

Security of discrete-modulated continuous-variable quantum key distribution

Stefan Bäuml¹, Carlos Pascual-García¹, Victoria Wright¹, Omar Fawzi², and Antonio Acín^{1,3}

¹ICFO-Institut de Ciències Fòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain.

²Université de Lyon, Inria, ENS de Lyon, UCBL, LIP, F-69342, Lyon Cedex 07, France.

³ICREA - Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain.

Continuous variable quantum key distribution with discrete modulation has the potential to provide information-theoretic security using widely available optical elements and existing telecom infrastructure. While their implementation is significantly simpler than that for protocols based on Gaussian modulation, proving their finite-size security against coherent attacks poses a challenge. In this work we prove finite-size security against coherent attacks for a discrete-modulated quantum key distribution protocol involving four coherent states and heterodyne detection. To do so, and contrary to most of the existing schemes, we first discretize all the continuous variables generated during the protocol. This allows us to use the entropy accumulation theorem, a tool that has previously been used in the setting of discrete variables, to construct the finite-size security proof. We then compute the corresponding finite-key rates through semi-definite programming and under a photon-number cutoff. Our analysis provides asymptotic rates in the range of $0.1 - 10^{-4}$ bits per round for distances up to hundred kilometres, while in the finite case and for realistic parameters, we get of the order of 10 Gbits of secret key after $n \sim 10^{11}$ rounds and distances of few tens of kilometres.

Contents

1	Introduction	2
2	Preliminaries	4
2.1	Basic notations	4
2.2	Security definition	5
3	The QKD protocol	5
3.1	The hypothetical QKD protocol	6
3.2	The physical QKD protocol	9
4	Security of the QKD protocol	10
4.1	Soundness	10
4.1.1	Reduction to Collective Attacks via Entropy Accumulation	11
4.2	Completeness	15
4.3	The Min-Tradeoff Function	16
4.3.1	Removing the dependence on the $\hat{E}$ subsystem	17
4.3.2	Finding an affine crossover min-tradeoff function	18
4.3.3	Optimisation of the crossover min-tradeoff function	19
4.4	Asymptotic Rates	21
5	Numerical implementation and results	22
6	Discussion	25
	Acknowledgements	27
A	Proof of Lemma 3	28
B	Proof of Lemma 4	30
C	Upper bounding the classical smooth max entropy	30
D	Perturbative analysis for finite-key distillation	32
	References	33

1 Introduction

Arguably one of the most technologically advanced applications of quantum information theory nowadays is quantum key distribution (QKD), which allows two honest parties, Alice and Bob, to obtain a cryptographic key, the security of which is guaranteed by the laws of quantum physics. Whereas QKD was originally conceived in a setting involving discrete variables [1, 2, 3], e.g. requiring the generation, or at least approximation, of single photon states, there exist a number of protocols based on continuous variable systems, such as squeezed or coherent states [4, 5, 6, 7]. These protocols, known as continuous variable quantum key distribution (CVQKD), provide a number of advantages over discrete variable quantum key distribution (DVQKD) in terms of implementation using present-day telecom infrastructure.

The security of DVQKD has been proven both in theory and in realistic implementations using diverse approaches, see for instance [8, 9, 10, 11, 12]. Different security proofs have also been provided for CVQKD, many of which make use of a particular feature of the protocol, namely that the quantum states sent from Alice to Bob are chosen according to a Gaussian distribution. Such protocols are also known as Gaussian modulated CVQKD protocols. An important ingredient when proving security of Gaussian modulated CVQKD against collective attacks is the extremality of Gaussian states [13]. Gaussian extremality implies that, for a given covariance matrix of Alice and Bob's system, the maximum over the Holevo quantity in the Devetak-Winter formula for the key rate [14], which involves an optimisation over Eve's full Fock space, is attained by the corresponding Gaussian state. Combining this with the fact that, in the case of Gaussian modulation, the covariance matrix of Alice and Bob's system can be directly computed from the observed statistics [15], security against collective attacks has been shown for Gaussian modulated coherent and squeezed states protocols involving both homodyne and heterodyne detection [16, 17, 18, 19]. Security against general attacks has been shown for protocols using coherent [20, 21, 19, 22] as well as squeezed states [23, 24]. The main tools that have been used are the de Finetti Theorem [20, 22], postselection techniques [25, 21] and entropic uncertainty rela-

tions [23, 24].

Unfortunately, the implementation of CVQKD protocols with Gaussian modulation turns out to be challenging because it is never achieved in practice [26], and is in fact often approximated by finite sets of states. A discrete modulation therefore significantly simplifies the preparation of states but also the error correction part, as much simpler reconciliation schemes can be used [27]. Discrete-modulated protocols involve Alice sending coherent states taken from a typically small set, e.g. containing two or four states, according to some distribution, to Bob, who then applies a homo- or heterodyne measurement and discretises his outcome. Despite their simplicity, less is known about the security of such schemes.

The main challenge is that, unlike in the case of Gaussian modulation, the first and second moment of Alice and Bob's state are generally not sufficient to determine Eve's information, as one cannot invoke Gaussian extremality. Nevertheless, security against collective attacks has been shown in a number of scenarios. In [27], security was proven for a limited class of transmission channels. For a protocol using Gaussian modulation for parameter estimation and discrete modulation for key generation, which requires decoy states, security against collective attacks and general security in the asymptotic limit has been shown in [28]. The authors of [29] apply an optimisation over possible covariance matrices of Alice and Bob's state as well as a reduction to the Gaussian optimality method to show security against collective attacks in the asymptotic limit. Higher key rates, which are secure against collective attacks in the asymptotic limit, are obtained by [30], which use an optimisation over all possible density matrices of Alice and Bob's state that are compatible with the observed statistics and without invoking the arguments of Gaussian optimality, but using a cutoff assumption that limits the number of photons in the state. This assumption was removed in [31] for the asymptotic case, and in [32] for the finite-size case; however both works only consider collective attacks. Security of discrete modulated coherent state protocols with homodyne or heterodyne detection against collective attacks was also shown by [33, 34] in the asymptotic case, and by [35] in the finite-size case. In the limit of a high number of coherent states, asymptotic security against collective

attacks was shown in [36]. In the setting of collective Gaussian attacks, the security of discrete modulated coherent state protocols with heterodyne detection, for any number of coherent states, has also been proven in the finite-size regime [37]. Finally, finite-size security against general attacks has been shown for a protocol involving a discrete modulation using two coherent states [38, 39, 40].

In this work, we consider a protocol involving a discrete modulation using four coherent states [41], and heterodyne detection [6], which is closely related to the protocol presented in [30]. The main difference with respect to [30] is that all the information generated by the protocol, for key generation and parameter estimation, is discretised, in a similar way as was done in the entropic uncertainty relation approach of [23, 24]. This allows us to prove security against general attacks, as well as finite block sizes, using the entropy accumulation theorem (EAT) [42, 43], which has previously been used to prove the security of device independent quantum key distribution against general attacks [44, 45, 46].

The EAT is a powerful tool that allows one to lower bound the conditional smooth min-entropy, a quantity that can be used to quantify the amount of secret key obtainable from a (generally unstructured) classical-quantum (cq) state by means of privacy amplification, using hash functions [11]. This is in fact the relevant situation in QKD protocols, since a cq-state is produced where Alice and Bob hold classical information, resulting in our case from Alice's preparation and Bob's measurements, whereas Eve's system remains quantum. The EAT requires the cq-state being the result of a sequence of maps, known as EAT channels, each of which provides classical outputs and side-information, while also passing on a quantum system to the next map. The lower bound on the conditional smooth min-entropy is in terms of a so-called 'min-tradeoff function', mapping the observed statistic of classical outputs of the EAT channels to a real number which cannot exceed the single round conditional von Neumann entropies of any of the EAT channels.

A major challenge, when applying the EAT in security proofs for QKD, is that the EAT channels need to fulfill a Markov condition, ensuring that in each round, given all past-side information, there are no new correlations between pre-

vious outcomes and the new side information. As information used for parameter estimation is obtained from measurements by Alice and Bob on systems which Eve could potentially have correlated in a way incompatible with the Markov condition, the EAT cannot be applied to the QKD protocol directly. Rather, a hypothetical EAT process is introduced which produces the same marginal states on the subsystems relevant to the security proof, and the smooth min-entropy of the QKD protocol is lower bounded using a combination of chain rules, as well as a min-tradeoff function corresponding to the EAT process [44, 45, 46].

As was recently pointed out by the authors of [47], another issue arises when applying the EAT in device dependent prepare-and-measurement protocols. Such protocols can be translated into entanglement based protocols, where Alice, instead of randomly sending states, prepares an entangled state, part of which is sent to Bob via an insecure channel, while the remaining part is kept in Alice's lab. Alice and Bob then perform measurements on their respective parts. The issue which arises is that the statistics obtained from the measurements is not sufficient to certify that the state between Alice and Bob is entangled, requiring additional constraints on Alice's marginal in the final key rate optimisation, which are incompatible with the EAT. We overcome this issue by adding an additional tomography performed by Alice in randomly chosen rounds, thus ensuring that Alice and Bob's measurement statistics are sufficient to certify entanglement between Alice and Bob.

Having overcome these challenges, we are able to derive a min-tradeoff function using the numerical approach presented in [48], which requires a photon number cutoff assumption. It involves a linearisation of the objective function and the use of duality, finally reducing the problem to a semi-definite programming optimisation, which can be efficiently handled numerically. Our numerical analysis also suggests that the values of the key rate do not significantly vary once the cutoff becomes large enough. Using this approach, we are able to obtain asymptotic rates in the range of $0.1 - 10^{-4}$ bits per round for distances up to hundred kilometres. In the finite setting, and for realistic parameters, we get of the order of 10 Gbits of secret key after $n \sim 10^{11}$ rounds and distances

of few tens of kilometres.

After most of the work that went into this result was completed, a generalised version of the EAT has been presented [49, 50], which offers an alternative method of overcoming the challenges to prove the security of device-dependent prepare-and-measure protocols mentioned in the previous two paragraphs. In another recent result, the authors of [32] have overcome the photon number cutoff assumption on Bob's state needed to compute the min-tradeoff function that defines the asymptotic rates by means of adding an additional energy test, as well as a dimension reduction technique presented in [31]. Their proof also works in the finite setting, albeit only against collective attacks.

2 Preliminaries

2.1 Basic notations

In this section we introduce some definitions and concepts we use throughout the paper. For a Hilbert space $\mathcal{H}_A$, we denote by $\mathcal{D}(\mathcal{H}_A)$ the set of density operators, i.e. positive semidefinite operators with unit trace, ρ_A , acting on quantum system A . Sometimes it will be convenient to consider subnormalised states, i.e. states with $\text{Tr}[\rho] \leq 1$, in which case we use the notation $\mathcal{D}_{\leq}(\mathcal{H}_A)$. The notation $\mathcal{H}_{AB}$ denotes a tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, and ρ_{AB} the corresponding bipartite density operator. Classical random variables X , taking values $\{x\}$ according to the distribution $\{p_x\}$ can be expressed as density operators as $\rho_X = \sum_x p_x |x\rangle \langle x|_X$. By XY we denote the Cartesian product of random variables X and Y . Further, we will be using the notation $A_1^n = A_1 A_2 \dots A_n$ and $X_1^n = X_1 X_2 \dots X_n$ for quantum and classical systems. We express cq states using the notation $\rho_{XA} = \sum_x p_x |x\rangle \langle x|_X \otimes \rho_A^x$. For a cq state $\rho_{CQ} = \sum_c p(c) |c\rangle \langle c| \otimes \rho_c$, an event Ω is defined as a subset of the elements $\{c\}$. The conditional state is then given by $\rho_{CQ|\Omega} = \frac{1}{\text{Pr}_\rho[\Omega]} \sum_{c \in \Omega} p(c) |c\rangle \langle c| \otimes \rho_c$ where $\text{Pr}_\rho[\Omega] := \sum_{c \in \Omega} p(c)$. When the state ρ is clear from the context, we use $\text{Pr}[\Omega]$ in place of $\text{Pr}_\rho[\Omega]$.

For two subnormalised states $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H}_A)$, we define the generalised fidelity

$$F(\rho, \sigma) = \left(\text{Tr}[\sqrt{\sqrt{\rho}\sqrt{\sigma}}] + \sqrt{(1 - \text{Tr}[\rho])(1 - \text{Tr}[\sigma])} \right)^2, \quad (1)$$

the generalised trace distance

$$\Delta(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1 + \frac{1}{2} |\text{Tr}[\rho - \sigma]|, \quad (2)$$

as well as the purified distance

$$P(\rho, \sigma) = \sqrt{1 - F(\rho, \sigma)}. \quad (3)$$

The generalised trace distance and the purified distance are metrics on $\mathcal{D}_{\leq}(\mathcal{H}_A)$. They are related by the Fuchs-van de Graaf inequality

$$\Delta(\rho, \sigma) \leq P(\rho, \sigma) \leq \sqrt{2\Delta(\rho, \sigma) - \Delta(\rho, \sigma)^2} \leq \sqrt{2\Delta(\rho, \sigma)}. \quad (4)$$

In this work we make use of a number of entropic quantities. In addition to the well known von Neumann entropy, $H(A)_\rho = H(\rho_A) = -\text{Tr}[\rho_A \log \rho_A]$, the conditional von Neumann entropy, $H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$, as well as the Umegaki relative entropy,

$$D(\rho||\sigma) = \frac{1}{\text{Tr}[\rho]} \text{Tr}[\rho(\log \rho - \log \sigma)], \quad (5)$$

when $\text{supp}(\rho) \subset \text{supp}(\sigma)$ and $+\infty$ otherwise, for positive semidefinite ρ and σ , we make use of min and max conditional entropies, defined for a subnormalised quantum state $\rho_{AB} \in \mathcal{D}_{\leq}(\mathcal{H}_{AB})$ by [51],

$$H_{\min}(A|B)_\rho = \sup_{\sigma_B \in \mathcal{D}_{\leq}(\mathcal{H}_B)} \sup \{ \lambda \in \mathbb{R} : \rho_{AB} \leq \exp(-\lambda) \mathbb{1}_A \otimes \sigma_B \}, \quad (6)$$

$$H_{\max}(A|B)_\rho = \max_{\sigma_B \in \mathcal{D}_{\leq}(\mathcal{H}_B)} \log F(\rho_{AB}, \mathbb{1}_A \otimes \sigma_B). \quad (7)$$

For $\epsilon \geq 0$, we can then define the smooth min and max entropies as [51],

$$H_{\min}^\epsilon(A|B)_\rho = \max_{\tilde{\rho} \in \mathcal{B}^\epsilon(\rho_{AB})} H_{\min}(A|B)_{\tilde{\rho}}, \quad (8)$$

$$H_{\max}^\epsilon(A|B)_\rho = \min_{\tilde{\rho} \in \mathcal{B}^\epsilon(\rho_{AB})} H_{\max}(A|B)_{\tilde{\rho}}, \quad (9)$$

where $\mathcal{B}^\epsilon(\rho_A)$ is the ϵ -ball around a state ρ_A in terms of purified distance, i.e. the set of subnormalised states $\tau \in \mathcal{D}_{\leq}(\mathcal{H}_A)$ such that $P(\tau, \rho) \leq \epsilon$. For parameter $a \in (1, 2)$, let us further define the sandwiched Rényi divergence [52, 53] for a quantum state ρ and positive semidefinite σ as

$$D_a(\rho||\sigma) = \frac{1}{a-1} \log \text{Tr} \left[\left(\sigma^{-\frac{a-1}{2a}} \rho \sigma^{-\frac{a-1}{2a}} \right)^a \right], \quad (10)$$

when $\text{supp}(\rho) \subset \text{supp}(\sigma)$ and $+\infty$ otherwise, and the conditional Rényi entropy as

$$H_a^\uparrow(A|B)_\rho = \inf_{\sigma_B \in \mathcal{D}_\leq(\mathcal{H}_B)} D_a(\rho_{AB} \| \mathbb{1}_A \otimes \sigma_B). \quad (11)$$

2.2 Security definition

When two parties, Alice and Bob, wish to communicate in perfect secrecy in the presence of a quantum eavesdropper Eve, they need to perform a QKD protocol, typically consisting of n rounds of quantum communication and local measurements, followed by classical post-processing steps involving parameter estimation, error correction and privacy amplification. An instance of a QKD protocol may be aborted if certain tests included in the protocol, such as parameter estimation, fail, or if a subprotocol, such as error correction aborts. If the protocol does not abort, the goal is to obtain a state close to a so-called perfect classical-classical-quantum (ccq) state of the form

$$\rho_{K_A K_B E}^{\text{perfect ccq}} = \frac{1}{d} \sum_{x=0}^{d-1} |xx\rangle \langle xx|_{K_A K_B} \otimes \rho_E, \quad (12)$$

where Alice and Bob's systems are classical, whereas Eve's system may be quantum. Such a state corresponds to $\log d$ bits of an ideal classical key between Alice and Bob which is secret in that it is completely uncorrelated from Eve, even if Eve is allowed to possess a quantum system. And it is correct in the sense that Alice and Bob's systems are perfectly classically correlated.

A proof of security of a QKD protocol then involves two parts: Firstly, it has to be shown that it results in a state that is sound, i.e. close to a perfect ccq-state. Formally, for $\epsilon^{\text{sou}} > 0$, a QKD protocol is said to be ϵ^{sou} -sound, if it results in a state $\rho_{K_A K_B E}^{\text{QKD}}$, such that if we condition on the event Ω_{NA} of not aborting the protocol it holds

$$\Pr_{\rho^{\text{QKD}}}[\Omega_{\text{NA}}] \frac{1}{2} \left\| \rho_{K_A K_B E}^{\text{QKD}} |_{\Omega_{\text{NA}}} - \rho_{K_A K_B E}^{\text{perfect ccq}} \right\|_1 \leq \epsilon^{\text{sou}}. \quad (13)$$

As we wish to treat the error correction protocol separately from the the remaining protocol, it is convenient to split the soundness property into a secrecy and correctness part. Namely, let $\epsilon^{\text{sec}} > 0$ and $\epsilon^{\text{cor}} > 0$. A QKD protocol is said to be ϵ^{sec} -secret if

$$\Pr_{\rho^{\text{QKD}}}[\Omega_{\text{NA}}] \frac{1}{2} \left\| \rho_{K_A E}^{\text{QKD}} |_{\Omega_{\text{NA}}} - \rho_{K_A E}^{\text{perfect ccq}} \right\|_1 \leq \epsilon^{\text{sec}}. \quad (14)$$

The protocol is further said to be ϵ^{cor} -correct if

$$\Pr_{\rho^{\text{QKD}}}[K_A \neq K_B \wedge \Omega_{\text{NA}}] \leq \epsilon^{\text{cor}}. \quad (15)$$

If the protocol is both ϵ^{sec} -secret and ϵ^{cor} -correct, it is $\epsilon^{\text{sec}} + \epsilon^{\text{cor}}$ -sound. The second part of a security proof is to show completeness, meaning that there is an honest implementation, i.e. an implementation without presence of Eve, that does succeed, i.e. does not abort, with high probability. Formally, for $\epsilon^{\text{com}} > 0$, we say that a QKD protocol is ϵ^{com} -complete, if

$$1 - \Pr_{\text{hon}}[\Omega_{\text{NA}}] \leq \epsilon^{\text{com}}, \quad (16)$$

where the subscript "hon" refers to the fact that we compute the probability with respect to the honest implementation specified by the protocol.

3 The QKD protocol

The QKD protocol we consider is based on the four coherent-state protocol using heterodyne detection described in [30]. However, we perform a discretisation of Bob's measurement outputs in both key and parameter rounds rather than just in key rounds. Our protocol also differs from the one presented in [30] in that we do not include post-selection. In each round of the protocol, Alice prepares one of four coherent states $|\varphi_x\rangle \in \{|\alpha\rangle, |-\alpha\rangle, |i\alpha\rangle, |-i\alpha\rangle\}$ for some predetermined $\alpha \in \mathbb{R}$, with probability $\frac{1}{4}$. The state is then sent to Bob via a noisy channel that is potentially compromised by Eve. Bob then performs a heterodyne measurement.

We will prove security using an equivalent entanglement-based QKD protocol. Such a protocol can be defined by the source replacement scheme [3, 15, 54, 55]. Namely, in each round $i = 1, \dots, n$, Alice prepares an independent copy of the pure state

$$|\psi\rangle_{AA'} = \frac{1}{2} \sum_{x=0}^3 |x\rangle_A |\varphi_x\rangle_{A'}. \quad (17)$$

Alice sends the A' subsystem to Bob via a noisy quantum channel, keeping the A subsystem. Alice and Bob then both perform measurements on their respective subsystems.

Whereas this kind of introduction of an entanglement-based protocol is commonly used when proving security of prepare-and-measure protocols, we face an additional challenge when

combining this approach with the EAT [47]. Namely, unlike in a device independent setting, the statistics obtained from Alice and Bobs measurement, even in an honest implementation, is not sufficient to certify entanglement of eq. (17). In fact, as in the entanglement-based version Alice only implements one measurement in the computational basis, the statistics produced by the protocol can equally be explained by the separable state

$$\rho_{AA'} = \frac{1}{4} \sum_{x=0}^3 |x\rangle \langle x|_A \otimes |\varphi_x\rangle \langle \varphi_x|_{A'}. \quad (18)$$

This is why to derive a positive secret-key rate, one includes a constraint on the marginal of Alice's state in the optimisation for the key rate [30]. Namely, the marginal is required to take the form

$$\rho_A = \frac{1}{4} \sum_{x,y=0}^3 \langle \varphi_y | \varphi_x \rangle |x\rangle \langle y|_A, \quad (19)$$

which is not satisfied by the separable state (18). The challenge then is to express such a constraint in terms of a distribution obtained by statistical analysis, which is required when applying the EAT.

We overcome this challenge by considering a hypothetical version of our protocol, where, in randomly chosen rounds, Alice performs tomographic measurements of her marginal on A and, in the end, verifies whether the obtained statistics are compatible with her marginal being equal to eq. (19). If this is not the case, the protocol gets aborted. As the data obtained in the tomography rounds is not used for key generation, the key rate obtained in this hypothetical protocol is never larger than the key rate obtained in the physically implemented protocol, where Alice performs no tomography. Also, as we are in a device dependent setting, where we can assume Alice's state preparation to be perfect, the only scenario under which the hypothetical protocol aborts after the tomography test is due to imperfect tomography, the probability of which becomes negligible for large enough n . In the following, we use the term 'hypothetical QKD protocol' when we consider the entanglement-based protocol including tomography and 'physical QKD protocol' when referring to the entanglement-based protocol which does not include tomography. By the source-replacement scheme the latter is equivalent to the prepare-and-measure protocol that is

actually performed in the laboratory by Alice and Bob.

When the state eq. (17) is sent from Alice to Bob, we assume that Eve can attack the channel used to send the A' subsystem coherently. This is equivalent to a scenario where Alice initially prepares all n independent and identically distributed (iid) copies of the state (17), which are then acted upon a channel $\mathcal{N}_{A'_1 \rightarrow B'_1}$. Let $\mathcal{U}_{A'^n \rightarrow B'^n E}$ be an isometric extension of the channel and let us define

$$|\Psi\rangle_{A'_1 B'_1 E} = \text{id}_{A^n} \otimes \mathcal{U}_{A'^n \rightarrow B'^n E} |\psi\rangle_{A'_1 A'^n}^{\otimes n}. \quad (20)$$

It has to be assumed that the E subsystem goes to Eve. Alice and Bob are left with the mixed state

$$\rho_{A'_1 B'_1} = \text{Tr}_E[\Psi_{A'_1 B'_1 E}]. \quad (21)$$

3.1 The hypothetical QKD protocol

We now describe a round of the hypothetical QKD protocol in detail. Let $0 \leq p^{\text{key}} \leq 1$, $0 \leq p^{\text{PE}} \leq 1$ and $0 \leq p^{\text{tom}} \leq 1$, where $p^{\text{key}} + p^{\text{PE}} + p^{\text{tom}} = 1$, be the respective probabilities for a given round being used for key generation, parameter estimation and tomography of Alice's marginal. For each round $i = 1, \dots, n$, Alice and Bob perform the following steps:

(1) *Alice's Measurement:* Alice uses a random number generator to create a random variable R_i , taking values $R_i = 0, 1, 2$ with respective probabilities p^{key} , p^{PE} and p^{tom} . If $R_i = 0$, the round is used for key generation. For $R_i = 1$, the round is employed for parameter estimation. In both cases Alice performs a projective measurement $\{|x\rangle \langle x|\}_{x=0}^3$ on subsystem A_i . If $R_i = 2$, Alice performs a tomography, using an informationally complete (IC) measurement defined by a Positive-Operator-Valued-Measure (POVM) $\{\Gamma_{x'}\}_{x'=0}^{15}$ on her subsystem. The outcome of Alice's measurement is described by a random variable X_i , taking values x_i . We define, for the sake of convenience,

the random variables

$$\hat{X}_i = \begin{cases} x_i & \text{if } R_i = 0, \\ \perp & \text{else.} \end{cases} \quad (22)$$

$$\tilde{X}_i = \begin{cases} x_i & \text{if } R_i = 1, \\ \perp & \text{else.} \end{cases} \quad (23)$$

$$X'_i = \begin{cases} x_i & \text{if } R_i = 2, \\ \perp & \text{else.} \end{cases} \quad (24)$$

The random variable R_i is then sent to Bob via an authenticated channel.

(2) *Bob's Measurement*: Bob performs a heterodyne measurement on subsystem B_i . From the outcome, Bob obtains a continuous random variable Y_i , taking values $y_i \in \mathbb{C}$. Again, it will be convenient to define

$$\hat{Y}_i = \begin{cases} y_i & \text{if } R_i = 0, \\ \perp & \text{else.} \end{cases} \quad (25)$$

$$\tilde{Y}_i = \begin{cases} y_i & \text{if } R_i = 1, \\ \perp & \text{else.} \end{cases} \quad (26)$$

(3) *Discretisation*: Bob discretises his heterodyne outcomes. For key rounds, let $\hat{y}_i = |\hat{y}_i|e^{i\hat{\theta}_i}$ for $\hat{\theta}_i \in [-\frac{\pi}{4}, \frac{7\pi}{4})$. Bob then creates a random variable

$$\hat{Z}_i = \begin{cases} 0 & \text{if } \hat{\theta}_i \in [-\frac{\pi}{4}, \frac{\pi}{4}) \\ 1 & \text{if } \hat{\theta}_i \in [\frac{\pi}{4}, \frac{3\pi}{4}) \\ 2 & \text{if } \hat{\theta}_i \in [\frac{3\pi}{4}, \frac{5\pi}{4}) \\ 3 & \text{if } \hat{\theta}_i \in [\frac{5\pi}{4}, \frac{7\pi}{4}) \\ \perp & \text{else,} \end{cases} \quad (27)$$

where $\hat{Z}_i = \perp$ is taken for non-key rounds. For parameter estimation rounds, Bob defines a discretisation given by an amplitude Δ and modules of length δ , such that $\Delta/\delta \in \mathbb{N}$. Let $j \in \{0, 1, 2, 3\}$ and $k \in \{0, \dots, \frac{\Delta}{\delta} - 1\}$ and let $\tilde{y}_i = |\tilde{y}_i|e^{i\tilde{\theta}_i}$. Bob then creates a random variable $\tilde{Z}_i$ according to

$$\tilde{Z}_i = \begin{cases} j + 4k & \text{if } \tilde{\theta}_i \in [\frac{\pi}{4}(2j-1), \frac{\pi}{4}(2j+1)) \\ & |\tilde{y}_i| \in [\delta k, \delta(k+1)), \\ j + 4\frac{\Delta}{\delta} & \text{if } \tilde{\theta}_i \in [\frac{\pi}{4}(2j-1), \frac{\pi}{4}(2j+1)) \\ & |\tilde{y}_i| \in [\Delta, \infty), \\ \perp & \text{else.} \end{cases} \quad (28)$$

Summarising steps (1) - (3), round i of the protocol has taken as inputs quantum systems $A_i B_i$ of the initial state (21), and created discrete classical random variables $\hat{X}_i$ and $\hat{Z}_i$ for key generation, $\tilde{X}_i$ and $\tilde{Z}_i$ to be used for parameter estimation, as well as X'_i to be used for tomography of Alice's marginal state. Let us define

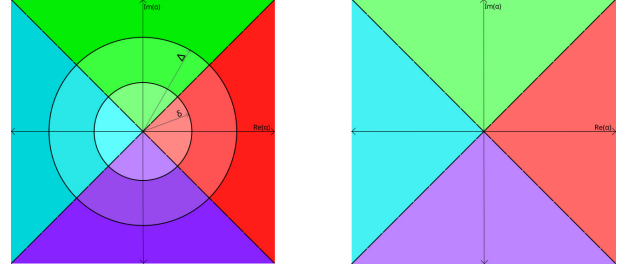


Figure 1: Discretisations of phase space by Bob for parameter estimation rounds (left) and key generation rounds (right). In this figure, the modulation for parameter estimation in phases and amplitudes is given by $\Delta/\delta = 2$, with the outmost modules extending to infinity.

$O_i := \tilde{X}_i X'_i \hat{Z}_i \tilde{Z}_i$ as the ‘output’ and $S_i := R_i$ as the ‘side information’. Let us further define $C_i = \tilde{X}_i \tilde{Z}_i X'_i$ as all the information used in statistical analysis. The reason we define O_i , S_i and C_i in this way is that we will later use these random variables when applying the EAT. The EAT requires the statistical analysis variable C_i to be obtainable from a simple read-out of ‘output’ and ‘side-information’ variables O_i and S_i . On the other hand, we cannot include $\tilde{X}_i$, $\tilde{Z}_i$ or X'_i into S_i because of the Markov condition, eq. (55). We therefore have to include them into O_i despite the fact that $\tilde{X}_i$ has to be communicated classically, and treat $\tilde{X}_i$ as additional side-information when applying Proposition 1.

Any round i of the protocol can then be described by a channel

$$\mathcal{M}^{\text{QKD}} : A_i B_i \rightarrow \hat{X}_i O_i S_i C_i. \quad (29)$$

After n rounds, the relevant systems of Alice, Bob and Eve are in the state

$$\sigma_{\hat{X}_1^n O_1^n C_1^n S_1^n E}^{\text{QKD}} = \text{id}_E \otimes \mathcal{M}^{\text{QKD} \otimes n} (\Psi_{A_1^n B_1^n E}). \quad (30)$$

The next step is to perform parameter estimation. To that purpose, Alice sends $\tilde{X}_1^n$ to Bob, who then performs the parameter estimation protocol deciding whether the protocol gets aborted or not. Further, Alice uses $X_1'^n$, obtained from her tomographic measurement, to reconstruct her marginal state. If the reconstructed state is not equal to the expected one up to a certain margin of confidence, the protocol is aborted. Alice informs Bob of her decision, and in the latter case the protocol is aborted.

In order to formalise the decision to abort, we need to introduce some notation. Let us de-

note by $\mathcal{C}$ the alphabet of all possible values of $c_i = (\tilde{x}_i, \tilde{z}_i, x'_i)$ that can occur in the protocol. Such values are $c_i = (\perp, \perp, \perp)$ in key rounds, $(x, z, \perp)$ with $x \in \{0, \dots, 3\}$ and $z \in \{0, \dots, m-1\}$, (where $m = 4\Delta/\delta + 4$ is the total number of modules) in parameter estimation rounds, as well as $(\perp, \perp, x')$ with $x' \in \{0, \dots, 15\}$ in tomography rounds. It will also be convenient to define by $\tilde{\mathcal{C}}$ the alphabet of all possible values c_i can take in parameter estimation and tomography rounds, only. For a given string $c_1^n \in \mathcal{C}^n$, we denote by $\text{freq}_{c_1^n} \in \mathcal{P}_{\mathcal{C}}$ the probability distribution corresponding to the frequency of symbols $c \in \mathcal{C}$ in c_1^n , defined by $\text{freq}_{c_1^n}(c) = |\{i : c_i = c\}|/n$.

In order to decide whether or not to abort, Alice and Bob need to benchmark their obtained statistics, given by a frequency distribution $\text{freq}_{c_1^n}$, against a distribution $p_0 \in \mathcal{P}_{\mathcal{C}}$, which can be obtained in an honest implementation of the protocol. Let p_0^{sim} be the distribution of parameter estimation random variables $(\tilde{X}, \tilde{Z})$ in the honest setting with no attack and p_0^{tom} be the distribution of the tomography random variable X' . Now, we define

$$p_0(x, z, \perp) = p^{\text{PE}} p_0^{\text{sim}}(x, z), \quad (31)$$

$$p_0(\perp, \perp, x') = p^{\text{tom}} p_0^{\text{tom}}(x'), \quad (32)$$

$$p_0(\perp, \perp, \perp) = 1 - \sum_{x,z} p_0(x, z, \perp) - \sum_{x'} p_0(\perp, \perp, x') \quad (33)$$

for $x \in \{0, \dots, 3\}$, $z \in \{0, \dots, m-1\}$, and $x' \in \{0, \dots, 15\}$. We will provide an explicit form of the p_0^{sim} and p_0^{tom} we use in Section 5.

In order to compare the two distributions $\text{freq}_{c_1^n}$ and $p_0 \in \mathcal{P}_{\mathcal{C}}$, we need to introduce figures of merit, which quantify the suitability of the distributions for key generation. For now, let us only assume that these figures of merit are given by affine functions $f^{\text{PE}} : \mathcal{P}_{\mathcal{C}} \rightarrow \mathbb{R}$ and $f^{\text{tom}} : \mathcal{P}_{\mathcal{C}} \rightarrow \mathbb{R}$ of the form

$$f^{\text{PE}}(p) = \sum_{x=0}^3 \sum_{z=0}^{m-1} h_{x,z,\perp} p(x, z, \perp), \quad (34)$$

$$f^{\text{tom}}(p) = \sum_{x'=0}^{15} h_{\perp,\perp,x'} p(\perp, \perp, x'), \quad (35)$$

for some coefficients $h_{x,z,\perp}, h_{\perp,\perp,x'} \in \mathbb{R}$. We will provide an explicit form of the functions later, together with the methodology that compares $\text{freq}_{c_1^n}$ and p_0 . For the moment, let us note that a reduced number of binnings (i.e., employing as

few modules as possible for parameter estimation) decreases the overall differences of distributions $\text{freq}_{c_1^n}$ and p_0 in the case of f^{PE} . This is eventually reflected as an increase in p^{key} , since fewer rounds must be spent on bounding the differences between said distributions. We now define the respective sets of probabilities for which we do not abort after parameter estimation or tomography as

$$\mathcal{P}_{\Omega_{\text{PE}}} := \left\{ p \in \mathcal{P}_{\mathcal{C}} : f^{\text{PE}}(p) \geq f^{\text{PE}}(p_0) - \delta_{\text{PE}}^{\text{tol}} \right\}, \quad (36)$$

$$\mathcal{P}_{\Omega_{\text{tom}}} := \left\{ p \in \mathcal{P}_{\mathcal{C}} : f^{\text{tom}}(p) \geq f^{\text{tom}}(p_0) - \delta_{\text{tom}}^{\text{tol}} \right\}, \quad (37)$$

for some $\delta_{\text{PE}}^{\text{tol}}, \delta_{\text{tom}}^{\text{tol}} > 0$. Let us also define $\delta^{\text{tol}} = \delta_{\text{PE}}^{\text{tol}} + \delta_{\text{tom}}^{\text{tol}}$, as well as the events of passing the parameter estimation and the tomography test as

$$\Omega_{\text{PE}} := \left\{ c_1^n \in \mathcal{C}^n : \text{freq}_{c_1^n} \in \mathcal{P}_{\Omega_{\text{PE}}} \right\}, \quad (38)$$

$$\Omega_{\text{tom}} := \left\{ c_1^n \in \mathcal{C}^n : \text{freq}_{c_1^n} \in \mathcal{P}_{\Omega_{\text{tom}}} \right\}, \quad (39)$$

$$\Omega_{\text{EA}} = \Omega_{\text{PE}} \cap \Omega_{\text{tom}}. \quad (40)$$

Assuming that Alice and Bob do not abort after parameter estimation or tomography, they perform an error correction protocol using reverse reconciliation. The information exchanged between Alice and Bob in this step is denoted L , and at the end of this step Alice computes a string $\bar{X}_1^n$. In order to check that the error correction was successful, Bob chooses a random hash function H and sends to Alice a description of H as well as the value $H' = H(\hat{Z}_1^n)$. Whenever $H(\hat{Z}_1^n) \neq H(\bar{X}_1^n)$, the protocol is aborted. Let us denote by H and H' the register containing the description and value of the hash function, respectively. Formally, we define the event of passing the error correction step as

$$\Omega_{\text{EC}} = [H(\hat{Z}_1^n) = H(\bar{X}_1^n)]. \quad (41)$$

We assume that there is a small probability, upper bounded by $\epsilon_{\text{EC}} > 0$, of the error correction being passed by mistake. For any $\hat{z}_1^n \neq \bar{x}_1^n$, $\Pr[H(\hat{z}_1^n) = H(\bar{x}_1^n)] \leq \epsilon_{\text{EC}}$, where $\Pr$ here is over the choice of H . Further we assume that the probability of not passing the error correction in an honest implementation is upper bounded by $\Pr_{\text{hon}}[H(\hat{Z}_1^n) \neq H(\bar{X}_1^n)] \leq \epsilon_{\text{EC}}^c$, for some $\epsilon_{\text{EC}}^c > 0$.

Protocol 1 Hypothetical QKD protocol

1. Alice prepares state $|\Psi\rangle_{AA'}^{\otimes n}$ given by eq. (17) and sends subsystems A_1^n to Bob via a noisy channel, and Bob receives subsystems B_1^n . The total state is then given by eq. (20).
 2. For round $i = 1, \dots, n$ the following steps are performed:
 - (a) Alice chooses $R_i \in \{0, 1, 2\}$ according to $(p^{\text{key}}, p^{\text{PE}}, p^{\text{tom}})$ and sends R_i to Bob.
 - (b) If $R_i = 0$, Alice measures A_i in a computational basis and stores output in $\hat{X}_i$. Bob measures B_i using a heterodyne measurement, discretises according to eq. (27), and stores the result in $\hat{Z}_i$.
 - (c) If $R_i = 1$, Alice measures A_i in computational basis and stores output in $\tilde{X}_i$. Bob measures B_i using heterodyne measurements, discretises according to eq. (28), and stores the result in $\tilde{Z}_i$.
 - (d) If $R_i = 2$, Alice measures A_i using an informationally complete measurement and stores output in X'_i .
 3. Alice and Bob use $\tilde{X}_1^n \tilde{Z}_1^n$ for parameter estimation. Alice uses X_1^n for the tomography test. If either fails the protocol is aborted.
 4. If the protocol has not been aborted, error correction is performed using reverse reconciliation. If error correction fails, the protocol is aborted. Otherwise privacy amplification is performed resulting in the final key.
-

Now, we define the event of not aborting the protocol after either parameter estimation, tomography or error correction as

$$\Omega_{\text{NA}} = \Omega_{\text{EA}} \cap \Omega_{\text{EC}}. \quad (42)$$

We note that Ω_{EA} only depends on the C_1^n registers, whereas Ω_{EC} depends on the $HH'\tilde{X}_1^n$ registers. The description of the protocol together with an attack of Eve leads to the final state

$$\begin{aligned} & \sigma_{\tilde{X}_1^n O_1^n S_1^n C_1^n HH'LE}^{\text{QKD}} \\ &= \Pr[\Omega_{\text{NA}}] \sigma_{\tilde{X}_1^n O_1^n S_1^n C_1^n HH'LE}^{\text{QKD}}|_{\Omega_{\text{NA}}} \\ &+ (1 - \Pr[\Omega_{\text{NA}}]) \sigma_{\tilde{X}_1^n O_1^n S_1^n C_1^n HH'LE}^{\text{QKD}}|_{\neg\Omega_{\text{NA}}}. \end{aligned}$$

At this point, if Alice and Bob did not abort, they proceed with a privacy amplification protocol (e.g. via two-universal hash functions) to distill the final, secret key.

3.2 The physical QKD protocol

Finally, let us describe the physical QKD protocol, which is the equivalent entanglement-based version of the prepare-and-measure protocol that is actually performed in a realistic implementation. The physical QKD protocol is essentially equal to the hypothetical protocol except for the following two differences: Firstly, in step (1) of the protocol, whenever $R_i = 2$ Alice does not perform tomography—the round is simply discarded. Hence, the random variable X'_i will be either $\perp$ or undefined. Secondly, as a consequence of not performing the tomography, abortion or non-abortion at the end of the protocol will only be determined by parameter estimation and error correction. I.e., the event Ω_{NA} will be replaced by $\Omega_{\text{NA}}^{\text{phys}} = \Omega_{\text{PE}} \cap \Omega_{\text{EC}}$.

In principle, a more efficient physical protocol can be obtained if no round is discarded. That is, if we set $p^{\text{tom}} = 0$. In our case, however, a non-zero value of p^{tom} is chosen because the comparison of the two protocols, hypothetical and physical, results in a much simpler analysis when using the same values for the probabilities p^{key} , p^{PE} and p^{tom} in both protocols. It is nevertheless worth noting that the value taken for p^{tom} below is very small, so the possible impact on the key rate does not represent a significant loss.

All in all, the physical protocol is composed by the following steps

Protocol 2 Physical QKD protocol

1. Alice prepares state $|\Psi\rangle_{AA'}^{\otimes n}$ given by eq. (17) and sends subsystems A_i^n to Bob via a noisy channel, and Bob receives subsystems B_1^n . The total state is then given by eq. (20).
 2. For round $i = 1, \dots, n$ the following steps are performed:
 - (a) Alice chooses $R_i \in \{0, 1, 2\}$ according to $(p^{\text{key}}, p^{\text{PE}}, p^{\text{tom}})$ and sends R_i to Bob.
 - (b) If $R_i = 0$, Alice measures A_i in a computational basis and stores output in $\hat{X}_i$. Bob measures B_i using a heterodyne measurement, discretises according to eq. (27), and stores the result in $\hat{Z}_i$.
 - (c) If $R_i = 1$, Alice measures A_i in a computational basis and stores output in $\tilde{X}_i$. Bob measures B_i using a heterodyne measurement, discretises according to eq. (28), and stores the result in $\tilde{Z}_i$.
 - (d) If $R_i = 2$, the round is not used.
 3. Alice and Bob use $\tilde{X}_1^n \tilde{Z}_1^n$ for parameter estimation. If it fails, the protocol is aborted.
 4. If the protocol has not been aborted, error correction is performed using reverse reconciliation. If error correction fails, the protocol is aborted. Otherwise privacy amplification is performed resulting in the final key.
-

4 Security of the QKD protocol

In this section we show the security against coherent attacks of the physical QKD protocol. The proof consists of two parts: Firstly, in Subsection 4.1 we show the soundness of the protocol and provide a lower bound on the key rate. We first show the soundness of the hypothetical protocol, which we then show implies the soundness of the physical protocol. The soundness proof of the hypothetical protocol is based on the EAT and depends on the choice of a min-tradeoff function of a particular form. Secondly, in Subsection 4.2, we show the completeness of the physical QKD

protocol, i.e. that there is a nonzero probability of it not being aborted. Finally, in Subsection 4.3, we show how a suitable min-tradeoff function can be derived from the numerical approach presented by [30, 48].

4.1 Soundness

In this section we provide a lower bound on the achievable key rate $r^{\text{phys}} = \ell/n$, where ℓ is the length of the key and n the number of rounds, conditioned on the event $\Omega_{\text{NA}}^{\text{phys}}$ of not aborting the physical QKD protocol. Such a lower bound can be obtained from Proposition 1 below, which is based on the leftover-hash Lemma [51]. To derive this proposition, we make the following

Bounded energy assumption: the attack by the eavesdropper involves states of finite energy.

Under this assumption, Eve's attack can be arbitrarily approximated by an attack using finite-dimensional systems. Or, in other words, we can assume Eve's attack to be defined in a Hilbert space of arbitrary finite dimension. This allows us to use the original version of the leftover hash lemma. We leave for future work how to generalise the security proof to infinite Hilbert space dimensions, for which a version of the leftover-hash Lemma has been derived in [56, 57].

Proposition 1 [11, 51] *Let $\epsilon^{\text{phys}}, \epsilon_{\text{EC}} \geq 0$. Let further leak_{EC} be the amount of information lost to Eve during error correction. Then Alice and Bob are able to extract a key of length,*

$$\ell \leq H_{\min}^{\epsilon^{\text{phys}}}(\hat{Z}_1^n | R_1^n \tilde{X}_1^n E)_{\sigma^{\text{phys, QKD}}|_{\Omega_{\text{NA}}^{\text{phys}}}} - \text{leak}_{\text{EC}} - 2 \log \frac{1}{\epsilon^{\text{phys}}}, \quad (43)$$

which is $3\epsilon^{\text{phys}} + \epsilon_{\text{EC}}$ -sound, in the sense that

$$\Pr_{\sigma^{\text{phys, QKD}}}[\Omega_{\text{NA}}^{\text{phys}}] \frac{1}{2} \|\sigma^{\text{phys, QKD}}|_{\Omega_{\text{NA}}^{\text{phys}}} - \sigma^{\text{perfect ccq}}\|_1 \leq 3\epsilon^{\text{phys}} + \epsilon_{\text{EC}}.$$

In order to apply Proposition 1, we need to lower bound the smooth min-entropy using the EAT. However, as noted in the introduction, we are unable to apply the EAT directly to our physical QKD protocol due to the need to characterise Alice's marginal system in a prepare-and-measure scenario. To that purpose we first consider the hypothetical QKD protocol that includes additional tomography measurements. However, due

to issues with the Markov condition, we will not be able to directly apply the EAT to our hypothetical protocol either. Instead, we will make use of various chain rules for smooth entropies in order to relate the output of our hypothetical QKD protocol to that of a series of n EAT channels, which we call the ‘EAT process’, and then apply the EAT to the EAT process, while dealing with the remaining terms separately.

We begin by considering the hypothetical QKD protocol. Let $n \in \mathbb{N}$ and $\epsilon > 0$. Conditioned on not aborting, the hypothetical QKD protocol results in the state $\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}$. By application of chain rules for smooth entropies (eq. (6.63) and eq. (6.56) in [51]), it holds

$$\begin{aligned} & H_{\min}^{\epsilon}(\hat{Z}_1^n | R_1^n \tilde{X}_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \geq H_{\min}^{\epsilon/4}(\hat{Z}_1^n \tilde{X}_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \quad - H_{\max}^{\epsilon/4}(\tilde{X}_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} - 2\Gamma(\epsilon/4), \end{aligned} \quad (44)$$

where $\Gamma(x) := -\log(1 - \sqrt{1 - x^2})$. By another application of a chain rule (eq. (6.57) in [51]), we obtain

$$\begin{aligned} & H_{\min}^{\epsilon/4}(\hat{Z}_1^n \tilde{X}_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \geq H_{\min}^{\epsilon/16}(\tilde{X}_1^n X_1^m \hat{Z}_1^n \tilde{Z}_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} - 3\Gamma(\epsilon/16) \\ & \quad - H_{\max}^{\epsilon/16}(X_1^m \tilde{Z}_1^n | \hat{Z}_1^n \tilde{X}_1^n R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}. \end{aligned} \quad (45)$$

We can now apply the same argument as used in [42] to upper bound the max entropy terms in eqs. (44) and (45). We begin by upper bounding the term $H_{\max}^{\epsilon/4}(\tilde{X}_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}$ in eq. (44). We note that by the strong subadditivity of the smooth max entropy [51], it holds

$$H_{\max}^{\epsilon/4}(\tilde{X}_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \leq H_{\max}^{\epsilon/4}(\tilde{X}_1^n | R_1^n)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}, \quad (46)$$

where the r.h.s. only involves classical registers. We further note that $\tilde{X}_i = \perp$, unless $R_i = 1$, which happens with probability p^{PE} , in which case $\tilde{X}_i$ takes a value in $\{0, \dots, 3\}$. Introducing a binary random variable $\tilde{R}_i$ that takes value 1 when Alice’s random variable is $R_i = 1$, and takes value 0 when $R_i = 0$ or $R_i = 2$, we can apply the data processing inequality and Lemma 6 in Ap-

pendix C, showing that

$$\begin{aligned} & H_{\max}^{\epsilon/4}(\tilde{X}_1^n | R_1^n)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \leq H_{\max}^{\epsilon/4}(\tilde{X}_1^n | \tilde{R}_1^n)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \leq np^{\text{PE}} \log 5 + \sqrt{\frac{n}{2} \ln \frac{32}{\epsilon^2 \Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]}} \log 5. \end{aligned} \quad (47)$$

In a similar way we can provide an upper bound on the term $H_{\max}^{\epsilon/16}(X_1^m \tilde{Z}_1^n | \hat{Z}_1^n \tilde{X}_1^n R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}$ in eq. (45). Again, it holds by strong subadditivity,

$$\begin{aligned} & H_{\max}^{\epsilon/16}(X_1^m \tilde{Z}_1^n | \hat{Z}_1^n \tilde{X}_1^n R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \leq H_{\max}^{\epsilon/16}(X_1^m \tilde{Z}_1^n | R_1^n)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}, \end{aligned} \quad (48)$$

where the r.h.s. is classical. Further, it holds that $X'_i \tilde{Z}_i = \perp\perp$, unless $R_i = 1$ or $R_i = 2$, which happens with probability $p^{\text{PE}} + p^{\text{tom}} = 1 - p^{\text{key}}$. In this case $X'_i \tilde{Z}_i$ takes a value in $\{0, \dots, 15, \perp\} \times \{0, \dots, m-1, \perp\}$. Let us again introduce a binary random variable $\bar{R}_i$, taking value 1 when $R_i = 1$ or $R_i = 2$ and value 0 when $R_i = 0$. We can now apply Lemma 6, identifying the pair $X'_i \tilde{Z}_i$ with X_i and the value $\perp\perp$ with $\perp$, and obtain

$$\begin{aligned} & H_{\max}^{\epsilon/16}(X_1^m \tilde{Z}_1^n | R_1^n)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \leq H_{\max}^{\epsilon/16}(X_1^m \tilde{Z}_1^n | \bar{R}_1^n)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \leq n(1 - p^{\text{key}}) \log(17(m+1)) \\ & \quad + \sqrt{\frac{n}{2} \ln \frac{512}{\epsilon^2 \Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]}} \log(17(m+1)). \end{aligned} \quad (49)$$

What remains to be done is to lower bound the term $H_{\min}^{\epsilon/16}(O_1^n | R_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}$ in eq. (45) using the EAT.

4.1.1 Reduction to Collective Attacks via Entropy Accumulation

In order to apply the EAT, we wish to condition on an event that is only defined on the statistical information C_1^n , where we recall that $C_i = \tilde{X}_i \tilde{Z}_i X'_i$ is given by classical information extracted from the outputs $O_i = \tilde{X}_i X'_i \hat{Z}_i \tilde{Z}_i$ and the side information $S_i = R_i$. For such conditioning, note that we can write $\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}$ as $(\sigma^{\text{QKD}}|_{\Omega_{\text{EA}}})|_{\Omega_{\text{EC}}}$ where the probability of the event Ω_{EC} with respect to the state $\sigma^{\text{QKD}}|_{\Omega_{\text{EA}}}$ is

given by $\Pr_{\sigma^{\text{QKD}}[\Omega_{\text{EC}}|\Omega_{\text{EA}}]}$. Now we use Lemma B.5 of [42] to show that for $a \in (1, 2)$

$$H_{\min}^{\epsilon/16}(O_1^n|S_1^n E)_{\sigma^{\text{QKD}}|\Omega_{\text{NA}}} \geq H_a^\uparrow(O_1^n|S_1^n E)_{\sigma^{\text{QKD}}|\Omega_{\text{NA}}} - \frac{\Gamma(\epsilon/16)}{a-1} \quad (51)$$

$$\geq H_a^\uparrow(O_1^n|S_1^n E)_{\sigma^{\text{QKD}}|\Omega_{\text{EA}}} - \frac{\Gamma(\epsilon/16)}{a-1} - \frac{a}{a-1} \log \left(\frac{1}{\Pr_{\sigma^{\text{QKD}}[\Omega_{\text{EC}}|\Omega_{\text{EA}}]}} \right). \quad (52)$$

In order to apply the EAT, we now consider the EAT process, which results in the same marginal state $\sigma_{O_1^n S_1^n E}^{\text{QKD}}|_{\Omega_{\text{EA}}}$, hence the same value for $H_a^\uparrow(O_1^n|S_1^n E)_{\sigma^{\text{QKD}}|\Omega_{\text{EA}}}$, as the hypothetical QKD protocol. The EAT process will be closely related to the hypothetical QKD protocol, however it will not include the output $\hat{X}_i$, which is not necessary in this context. Also, the EAT process will not include an error correction or privacy amplification protocol.

We begin by defining our EAT channels. To that purpose, we take the channel $\mathcal{M}^{\text{QKD}}$ defined in eq. (29); however we omit the output of $\hat{X}_i$, resulting in a channel

$$\mathcal{M}^{\text{EAT}} : A_i B_i \rightarrow O_i C_i S_i, \quad (53)$$

which performs steps (1) - (3) of the hypothetical QKD protocol, but in the end does not output Alice's key system $\hat{X}_i$. It is easy to see that C_i can be obtained by readout of classical information contained in O_i and S_i . As it only contains discretised information, O_i is finite dimensional. Further defining $Q_i := A_{i+1}^n B_{i+1}^n$, we can now define a channel $\mathcal{M}_i^{\text{EAT}} : Q_{i-1} \rightarrow Q_i O_i S_i C_i$ by

$$\mathcal{M}_i^{\text{EAT}} := \text{id}_{Q_i} \otimes \mathcal{M}_{A_i B_i \rightarrow O_i S_i C_i}^{\text{EAT}}. \quad (54)$$

In order to apply the EAT, we still have to show that the Markov condition

$$O_1^{i-1} \leftrightarrow S_1^{i-1} E \leftrightarrow S_i \quad (55)$$

or, equivalently $I(O_1^{i-1} : S_i | S_1^{i-1} E) = 0$, is fulfilled for all $i \in \{1, \dots, n\}$. In our case this holds trivially as $S_i = R_i$ is obtained by a local random number generator, which is used independently in each round.

We can now define our EAT process as a concatenation of EAT channels $\mathcal{M}_n^{\text{EAT}} \circ \dots \circ \mathcal{M}_1^{\text{EAT}}$,

yielding the following state,

$$\begin{aligned} & \sigma_{O_1^n S_1^n C_1^n E}^{\text{EAT}} \\ &= \text{id}_E \otimes \left(\mathcal{M}_n^{\text{EAT}} \circ \dots \circ \mathcal{M}_1^{\text{EAT}} \right) \left(\Psi_{A_1^n B_1^n E} \right) \end{aligned} \quad (56)$$

$$= \text{id}_E \otimes \mathcal{M}^{\text{EAT} \otimes n} \left(\Psi_{A_1^n B_1^n E} \right) \quad (57)$$

$$= \text{Tr}_{\hat{X}_1^n} \left[\text{id}_E \otimes \mathcal{M}^{\text{QKD} \otimes n} \left(\Psi_{A_1^n B_1^n E} \right) \right]. \quad (58)$$

The EAT process then concludes with Alice and Bob using C_1^n to perform the tomography test as well as parameter estimation. This is done in the same way as in the hypothetical QKD protocol. Consequently it holds,

$$\sigma_{O_1^n S_1^n E}^{\text{QKD}}|_{\Omega_{\text{EA}}} = \sigma_{O_1^n S_1^n E}^{\text{EAT}}|_{\Omega_{\text{EA}}}, \quad (59)$$

$$H_a^\uparrow(O_1^n|S_1^n E)_{\sigma^{\text{QKD}}|\Omega_{\text{EA}}} = H_a^\uparrow(O_1^n|S_1^n E)_{\sigma^{\text{EAT}}|\Omega_{\text{EA}}}. \quad (60)$$

Hence, it will be sufficient to lower bound the r.h.s. of eq. (60) using the EAT. We can now define a *min-tradeoff function* as a function $f : \mathcal{P}_{\mathcal{C}} \rightarrow \mathbb{R}$ such that for all $i = 1, \dots, n$ it holds

$$f(p) \leq \inf_{|\rho\rangle \in \Sigma_i(p)} H(O_i | S_i \tilde{E})_{\rho^{\text{EAT},i}}, \quad (61)$$

where $\tilde{E}$ can be chosen isomorphic to Q_{i-1} , and we have defined

$$\Sigma_i(p) = \left\{ |\rho\rangle_{Q_{i-1} \tilde{E}} \in \mathcal{H}_{Q_{i-1} \tilde{E}} : \langle c | \rho_{C_i}^{\text{EAT},i} | c \rangle \equiv p(c) \right\}, \quad (62)$$

for a state

$$\begin{aligned} \rho_{O_i S_i C_i Q_i \tilde{E}}^{\text{EAT},i} &= \text{id}_{\tilde{E}} \otimes \mathcal{M}_i^{\text{EAT}}(\rho_{Q_{i-1} \tilde{E}}) \\ &= \text{id}_{Q_i \tilde{E}} \otimes \mathcal{M}_{A_i B_i \rightarrow O_i S_i C_i}^{\text{EAT}} \left(\rho_{A_i B_i Q_i \tilde{E}} \right). \end{aligned} \quad (63)$$

Here $\equiv$ stands for equality for all $c \in \mathcal{C}$. Note that $|\rho\rangle$ can be chosen pure by strong subadditivity as remarked in [42].

In the following we will consider the case where $f(p) = f^{\text{PE}}(p) + f^{\text{tom}}(p) + \text{const}$, with affine functions f^{PE} and f^{tom} of the form given by eqs. (34,35), respectively. In that case, it holds for all $c_1^n \in \Omega_{\text{EA}}$ that $f(\text{freq}_{c_1^n}) \geq f(p_0) - \delta^{\text{tol}}$. We can then formulate the entropy accumulation theorem, given by [43, Proposition V.3], in the following way:

Proposition 2 [43] *Let $n \in \mathbb{N}$. Let p_0 be given by eqs. (31,32). Let Ω_{EA} be the event defined by eqs.*

(36-40) for some $\delta_{\text{PE}}^{\text{tol}}, \delta_{\text{tom}}^{\text{tol}} > 0$, $\delta^{\text{tol}} = \delta_{\text{PE}}^{\text{tol}} + \delta_{\text{tom}}^{\text{tol}}$, and an affine min-tradeoff function f such that $f(p) = f^{\text{PE}}(p) + f^{\text{tom}}(p) + \text{const.}$ Then for $a \in (1, 2)$, and a set of registers $O_1^n S_1^n E$ fulfilling the Markov chain (55),

$$\begin{aligned} & H_a^\uparrow(O_1^n | S_1^n E)_{\sigma^{\text{EAT}}|_{\Omega_{\text{EA}}}} \\ & \geq n f(p_0) - n \left(\delta^{\text{tol}} + \frac{(a-1) \ln 2}{2} V^2 \right) \\ & \quad - \frac{a}{a-1} \log \frac{1}{\Pr_{\sigma^{\text{EAT}}}[\Omega_{\text{EA}}]} - n(a-1)^2 K_a, \end{aligned} \quad (64)$$

where we have defined

$$\begin{aligned} V &= \sqrt{\text{Var}(f) + 2 + \log(2d_O^2 + 1)}, \\ K_a &= \frac{2^{(a-1)(2 \log d_O + \max(f) - \min_\Sigma(f))}}{6(2-a)^3 \ln 2} \\ & \quad \times \ln^3 \left(2^{2 \log d_O + \max(f) - \min_\Sigma(f)} + e^2 \right), \end{aligned} \quad (65) \quad (66)$$

where $\max(f) = \max_{p \in \mathcal{P}_C} f(p)$ and $\min_\Sigma(f) = \min_{p: \Sigma \neq \emptyset} f(p)$ and $\text{Var}(f)$ denotes the variance of f .

We will now use Proposition 2 to show the soundness of the hypothetical protocol. For that purpose, we need the following Lemma, which formalises the intuition that in order to upper bound the probability of aborting after tomography, we have to choose the corresponding tolerance parameter large enough.

Lemma 1 *Let $n \in \mathbb{N}$ and $\epsilon^{\text{tom}} \in (0, 1)$. Let us assume it holds*

$$\delta_{\text{tom}}^{\text{tol}} \geq 2 \sqrt{\log \left(\frac{n}{\epsilon^{\text{tom}}} \right) \sum_{i=1}^{16} \frac{\gamma'_i c_i'^2}{n}} + \frac{3D'}{n} \log \frac{n}{\epsilon^{\text{tom}}}, \quad (67)$$

where we have defined for, $i \in \{1, \dots, 16\}$,

$$\gamma'_i := \frac{\pi'_i (1 - \sum_{j=1}^i \pi'_j)}{1 - \sum_{j=1}^{i-1} \pi'_j}, \quad (68)$$

$$c'_i := h'_i - \frac{\sum_{j=i+1}^{17} h'_j \pi'_j}{1 - \sum_{j=1}^i \pi'_j}, \quad (69)$$

$$D' := \max_{i,j \in \{1, \dots, 17\}} |h'_i - h'_j|, \quad (70)$$

for $\pi'_{x'+1} = p_0(\perp, \perp, x')$, and $h'_{x'+1} = h_{\perp, \perp, x'}$, for $x' = 0, \dots, 15$, as well as $\pi'_{17} = 1 - \sum_{i=1}^{16} \pi'_i$, and $h'_{17} = 0$. Then it holds

$$\Pr_{\sigma^{\text{QKD}}}[\neg \Omega_{\text{tom}}] \leq \epsilon^{\text{tom}}. \quad (71)$$

Proof. As the tomography is performed entirely within Alice's lab, with no influence of Eve or the noisy channel, we can restrict our attention to an honest implementation of the protocol, i.e. $\Pr_{\sigma^{\text{QKD}}}[\neg \Omega_{\text{tom}}] = \Pr_{\text{hon}}[\neg \Omega_{\text{tom}}]$. Let us assume an honest application gives us the distribution

$$\begin{aligned} p_0(\perp, \perp, x') &= (1 - p^{\text{key}}) \tilde{p}_0(\perp, \perp, x') \\ &= p^{\text{tom}} p_0^{\text{tom}}(x'), \end{aligned} \quad (72)$$

for $x' \in \{0, \dots, 15\}$. Let us further assume that Alice and Bob observe some frequency distribution $\text{freq}_{c_1^n}$. Recalling the definition of the event Ω_{tom} , we note that it holds

$$\begin{aligned} & \Pr_{\text{hon}}[\Omega_{\text{tom}}] \\ & \geq \Pr_{\text{hon}} \left[\left| f^{\text{tom}}(\text{freq}_{c_1^n}) - f^{\text{tom}}(p_0) \right| \leq \delta_{\text{tom}}^{\text{tol}} \right]. \end{aligned} \quad (73)$$

An honest implementation of the protocol corresponds to n independent multinoulli trials with parameter p_0 . In order to provide lower bounds we can therefore make use of a concentration result provided by Proposition 2 of [58]. Namely, it holds with probability $1 - \epsilon^{\text{tom}}$ that

$$\begin{aligned} & \left| f^{\text{tom}}(\text{freq}_{c_1^n}) - f^{\text{tom}}(p_0) \right| = \left| (\hat{\pi}' - \pi')^T h' \right| \\ & \leq 2 \sqrt{\log \left(\frac{n}{\epsilon^{\text{tom}}} \right) \sum_{i=1}^{16} \frac{\gamma'_i c_i'^2}{n}} + \frac{3D'}{n} \log \frac{n}{\epsilon^{\text{tom}}}, \end{aligned} \quad (74)$$

where $\hat{\pi}'_{x'+1} = \text{freq}_{c_1^n}(\perp, \perp, x')$ for $x' = 0, \dots, 15$, as well as $\hat{\pi}'_{17} = 1 - \sum_{i=1}^{16} \hat{\pi}'_i$. Hence, if we choose the tolerance parameter $\delta_{\text{tom}}^{\text{tol}}$ fulfilling eq. (67), we obtain the desired bound

$$\Pr_{\text{hon}} \left[\left| f^{\text{tom}}(\text{freq}_{c_1^n}) - f^{\text{tom}}(p_0) \right| \leq \delta_{\text{tom}}^{\text{tol}} \right] \geq 1 - \epsilon^{\text{tom}}, \quad (75)$$

finishing the proof. $\blacksquare$

In order to show the soundness of the physical QKD protocol, which does not include tomography of Alice's marginal system, we also need the following Lemma, which relates the smooth min entropies of the physical and hypothetical protocol.

Lemma 2 *Let the smoothing parameters $\epsilon \in (0, 1 - \sqrt{2 \Pr_{\sigma^{\text{QKD}}}[\neg \Omega_{\text{tom}} | \Omega_{\text{EC}} \cap \Omega_{\text{PE}}]})$ and $\epsilon^{\text{phys}} \in (\epsilon + \sqrt{2 \Pr_{\sigma^{\text{QKD}}}[\neg \Omega_{\text{tom}} | \Omega_{\text{EC}} \cap \Omega_{\text{PE}}]}, 1)$. Then it holds*

$$\begin{aligned} & H_{\min}^{\epsilon^{\text{phys}}}(\hat{Z}_1^n | S_1^n \tilde{X}_1^n E)_{\sigma^{\text{phys, QKD}}|_{\Omega_{\text{NA}}^{\text{phys}}}} \\ & \geq H_{\min}^{\epsilon}(\hat{Z}_1^n | S_1^n \tilde{X}_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}. \end{aligned} \quad (76)$$

Proof. We begin by noting that Alice's tomography in the hypothetical protocol does not change the $\hat{Z}_1^n S_1^n \tilde{X}_1^n E$ subsystems of the final state, i.e.

$$\sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n H H' \tilde{X}_1^n \tilde{Y}_1^n E}^{\text{phys, QKD}} = \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n H H' \tilde{X}_1^n \tilde{Y}_1^n E}^{\text{QKD}} \quad (77)$$

This is by non-signalling. As the events Ω_{EC} and Ω_{PE} , defined by eqs. (38) and (41), respectively, depend only on the systems $H H' \hat{Z}_1^n \tilde{X}_1^n$ and $\tilde{Y}_1^n \tilde{X}_1^n$, it also holds

$$\begin{aligned} \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n H H' \tilde{X}_1^n \tilde{Y}_1^n E}^{\text{phys, QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}} \\ = \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n H H' \tilde{X}_1^n \tilde{Y}_1^n E}^{\text{QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}}. \end{aligned} \quad (78)$$

Using the triangle inequality, it follows that

$$\begin{aligned} & \left\| \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}} \cap \Omega_{\text{tom}}} - \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{phys, QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}} \right\|_1 \\ & \leq \left\| \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}} \cap \Omega_{\text{tom}}} - \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}} \right\|_1 \\ & \quad + \left\| \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}} - \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{phys, QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}} \right\|_1 \\ & \leq 2 \Pr_{\sigma^{\text{QKD}}}[\neg \Omega_{\text{tom}} | \Omega_{\text{EC}} \cap \Omega_{\text{PE}}]. \end{aligned} \quad (79)$$

Hence by eq. (4) it holds

$$\begin{aligned} & P \left(\sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}} \cap \Omega_{\text{tom}}}, \sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{phys, QKD}}|_{\Omega_{\text{EC}} \cap \Omega_{\text{PE}}} \right) \\ & \leq \sqrt{2 \Pr_{\sigma^{\text{QKD}}}[\neg \Omega_{\text{tom}} | \Omega_{\text{EC}} \cap \Omega_{\text{PE}}]}. \end{aligned} \quad (80)$$

Further, by the triangle inequality, the ϵ -ball (in terms of purified distance) around $\sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{NA}}}$ is contained in the ϵ^{phys} -ball around $\sigma_{\hat{Z}_1^n S_1^n \tilde{X}_1^n E}^{\text{QKD}}|_{\Omega_{\text{NA}}^{\text{phys}}}$, implying eq. (76). $\blacksquare$

We are now ready to prove the soundness physical QKD protocol, which is our main result.

Theorem 1 (Soundness) *Let $n \in \mathbb{N}$. Let $\epsilon_{\text{NA}}^{\text{phys}}, \epsilon^{\text{tom}}, \epsilon_{\text{EC}} \in (0, 1)$ such that $\epsilon^{\text{tom}} < \frac{1}{2} \epsilon_{\text{NA}}^{\text{phys}}$. Let $\epsilon \in \left(0, 1 - \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}\right)$, and define $\epsilon^{\text{phys}} = \epsilon + \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}$. Let $0 \leq p^{\text{key}}, p^{\text{PE}}, p^{\text{tom}} \leq 1$ such that $p^{\text{key}} + p^{\text{PE}} + p^{\text{tom}} = 1$. Let f be an affine min-tradeoff function of the form $f(p) = f^{\text{PE}}(p) + f^{\text{tom}}(p) + \text{const}$. Let p_0 be given by eqs. (31, 32). Let $\delta_{\text{PE}}^{\text{tol}} > 0$ and define*

$$\delta_{\text{tom}}^{\text{tol}} = 2 \sqrt{\log \left(\frac{n}{\epsilon^{\text{tom}}} \right) \sum_{i=1}^{16} \frac{\gamma'_i c_i^2}{n}} + \frac{3D'}{n} \log \frac{n}{\epsilon^{\text{tom}}}. \quad (81)$$

Let further $\Omega_{\text{PE}}, \Omega_{\text{tom}}, \Omega_{\text{EC}}$, and Ω_{NA} be defined by eqs. (38-42). Let leak_{EC} be the amount of information leaked during error correction. Then, if $\Pr_{\sigma^{\text{phys, QKD}}}[\Omega_{\text{PE}} \cap \Omega_{\text{EC}}] \geq \epsilon_{\text{NA}}^{\text{phys}}$, for any $a \in (1, 2)$, the physical QKD protocol provides an $3\epsilon^{\text{phys}} + \epsilon_{\text{EC}}^{\text{sound}}$ key at rate $r^{\text{phys}} = \ell/n$ with

$$\begin{aligned} & r^{\text{phys}}|_{\Omega_{\text{NA}}^{\text{phys}}} \\ & \geq f(p_0) - \delta_{\text{PE}}^{\text{tol}} - \delta_{\text{tom}}^{\text{tol}} - \frac{(a-1) \ln 2}{2} V^2 \\ & \quad - (a-1)^2 K_a - p^{\text{PE}} \log 5 \\ & \quad - (1 - p^{\text{key}}) \log(17(m+1)) \\ & \quad - \frac{1}{\sqrt{n}} \left[\sqrt{\frac{1}{2} \ln \frac{32}{\epsilon^2 (\epsilon_{\text{NA}}^{\text{phys}} - \epsilon^{\text{tom}})}} \log 5 \right. \\ & \quad \left. + \sqrt{\frac{1}{2} \ln \frac{512}{\epsilon^2 (\epsilon_{\text{NA}}^{\text{phys}} - \epsilon^{\text{tom}})}} \log(17(m+1)) \right] \\ & \quad - \frac{1}{n} \left[\frac{\Gamma(\epsilon/16)}{a-1} + \frac{a}{a-1} \log \frac{1}{\epsilon_{\text{NA}}^{\text{phys}} - \epsilon^{\text{tom}}} \right. \\ & \quad \left. + \text{leak}_{\text{EC}} + 2 \log \frac{1}{\epsilon^{\text{phys}}} + 2\Gamma(\epsilon/4) + 3\Gamma(\epsilon/16) \right]. \end{aligned} \quad (82)$$

Proof. We begin by lower bounding $H_{\min}^{\epsilon}(\hat{Z}_1^n | S_1^n \tilde{X}_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}}$ using eqs. (44 - 52). We then note that by eq. (59) it holds $H_{\min}^{\epsilon}(O_1^n | S_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{EA}}}} = H_{\min}^{\epsilon}(O_1^n | S_1^n E)_{\sigma^{\text{EAT}}|_{\Omega_{\text{EA}}}}$, allowing us to apply Proposition 2. Since $\Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{EC}} | \Omega_{\text{EA}}] \Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{EA}}] = \Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]$, the terms $\frac{a}{a-1} \log \frac{1}{\Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{EC}} | \Omega_{\text{EA}}]}$ and $\frac{a}{a-1} \log \frac{1}{\Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{EA}}]}$ in eqs (52) and (64), can be merged, resulting in a term that depends only on $\Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]$. Formally, we obtain

$$\begin{aligned} & H_{\min}^{\epsilon}(\hat{Z}_1^n | S_1^n \tilde{X}_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \\ & \geq n \left[f(p_0) - \delta_{\text{PE}}^{\text{tol}} - \delta_{\text{tom}}^{\text{tol}} - \frac{(a-1) \ln 2}{2} V^2 \right. \\ & \quad - (a-1)^2 K_a - p^{\text{PE}} \log 5 \\ & \quad - (1 - p^{\text{key}}) \log(17(m+1)) \left. \right] \\ & \quad - \sqrt{n} \left[\sqrt{\frac{1}{2} \ln \frac{32}{\epsilon^2 \Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]}} \log 5 \right. \\ & \quad \left. + \sqrt{\frac{1}{2} \ln \frac{512}{\epsilon^2 \Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]}} \log(17(m+1)) \right] \\ & \quad - \frac{\Gamma(\epsilon/16)}{a-1} - \frac{a}{a-1} \log \frac{1}{\Pr_{\sigma^{\text{QKD}}}[\Omega_{\text{NA}}]} \\ & \quad - 2\Gamma(\epsilon/4) - 3\Gamma(\epsilon/16). \end{aligned} \quad (83)$$

Further, it holds

$$\begin{aligned} & \Pr_{\sigma_{\text{QKD}}}[\Omega_{\text{NA}}] \\ &= \Pr_{\sigma_{\text{QKD}}}[\Omega_{\text{PE}} \cap \Omega_{\text{EC}}] \\ &\quad - \Pr_{\sigma_{\text{QKD}}}[\Omega_{\text{PE}} \cap \Omega_{\text{EC}} \cap \neg\Omega_{\text{tom}}] \\ &\geq \Pr_{\sigma_{\text{QKD}}}[\Omega_{\text{PE}} \cap \Omega_{\text{EC}}] - \Pr_{\sigma_{\text{QKD}}}[\neg\Omega_{\text{tom}}]. \end{aligned} \quad (84)$$

By eq. (81) and Lemma 1, we can bound $\Pr_{\sigma_{\text{QKD}}}[\neg\Omega_{\text{tom}}] \leq \epsilon^{\text{tom}}$. Further, by eq. (77), it holds that $\Pr_{\sigma^{\text{phys}}, \text{QKD}}[\Omega_{\text{PE}} \cap \Omega_{\text{EC}}] = \Pr_{\sigma_{\text{QKD}}}[\Omega_{\text{PE}} \cap \Omega_{\text{EC}}]$. Hence, by assumption, it holds

$$\Pr_{\sigma_{\text{QKD}}}[\Omega_{\text{NA}}] \geq \epsilon_{\text{NA}}^{\text{phys}} - \epsilon^{\text{tom}}. \quad (85)$$

Now we can apply Lemma 2 to show that for our choice of ϵ , and $\epsilon^{\text{phys}} = \epsilon + \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}$, it holds

$$\begin{aligned} & H_{\min}^{\epsilon^{\text{phys}}}(\hat{Z}_1^n | S_1^n \tilde{X}_1^n E)_{\sigma^{\text{phys}}, \text{QKD}}|_{\Omega_{\text{NA}}^{\text{phys}}} \\ & \geq H_{\min}^{\epsilon}(\hat{Z}_1^n | S_1^n \tilde{X}_1^n E)_{\sigma^{\text{QKD}}|_{\Omega_{\text{NA}}}} \end{aligned} \quad (86)$$

and apply Proposition 1, finishing the proof. ■

4.2 Completeness

In this section we show that the physical QKD protocol is complete, i.e. we provide a lower bound on the probability $\Pr_{\text{hon}}[\Omega_{\text{NA}}^{\text{phys}}]$ of an honest application not aborting.

Theorem 2 (Completeness) *Let $\epsilon_{\text{PE}}^c \in (0, 1)$. Let Ω_{PE} be as defined in eq. (38), with*

$$\delta_{\text{PE}}^{\text{tol}} = 2\sqrt{\log\left(\frac{n}{\epsilon_{\text{PE}}^c}\right) \sum_{i=1}^{4m} \frac{\gamma_i c_i^2}{n} + \frac{3D}{n} \log \frac{n}{\epsilon_{\text{PE}}^c}}, \quad (87)$$

where we have defined an ordering $(x, z, \perp) \rightarrow i$ by $(0, 0, \perp) \rightarrow 1, (0, 1, \perp) \rightarrow 2, \dots, (0, m-1, \perp) \rightarrow m, (1, 0, \perp) \rightarrow m+1, (1, 1, \perp) \rightarrow m+2, \dots, (3, m-1, \perp) \rightarrow 4m$. We then set $\pi_i = p_0(x, z, \perp)$, $\hat{\pi}_i = \text{freq}_{c_1^n}(x, z, \perp)$, $h_i = h_{x,z,\perp}$ for $i = 1, \dots, 4m$, as well as $\pi_{4m+1} := 1 - \sum_{i=1}^{4m} \pi_i$, $\hat{\pi}_{4m+1} := 1 - \sum_{i=1}^{4m} \hat{\pi}_i$ and $h_{4m+1} = 0$. Let us further define

$$\gamma_i := \frac{\pi_i \left(1 - \sum_{j=1}^i \pi_j\right)}{1 - \sum_{j=1}^{i-1} \pi_j}, \quad (88)$$

$$c_i := h_i - \frac{\sum_{j=i+1}^{4m+1} h_j \pi_j}{1 - \sum_{j=1}^i \pi_j}, \quad (89)$$

for $i = 1, \dots, 4m$, as well as $D := \max_{i,j \in \{1, \dots, 4m+1\}} |h_i - h_j|$. Let further $\epsilon_{\text{EC}}^c \in (0, 1)$ be a suitable completeness parameter for the error correction protocol used. Then the physical QKD protocol is $\epsilon_{\text{PE}}^c + \epsilon_{\text{EC}}^c$ -complete, i.e. $\Pr_{\text{hon}}[\Omega_{\text{NA}}^{\text{phys}}] \geq 1 - \epsilon_{\text{PE}}^c - \epsilon_{\text{EC}}^c$.

Proof. We consider the following honest implementation: We apply the physical QKD protocol, as described in Section 3, where the noisy channel $\mathcal{N}_{A_1^n \rightarrow B_1^n}$ is given by n iid uses of a phase-invariant Gaussian channel with transmittance η and excess noise ξ , however without an attack by Eve. By simulating this channel we obtain a distribution $p_0^{\text{sim}}(x, z)$, which depends on η and ξ and is given by eq. (156), and set $p_0(x, z, \perp) = p^{\text{PE}} p_0^{\text{sim}}(x, z)$ for $x \in \{0, \dots, 3\}$, $z \in \{0, \dots, m-1\}$. We note that the protocol can abort after parameter estimation or error correction. By the union bound it holds

$$1 - \Pr_{\text{hon}}[\Omega_{\text{NA}}^{\text{phys}}] \leq 1 - \Pr_{\text{hon}}[\Omega_{\text{PE}}] + 1 - \Pr_{\text{hon}}[\Omega_{\text{EC}}]. \quad (90)$$

We begin by considering abortion after parameter estimation. Let us assume an honest application gives us $p_0(x, z, \perp)$, for $x \in \{0, \dots, 3\}$, $z \in \{0, \dots, m-1\}$ according to eq. (31). Let us further assume that Alice and Bob observe some frequency distribution $\text{freq}_{c_1^n}$. Recalling the definition of the event Ω_{PE} , we note that it holds

$$\begin{aligned} & \Pr_{\text{hon}}[\Omega_{\text{PE}}] \\ & \geq \Pr_{\text{hon}} \left[\left| f^{\text{PE}}(\text{freq}_{c_1^n}) - f^{\text{PE}}(p_0) \right| \leq \delta_{\text{PE}}^{\text{tol}} \right]. \end{aligned} \quad (91)$$

An honest implementation of the protocol corresponds to n independent multinoulli trials with parameter p_0 . In order to provide lower bounds we can again make use of the concentration result provided by Proposition 2 of [58].

Let now $\epsilon_{\text{PE}}^c \in (0, 1)$. By Proposition 2 of [58], it then holds with probability $1 - \epsilon_{\text{PE}}^c$ that

$$\begin{aligned} & \left| f^{\text{PE}}(\text{freq}_{c_1^n}) - f^{\text{PE}}(p_0) \right| = \left| (\hat{\pi} - \pi)^T h \right| \\ & \leq 2\sqrt{\log\left(\frac{n}{\epsilon_{\text{PE}}^c}\right) \sum_{i=1}^{4m} \frac{\gamma_i c_i^2}{n} + \frac{3D}{n} \log \frac{n}{\epsilon_{\text{PE}}^c}}. \end{aligned} \quad (92)$$

Hence, if we choose the tolerance parameter $\delta_{\text{PE}}^{\text{tol}}$ as in eq. (87), we obtain the desired completeness bound

$$\Pr_{\text{hon}} \left[\left| f^{\text{PE}}(\text{freq}_{c_1^n}) - f^{\text{PE}}(p_0) \right| \leq \delta_{\text{PE}}^{\text{tol}} \right] \geq 1 - \epsilon_{\text{PE}}^c. \quad (93)$$

Finally, we have to consider the error correction. Let $\epsilon_{\text{EC}}^c \in (0, 1)$, such that $1 - \Pr_{\text{hon}}[\Omega_{\text{EC}}] \leq \epsilon_{\text{EC}}^c$, i.e. the error correction is assumed to abort with probability at most ϵ_{EC}^c , finishing the proof. ■

For a given empirical distribution p_0 and a suitable choice of a min-tradeoff function, Theorems 1 and 2 combined show the security of the finite round physical QKD protocol. In order to obtain the best finite size key rate, for given $n \in \mathbb{N}$, as well as some choice for the parameters $\epsilon_{\text{NA}}^{\text{phys}}, \epsilon^{\text{tom}}, \epsilon_{\text{EC}}, \epsilon_{\text{EC}}^c, \epsilon_{\text{PE}}^c \in (0, 1)$, such that $\epsilon^{\text{tom}} \ll \frac{1}{2}\epsilon_{\text{NA}}^{\text{phys}}$, as well as $\epsilon \in \left(0, 1 - \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}\right)$ and $\epsilon^{\text{phys}} = \epsilon + \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}$, we can set the tolerance parameters as in eqs. (81,87), and maximise the key rate given by (82) over probabilities $0 \leq p^{\text{key}}, p^{\text{PE}}, p^{\text{tom}} \leq 1$ such that $p^{\text{key}} + p^{\text{PE}} + p^{\text{tom}} = 1$, and over $a \in (1, 2)$. We note that in order to get a non-trivial result, we will have to choose ϵ_{PE}^c and ϵ_{EC}^c such that the success probability of the honest implementation meets the threshold $\epsilon_{\text{NA}}^{\text{phys}}$ used in Theorem 1, i.e. we need $1 - \epsilon_{\text{PE}}^c - \epsilon_{\text{EC}}^c > \epsilon_{\text{NA}}^{\text{phys}}$.

4.3 The Min-Tradeoff Function

The main task now is to find a min-tradeoff function f that provides a non-trivial bound for our protocol. As we will choose the number of key rounds to be significantly larger than the number of test rounds (i.e. rounds used for parameter estimation or tomography), it will be convenient to use the infrequent sampling framework introduced in [43], in which the statistical analysis only includes outputs in test rounds. To that purpose, we divide $\mathcal{M}_i^{\text{EAT}}$ into a key part, incorporating its action in key rounds; and a test part, incorporating its action in parameter estimation and tomography rounds, $\mathcal{M}_i^{\text{EAT}, \text{key}} : Q_{i-1} \rightarrow Q_i O_i S_i$ and $\mathcal{M}_i^{\text{EAT}, \text{test}} : Q_{i-1} \rightarrow Q_i O_i S_i C_i$, such that

$$\mathcal{M}_i^{\text{EAT}}(\cdot) = p^{\text{key}} \mathcal{M}_i^{\text{EAT}, \text{key}}(\cdot) \otimes |\perp\rangle \langle \perp|_{C_i} + (1 - p^{\text{key}}) \mathcal{M}_i^{\text{EAT}, \text{test}}(\cdot). \quad (94)$$

Let us now define a *crossover min-tradeoff function* [43] as a function $g : \mathcal{P}_{\tilde{\mathcal{C}}} \rightarrow \mathbb{R}$ such that for all $i = 1, \dots, n$ and $\tilde{p} \in \mathcal{P}_{\tilde{\mathcal{C}}}$ it holds

$$g(\tilde{p}) \leq \inf_{|\rho\rangle \in \tilde{\Sigma}_i(\tilde{p})} H(O_i | S_i \tilde{E})_{\rho^{\text{EAT}, i}}, \quad (95)$$

where $\tilde{E}$ can be chosen isomorphic to Q_{i-1} , and we have defined

$$\tilde{\Sigma}_i(\tilde{p}) = \left\{ |\rho\rangle_{Q_{i-1} \tilde{E}} \in \mathcal{H}_{Q_{i-1} \tilde{E}} : \langle c | \rho_{C_i}^{\text{EAT}, \text{test}, i} | c \rangle \equiv \tilde{p}(c) \right\}, \quad (96)$$

for states

$$\begin{aligned} \rho_{O_i S_i C_i Q_i \tilde{E}}^{\text{EAT}, \text{test}, i} &= \text{id}_{\tilde{E}} \otimes \mathcal{M}_i^{\text{EAT}, \text{test}}(\rho_{Q_{i-1} \tilde{E}}) \\ &= \text{id}_{Q_i \tilde{E}} \otimes \mathcal{M}_{A_i B_i \rightarrow O_i S_i C_i}^{\text{EAT}, \text{test}}(\rho_{A_i B_i Q_i \tilde{E}}). \end{aligned} \quad (97)$$

Further, it holds for all $i = 1, \dots, n$,

$$\inf_{|\rho\rangle \in \tilde{\Sigma}_i(\tilde{p})} H(O_i | S_i \tilde{E})_{\rho^{\text{EAT}, i}} \geq \inf_{\substack{|\rho\rangle \in \mathcal{H}_{\tilde{E}} \\ |\rho\rangle \in \tilde{\Sigma}_{\tilde{E}}(\tilde{p})}} H(O | S \hat{E})_{\rho^{\text{EAT}}}, \quad (98)$$

where we have defined the states $\rho_{OSC\hat{E}}^{\text{EAT}} = \text{id}_{\hat{E}} \otimes \mathcal{M}_{AB \rightarrow OSC}^{\text{EAT}}(\rho_{AB\hat{E}})$ and $\rho_{OSC\hat{E}}^{\text{EAT}, \text{test}} = \text{id}_{\hat{E}} \otimes \mathcal{M}_{AB \rightarrow OSC}^{\text{EAT}, \text{test}}(\rho_{AB\hat{E}})$, as well as the set $\tilde{\Sigma}_{\tilde{E}}(\tilde{p}) = \left\{ |\rho\rangle_{AB\hat{E}} \in \mathcal{H}_{AB\hat{E}} : \langle c | \rho_C^{\text{EAT}, \text{test}} | c \rangle \equiv \tilde{p}(c) \right\}$. We can therefore relax the problem to finding a function $g : \mathcal{P}_{\tilde{\mathcal{C}}} \rightarrow \mathbb{R}$ such that

$$g(\tilde{p}) \leq \inf_{\substack{|\rho\rangle \in \mathcal{H}_{\tilde{E}} \\ |\rho\rangle \in \tilde{\Sigma}_{\tilde{E}}(\tilde{p})}} H(O | S \hat{E})_{\rho^{\text{EAT}}}. \quad (99)$$

According to Lemma V.5 of [43], we translate our crossover min-tradeoff function g into a min-tradeoff function f via the definition

$$f(\delta_c) = \max(g) + \frac{g(\delta_c) - \max(g)}{1 - p^{\text{key}}} \quad \forall c \in \tilde{\mathcal{C}}, \quad (100)$$

$$f(\delta_{(\perp, \perp, \perp)}) = \max(g), \quad (101)$$

where δ_c denotes the distribution that equals 1 for c and 0 everywhere else. Further, $\max(g) = \max_{\tilde{p} \in \mathcal{P}_{\tilde{\mathcal{C}}}} g(\tilde{p})$ and $\min(g) = \min_{\tilde{p} \in \mathcal{P}_{\tilde{\mathcal{C}}}} g(\tilde{p})$. If p is of the form $p(c) = (1 - p^{\text{key}})\tilde{p}(c)$ for $c \in \tilde{\mathcal{C}}$ and $p(\perp, \perp, \perp) = p^{\text{key}}$, it holds $f((1 - p^{\text{key}})\tilde{p}) = g(\tilde{p})$ for all $\tilde{p} \in \mathcal{P}_{\tilde{\mathcal{C}}}$. Further it holds

$$\max(f) = \max(g), \quad (102)$$

$$\min_{\Sigma}(f) \geq \min(g), \quad (103)$$

$$0 \leq \text{Var}(f) \leq \frac{1}{1 - p^{\text{key}}} (\max(g) - \min(g))^2. \quad (104)$$

Hence we can upper bound the expressions in eqs. (65,66) by

$$\begin{aligned}
V &\leq \tilde{V} = \sqrt{\frac{1}{1-p^{\text{key}}} (\max(g) - \min(g))^2 + 2} \\
&\quad + \log(2d_O^2 + 1), \quad (105) \\
K_a &\leq \tilde{K}_a = \frac{2^{(a-1)(2\log d_O + \max(g) - \min(g))}}{6(2-a)^3 \ln 2} \\
&\quad \times \ln^3 \left(2^{2\log d_O + \max(g) - \min(g)} + e^2 \right). \quad (106)
\end{aligned}$$

In what follows, we provide a crossover minimum tradeoff function for our choice of EAT channels $\{\mathcal{M}_i^{\text{EAT}}\}_{i=1}^n$, lower bounding the r.h.s. of (95). We begin by noting that, by the chain rule for the von Neumann entropy [59], it holds

$$\begin{aligned}
&\inf_{\substack{\mathcal{H}_{\hat{E}} \simeq \mathcal{H}_{Q_1} \\ |\rho\rangle \in \Sigma_{\hat{E}}(\tilde{p})}} H(O|\hat{S}\hat{E})_{\rho^{\text{EAT}}} \\
&\geq \inf_{\substack{\mathcal{H}_{\hat{E}} \simeq \mathcal{H}_{Q_1} \\ |\rho\rangle \in \Sigma_{\hat{E}}(\tilde{p})}} \left(H(\hat{Z}|\hat{S}\hat{E})_{\rho^{\text{EAT}}} + H(\tilde{Z}\tilde{X}X'|\hat{Z}\hat{S}\hat{E})_{\rho^{\text{EAT}}} \right) \quad (107)
\end{aligned}$$

$$\geq \inf_{\substack{\mathcal{H}_{\hat{E}} \simeq \mathcal{H}_{Q_1} \\ |\rho\rangle \in \Sigma_{\hat{E}}(\tilde{p})}} H(\hat{Z}|\hat{S}\hat{E})_{\rho^{\text{EAT}}}, \quad (108)$$

where we have used that, as $\tilde{Z}\tilde{X}X'$ is classical, there cannot be any entanglement across the $\tilde{Z}\tilde{X}X' : \hat{Z}\hat{S}\hat{E}$ partition, hence the second term in (107) has to be non-negative. Let us now define $g : \mathcal{P}_{\tilde{C}} \rightarrow \mathbb{R}$,

$$g(\tilde{p}) := \inf_{\substack{\mathcal{H}_{\hat{E}} \simeq \mathcal{H}_{Q_1} \\ |\rho\rangle \in \Sigma_{\hat{E}}(\tilde{p})}} H(\hat{Z}|\hat{S}\hat{E})_{\rho^{\text{EAT}}}, \quad (109)$$

which can serve as a crossover min-tradeoff function for EAT channels $\{\mathcal{M}_i^{\text{EAT}}\}$. In order to obtain an efficiently numerically computable crossover min-tradeoff function, we now make use of the framework presented in [48] to remove the dependency on Eve's subsystem.

4.3.1 Removing the dependence on the $\hat{E}$ subsystem

The idea is to consider a coherent version of a round of the protocol leading to Bob's raw key $\hat{Z}$. Namely, Alice and Bob's measurements are performed in a coherent fashion, i.e. by means of isometries acting on the system to be measured and adding a quantum register containing the quantum information which, once dephased,

will provide the measurement result, but not yet dephasing it. Alice and Bob then publicly announce partial information about their measurement outcomes, while keeping part of the information stored coherently. From that information they decide whether they use the round for key generation, parameter estimation or tomography of Alice's part. If they use the round for key generation, in the case of reverse reconciliation, Bob applies a key map to his coherently stored measurement outcomes, which provides a coherent key register. The key can then be obtained by means of a so-called pinching operation, i.e. a measurement that dephases the key register.

As all steps of the protocol before the pinching are performed coherently, we can express the outcome as a pure state, allowing us to apply Theorem 1 in [60], which removes the dependence on the $\hat{E}$ subsystem. In order to formulate our result, we need to introduce the CP map $\mathcal{G} : AB \rightarrow AB\hat{Z}$ that describes the coherent version of the protocol. This map is given by a single Kraus operator

$$G = \mathbb{1}_A \otimes \sum_{z=0}^3 \sqrt{R_B^z} \otimes |z\rangle_{\hat{Z}}, \quad (110)$$

where we have defined the region operators

$$R_B^z = \frac{1}{\pi} \int_0^\infty \int_{\frac{\pi}{4}(2z-1)}^{\frac{\pi}{4}(2z+1)} \gamma |\gamma e^{i\theta}\rangle \langle \gamma e^{i\theta}| d\theta d\gamma, \quad (111)$$

for $z \in \{0, 1, 2, 3\}$.

Furthermore, we define the pinching operation $\mathcal{Z} : \hat{Z} \rightarrow \hat{Z}$, defined by Kraus operators

$$Z_j = |j\rangle \langle j|_{\hat{Z}} \otimes \mathbb{1}, \quad (112)$$

for $j \in \{0, 1, 2, 3\}$, and the identity is extended to all registers other than $\hat{Z}$. It then holds

Lemma 3 *The crossover min-tradeoff function defined by eq. (109) can be reformulated as follows*

$$\begin{aligned}
g(\tilde{p}) &= \inf_{\substack{\mathcal{H}_{\hat{E}} \simeq \mathcal{H}_{Q_1} \\ |\rho\rangle \in \Sigma_{\hat{E}}(\tilde{p})}} H(\hat{Z}|\hat{S}\hat{E})_{\rho^{\text{EAT}}} \\
&= p^{\text{key}} \inf_{\rho \in \Sigma(\tilde{p})} D(\mathcal{G}(\rho_{AB}) || \mathcal{Z}(\mathcal{G}(\rho_{AB}))), \quad (113)
\end{aligned}$$

where we have defined the set $\Sigma(\tilde{p}) = \{\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB}) : \langle c | \mathcal{M}^{\text{EAT, test}}(\rho)_C | c \rangle \equiv \tilde{p}(c)\}$, which is independent of the reference system.

The proof of Lemma 3 goes along the line of the discussion in [30] and can be found in Appendix A. Let us note that, by definition of the protocol, it holds $\dim(A) = 4$, but the dimension of Bob's states can be infinite. We address this problem by invoking again the bounded-energy assumption, implying that the solution to (113) can be arbitrarily well approximated by taking a large enough finite dimension. We then take d_B to be arbitrary but finite. Under said assumption, the set $\Sigma(\tilde{p})$ is compact and, as the objective is continuous [48], a minimum is attained in eq. (113). Following [61] we can show

Lemma 4 For a given $0 < p^{\text{key}} \leq 1$,

$$g(\tilde{p}) = p^{\text{key}} \min_{\rho \in \Sigma(\tilde{p})} D(\mathcal{G}(\rho_{AB}) || \mathcal{Z}(\mathcal{G}(\rho_{AB}))) \quad (114)$$

is a convex function on $\mathcal{P}_{\tilde{c}}$.

The proof can be found in Appendix B.

4.3.2 Finding an affine crossover min-tradeoff function

We note that, for any given distribution $\tilde{p} \in \mathcal{P}_{\tilde{c}}$, eq. (114) is a convex optimisation problem with semidefinite constraints. As the objective is not affine, however, it is not a semidefinite program (SDP). Also, the dependence of g on the distribution $\tilde{p}$ is hidden in the constraints. We will now follow the steps taken in [48] and perform a first order Taylor expansion (around some state $\tilde{\rho}_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$) providing a lower bound on the optimisation problem in eq. (114). The resulting expression contains an SDP with a linear objective.

We then consider the dual of the SDP, which for any dual feasible point provides an affine lower bound on the original SDP. By the nature of duality, which roughly speaking incorporates the constraints into the objective, the objective of the dual problem will explicitly depend on $\tilde{p}$ in an affine way, as will the entire expression lower bounding the optimising problem in eq. (114). Thus, for any given state $\tilde{\rho}_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$, as well as any dual feasible point, we can obtain an affine crossover min-tradeoff function.

To begin with, let us explicitly consider the optimisation problem in eq. (114). Let $m = 4\Delta/\delta + 4$ represent the total number of modules

in Bob's discretisation. For a probability distribution $\tilde{p} \in \mathcal{P}_{\tilde{c}}$, the optimisation takes the form

$$\begin{aligned} \min_{\rho_{AB}} & D(\mathcal{G}(\rho_{AB}) || \mathcal{Z}(\mathcal{G}(\rho_{AB}))) \\ \text{s.t. } & \rho_{AB} \geq 0, \text{Tr}[\rho_{AB}] = 1, \\ & \forall x \in \{0, 1, 2, 3\}, \forall z \in \{0, \dots, m-1\} : \\ & \tilde{p}^{\text{PE}} \text{Tr} \left[\left(|x\rangle \langle x|_A \otimes \tilde{R}_B^z \right) \rho_{AB} \right] = \tilde{p}(x, z, \perp), \\ & \forall x' \in \{0, \dots, 15\} : \\ & \tilde{p}^{\text{tom}} \text{Tr} [\Gamma_{x'} \rho_A] = \tilde{p}(\perp, \perp, x'), \end{aligned} \quad (115)$$

where we have defined $\tilde{p}^{\text{PE}} = \frac{p^{\text{PE}}}{1-p^{\text{key}}}$ and $\tilde{p}^{\text{tom}} = 1 - \tilde{p}^{\text{PE}}$. Further, $\tilde{R}^z$ are region operators defined in analogy to (111), but with the discretisation used for parameter estimation given by eq. (28). Regarding the constraints, the region operators related to parameter estimation add up to the identity matrix, so that there is no need to impose the constraint $\text{Tr}[\rho_{AB}] = 1$. We now closely follow [48] to lower bound eq. (115). For brevity, let us define $r(\rho) := D(\mathcal{G}(\rho) || \mathcal{Z}(\mathcal{G}(\rho)))$. By the properties of the pinching quantum channel, this expression can be rewritten without loss of generality in terms of von Neumann entropies

$$r(\rho) = H(\mathcal{Z}(\mathcal{G}(\rho))) - H(\mathcal{G}(\rho)). \quad (116)$$

Using the methodology of [62], it is possible to apply here a facial reduction to reformulate the maps $\mathcal{Z}$ and $\mathcal{G}$ into maps which are strictly positive definite; this does not only assure that the new objective function is differentiable for any $\rho > 0$, but also reduces the dimension of both maps which simplifies the subsequent numerical analysis. This process can be seen as a unitary transformation, such that ¹

$$\mathcal{G}(\rho) = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} \tilde{\mathcal{G}}(\rho) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U^\dagger \\ V^\dagger \end{bmatrix}, \quad (117)$$

where $\tilde{\mathcal{G}}(\rho) > 0$ for $\rho > 0$. A similar procedure follows for $\mathcal{Z}(\mathcal{G}(\rho))$, resulting in a new map $\tilde{\mathcal{Z}}(\rho) > 0$. Hence, by taking advantage of the fact that the von Neumann entropy is invariant under unitary transformations, we arrive at a simpler objective function

$$r(\rho) = H(\tilde{\mathcal{Z}}(\rho)) - H(\tilde{\mathcal{G}}(\rho)). \quad (118)$$

¹For the numerical implementation, this decomposition can be obtained in MATLAB by using the function `rank`.

With the maps $\tilde{\mathcal{Z}}, \tilde{\mathcal{G}}$ the matrix gradient $\nabla r(\rho)$ is now given by

$$\nabla r(\rho)^T = [\tilde{\mathcal{G}}^\dagger(\log \tilde{\mathcal{G}}(\rho)) + \tilde{\mathcal{G}}^\dagger(\mathbb{1})] - [\tilde{\mathcal{Z}}^\dagger(\log \tilde{\mathcal{Z}}(\rho)) + \tilde{\mathcal{Z}}^\dagger(\mathbb{1})]. \quad (119)$$

Let now $\tilde{p} \in \mathcal{P}_{\tilde{\mathcal{C}}}$ and $\rho_{\tilde{p}}^* \in \Sigma(\tilde{p})$ be the minimiser of (115). For any $\tilde{\rho} \in \mathcal{D}(\mathcal{H}_{AB})$, it then holds

$$\begin{aligned} \frac{g(\tilde{p})}{p^{\text{key}}} &= r(\rho_{\tilde{p}}^*) \\ &\geq r(\tilde{\rho}) + \text{Tr}[(\rho_{\tilde{p}}^* - \tilde{\rho})^T \nabla r(\tilde{\rho})] \end{aligned} \quad (120)$$

$$\begin{aligned} &\geq r(\tilde{\rho}) - \text{Tr}[\tilde{\rho}^T \nabla r(\tilde{\rho})] \\ &\quad + \min_{\sigma \in \Sigma(\tilde{p})} \text{Tr}[\sigma^T \nabla r(\tilde{\rho})], \end{aligned} \quad (121)$$

where the first inequality is due to the fact that r is a convex, differentiable function over the convex set $\mathcal{D}(\mathcal{H}_{AB})$, hence it can be lower bounded by its first order Taylor expansion at $\tilde{\rho}$ (see e.g. [63] p.69), and the second inequality is due to the fact that $\rho_{\tilde{p}}^* \in \Sigma(\tilde{p})$. For any $\tilde{\rho} \in \mathcal{D}(\mathcal{H}_{AB})$ and $\tilde{p}$, the optimisation problem in eq. (121) is an SDP in standard form, explicitly given by

$$\min_{\sigma_{AB}} \text{Tr}[\sigma^T \nabla r(\tilde{\rho})] \quad (122)$$

s.t. $\sigma_{AB} \geq 0$,

$\forall x \in \{0, 1, 2, 3\}, \forall z \in \{0, \dots, m-1\}$:

$$\tilde{p}^{\text{PE}} \text{Tr}[(|x\rangle\langle x|_A \otimes \tilde{R}_B^z) \sigma_{AB}] = \tilde{p}(x, z, \perp),$$

$\forall x' \in \{0, \dots, 15\}$:

$$\tilde{p}^{\text{tom}} \text{Tr}[\Gamma_{x'} \sigma_A] = \tilde{p}(\perp, \perp, x').$$

The dual problem of the SDP (122) takes the form

$$\max_{\vec{v} \in \Sigma_{\tilde{p}}^*} \ell_{\tilde{p}}(\vec{v}), \quad (123)$$

where the dual objective is given by

$$\ell_{\tilde{p}}(\vec{v}) = \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{xz} \frac{\tilde{p}(x, z, \perp)}{\tilde{p}^{\text{PE}}} + \sum_{x'=0}^{15} \nu_{x'} \frac{\tilde{p}(\perp, \perp, x')}{\tilde{p}^{\text{tom}}}, \quad (124)$$

which is affine with respect to $\tilde{p}$. Further, the set $\Sigma_{\tilde{p}}^*$ is defined as

$$\begin{aligned} \Sigma_{\tilde{p}}^* &= \left\{ \vec{v} \in \mathbb{R}^{4m+16} : \right. \\ &\quad \nabla r(\rho) - \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{xz} (|x\rangle\langle x|_A \otimes \tilde{R}_B^z)^T \\ &\quad \left. - \sum_{x'=0}^{15} \nu_{x'} \Gamma_{x'}^T \geq 0 \right\} \end{aligned} \quad (125)$$

which is independent of $\tilde{p}$. By weak duality it then holds

$$\begin{aligned} g(\tilde{p}) &= p^{\text{key}} r(\rho_{\tilde{p}}^*) \\ &\geq p^{\text{key}} \left(r(\tilde{\rho}) - \text{Tr}[\tilde{\rho}^T \nabla r(\tilde{\rho})] + \max_{\vec{v} \in \Sigma_{\tilde{p}}^*} \ell_{\tilde{p}}(\vec{v}) \right) \end{aligned} \quad (126)$$

$$\geq p^{\text{key}} \left(r(\tilde{\rho}) - \text{Tr}[\tilde{\rho}^T \nabla r(\tilde{\rho})] + \ell_{\tilde{p}}(\vec{v}) \right) \quad (127)$$

$$=: \tilde{g}_{\vec{v}, \tilde{p}}(\tilde{p}) \quad (128)$$

for any $\tilde{\rho} \in \mathcal{D}(\mathcal{H}_{AB})$ and any $\vec{v} \in \Sigma_{\tilde{p}}^*$. We note that for any such choice of $\tilde{\rho}, \vec{v}$, the function $\tilde{g}_{\vec{v}, \tilde{p}} : \mathcal{P}_{\tilde{\mathcal{C}}} \rightarrow \mathbb{R}$ is an affine crossover min-tradeoff function.

4.3.3 Optimisation of the crossover min-tradeoff function

In this section we describe how we can numerically obtain almost optimal, i.e. optimal up to numerical imprecision, choices for our parameters $\tilde{\rho}$ and $\vec{v}$ in the crossover min-tradeoff function (128), for a given distribution $\tilde{p}_0 \in \mathcal{P}_{\tilde{\mathcal{C}}}$. The distribution will be of the form $\tilde{p}_0(x, z, \perp) = \tilde{p}^{\text{PE}} p_0^{\text{sim}}(x, z)$, for all $x \in \{0, 1, 2, 3\}$ and $z \in \{0, \dots, m-1\}$, where $p_0^{\text{sim}}(x, z)$ is a distribution obtained by simulating an honest implementation of the physical QKD protocol. Similarly, $\tilde{p}_0(\perp, \perp, x') = \tilde{p}^{\text{tom}} p_0^{\text{tom}}(x')$ for all $x' \in \{0, \dots, 15\}$, where $p_0^{\text{tom}}(x')$ is the distribution obtained in the hypothetical tomography. For the explicit form of $p_0^{\text{sim}}(x, z)$ and $p_0^{\text{tom}}(x')$, given by a simulation of the hypothetical QKD protocol, see Section 5.

We note that whereas the choices for $\tilde{\rho}$ and $\vec{v}$ will only be optimal up to numerical imprecision, it is possible to analytically confirm their feasibility, i.e. that $\tilde{\rho} \in \mathcal{D}(\mathcal{H}_{AB})$ and $\vec{v} \in \Sigma_{\tilde{p}}^*$. Thus we can analytically verify that the corresponding function $\tilde{g}_{\vec{v}, \tilde{p}}$ is indeed a valid crossover min-tradeoff function.

Our numerical method now works as follows: We begin with some $\tilde{\rho}^{(0)} \in \mathcal{D}(\mathcal{H}_{AB})$ and, for $i = 1, \dots, n^{\text{iter}}$, where $n^{\text{iter}} \in \mathbb{N}$, iteratively compute

$$\begin{aligned}
\Delta\tilde{\rho}^{(i)} &= \arg \min_{\sigma_{AB}} \text{Tr} \left[\sigma^T \nabla r(\tilde{\rho}^{(i-1)}) \right] \\
\text{s.t. } \sigma_{AB} &\geq 0, \\
\forall x \in \{0, 1, 2, 3\}, \forall z \in \{0, \dots, m-1\} : \\
\text{Tr} \left[(|x\rangle\langle x|_A \otimes \tilde{R}_B^z) \sigma_{AB} \right] &= p_0^{\text{sim}}(x, z), \\
\forall x' \in \{0, \dots, 15\} : \\
\text{Tr} [\Gamma_{x'} \sigma_A] &= p_0^{\text{tom}}(x').
\end{aligned} \tag{129}$$

Once this SDP is solved and $\Delta\tilde{\rho}^{(i)}$ is known, the value of the relative entropy is minimized according to

$$\min_{\kappa \in (0,1)} r(\tilde{\rho}^{(i-1)} + \kappa \Delta\tilde{\rho}^{(i)}). \tag{130}$$

Such minimization can be computed in MATLAB with the function `fminbnd`. Then, we set a new density matrix $\tilde{\rho}^{(i)} = \tilde{\rho}^{(i-1)} + \kappa^* \Delta\tilde{\rho}^{(i)}$, with the optimal coefficient κ^* , and repeat the optimisation (129). After n^{iter} we set $\tilde{\rho}_0 = \tilde{\rho}^{(n^{\text{iter}})}$.

The numerical computation of the dual of (122) requires to take into account the difference in the numerical representation of the states and operators with respect to their analytical values, which leads to a violation of the constraints due to the computational limitations of the computers. According to Theorem 3 of [48], this error may be taken into account by introducing a new parameter ε' , which takes the absolute value of the maximal such error, and expands the feasible set to provide a lower bound while preserving the reliability of the approach. With this methodology, the dual takes the form [63],

$$\max_{(\vec{\nu}, \vec{\mu}) \in \tilde{\Sigma}_{\tilde{\rho}_0}^*} \ell_{\tilde{\rho}_0, \varepsilon'}^0(\vec{\nu}, \vec{\mu}), \tag{131}$$

where the dual objective is given by

$$\begin{aligned}
\ell_{\tilde{\rho}_0, \varepsilon'}^0(\vec{\nu}, \vec{\mu}) &= \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{xz} p_0^{\text{sim}}(x, z) \\
&+ \sum_{x'=0}^{15} \nu'_{x'} p_0^{\text{tom}}(x') \\
&- \varepsilon' \sum_{z'=1}^{4m+16} \mu_{z'},
\end{aligned} \tag{132}$$

with the set $\tilde{\Sigma}_{\tilde{\rho}_0}^*$ defined as

$$\begin{aligned}
\tilde{\Sigma}_{\tilde{\rho}_0}^* &= \left\{ (\vec{\nu}, \vec{\mu}) \in (\mathbb{R}^{4m+16}, \mathbb{R}^{4m+16}) : -\vec{\mu} \leq \vec{\nu} \leq \vec{\mu}, \right. \\
&\quad \nabla r(\tilde{\rho}_0) - \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{xz} \left(|x\rangle\langle x|_A \otimes \tilde{R}_B^z \right)^T \\
&\quad \left. - \sum_{x'=0}^{15} \nu'_{x'} \Gamma_{x'}^T \geq 0 \right\}.
\end{aligned} \tag{133}$$

From this maximization, as well as a fixed value ε' taken according to the maximal numerical error at the constraints, we obtain $\vec{\nu}_0$, and note that $\vec{\nu}_0 \in \tilde{\Sigma}_{\tilde{\rho}_0}^*$. This allows us to define our crossover min-tradeoff function as

$$\begin{aligned}
\tilde{g}_0(\tilde{p}) &:= \tilde{g}_{\vec{\nu}_0, \tilde{\rho}_0}(\tilde{p}) \\
&= p^{\text{key}} \left(r(\tilde{\rho}_0) - \text{Tr} \left[\tilde{\rho}_0^T \nabla r(\tilde{\rho}_0) \right] + \ell_{\tilde{p}}(\vec{\nu}_0) \right) \\
&= p^{\text{key}} \left(G_0 + \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{0,xz} \frac{\tilde{p}(x, z, \perp)}{\tilde{p}^{\text{PE}}} \right. \\
&\quad \left. + \sum_{x'=0}^{15} \nu'_{0,x'} \frac{\tilde{p}(\perp, \perp, x')}{\tilde{p}^{\text{tom}}} \right),
\end{aligned} \tag{134}$$

with a constant

$$G_0 := r(\tilde{\rho}_0) - \text{Tr} \left[\tilde{\rho}_0^T \nabla r(\tilde{\rho}_0) \right]. \tag{136}$$

In order to compute the higher order terms of the EAT, we need to find $\max(\tilde{g}_0) = \max_{\tilde{p} \in \mathcal{P}_{\tilde{c}}} \tilde{g}_0(\tilde{p})$ and $\min(\tilde{g}_0) = \min_{\tilde{p} \in \mathcal{P}_{\tilde{c}}} \tilde{g}_0(\tilde{p})$. We note that, as $\mathcal{P}_{\tilde{c}}$ is convex and $\tilde{g}_0$ is affine, we can restrict to the extreme points of $\mathcal{P}_{\tilde{c}}$. Namely we get

$$\max(\tilde{g}_0) = p^{\text{key}} G_0 + p^{\text{key}} \max(\nu_0), \tag{137}$$

$$\min(\tilde{g}_0) = p^{\text{key}} G_0 + p^{\text{key}} \min(\nu_0), \tag{138}$$

where we have defined

$$\max(\nu_0) := \max \left(\left\{ \frac{\nu_{0,xz}}{\tilde{p}^{\text{PE}}} \right\}_{(x,z)=(0,0)}^{(3,m-1)} \cup \left\{ \frac{\nu'_{0,x'}}{\tilde{p}^{\text{tom}}} \right\}_{x'=0}^{15} \right) \tag{139}$$

$$\min(\nu_0) := \min \left(\left\{ \frac{\nu_{0,xz}}{\tilde{p}^{\text{PE}}} \right\}_{(x,z)=(0,0)}^{(3,m-1)} \cup \left\{ \frac{\nu'_{0,x'}}{\tilde{p}^{\text{tom}}} \right\}_{x'=0}^{15} \right) \tag{140}$$

In the case where the minimisers are non-positive and the maximisers are non-negative, we can upper bound

$$\max(\tilde{g}_0) - \min(\tilde{g}_0) \leq \max(\nu_0) - \min(\nu_0), \tag{141}$$

which is independent of p^{key} . Finally, we can introduce the min-tradeoff function induced by our crossover min-tradeoff function $\tilde{g}_0$, given by eq. (135), via eqs. (100,101).

$$f(p) = \sum_{c \in \tilde{\mathcal{C}}} p(c) \left(\max(\tilde{g}_0) + \frac{\tilde{g}_0(\delta_c) - \max(\tilde{g}_0)}{1 - p^{\text{key}}} \right) + p(\perp, \perp, \perp) \max(\tilde{g}_0) \quad (142)$$

$$= \max(\tilde{g}_0) + \sum_{c \in \tilde{\mathcal{C}}} \frac{p(c) (\tilde{g}_0(\delta_c) - \max(\tilde{g}_0))}{1 - p^{\text{key}}} \quad (143)$$

$$= p^{\text{key}} (G_0 + \max(\nu_0)) + \sum_{x=0}^3 \sum_{z=0}^{m-1} p^{\text{key}} \frac{\nu_{0,xz}/\tilde{p}^{\text{PE}} - \max(\nu_0)}{1 - p^{\text{key}}} p(x, z, \perp) + \sum_{x'=0}^{15} p^{\text{key}} \frac{\nu'_{0,x'}/\tilde{p}^{\text{tom}} - \max(\nu_0)}{1 - p^{\text{key}}} p(\perp, \perp, x'). \quad (144)$$

Let us now observe that we can split the min-tradeoff function according to a constant term (since it does not depend on the probabilities) and the previously-defined functions (34-35) for parameter estimation and tomography

$$f(p) = \text{const} + f^{\text{PE}}(p) + f^{\text{tom}}(p), \quad (145)$$

with

$$f^{\text{PE}}(p) = \sum_{x=0}^3 \sum_{z=0}^{m-1} h_{x,z,\perp} p(x, z, \perp), \quad (146)$$

$$f^{\text{tom}}(p) = \sum_{x'=0}^{15} h_{\perp,\perp,x'} p(\perp, \perp, x'). \quad (147)$$

The affine coefficients of the functions, given in terms of the crossover min-tradeoff function (144), provide the means for calculating the statistical deviations $\delta_{\text{PE}}^{\text{tol}}$ and $\delta_{\text{tom}}^{\text{tol}}$. Thanks to the affine structure of 145, we can use $f(p)$ as a min-tradeoff function in Theorem 1. When applying such Theorem, we evaluate the min-tradeoff function in our simulated honest distribution, $f(p_0)$. However, we use the properties of the distribution p_0 to reformulate our function in more convenient shape

$$f(p_0) = p^{\text{key}} G_0 + \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{0,xz} p_0^{\text{sim}}(x, z) + \sum_{x'=0}^{15} \nu'_{0,x'} p_0^{\text{tom}}(x') \quad (148)$$

$$= \tilde{g}_0(\tilde{p}_0). \quad (149)$$

4.4 Asymptotic Rates

With our choice of a min-tradeoff function $f(p)$, we can now compute the asymptotic key rate in Theorem 1, and show that we can achieve soundness and completeness in the asymptotic limit.

Let $n \in \mathbb{N}$. We begin by noting that, for fixed $m \in \mathbb{N}$, our numerically obtained values $\nu_{0,xz}$, for $x = 0, \dots, 3, z = 0, \dots, m-1$, and $\nu'_{0,x'}$, for $x' = 0, \dots, 15$, are constant, i.e. independent of n . Let us consider some fixed values for the parameters $\epsilon_{\text{NA}}^{\text{phys}}, \epsilon^{\text{tom}}, \epsilon_{\text{EC}}, \epsilon_{\text{EC}}^c, \epsilon_{\text{PE}}^c \in (0, 1)$, such that $\epsilon^{\text{tom}} < \frac{1}{2} \epsilon_{\text{NA}}^{\text{phys}}$, as well as $\epsilon \in \left(0, 1 - \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}\right)$, $\epsilon^{\text{phys}} = \epsilon + \sqrt{2\epsilon^{\text{tom}}/\epsilon_{\text{NA}}^{\text{phys}}}$. Let us also consider some constant $0 \leq \tilde{p}^{\text{PE}} \leq 1$, and $\tilde{p}^{\text{tom}} = 1 - \tilde{p}^{\text{PE}}$.

Since $n \rightarrow \infty$, we can select any scaling such that all rounds tend asymptotically to be spent on key generation, and the finite size effects are reduced. For simplicity, let us then take $a = 1 + n^{-3/4}$, as well as $p^{\text{key}} = 1 - n^{-\frac{1}{2}}$ and $p^{\text{PE}} = \tilde{p}^{\text{PE}} n^{-\frac{1}{2}}$, implying $p^{\text{tom}} = \tilde{p}^{\text{tom}} n^{-\frac{1}{2}}$. It then holds that $h_{x,z,\perp} = \mathcal{O}(n^{\frac{1}{2}})$ and $h_{\perp,\perp,x'} = \mathcal{O}(n^{\frac{1}{2}})$, as well as $p_0(x, z, \perp) = \mathcal{O}(n^{-\frac{1}{2}})$ and $p_0(\perp, \perp, x') = \mathcal{O}(n^{-\frac{1}{2}})$, for all $x \in \{0, \dots, 3\}$, $z \in \{0, \dots, m-1\}$ and $x' \in \{0, \dots, 15\}$. Hence, for all $i = 1, \dots, 4m$, the quantities defined in eqs. (88-89) scale as follows: $\gamma_i = \mathcal{O}(n^{-\frac{1}{2}})$, $c_i = \mathcal{O}(n^{\frac{1}{2}})$, and $D = \mathcal{O}(n^{\frac{1}{2}})$. Consequently, in order to fulfill eq. (87), we have to choose $\delta_{\text{PE}}^{\text{tol}} = \mathcal{O}((\log n)^{\frac{1}{2}} n^{-\frac{1}{4}})$. Similarly, it can be shown that that we need $\delta_{\text{tom}}^{\text{tol}} = \mathcal{O}((\log n)^{\frac{1}{2}} n^{-\frac{1}{4}})$, in order to satisfy eq. 81).

As for the remaining higher order terms in eq. (82), we note that by eqs. (105,106,141) it holds

$$V \leq \tilde{V} = \mathcal{O}\left(n^{\frac{1}{4}}\right) \quad (150)$$

$$K_a \leq \tilde{K}_a = \mathcal{O}(1). \quad (151)$$

Although $\tilde{V}$ and $\tilde{K}_a$ do not decrease with the number of rounds, we note that in (82) they ap-

pear multiplied by $a - 1$. This eventually leads to

$$(a - 1)\tilde{V} = \mathcal{O}\left(n^{-\frac{1}{2}}\right), \quad (152)$$

$$(a - 1)^2\tilde{K}_a = \mathcal{O}\left(n^{-\frac{3}{2}}\right). \quad (153)$$

Hence, all remaining higher order terms, except $\frac{1}{n}\text{leak}_{\text{EC}}$, which we keep open, scale as $\mathcal{O}(n^{-\frac{1}{4}})$ or less. Further, the term $f(p_0)$ in Theorem 1, given by eq. (148), only depends on n via the prefactor p^{key} . In summary, we can obtain the following bound on the asymptotic key rate.

Theorem 3 (Asymptotic rate) *For the above mentioned values of the parameters, it holds*

$$\begin{aligned} r^{\text{phys}}|_{\Omega_{\text{NA}}^{\text{phys}}} &\geq G_0 + \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{0,xz} p_0^{\text{sim}}(x, z) \\ &+ \sum_{x'=0}^{15} \nu'_{0,x'} p_0^{\text{tom}}(x') - \frac{1}{n} \text{leak}_{\text{EC}} \\ &+ \mathcal{O}((\log n)^{\frac{1}{2}} n^{-\frac{1}{4}}) \end{aligned} \quad (154)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} r^{\text{phys}}|_{\Omega_{\text{NA}}^{\text{phys}}} &\geq G_0 + \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{0,xz} p_0^{\text{sim}}(x, z) \\ &+ \sum_{x'=0}^{15} \nu'_{0,x'} p_0^{\text{tom}}(x') \\ &- \lim_{n \rightarrow \infty} \frac{1}{n} \text{leak}_{\text{EC}}. \end{aligned} \quad (155)$$

5 Numerical implementation and results

In order to show that our approach produces non-trivial key rates in a realistic implementation, we consider the same scenario that was used in [30]. Namely, we simulate an experiment in which Alice and Bob are linked by an optical fibre of length D with excess noise ξ , transmittance $\eta = 10^{-\omega D/10}$ and an attenuation of $\omega = 0.2$ dB/km. This provides us with a simulated distribution that can be computed efficiently using MATLAB

$$p_0^{\text{sim}}(x, z) = \int_{\tilde{\mathcal{R}}_z} \frac{\gamma \exp\left(\frac{-|\gamma e^{i\theta} - \sqrt{\eta} \varphi_x|^2}{1 + \eta \xi/2}\right)}{4\pi(1 + \eta \xi/2)} d\theta d\gamma, \quad (156)$$

where $\tilde{\mathcal{R}}_z$ represents the fragment of the phase space corresponding to each module $z \in$

$\{0, \dots, m - 1\}$, defined according to the intervals described in (28), and $\varphi_x \in \{\alpha, i\alpha, -\alpha, -i\alpha\}$ are the coherent state amplitudes used by Alice with $\alpha \in \mathbb{R}$. The region operators for the constraints $\tilde{R}_B^z$ are given by the same intervals as in (156)

$$\tilde{R}_B^z = \frac{1}{\pi} \int_{\tilde{\mathcal{R}}_z} \gamma |\gamma e^{i\theta}\rangle \langle \gamma e^{i\theta}| d\theta d\gamma, \quad (157)$$

while their numerical implementation requires to switch to the Fock basis. This is done via the inner product [64],

$$\langle \gamma e^{i\theta} | k \rangle = \frac{\gamma^k e^{-\gamma^2/2} e^{-ik\theta}}{\sqrt{k!}}. \quad (158)$$

For the hypothetical tomography, we choose an IC POVM $\{\Gamma_{x'}\}_{x'=0}^{15}$, which completely describes Alice's marginal with a probability distribution

$$p_0^{\text{tom}}(x') = \frac{1}{4} \sum_{x,y=0}^3 \langle \varphi_y | \varphi_x \rangle \text{Tr}[\Gamma_{x'} | x\rangle \langle y|_A]. \quad (159)$$

As argued, under the bounded-energy assumption, the computation of the trade-off requires solving the optimisation for arbitrary finite dimension d_B . At the moment we are unable to do this, so we truncate operators by introducing a photon number cutoff N_c . That is, we impose that all operators, when expressed in the Fock basis, involve terms having at most N_c photons, which implies that $d_B = N_c + 1$. We solve the optimisation for increasing values of N_c and we always observe that the obtained trade-offs numerically converge, see for instance Fig 2 below. A value of $N_c = 15$ provides a good balance between the execution time of the solver and the reliability of the numerics, which is consistent with what was previously observed in [30, 62, 29]. Based on all the obtained numerical evidence, we make the following

Numerical convergence assumption: The derived numerical trade-offs for the considered cut-offs provide reliable approximations to the trade-off for arbitrary finite dimension.

In our view, this assumption is quite plausible in the considered setup, as the amplitude of the states detected by Bob decreases with the channel losses, which eventually means a decreasing average number of received photons.

With all the elements of the optimisation defined, we minimize the SDP (129) according to

the Frank-Wolfe algorithm. For this process we use the toolbox YALMIP [65] together with the interior point solver SDPT3 [66, 67]. Once the suboptimal bound is obtained, we compute the dual (131) using the optimization software CVX [68, 69], since it provides slightly better results than YALMIP and SDPT3. For the iterative process of the Frank-Wolfe algorithm, we set a stopping criterion based on calculating the lower bound (126) every 15 iterations of the minimization (129); if the relative difference between the upper bound given by minimising (118) and the reliable lower bound is smaller than a 2%, the algorithm stops the optimization. If this margin is not reached, the algorithm continues until a total of 300 iterations are performed. Using this approach we obtain $\tilde{\rho}_0$ and $\tilde{\nu}_0$, the feasibility of which can be checked analytically. This allows us to obtain a crossover min-tradeoff function $\tilde{g}_0$ via eq. (134). By Theorem 3, we observe the asymptotic rate

$$r_\infty \geq G_0 + \sum_{x=0}^3 \sum_{z=0}^{m-1} \nu_{0,xz} p_0^{\text{sim}}(x, z) + \sum_{x'=0}^{15} \nu'_{0,x'} p_0^{\text{tom}}(x') - \lim_{n \rightarrow \infty} \frac{1}{n} \text{leak}_{\text{EC}}. \quad (160)$$

For the classical information leaked during error correction, we can assume an honest, iid implementation of the protocol. We introduce a parameter f that quantifies the error correction efficiency with respect to the ideal Shannon limit, so that we write

$$\frac{1}{n} \text{leak}_{\text{EC}} \leq p^{\text{key}} (1 + f) H(\hat{Z}|\hat{X}) \quad (161)$$

where $\hat{X}$, $\hat{Z}$ represent the key string bits (after removing the symbol $\perp$) of Alice and Bob respectively. On the other hand, the parameter p^{key} comes from the fact that only the signals coming from key rounds require error correction. The error correction efficiency may depend on several factors, such as the chosen code, the block size or the form of the probability distribution between Alice and Bob. Here we take f from values that range from 0% to 5% as a showcase of the potential results for our scheme given diverse efficiencies for error correction. The Shannon term $H(\hat{Z}|\hat{X})$ can be computed numerically according

to the distribution (156) adapted for the modulation of the key rounds, namely

$$p_0^{\text{EC}}(x, z) = \int_0^\infty \int_{\frac{\pi}{4}(2z-1)}^{\frac{\pi}{4}(2z+1)} \frac{\gamma \exp\left(\frac{-|\gamma e^{i\theta} - \sqrt{\eta} \varphi_x|^2}{1 + \eta \xi/2}\right)}{4\pi(1 + \eta \xi/2)} d\theta d\gamma. \quad (162)$$

With the error correction cost, it is not only possible to calculate the asymptotic secret key rate, but also optimise the amplitude α that Alice chooses for her coherent states, which is not attainable with only the results from the SDP.

To test the accuracy of our approach, we compared our results for different values of the cutoff, as well as with the so far standard method of computing the asymptotic key rate based on performing parameter estimation with moments of the quadrature operators [30]. The comparison can be found in Figure 2 where one can see that, while using moments to constrain the state shared by Alice and Bob produces better rates for distances shorter than 15 km, both approaches provide comparable results for larger distances. Note that computing moments and coarse-grained probabilities are different ways of discretising the information contained in a CV distribution. These results show that taking moments is better for short distances, albeit the two approaches lead to almost the same values when losses become large. Moreover, as announced, in both cases the asymptotic key rates seem to saturate when increasing the cutoff value. We verified such hypothesis at the inset of Figure 2, where it can be observed that the curves for our modulation converge to the same values.

Figure 3 shows the asymptotic key rates according to eq. (160), where a modulation $(\Delta, \delta) = (0.9, 0.9)$ was employed together with a cutoff $N_c = 15$ for ideal error correction $f = 0\%$. For distances below 150 km, the algorithm typically needs 120 or less iterations to converge. For larger distances, it was necessary to reach the limit of 300 iterations before using eq. (123) to obtain the reliable lower bound. On the other hand, we observed a numerical error at the constraints ε' typically between 10^{-10} for small lengths and 10^{-15} for very long distances, which ensures both the reliability of the code and the tightness of the key rates.

Switching to the finite-size regime, we can use

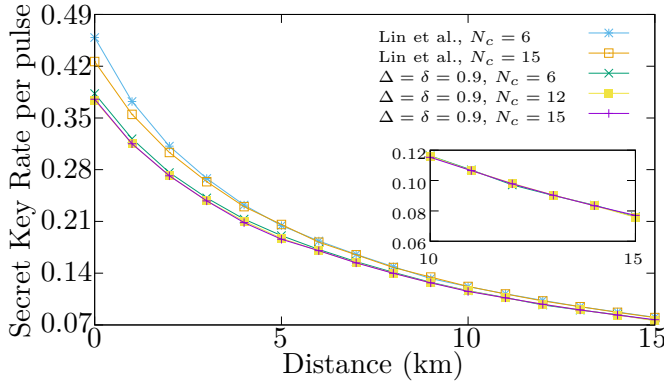


Figure 2: Asymptotic secret key generation rate with $\xi = 2\%$ according to the parameter estimation described in Lin et al. [30] and our modulation $(\Delta, \delta) = (0.9, 0.9)$, both with ideal error correction (i.e., Shannon limit) $f = 0\%$, and diverse values for the cutoff N_c . The amplitude, taken to be the same for all curves, was optimised with respect to the distance. The inset shows the convergence of our modulation.

our crossover min-tradeoff function $\tilde{g}_0$ in Theorem 1 to observe the finite key generation rates. Choosing the parameters $\xi = 1\%$, $f = 1\%$, $N_c = 12$, $\epsilon = 10^{-10}$, $\epsilon_{\text{NA}}^{\text{phys}} = 10^{-4}$, $\epsilon^{\text{tom}} = 10^{-10}$ and $\epsilon_{\text{PE}}^c = 10^{-10}$ together with a grid search optimisation over a and p^{key} , we obtain non-zero key rates for $n \geq 10^{12}$ rounds and distances $D \geq 15$ km. The outcomes of this process are illustrated in Figure 4, where the curves represent different values for the finite key generation rates with respect to the number of rounds taken for the protocol.

In this regard, the simplest approach to perform the finite-key analysis is to use the resulting data from the asymptotic regime, particularly the dual point (131), in order to build the min-tradeoff function. However, this leads in general to suboptimal results for the finite case since the calculated dual point is optimal only in the asymptotic regime—the dual variables appear in the correction terms of the finite-key rate, whose optimization is not included in the Frank-Wolfe method. For instance, $\tilde{K}_a$ scales exponentially with the spread of the min-tradeoff function, which depends on the dual variables according to (139-140). Therefore, the dual variables severely affect the finite-key rates. In order to ameliorate this inconvenience, we make use of a perturbative analysis based on genetic algorithms, which reduces the value of the dual variables while preserving a reasonable performance

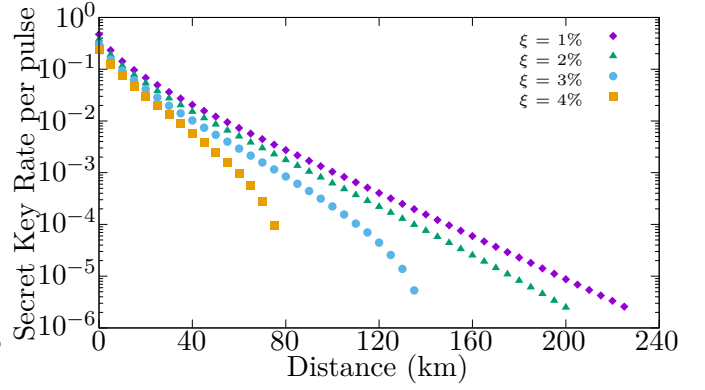


Figure 3: Asymptotic secret key generation rate according to (160) in terms of distances D and excess noise ξ with a modulation $(\Delta, \delta) = (0.9, 0.9)$, ideal error correction $f = 0\%$ and cutoff $N_c = 15$. The amplitude of the coherent states was optimized with respect to the distance.

for any block sizes n . We defer the details of the method to Appendix D, and refer to the complete code available in [70].

We also note that the overall numerical performance of our code enables us to derive the asymptotic secret key in the order of minutes with a reasonable value of the cutoff, $N_c = 12$. Although the perturbative analysis described here increases the overhead of the computations, a fine-tuned implementation can increase the efficiency of the code and perform the complete finite-size analysis for a given distance in a few minutes—such that it can be used in real, on-demand applications.

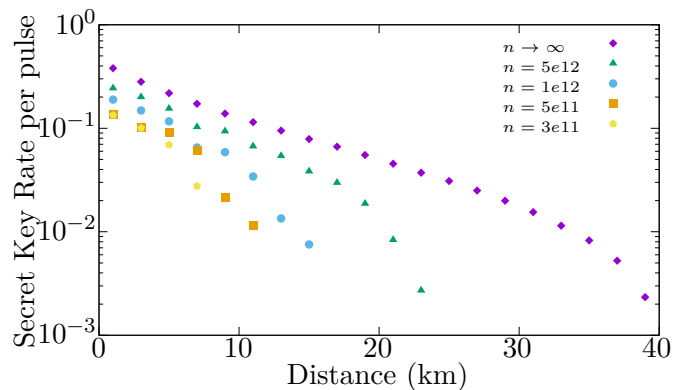


Figure 4: Finite-size secret key generation rate according to (82) for $\xi = 1\%$, $f = 1\%$, $N_c = 12$, n rounds for the protocol and a modulation $(\Delta, \delta) = (0.9, 0.9)$. The parameters a , p^{key} and $\tilde{p}^{\text{PE}}$ were optimised according to a grid search, and we set $\epsilon = 10^{-10}$, $\epsilon_{\text{NA}}^{\text{phys}} = 10^{-4}$, $\epsilon^{\text{tom}} = 10^{-10}$ and $\epsilon_{\text{PE}}^c = 10^{-10}$.

Finally, we explore the impact of error correc-

tion efficiency in the observed key rates. It is well known that in standard CVQKD protocols, as considered in this work, the value of Alice's and Eve's conditional entropies on Bob's results, $H(\hat{Z}|\hat{X})$ and $H(\hat{Z}|\hat{E})$, are very close, especially for large distances. Hence, a non-zero value of f severely affects the possibility of having non-zero key rates. To study this, we plot the finite key rates for blocks of size $n = 5 \times 10^{12}$ as a function of the error correction efficiency in Fig. 5. As it can be seen, small values of f , or in other words, error correction codes with efficiency very close to the Shannon limit, are necessary to generate a secret key for distances beyond 20 km.

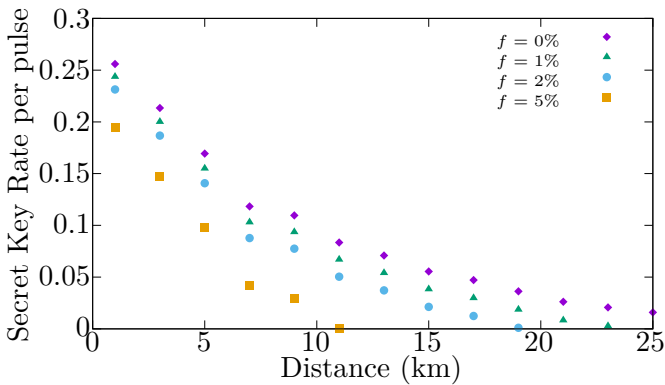


Figure 5: Finite-size secret key generation rates for $n = 5 \times 10^{12}$ rounds and different values of the error correction efficiency f . The parameters were taken to be the same as in Figure 4.

6 Discussion

In this work we have provided a security proof against arbitrary general attacks for a discrete modulated CVQKD protocol, in which Alice prepares four coherent states and Bob performs heterodyne measurements. The proof exploits the fact that the information used for parameter estimation consists of coarse grained probabilities of the generated continuous measurement outcomes instead of moments, as is the case in most of previous approaches of CVQKD. As shown, this hardly affects the asymptotic key rates, but significantly simplifies the security analysis, as one can employ methods originally introduced in the context of DVQKD, such as the EAT.

Despite the simplifications, the application of the EAT, in its original form [42, 43], to the considered prepare-and-measure QKD protocol

is not straightforward. The challenging aspect of this has been the fact that, in order to describe the QKD protocol as a sequence of EAT channels, where Eve's reference system cannot be updated by the EAT channels, we had to use an entanglement-based version of the protocol when applying the EAT. Whereas any prepare-and-measure QKD protocol can be easily transformed to an entanglement-based protocol using the source replacement scheme, we have been faced with the issue that the minimisation defining our min-tradeoff function in eq (61) can be constrained only in terms of the observed statistics from Alice and Bob's measurements in parameter estimation rounds, i.e. by a distribution of the classical output C_1^n . Such constraints are sufficient to obtain a nonzero key rate in DIQKD protocols, as has been considered in [44, 45, 46]. This is due to the fact that in device independent settings the observed statistics alone has to be sufficient to certify an entangled state between Alice and Bob, which is a prerequisite to obtain secure key. In device dependent settings, such as the one we have considered here, however, the observed statistics from Alice and Bob's parameter estimation rounds does not necessarily suffice to certify entanglement. Consequently, if we only use statistics from parameter estimation rounds, the bound on the key rate becomes trivial. We have overcome this issue by considering a hypothetical protocol in which Alice uses some randomly chosen rounds to perform a state tomography on her marginal state, the outcome of which is included in C_1^n . Thus, the observed statistics becomes sufficient to obtain nontrivial bounds on the key rate.

The introduction of the hypothetical tomography poses some additional challenges in the finite size security proof. In particular there is a possibility that the tomography test does not pass, in which case the hypothetical protocol would abort. In order to ensure that this only happens with negligible probability, we have introduced a tolerance parameter $\delta_{\text{tom}}^{\text{tol}}$ in Lemma 1, which has to be subtracted from our key rate. Also, in order to prove security of the physical protocol, it is necessary to show that the raw key states obtained in the hypothetical and physical protocol do not differ by too much and adapt the smoothing parameter accordingly. We have done this in Lemma 2, again at the cost of a reduction of the

key rate. We note that other approaches, such as using the Asymptotic Equipartition Property [19, 35, 32], provide higher finite key rates without the need to add a virtual tomography—such frameworks prove to be simpler, and the numerical key rates (via e.g. [32, Theorem 6]) can be calculated faster, but they are limited to the case of collective attacks, which is surpassed in this work.

As mentioned in the introduction, after much of the work going into this result was finished, a generalised version of the EAT has been presented in [49], known as Generalised EAT (GEAT). In contrast to the original EAT, the new version allows for Eve’s reference system to be updated, while also relaxing the Markov condition to a non-signalling condition. Using GEAT, it is possible to express a prepare-and-measure QKD protocol directly into a sequence of EAT channels, without the need to use an entanglement based version of the protocol, as was shown in [50]. When using this new method, there is no need to introduce a hypothetical tomography, hence our Lemmas 1 and 2 would not be needed and higher key rates may be expected. The only caveat when using GEAT is that it makes an additional assumption on Eve’s attack, namely only allowing her to have one quantum system at a time. This condition can be enforced by Alice waiting for Bob to confirm he has received a state before sending the next one [50], which might not always be practical. Our method of applying the original EAT does not need this assumption. We therefore believe that our current method of using an hypothetical protocol with tomography of Alice’s marginal, is of interest not only for discrete modulated CVQKD, but for proving security of device dependent QKD in settings where the condition that Eve only holds one system at a time is not practical.

Our security analysis can be improved in several directions. As mentioned, being a prepare-and-measure protocol, it is natural to consider the application of GEAT. This may not only provide larger finite-key generation rates, but also allow one to study variants of the protocol using homodyne measurements, which we were unable to accommodate within our security analysis. Another related question is to analyse using GEAT how the obtained rates vary with the number of states prepared by Alice and, in particular, how

they approximate the rates of Gaussian modulated protocols.

The derived key rates are valid under the bounded-energy assumption, stating that Eve’s attack involves states of bounded energy, and a numerical-convergence assumption, stating that the numerical curves obtained for increasing number of photons are very close to the trade-off for arbitrary finite dimension. The first assumption is physically realistic and implies that the states in the protocol can be arbitrarily well approximated by states in a finite dimensional Hilbert space of large enough dimension [71]. This allows the use of EAT since, whereas this theorem does not require an explicit bound on the Hilbert space dimension of Eve, all Hilbert space dimensions are assumed finite [42]. It would be interesting to remove this assumption using recent advances towards a generalisation of the EAT to infinite dimensional Hilbert spaces [72].

The second assumption seems quite plausible in the considered setup, as Alice first prepares coherent states with a small average number of photons that are later sent through a lossy channel. Yet, it is interesting to study how to remove the cutoff in the computation of the asymptotic key rates. This has been achieved for collective attacks in the case where the information used in parameter estimation is made of moments of Bob’s quadratures [31, 32]. The idea in [31] is to introduce a cutoff parameter that depends on the expectation values obtained in parameter estimation and replace the infinite dimensional optimization by a finite dimensional one, plus a correction term, both of which depend on the cutoff parameter. Combining such an approach with the EAT, while possible in principle, is hindered by the dependence of the cutoff parameter on the observed statistics, which has to be taken into account when defining a min-tradeoff function. Namely, the correction term, which is non-affine in the cutoff parameter would have to be included in the min-tradeoff function, and the constraints of the optimisation from which we obtain our min-tradeoff function would contain non-affine terms in the cutoff parameter, greatly complicating the derivation of an affine min-tradeoff function. We leave for future work a complete analysis of how to adapt this framework, and overcome these limitations.

Besides, while presented for a specific proto-

col consisting of four coherent states, our security proof can be adapted to any other constellation of coherent states. It deserves further investigation to study how the key rate changes when using more states and whether and how one can approximate the rate of protocols using Gaussian modulation. It is in fact expected that, as it happens for four coherent states (see Fig 2), the rates obtained when using coarse-grained probabilities will be very close to those obtained when using moments [30]. Moreover, our approach provides a wide framework for the security analysis of CVQKD since the EAT is naturally device independent. Thus, we build a finite size security proof from the asymptotic regime via a min-tradeoff function without making any assumptions on the attack besides the Markov condition (55), here trivially satisfied. This is a fact of importance, provided that the optimal attack for DM CVQKD is not known.

Finally, it is worth noting that, for excess noise $\xi \geq 0.01$, both our proof and the one in [38, 39, 40] require block sizes of the order of $10^{11} - 10^{12}$ to obtain a positive key rate, which are significantly larger than what needed for DVQKD. It is an interesting open question to understand whether a different proof strategy (e.g. the GEAT) can improve this substantially or if this is an intrinsic requirement of CVQKD. Comparing our results with [40], for $\xi = 0.01$ and $n = 10^{12}$, our method provides higher rates for distances around 10 km, although different modelling of the error correction efficiency [73] complicates a direct comparison of the rates.

To conclude, we provide a security proof for CVQKD protocols in which the information in parameter estimation consists of coarse-grained probabilities instead of moments, as done so far. The analysis consist of two main ingredients: (i) the application of EAT including a local tomography process to derive the finite-key rates (ii) the computation of the asymptotic key rates using the formalism of [30] for increasing number of photons. Our work therefore shows that use of coarse-grained probabilities in parameter estimation opens new avenues to prove the security of discrete modulated CVQKD protocols, as well-established methods developed in DVQKD can be applied in a rather straightforward way without any significant impact on the obtained key rates.

Acknowledgements

We would like to thank Rotem Arnon-Friedman, Ian George, Shouvik Ghorai, Min-Hsiu Hsieh, Florian Kanitschar, Anthony Leverrier, Rotem Liss, Bill Munro, Gelo Noel Tabia, Enky Oudot, Stefano Pironio, Toshihiko Sasaki, Ernest Tan, Thomas Van Himbeeck and Shin-Ichiro Yamano for insightful discussions. We would also like to thank two anonymous referees, for conferences QCrypt and QIP, for their insightful comments. This work is supported by the ERC (AdG CERQUTE, grant agreement No. 834266, and StG AlgoQIP, grant agreement No. 851716), the Government of Spain (FUNQIP, NextGen Funds and Severo Ochoa CEX2019-000910-S), Fundació Cellex, Fundació Mir-Puig, Generalitat de Catalunya (CERCA and the postdoctoral fellowship programme Beatriu de Pinós), the AXA Chair in Quantum Information Science, European Union's Horizon 2020 research and innovation programme under grant agreements No. 820466 (project CiviQ), No. 101114043 (project QSNP), No. 101017733 (project Veriqtas) within the QuantERA II Programme, and No. 801370 (2019 BP 00097) within the Marie Skłodowska-Curie Programme.

A Proof of Lemma 3

Let $\mathcal{H}_{\hat{E}}$ be a Hilbert space. We begin with a pure state $|\rho\rangle_{AB\hat{E}} \in \mathcal{H}_{AB\hat{E}}$. Alice and Bob's measurements are then performed coherently by means of a series of isometries. Alice's measurement, as well as the random number generator determining what the round is used for, are described by

$$W_{ARX \leftarrow A}^A = \sum_{r=0}^2 \sqrt{p_r} \left(\sum_{x_r} \sqrt{P_A^{x_r}} \otimes |r\rangle_R \otimes |x_r\rangle_X \right), \quad (163)$$

where $p_0 = p^{\text{key}}, p_1 = p^{\text{PE}}, p_2 = p^{\text{tom}}$, and $P_A^{x_r}$ denotes the x -th POVM element applied by Alice when her random bit provides an outcome r

$$\begin{aligned} \{P_A^{x_0}\}_{x_0} &= \{P_A^{x_1}\}_{x_1} = \{|x\rangle\langle x|_A\}_{x=0}^3, \\ \{P_A^{x_2}\}_{x_2} &= \{\Gamma_A^x\}_{x=0}^{15}. \end{aligned}$$

Note that the POVM elements are given by a square root to preserve the isometric characteristics of the measurement. Register R will announce whether the bit will be used for the generation of the key, parameter estimation or tomography, whereas X stores the result of Alice's measurement. Bob will perform a heterodyne measurement, which he will later discretise according to the goal of the round. Such measurement is given by the isometry

$$W_{BY \leftarrow B}^B = \int d^2y \sqrt{\frac{|y\rangle\langle y|_B}{\pi}} \otimes |y\rangle_Y. \quad (164)$$

Where the integral is given by the fact that coherent states form a continuous basis. The classical communication of R between Alice and Bob, which is wiretapped by Eve, can be expressed coherently by adding ancillary subsystems followed by CNOTs.

$$V_{[R] \leftarrow R}^{\text{c1}} = U_{R:R'R''}^{\text{CNOT}} |00\rangle_{R'R''}, \quad (165)$$

where R' is distributed to Bob and R'' to Eve, and we have introduced the simplifying notation $[R] := RR'R''$. Furthermore, $U_{R:R'R''}^{\text{CNOT}}$ is the unitary describing a double CNOT taking R as control and R' and R'' as targets. Now that Bob and Alice have made their public announcements, we have to apply a new isometry where Bob discretises the key,

$$V_{RY\hat{Z} \leftarrow RY}^K = |0\rangle\langle 0|_R \otimes \sum_{z=0}^3 \sqrt{R_Y^z} \otimes |z\rangle_{\hat{Z}} + (|1\rangle\langle 1|_R + |2\rangle\langle 2|_R) \otimes \mathbb{1}_Y \otimes |\perp\rangle_{\hat{Z}}. \quad (166)$$

Here, the set $\{R_Y^z\}_{z=0}^3$ represents the region operators for the discretisation in key rounds, whose definitions are given in (111). The state that results after applying all the isometries is then given by (where we have omitted identities on systems not involved)

$$|\omega\rangle_{ABXY\hat{Z}[R]\hat{E}} = V^K V^{\text{c1}} W^B W^A |\rho\rangle_{AB\hat{E}}. \quad (167)$$

Finally, the key register is dephased by a pinching map $\mathcal{Z}' : \hat{Z} \rightarrow \hat{Z}$, defined with the Kraus operators

$$Z_j = |j\rangle\langle j|_{\hat{Z}} \otimes \mathbb{1}, \quad (168)$$

for $j \in \{0, 1, 2, 3, \perp\}$. Note that this is same definition as in (112), albeit here with the symbol $\perp$ included. We can now apply Theorem 1 from [60], to show that

$$H(\hat{Z}|R''\hat{E})_{\hat{\mathcal{Z}}(\omega)} = D(\omega_{ABXY\hat{Z}RR'}) || \mathcal{Z}'(\omega_{ABXY\hat{Z}RR'}) \quad (169)$$

Now, the r.h.s. does no longer depend on $\hat{E}$. Let us also observe that in the marginal $\omega_{ABXY\hat{Z}RR'[P]}$ registers RR' have decohered due to traceout of R'' . We can then reformulate it as

$$\omega_{ABXY\hat{Z}RR'} = p^{\text{key}} |00\rangle\langle 00|_{RR'} \otimes \omega_{ABXY\hat{Z}}^{\text{key}} + (1 - p^{\text{key}}) (|11\rangle\langle 11|_{RR'} + |22\rangle\langle 22|_{RR'}) \otimes \omega_{ABXY}^{\perp} \otimes |\perp\rangle\langle \perp|_{\hat{Z}}, \quad (170)$$

where p^{key} denotes the probability that Alice will use the round for the generation of the key. The state (170) has a cq structure, so that by the properties of the relative entropy on cq-states [59], we can simplify (169) by splitting the state according to the classical registers RR' . Moreover, the state $\omega_{ABXY}^\perp \otimes |\perp\rangle\langle\perp|_{\hat{Z}}$ is invariant under the pinching, so that the relative entropy for such state is zero. The whole process adds up to the equality

$$D(\omega_{ABXYRR'\hat{Z}} || \mathcal{Z}'(\omega_{ABXYRR'\hat{Z}})) = p^{\text{key}} D(\omega_{ABXY\hat{Z}}^{\text{key}} || \mathcal{Z}(\omega_{ABXY\hat{Z}}^{\text{key}})), \quad (171)$$

where we have also substituted $\mathcal{Z}'$ for $\mathcal{Z}$ since we have removed $\perp$ from the key register $\hat{Z}$. The explicit form of the key state $\omega_{ABXY\hat{Z}}^{\text{key}}$ is then given by

$$\omega_{ABXY\hat{Z}}^{\text{key}} = \frac{1}{p^{\text{key}}} \text{Tr}_{R''\hat{E}} \left[\langle 00 |_{RR'} \omega_{ABXY[R]\hat{Z}\hat{E}} | 00 \rangle_{RR'} \right]. \quad (172)$$

Following the arguments provided in Appendix A of [30], we can further simplify (171). First of all, the reduction of the state (167) according to the properties of the relative entropy for cq-states has suppressed the sum over r at (163), leaving only the term related to the key generation, namely $r = 0$. Hence, Alice's operator for the key state is given by

$$W_{AX \leftarrow A}'^A = \sum_{x=0}^3 |x\rangle\langle x|_A \otimes |x\rangle_X, \quad (173)$$

where we used the fact that Alice's POVM elements in A are projectors, so that the square root can be removed. This operator now merely copies and projects the information stored in A to the new register X , which effectively represents an isometry that is invariant under the pinching (since both registers are not related to the key register $\hat{Z}$). Hence we can simplify this isometry by removing X , and the final operator for Alice will be a mere identity in A .

As for the key rounds, Bob only needs to obtain the discretised key variable $\hat{Z}$. Thus, he can group the POVM elements corresponding to a particular value of $\hat{Z}$, forming a *coarse grained* POVM $\{R_B^i\}_{i=0}^3$ that acts directly on register B , and is given by the region operators defined in (111). Hence, register Y is not necessary and Bob's measurement and discretisation will thus be given by

$$W_{B\hat{Z} \leftarrow B}'^B = \sum_{z=0}^3 \sqrt{R_B^z} \otimes |z\rangle_{\hat{Z}}. \quad (174)$$

Now, the simplified maps for Alice and Bob are combined to provide the CP map $\mathcal{G} : AB \rightarrow AB\hat{Z}$ that represents the postprocessing, which as shown in (110) is given by the superoperator

$$G = W'^A \otimes W'^B = \mathbb{1}_A \otimes \sum_{z=0}^3 \sqrt{R_B^z} \otimes |z\rangle_{\hat{Z}}. \quad (175)$$

We can now conclude with the redefinition of the relative entropy at (171) in terms of the postprocessing map,

$$D(\omega_{ABXY\hat{Z}}^{\text{key}} || \mathcal{Z}(\omega_{ABXY\hat{Z}}^{\text{key}})) = D(\mathcal{G}(\rho_{AB}) || \mathcal{Z}(\mathcal{G}(\rho_{AB}))). \quad (176)$$

By definition, register R'' is identical to register S in (109) and for any $\mathcal{H}_{\hat{E}}$ and any $|\rho\rangle_{AB\hat{E}} \in \mathcal{H}_{AB\hat{E}}$ it holds $\mathcal{Z}(\omega)_{\hat{Z}S\hat{E}} = (\text{id}_{\hat{E}} \otimes \mathcal{M}^{\text{EAT}}(\rho))_{\hat{Z}S\hat{E}}$, for ω defined as in (167). Combining the equations (169,171,176), we obtain that for all $\tilde{p} \in \mathcal{P}_{\tilde{C}}$,

$$g(\tilde{p}) = \inf_{\rho \in \Sigma(\tilde{p})} H(\hat{Z} | R'' \hat{E})_{\mathcal{Z}(\omega)} = p^{\text{key}} \inf_{\rho \in \Sigma(\tilde{p})} D(\mathcal{G}(\rho_{AB}) || \mathcal{Z}(\mathcal{G}(\rho_{AB}))), \quad (177)$$

which finishes the proof.

B Proof of Lemma 4

Let $p, q \in \mathcal{P}_{\tilde{\mathcal{C}}}$, $0 \leq \lambda \leq 1$. Without loss of generality we can assume that $\Sigma(p)$ is not empty. Then there exist states $\rho_{AB} \in \Sigma(p)$ and $\tau_{AB} \in \Sigma(q)$ such that

$$p^{\text{key}} D(\mathcal{G}(\rho_{AB}) \| \mathcal{Z}(\mathcal{G}(\rho_{AB}))) = g(p), \quad (178)$$

$$p^{\text{key}} D(\mathcal{G}(\tau_{AB}) \| \mathcal{Z}(\mathcal{G}(\tau_{AB}))) = g(q). \quad (179)$$

Let us now consider the flag state

$$\omega_{ABF} = \lambda \rho_{AB} \otimes |0\rangle \langle 0|_F + (1 - \lambda) \tau_{AB} \otimes |1\rangle \langle 1|_F. \quad (180)$$

It then holds [59]

$$\begin{aligned} p^{\text{key}} D(\mathcal{G} \otimes \text{id}_F(\omega_{ABF}) \| \mathcal{Z} \circ \mathcal{G} \otimes \text{id}_F(\omega_{ABF})) &= \lambda p^{\text{key}} D(\mathcal{G}(\rho_{AB}) \| \mathcal{Z}(\mathcal{G}(\rho_{AB}))) \\ &\quad + (1 - \lambda) p^{\text{key}} D(\mathcal{G}(\tau_{AB}) \| \mathcal{Z}(\mathcal{G}(\tau_{AB}))) \\ &= \lambda g(p) + (1 - \lambda) g(q), \end{aligned} \quad (181)$$

As tracing out the flag system F cannot increase the relative entropy, it holds

$$D(\mathcal{G}(\omega_{AB}) \| \mathcal{Z}(\mathcal{G}(\omega_{AB}))) \leq D(\mathcal{G} \otimes \text{id}_F(\omega_{ABF}) \| \mathcal{Z} \circ \mathcal{G} \otimes \text{id}_F(\omega_{ABF})). \quad (182)$$

Let now $c \in \tilde{\mathcal{C}}$. It then holds

$$\begin{aligned} \langle c | \text{Tr}_{OS} [\mathcal{M}^{\text{EAT, test}}(\omega_{AB})] | c \rangle &= \lambda \langle c | \text{Tr}_{OS} [\mathcal{M}^{\text{EAT, test}}(\rho_{AB})] | c \rangle + (1 - \lambda) \langle c | \text{Tr}_{OS} [\mathcal{M}^{\text{EAT, test}}(\tau_{AB})] | c \rangle \\ &= \lambda p(c) + (1 - \lambda) q(c). \end{aligned} \quad (183)$$

This implies that $\omega_{AB} \in \Sigma(\lambda p + (1 - \lambda) q)$. By definition of g , and eqs. (182) and (181), it then holds

$$\begin{aligned} g(\lambda p + (1 - \lambda) q) &\leq p^{\text{key}} D(\mathcal{G}(\omega_{AB}) \| \mathcal{Z}(\mathcal{G}(\omega_{AB}))) \\ &\leq p^{\text{key}} D(\mathcal{G} \otimes \text{id}_F(\omega_{ABF}) \| \mathcal{Z} \circ \mathcal{G} \otimes \text{id}_F(\omega_{ABF})) \\ &= \lambda g(p) + (1 - \lambda) g(q), \end{aligned} \quad (184)$$

finishing the proof.

C Upper bounding the classical smooth max entropy

Let $n \in \mathbb{N}$, and for $1 = 1, \dots, n$, let Y_i be a binary classical random variable such that $P_{Y_i}(1) = p$ and $P_{Y_i}(0) = 1 - p$. Further, define classical random variable X_i such that $X_i = \perp$ if $Y_i = 0$. Otherwise the values are chosen from an alphabet $\mathcal{X}$ such that $|\mathcal{X} \cup \{\perp\}| = d$. We use the operator representation to describe the joint state as

$$\begin{aligned} \rho_{X_1^n Y_1^n} &= \sum_{x_1, \dots, x_n \in \mathcal{X} \cup \{\perp\}} \sum_{y_1, \dots, y_n=0}^1 P_{X_1^n Y_1^n}(x_1, \dots, x_n, y_1, \dots, y_n) \\ &\quad \times |x_1, \dots, x_n\rangle \langle x_1, \dots, x_n|_{X_1^n} \otimes |y_1, \dots, y_n\rangle \langle y_1, \dots, y_n|_{Y_1^n}, \end{aligned} \quad (185)$$

etc.

Lemma 5 *For any $\epsilon > 0$ it holds*

$$H_{\max}^{\epsilon}(X_1^n | Y_1^n)_{\rho} \leq np \log d + \sqrt{\frac{n}{2} \ln \frac{2}{\epsilon^2}} \log d. \quad (186)$$

Proof.

Let $\epsilon > 0$ and define $\delta := \left(\frac{\ln 2 - 2 \ln \epsilon}{2n}\right)^{\frac{1}{2}}$. We can divide the sum in eq. (185) into a part with up to $\lfloor n(p + \delta) \rfloor$ terms with $Y_i = 1$, hence non-trivial X_i , and a part with more than $\lfloor n(p + \delta) \rfloor$ such terms, $\rho_{X_1^n Y_1^n} = \rho'_{X_1^n Y_1^n} + \rho''_{X_1^n Y_1^n}$, where

$$\begin{aligned} \rho'_{X_1^n Y_1^n} &= \sum_{x_1, \dots, x_n \in \mathcal{X} \cup \{\perp\}} \sum_{\substack{y_1, \dots, y_n \in \{0,1\}^n \\ \sum_i y_i \leq \lfloor n(p+\delta) \rfloor}} P_{X_1^n Y_1^n}(x_1, \dots, x_n, y_1, \dots, y_n) \\ &\quad \times |x_1, \dots, x_n\rangle \langle x_1, \dots, x_n|_{X_1^n} \otimes |y_1, \dots, y_n\rangle \langle y_1, \dots, y_n|_{Y_1^n}, \end{aligned} \quad (187)$$

$$\begin{aligned} \rho''_{X_1^n Y_1^n} &= \sum_{x_1, \dots, x_n \in \mathcal{X} \cup \{\perp\}} \sum_{\substack{y_1, \dots, y_n \in \{0,1\}^n \\ \sum_i y_i > \lfloor n(p+\delta) \rfloor}} P_{X_1^n Y_1^n}(x_1, \dots, x_n, y_1, \dots, y_n) \\ &\quad \times |x_1, \dots, x_n\rangle \langle x_1, \dots, x_n|_{X_1^n} \otimes |y_1, \dots, y_n\rangle \langle y_1, \dots, y_n|_{Y_1^n}, \end{aligned} \quad (188)$$

Let us define

$$\kappa := \text{Tr} \left[\rho''_{X_1^n Y_1^n} \right] = \sum_{k=\lfloor n(p+\delta) \rfloor + 1}^n p^k (1-p)^{n-k} \binom{n}{k}, \quad (189)$$

and note that by Hoeffding's inequality, it holds $\kappa \leq e^{-2n\delta^2} \leq \frac{\epsilon^2}{2}$. By [51], Lemma 3.17, it then holds for the purified distance

$$P(\rho_{X_1^n Y_1^n}, \rho'_{X_1^n Y_1^n}) \leq \sqrt{\left\| \rho_{X_1^n Y_1^n} - \rho'_{X_1^n Y_1^n} \right\|_1 + \text{Tr} \left(\rho_{X_1^n Y_1^n} - \rho'_{X_1^n Y_1^n} \right)} = \sqrt{2\kappa} \leq \epsilon \quad (190)$$

hence ρ' is in the ϵ -ball around ρ . Consequently, as only the non trivial X_i contribute to the max entropy, it holds

$$H_{\max}^\epsilon(X_1^n | Y_1^n)_\rho \leq H_{\max}(X_1^n | Y_1^n)_{\rho'} \leq \log d^{\lfloor n(p+\delta) \rfloor}. \quad (191)$$

Inserting our choice for δ completes the proof. $\blacksquare$

Now, let's add conditioning on an event Ω that occurs with probability $p_\Omega > 0$. We can express the state (185) as $\rho_{X_1^n Y_1^n} = \text{Pr}[\Omega] \rho_{X_1^n Y_1^n} |_\Omega + (1 - \text{Pr}[\Omega]) \rho_{X_1^n Y_1^n} |_{-\Omega}$.

Lemma 6 For any $\epsilon > 0$ and $0 < p_\Omega \leq 1$ it holds

$$H_{\max}^\epsilon(X_1^n | Y_1^n)_{\rho|_\Omega} \leq np \log d + \sqrt{\frac{n}{2} \ln \frac{2}{\epsilon^2 \text{Pr}[\Omega]}} \log d. \quad (192)$$

Proof.

Let $\epsilon > 0$ and define $\delta := \left(\frac{\ln 2 - \ln p_\Omega - 2 \ln \epsilon}{2n}\right)^{\frac{1}{2}}$. Again, we divide $\rho_{X_1^n Y_1^n} |_\Omega = \rho'_{X_1^n Y_1^n} |_\Omega + \rho''_{X_1^n Y_1^n} |_\Omega$, where

$$\begin{aligned} \rho'_{X_1^n Y_1^n} |_\Omega &= \sum_{x_1, \dots, x_n \in \mathcal{X} \cup \{\perp\}} \sum_{\substack{y_1, \dots, y_n \in \{0,1\}^n \\ \sum_i y_i \leq \lfloor n(p+\delta) \rfloor}} P_{X_1^n Y_1^n}(x_1, \dots, x_n, y_1, \dots, y_n | \Omega) \\ &\quad \times |x_1, \dots, x_n\rangle \langle x_1, \dots, x_n|_{X_1^n} \otimes |y_1, \dots, y_n\rangle \langle y_1, \dots, y_n|_{Y_1^n}, \end{aligned} \quad (193)$$

$$\begin{aligned} \rho''_{X_1^n Y_1^n} |_\Omega &= \sum_{x_1, \dots, x_n \in \mathcal{X} \cup \{\perp\}} \sum_{\substack{y_1, \dots, y_n \in \{0,1\}^n \\ \sum_i y_i > \lfloor n(p+\delta) \rfloor}} P_{X_1^n Y_1^n}(x_1, \dots, x_n, y_1, \dots, y_n | \Omega) \\ &\quad \times |x_1, \dots, x_n\rangle \langle x_1, \dots, x_n|_{X_1^n} \otimes |y_1, \dots, y_n\rangle \langle y_1, \dots, y_n|_{Y_1^n}, \end{aligned} \quad (194)$$

Let us define

$$\kappa := \text{Tr} \left[\rho''_{X_1^n Y_1^n} | \Omega \right] \quad (195)$$

$$= \sum_{k=\lfloor n(p+\delta) \rfloor + 1}^n \Pr(|\{i : Y_i = 1\}| = k | \Omega) \quad (196)$$

$$= \frac{1}{\Pr[\Omega]} \sum_{k=\lfloor n(p+\delta) \rfloor + 1}^n \Pr(|\{i : Y_i = 1\}| = k \cap \Omega) \quad (197)$$

$$\leq \frac{1}{\Pr[\Omega]} \sum_{k=\lfloor n(p+\delta) \rfloor + 1}^n p^k (1-p)^{n-k} \binom{n}{k} \quad (198)$$

$$= \frac{1}{\Pr[\Omega]} \Pr[k > n(p+\delta)]. \quad (199)$$

By Hoeffding's inequality, it holds $\kappa \leq \frac{e^{-2n\delta^2}}{p\Omega} \leq \frac{\epsilon^2}{2}$. By [51], Lemma 3.17, it then holds for the purified distance

$$\Pr(\rho_{X_1^n Y_1^n} | \Omega, \rho'_{X_1^n Y_1^n} | \Omega) \leq \sqrt{\left\| \rho_{X_1^n Y_1^n} | \Omega - \rho'_{X_1^n Y_1^n} | \Omega \right\|_1 + \text{Tr} \left(\rho_{X_1^n Y_1^n} | \Omega - \rho'_{X_1^n Y_1^n} | \Omega \right)} = \sqrt{2\kappa} \leq \epsilon \quad (200)$$

hence $\rho' | \Omega$ is in the ϵ -ball around $\rho | \Omega$. Consequently, as only the non-trivial X_i contribute to the max entropy, it holds

$$H_{\max}^\epsilon(X_1^n | Y_1^n)_{\rho | \Omega} \leq H_{\max}(X_1^n | Y_1^n)_{\rho' | \Omega} \leq \log d^{\lfloor n(p+\delta) \rfloor}. \quad (201)$$

Inserting our choice for δ completes the proof.

■

D Perturbative analysis for finite-key distillation

The framework presented in (4.3) provides a method to derive finite secret key rates via Frank-Wolfe and the dual problem (131) using a min-tradeoff function. This technique has the drawback that the dual variables appear in the correction terms of the finite key rate (82), and thus these corrections are not directly optimised. Eventually, this poses a problem that harms the performance of our methodology, especially for small numbers of rounds n . We overcome this obstacle by employing a perturbative analysis via genetic algorithms—for the original dual objective (131), we apply a modification

$$\ell_{\tilde{p}_0, \epsilon'}^0(\vec{\nu}, \vec{\mu}) \rightarrow \ell_{\tilde{p}_0, \epsilon'}^0(\vec{\nu}, \vec{\mu}) - \zeta_{\tilde{p}_0}(\vec{\kappa}) \quad (202)$$

given by a perturbative term

$$\zeta_{\tilde{p}_0}(\vec{\kappa}) = \kappa_0 \|\tilde{p}_0\|_1 + \kappa_1 \|\tilde{p}_0\|_2 + \kappa_2 \|\tilde{p}_0\|_\infty. \quad (203)$$

Let us observe that, for $\vec{\kappa} \in \mathbb{R}_+^3$, the perturbed dual objective serves as a lower bound for the original one. Moreover, the perturbation acts as a term that reduces the spread of the dual variables when the new objective dual is employed, and since both maximisations are executed under the same set of constraints, we can solve the SDP given by the perturbed objective function, and insert the solution in the original version (131) to build the min-tradeoff function. In order to derive useful values for $\vec{\kappa}$ that balance a minimised value for the dual variables with an increased performance in the finite-key analysis, we use a step inspired by genetic algorithms. The method goes as follows.

Algorithm 1 Genetic subroutine

For round $m \in \{1, \dots, 5\}$ perform the following steps:

1. Generate a set of random vectors $\{\vec{\kappa}_j\}_{j=1}^{100}$, with coefficients ranging between 0 and 10^{-3} .
2. For every vector, define a perturbed objective dual as (202) for the maximisation (131).
3. Solve the resulting SDP, and use the solution to calculate both the min-tradeoff function and the finite secret key rate according to Theorem 1.
4. Record as S_m the highest finite key rate achieved for the iteration.
5. Discard the vectors that provide finite rates below the percentile 10, and those whose value for the spread of the min-tradeoff function is above the percentile 95.
6. Combine randomly the remaining vectors in 90 pairs $(\vec{\kappa}_i, \vec{\kappa}_j)$, with the associated finite rates (F_i, F_j) , and create a new population of vectors.
7. For each pair $(\vec{\kappa}_i, \vec{\kappa}_j)$, create a vector $\vec{\kappa}_k$ by means of genetic crossings. For $l \in \{0, 1, 2\}$, every entry $\kappa_{l,k}$ of the new vector is evaluated with the following procedure:
 - Generate a random value $p \in [0, 1]$ using a uniform distribution. If $p > 0.9$, assign to κ_k^l a random value between 10^{-8} and 10^{-1} .
 - Otherwise, draw a value from a binomial distribution with a bias $F_i/(F_i + F_j)$ towards the zero. If the value is zero, evaluate $\kappa_{l,k} := \kappa_{l,i}$, and otherwise $\kappa_{l,k} := \kappa_{l,j}$.
8. Complete the new population by adding the 10 best performing vectors (in terms of finite key rates) from the previous round.
9. Start over the routine with the new set of vectors.

Once this subroutine is complete, the final secret key rate is given by the maximum value in $\{S_1, \dots, S_5\}$. We note that this approach is purely heuristic, and it can be further improved by finely adjusting the number of iterations and the range of values for the coefficients. Nevertheless, it provides a proper framework to derive reliable, high key rates in the finite setting. In particular, it enables us to reduce the value of the coefficients ϵ , ϵ_{PE}^c and ϵ^{tom} without affecting noticeably the results of our method.

References

- [1] Charles H. Bennett and Gilles Brassard. “Quantum Cryptography: Public Key Distribution and Coin Tossing”. In Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing. Bangalore, India (1984). IEEE Computer Society Press, New York.
- [2] Artur K. Ekert. “Quantum cryptography based on bell’s theorem”. *Phys. Rev. Lett.* **67**, 661–663 (1991).
- [3] Charles H. Bennett, Gilles Brassard, and N. David Mermin. “Quantum cryptography without bell’s theorem”. *Phys. Rev. Lett.* **68**, 557–559 (1992).
- [4] N. J. Cerf, M. Lévy, and G. Van Assche. “Quantum distribution of gaussian keys using squeezed states”. *Phys. Rev. A* **63**, 052311 (2001).
- [5] Frédéric Grosshans and Philippe Grangier. “Continuous variable quantum cryptography using coherent states”. *Phys. Rev. Lett.* **88**, 057902 (2002).
- [6] Christian Weedbrook, Andrew M. Lance, Warwick P. Bowen, Thomas Symul, Timothy C. Ralph, and Ping Koy Lam. “Quantum cryptography without switching”. *Phys. Rev. Lett.* **93**, 170504 (2004).
- [7] Raúl García-Patrón and Nicolas J. Cerf.

- “Continuous-variable quantum key distribution protocols over noisy channels”. *Phys. Rev. Lett.* **102**, 130501 (2009).
- [8] Peter W. Shor and John Preskill. “Simple proof of security of the bb84 quantum key distribution protocol”. *Phys. Rev. Lett.* **85**, 441–444 (2000).
- [9] Daniel Gottesman, H-K Lo, Norbert Lutkenhaus, and John Preskill. “Security of quantum key distribution with imperfect devices”. In *International Symposium on Information Theory*, 2004. ISIT 2004. Proceedings. Page 136. IEEE (2004).
- [10] Renato Renner, Nicolas Gisin, and Barbara Kraus. “Information-theoretic security proof for quantum-key-distribution protocols”. *Physical Review A* **72**, 012332 (2005).
- [11] Renato Renner. “Security of quantum key distribution”. *International Journal of Quantum Information* **06**, 1–127 (2008).
- [12] Masato Koashi. “Simple security proof of quantum key distribution based on complementarity”. *New Journal of Physics* **11**, 045018 (2009).
- [13] Michael M. Wolf, Geza Giedke, and J. Ignacio Cirac. “Extremality of gaussian quantum states”. *Phys. Rev. Lett.* **96**, 080502 (2006).
- [14] Igor Devetak and Andreas Winter. “Distillation of secret key and entanglement from quantum states”. *Proceedings of the Royal Society A: Mathematical, Physical and engineering sciences* **461**, 207–235 (2005).
- [15] Frédéric Grosshans, Nicolas J. Cerf, Jérôme Wenger, Rosa Tualle-Brouri, and Philippe Grangier. “Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables”. *Quantum Info. Comput.* **3**, 535–552 (2003).
- [16] Miguel Navascués, Frédéric Grosshans, and Antonio Acín. “Optimality of gaussian attacks in continuous-variable quantum cryptography”. *Phys. Rev. Lett.* **97**, 190502 (2006).
- [17] Raúl García-Patrón and Nicolas J. Cerf. “Unconditional optimality of gaussian attacks against continuous-variable quantum key distribution”. *Phys. Rev. Lett.* **97**, 190503 (2006).
- [18] Anthony Leverrier and Philippe Grangier. “Simple proof that gaussian attacks are optimal among collective attacks against continuous-variable quantum key distribution with a gaussian modulation”. *Phys. Rev. A* **81**, 062314 (2010).
- [19] Anthony Leverrier. “Composable security proof for continuous-variable quantum key distribution with coherent states”. *Phys. Rev. Lett.* **114**, 070501 (2015).
- [20] R. Renner and J. I. Cirac. “de finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography”. *Phys. Rev. Lett.* **102**, 110504 (2009).
- [21] Anthony Leverrier, Raúl García-Patrón, Renato Renner, and Nicolas J. Cerf. “Security of continuous-variable quantum key distribution against general attacks”. *Phys. Rev. Lett.* **110**, 030502 (2013).
- [22] Anthony Leverrier. “Security of continuous-variable quantum key distribution via a gaussian de finetti reduction”. *Phys. Rev. Lett.* **118**, 200501 (2017).
- [23] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner. “Continuous variable quantum key distribution: Finite-key analysis of composable security against coherent attacks”. *Phys. Rev. Lett.* **109**, 100502 (2012).
- [24] Fabian Furrer. “Reverse-reconciliation continuous-variable quantum key distribution based on the uncertainty principle”. *Phys. Rev. A* **90**, 042325 (2014).
- [25] Matthias Christandl, Robert König, and Renato Renner. “Postselection technique for quantum channels with applications to quantum cryptography”. *Phys. Rev. Lett.* **102**, 020504 (2009).
- [26] Cosmo Lupo. “Towards practical security of continuous-variable quantum key distribution”. *Phys. Rev. A* **102**, 022623 (2020).
- [27] Anthony Leverrier and Philippe Grangier. “Unconditional security proof of long-distance continuous-variable quantum key distribution with discrete modulation”. *Phys. Rev. Lett.* **102**, 180504 (2009).
- [28] Anthony Leverrier and Philippe Grangier. “Continuous-variable quantum-key-distribution protocols with a non-gaussian modulation”. *Phys. Rev. A* **83**, 042312 (2011).
- [29] Shouvik Ghorai, Philippe Grangier, Eleni Diamanti, and Anthony Leverrier. “Asymp-

- otic security of continuous-variable quantum key distribution with a discrete modulation”. *Phys. Rev. X* **9**, 021059 (2019).
- [30] Jie Lin, Twesh Upadhyaya, and Norbert Lütkenhaus. “Asymptotic security analysis of discrete-modulated continuous-variable quantum key distribution”. *Phys. Rev. X* **9**, 041064 (2019).
- [31] Twesh Upadhyaya, Thomas van Himbeeck, Jie Lin, and Norbert Lütkenhaus. “Dimension reduction in quantum key distribution for continuous- and discrete-variable protocols”. *PRX Quantum* **2**, 020325 (2021).
- [32] Florian Kanitschar, Ian George, Jie Lin, Twesh Upadhyaya, and Norbert Lütkenhaus. “Finite-size security for discrete-modulated continuous-variable quantum key distribution protocols”. *PRX Quantum* **4**, 040306 (2023).
- [33] Aurélie Denys, Peter Brown, and Anthony Leverrier. “Explicit asymptotic secret key rate of continuous-variable quantum key distribution with an arbitrary modulation”. *Quantum* **5**, 540 (2021).
- [34] Wen-Bo Liu, Chen-Long Li, Yuan-Mei Xie, Chen-Xun Weng, Jie Gu, Xiao-Yu Cao, Yu-Shuo Lu, Bing-Hong Li, Hua-Lei Yin, and Zeng-Bing Chen. “Homodyne detection quadrature phase shift keying continuous-variable quantum key distribution with high excess noise tolerance”. *PRX Quantum* **2**, 040334 (2021).
- [35] Cosmo Lupo and Yingkai Ouyang. “Quantum key distribution with nonideal heterodyne detection: Composable security of discrete-modulation continuous-variable protocols”. *PRX Quantum* **3**, 010341 (2022).
- [36] Eneet Kaur, Saikat Guha, and Mark M. Wilde. “Asymptotic security of discrete-modulation protocols for continuous-variable quantum key distribution”. *Phys. Rev. A* **103**, 012412 (2021).
- [37] Panagiotis Papanastasiou and Stefano Pirandola. “Continuous-variable quantum cryptography with discrete alphabets: Composable security under collective gaussian attacks”. *Phys. Rev. Res.* **3**, 013047 (2021).
- [38] Takaya Matsuura, Kento Maeda, Toshihiko Sasaki, and Masato Koashi. “Finite-size security of continuous-variable quantum key distribution with digital signal processing”. *Nature communications* **12**, 1–13 (2021).
- [39] Shinichiro Yamano, Takaya Matsuura, Yui Kuramochi, Toshihiko Sasaki, and Masato Koashi. “Finite-size security proof of binary-modulation continuous-variable quantum key distribution using only heterodyne measurement”. *Physica Scripta* **99**, 025115 (2024).
- [40] Takaya Matsuura, Shinichiro Yamano, Yui Kuramochi, Toshihiko Sasaki, and Masato Koashi. “Refined finite-size analysis of binary-modulation continuous-variable quantum key distribution”. *Quantum* **7**, 1095 (2023).
- [41] Ryo Namiki and Takuya Hirano. “Security of quantum cryptography using balanced homodyne detection”. *Phys. Rev. A* **67**, 022308 (2003).
- [42] Frédéric Dupuis, Omar Fawzi, and Renato Renner. “Entropy accumulation”. *Communications in Mathematical Physics* **379**, 867–913 (2020).
- [43] Frédéric Dupuis and Omar Fawzi. “Entropy accumulation with improved second-order term”. *IEEE Transactions on information theory* **65**, 7596–7612 (2019).
- [44] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. “Practical device-independent quantum cryptography via entropy accumulation”. *Nature communications* **9**, 459 (2018).
- [45] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. “Simple and tight device-independent security proofs”. *SIAM Journal on Computing* **48**, 181–225 (2019).
- [46] Ernest Y.-Z. Tan, Pavel Sekatski, Jean-Daniel Bancal, René Schwonnek, Renato Renner, Nicolas Sangouard, and Charles C.-W. Lim. “Improved DIQKD protocols with finite-size analysis”. *Quantum* **6**, 880 (2022).
- [47] Ian George, Jie Lin, Thomas van Himbeeck, Kun Fang, and Norbert Lütkenhaus. “Finite-key analysis of quantum key distribution with characterized devices using entropy accumulation” (2022). [arXiv:2203.06554](https://arxiv.org/abs/2203.06554).
- [48] Adam Winick, Norbert Lütkenhaus, and Patrick J. Coles. “Reliable numerical key rates for quantum key distribution”. *Quantum* **2**, 77 (2018).
- [49] Tony Metger, Omar Fawzi, David Sutter,

- and Renato Renner. “Generalised entropy accumulation”. In 2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS). Pages 844–850. IEEE (2022).
- [50] Tony Metger and Renato Renner. “Security of quantum key distribution from generalised entropy accumulation”. *Nature Communications* **14**, 5272 (2023).
- [51] Marco Tomamichel. “Quantum information processing with finite resources: Mathematical foundations”. Volume 5. Springer. (2015). url: <https://link.springer.com/book/10.1007/978-3-319-21891-5>.
- [52] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. “On quantum Rényi entropies: A new generalization and some properties”. *Journal of Mathematical Physics* **54**, 122203 (2013).
- [53] Mark M Wilde, Andreas Winter, and Dong Yang. “Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched rényi relative entropy”. *Communications in Mathematical Physics* **331**, 593–622 (2014).
- [54] Marcos Curty, Maciej Lewenstein, and Norbert Lütkenhaus. “Entanglement as a precondition for secure quantum key distribution”. *Phys. Rev. Lett.* **92**, 217903 (2004).
- [55] Agnes Ferenczi and Norbert Lütkenhaus. “Symmetries in quantum key distribution and the connection between optimal attacks and optimal cloning”. *Phys. Rev. A* **85**, 052310 (2012).
- [56] Fabian Furrer. “Security of continuous-variable quantum key distribution and aspects of device-independent security”. PhD thesis. Gottfried Wilhelm Leibniz Universität Hannover. (2012).
- [57] Mario Berta, Fabian Furrer, and Volkher B. Scholz. “The smooth entropy formalism for von Neumann algebras”. *Journal of Mathematical Physics* **57** (2015).
- [58] Shipra Agrawal and Randy Jia. “Optimistic Posterior Sampling for Reinforcement Learning: Worst-Case Regret Bounds”. *Mathematics of Operations Research* **48**, 363–392 (2023).
- [59] Mark M Wilde. “Quantum information theory”. Cambridge University Press. (2013).
- [60] Patrick J. Coles. “Unification of different views of decoherence and discord”. *Phys. Rev. A* **85**, 042103 (2012).
- [61] Wen-Zhao Liu, Ming-Han Li, Sammy Ragy, Si-Ran Zhao, Bing Bai, Yang Liu, Peter J Brown, Jun Zhang, Roger Colbeck, Jingyun Fan, et al. “Device-independent randomness expansion against quantum side information”. *Nature Physics* **17**, 448–451 (2021).
- [62] Hao Hu, Jiyoung Im, Jie Lin, Norbert Lütkenhaus, and Henry Wolkowicz. “Robust Interior Point Method for Quantum Key Distribution Rate Computation”. *Quantum* **6**, 792 (2022).
- [63] Stephen Boyd and Lieven Vandenbergh. “Convex optimization”. Cambridge University Press. (2004).
- [64] Stephen M. Barnett and Paul M. Radmore. “Methods in Theoretical Quantum Optics”. Oxford University Press. (2002).
- [65] J. Löfberg. “Yalmip : A toolbox for modeling and optimization in matlab”. In Proceedings of the CACSD Conference. Taipei, Taiwan (2004).
- [66] M. J. Todd K. C. Toh and R. H. Tütüncü. “Sdpt3 — a matlab software package for semidefinite programming, version 1.3”. *Optimization Methods and Software* **11**, 545–581 (1999).
- [67] R. H. Tütüncü, K. C. Toh, and M. J. Todd. “Solving semidefinite-quadratic-linear programs using sdpt3”. *Mathematical programming Ser.B* **95**, 189–217 (2003).
- [68] Michael Grant and Stephen Boyd. “CVX: Matlab software for disciplined convex programming, version 2.1”. <http://cvxr.com/cvx> (2014).
- [69] Michael Grant and Stephen Boyd. “Graph implementations for nonsmooth convex programs”. In V. Blondel, S. Boyd, and H. Kimura, editors, Recent Advances in Learning and Control. Pages 95–110. Lecture Notes in Control and Information Sciences. Springer-Verlag Limited (2008).
- [70] Carlos Pascual-Garcia. “Discrete Modulated CVQKD via the EAT”. https://github.com/CPascualGarcia/DiscreteModulatedCVQKD_EAT.git (2024).
- [71] Marco Tomamichel and Anthony Leverrier. “A largely self-contained and complete se-

- curity proof for quantum key distribution”. [Quantum](#) **1**, 14 (2017).
- [72] Omar Fawzi, Li Gao, and Mizanur Rahman. “Asymptotic equipartition theorems in von neumann algebras” (2023). [arXiv:2212.14700](#).
- [73] Anthony Leverrier. “Information reconciliation for discretely-modulated continuous-variable quantum key distribution” (2023). [arXiv:2310.17548](#).