allows to send qubits but is insecure (under the control of the adversary). Alice and Bob want to agree on a secret key by communicating only over these channels such that even a computationally unlimited adversary Eve that eavesdrops on the classical channel and controls the quantum channel cannot learn anything about the key. (But Eve is allowed to disrupt the communication.)

The basic idea of quantum key exchange is the following: If Alice sends to Bob qubits encoded in a random basis (unknown to Eve), then if Eve measures the qubits she will

necessarily introduce disturbances. Then Alice and Bob perform some checks on the qubits received by Bob, and if Eve eavesdropped, we may expect some of these checks to fail and Alice and Bob will abort the protocol. Otherwise, Alice and Bob use the transmitted qubits to derive a shared secret key.

There are various desirable properties that a QKD protocol should have:

- *Provable security.* It should be possible to actually prove the security of the protocol. This is a must, otherwise we do not gain much over the classical key exchange protocols.
- *Error tolerance.* The key exchange protocol should work even if the communication channel is noisy (introduces errors). This is difficult because a noisy channel also introduces disturbances that look similar to those introduced by an eavesdropper. So if Alice and Bob abort whenever there is a disturbance, the protocol will never succeed. If they choose not to abort, Eve may learn some information.
- *Realisability.* The protocol should not need to use a quantum computer. It should be executable using only simple operations like sending polarised photons and measuring the polarisation.
- *Arbitrary distance.* The key exchange protocol should work over an arbitrary distance. In realistic channels, the noise increases with the distance. From some distance on, the noise is too large to make key exchange possible. One solution is to add relays on the way that correct errors or perform other computations, but these relays should not be assumed to be secure (they might be under the control of Eve). Quantum error correction can be used in untrusted relays, but this needs a quantum computer.

The (rough) state of the art is listed in the following table:

	BB84 and others	Lo-Chau	this lecture
Provable security	yes	yes	yes
Error tolerance	yes	yes	no
Realisability	yes	no	no
Arbitrary distance	no	yes	no

Here BB84 and other stands for most of the currently investigated protocols of which (variations of) BB84 [BB84] are the most well-known. Lo-Chau stands for the protocol proposed in [LC99].

In this lecture, we analyse a simplification of the Lo-Chau protocol that does not need to use quantum error correction.

Most research today concentrates on trying to improve the range (distance) of QKD protocol with available technology. Current records lie in the order of 250 km [SWV$^+$09], and about 140 km through a wireless connection [SMWF$^+$07].

QKDSecDef

Definition 36 (Security of QKD) *Let a QKD protocol π be given. Let $n \in \mathbb{N}$. Let $\varepsilon > 0$.*

Let an adversary Eve be given (that has full control over the quantum channel between Alice and Bob, but can only listen to but not modify the classical channel between Alice and Bob). Then let $\rho^{\text{Real}}_{ABE} \in S(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E)$ *be the density operator describing the joint state of Alice's, Bob's and Eve's system in the case that Alice and Bob do not abort. Here* $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^{2^n}$ *because Alice's and Bob's final state consist of an* n*-bit key, and* $\mathcal{H}_E$ *is some arbitrary Hilbert space defined by Eve.*

Let $S_{\text{Ideal}} \subseteq S(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E)$ *be the set of all states of the form*

$$\Big(\sum_{k \in \{0,1\}^n} 2^{-n} (|k\rangle\langle k| \otimes |k\rangle\langle k|) \Big) \otimes \rho_E, \qquad \rho_E \in S(\mathcal{H}_E).$$

By P_{success} *denote the probability that Alice and Bob do not abort the protocol and thus output a key (given a particular adversary Eve).*

We say that π *is* ε*-secure if the following holds: For every adversary Eve, we have that*

$$\exists \rho^{\text{Ideal}}_{ABE} \in S_{\text{Ideal}} : \qquad \text{TD}(\rho^{\text{Real}}_{ABE}, \rho^{\text{Ideal}}_{ABE}) \cdot P_{\text{success}} \leq \varepsilon.$$

Intuitively this means that the keys output by Alice and Bob are the same with high probability, that these keys are almost uniformly distributed, and that Eve's information is almost independent of that key.

Bell

Definition 37 (Bell states) *The four* Bell states *are:*

$$\begin{aligned}
|\beta_{00}\rangle &= \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle \\
|\beta_{01}\rangle &= \frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle \\
|\beta_{10}\rangle &= \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle \\
|\beta_{11}\rangle &= \frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|10\rangle
\end{aligned}$$

The four Bell states form a basis of $\mathbb{C}^4$.

As a shorthand, we write $|\widetilde{xy}\rangle$ with $x, y \in \{0,1\}^n$ for the state $|\beta_{x_1 y_1}\rangle \otimes |\beta_{x_2 y_2}\rangle \otimes |\beta_{x_3 y_3}\rangle \otimes \ldots \otimes |\beta_{x_n y_n}\rangle$. In particular, $|\widetilde{0 \ldots 0}\rangle = |\beta_{00}\rangle^{\otimes n} = 2^{-n/2} \sum_{x \in \{0,1\}^n} |x\rangle \otimes |x\rangle$. (Note: we implicitly assume here that the qubits are reordered that the first qubits of each Bell state come before all the second qubits.)

The states $|\widetilde{xy}\rangle$ with $x, y \in \{0,1\}^n$ form a basis of $\mathbb{C}^{2^{2n}}$ ($2n$-qubit systems).

We will analyze the following QKD protocol:

QKDProto

Definition 38 (QKD protocol)
Parameters:

- *m: Number of qubits exchanged over the channel.*
- *q: Number of qubit pairs checked during Bell test* ($q < n$).
- *n: Length of raw key* ($n = m - q$).

- *t: Maximum number of errors in raw key.*
- *H: Parity check matrix of a linear binary error correcting code with n bit codewords, correcting t errors. (See Definition 41.)*
- *k: Bitlength of unencoded messages in that code. (H is a $\mathbb{F}_2^{(n-k)\times n}$ matrix.)*
- *ℓ: The length of the final key.*
- *s: the length of the seed of the universal hash function.*
- *$F : \{0,1\}^s \times \{0,1\}^n \to \{0,1\}^\ell$: a universal hash function. (See Definition 44.)*

Protocol:

Step 1. Alice prepares the $2m$-qubit state $|\widetilde{00}\rangle$ and sends the second half of each qubit pair to Bob over the insecure quantum channel. (We call the joint state of Alice, Bob, Eve after this state ρ_{init}.)

Step 2. Alice and Bob perform the "Bell test" (see Definition 39 below). (This reduces the number of qubit pairs from m to n. We call the joint state ρ_{test}.)

Step 3. Alice and Bob measure their respective n-qubit quantum systems in the computational basis. Call the measurement outcomes K_A, K_B. ("Raw keys.". We call the joint state ρ_{raw}.)

Step 4. Alice sends $\sigma := HK_A$ to Bob (over the authenticated channel). Bob finds e with $He = \sigma + HK_B$ and $|e| \leq t$. Then Bob updates his key to be $K'_B := K_B \oplus e$. (Such an e is unique and can efficiently be found if it exists by definition of error correcting codes. We call the joint state ρ_{corr}.))

Step 5. Privacy amplification: Alice picks $S \in \{0,1\}^s$ and sends S to Bob. Alice computes $K''_A := F(S, K_A)$, Bob computes $K''_B := F(S, K'_B)$. K''_A and K''_B are the final key. (If all goes well, $K''_A = K''_B$.)

We claim that for suitable choices of parameters, this protocol is a secure QKD protocol in the sense of Definition 36. We now proceed to analyze the protocol step by step. After Step 1, Alice and Bob have m qubits each, but besides that, we make not claims about the structure of ρ_{init}. (Since the communication went over the insecure channel, Eve could have modified it arbitrarily.)

12.1 Bell test

BellTest

We now describe Step 2 in more detail, and analyze what we can say about the state ρ_{test} after that step.

Definition 39 *Let a state $\rho \in S(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E)$ be given with $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^m$. Let $q \in \mathbb{N}$, $q \leq m$. The* Bell test *is the following procedure:*

- *Choose q distinct indices $i_1, \ldots, i_q \in \{1, \ldots, m\}$.*
- *For each index i, measure the i-th Alice-Bob qubit pair of ρ using one of the following measurements:*
  - *$P_{yes} := |\beta_{00}\rangle\langle\beta_{00}| + |\beta_{01}\rangle\langle\beta_{01}|$ and $P_{no} := 1 - P_{yes}$. (I.e., we check that the state is not $|\beta_{10}\rangle$ or $|\beta_{11}\rangle$.)*

– $P_{yes} := |\beta_{00}\rangle\langle\beta_{00}| + |\beta_{10}\rangle\langle\beta_{10}|$ *and* $P_{no} := 1 - P_{yes}$. *(I.e., we check that the state is not* $|\beta_{01}\rangle$ *or* $|\beta_{11}\rangle$*.)*

- *If this measurement returned no, abort.*

Note that this test cannot be directly implemented by Alice and Bob because it performs measurements on the joint state of Alice and Bob that cannot be implemented locally. On the exercise sheet, however, we devise an equivalent test that can be implemented with local operations and classical communication (the latter will then be performed through the authenticated channel).

For the analysis, we fix the following notation:

TildeNotation

For $x, y \in \{0,1\}^m$, by $|xy|$ we denote the number of bitpairs in xy that are not 00. More precisely, $|xy| = |\{i : x_i \neq 0 \vee y_i \neq 0\}|$.

BellTestAna

Let P_{ok} be the projector $\sum_{|xy|\leq t} |\widetilde{xy}\rangle\langle\widetilde{xy}| \otimes I_E$ (where I_E is the identity on Eve's system $\mathcal{H}_E$). That is, intuitively P_{ok} projects onto states that have at most t wrong qubit pairs. For notational convenience, we write $P_{ok}(\rho) := P_{ok}\rho P_{ok}^\dagger$.

Let T denote the (not trace-preserving) quantum operation describing the Bell test. More precisely, given a state $\rho \in S(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E)$, $T(\rho) := p\tilde{\rho}$ where $\tilde{\rho}$ is the state after passing the Bell test and p is the probability of passing the Bell test. Note that $\tilde{\rho} = \frac{T(\rho)}{\operatorname{tr} T(\rho)}$.[2]

Recall that ρ_{init} is the state that Alice and Bob hold in the QKD protocol before the Bell test. If $\rho_{init} = |\widetilde{0\ldots0}\rangle\langle\widetilde{0\ldots0}|$ (i.e., Eve has not disturbed the state), then the Bell test passes with probability 1. If $\rho_{init} = |\widetilde{xy}\rangle\langle\widetilde{xy}|$ where for more than t indices i, $x_i y_i \neq 00$ (i.e., Eve has disturbed a lot), the Bell test passes with probability at most $\delta_q := (1 - \frac{t+1}{2m})^q$. Note that for $t = 0$, even for $q = m$, this does not converge to 0, so we cannot use this test to ensure that there are no errors in the state. However, if t is a fixed fraction of m, δ_q converges exponentially fast to 0 for $m, q \to \infty$.

Lemma 10 *Let a state* $\rho \in S(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E)$ *be given with* $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^m$. *Let* $q \in \mathbb{N}$, $q \leq m$.

Let $\tilde{\rho} := \frac{T(\rho)}{\operatorname{tr} T(\rho)}$ *(the state after passing the Bell test). Let* $P_{success} := \operatorname{tr} T(\rho)$ *(the probability of passing the Bell test).*

Then $\operatorname{tr} P_{ok}(\tilde{\rho}) \geq \frac{\operatorname{tr} T(\rho) - \delta_q}{\operatorname{tr} T(\rho)}$. *That is, the (hypothetical) test whether* $\tilde{\rho}$ *indeed has at most* t *bad qubits will fail with probability at most* $\frac{\delta_q}{P_{success}}$.

In the following, let *t-Error* denote the set of states $|\Psi\rangle$ that are a superposition of states $|\tilde{x}\rangle$ with $|x| \leq t$, and with an arbitrary state on Eve's side. (In other words, in $|\Psi\rangle$, at most t bad qubit pairs occur.) Formally,

$$t\text{-}Error := span\{|\tilde{x}\rangle \otimes |\Psi^E\rangle : |x| \leq t, |\Psi^E\rangle \text{ arbitrary}\}.$$

[2] This encoding of the Bell test is analogous to $P_{ok}(\rho)$ where also both the post-measurement state and the probability are encoded in the operator $P_{ok}(\rho)$ of trace ≤ 1.

Let $S_{\text{Ideal}}^{t\text{-}Error}$ be the set of all states that are mixtures of states in *t-Error*. Formally,

$$S_{\text{Ideal}}^{t\text{-}Error} := \Big\{ \sum_i p_i |\widetilde{\Psi_i}\rangle\langle\widetilde{\Psi_i}| : \sum_i p_i = 1, \forall i.\, p_i \geq 0, \forall i.\, |\Psi_i\rangle \in t\text{-}Error \Big\}.$$

We have

Lemma 11 *For any adversary Eve, there exists a state $\rho_{test}^{ideal} \in S_{\text{Ideal}}^{t\text{-}Error}$ such that $\text{TD}(\rho_{test}, \rho_{test}^{ideal}) \cdot P_{success} \leq \sqrt{\delta_q}$. Here $P_{success}$ is the probability of passing the Bell test.*

To see this, note that $\rho_{test} = \frac{T(\rho_{init})}{\operatorname{tr} T(\rho_{init})}$ and $P_{success} = \operatorname{tr} T(\rho_{init})$, and that $S_{\text{Ideal}}^{t\text{-}Error}$ is the set of all states $\sum_i p_i |\Psi_i\rangle$ with $|\Psi_i\rangle \in \operatorname{im} P_{ok}$. Then Lemma 8 implies $\text{TD}(\tilde{\rho}, \rho') \leq \sqrt{\frac{\delta_q}{P_{success}}}$ from we $\text{TD}(\rho_{test}, \rho_{test}^{ideal}) \cdot P_{success} \leq \sqrt{\delta_q}$ immediately follows.

Note: compare this lemma with the definition of secure QKD schemes (Definition 36). Basically, the lemma shows that the protocol until Step 2 is a secure QKD-protocol, except that the set of ideal state S_{Ideal} is replaced by $S_{\text{Ideal}}^{t\text{-}Error}$ which consists of states where Alice and Bob have Bell pairs with at most t errors. So basically, we have shown that we have a $\sqrt{\delta_q}$-secure "t-error Bell state distribution protocol".

12.2 Measuring the raw key

RawKey

In Step 3, Alice and Bob measure the raw key K_A, K_B. Note that the raw key is not necessarily a good key yet. For example, we do not have the guarantee that $K_A = K_B$, and Eve might have partial information about K_A or K_B.

But the raw key is not all bad. In this section we analyze what useful properties it does have.

RawKeyKeyDiff

Lemma 12 *If $\rho_{test} \in S_{\text{Ideal}}^{t\text{-}Error}$, then, after Step 3, with probability 1 we have $|K_A \oplus K_B| \leq t$.*

This is shown on a homework sheet. Since Lemma 11 guarantees that ρ_{test} will be close to $S_{\text{Ideal}}^{t\text{-}Error}$, we know that $|K_A \oplus K_B| \leq t$ with high probability.

RawKeyGuess

Lemma 13 *If $\rho_{test} \in S_{\text{Ideal}}^{t\text{-}Error}$, then, for any algorithm that accesses only Eve's state in ρ_{raw} and outputs a guess K_E of Alice's key, we have $\Pr[K_A = K_E] \leq (3n+1)^t 2^{-n}$.*

This is shown in the practice session.

Lemma 13 allows us to quantify the so-called min-entropy of the raw key. The min-entropy is a measure of uncertainty that has a lot of importance in cryptography, and is defined as follows:

MinEnt

Definition 40 (Min-entropy) *Let $\mathcal{H}_K \otimes \mathcal{H}_B \otimes \mathcal{H}_E$ with $\mathcal{H}_K = \mathbb{C}^{\mathbf{K}}$ be a tripartite system. Let ρ be a cqq-state*[3] *on this system. ($\mathcal{H}_K$ represents the part of the state*

[3] A *cqq-state* means a state that is classical (c) in the first component, and potentially quantum (q) in the second and third. More precisely, $\rho = \sum_x p_i |x\rangle\langle x| \otimes \rho_i$ with $\rho_i \in S(\mathcal{H}_B \otimes \mathcal{H}_E)$.

containing the key, $\mathcal{H}_E$ Eve's part of the system, and $\mathcal{H}_B$ any other parts of the system not belonging to the key or Eve.)

We define the min-entropy as:

$$H_\infty(K|E)_\rho := -\log \max_M \Pr[K = M \text{ given } \rho].$$

Here M ranges over arbitrary quantum algorithms with input in $\mathcal{H}_E$ and classical output.

By "$\Pr[K = M$ given $\rho]$" we mean the following probability: Measure the system $\mathcal{H}_K$ in the computational basis. Run the algorithm M on system $\mathcal{H}_E$. Then the outcomes of the two measurements are equal.

The intuition behind this definition is again that $2^{-H_\infty(K|E)_\rho}$ is the maximum probability that Eve guesses the key K (contained in $\mathcal{H}_K$) while having access to the system $\mathcal{H}_E$. Notice that the definition assumes that the subsystem $\mathcal{H}_K$ contains a classical key. The definition of min-entropy generalizes to the case that all systems contain quantum data. However, in that case the definition is considerably less intuitive.

RawKeyEnt

By definition of H_∞, we can restate Lemma 13 as follows:

Lemma 14 *If $\rho_{test} \in S_{\text{Ideal}}^{t\text{-}Error}$ then*

$$H_\infty(K_A|E)_{\rho_{raw}} \geq -\log\big((3n+1)^t 2^{-n}\big) = n - t\log(3n+1).$$

(Here in slight abuse of notation we write K_A for Alice's subsystem.)

RawKeyAna

Let

$$S_{\text{Ideal}}^{raw} := \{\rho : K_A, K_B \text{ are classical in } \rho,\ H_\infty(K_A|E)_\rho \geq n - t\log(3n+1), \\ |K_A \oplus K_B| \leq t \text{ in } \rho\}$$

From Lemma 13 and Lemma 14 we immediately get:

Lemma 15 *If $\rho_{test} \in S_{\text{Ideal}}^{t\text{-}Error}$ then $\rho_{raw} \in S_{\text{Ideal}}^{raw}$.*

And combining this with Lemma 11, we get

Lemma 16 *For any adversary Eve, there exists a state $\rho_{raw}^{ideal} \in S_{\text{Ideal}}^{raw}$ such that $\text{TD}(\rho_{raw}, \rho_{raw}^{ideal}) \cdot P_{success} \leq \sqrt{\delta_q}$. Here $P_{success}$ is the probability of passing the protocol up to this point.*

(We used here implicitly that the trace distance can only decrease under quantum operations (Lemma 7) and that Step 3 is a quantum operation. We also used that the probability of success $P_{success}$ does not change after the Bell test any more, since the protocol can only abort during the Bell test.)

Basically, Lemma 16 shows that until Step 3, we have a $\sqrt{\delta_q}$-secure "leaky and not-exactly-the-same key distribution protocol".

12.3 Error-correction

In Step 4, make sure that Alice and Bob have the same key. To understand this step, we need some basics about error correcting codes, first.

ECC

Definition 41 (Error correcting code) *A (linear binary)* error correcting code *with codewords of length n, messages of length m, and correcting t errors consists of the following parts:*

- *A polynomial-time encoding algorithm that maps $m \in \{0,1\}^k$ into a codeword $c \in \{0,1\}^n$. Since the code is linear, $c = Gm$ for some fixed matrix $G \in \mathbb{F}_2^{n \times k}$ (the "generator matrix"). Let C be the set of all codewords, i.e., the image of G.*
- *A polynomial-time decoding algorithm that maps a codeword $c \in C$ to the original message m with $c = Gm$. (This can be done, e.g., by solving the linear equation system $c = Gm$ for unknown m.)*
- *A parity check matrix H, that is defined such that $Hc = 0$ iff $c \in C$. We call Hc' the syndrome of c'. Since $H(c \oplus e) = He$ for $c \in C$, the syndrome of a codeword with errors e only depends on the errors e, not on c.*
- *A polynomial-time error correction algorithm. Given a $c' \in \{0,1\}^n$ such that there exists a c with $|c \oplus c'| \leq t$, the algorithm finds c. (Think of c' as a codeword with errors.)*

Note that if H is a parity check matrix of a code that corrects t errors, then we have the following property: Given $\sigma = He$ for some e with $|e| \leq t$, we can efficiently find e. This is done as follows: Find some c' with $Hc' = \sigma$. Then $H(c' \oplus e) = 0$, hence $c := c' \oplus e$ is a valid codeword, and $|c \oplus c'| \leq t$. Thus the decoding algorithm applied to c' returns c. And thus we can compute $e := c \oplus c'$.

QKDCorr

This is what we do in Step 4. Assume that $|K_A \oplus K_B| \leq t$. Then Bob searches an e such that $He = H(K_A \oplus K_B)$. This will be $e = K_A \oplus K_B$. Hence $K'_B = K_A$. Thus we have:

Lemma 17 *If $\rho_{raw} \in S^{raw}_{\text{Ideal}}$, then in ρ_{corr} we have $K'_B = K_A$. (With slight abuse of notation, we now refer to Bob's system by K'_B.*

Thus Step 4 makes sure that Alice and Bob have the same key. However, sending σ over the network means that Eve learns additional information about the key. The following fact about min-entropy shows that Eve cannot learn more than $n - k$ bits (where $n - k$ is the length of σ).

Chain

Lemma 18 (Chain rule) *For any density operator ρ, $H_\infty(X|YE)_\rho \geq H_\infty(XY|E)_\rho - \ell$ if Y is an ℓ-bit system.*[4]

QKDCorrAna

From Lemma 18, we can conclude that the min-entropy decreases at most by $n - k$, hence:

[4]This even holds if X and Y are not classical. Notice that our definition of H_∞ only allows to talk about classical X, Y, but more general definitions exist [Ren05].

Lemma 19 *If* $\rho_{raw} \in S^{raw}_{\text{Ideal}}$, *then* $H_\infty(K_A|E)_{\rho_{corr}} \geq H_\infty(K_A|E)_{\rho_{raw}} - (n-k) \geq k - t\log(3n+1)$.

Let

$$S^{corr}_{\text{Ideal}} := \{\rho : K_A, K'_B \text{ are classical in } \rho,\ H_\infty(K_A|E)_\rho \geq k - t\log(3n+1),\ K_A = K'_B \leq t \text{ in } \rho\}$$

From Lemma 17 and Lemma 19 we then immediately have:

Lemma 20 *If* $\rho_{raw} \in S^{t\text{-}Error}_{\text{Ideal}}$ *then* $\rho_{corr} \in S^{corr}_{\text{Ideal}}$.

And combining this with Lemma 16, we get

Lemma 21 *For any adversary Eve, there exists a state* $\rho^{ideal}_{corr} \in S^{corr}_{\text{Ideal}}$ *such that* $\text{TD}(\rho_{corr}, \rho^{ideal}_{corr}) \cdot P_{success} \leq \sqrt{\delta_q}$. *Here* $P_{success}$ *is the probability of passing the protocol up to this point.*

(We used here implicitly that the trace distance can only decrease under quantum operations (Lemma 7) and that Step 4 is a quantum operation on the joint state of Alice, Bob, Eve. We also used that the probability of success $P_{success}$ does not change after the Bell test any more, since the protocol can only abort during the Bell test.)

Basically, Lemma 21 shows that until Step 4, we have a $\sqrt{\delta_q}$-secure "leaky key distribution protocol".

12.4 Privacy amplification

PrivAmp

After Step 4, Alice and Bob have the same key (with high probability), but Eve might still have partial information about that key. To get rid of the remaining knowledge of Eve, Alice and Bob perform "privacy amplification". The idea here is to apply a function F to the key such that, if Eve has only partial knowledge of the input K_A to F, then Eve has (close to) no knowledge about the output of F. That is, F should transform something weakly random into something (close to) uniformly random. The main tool is a so-called strong quantum randomness extractor (or simple strong quantum extractor).

RandExtQ

Definition 42 (Strong quantum extractor) *A function* $F : S \times X \to Y$ *is a* strong (k, ε)-quantum extractor *iff the following holds:*

Consider a multi-partite quantum system $\mathcal{H}_X \otimes \mathcal{H}_E$ *with* $\mathcal{H}_X = \mathbb{C}^X$. *Let* $\mathcal{H}_S := \mathbb{C}^S$, $\mathcal{H}_Y := \mathbb{C}^Y$. *Consider a cq-state* ρ *(i.e.,* ρ *is of the form* $\rho = \sum_x p_x |x\rangle\langle x| \otimes \rho_x$*). Assume that* $H_\infty(K|E)_\rho \geq k$.

Let

$$\rho_{extr} := \sum_{x,s} \tfrac{1}{|S|} p_x |F(s,x)\rangle\langle F(s,x)| \otimes \rho_x \otimes |s\rangle\langle s| \in S(\mathcal{H}_Y \otimes \mathcal{H}_E \otimes \mathcal{H}_S).$$

That is, ρ_{extr} *is the result of adding a register* S *containing a random value (the seed) to* ρ *and then replacing* X *by* $F(S, X)$.

Let

$$\rho_{perf} := \left(\sum_y \tfrac{1}{|Y|}|y\rangle\langle y|\right) \otimes \left(\sum_x p_x \rho_x\right) \otimes \left(\sum_r \tfrac{1}{|R|}|r\rangle\langle r|\right) \in S(\mathcal{H}_Y \otimes \mathcal{H}_E \otimes \mathcal{H}_S).$$

That is, ρ_{perf} is the result of adding a register S containing a random value (the seed) to ρ and then replacing X by a random value from Y.

Then $\mathrm{TD}(\rho_{extr}, \rho_{perf}) \leq \varepsilon$.

Intuitively, this means that as long as $H_\infty(K|E)_\rho \geq k$, one cannot distinguish between $F(S, X)$ and uniformly random Y, even given E and S.

RandExtC

For comparison, here is the definition of a classical strong extractor:

Definition 43 (Strong extractor) *A function $F : S \times X \to Y$ is a* strong (k, ε)-extractor *iff the following holds:*

Consider random variables $X \in X$ and E with $H_\infty(X|E) \geq k$. Let $S \in S$ and $Y \in Y$ be uniformly random and independent of each other and X, E.

Then

$$\mathrm{SD}\Big((F(S, X), E, S);\ (Y, E, S)\Big) \leq \varepsilon.$$

This is the same as the strong quantum extractor, except that now all registers are classical (even E), which makes notation much simpler. In particular, a strong (k, ε)-quantum extractor is a strong (k, ε)-extractor.

UHF

Examples for strong quantum extractors are so-called universal hash functions (a.k.a. two-universal hash functions):

Definition 44 (Universal hash function) *A function $f : S \times X \to Y$ is a* universal hash function (UHF) *iff for all $x, y \in X$ with $x \neq y$, we have that*

$$\Pr[f(s, x) = f(s, y) : s \xleftarrow{\$} S] \leq \tfrac{1}{|Y|}.$$

Here $s \xleftarrow{\$} S$ means that s is uniformly randomly chosen from S.

LHL

Universal hash functions are known to be strong extractors, even in the quantum case:

Lemma 22 (Leftover hash lemma, quantum-variant) *Let $f : S \times X \to Y$ be a universal hash function with $|Y| \leq 2^\ell$. Let $k \geq 0$. Let $\varepsilon := 2^{-\frac{1}{2}(k-\ell)-1}$. Then f is a strong (k, ε)-quantum extractor.*

PrivAmpAna

We can now analyze Step 5. Before Step 5, we have the state ρ_{corr}. If $\rho_{corr} \in S^{corr}_{\mathrm{Ideal}}$, then $K_A = K'_B$ in ρ_{corr}, and thus also $K''_A = K''_B$ in ρ_{priv}.

Furthermore, if $\rho_{corr} \in S^{corr}_{\mathrm{Ideal}}$, then $H_\infty(K_A|E)_{\rho_{corr}} \geq k - t\log(3n+1) =: d$. Let $\rho^{AE}_{corr} := \mathrm{tr}_B\, \rho_{corr}$ and $\rho^{AE}_{priv} := \mathrm{tr}_B\, \rho_{priv}$. Note that ρ_{priv} differs from ρ_{corr} besides other things in that the seed S is now added to Eve's state E. Then $H_\infty(K_A|E)_{\rho^{AE}_{corr}} \geq d$, and ρ^{AE}_{priv} is the state ρ_{extr} from Definition 42 (if we set $\rho := \rho^{AE}_{corr}$ in that definition).

Since F is a strong (d,γ)-quantum extractor by Lemma 22 for $\gamma := 2^{-\frac{1}{2}(d-\ell)-1}$, it follows by Definition 42 that ρ^{AE}_{priv} has statistical distance γ from a state of the form $\rho^{AE}_{ideal} := \left(\sum_y \frac{1}{|Y|}|y\rangle\langle y|\right) \otimes \rho_E$ for some ρ_E. (ρ_E here contains the second and the third tensor factor of ρ_{priv} from Definition 42.)

Let $\mathcal{E}$ denote the quantum operation that copies the (classical) register A into a register B. Since $K''_A = K''_B$ in ρ_{priv} (still assuming $\rho_{corr} \in S^{corr}_{\text{Ideal}}$), we have that $\rho_{priv} = \mathcal{E}(\rho^{AE}_{priv})$. Let $\rho_{ideal} := \mathcal{E}(\rho^{AE}_{ideal})$. Since $\mathrm{TD}(\rho^{AE}_{priv}, \rho^{AE}_{ideal}) \leq \gamma$, with Lemma 7 we get $\mathrm{TD}(\rho_{priv}, \rho_{ideal}) = \mathrm{TD}(\mathcal{E}(\rho^{AE}_{priv}), \mathcal{E}(\rho^{AE}_{ideal})) \leq \gamma$. Furthermore, note that $\rho_{ideal} = \left(\sum_y \frac{1}{|Y|}|y\rangle\langle y| \otimes |y\rangle\langle y|\right) \otimes \rho_E \in S_{\text{Ideal}}$.

Thus we have:

Lemma 23 *If $\rho_{corr} \in S^{corr}_{\text{Ideal}}$, then there is a $\rho_{ideal} \in S_{\text{Ideal}}$ with $\mathrm{TD}(\rho_{priv}, \rho_{ideal}) \leq \gamma$.*

Combining this with Lemma 21 we get

Lemma 24 *For any adversary Eve, there exists a state $\rho^{ideal} \in S_{\text{Ideal}}$ such that $\mathrm{TD}(\rho_{priv}, \rho^{ideal}) \cdot P_{success} \leq \sqrt{\delta_q} + \gamma$. Here $P_{success}$ is the probability of passing the protocol up to this point.*

QKDWrapup

Since ρ_{priv} is the final state of the protocol from Definition 38, and $\delta_q = (1 - \frac{t+1}{2m})^q$ and $\gamma = 2^{-\frac{1}{2}(k - t\log(3n+1) - \ell) - 1}$, we immediately get:

Theorem 4 (Security of QKD) *The protocol from Definition 38 is ε-secure in the sense of Definition 36 for*

$$\varepsilon := \left(1 - \tfrac{t+1}{2m}\right)^{q/2} + 2^{-\frac{1}{2}(k - t\log(3n+1) - \ell) - 1}.$$

Further reading: [NC00, Section 12.6]. (However, things are a very vague there.)

13 Shor's algorithm

Note: The following section will contain only a simplified exposition that is not complete but will give the rough idea of how to factor integers using quantum computers.

Fact

Definition 45 (Factoring problem) *Given a non-prime integer $m > 1$, find an integer $d \mid m$ with $d \neq 1$, $d \neq m$ (a non-trivial divisor).*

Period

Definition 46 (Period finding problem) *Let f be a periodic function. I.e., there is some p such that $f(x) = f(x+p)$ for all x.*[5]

Find p.

FactFromPeriod

[5]For simplicity, we implicitly also assume that $f(x) \neq f(y)$ if $x - y$ is not a multiple of p. Then the period is uniquely determined.

Lemma 25 (Reducing factoring to period finding) *Given an oracle that solves the period finding problem (for functions of the form $f_a(x) := a^x \bmod N$) we can solve the factoring problem with probability at least $\frac{1}{4}$ in polynomial-time using a single query to the period finding oracle.*

The idea of the reduction is, for random a, to find the smallest r such that $a^r \equiv 1 \pmod N$, and then to compute $\gcd(a^{r/2}+1, m)$ and $\gcd(a^{r/2}-1, m)$ for random x. With probability at least $\frac{1}{4}$, r will be even and one of the two gcds will be a non-trivial divisor of m.

DFT

Definition 47 (Discrete Fourier transform) *The* discrete Fourier transform (DFT) *is a linear transformation on $\mathbb{C}^N$ represented by the matrix $D_N = \frac{1}{\sqrt{N}} \left((e^{2i\pi kl/N})\right)_{kl} \in \mathbb{C}^{N\times N}$.*

Note that since $2i\pi kl/N$ is an imaginary number, all entries of D_N have absolute value 1.

Lemma 26 (Properties of the discrete Fourier transform)

- *The discrete Fourier transform D_N is unitary.*
- Frequency analysis: *Given a vector x which is p-periodic (i.e., $x_i = x_{i+p \bmod N}$ for all i; a special case would be a vectors with 1's at every p-th position), $D_N x$ has entries (non-zero values) on the multiples of N/p.*[6] *Note that N/p intuitively represents the frequency of x.*

DFTAlgo

Theorem 5 (Realising the discrete Fourier transform) *There is a quantum algorithm taking an n qubit state $|\Psi\rangle$ as input and returning $D_N|\Psi\rangle$ where D_N is the Fourier transform on $\mathbb{C}^N$ with $N = 2^n$. This algorithm runs in polynomial time in n.*

Shor

Theorem 6 (Shor's algorithm for period finding) *Assume that $f(n)$ can be computed in polynomial time. Assume that f is periodic and that an upper bound on the period of f is known.*

Then there is a polynomial-time quantum algorithm that returns the period of f.

The algorithm roughly goes as follows: Let $N = 2^n$ be sufficiently larger than the period of f. The algorithm starts with a quantum state $|0\rangle|0\rangle \in \mathcal{H}_X \otimes \mathcal{H}_Y$, the first system $\mathcal{H}_X := \mathbb{C}^N$ encoding integers $\{0, \ldots, N-1\}$, and the second system $\mathcal{H}_Y$ encoding outputs of f. It applies the Hadamard transform to every qubit of $\mathcal{H}_X$. This results in the state $|\Psi_1\rangle \propto \sum_{x\in\{0,1\}^n} |x\rangle|0\rangle$ ($\propto$ means equal up to a normalization factor). We can implement the unitary transformation U that takes $|x\rangle|y\rangle$ to $|x\rangle|y \oplus f(x)\rangle$. By applying U to $|\Psi_1\rangle$, we get $|\Psi_2\rangle \propto \sum_{x\in\{0,1\}^n} |x\rangle|f(x)\rangle$. We then measure the system $\mathcal{H}_B$ in the

[6] If $p \nmid N$, this holds only approximately. In this exposition, we will not formulate exact bounds for the approximation.

computational basis. This results in a measurement outcome $y = f(x')$ for some x'. The state after this measurement is $|\Psi_3\rangle \propto \sum_x |x\rangle$ where the sum ranges over all x with $f(x) = y = f(x')$, i.e., $x = x' + kp$ where p is the period of f. Hence $|\Psi_3\rangle$ is p-periodic. Thus, if we apply the Fourier transform D_N, we get a vector $D_N|\Psi_3\rangle$ which has entries on multiples of N/p (approximately). If we measure the system in the computational basis, we get a multiple of N/p. From this we can compute an approximate divider of p. Additional work needs to be done to recover the exact value of p from this, but this is a classical computation and omitted here.

Definition 48 (Discrete logarithm problem) *Let G be a (multiplicative) group and g a generator. Given $y \in G$, find x with $g^x = y$. (That value x is called the discrete logarithm* $\operatorname{dlog} y$ *of y.)*

DlogAlgo

Theorem 7 *Assume a group G with generator g in which exponentiation is feasible in polynomial-time. There is a polynomial-time quantum algorithm that returns* $\operatorname{dlog} a$ *on input of $a \in G$.*

Further reading: [NC00, Sections 5.1–5.3]

14 Lattice-based cryptography

In this section, we introduce one example of so-called *lattice-based cryptography* that is a candidate for classical cryptography secure against attacks using quantum computers.

14.1 Learning with errors

We first introduce a computational problem that forms the basis of the cryptosystem described in the next section.

We warm up with a slightly simpler to explain problem:

BinCompLWE

Informally, the binary computational LWE problem is, given a publicly known binary matric A, to find s given $As + e$ (where e is an error vector with "few" 1's).

Definition 49 (Binary computational LWE problem) *Fix parameters $n, m > 0$ (integers) and $p \in [0, 1]$. Let $A \xleftarrow{\$} \{0,1\}^{m \times n}$ (a uniformly random binary $m \times n$-matrix) and $s \xleftarrow{\$} \{0,1\}^n$ (a uniformly random binary n-vector), and let $e \in \{0,1\}^m$ be chosen by independently letting each e_i be 1 with probability p.*

The task of the binary decisional LWE problem (with parameters n, m, p) is to compute s given $A, b := As + e$.

It is generally believed that this problem is hard for suitable parameters. (I.e., the probability of guessing s is exponentially small for a polynomial-time adversary.)

BinDecLWE

Instead of asking an adversary to find s, we ask the simpler question: Is the vector b indeed of the form $As + e$, or is it just a random vector? (This is simpler, at least for

relevant parameter choices, since if you can find s from $As + e$, you can also tell whether you got $As + e$ by just checking whether you do find s.)

Definition 50 (Binary decisional LWE problem) *Fix parameters $n, m > 0$ (integers) and a $p \in [0, 1]$. Let $A \xleftarrow{\$} \{0,1\}^{m \times n}$ (a uniformly random binary $m \times n$-matrix) and $s \xleftarrow{\$} \{0,1\}^n$ (a uniformly random binary n-vector), and let $e \in \{0,1\}^m$ be chosen by independently letting each e_i be 1 with probability p. Let $r \xleftarrow{\$} \{0,1\}^m$.*

The task of the binary decisional LWE problem (with parameters n, m, p) is to distinguish the following two data:

- *A, $As + e$*
- *A, r*

It is generally believed that this problem is hard for suitable parameters. (I.e., the probability of guessing right is exponentially close to random guessing for a polynomial-time adversary.)

DecLWE

The binary LWE problem considered the problem of guessing a bitstring given A and $As + e$.

However, there is no reason per se to consider only bitstrings. Instead, we can fix an additional parameter q, and perform all operations modulo q. That is, A and s contain elements of $\mathbb{Z}_q$. And e is a vector consisting of "small" numbers in $\mathbb{Z}_q$. (By "small" we intuitively mean close to 0. So for example 1 would be small, but $q - 1 \equiv -1$ would also be small, but $q/2$ would not be.) Since now the errors e_i are not just 0 or 1, we need to specify a distribution χ that tells us how e_i is distributed. (We think of χ as a distribution that gives 0 or small values in $\mathbb{Z}_q$ with high probability.)

Definition 51 (Decisional LWE problem) *Fix parameters $n, m, q > 0$ (integers) and a distribution χ over $\mathbb{Z}_q$ Let $A \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$ (a uniformly random $m \times n$-matrix) and $s \xleftarrow{\$} \mathbb{Z}_q^n$ (a uniformly random n-vector), and let $e \leftarrow \chi^m$ (i.e., e consists of m values independently chosen from χ). Let $r \xleftarrow{\$} \mathbb{Z}^m$.*

The task of the decisional LWE problem (with parameters n, m, q, χ) is to distinguish the following two data:

- *A, $b := As + e$*
- *A, r*

It is generally believed that this problem is hard for suitable parameters, with some distribution χ of small values.

CompLWE

Of course, there is also a computational LWE problem in the general case. Basically, a combination of Definition 49 and Definition 51. We omit the details here.

14.2 Regev's cryptosystem

Regev

We now describe how to build a public key encryption scheme based on LWE. It is not the most efficient (it encrypts each bit separately) but it contains important ideas used in many modern lattice-based encryption schemes.

In the following we interpret the elements of $\mathbb{Z}_q$ as integers $\{-\lceil q/2\rceil + 1, \ldots, \lfloor q/2\rfloor\}$ (instead of, as usual, as integers $\{0, \ldots, q-1\}$). This is relevant whenever we say something like "$|x| \leq n$ as integers" for some $x \in \mathbb{Z}_q$.

And $\cdot$, applied to two vectors, is the inner product.

Definition 52 (Regev's cryptosystem) *Regev's cryptosystem is a public key encryption scheme with message space* $\{0,1\}$.

- ***Parameters.*** *The parameters* n, m, q, χ *from Definition 51. We assume that* χ *is chosen in a way such that* $|x \cdot e| \geq q/4$ *some small probability* p_{error} *when* $x \xleftarrow{\$} \{0,1\}^m$ *and* $e \xleftarrow{\$} \chi^m$.
- ***Key generation.*** *Generate* A, b, s *as in Definition 51. The secret key is* s. *The public key is* (A, b).
- ***Encryption.*** *To encrypt* $\mu \in \{0,1\}$, *pick* $x \xleftarrow{\$} \{0,1\}^m$. *Let* $c_1 := A^T x$ *and* $c_2 := x \cdot b + \mu\lfloor q/2\rfloor$ *(all calculated in* $\mathbb{Z}_q$*).*
- ***Decryption.*** *To decrypt* (c_1, c_2), *we compute* $t := c_2 - c_1 \cdot s$. *If* $|t| < q/4$ *(where* t *is interpreted as an integer, see above), return message* 0, *otherwise return message* 1.

Lemma 27 (Correctness) *Decrypting the encryption of a message* μ *returns* μ *with probability at least* $1 - P_{error}$.

RegevCPA

Lemma 28 (IND-CPA security (informal)) *If the decisional LWE problem is hard (for the parameters used in Definition 52) and if* $m - n \log q$ *is large enough (superlogarithmic), then an encryption of* 0 *and an encryption of* 1 *are indistinguishable.*

Further reading: [Pei16] gives an extensive overview of the basics of lattice-based cryptography. (Section 4.2 and 5.2 cover the material given here.) Regev's cryptosystem was originally proposed in [Reg09] together with an formal investigation of the hardness of LWE.

15 Zero-knowledge proofs

A zero-knowledge proof is, intuitively speaking, a protocol in which a prover P is able to convince a verifier V of the truth of a statement x in such a way that the verifier learns nothing (except, of course, the fact that x is true).

More formally, we first fix a relation R. If $(x, w) \in R$, we say that w is a witness for the statement x. We defined the language L_R of true statements as follows:

$$L_R := \{x : \exists w.(x, w) \in R\}.$$

In other words, $x \in L_R$ iff there is a witness for x.

ProofSys

We first define what it means for (P, V) to form a proof system (in the classical case). For this, we first introduce the following notation: For two machines A, B, $\langle A(a), B(b)\rangle$ denotes the output of B after an interaction of A and B where A gets input a and B gets input b.

Definition 53 (Proof systems) *We call a pair (P, V) of interactive machines a* proof *or* proof system *for the relation R with* soundness-error ε *iff the following two conditions are fulfilled:*

- Completeness: *For any $(x, w) \in R$, we have that* $\Pr[\langle P(x, w), V(x)\rangle = 1] = 1$. *(I.e., when the prover gets a valid witness w for x, then he manages to convince V of the truth of x.)*[7]
- Soundness: *For any (potentially computationally unlimited) machine P^*, and for any $x \notin L_R$, we have* $\Pr[\langle P^*(), V(x)\rangle = 1] \leq \varepsilon$. *(I.e., except for probability ε, no prover can convince V of a wrong statement x.)*

We can now define what it means that the verifier does not learn anything: ZK

Definition 54 (Zero-knowledge) *A pair (P, V) of interactive machines is* statistical zero-knowledge *if for any polynomial-time*[8] *V^* there exists a polynomial-time S and a negligible μ such that for all $(x, w) \in R$ and all $z \in \{0,1\}^*$, we have*

$$\mathrm{SD}(\langle P(x, w), V^*(x, z)\rangle, S(x, z)) \leq \mu(|x|).$$

(I.e., the simulator can simulate anything V^* learns without knowing the witness w.)

An example for a zero-knowledge proof is the following: GIZK

Definition 55 (Graph isomorphism) *The relation R_{GI} is defined as follows: $(x, w) \in R_{GI}$ iff $x = (G_1, G_2)$ and $w = \phi$ where G_1, G_2 are graphs and $\phi : G_1 \to G_2$ is a graph isomorphism.*

Definition 56 (Graph isomorphism proof system) *Let* GIP *denote the following protocol between machine P and V:*

- *P gets inputs $x = (G_1, G_2)$ and $w = \phi$.*
- *V gets input x.*
- *P picks a uniformly random permutation ψ_1 on the vertices of G_1 and computes $H := \psi_1(G_1)$. (Notice that now $\psi_1 : G_1 \to H$ is an isomorphism.)*
- *P sends H to V.*
- *V picks $i \in \{1, 2\}$ uniformly and sends i to P.*
- *P computes $\psi_2 := \psi_1 \circ \phi^{-1}$ and sends $\psi_i : G_i \to H$.*
- *V checks whether $\psi_i : G_1 \to H$ is an isomorphism. If so, V outputs* 1.

Theorem 8 GIP *is a statistical zero-knowledge proof system.*

We now present the definitions of zero-knowledge proofs for the quantum case. For two quantum or classical machines A, B, $\langle A(a), B(b)\rangle$ denotes the quantum state of B

[7] Of course, one could also relax this condition and allow a certain error in the completeness instead of requiring probability 1. For simplicity, we stick to the present definition.

[8] In this section, we will call a machine polynomial-time its running-time is bounded by a polynomial in the length of its *first* argument.

(or, if B is classical, its output) after an interaction of A and B where A gets input a and B gets input b. Here a and b may be classical values or density operators.

The definition of being a proof system (i.e., completeness and soundness) is word for word the same as in the classical case (Definition 53), except that P^* is allowed to be a quantum machine.

More interesting is the definition of statistical quantum zero-knowledge: QZK

Definition 57 (Quantum zero-knowledge) *A pair (P, V) of interactive machines is* statistical quantum zero-knowledge *if for any polynomial-time* quantum-machine V^* *there exists a polynomial-time* quantum-machine S *and a negligible μ such that for all $(x, w) \in R$ and all density operators ρ, we have*

$$\mathrm{TD}(\langle P(x,w), V^*(x,\rho)\rangle, S(x,\rho)) \leq \mu(|x|).$$

(I.e., the simulator can simulate anything V^* learns without knowing the witness w.)

Note that in this case, the "auxiliary input" that V^* gets (called z in the case of Definition 54) is a quantum state.

To show that GIP is statistical QZK, we need to construct a suitable simulator S. QZKProblem However, it turns out that the construction from the classical case does not directly carry over. The reason is that the simulator in the classical case uses rewinding: It tries to produce a simulation, and, if it fails, it tries again. In the quantum case, trying again is not necessarily an option, because the first try may have destroyed the input state ρ, so the second try will fail.

What does work, however, is constructing a polynomial-time simulator S_1 that tries to produce a simulation and either produces a perfect simulation or aborts, and that aborts with probability exactly $\frac{1}{2}$. QAbortSim (More precisely, if $(x, w) \in R$ and ρ' is the state output by the simulator $S_1(x, \rho)$, then $\operatorname{tr} P_\perp \rho' = \frac{1}{2}$ and $P_\perp \rho' P_\perp / \operatorname{tr} P_\perp \rho' P_\perp = \langle P(x, w), V^*(x, \rho)\rangle$ where $P_\perp$ projects on the state denoting abort.)

The construction of this simulator is analogous to the classical case and the proof that it produces a perfect simulation with probability $\frac{1}{2}$ also follows very closely the lines of the proof in the classical case.

To produce a simulator S in the sense of Definition 57 from S_1, we cannot directly QRewind follow the classical proof. Instead, we use the following lemma:

Lemma 29 (Quantum rewinding lemma [Wat09]) *Let Q be a unitary operation from $\mathcal{H}_{in} \otimes \mathcal{H}_{anc}$ to $\mathcal{H}_{out} \otimes \mathcal{H}_{succ}$ with $\mathcal{H}_{succ} = \mathbb{C}^2$. (This implies that $\dim \mathcal{H}_{in} \otimes \mathcal{H}_{anc} = \dim \mathcal{H}_{out} \otimes \mathcal{H}_{succ}$ since a unitary operation is a square matrix.)*

Assume that there is a value $p \leq \frac{1}{2}$ such that for any $|\Psi\rangle \in \mathcal{H}_{in}$, we have that applying Q to $|\Psi\rangle \otimes |0\rangle$ and then measuring $\mathcal{H}_{succ}$ in the computational basis gives outcome 1 (success) with probability p (not $\geq p$). Let $|\tilde{\phi}_{succ}\rangle$ denote the post measurement state in $\mathcal{H}_{out}$ in that case.

Consider the following algorithm (depending on a parameter q):

1. *Let $|\Psi\rangle$ denote the input of the algorithm (in $\mathcal{H}_{in}$)*

2. *Initialize* $\mathcal{H}_{anc}$ *with* $|0\rangle$.
3. *Apply* Q.
4. *Measure* $\mathcal{H}_{succ}$ *in the computational basis.*
5. *If the outcome is* 1*, exit (successfully).*
6. *Apply* $Q^\dagger$.
7. *Apply* FLIP *to* $\mathcal{H}_{anc}$ *where* $\text{FLIP}\,|0\rangle := |0\rangle$ *and* $\text{FLIP}\,|x\rangle := -|x\rangle$ *for* $x \neq 0$.
8. *Go to 3. (But at most* q *times.)*

Then for a suitable $q \in poly(1/p)$*, we have that*

- *The probability that* R *exits successfully is overwhelming.*
- *The post measurement state in* $\mathcal{H}_{out}$ *in that case is* $|\tilde{\phi}_{succ}\rangle$.

QZKAna

This lemma can be used to construct the simulator S from S_1: First, we purify S_1 (i.e., replace measurements by CNOTs on ancilla qubits in $\mathcal{H}_{anc}$ initialized with $|0\rangle$), resulting in Q. Then S runs R and outputs the state $|\tilde{\phi}_{succ}\rangle$.

Notice that in the classical case, it is sufficient that S_1 succeeds with probability $\geq p$ (possibly dependent on the auxiliary input z), while in the quantum case, we need that the simulator S_1 succeeds with a probability p that is independent of the auxiliary input ρ.

Notice further that the above lemma only covers the case where the simulation is perfect. There is a variant of that lemma which also covers the case where S_1 produces a state that has negligible trace distance from $\langle P(x,w), V^*(x,\rho)\rangle$. This allows to cover a wider range of protocols and even protocols that are only computationally QZK.

Further reading: An overview over zero-knowledge proofs in the classical case can be found in [Gol01, Chapter 4]. For quantum zero-knowledge, see [Wat09].

16 A physical view on quantum mechanics

Physical

Note: In this section, we will use mathematics in a non-rigorous way. That is, we will *implicitly* assume that functions are continuous or differentiable whenever needed, that Dirac deltas[9] can be treated like ordinary functions, and more. This is common in theoretical physics.

Throughout this lecture, we have been using an abstraction of quantum mechanics that represents physical systems as consisting of a finite set of classical states (e.g., $|0\rangle$, $|1\rangle$, etc.) that can occur in superposition and on which we can perform various quantum operations. However, it is not immediately obvious how this related to the physical world. For example the location of a particle is a continuous variable. How is the speed of a particle modeled? How do forces between particles come into play in determining the behavior of particles? In this section, we will shed some light on these question. However, we can only give an idea here, for deeper understanding a full course or textbook on quantum mechanics is needed.

[9]The Dirac delta δ is a function $\delta(x)$ such that $\delta(0) = \infty$ and $\delta(x) = 0$ elsewhere. It can be informally seen as a limit of functions δ_n where $\int_{-\infty}^{\infty} \delta_n(x)dx = 1$ for all n, and $\delta_n(x) \to 0$ for all $x \neq 0$. I.e., a limit of functions that get more and more concentrated around 0.

The wave function. We start by discussing how the state of a single particle can be represented. Classically, a particle can have a single position $x \in \mathbb{R}^3$ at any given time $t \in \mathbb{R}$. Thus, classically, we would describe the time evolution of a particle by a function $x : \mathbb{R} \to \mathbb{R}^3$ such that $x(t)$ is the location of the particle at time t. In a quantum setting, however, the location of a particle is not determined, the particle can be in a superposition of many different locations. Thus, at any time, the state of the particle is described by a function $\psi : \mathbb{R}^3 \to \mathbb{C}$, where $\psi(x)$ is the amplitude of the state being at location x. Or, to model the fact that the location may depend on the time, we add another parameter: $\psi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{C}$, where $\psi(x, t)$ is the amplitude of the state being at location x at time t.

In the following, to keep things simpler, we will consider only one-dimensional space, i.e., the particle can be found somewhere on a line. Then $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, and $\psi(x, t)$ denotes the amplitude of the particle being at position $x \in \mathbb{R}$ at time t.

Definition 58 (Wave function) *A (one-dimensional one-particle time-dependent)* wave function *is a function $\psi : \mathbb{R}\times\mathbb{R} \to \mathbb{C}$ such that for all t, we have $\int_{-\infty}^{\infty} |\psi(x, t)|^2 dx = 1$.*

The wave function tells us where the particle can be found with what probability. For example, if we measure (at time t) whether the particle is somewhere in the interval $[a, b]$, the probability of yes is $\int_a^b |\psi(x, t)|^2 dx$. But the wave function encodes more then just the probability. For example, the phase of $\psi(x, t)$ encodes additional information such as the momentum of the particle, and is relevant for many quantum mechanical effects such as interference.

Time evolution. Now in the classical setting, the position of a particle at a certain point, together with it's velocity, determines it's future evolution. (Assuming that we know the potential in which the particle moves.) Namely, if the potential is $V(x, t)$, then $-\frac{\partial V(x,t)}{\partial x}$ is the force that acts on the particle at position x. And hence, if the particle has mass m, $-\frac{1}{m}\frac{\partial V(x,t)}{\partial x}$ is the acceleration of the particle. We can write this as a differential equation: $\frac{\partial^2 x(t)}{\partial t^2} = -\frac{1}{m}\frac{\partial V(x,t)}{\partial x}$. This determines $x(t)$ for all $t > t_0$ if x and $\frac{\partial x}{\partial t}$ are given at $t = t_0$.

Similarly, in the quantum case, if $\psi(x, t_0)$ is given for all x, then there is a differential equation that determines the future evolution of the particle.

Definition 59 (Schrödinger equation) *The (time-dependent one-dimensional one-particle)* Schrödinger equation *for a particle of mass m in a potential $V : \mathbb{R} \to \mathbb{R}$ is given by*

$$i\hbar\frac{\partial\psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x, t)}{\partial x^2} + V(x, t).$$

Here $\hbar$ is the so-called reduced Planck constant*, $\hbar \approx 6.62606957 \cdot 10^{-34} Js$ (Joule seconds).*

So, given an initial state $\psi(x, t_0)$ for all x, and the potential $V(x, t)$ for all x, t, we can decide what $\psi(x, t)$ is for all $t \geq t_0$ by solving the Schrödinger equation.

Before we have a look at this, let us have a look at the components of the Schrödinger equation. The term $-\hbar^2 \frac{\partial^2 \psi(x,t)}{\partial x^2}$ turns out to be the squared momentum of the particle (for reasons that are beyond the scope of this exposition). And since the momentum is the velocity times the mass, $-\frac{\hbar^2}{m^2} \frac{\partial^2 \psi(x,t)}{\partial x^2}$ is the squared velocity v^2. And the kinetic energy K is $K = \frac{mv^2}{2}$. Hence the kinetic energy is described by the term $-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2}$. And since we add to that the potential, the right hand side of the Schrödinger is nothing else but the energy of the particle at position x and time t. The operator that computes the energy given the wave function is called the *Hamiltonian* $\hat{H}(t)$, in our case $\hat{H}(t)\psi := -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x,t)\psi(x,t)$.[10] So, the Schrödinger equation tells us that the differential of ψ (in the variable t) is $-\frac{i}{\hbar}\hat{H}(t)\psi(x,t)$, i.e., the energy times $-\frac{i}{\hbar}$. If $\hat{H}(t)\psi(x,t)$ were a constant E, we could then easily solve the differential equation $\frac{\partial \psi(x,t)}{\partial t} = -\frac{iE}{\hbar}$ and see that $\psi(x,t) = \psi(x,t_0)e^{-iEt/\hbar}$. Since $e^{-iEt/\hbar}$ as a function of t has an oscillating behavior (if you don't know what this function looks like, I recommend to visualize it), this tells us that the wave function oscillates over time, and oscillates faster if the energy E is higher. (This matches the fact that light with higher frequency has higher energy.)

Now, in principle, we can derive the behavior of a particle in any potential. (And easily generalize this to many particle systems.) However, in practice the Schrödinger equation is quite complex to solve, and therefore physicists often use the so-called time-independent Schrödinger equation to get a better understanding of the system and to solve the time-dependent case, as described in the following.

Time-independent Schrödinger equation. We now consider the special case where the potential does not depend on the time t, i.e., $V(x,t) = V(x)$ for all x, t. In this case, the Hamiltonian does not depend on t either, and we write $\hat{H}\psi := -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x)$. Assume further that we are given $\psi_0(x) := \psi(x,t_0)$ for some t_0, i.e., the initial state at time t_0. It turns out that then we can then write ψ_0 as a linear combination of eigenvectors of $\hat{H}$ (possibly an infinite number of them). More precisely, we first solve the so-called time-independent Schrödinger equation:

Definition 60 (Time-independent Schrödinger equation) *The (one-particle one-dimensional)* time-independent Schrödinger equation *for a time-independent potential $V(x)$ is given by $\hat{H}\psi_0 = E\psi_0$ or equivalently $-\frac{\hbar^2}{2m} \frac{\partial^2 \psi_0(x)}{\partial x^2} + V(x)\psi_0(x) = E\psi_0(x)$. In this equation, $\hbar, m, V$ are given, and $E \in \mathbb{R}$ and $\psi_0 : \mathbb{R} \to \mathbb{C}$ are to be found.*

What does the time-independent Schrödinger equation have to do with the time-dependent Schrödinger equation from above? Consider a solution ψ_0, E to the time-

[10] In more general cases (e.g., multiple particles etc.), the Schrödinger equation is suitably generalized as follows: Let M be the set of all possible classical states of the system. (E.g., for two particles in three-dimensional space, $M := \mathbb{R}^3 \times \mathbb{R}^3$.) A wave function is $\psi : M \times \mathbb{R} \to \mathbb{C}$, and the Hamiltonian $\hat{H}(t)$ is an operator that, given a function $\psi : M \to \mathbb{C}$ returns a function $E : M \to \mathbb{R}$ where $E(m)$ is the energy at "position" m. ($\hat{H}(t)$ operates on ψ for each t individually.) And then the time evolution is then in this generic setting described via the Schrödinger equation $i\hbar \frac{\partial \psi(x,t)}{\partial t} = \hat{H}(t)\psi(x,t)$.

independent Schrödinger equation. It is then easy to verify that $\psi(x,t) := \psi_0(x)e^{-iEt/\hbar}$ satisfies the time-dependent Schrödinger equation (try it!).

So, generally, if we can write $\psi_0(x) = \sum_E \alpha_E \psi_E(x)$ or $\psi_0(x) = \int \alpha_E \psi_E(x) dE$ where ψ_E, E are solutions to the time-independent Schrödinger equation, then $\psi(x,t) := \sum_E \alpha_E \psi_E(x) e^{-iEt/\hbar}$ or $\psi(x,t) := \int \alpha_E \psi_E(x) e^{-iEt/\hbar} dE$ is the solution of the time-dependent Schrödinger equation with initial condition $\psi(x,t_0) = \psi_0(x)$.

Thus, finding solutions to the time-independent Schrödinger equation is crucial for the understanding of the time-evolution of a system with unchanging potential (or, in the more general case, of a system with unchanging Hamiltonian $\hat{H}$).

Example: Infinite potential well. We now apply what we have learned about the Schrödinger equation to a simple setting, the "infinite potential well" or "infinite square well". In this example, we assume that the potential is zero within a certain area, and infinite[11] outside of that area. That is,

$$V(x) = \begin{cases} 0, & (0 < x < L) \\ \infty, & (\text{otherwise}) \end{cases}$$

This corresponds to a particle inside a (one-dimensional) box, it can freely move inside, but can never leave the box.

First, consider a classical setting: A classical particle will be able to have any position inside the box, and can have arbitrary speed v (until it hits the wall, whereupon the speed becomes $-v$). In particular, the kinetic energy $E = \frac{mv^2}{2}$ can take any non-negative value.

Now intuitively, we would expect that in the quantum setting, the possible states of the particle in the well would be any superposition of the classical possibilities (i.e., $\int_0^L \int_0^\infty \alpha_{x,E} |x,E\rangle \, dE \, dx$ where $|x,E\rangle$ stands for a wave function where the particle has location x and kinetic energy E). Yet, we will see that the situation is quite different in the quantum setting!

As explained above, in the quantum setting, we first need to find solutions to the time-independent Schrödinger equation. I.e., we need to solve $-\frac{\hbar^2}{2m}\frac{\partial^2 \psi_0(x)}{\partial x^2} + V(x)\psi_0(x) = E\psi_0(x)$.

We will assume $E > 0$. (A similar calculation shows that there is no solution with $E < 0$.) First, since $V(x) = \infty$ for $x \notin (0,L)$, we have $\psi_0(x) = 0$ for $x \notin (0,L)$. (We do not give a rigorous mathematical argument for this, but intuitively/physically, we expect the particle not to be at a place of infinite potential.) Furthermore, standard techniques for solving differential equations show that the solutions of the equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2 \psi_0(x)}{\partial x^2} = E\psi_0(x) \qquad (*)$$

are of the form $A\sin(kx) + B\cos(kx)$ with $k := \frac{\sqrt{2mE}}{\hbar}$. Since $V(x) = 0$ for $x \in (0,L)$, we have that on $(0,L)$, the solutions to the Schrödinger equation coincide with the solutions

[11] Allowing infinite potentials is, of course, a contradiction to the fact that we consider V to be a function $\mathbb{R} \to \mathbb{R}$, and not well-defined. A more rigorous treatment could, e.g., consider a sequence of potentials which converges to 0 in $(0,L)$ and to ∞ outside.

of $(*)$. Since furthermore, the solutions to the Schrödinger equation have to satisfy $\psi_0(0) = 0$, we have $B = 0$. And since $\psi_0(L) = 0$, we need $\sin(kx) = 0$. This implies that $k = n\pi/L$ for integers $n \geq 0$. Furthermore, $k = 0$ is excluded because then $\psi_0(x) = A \cdot 0$, which cannot satisfy $\langle \psi_0, \psi_0 \rangle = \int_{-\infty}^{\infty} \psi_0^*(x)\psi_0(x)\,dx = 1$. So the Schrödinger equation only has solutions for $E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2}{2mL^2} n^2 =: E_n$ with $n > 0$. And in each such case, $\psi_0 = |n\rangle := A_n \sin(\frac{n\pi x}{L})$ for a suitable normalization factor A_n.

So, summarizing: All solutions $\psi(x,t)$ to the time-dependent Schrödinger equation are of the form $\psi(x,t_0) = \sum_{n\geq 1} \alpha_n |n\rangle$, and then $\psi(x,t) = \sum_{n\geq 1} \alpha_n e^{-iE_n t/\hbar} |n\rangle$. This fully describes all possible time-evolutions of the state of the particle.

In particular, we see that the energy of the particle will always be a multiple of $\frac{\hbar^2\pi^2}{2mL^2}$. That is, only specific energies are possible! This is in stark contrast to the classical case where any $E > 0$ is possible. (Of course, we can have a superposition of different energies, so that the average energy is any possible value. But if we were to measure the energy of the system, we would always get one of the values E_n.)

This energy quantisation is not an artifact of our infinite potential. It also occurs in the similar but more complicated analysis of an electron in the electric field of the nucleus of an atom. There the electron will also be able to only take certain energies. (This is the reason why photons emitted from atoms can have only certain energies – the energies of the photons correspond to the differences between the different energy levels.)

Note also that $E = 0$ is also excluded. In other words, the kinetic energy of a particle in a box can never be zero – the particle never rests. This is related to Heisenberg's uncertainty relation: since we know where the particle is (in the box), there must be a certain small uncertainty about its momentum and hence its velocity. So the velocity cannot be zero.

Link to our formalism. So if in "real" physics, the state of a system is described as a wave function, why do we treat quantum mechanical systems so differently in quantum information theory (i.e., in this lecture)? Namely, we treat quantum states as elements from a finite dimensional Hilbert space. And operations on these are unitary transformations. In fact, this is not really different from the wave function formalism. What we do is simply to give names to individual orthogonal solutions of the Schrödinger equation (e.g., we write the wave function corresponding to energy E_n as $|n\rangle$). And it turns out that for any Hamiltonian (which must be a Hermitian operator), the Schrödinger equation then predicts a unitary time evolution on some initial wave function $|\Psi\rangle$. (I.e., for any Hamiltonian $\hat{H}$ and any t, there is a unitary transformation U such that for any solution ψ of the time-dependent Schrödinger equation, $|\psi_t\rangle = U|\psi_0\rangle$ where $|\psi_t\rangle := \psi(\cdot, t)$.) So our formalism captures the laws of quantum mechanics without describing details that are not important for our specific case.

A Linear Algebra

In the following, we refresh the basic definitions from linear algebra that will be needed during the course. In all definitions, we will restrict our attention to the finite dimensional

case only.

Hilb

Definition 61 (Hilbert space) *The n-dimensional* Hilbert space *is $\mathbb{C}^n$, the n-dimensional complex vector space.*[12]*.*

$\mathbb{C}^n$ is endowed with the following inner product*:*

$$\langle \Psi, \Phi \rangle := \sum_{i=1}^{n} \Psi_i^* \Phi_i$$

where x^* is the complex conjugate of x.[13]

The (Euclidean) norm $\|\cdot\|$ *is defined by*

$$\|\Psi\| := \sqrt{\langle \Psi, \Psi \rangle} = \sqrt{\sum_{i=1}^{n} \Psi_i^* \Psi_i} = \sqrt{\sum_{i=1}^{n} |\Psi_i|^2}.$$

We call two vectors Ψ and Φ orthogonal *if $\langle \Psi, \Phi \rangle = 0$. We call Ψ orthogonal to a subspace $V \subseteq \mathbb{C}^n$ if Ψ is orthogonal to all $x \in V$.*

Furthermore, we call a vector *normalised* if $\|\Psi\| = 1$, and we call a *set* of vectors *orthogonal* if they are pairwise orthogonal, and we call a set of vectors *orthonormal* if they are all normalised and pairwise orthogonal.

ConjTrans

Definition 62 (Conjugate transpose) *Given a matrix $M \in \mathbb{C}^{n \times m}$, we define $M^\dagger$ as the complex conjugate of the transposition of M, i.e., $(M^\dagger)_{ij} = (M_{ji})^*$. (This is the analogue of transposition.)*

We have $(M^\dagger)^\dagger = M$ and $\langle Mx, y \rangle = \langle x, M^\dagger y \rangle$ (and vice-versa).

Dirac

Definition 63 (Dirac notation) *In the* Dirac notation*, a vector Ψ in $\mathbb{C}^n$ is written $|\Psi\rangle$. By $\langle \Psi|$ we denote the function mapping $|\Phi\rangle$ to $\langle \Psi, \Phi \rangle$ (or equivalently: $\langle \Psi|$ is the row vector $|\Psi\rangle^\dagger$).*

In particular, we can now write $\langle \Psi | \Phi \rangle$ for the inner product $\langle \Psi, \Phi \rangle$. And for the projection P_V onto $V = \operatorname{span} \Psi$ we write $P_V = |\Psi\rangle\langle\Psi|$. (Try it out and evaluate $P_V|\Phi\rangle$!)

Trace

Definition 64 (Trace) *The* trace $\operatorname{tr} M$ *of a matrix $M \in \mathbb{C}^{n \times n}$ is $\sum_i M_{ii}$.*

The trace can also be computed as $\sum_i \langle i|M|i\rangle$ for any orthonormal basis $|1\rangle, \ldots, |n\rangle$ of $\mathbb{C}^n$.

Herm

Definition 65 (Hermitian matrices) *A matrix $M \in \mathbb{C}^{n \times n}$ is called Hermitian, if $M = M^\dagger$. (This is the analogue of symmetric matrices.)*

A Hermitian matrix M can be diagonalised, i.e., there is an orthonormal basis $|1\rangle, \ldots, |n\rangle$ such that $M = \sum_i \lambda_i |i\rangle\langle i|$ where λ_i are the eigenvalues of M.

PosMat

[12] Or any complex vector space isomorphic to $\mathbb{C}^n$

[13] I.e., $(a + bi)^* = a - bi$.

Definition 66 (Positive matrices) *A matrix $M \in \mathbb{C}^{n\times n}$ is* positive *if for all $|\Psi\rangle \in \mathbb{C}^n$ we have $\langle\Psi|M|\Psi\rangle \geq 0$.*

Note that positive is meant in the sense of positive semidefinite (or nonnegative), i.e., we allow, e.g., $M = 0$.

A positive Hermitian matrix has only nonnegative eigenvalues $\lambda_i \geq 0$.

AbsMat

Definition 67 (Absolute value of a matrix) *For a positive Hermitian matrix M, let $\sqrt{M}$ be the positive matrix satisfying $(\sqrt{M})^\dagger(\sqrt{M}) = M$. For a (not necessarily positive or Hermitian) matrix M, we define $|M| := \sqrt{M^\dagger M}$.*

The matrix $|M|$ is always positive Hermitian. For a positive Hermitian matrix M, we have $|M| = M$. For a diagonal matrix M, we get $|M|$ by taking the absolute value of every element on the diagonal.

For a positive Hermitian M, we can compute $\sqrt{M}$ by first diagonalising M as $UDU^\dagger$ (with unitary U and diagonal D), and then computing $\sqrt{D}$ (by taking the square root of each diagonal element individually) and then computing $\sqrt{M} = U\sqrt{D}U^\dagger$. Since for a matrix M, we have that $M^\dagger M$ is positive Hermitian, we can use this procedure to compute $|M|$.

Unitary

Definition 68 (Unitary matrices) *A matrix $M \in \mathbb{C}^{n\times n}$ is unitary if $M^\dagger M = MM^\dagger = I$ where I is the identity matrix. (Unitary matrices are the analogue to rotation matrices.)*

Note: If M is unitary, then $\|Mx\| = \|x\|$ and $\langle Mx, My\rangle = \langle x, y\rangle$.

Proj

Definition 69 (Projections) *A matrix $M \in \mathbb{C}^{n\times n}$ is a* projection *if for all x we have $MMx = Mx$ (or equivalently, $MM = M$).*

The orthogonal projection *P_V onto a subspace $V \subseteq \mathbb{C}^n$ is defined by $P_V(u + v) = v$ where $v \in V$ and u is orthogonal to V. (Note that any state $x \in \mathbb{C}^n$ can be represented uniquely as such a sum $x = u + v$.)*

For a one-dimensional subspace $V = \text{span}\{v\}$ with $\|v\| = 1$, we have that $P_V x = v\langle v, x\rangle$.

SingVal

Lemma 30 (Singular value decomposition) *For any square matrix $A \in \mathbb{C}^{n\times n}$, there are unitary matrices $U, V \in \mathbb{C}^{n\times n}$ and a diagonal matrix $D \in \mathbb{C}^{n\times n}$ with only nonnegative real entries such that $A = UDV$.*

Tensor

Definition 70 (Tensor product) *Given two Hilbert spaces $\mathbb{C}^n, \mathbb{C}^m$ with orthonormal bases $B_1 = \{|i\rangle\}, B_2 = \{|j\rangle\}$, the* tensor product *(or* Kronecker product*) $\mathbb{C}^n \otimes \mathbb{C}^m$ is the Hilbert space $\mathbb{C}^{nm}$ with basis $B_1 \times B_2 = \{|i, j\rangle\}$.*[14]

[14]There exists a more general category theoretical definition using a universal property, but for our purposes this specialisation is sufficient.

Given two vectors $|\Psi_1\rangle = \sum_i \alpha_i |i\rangle \in \mathbb{C}^n$ *and* $|\Psi_2\rangle = \sum_j \beta_j |j\rangle \in \mathbb{C}^m$*, their tensor product is given by*

$$|\Psi_1\rangle \otimes |\Psi_2\rangle = \sum_{i,j} \alpha_i \beta_j |i,j\rangle \in \mathbb{C}^n \otimes \mathbb{C}^m.$$

Given two linear operations $M_1 : \mathbb{C}^n \to \mathbb{C}^n$ *and* $M_2 : \mathbb{C}^m \to \mathbb{C}^m$*, we define the linear operation* $M_1 \otimes M_2$ *to be the unique linear operation satisfying*

$$(M_1 \otimes M_2)|i,j\rangle = (M_1|i\rangle) \otimes (M_2|j\rangle).$$

Further reading: [NC00, Section 2.1]

References

[BB84] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public-key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing 1984*, pages 175–179. IEEE Computer Society, 1984.

[Gol01] Oded Goldreich. *Foundations of Cryptography – Volume 1 (Basic Tools).* Cambridge University Press, August 2001. Previous version online available at `http://www.wisdom.weizmann.ac.il/~oded/frag.html`.

[LC99] H. K. Lo and H. F. Chau. Unconditional Security of Quantum Key Distribution over Arbitrarily Long Distances. *Science*, 283(5410):2050, 1999. Online available at `http://arxiv.org/abs/quant-ph/9803006`.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000.

[Pei16] Chris Peikert. A decade of lattice cryptography. *Foundations and Trends in Theoretical Computer Science*, 10(4):283–424, 2016. `https://web.eecs.umich.edu/~cpeikert/pubs/lattice-survey.pdf`.

[Reg09] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34:1–34:40, September 2009.

[Ren05] Renato Renner. *Security of Quantum Key Distribution.* PhD thesis, ETH Zurich, September 2005. Available at `http://arxiv.org/abs/quant-ph/0512258v2`.

[SMWF+07] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, A. Zeilinger, and H. Weinfurter. Experimental Demonstration of Free-Space Decoy-State Quantum Key Distribution over 144 km. *Physical Review Letters*, 98(1):10504, 2007. Online available at `http://www.quantum.at/uploads/media/PRL_98__010504__2007_.pdf`.

[SWV+09] D. Stucki, N. Walenta, F. Vannel, R. T. Thew, N. Gisin, H. Zbinden, S. Gray, C. R. Towery, and S. Ten. High rate, long-distance quantum key distribution over 250 km of ultra low loss fibres. *New Journal of Physics*, 11(7):075003, 2009.

[Unr21] Dominique Unruh. Lecture "Quantum Cryptography", spring 2021. Webpage is `https://kodu.ut.ee/~unruh/courses/qc/2021/`.

[Wat09] John Watrous. Zero-knowledge against quantum attacks. *SIAM J. Comput.*, 39(1):25–58, 2009.

[Wik] Wikipedia contributors. Wikipedia, the free encyclopedia (english edition). `http://en.wikipedia.org`.

Index