

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’ (RRDPS) 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 show that the RRDPS protocol is equivalent to RRDPS with basis permutation and phase flips performed by Alice and Bob; this causes a symmetrisation of Eve’s state. 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 non-asymptotic key size, the trace distance between the real state and a state in which no leakage exists. We invoke post-selection in order to go from qudit-wise attacks to general attacks. For asymptotic key size we obtain a bound on the trace distance based on the von Neumann entropy.

Our asymptotic result for the privacy amplification is sharper than existing bounds. At low qudit dimension, even our non-asymptotic result is sharper than existing asymptotic 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 -dimensional 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 $1 - h(\beta) - I_{AE}$, where h is the binary entropy function and β the bit error rate.

1.2 Prior work on the security of RRDPS

The security of RRDPS has been discussed in a number of papers [14–17]. The original RRDPS paper gives an asymptotic upper bound for the privacy amplification,

$$I_{AE} \leq h\left(\frac{1}{d-1}\right) \quad (1)$$

(Eq. 5 in [14] with photon number set to 1). The security analysis in [14] is based on the Shor-Preskill proof technique [11] and an estimate of the phase error. It is not known how tight the bound (1) is. Ref. [15] follows [14] and does a more accurate computation of phase error rate, tightening the $1/(d-1)$ in (1) to $1/d$. In [16] Sasaki and Koashi add noise-dependence to their analysis and claim a bound

$$I_{\text{AE}} \leq h\left(\frac{2\beta}{d-2}\right) \quad \text{for } \beta \leq \frac{1}{2} \cdot \frac{d-2}{d-1} \quad (2)$$

and $I_{\text{AE}} \leq h\left(\frac{1}{d-1}\right)$ for $\beta \in \left[\frac{1}{2} \cdot \frac{d-2}{d-1}, \frac{1}{2}\right]$. (See Section 5). The analysis in [17] considers only intercept-resend attacks, and hence puts a *lower bound* on Eve's potential knowledge, $I_{\text{AE}} \geq 1 - h\left(\frac{1}{2} + \frac{1}{d}\right) = \mathcal{O}(1/d^2)$.¹

1.3 Contributions and outline

In this paper we give a security proof of RRDPS. We give a bound on the required amount of privacy amplification. We use 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 show that the RRDPS protocol is equivalent to a protocol that contains an additional randomisation step by Alice and Bob. The randomisation consists of phase flips and a permutation of the basis states. We construct an EPR variant of RRDPS-with-randomisation; it is equivalent to RRDPS if Alice creates the EPR pair and immediately does her measurement. The effect of the randomisation is that Alice and Bob's entangled state after Eve's attack on the EPR pair is symmetrised and can be described using just three real degrees of freedom.
- 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 .
- 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 (after PA) of the full QKD key 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 go from qudit-wise attacks to general attacks by using the post-selection technique. This inflicts a penalty $(d^4 - 1) \log(n + 1)$ on the amount of privacy amplification.
- We compute the von Neumann mutual information between one ancilla state and Alice's secret bit. This provides a bound on the PA 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. These results give some insight into simple attacks that Eve can launch against individual qudits.

In Section 2 we introduce notation and post-selection. In Section 3 we briefly summarise the RRDPS scheme, and discuss the attacker model. Section 4 states the main result: the amount of privacy amplification needed for RRDPS to be secure, (i) at finite key length and (ii) asymptotically. Section 5 compares our results to previous bounds. The remainder of the paper builds towards the proof of the main results. In Section 6 we show that the randomisation step does not modify RRDPS, and we introduce the EPR version of the protocol. In Section 7 we impose

¹ Ref. [17] gives a min-entropy of $-\log\left(\frac{1}{2} + \frac{1}{d}\right)$, which translates to Shannon entropy $h\left(\frac{1}{2} + \frac{1}{d}\right)$.

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 8 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 9 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 10 we prove the asymptotic result by computing Eve’s knowledge in terms of von Neumann entropy. The appendix provides supplementary information about the leakage in terms of min-entropy loss and accessible information.

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 $H(X)$. 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 $H_{\min}(X) = -\log \max_{x \in \mathcal{X}} \Pr[X = x]$, and the conditional min-entropy is $H_{\min}(X|Y) = -\log \mathbb{E}_y \max_{x \in \mathcal{X}} \Pr[X = x|Y = y]$. 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 $\|A\|_1 = \text{tr} \sqrt{A^\dagger A}$. The trace distance between matrices ρ and σ is denoted as $D(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1$; it is a generalisation of the statistical distance and represents the maximum possible advantage one can have in distinguishing ρ from σ . The von Neumann entropy of a mixed state ρ is denoted as $S(\rho)$ and equals $-\text{tr} \rho \log \rho$.

Consider a bipartite system ‘XE’ where X is a uniform classical random variable and Eve’s part ‘E’ depends on X . The combined quantum-classical state is $\rho^{\text{XE}} = \mathbb{E}_x |x\rangle\langle x| \otimes \rho^{\text{E}}(x)$. The individual parts are in state $\rho^{\text{X}} = \mathbb{E}_x |x\rangle\langle x|$ and $\rho^{\text{E}} = \mathbb{E}_x \rho^{\text{E}}(x)$ respectively. The statistical distance between X and a uniform variable given $\rho(X)$ is a measure of the security of X given ρ . This distance is given by [18]

$$\mathcal{D}(X|\rho^{\text{E}}(X)) \stackrel{\text{def}}{=} D(\rho^{\text{XE}}, \rho^{\text{X}} \otimes \rho^{\text{E}}) = \frac{1}{2} \|\rho^{\text{XE}} - \rho^{\text{X}} \otimes \rho^{\text{E}}\|_1 \quad (3)$$

i.e. the distance between the true quantum-classical state and a state in which Eve’s part is decoupled from X . If the distance is ε , then it is said that X is ε -secure. Statements like (3) that are stated in terms of statistical distance have the advantage of being universally composable [18]. The term Privacy Amplification is abbreviated as PA.

2.2 Post-selection

In a *collective attack* Eve acts on individual qudits. This is not the most general attack. For protocols that obey permutation symmetry, a post-selection argument [19] can be used to show that ε -security against collective attacks implies ε' -security against general attacks, with $\varepsilon' = \varepsilon(n+1)^{d^4-1}$, where d is the dimension of the qudit space. Hence, by paying a price in terms of privacy amplification, e.g. changing the usual privacy amplification term $2 \log \frac{1}{\varepsilon}$ to $2 \log \frac{1}{\varepsilon} + 2(d^4 - 1) \log(n+1)$, one can ‘buy’ security against general attacks.

3 The RRDPS scheme

We briefly review the RRDPS scheme [14]. For proof-technical reasons we explicitly include a channel monitoring procedure (step 4); our proof technique needs this step in the non-saturated regime. Step 4 can be omitted if Alice and Bob decide to perform privacy amplification as if Eve causes 50% noise in every qudit.

3.1 The RRDPS protocol

The dimension of the qudit space is d . The basis states² 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 number of qudits is denoted as n . We introduce a system parameter L denoting a list length and system parameters $\tilde{\beta} \in [0, \frac{1}{2}]$, $\eta \ll 1$ related to the tolerated noise level. 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

$$|\mu_a\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} (-1)^{a_t} |t\rangle \quad (4)$$

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 $(M_{ks}^{(r)})_{k \in \{0, \dots, d-1\}, s \in \{0, 1\}}$,

$$M_{ks}^{(r)} = \frac{1}{2} |\Psi_{ks}^{(r)}\rangle \langle \Psi_{ks}^{(r)}| \quad |\Psi_{ks}^{(r)}\rangle = \frac{|k\rangle + (-1)^s |k+r\rangle}{\sqrt{2}}. \quad (5)$$

The result of the measurement $\mathcal{M}^{(r)}$ on $|\mu_a\rangle$ is an integer $k \in \{0, \dots, d-1\}$ and a bit s which equals $a_k \oplus a_{k+r}$ if there is no noise/interference.³

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 .

Steps 1–3 are repeated N times.

4. Alice selects a random subset $\mathcal{L} \subset [N]$, with $|\mathcal{L}| = L$. For the rounds indicated by \mathcal{L} , Alice and Bob publicly compare their values of s' and s . They continue the protocol only if the number of occurrences $s \neq s'$ is smaller than $\tilde{\beta}L$.
5. Finally, on the remaining $n = N - L$ bits Alice and Bob carry out the standard procedures of information reconciliation and Privacy Amplification. After PA the size of the key is ℓ bits.

If step 4 is not performed, Alice and Bob have to assume that Eve learns as much as when causing bit error rate $\frac{1}{2}$. This mode of operation (without monitoring) was proposed in the original RRDPS paper [14].

If Eve causes bit error probability exceeding β (with $\beta > \tilde{\beta}$), her probability of passing step 4 is exponentially small. Applying the Hoeffding inequality yields an upper bound on the probability of $\exp[-2L(\beta - \tilde{\beta})^2]$. Let $\eta \ll 1$ be a security parameter. By setting $\tilde{\beta} \leq \beta - \sqrt{\frac{1}{2L} \ln \frac{1}{\eta}}$, we can make sure that an Eve who causes bit error probability exceeding β fails the test except with probability η .

In order for \mathcal{L} to be statistically representative, L needs to be at least of order $\log \ell$ [20]. We will assume $L > \log \ell$.

² 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.

³ 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 $|\Psi_{\ell u}^{(r)}\rangle$, where r accidentally equals Bob's r , then there is a 50% probability that Bob's measurement $\mathcal{M}^{(r)}$ yields $k \neq \ell$ with random s .

3.2 Attacker model

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 will first analyze attacks in which Eve couples an ancilla to each EPR pair individually. Then we invoke the *post-selection* method [19] to cover general attacks.

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.)

4 Main results

4.1 Non-asymptotic result

Our first result is a non-asymptotic bound on the secrecy of the QKD key z .

Theorem 1 *Let $\mathbf{r} = (r_1, \dots, r_n)$ be the values of the parameter r in n rounds of RRDPS, and similarly $\mathbf{k} = (k_1, \dots, k_n)$. Let $z \in \{0, 1\}^\ell$ be the QKD key derived from the n rounds. Let u be the (public) random seed used in the privacy amplification. Let $\beta \in [0, \frac{1}{2}]$. Consider a collective attack such that Eve's probability of causing a bit flip, averaged per qudit, does not exceed β . At given \mathbf{r}, \mathbf{k} let $\rho^{\text{ZUE}}(\mathbf{r}, \mathbf{k})$ denote the quantum-classical state of the variables Z, U and Eve's subsystem 'E', which consists of all n ancillas. The security of Z given R, K, U and Eve's quantum information can be expressed as*

$$\frac{1}{2} \|\rho^{\text{ZUE}}(\mathbf{r}, \mathbf{k}) - \rho^{\text{Z}}(\mathbf{r}, \mathbf{k}) \otimes \rho^{\text{UE}}(\mathbf{r}, \mathbf{k})\|_1 < \frac{1}{2} \sqrt{2^{\ell-n(1-2\log T)}} \quad (6)$$

where T is given by

$$\beta \leq \beta_* : \quad T = 2\beta + \sqrt{1 - 2\beta} \left[\sqrt{1 - 2\beta} \frac{d-1}{d-2} + \frac{\sqrt{2\beta}}{\sqrt{d-2}} \right] \quad (7)$$

$$\beta \geq \beta_* : \quad T = 2\beta_* + \sqrt{1 - 2\beta_*} \left[\sqrt{1 - 2\beta_*} \frac{d-1}{d-2} + \frac{\sqrt{2\beta_*}}{\sqrt{d-2}} \right] \quad (8)$$

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

$$\beta_* = \frac{x_d/2}{1 + x_d}, \quad (9)$$

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

$$\left(1 - \frac{x}{d-2}\right)^{\frac{1}{2}} + \left(1 + \frac{1}{d-2}\right) \left(1 - \frac{x}{d-2}\right)^{-\frac{1}{2}} + \frac{1}{\sqrt{d-2}} \left(\sqrt{x} - \frac{1}{\sqrt{x}}\right) - 2 = 0. \quad (10)$$

The proof is given in Section 9, after several sections that prepare the ground. Theorem 1 holds for attacks in which Eve couples an ancilla to each individual EPR pair (though she may later act in any way whatsoever on the whole set of n ancillas). As explained in Section 2.2, by a post-selection argument security against qudit-wise attacks implies security against *general attacks*, but with a less favourable security parameter. In the case of general attacks, we have to multiply the right hand side of (6) by $(n+1)^{d^4-1}$. Hence, in order to obtain ε -security we have to set $\ell = n(1 - 2 \log T) - (d^4 - 1) \log(n+1) - 2 \log \frac{1}{\varepsilon} + 2$.

Let Alice and Bob use an error-correcting code with codeword size n and syndrome size σ . The information reconciliation leaks σ bits of information. (Or consumes σ bits of key material, depending on the information reconciliation procedure). It holds that $\sigma > nh(\tilde{\beta})$. Asymptotically $\tilde{\beta} \rightarrow \beta$ and $\sigma \rightarrow nh(\beta)$. The QKD key generation rate is $(\ell - \sigma)/N$, where $N = n + L$ (see Section 3.1), with L the number of qubits spent on channel monitoring (if monitoring is performed at all).

$$\text{rate} = \frac{n}{n+L}(1 - 2 \log T) - \frac{\sigma}{n+L} - (d^4 - 1) \frac{\log(n+1)}{n+L} - \frac{2}{n+L} \log \frac{1}{\varepsilon} + \frac{2}{n+L} \quad (11)$$

$$\geq \left(1 - \frac{L}{n}\right)(1 - 2 \log T) - \frac{\sigma}{n} - (d^4 - 1) \frac{\log(n+1)}{n} - \frac{2}{n} \log \frac{1}{\varepsilon}. \quad (12)$$

The achieved security level is $\max\{\varepsilon, \eta\}$, where η is the probability that the number of bit errors is smaller than $\tilde{\beta}L$ when Eve causes bit error probability larger than β (see Section 3.1).⁴ It is advantageous to set $\eta = \varepsilon$.

4.2 Asymptotic result

For asymptotically large n , it has been shown [21], using the properties of smooth Rényi entropies, that $\|\rho^{\text{ZUE}} - \rho^{\text{Z}} \otimes \rho^{\text{UE}}\|_1 \leq \sqrt{2^{\ell-n(1-I_{\text{AE}})}}$, where I_{AE} is the single-qudit von Neumann information leakage, $I_{\text{AE}} \stackrel{\text{def}}{=} S(E) - S(E|S')$. 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:*

$$\beta \leq \beta_0 : \quad I_{\text{AE}} = (1 - 2\beta)h\left(\frac{1}{d-2} \cdot \frac{2\beta}{1-2\beta}\right) \quad (13)$$

$$\beta \geq \beta_0 : \quad I_{\text{AE}} = (1 - 2\beta_0)h\left(\frac{1}{d-2} \cdot \frac{2\beta_0}{1-2\beta_0}\right). \quad (14)$$

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

$$\beta_0 = \frac{1}{2} \left[1 + \frac{1}{(d-2)(1-y_d)} \right]^{-1} \quad (15)$$

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

The proof is given in Section 10. The formulation of our main results in terms of statistical distance ensures that the results are Universally Composable. In Section 10 we will see that Theorem 2 is sharper than (2) and hence allows for a higher QKD key generation rate.

⁴ If this rare event occurs, we have no other bound than $\frac{1}{2} \|\rho^{\text{ZUE}} - \rho^{\text{Z}} \otimes \rho^{\text{UE}}\|_1 \leq 1$.

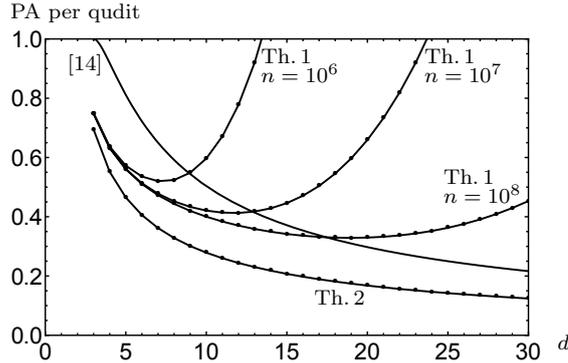


Fig. 1 Saturated PA per qudit as a function of d . Comparison of [14] and our results (Theorem 1 and Theorem 2). Our non-asymptotic result is shown for several values of n .

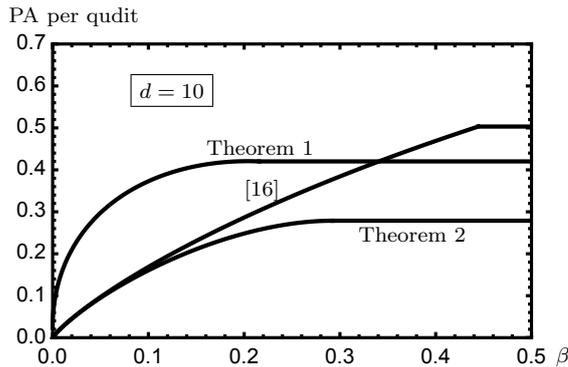


Fig. 2 Amount of Privacy Amplification per qudit as a function of β , for $d = 10$. Comparison of our Theorem 1 ($n = 10^7$) and Theorem 2 versus the PA of [16], which equals $h(\frac{2\beta}{d-2})$ below saturation and $h(\frac{1}{d-1})$ above saturation.

5 Comparison with previous analyses

5.1 Phase error

Sasaki and Koashi [16] provided an upper bound on the PA equal to $h(e^{\text{ph}})$, where e^{ph} is the phase error rate. They derived a relation between the phase error rate and the bit error rate, $e^{\text{ph}} \leq \inf_{\lambda \geq 0} [\lambda\beta + \max\{\Omega_-(\nu, \lambda), \Omega_+(\nu, \lambda)\}]$, where ν is the photon number and Ω_{\pm} are functions which for $\nu = 1$ reduce to $\Omega_-(1, \lambda) = 0$ and $\Omega_+(1, \lambda) = \frac{1}{d-1} - \lambda \frac{d-2}{2(d-1)}$. At $\nu = 1$ the optimal λ is $\frac{2}{d-2}$, yielding $e^{\text{ph}} \leq \frac{2\beta}{d-2}$ and thus an upper bound of $h(\frac{2\beta}{d-2})$ on the PA.

5.2 Comparison

We first compare our asymptotic result (Theorem 2) to the asymptotic $h(\frac{2\beta}{d-2})$ of [16]. For all $\beta \in [0, \beta_0]$ and $d > 2$ it holds that

$$(1 - 2\beta)h(\frac{2\beta}{(d-2)(1-2\beta)}) \leq h(\frac{2\beta}{d-2}). \quad (16)$$

This is verified as follows. Let $p_0 = 2\beta$, $p_1 = 1 - 2\beta$, $x_0 = 0$, $x_1 = \frac{2\beta}{(d-2)(1-2\beta)}$. The left hand side of (16) can be expressed as $p_0h(x_0) + p_1h(x_1)$, and the right hand side as $h(p_0x_0 + p_1x_1)$. Because h is concave we have $\mathbb{E}h(\dots) \leq h(\mathbb{E}\dots)$.

Thus our von Neumann result is sharper than [16]. It is difficult to pinpoint what causes the difference in tightness.

Note too that our saturation occurs at lower β than in [16], especially for small d .

Our Theorem 1 is *non-asymptotic*; we cannot compare it to previous results since the previous results are for the asymptotic regime.

Figs.1 and 2 show plots of the PA per qudit. In Fig.1 the post-selection ‘price’ proportional to $d^4 \frac{\log n}{n}$ is clearly visible; for large d the cost is prohibitive. Interestingly, at small d our non-asymptotic result for the saturated PA is sharper than the asymptotic $h(\frac{1}{d-1})$ [14,16].

6 Symmetrised EPR version of the protocol

6.1 RRDPS is equivalent to RRDPS with random permutations

We show that inserting a symmetrisation step into RRDPS does not affect the protocol. More specifically, the following protocol is equivalent to RRDPS steps 1 to 3. (For brevity we do not explicitly write down the channel monitoring, information reconciliation and privacy amplification.)

- S1** Alice picks a random $a \in \{0, 1\}^d$ and a random permutation π .
She prepares $|\mu_a\rangle = \frac{1}{\sqrt{d}} \sum_t (-1)^{a_t} |t\rangle$.
- S2** Alice performs the permutation π on the state $|\mu_a\rangle$. She sends the result to Bob. After pausing for a while, she sends π to Bob.
- S3** Eve does something with the state, without knowing π . Then she sends the result to Bob.
- S4** Bob receives a state and stores it until he receives π . Bob applies π^{-1} to the state.
- S5** Bob picks a random $r \in \{1, \dots, d-1\}$ and does the $\mathcal{M}^{(r)}$ POVM. The result is an index $k \in \{0, \dots, d-1\}$ and a bit $s = a_k \oplus a_{k+r}$. He computes $\ell = k + r \bmod d$. He announces k, ℓ .
- S6** Alice computes $s' = a_k \oplus a_\ell$.

The equivalence is shown as follows. After step S2, the state is $\frac{1}{\sqrt{d}} \sum_t (-1)^{a_t} |\pi(t)\rangle$
 $= \frac{1}{\sqrt{d}} \sum_\tau (-1)^{a_{\pi^{-1}(\tau)}} |\tau\rangle = |\mu_{\pi^{-1}(a)}\rangle$. Hence Alice’s process {state preparation followed by π } can be replaced by {acting with π^{-1} on a followed by state preparation}. Similarly, Bob’s process {apply π^{-1} to state; pick random r ; do $\mathcal{M}^{(r)}$; send k, ℓ } has exactly the same effect as {pick random r ; do $\mathcal{M}^{(r)}$; apply π to k, ℓ ; send $\pi(k), \pi(\ell)$ }. Next, Bob’s computation of $\pi(k), \pi(\ell)$ can be moved to Alice. Then, Alice’s actions {pick random a' ; send $\pi^{-1}(a)$ to state preparation; send a to step S6} can be replaced by {pick random a' ; send a' to state preparation; send $\pi(a)$ to step S6}. Finally, in step S6 we use $\pi(a)_{\pi(k)} = a_k$ and $\pi(a)_{\pi(\ell)} = a_\ell$.

Remark. In step S3 it is crucial that Eve does not know π at the moment of her manipulation of the state. This will allow us to derive a symmetrised form of the density matrix in Section 6.3.

6.2 RRDPS is equivalent to RRDPS with random phase flips

Analogous with Section 6.1, it can be seen that adding an extra phase-flipping step to RRDPS does not affect RRDPS. Consider the following protocol.

- F1** Alice picks a random $a \in \{0, 1\}^d$ and a random $c \in \{0, 1\}^d$. She prepares $|\mu_a\rangle = \frac{1}{\sqrt{d}} \sum_t (-1)^{a_t} |t\rangle$.
- F2** Alice performs the phase flips on the state $|\mu_a\rangle$, according to the rule $|t\rangle \rightarrow (-1)^{c_t} |t\rangle$ for basis states. She sends the result to Bob. After pausing for a while, she sends c to Bob.
- F3** Eve does something with the state, without knowing c . Then she sends the result to Bob.
- F4** Bob receives a state and stores it until he receives c . Bob applies phase flips c to the state.
- F5** Bob picks a random $r \in \{1, \dots, d-1\}$ and does the $\mathcal{M}^{(r)}$ POVM. The result is an index $k \in \{0, \dots, d-1\}$ and a bit $s = a_k \oplus a_{k+r}$. He computes $\ell = k + r \bmod d$. He announces k, ℓ .
- F6** Alice computes $s' = a_k \oplus a_\ell$.

The equivalence to RRDPS is seen as follows. After step F2 the state is $|\mu_{a \oplus c}\rangle$. Hence Alice’s process {pick random a ; prepare state; flip with c } is equivalent to {pick random a ; flip with c ; prepare state}. Similarly, Bob’s process {flip with c ; pick random r ; do $\mathcal{M}^{(r)}$ } is equivalent to {pick random r ; do $\mathcal{M}^{(r)}$; change s to $s \oplus c_k \oplus c_\ell$ }. This holds because in the first case Bob obtains $s = (a \oplus c)_k \oplus (a \oplus c)_\ell = (a_k \oplus a_\ell) \oplus c_k \oplus c_\ell$. Furthermore, Alice’s steps {pick random a ; send a to computation of s' and flipped a to state preparation} are equivalent to {pick random a' ; send flipped a to computation of s' and a' to state preparation}. The final effect of these

transformations of the ‘F’ protocol is that (i) there is no physical phase flipping at all, (ii) Bob needs no quantum memory, and (iii) Alice and Bob both obtain a secret bit $(a_k \oplus a_\ell) \oplus c_k \oplus c_\ell$; though not equal to $a_k \oplus a_\ell$, it is statistically the same.

6.3 EPR version

We introduce a protocol based on EPR pairs that is equivalent to the combined ‘S’ and ‘F’ protocols, and hence also equivalent to RRDPS.

E1 A maximally entangled two-qudit state is prepared.

$$|\alpha_0\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} |tt\rangle. \quad (17)$$

One qudit (‘A’) is intended for Alice, and one (‘B’) for Bob.

E2 Eve does something with the EPR pair. Then Alice and Bob each receive their own qudit.

E3 Alice and Bob pick a random permutation π . They both apply π to their own qudit. Then they forget π .

E4 Alice and Bob pick a random string $c \in \{0, 1\}^d$. They both apply phase flips $|t\rangle \rightarrow (-1)^{c_t} |t\rangle$ to their own qudit. Then they forget c .

E5 Alice performs a POVM $\mathcal{Q} = (Q_z)_{z \in \{0, 1\}^d}$ on her own qudit, where

$$Q_z = \frac{d}{2^d} |\mu_z\rangle \langle \mu_z|. \quad (18)$$

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

E6 Bob picks a random integer $r \in \{1, \dots, d-1\}$ and performs the POVM measurement $\mathcal{M}^{(r)}$ on his qudit. The result of the measurement is an integer $k \in \{0, \dots, d-1\}$ and a bit s . Bob computes $\ell = k + r \bmod d$. Bob announces k, ℓ .

E7 Alice computes $s' = a_k \oplus a_\ell$.

The equivalence to the protocol in Section 6.1 is seen as follows. First, let Alice be the origin of the EPR pair, and let her perform \mathcal{Q} as soon as she has created the EPR pair. This process is equivalent to preparing a qudit state $|\mu_a\rangle$ with random a . The only difference is that the EPR protocol allows Eve to couple her ancilla to the AB system instead of only the B system. Hence the EPR version *overestimates* Eve’s power. Security of the EPR version implies security of the original RRDPS.⁵ Furthermore, the permutations and phase flips in steps E3,E4 cancel out exactly like in protocols ‘S’ and ‘F’.

Remark: The protocol equivalences in Sections 6.1–6.3 can be nicely visualised using diagrammatic techniques [22]. We do not show the protocol diagrams in this paper.

Lemma 1 *The hermitian matrices Q_z as defined in (18) form a POVM, i.e. $\sum_{z \in \{0, 1\}^d} Q_z = \mathbb{1}$.*

Proof:

$$\sum_z |\mu_z\rangle \langle \mu_z| = \sum_z \frac{1}{d} \sum_{t, t'=0}^{d-1} (-1)^{z_{t'} + z_t} |t\rangle \langle t'| = \frac{1}{d} \sum_{t, t'=0}^{d-1} |t\rangle \langle t'| \sum_z (-1)^{z_{t'} + z_t}.$$

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 RRDPS.

⁵ In Appendix A it will turn out that Eve’s optimal attack is achieved by acting on Bob’s qudit only; hence the EPR version is fully equivalent to original RRDPS.

6.4 Effect of the random transforms: state symmetrisation

Let $\rho^{\text{AB}} = |\alpha_0\rangle\langle\alpha_0|$ denote the pure EPR state of Alice and Bob, and let $\hat{\rho}^{\text{AB}}$ be the mixed state of the AB system after Eve's manipulation in step E2. We write

$$\hat{\rho}^{\text{AB}} = \sum_{t,t',\tau,\tau' \in \{0,\dots,d-1\}} \hat{\rho}_{\tau\tau'}^{tt'} |t, t'\rangle\langle\tau, \tau'|, \quad (19)$$

with $\hat{\rho}_{\tau\tau'}^{tt'} = (\hat{\rho}_{\tau\tau'}^{tt'})^*$ and $\sum_{tt'} \hat{\rho}_{tt'}^{tt'} = 1$. The effect of step E3 is that the AB state gets averaged over all permutations, i.e. we get the following mapping

$$\hat{\rho}^{\text{AB}} \mapsto \tilde{\rho}^{\text{AB}} \stackrel{\text{def}}{=} \frac{1}{d!} \sum_{\pi} \sum_{t,t',\tau,\tau'} \hat{\rho}_{\pi(\tau),\pi(\tau')}^{\pi(t),\pi(t')} |t, t'\rangle\langle\tau, \tau'| \quad (20)$$

$$\stackrel{\text{def}}{=} \sum_{t,t',\tau,\tau'} \tilde{\rho}_{\tau\tau'}^{tt'} |t, t'\rangle\langle\tau, \tau'|. \quad (21)$$

Here the parameters $\tilde{\rho}_{\tau\tau'}^{tt'}$ are invariant under simultaneous permutation of the four indices, i.e. $\tilde{\rho}_{\pi(\tau),\pi(\tau')}^{\pi(t),\pi(t')} = \tilde{\rho}_{\tau\tau'}^{tt'}$ for all π, t, t', τ, τ' . The consequence is that $\tilde{\rho}^{\text{AB}}$ contains only a few degrees of freedom, namely the constants $\tilde{\rho}_{ss}^{ss}, \tilde{\rho}_{st}^{ss}, \tilde{\rho}_{ts}^{ss}, \tilde{\rho}_{tt}^{ss}, \tilde{\rho}_{st}^{st}, \tilde{\rho}_{ts}^{st}, \tilde{\rho}_{tu}^{ss}, \tilde{\rho}_{su}^{st}, \tilde{\rho}_{us}^{st}, \tilde{\rho}_{us}^{st}, \tilde{\rho}_{uv}^{st}$, where s, t, u, v are mutually distinct.

Next, the random phase flips reduce the degrees of freedom even further. Let F_c be the phase flip operator.

$$\bar{\rho}^{\text{AB}} \stackrel{\text{def}}{=} \mathbb{E}_{c \in \{0,1\}^d} F_c \tilde{\rho}^{\text{AB}} F_c^\dagger \quad (22)$$

$$= \mathbb{E}_c \sum_{tt'\tau\tau'} \tilde{\rho}_{\tau\tau'}^{tt'} (-1)^{c_t+c_{t'}+c_\tau+c_{\tau'}} |t, t'\rangle\langle\tau, \tau'| \quad (23)$$

$$= \sum_{tt'\tau\tau'} |t, t'\rangle\langle\tau, \tau'| \tilde{\rho}_{\tau\tau'}^{tt'} \mathbb{E}_c (-1)^{c_t+c_{t'}+c_\tau+c_{\tau'}} \quad (24)$$

$$\stackrel{\text{def}}{=} \sum_{tt'\tau\tau'} |t, t'\rangle\langle\tau, \tau'| \bar{\rho}_{\tau\tau'}^{tt'}. \quad (25)$$

From (24) we see that any time index that occurs an odd number of times will be wiped out, i.e. $\mathbb{E}_c (-1)^{c_t} = 0$. The only surviving degrees of freedom are the four constants $\bar{\rho}_{\bullet\bullet}^{\bullet\bullet}, \bar{\rho}_{\circ\circ}^{\circ\circ}, \bar{\rho}_{\bullet\circ}^{\bullet\circ}, \bar{\rho}_{\circ\bullet}^{\circ\bullet}$, where \bullet and \circ denote distinct arbitrary indices. Note that these constants are real-valued. We can now write

$$\bar{\rho}^{\text{AB}} = \bar{\rho}_{\bullet\bullet}^{\bullet\bullet} \sum_t |tt\rangle\langle tt| + \bar{\rho}_{\circ\circ}^{\circ\circ} \sum_{[t\tau]} |tt\rangle\langle\tau\tau| + \bar{\rho}_{\bullet\circ}^{\bullet\circ} \sum_{[tt']} |tt'\rangle\langle tt'| + \bar{\rho}_{\circ\bullet}^{\circ\bullet} \sum_{[tt']} |tt'\rangle\langle t't|. \quad (26)$$

Furthermore, the requirement $\text{tr} \bar{\rho}^{\text{AB}} = 1$ imposes the constraint $d\bar{\rho}_{\bullet\bullet}^{\bullet\bullet} + d(d-1)\bar{\rho}_{\circ\circ}^{\circ\circ} = 1$, reducing the number of degrees of freedom to three.

7 Imposing the noise constraint

The channel monitoring restricts the ways in which Eve can alter the AB state. We will determine the most general allowed $\bar{\rho}^{\text{AB}}$ that is compatible with bit error rate β . (We will later see that it is optimal for Eve to cause the same bit error rate in all rounds. This is due to the concavity of the leakage as a function of the error rate.) 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 $\bar{\rho}^{\text{AB}}$, 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*

$$P_{aks|r} = \frac{1}{2d2^d} + \frac{1}{2 \cdot 2^d} (\bar{\rho}_{\circ\circ}^{\circ\circ} + \bar{\rho}_{\bullet\bullet}^{\bullet\bullet}) (-1)^{s+a_k+a_{k+r}}. \quad (27)$$

Proof: $P_{aks|r} = \text{tr}(Q_a \otimes M_{ks}^{(r)}) \bar{\rho}^{\text{AB}}$
 $= \text{tr}(\frac{1}{2^d} \sum_{\ell\ell'} (-1)^{a_\ell + a_{\ell'}} |\ell\rangle\langle\ell'| \otimes \frac{1}{2} \frac{|k\rangle + (-1)^s |k+r\rangle}{\sqrt{2}} \frac{\langle k| + (-1)^s \langle k+r|}{\sqrt{2}}) \sum_{tt'\tau\tau'} \bar{\rho}_{\tau\tau'}^{tt'} |t\rangle\langle\tau| \otimes |t'\rangle\langle\tau'|$
 $= \frac{1}{2^{2d}} \sum_{tt'\tau\tau'} \bar{\rho}_{\tau\tau'}^{tt'} (-1)^{a_t + a_\tau} [\delta_{t'k} + (-1)^s \delta_{t',k+r}] [\delta_{\tau'k} + (-1)^s \delta_{\tau',k+r}]$
 $= \frac{1}{2^{2d}} \sum_{t\tau} (-1)^{a_t + a_\tau} [\bar{\rho}_{\tau k}^{tk} + \bar{\rho}_{\tau, k+r}^{t, k+r} + (-1)^s \bar{\rho}_{\tau, k+r}^{tk} + (-1)^s \bar{\rho}_{\tau k}^{t, k+r}]$. We use $\bar{\rho}_{\tau\ell}^{t\ell} = \delta_{t\ell} \delta_{\tau\ell} \bar{\rho}_{\bullet\bullet}^{\bullet\bullet} + \delta_{\tau t} (1 - \delta_{t\ell}) \bar{\rho}_{\bullet\circ}^{\bullet\circ}$ for the first two terms, setting $\ell = k$ and $\ell = k+r$. Since $k+r \neq k$ we write $\bar{\rho}_{\tau, k+r}^{tk} = \delta_{tk} \delta_{\tau, k+r} \bar{\rho}_{\circ\circ}^{\bullet\bullet} + \delta_{t, k+r} \delta_{\tau k} \bar{\rho}_{\circ\bullet}^{\bullet\circ}$, and similarly for $\bar{\rho}_{\tau k}^{t, k+r}$. Finally we use $\bar{\rho}_{\bullet\bullet}^{\bullet\bullet} + (d-1) \bar{\rho}_{\circ\circ}^{\bullet\bullet} = 1/d$. (See end of Section 6.4.) \square

We now impose the constraint that a bit error occurs with probability β ,

$$\Pr[S = A_K \oplus A_{K+R}] = 1 - \beta. \quad (28)$$

Here the random variables are A , R , K , and S .

Theorem 3 *The constraint (28) can only be satisfied by a density function of the form*

$$\bar{\rho}^{\text{AB}} = (1 - 2\beta - V) |\alpha_0\rangle\langle\alpha_0| + V \frac{1}{d} \sum_{tt'} |tt'\rangle\langle t't| + (2\beta - \mu) \frac{1}{d^2} + \mu \frac{1}{d} \sum_t |tt\rangle\langle tt| \quad (29)$$

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

$$\bar{\rho}_{\tau\tau'}^{tt'} = \frac{1 - 2\beta - V}{d} \delta_{t't} \delta_{\tau'\tau} + \frac{V}{d} \delta_{\tau t'} \delta_{\tau' t} + \frac{2\beta - \mu}{d^2} \delta_{\tau t} \delta_{\tau' t'} + \frac{\mu}{d} \delta_{t't} \delta_{\tau t} \delta_{\tau' t}. \quad (30)$$

Proof: We write $\Pr[S = A_K \oplus A_{K+R}] = \sum_{akrs} \frac{1}{d-1} P_{aks|r} \delta_{s, a_k \oplus a_{k+r}}$ and use Lemma 2. This yields $\Pr[S = A_K \oplus A_{K+R}] = \frac{1}{2} + \frac{d}{2} (\bar{\rho}_{\circ\circ}^{\bullet\bullet} + \bar{\rho}_{\bullet\circ}^{\bullet\circ})$. The constraint (28) can only be satisfied by setting $\bar{\rho}_{\circ\circ}^{\bullet\bullet} + \bar{\rho}_{\bullet\circ}^{\bullet\circ} = \frac{1-2\beta}{d}$. We choose $\bar{\rho}_{\circ\circ}^{\bullet\bullet}, \bar{\rho}_{\bullet\circ}^{\bullet\circ}$ as the two independent degrees of freedom and re-parametrise them as $\bar{\rho}_{\circ\circ}^{\bullet\bullet} = (1 - 2\beta - V)/d$ and $\bar{\rho}_{\bullet\circ}^{\bullet\circ} = (2\beta - \mu)/d^2$, where $\mu, V \in \mathbb{R}$ are the new independent degrees of freedom. Substitution into (26) yields (29). \square

Theorem 3 shows that (at fixed β) there are still two degrees of freedom, μ and V , in Eve's manipulation of the EPR pair. This differs from standard qubit-wise QKD, where the bit error probability completely fixes Eve's ancilla state.

8 Purification

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

8.1 The purified state and its properties

We introduce the following notation,

$$|\alpha_j\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{d}} \sum_t e^{i\frac{2\pi}{d} jt} |tt\rangle, \quad j \in \{0, \dots, d-1\} \quad (31)$$

$$|D_{tt'}^\pm\rangle \stackrel{\text{def}}{=} \frac{|tt'\rangle \pm |t't\rangle}{\sqrt{2}} \quad t < t'. \quad (32)$$

Lemma 3 *The $\bar{\rho}^{\text{AB}}$ given in (29) has the following orthonormal eigensystem,*

$$\begin{aligned} |\alpha_0\rangle & \text{ with eigenvalue } \lambda_0 \stackrel{\text{def}}{=} \frac{2\beta - \mu}{d^2} + \frac{\mu + V}{d} + 1 - 2\beta - V \\ |\alpha_j\rangle \quad j \in \{1, \dots, d-1\} & \text{ with eigenvalue } \lambda_1 \stackrel{\text{def}}{=} \frac{2\beta - \mu}{d^2} + \frac{\mu + V}{d}. \\ |D_{tt'}^\pm\rangle \quad (t < t') & \text{ with eigenvalue } \lambda_\pm \stackrel{\text{def}}{=} \frac{2\beta - \mu}{d^2} \pm \frac{V}{d} \end{aligned} \quad (33)$$

Proof: The term proportional to $\mathbb{1}$ in (29) 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 $|D_{tt'}^\pm\rangle$. We have $\langle\alpha_0|D_{tt'}^\pm\rangle = 0$ and $\langle uu|D_{tt'}^\pm\rangle = 0$. Hence the $(1 - 2\beta - V)$ -term and the μ -term in (29) yield zero when acting on $|D_{tt'}^\pm\rangle$. Furthermore $\sum_{uu'}|uu'\rangle\langle u'u|D_{tt'}^+\rangle = \sum_{uu'}|uu'\rangle\frac{\delta_{ut}\delta_{u't'}+\delta_{u't}\delta_{u't'}}{\sqrt{2}} = |D_{tt'}^+\rangle$. Similarly, $\sum_{uu'}|uu'\rangle\langle u'u|D_{tt'}^-\rangle = \sum_{uu'}|uu'\rangle\frac{\delta_{ut}\delta_{u't'}-\delta_{u't}\delta_{u't'}}{\sqrt{2}}\text{sgn}(u-u') = -|D_{tt'}^-\rangle$. \square

In diagonalised form the $\bar{\rho}^{\text{AB}}$ is given by

$$\bar{\rho}^{\text{AB}} = \lambda_0|\alpha_0\rangle\langle\alpha_0| + \lambda_1\sum_{j=1}^{d-1}|\alpha_j\rangle\langle\alpha_j| + \lambda_+\sum_{tt':t<t'}|D_{tt'}^+\rangle\langle D_{tt'}^+| + \lambda_-\sum_{tt':t<t'}|D_{tt'}^-\rangle\langle D_{tt'}^-|. \quad (34)$$

The purification is

$$\begin{aligned} |\Psi^{\text{ABE}}\rangle &= \sqrt{\lambda_0}|\alpha_0\rangle \otimes |E_0\rangle + \sqrt{\lambda_1}\sum_{j=1}^{d-1}|\alpha_j\rangle \otimes |E_j\rangle \\ &\quad + \sqrt{\lambda_+}\sum_{tt':t<t'}|D_{tt'}^+\rangle \otimes |E_{tt'}^+\rangle + \sqrt{\lambda_-}\sum_{tt':t<t'}|D_{tt'}^-\rangle \otimes |E_{tt'}^-\rangle. \end{aligned} \quad (35)$$

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

8.2 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 $k \in \{0, \dots, d-1\}$, $s \in \{0, 1\}$, Eve's state is given by*

$$\sigma_{as}^{rk} = \text{tr}_{\text{AB}}\left[|\Psi^{\text{ABE}}\rangle\langle\Psi^{\text{ABE}}|\frac{Q_a \otimes M_{ks}^{(r)} \otimes \mathbb{1}}{P_{aks|r}}\right]. \quad (36)$$

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 $(Q_a \otimes M_{ks}^{(r)} \otimes \mathbb{1})|\Psi^{\text{ABE}}\rangle$. It is easily verified that the normalisation in (36) is correct: taking the trace in E-space yields $\text{tr}_{\text{AB}}\text{tr}_{\text{E}}[|\Psi^{\text{ABE}}\rangle\langle\Psi^{\text{ABE}}|Q_a \otimes M_{ks}^{(r)} \otimes \mathbb{1}] = \text{tr}_{\text{AB}}\bar{\rho}^{\text{AB}}Q_a \otimes M_{ks}^{(r)} = P_{aks|r}$. \square

Lemma 5 *It holds that*

$$\frac{d}{2d}\sum_{\substack{a_0 \cdots a_{d-1} \\ \text{without } a_k, a_{k+r}}}|\mu_a\rangle\langle\mu_a| = \frac{1}{4}\mathbb{1} + \frac{1}{4}(-1)^{a_k+a_{k+r}}\left(|k\rangle\langle k+r| + |k+r\rangle\langle k|\right) \quad (37)$$

$$= M_{k, a_k \oplus a_{k+r}}^{(r)} + \frac{1}{4}\sum_{t: t \neq k, k+r}|t\rangle\langle t|. \quad (38)$$

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} . \square

Lemma 6 *It holds that*

$$\mathbb{E}_{a: a_k \oplus a_{k+r} = s'}|\mu_a\rangle\langle\mu_a| = \frac{1}{d} + (-1)^{s'}\frac{|k\rangle\langle k+r| + |k+r\rangle\langle k|}{d}. \quad (39)$$

Proof: We have $\mathbb{E}_{a:a_k \oplus a_{k+r}=s'} |\mu_a\rangle\langle\mu_a| = 2^{-(d-1)} \sum_{a_k} \sum_{a_{k+r}} \delta_{a_k \oplus a_{k+r}, s'}$.
 $\sum_{a \text{ without } a_k, a_{k+r}} |\mu_a\rangle\langle\mu_a|$. For the rightmost summation we use Lemma 5. Performing the \sum_{a_k}
and $\sum_{a_{k+r}}$ summations yields (39). \square

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

$$\sigma_{s'}^{rk} = \mathbb{E}_{s,a:a_k \oplus a_{k+r}=s'} [\sigma_{as}^{rk}]. \quad (40)$$

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 $\sigma_{s'}^{rk}$. We define, for $t, t' \in \{0, \dots, d-1\}$, non-normalised vectors $|w_{tt'}\rangle$ in Eve's Hilbert space as

$$|w_{tt'}\rangle \stackrel{\text{def}}{=} \langle tt' | \Psi^{\text{ABE}} \rangle. \quad (41)$$

Furthermore we define angles α and φ as

$$\cos 2\alpha \stackrel{\text{def}}{=} \frac{\langle w_{kk} | w_{k+r, k+r} \rangle}{\langle w_{kk} | w_{kk} \rangle}, \quad \cos 2\varphi \stackrel{\text{def}}{=} \frac{\langle w_{k, k+r} | w_{k+r, k} \rangle}{\langle w_{k, k+r} | w_{k, k+r} \rangle} \quad (42)$$

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

$$\frac{|w_{kk}\rangle}{\sqrt{\langle w_{kk} | w_{kk} \rangle}} = \cos \alpha |A\rangle + \sin \alpha |B\rangle \quad (43)$$

$$\frac{|w_{k+r, k+r}\rangle}{\sqrt{\langle w_{k+r, k+r} | w_{k+r, k+r} \rangle}} = \cos \alpha |A\rangle - \sin \alpha |B\rangle \quad (44)$$

$$\frac{|w_{k, k+r}\rangle}{\sqrt{\langle w_{k, k+r} | w_{k, k+r} \rangle}} = \cos \varphi |C\rangle + \sin \varphi |D\rangle \quad (45)$$

$$\frac{|w_{k+r, k}\rangle}{\sqrt{\langle w_{k+r, k} | w_{k+r, k} \rangle}} = \cos \varphi |C\rangle - \sin \varphi |D\rangle. \quad (46)$$

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 $\sigma_{s'}^{rk}$ are given by*

$$\xi_0 \stackrel{\text{def}}{=} \frac{d}{2} \cdot \frac{\lambda_+ + \lambda_-}{2} \quad (47)$$

$$\xi_1 \stackrel{\text{def}}{=} \frac{d}{2} (\lambda_1 + \lambda_-) = \beta - \frac{d}{2} (\frac{d}{2} - 1) (\lambda_+ + \lambda_-) \quad (48)$$

$$\xi_2 \stackrel{\text{def}}{=} \frac{d}{2} (\lambda_1 + 2 \frac{\lambda_0 - \lambda_1}{d} + \lambda_+) = 1 - \beta - \frac{d}{2} (\frac{d}{2} - 1) (\lambda_+ + \lambda_-) \quad (49)$$

and the diagonal representation of $\sigma_{s'}^{rk}$ is

$$\begin{aligned} \sigma_{s'}^{rk} = & \xi_0 \sum_{\substack{t \in \{0, \dots, d-1\} \\ t \neq k, t \neq k+r}} \left(\frac{|w_{tk}\rangle\langle w_{tk}|}{\langle w_{tk} | w_{tk} \rangle} + \frac{|w_{t, k+r}\rangle\langle w_{t, k+r}|}{\langle w_{t, k+r} | w_{t, k+r} \rangle} \right) \\ & + \xi_2 \frac{[\sqrt{\xi_2 - \frac{d}{2} \lambda_+} |A\rangle + (-1)^{s'} \sqrt{\frac{d}{2} \lambda_+} |C\rangle][\dots]^\dagger}{\xi_2} \\ & + \xi_1 \frac{[\sqrt{\xi_1 - \frac{d}{2} \lambda_-} |B\rangle - (-1)^{s'} \sqrt{\frac{d}{2} \lambda_-} |D\rangle][\dots]^\dagger}{\xi_1} \end{aligned} \quad (50)$$

Proof: We have

$$\begin{aligned}\sigma_s^{rk} &= \text{tr}_{\text{AB}} |\Psi^{\text{ABE}}\rangle \langle \Psi^{\text{ABE}} | \mathbb{E}_{a:a_k \oplus a_{k+r}=s'} Q_a \otimes \mathbb{E}_{s|s'} \frac{M_{ks}^{(r)}}{P_{aks|r}} \otimes \mathbb{1} \\ &= d2^d \text{tr}_{\text{AB}} |\Psi^{\text{ABE}}\rangle \langle \Psi^{\text{ABE}} | [\mathbb{E}_{a:a_k \oplus a_{k+r}=s'} Q_a] \otimes \left[\sum_s M_{ks}^{(r)} \right] \otimes \mathbb{1}.\end{aligned}\quad (51)$$

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

$$|w_{tt}\rangle = \sqrt{\lambda_0/d} |E_0\rangle + \sqrt{\lambda_1/d} \sum_{j=1}^{d-1} (e^{i\frac{2\pi}{d}})^{jt} |E_j\rangle \quad (52)$$

$$\langle w_{tt} | w_{tt} \rangle = \lambda_1 + \frac{\lambda_0 - \lambda_1}{d}, \quad (53)$$

and for $t \neq t'$ we have

$$|w_{tt'}\rangle = \sqrt{\lambda_+/2} |E_{(tt')}^+\rangle + \text{sgn}(t' - t) \sqrt{\lambda_-/2} |E_{(tt')}^-\rangle \quad (54)$$

$$\langle w_{tt'} | w_{tt'} \rangle = (\lambda_+ + \lambda_-)/2. \quad (55)$$

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

$$\langle w_{tt} | w_{tt'} \rangle = 0, \quad \langle w_{tt} | w_{t't} \rangle = 0 \quad (56)$$

$$\langle w_{tt} | w_{t't'} \rangle = \frac{\lambda_0 - \lambda_1}{d}, \quad \langle w_{t't'} | w_{t't} \rangle = \frac{\lambda_+ - \lambda_-}{2}. \quad (57)$$

We get

$$\cos 2\alpha = 1 - \frac{d\lambda_1}{\lambda_0 + (d-1)\lambda_1}, \quad \cos 2\varphi = 1 - \frac{2\lambda_-}{\lambda_+ + \lambda_-} \quad (58)$$

After some tedious algebra the result (50) follows. \square

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*

$$\begin{aligned}\frac{\sigma_0^{rk} + \sigma_1^{rk}}{2} &= \sum_{\substack{t \in \{0, \dots, d-1\} \\ t \neq k, t \neq k+r}} \xi_0 \cdot \left(\frac{|w_{tk}\rangle\langle w_{tk}|}{\langle w_{tk} | w_{tk} \rangle} + \frac{|w_{t,k+r}\rangle\langle w_{t,k+r}|}{\langle w_{t,k+r} | w_{t,k+r} \rangle} \right) \\ &\quad + (\xi_2 - \frac{d}{2}\lambda_+) |A\rangle\langle A| + \frac{d}{2}\lambda_+ |C\rangle\langle C| + (\xi_1 - \frac{d}{2}\lambda_-) |B\rangle\langle B| + \frac{d}{2}\lambda_- |D\rangle\langle D|.\end{aligned}$$

Proof: Follows directly from Theorem 4 by discarding the terms in (50) that contain $(-1)^{s'}$ (the AC and BD crossterms). \square

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

$$\begin{aligned}\frac{\sigma_0^{rk} - \sigma_1^{rk}}{2} &= \frac{1}{2} \sqrt{d\lambda_+} \sqrt{d\lambda_- + 2(1-\beta) - \frac{d^2}{2}(\lambda_+ + \lambda_-)} (|A\rangle\langle C| + |C\rangle\langle A|) \\ &\quad - \frac{1}{2} \sqrt{d\lambda_-} \sqrt{d\lambda_+ + 2\beta - \frac{d^2}{2}(\lambda_+ + \lambda_-)} (|B\rangle\langle D| + |D\rangle\langle B|).\end{aligned}\quad (59)$$

Proof: Using Theorem 4, we see everything except the AC and BD crossterms cancel from (50). \square

9 Statistical distance; proof of Theorem 1

Now that we have described Eve's most general allowed state, and how it is connected to Alice's secret bit s' , it is finally 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 $\mathbf{r} = (r_1, \dots, r_n)$, $\mathbf{k} = (k_1, \dots, k_n)$. Let $x = (s'_1, \dots, s'_n)$. Let $z \in \{0, 1\}^\ell$ be the QKD key obtained by applying privacy amplification to x , i.e. $z = \text{Ext}(x, u)$, where Ext is a universal hash function (UHF) and $u \in \mathcal{U}$ is public randomness. We write $\mathbb{E}_u[\dots] = \frac{1}{|\mathcal{U}|} \sum_u (\dots)$ and $\mathbb{E}_x[\dots] = 2^{-n} \sum_{x \in \{0, 1\}^n} (\dots)$. At given (\mathbf{r}, \mathbf{k}) the quantum-classical state describing Z, U , and Eve's system 'E' is given by

$$\rho^{\text{ZUE}}(\mathbf{r}, \mathbf{k}) = \sum_z \mathbb{E}_u |zu\rangle \langle zu| \otimes \mathbb{E}_x \delta_{z, \text{Ext}(u, x)} \bigotimes_{i=1}^n \sigma_{x_i}^{r_i k_i}. \quad (60)$$

The state of the 'Z' and 'UE' subsystems is

$$\rho^Z(\mathbf{r}, \mathbf{k}) = \text{tr}_{\text{UE}} \rho^{\text{ZUE}}(\mathbf{r}, \mathbf{k}) = 2^{-\ell} \sum_z |z\rangle \langle z| \quad (61)$$

$$\rho^{\text{UE}}(\mathbf{r}, \mathbf{k}) = \text{tr}_Z \rho^{\text{ZUE}}(\mathbf{r}, \mathbf{k}) = \mathbb{E}_u |u\rangle \langle u| \otimes \omega_{\text{av}}(\mathbf{r}, \mathbf{k}) \quad (62)$$

$$\omega_{\text{av}}(\mathbf{r}, \mathbf{k}) \stackrel{\text{def}}{=} \bigotimes_{i=1}^n \frac{\sigma_0^{r_i k_i} + \sigma_1^{r_i k_i}}{2}. \quad (63)$$

Note that ω_{av} does not depend on u . For notational brevity we stop explicitly mentioning the \mathbf{r}, \mathbf{k} dependence from this point on. From (60)–(62) we get

$$\rho^{\text{ZUE}} - \rho^Z \otimes \rho^{\text{UE}} = 2^{-\ell} \sum_z \mathbb{E}_u |zu\rangle \langle zu| \otimes \left\{ 2^\ell \mathbb{E}_x \delta_{z, \text{Ext}(u, x)} \bigotimes_{i=1}^n \sigma_{x_i}^{r_i k_i} - \omega_{\text{av}} \right\} \quad (64)$$

$$\stackrel{\text{def}}{=} 2^{-\ell} \sum_z \mathbb{E}_u |zu\rangle \langle zu| \otimes \Delta_{zu} \quad (65)$$

Because of the zu block structure we have

$$\|\rho^{\text{ZUE}} - \rho^Z \otimes \rho^{\text{UE}}\|_1 = 2^{-\ell} \sum_z \mathbb{E}_u \|\Delta_{zu}\|_1. \quad (66)$$

Lemma 7 *It holds that*

$$2^{-\ell} \sum_z \mathbb{E}_u \|\Delta_{zu}\|_1 \leq \text{tr} \sqrt{2^{-\ell} \sum_z \mathbb{E}_u \Delta_{zu}^2}. \quad (67)$$

Proof: $2^{-\ell} \sum_z \mathbb{E}_u \|\Delta_{zu}\|_1 = 2^{-\ell} \sum_z \mathbb{E}_u \text{tr} \sqrt{\Delta_{zu}^2} = \text{tr} 2^{-\ell} \sum_z \mathbb{E}_u \sqrt{\Delta_{zu}^2}$. We apply Jensen's inequality for operator-concave functions. \square

Lemma 8 *It holds that*

$$2^{-\ell} \sum_z \mathbb{E}_u \Delta_{zu}^2 = \frac{2^\ell - 1}{2^n} \bigotimes_{i=1}^n \frac{(\sigma_0^{r_i k_i})^2 + (\sigma_1^{r_i k_i})^2}{2}. \quad (68)$$

Proof: From the definition of Δ_{zu} and ω_{av} we get

$$\begin{aligned} 2^{-\ell} \sum_z \mathbb{E}_u \Delta_{zu}^2 &= \frac{2^\ell}{2^{2n}} \sum_{xyz} \mathbb{E}_u \delta_{z, \text{Ext}(x, u)} \delta_{z, \text{Ext}(y, u)} \bigotimes_{i=1}^n \sigma_{x_i}^{r_i k_i} \sigma_{y_i}^{r_i k_i} + \omega_{\text{av}}^2 \\ &\quad - \omega_{\text{av}} \frac{1}{2^n} \sum_{xz} \mathbb{E}_u \delta_{z, \text{Ext}(x, u)} \bigotimes_{i=1}^n \sigma_{x_i}^{r_i k_i} - \left(\frac{1}{2^n} \sum_{xz} \mathbb{E}_u \delta_{z, \text{Ext}(x, u)} \bigotimes_{i=1}^n \sigma_{x_i}^{r_i k_i} \right) \omega_{\text{av}}. \end{aligned} \quad (69)$$

We split the \sum_{xy} sum into a sum with $y = x$ and a sum with $y \neq x$. Then we use $\sum_z \delta_{z, \text{Ext}(x, u)} = 1$ and $\sum_z \mathbb{E}_u \delta_{z, \text{Ext}(x, u)} \delta_{z, \text{Ext}(y, u)} = 2^{-\ell}$ 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 $2^{-n} \sum_x \bigotimes_i \sigma_{x_i}^{r_i k_i} = \omega_{\text{av}}$, most of the terms cancel and (68) is what remains. \square

Lemma 9 *It holds that*

$$\begin{aligned} \frac{(\sigma_0^{rk})^2 + (\sigma_1^{rk})^2}{2} &= \sum_{\substack{t \in \{0, \dots, d-1\} \\ t \neq k, t \neq \ell}} \xi_0^2 \left(\frac{|w_{tk}\rangle\langle w_{tk}|}{\langle w_{tk}|w_{tk}\rangle} + \frac{|w_{t\ell}\rangle\langle w_{t\ell}|}{\langle w_{t\ell}|w_{t\ell}\rangle} \right) + \xi_1 (\xi_1 - \frac{d}{2}\lambda_-) |B\rangle\langle B| \\ &+ \xi_1 \frac{d}{2}\lambda_- |D\rangle\langle D| + \xi_2 (\xi_2 - \frac{d}{2}\lambda_+) |A\rangle\langle A| + \xi_2 \frac{d}{2}\lambda_+ |C\rangle\langle C| \end{aligned}$$

with ξ_0, ξ_1, ξ_2 as defined in Theorem 4.

Proof: Follows directly from Theorem 4. \square

Lemma 10 *The statistical distance between the real and decoupled state can be bounded as*

$$\frac{1}{2} \|\rho^{\text{ZUE}} - \rho^{\text{Z}} \otimes \rho^{\text{UE}}\|_1 < \frac{1}{2} \sqrt{2^{\ell-n} T^n} \quad (70)$$

$$T \stackrel{\text{def}}{=} 2(d-2)\xi_0 + \sqrt{\xi_2(\xi_2 - \frac{d}{2}\lambda_+)} + \sqrt{\xi_2 \frac{d}{2}\lambda_+} + \sqrt{\xi_1(\xi_1 - \frac{d}{2}\lambda_-)} + \sqrt{\xi_1 \frac{d}{2}\lambda_-}. \quad (71)$$

Proof:

Substitution of Lemma 8 into Lemma 7 gives $\|\rho^{\text{ZUE}} - \rho^{\text{Z}} \otimes \rho^{\text{UE}}\|_1 \leq \sqrt{\frac{2^\ell - 1}{2^n}} \prod_{i=1}^n \text{tr} \sqrt{\frac{(\sigma_0^{r_i k_i})^2 + (\sigma_1^{r_i k_i})^2}{2}}$.

The trace does not depend on the actual value of r_i and k_i . We define $T = \text{tr} \sqrt{(\sigma_0^{rk})^2 + (\sigma_1^{rk})^2} / \sqrt{2}$ for arbitrary r, k . From Lemma 9 we obtain (71). Finally we use $2^\ell - 1 < 2^\ell$. \square

Remark. We are able to derive a tight bound because the expression $\text{tr} \sqrt{\sigma_0^2 + \sigma_1^2}$ is easy to compute without applying any inequalities.

Since Eve is still free to choose the parameters μ and V (or, equivalently, λ_+ and λ_-) she can choose them such that the trace distance is maximized.

Theorem 5 *Eve's choice that maximizes $\|\rho^{\text{ZUE}} - \rho^{\text{Z}} \otimes \rho^{\text{UE}}\|_1$ is given by*

$$\beta \leq \beta_* : \quad T = 2\beta + \sqrt{1-2\beta} \left[\sqrt{1-2\beta} \frac{d-1}{d-2} + \frac{\sqrt{2\beta}}{\sqrt{d-2}} \right] \quad (72)$$

$$\text{at } \lambda_- = 0, \quad \lambda_+ = \frac{4\beta}{d(d-2)} \quad (73)$$

$$\beta \geq \beta_* : \quad T = 2\beta_* + \sqrt{1-2\beta_*} \left[\sqrt{1-2\beta_*} \frac{d-1}{d-2} + \frac{\sqrt{2\beta_*}}{\sqrt{d-2}} \right] \quad (74)$$

$$\text{at } \lambda_- = \frac{4\beta_*(\beta - \beta_*)}{d(d-2)(1-2\beta_*)}, \quad \lambda_+ = \frac{4\beta_*(1-\beta-\beta_*)}{d(d-2)(1-2\beta_*)}. \quad (75)$$

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

$$\beta_* = \frac{x_d/2}{1+x_d}, \quad (76)$$

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

$$\left(1 - \frac{x}{d-2}\right)^{\frac{1}{2}} + \frac{d-1}{d-2} \left(1 - \frac{x}{d-2}\right)^{-\frac{1}{2}} + \frac{1}{\sqrt{d-2}} \left(\sqrt{x} - \frac{1}{\sqrt{x}}\right) - 2 = 0. \quad (77)$$

Proof: We start from (71). At $\beta = \frac{1}{2}$ 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

$$T_{\max}^{\beta=\frac{1}{2}} = \zeta(q, d) \stackrel{\text{def}}{=} q + \sqrt{1-q} \left(\sqrt{1 - \frac{d-1}{d-2}q} + \frac{\sqrt{q}}{\sqrt{d-2}} \right). \quad (78)$$

On the other hand, we note that substitution of (73) into (71) yields (72), which is precisely of the form $\zeta(q, d)$ if we identify $2\beta \equiv q$. Hence, at some $\beta < \frac{1}{2}$ it is already possible to achieve $T = T_{\max}^{\beta=1/2}$, i.e. we have saturation. We note that substitution of (75) into (71) yields (74). 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 (77) by setting $x = 2\beta/(1 - 2\beta)$.⁶ \square

The upper bound on the amount of information that Eve has about S' is $2 \log T$. This is a concave function of β (see Fig. 3). Hence there is no advantage for Eve to cause different error rates in different rounds. For Eve it is optimal to cause error rate β in every round.

This concludes the proof of Theorem 1.

The optimal λ_+, λ_- are plotted in Fig. 5 (Appendix B). The expression $2 \log T$ is plotted in Fig. 3.

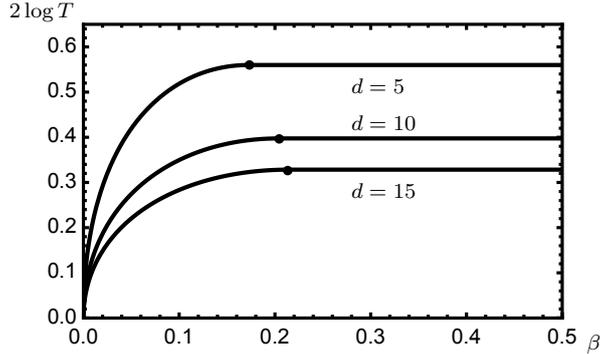


Fig. 3 Leakage $2 \log T$ as a function of the bit error rate for $d = 5$, $d = 10$ and $d = 15$. (This does not include the post-selection term). A dot indicates the saturation point β_* .

Lemma 11 *The large- d asymptotics of the saturation value β_* is given by*

$$\beta_* = \frac{1}{4} - \frac{1}{8\sqrt{d-2}} - \mathcal{O}\left(\frac{1}{(d-2)^{3/2}}\right), \quad (79)$$

which yields

$$T = 1 + \frac{1}{2\sqrt{d-2}} - \mathcal{O}\left(\frac{1}{d-2}\right) \quad (80)$$

$$\|\rho^{\text{ZUE}} - \rho^{\text{Z}} \otimes \rho^{\text{UE}}\|_1 \leq 2^{-\frac{1}{2}n[1 - \frac{1}{\sqrt{d-2}\ln 2} + \mathcal{O}(\frac{1}{d-2}) - \frac{\epsilon}{n}]}. \quad (81)$$

Proof: We set $x_d = 1 - 1/\sqrt{d-2} + a/(d-2)$, where a is supposedly of order 1, and substitute this into (77). This yields $a = \frac{1}{2} + \mathcal{O}(1/\sqrt{d-2})$, which is indeed of order 1. Substitution of x_d into (76) gives (79), and substitution of β_* into (74) gives (80). Finally, substitution of (80) into Lemma 10 yields (81). \square

10 Von Neumann entropy; Proof of 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.⁷ We denote

⁶ After some rewriting it can be seen that (77) 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.

⁷ By applying Jensen's inequality once more to Lemma 7, 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.

the leakage from Alice to Eve, in terms of von Neumann entropy, as I_{AE} . It is given by

$$\begin{aligned}
I_{\text{AE}} &= S(\sigma_{S'}^{RK}|RK) - S(\sigma_{S'}^{RK}|RKS') \\
&= \mathbb{E}_{rk}[S(\sigma_{S'}^{rk}) - S(\sigma_{S'}^{rk}|S')] \\
&= \mathbb{E}_{rk} \left[S\left(\frac{\sigma_0^{rk} + \sigma_1^{rk}}{2}\right) - \frac{S(\sigma_0^{rk}) + S(\sigma_1^{rk})}{2} \right] \\
&= S\left(\frac{\sigma_0^{rk} + \sigma_1^{rk}}{2}\right) - \frac{S(\sigma_0^{rk}) + S(\sigma_1^{rk})}{2} \quad r, k \text{ arbitrary.}
\end{aligned} \tag{82}$$

In the last line we used that the eigenvalues of σ_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*

$$\beta \leq \beta_0 : \quad I_{\text{AE}} = (1 - 2\beta)h\left(\frac{1}{d-2} \cdot \frac{2\beta}{1-2\beta}\right) \tag{83}$$

$$\text{at } \lambda_- = 0, \quad \lambda_+ = \frac{4\beta}{d(d-2)} \tag{84}$$

$$\beta \geq \beta_0 : \quad I_{\text{AE}} = (1 - 2\beta_0)h\left(\frac{1}{d-2} \cdot \frac{2\beta_0}{1-2\beta_0}\right) \tag{85}$$

$$\text{at } \lambda_- = \frac{4\beta_0(\beta - \beta_0)}{d(d-2)(1-2\beta_0)}, \quad \lambda_+ = \frac{4\beta_0(1 - \beta - \beta_0)}{d(d-2)(1-2\beta_0)}. \tag{86}$$

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

$$\beta_0 = \frac{1}{2} \left[1 + \frac{1}{(d-2)(1-y_d)} \right]^{-1} \tag{87}$$

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

Proof: We start from (82). 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 (82) comes entirely from the $|A\rangle, |B\rangle, |C\rangle, |D\rangle$ subspace,

$$\begin{aligned}
I_{\text{AE}} &= \xi_1 \log \xi_1 + \xi_2 \log \xi_2 - (\xi_2 - \frac{d}{2}\lambda_+) \log(\xi_2 - \frac{d}{2}\lambda_+) - \frac{d}{2}\lambda_+ \log(\frac{d}{2}\lambda_+) \\
&\quad - (\xi_1 - \frac{d}{2}\lambda_-) \log(\xi_1 - \frac{d}{2}\lambda_-) - \frac{d}{2}\lambda_- \log(\frac{d}{2}\lambda_-) \\
&= \xi_1 h\left(\frac{d}{2} \cdot \frac{\lambda_-}{\xi_1}\right) + \xi_2 h\left(\frac{d}{2} \cdot \frac{\lambda_+}{\xi_2}\right).
\end{aligned} \tag{88}$$

We note that (88) is invariant under the transformation $(\beta \rightarrow 1 - \beta; \lambda_+ \leftrightarrow \lambda_-)$. At $\beta = 1/2$ we must hence have $\lambda_+ = \lambda_- = \lambda$.

$$I_{\text{AE}}^{\beta=\frac{1}{2}} = g(d, \lambda) \stackrel{\text{def}}{=} [1 - d(d-2)\lambda] \cdot h\left(\frac{d\lambda}{1 - d(d-2)\lambda}\right). \tag{89}$$

At $\beta = \frac{1}{2}$, the largest leakage that Eve can cause is $\max_{\lambda} g(d, \lambda) = g(d, \lambda_*)$.⁸ Next we note that substitution of (86) into (88) yields (85); this has the same form as $g(d, \lambda)$ (89) 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

⁸ $\frac{\partial^2 g(d, \lambda)}{\partial \lambda^2} = -\frac{d}{\lambda} [1 - d(d-2)\lambda]^{-1} [1 - d(d-1)\lambda]^{-1}$, 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_* .

constant at $I_{\text{AE}}^{\max}(\beta) = g(d, \lambda_*)$ on the interval $\beta \in [\beta_0, \frac{1}{2}]$. The value $g(d, \lambda_*)$ precisely equals (85). Next we determine the value of β_0 . Demanding $\partial g(d, \lambda)/\partial \lambda = 0$ at $\lambda = \lambda_*$ yields

$$\log \frac{[1 - d(d-1)\lambda_*]^{d-1}}{[1 - d(d-2)\lambda_*]^{d-2} \lambda_* d} = 0. \quad (90)$$

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 (87), 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_{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 (84) into (88) yields (83). \square

This concludes the proof of theorem 2.

Note that the leakage I_{AE} is a concave function of β . Hence it is optimal for Eve to cause error rate β in every round.

Remark. From $y > 0$ and (87) it follows that $\beta_0 < \frac{1}{2} \cdot \frac{d-2}{d-1}$.

Fig. 4 shows the von Neumann leakage for three values of d . The optimal λ_+, λ_- are plotted in Fig. 5 (Appendix B).

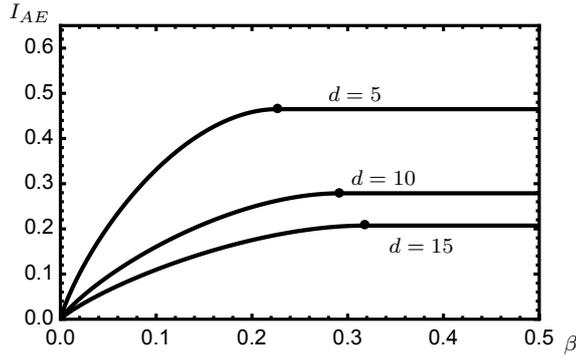


Fig. 4 Leakage I_{AE} in terms of von Neumann entropy (Theorem 2) as a function of the bit error rate, for $d = 5$, $d = 10$ and $d = 15$. A dot indicates the saturation point β_0 .

Lemma 12 The large- d asymptotics of the I_{AE} is given by

$$\beta \leq \beta_0 : I_{\text{AE}} = \frac{2\beta}{d-2} \log \frac{(d-2)(1-2\beta)e}{2\beta} + \mathcal{O}(d^{-2}) \quad (91)$$

$$\beta \geq \beta_0 : I_{\text{AE}} = \frac{\log d}{d} + \mathcal{O}\left(\frac{\log \log d}{d}\right). \quad (92)$$

Proof: The result for $\beta < \beta_0$ follows by doing a series expansion of (83) 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 $\alpha \cdot \left\{ \left(1 - \frac{\ln[(d-1)/\alpha]}{d-1}\right)^{d-1} \frac{d-1}{\alpha} \right\} = \ln \frac{d-1}{\alpha}$. Using the fact that $\lim_{n \rightarrow \infty} (1 - x/n)^n = e^{-x}$ we see that the expression $\{\cdot\}$ 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 (87) gives $1 - 2\beta_0 = \frac{1}{\ln d} + \mathcal{O}\left(\frac{\ln \ln d}{(\ln d)^2}\right)$. Substituting this result for $1 - 2\beta_0$ into (85) finally yields (92). \square

11 Discussion

We remark on the optimal attack. The $\bar{\rho}^{\text{AB}}$ 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. 96 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. Eve's 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 [17].

Acknowledgements

We thank Serge Fehr for useful discussions. We thank the anonymous reviewers of Eurocrypt 2018 for their comments about symmetrisation. Part of this research was funded by NWO (CHIST-ERA project ID_IOT).

A Details of Eve's unitary operation

In Theorem 7 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 7 *The operation that maps the pure EPR state to $|\Psi^{\text{ABE}}\rangle$ (35) 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 (35) can be written as the following mapping,

$$U\left(|t\rangle_{\text{B}} \otimes |E_0\rangle_{\text{E}}\right) = |\Omega_t\rangle, \quad (93)$$

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

$$|\Omega_t\rangle \stackrel{\text{def}}{=} \sqrt{\lambda_0}|t\rangle|E_0\rangle + \sqrt{\lambda_1}|t\rangle \sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}jt}|E_j\rangle + \sqrt{\frac{d\lambda_+}{2}} \sum_{t':t' \neq t} |t'\rangle|E_{(tt')}^+\rangle + \sqrt{\frac{d\lambda_-}{2}} \sum_{t':t' \neq t} |t'\rangle|E_{(tt')}^-\rangle \text{sgn}(t' - t). \quad (94)$$

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. (93,94) 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 $|t\rangle_{\text{B}} \otimes |\varepsilon\rangle_{\text{E}}$, 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 $|\Omega_t\rangle$, $t \in \{0, \dots, d-1\}$. The dimension of the BE space is d^2 and allows us to make such a choice of $d(d^2 - 1)$ vectors. \square

Theorem 8 *Let Alice send the state $|\mu_a\rangle$ to Bob. Let Eve apply the unitary operation U (specified in the proof of Theorem 7) to this state and her ancilla. The result can be written as*

$$U\left(|\mu_a\rangle \otimes |E_0\rangle\right) = \frac{1}{\sqrt{d}} \sum_{t=0}^{d-1} (-1)^{a_t} |t\rangle \otimes |A_t^a\rangle, \quad (95)$$

$$|A_t^a\rangle \stackrel{\text{def}}{=} \sqrt{\lambda_0}|E_0\rangle + \sqrt{\lambda_1} \sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}jt}|E_j\rangle + \sqrt{\frac{d}{2}} \sum_{t':t' \neq t} (-1)^{a_t + a_{t'}} \left[\sqrt{\lambda_+} |E_{(tt')}^+\rangle + \sqrt{\lambda_-} \text{sgn}(t' - t) |E_{(tt')}^-\rangle \right]. \quad (96)$$

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 $U(|\mu_a\rangle|E_0\rangle) = (1/\sqrt{d})\sum_t(-1)^{at}|\Omega_t\rangle$ and we substitute (94). Re-labeling of summation variables yields (95,96). 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 $\tilde{\rho}^{AB}$. For $\tau \neq t$ the inner product $\langle A_\tau^a|A_t^a\rangle$ yields

$$\lambda_0 + \lambda_1 \sum_{j=1}^{d-1} e^{i\frac{2\pi}{d}j(t-\tau)} + \frac{d}{2} \sum_{t' \neq t} \sum_{\tau' \neq \tau} (-1)^{a_t+a_{t'}+a_\tau+a_{\tau'}} \delta_{t'\tau} \delta_{\tau't} [\lambda_+ + \lambda_- \text{sgn}(t'-t) \text{sgn}(\tau'-\tau)]. \quad (97)$$

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 (97) set the phase $(-1)^{\dots}$ to 1 and $\text{sgn}(t'-t)\text{sgn}(\tau'-\tau) = \text{sgn}(\tau-t)\text{sgn}(t-\tau) = -1$. Finally we use $\lambda_0 - \lambda_1 = 1 - 2\beta - V$ and $\lambda_+ - \lambda_- = 2V/d$. \square Theorem 8 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 3 *The pure state (95) in Bob and Eve's space gives rise to the following mixed state ρ_a^B in Bob's subsystem,*

$$\rho_a^B = (1 - 2\beta)|\mu_a\rangle\langle\mu_a| + 2\beta\frac{\mathbb{1}}{d}. \quad (98)$$

Proof: Follows directly from (95) 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 Min-entropy and accessible entropy

By way of supplementary information we present a number of results about simple attacks on individual qudits. This teaches us which kind of qubit-wise measurement is informative for Eve. Furthermore, 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. It is possible to give a composable security proof based on the min-entropy result and post-selection, but this yields a lower rate than Theorem 1.

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 turns out to be smaller than the leakage of Theorems 1 and 2.

B.1 (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 $\rho^{AB} = \mathbb{E}_{x \in \mathcal{X}} |x\rangle\langle x| \otimes \rho_x$ with the $|x\rangle$ forming an orthonormal basis. The min-entropy of the classical RV X given part 'B' of the system is [23]

$$H_{\min}(X|\rho_X) = -\log \max_{\mathcal{M}} \mathbb{E}_{x \in \mathcal{X}} \text{tr}[M_x \rho_x]. \quad (99)$$

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 condition⁹ [24]

$$\forall_{x \in \mathcal{X}} : \Lambda - \rho_x \geq 0, \quad (100)$$

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

$$H_{\min}(X|Y, \rho_X(Y)) = -\log \mathbb{E}_{y \in \mathcal{Y}} \max_{\mathcal{M}} \mathbb{E}_{x \in \mathcal{X}} \text{tr}[M_x \rho_x(y)]. \quad (101)$$

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

Then

$$X \sim (p_0, p_1) : \quad H_{\min}(X|Y, \rho_X(Y)) = -\log \left(\frac{1}{2} + \frac{1}{2} \mathbb{E}_y \text{tr} \left\| p_0 \rho_0(y) - p_1 \rho_1(y) \right\|_1 \right). \quad (102)$$

⁹ Ref. [24] specifies a second condition, namely $\Lambda^\dagger = \Lambda$. However, the hermiticity of Λ already follows from the condition (100).

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

$$H(X|\rho_X) = \min_{\mathcal{M}} H(X|\mathcal{M}(\rho_X)). \quad (103)$$

The ‘accessible information’ is defined as the mutual information $H(X) - H(X|\rho_X)$. In contrast to the min-entropy case, there is no simple test analogous to (100) which tells you whether a local minimum in (103) is a global minimum.

B.2 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 B.1). Eq. (102) 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 β_{sat} ,

$$\beta_{\text{sat}} \stackrel{\text{def}}{=} \frac{1}{4} \cdot \frac{d-2}{d-1}. \quad (104)$$

Again we optimize λ_+ and λ_- .

Lemma 13 For all r, k the choice for λ_+ and λ_- that maximizes the trace distance $\frac{1}{2} \|\sigma_0^{rk} - \sigma_1^{rk}\|_1$ is

$$\lambda_+ = \frac{4\beta}{d(d-2)} \quad \lambda_- = 0 \quad \text{for } \beta < \beta_{\text{sat}} \quad (105)$$

$$\lambda_+ = \frac{4\beta_{\text{sat}}}{d(d-2)} - \frac{2(\beta - \beta_{\text{sat}})}{d^2} \quad \lambda_- = \frac{2(\beta - \beta_{\text{sat}})}{d^2} \quad \text{for } \beta \geq \beta_{\text{sat}}. \quad (106)$$

which gives

$$\frac{1}{2} \|\sigma_0^{rk} - \sigma_1^{rk}\|_1 = \begin{cases} \frac{1}{\sqrt{d-1}} \frac{\sqrt{\beta}}{\beta_{\text{sat}}} \sqrt{2\beta_{\text{sat}} - \beta} & \text{for } \beta < \beta_{\text{sat}} \\ \frac{1}{\sqrt{d-1}} & \text{for } \beta \geq \beta_{\text{sat}}. \end{cases} \quad (107)$$

Proof: From Corollary 2 it is easy to see that

$$\begin{aligned} \frac{1}{2} \|\sigma_0^{rk} - \sigma_1^{rk}\|_1 &= \sqrt{d\lambda_-} \sqrt{d\lambda_+ + 2\beta - \frac{d^2}{2}(\lambda_+ + \lambda_-)} \\ &\quad + \sqrt{d\lambda_+} \sqrt{d\lambda_- + 2(1-\beta) - \frac{d^2}{2}(\lambda_+ + \lambda_-)}. \end{aligned} \quad (108)$$

In Appendix C we derive the λ_+ , λ_- that maximize (108) while keeping all eigenvalues non-negative. \square

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 (106) is shown in a simplified form one can manipulate it to the same form as (75) and (86) with β_{sat} instead of β_* or β_0 .

Fig. 5 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 9 The min-entropy of the bit S' given R, K and the state $\sigma_{S'}^{RK}$ is

$$\beta < \beta_{\text{sat}} : \quad H_{\min}(S'|RK\sigma_{S'}^{RK}) = -\log\left(\frac{1}{2} + \frac{1}{2\sqrt{d-1}} \frac{\sqrt{\beta}}{\beta_{\text{sat}}} \sqrt{2\beta_{\text{sat}} - \beta}\right) \quad (109)$$

$$\beta \geq \beta_{\text{sat}} : \quad H_{\min}(S'|RK\sigma_{S'}^{RK}) = -\log\left(\frac{1}{2} + \frac{1}{2\sqrt{d-1}}\right). \quad (110)$$

Proof: Eq. (102) with X uniform, $X \rightarrow S'$, $Y \rightarrow (R, K)$ becomes

$$\begin{aligned} H_{\min}(S'|RK\sigma_{s'}^{RK}) &= -\log\left(\frac{1}{2} + \frac{1}{2} \mathbb{E}_{rk} \left\| \frac{1}{2} \sigma_0^{rk} - \frac{1}{2} \sigma_1^{rk} \right\|_1\right) \\ &= -\log\left(\frac{1}{2} + \frac{1}{4} \|\sigma_0^{rk} - \sigma_1^{rk}\|_1\right) \quad (r, k \text{ arbitrary}). \end{aligned} \quad (111)$$

In the last step we omitted the expectation over r and k since the trace distance does not depend on r, k . Substitution of (107) into (111) gives the end result. \square

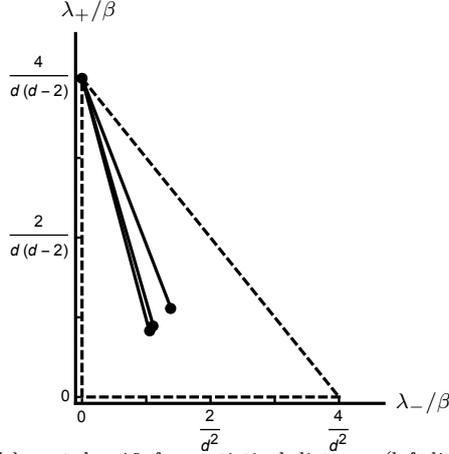


Fig. 5 Optimal choice of λ_+ and λ_- at $d = 10$ for statistical distance (left line), min-entropy (middle line) and von Neumann entropy (right line). The dashed triangle represents the region for which the eigenvalues λ_+ , λ_- and λ_1 are non-negative. The black dots indicate the optimum at $\beta = \frac{1}{2}$ (dots inside the triangle) and $\beta \leq \beta_*, \beta_{\text{sat}}, \beta_0$ (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 $\beta > 2\beta_{\text{sat}}$.

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

$$T_0^{rk} = \frac{1}{2} \left(\mathbb{1} + |A\rangle\langle C| + |C\rangle\langle A| - |B\rangle\langle D| - |D\rangle\langle B| \right) ; \quad T_1^{rk} = \mathbb{1} - T_0^{rk}. \quad (112)$$

Proof: The trace distance in Lemma 13 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 (112) satisfies the test (100). \square

As expected, the min-entropy loss decreases as the dimension of the Hilbert space grows. We see that the entropy loss saturates at $\beta = \beta_{\text{sat}}$; hence RRDPS is secure up to arbitrarily high noise levels. Fig. 6 shows the min-entropy leakage as a function of β .

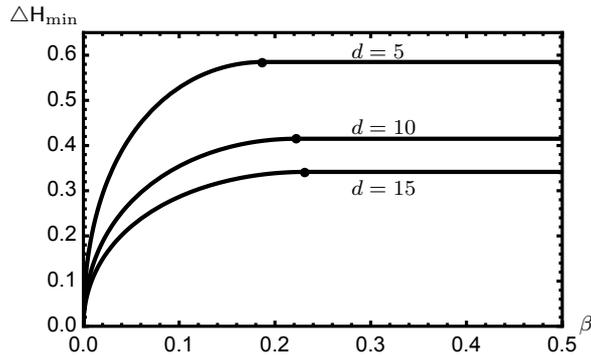


Fig. 6 Min-entropy leakage as a function of the bit error rate for $d = 5$, $d = 10$ and $d = 15$. A dot indicates the saturation point β_{sat} .

B.3 Accessible Shannon information

Lemma 14 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

$$\mathsf{H}(X|\rho_{XY}) = \min_{\text{POVM } \mathcal{M}=(M_m)_{m \in \mathcal{X}}} \mathbb{E}_{x \in \mathcal{X}} \mathsf{H}\left(\{\text{tr } M_m \mathbb{E}_{y|x} \rho_{xy}\}_{m \in \mathcal{X}}\right). \quad (113)$$

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 $\Pr[z|x] = \mathbb{E}_{y|x} \Pr[z|xy] = \mathbb{E}_{y|x} \text{tr } M_z \rho_{xy}$. \square

Corollary 5 *It holds that*

$$\mathsf{H}(S'|RK\sigma_{AS}^{RK}) = \mathbb{E}_{rk} \min_{\mathcal{G}^{rk}=(G_0^{rk}, G_1^{rk})} \mathbb{E}_{s'} h(\text{tr } G_m^{rk} \sigma_{s'}^{rk}), \quad m \in \{0,1\} \text{ arbitrary}. \quad (114)$$

Proof: Application of Lemma 14 yields

$$\begin{aligned} \mathsf{H}(S'|RK\sigma_{AS}^{RK}) &= \mathbb{E}_{rk} \min_{\mathcal{G}^{rk}=(G_0^{rk}, G_1^{rk})} \mathbb{E}_{s'} H(\{\text{tr } G_m^{rk} \mathbb{E}_{as|s'} \sigma_{as}^{rk}\}_{m \in \{0,1\}}) \\ &= \mathbb{E}_{rk} \min_{\mathcal{G}^{rk}=(G_0^{rk}, G_1^{rk})} \mathbb{E}_{s'} H(\{\text{tr } G_m^{rk} \sigma_{s'}^{rk}\}_{m \in \{0,1\}}) \end{aligned} \quad (115)$$

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 $h(1-p) = h(p)$. \square

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 $\text{tr } G_{s'}^{rk} \sigma_{s'}^{rk}$ minimizes the Shannon entropy.

Theorem 10 *The Shannon entropy of Alice's bit S' given the state σ_{AS}^{RK} , R and K is:*

$$\beta < \beta_{\text{sat}} : \quad \mathsf{H}(S'|RK\sigma_{AS}^{RK}) = h\left(\frac{1}{2} + \frac{1}{2\sqrt{d-1}} \frac{\sqrt{\beta}}{\beta_{\text{sat}}} \sqrt{2\beta_{\text{sat}} - \beta}\right). \quad (116)$$

$$\beta \geq \beta_{\text{sat}} : \quad \mathsf{H}(S'|RK\sigma_{AS}^{RK}) = h\left(\frac{1}{2} + \frac{1}{2\sqrt{d-1}}\right). \quad (117)$$

Proof: The min-entropy result (109,110) can be written as $\mathsf{H}_{\min}(S'|RK\sigma_{AS}^{RK}) = -\log \text{tr } T_{s'}^{rk} \sigma_{s'}^{rk}$, so we already have an expression for $\text{tr } T_{s'}^{rk} \sigma_{s'}^{rk}$. Substitution of \mathcal{T}^{rk} for \mathcal{G}^{rk} in (114) yields the result. \square

Since the optimal POVM for min- and Shannon entropy are the same, saturation occurs at the same point ($\beta = \beta_{\text{sat}}$). Fig 7 shows the Shannon entropy leakage (mutual information) $I_{\text{AE}} = 1 - \mathsf{H}(S'|RK\sigma_{AS}^{RK})$ as a function of β .

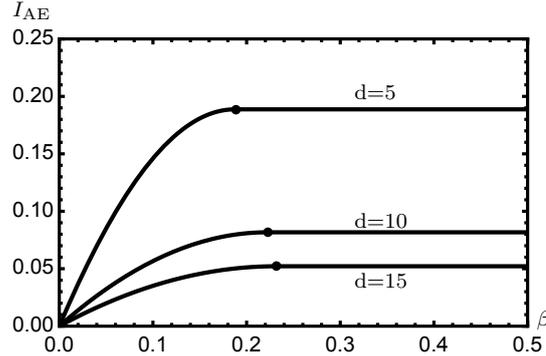


Fig. 7 Accessible Shannon entropy as a function of β for $d = 5$, $d = 10$ and $d = 15$. A dot indicates the saturation point β_{sat} .

C Optimization for the min-entropy

Here we prove that (105,106) maximizes (108). We first show that (108) is concave and obtain the optimum for $\beta \geq \beta_{\text{sat}}$. Then we take into account the constraints on the eigenvalues and derive the optimum for $\beta < \beta_{\text{sat}}$.

Unconstrained optimization. For notational convenience we define

$$w_1 = \sqrt{d\lambda_+ + 2\beta - \frac{d^2}{2}(\lambda_+ + \lambda_-)} \quad (118)$$

$$w_2 = \sqrt{d\lambda_- + 2(1-\beta) - \frac{d^2}{2}(\lambda_+ + \lambda_-)}. \quad (119)$$

This allows us to formulate everything in terms of λ_+ and λ_- . Eq. (108) becomes

$$\frac{1}{2} \left\| \sigma_0^{rk} - \sigma_1^{rk} \right\|_1 = \sqrt{d\lambda_-} w_1 + \sqrt{d\lambda_+} w_2. \quad (120)$$

Next we compute the derivatives,

$$\frac{\partial}{\partial \lambda_+} \left\| \sigma_0^{rk} - \sigma_1^{rk} \right\|_1 = -\frac{d^2}{2} \frac{\sqrt{\lambda_+}}{w_2} + \frac{w_2}{\sqrt{\lambda_+}} + \left(d - \frac{d^2}{2}\right) \frac{\sqrt{\lambda_-}}{w_1}. \quad (121)$$

$$\frac{\partial}{\partial \lambda_-} \left\| \sigma_0^{rk} - \sigma_1^{rk} \right\|_1 = -\frac{d^2}{2} \frac{\sqrt{\lambda_-}}{w_1} + \frac{w_1}{\sqrt{\lambda_-}} + \left(d - \frac{d^2}{2}\right) \frac{\sqrt{\lambda_+}}{w_2}. \quad (122)$$

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

$$\left(\frac{1}{2}d^2 - d\right)(\lambda_-^2 - \lambda_+^2) + 2\beta\lambda_+ - 2(1-\beta)\lambda_- = 0. \quad (123)$$

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

$\frac{w_2}{w_1} = \text{expression}$. Equating these two expressions gives us another hyperbola,

$$\left(d^2\lambda_+ + \frac{d^2}{2}\lambda_- - d\lambda_- - 2(1-\beta)\right) \left(d^2\lambda_- + \frac{d^2}{2}\lambda_+ - d\lambda_+ - 2\beta\right) - \lambda_- \lambda_+ \left(d - \frac{d^2}{2}\right) = 0. \quad (124)$$

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

$$\lambda_+ = 0 ; \quad \lambda_- = \frac{4(1-\beta)}{d(d-2)} \quad (125)$$

$$\lambda_+ = \frac{4\beta}{d(d-2)} ; \quad \lambda_- = 0 \quad (126)$$

$$\lambda_+ = \frac{1}{2d(d-1)} + \frac{1-2\beta}{d^2} ; \quad \lambda_- = \frac{1}{2d(d-1)} - \frac{1-2\beta}{d^2} \quad (127)$$

$$\lambda_+ = \frac{2+d(1-2\beta)}{2d^2} ; \quad \lambda_- = \frac{2-d(1-2\beta)}{2d^2}. \quad (128)$$

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 (121,122) 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 (128) into the derivatives does not yield two zeros. Expression (127) 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 ,

$$\lambda_+ = 0 ; \quad \lambda_- = \frac{4\beta}{d^2} \Rightarrow \frac{1}{2} \left\| \sigma_0^{rk} - \sigma_1^{rk} \right\|_1 = 0. \quad (129)$$

$$\lambda_- = 0 ; \quad \lambda_+ = \frac{4\beta}{d(d-2)} \Rightarrow \frac{1}{2} \left\| \sigma_0^{rk} - \sigma_1^{rk} \right\|_1 = 2\sqrt{2} \frac{\sqrt{\beta(d-2) - 2\beta^2(d-1)}}{d-2}. \quad (130)$$

Clearly (130) is the larger of the two and therefore the optimal choice. \square

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. *Phys. 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. Z. Zhang, X. Yuan, Z. Cao, and X. Ma. Practical round-robin differential-phase-shift quantum key distribution. *New Journal of Physics*, 19:033013, 2017.
16. T. Sasaki and M. Koashi. A security proof of the round-robin differential phase shift quantum key distribution protocol based on the signal disturbance. *Quantum Science and Technology*, 2(2):024006, 2017.
17. B. Škorić. A short note on the security of Round-Robin Differential Phase-Shift QKD, 2017. <https://eprint.iacr.org/2017/052.pdf>.
18. R. König R. Renner. Universally composable privacy amplification against quantum adversaries. *Proc. of TCC, LNCS 3378*, 2005.
19. M. Christandl, R. König, and R. Renner. Postselection technique for quantum channels with applications to quantum cryptography. *Phys.Rev.Lett.*, 102:020504, 2009.
20. Hoi-Kwong Lo. Efficient quantum key distribution scheme and a proof of its unconditional security. *Journal of Cryptology*, 18(2):133–165, 2005.
21. R. Renner. *Security of quantum key distribution*. PhD thesis, ETH Zürich, 2005.
22. B. Coecke and A. Kissinger. *Picturing quantum processes*. Cambridge University Press, 2017.
23. R. König, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. *IEEE Trans.Inf.Th.*, 55(9):4337–4347, 2009.
24. A.S. Holevo. Statistical decision theory for quantum systems. *Journal of multivariate analysis*, 3:337–394, 1973.