$$= \sum_{j,k} \alpha_j \bar{\alpha}_k \langle j|k\rangle \mathrm{tr}(E_i|\psi_k\rangle\langle\psi_j|) = \sum_j |\alpha_j|^2 \mathrm{tr}(E_i|\psi_j\rangle\langle\psi_j|)\,,$$

which coincides with $\mathrm{tr}(E_i\rho_B)$, proving the claim. □

It is easy to see that applying a *unitary matrix* to the part that is traced out does not change the outcome of the partial trace, and, similarly, *measuring* the part that is traced out but ignoring the outcome of the measurement does not change the outcome of the partial trace. Thus, one cannot influence (part of) a state by "doing something" to another.

A physical process is most generally described by a linear mapping $\mathfrak{T}$ characterized as follows; we will see later (see Proposition 4.9) that even a measurement can be understood as such a general quantum operation.

Definition 4.3. *A* **quantum operation** $\mathfrak{T} : \mathcal{D}(\mathcal{H}_A) \to \mathcal{D}(\mathcal{H}_{A'})$ *is specified by two "default" density matrices* $\tau_E \in \mathcal{D}(\mathcal{H}_E)$ *and* $\tau_{A'} \in \mathcal{D}(\mathcal{H}_{A'})$ *(like* $|0\cdots0\rangle\langle0\cdots0|$*) and by a unitary matrix* $U \in \mathcal{U}(\mathcal{H}_A \otimes \mathcal{H}_E \otimes \mathcal{H}_{A'})$*, and acts as*

$$\mathfrak{T}(\rho_A) = \mathrm{tr}_{AE}\big(U(\rho_A \otimes \tau_E \otimes \tau_{A'})U^\dagger\big) \in \mathcal{D}(\mathcal{H}_{A'})\,.$$

It follows from previous observations that any quantum operation as defined above is trace preserving and completely positive, where the latter means that its natural extension $\mathfrak{T} \otimes \mathbb{I}_B$ maps positive semi-definite operators (acting on $\mathcal{H}_A \otimes \mathcal{H}_B$) to positive semi-definite operators (acting on $\mathcal{H}_{A'} \otimes \mathcal{H}_B$) for any $\mathcal{H}_B$. Furthermore, one can show that any **completely-positive trace-preserving** (in short: **CPTP**) map $\mathfrak{T}$ from $\mathcal{E}nd(\mathcal{H}_A)$ to $\mathcal{E}nd(\mathcal{H}_{A'})$ is a quantum operation in the above sense. The equivalence of quantum operations (as defined above) and CPTP maps is knows as *Stinespring's Dilation Theorem*.

4.2 Distance Between States

Since the density matrix uniquely described the behavior of a quantum system, two systems whose respective states are given by the same density matrix behave in exactly the same way. Now we want to be able to say that if the density matrices of two states are *close* then the states behave *similarly* and are hard to distinguish. For measuring the closeness of two density matrices we introduce a suitable distance for density matrices.

Definition 4.4. *The* **trace norm** *of a Hermitian matrix* A *is defined as* $\|A\|_{tr} := \mathrm{tr}|A|$*, where* $|A| := \sqrt{A^\dagger A}$ *is the positive semi-definite square root of* $A^\dagger A$*. The* **trace distance** *of two density matrices* ρ *and* σ *in* $\mathcal{D}(\mathcal{H})$ *is defined as* $\delta(\rho,\sigma) := \frac{1}{2}\|\rho - \sigma\|_{tr}$*.*[14]

It is not hard to see that $\|A\|_{tr}$ can also be understood as $\|A\|_{tr} = \sum_i |\lambda_i|$, where the λ_i's are the (real) eigenvalues of A.

Theorem 4.5. *Let* P *and* Q *be the probability distributions obtained by measuring two respective quantum states* $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ *using POVM* $E = \{E_i\}_i$*. Then the statistical distance between* P *and* Q *is upper bounded by the trace distance between* ρ *and* σ*:* $\mathrm{SD}(P,Q) \le \delta(\rho,\sigma)$*.*

Proof. Note that

$$\mathrm{SD}(P,Q) = \frac{1}{2}\sum_i \big|P(i) - Q(i)\big| = \frac{1}{2}\sum_i \big|\mathrm{tr}\big(E_i\rho\big) - \mathrm{tr}\big(E_i\sigma\big)\big| = \frac{1}{2}\sum_i \big|\mathrm{tr}\big(E_i(\rho-\sigma)\big)\big|\,,$$

[14] The factor $\frac{1}{2}$ is for normalization purposes: it ensures that $\delta(\rho,\sigma) \le 1$.

where the matrix E_i is positive semi-definite, and as such has non-negative diagonal elements, and $\sum_i E_i = \mathbb{I}$. By considering $\rho - \sigma$ to be in diagonal form (which we may do without loss of generality), it follows that the above is bounded by

$$\leq \frac{1}{2}\sum_i \mathrm{tr}\big(E_i|\rho-\sigma|\big) = \frac{1}{2}\mathrm{tr}\big(\mathbb{I}|\rho-\sigma|\big) = \delta(\rho,\sigma)\,,$$

which was to be shown. □

From this, from Theorem 4.6 below, and from Lemma B.2, it follows that any physical processing of ρ respectively σ allows to distinguish ρ and σ with advantage at most $\delta(\rho,\sigma)$. In other words, if $\delta(\rho,\sigma)$ is small then the quantum state ρ behaves exactly as σ except with small "error" probability.

Theorem 4.6. *For any quantum operation (i.e., CPTP map) $\mathfrak{T} : \mathcal{D}(\mathcal{H}_A) \to \mathcal{D}(\mathcal{H}_{A'})$ and for any density matrices $\sigma, \rho \in \mathcal{D}(\mathcal{H}_A)$*

$$\delta(\mathfrak{T}(\rho),\mathfrak{T}(\sigma)) \leq \delta(\rho,\sigma)\,.$$

The proof is given as an exercise.

In case of *pure* states, the trace distance is determined by their inner product:

Proposition 4.7. *For $|\varphi\rangle, |\psi\rangle \in \mathcal{H}$: $\delta\big(|\varphi\rangle\langle\varphi|, |\psi\rangle\langle\psi|\big) = \sqrt{1-|\langle\varphi|\psi\rangle|^2}$.*

Proof. We can choose an orthonormal basis $\{|0\rangle, |1\rangle, \cdots, |d-1\rangle\}$ of $\mathcal{H}$ with $|\varphi\rangle = \omega|0\rangle$ and $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ such that α_0 and α_1 are real (and $|\omega| = 1$ and $|\alpha_0|^2 + |\alpha_1|^2 = \alpha_0^2 + \alpha_1^2 = 1$). As both sides of the equation to be proven are invariant under multiplying $|\varphi\rangle$ with $\bar{\omega}$, we may assume without loss of generality that $\omega = 1$. It follows that $1 - |\langle\varphi|\psi\rangle|^2 = 1 - \alpha_0^2 = \alpha_1^2$ and thus $\sqrt{1-|\langle\varphi|\psi\rangle|^2} = |\alpha_1|$. On the other hand, when expressing $|\varphi\rangle\langle\varphi| - |\psi\rangle\langle\psi|$ in this basis we see that

$$\delta\big(|\varphi\rangle\langle\varphi|, |\psi\rangle\langle\psi|\big) = \frac{1}{2}\mathrm{tr}\big(\big||\varphi\rangle\langle\varphi| - |\psi\rangle\langle\psi|\big|\big) = \frac{1}{2}\mathrm{tr}\left(\begin{vmatrix} 1-\alpha_0^2 & -\alpha_0\alpha_1 \\ -\alpha_0\alpha_1 & -\alpha_1^2 \end{vmatrix}\right) = \frac{1}{2}\mathrm{tr}\left(\begin{vmatrix} \alpha_1^2 & -\alpha_0\alpha_1 \\ -\alpha_0\alpha_1 & -\alpha_1^2 \end{vmatrix}\right)$$

where the matrix in the right-hand-side expression has eigenvalues $\pm\alpha_1$ (which can easily be seen by computing its characteristic polynomial). It follows that $\delta\big(|\varphi\rangle\langle\varphi|, |\psi\rangle\langle\psi|\big) = |\alpha_1|$. □

4.3 Combining Classical and Quantum Information

We will consider *mixed* systems, which are partly classical and partly quantum, i.e., which contain classical and quantum information. In order to have one common language to talk about the classical, the quantum and the joint system, we adopt the Hilbert space formalism also for classical systems, and describe classical information as quantum states. For a finite set $\mathcal{X}$, we identify the (classical) element $x \in \mathcal{X}$ with the (pure) quantum state $|x\rangle \in \mathcal{H}_X$, where $\mathcal{H}_X = \mathbb{C}^{|\mathcal{X}|}$ and $\{|x\rangle\}_{x\in\mathcal{X}}$ is a fixed orthonormal basis of $\mathcal{H}_X$, typically the canonical basis. Note that such an "encoding" can be "decoded" simply by measuring $|x\rangle$ in the basis $\{|x\rangle\}_{x\in\mathcal{X}}$: as a result, x is observed with probability 1.

Consider now a random variable X over $\mathcal{X}$ with distribution P_X. Using the above "encoding", the classical-system-made-quantum is in state $|x\rangle$ with probability $P_X(x)$. As discussed in Section 1.6, such a randomized state is described by the density matrix[15]

$$\rho_X = \sum_x P_X(x)|x\rangle\langle x|\,.$$

[15] Note that we are using the same name, X, for the classical random variable and for the corresponding quantum system.

Conversely, any state ρ_X of such a form is called **classical** (with respect to $\{|x\rangle\}_x$). Note that for two classical states $\rho_X = \sum_x P_X(x)|x\rangle\langle x|$ and $\rho'_X = \sum_x P'_X(x)|x\rangle\langle x|$, the trace distance $\delta(\rho_X, \rho'_X)$ coincides with the statistical distance $\mathrm{SD}(P_X, P'_X)$ between the corresponding distributions P_X and P'_X.

Now, we consider a mixed system, consisting of a random variable X and a quantum system E, where the state of E depends on X in that if X takes on value x then the state of E is given by density matrix $\rho_{\mathsf{E}|X=x} \in \mathcal{D}(\mathcal{H}_{\mathsf{E}})$. Similar to above, this mixed system can then be described by the density matrix

$$\rho_{X\mathsf{E}} = \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_{\mathsf{E}|X=x}\,.$$

In this case we say that the quantum state $\rho_{X\mathsf{E}}$ has a **classical** X (with respect to $\{|x\rangle\}_x$). The quantum system E alone is then described by

$$\rho_{\mathsf{E}} = \mathrm{tr}_X(\rho_{X\mathsf{E}}) = \sum_x P_X(x)\rho_{\mathsf{E}|X=x}\,.$$

More generally, for any event $\mathcal{E}$, well-defined by $P[\mathcal{E}|X=x]$ for any $x \in \mathcal{X}$, we write[16]

$$\rho_{X\mathsf{E}|\mathcal{E}} = \sum_x P_{X|\mathcal{E}}(x)|x\rangle\langle x| \otimes \rho_{\mathsf{E}|X=x} \qquad \text{and} \qquad \rho_{\mathsf{E}|\mathcal{E}} = \mathrm{tr}_X(\rho_{X\mathsf{E}|\mathcal{E}}) = \sum_x P_{X|\mathcal{E}}(x)\rho_{\mathsf{E}|X=x}\,,$$

which describe the mixed and the single quantum system *given that* $\mathcal{E}$ *occurs*.

This formalism naturally extends to states that depend on several, possibly dependent, random variables X, Y etc. To simplify notation, we often write ρ^x_{E} instead of $\rho_{\mathsf{E}|X=x}$.[17]

Lemma 4.8. *For* $\rho_{X\mathsf{E}}, \rho'_{X\mathsf{E}} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_{\mathsf{E}})$ *with the* same *classical* X *(in that* $P_X = P'_X$*),*

$$\delta\big(\rho_{X\mathsf{E}}, \rho'_{X\mathsf{E}}\big) = \sum_x P_X(x)\delta\big(\rho_{\mathsf{E}|X=x}, \rho'_{\mathsf{E}|X=x}\big) = E_{x\leftarrow X}\big[\delta\big(\rho_{\mathsf{E}|X=x}, \rho'_{\mathsf{E}|X=x}\big)\big]\,.$$

Proof. Note that $\rho_{X\mathsf{E}} - \rho'_{X\mathsf{E}} = \sum_x P_X(x)|x\rangle\langle x| \otimes (\rho^x_{\mathsf{E}} - \rho'^x_{\mathsf{E}})$ is a block-diagonal matrix with the (diagonalizable) blocks $P_X(x)(\rho^x_{\mathsf{E}} - \rho'^x_{\mathsf{E}})$ on the diagonal. Thus, by grouping the summation over the (norms of the) eigenvalues for each block, factoring out $P_X(x)$, and summing over the blocks, we get that indeed $\delta\big(\rho_{X\mathsf{E}}, \rho'_{X\mathsf{E}}\big) = \sum_x P_X(x)\delta(\rho^x_{\mathsf{E}}, \rho'^x_{\mathsf{E}})$. □

It is easy to see that the random variable X is *independent* of the state of the quantum system E in that $\rho^x_{\mathsf{E}} = \rho^{x'}_{\mathsf{E}}$ (and thus $= \rho_{\mathsf{E}}$) for all $x, x' \in \mathcal{X}$, if and only if

$$\rho_{X\mathsf{E}} = \rho_X \otimes \rho_{\mathsf{E}}\,.$$

This in particular implies that no information on X can be obtained from having access to the quantum system E. Similarly, X is *random and independent* of the state of E if and only if

$$\rho_{X\mathsf{E}} = \rho_{U\mathsf{E}} = \rho_U \otimes \rho_{\mathsf{E}}$$

where U is independently uniformly distributed over $\mathcal{X}$, and thus $\rho_U = \frac{1}{|\mathcal{X}|}\mathbb{I}_X$. This is the situation which we typically want to achieve in cryptography, where X is intended to be used as a cryptographic key and E collects the information an attacker may have.

We conclude by showing that a measurement can also be understood as a quantum operation.

[16] Note that this notation is consistent when considering the event $\mathcal{E} = [X = x]$.

[17] This may cause some ambiguity when E depends on several random variables; however, the meaning should always be clear from the context.

Proposition 4.9. *Let D be a quantum system and $\{|x\rangle\}_{x\in\mathcal{X}}$ an orthonormal basis of $\mathcal{H}_{\mathsf{D}}$ where $\mathcal{X} = \{0,\dots,d-1\}$. Consider the unitary matrix $U_{X\mathsf{D}} \in \mathcal{U}(\mathcal{H}_X \otimes \mathcal{H}_{\mathsf{D}})$, where $\mathcal{H}_X = \mathcal{H}_{\mathsf{D}}$, such that $U_{X\mathsf{D}}|x,y\rangle = |x+y\rangle|y\rangle$ (where the addition is modulo d). Then, for any density matrix $\sigma_{\mathsf{DE}} \in \mathcal{D}(\mathcal{H}_{\mathsf{D}} \otimes \mathcal{H}_{\mathsf{E}})$, the state*

$$\rho_{X\mathsf{E}} = \mathrm{tr}_{\mathsf{D}}\big((U_{X\mathsf{D}} \otimes \mathbb{I}_{\mathsf{E}})(|0\rangle\langle 0| \otimes \sigma_{\mathsf{DE}})(U_{X\mathsf{D}} \otimes \mathbb{I}_{\mathsf{E}})^\dagger\big)$$

has classical X, and P_X equals the distribution of the outcome of measuring subsystem D of σ_{DE} in basis $\{|x\rangle\}_{x\in\mathcal{X}}$ and $\rho_{\mathsf{E}|X=x}$ equals the state to which E collapses when x is observed.

Proof. To simplify notation, we skip some of the indices. First, we consider the case of an "empty" E, i.e. $\sigma \in \mathcal{D}(\mathcal{H}_{\mathsf{D}})$. Due to linearity, it suffices to show the claim for a *pure* state $\sigma = |\varphi\rangle\langle\varphi|$, where we can write $|\varphi\rangle = \sum_x \alpha_x|x\rangle$. Then,

$$\begin{aligned}\mathrm{tr}_{\mathsf{D}}\big(U(|0\rangle\langle 0| \otimes \sigma)U^\dagger\big) &= \sum_{x,y} \alpha_x\bar{\alpha}_y \mathrm{tr}_{\mathsf{D}}\big(U(|0\rangle\langle 0| \otimes |x\rangle\langle y|)U^\dagger\big) = \sum_{x,y} \alpha_x\bar{\alpha}_y \mathrm{tr}_{\mathsf{D}}\big(U|0,x\rangle\langle 0,x|U^\dagger\big)\\ &= \sum_{x,y} \alpha_x\bar{\alpha}_y \mathrm{tr}_{\mathsf{D}}\big(|x\rangle|x\rangle\langle y|\langle y|\big) = \sum_{x,y} \alpha_x\bar{\alpha}_y \langle y|x\rangle|x\rangle\langle y| = \sum_x |\alpha_x|^2|x\rangle\langle x|\,.\end{aligned}$$

This is a classical state (with respect to $\{|x\rangle\}_x$), and $P_X(x) = |\alpha_x|^2$ is the probability of observing x when measuring σ in basis $\{|x\rangle\}_x$.

In case of a composite state $\sigma \in P(\mathcal{H}_{\mathsf{D}} \otimes \mathcal{H}_{\mathsf{E}})$, it still suffices to show the claim for a *pure* state $\sigma = |\varphi\rangle\langle\varphi|$, where now $|\varphi\rangle = \sum_x \alpha_x|x\rangle|\psi_x\rangle$. Similar to above, except that $|\psi_x\rangle\langle\psi_y|$ is carried along, it follows that

$$\rho_{X\mathsf{E}} = \sum_x |\alpha_x|^2|x\rangle\langle x| \otimes |\psi_x\rangle\langle\psi_x|$$

where $p_x = \sum_{i,j} \alpha_{x,i}\bar{\alpha}_{x,j} = \big|\sum_i \alpha_{x,i}\big|^2$ equals the probability to observe x when measuring the first part of σ, and $\rho_{\mathsf{E}|X=x} = |\psi_x\rangle\langle\psi_x|$ equals the state to which the second part of σ collapses when x is observed. □

5 Privacy Amplification

In this section we show how a *weak* (or *raw*) key X, about which an adversary has some but limited quantum information, can be transformed into a *strong* key K which looks essentially random to the adversary. The transformation is fully public, no secrecy is involved there. Such a process of "amplifying privacy" is called **privacy amplification**. The assumption that the adversary has only *limited* information on the weak key X will be expressed here by essentially requiring that the size of the adversary's quantum state, measured in qubits, is smaller than the a priory collision entropy $\mathrm{H}_2(X)$ or the min-entropy $\mathrm{H}_\infty(X)$ of X; we refer to Appendix B.5 for the definitions of $\mathrm{H}_2(X)$ and $\mathrm{H}_\infty(X)$ (as well as of the conditional versions).

Definition 5.1. *A function $f : \mathcal{S} \times \mathcal{X} \to \mathcal{W}$ is called* **universal** *if*

$$P\big[f(S,x) = f(S,x')\big] \leq \frac{1}{|\mathcal{W}|}$$

for any $x \neq x' \in \mathcal{X}$ and for S uniformly distributed over $\mathcal{S}$.

Although a universal function does not have to be hashing, i.e. compressing, they are usually referred to as universal *hash* functions. Examples of universal (hash) functions are

$$f : \{0,1\}^{\ell\times n} \times \{0,1\}^n \to \{0,1\}^\ell,\ (A,x) \mapsto Ax$$