Security proof for Round Robin Differential Phase Shift QKD

Daan Leermakers and Boris Škorić
TU Eindhoven
d.leermakers.1@tue.nl, b.skoric@tue.nl

We give a security proof of the 'Round Robin Differential Phase Shift' Quantum Key Distribution scheme, and we give a tight bound on the required amount of privacy amplification. Our proof consists of the following steps. We construct an EPR variant of the scheme. We identify Eve's optimal way of coupling an ancilla to an EPR qudit pair under the constraint that the bit error rate between Alice and Bob should not exceed a value β. As a function of β we derive, for finite key size, the trace distance between the real state and a state in which no leakage exists. For asymptotic key size we obtain a bound on the trace distance by computing the von Neumann entropy. Our asymptotic result for the privacy amplification is sharper than existing bounds.

1 Introduction

1.1 Quantum Key Distribution and the RRDPS scheme

Quantum-physical information processing is different from classical information processing in several remarkable ways. Performing a measurement on an unknown quantum state typically destroys information; It is impossible to clone an unknown state by unitary evolution [1]; Quantum entanglement is a form of correlation between subsystems that does not exist in classical physics. Numerous ways have been devised to exploit these quantum properties for security purposes [2]. By far the most popular and well studied type of protocol is Quantum Key Distribution (QKD). QKD was first proposed in a famous paper by Bennett and Brassard in 1984 [3]. Given that Alice and Bob have a way to authenticate classical messages to each other (typically a short key), and that there is a quantum channel from Alice to Bob, QKD allows them to create a random key of arbitrary length about which Eve knows practically nothing. BB84 works with two conjugate bases in a two-dimensional Hilbert space. Many QKD variants have since been described in the literature [4–9], using e.g. different sets of qubit states, EPR pairs, qudits instead of qubits, or continuous variables. Furthermore, various proof techniques have been developed [10–13].
In 2014, Sasaki, Yamamoto and Koashi introduced *Round-Robin Differential Phase-Shift* (RRDPS) [14], a QKD scheme based on d-dimenional qudits. It has the advantage that it is very noise resilient while being easy to implement using photon pulse trains and interference measurements. One of the interesting aspects of RRDPS is that it is possible to omit the monitoring of signal disturbance. Even at high disturbance, Eve can obtain little information I_{AE} about Alice's secret bit. The value of I_{AE} determines how much privacy amplification is needed. As a result of this, the maximum possible QKD rate (the number of

actual key bits conveyed per quantum state) is $1 - h(\beta) - I_{\rm AE}$, where h is the binary entropy function and β the bit error rate.[1]

1.2 Prior work on the security of RRDPS

The security of RRDPS has been discussed in a number of papers [14, 16–18]. The original RRDPS paper gives an upper bound for asymptotic key length,

$$I_{\rm AE} \leq h(\frac{1}{d-1}) \tag{1}$$

(Eq. 5 in [14] with photon number set to 1). The security analysis in [14] is based on an entropic inequality for non-commuting measurements. There are two issues with this analysis. First, the proof is not written out in detail. Second, it is not known how tight the bound is.
Ref. [16] follows [14] and does a more accurate computation of phase error rate, tightening the $1/(d-1)$ in (1) to $1/d$. In [17] Sasaki and Koashi add β-dependence to their analysis and claim a bound

$$I_{\rm AE} \leq h(\frac{2\beta}{d-2}) \qquad \text{for } \beta \leq \frac{1}{2} \cdot \frac{d-2}{d-1} \tag{2}$$

and $I_{\rm AE} \leq h(\frac{1}{d-1})$ for $\beta \in [\frac{1}{2} \cdot \frac{d-2}{d-1}, \frac{1}{2}]$. The analysis in [18] considers only intercept-resend attacks, and hence puts a *lower bound* on Eve's potential knowledge, $I_{\rm AE} \geq 1 - h(\frac{1}{2} + \frac{1}{d}) = \mathcal{O}(1/d^2)$.[2]

1.3 Contributions and outline

In this paper we give a security proof of RRDPS in the case of coherent attacks. We give a bound on the required amount of privacy amplification. We adopt a proof technique inspired by [11], [13] and [10]. We consider the case where Alice and Bob *do monitor the channel* (i.e. they are able to tune the amount of privacy amplification (PA) as a function of the observed bit error rate) as well as the saturated regime where the leakage does not depend on the amount of noise.

- We construct an EPR variant of RRDPS; it is equivalent to RRDPS if Alice creates the EPR pair and immediately does her measurement.[3] We identify Eve's optimal way of coupling an ancilla to an EPR qudit pair under the constraint that the bit error rate between Alice and Bob does not exceed some value β.

[1] Monitoring of signal disturbance induces a small penalty on the QKD rate. However, the number of qubits that needs to be discarded is only logarithmic in the length of the derived key [15] and hence we will ignore the penalty.
[2] Ref. [18] gives a min-entropy of $-\log(\frac{1}{2} + \frac{1}{d})$, which translates to Shannon entropy $h(\frac{1}{2} + \frac{1}{d})$.
[3] This is similar to the Shor-Preskill technique [11].

– We consider an attack where Eve applies the above coupling to each EPR qudit-pair individually. We compute an upper bound on the statistical distance of the full QKD key (after PA) from uniformity, conditioned on Eve's ancilla states. From this we derive how much privacy amplification is needed. The result does not depend on the way in which Eve uses her ancillas, i.e. she may apply a postponed coherent measurement on the whole system of ancillas.
– We compute the von Neumann mutual information between one ancilla state and Alice's secret bit. This provides a bound on the leakage in the asymptotic (long key) regime [12]. Our result is sharper than [14].
– We provide a number of additional results by way of supplementary information. (i) We show that Eve's ancilla coupling can be written as a unitary operation on the Bob-Eve system. This means that the attack can be executed even if Eve has no access to Alice's qudit; this is important especially in the reduction from the EPR version to the original RRDPS. (ii) We compute the min-entropy of one secret bit given the corresponding ancilla. (iii) We compute the accessible information (mutual Shannon entropy) of one secret bit given the corresponding ancilla. The min-entropy and accessible-information results are relevant for collective attacks.

In Section 2 we introduce notation and briefly summarise the RRDPS scheme, the attacker model, and extraction of classical information from (mixed) quantum states. Section 3 states the main result: the amount of privacy amplification needed for RRDPS to be secure, (i) at finite key length and (ii) asymptotically. The remainder of the paper builds towards the proof of these results, and provides supplementary information about the leakage in terms of min-entropy loss and accessible (Shannon) information.
In Section 4 we introduce an EPR version of RRDPS. In Section 5 we impose the constraint that Eve's actions must not cause a bit error rate higher than β, and determine which mixed states of the Alice-Bob system are still allowed. There are only two scalar degrees of freedom left, which we denote as μ and V. In Section 6 we do the purification of the Alice-Bob mixed state, thus obtaining an expression for the state of Eve's ancilla. Although the ancilla space has dimension d^2, we show that only a four-dimensional subspace is relevant for the analysis. In Section 7 we prove the non-asymptotic main result by deriving an upper bound on the statistical distance between the distribution of the QKD key and the uniform distribution, conditioned on Eve's ancillas. In Section 8 we prove the asymptotic result by computing Eve's knowledge in terms of von Neumann entropy. In Section 9 we study collective attacks. Section 10 compares our results to previous bounds.

2 Preliminaries

2.1 Notation and terminology

Classical Random Variables (RVs) are denoted with capital letters, and their realisations with lowercase letters. The probability that a RV X takes value x

is written as $\Pr[X = x]$. The expectation with respect to RV X is denoted as $\mathbb{E}_x f(x) = \sum_{x\in\mathcal{X}} \Pr[X = x] f(x)$. The constrained sum $\sum_{t,t':t\neq t'}$ is abbreviated as $\sum_{[tt']}$ and $\mathbb{E}_{u,v:u\neq v}$ as $\mathbb{E}_{[uv]}$. The Shannon entropy of X is written as $\mathsf{H}(X)$. Sets are denoted in calligraphic font. The notation 'log' stands for the logarithm with base 2. The min-entropy of $X \in \mathcal{X}$ is $\mathsf{H}_{\min}(X) = -\log \max_{x\in\mathcal{X}} \Pr[X = x]$, and the conditional min-entropy is $\mathsf{H}_{\min}(X|Y) = -\log \mathbb{E}_y \max_{x\in\mathcal{X}} \Pr[X = x|Y = y]$. The notation h stands for the binary entropy function $h(p) = p \log \frac{1}{p} + (1 - p) \log \frac{1}{1-p}$. Bitwise XOR of binary strings is written as '$\oplus$'. The Kronecker delta is denoted as δ_{ab}. For quantum states we use Dirac notation. The notation 'tr' stands for trace. The Hermitian conjugate of an operator A is written as $A^\dagger$. When A is a complicated expression, we sometimes write $(A + \text{h.c.})$ instead of $A + A^\dagger$. The complex conjugate of z is denoted as z^*. We use the Positive Operator Valued Measure (POVM) formalism. A POVM $\mathcal{M}$ consists of positive semidefinite operators, $\mathcal{M} = (M_x)_{x\in\mathcal{X}}$, $M_x \geq 0$, and satisfies the condition $\sum_x M_x = \mathbb{1}$. The trace norm of A is $\|A\|_1 = \operatorname{tr}\sqrt{A^\dagger A}$. The trace distance between matrices ρ and σ is denoted as $\frac{1}{2}\|\rho - \sigma\|_1$.

2.2 (Min-)entropy of a classical variable given a quantum state

The notation $\mathcal{M}(\rho)$ stands for the classical RV resulting when $\mathcal{M}$ is applied to mixed state ρ. Consider a bipartite system 'AB' where the 'A' part is classical, i.e. the state is of the form $\rho^{\mathrm{AB}} = \mathbb{E}_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \rho_x$ with the $|x\rangle$ forming an orthonormal basis. The min-entropy of the classical RV X given part 'B' of the system is [19]

$$\mathsf{H}_{\min}(X|\rho_X) = -\log \max_{\mathcal{M}} \mathbb{E}_{x\in\mathcal{X}} \operatorname{tr}[M_x \rho_x]. \tag{3}$$

Here $\mathcal{M} = (M_x)_{x\in\mathcal{X}}$ denotes a POVM. Let $\Lambda \stackrel{\text{def}}{=} \sum_x \rho_x M_x$. If a POVM can be found that satisfies the condition[4] [20]

$$\forall_{x\in\mathcal{X}} : \; \Lambda - \rho_x \geq 0, \tag{4}$$

then there can be no better POVM for guessing X (but equally good POVMs may exist). For states that also depend on a classical RV $Y \in \mathcal{Y}$, the min-entropy of X given the quantum state and Y is

$$\mathsf{H}_{\min}(X|Y, \rho_X(Y)) = -\log \mathbb{E}_{y\in\mathcal{Y}} \max_{\mathcal{M}} \mathbb{E}_{x\in\mathcal{X}} \operatorname{tr}[M_x \rho_x(y)]. \tag{5}$$

A simpler expression is obtained when X is a binary variable. Let $X \in \{0, 1\}$. Then

$$\begin{aligned} &X \sim (p_0, p_1): \\ &\mathsf{H}_{\min}\left(X|Y, \rho_X(Y)\right) = -\log\left(\frac{1}{2} + \frac{1}{2}\mathbb{E}_y \operatorname{tr}\left\| p_0\rho_0(y) - p_1\rho_1(y) \right\|_1\right). \end{aligned} \tag{6}$$

[4] Ref. [20] specifies a second condition, namely $\Lambda^\dagger = \Lambda$. However, the hermiticity of Λ already follows from the condition (4).

This generalizes in a straightforward manner for states that depend on multiple classical RVs. The Shannon entropy of a classical variable given a measurement on a quantum state is given by

$$\mathsf{H}(X|\rho_X) = \min_{\mathcal{M}} \mathsf{H}(X|\mathcal{M}(\rho_X)). \tag{7}$$

The 'accessible information' is defined as the mutual information $\mathsf{H}(X) - \mathsf{H}(X|\rho_X)$. In contrast to the min-entropy case, there is no simple test analogous to (4) which tells you whether a local minimum in (7) is a global minimum.

2.3 The RRDPS scheme in a nutshell

The dimension of the qudit space is d. The basis states[5] are denoted as $|t\rangle$, with time indices $t \in \{0, \dots, d-1\}$. Whenever we use notation "$t_1 + t_2$" it should be understood that the addition of time indices is modulo d. The RRDPS scheme consists of the following steps.

1. Alice generates a random bitstring $a \in \{0,1\}^d$. She prepares the single-photon state

$$|\mu_a\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} (-1)^{a_t} |t\rangle \tag{8}$$

   and sends it to Bob.
2. Bob chooses a random integer $r \in \{1, \dots, d-1\}$. Bob performs a POVM measurement $\mathcal{M}^{(r)}$ described by a set of $2d$ operators $(M_{ks}^{(r)})_{k \in \{0,\dots,d-1\}, s \in \{0,1\}}$,

$$M_{ks}^{(r)} = \frac{1}{2} |\Psi_{ks}^{(r)}\rangle\langle\Psi_{ks}^{(r)}| \qquad |\Psi_{ks}^{(r)}\rangle = \frac{|k\rangle + (-1)^s |k+r\rangle}{\sqrt{2}}. \tag{9}$$

   The result of the measurement $\mathcal{M}^{(r)}$ on $|\mu_a\rangle$ is an random integer $k \in \{0, \dots, d-1\}$ and a bit $s = a_k \oplus a_{k+r}$.[6]
3. Bob announces k and r over a public but authenticated channel. Alice computes $s = a_k \oplus a_{k+r}$. Alice and Bob now have a shared secret bit s.

This procedure is repeated multiple times.
To detect eavesdropping, Alice and Bob can compare a randomly selected fraction of their secret bits. If this comparison is not performed, Alice and Bob have to assume that Eve learns as much as when causing bit error rate $\beta = \frac{1}{2}$. This mode of operation (without monitoring) was proposed in the original RRDPS paper [14].
Finally, on the remaining bits Alice and Bob carry out the standard procedures of information reconciliation and privacy amplification.

---

[5] The physical implementation [14] is a *pulse train*: a photon is split into d coherent pieces which are released at different, equally spaced, points in time.

[6] The phase $(-1)^{a_k \oplus a_{k+r}}$ is the phase of the field oscillation in the $(k+r)$'th pulse relative to the k'th. The measurement $\mathcal{M}^{(r)}$ is an interference measurement where one path is delayed by r time units.

The security of RRDPS is intuitively understood as follows. A measurement in a d-dimensional space cannot extract more than $\log d$ bits of information. The state $|\mu_a\rangle$, however, contains $d-1$ pieces of information, which is a lot more than $\log d$. Eve can learn only a fraction of the string a embedded in the qudit. Furthermore, what information she has is of limited use, because she cannot force Bob to select specific phases. (i) She cannot force Bob to choose a specific r value. (ii) Even if she feeds Bob a state of the form $|\Psi^{(r)}_{\ell u}\rangle$, where r accidentally equals Bob's r, then there is a $\frac{1}{2}$ probability that Bob's measurement $\mathcal{M}^{(r)}$ yields $k \neq \ell$ *with random* s.

2.4 Attacker model; channel monitoring

There is a quantum channel from Alice to Bob. There is an authenticated but non-confidential classical channel between Alice and Bob. We allow Eve to attack individual qudit positions in any way allowed by the laws of quantum physics, e.g. using unbounded quantum memory, entanglement, lossless operations, arbitrary POVMs, arbitrary unitary operators etc. All bit errors observed by Alice and Bob are assumed to be caused by Eve. Eve cannot influence the random choices of Alice and Bob, nor the state of their (measurement) devices. There are no side channels. This is the standard attacker model for quantum-cryptographic schemes. We consider the following channel monitoring technique. Alice and Bob test the bit error rate *for each combination* (a, k) *separately*, demanding that for each (a, k) the observed bit error rate does not exceed $\tilde{\beta} < \beta$.[7] Furthermore they test if k is uniform for every a. Since Eve has no control over r, passing these tests implies that for all (a, k, r) the bit error probability does not exceed β with overwhelming probability.[8]

The number of 'sacrificed' qudits required to implement all the tests on the bit error rate is of order $2^d \cdot d \cdot \log \kappa$, where κ is the length of the final key [15]. We will assume that n is chosen sufficiently large to ensure $d2^d \log \kappa \ll n$.

We will analyze an attack in which Eve couples an ancilla to each EPR pair individually *in the same way*, i.e. causing the same bit error probability (β). This looks like a serious restriction on Eve. However, it will turn out (Section 7) that the leakage is a concave function of β, which means that it is sub-optimal for Eve to use different ways of coupling for different EPR pairs.

We will see that the leakage becomes constant when β reaches a saturation point. If Alice and Bob are willing to tolerate such a noise level, then channel monitoring is no longer necessary for determining the leakage; they just assume that the maximum possible leakage occurs. (Monitoring is still necessary to determine which error-correcting code should be applied.)

[7] The gap between $\tilde{\beta}$ and β must be properly chosen as a function of the number of samples and the required confidence level.

[8] Any statistical uncertainty about the uniformity of k due to finite sample size can be represented as a statistical distance $\delta_{\rm sampl}$ between the real state and the state that satisfies the constraints. The parameter $\delta_{\rm sampl}$ will then appear, via the triangle inequality, as an additional term in the expression for the trace distance in Theorems 1 and 2.

Note that for large d it becomes impractical to determine the bit error rate for each combination (a, k) individually due to the exponential factor 2^d; the saturation value of the leakage should be assumed.

3 Main results

Our first result is a non-asymptotic bound on the trace distance between the real state and a state in which there exists no leakage.

Theorem 1. *Let $\mathbf{r} = (r_1, \ldots, r_n)$ be the values of the parameter r in n rounds of RRDPS, and similarly $\mathbf{k} = (k_1, \ldots, k_n)$. Let $\rho(\mathbf{r}, \mathbf{k})$ be the quantum-classical state describing Eve's ancillas as well as the ℓ-bit QKD key derived from the n-bit noisy secret $s'_1, \ldots, s'_n$. Let $\rho_{\mathrm{id}}(\mathbf{r}, \mathbf{k})$ be an 'ideal' state in which Eve's ancillas are completely decoupled from the QKD key. The statistical distance between these two states can be bounded as*

$$\frac{1}{2}\|\rho(\mathbf{r}, \mathbf{k}) - \rho_{\mathrm{id}}(\mathbf{r}, \mathbf{k})\|_1 < \frac{1}{2}\sqrt{2^{\ell - n(1-2\log T)}} \tag{10}$$

where T is given by

$$\beta \le \beta_* : \quad T = 2\beta + \sqrt{1-2\beta}\Big[\sqrt{1-2\beta\frac{d-1}{d-2}} + \frac{\sqrt{2\beta}}{\sqrt{d-2}}\Big] \tag{11}$$

$$\beta \ge \beta_* : \quad T = 2\beta_* + \sqrt{1-2\beta_*}\Big[\sqrt{1-2\beta_*\frac{d-1}{d-2}} + \frac{\sqrt{2\beta_*}}{\sqrt{d-2}}\Big] \tag{12}$$

and β_* is a saturation value that depends on d as

$$\beta_* = \frac{x_d/2}{1+x_d}, \tag{13}$$

where x_d is the solution on $(0, 1)$ of the equation

$$(1 - \frac{x}{d-2})^{\frac{1}{2}} + (1 + \frac{1}{d-2})(1 - \frac{x}{d-2})^{-\frac{1}{2}} + \frac{1}{\sqrt{d-2}}(\sqrt{x} - \frac{1}{\sqrt{x}}) - 2 = 0. \tag{14}$$

For asymptotically large n, it has been shown [21], using the properties of smooth Rényi entropies, that

$$\frac{1}{2}\|\rho - \rho_{\mathrm{id}}\|_1 \le \frac{1}{2}\sqrt{2^{\ell - n(1-I_{\mathrm{AE}})}}, \tag{15}$$

where I_{AE} is the single-qudit von Neumann information leakage $S(E) - S(E|S')$. Here 'E' stands for Eve's ancilla state and S' is Alice's secret bit.
Our second result is a computation of the von Neumann leakage I_{AE} for RRDPS.

Theorem 2. *The information leakage about the secret bit S' given R, K and Eve's quantum state, in terms of von Neumann entropy, is given by:*

$$\beta \leq \beta_0 : \quad I_{\rm AE} = (1-2\beta) h(\frac{1}{d-2} \cdot \frac{2\beta}{1-2\beta}) \tag{16}$$

$$\beta \geq \beta_0 : \quad I_{\rm AE} = (1-2\beta_0) h(\frac{1}{d-2} \cdot \frac{2\beta_0}{1-2\beta_0}). \tag{17}$$

Here β_0 is a saturation value (different from β_*) given by

$$\beta_0 = \frac{1}{2}\Big[1 + \frac{1}{(d-2)(1-y_d)}\Big]^{-1} \tag{18}$$

where y_d is the unique positive root of the polynomial $y^{d-1} + y - 1$.

The theorems are proven in Sections 7 and 8. The formulation of the security in terms of statistical distance ensures that the results are Universally Composable. In Section 8 we will see that Theorem 2 is sharper than (2) and hence allows for a higher QKD rate ℓ/n.

4 EPR version of the protocol

We follow the standard approach and re-formulate the protocol using EPR pairs. This will make it easier and more intuitive to describe the most general attack that Eve can perform.

E1 A maximally entangled two-qudit state is prepared.

$$|\alpha_0\rangle \stackrel{\rm def}{=} \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} |tt\rangle. \tag{19}$$

One qudit goes to Alice, and one to Bob.

E2 Eve does something with the EPR pair. Then Bob receives the 'B' qudit.

E3 Alice performs a POVM $\mathcal{Q} = (Q_z)_{z\in\{0,1\}^d}$ on her own qudit, where

$$Q_z = \frac{d}{2^d} |\mu_z\rangle\langle\mu_z|. \tag{20}$$

This results in a measured string $a \in \{0,1\}^d$.

E4 Bob picks a random integer $r \in \{1, \ldots, d-1\}$ and performs the POVM measurement $\mathcal{M}^{(r)}$ on his qudit. The result of the measurement is an integer $k \in \{0, \ldots, d-1\}$ and a bit s.

E5 Bob announces r and k. Alice computes $s' = a_k \oplus a_{k+r}$.

Note. If Alice is the party who prepares the EPR pair, and she immediately does the $\mathcal{Q}$ measurement, then the EPR version of the protocol reduces to the original RRDPS protocol, i.e. Alice measures a random string $a \in \{0,1\}^d$, while a single qudit in state $|\mu_a\rangle$ is sent to Bob. The only difference is that we allow Eve to couple her ancilla to the AB system instead of only the B system. (In Appendix A it will however turn out that Eve's attack is achieved by acting on Bob's qudit only; hence our results do not overestimate Eve's knowledge.)

Lemma 1. *The hermitian matrices Q_z as defined in (20) form a POVM, i.e.* $\sum_{z\in\{0,1\}^d} Q_z = \mathbb{1}$.

<u>*Proof:*</u>
$\sum_z |\mu_z\rangle\langle\mu_z| = \sum_z \frac{1}{d}\sum_{t,t'=0}^{d-1}(-1)^{z_{t'}+z_t}|t\rangle\langle t'| = \frac{1}{d}\sum_{t,t'=0}^{d-1}|t\rangle\langle t'|\sum_z(-1)^{z_{t'}+z_t}$.
Using $\sum_z(-1)^{z_{t'}+z_t} = 2^d\delta_{tt'}$ we get $\sum_z |\mu_z\rangle\langle\mu_z| = \frac{2^d}{d}\sum_t |t\rangle\langle t| = \frac{2^d}{d}\mathbb{1}$. $\square$
Alice and Bob's measurements can be carried out in the opposite order. It is not important whether $\mathcal{Q}$ is practical or not; it is a theoretical construct which allows us to build an EPR version of the protocol equivalent to the original protocol.

5 Imposing the noise constraint

Let ρ^{AB} denote the pure EPR state of Alice and Bob. The channel monitoring restricts the ways in which Eve can alter the state from ρ^{AB} to $\tilde{\rho}^{\text{AB}}$ without being detected. We will determine the most general allowed $\tilde{\rho}^{\text{AB}}$ that is compatible with bit error rate *exactly* β for all values of (a,k,r,s).
After Eve's actions in step E2, Alice and Bob have a bipartite mixed state $\tilde{\rho}^{\text{AB}}$ that can be represented in its most general form as

$$\tilde{\rho}^{\text{AB}} = \sum_{t,t',\tau,\tau'\in\{0,\ldots,d-1\}} \rho_{\tau\tau'}^{tt'}|t,t'\rangle\langle\tau,\tau'|, \tag{21}$$

with $\rho_{tt'}^{\tau\tau'} = (\rho_{\tau\tau'}^{tt'})^*$ and $\sum_{tt'}\rho_{tt'}^{tt'} = 1$. We introduce the notation $P_{aks|r} = \Pr[A=a, K=k, S=s|R=r]$.

Lemma 2. *Let Alice and Bob's bipartite state be given by (21) and let them perform the measurements $\mathcal{Q}$ and $\mathcal{M}^{(r)}$ respectively. At given r, the joint probability of the outcomes a,k,s is given by*

$$P_{aks|r} = \frac{1}{4\cdot 2^d}\sum_{t\tau}(-1)^{a_t+a_\tau}[\rho_{\tau k}^{tk} + \rho_{\tau,k+r}^{t,k+r} + (-1)^s(\rho_{\tau,k+r}^{tk} + \rho_{\tau k}^{t,k+r})]. \tag{22}$$

<u>*Proof:*</u> $P_{aks|r} = \text{tr}\,(Q_a\otimes M_{ks}^{(r)})\tilde{\rho}^{\text{AB}}$
$= \text{tr}\,(\frac{1}{2^d}\sum_{\ell\ell'}(-1)^{a_\ell+a_{\ell'}}|\ell\rangle\langle\ell'|\otimes\frac{1}{2}\frac{|k\rangle+(-1)^s|k+r\rangle}{\sqrt{2}}\frac{\langle k|+(-1)^s\langle k+r|}{\sqrt{2}})\sum_{tt'\tau\tau'}\rho_{\tau\tau'}^{tt'}|t\rangle\langle\tau|\otimes|t'\rangle\langle\tau'|$
$= \frac{1}{2^d 4}\sum_{tt'\tau\tau'}\rho_{\tau\tau'}^{tt'}(-1)^{a_t+a_\tau}[\delta_{t'k}+(-1)^s\delta_{t',k+r}][\delta_{\tau'k}+(-1)^s\delta_{\tau',k+r}]$
$= \frac{1}{2^d 4}\sum_{t\tau}(-1)^{a_t+a_\tau}[\rho_{\tau k}^{tk}+\rho_{\tau,k+r}^{t,k+r}+(-1)^s\rho_{\tau,k+r}^{tk}+(-1)^s\rho_{\tau k}^{t,k+r}]$. $\square$
We now impose the constraint that the event $s\neq s'$ occurs with probability β for all combinations (a,k,r).

Theorem 3. *The constraint*

$$\forall_{a,k,s,r}: P_{aks|r} = \frac{1}{2^d d}\Big[\delta_{s,a_k\oplus a_{k+r}}(1-\beta) + (1-\delta_{s,a_k\oplus a_{k+r}})\beta\Big]$$

can only be satisfied by a density function of the form

$$\tilde{\rho}^{\mathrm{AB}} = (1-2\beta-V)|\alpha_0\rangle\langle\alpha_0| + V\frac{1}{d}\sum_{tt'}|tt'\rangle\langle t't| + (2\beta-\mu)\frac{\mathbb{1}}{d^2} + \mu\frac{1}{d}\sum_t |tt\rangle\langle tt| \quad (23)$$

with $\mu, V \in \mathbb{R}$. *Written componentwise,*

$$\rho^{tt'}_{\tau\tau'} = \frac{1-2\beta-V}{d}\delta_{t't}\delta_{\tau'\tau} + \frac{V}{d}\delta_{\tau t'}\delta_{\tau' t} + \frac{2\beta-\mu}{d^2}\delta_{\tau t}\delta_{\tau' t'} + \frac{\mu}{d}\delta_{t't}\delta_{\tau t}\delta_{\tau' t}. \quad (24)$$

<u>Proof:</u> In expression (22) we distinguish two cases. (i) In the terms without s-dependence we have to make sure that the factor $(-1)^{a_t+a_\tau}$ vanishes. This requires setting $\tau = t$, i.e. $\rho^{tt'}_{\tau\tau'} = \alpha\delta_{\tau t}$, where α is allowed to depend on t; however, α cannot depend on t' and τ' (other than via $\delta_{t'\tau'}$) since then $P_{aks|r}$ would depend on k and $k+r$. (ii) In the terms containing $(-1)^s$ we have to make sure that $(-1)^{a_t+a_\tau} = (-1)^{a_k+a_{k+r}}$. This requires setting $\rho^{tt'}_{\tau\tau'} \propto \delta_{tt'}\delta_{\tau\tau'}$ or $\rho^{tt'}_{\tau\tau'} \propto \delta_{t\tau'}\delta_{\tau t'}$, where the proportionality constant can not depend on t or τ other than via $\delta_{t\tau}$. Combining these two cases we get the general expression

$$\begin{aligned}\rho^{tt'}_{\tau\tau'} &= f_t\delta_{t\tau}\delta_{t'\tau'}(1-\delta_{tt'}) + c\delta_{tt'}\delta_{\tau\tau'}(1-\delta_{t\tau}) \\ &\quad + e\delta_{t\tau'}\delta_{\tau t'}(1-\delta_{t\tau}) + g\delta_{tt'}\delta_{t\tau}\delta_{t\tau'} \end{aligned} \quad (25)$$

with $f_t, c, e, g \in \mathbb{R}$ and $gd + (d-1)\sum_t f_t = 1$. (The latter in order to ensure that the trace equals 1.) Substitution into (22) yields

$$P_{aks|r} = \frac{1}{4\cdot 2^d}[2\sum_t f_t - f_k - f_{k+r} + 2g + 2(c+e)(-1)^{s+a_k+a_{k+r}}]. \quad (26)$$

In order to remove the dependence on k and $k+r$ we have to set $f_t = f$, i.e. constant. Furthermore we have to set $c + e = (1-2\beta)/d$ in order to satisfy the noise constraint. Finally we reparametrise our constants as $\mu = 2\beta - d^2 f$, $V = de$. $\square$

Theorem 3 shows that (at fixed β) there are only two degrees of freedom, μ and V, in Eve's manipulation of the EPR pair.

6 Purification

According to the attacker model we have to assume that Eve has the purification of the state $\tilde{\rho}^{\mathrm{AB}}$. The purification contains all information that exists outside the AB system.

6.1 The purified state and its properties

We introduce the following notation,

$$|\alpha_j\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{d}}\sum_t e^{i\frac{2\pi}{d}jt}|tt\rangle, \qquad j \in \{0,\ldots,d-1\} \quad (27)$$

$$|D^{\pm}_{tt'}\rangle \stackrel{\text{def}}{=} \frac{|tt'\rangle \pm |t't\rangle}{\sqrt{2}} \qquad t < t'. \quad (28)$$

Lemma 3. *The* $\tilde{\rho}^{\mathrm{AB}}$ *given in (23) has the following orthonormal eigensystem,*

$$\begin{aligned} |\alpha_0\rangle \ \ \textit{with eigenvalue}\ \lambda_0 &\stackrel{\text{def}}{=} \frac{2\beta-\mu}{d^2} + \frac{\mu+V}{d} + 1 - 2\beta - V \\ |\alpha_j\rangle \quad j \in \{1,\ldots,d-1\} \ \ \textit{with eigenvalue}\ \lambda_1 &\stackrel{\text{def}}{=} \frac{2\beta-\mu}{d^2} + \frac{\mu+V}{d}. \\ |D^{\pm}_{tt'}\rangle \quad (t<t') \ \ \textit{with eigenvalue}\ \lambda_{\pm} &\stackrel{\text{def}}{=} \frac{2\beta-\mu}{d^2} \pm \frac{V}{d} \end{aligned} \tag{29}$$

Proof: The term proportional to $\mathbb{1}$ in (23) yields a contribution $(2\beta-\mu)/d^2$ to each eigenvalue. First we look at $|\alpha_j\rangle$. We have $\langle\alpha_0|\alpha_j\rangle = \delta_{j0}$. Furthermore $\langle t't|\alpha_j\rangle = \delta_{t't} e^{i\frac{2\pi}{d}jt}/\sqrt{d}$, which gives $(\sum_{tt'} |tt'\rangle\langle t't|)|\alpha_j\rangle = |\alpha_j\rangle$. Similarly we have $(\sum_t |tt\rangle\langle tt|)|\alpha_j\rangle = |\alpha_j\rangle$. Next we look at $|D^{\pm}_{tt'}\rangle$. We have $\langle\alpha_0|D^{\pm}_{tt'}\rangle = 0$ and $\langle uu|D^{\pm}_{tt'}\rangle = 0$. Hence the $(1-2\beta-V)$-term and the μ-term in (23) yield zero when acting on $|D^{\pm}_{tt'}\rangle$. Furthermore $\sum_{uu'} |uu'\rangle\langle u'u|D^{+}_{tt'}\rangle = \sum_{uu'} |uu'\rangle \frac{\delta_{ut}\delta_{u't'}+\delta_{ut'}\delta_{u't}}{\sqrt{2}}$ $= |D^{+}_{tt'}\rangle$. Similarly, $\sum_{uu'} |uu'\rangle\langle u'u|D^{-}_{tt'}\rangle = \sum_{uu'} |uu'\rangle \frac{\delta_{ut}\delta_{u't'}-\delta_{ut'}\delta_{u't}}{\sqrt{2}} \mathrm{sgn}(u-u')$ $= -|D^{-}_{tt'}\rangle$. □

In diagonalised form the $\tilde{\rho}^{\mathrm{AB}}$ is given by

$$\begin{aligned} \tilde{\rho}^{\mathrm{AB}} = \lambda_0|\alpha_0\rangle\langle\alpha_0| + \lambda_1 \sum_{j=1}^{d-1} |\alpha_j\rangle\langle\alpha_j| + \lambda_+ \sum_{tt':t<t'} |D^{+}_{tt'}\rangle\langle D^{+}_{tt'}| \\ +\lambda_- \sum_{tt':t<t'} |D^{-}_{tt'}\rangle\langle D^{-}_{tt'}|. \end{aligned} \tag{30}$$

The purification is

$$\begin{aligned} |\Psi^{\mathrm{ABE}}\rangle = \sqrt{\lambda_0}|\alpha_0\rangle \otimes |E_0\rangle + \sqrt{\lambda_1} \sum_{j=1}^{d-1} |\alpha_j\rangle \otimes |E_j\rangle \\ +\sqrt{\lambda_+} \sum_{tt':t<t'} |D^{+}_{tt'}\rangle \otimes |E^{+}_{tt'}\rangle + \sqrt{\lambda_-} \sum_{tt':t<t'} |D^{-}_{tt'}\rangle \otimes |E^{-}_{tt'}\rangle. \end{aligned} \tag{31}$$

where we have introduced orthonormal basis states $|E_j\rangle$, $|E^{\pm}_{tt'}\rangle$ in Eve's Hilbert space. In Appendix A we give more details on Eve's unitary operation.

6.2 Eve's state

Eve waits for Alice and Bob to perform their measurements and reveal k and r.

Lemma 4. *After Alice has measured* $a \in \{0,1\}^d$ *and Bob has measured* $k \in \{0,\ldots,d-1\}$, $s \in \{0,1\}$, *Eve's state is given by*

$$\sigma^{rk}_{as} = \mathrm{tr}_{\mathrm{AB}}\Big[|\Psi^{\mathrm{ABE}}\rangle\langle\Psi^{\mathrm{ABE}}| \frac{Q_a \otimes M^{(r)}_{ks} \otimes \mathbb{1}}{P_{aks|r}}\Big] \tag{32}$$

with $P_{aks|r}$ *as defined in Theorem 3.*

Proof: The POVM elements Q_a and $M_{ks}^{(r)}$ are proportional to projection operators. Hence the tripartite ABE pure state after the measurement is proportional to $(Q_a \otimes M_{ks}^{(r)} \otimes \mathbb{1})|\Psi^{\rm ABE}\rangle$. It is easily verified that the normalisation in (32) is correct: taking the trace in E-space yields $\mathrm{tr}_{\rm AB}\mathrm{tr}_{\rm E}|\Psi^{\rm ABE}\rangle\langle\Psi^{\rm ABE}|Q_a \otimes M_{ks}^{(r)} \otimes \mathbb{1}$ $= \mathrm{tr}_{\rm AB}\, \tilde\rho^{\rm AB} Q_a \otimes M_{ks}^{(r)} = P_{aks|r}$. □

Lemma 5. *It holds that*

$$\frac{d}{2^d} \sum_{\substack{a_0\cdots a_{d-1} \\ \text{without } a_k, a_{k+r}}} |\mu_a\rangle\langle\mu_a| = \frac{1}{4}\mathbb{1} + \frac{1}{4}(-1)^{a_k+a_{k+r}}\Big(|k\rangle\langle k+r| + |k+r\rangle\langle k|\Big) \tag{33}$$

$$= M^{(r)}_{k,a_k\oplus a_{k+r}} + \frac{1}{4}\sum_{t:\ t\neq k,k+r} |t\rangle\langle t|. \tag{34}$$

Proof: We have $|\mu_a\rangle\langle\mu_a| = \frac{1}{d}\mathbb{1} + \frac{1}{d}\sum_{[t\tau]}|t\rangle\langle\tau|(-1)^{a_t+a_\tau}$. Summation of the $\frac{1}{d}\mathbb{1}$ term is trivial and yields $2^{d-2}\cdot\frac{1}{d}\mathbb{1}$. In the summation of the factor $(-1)^{a_t+a_\tau}$ in the second term, any summation $\sum_{a_t}(-1)^{a_t}$ yields zero. The only nonzero contribution arises when $t = k, \tau = k+r$ or $t = k+r, \tau = k$; the a-summation then yields a factor 2^{d-2}. □

Lemma 6. *It holds that*

$$\mathbb{E}_{a:a_k\oplus a_{k+r}=s'}|\mu_a\rangle\langle\mu_a| = \frac{\mathbb{1}}{d} + (-1)^{s'}\frac{|k\rangle\langle k+r| + |k+r\rangle\langle k|}{d}. \tag{35}$$

Proof: We have $\mathbb{E}_{a:a_k\oplus a_{k+r}=s'}|\mu_a\rangle\langle\mu_a| = 2^{-(d-1)}\sum_{a_k}\sum_{a_{k+r}}\delta_{a_k\oplus a_{k+r},s'}\cdot$ $\sum_{a \text{ without } a_k,a_{k+r}}|\mu_a\rangle\langle\mu_a|$. For the rightmost summation we use Lemma 5. Performing the $\sum_{a_k}$ and $\sum_{a_{k+r}}$ summations yields (35). □

Eve's task is to guess Alice's bit $s' = a_k \oplus a_{k+r}$ from the mixed state σ^{rk}_{as}, where Eve does not know a and s. We define

$$\sigma^{rk}_{s'} = \mathbb{E}_{s,a:a_k\oplus a_{k+r}=s'}[\sigma^{rk}_{as}]. \tag{36}$$

This represents Eve's ancilla state given some value of Alice's bit s'. Next we introduce notations that are useful for understanding the structure of $\sigma^{rk}_{s'}$. We define, for $t, t' \in \{0,\ldots,d-1\}$, non-normalised vectors $|w_{tt'}\rangle$ in Eve's Hilbert space as

$$|w_{tt'}\rangle \stackrel{\rm def}{=} \langle tt'|\Psi^{\rm ABE}\rangle. \tag{37}$$

Furthermore we define angles α and φ as

$$\cos 2\alpha \stackrel{\rm def}{=} \frac{\langle w_{kk}|w_{k+r,k+r}\rangle}{\langle w_{kk}|w_{kk}\rangle}, \quad \cos 2\varphi \stackrel{\rm def}{=} \frac{\langle w_{k,k+r}|w_{k+r,k}\rangle}{\langle w_{k,k+r}|w_{k,k+r}\rangle} \tag{38}$$

and vectors $|A\rangle, |B\rangle, |C\rangle, |D\rangle$

$$\frac{|w_{kk}\rangle}{\sqrt{\langle w_{kk}|w_{kk}\rangle}} = \cos\alpha|A\rangle + \sin\alpha|B\rangle \tag{39}$$

$$\frac{|w_{k+r,k+r}\rangle}{\sqrt{\langle w_{k+r,k+r}|w_{k+r,k+r}\rangle}} = \cos\alpha|A\rangle - \sin\alpha|B\rangle \tag{40}$$

$$\frac{|w_{k,k+r}\rangle}{\sqrt{\langle w_{k,k+r}|w_{k,k+r}\rangle}} = \cos\varphi|C\rangle + \sin\varphi|D\rangle \tag{41}$$

$$\frac{|w_{k+r,k}\rangle}{\sqrt{\langle w_{k+r,k}|w_{k+r,k}\rangle}} = \cos\varphi|C\rangle - \sin\varphi|D\rangle. \tag{42}$$

The $|A\rangle$, $|B\rangle$, $|C\rangle$, $|D\rangle$ are mutually orthogonal, and also orthogonal to any vector $|w_{tt'}\rangle$ $(t' \neq t)$ with $\{t,t'\} \neq \{k, k+r\}$.

Theorem 4. *The eigenvalues of* $\sigma^{rk}_{s'}$ *are given by*

$$\xi_0 \stackrel{\text{def}}{=} \frac{d}{2} \cdot \frac{\lambda_+ + \lambda_-}{2} \tag{43}$$

$$\xi_1 \stackrel{\text{def}}{=} \tfrac{d}{2}(\lambda_1 + \lambda_-) = \beta - \tfrac{d}{2}(\tfrac{d}{2} - 1)(\lambda_+ + \lambda_-) \tag{44}$$

$$\xi_2 \stackrel{\text{def}}{=} \tfrac{d}{2}(\lambda_1 + 2\frac{\lambda_0 - \lambda_1}{d} + \lambda_+) = 1 - \beta - \tfrac{d}{2}(\tfrac{d}{2} - 1)(\lambda_+ + \lambda_-) \tag{45}$$

and the diagonal representation of $\sigma^{rk}_{s'}$ *is*

$$\begin{aligned}\sigma^{rk}_{s'} = \xi_0 &\sum_{\substack{t\in\{0,\ldots,d-1\} \\ t\neq k, t\neq k+r}} \Big(\frac{|w_{tk}\rangle\langle w_{tk}|}{\langle w_{tk}|w_{tk}\rangle} + \frac{|w_{t,k+r}\rangle\langle w_{t,k+r}|}{\langle w_{t,k+r}|w_{t,k+r}\rangle}\Big) \\ &+ \xi_2 \frac{[\sqrt{\xi_2 - \frac{d}{2}\lambda_+}|A\rangle + (-1)^{s'}\sqrt{\frac{d}{2}\lambda_+}|C\rangle][\cdots]^\dagger}{\xi_2} \\ &+ \xi_1 \frac{[\sqrt{\xi_1 - \frac{d}{2}\lambda_-}|B\rangle - (-1)^{s'}\sqrt{\frac{d}{2}\lambda_-}|D\rangle][\cdots]^\dagger}{\xi_1}\end{aligned} \tag{46}$$

<u>*Proof:*</u> We have

$$\begin{aligned}\sigma^{rk}_{s'} &= \mathrm{tr}_{\mathrm{AB}}|\Psi^{\mathrm{ABE}}\rangle\langle\Psi^{\mathrm{ABE}}|\mathbb{E}_{a:a_k\oplus a_{k+r}=s'} Q_a \otimes \mathbb{E}_{s|s'} \frac{M^{(r)}_{ks}}{P_{aks|r}} \otimes \mathbb{1} \\ &= d2^d\ \mathrm{tr}_{\mathrm{AB}}|\Psi^{\mathrm{ABE}}\rangle\langle\Psi^{\mathrm{ABE}}|[\mathbb{E}_{a:a_k\oplus a_{k+r}=s'} Q_a] \otimes [\sum_s M^{(r)}_{ks}] \otimes \mathbb{1}.\end{aligned} \tag{47}$$

We use Lemma 6 to evaluate the $\mathbb{E}_a$ factor. We use $\sum_s M^{(r)}_{ks} = \frac{1}{2}|k\rangle\langle k| + \frac{1}{2}|k+r\rangle\langle k+r|$. This allows us to write everything in terms of $|w_{tt'}\rangle$ states. For $t = t'$ we have

$$|w_{tt}\rangle = \sqrt{\lambda_0/d}|E_0\rangle + \sqrt{\lambda_1/d}\sum_{j=1}^{d-1}(e^{i\frac{2\pi}{d}})^{jt}|E_j\rangle \tag{48}$$

$$\langle w_{tt}|w_{tt}\rangle = \lambda_1 + \frac{\lambda_0 - \lambda_1}{d}, \tag{49}$$

and for $t \neq t'$ we have

$$|w_{tt'}\rangle = \sqrt{\lambda_+/2}|E^+_{(tt')}\rangle + \mathrm{sgn}(t'-t)\sqrt{\lambda_-/2}|E^-_{(tt')}\rangle \tag{50}$$

$$\langle w_{tt'}|w_{tt'}\rangle = (\lambda_+ + \lambda_-)/2. \tag{51}$$

The following properties hold ($t \neq t'$)

$$\langle w_{tt}|w_{tt'}\rangle = 0 \ , \ \langle w_{tt}|w_{t't}\rangle = 0 \tag{52}$$

$$\langle w_{tt}|w_{t't'}\rangle = \frac{\lambda_0 - \lambda_1}{d} \ , \ \langle w_{tt'}|w_{t't}\rangle = \frac{\lambda_+ - \lambda_-}{2}. \tag{53}$$

We get

$$\cos 2\alpha = 1 - \frac{d\lambda_1}{\lambda_0 + (d-1)\lambda_1}, \quad \cos 2\varphi = 1 - \frac{2\lambda_-}{\lambda_+ + \lambda_-} \tag{54}$$

After some tedious algebra the result (46) follows. □

Note that the σ_0^{rk} and σ_1^{rk} have the same set of eigenvalues: $2(d-2)$ times ξ_0, and once ξ_1 and ξ_2.

Corollary 1. *It holds that*

$$\frac{\sigma_0^{rk} + \sigma_1^{rk}}{2} = \sum_{\substack{t\in\{0,\ldots,d-1\} \\ t\neq k, t\neq k+r}} \xi_0 \cdot \Big(\frac{|w_{tk}\rangle\langle w_{tk}|}{\langle w_{tk}|w_{tk}\rangle} + \frac{|w_{t,k+r}\rangle\langle w_{t,k+r}|}{\langle w_{t,k+r}|w_{t,k+r}\rangle}\Big)$$
$$+(\xi_2 - \tfrac{d}{2}\lambda_+)|A\rangle\langle A| + \tfrac{d}{2}\lambda_+|C\rangle\langle C| + (\xi_1 - \tfrac{d}{2}\lambda_-)|B\rangle\langle B| + \tfrac{d}{2}\lambda_-|D\rangle\langle D|.$$

Proof: Follows directly from Theorem 4 by discarding the terms in (46) that contain $(-1)^{s'}$ (the AC and BD crossterms). □

Corollary 2. *The difference between σ_0^{rk} and σ_1^{rk} can be written as*

$$\frac{\sigma_0^{rk} - \sigma_1^{rk}}{2} = \frac{1}{2}\sqrt{d\lambda_+}\sqrt{d\lambda_- + 2(1-\beta) - \frac{d^2}{2}(\lambda_+ + \lambda_-)}\Big(|A\rangle\langle C| + |C\rangle\langle A|\Big)$$
$$-\frac{1}{2}\sqrt{d\lambda_-}\sqrt{d\lambda_+ + 2\beta - \frac{d^2}{2}(\lambda_+ + \lambda_-)}\Big(|B\rangle\langle D| + |D\rangle\langle B|\Big). \tag{55}$$

Proof: Using Theorem 4, we see everything except the AC and BD crossterms cancel from (46). □

7 Statistical distance

Now that we have described Eve's most general allowed state, and how it is connected to Alice's secret bit s', it is time to prove Theorem 1.

Let r_i be the 'r'-value in round i and similarly k_i, s'_i. We use the notation $\mathbf{r} = (r_1, \ldots, r_n)$, $\mathbf{k} = (k_1, \ldots, k_n)$. Let $x = (s'_1, \ldots, s'_n)$. Let $z \in \{0,1\}^\ell$ be the QKD key obtained by applying privacy amplification to x, i.e. $z = \mathtt{Ext}(x, u)$, where $\mathtt{Ext}$ is a universal hash function (UHF) and $u \in \mathcal{U}$ is public randomness. Ideally

Eve's n-ancilla state state would be decoupled from z; this can be represented as a combined quantum-classical state of the form

$$\rho_{\mathrm{ideal}}(\mathbf{r},\mathbf{k}) \stackrel{\mathrm{def}}{=} \frac{1}{2^\ell}\sum_{z\in\{0,1\}^\ell}|z\rangle\langle z|\otimes\frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}}|u\rangle\langle u|\otimes\omega^{\mathrm{E}}(\mathbf{r},\mathbf{k}) \tag{56}$$

for some state ω^{E}. In reality, however, Eve's n-ancilla state has some weak dependence on z. The combined state in reality is given by

$$\rho(\mathbf{r},\mathbf{k}) = \frac{1}{2^\ell}\sum_{z\in\{0,1\}^\ell}|z\rangle\langle z|\otimes\frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}}|u\rangle\langle u|\otimes\frac{2^\ell}{2^n}\sum_{x\in\{0,1\}^n}\delta_{z,\mathtt{Ext}(x,u)}\bigotimes_{i=1}^{n}\sigma_{x_i}^{r_ik_i}. \tag{57}$$

We will prove an upper bound on the distance between ρ and ρ_{ideal} for a specific choice of ω^{E}, namely an average matrix σ_{av} defined as

$$\sigma_{\mathrm{av}}(\mathbf{r},\mathbf{k}) \stackrel{\mathrm{def}}{=} \bigotimes_{i=1}^{n}\frac{\sigma_0^{r_ik_i}+\sigma_1^{r_ik_i}}{2}. \tag{58}$$

For this specific choice we write 'ρ_{id}' instead of the general ρ_{ideal}. Furthermore we define difference matrices $\mathit{\Delta}$ as

$$\begin{aligned}\rho(\mathbf{r},\mathbf{k})-\rho_{\mathrm{id}}(\mathbf{r},\mathbf{k}) &= \mathbb{E}_{zu}|z\rangle\langle z|\otimes|u\rangle\langle u|\otimes\mathit{\Delta}(z,u,\mathbf{r},\mathbf{k})\\ \mathit{\Delta}(z,u,\mathbf{r},\mathbf{k}) &\stackrel{\mathrm{def}}{=} \frac{2^\ell}{2^n}\sum_{x\in\{0,1\}^n}\delta_{z,\mathtt{Ext}(x,u)}\bigotimes_{i=1}^{n}\sigma_{x_i}^{r_ik_i}-\sigma_{\mathrm{av}}(\mathbf{r},\mathbf{k}).\end{aligned} \tag{59}$$

Lemma 7. *It holds that*

$$\|\rho(\mathbf{r},\mathbf{k})-\rho_{\mathrm{id}}(\mathbf{r},\mathbf{k})\|_1 = \mathbb{E}_{zu}\|\mathit{\Delta}(z,u,\mathbf{r},\mathbf{k})\|_1. \tag{60}$$

Proof: This follows from the block structure of $\rho-\rho_{\mathrm{id}}$. The list of eigenvalues of $\rho-\rho_{\mathrm{id}}$ is obtained by combining the individual eigenvalue lists of the $\mathit{\Delta}(z,u,\mathbf{r},\mathbf{k})$ for all combinations (z,u). □

Lemma 8. *It holds that*

$$\mathbb{E}_{zu}\|\mathit{\Delta}(z,u,\mathbf{r},\mathbf{k})\|_1 \leq \mathrm{tr}\sqrt{\mathbb{E}_{zu}\mathit{\Delta}^2(z,u,\mathbf{r},\mathbf{k})}. \tag{61}$$

Proof: $\mathbb{E}_{zu}\|\mathit{\Delta}(z,u,\mathbf{r},\mathbf{k})\|_1 = \mathbb{E}_{zu}\mathrm{tr}\sqrt{\mathit{\Delta}^2(z,u,\mathbf{r},\mathbf{k})} = \mathrm{tr}\,\mathbb{E}_{zu}\sqrt{\mathit{\Delta}^2(z,u,\mathbf{r},\mathbf{k})}$. We apply Jensen's inequality. □

Lemma 9. *It holds that*

$$\mathbb{E}_{zu}\mathit{\Delta}^2(z,u,\mathbf{r},\mathbf{k}) = \frac{2^\ell-1}{2^n}\bigotimes_{i=1}^{n}\frac{(\sigma_0^{r_ik_i})^2+(\sigma_1^{r_ik_i})^2}{2}. \tag{62}$$

Proof: From the definition (59) of Δ we get

$$\mathbb{E}_{zu}\Delta^2(z,u,\mathbf{r},\mathbf{k}) = \frac{2^{2\ell}}{2^{2n}}\sum_{xy}\mathbb{E}_{zu}\delta_{z,\mathtt{Ext}(x,u)}\delta_{z,\mathtt{Ext}(y,u)}\bigotimes_{i=1}^{n}\sigma_{x_i}^{r_ik_i}\sigma_{y_i}^{r_ik_i}+\sigma_{\rm av}^2$$
$$-\sigma_{\rm av}\frac{2^\ell}{2^n}\sum_x\mathbb{E}_{zu}\delta_{z,\mathtt{Ext}(x,u)}\bigotimes_{i=1}^{n}\sigma_{x_i}^{r_ik_i}$$
$$-\Big(\frac{2^\ell}{2^n}\sum_x\mathbb{E}_{zu}\delta_{z,\mathtt{Ext}(x,u)}\bigotimes_{i=1}^{n}\sigma_{x_i}^{r_ik_i}\Big)\sigma_{\rm av}. \tag{63}$$

We split the $\sum_{xy}$ sum into a sum with $y=x$ and a sum with $y\neq x$. Then we use $\sum_z\delta_{z,\mathtt{Ext}(x,u)}=1$ and $\sum_z\mathbb{E}_u\delta_{z,\mathtt{Ext}(x,u)}\delta_{z,\mathtt{Ext}(y,u)}=2^{-\ell}$ for $y\neq x$. The latter is the defining property of UHFs. Then we rewrite $\sum_{xy:\,y\neq x}$ as $\sum_{xy}-\sum_{xy}\delta_{xy}$. Finally, after applying $2^{-n}\sum_x\bigotimes_i\sigma_{x_i}^{r_ik_i}=\sigma_{\rm av}$, most of the terms cancel and (62) is what remains. □

Lemma 10. *It holds that*

$$\frac{(\sigma_0^{rk})^2+(\sigma_1^{rk})^2}{2} = \sum_{\substack{t\in\{0,\ldots,d-1\}\\ t\neq k,t\neq\ell}}\xi_0^2\Big(\frac{|w_{tk}\rangle\langle w_{tk}|}{\langle w_{tk}|w_{tk}\rangle}+\frac{|w_{t\ell}\rangle\langle w_{t\ell}|}{\langle w_{t\ell}|w_{t\ell}\rangle}\Big)+\xi_1(\xi_1-\tfrac{d}{2}\lambda_-)|B\rangle\langle B|$$
$$+\xi_1\tfrac{d}{2}\lambda_-|D\rangle\langle D|+\xi_2(\xi_2-\tfrac{d}{2}\lambda_+)|A\rangle\langle A|+\xi_2\tfrac{d}{2}\lambda_+|C\rangle\langle C|.$$

Proof: Follows directly from Theorem 4. □

Lemma 11. *The statistical distance between the real and ideal state can be bounded as*

$$\|\rho(\mathbf{r},\mathbf{k})-\rho_{\rm id}(\mathbf{r},\mathbf{k})\|_1<\sqrt{2^{\ell-n}}T^n \tag{64}$$
$$T\stackrel{\rm def}{=}2(d-2)\xi_0+\sqrt{\xi_2(\xi_2-\tfrac{d}{2}\lambda_+)}+\sqrt{\xi_2\tfrac{d}{2}\lambda_+}+\sqrt{\xi_1(\xi_1-\tfrac{d}{2}\lambda_-)}+\sqrt{\xi_1\tfrac{d}{2}\lambda_-}. \tag{65}$$

Proof: Substitution of Lemma 9 into Lemma 8 into Lemma 7 gives $\|\rho(\mathbf{r},\mathbf{k})-\rho_{\rm id}(\mathbf{r},\mathbf{k})\|_1\leq\sqrt{\frac{2^\ell-1}{2^n}}\prod_{i=1}^n\mathrm{tr}\sqrt{\frac{(\sigma_0^{r_ik_i})^2+(\sigma_1^{r_ik_i})^2}{2}}$. The trace does not depend on the actual value of r_i and k_i. We define $T=\mathrm{tr}\sqrt{(\sigma_0^{rk})^2+(\sigma_1^{rk})^2}/\sqrt{2}$ for arbitrary r,k. From Lemma 10 we obtain (65). Finally we use $2^\ell-1<2^\ell$. □

Corollary 3. *Let ε be a small constant. The distance $\|\rho(\mathbf{r},\mathbf{k})-\rho_{\rm id}(\mathbf{r},\mathbf{k})\|_1$ can be made equal to ε by setting $\ell/n=1-2\log T-\frac{2}{n}\log\frac{1}{\varepsilon}$.*

Remark. Corollary 3 provides a tighter bound on the QKD rate than similar statements based on Rényi-2 entropy. We are able to compute the square root in $\mathrm{tr}\sqrt{\sigma_0^2+\sigma_1^2}$, whereas in Rényi-2 entropy Jensen's inequality is used to bound the trace as $\sqrt{\text{dimension}}\sqrt{\mathrm{tr}\,\sigma_0^2+\mathrm{tr}\,\sigma_1^2}$.

Since Eve is still free to choose the parameters μ and V (or, equivalently, λ_+ and λ_-) she can choose them such that $\|\rho(\mathbf{r},\mathbf{k})-\rho_{\rm id}(\mathbf{r},\mathbf{k})\|_1$ is maximized.

Theorem 5. *Eve's choice that maximises $\|\rho(\mathbf{r},\mathbf{k})-\rho_{\rm id}(\mathbf{r},\mathbf{k})\|_1$ is given by*

$$\beta\leq\beta_*: \quad T = 2\beta+\sqrt{1-2\beta}\Big[\sqrt{1-2\beta\frac{d-1}{d-2}}+\frac{\sqrt{2\beta}}{\sqrt{d-2}}\Big] \tag{66}$$

$$\textit{at } \lambda_-=0, \quad \lambda_+=\frac{4\beta}{d(d-2)} \tag{67}$$

$$\beta\geq\beta_*: \quad T = 2\beta_*+\sqrt{1-2\beta_*}\Big[\sqrt{1-2\beta_*\frac{d-1}{d-2}}+\frac{\sqrt{2\beta_*}}{\sqrt{d-2}}\Big] \tag{68}$$

$$\textit{at } \lambda_-=\frac{4\beta_*(\beta-\beta_*)}{d(d-2)(1-2\beta_*)}, \quad \lambda_+=\frac{4\beta_*(1-\beta-\beta_*)}{d(d-2)(1-2\beta_*)}. \tag{69}$$

Here β_* is a saturation value that depends on d as follows,

$$\beta_*=\frac{x_d/2}{1+x_d}, \tag{70}$$

where x_d is the solution on $(0,1)$ of the equation

$$(1-\frac{x}{d-2})^{\frac12}+\frac{d-1}{d-2}(1-\frac{x}{d-2})^{-\frac12}+\frac{1}{\sqrt{d-2}}(\sqrt{x}-\frac{1}{\sqrt{x}})-2=0. \tag{71}$$

<u>Proof:</u> We start from (65). At $\beta=\frac12$ the expression for T is symmetric in λ_+ and λ_-. Hence the overall maximum achievable at any β lies at $\lambda_+=\lambda_-=\frac{q}{d(d-2)}$ for some as yet unknown q. We have

$$T_{\rm max}^{\beta=\frac12}=\zeta(q,d)\stackrel{\rm def}{=}q+\sqrt{1-q}\Big(\sqrt{1-\frac{d-1}{d-2}q}+\frac{\sqrt{q}}{\sqrt{d-2}}\Big). \tag{72}$$

On the other hand, we note that substitution of (67) into (65) yields (66), which is precisely of the form $\zeta(q,d)$ if we identify $2\beta\equiv q$. Hence, at some $\beta<\frac12$ it is already possible to achieve $T=T_{\rm max}^{\beta=1/2}$, i.e. we have saturation. We note that substitution of (69) into (65) yields (68). The saturation value β_* is found by solving $\partial\zeta(2\beta,d)/\partial\beta=0$; after some simplification, this equation can be rewritten as (71) by setting $x=2\beta/(1-2\beta)$.[9] $\square$

This concludes the proof of Theorem 1.

The upper bound on the amount of information Eve has about S' $(2\log T)$ is plotted in Fig. 1. The optimal λ_+,λ_- are shown in Fig. 3 (Section 9).

Lemma 12. *The large-d asymptotics of the saturation value β_* is given by*

$$\beta_*=\frac14-\frac{1}{8\sqrt{d-2}}-\mathcal{O}(\frac{1}{(d-2)^{3/2}}), \tag{73}$$

[9] After some rewriting it can be seen that (71) is equivalent to a complicated 6th order polynomial equation. We have not yet been able to prove that the solution on $(0,1)$ is unique. Our numerical solutions however indicate that this is the case.

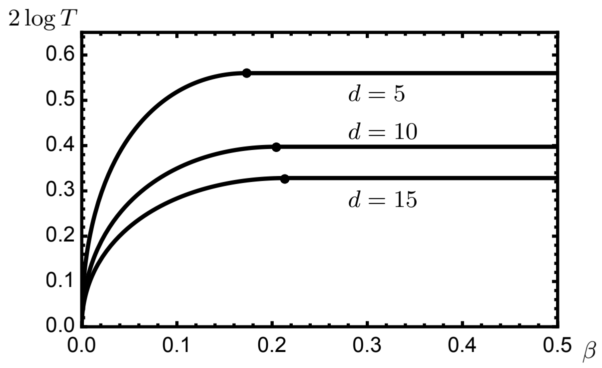

Fig. 1. *Upper bound on the information leakage as a function of the bit error rate for $d = 5$, $d = 10$ and $d = 15$ (Theorem 1). A dot indicates the saturation point β_*.*

which yields

$$T = 1 + \frac{1}{2\sqrt{d-2}} - \mathcal{O}(\frac{1}{d-2}) \tag{74}$$

$$\|\rho(\mathbf{r},\mathbf{k}) - \rho_{\rm id}(\mathbf{r},\mathbf{k})\|_1 \leq 2^{-\frac{1}{2}n[1-\frac{1}{\sqrt{d-2}\ln 2}+\mathcal{O}(\frac{1}{d-2})-\frac{\ell}{n}]}. \tag{75}$$

Proof: We set $x_d = 1 - 1/\sqrt{d-2} + a/(d-2)$, where a is supposedly of order 1, and substitute this into (71). This yields $a = \frac{1}{2} + \mathcal{O}(1/\sqrt{d-2})$, which is indeed of order 1. Substitution of x_d into (70) gives (73), and substitution of β_* into (68) gives (74). Finally, substitution of (74) into Lemma 11 yields (75). □

Note. If Eve applies different β to different qudits then the T^n in Lemma 11 becomes a product of different T values, each associated with a different β. The total leakage is then a weighted average of $2\log T$ values; as can be seen from Fig. 1 the leakage is a concave function of β, i.e. it is better for Eve to stick to a single β.

8 Von Neumann entropy

Here we prove Theorem 2. Using smooth Rényi entropies it was shown in [12] that, in the large n limit, the von Neumann leakage per qubit is the relevant quantity for determining the required amount of PA.[10] We denote the leakage from Alice to Eve, in terms of von Neumann entropy, $I_{\rm AE}$. It is given by

$$\begin{aligned} I_{\rm AE} &= S(\sigma_{S'}^{RK}|RK) - S(\sigma_{S'}^{RK}|RKS') \\ &= \mathbb{E}_{rk}[S(\sigma_{S'}^{rk}) - S(\sigma_{S'}^{rk}|S')] \end{aligned}$$

[10] By applying Jensen's inequality once more to lemma 8, we can move the trace into the square root and get an expression which is equivalent to lemma 4.4 in [21]. After this point the proof structure from [21] can be followed. Thus the Von Neumann leakage is also an asymptotic case of our statistical distance result Theorem 1.

$$= \mathbb{E}_{rk} \left[S(\frac{\sigma_0^{rk} + \sigma_1^{rk}}{2}) - \frac{S(\sigma_0^{rk}) + S(\sigma_1^{rk})}{2} \right]$$

$$= S\Big(\frac{\sigma_0^{rk} + \sigma_1^{rk}}{2}\Big) - \frac{S(\sigma_0^{rk}) + S(\sigma_1^{rk})}{2} \quad r, k \text{ arbitrary.} \tag{76}$$

In the last line we used that the eigenvalues of $\sigma_{s'}^{rk}$ and $\sigma_0^{rk} + \sigma_1^{rk}$ do not actually depend on r and k. Again λ_+ and λ_- can be optimized to Eve's advantage.

Theorem 6. *Eve's choice that maximizes the von Neumann leakage is given by*

$$\beta \leq \beta_0 : \quad I_{\rm AE} = (1-2\beta) h(\frac{1}{d-2} \cdot \frac{2\beta}{1-2\beta}) \tag{77}$$

$$\text{at } \lambda_- = 0, \quad \lambda_+ = \frac{4\beta}{d(d-2)} \tag{78}$$

$$\beta \geq \beta_0 : \quad I_{\rm AE} = (1-2\beta_0) h(\frac{1}{d-2} \cdot \frac{2\beta_0}{1-2\beta_0}) \tag{79}$$

$$\text{at } \lambda_- = \frac{4\beta_0(\beta-\beta_0)}{d(d-2)(1-2\beta_0)}, \quad \lambda_+ = \frac{4\beta_0(1-\beta-\beta_0)}{d(d-2)(1-2\beta_0)}. \tag{80}$$

Here β_0 is a saturation value that depends on d as follows,

$$\beta_0 = \frac{1}{2} \Big[1 + \frac{1}{(d-2)(1-y_d)}\Big]^{-1} \tag{81}$$

where y_d is the unique positive root of the polynomial $y^{d-1} + y - 1$.

<u>*Proof:*</u> We start from (76). We note that the eigenvalue set of $(\sigma_0^{rk} + \sigma_1^{rk})/2$ largely coincides with that of σ_0^{rk} and σ_1^{rk} (Theorem 4 and Corollary 1). What remains of (76) comes entirely from the $|A\rangle, |B\rangle, |C\rangle, |D\rangle$ subspace,

$$\begin{aligned} I_{\rm AE} &= \xi_1 \log \xi_1 + \xi_2 \log \xi_2 - (\xi_2 - \tfrac{d}{2}\lambda_+) \log(\xi_2 - \tfrac{d}{2}\lambda_+) - \tfrac{d}{2}\lambda_+ \log(\tfrac{d}{2}\lambda_+) \\ &\quad -(\xi_1 - \tfrac{d}{2}\lambda_-) \log(\xi_1 - \tfrac{d}{2}\lambda_-) - \tfrac{d}{2}\lambda_- \log(\tfrac{d}{2}\lambda_-) \\ &= \xi_1 h(\frac{d}{2} \cdot \frac{\lambda_-}{\xi_1}) + \xi_2 h(\frac{d}{2} \cdot \frac{\lambda_+}{\xi_2}). \end{aligned} \tag{82}$$

We note that (82) is invariant under the transformation $(\beta \to 1-\beta; \lambda_+ \leftrightarrow \lambda_-)$. At $\beta = 1/2$ we must hence have $\lambda_+ = \lambda_- = \lambda$.

$$I_{\rm AE}^{\beta=\frac{1}{2}} = g(d, \lambda) \stackrel{\rm def}{=} [1 - d(d-2)\lambda] \cdot h\Big(\frac{d\lambda}{1-d(d-2)\lambda}\Big). \tag{83}$$

At $\beta = \frac{1}{2}$, the largest leakage that Eve can cause is $\max_\lambda g(d, \lambda) = g(d, \lambda_*)$.[11] Next we note that substitution of (80) into (82) yields (79); this has the same

[11] $\frac{\partial^2 g(d,\lambda)}{\partial \lambda^2} = -\frac{d}{\lambda}[1-d(d-2)\lambda]^{-1}[1-d(d-1)\lambda]^{-1}$, hence g is a concave function of λ on the interval $\lambda \in [0, \frac{1}{d(d-1)}]$, which interval coincides with the region allowed by the constraints on μ, V. The function g has a single maximum at some point λ_*.

form as $g(d,\lambda)$ (83) if we make the identification $\lambda d(d-2) = 2\beta_0$. Moreover, by setting $\beta_0 = \frac{1}{2}\lambda_* d(d-2)$, Eve achieves the overall maximum leakage $g(d,\lambda_*)$ already at a value of β smaller than $\frac{1}{2}$. Since the maximum leakage cannot decrease with β, this implies that the maximum leakage saturates at $\beta = \beta_0$ and stays constant at $I_{\rm AE}^{\max}(\beta) = g(d,\lambda_*)$ on the interval $\beta \in [\beta_0, \frac{1}{2}]$. The value $g(d,\lambda_*)$ precisely equals (79). Next we determine the value of β_0. Demanding $\partial g(d,\lambda)/\partial\lambda = 0$ at $\lambda = \lambda_*$ yields

$$\log \frac{[1-d(d-1)\lambda_*]^{d-1}}{[1-d(d-2)\lambda_*]^{d-2}\lambda_* d} = 0. \tag{84}$$

This is equivalent to the polynomial equation $y^{d-1} + y - 1 = 0$ with $y \in [0,1]$ if we make the identification $y = 1 - \frac{\lambda_* d}{1-\lambda_* d(d-2)} = \frac{1-\lambda_* d(d-1)}{1-\lambda_* d(d-2)}$. (It is readily seen that $\lambda_* \in [0, \frac{1}{d(d-1)}]$ implies $y \in [0,1]$.) This precisely matches (81), because of the optimal choice $\beta_0 = \frac{1}{2}\lambda_* d(d-2)$. By Descartes' rule of signs, the function $y^{d-1} + y - 1$ has exactly one positive root.
When β is decreased below β_0, the location (λ_-, λ_+) of the maximum of the stationary point of $I_{\rm AE}$ leaves the 'allowed' triangular region; this happens at a corner of the triangle, $\lambda_- = 0$, $\lambda_+ = \frac{4\beta}{d(d-2)}$. For $\beta < \beta_0$ this corner yields the highest achievable leakage. Substitution of (78) into (82) yields (77). □
This concludes the proof of theorem 2.

Remark. From $y > 0$ and (81) it follows that $\beta_0 < \frac{1}{2} \cdot \frac{d-2}{d-1}$.
Fig. 2 shows the von Neumann mutual information for three values of d. The optimal λ_+,λ_- are plotted in Fig. 3 (Section 9).

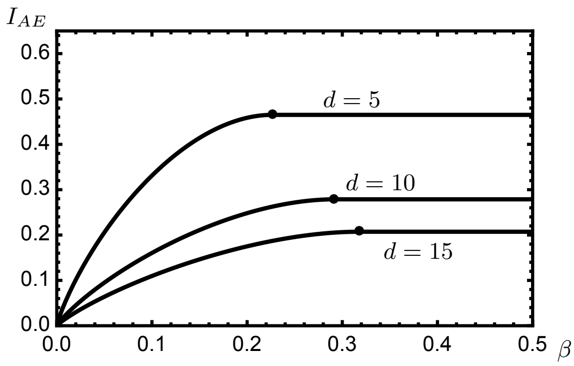

Fig. 2. *Mutual information between Alice and Eve in terms of von Neumann entropy as a function of the bit error rate, for $d = 5$, $d = 10$ and $d = 15$ (Theorem 2). A dot indicates the saturation point β_0.*

Lemma 13. *The large-d asymptotics of the $I_{\rm AE}$ is given by*

$$\beta \le \beta_0 : I_{\rm AE} = \frac{2\beta}{d-2} \log \frac{(d-2)(1-2\beta)e}{2\beta} + \mathcal{O}(d^{-2}) \tag{85}$$

$$\beta \geq \beta_0 : I_{\rm AE} = \frac{\log d}{d} + \mathcal{O}(\frac{\log\log d}{d}). \tag{86}$$

Proof: The result for $\beta < \beta_0$ follows by doing a series expansion of (77) in the small parameter $1/(d-2)$. For $\beta > \beta_0$ we study the equation $y^{d-1} = 1-y$. Let us try a solution of the form $y = 1 - \frac{\ln[(d-1)/\alpha]}{d-1}$ for some unknown α. This yields $\alpha \cdot \{(1-\frac{\ln[(d-1)/\alpha]}{d-1})^{d-1}\frac{d-1}{\alpha}\} = \ln\frac{d-1}{\alpha}$. Using the fact that $\lim_{n\to\infty}(1-x/n)^n = e^{-x}$ we see that the expression $\{\cdots\}$ is close to 1 if it holds that $\ln\frac{d-1}{\alpha} \ll d-1$, and that the equation is then satisfied by $\alpha = \mathcal{O}(\ln d)$, which is indeed consistent with $\ln\frac{d-1}{\alpha} \ll d-1$. Substituting $\alpha = \mathcal{O}(\ln d)$ into the expression for y and then into (81) gives $1-2\beta_0 = \frac{1}{\ln d} + \mathcal{O}(\frac{\ln\ln d}{[\ln d]^2})$. Substituting this result for $1-2\beta_0$ into (79) finally yields (86). $\square$

9 Collective attacks

By way of supplementary information we present a number of results about collective attacks. These are attacks on individual qudits, i.e. Eve performs the same measurement on every individual ancilla that she holds. First, this teaches us which kind of measurement is informative for Eve. Second, it quantifies the gap between what is provable for general attacks and what is provable for more restricted attacks. We compute leakage in terms of min-entropy loss and in terms of accessible (Shannon) information. Since min-entropy is a very conservative measure we will see that the min-entropy loss exceeds the leakage found in Theorems 1 and 2. The main interest is in Eve's measurement itself. The accessible information is the relevant quantity when Eve's quantum memory is short-lived, forcing her to perform a measurement on her ancillas before she has observed Alice and Bob's usage of the QKD key. As expected, the accessible information will turn out to be smaller than the leakage of Theorems 1 and 2.

9.1 Min-entropy

Eve's ability to distinguish between the cases $s' = 0$ and $s' = 1$ depends on the distance between σ_0^{rk} and σ_1^{rk} (see Section 2.2). Eq. (6) with $p_0 = \frac{1}{2}$, $p_1 = \frac{1}{2}$ tells us that the relevant quantity is $\|\sigma_0^{rk} - \sigma_1^{rk}\|_1$. For notational convenience we define the value $\beta_{\rm sat}$,

$$\beta_{\rm sat} \stackrel{\rm def}{=} \frac{1}{4}\cdot\frac{d-2}{d-1}. \tag{87}$$

Again we optimize λ_+ and λ_-.

Lemma 14. *For all r and k, the choice for λ_+ and λ_- that maximizes the trace distance $\frac{1}{2}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1$ is*

$$\lambda_+ = \tfrac{4\beta}{d(d-2)} \qquad \lambda_- = 0 \qquad \textit{for } \beta < \beta_{\rm sat} \tag{88}$$

$$\lambda_+ = \tfrac{4\beta_{\rm sat}}{d(d-2)} - \tfrac{2(\beta-\beta_{\rm sat})}{d^2} \qquad \lambda_- = \tfrac{2(\beta-\beta_{\rm sat})}{d^2} \qquad \textit{for } \beta \geq \beta_{\rm sat}. \tag{89}$$

which gives

$$\frac{1}{2}\left\|\sigma_0^{rk}-\sigma_1^{rk}\right\|_1 = \begin{cases} \frac{1}{\sqrt{d-1}}\frac{\sqrt{\beta}}{\beta_{\rm sat}}\sqrt{2\beta_{\rm sat}-\beta} \; \textit{for} & \beta<\beta_{\rm sat} \\ \frac{1}{\sqrt{d-1}} \; \textit{for} & \beta\geq\beta_{\rm sat}. \end{cases} \tag{90}$$

<u>Proof:</u> From corollary 2 it is easy to see that

$$\begin{aligned} \frac{1}{2}\left\|\sigma_0^{rk}-\sigma_1^{rk}\right\|_1 &= \sqrt{d\lambda_-}\sqrt{d\lambda_+ + 2\beta - \frac{d^2}{2}(\lambda_+ + \lambda_-)} \\ &\quad + \sqrt{d\lambda_+}\sqrt{d\lambda_- + 2(1-\beta) - \frac{d^2}{2}(\lambda_+ + \lambda_-)}. \end{aligned} \tag{91}$$

In Appendix B we derive the λ_+, λ_- that maximize (91) while keeping all eigenvalues non-negative. □

Remark. The optimal choice for λ_+,λ_- has the same form for all three optimizations that we have performed. The only difference is the saturation value. Although (89) is shown in a simplified form one can manipulate it to the same form as (69) and (80) with $\beta_{\rm sat}$ instead of β_* or β_0.

Fig. 3 shows the optimal λ_+ and λ_- together with the constraints on the λ parameters for all three optimizations. The lower dots in the figure correspond to $\beta=\frac{1}{2}$. For all three information measures the optimum moves towards the top corner of the triangle for decreasing β. For β values below the saturation point the optimum is the top corner, with $\lambda_-=0$ and $\lambda_1=0$.

Knowing the optimal values for λ_+ and λ_-, we compute the min-entropy leakage.

Theorem 7. *The min-entropy of the bit S' given R, K and the state $\sigma_{S'}^{RK}$ is*

$$\beta<\beta_{\rm sat}: \quad \mathsf{H}_{\rm min}(S'|RK\sigma_{S'}^{RK}) = -\log\left(\frac{1}{2}+\frac{1}{2\sqrt{d-1}}\frac{\sqrt{\beta}}{\beta_{\rm sat}}\sqrt{2\beta_{\rm sat}-\beta}\right) \tag{92}$$

$$\beta\geq\beta_{\rm sat}: \quad \mathsf{H}_{\rm min}(S'|RK\sigma_{S'}^{RK}) = -\log\left(\frac{1}{2}+\frac{1}{2\sqrt{d-1}}\right). \tag{93}$$

<u>Proof:</u> Eq. (6) with X uniform, $X\to S'$, $Y\to(R,K)$ becomes

$$\begin{aligned} \mathsf{H}_{\rm min}(S'|RK\sigma_{s'}^{RK}) &= -\log\left(\frac{1}{2}+\frac{1}{2}\mathbb{E}_{rk}\left\|\frac{1}{2}\sigma_0^{rk}-\frac{1}{2}\sigma_1^{rk}\right\|_1\right) \\ &= -\log\left(\frac{1}{2}+\frac{1}{4}\left\|\sigma_0^{rk}-\sigma_1^{rk}\right\|_1\right) \qquad (r,k \text{ arbitrary}). \end{aligned} \tag{94}$$

In the last step we omitted the expectation over r and k since the trace distance does not depend on r, k. Substitution of (90) into (94) gives the end result. □

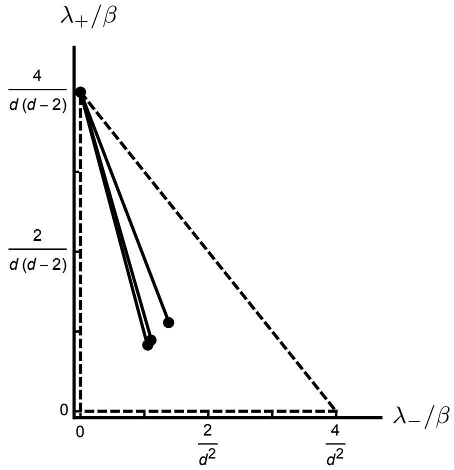

Fig. 3. *Optimal choice of λ_+ and λ_- at $d = 10$ for statistical distance (left line), min-entropy (middle line) and von Neumann entropy (right line). The dashed triangle represents the region for which the eigenvalues λ_+, λ_- and λ_1 are non-negative. The black dots indicate the optimum at $\beta = \frac{1}{2}$ (dots inside the triangle) and $\beta \leq \beta_*, \beta_{\rm sat}, \beta_0$ (upper corner of the triangle). Not shown in this plot is the $\lambda_0 \geq 0$ constraint which cuts off the upper left corner of the triangle for $\beta > 2\beta_{\rm sat}$.*

Corollary 4. *Eve's optimal POVM $\mathcal{T}^{rk} = (T_0^{rk}, T_1^{rk})$ for maximising the min-entropy leakage is given by*

$$T_0^{rk} = \frac{1}{2}\Big(\mathbb{1} + |A\rangle\langle C| + |C\rangle\langle A| - |B\rangle\langle D| - |D\rangle\langle B|\Big) \quad ; \quad T_1^{rk} = \mathbb{1} - T_0^{rk}. \tag{95}$$

Proof: The trace distance in Lemma 14 is the sum of the positive eigenvalues of $\sigma_0^{rk} - \sigma_1^{rk}$. In the space spanned by $|A\rangle, |B\rangle, |C\rangle, |D\rangle$, the optimal T_0 consists of the projection onto the space spanned by the eigenvectors corresponding to the positive eigenvalues. These eigenvectors are $|v_1\rangle = \frac{|A\rangle + |C\rangle}{\sqrt{2}}$ and $|v_2\rangle = \frac{|D\rangle - |B\rangle}{\sqrt{2}}$. The matrix that projects onto them is $|v_1\rangle\langle v_1| + |v_2\rangle\langle v_2| = \frac{1}{2}|A\rangle\langle A| + \frac{1}{2}|B\rangle\langle B| + \frac{1}{2}|C\rangle\langle C| + \frac{1}{2}|D\rangle\langle D| + |A\rangle\langle C| + |C\rangle\langle A| - |B\rangle\langle D| - |D\rangle\langle B|$. In order to satisfy the constraint $T_0 + T_1 = \mathbb{1}$ and symmetry, half the identity matrix in the remaining $d^2 - 4$ dimensions has to be added to T_0. We mention, without showing it, that (95) satisfies the test (4). □

As expected, the min-entropy loss decreases as the dimension of the Hilbert space grows. We see that the entropy loss saturates at $\beta = \beta_{\rm sat}$; hence RRDPS is secure up to arbitrarily high noise levels. Fig. 4 shows the min-entropy leakage as a function of β.

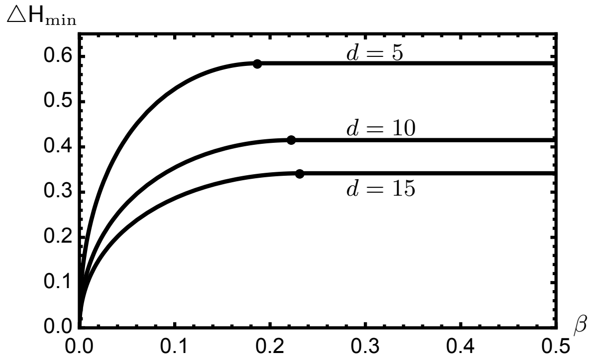

Fig. 4. *Min-entropy leakage as a function of the bit error rate for* $d = 5$*,* $d = 10$ *and* $d = 15$*. A dot indicates the saturation point* β_{sat}*.*

9.2 Accessible Shannon information

Lemma 15. *Let* $X \in \mathcal{X}$ *be a uniformly distributed random variable. Let* $Y \in \mathcal{Y}$ *be a random variable. Let* ρ_{xy} *be a quantum state coupled to the classical* x, y*. The Shannon entropy of* X *given a state* ρ_{XY} *that has to be measured (for unknown* X *and* Y*) is given by*

$$\mathsf{H}(X|\rho_{XY}) = \min_{POVM\ \mathcal{M}=(M_m)_{m\in\mathcal{X}}} \mathbb{E}_{x\in\mathcal{X}} \mathsf{H}\Big(\{\operatorname{tr} M_m \mathbb{E}_{y|x}\rho_{xy}\}_{m\in\mathcal{X}}\Big). \tag{96}$$

Proof: We have $\mathsf{H}(X|\rho_{XY}) = \min_{\mathcal{M}} \mathsf{H}(X|Z)$, where Z is the outcome of the POVM measurement $\mathcal{M}$. Z is a classical random variable that depends on X and Y. We can write $\mathsf{H}(X|Z) = \mathsf{H}(X) - \mathsf{H}(Z) + \mathsf{H}(Z|X)$. Since X is uniform, and Z is an estimator for X, the Z is uniform as well. Thus we have $\mathsf{H}(X) - \mathsf{H}(Z) = 0$, which yields $\mathsf{H}(X|\rho_{XY}) = \min_{\mathcal{M}} \mathsf{H}(Z|X) = \min_{\mathcal{M}} \mathbb{E}_x \mathsf{H}(Z|X = x)$. The probability $\Pr[z|x]$ is given by $\Pr[z|x] = \mathbb{E}_{y|x}\Pr[z|xy] = \mathbb{E}_{y|x}\operatorname{tr} M_z\rho_{xy}$. □

Corollary 5. *It holds that*

$$\mathsf{H}(S'|RK\sigma_{AS}^{RK}) = \mathbb{E}_{rk} \min_{\mathcal{G}^{rk}=(G_0^{rk},G_1^{rk})} \mathbb{E}_{s'} h(\operatorname{tr} G_m^{rk}\sigma_{s'}^{rk}), \quad m \in \{0,1\} \text{ arbitrary.} \tag{97}$$

Proof: Application of Lemma 15 yields

$$\begin{aligned}\mathsf{H}(S'|RK\sigma_{AS}^{RK}) &= \mathbb{E}_{rk} \min_{\mathcal{G}^{rk}=(G_0^{rk},G_1^{rk})} \mathbb{E}_{s'} H(\{\operatorname{tr} G_m^{rk}\mathbb{E}_{as|s'}\sigma_{as}^{rk}\}_{m\in\{0,1\}}) \\ &= \mathbb{E}_{rk} \min_{\mathcal{G}^{rk}=(G_0^{rk},G_1^{rk})} \mathbb{E}_{s'} H(\{\operatorname{tr} G_m^{rk}\sigma_{s'}^{rk}\}_{m\in\{0,1\}}) \end{aligned} \tag{98}$$

where in the last step we used the definition of $\sigma_{s'}^{rk}$. Finally, the Shannon entropy of a binary variable is given by the binary entropy function h, where $h(1-p) = h(p)$. □

From Corollary 5 we see that the POVM $\mathcal{T}^{rk}$ associated with the min-entropy also optimizes the Shannon entropy: maximizing the guessing probability $\operatorname{tr} G_{s'}^{rk}\sigma_{s'}^{rk}$ minimizes the Shannon entropy.

Theorem 8. *The Shannon entropy of Alice's bit S' given the state σ_{AS}^{RK}, R and K is:*

$$\beta < \beta_{\rm sat}: \qquad \mathsf{H}(S'|RK\sigma_{AS}^{RK}) = h\left(\frac{1}{2} + \frac{1}{2\sqrt{d-1}}\frac{\sqrt{\beta}}{\beta_{\rm sat}}\sqrt{2\beta_{\rm sat}-\beta}\right). \quad (99)$$

$$\beta \geq \beta_{\rm sat}: \qquad \mathsf{H}(S'|RK\sigma_{AS}^{RK}) = h\left(\frac{1}{2} + \frac{1}{2\sqrt{d-1}}\right). \quad (100)$$

Proof: The min-entropy result (92,93) can be written as $\mathsf{H}_{\rm min}(S'|RK\sigma_{S'}^{RK}) = -\log \operatorname{tr} T_{s'}^{rk}\sigma_{s'}^{rk}$, so we already have an expression for $\operatorname{tr} T_{s'}^{rk}\sigma_{s'}^{rk}$. Substitution of $\mathcal{T}^{rk}$ for $\mathcal{G}^{rk}$ in (97) yields the result. □

Since the optimal POVM for min- and Shannon entropy are the same, saturation occurs at the same point ($\beta = \beta_{\rm sat}$). Fig 5 shows the Shannon entropy leakage (mutual information) $I_{\rm AE} = 1 - \mathsf{H}(S'|RK\sigma_{AS}^{RK})$ as a function of β.

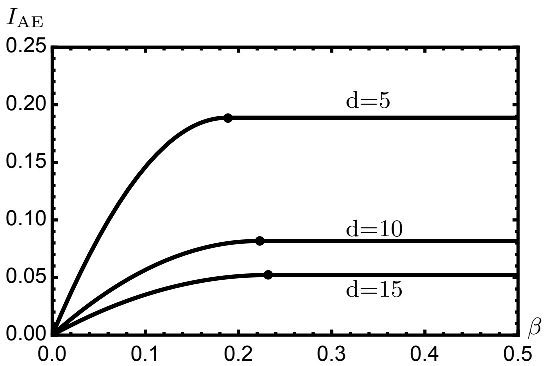

Fig. 5. *Accessible Shannon entropy as a function of β for $d = 5$, $d = 10$ and $d = 15$. A dot indicates the saturation point $\beta_{\rm sat}$.*

10 Discussion

10.1 Comparison with previous analyses

Our Theorem 1 is non-asymptotic; we cannot compare it to previous results since the previous results are for the asymptotic regime. Figs. 6 and 7 show our results versus previous bounds on the leakage. It is clear that our on Neumann result is sharper than [17] for all β and d. Interestingly, our non-asymptotic result for the saturated leakage is sharper than the asymptotic [17] for $d \leq 22$. Note too that saturation occurs at lower β (especially for small d) than reported in [17].

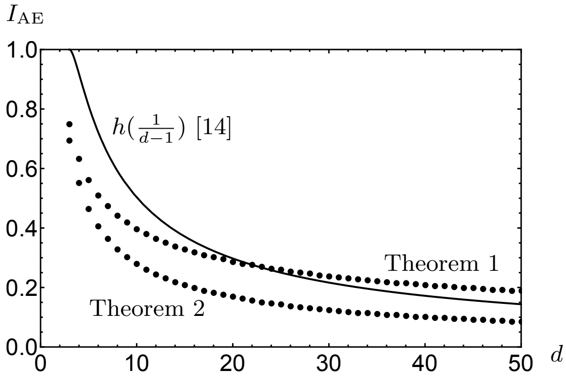

Fig. 6. *Saturated leakage as a function of d. Comparison of [14] and our results (Theorem 1 and Theorem 2).*

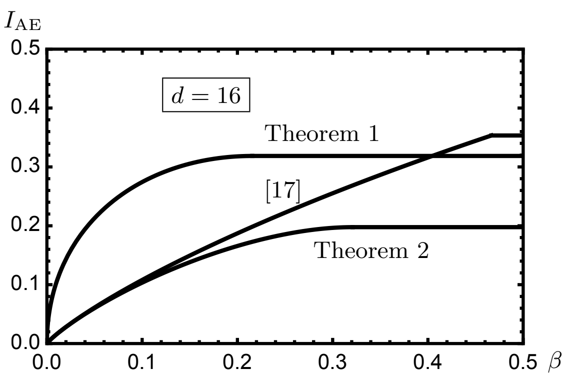

Fig. 7. *Leakage as a function of β, for $d = 16$. Comparison of our Theorem 1 and Theorem 2 versus [17], $h(\frac{2\beta}{d-2})$ below and $h(\frac{1}{d-1})$ above saturation.*

10.2 Remarks on the optimal attack

The $\tilde{\rho}^{\rm AB}$ mixed state allowed by the noise constraint has two degrees of freedom, μ and V. While this is more than the zero degrees of freedom in the case of qubit-based QKD [12], it is still a small number, given the dimension d^2 of the Hilbert space.
Eve's attack has an interesting structure. Eve entangles her ancilla with Bob's qudit. Bob's measurement affects Eve's state. When Bob reveals r, k, Eve knows which 4-dimensional subspace is relevant. However, the basis state $|k\rangle$ in Bob's qudit is coupled to $|A_k^a\rangle$ in Eve's space (see appendix A), which is spanned by $d-1$ different basis vectors $|E^+_{(kt')}\rangle$ (Eq. 104 with $\lambda_1 = 0$, $\lambda_- = 0$), each carrying different phase information $a_k \oplus a_{t'}$. Only one out of $d-1$ carries the information she needs, and she cannot select which one to read out. Her problem is aggravated by the fact that the $|A_t^a\rangle$ vectors are not orthogonal (except at $\beta = \frac{1}{2}$). Note that this entanglement-based attack is far more powerful than the intercept-resend attack studied in [18].

A Details of Eve's unitary operation

In Theorem 9 below we show that Eve does not have to touch Alice's qudit. Hence the attacks that we are describing here can also be carried out in the original (non-EPR) protocol, where Eve gets access only to the qudit state sent to Bob.

Theorem 9. *The operation that maps the pure EPR state to $|\Psi^{\rm ABE}\rangle$ (31) can be represented as a unitary operation on Bob's subsystem and Eve's ancilla.*

Proof: Let Eve's ancilla have initial state $|E_0\rangle$. The transition from the pure EPR state to (31) can be written as the following mapping,

$$U\Big(|t\rangle_{\rm B} \otimes |E_0\rangle_{\rm E}\Big) = |\Omega_t\rangle, \tag{101}$$

where $|\Omega_t\rangle$ is a state in the BE system defined as

$$\begin{aligned} |\Omega_t\rangle \stackrel{\rm def}{=} \sqrt{\lambda_0}|t\rangle|E_0\rangle + \sqrt{\lambda_1}|t\rangle \sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}jt}|E_j\rangle + \sqrt{\frac{d\lambda_+}{2}} \sum_{t':t'\neq t} |t'\rangle|E^+_{(tt')}\rangle \\ +\sqrt{\frac{d\lambda_-}{2}} \sum_{t':t'\neq t} |t'\rangle|E^-_{(tt')}\rangle \mathrm{sgn}(t'-t). \end{aligned} \tag{102}$$

The notation (tt') indicates ordering of t and t' such that the smallest index occurs first. It holds that $\langle\Omega_t|\Omega_\tau\rangle = \delta_{t\tau}$. Eqs. (101,102) show that the attack can be represented as an operation that does not touch Alice's subsystem. Next we have to prove that the mapping is unitary. The fact that $\langle\Omega_t|\Omega_\tau\rangle = \delta_{t\tau}$ shows that orthogonality in Bob's space is correctly preserved. In order to demonstrate full preservation of orthogonality we have to define the action of the operator U

on states of the form $|t\rangle_{\mathrm{B}} \otimes |\varepsilon\rangle_{\mathrm{E}}$, where $|\varepsilon\rangle$ is one of Eve's basis vectors orthogonal to $|E_0\rangle$, in such a way that the resulting states are mutually orthogonal and orthogonal to all $|\Omega_t\rangle$, $t \in \{0, \ldots, d-1\}$. The dimension of the BE space is d^3 and allows us to make such a choice of $d(d^2-1)$ vectors. $\square$

Theorem 10. *Let Alice send the state $|\mu_a\rangle$ to Bob. Let Eve apply the unitary operation U (specified in the proof of Theorem 9) to this state and her ancilla. The result can be written as*

$$U\Big(|\mu_a\rangle \otimes |E_0\rangle\Big) = \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} (-1)^{a_t} |t\rangle \otimes |A_t^a\rangle, \tag{103}$$

$$\begin{aligned} |A_t^a\rangle \stackrel{\text{def}}{=} & \sqrt{\lambda_0}|E_0\rangle + \sqrt{\lambda_1} \sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}jt} |E_j\rangle \\ & + \sqrt{\frac{d}{2}} \sum_{t':t'\neq t} (-1)^{a_t + a_{t'}} \Big[\sqrt{\lambda_+}|E^+_{(tt')}\rangle + \sqrt{\lambda_-}\mathrm{sgn}(t'-t)|E^-_{(tt')}\rangle\Big]. \end{aligned} \tag{104}$$

The states $|A_t^a\rangle$ are normalised and satisfy $\forall_{t\tau:\tau\neq t}$ $\langle A_\tau^a|A_t^a\rangle = (1-2\beta)$.

Proof: We start from $U(|\mu_a\rangle|E_0\rangle) = (1/\sqrt{d})\sum_t (-1)^{a_t}|\Omega_t\rangle$ and we substitute (102). Re-labeling of summation variables yields (103,104). The norm $\langle A_t^a|A_t^a\rangle$ equals $\lambda_0 + (d-1)\lambda_1 + \frac{d(d-1)}{2}\lambda_+ + \frac{d(d-1)}{2}\lambda_-$, which equals 1 since this is also equal to the trace of $\tilde{\rho}^{\mathrm{AB}}$. For $\tau \neq t$ the inner product $\langle A_\tau^a|A_t^a\rangle$ yields

$$\begin{aligned} \lambda_0 & + \lambda_1 \sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}j(t-\tau)} \\ & + \frac{d}{2} \sum_{t'\neq t} \sum_{\tau'\neq\tau} (-1)^{a_t + a_{t'} + a_\tau + a_{\tau'}} \delta_{t'\tau}\delta_{\tau' t}[\lambda_+ + \lambda_- \mathrm{sgn}(t'-t)\mathrm{sgn}(\tau'-\tau)]. \end{aligned} \tag{105}$$

We use $\sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}j(t-\tau)} = d\delta_{\tau t} - 1 = -1$. Furthermore the Kronecker deltas in (105) set the phase $(-1)^{\cdots}$ to 1 and $\mathrm{sgn}(t'-t)\mathrm{sgn}(\tau'-\tau) = \mathrm{sgn}(\tau-t)\mathrm{sgn}(t-\tau) = -1$. Finally we use $\lambda_0 - \lambda_1 = 1 - 2\beta - V$ and $\lambda_+ - \lambda_- = 2V/d$. $\square$

Theorem 10 reveals an intuitive picture. In the noiseless case ($\beta = 0$) it holds that $\forall_t$ $|A_t^a\rangle = |E_0\rangle$, i.e. Eve does nothing, resulting in the factorised state $|\mu_a\rangle|E_0\rangle$. In the case of extreme noise ($\beta = \frac{1}{2}$) we have $\langle A_t^a|A_\tau^a\rangle = \delta_{t\tau}$, which corresponds to a maximally entangled state between Bob and Eve.

Corollary 6. *The pure state (103) in Bob and Eve's space gives rise to the following mixed state ρ_a^{B} in Bob's subsystem,*

$$\rho_a^{\mathrm{B}} = (1-2\beta)|\mu_a\rangle\langle\mu_a| + 2\beta\frac{\mathbb{1}}{d}. \tag{106}$$

Proof: Follows directly from (103) by tracing out Eve's space and using the inner product $\langle A_\tau^a|A_t^a\rangle = (1-2\beta)$ for $\tau \neq t$. $\square$

From Bob's point of view, what he receives is a mixture of the $|\mu_a\rangle$ state and the fully mixed state. The interpolation between these two is linear in β. Note that the parameters μ, V are not visible in ρ_a^{B}.

B Optimization for the min-entropy

Here we prove that (88,89) maximizes (91). We first show that (91) is concave and obtain the optimum for $\beta \geq \beta_{\rm sat}$. Then we take into account the constraints on the eigenvalues and derive the optimum for $\beta < \beta_{\rm sat}$.

Unconstrained optimization. For notational convenience we define

$$w_1 = \sqrt{d\lambda_+ + 2\beta - \frac{d^2}{2}(\lambda_+ + \lambda_-)} \tag{107}$$

$$w_2 = \sqrt{d\lambda_- + 2(1-\beta) - \frac{d^2}{2}(\lambda_+ + \lambda_-)}. \tag{108}$$

This allows us to formulate everything in terms of λ_+ and λ_-. Eq. (91) becomes

$$\frac{1}{2}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = \sqrt{d\lambda_-}w_1 + \sqrt{d\lambda_+}w_2. \tag{109}$$

Next we compute the derivatives,

$$\frac{\partial}{\partial\lambda_+}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = -\frac{d^2}{2}\frac{\sqrt{\lambda_+}}{w_2} + \frac{w_2}{\sqrt{\lambda_+}} + (d - \frac{d^2}{2})\frac{\sqrt{\lambda_-}}{w_1}. \tag{110}$$

$$\frac{\partial}{\partial\lambda_-}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = -\frac{d^2}{2}\frac{\sqrt{\lambda_-}}{w_1} + \frac{w_1}{\sqrt{\lambda_-}} + (d - \frac{d^2}{2})\frac{\sqrt{\lambda_+}}{w_2}. \tag{111}$$

Setting both these derivatives to zero yields a stationary point of the function. Setting $w_1\sqrt{\lambda_+}\frac{\partial}{\partial\lambda_+}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 - w_2\sqrt{\lambda_-}\frac{\partial}{\partial\lambda_-}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1$ to zero gives $\lambda_+ w_1^2 - \lambda_- w_2^2 = 0$, which describes a hyperbola

$$(\tfrac{1}{2}d^2 - d)(\lambda_-^2 - \lambda_+^2) + 2\beta\lambda_+ - 2(1-\beta)\lambda_- = 0. \tag{112}$$

Next, the equations $\sqrt{\lambda_+}w_1 w_2\frac{\partial}{\partial\lambda_+}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = 0$ and $\sqrt{\lambda_-}w_1 w_2\frac{\partial}{\partial\lambda_-}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = 0$ can both easily be written in the form $\frac{w_2}{w_1} = expression$. Equating these two expressions gives us another hyperbola,

$$\left(d^2\lambda_+ + \frac{d^2}{2}\lambda_- - d\lambda_- - 2(1-\beta)\right)\left(d^2\lambda_- + \frac{d^2}{2}\lambda_+ - d\lambda_+ - 2\beta\right) - \lambda_-\lambda_+(d - \frac{d^2}{2}) = 0. \tag{113}$$

The stationary point lies at the crossing of these two hyperbolas. There are four crossing points,

$$\lambda_+ = 0\,; \qquad \lambda_- = \frac{4(1-\beta)}{d(d-2)} \tag{114}$$

$$\lambda_+ = \frac{4\beta}{d(d-2)} \quad ; \quad \lambda_- = 0 \tag{115}$$

$$\lambda_+ = \frac{1}{2d(d-1)} + \frac{1-2\beta}{d^2} \quad ; \quad \lambda_- = \frac{1}{2d(d-1)} - \frac{1-2\beta}{d^2} \tag{116}$$

$$\lambda_+ = \frac{2 + d(1-2\beta)}{2d^2} \quad ; \quad \lambda_- = \frac{2 - d(1-2\beta)}{2d^2}. \tag{117}$$

In the steps above, we have multiplied our derivatives by λ_+, λ_-, w_1 and w_2; this has introduced spurious zeros that now need to be removed. From (110,111) it is easily seen that $\lambda_+ = 0$ and $\lambda_- = 0$ are never stationary points since the derivatives diverge near these values. Furthermore, we find that substitution of (117) into the derivatives does not yield two zeros. Expression (116) is the only stationary point. As the function value lies higher there than in other points, we conclude that $\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1$ is concave.

Constrained optimization. The optimization problem is constrained by the fact that the λ eigenvalues are non-negative. For $\beta \geq \beta_{\text{sat}}$ the stationary point satisfies the constraints and hence is the optimal choice for $\beta \geq \beta_{\text{sat}}$.

For $\beta < \beta_{\text{sat}}$ the stationary point has $\lambda_- < 0$, i.e. it lies outside the allowed region. Because of the concavity the highest function value which satisfies the constraints occurs at $\lambda_0 = 0$, $\lambda_1 = 0$, $\lambda_+ = 0$ or $\lambda_- = 0$. It is easily seen that $\lambda_0 \geq 0$ implies $\lambda_+ \leq \frac{1}{d-1} - \frac{2\beta}{d}$ and $\lambda_1 \geq 0$ implies $\lambda_+ \leq \frac{4\beta}{d(d-2)} - \frac{d}{d-2}\lambda_-$ and $\lambda_- \leq \frac{4\beta}{d^2} - \frac{d-2}{d}\lambda_+$. In the range $\beta < \beta_{\text{sat}}$ it holds that $\frac{4\beta}{d(d-2)} < \frac{1}{d-1} - \frac{2\beta}{d}$; hence the λ_0-constraint is irrelevant in this region. We get $\lambda_1 = 0$ when $\lambda_+ = \frac{4\beta}{d(d-2)} - \frac{d}{d-2}\lambda_-$. Substitution gives $\frac{1}{2}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 =$ $\frac{\sqrt{2}}{d-2}\sqrt{2(1-\beta) + d(1 - 2\beta + d\,(1 - 2\beta(d-1)\lambda_-))\,(d^2\lambda_- - 4\beta)}$ which has its maximum at $\lambda_- = 0$ for non-negative values of λ_-. So either $\lambda_- = 0$ or $\lambda_+ = 0$. This leaves two options for the maximum at low β,

$$\lambda_+ = 0 \quad ; \quad \lambda_- = \frac{4\beta}{d^2} \quad \Rightarrow \quad \frac{1}{2}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = 0. \tag{118}$$

$$\lambda_- = 0 \quad ; \quad \lambda_+ = \frac{4\beta}{d(d-2)} \quad \Rightarrow \quad \frac{1}{2}\left\|\sigma_0^{rk} - \sigma_1^{rk}\right\|_1 = 2\sqrt{2}\frac{\sqrt{\beta(d-2) - 2\beta^2(d-1)}}{d-2}. \tag{119}$$

Clearly (119) is the larger of the two and therefore the optimal choice. □

References

1. W.K. Wootters and W.H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.
2. A. Broadbent and C. Schaffner. Quantum cryptography beyond quantum key distribution. *Des. Codes Cryptogr.*, 78:351–382, 2016.
3. C.H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. *IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.
4. A.K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661 – 663, 1991.
5. C.H. Bennett. Quantum cryptography using any two nonorthogonal states. *Pys. Rev. Lett.*, 68(21):3121–3124, 1992.
6. T.C. Ralph. Continuous variable quantum cryptography. *Phys. Rev. A*, 61:010303, 1999.

7. M. Hillery. Quantum cryptography with squeezed states. *Phys. Rev. A*, 61:022309, 2000.
8. D. Gottesman and J. Preskill. Secure quantum key distribution using squeezed states. *Phys. Rev. A*, 63:022309, 2001.
9. K. Inoue, E. Waks, and Y. Yamamoto. Differential phase shift quantum key distribution. *Phys. Rev. Lett.*, 89(3):037902,1–3, 2002.
10. D. Bruß. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, 81(14):3018–3021, 1998.
11. P. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Phys.Rev.Lett.*, 85:441, 2000.
12. R. Renner, N. Gisin, and B. Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Phys.Rev.A*, 72:012332, 2005.
13. B. Kraus, N. Gisin, and R. Renner. Lower and upper bounds on the secret key rate for quantum key distribution protocols using one-way classical communication. *Phys.Rev.Lett.*, 95:080501, 2005.
14. T. Sasaki, Y. Yamamoto, and M. Koashi. Practical quantum key distribution protocol without monitoring signal disturbance. *Nature*, 509:475–478, May 2014.
15. Hoi-Kwong Lo. Efficient quantum key distribution scheme and a proof of its unconditional security. *Journal of Cryptology*, 2005.
16. Z. Zhang, X. Yuan, Z. Cao, and X. Ma. Round-robin differential-phase-shift quantum key distribution. `http://arxiv.org/abs/1505.02481v1`, 2015.
17. T. Sasaki and M. Koashi. A security proof of the round-robin differential phase shift quantum key distribution protocol based on the signal disturbance. `https://arxiv.org/abs/1701.08509`, 2017.
18. B. Škorić. A short note on the security of Round-Robin Differential Phase-Shift QKD, 2017.
19. R. König, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. *IEEE Trans.Inf.Th.*, 55(9):4337–4347, 2009.
20. A.S. Holevo. Statistical decision theory for quantum systems. *Journal of multivariate analysis*, 3:337–394, 1973.
21. R. König R. Renner. Universally composable privacy amplification against quantum adversaries. *Proc. of TCC, LNCS 3378*, 2005.