Joint-measurability and quantum communication with untrusted devices

Michele Masini[1], Marie Ioannou[2], Nicolas Brunner[2], Stefano Pironio[1], and Pavel Sekatski[2]

[1]Laboratoire d'Information Quantique, Université libre de Bruxelles (ULB), Belgium
[2]Department of Applied Physics, University of Geneva, Switzerland

21 March 2024

Photon loss represents a major challenge for the implementation of quantum communication protocols with untrusted devices, e.g. in the device-independent (DI) or semi-DI approaches. Determining critical loss thresholds is usually done in case-by-case studies. In the present work, we develop a general framework for characterizing the admissible levels of loss and noise in a wide range of scenarios and protocols with untrusted measurement devices. In particular, we present general bounds that apply to prepare-and-measure protocols for the semi-DI approach, as well as to Bell tests for DI protocols. A key step in our work is to establish a general connection between quantum protocols with untrusted measurement devices and the fundamental notions of channel extendibility and joint-measurability, which capture essential aspects of the communication and measurement of quantum information. In particular, this leads us to introduce the notion of partial joint-measurability, which naturally arises within quantum cryptography.

1 Introduction

Photon loss, which results from unavoidable absorption and scattering in optical channels, as well as limited detector efficiency, represents one of the key challenges for the implementation of long-distance quantum communication. This issue becomes even more critical in experiments and applications where quantum states are transmitted through untrusted channels and measured using untrusted apparatuses. This is common in quantum cryptography, most notably within the device-independent (DI) [1, 2] or semi-device-independent (SDI) [3, 4] approaches, but also in fundamental tests of quantum physics such as quantum Bell nonlocality [5, 6] and steering [7, 8].

In these situations, certifying quantum properties such as entanglement or randomness, and achieving quantum advantages, such as information-theoretic security, becomes impossible when photon losses exceed a certain threshold. The underlying reason is that the untrusted channels and measurement devices could potentially implement processes in which losses are not merely passive phenomena, but are influenced by the choice of the experimenter's measurement settings in a way that could, for instance, allow for blinding attacks [9, 10, 11, 12]. In such attacks, the eavesdropper might selectively and remotely cause detection events based on whether their own measurement matches that of the receiver. More generally, the most famous example of this is arguably the so-called "detection loophole" [13], that plagued experimental tests of quantum nonlocality for decades, before experiments could finally close it, and, eventually led to the celebrated loophole-free Bell experiments.

The level of admissible losses is generally highly dependent on the specific quantum communication protocol that is considered. Moreover, this tolerance also depends on the detailed aspects of the experimental setup, including the specific states that are intended to be transmitted through the channel and the expected measurements to be performed on them. An intense effort has been devoted to characterizing critical loss thresholds. This started in the context of Bell experiments (see e.g. [13, 14, 15, 16] and [17, 6] for reviews) as well as in steering test (see e.g. [18, 19, 20]). In turn, critical efficiencies for DI and SDI protocols have also been discussed (see e.g. [21, 22]). While all these works have played a significant role towards the first loophole-free Bell tests [23, 24, 25] and proof-of-principle implementations of DI protocols [26, 27, 28], it is fair to say that our understanding is currently limited to few specific cases.

In this paper, we develop a general framework for characterizing critical losses in quantum communication setups. This allows us to establish lower-bounds on admissible losses that are applicable to a wide range of experiments and protocols involving untrusted measurement devices. The key point of our work is to establish a connection with the notion of "joint-measurability", which captures the incompatibility of quantum measurements [29, 30, 31].

Our starting point is to introduce the concept of a "channel-measurement unit" (CMU), which repre-

Michele Masini: michele.masini@ulb.be
Nicolas Brunner: Nicolas.Brunner@unige.ch
Stefano Pironio: stefano.pironio@ulb.be
Pavel Sekatski: pavel.sekatski@gmail.com

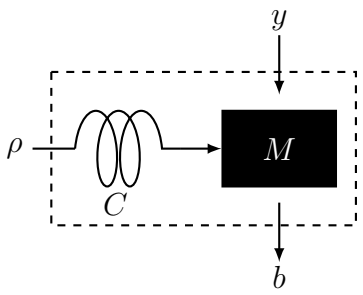

Figure 1: A channel-measurement unit (CMU) is composed of an untrusted channel C and an untrusted measurement device M. The CMU receives as input a quantum state ρ that is transmitted through the channel C and then measured by M. The classical input y denotes the choice of measurement (performed e.g. by the honest user), while its output is denoted b.

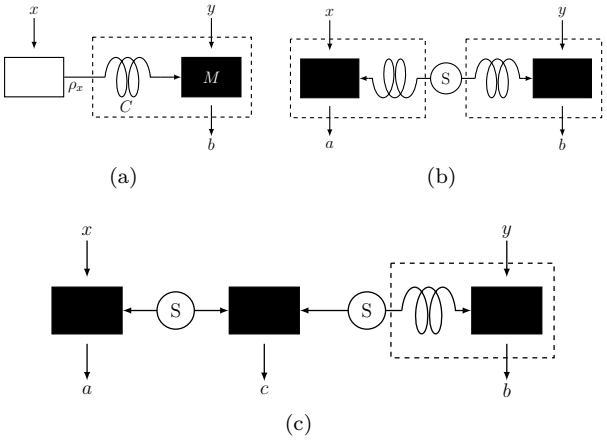

Figure 2: The CMU is a basic component of many experiments with untrusted devices, such as (a) a simple prepare-and-measure scenario, (b) A Bell test, (c) a quantum network featuring several sources and measurements

sents a fundamental component of quantum communication protocols with untrusted devices. Specifically, the CMU consists of an untrusted communication channel C followed by an untrusted measuring device M, as depicted in Fig. 1. The CMU serves as the basic building block for a broad range of setups. The simplest examples are SDI prepare-and-measure scenarios [32, 33], as in Fig. 2.a), and bipartite Bell experiment where a pair of entangled photons are sent to two separated CMUs (see Fig. 2.b.). These setups serve as a basis for SDI QKD protocols [3, 4, 22]) and DI protocols [1, 27] respectively. More generally, the CMU can also be seen as a part of more complex scenarios, e.g. involving quantum networks or quantum circuits (see Fig. 2.c).

We aim to determine the admissible level of loss in the CMU before it loses its ability to exhibit 'quantumness', i.e. when it becomes classically simulable. We provide several natural definitions for this, which are shown to be all equivalent to the condition that the effective POMVs implemented by the CMU, representing the combined effect of the channel and the measurement device, are jointly measurable. An equivalent condition, based on the channel extendibility of the CMU, is also particularly convenient to work with. Using these definitions, we provide several upper bounds on the loss threshold. Notably, these bounds typically depend only on the number of measurements implemented by the CMU, and on the noise level, but not on the details of the channel and/or the measurements.

More generally, our work motivates the investigation of joint measurability for sets of measurements that are not only noisy, the case most existing works have focused on [34, 35, 36, 37, 38, 39], but that also feature losses, which has been considered only in a limited number of previous works [40, 41, 42, 43].

We consider two different scenarios for taking losses into account in protocols with untrusted devices. The first consists of attributing a specific measurement outcome (an inconclusive 'no-click' outcome) to events where the photons are lost. We refer to this as the *no-click scenario*, corresponding e.g. to experiments involving single-photon detectors. The second consists of general optical scenarios where measurements may always produce a conclusive outcome, as, e.g., in continuous variable setups based on homodyne (or heterodyne) measurements. In this case, we consider a standard model of loss for the optical channel, corresponding to mixing the input system with a thermal state through a beam splitter of limited transmissivity. We refer to this as the *thermal-noise channel scenario*.

In the no-click scenario, we introduce general bounds on critical loss thresholds, which, in their simplest form, recover or improve on certain known bounds [14, 21]. In particular, we prove a bound on N-extendibility of channels combining white noise and loss. In the thermal-noise channel scenario, we point out that it is possible to apply directly existing bounds for channel extendibility [44] and joint-measurability [45], but we also provide in addition explicit strategies for achieving these bounds.

Finally, we examine in more detail the applications of our results to QKD. In particular, analogously to [21], we point out that in a QKD protocol where one of the parties uses a CMU, no key rate can be extracted already when the subset of measurements used to generate the key is jointly measurable. This leads us to introduce a weaker notion of joint measurability which we call 'partial joint-measurability'.

The paper is structured as follows. In Section 2, we introduce the CMU and make it clear that the concept of joint-measurability is appropriate for defining at what point a CMU loses its utility in a quantum protocol with untrusted devices and we relate it to other properties of the CMU, such as channel extendibility. In Section 3, we derive upper-bounds in

the no-click scenario and in Section 4 in the thermal-noise channel scenario. In Section 5, we discuss the applications of our bounds to QKD, define the notion of 'partial joint-measurability', and provide as an illustration upper-bounds on the key rate for various classes of DI QKD protocols. Finally, we conclude in Section 6.

2 Joint-measurability of a CMU

A CMU consists of a quantum channel C followed by a measurement device M, as in Fig. 1. The channel is described by CPTP map Φ, while the measurement is given by a set of POVMs $M_y = \{M_{b|y}\}_b$, where $M_{b|y} \geq 0$ and $\sum_b M_{b|y} = \mathbb{1}$ for all y, b. Note that we use the notation M_y to denote the POVMs and $M_{b|y}$ their elements.

From a black-box perspective, the CMU is characterized by the probabilities $P(b|y,\rho)$ of obtaining a measurement outcome b, given the measurement input y and the incoming quantum state ρ. Formally we have that

$$P(b|y,\rho) = \mathrm{Tr}\left[\Phi(\rho)\, M_{b|y}\right]. \tag{1}$$

The CMU can also be described through a set of effective POVMs with elements $M^{\Phi}_{b|y} = \Phi^*\left(M_{b|y}\right)$, where Φ^* is the dual channel of Φ. These POVMs describe the combined effect of the channel and of the measurements. The probabilities $P(b|y,\rho)$ can then also be written as

$$P(b|y,\rho) = \mathrm{Tr}\left[\rho\, M^{\Phi}_{b|y}\right]. \tag{2}$$

We would like to determine the admissible level of loss in the CMU, regardless of its usage in a larger experiment, and in a way that is independent of the state preparation, before it loses its ability to exhibit 'quantumness'. Since most experiments and protocols with untrusted devices require the distribution of entanglement and/or incompatible measurements, we could define this threshold using one of the two following natural conditions:

(1) The CMU can be replaced by an equivalent CMU for which the channel Φ is entanglement-breaking.

(2) The CMU can be replaced by an equivalent CMU for which the N measurements M_y are compatible, i.e., jointly measurable.

By 'equivalent' CMU, we mean one that may differ in implementation from the original CMU, but yields the same output probabilities $P(b|y,\rho)$ for all ρ, since from an operational and blackbox perspective, only these output probabilities are relevant.

As an example of the above definition, consider a standard bipartite nonlocality test. If either Alice's or Bob's CMU satisfies condition (1) or (2), it becomes clear that the entire experiment admits a local hidden-variable theory and no violation of a Bell inequality is possible. Similar conclusions can be drawn for steering experiments. Moreover, even protocols that do not directly rely on entanglement may require condition (1) to hold. For example, in a prepare-and-measure quantum key distribution (QKD) scheme where the channel connecting Alice to Bob is entanglement-breaking, no key rate can be extracted [46].

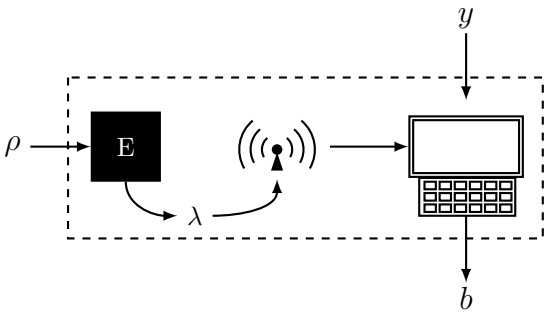

Figure 3: A CMU that satisfies conditions (3) and (4) can be implemented through a (single) quantum measurement E_λ at the entrance of the channel, whose classical output λ is then transmitted and processed to generate the output b depending on y.

More generally, the purpose of a CMU is to distribute quantum information from the entry of the channel to a remote measurement device where one of several measurements should be performed. Then an alternative, and seemingly more stringent, way to define the threshold at which a CMU becomes ineffective is the following:

(3) The CMU can be replaced by an equivalent CMU for which the channel Φ is a quantum-classical channel and the measurements M_y represent classical measurements on the output of this channel.

In a quantum experiment involving a CMU that satisfies condition (3), the quantum part of the experiment can be truncated just before the CMU, since all information processing that happens in the CMU is purely classical, see Fig. 3. This criterion is used, for instance, in [47] to determine when a routed Bell experiment can establish only short-range quantum correlations. Furthermore, in any cryptographic protocol where the CMU channel is a public channel and satisfies condition (3), an eavesdropper could have a perfect copy of the information sent in that channel since it is purely classical.

Although conditions (1), (2), (3) may appear distinct, they are actually all equivalent in an untrusted scenario, and equivalent to the following condition:

(4) The set of effective POVMs $M^{\Phi}_{b|y}$ of the CMU are jointly measurable.

The set of effective POVMs $M^{\Phi}_{b|y}$ is said to be jointly measurable [31] if there exists a *parent POVM* E with

elements E_λ, and probability distributions $p(b|y,\lambda)$ such that

$$M^{\Phi}_{b|y} = \int_\lambda d\lambda\, p(b|y,\lambda) E_\lambda\,. \tag{3}$$

Operationally, and as illustrated in Fig. 3, this means that the CMU can in principle be implemented via the following strategy. First, at the entrance of the channel, upon receiving the quantum state ρ, one measures the parent POVM E; note that this is a single (fixed) quantum measurement, hence independent of the input y. The outcome λ of the parent POVM is then transmitted through a classical channel to the remote measurement device, where a simple classical device generates the output b using the probability distribution $p(b|y,\lambda)$, which depends on y and on the classical outcome λ.

From the above operational interpretation of (4), it is immediate that conditions (3) and (4) are equivalent. It is also clear that conditions (2) and (4) are equivalent, since if the measurements M_y are jointly measurable, the same holds for the effective measurements M^{Φ}_y, and vice versa. Finally, conditions (1) and (4) are also equivalent, as shown in [48]. Indeed, Φ is entanglement-breaking if $\Phi(\rho) = \sum_\lambda \sigma_\lambda \operatorname{Tr}[E_\lambda \rho]$. But then $P(b|y,\rho) = \sum_\lambda \operatorname{Tr}[E_\lambda\rho] \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right] = \operatorname{Tr}\left[\rho M^{\Phi}_{b|y}\right]$, for $M^{\Phi}_{b|y} = \sum_\lambda \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right] E_\lambda$ which is of the form (3) if we define $P(b|y,\lambda) = \operatorname{Tr}\left[\sigma_\lambda M_{b|y}\right]$. This argument can be extended to infinite-dimensional quantum systems; in this case, an entanglement breaking channel is of the form $\Phi(\rho) = \int \mathrm{d}\lambda\, \sigma_\lambda \operatorname{Tr}\rho E_\lambda$ where the POVM is in general not atomic [49].

Note that condition (4) is not only sufficient for guaranteeing that a CMU is "useless", but also a necessary one; indeed, any set of measurements that is incompatible (in the sense of not being jointly measurable) can be used to demonstrate the effect of quantum steering [34, 35].

In the following, we will thus take condition (4) as our definition of when a CMU does not exhibit quantumness. We say that a CMU is *jointly-measurable* (*JM*) whenever condition (4) holds.

Finally, note that there are other possible conditions than (1) on the channel of a CMU that are equivalent to (4). For instance, a channel is said to be *incompatibility-breaking* if the set of measurements $\{\Phi^*(M_y)\}$ are compatible for any M_y [50]. It is said to be N-*extendable* if there exists another channel $\Phi_{1\to N}$, called the parent channel, which outputs N quantum system $S_1 \ldots, S_N$ such that

$$\operatorname{Tr}_{S_\mu|\mu\neq k} \Phi_{1\to N}(\rho) = \Phi(\rho) \tag{4}$$

for any output system S_k and any input state ρ. Let us recall that N is the number of measurements performed by our CMU. These notions give rise to the following additional conditions.

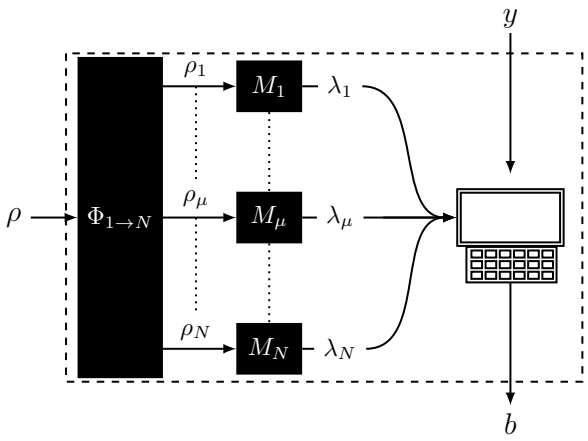

Figure 4: A CMU satisfying condition (6) can be implemented by splitting the incoming system in N systems, measuring each of the output systems with one of the N different measurements, and sending their classical outcomes to Bob.

(5) The CMU can be replaced by an equivalent CMU for which the channel Φ is incompatibility-breaking.

(6) The CMU can be replaced by an equivalent CMU for which the channel Φ is N-extendable.

The entanglement-breaking, incompatibility-breaking, and N-extendibility of a channel are in principle distinct properties. While every entanglement-breaking channel is N-extendable, and any N-extendable channel is incompatibility breaking, there exist incompatible channels that are not N-extendable, and N-extendable channels that are not entanglement-breaking. However, in the context of an untrusted CMU, the conditions (1), (5), and (6) are all equivalent to (4). This is because they are not statements about the properties of the channel in the actual realization of the CMU, but they are statements about an operationally equivalent CMU that reproduces the *combined* effect of the *channels* and of the *measurements* of the original CMU.

Specifically, since a classical channel is both incompatibility-breaking and N-extendable , we clearly have (3)⇒(5) and (3)⇒(6), and thus (4)⇒(5) and (4)⇒(6). It is also clear that (5)⇒(4). On the other hand, if (6) holds, i.e., the channel Φ is N-extendable, the effective POVMs $\Phi^*(M_{b|y})$ are jointly-measurable because they can be obtained by applying the channel $\Phi_{1\to N}$ on the input state, and then performing each of the N different measurement M_y for $y = 1, \ldots N$ on each of the N output systems, as illustrated in Fig. 4. Thus (6)⇒(4). We will make extensive use of this equivalence between the *JM* property of the CMU and N-extendibility in the remainder of the paper.

3 No-click CMUs

We first analyze a large class of CMUs taking as input quantum states ρ in a finite d-dimensional Hilbert space $\mathcal{H}$ and for which the N measurements $y = 1, \ldots, N$ yield a finite, discrete set of $L+1$ outcomes $b = 1, \ldots, L, \varnothing$, where the last outcome $\varnothing$ is a 'no-click' outcome happening when the measurement device fails to register an actual outcome, e.g. because the photon is lost. In addition to the loss, the measurements can also be affected by another type of noise.

In general, such a CMU is described by effective POVMs of the form

$$M_{b|y}^{\eta,\Phi} = \begin{cases} \eta\, \Phi^*\big(M_{b|y}\big) & \text{if } b \neq \varnothing \\ (1-\eta)\mathbb{1} & \text{if } b = \varnothing\,, \end{cases} \tag{5}$$

where $M_{b|y}$ describe the ideal POVMs, η is the *detecting efficiency*, and the CPTP map $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ models the additional noise channel (Φ^* is its adjoint). This CMU can be understood as a concatenation of the channel Φ and the *loss channel*

$$\Lambda_\eta(\rho) = \eta\rho + (1-\eta)\,|\varnothing\rangle\langle\varnothing| \tag{6}$$

acting on the state ρ before the measurements $\tilde{M}_y$ with elements

$$\tilde{M}_{b|y} = \begin{cases} M_{b|y} & \text{if } b \neq \varnothing \\ |\varnothing\rangle\langle\varnothing| & \text{if } b = \varnothing \end{cases} \tag{7}$$

are performed. Here, $|\varnothing\rangle$ is an additional state orthogonal to $\mathcal{H}$ describing the loss of a photon, and (7) extends the ideal POVMs so as to produce the no-click outcome whenever the state $|\varnothing\rangle$ is received. This allows one to write $M_{b|y}^{\eta,\Phi} = (\Lambda_\eta \circ \Phi)^*(\tilde{M}_{b|y})$. Note that whether the no-click outcome is deterministically assigned to another outcome or treated as an additional outcome does not impact the results, as this is a classical post-processing step. We will discuss again this scenario in Section 5.3.2 in the context of QKD.

For most of our applications, we will consider the noise channel Φ to be the *white-noise* channel

$$W_v(\rho) = v\,\rho + (1-v)\frac{\mathbb{1}}{d}\,, \tag{8}$$

with *visibility* v. In this case, the effective POVMs of the CMU read

$$M_{b|y}^{\eta,v} = \begin{cases} \eta\, v\, M_{b|y} + \eta\,(1-v) q_{b|y}\,\mathbb{1} & \text{if } b \neq \varnothing \\ (1-\eta)\mathbb{1} & \text{if } b = \varnothing\,, \end{cases} \tag{9}$$

where $q_{b|y} = \mathrm{Tr}\left[M_{b|y}\right]/d$.

Given the ideal POVMs $M_{b|y}$, the noise channel Φ and the efficiency η, the question of whether the effective POVMs (5) are *JM* can be formulated as a semidefinite program (SDP) [51, 31]. Furthermore, the threshold efficiency η_* below which they are *JM* can readily be determined by solving a single SDP since the above effective POVMs depend linearly on η.

In the following, we are interested in providing bounds on the critical values of η that hold for a large class of measurements M_y and noise models Φ.

3.1 Bounds on η for arbitrary sets of measurements implied by channel extendibility

In this section, we first collect some results on the N-extendibility of some families of channels. As discussed in the introduction, we then use them to imply that any CMU featuring such a channel and composed of N POVMs is *JM*. We start by introducing two general parametric families of channels.

A channel R_q with $q \in [0,1]$ is called *replacement* if it is of the form

$$R_q(\rho) = q\,\rho + (1-q)\sigma, \tag{10}$$

where σ is any fixed state (possibly orthogonal to $\mathcal{H}$). The loss channel Λ_η in Eq. (6) and the white-noise channel W_v in Eq. (8) are instances of replacement channels.

A channel Γ_v with $v \in [0,1]$ is called *convex noise* if it is of the form

$$\Gamma_v(\rho) = v\rho + (1-v)\Gamma_0(\rho), \tag{11}$$

for some Γ_0. An equivalent definition is to impose the properties $\Gamma_1 = \mathrm{id}$ and $p\,\Gamma_{v_1} + (1-p)\,\Gamma_{v_2} = \Gamma_{\bar{v}}$, where $\bar{v} = p\,v_1 + (1-p)v_2$. It is easy to see that all the replacement channels are convex noise channels. However, there are other prominent examples of convex noise channels, e.g. the dephasing channel

$$\Gamma_v(\rho) = v\rho + (1-v)\mathrm{diag}\,(\rho)\,, \tag{12}$$

where $\mathrm{diag}(\rho) = \sum_i \rho_{ii}\,|i\rangle\langle i|$ is the diagonal part of ρ in the computational basis.

With these definitions, we now formulate a few results on N-extendibility. The first one holds for all replacement channels.

Lemma 1 ([50])**.** *A replacement channel $R_q(\rho)$ as in Eq. (10) is N-extendable if $q \leq \frac{1}{N}$.*

Proof. For $q \leq \frac{1}{N}$ the channel

$$\begin{aligned} R_{1\to N}(\rho) = {} & q\,\rho\otimes\sigma^{\otimes(N-1)} + q\,\sigma\otimes\rho\otimes\sigma^{\otimes(N-2)} \\ & + \ldots + q\,\sigma^{\otimes(N-1)}\otimes\rho + (1-qN)\sigma^{\otimes N}\,. \end{aligned} \tag{13}$$

is well defined and is the parent of R_q. □

This directly implies that the loss channel Λ_η in Eq. (6) is N-extendable for $\eta \leq \frac{1}{N}$. A similar implication holds for the white-noise channel W_v in Eq. (8), however for this channel the following result gives a tighter bound.

Lemma 2 ([52]). *The white-noise channel W_v in Eq. (8) is N-extendable if*

$$v \leq \frac{N+d}{N(d+1)}. \tag{14}$$

Proof. This follows from results on universal quantum cloning. An optimal one-to-N universal cloning machine is a channel $\Phi_{1\to N}^{\mathrm{UCM}}$ from one to N qudits whose output resembles N copies of the input state as closely as possible for all possible input states [52, 53]. Furthermore, for each of the output systems, the marginal channels of Eq. (4) induced by the optimal universal cloning machine are precisely given by a white-noise channel W_v (implied by the symmetry of the task) with visibility saturating the inequality (14) [52]. More noise can always be added if v is lower (see Lemma 3 below). □

Now, following our original motivation we proceed to study the N-extendibility of concatenated channels. The next result is a very simple but useful observation.

Lemma 3. *Consider an N-extendable channel Φ and another arbitrary channel Γ. The concatenated channels*

$$\Gamma \circ \Phi \quad \textit{and} \quad \Phi \circ \Gamma \tag{15}$$

are N-extendable.

Proof. The parent channels are given by $\Gamma^{\otimes N} \circ \Phi_{1\to N}$ and $\Phi_{1\to N} \circ \Gamma$ respectively. □

Now consider a concatenation of channels with known properties.

Lemma 4. *Consider a convex noise channel Γ_v (Eq. 11) with $v \geq v_*$ and a replacement channel R_q (Eq. 10) with $q \geq q_*$, where above the threshold values the channels Γ_v and R_q are N-extendable. The concatenated channel*

$$\Phi_{q,v} = R_q \circ \Gamma_v \tag{16}$$

is N-extendable for

$$q \leq \frac{1-v_*}{(1-v)+(v-v_*)/q_*}. \tag{17}$$

Proof. First, let us write the concatenated channel explicitly as

$$\begin{aligned}\Phi_{q,v}(\rho) &= R_q\big(v\rho + (1-v)\Gamma_0(\rho)\big)\\ &= q\,v\,\rho + q(1-v)\Gamma_0(\rho) + (1-q)\sigma.\end{aligned} \tag{18}$$

By assumption, there exist parent channels $\Gamma_{1\to N}$ and $R_{1\to N}$ that reproduce the marginal channels Γ_{v_*} and R_{q_*} when tracing out $N-1$ systems. Let us now consider a mixture thereof $p\,\Gamma_{1\to N} + (1-p)\,R_{1\to N}$ for some parameter $p \in [0,1]$. It defines a parent channel which establishes the N-extendibility of $\mathcal{E}_p(\rho) = \mathrm{tr}_{N-1}(p\,\Gamma_{1\to N} + (1-p)R_{1\to N})(\rho)$ given by

$$\begin{aligned}\mathcal{E}_p(\rho) &= p\,\Gamma_{v_*}(\rho) + (1-p)R_{q_*}(\rho)\\ &= (p\,v_* + (1-p)q_*)\rho + p(1-v_*)\Gamma_0(\rho)\\ &\quad + (1-p)(1-q_*)\sigma.\end{aligned}$$

Comparing with Eq. (18) we conclude that $\Phi_{q,v}$ is identical to $\mathcal{E}_p$ if

$$\begin{aligned}1-q &= (1-p)(1-q_*)\\ q(1-v) &= p(1-v_*)\,.\end{aligned} \tag{19}$$

Solving the first equation for p gives $1-p = \frac{1-q}{1-q_*}$. Plugging this in the second equation implies

$$q = \frac{1-v_*}{(1-v)+(v-v_*)/q_*}\,. \tag{20}$$

The channel $\Phi_{q,v}$ is thus N-extendable when this equality is satisfied. To see that any channel with a lower transmission $q' \leq q$ is also N-extendable note that it can be decomposed as $\Phi_{q',v} = R_{q'/q} \circ \Phi_{q,v}$ and apply Lemma 3. □

Let us now come back to our channels of interest and apply the above Lemmas to obtain sufficient conditions for the CMUs with effective POVMs (5) and (9) to become *JM*. First, consider the concatenation of any channel Φ with the loss channel Λ_η and apply Lemmas 1 and 3 to obtain the following bound.

Upper bound 1 ([21]). *The concatenation of the loss channel with any quantum channel Φ*

$$\Lambda_\eta \circ \Phi(\rho) = \eta\,\Phi(\rho) + (1-\eta)\,|\varnothing\rangle\langle\varnothing| \tag{21}$$

is N-extendable, and thus the corresponding CMU with effective POVMs (5) is JM, if

$$\eta \leq \frac{1}{N}\,. \tag{22}$$

This is essentially a reformulation from the channel perspective of a result already obtained in [14] in the context of Bell experiments, in [21] for general one-sided-device-independent protocols, and in [50] where channels of the form $R_\eta \circ \Gamma(\rho) = \eta\,\Gamma(\rho) + (1-\eta)\sigma$ were shown to be N-incompatibility breaking for $\eta \leq \frac{1}{N}$.

Next, we also assume that the noise channel $\Phi = \Phi_v$ is convex and apply Lemmas 1 and 4 to get.

Upper bound 2. *The channel*

$$\Lambda_\eta \circ \Phi_v(\rho) = \eta\,\Phi_v(\rho) + (1-\eta)\,|\varnothing\rangle\langle\varnothing|\,, \tag{23}$$

corresponding to the concatenation of the loss channel and any convex noise channel Φ_v with $v \geq v_*$ where Φ_{v_*} is N-extendable, is N-extendable, and thus the corresponding CMU is JM, if

$$\eta \leq \frac{1-v_*}{(1-v)+N(v-v_*)}\,. \tag{24}$$

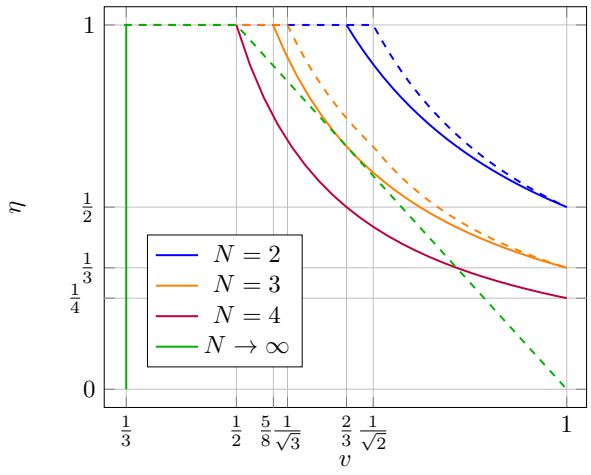

Figure 5: The full curves depict the lines below and on the left of which the channel $\Phi_{\eta,v}$ of the no-click white-noise qubit CMU is N-extendable according to Upper bound 3 for $N = 2, 3, 4$ and according to Lemma 2 for $N = \infty$. The dashed curves correspond to the additional assumption that the ideal measurements are binary and are determined according to Upper bound 6 (see section 3.2) for $N = 2, 3, 4$ and the bound (51) (see section 3.2) for $N = \infty$. Note that for the case $N = 4$ the two lines coincide.

Remarkably, this bound is useful even if we only know that the channel Φ_v becomes N-extendable at $v = 0$. Direct application of (24) shows that in this case $\Lambda_\eta \circ \Phi_v$ is N-extendable for

$$\eta \leq \frac{1}{(1-v)+Nv}. \tag{25}$$

This bound applies for example when $\Phi_v = \Gamma_v$ is the dephasing channel in Eq. (12).

Similarly, if Φ_v is a replacement channel itself (or becomes N-extendable for $v \leq \frac{1}{N}$ for another reason) the bound (24) implies that $\Lambda_\eta \circ \Phi_v$ is N-extendable for

$$\eta \leq \frac{1}{vN}. \tag{26}$$

Finally, we give special treatment to the case where $\Phi_v = W_v$ is the white-noise channel, as it will be important for applications. Combining the Upper bound 2 with Lemma 2 (giving $v_* = \frac{N+d}{N(d+1)}$) we get our last bound in this section.

Upper bound 3. *The concatenation of loss and white-noise*

$$\Lambda_\eta \circ W_v(\rho) = \eta\, v\, \rho + \eta(1-v)\frac{\mathbb{1}}{d} + (1-\eta)\,|\varnothing\rangle\langle\varnothing|\,, \tag{27}$$

is N-extendable, and thus the corresponding CMU with effective POVMs (9) is JM, if

$$\eta \leq \frac{d}{N\Big(v(d+1)-1\Big)}\,. \tag{28}$$

Note that we can make this bound dimension-independent by taking the worst-case value of the right-hand side of (28), $\eta \leq \min_{d\geq 1} d/\big(N(v(d+1)-1)\big) = \frac{1}{Nv}$, which coincides with the more general bound (26).

The fundamental trade-off between the visibility v and the detection efficiency η provided by the bound (28) is illustrated in Fig. 5 in the case of qubits.

3.2 Bounds on η for binary qubit measurements

The bounds obtained so far on η depend on generic parameters such as the number of measurements N, the level of noise v, or the dimension d of the underlying Hilbert space in the case of white noise. And as they follow from the N-extendibility of the CMU channel, they hold for any possible measurements M_y implemented by the CMU. We now take additionally into account information about the specific measurements M_y that define the CMU, focusing on the particular case of qubits.

First note the following variation of Upper bound 2.

Upper bound 2'. *Consider a no-click CMU defined in terms of a convex noise channel Φ_v with $v \geq v_*$ and ideal measurements M_y such that the effective POVMs $\Phi^*_{v_*}(M_y)$ are jointly measurable. Then the effective POVMs $\Phi^*_{\eta,v}(M_y)$ defining the no-click CMU in Eq. (5) are JM for*

$$\eta \leq \frac{1-v_*}{(1-v)+N(v-v_*)}\,. \tag{29}$$

This can be proven identically to the Lemma 4 by comparing the POVMs

$$\begin{aligned}\Phi^*_{\eta,v}(M_{b|y}) = \eta v M_{b|y} + \eta(1-v)\Phi^*_{1,0}(M_{b|y}) \\ + (1-\eta)\Phi^*_{0,1}(M_{b|y})\end{aligned} \tag{30}$$

with the family of POVMs

$$p\,\Phi^*_{1,v_*}(M_{b|y}) + (1-p)\Phi^*_{\frac{1}{N},1}(M_{b|y}), \tag{31}$$

which are *JM* by construction.

From now on we set $d = 2$ for the rest of the section. A binary qubit measurements is composed of two POVM elements $\{E_+, E_-\}$ of the form

$$E_\pm = \frac{1}{2}\big((1\pm\gamma)\mathbb{1} \pm \boldsymbol{m}\cdot\boldsymbol{\sigma}\big), \tag{32}$$

with $1-|\gamma| \leq \|\boldsymbol{m}\|$ guaranteed by positivity. If $\gamma = 0$ it is said to be unbiased, otherwise it is biased. A biased measurement can be decomposed as a mixture

$$\{E_+, E_-\} = \|\boldsymbol{m}\|\{M_+, M_-\} + (1-\|\boldsymbol{m}\|)\{R_+, R_-\}, \tag{33}$$

of a PVM (unbiased)

$$M_\pm = \frac{1}{2}(\mathbb{1} \pm \hat{\boldsymbol{m}}\cdot\boldsymbol{\sigma}) \tag{34}$$

with a unit vector $\hat{\boldsymbol{m}} = \frac{\boldsymbol{m}}{\|\boldsymbol{m}\|}$, and a "dummy measurement" $R_\pm = q_\pm \mathbb{1}$ with $q_\pm = \frac{1\pm\gamma-\|m\|}{2(1-\|m\|)}$ whose output is independent of the measured state. This observation leads to the following result.

Lemma 5. *Consider a channel Φ and a set of binary measurements $E_{\pm|y} = \frac{1}{2}\big((1 \pm \gamma_y)\mathbb{1} \pm \boldsymbol{m}_y \cdot \boldsymbol{\sigma}\big)$. The measurements set $\Phi^*(E_{\pm|y})$ is JM if the measurements $\Phi^*(M_{\pm|y})$ with $M_{\pm|y} = \frac{1}{2}\big(\mathbb{1} \pm \hat{\boldsymbol{m}}_y \cdot \boldsymbol{\sigma}\big)$ and $\hat{\boldsymbol{m}}_y = \frac{\boldsymbol{m}_y}{\|\boldsymbol{m}_y\|}$ are JM.*

Proof. By Eq. (33) the measurements $\Phi^*(E_{\pm|y})$ can be decomposed as

$$\Phi^*(E_{\pm|y}) = \|\boldsymbol{m}_y\|\Phi^*(M_{\pm|y}) + (1 - \|\boldsymbol{m}_y\|)R_{\pm|y} \quad (35)$$

where $R_{\pm|y} = \Phi^*(R_{\pm|y}) = q_{\pm|y}\mathbb{1}$ (the adjoint of a CPTP map is unital).

Since the measurements $\Phi^*(M_{\pm|y})$ are *JM* by assumption there exists a parent POVM E_λ simulating them. Then the measurements $\Phi^*(E_{\pm|y})$ are simulated by the following strategy. Perform the parent POVM E_λ. Upon receiving the setting y simulate $\Phi^*(M_{\pm|y})$ with probability $\|\boldsymbol{m}_y\|$, or with probability $1 - \|\boldsymbol{m}_y\|$ sample a random output from $q_{\pm|y}$. □

It implies that when discussing the *JM* of sets of *ideal* binary measurements subject to noise and loss, biased measurements can be ignored. That is, if we show that the set $\Phi^*_{\eta,v}(M_{\pm|y})$ with

$$M_{\pm|y} = \frac{1}{2}(\mathbb{1} \pm \hat{\boldsymbol{m}}_y \cdot \boldsymbol{\sigma}), \quad (36)$$

and $\|\hat{\boldsymbol{m}}_y\| = 1$ is *JM* it automatically implies that any set of (biased) measurements $\Phi^*_{\eta,v}(E_{\pm|y})$ with $\boldsymbol{m}_y$ parallel to $\hat{\boldsymbol{m}}_y$ is also *JM*. We will thus restrict our consideration to N measurements of the form (36) in the rest of the section. Furthermore, we now set the channel Φ_v to be the white-noise channel W_v mapping PVMs to

$$W_v^*(M_{\pm|y}) = \frac{1}{2}(\mathbb{1} \pm v\,\hat{\boldsymbol{m}}_y \cdot \boldsymbol{\sigma}), \quad (37)$$

which are also unbiased.

The question of the joint measurability of unbiased qubit measurements has been studied extensively, see e.g. section III.A in the review [54].

In particular, from [55] it follows that any set of $N = 2$ measurements of the form (37) is *JM* if and only if

$$v \le v_*(\hat{\boldsymbol{m}}_1, \hat{\boldsymbol{m}}_2) := \frac{2}{\|\hat{\boldsymbol{m}}_1 + \hat{\boldsymbol{m}}_2\| + \|\hat{\boldsymbol{m}}_1 - \hat{\boldsymbol{m}}_2\|}. \quad (38)$$

For $N = 3$, the measurements are *JM* if and only if [56, 57]

$$v \le v_*(\hat{\boldsymbol{m}}_1, \hat{\boldsymbol{m}}_2, \hat{\boldsymbol{m}}_3) := \frac{4}{\sum_{k=0}^{3} \|\boldsymbol{t}_k - \boldsymbol{t}_{FT}\|}, \quad (39)$$

where $\boldsymbol{t}_0 = \hat{\boldsymbol{m}}_1 + \hat{\boldsymbol{m}}_2 + \hat{\boldsymbol{m}}_3$, $\boldsymbol{t}_k = 2\hat{\boldsymbol{m}}_k - \boldsymbol{t}_0$ for $k \ge 1$, and $\boldsymbol{t}_{FT}$ is the so-called Fermat-Toricelli vector of the set $\{\hat{\boldsymbol{m}}_1, \hat{\boldsymbol{m}}_2, \hat{\boldsymbol{m}}_3\}$, i.e. $\boldsymbol{t}_{FT} = \mathrm{argmin}_{\boldsymbol{t}} \sum_{k=1}^{3} \|\hat{\boldsymbol{m}}_k - \boldsymbol{t}\|$

For any set of N unbiased qubit measurements, we now show the following result.

Lemma 6. *A set of N unbiased qubit measurements $M_{\pm|y} = \frac{1}{2}(\mathbb{1} \pm \boldsymbol{m}_y \cdot \boldsymbol{\sigma})$ specified by the vectors $\boldsymbol{m}_1, \ldots, \boldsymbol{m}_N$ with $\|\boldsymbol{m}_y\| \le 1$ is JM if*

$$\sum_{\boldsymbol{a}} \|\sum_{k=1}^{N} (-1)^{a_k} \boldsymbol{m}_k\| \le 2^N \quad (40)$$

where $\boldsymbol{a} = (a_1, \ldots, a_N) \in \{0,1\}^N$ runs through the 2^N bitstrings. (See also the relaxed condition in Eq. (48)).

Proof. Define the following set of 2^N vectors

$$\boldsymbol{w}_{\boldsymbol{a}} = (-1)^{a_1}\boldsymbol{m}_1 + (-1)^{a_2}\boldsymbol{m}_2 + \cdots + (-1)^{a_N}\boldsymbol{m}_N \quad (41)$$

labeled by the bitstrings $\boldsymbol{a} = (a_1, \ldots, a_N) \in \{0,1\}^N$ and denoted their normalized versions as $\hat{\boldsymbol{w}}_{\boldsymbol{a}} = \frac{\boldsymbol{w}_{\boldsymbol{a}}}{\|\boldsymbol{w}_{\boldsymbol{a}}\|}$, which define 2^N projective measurements. Using the probability distribution $p(\boldsymbol{a}) = \frac{\|\boldsymbol{w}_{\boldsymbol{a}}\|}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|}$ define the parent POVM composed of 2^{N+1} elements

$$E_{\pm,\boldsymbol{a}} = \frac{p(\boldsymbol{a})}{2}(\mathbb{1} \pm \hat{\boldsymbol{w}}_{\boldsymbol{a}} \cdot \boldsymbol{\sigma}) = \frac{1}{2}\left(p(\boldsymbol{a})\mathbb{1} \pm \frac{\boldsymbol{w}_{\boldsymbol{a}} \cdot \boldsymbol{\sigma}}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|}\right). \quad (42)$$

With the post-processing

$$p(b|y, \pm, \boldsymbol{a}) = \begin{cases} 1 & \pm 1 = \mathrm{sign}(b\,(-1)^{a_y}) \\ 0 & \text{otherwise} \end{cases} \quad (43)$$

for $b = \pm 1$, the parent POVMs can simulate all measurements of the form

$$\tilde{M}_{b|y} = \sum_{\pm,\boldsymbol{a}} E_{\pm,\boldsymbol{a}}\, p(b|y, \pm, \boldsymbol{a}) = \sum_{\boldsymbol{a}} E_{\mathrm{sign}(b(-1)^{a_y}),\boldsymbol{a}} \quad (44)$$

$$= \frac{1}{2}(\mathbb{1} + b\frac{2^N}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|}\boldsymbol{m}_k \cdot \boldsymbol{\sigma}). \quad (45)$$

These measurements are identical to $M_{b|y}$ if $\frac{2^N}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|} = 1$, in which case the proof is done. If this ratio is even larger $\frac{2^N}{\sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|} \ge 1$, one can mix some randomness into $\tilde{M}_{b|y}$ to complete the proof. □

Applying this Lemma to the measurements $W_v^*(M_{\pm|y})$ in Eq. (37) with $\boldsymbol{m}_y = v\,\hat{\boldsymbol{m}}_y$ gives the follwing condition for their joint measurability

$$v \le v_\star(\hat{\boldsymbol{m}}_1, \ldots \hat{\boldsymbol{m}}_N) := \frac{2^N}{\sum_{\boldsymbol{a}} \|\sum_{k=1}^{N} (-1)^{a_k} \hat{\boldsymbol{m}}_k\|} \quad (46)$$

Note that Eq. (38) is a particular case of Eq. (46) for $N = 2$, while it is unclear to us if this is also the case

for Eq. (39). An attentive reader also noticed that for general N we used $v_\star$ instead of v_* to indicate that the bound is potentially not tight.

The threshold visibilities for two, three and N measurements in Eqs. (38,39,46) can be used together with the Upper bound 2' to obtain sufficient conditions on the joint measurability of the corresponding CMUs $M^{\eta,v}_{b|y}$. The resulting bounds would depend on the specific choice of binary the qubit measurements described by the vectors $\hat{\boldsymbol{m}}_1, \ldots, \hat{\boldsymbol{m}}_N$.

Nevertheless, it is also possible to derive a simple bound for any set of N binary qubit measurements by noting that the rhs of Eq. (46) satisfies

$$v_*(\hat{\boldsymbol{m}}_1, \ldots, \hat{\boldsymbol{m}}_N) \geq \frac{1}{\sqrt{N}}. \tag{47}$$

To see this first use the concavity of the square root to obtain $\sum_{\boldsymbol{a}} \frac{1}{2^N} \|\boldsymbol{w}_{\boldsymbol{a}}\| = \sum_{\boldsymbol{a}} \frac{1}{2^N} \sqrt{\|\boldsymbol{w}_{\boldsymbol{a}}\|^2} \leq \sqrt{\frac{1}{2^N} \sum_{\boldsymbol{a}} \|\boldsymbol{w}_{\boldsymbol{a}}\|^2}$. Then in the last term simply compute $\sum_{\boldsymbol{a}} \| \sum_{k=1}^N (-1)^{a_k} \hat{\boldsymbol{m}}_k \|^2 = \sum_{\boldsymbol{a}} \sum_{kj=1}^N (-1)^{a_k + a_j} \hat{\boldsymbol{m}}_k \cdot \hat{\boldsymbol{m}}_j = \sum_{\boldsymbol{a}} \sum_{k=1}^N \hat{\boldsymbol{m}}_k \cdot \hat{\boldsymbol{m}}_k = 2^N N$, which implies $\sqrt{\sum_{\boldsymbol{a}} \frac{1}{2^N} \| \sum_{k=1}^N (-1)^{a_y} \hat{\boldsymbol{m}}_k \|^2} = \sqrt{\frac{1}{2^N} 2^N N} = \sqrt{N}$ and the inequality (47). This bound is, however, only saturable for $N = 2$ and 3 by the measurements corresponding to MUBs. Remarkably the same arguments imply $\frac{1}{2^N} \sum_{\boldsymbol{a}} \| \sum_{k=1}^N (-1)^{a_k} \boldsymbol{m}_k \| \leq \sqrt{\sum_k \|\boldsymbol{m}_k\|^2}$, hence

$$\sum_{k=1}^N \|\boldsymbol{m}_k\|^2 \leq 1 \tag{48}$$

can be used as a relaxation of the condition (40) in Lemma 6.

To conclude this section we combine the Upper bound 2', the Lemmas 5 with 6, and the bound (47) to obtain the following simple sufficient condition for joint measurability of our CMUs.

Upper bound 6. *The no-click CMU implementing N binary qubit measurements (9,36) with white-noise visibility v and detection efficiency η is JM if*

$$\eta \leq \frac{1}{\sqrt{N}\left((\sqrt{N}+1)v - 1\right)}. \tag{49}$$

It is easy to see that for $N \geq 4$ this condition does not bring any improvement over the general upper bound (28), which reduces in the case $d = 2$ to

$$\eta \leq \frac{2}{N\left(3v - 1\right)}. \tag{50}$$

In contrast, for $N = 2$ and 3, it improves on this bound, as illustrated in Fig. 5. We can also demonstrate numerically that in these cases the bound (49) is tight. Indeed, selecting σ_z and σ_x as the two measurement directions in the case $N = 2$, and additionally σ_y in the case $N = 3$, we can determine, for any fixed v, the maximal value of η such that the resulting effective POVMs are *JM* by solving a single SDP, as noted below Eq. (5). We observed that for all the tested values v, the maximal transmission η found by the SDP matches with the upper bound of Eq. (49) up to numerical precision.

3.3 Bound on η independent of N

Finally, let us mention the asymptotic case where the number of measurements N is unbounded. In [43] a necessary and sufficient condition for the joint measurability of all d-dimensional projective measurements subject to white noise and loss was derived. In the case of qubits, this condition takes a simple form

$$\eta \leq 2(1 - v), \tag{51}$$

depicted in Fig. 5. This bound improves over the upper bound (50) whenever $(1 - v)(3v - 1) > \frac{1}{N}$. By the Lemma 5 this result is promoted to all binary measurements.

For all POVMs, a sufficient condition is given by $\eta \leq (1 - v)^{d-1}$ [42]. When N is large these bounds can be tighter and used instead of the bound (28).

4 Thermal-noise channel scenario

The bounds obtained in the previous section are based on the assumption that the measurements of the CMU have a finite number of discrete outcomes and that a photon loss can be modeled as a no-click outcome. However, some common quantum optics measurements are not of that form. Quadrature measurements are an example of measurements with an infinite, continuous set of outcomes, and which always produce an outcome, even when the photon is lost.

In this section, we thus make no hypothesis at all on the measurement device and we assume that losses in the channel are modeled, as is usual in quantum optics, through a thermal-noise (or attenuator) channel [58, 59] characterized through the *transmittance* η and the *excess noise* ϵ. Specifically, the thermal-noise channel can be viewed as combining on a beam-splitter of transmittance η, the incoming state ρ and a thermal state τ_ν with mean photon number $\nu = (\eta\epsilon)/(2(1 - \eta))$, as illustrated in Fig. 6. The resulting CPTP map is

$$\Phi_{\eta,\epsilon}(\rho) = \mathrm{Tr}_2\left(U_\eta \rho \otimes \tau_\nu U_\eta^\dagger\right), \tag{52}$$

where

$$\tau_\nu = \frac{1}{\pi\nu} \int_{\mathbb{C}} d^2\beta e^{-|\beta|^2/\nu} |\beta\rangle\langle\beta|, \tag{53}$$

is a thermal state written in the basis of coherent states $|\beta\rangle$, and

$$U_\eta = e^{\arccos(\sqrt{\eta})(\hat{a}^\dagger \hat{b} - \hat{a}\hat{b}^\dagger)}, \tag{54}$$

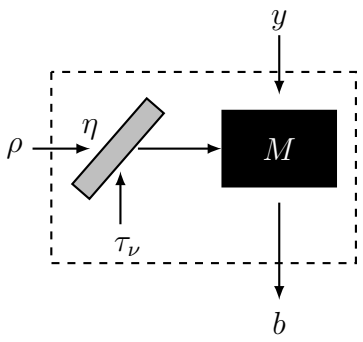

Figure 6: CMU where the channel is a thermal-noise channel which can be viewed as combining ρ with a thermal state τ_ν at a beam-splitter of transmittance η.

represent the action of the beam-splitter, where $\hat{a}$ and $\hat{b}$ are annihilation operators acting on the input state ρ and the thermal state τ_ν, respectively.

Upper bound 7. *The channel $\Phi_{\eta,\epsilon}$ of a thermal-noise CMU is N-extendable, and thus the corresponding CMU is JM irrespective of the measurements M_y, if*

$$\eta \leq \frac{1}{N(1-\epsilon/2)}\,. \tag{55}$$

Proof. Consider the channel $\Phi_{1\to N} = \Phi_{\mathrm{BS}_N} \circ \Phi_{\mathrm{Amp}}$ corresponding to first amplifying the input state ρ with a gain $G > 1$ and then splitting it using a symmetric N-mode beam-splitter, as illustrated in Fig. 7. We now show that if we trace out $N-1$ output modes at the exit of this channel, we get the thermal-noise channel $\Phi_{\eta,\epsilon}$ by choosing appropriately G.

The thermal-noise channel, the amplifier, and the beam-splitter are Gaussian channels, and their action can be fully described in terms of two real matrices, the scaling and noise matrices X and Y [58]. These matrices capture the evolution of the characteristic function of the system, regardless of whether the initial state is Gaussian or not. For the thermal-noise channel, the amplifier, and the channel corresponding to tracing out all but one ouptput mode of a symmetric N-mode beam-splitter, these matrices are respectively

$$X_{\mathrm{Th}} = \sqrt{\eta}\mathbb{1}, \qquad Y_{\mathrm{Th}} = (1-\eta+\epsilon\eta)\mathbb{1}\,, \tag{56}$$

$$X_{\mathrm{Amp}} = \sqrt{G}\mathbb{1}_2, \qquad Y_{\mathrm{Amp}} = (G-1)\,\mathbb{1}_2, \tag{57}$$

and

$$X_{\mathrm{BS}} = \sqrt{\frac{1}{N}}\mathbb{1}_2, \qquad Y_{\mathrm{BS}} = \frac{N-1}{N}\mathbb{1}_2,. \tag{58}$$

In the Gaussian systems framework, the channel obtained by applying two channels in succession, each described by the matrices $X_{A,B}$ and $Y_{A,B}$, is described by the matrices

$$X_{B\circ A} = X_B X_A, \tag{59}$$

$$Y_{B\circ A} = X_B^T Y_A X_B + Y_B. \tag{60}$$

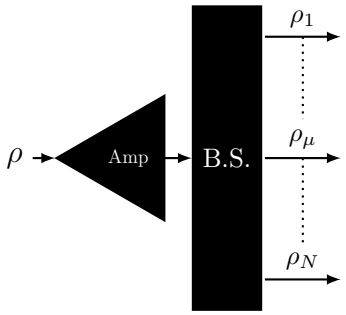

Figure 7: Strategy for establishing the N-extendibility of the thermal-noise-channel. The initial state ρ is amplified and split into N modes using a beam splitter.

We thus have

$$X_{\mathrm{BS}\circ\mathrm{Amp}} = X_{\mathrm{BS}}X_{\mathrm{Amp}} = \sqrt{\frac{G}{N}}\mathbb{1}_2\,, \tag{61}$$

$$Y_{\mathrm{BS}\circ\mathrm{Amp}} = X_{\mathrm{BS}}^T Y_{\mathrm{Amp}} X_{\mathrm{BS}} + Y_{\mathrm{BS}} = \frac{G-N-2}{N}\mathbb{1}_2\,.$$

These matrices are equal to the target matrices X_{Th} and Y_{Th} for $G = 1/(1-\epsilon/2)$ and

$$\eta = \frac{1}{N(1-\epsilon/2)}. \tag{62}$$

Of course, lower values of η are also possible by simply attenuating the overall channel output, resulting in the bound (55). □

We note that the above result was also obtained in [44], where a necessary and sufficient condition for the N-extendibility of single-mode Gaussian channels is presented. The above proof, however, provides an explicit strategy for establishing the N-extendibility of the thermal-noise channel.

While we considered a single-mode thermal-noise channel $\Phi_{\eta,\epsilon}$, the argument extends to n-mode scenarios, where the global noise channel is the product $\bar{\Phi}_{\eta,\epsilon} = \bigotimes_{i=1} \Phi_{(\eta,\epsilon)}$. It is easy to see that this channel $\bar{\Phi}_{\eta,\epsilon}$ is N-extendable if $\Phi_{\eta,\epsilon}$ is. Therefore the Upper bound 7 also holds in this scenario.

The bound (55) is completely generic as it depends only on the number of measurements N and the excess noise ϵ, and holds for any possible measurements M_y implemented by the CMU. A much more stringent bound can be obtained if we assume that the measurements M_y are Gaussian measurements, which can be implemented by first applying a Gaussian channel to the state and then performing homodyne detection [60].

Upper bound 8 ([45])**.** *A CMU consisting of a thermal-noise channel followed by Gaussian measurements is JM if*

$$\eta \leq \frac{1}{2-\epsilon}\,. \tag{63}$$

Remarkably, the above bound on the transmissivity η is independent of the total number N of measurements performed. It follows from the fact that the thermal-noise channel is incompatibility breaking for Gaussian measurements as shown in [45]. We provide a simple proof of this bound in Appendix A in the case of pure homodyne measurements.

5 Applications to QKD protocols

Our approach allows us to put bounds on the performance of any QKD protocol involving a CMU, in particular within the DI or SDI approaches. Consider a protocol in which one of the parties, say Bob, generates their copy of the secret key by performing an untrusted measurement on a quantum state received from an untrusted channel, i.e., the key is obtained by post-processing the input y and output b of a CMU. Then no key can be extracted from the protocol if the effective POVMs of the CMU are JM. This follows from the fact that in any CMU that is JM, Alice and Bob cannot certify the presence of entanglement, which is a necessary condition for the security of QKD [46].

This result can also easily be directly proven. Let $B' = BY$ where B and Y are the random variables describing the input and output of Bob's measuring device, let A denote Alice's registers used in the QKD protocol and $\mathcal{E}$ be Eve's classical register (potentially obtained by measuring a quantum system she holds). It is known that for arbitrary post-processing of A and B', the key rate that can be extracted is upper-bounded by the intrinsic information [61]

$$I(A:B'\downarrow\mathcal{E}) \equiv \min_{\mathcal{E}\to\mathcal{F}} I(A:B'|\mathcal{F}), \tag{64}$$

where $\mathcal{F}$ runs through all post-processing of Eve's register $\mathcal{E}$ and $I(A:B'|\mathcal{F}) = \sum_f p(f) I(A:B'|\mathcal{F}=f)$ is the conditional mutual information.

Now consider the case where Bob's CMU is JM. Then, as Eve controls the public channel of the CMU, we can assume that she implements herself the parent POVM E, obtains the output Λ, keeps a copy for her, and sends the other copy to Bob's apparatus that uses it to determine B given Y. Hence, the probability distribution conditional on Eve's register $\mathcal{E} = \Lambda$ factorizes

$$p(A=a, B'=by|\mathcal{E}=\lambda) = p(a|\lambda)p(y)p(b|y,\lambda) \tag{65}$$

and $I(A:B'\downarrow\mathcal{E}) \leq I(A:B'|\mathcal{E}) = 0$, proving that no key rate can be extracted.

We conclude from the above discussion that in any DI or SDI QKD protocol where Bob's device is untrusted and his inputs and outputs are used to establish the final key, no key can be extracted if the losses and noise satisfy the various bounds presented above. However, these bounds can be further improved by using the concept of partial-joint-measurability, which we introduce in the next section.

5.1 Partial-joint measurability

In many QKD protocols, the key on Bob's side is obtained by post-processing the input y and output b only of a subset $\mathcal{K} \subset \{1,\dots,N\}$ of cardinality $K = |\mathcal{K}| < N$ of all possible measurement inputs, while other measurements are solely used for parameter estimation. In this situation, we can weaken the notion of JM through the following definition.

We say that the POVMs M_y for $y \in \{1,\dots,N\}$ are $\mathcal{K}$-JM, where $\mathcal{K}$ is subset of $\{1,\dots,N\}$, if there exists *i)*, a quantum instrument $I = \{I_\lambda\}_\lambda$ with classical measurement outcomes λ, *ii)* probability distributions $p(b|y,\lambda)$, and *iii)* measurement $M_{y,\lambda}$ such that, for any state ρ, the probabilities $p(b|y,\rho) = \mathrm{Tr}\left[\rho M_{b|y}\right]$ can be written as

$$P(b|y,\rho) = \begin{cases} \sum_\lambda p(b|y,\lambda)\,\mathrm{Tr}\left[I_\lambda(\rho)\right] & \text{if } y\in\mathcal{K} \\ \sum_\lambda \mathrm{Tr}\left[I_\lambda(\rho) M_{b|y,\lambda}\right] & \text{if } y\notin\mathcal{K}. \end{cases} \tag{66}$$

Operationally, this means that instead of carrying out the original POVMs M_y on the state ρ, one can instead first perform the parent instrument I on the state ρ, obtaining a classical output λ and a quantum output $I_\lambda(\rho)$. If $y \in \mathcal{K}$, then one only uses the classical output λ to generate the outcome b with probability $p(b|y,\lambda)$, as in the case of full JM. If $y \notin \mathcal{K}$, then one generates b by carrying out on the state $I_\lambda(\rho)$ a measurement defined by the operators $\{M_{b|y,\lambda}\}$, which depend on y, but also on λ.

Since the relations (66) should hold for any input state ρ, they can be equivalently formulated as

$$M_{b|y} = \begin{cases} \sum_\lambda p(b|y,\lambda) I^*_\lambda(\mathbb{1}) & \text{if } y\in\mathcal{K} \\ \sum_\lambda I^*_\lambda(M_{b|y,\lambda}) & \text{if } y\notin\mathcal{K} \end{cases}, \tag{67}$$

where I^*_λ denotes the adjoint of I_λ and $I^*_\lambda(\mathbb{1}) = E_\lambda$ defines a parent POVM as in the case of full JM.

In the context of QKD applications, when a CMU is $\mathcal{K}$-JM, we can assume that the parent instrument I_λ is performed by Eve, which can store a copy of its classical output λ. If Bob's register $B' = BY$ used to distill the final key only depends on the inputs $y \in \mathcal{K}$ (while the other inputs may still be used for parameter estimation) the intrinsic information is zero

$$I(A:B'\downarrow\mathcal{E}) = 0\,. \tag{68}$$

This is because, analogously to the discussion leading to Eq. (65), the probability distribution of Alice's and Bob's random variable factorizes $p(A,B'|\mathcal{E}=\lambda) = p(A|\mathcal{E}=\lambda)p(B'|\mathcal{E}=\lambda)$ when conditioned on λ, which is held by Eve. Hence, if the CMU gives rise to $\mathcal{K}$-JM measurements, no key can be distilled from the outputs of any measurements in $\mathcal{K}$.

Note that in the case where $K = N-1$, partial joint-measurability becomes equivalent to full joint-measurability. This is because, for $K = N-1$, there is only a single measurement setting $y \notin \mathcal{K}$. Therefore, there is no need to send any quantum state to Bob's

device as this single measurement can immediately be performed after the parent instrument. The concatenation of the parent instrument and this subsequent measurement represents a parent POVM providing a classical outcome for all values of y, i.e., it implies that the set of Bob's POVM is fully JM.

In the case of the no-click scenario, the following result relates the full joint-measurability of a subset $\mathcal{K}$ of the measurements to the partial joint-measurability of the full set of measurements.

Upper bound 9. *Consider a no-click CMU with an arbitrary number N of inputs and assume that for a subset $\mathcal{K} \subseteq \{1, \ldots, N\}$, the measurements $\{\Phi_\eta^*(M_y)\}_{y\in\mathcal{K}}$ are JM for a detection efficiency η. Then the full CMU is $\mathcal{K}$-JM for detection efficiency $\eta' \leq \eta/(1+\eta)$.*

Proof. By assumption, the set $\{\Phi_\eta^*(M_y)\}_{y\in\mathcal{K}}$ is *JM*, which means that there exists a parent instrument $\mathcal{E}$ that simulates it. Define then the following strategy. With probability q, perform the parent POVM $\mathcal{E}$ and follow the corresponding simulation scheme if $y \in \mathcal{K}$, while output no-click $\varnothing$ if $y \notin \mathcal{K}$. With probability $1-q$, leave the system untouched, output no-click $\varnothing$ if $y \in \mathcal{K}$, and perform the measurement $\Phi^*(M_y)$ if $y \notin \mathcal{K}$. On average this strategy realizes the POVMs with elements

$$\tilde{M}_{b|y} = \begin{cases} q\eta\Phi^*(M_{b|y}) & \text{if } b \neq \varnothing \\ (q(1-\eta) + (1-q))\,\mathbb{1} & \text{if } b = \varnothing \end{cases} \tag{69}$$

if $y \in \mathcal{K}$, and

$$\tilde{M}_{b|y} = \begin{cases} (1-q)\Phi^*(M_{b|y}) & \text{if } b \neq \varnothing \\ q\mathbb{1} & \text{if } b = \varnothing \end{cases} \tag{70}$$

if $y \notin \mathcal{K}$. These measurements have the form of the no-click CMU (5) if $q\eta = (1-q) = \eta'$, which implies $\eta' = \eta/(1+\eta)$. □

As an illustration, if we apply the above construction to the bound (49), we obtain that a white-noise no-click CMU implementing the qubit measurements (36) is $\mathcal{K}$-JM if

$$\eta \leq \frac{1}{v(K+\sqrt{K})}, \tag{71}$$

which improves the original bound when $K \leq N-1$.

Note that by construction, the condition of partial joint measurability is weaker than that of joint measurability. The bound (71) allows us to illustrate this difference explicitly. Indeed, we mentioned below the bound (49) that for $N = 3$, it is tight. In particular, the three MUBs measurements $\sigma_x, \sigma_y, \sigma_z$ are incompatible if $\eta > 2/(9v-3)$. However, from (71) these three measurements are 1-*JM* up to $\eta \leq 1/(2v)$. For instance, for $v = 1$, the three MUB measurements are incompatible but 1-*JM* in the region $\frac{1}{3} < \eta \leq \frac{1}{2}$.

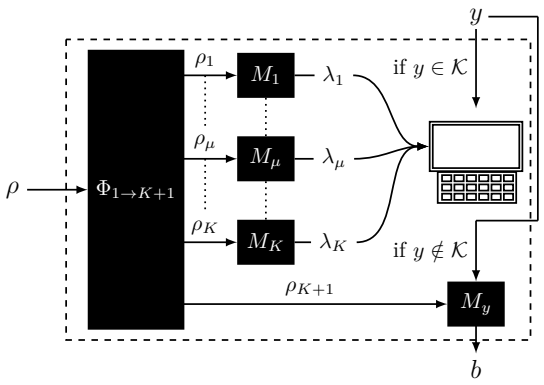

Figure 8: Consider a CMU that admits an implementation with a channel Φ that is $K+1$-extendable. Perform each of the K measurements M_y for $y \in \mathcal{K}$ in the first K extensions, resulting in outcomes $\lambda_1, \ldots, \lambda_K$, and forward the $K+1$th extension to the measurement device. Then by simply outputting λ_y if $y \in \mathcal{K}$ and performing the measurement M_y on the $K+1$th extension if $y \notin \mathcal{K}$ defines a CMU with effective POVMs that are $\mathcal{K}$-JM.

The construction above can also be applied to, e.g., the generic white-noise no-click bound (28). However in this case a better improvement can be obtained using the following relation between partial joint measurability and channel extendibility.

Lemma 7. *If a channel Φ is $(K+1)$-extendable, then the effective POVMs $\Phi^*(M_y)$ are $\mathcal{K}$-JM for any subset $\mathcal{K} \subseteq \{1, \ldots, N\}$ of K measurements.*

To see this it is sufficient to apply, as illustrated in Fig. 8, the extended channel $\Phi_{1\to K+1}$, perform the intended measurements M_y for y in $\mathcal{K}$ on the first K systems and leave the last one untouched. This defines a quantum instrument that satisfies all the required properties.

The above observation can be applied to the bounds (28) and (55) and implies that the white-noise no-click CMU is $\mathcal{K}$-JM if

$$\eta \leq \frac{d}{(K+1)\,(v(d+1)-1)}, \tag{72}$$

and that the thermal-noise CMU is $\mathcal{K}$-JM if

$$\eta \leq \frac{1}{(K+1)(1-\epsilon/2)}, \tag{73}$$

which both improve (or equal) the original bounds when $K \leq N-1$.

In the context of DIQKD or semi-DI QKD where the key is generated from a subset $\mathcal{K}$ of the measurements, the bounds (71), (72), and (73) can be used to strengthen the critical levels of losses and noises beyond which no key can be generated whenever $K \leq N-1$. For instance, in the thermal-noise channel scenario, no key can be generated if $\eta < 1/(2-\epsilon)$ whenever the key is generated from a single measurement, independently of the total number N of measurements used in the protocol, including those used

for parameter estimation. While in the white-noise no-click scenario, we deduce from (72) that no key can be generated for $K = 1$ if $\eta \leq 1/(3v - 1)$, corresponding, e.g., to $\eta \leq 0.5235$ for $v = 0.97$.

5.2 Convex combination attacks for public communication from Bob to Alice

As we explained above in any QKD protocol involving Alice and Bob and where Bob's quantum device is an untrusted CMU, no security can be guaranteed if the CMU is (partially)-*JM*, since Eve can then have full information about Bob's output, i.e., Alice and Bob are decorrelated conditioned on Eve's knowledge. This is true independently of the protocol used to distill the shared key from Alice and Bob's generated data. However, as Eve's *JM* attack targets Bob's device, one would intuitively expect protocols involving one-way public communication from Bob to Alice for key distillation to be more sensitive to such attacks. Indeed, for such protocols, the maximum achievable key rate is given by the Devetak-Winter formula [62, 63]

$$r_\infty = H(B|E) - H(B|A) \tag{74}$$

where $H(X|Y)$ is the von Neumann entropy of X conditioned on Y. This key rate is zero as long as Eve's uncertainty $H(B|E)$ on Bob's symbols is lower than Alice's uncertainty $H(B|A)$, thus Eve's does not necessarily need to have full information about Bob's raw key for the protocol to be insecure.

This leads us to consider a class of *convex combination attacks* directly inspired by a similar strategy proposed in [64] in the context of DIQKD[1]. In such attacks, Eve employs a mixture of two strategies: either, with probability p, she implements a lossy and noisy version of the expected measurements performed by Bob, in which case she exploits the fact that this implementation is (partially) *JM* to get full information about Bob's outcome. Or with probability $1 - p$, she replaces Bob's CMU with an ideal, loss-free, and noise-free CMU, in which case she has no information about Bob's outcome. This mixed strategy allows her to reproduce the actual CMU used by Bob beyond the admissible level of losses and noises derived in the previous sections, while at the same time providing her with enough information to make $r_\infty = H(B|E) - H(B|A) = 0$.

[1]In the case of DIQKD, our attacks improve on the ones of [64] because they take into account not only losses but also white-noise. However, we can further improve these attacks by targeting both CMUs of a DIQKD protocol, see next subsection. We also remark that it is important to consider separately the case where a no-click outcome is kept as a distinct outcome in Alice and Bob's post-processing and the case where it is discarded. In [64] only the former case is analyzed. An example of a DIQKD protocol where binning of the no-click outcome leads to a positive key rate for a detection efficiency that is lower than the threshold computed in [64] can be found in [82].

We consider such attacks only in the no-click scenario. Indeed, they cannot straightforwardly be applied to CMUs with a thermal-noise channel, since such channels are not closed under convex combinations. Specifically, we show the following in Appendix B.

Lemma 8. *A thermal-noise channel $\Phi_{\eta,\epsilon}$ that is not N-extendable cannot be decomposed as*

$$\Phi_{\eta,\epsilon} = p\,\Phi_G + (1-p)\Phi' \qquad (0 < p < 1) \tag{75}$$

where Φ_G is a N-extendable Gaussian channel and Φ' is an arbitrary channel.

We prove this observation by showing that the Q-function of the channel $\Phi' = (\Phi_{\eta,\epsilon} - p\,\Phi_G)/(1-p)$ (for a coherent state input) can not be positive for such a decomposition.

In the following, we thus focus on QKD protocols with one-way communication from Bob to Alice and involving a white-noise no-click CMU represented by effective POVMs $\{M_y^{\eta,v}\}$ of the form (9). We assume that this set of POVMs is neither *JM*, nor $\mathcal{K}$-JM, where $\mathcal{K}$ is the set of inputs used to establish the raw key, since otherwise we have already shown that no key can be extracted.

Consider now an attack, where, with probability p, Eve uses a strategy S_* and, with probability $1 - p$, a strategy S_1. In the former case, Bob's untrusted measurement device implements the POVMs $M_y^{\eta^*,v^*}$, where η^* and v^* are chosen such that these POVMs are $\mathcal{K} - JM$ so that Eve has full information about Bob's outcome, i.e., $H(B|E,S_*) = 0$. In the latter case, Bob's measurement device implements the ideal POVMs $M_y^{1,1}$, which may prevent Eve from having any non-trivial information about his outcome. Conservatively, we can upper-bound Eve's information by $H(B|E,S_1) \leq H_{1,1}(B)$, where $H_{1,1}(B)$ is the entropy of Bob's outcome, averaged over the inputs $y \in \mathcal{K}$, when Bob performs ideal POVMs $M_y^{1,1}$.

Altogether Eve's information on Bob's outcome is then given by

$$\begin{aligned} H(B|E) &= pH(B|E,S_*) + (1-p)H(B|E,S_1) \\ &\leq (1-p)H_{1,1}(B)\,, \end{aligned} \tag{76}$$

hence the key rate is upper-bounded by

$$r_\infty \leq (1-p)H_{1,1}(B) - H_{\eta,v}(B|A)\,, \tag{77}$$

where p remains to be determined, and where $H_{\eta,v}(B|A)$ is computed from the actual correlations between Alice and Bob's outcomes, characterized by an efficiency η and visibility v.

This attack must reproduce on average Bob's actual POVMs $M_{b|y}^{\eta,v}$. Thus, it must satisfy

$$M_{b|y}^{\eta,v} = pM_{b|y}^{\eta_*,v_*} + (1-p)M_{b|y}^{1,1}, \tag{78}$$

which is equivalent to the following system of equations

$$\begin{aligned} p(1-\eta_* v_*) &= 1-\eta v \\ p\,(1-\eta_*) &= 1-\eta\,. \end{aligned} \tag{79}$$

We have seen in the previous sections, that the POVMs $M_y^{\eta^*,v^*}$ are $\mathcal{K}$-JM if η^* satisfies a bound $\eta^* \leq \hat{\eta}_{N,K,d}(v^*)$ which depends on the visibility v^*, the number of measurements N or K, the Hilbert space dimension d, and other properties of the measurements. Eve will pick the transmission value η^* which saturates this bound

$$\eta_* = \hat{\eta}_{N,K,d}(v_*). \tag{80}$$

Inserting this value for η_* in Eq. (79), we can then determine p (and v_*) as a function of Bob's actual CMU parameters η and v, and the obtained value of p can be used in (77) to get a bound on the key rate. We now illustrate this using several of the bounds obtained in the previous sections.

Arbitrary measurements in dimension d. Using the bound (72) and setting $\eta_* = \frac{d}{(K+1)(v_*(d+1)-1)}$ to the value saturating this bound, by solving Eq. (79) we find that $p = p'_{K,N,d}$ with

$$\begin{aligned} p'_{K,N,d}(\eta,v) =& \frac{\eta(1-v)+d(1-\eta v)}{d} \\ &\times \max\left\{\frac{K+1}{K}, \frac{N}{N-1}\right\}. \end{aligned} \tag{81}$$

Here we used the fact that if all the N measurements are used to extract the key one simply replaces $K+1$ with N. An interesting special case is the one where the key is extracted from an extremely large number of measurements. A bound on the key rate independent of N and K can be obtained from Eq. (81), by taking the limit of large K, N. We find

$$\lim_{K,N\to\infty} p'_{K,N,d}(\eta,v) = 1 - \frac{\eta(v(d+1)-1)}{d}, \tag{82}$$

and thus by Eq. (77) the key rate $r_\infty = 0$ is zero if $\frac{\eta(v(d+1)-1)}{d} \leq \frac{H_{\eta,v}(B|A)}{H_{1,1}(B)}$.

A few binary measurements in dimension 2. In the case $d = 2$, the value (81) based on the generic bound (72) can be used. However, in the specific case of qubit measurement of the form (36) and $K = N = 2$ or $K = N = 3$, it is more advantageous to use the bound (50). Setting η_* to saturate this bound we find that p is given by

$$p''_N(\eta,v) = \frac{N(1-\eta v)+\eta\sqrt{N}(1-v)}{N-1}\,. \tag{83}$$

Note that for any specific set of N qubit measurements, one can of course use a tighter *JM* condition based on Upper bound 2' and (46) to derive a better attack.

All binary measurements in dimension 2. Finally, as discussed in section 3.3 the set of all qubit PVMs subject to noise and loss become *JM* for $\eta \leq 2(1-v)$. Setting $\eta_* = 2(1-v_*)$ in 79 we find p to be equal to

$$p'''(\eta,v) = 1-\eta v + \sqrt{\eta(1-v)(2-\eta(1+v))}\,. \tag{84}$$

For the case of a qubit ($d = 2$) with projective measurements, which is the most studied in the literature, we have thus obtained three different expressions for p leading to different bounds on the key rate in Eq. (77). In the following table, we compare these expressions and indicate which one gives a better attack depending on the values N and K.

K \ N	2	3	≥ 4
1	Eq. (83)	Eq. (81)	Eq. (81)
2	Eq. (83)	Eq. (83)	Eq. (81)
3	x	Eq. (83)	Eqs. (81) or (84)
≥ 4	x	x	Eqs. (81) or (84)
$\to\infty$	x	x	Eq. (84)

Table 1: In the case of N binary measurements (ideally) on a qubit, K of which are used for key extraction, the table gives the optimal attack parameter p among the expressions $p'_{K,N,2}(\eta,v), p''_N(\eta,v)$ or $p'''(\eta,v)$ in Eqs. (81 ,83, 84) derived by using different JM bounds discussed in the previous section. In addition to the table, we find that for $v = 1$ the maximal value is always $p = (1-\eta)\max\left\{\frac{K+1}{K}, \frac{N}{N-1}\right\}$.

5.2.1 Application to BB84/CHSH-type protocols

We now illustrate the above approach on a family of qubit protocols where the key on Bob's side is based on the output $b \in \{0,1\}$ of a single measurement, which we assume in the honest implementation to be a σ_z measurement. We also assume that in the key generation rounds, the state entering Bob's CMU is the σ_z eigenstate $|0\rangle$ or $|1\rangle$, depending on Alice's key variable $a = 0$ or $a = 1$. These are the only features of the protocol that we need to apply a bound on the key rate using our approach.

Protocols satisfying the above conditions are the one-sided-DI version of the original BB84 protocol, where the source is fully characterized [66], semi-DI versions of it, where Alice's device is only trusted to prepare qubit states [67], the semi-DI CHSH prepare-and-measure protocol of [4], entanglement-based protocols such as the one-sided-DI protocols based on steering of [68, 69], fully DI CHSH-based protocols [70], or their routed version [71].

As $K = 1$ and $d = 2$, we have from (81), $p = 2+\eta(1-3v)$, while $H_{1,1}(B) = 1$. To compute $H_\eta(B|A)$, we can consider two possibilities. Either

non-detection events are treated as distinct outcomes, i.e., $b \in \{0, 1, \varnothing\}$. Or they are grouped with one of the other possible outcomes. In the first case, we have $H_{\eta,v}(B|A) \geq \eta\, h\left(\frac{1+v}{2}\right) + h(\eta)$, so that the key rate is upper-bounded as

$$r_\infty \leq \eta(3v-1) - 1 - \eta\, h\left(\frac{1+v}{2}\right) - h(\eta)\,. \tag{85}$$

In the case where Bob bins his no-click outcome with one of the other outcomes, we have $H(B|A) \geq \frac{1}{2}\left(h\left(\frac{\eta}{2}(1+v)\right) + h\left(\frac{\eta}{2}(1-v)\right)\right)$ and the key rate is upper-bounded as

$$r_\infty \leq \eta(3v-1) - 1 - \frac{1}{2}\left(h\left(\frac{\eta}{2}(1+v)\right) + h\left(\frac{\eta}{2}(1-v)\right)\right)\,.$$

As an illustration, for a visibility $v = 1$, we find thresholds of $\eta = 83.0\%$ in the case where Bob does not bin and $\eta = 71.6\%$ in the case where he bins, and for a visibility $v = 0.97$, we find thresholds of $\eta = 87\%$ and $\eta = 76\%$, respectively.

5.2.2 Application to receiver-device-independent protocols

As a second application, we consider the class of so-called receiver device-independent (RDI) protocols [22, 72], which can be seen as an extension of the B92 protocol [73] to the case of many states and measurements. These are prepare-and-measure protocols as outlined in Fig. 2(a). Here, the measuring device is completely untrusted, while the preparing device is assumed to produce N pure states whose Gram matrix is fixed. We now provide a brief description of the protocol with an intuitive motivation for the ideal case (no noise), an interested reader is invited to consult the original papers for the full details. We will consider the version of the protocol where Alice prepares qubit states

$$|\psi_x\rangle = \cos(\theta/2)\,|0\rangle + e^{\frac{i2\pi}{N}x}\sin(\theta/2)\,|1\rangle\,, \tag{86}$$

for some choice of $\theta \in [0, \frac{\pi}{2}]$ and a value $x \in \{0, \dots, N-1\}$ that she choses at random, analyzed in [72] in detail. Bob's ideal measurements are then $M_{b|y} = \left|\psi_y^\perp\right\rangle\!\left\langle\psi_y^\perp\right|, |\psi_y\rangle\langle\psi_y|$ for $b = 0, 1$ respectively, with the setting $y \in \{0, \dots, N-1\}$ chosen at random. After his measurement, Bob tells Alice to reject the round if $b \neq 0$. At this point, since $\langle\psi_y|\psi_y^\perp\rangle = 0$ both parties know that their values x and y must satisfy $x \neq y$ in all the rounds which are not rejected. Next, Alice reveals an unordered pair of values (x, x') where x is the value she encoded in the state and x' is a different value chosen at random. This pair is only accepted by Bob if $y \in \{x, x'\}$. Since there are only two possible values left, and Bob can exclude one of them (from $x \neq y$), Alice and Bob are left with a perfectly correlated bit after the sifting. This gives the idea of the protocol. We now explicitly compute the correlations observed by Alice and Bob.

The probability to observe the outcome $b = 0$ depends on the state $|\psi_x\rangle$ sent by Alice and is given by

$$p^{(\eta,v)}(0|x,y) = \eta\, v |\langle\psi_x|\psi_y^\perp\rangle|^2 + \frac{\eta(1-v)}{2} \tag{87}$$

where $|\langle\psi_x|\psi_y^\perp\rangle|^2 = \sin^2(\theta)\sin^2(\frac{\pi(x-y)}{N})$. From Bob's perspective (knowing y), a round is successful if $b = 0$, and if $x = y$ (happening with probability $\frac{1}{N}$) or if $x' = y$ (for some $x \neq y$), summing the two gives

$$p(\text{succ}|y) = \frac{1}{N}p(0|y,y) + \frac{1}{N(N-1)}\sum_{x\neq y} p(0|x,y). \tag{88}$$

Combining with Eq. (87) we find that the success probability is independent of y and given by

$$p^{(\eta,v)}(\text{succ}) = \frac{\eta v \sin^2(\theta)}{2(N-1)} + \frac{\eta(1-v)}{N} \tag{89}$$

Eve's strategy consists of replacing the CMU with a jointly measurable one, given by the parameters (η_*, v_*), with probability $p = p(\text{attack})$, or with the ideal one with probability $1-p$. In fact, the probability that the round is successful depends on her choice. The conditional probability of the attack reads

$$\bar{p} = p(\text{attack}|\text{succ}) = \frac{p^{(\eta_*,v_*)}(\text{succ})\, p}{p^{(\eta,v)}(\text{succ})}, \tag{90}$$

and implies $H(B|E, \text{succ}) = (1-\bar{p})$. With the help of Eqs. (79) we see that

$$p^{(\eta_*,v_*)}(\text{succ})\, p = \frac{p\eta_* v_* \sin^2(\theta)}{2(N-1)} + \frac{p\eta_*(1-v_*)}{N} \tag{91}$$

$$= p^{(\eta,v)}(\text{succ}) - \frac{(1-p)\sin^2(\theta)}{2(N-1)} \tag{92}$$

and thus $\bar{p} = 1 - (1-p)\frac{\frac{\sin^2(\theta)}{2(N-1)}}{p^{(\eta,v)}(\text{succ})}$.

When Eve replaces, with probability p, the CMU with a *JM* one, her parent POVM predicts, according to the strategies detailed in Section 3 that a no-click outcome $b = \{0, 1\}$ will be obtained for Bob for only one of his possible measurements y. This implies that when Bob announces that he obtained the value $b = 0$, Eve can infer which input y was used by Bob and hence her entropy about Bob's raw key is 0. Analogously to Eq. (76), we thus have $H(B|E, \text{succ}) \leq (1-\bar{p})H_{1,1}(B|\text{succ}) \leq 1 - \bar{p}$ or

$$H(B|E, \text{succ}) \leq (1-p)\frac{\frac{\sin^2(\theta)}{2(N-1)}}{p^{(\eta,v)}(\text{succ})}. \tag{93}$$

While this formula is affected by post-selection, we see that the optimal strategy for Eve is still the one maximizing p such that the noisy CMU is jointly measurable and we can use the formulas derived in the last section and summarized in Table. 1 (in our case

$K = N$ as all of Bob's measurements are used for the key).

Finally, to compute the key rate, we also need the expression of $H(B|A, \text{succ})$. For a successful round where Alice prepared the state x and also revealed x', Bob must have chosen the measurements with $y = x$ or $y = x'$. For Alice the likelihood of these events is

$$p(y = x|A, \text{succ}) = \frac{p^{(\eta,v)}(0|x, x)}{p^{(\eta,v)}(0|x, x) + p^{(\eta,v)}(0|x, x')} \tag{94}$$

$$p(y = n|A, \text{succ}) = \frac{p^{(\eta,v)}(0|x, x')}{p^{(\eta,v)}(0|x, x) + p^{(\eta,v)}(0|x, x')} \tag{95}$$

The entropy of this distribution (as well as the success probability) depends on the pair $\{x, x'\}$. The entropy is minimal if $\sin^2(\pi\frac{x-x'}{N}) = 1$ in which case it is equal to the rhs of

$$H(B|A, \text{succ}) \geq h\left(\frac{1-v}{2(1-v\cos^2(\theta))}\right). \tag{96}$$

This gives a general upper bound on the key rate with one-way error correction

$$\begin{aligned} r &\leq p(\text{succ})\Big(H(B|E, \text{succ}) - H(B|A, \text{succ})\Big) \qquad (97)\\ &\leq (1-p)\frac{\sin^2(\theta)}{2(N-1)} - p^{(\eta,v)}(\text{succ})h\left(\frac{1-v}{2(1-v\cos^2(\theta))}\right) \end{aligned}$$

To apply the bound we now take the limit of many measurements $N \to \infty$, where it was shown [22] that the key rate remains positive for arbitrary losses provided that the visibility is perfect $v = 1$ and $N > \frac{1}{\eta}$. Since the ideal measurements of Bob are qubit PVMs we can use the Eq. (84) for the strategy of Eve that maximizes $(1-p)$. We find that for $\eta = 10\%, 1\%, 0.1\%$ the key rate is zero if $v \leq 95\%, 99.5\%, 99.95\%$ respectively.

With the same approach tighter bounds can be derived on the RDI protocol by computing the entropy $H(B|A, \text{succ})$ explicitly, and deriving a tighter *JM* condition for the specific set of the CMUs stemming from the measurements $M_{b|y} = \{|\psi_y\rangle\langle\psi_y|, |\psi_y^\perp\rangle\langle\psi_y^\perp|\}$ specified by Eq. (86).

5.3 Convex combination attacks for DIQKD protocols

The convex combination attacks introduced in the previous section are asymmetric as they apply to generic QKD protocols where Bob holds a CMU, but where Alice may hold, e.g., a trusted preparation device. Fully DIQKD protocols, however, are symmetric in the sense that both parties hold a CMU. Exploiting this feature, we analyze in this section improved convex combination attacks for DIQKD protocols that lead to more stringent constraints on the key rate. Furthermore, again due to the symmetry of the protocol, they can be applied to situations where the public communication for key distillation goes from Bob to Alice, as in the previous section, or from Alice to Bob. For simplicity, we consider explicitly only the former case. But we note that in many cases the measurements and post-processing performed by Alice and Bob in DIQKD are essentially identical (up to e.g. local unitaries), implying that the same bounds on the key rate would be obtained by considering the public communication in the other direction.

Our convex combination attacks, in their simplest version, are a mixture of strategies where Eve replaces either Alice's or Bob's CMU with a full *JM* for a portion of the time, and allowing the devices to operate ideally for the remainder. In the former case, the correlations between Alice and Bob are local, while they are non-local in the latter case. This class of attacks thus fits in the same framework as those proposed in [74, 75] involving a convex combination of local and non-local correlations. Such local/non-local mixtures are obtained in [74, 75] through analytical or numerical (linear programming) results tailored to specific quantum correlations. However, the analytical results are limited to highly specific cases and linear programming methods become increasingly challenging when dealing with a large number of measurements and outputs. In contrast, the approach based on joint measurability is generic, as it depends only on rough properties of the quantum protocol, and it can provide bounds even in scenarios with many measurements. The downside is that it may lead to bounds on the key rate less stringent than the ones of [74, 75] when applied to the specific correlations for which their methods are tailored.

Note, however, that more generally, we will consider attacks where Eve replaces Bob's CMU with a $\mathcal{K}$-*JM* one with some probability. This gives rise to a class of attacks different from the one considered in [74, 75], since correlations arising from a $\mathcal{K}$-*JM* POVMs are not necessarily local, yet, are sufficient for Eve to have full information about Bob's key-generating outcomes. This may compensate for the fact that our joint measurability bounds are not tailored to specific correlations. In contrast, it is important to note that Alice's device cannot be replaced with a $\mathcal{K}$-*JM* one, as this does not allow Eve to guess Bob's key generating outcomes in general.

Let us now describe our attacks in more detail. Consider a generic DIQKD scenario consisting of a central source distributing bipartite states ρ, a CMU for Alice, and a CMU for Bob, as depicted in Fig. 2.b. For concreteness, we assume that both Alice and Bob have a white-noise no-click CMU characterized by the effective POVMs (9) with detector efficiency η and white-noise visibility v. Therefore, together they implement the joint effective POVMs $M_{a|x}^{\eta,v} \otimes M_{b|y}^{\eta,v}$ on the incoming state ρ. We further denote by N_A the

number of measurements on Alice's side, and by N_B the overall number of measurements on Bob's side. Furthermore, let us assume that the key is only extracted from a subset of Bob's measurements with the inputs $y \in \mathcal{K}_B$.

Eve's attack is based on the following convex decomposition of Alice and Bob's joint POVMs

$$\begin{aligned} M^{\eta,v}_{a|x} \otimes M^{\eta,v}_{b|y} = p_A M^{\eta_A,v_A}_{a|x} \otimes M^{1,1}_{b|y} + p_B M^{1,1}_{a|x} \otimes M^{\eta_B,v_B}_{b|y} \\ + q M^{\eta',0}_{a|x} \otimes M^{\eta',0}_{b|y} \\ + (1-p_A-p_B-q) M^{1,1}_{a|x} \otimes M^{1,1}_{b|y}\,, \end{aligned} \tag{98}$$

where $\eta' = \eta(1-v)/(1-\eta v)$, $q = (1-\eta v)^2$, and η_A, v_A, η_B, v_B, p_A, and p_B are free parameters that Eve can chose. Using the explicit form for the white-noise no-click CMU POVMs (9), it is readily verified that this decomposition is valid provided that these parameters satisfy the following system of equations

$$2\eta v(1-\eta v) = p_A(1-\eta_A v_A) + p_B(1-\eta_B v_B) \tag{99}$$

$$\eta^2(1-v)v = \eta_A p_A\,(1-v_A) = \eta_B p_B\,(1-v_B) \tag{100}$$

$$\eta v(1-\eta) = p_A\,(1-\eta_A) = p_B\,(1-\eta_B)\,. \tag{101}$$

The first constraint (99) is linearly redundant and can be omitted. The remaining ones are equivalent to

$$\eta v(1-\eta v) = p_A\,(1-\eta_A v_A) = p_B\,(1-\eta_B v_B) \tag{102}$$

$$\eta v(1-\eta) = p_A\,(1-\eta_A) = p_B\,(1-\eta_B)\,. \tag{103}$$

Here, one sees that the tuples of parameters (p_A, η_A, v_A) and (p_B, η_A, v_A) are subject to independent constraints, leaving one free parameter for each tuple.

The convex decomposition (98) can be exploited by Eve as follows. With probability $1-p_A-p_B-q$, Alice's and Bob's CMUs behave ideally as they implement the ideal POVMs $M^{1,1}_{a|x} \otimes M^{1,1}_{b|y}$. In this case, we assume that Eve has no extra information about Bob's outcomes, i.e. her conditional entropy is only bounded by $H(B|E) \leq H_{1,1}(B)$, where $H_{1,1}(B)$ the entropy of Bob's outcomes averaged over the inputs $y \in \mathcal{K}_B$ in the ideal (noise-free and loss-free) situation. With probability q, the CMUs of Alice and Bob implement the POVMs $M^{\eta',0}_{a|x} \otimes M^{\eta',0}_{b|y}$, which are maximally noisy, i.e., the visibility is zero. In this case, the correlations generated by Alice and Bob are local. Since Eve can further decompose these correlations as a mixture of deterministic correlations, she has full information about Bob's outcomes, i.e., $H(B|E) = 0$. With probability p_A, Alice's CMU implements the POVMs $M^{\eta_A,v_A}_{a|x}$. If we choose η_A and v_A such that these POVMs are (full) JM, then, again, Alice and Bob's correlations are local, and thus $H(B|E) = 0$. Finally, with probability p_B, Bob's CMU implements the POVMs $M^{\eta_B,v_B}_{b|y}$. If we now choose η_B and v_B such that these POVMs are $\mathcal{K}_B$-JM, Eve has again full information about Bob's outcomes for the inputs $y \in \mathcal{K}_B$, i.e., $H(B|E) = 0$, since she can implement the parent instrument and predict all these outcomes. Note that in this last case, though, the correlations between Alice and Bob, which involve also the inputs $y \notin \mathcal{K}$, are not necessarily local. Summarizing, by our attack the key rate is upper-bounded as

$$\begin{aligned} r_\infty &\leq t H_{1,1}(B) - H_{\eta,v}(B|A)\,, \\ &\text{with} \\ t &= 1-q-p_A-p_B = 2\eta v - \eta^2 v^2 - p_A - p_B\,. \end{aligned} \tag{104}$$

Hence to find the tightest upper-bound, we must maximize the probabilities p_A and p_B subject to the constraints that the POVMs M^{η_A,v_A}_x are full JM and the POVMs M^{η_B,v_B}_y are $\mathcal{K}$-JM. For this, it is optimal for Eve to set the efficiency $\eta_A = \hat{\eta}_{N_A,d}(v_A)$ saturating one of the bounds for JM that we have derived in the previous sections and similarly setting $\eta_B = \hat{\eta}_{N_B,K_B,d}(v_B)$ saturating one of the bounds for $\mathcal{K}_B$-JM. Then, the optimal values of p_A and p_B are given by the solutions of the Eqs. (102,103), i.e

$$\begin{aligned} \frac{p_A}{\eta v}\,(1-\hat{\eta}_{N_A,d}(v_A)\,v_A) &= \frac{p_B}{\eta v}\,(1-\hat{\eta}_{N_B,K_B,d}(v_B)v_B) \\ &= 1-\eta v \\ \frac{p_A}{\eta v}\,(1-\hat{\eta}_{N_A,d}(v_A)) &= \frac{p_B}{\eta v}\,(1-\hat{\eta}_{N_B,K_B,d}(v_B)) \\ &= 1-\eta. \end{aligned} \tag{105}$$

Remarkably, these equations become equivalent to those in (79) by the following parameter change $(p_{A(B)}, \eta_{A(B)}, v_{A(B)}) \to (\eta v\, p, \eta_*, v_*)$. It follows that the expressions $\eta\, v\, p'_{K,N,d}(\eta, v), \eta\, v\, p''_N(\eta, v)$ or $\eta\, v\, p'''(\eta, v)$ in Eqs. (81 ,83, 84) can be readily used to find $p_{A(B)}$ in different cases.

In particular, our bound (104) guarantees that the key rate is zero if one can choose the values p_A and p_B such that

$$p_A + p_B \geq 2\eta v - \eta^2 v^2 - \frac{H_{\eta,v}(B|A)}{H_{1,1}(B)} \implies r_\infty = 0. \tag{106}$$

Noticing that the term $\frac{H_{\eta,v}(B|A)}{H_{1,1}(B)}$ is always positive one obtains a weaker but correlation-independent condition for zero key rate

$$p_A + p_B \geq 2\eta v - \eta^2 v^2 \implies r_\infty = 0. \tag{107}$$

We now illustrate this bound on the three specific attack strategies already discussed in Sec. 5.2.

Arbitrary measurements in dimension d. When we make no assumption on the measurements performed except their dimensionality, we can use bounds (28) and (72), and we obtain that the probabilities p_A and p_B must be

$$p_A = \eta\, v\, p'_{N_A-1,N_A,d}(\eta, v) \tag{108}$$

$$p_B = \eta\, v\, p'_{K_B,N_B,d}(\eta, v)\,, \tag{109}$$

with the function $p'_{K,N,d}(\eta, v)$ defined in equation (81). In particular, no key can be extracted (Eq. 107), whenever

$$\eta \leq \frac{(T_{N_A,K_B,N_B} - 2)\, d}{(T_{N_A,K_B,N_B} - 1)\, d\, v - T_{N_A,K_B,N_B}(1-v)},$$

where we introduced $T_{N_A,K_B,N_B} = \frac{N_A}{N_A-1} + \min\left\{\frac{K_B+1}{K_B}, \frac{N_B}{N_B-1}\right\}$ to shorten the equation. This bound becomes $\eta \leq 1 - \frac{1}{T_{N_A,K_B,N_B}-1}$ for $v = 1$, and $v \leq \frac{T_{N_A,K_B,N_B}+d(T_{N_A,K_B,N_B}-2)}{T_{N_A,K_B,N_B}+d(T_{N_A,K_B,N_B}-1)}$ for $\eta = 1$.

A few binary measurements on a qubit. When Alice and Bob use N_A, N_B (respectively) binary qubit measurements we can use the bound in Eq. (49) to obtain

$$p_{A(B)} = \eta\, v\, p''_{N_{A(B)}}(\eta, v), \tag{110}$$

with the function $p''_N(\eta, v)$ defined in Eq. (83). In particular, for $N_A = N_B = N$ we find that no key can be extracted (Eq. (107)) whenever

$$\eta \leq \frac{2(N+1)v + 4(1-v)\sqrt{N}}{(N-1)^2 v^2 + 4N(2v-1)},$$

which becomes $v \leq \frac{2}{\sqrt{N}+1}$ for $\eta = 1$.

All binary measurements on a qubit. Finally, in the same situation but with an unrestricted number of measurements for Alice and Bob, we can use the bound of Eq. (51) valid for the set of all qubit PVMs to obtain

$$p_{A(B)} = \eta\, v\, p'''(\eta, v), \tag{111}$$

with $p'''(\eta, v)$ defined in Eq. (84). In particular, we find that no key can be extracted (Eq. (107)) if

$$\eta \leq \frac{8(1-v)}{4 - 3v^2},$$

which becomes $v \leq \frac{2}{3}$ for $\eta = 1$.

Note that different bounds can be used on the sides of Alice and Bob. Furthermore, the attack we presented can be easily improved with additional knowledge about the measurements performed by the parties, as in this case one can use tighter (partial)-JM bounds specific to a given CMU.

Recall also that in the most common case where Alice and Bob perform qubit PVMs, the comparison of the three bounds $p'_{K,N,2}(\eta, v)$, $p''_N(\eta, v)$ and $p'''(\eta, v)$ was given in Table 1. It can be readily used to find the optimal attack of Eve in different cases. With its help, let us now explore the performance of popular DIQKD protocols in noisy conditions and determine the maximum achievable key rate. The challenge here is to compute the terms $H_{\eta,v}(B|A)$ and $H_{1,1}(B)$.

5.3.1 Application to qubit protocols without binning

Consider the protocols where Alice and Bob share a two-qubit state

$$|\Psi_\theta\rangle = \cos(\theta)\,|00\rangle + \sin(\theta)\,|11\rangle\,. \tag{112}$$

When the state is not maximally entangled Alice and Bob can only establish perfect correlation when they both measure in the Z basis. In this case, we thus assume that the key is only extracted from this measurement and set $K_B = 1$. The entropy of the key-generating outcome is then given by

$$H_{1,1}(B) = h\left(\cos^2(\theta)\right) \tag{113}$$

for the ideal implementation. The conditional entropy term $H_{\eta,v}(B|A)$ is always larger than (equal for optimally aligned measurements)

$$H_{\eta,v}(B|A) = \eta(1-\eta) + h(\eta) + \eta^2\, h\left(\frac{1+v^2}{2}\right), \tag{114}$$

where $h(x)$ is the binary entropy.

Plugging these expressions into Eq. (104) gives

$$r_\infty \leq h\big(\cos^2(\theta)\big)(2\eta v - p_A - p_B - \eta^2 v^2) - H_{\eta,v}(B|A). \tag{115}$$

One notices that $h(\cos^2(\theta))$ is the only term depending on the parameter θ, and it is straightforward to see that the optimal protocol (where the key rate bound is the less stringent) consists of preparing the maximally entangled state. We can thus set $\theta = \frac{\pi}{4}$ for which $H_{1,1}(B) = 1$. To find the best bound it remains to maximize $p_A + p_B$ with the guidance of Table 1 and the functions $p'_{K,N,d}(\eta, v), p''_N(\eta, v)$ or $p'''(\eta, v)$ defined in Eqs. (81 ,83, 84). To do so we consider three situations.

(i) $(N_A, N_B, K_B) = (3, 2, 1)$ is the setting of the DIQKD CHSH protocol introduced in [76, 1]. Here the optimal bound reads

$$r_\infty \leq \eta v\big(2 - p''_3(\eta, v) - p''_2(\eta, v)\big) - \eta^2 v^2 - H_{\eta,v}(B|A). \tag{116}$$

(ii) $(N_A, N_B, K_B) = (\infty, \infty, 1)$ corresponds to the setting where only one measurement is used for key generation, but the number of test measurements is not restricted. Here the bounds

$$r_\infty \leq \eta v\big(2 - p'_{1,\infty,2}(\eta, v) - p'''(\eta, v)\big) - \eta^2 v^2 - H_{\eta,v}(B|A), \tag{117}$$

are to be used depending on the values of η and v.

(iii) $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ is the scenario with no restriction on the number of measurements. This is a meaningful scenario as for the maximally entangled

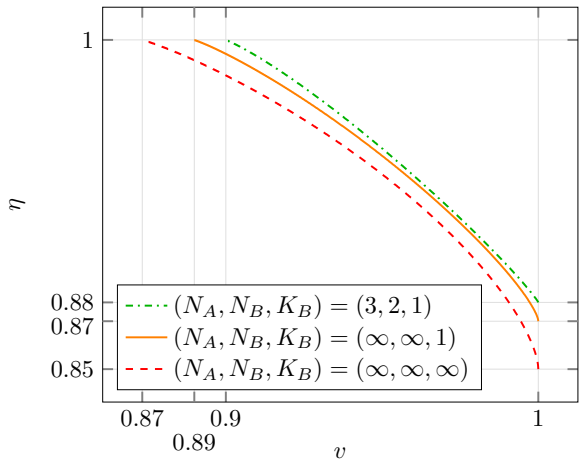

Figure 9: For the parameters (η, v) below the lines the key rate r_∞ is guaranteed to be zero. The lines show when the right-hand sides of the corresponding equation fall to zero. (From top to bottom) for Eq. (116) : *(i)* $(N_A, N_B, K_B) = (3, 2, 1)$ dash-dotted line , Eq. (117) : *(ii)* $(N_A, N_B, K_B) = (\infty, \infty, 1)$ full line, and Eq. (118) : *(iii)* $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ dashed line. This applies to DIQKD protocols based on two-qubit states, dichotomic measurements (ideally), no-binning of the no-click outcome, and one-way public communication from Bob to Alice.

Setting (N_A, N_B, K_B)	Threshold η (at $v = 1$)	Threshold v (at $\eta = 1$)
(3,2,1)	88.3 %	89.8 %
(∞,∞,1)	87.4 %	88.8 %
(∞,∞,∞)	85.3 %	87.1 %

Table 2: Threshold efficiency η and visibility v for varions number of measurements (N_A, N_B, K_B). This applies to DIQKD protocols with the maximally entangled two-qubit state, dichotomic measurements (ideally), no-binning of the no-click outcome, and one-way public communication from Bob to Alice.

state Alice and Bob can have perfect correlation for any number of measurements. Here the bound

$$r_\infty \leq \eta v\big(2 - 2\,p'''(\eta, v)\big) - \eta^2 v^2 - H_{\eta,v}(B|A) \quad (118)$$

is optimal.

In Fig. 9 we plot the lines in the (η, v) plane below which the bounds of Eqs. (116-118) guarantee a zero key rate. The threshold values of η and v (for $v, \eta = 1$ respectively) are reported in Table. 2

5.3.2 Application to qubit protocols with binning

A common strategy in the DIQKD protocols is to bin non-click outcomes of Bob with the most likely among the other two outcomes. After such binning the key rate bound in the Eq. (115) remains valid upon replacing the conditional entropy term with

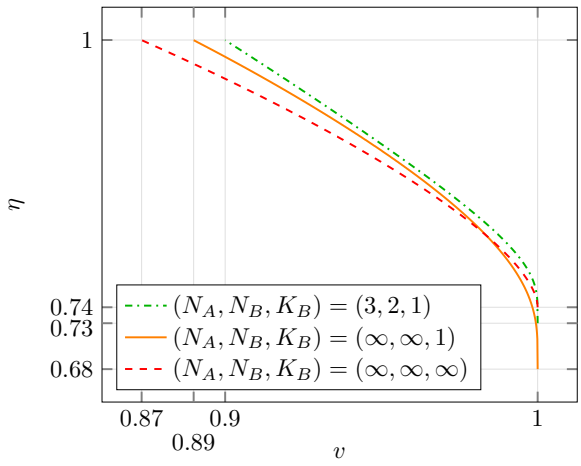

Figure 10: For the parameters (η, v) below the lines the key rate r_∞ was found to be zero after a maximization over θ. The lines show when the right-hand sides of the corresponding equation fall to zero. (From top to bottom) for Eq. (120) : *(i')* $(N_A, N_B, K_B) = (3, 2, 1)$ dash-dotted line , and Eq. (121) : *(ii')* $(N_A, N_B, K_B) = (\infty, \infty, 1)$ full line. For the curve, Eq. (122) : *(iii')* $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ dashed line, no maximization over θ is needed as it only makes sense to consider the maximally entangled state $\theta = \frac{\pi}{4}$. This applies to DIQKD protocols based on two-qubit states, dichotomic measurements (ideally), binning of the no-click outcomes, and one-way public communication from Bob to Alice.

Setting (N_A, N_B, K_B)	Threshold η (at $v = 1$)	Threshold v (at $\eta = 1$)
(3,2,1)	72.7 %	89.8 %
(∞,∞,1)	68.3 %	88.8 %
(∞,∞,∞)	74.2 %	87.1 %

Table 3: Threshold efficiency η and visibility v for different number of measurements (N_A, N_B, K_B). This applies to DIQKD protocols with the partially entangled two-qubit state, dichotomic measurements (ideally), binning of the no-click outcomes, and one-way public communication from Bob to Alice.

$$\begin{aligned} H^\theta_{\eta,v}(B|A) &= (1-\eta)\, h\left(\frac{\eta}{2}(1 - v\cos(2\theta))\right) \\ &+ \frac{\eta}{2}(1 - v\cos(2\theta))\; h\left(\eta\left(1 - \frac{1-v^2}{2(1-v\cos(2\theta))}\right)\right) \\ &+ \frac{\eta}{2}(1 + v\cos(2\theta))\; h\left(\eta\frac{1-v^2}{2(1+v\cos(2\theta))}\right). \end{aligned} \quad (119)$$

Note that in general this reduces Alice's entropy, i.e., $H^\theta_{\eta,v}(B|A)$ here above is lower or equal to $H_{\eta,v}(B|A)$ in Eq. (114) since the binning channel applied by Bob is two-to-one. It follows that the key rate bounds that we obtain using binning are less stringent than without. With the new expression for $H^\theta_{\eta,v}(B|A)$ let us consider the following settings.

(i') For $(N_A, N_B, K_B) = (3, 2, 1)$ the optimal bound reads

$$r_\infty \leq h\big(\cos^2(\theta)\big)\big(\eta v(2 - p_3''(\eta, v) - p_2''(\eta, v)) - \eta^2 v^2\big) - H^\theta_{\eta,v}(B|A). \quad (120)$$

(ii') For $(N_A, N_B, K_B) = (\infty, \infty, 1)$ the optimal bound reads

$$r_\infty \leq h\big(\cos^2(\theta)\big)\big(\eta v(2 - p'_{1,\infty,2}(\eta, v) - p'''(\eta, v)) - \eta^2 v^2\big) - H^\theta_{\eta,v}(B|A) \quad (121)$$

(iii') For $(N_A, N_B, K_B) = (\infty, \infty, \infty)$ we only consider the case of the maximally entangled state $\theta = \frac{\pi}{4}$, as it is the only one where Alice and Bob can have perfect correlations in more than one measurement basis. We then get the optimal bound

$$r_\infty \leq \eta v\big(2 - 2\, p'''(\eta, v)\big) - \eta^2 v^2 - H^{\frac{\pi}{4}}_{\eta,v}(B|A). \quad (122)$$

To derive state-independent upper-bounds one needs to maximize the right-hand side of the above expressions with respect to θ. Here, both quantities $H_{1,1}(B)$ and $H^\theta_{\eta,v}(B|A)$ have a nontrivial dependence on the angle θ, and this maximization has to be done for all η and v. In Fig. 10, we display the thresholds that we found after maximizing the key rate over the angle θ in a heuristical way. Specifically, we used the Nelder–Mead technique [77]. In Table. 3 we give the corresponding threshold values.

6 Discussion

We have developed a general framework for characterizing the impact of losses in quantum communication setups with untrusted measurement devices. A key element in our work is the concept of a CMU, which should be viewed as a set of effective measurements, characterizing both the imperfections of the channel and the subsequent measurement apparatus. We investigated the joint measurability of this set, establishing a direct connection between this fundamental notion and quantum communication, in particular QKD.

Our work motivates the investigation of the (in)compatibility of sets of POVM under the combined effect of noise and loss, of which we discussed several cases. In particular, we have derived a sufficient condition for the compatibility of sets of N unbiased qubit measurements (Lemma 6). We have also seen that results on channel extendibility are very useful in this context, as they imply compatibility of CMUs solely based on the noise channel and the number of measurements that compose it. We derived a general result on the extendibility of concatenated noise channels. Our results can be applied to any protocol or experimental scenario with untrusted measurement devices, including DI and SDI protocols, providing insight into the admissible trade-off between noise and loss. Obtaining more results in this direction would be desirable.

In the context of QKD, the notion of joint measurability of a CMU leads to necessary criteria for extracting a secure key. The analysis of QKD protocols also led us to introduce the concept of partial joint measurability. We showed that this notion differs from standard joint measurability, in the sense that there exist sets of measurements that are partially jointly measurable but not (fully) jointly measurable. It would be interesting to investigate this concept further, in particular how to cast partial joint measurability as a semi-definite program. We note that this concept as we defined it differs from the notions discussed in previous works, where sets of POVMs that are incompatible can feature subsets of POVMs that are jointly measurable [78, 79, 31]. In contrast, our definition is closely related to the notion of no-exclusivity introduced recently for quantum instruments [80], and to the notion of weak-compatibility introduced earlier in the same context [81]. In particular, the notion of partial joint-measurability that we introduced for a channel followed by a set of POVMs can be generalized to an analogous notion of partial no-exclusivity for a channel followed by a set of instruments. Lemma 7 then can naturally be extended to this setting.

Finally, as an outlook note that the intuition behind the notion of partial joint measurability can be used to improve the class of attacks on DIQKD considered in [74, 75]. These attacks are based on decompositions of the correlations observed by Alice and Bob $p(a, b|x, y) = q\, p_L(a, b|x, y) + (1-q) p_{NL}(a, b|x, y)$ into a mixture of local $p_L(a, b|x, y)$ and non-local $p_{NL}(a, b|x, y)$ correlations. In fact, the locality of correlation $p_L(a, b|x, y)$ is sufficient but not necessary for Eve to predict the key. Since she only needs to predict the outcomes of the key-generating measurements, we introduced here attacks where $p(a, b|x, y)$ is decomposed into a mixture of non-local quantum correlations and partly-*JM* $p_{JM}(a, b|x, y)$ correlations. The partly-*JM* correlations can be associated with a quantum-classical state, where only the classical register is used to simulate the outputs of the key-generating measurements, whereas for the other measurements the quantum register is used, hence can display non-locality.

Acknowledgments

We thank Nicolas Cerf for fruitful discussions. P.S., M.I., and N.B. acknowledge funding from the Swiss National Science Foundation (project 192244) and by the Swiss Secretariat for Education, Research and Innovation (SERI) under contract number UeM019-3. M.M. and S.P. acknowledge funding from the VERIqTAS project within the QuantERA II Programme

that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101017733 and the F.R.S-FNRS Pint-Multi programme under Grant Agreement R.8014.21, from the European Union's Horizon Europe research and innovation programme under the project "Quantum Security Networks Partnership" (QSNP, grant agreement No 101114043), from the F.R.S-FNRS through the PDR T.0171.22, from the FWO and F.R.S.-FNRS under the Excellence of Science (EOS) programme project 40007526, from the FWO through the BeQuNet SBO project S008323N, from the Belgian Federal Science Policy through the contract RT/22/BE-QCI and the EU "BE-QCI" programme. S.P. is a Research Director of the Fonds de la Recherche Scientifique – FNRS.

Funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them.

A Proof of the bound (63) in the case of homodyne measurements

Homodyne measurements corresponding to the observables

$$\hat{X}(\theta) = \cos\theta\hat{X} + \sin\theta\hat{P}, \tag{123}$$

specified by the angle θ and where $\hat{X}$ and $\hat{P}$ are the position and momentum quadrature operators

$$\hat{X} = \frac{\hat{a}+\hat{a}^\dagger}{\sqrt{2}}, \quad \text{and} \quad \hat{P} = \frac{\hat{a}-\hat{a}^\dagger}{i\sqrt{2}}. \tag{124}$$

We now show that when the measurements M_y in a thermal-noise CMU correspond to the above observables (with the angle $\theta(y)$ depending on y), the resulting effective POVMs are JM independently of the total number N of such measurements.

In a thermal-noise model, the annihilation operators follow the input-output relation

$$\hat{a}_{\text{out}} = \sqrt{\eta}\hat{a}_{\text{in}} + \sqrt{1-\eta}a_{\tau_\nu}, \tag{125}$$

where a_{τ_ν} is the annihilation operator acting on the thermal state that is combined with the input state in our noise model. We can use the latter relation to express the effective CMU quadratures that are implemented by the CMU as

$$\hat{X}_{\text{out}}(\theta) = \sqrt{\eta}\hat{X}_{\text{in}}(\theta) + \sqrt{1-\eta}\hat{X}_{\tau_\nu}(\theta). \tag{126}$$

We now introduce a single parent POVM that simulates the results of the measurements of such quadratures for any value of θ. This parent POVM consists in performing first a heterodyne measurement on the input state, i.e, we split the initial mode into two modes using a 50/50 beam splitter and measuring the X quadrature of one output mode and the P quadrature of the other output mode. The obtained classical outcomes, x and p, are then sent to the measurement device, where given the input angle θ, it outputs

$$x_\theta = G(\cos\theta x + \sin\theta p) + \Delta, \tag{127}$$

where G is a constant real number, while Δ is a random real number with centered Gaussian distribution of variance σ^2.

In this equivalent CMU, the input-output relations for the annihilation operators after the 50/50 beam splitter are

$$\hat{a}_1 = \frac{\hat{a}_{\text{in}}+\hat{v}}{\sqrt{2}}, \quad \text{and} \quad \hat{a}_2 = \frac{\hat{a}_{\text{in}}-\hat{v}}{\sqrt{2}}, \tag{128}$$

where $\hat{v}$ is the annihilation operator of the vacuum state entering the dark port of our beam splitter and $\hat{a}_1$ and $\hat{a}_2$ are the annihilation operators of the two output modes. The quadrature $\hat{X}$ measured in the first output mode and the quadrature $\hat{P}$ measured in the second output mode can then be written as

$$\hat{X}_1 = \frac{\hat{X}_{\text{in}}+\hat{X}_v}{\sqrt{2}}, \quad \text{and} \quad \hat{P}_2 = \frac{\hat{P}_{\text{in}}-\hat{P}_v}{\sqrt{2}}. \tag{129}$$

The probability distribution of the final output x_θ of the CMU after the post-processing (127) is then fully characterized by the moments $\langle x_\theta^n\rangle$. This strategy simulates the original CMU if these moments coincide with the moments $\langle\hat{X}(\theta)^n\rangle_{\text{th}}$ of the distribution of the effective observables (126). We now determine such moments.

We first compute the target moments.

$$\langle\hat{X}(\theta)^n\rangle_{\text{th}} = \left\langle\left(\sqrt{\eta}\hat{X}_{\text{in}}\theta) + \sqrt{1-\eta}\hat{X}_{\tau_\nu}(\theta)\right)^n\right\rangle \tag{130}$$

$$= \sum_{k=0}^{n}\binom{n}{k}\sqrt{\eta}^{n-k}\langle\hat{X}_{\text{in}}(\theta)^{n-k}\rangle\sqrt{1-\eta}^{k}\langle\hat{X}_{\tau_\nu}(\theta)^k\rangle \tag{131}$$

$$= \sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k}\sqrt{\eta}^{n-2k}\langle\hat{X}_{\text{in}}(\theta)^{n-2k}\rangle (1-\eta)^k\langle\hat{X}_{\tau_\nu}(\theta)^{2k}\rangle \tag{132}$$

$$= \sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k}\sqrt{\eta}^{n-2k}\langle\hat{X}_{\text{in}}(\theta)^{n-2k}\rangle (2k-1)!!\left(\frac{1}{2}(1-\eta+\epsilon\eta)\right)^k \tag{133}$$

where we used the fact that the odd moments of the thermal state in Eq. (53) are zero, while the even ones are

$$(2k-1)!!\left(\frac{1}{2}+\frac{\epsilon\eta}{2(1-\eta)}\right)^k. \tag{134}$$

We now determine the moments following from the

JM strategy.

$$\langle x_\theta^n\rangle = \langle\Big(G(\cos\theta\hat{X}_1+\sin\theta\hat{P}_2)+\Delta\Big)^n\rangle \tag{135}$$

$$= \left\langle\left(\frac{G}{\sqrt{2}}(\hat{X}_{\rm in}(\theta)+\hat{X}_v(-\theta))+\Delta\right)^n\right\rangle \tag{136}$$

$$= \sum_{k=0}^{n}\binom{n}{k}\left(\frac{G}{\sqrt{2}}\right)^{n-k}\langle X_{\rm in}(\theta)^{n-k}\rangle \left\langle\left(\frac{G}{\sqrt{2}}\hat{X}_v(-\theta)+\Delta\right)^k\right\rangle \tag{137}$$

$$= \sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k}\left(\frac{G}{\sqrt{2}}\right)^{n-2k}\langle X_{\rm in}(\theta)^{n-2k}\rangle \left\langle\left(\frac{G}{\sqrt{2}}\hat{X}_v(-\theta)+\Delta\right)^{2k}\right\rangle, \tag{138}$$

where in the last line we used the fact that the sum of two random variables with Gaussian central distribution has zero odd moments. Now, we have

$$\left\langle\left(\frac{G}{\sqrt{2}}\hat{X}_v(-\theta)+\Delta\right)^{2k}\right\rangle =$$

$$= \sum_{l=0}^{2k}\binom{2k}{l}\left(\frac{G}{\sqrt{2}}\right)^{2k-l}\langle\hat{X}_v(-\theta)^{2k-l}\rangle\langle\Delta^l\rangle \tag{139}$$

$$= \sum_{l=0}^{k}\binom{2k}{2l}\left(\frac{G}{\sqrt{2}}\right)^{2(k-l)}\left(\frac{1}{2}\right)^{k-l}\sigma^{2l} \tag{140}$$

$$(2(k-l)-1)!!(2l-1)!!, \tag{141}$$

where we used again that the odd moments are zero and that the even ones are

$$\langle\hat{X}_v(-\theta)^{2(k-l)}\rangle = \left(\frac{1}{2}\right)^{k-l}(2(k-l)-1)!!, \tag{142}$$

$$\langle\Delta^l\rangle = \sigma^{2l}(2l-1)!!. \tag{143}$$

Using $(2n)! = (2n)!!(2n-1)!!$ and $(2n)!! = 2^n n!$, we get

$$\binom{2k}{2l}(2(k-l)-1)!!(2l-1)!! = \binom{k}{l}(2k-1)!!, \tag{144}$$

and thus

$$\left\langle\left(\frac{G}{\sqrt{2}}\hat{X}_v(-\theta)+\Delta\right)^{2k}\right\rangle = (2k-1)!!\left(\frac{G^2}{4}+\sigma^2\right)^k. \tag{145}$$

We finally obtain that

$$\langle x_\theta^n\rangle = \sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k}\left(\frac{G}{\sqrt{2}}\right)^{n-2k}\langle X_{\rm in}(\theta)^{n-2k}\rangle (2k-1)!!\left(\frac{G^2}{4}+\sigma^2\right)^k. \tag{146}$$

It thus follows that $\langle\hat{X}(\theta)^n\rangle_{\rm th} = \langle x_\theta^n\rangle$ for all n if

$$G=\sqrt{2\eta}, \quad \text{and} \quad \sigma^2=\frac{1}{2}(1-2\eta+\epsilon\eta). \tag{147}$$

Since $\sigma^2\geq 0$, we conclude that this strategy works as long as (63) holds.

B Proof of Lemma 8

In this section, we prove Lemma 8 in the main text. But first, we recall some elements of the formalism of Gaussian channels, states, and their phase-space representation.

The Wigner function of a single-mode Gaussian state with the covariance matrix V and displacement vector $\boldsymbol{\mu}$ is given by

$$W(\alpha) = \frac{1}{\sqrt{\det(\pi V/2)}}\exp\left(-2(\boldsymbol{\alpha}-\boldsymbol{\mu})^T V^{-1}(\boldsymbol{\alpha}-\boldsymbol{\mu})\right). \tag{148}$$

with $\boldsymbol{\alpha} = \binom{\mathrm{Re}\,\alpha}{\mathrm{Im}\,\alpha}$. By definition, the pair $V, \boldsymbol{\mu}$ thus uniquely specifies a Gaussian state. In particular, for a coherent state $|\gamma\rangle$ we have $\boldsymbol{\mu} = \boldsymbol{\gamma} = \binom{\mathrm{Re}\,\gamma}{\mathrm{Im}\,\gamma}$ and $V = \mathbb{1}$.

A single mode Gaussian channel Φ maps Gaussian states to Gaussian states. It can be represented by matrices X, Y and a vector $\boldsymbol{\delta}$, which transform the Wigner function as follows [59, 44]

$$\boldsymbol{\mu} \mapsto X\boldsymbol{\mu}+\boldsymbol{\delta} \tag{149}$$

$$V \mapsto XVX^T+Y \tag{150}$$

Therefore, if a Gaussian channel acts on a coherent state, the resulting Wigner function of state $\Phi(|\gamma\rangle\langle\gamma|)$ is Gaussian and characterized by the pair $V' = XX^T+Y$ and $\boldsymbol{\mu}' = X\boldsymbol{\gamma}+\boldsymbol{\delta}$.

Finally, let us compute the Q-function corresponding to this state. It is well known that it is the convolution of the Wigner function with a Gaussian

$$Q(\alpha) = \frac{2}{\pi}\int \mathrm{d}^2\beta\, W(\beta)e^{-2|\alpha-\beta|^2}. \tag{151}$$

Hence the Q-function is also a Gaussian with the displacement vector $\boldsymbol{\mu}_Q = \boldsymbol{\mu}'$ and covariance matrix $V_Q = V'+\mathbb{1}$ (note that a convolution of two distributions describes a random variable given by the sum of the random variables described by the distributions). Thus if a Gaussian channel (X, Y, δ) acts on a coherent state $|\gamma\rangle$, the Q-function of the finals state is Gaussian with

$$\boldsymbol{\mu}_Q = X\boldsymbol{\gamma}+\boldsymbol{\delta} \tag{152}$$

$$V_Q = XX^T+Y+\mathbb{1}. \tag{153}$$

Now let us come back to the Lemma that we want to prove.

Lemma 8. *A thermal-noise channel $\Phi_{\eta,\epsilon}$ that is not N-extendable cannot be decomposed as*

$$\Phi_{\eta,\epsilon} = p\,\Phi_G + (1-p)\Phi' \qquad (0<p<1) \tag{154}$$

where Φ_G is a N-extendable Gaussian channel and Φ' is an arbitrary channel.

Proof. To prove this result we use the following Lemma, which is demonstrated below.

Lemma 9. *The thermal-noise channel $\Phi_{\eta,\epsilon}$ can admit a decomposition*

$$\Phi_{\eta,\epsilon} = p\,\Phi_G + (1-p)\Phi' \tag{155}$$

with a Gaussian channel Φ_G characterized by the matrices X and Y (see e.g.[59]), a positive trace preserving map Φ', and some $p>0$, only if

$$X = \sqrt{\eta}\mathbb{1} \qquad Y \leq (1-\eta+\eta\varepsilon)\mathbb{1}. \tag{156}$$

Let us now take a thermal channel $\Phi_{\eta,\epsilon}$ which is not N-extendable, that is with

$$\eta(1-\epsilon/2) > \frac{1}{N} \tag{157}$$

accordingly to [44] (note that on the level of the channel, this is a necessary and sufficient condition). Consider any Gaussian channel Φ_G appearing in the decomposition of Eq. (155). On the one hand we know that it is N-extendable if and only if [44]

$$\sqrt{\det Y} \geq 1 - \frac{1}{N} + \left|\det X - \frac{1}{N}\right|. \tag{158}$$

On the other, Lemma 1 implies that $\sqrt{\det Y} \leq (1-\eta+\eta\epsilon)$ and $\det X = \eta$ must hold. By using these identities together with the bound in Eq. (157), it is easy to see that the condition of Eq. (158) can never be satisfied. Hence, the decomposition in Eq. (155) is impossible. □

Proof of Lemma 9. First note that the case $\eta = 0$ is trivial, we thus assume $\eta > 0$.

Consider the map $\tilde{\Phi} = (1-p)\Phi' = \Phi_{\eta,\epsilon} - p\Phi_G$, which must be at least positive for the decomposition to make physical sense. Let us act with $\tilde{\Phi}$ on a coherent state $|\gamma\rangle$, the Q-function of final state $\tilde{\Phi}(|\gamma\rangle\langle\gamma|)$ is given by

$$\tilde{Q}_\gamma(\alpha) = Q^{\eta,\epsilon}_\gamma(\alpha) - p\,Q^G_\gamma(\alpha), \tag{159}$$

where $Q^{\eta,\epsilon}_\gamma$ and Q^G_γ are the Q-functions associated to the states $\Phi_{\eta,\epsilon}(|\gamma\rangle\langle\gamma|)$ and $\Phi_G(|\gamma\rangle\langle\gamma|)$. The Q-function of a state is proportional to the probability density of the output of the heterodyne measurement. Therefore, if $\tilde{\Phi}$ is a positive channel, the function $\tilde{Q}_\gamma(\alpha)$ must be positive for all α and γ, i.e. we must have

$$Q^{\eta,\epsilon}_\gamma(\alpha) \geq p\,Q^G_\gamma(\alpha) \tag{160}$$

for some nonzero p and all α and γ. This is equivalent to requiring that the ratio $\frac{Q^G_\gamma(\alpha)}{Q^{\eta,\epsilon}_\gamma(\alpha)}$ is bounded. Since both distributions are Gaussian this is equivalent to the following bound

$$(\boldsymbol{\alpha}-\boldsymbol{\mu}_{th})^T V_{th}^{-1}(\boldsymbol{\alpha}-\boldsymbol{\mu}_{th})^T - (\boldsymbol{\alpha}-\boldsymbol{\mu}_G)^T V_G^{-1}(\boldsymbol{\alpha}-\boldsymbol{\mu}_G) < \infty, \tag{161}$$

where for the thermal channel $\Phi_{\eta,\epsilon}$ we have $\boldsymbol{\mu}_{th} = \sqrt{\eta}\boldsymbol{\gamma}$ and $V_{th} = (2+\epsilon\eta)\mathbb{1}$, while for a general Gaussian channel $V_G = XX^T + Y + \mathbb{1}$ and $\boldsymbol{\mu}_G = X\boldsymbol{\gamma} + \boldsymbol{\delta}$ accordingly to Eqs. (152,153). This condition is only nontrivial in the limit where $|\alpha|, |\gamma|$ or both go to infinity. In this limit, $\boldsymbol{\delta}$ plays no role, and we can thus ignore it.

Plugging the values of $\boldsymbol{\mu}_{th}$, V_{th} and $\boldsymbol{\mu}_G$ in Eq. (161) we obtain

$$\frac{|\alpha - \sqrt{\eta}\gamma|^2}{2+\eta\epsilon} - (\boldsymbol{\alpha} - X\boldsymbol{\gamma})^T (V_G)^{-1}(\boldsymbol{\alpha} - X\boldsymbol{\gamma}) \leq \infty. \tag{162}$$

This is a bilinear form in the varibles $\boldsymbol{\xi} = (\mathrm{Re}\alpha, \mathrm{Im}\alpha, \mathrm{Re}\gamma, \mathrm{Im}\gamma)^T$ and can be expressed as

$$\boldsymbol{\xi}^T M \boldsymbol{\xi} \leq \infty \tag{163}$$

$$M = \left(\begin{array}{c|c} \frac{\mathbb{1}}{2+\eta\epsilon} - V_G^{-1} & -\frac{\sqrt{\eta}\mathbb{1}}{2+\eta\epsilon} + X^T V_G^{-1} \\ \hline -\frac{\sqrt{\eta}\mathbb{1}}{2+\eta\epsilon} + V_G^{-1} X & \frac{\eta\mathbb{1}}{2+\eta\epsilon} + X^T V_G^{-1} X \end{array}\right), \tag{164}$$

or simply $M \leq 0$. The positivity of this matrix M remains unchanged if we multiply it by the diagonal matrix

$$\mathrm{diag}(\sqrt{2+\eta\epsilon}, \sqrt{2+\eta\epsilon}, -\frac{\sqrt{2+\eta\epsilon}}{\sqrt{\eta}}, -\frac{\sqrt{2+\eta\epsilon}}{\sqrt{\eta}}) \tag{165}$$

from the left and from the right. This allows us to rewrite the condition as

$$\left(\begin{array}{c|c} \mathbb{1} - W^{-1} & \mathbb{1} - Z^T W^{-1} \\ \hline \mathbb{1} - W^{-1} Z & \mathbb{1} - Z^T W^{-1} Z \end{array}\right) \leq 0 \tag{166}$$

or

$$\mathbb{1}_4 \leq \left(\begin{array}{c|c} W^{-1} & Z^T W^{-1} \\ \hline W^{-1} Z & Z^T W^{-1} Z \end{array}\right) \tag{167}$$

for $W^{-1} = (2+\eta\epsilon)V_G^{-1}$ and $Z = X/\sqrt{\eta}$.

We now consider two cases. First, assume that the matrix X (and hence Z) is invertible. Multiplying the Eq. (167) with $\mathrm{diag}(\mathbb{1}|Z^{-1})$ from the right and $\mathrm{diag}(\mathbb{1}|(Z^{-1})^T)$ from the left does not change the positivity of a matrix, and allows us to cast the inequality in the form

$$\left(\begin{array}{c|c} \mathbb{1} & (Z^{-1})^T \\ \hline Z^{-1} & (Z^{-1})^T Z \end{array}\right) \leq \left(\begin{array}{c|c} W^{-1} & W^{-1} \\ \hline W^{-1} & W^{-1} \end{array}\right). \tag{168}$$

The right hand side is block-diagonal $\binom{1}{1}\binom{1}{1}^T \otimes W^{-1}$. And since the inequality must hold in both blocks, we get

$$(\mathbb{1} + {Z^{-1}}^T)(\mathbb{1} + Z^{-1}) \leq 4W^{-1} \tag{169}$$

$$(\mathbb{1} - {Z^{-1}}^T)(\mathbb{1} - Z^{-1}) \leq 0. \tag{170}$$

The second inequality implies $\mathbb{1} - Z^{-1} = 0$, or $X = \sqrt{\eta}\mathbb{1}$. This reduces the fist one to $W^{-1} \geq \mathbb{1}$, or $W \leq \mathbb{1}$. Recalling that $W = V_G/(2+\eta\epsilon)$ with $V_G = XX^T + Y + 1 = \eta\mathbb{1} + Y + 1$ we obtain

$$\frac{(1+\eta)\mathbb{1} + Y}{2 + \eta\epsilon} \leq \mathbb{1}, \tag{171}$$

leading to

$$Y \leq (1 - \eta + \eta\epsilon)\mathbb{1}. \tag{172}$$

This proves the Lemma for an invertible X.

Next, consider the remaining case where X is not invertible. There is thus a normalized vector $\boldsymbol{v}$ such that $X\boldsymbol{v} = Z\boldsymbol{v} = 0$. Multiplying the Eq (167) with $(0, 0|\boldsymbol{v})$ from the right and $(0, 0|\boldsymbol{v})^T$ from the left we find

$$1 \leq 0. \tag{173}$$

X must thus be invertible, which concludes the proof. □

Note that the Lemma is most probably also true with an if and only statement. The if direction could be shown by an explicit decomposition, noting that the thermal-noise channel can be viewed as a loss channel followed by a random displacement with a normally distributed complex amplitude.

References

[1] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-independent security of quantum cryptography against collective attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[2] I. W. Primaatmaja, K. T. Goh, E. Y.-Z. Tan, J. T.-F. Khoo, S. Ghorai, and C. C.-W. Lim, Security of device-independent quantum key distribution protocols: a review, Quantum **7**, 932 (2023).

[3] M. Pawłowski and N. Brunner, Semi-device-independent security of one-way quantum key distribution, Phys. Rev. A **84**, 010302 (2011).

[4] E. Woodhead and S. Pironio, Secrecy in prepare-and-measure clauser-horne-shimony-holt tests with a qubit bound, Phys. Rev. Lett. **115**, 150501 (2015).

[5] J. S. Bell, On the einstein podolsky rosen paradox, Physics Physique Fizika **1**, 195 (1964).

[6] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[7] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, entanglement, nonlocality, and the einstein-podolsky-rosen paradox, Phys. Rev. Lett. **98**, 140402 (2007).

[8] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020).

[9] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Hacking commercial quantum cryptography systems by tailored bright illumination, Nature Photonics **4**, 686 (2010).

[10] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Full-field implementation of a perfect eavesdropper on a quantum cryptography system, Nature Communications **2**, 1 (2011).

[11] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs, Device calibration impacts security of quantum key distribution, Phys. Rev. Lett. **107**, 110501 (2011).

[12] A. N. Bugge, S. Sauge, A. M. M. Ghazali, J. Skaar, L. Lydersen, and V. Makarov, Laser damage helps the eavesdropper in quantum cryptography, Phys. Rev. Lett. **112**, 070503 (2014).

[13] P. H. Eberhard, Background level and counter efficiencies required for a loophole-free einstein-podolsky-rosen experiment, Phys. Rev. A **47**, R747 (1993).

[14] S. Massar and S. Pironio, Violation of local realism vs detection efficiency, Phys. Rev. A **68**, 062109 (2003), arXiv:quant-ph/0210103.

[15] T. Vértesi, S. Pironio, and N. Brunner, Closing the detection loophole in bell experiments using qudits, Phys. Rev. Lett. **104**, 060401 (2010).

[16] C. Branciard, Detection loophole in bell experiments: How postselection modifies the requirements to observe nonlocality, Phys. Rev. A **83**, 032123 (2011).

[17] J.-Å. Larsson, Loopholes in bell inequality tests of local realism, J. Phys. A **47**, 424003 (2014).

[18] B. Wittmann, S. Ramelow, F. Steinlechner, N. K. Langford, N. Brunner, H. M. Wiseman, R. Ursin, and A. Zeilinger, Loophole-free einstein–podolsky–rosen experiment via quantum steering, New J. Phys. **14**, 053030 (2012).

[19] A. J. Bennet, D. A. Evans, D. J. Saunders, C. Branciard, E. G. Cavalcanti, H. M. Wiseman, and G. J. Pryde, Arbitrarily loss-tolerant einstein-podolsky-rosen steering allowing a demonstration over 1 km of optical fiber with no detection loophole, Phys. Rev. X **2**, 031003 (2012).

[20] V. Srivastav, N. H. Valencia, W. McCutcheon, S. Leedumrongwatthanakun, S. Designolle, R. Uola, N. Brunner, and M. Malik, Quick quantum steering: Overcoming loss and noise with qudits, Phys. Rev. X **12**, 041023 (2022).

[21] A. Acín, D. Cavalcanti, E. Passaro, S. Pironio, and P. Skrzypczyk, Necessary detection efficiencies for secure quantum key distribution and bound randomness, Phys. Rev. A **93**, 012319 (2016).

[22] M. Ioannou, M. A. Pereira, D. Rusca, F. Grünenfelder, A. Boaron, M. Perrenoud, A. A. Abbott, P. Sekatski, J.-D. Bancal, N. Maring, *et al.*, Receiver-device-independent quantum key distribution, Quantum **6**, 718 (2022).

[23] B. Hensen *et al.*, Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres, Nature **526**, 682 (2015).

[24] L. K. Shalm *et al.*, Strong loophole-free test of local realism, Phys. Rev. Lett. **115**, 250402 (2015).

[25] M. Giustina *et al.*, Significant-loophole-free test of bell's theorem with entangled photons, Phys. Rev. Lett. **115**, 250401 (2015).

[26] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by bell's theorem, Nature **464**, 1021 (2010).

[27] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J.-D. Bancal, Experimental quantum key distribution certified by bell's theorem, Nature **607**, 682 (2022).

[28] W. Zhang, T. van Leent, K. Redeker, R. Garthoff, R. Schwonnek, F. Fertig, S. Eppelt, W. Rosenfeld, V. Scarani, C. C.-W. Lim, and H. Weinfurter, A device-independent quantum key distribution system for distant users, Nature **607**, 687 (2022).

[29] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A **49**, 123001 (2016).

[30] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum Measurement* (Springer International Publishing, 2016).

[31] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Colloquium: Incompatible measurements in quantum information science, Rev. Mod. Phys. **95**, 011003 (2023).

[32] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Device-independent tests of classical and quantum dimensions, Phys. Rev. Lett. **105**, 230501 (2010).

[33] Y. Wang, I. W. Primaatmaja, E. Lavie, A. Varvitsiotis, and C. C. W. Lim, Characterising the correlations of prepare-and-measure quantum networks, npj Quantum Information **5**, 10.1038/s41534-019-0133-3 (2019).

[34] M. T. Quintino, T. Vértesi, and N. Brunner, Joint measurability, einstein-podolsky-rosen steering, and bell nonlocality, Phys. Rev. Lett. **113**, 160402 (2014).

[35] R. Uola, T. Moroder, and O. Gühne, Joint Measurability of Generalized Measurements Implies Classicality, Phys. Rev. Lett. **113**, 160403 (2014), publisher: American Physical Society.

[36] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, Phys. Rev. A **92**, 022115 (2015).

[37] J. Bavaresco, M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, Most incompatible measurements for robust steering tests, Phys. Rev. A **96**, 022110 (2017).

[38] S. Designolle, P. Skrzypczyk, F. Fröwis, and N. Brunner, Quantifying measurement incompatibility of mutually unbiased bases, Phys. Rev. Lett. **122**, 050402 (2019).

[39] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New J. Phys. **21**, 113053 (2019).

[40] P. Skrzypczyk and D. Cavalcanti, Loss-tolerant einstein-podolsky-rosen steering for arbitrary-dimensional states: Joint measurability and unbounded violations under losses, Phys. Rev. A **92**, 022354 (2015).

[41] M. Ioannou, P. Sekatski, S. Designolle, B. D. M. Jones, R. Uola, and N. Brunner, Simulability of high-dimensional quantum measurements, Phys. Rev. Lett. **129**, 190401 (2022).

[42] P. Sekatski, F. Giraud, R. Uola, and N. Brunner, Unlimited one-way steering, Phys. Rev. Lett. **131**, 110201 (2023).

[43] P. Sekatski, Compatibility of projective measurements subject to white noise and loss, Phys. Rev. A **109**, 022215 (2024).

[44] L. Lami, S. Khatri, G. Adesso, and M. M. Wilde, Extendibility of bosonic gaussian states, Phys. Rev. Lett. **123**, 050501 (2019).

[45] S. Rahimi-Keshari, M. Mehboudi, D. De Santis, D. Cavalcanti, and A. Acín, Verification of joint measurability using phase-space quasiprobability distributions, Phys. Rev. A **104**, 042212 (2021).

[46] M. Curty, M. Lewenstein, and N. Lütkenhaus, Entanglement as a Precondition for Secure Quantum Key Distribution, Phys. Rev. Lett. **92**, 217903 (2004), publisher: American Physical Society.

[47] E. P. Lobo, J. Pauwels, and S. Pironio, Certifying long-range quantum correlations through routed bell tests, arXiv preprint arXiv:2310.07484 (2023).

[48] M. F. Pusey, Verifying the quantumness of a channel with an untrusted device, J. Opt. Soc. Am. B 10.1364/JOSAB.32.000A56 (2015).

[49] A. Holevo, M. Shirokov, and R. Werner, Separability and entanglement-breaking in infinite dimensions, arXiv preprint quant-ph/0504204 (2005).

[50] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz, Incompatibility breaking quantum channels, J. Phys. A **48**, 435301 (2015).

[51] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Measurements incompatible in quantum the-

ory cannot be measured jointly in any other no-signaling theory, Phys. Rev. Lett. **103**, 230402 (2009).

[52] R. F. Werner, Optimal cloning of pure states, Phys. Rev. A **58**, 1827 (1998).

[53] M. Keyl and R. F. Werner, Optimal cloning of pure states, testing single clones, J. Math. Phys. **40**, 3283 (1999).

[54] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Colloquium: Incompatible measurements in quantum information science, Rev. Mod. Phys. **95**, 011003 (2023).

[55] P. Busch, Unsharp reality and joint measurements for spin observables, Phys. Rev. D **33**, 2253 (1986).

[56] R. Pal and S. Ghosh, Approximate joint measurement of qubit observables through an arthur–kelly model, J. Phys. A **44**, 485303 (2011).

[57] S. Yu and C. Oh, Quantum contextuality and joint measurement of three observables of a qubit, arXiv preprint arXiv:1312.6470 (2013).

[58] R. Garcia-Patron Sanchez, Quantum information with optical continuous variables: from bell tests to key distribution (2007).

[59] A. Serafini, *Quantum continuous variables: a primer of theoretical methods* (CRC press, 2017).

[60] J. Kiukas and J. Schultz, Informationally complete sets of gaussian measurements, J. Phys. A **46**, 485303 (2013).

[61] R. Renner and S. Wolf, New bounds in secret-key agreement: The gap between formation and secrecy extraction, in *Advances in Cryptology—EUROCRYPT 2003: International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland, May 4–8, 2003 Proceedings 22* (Springer, 2003) pp. 562–577.

[62] I. Devetak and A. Winter, Distillation of secret key and entanglement from quantum states, Proc. R. Soc. Lond. A **461**, 207 (2005).

[63] R. Renner, N. Gisin, and B. Kraus, Information-theoretic security proof for quantum-key-distribution protocols, Phys. Rev. A **72**, 012332 (2005).

[64] J. Kołodyński, A. Máttar, P. Skrzypczyk, E. Woodhead, D. Cavalcanti, K. Banaszek, and A. Acín, Device-independent quantum key distribution with single-photon sources, Quantum **4**, 260 (2020).

[65] In the case of DIQKD, our attacks improve on the ones of [64] because they take into account not only losses but also white-noise. However, we can further improve these attacks by targeting both CMUs of a DIQKD protocol, see next subsection. We also remark that it is important to consider separately the case where a no-click outcome is kept as a distinct outcome in Alice and Bob's post-processing and the case where it is discarded. In [64] only the former case is analyzed. An example of a DIQKD protocol where binning of the no-click outcome leads to a positive key rate for a detection efficiency that is lower than the threshold computed in [64] can be found in [82].

[66] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, One-sided device-independent qkd and position-based cryptography from monogamy games, in *Advances in Cryptology–EUROCRYPT 2013: 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26-30, 2013. Proceedings 32* (Springer, 2013) pp. 609–625.

[67] E. Woodhead, Semi device independence of the bb84 protocol, New J. Phys. **18**, 055010 (2016).

[68] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering, Phys. Rev. A **85**, 010301 (2012).

[69] M. Masini and S. Sarkar, One-sided di-qkd secure against coherent attacks over long distances, arXiv preprint arXiv:2403.11850 (2024).

[70] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-independent security of quantum cryptography against collective attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[71] T. Leroy, E. Peter Lobo, J. Pauwels, and S. Pironio, Device-independent quantum key distribution based on routed bell experiments (2024).

[72] M. Ioannou, P. Sekatski, A. A. Abbott, D. Rosset, J.-D. Bancal, and N. Brunner, Receiver-device-independent quantum key distribution protocols, New J. Phys. **24**, 063006 (2022).

[73] C. H. Bennett, Quantum cryptography using any two nonorthogonal states, Phys. Rev. Lett. **68**, 3121 (1992).

[74] M. Farkas, M. Balanzó-Juandó, K. Łukanowski, J. Kołodyński, and A. Acín, Bell nonlocality is not sufficient for the security of standard device-independent quantum key distribution protocols, Phys. Rev. Lett. **127**, 050503 (2021).

[75] K. Łukanowski, M. Balanzó-Juandó, M. Farkas, A. Acín, and J. Kołodyński, Upper bounds on key rates in device-independent quantum key distribution based on convex-combination attacks, Quantum **7**, 1199 (2023).

[76] A. Acin, S. Massar, and S. Pironio, Efficient quantum key distribution secure against no-signalling eavesdroppers, New J. Phys. **8**, 126 (2006).

[77] J. A. Nelder and R. Mead, A Simplex Method for Function Minimization, The Computer Journal **7**, 308 (1965).

[78] T. Heinosaari, D. Reitzner, and P. Stano, Notes

on joint measurability of quantum observables, Foundations of Physics **38**, 1133 (2008).

[79] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity, Physics Reports **506**, 1 (2011).

[80] F. Buscemi, K. Kobayashi, S. Minagawa, P. Perinotti, and A. Tosini, Unifying different notions of quantum incompatibility into a strict hierarchy of resource theories of communication, Quantum **7**, 1035 (2023).

[81] G. M. D'Ariano, P. Perinotti, and A. Tosini, Incompatibility of observables, channels and instruments in information theories, J. Phys. A **55**, 394006 (2022).

[82] M. Masini, S. Pironio, and E. Woodhead, Simple and practical DIQKD security analysis via BB84-type uncertainty relations and Pauli correlation constraints, Quantum **6**, 843 (2022).