Quantum non-locality Fundamentals and Applications in Quantum Information Science

BRUNNER, Nicolas

Abstract

This thesis is devoted to the study of quantum non-locality, in the framework of quantum information science. The research presented here includes both theoretical works on some fundamental concepts, such as the measurement process and quantum correlations, as well as some applications in modern quantum optics. As it is often the case in quantum information science, the joint study of fundamental notions and applications is very productive. While a better understanding of the fundamental concepts may lead to new applications, it happens also that applications shed new light on some fundamental notions...

Reference


URN : urn:nbn:ch:unige-7237
DOI : 10.13097/archive-ouverte/unige:723
Quantum non-locality

Fundamentals and Applications in Quantum Information Science

Thèse

présentée à la Faculté des sciences de l’Université de Genève
pour obtenir le grade de Docteur ès sciences, mention physique

par

Nicolas BRUNNER

de

Genève (GE)

Thèse N° 3910

Genève
Atelier de reproduction de la Section de physique
2007
"Quantum Non-Localité
Fundamentals and Applications in Quantum Information Science"

La Faculté des sciences, sur le préavis de Messieurs N. Gisin, professeur ordinaire et directeur de thèse (Groupe de physique appliquée), C. Simon, docteur (Groupe de physique appliquée), S. Popescu, professeur (University of Bristol – Department of Mathematics – Bristol, United Kingdom), A. Acin, professeur (Institut de Ciències Fotòniques – Castelldefels (Barcelona), Spain), et V. Scarani, National University of Singapore – Faculty of Science – Department of Physics – Quantum Information Science and Technology Group – Singapore), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 25 juin 2007

Thèse - 3910 -

Le Doyen, Jean-Marc Triscone
Scientists have many, many failures and then they get a breakthrough.  
But it’s all based on the failures.

David Lynch (1946)
Abstract

This thesis is devoted to the study of quantum non-locality, in the framework of quantum information science. The research presented here includes both theoretical works on some fundamental concepts, such as the measurement process and quantum correlations, as well as some applications in modern quantum optics. As it is often the case in quantum information science, the joint study of fundamental notions and applications is very productive. While a better understanding of the fundamental concepts may lead to new applications, it happens also that applications shed new light on some fundamental notions.

First we will point out a nice analogy between two apparently unrelated topics in physics, namely the abstract theory of weak measurements and polarization effects in telecom fiber networks. The quantum formalism is shown to simplify the calculations of polarization effects in telecom networks. Furthermore our analogy clarifies the concept of weak measurements, often thought to be a weirdness of theorists. Based on this analogy, we experimentally demonstrate slow and fast light in a simple optical network. We present direct measurements of superluminal group velocities as well as signal velocities. Our results clearly illustrate the difference between group and signal velocities.

Then we explore both concepts of entanglement and non-locality. Using extensively Bell inequalities, in particular their geometrical representation, we address a wide class of problems, such as the simulation of entanglement, the detection loophole and cryptography. We present results for the simulation of entanglement with non-local resources. Our results strongly suggest that entanglement and non-locality are different resources. We discuss also the detection loophole in asymmetric Bell tests. Threshold detection efficiencies are derived for the case where one of the two particles is measured with a perfect efficiency, as for example in entangled atom-photon systems. According to our results the detection loophole might be closed for well-separated atom and photon with today’s technology. A work in progress addresses the problem of detecting the dimension of the Hilbert of some given quantum correlations. Contributions of this thesis include new security proofs for quantum cryptography.

Finally we present two applications in quantum information science. Both come from the so-called time-bins encoding of quantum information. The first is a novel scheme for the Bell-state measurement (BSM), which plays a crucial role in many quantum information protocols. This BSM is implemented with linear optics only and can distinguishing three out the four Bell states, an improvement compared to previous schemes. The second application is a fast and simple quantum key distribution (QKD) protocol.
Résumé

Cette thèse est dédiée à l’étude de la non-localité quantique, dans le cadre de la science de l’information quantique. Les recherches présentées ici incluent des travaux théoriques sur certains concepts fondamentaux de la physique quantique, comme la théorie de la mesure et les corrélations quantiques, ainsi que des applications en optique quantique moderne. Comme c’est souvent le cas en science de l’information quantique, l’étude conjointe des notions fondamentales ainsi que des applications est très fructueuse. D’une part, une meilleure compréhension de la théorie conduit souvent à de nouvelles applications. D’autre part, certaines applications mettent en lumière des concepts fondamentaux parfois mal compris. Cette thèse s’inscrit parfaitement dans cette optique.

Tout d’abord nous mettrons en évidence une analogie entre deux sujets de la physique moderne qui, a priori, ne semblent avoir aucun lien entre eux. Il s’agit de la théorie des mesures faibles en mécanique quantique, et des effets de polarisations dans les réseaux optiques. Premièrement, nous montrerons comment le formalisme quantique des mesures faibles avec post-sélection, simplifie le calcul des effets de polarisation dans les fibres optique. De plus cette analogie permettra de comprendre les mesures faibles, souvent considérées comme une curiosité de théoriciens, sous un nouveau jour. Ensuite, nous profiterons de cette analogie pour créer une expérience dite de slow and fast light. Dans cette expérience nous démontrerons le contrôle de la vitesse de groupe d’une impulsion lumineuse à travers une fibre optique. En particulier une situation de propagation supraluminique sera discutée. Ici nous insisterons sur la différence entre les concepts de vitesse de groupe et vitesse de signal.

Dans une seconde partie, nous explorerons les concepts d’intrication et de non-localité. A l’aide des inégalités de Bell, en particulier de leur représentation géométrique, nous adresserons une large classe de problèmes comme la simulation de l’intrication, le detection loophole et la cryptographie. Tout ces sujets sont étroitement liés. Concernant la simulation de l’intrication par des ressources non-locales, notre résultat principal suggère que l’intrication et la non-localité sont véritablement deux ressources différentes. Nous discuterons également du detection loophole dans des expériences de Bell asymétriques. Les efficacités minimales de détection sont dérivées, en particulier dans le cas où l’une des deux particules est détectée à chaque fois, comme par exemple dans un système d’atome-photon intriqués. D’après nos résultats, une expérience fermant le detection loophole serait envisageable avec les technologies actuelles. Un travail en cours présentera quelques résultats sur la détection de la dimension de l’espace de Hilbert pour des corrélations quantiques. Des contributions de cette thèse évoquent de nouvelles preuves de sécurité pour la cryptographie quantique.

Finalement nous présenterons encore deux applications en science de l’information quantique. Chacune utilise le codage de l’information en temps (time-bins).
La première est une nouveau schéma de la mesure de Bell, un ingrédient clé dans beaucoup de protocoles de l’information quantique. La particularité de cette mesure de Bell est qu’elle permet de distinguer trois parmi les quatre états de Bell, ce qui constitue une amélioration par rapport au schéma traditionnel. La seconde application est un protocole de distribution de clé quantique à la fois rapide et simple.
# Contents

Abstract  
Résumé

## I English version

1 Introduction  

2 Telecom networks as weak quantum measurements with post-selection  
  2.1 PMD performs polarization measurements  
  2.2 Post-selection and weak value  
    2.2.1 Post-selection on a pure state  
    2.2.2 Finite PDL and general networks  
  2.3 Conclusion

3 Superluminal group velocity in an optical fiber  
  3.1 Tailoring the speed of light  
  3.2 Slow and fast-light in an optical fiber  
  3.3 Experimental setup and results  
  3.4 Conclusion

4 Three-Bell-State Measurement  
  4.1 Bell measurement  
  4.2 Implementations of the three-Bell-state measurement  
    4.2.1 Time-bins implementation  
    4.2.2 Polarization implementation  
  4.3 Teleportation experiment  
  4.4 Conclusion

5 Bell-type inequalities for non-local resources  
  5.1 Bell inequalities, polytopes, and non-local machines
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>Entanglement and non-locality are different resources</td>
<td>26</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Three settings on each side</td>
<td>26</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Extension to more settings</td>
<td>28</td>
</tr>
<tr>
<td>5.3</td>
<td>Bell-type inequalities for stronger non-local resources</td>
<td>28</td>
</tr>
<tr>
<td>5.4</td>
<td>Outlook</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>Detection loophole in asymmetric Bell experiments</td>
<td>30</td>
</tr>
<tr>
<td>6.1</td>
<td>Detection loophole in asymmetric Bell experiments</td>
<td>31</td>
</tr>
<tr>
<td>6.2</td>
<td>Threshold detection efficiencies</td>
<td>32</td>
</tr>
<tr>
<td>6.3</td>
<td>Experimental feasibility and outlook</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>Hilbert space dimension witnesses</td>
<td>35</td>
</tr>
<tr>
<td>7.1</td>
<td>CGLMP</td>
<td>36</td>
</tr>
<tr>
<td>7.2</td>
<td>Linear dimension witnesses</td>
<td>36</td>
</tr>
<tr>
<td>7.3</td>
<td>Outlook</td>
<td>38</td>
</tr>
<tr>
<td>8</td>
<td>Other contributions</td>
<td>39</td>
</tr>
<tr>
<td>8.1</td>
<td>Fast and simple one-way quantum key distribution</td>
<td>39</td>
</tr>
<tr>
<td>8.2</td>
<td>Device-independent quantum cryptography</td>
<td>41</td>
</tr>
<tr>
<td>II</td>
<td>Version française</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>Introduction</td>
<td>46</td>
</tr>
<tr>
<td>10</td>
<td>Réseau télécom vus comme des mesures faibles avec post-sélection</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>Slow- et fast-light dans une fibre optique</td>
<td>52</td>
</tr>
<tr>
<td>12</td>
<td>Mesure de Bell</td>
<td>54</td>
</tr>
<tr>
<td>13</td>
<td>Inégalité de Bell pour des ressources non-locales</td>
<td>55</td>
</tr>
<tr>
<td>14</td>
<td>Loophole de détection pour les tests de Bell asymétriques</td>
<td>57</td>
</tr>
<tr>
<td>15</td>
<td>Détecter la dimension de l’espace de Hilbert</td>
<td>59</td>
</tr>
<tr>
<td>16</td>
<td>Contributions en cryptographie quantique</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Bibliography / Bibliographie</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Publication list / Liste des publications</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Acknowledgements / Remerciements</td>
<td>69</td>
</tr>
</tbody>
</table>
III Published articles / 
Articles publiés 70
Part I

English version
Chapter 1

Introduction

A few years after the discovery of modern quantum mechanics, it was realized that, though the theory had an unbelievable predictive power, it contained very counterintuitive features. These strange properties were at the core of the famous dialogue between Niels Bohr and Albert Einstein. Initially Einstein tried to argue that quantum mechanics is inconsistent, but he later reformulated his argument towards demonstrating that quantum mechanics is incomplete. This argument has been made famous through the seminal paper of Einstein, Podolsky and Rosen (EPR) [1]. EPR pointed out that entanglement is in logical conflict with locality.

For many years the EPR debate went on, but was merely considered as a philosophical issue by the vast majority of the scientific community. However in 1964, the discussion was revived by John Bell, who came out with a way of testing the completeness of quantum mechanics: the famous Bell inequality [2]. In fact this discovery, which certainly ranks among the most profound discoveries of science, is not specifically about quantum mechanics. It is rather a general proof that there is an upper limit to the correlations of distant events, if one assumes the principle of locality.

Indeed Bell’s work triggered renewed interest for quantum non-locality. A bunch of physicists underwent the challenging task to experimentally test quantum non-locality in the light of Bell’s result [3]. This research culminated in 1982 when Alain Aspect and co-workers performed the first Bell test with space-like separated measurements [4]. Indeed there results confirmed the quantum prediction.

So in half of a century, quantum non-locality evolved from a mere philosophical issue to an experimentally demonstrated feature of quantum mechanics. No matter how fascinating this property is, the vast majority of the physicists did not pay much attention to it. The reason of this disinterest is that a kind of general consensus had been established that non-locality was nothing but a laboratory curiosity, that would certainly find some rational explanation one day.

In the nineties however the situation changed dramatically. Thanks to some visionary physicists and information scientists, it was realized that entanglement is not only a fascinating concept, but can in fact be useful. It can even accomplish tasks which are impossible classically! It is worth mentioning Arthur Ekert’s breakthrough [5]: quantum non-locality can be exploited to distribute a cryptographic
key between two distant partners. Furthermore the secrecy of the key could be verified with a Bell inequality, and guaranteed against an eavesdropper only limited by the laws of physics! This was the beginning of the science of quantum information, which in less than twenty years, has tremendously developed [6].

Until recently, entanglement and non-locality have been thought as being two different aspects of the same phenomena. However the truth is by far more interesting! It was realized some years ago, that non-locality could be studied for its own sake, without invoking quantum mechanics at all. This conceptual breakthrough has its origins in a seminal paper by Popescu and Rohrlich [7, 8], which showed that quantum correlations are only part of the game; there are correlations more powerful than those of quantum mechanics that do not allow the transmission of information faster than light. Such correlations are said to be no-signaling. Remarkably it turned out that many of the fascinating properties of quantum physics, such as no-cloning, monogamy or uncertainty relations, are in fact general properties of any no-signaling theory [9]. Furthermore the new tools developed for the study of non-locality, allow to quantify the power of quantum correlations, which is indeed a central issue in quantum information science.

This thesis includes many different works, ranging from fundamental topics in quantum theory to applied modern quantum optics. Indeed some of these parts may seem quite unrelated to each other at first sight, but there are all in some sense different manifestations of the same phenomenon: quantum non-locality. The spirit of this research is that a better understanding of the fundamental concepts of quantum theory may lead to new applications. From a fundamental point of view, the different works presented here study two main topics in modern quantum physics: the measurement process and quantum correlations.

**Measurements in Quantum Mechanics**

The measurement process is central in quantum physics, because it translates the Hilbert space into classical language. In this thesis we will evoke two modern concepts of quantum measurements: weak measurements [10] and positive-operator-valued-measures (POVM) [11].

*Weak measurements.* First we will point out a nice analogy between the abstract theory of weak measurements and polarization effects in telecom fiber networks (Chap. 2), two topics in modern physics which seem quite unrelated at first glance. In particular we will show how the quantum formalism simplifies the calculations for the combined effect of two polarization effects in fiber optics. Furthermore our analogy sheds new light on the concept of weak measurements which was often thought to be a weirdness of theorists. It is important to mention that entanglement plays a crucial role here. In fact the degree of entanglement between the pointer and the physical system to be measured will characterize the strength of the quantum measurement: strong entanglement will lead to a strong measurement, while an almost separable state will lead to a weak measurement. Then, based on this analogy, we will experimentally demonstrate slow and fast light in an optical fiber (Chap. 3). The results include direct measurements of superluminal group velocities as well as signal velocities. It is stressed that though the group velocity may exceed the speed of light $c$, the signal velocity, the *information* speed, remains constant and smaller.
POVM. We present a novel scheme for the Bell-state measurement (BSM), which is a key ingredient in many quantum information protocols [12, 13]. Our scheme is a linear optics implementation of a 21-outcomes POVM. This POVM allows to distinguish three-out-of-four Bell states, contrary to previous linear optics scheme which only discriminate two Bell states [14]. This novel BSM achieves the optimal overall efficiency of 50% [15]. It was also successfully implemented in a recent teleportation experiment.

Quantum correlations and Bell inequalities

Through the study of non-locality, in particular with the tool of Bell inequalities we address four main problematics. Indeed there are many connections between them.

Entanglement and non-locality. By exploring the geometrical representation of Bell inequalities [8], we present some results which point out differences between the concepts of entanglement and non-locality. Our main tools are polytopes and non-local machines. First we will prove that some non-maximally entangled states require a strictly larger amount of non-local resources than maximally entangled ones, in order to be simulated (Chap. 5). This suggests that entanglement and non-locality are different resources, an hypothesis which has been formulated in previous works.

Detection loophole. The problem of closing the detection loophole in asymmetric Bell tests is addressed (Chap. 6). In particular we will show that if one of the particles is always detected, an efficiency as low as 43% can be tolerated for the other particle. In analogy to the symmetric case, we find that non-maximally entangled states help: they can tolerate a smaller detection efficiency than the maximally entangled one. So, in some sense they contain more non-locality than maximally entangled states, since a smaller detection efficiency suffices to reveal their non-locality. This confirms the intuition of Chap. 5. Finally we also discuss practical implementations. In the light of our results, closing the detection loophole in asymmetric Bell tests seems realistic with today’s technology.

Detecting the Hilbert space dimension. We will discuss the problem of testing the dimensionality of some quantum correlations (Chap. 7). This is a new problematic. The main questions will be presented, as well as some first results. Again the geometry of Bell inequalities will be our main tool.

Cryptography. Though quantum cryptography does not require entanglement in practice, it is a key ingredient in most security proofs of QKD [16]. Contributions of this thesis go to the security proof of QKD against a post-quantum Eve, as well as to a device-independent security proof of quantum cryptography (Chap. 8). In both of these results, the security of the protocol is guaranteed through the violation of a Bell inequality. Finally a last contribution presents a novel protocol for practical QKD, which has both advantages of being fast and simple (Chap. 8).
Each chapter of this manuscript gives a summary of some particular result. The chapters are self-contained and can be read individually. Most of the results presented here have been published, and the corresponding papers can be found at the end of this thesis. Throughout the text these articles are referenced alphabetically, while other articles are referred numerically. Indeed the more interested reader should go to the published articles for more details.
Chapter 2

Telecom networks as weak quantum measurements with post-selection

Several times in the history of science, it appeared that different scientists working in different fields with different motivations discovered the same thing. A famous example is the connection between Einstein’s theory of general relativity and differential geometry. Indeed both fields benefited from this connection: physics was provided with a powerful tool, while mathematics gained in popularity because it proved to be useful. In a similar spirit this section points out an analogy between the abstract theory of weak quantum measurements followed by post-selection and the physics of telecom networks. More precisely it is shown how the quantum formalism of weak measurements simplifies the mathematical study of polarization effects in optical networks.

With the introduction of long-distance optical networks, telecommunication systems are becoming more and more complex. Typically a modern systems includes several different components such as couplers, wavelength division multiplexing (WDM), optical isolators and circulators, which are located between the trunks of optical fibers. One of the most problematic aspects of such networks, are the optical distortions induced by the combination of two polarization effects: polarization-mode-dispersion (PMD) and polarizing-dependent-losses (PDL) \[17\]. The first, PMD, is important in optical fibers. It is due to the fiber’s birefringence; the propagation speed of an optical pulse depends on its state of polarization, in particular the fastest and the slowest modes correspond to orthogonal polarizations. The second effect, PDL, is significant in the optical components placed between the fibers. Typically these components have no PMD while the fibers have negligible PDL. Thus an optical network can be described as a concatenation of trunks, alternating PMD and PDL.

The ”standard” formalism for addressing the effect of PMD and PDL is rather complicated. Calculations, based on Jones vectors \[18\], are lengthy and must be done numerically for realistic networks. Remarkably it has been shown that the combined effect of PMD and PDL can give rise to interesting phenomena, such as
anomalous dispersion [17], even for very simple networks, for example when a PDL element is sandwiched between two PMD elements.

Here we show that optical networks can be seen as weak quantum measurements with post-selection [A]. The analogy basically works in two parts. First we show that PMD performs polarization measurements, and that the telecom limit of PMD corresponds exactly to the regime of a weak measurement of polarization. This is the content of section 2.1. Then we show that PDL provides the necessary post-selection. The concept of post-selection, which may seem quite artificial in the approach of quantum theorists, is here introduced in a very natural way: one only considers those photons that have not been lost in the network! Note that this is non-trivial physics since the post-selection depends on the measured degree of freedom, namely polarization. The case of post-selection on some pure state is treated in section 2.2. Section 2.2.2 presents post-selection on a mixed state as well as the extension of our analogy to general networks. Finally a brief conclusion is given in section 2.3.

2.1 PMD performs polarization measurements

In standard optics, an optical fiber is represented by a channel supporting two orthogonal polarization eigenmodes. Because of birefringence, those two modes propagate at different velocities in the fiber. To quantify this effect, telecom physicists use the concept of differential-group-delay (DGD) which is the time separating the arrival of the slow and the fast eigenmodes. On a monochromatic wave of frequency $\omega$, PMD is represented by a unitary operation of the form $U(\delta\tau, \hat{z}) = \exp(-\frac{i}{2}\delta\tau\omega\hat{z} \cdot \vec{\sigma})$, where $\hat{z}$ is the birefringence axis and $\delta\tau$ is the DGD.

Now let us consider a gaussian optical pulse, with a coherence time $t_c$ and a central frequency $\omega_0$, passing through a fiber with a birefringence along the $z$ axis. Thus the fiber’s eigenmodes are the eigenstates of $\sigma_z$, i.e. $|H\rangle$ and $|V\rangle$. The polarization of the pulse is $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$ where $|\alpha|^2 + |\beta|^2 = 1$. So our initial state reads

$$|\Psi_{IN}\rangle = g(t) \otimes |\psi\rangle$$

(2.1)

where $g(t)$ is defined such that $G(t) \equiv g(t)^2 = \frac{1}{t_c \sqrt{2\pi}} e^{-t^2/2t_c^2}$ is a probability distribution. At the fiber’s output the state evolves to

$$|\Psi_{PMD}\rangle = U(\delta\tau, \hat{z})|\Psi_{IN}\rangle = \tilde{\alpha}g_- (t)|H\rangle + \tilde{\beta}g_+ (t)|V\rangle$$

(2.2)

where $g_\pm \equiv g(t \pm \frac{\delta\tau}{2})$ and $\tilde{\alpha} \equiv e^{i\frac{\delta\tau}{2} \omega_0}\alpha$, $\tilde{\beta} \equiv e^{-i\frac{\delta\tau}{2} \omega_0}\beta$. So the effect of PMD is a spatial separation of the two fiber’s eigenmodes as well as a rotation of the polarization state around the $z$ axis.

At the fiber’s output one can then measure the time of arrival (TOA) of the pulse. If $t_c << \delta\tau$, then the two eigenmodes are spatially well separated (see Fig. 2.1). Thus the TOA gives full information about the polarization state. Note also that in this case the wavefunction of the light pulse is very disturbed; it is a collapse as in the usual picture of a quantum (strong) measurement.
Figure 2.1: PMD performs polarization measurements. The ratio $\delta \tau / t_c$ characterizes the strength of the measurement. In case of a strong measurement, the TOA provides whole information about the polarization state. In case of a weak measurement the TOA provides only gives only partial information about the polarization state. The telecom limit of PMD corresponds to a weak measurement.

However in telecom optics, it is more common to have the opposite limit, $t_c >> \delta \tau$. In this case the TOA measurement provides only partial information on the polarization. In this case the measurement is said to be weak, in the sense meant by quantum theorists [10]. Note also that the wavefunction is only modified a little bit by the measurement process. So there is clearly a trade-off between disturbance of the system and the information one can get from the measurement.

Note also that until the particle is detected by the TOA measurement, the process is completely reversible. Mathematically this is clear, since PMD is a unitary operation. Thus we say that PMD achieves only the pre-measurement - the final collapse being done when the TOA of the particle is measured. The wavefunction of the light pulse plays the role of a measurement pointer, which during the process of pre-measurement (PMD) is coupled to the degree of freedom that will be measured, i.e. polarization. This coupling creates entanglement between the pointer and the state of polarization. Indeed in the case of a strong measurement, $g-(t) g+(t) \approx 0$ and the global state is entangled. In the case of a weak measurement this entanglement is very weak, $g-(t) \approx g+(t)$, and the state is almost separable.

2.2 Post-selection and weak value

2.2.1 Post-selection on a pure state

To complete the analogy, PDL provides the necessary post-selection. At the end of the fiber we place a polarizer, which is simply an infinite PDL. This corresponds to doing a post-selection on a pure state $\phi = \mu |H\rangle + \nu |V\rangle$. So now we have to project the state $\Psi_{PMD}$ onto polarization state $|\phi\rangle$, which gives
\[ |\Psi_{OUT}\rangle = |\phi\rangle \langle \phi| \Psi_{PMD} \rangle = (\hat{\alpha} \tilde{\mu} g_-(t) + \tilde{\beta} \tilde{\nu} g_+(t))|\phi\rangle \equiv f(t)|\phi\rangle. \tag{2.3} \]

Clearly, \( f(t) \) is the post-selected component of the field. At the end we measure the intensity \( I = |f(t)|^2 = |A|^2 G_-(t) + |B|^2 G_+(t) + 2\text{Re}(\bar{A}B)g_-(t)g_+(t) \). The mean TOA is given by

\[
\langle t \rangle = \frac{\int_{-\infty}^{\infty} tI(t)dt}{\int_{-\infty}^{\infty} I(t)dt} = \ldots = \frac{\delta \tau}{2} \frac{|A|^2 - |B|^2}{|A|^2 + |B|^2 + 2\text{Re}(\bar{A}B) e^{-\frac{1}{2}(\delta \tau/t_c)^2}}. \tag{2.4} \]

An important feature of equation (2.4) is its dependance on the strength of the measurement (i.e. in \( \delta \tau/t_c \)). Now we continue our analysis, both in the case of strong and weak measurement.

In the limit of a strong measurement, \( t_c \ll \delta \tau \), the overlap \( g_+(t)g_-(t) \approx 0 \), so the detected intensities correspond to two well-separated gaussians, \( I(t) = |A|^2 G_-(t) + |B|^2 G_+(t) \). A detection in \( G_- \) corresponds to the fast eigenmode, i.e. the H polarization, so the probability that the polarization was H given the preparation and the post-selection is simply the integral of \( G_- \) normalized to the total intensity \( \sum_{K=H,V} \text{Prob}(\phi|K)\text{Prob}(K|\psi) \), which is known as the Aharonov-Bergman-Lebovitz (ABL) rule \([19]\), which is nothing but the classical rule for probability of sequential events. Of course, \( \text{Prob}(V) = 1 - \text{Prob}(H) \). One can also compute the mean time of arrival \( \langle t \rangle = \text{Prob}(H) \frac{\delta \tau}{2} + \text{Prob}(V)(-\frac{\delta \tau}{2}) \). Since the mean value of \( \sigma_z \) is simply \( \langle \sigma_z \rangle = \text{Prob}(H) - \text{Prob}(V) \), we have

\[
\langle t \rangle = \frac{\delta \tau}{2} \langle \sigma_z \rangle. \tag{2.6} \]

Equation (2.6) remains valid in the regime of a weak measurement. When \( \delta \tau \ll t_c \) equation (2.4) becomes

\[
\langle t \rangle = \frac{\delta \tau}{2} \frac{|A|^2 - |B|^2}{|A|^2 + |B|^2 + 2\text{Re}(\bar{A}B)} = \frac{\delta \tau}{2} \text{Re}\left(\frac{A - B}{A + B}\right). \tag{2.7} \]

Noticing that \( A + B = \hat{\alpha} \hat{\mu} + \tilde{\beta} \tilde{\nu} = \langle \phi|\psi \rangle \) and \( A - B = \hat{\alpha} \hat{\mu} - \tilde{\beta} \tilde{\nu} = \langle \phi|\sigma_z|\psi \rangle \). Using (2.6), we find

\[
\langle \sigma_z \rangle_w = \text{Re}\left(\frac{\langle \phi|\sigma_z|\psi \rangle}{\langle \phi|\psi \rangle}\right), \tag{2.8} \]
which is exactly the weak value of $\sigma_z$ when the post-selection is done on a pure state $|\phi\rangle$, according to Aharonov and Vaidman [10]. Note that $\langle \sigma_z \rangle_w$ can reach arbitrarily large values, leading to an apparently paradoxical situation since the eigenvalues of $\sigma_z$ are $\pm 1$. But there is no paradox at all since $\langle \sigma_z \rangle_w > 1$ simply means $\langle t \rangle > \frac{\delta \tau}{2}$. This situation is reached by post-selecting on a state $|\phi\rangle$ nearly orthogonal to $|\psi\rangle$. These are very rare events; the shape $F(t)$ of the pulse is strongly distorted, and it is not astonishing that its center of mass could be found far away from its expected position in the absence of post-selection. So we recover the phenomenon of anomalous dispersion with the quantum formalism.

Note that post-selection plays a crucial part in this effect. Without post-selection no interference between $g_-(t)$ and $g_+(t)$ could happen, since they correspond to orthogonal polarization modes. But post-selection projects these two modes onto a single polarization state, which then allows for interferences between $g_-(t)$ and $g_+(t)$.

### 2.2.2 Finite PDL and general networks

The quantum formalism of weak measurements also allows to describe post-selection on a mixed state, which in our analogy correspond to place a device with some finite PDL at the end of the birefringent fiber. Mathematically such device is described by a non-unitary operation of the form

$$F(\mu, \hat{n}) = e^{-\mu/2}e^{\mu\sigma_n/2}$$

(2.9)

where $\sigma_n = \hat{n} \cdot \hat{\sigma}$. The most and least attenuated states, respectively $| - \hat{n}\rangle$ and $| + \hat{n}\rangle$, are orthogonal. The attenuation between them, expressed in dB, is $10 \log_{10}(e^{2\mu})$. Mathematically speaking the PDL operator is not a projective measurement, but a POVM [11]. PDL acts as a sort of filter, for example in the unambiguous discrimination of non-orthogonal quantum states [20]. After some calculations one finds the following expression for the mean TOA in the case of a weak measurement of $\sigma_z$

$$\langle \sigma_z \rangle_w = \frac{\langle \sigma_z \rangle_\psi + \gamma n_z}{1 + \gamma \hat{n} \cdot \hat{\sigma}} = Re\left(\frac{\langle \psi | F^\dagger F \sigma_z | \psi \rangle}{\langle \psi | F^\dagger F | \psi \rangle}\right) = Re\left(\frac{\langle \psi | F^2 \sigma_z | \psi \rangle}{\langle \psi | F^2 | \psi \rangle}\right).$$

(2.10)

This is exactly the expression given by quantum theorists for the mean value of $\sigma_z$ when post-selection is done on the mixed state $\rho = \frac{1}{F(H)}F^2$. Note that $F[H]$ and $F[V]$ correspond to principal states of polarization (PSP) as defined in ref. [17].

In general an optical network is represented by an arbitrary concatenation of PMD and PDL elements. Our approach can be generalized to describe the combination of PMD and PDL in such networks. Here we will briefly discuss the case of a three trunk network PMD-PDL-PMD, represented by the operator $T = U(\delta \tau_2, \hat{m})F(\mu, \hat{n})U(\delta \tau_1, \hat{z})$. Following a similar calculation than above, one finds
\[ \langle t \rangle = \frac{\delta \tau_1}{2} \text{Re} \left( \frac{\langle \psi | F^2 \sigma_z | \psi \rangle}{\langle \psi | F^2 | \psi \rangle} \right) + \frac{\delta \tau_2}{2} \frac{\langle \Psi_F | \sigma_m | \Psi_F \rangle}{\langle \Psi_F | \Psi_F \rangle} , \]

where \( | \Psi_F \rangle = F|\psi\rangle/\sqrt{\langle F | F \rangle_{\psi_0}} \) is the filtered state. This result is intuitively clear: the first term is the weak value obtained by forgetting about the second PMD element; the second term is just the mean value of \( \sigma_m \) on the filtered state \( | \Psi_F \rangle \) by forgetting the first PMD element.

Indeed formula (2.11) generalizes to arbitrary networks. This example also shows that the formalism of weak measurements considerably simplifies the calculation of the combined effect of PMD and PDL by adding an intuitive meaning to the formula of the mean TOA.

### 2.3 Conclusion

We have pointed out a strong analogy between the abstract theory of weak measurements with post-selection and polarization effects in optical fiber telecom networks. Furthermore, the quantum formalism of weak measurements was shown to describe, in a much simpler way than in standard optics, the combined effect of PMD and PDL. On the other hand, weak measurements with post-selection, often thought as weirdness for theorists, appear to be useful in describing important polarization effects in telecom networks. From a more general point of view this work might open the way for some fruitful collaborations between quantum theorists and telecom ingeneers, two group of physicists which have almost no connections.

Finally it is worth mentioning that this work closes a loop of analogies. On the one hand, in Ref. [21], Gisin and Go stressed the analogy between PMD-PDL systems and the mixing and decay of kaons. On the other hand, Ref [22] pointed out that adiabatic measurements in metastable systems are a kind of weak measurement; an example of these being kaons. So, by relating weak measurements and PMD-PDL networks, we close the loop.
Chapter 3

Superluminal group velocity in an optical fiber

In his famous theory of special relativity, Albert Einstein stated that there exists in Nature a limit on the speed at which any object can travel, namely the speed of light. A century later, many experimental papers have been published invoking "superluminal velocities" or "faster than light speeds" [23]. Nevertheless all of these experiments are in perfect agreement with Einstein’s causality. This apparent contradiction is resolved once the concept of group velocity is clearly defined. The group velocity is the speed at which the center of mass of an object travels. Remarkably, neither the energy nor the information travel, in general, at the group velocity, thus there is no paradox at all observing faster than light group velocities. Since the seminal work of Sommerfeld, extended and completed by Brillouin [24], it is known that information travels at the signal velocity, defined as the speed of the front of a square pulse. This velocity cannot exceed $c$. The fact that no modification of the group velocity can increase the speed at which information is transmitted has been directly demonstrated in a recent experiment [25].

In this chapter we present a slow and fast-light experiment [B] based on the results of the previous chapter, namely the description of PMD-PDL systems in terms of weak measurements with post-selection [A]. Direct measurements of superluminal group velocities and signal velocities in an optical fiber are performed. The organization of this chapter is as follows. After a brief review of the concepts of slow and fast-light in section 3.1, we give in section 3.2 a description of our experiment in the light of the results of chapter 2. Section 3.3 presents the experimental setup and the results. Some final comments are given in section 3.4.

3.1 Tailoring the speed of light

When a light pulse, sharply peaked in the frequency domain, travels through a medium with dispersion $\omega(k)$, the speed of its center of mass is the group velocity $v_g$ of the medium for the central frequency [26]. It is usual to define a group index of refraction $n_g = \frac{c}{v_g}$ which can be derived from the standard index of refraction $n(\omega)$
\[ n_g(\omega) = n(\omega) + \omega \frac{dn}{d\omega}(\omega) \ . \tag{3.1} \]

Equation (3.1) shows that the group index can become smaller than one, leading to a group velocity larger than \( c \), when \( \frac{dn}{d\omega} \) is negative, i.e. when the refractive index decreases rapidly with frequency. Such superluminal group velocities have been observed in numerous experiments conducted on atomic vapours and particle tunnelling. In fact the group velocity can even become negative when \( \omega \frac{dn}{d\omega}(\omega) < -n(\omega) \) [27]. In this case the pulse is reshaped in the medium, such that the center of mass of the output pulse exits the medium before the center of mass of the input pulse has entered it. This has been experimentally demonstrated [28].

On the other hand when \( \frac{dn}{d\omega} \) becomes very large, the group velocity is considerably reduced. This is slowlight. Recent experiments have reported speeds of order of a few meters per second [29]. It has even been possible to store the light in different kind of mediums for some microseconds [30, 31]. Most slowlight experiments rely on a technique called electromagnetically induced transparency (EIT) [32, 33], which requires complex technology. In principle EIT can be achieved without any losses, contrary to our approach.

### 3.2 Slow and fast-light in an optical fiber

![Figure 3.1: A birefringent fiber is sandwiched between two polarizers. The scenario of weak measurements is recovered: (1) pre-selection, (2) weak measurement, (3) post-selection.](image)

Our experiment consists of a birefringent fiber sandwiched between two polarizers. The first polarizer controls the pre-selected polarization state and the second one the post-selected state. As shown in the previous chapter the mean TOA of a light pulse sent through this system is given by

\[ \langle t \rangle = \frac{\delta \tau}{2} \langle \sigma_z \rangle_w = \frac{\delta \tau}{2} \frac{\langle \phi \vert \sigma_z \vert \psi \rangle}{\langle \phi \vert \psi \rangle} \ . \tag{3.2} \]

The refractive index of the fiber is \( n \approx \frac{3}{2} \) for the telecom frequencies we are using here (\( \lambda = 1550\text{nm} \)). When both polarizers are aligned, i.e. \( \vert \phi \rangle = \vert \psi \rangle \), the pulse propagates at speed \( v_g = \frac{c}{n} = \frac{L}{t_f} \) where \( t_f \) is the free propagation time. However
when the polarizers are not aligned anymore, the center of mass of the pulse exits the medium after a time \( t = t_f + \langle t \rangle \), with \( \langle t \rangle \) the mean TOA given in equation (3.2). If the deformation of the pulse is weak, the group velocity is still the speed of the center of mass, now given by

\[
v_g = \frac{L}{t_f + \langle t \rangle}.
\]

Thus the group velocity can be tuned from very large, or even negative, \( (\langle t \rangle \to -\infty) \) to very small \( (\langle t \rangle \to \infty) \) — although in these limiting situations the pulse is usually strongly distorted, so that our reasoning breaks down. Note also that the response function \( G(\omega) \) of our setup is derived in Ref. [B]. Its connection with weak values is clearly pointed out.

Previous slow and fast light experiments exploited the dispersive properties of some well-chosen media. Therefore most of these experiments were designed either for slow light, or for fast light, but not both in the same experiment. From this point of view, our experiment is completely different. Here we create an artificial dispersion in the optical fiber with the help of the combined effect of PMD and PDL. So we can tune the group velocity from slow to fast light by merely rotating a polarizer. Thus our experiment requires neither complex equipment, nor cryogenic temperatures. On the contrary it is tabletop and uses only standard telecom devices.

### 3.3 Experimental setup and results

The experimental setup is sketched out in fig. 3.2. As a source and for detection we use an optical time domain reflectometer (OTDR). This is a telecom instrument designed to measure loss profiles of fibers: it sends short laser pulses and analyzes the amount of back-scattered light as a function of time. Here we use a commercial prototype OTDR working in photon counting mode at the telecom wavelength \( \lambda = 1.55 \mu \text{m} \) [34]. It contains a pulsed DFB laser \( (\Delta \nu \approx 500\text{MHz} \), pulse duration 2 ns) and a gated peltier-cooled InGaAs photon counter (gate duration 2 ns). The photon counting OTDR is well adapted to this experiment since it allows to monitor the optical pulse, even in the presence of strong absorption. The birefringent fiber is a polarization maintaining (PM) fiber of length \( L = 1.5 \text{ m} \), with a DGD \( \delta \tau = 2.66 \text{ ns} \).
ps, measured with an interferometric low-coherence method [35]. The fiber is placed between two polarizing beam splitter (PBS) cubes with specified extinction ratios of 50 dB. Both cubes are mounted on rotational stages, allowing a precise alignment of pre- and post-selected polarization states. Furthermore a micrometer step motor permits to slightly change the length of the fiber. This allows alignment of the post-selection also in the $x - y$ plane of the Poincaré sphere, since polarization is rotating around the $z$ axis in the fiber.

The crucial part of the experiment is the alignment of both PBS cubes. The input polarizer has to be aligned at 45° relative to the fiber’s axis. This is done by injecting incoherent light (from a LED) into the system and minimizing the degree of polarization at the output of the PM fiber. For the post-selection, the position of the second PBS cube and the length of the fiber are adjusted such that transmission through the system is minimum.

![Figure 3.3](image)

Figure 3.3: (a) Successive measurements with slight change of the post-selection. (b) Fast and slow light measurements. The vertical line has travelled at the speed of light. The fast pulse shows superluminal group velocity.

Fig. 3.3(a) shows the data of a sequence of successive measurements on a dB scale. The largest curve is the reference pulse, i.e. the pulse in the normal regime, without any specific alignment of the PBS cubes. Between each measurement the post-selection is slightly changed, in order to decrease transmission: we observe that the lower the transmission, the higher the group velocity, in agreement with the theory. In addition, Fig. 3.3(a) clearly illustrates the difference between group and signal velocities [24]. In fact, even though the group velocity is higher for each successive curve, the signal velocity remains constant and equal to $c/n_f$, since the front parts of all pulses are strictly identical. To reach superluminal group velocities, much longer input pulses are needed. We replace the OTDR source by an external source: a DFB laser ($\Delta\nu \simeq 2$ MHz) working in CW mode, modulated by an external electro optic modulator (EOM). This source creates nearly gaussian pulses with a coherence time of about 50ns. The OTDR is still used for detection and triggers the modulator. Results are presented in Fig. 3.3(b). The pulse on the right shows clearly superluminal group velocity. By fitting the position of the maximum of the output pulse, we find $W \simeq -3500$, consistent with superluminal but non-negative group velocity, as expected.
3.4 conclusion

We have demonstrated a simple method for creating slow and fast light in a telecom fiber. This experiment was designed in the light of the connection established previously between weak measurements and polarization effects in optical networks. The results included direct measurements of superluminal group velocities and of signal velocities. Group and signal velocities were clearly shown to be different. Furthermore the signal velocity remained equal to $\sim \frac{c}{n}$ for any value of the group velocity, in particular for superluminal group velocities.
Chapter 4

Three-Bell-State Measurement

The Bell state measurement (BSM) is a key ingredient of many quantum communication and information processing protocols, such as quantum teleportation [12, 36, 37, 38], entanglement swapping [39, 40], dense coding [13, 41] and fault tolerant quantum computing. BSM have been successfully achieved on continuous variable systems [42, 43] and using nonlinear optics [44]. For qubits though, it has been shown that, using linear optics only, a perfect BSM cannot be implemented. Two cases should be distinguished: if one is allowed to use a large number of auxiliary photons, the probability of success of the BSM can be made arbitrarily close to one. This is the famous result of KLM [45]. However, if all of those auxiliary modes are in the vacuum state, the maximal probability of success of the BSM is limited to 1/2 [15]. Note that this proof only bounds the probability of success of the BSM, but does not limit the number of Bell states that can be discriminated. Up to date, all practical implementations of qubits BSM with linear optics detect perfectly two out of four equiprobable Bell states and are thus optimal.

Here we present a novel BSM which discriminates three out of the four Bell states and still achieves the maximal overall success rate of 50% [C,D]. After a brief introduction on the Bell measurement in section 4.1, we present the three-Bell-state measurement, first in a time-bins implementation, second in a polarization implementation. Section 4.3 reviews a teleportation experiment, which successfully implemented the new BSM. Section 4.4 gives a short a conclusion.

4.1 Bell measurement

The Bell states are the four maximally entangled states

$$\psi^\pm = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \quad \phi^\pm = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle).$$

(4.1)

They form an orthogonal basis for quantum states of two qubits. A BSM is an operation that discriminates between the four Bell states. Each state is mapped to a list of detection patterns. Whenever a pattern is associated to one of the Bell states only, the BSM is successful. The standard way of performing a BSM with photonics qubits is using a 50/50 beam-splitter (BS). Each photon of the Bell state enters in
a different port of the BS. Then the Hong-Ou-Mandel bunching effect [46] provides
the necessary two-photon interference and discriminates between $|\psi_+\rangle$, $|\psi_-\rangle$ and
$|\phi_\pm\rangle$. In the two first cases the BSM is conclusive and thus the overall probability
of success is $1/2$, reaching the bound of [15]. This setup has been used in many
experiments, for polarization encoding as well as time-bins qubits.

4.2 Implementations of the three-Bell-state measurement

4.2.1 Time-bins implementation

In this section we describe the new BSM in a time-bins implementation. A review
about time-bins qubits encoding is beyond the scope of this thesis, but can be found
in [47].

To get the new BSM, we replace the beam-splitter from the standard BSM, by
an unbalanced Mach-Zehnder interferometer (Fig. 4.1). The path length difference
corresponds to one time-bin. For a single incoming photon in mode $a$ or $b$ the output
state reads

$$a_j^\dagger \rightarrow \frac{1}{\sqrt{2}}(c_j + id_{j+1}) \rightarrow \frac{1}{2}(e_j^\dagger + if_j^\dagger + i f_{j+1}^\dagger - e_{j+1}^\dagger)\quad (4.2)$$

$$b_j^\dagger \rightarrow \frac{1}{\sqrt{2}}(d_{j+1} + ic_j^\dagger) \rightarrow \frac{1}{2}(f_j^\dagger + ie_j^\dagger + i e_{j+1}^\dagger - f_{j+1}^\dagger)\quad (4.3)$$

where $x_j^\dagger$ represents a photon in mode $x$ in time-bin $j$ ($j = 0, 1$) and the additional
phase $\delta$ has been set to 0.

Using the same notation the four incoming Bell states are

$$\phi_\pm = \frac{1}{\sqrt{2}}(a_0^\dagger b_0^\dagger \pm a_1^\dagger b_1^\dagger),$$

$$\psi_\pm = \frac{1}{\sqrt{2}}(a_0^\dagger b_1^\dagger \pm a_1^\dagger b_0^\dagger).$$

Thus inserting (4.2) in these expressions one can easily compute the detection patterns for each Bell state, which are given in table (4.1).

Because of the new interferometer there are now three different times of arrival
(0,1,2) at the two detectors. Overall there are 21 different detection patterns, all
appearing, at least once, in table (4.1). Those patterns appearing only once allow

Figure 4.1: Time-bins implementation of the three-Bell-state measurement.
| \phi_+ \rangle | D_1 \rangle | D_2 \rangle | P \\ 
00 & 22 & 0 & 2 & 1 \\
00 & 22 & 0 & 2 & 1 \\
1/16 & 1/16 & 1/16 & 1/8 & 1/8 & 1/2 \\

| \phi_- \rangle | D_1 \rangle | D_2 \rangle | P \\ 
00 & 22 & 0 & 2 & 11 \\
00 & 22 & 0 & 2 & 11 \\
1/16 & 1/16 & 1/16 & 1/4 & 1/4 & 1/8 & 1/8 \\

| \psi_+ \rangle | D_1 \rangle | D_2 \rangle | P \\ 
01 & 12 & 1 & 0 & 2 & 1 \\
01 & 12 & 0 & 1 & 1 & 2 \\
1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\

| \psi_- \rangle | D_1 \rangle | D_2 \rangle | P \\ 
11 & 20 & 2 & 0 \\
11 & 20 & 0 & 2 \\
1/4 & 1/4 & 1/8 & 1/8 & 1/8 & 1/8 \\

Table 4.1: Detection patterns for the four Bell states. A "0" on line \( D_1 \) means that a photon has been detected in detector \( D_1 \) in time-bin "0". The probability of each pattern is shown on line \( P \). Patterns in bold character appear only once in the table and thus identify a Bell state. The other patterns are associated to two Bell states. In those cases the BSM is inconclusive.

to discriminate a Bell state, while the others lead to an inconclusive result. The \( |\phi^+\rangle \) and \( |\psi^-\rangle \) states are detected with probability one half, and the \( |\psi^+\rangle \) state with probability one. This leads to an overall success rate of 50% and is thus optimal according to [15]. However in a more realistic experimental environment, one has to include the dead time of the detectors. This imposes to consider only patterns where one photon is detected in each detector. In this case the probabilities of discriminating \( |\phi^+\rangle \), \( |\psi^-\rangle \) and \( |\psi^+\rangle \) are reduced to \( 1/2 \), \( 1/4 \) and \( 1/2 \) respectively. This leads to an overall probability of success of \( 5/16 \) for the BSM. Note that this is still larger than the usual success rate of 25% for a simple BS used with two current-day detectors.

4.2.2 Polarization implementation

From a mathematical point of view, the three-Bell-state measurement is a 21-outcomes POVM [11]. It can basically be implemented with any type of encoding. In this section we present the polarization implementation. The setup is sketched in Fig. 4.2. The main difference with the time-bins implementation is the number of detectors. Here, four detectors are needed, while the time-bins setup required only two of them. However in this case, detectors dead time do not penalize the measurement because there will never be detection patterns where two photons end
up in the same detector. Thus a polarization implementation of this three-Bell-state measurement would lead to 50% efficiency with today’s technology.

Figure 4.2: Polarization implementation of the three-Bell-state measurement. The shaded cubes represent beam-splitters and the white cubes are polarizing-beam-splitters (PBS).

4.3 Teleportation experiment

In order to demonstrate a successful three-Bell-state measurement, J. van Houwelingen and A. Beveratos performed a teleportation experiment [C,D]. They used the time-bins implementation. A schematic of the experimental setup is shown in Fig. 4.3. Alice prepares the photon to be teleported in the state $|\psi_A\rangle = |0\rangle + e^{i\alpha}|1\rangle$. The entangled photon pair is prepared in the $|\phi^+\rangle$ Bell state. Then Charlie performs the BSM. Finally, to verify teleportation indeed occurred, Bob measures interference fringes whenever the BSM was successful.

Figure 4.3: Experimental setup. The novel part is Charlie’s interferometer, which achieves the three-Bell-state measurement.

Results are presented in Fig. 4.4. They obtained an overall fidelity of 67%, which after noise substraction gave a fidelity of 76%. A nice feature of the experiment is that, for the BSM, projections on the three distinguishable Bell state were observed.
Depending on which Bell state came out at the BSM, the interference fringes measured on Bob’s side are shifted with the correct phase. Here this phase is 0 for projections on $|\phi^+\rangle$ and $|\psi^-\rangle$ and $\pi$ for projection on $|\psi^+\rangle$ as expected.

![Interference fringes obtained by scanning Bob’s interferometer. As expected there is $\pi$ phase between $|\psi^+\rangle$, and $\{|\phi^+\rangle, |\psi^-\rangle\}$](image)

**4.4 Conclusion**

In this chapter we have presented a novel BSM, which allows one to distinguish three out of four Bell states, while still reaching the maximal overall efficiency of 50% for linear optics. A recent teleportation experiment included this three-Bell-state measurement.

From a practical point of view, this new BSM is unfortunately not very interesting. Since it needs one additional interferometer, compared to the standard BSM (which requires only a simple beam-splitter), the new BSM introduces excess losses (3 dB in the experiment). Indeed these losses reduce the count rate, which is an important drawback for practical implementations.

Nevertheless this new BSM remains interesting from a fundamental point of view, since it illustrates the power of generalized measurements, so-called POVM’s.
Chapter 5

Bell-type inequalities for non-local resources

One of the most striking properties of quantum mechanics is non-locality. It is well known that two separated observers, each holding half of an entangled quantum state and performing appropriate measurements, share correlations which are non-local, in the sense that they violate a Bell inequality \[2\]. A key feature of entanglement is that it does not allow the two distant observers to send information to each other faster than light, i.e. correlation from measurements on quantum states are no-signaling.

A frustrating aspect of quantum mechanics is that it does not provide the slightest hint about the way it achieves this amazing non-locality! It is therefore interesting trying to reproduce, or simulate, the quantum correlations by other means. In order to do this one has first to choose some resource, which has to contain some non-locality, but can be described without invoking quantum theory. This approach also allows to quantify how powerful the non-local correlations of quantum mechanics are, which is a central issue in quantum information science.

A quite natural choice for such a non-local resource is indeed classical communication. Constructing no-signaling correlations with classical communication is quite tricky. One has to cleverly hide the communication to avoid signaling. Several works \[48, 49, 50\] underwent the task of estimating the amount of communication required to simulate the maximally entangled state of two qubits. These partial results were superseded in 2003, when Toner and Bacon \[51\] proved that the singlet can be simulated exactly using local variables plus one bit of communication per pair.

Recently another non-local resource has been proposed to study this problem: the PR-box. In 1994, Popescu and Rohrlich \[7, 8\] found out that there exist no-signaling correlations which are more non-local than those of quantum physics. In particular they pointed out a set of correlations that violate the Clauser-Horne-Shimony-Holt (CHSH) \[52\] Bell inequality up to its algebraic maximum; these correlations are now commonly known as the PR-box. Indeed the PR-box appears as a very natural resource to study the problem of simulating quantum correlations: it is intrinsically no-signaling and also more powerful than QM. The PR-box was then proven to be a
powerful resource for information theoretic tasks, such as communication complexity [53] and cryptography [54, 55]. It was also recently suggested that the PR-box is a unit of non-locality [56]. Note that a PR-box is a strictly weaker resource than a bit of communication [E]. Recently, Cerf et al. [57] presented a model using a single PR-box which simulates correlations from any projective measurement on the singlet. Degorre et al. [58] presented an elegant reformulation of both the Toner-Bacon model and the PR-box model of Cerf et al. as a distributed sampling problem.

It appears very natural to extend this study to other quantum states. However, this turns out to be very difficult, even for pure non-maximally entangled states of two qubits. In this chapter we show that some non-maximally entangled states can in fact not be simulated with a single use of a PR-box [E,F]. In other words, these states require a strictly larger amount of non-local resources than the maximally entangled state, in order to be simulated. This suggests that entanglement and non-locality are different resources. To demonstrate this result we found a Bell-type inequality which cannot be violated by any strategy which uses at most one PR-box. Then it was proven that this inequality is violated by some non-maximally entangled state.

This chapter is organized as follows. Section 5.1 reviews some useful concepts about Bell inequalities. The main result is presented in section 5.2. Section 5.3 studies other non-local resources. Finally an outlook is given in section 5.4.

5.1 Bell inequalities, polytopes, and non-local machines

In this section we review some useful tools. Let us consider a typical Bell test scenario. Two distant observers, Alice and Bob, share some quantum state. Each of them chooses between a set of measurements (settings) \( \{A_i\}_{i=1..N_A} \) and \( \{B_j\}_{j=1..N_B} \). The result of the measurement is noted \( r_A, r_B \). Here we will focus on dichotomic observables and we will restrict Alice and Bob to use the same number of settings, i.e. \( r_{A,B} \in \{0,1\} \) and \( N_A = N_B \equiv N \). An "experiment" is fully characterized by the family of \( 4N^2 \) probabilities \( P(r_A, r_B|A_i, B_j) \equiv P_{ij}(r_A, r_B) \) and can be seen as a point in a \( 4N^2 \)-dimensional probability space. As probabilities must satisfy

1. Positivity: \( P_{ij}(r_A, r_B) \geq 0 \quad \forall i, j, r_A, r_B \)
2. Normalization: \( \sum_{r_A, r_B=0,1} P_{ij}(r_A, r_B) = 1 \quad \forall i, j \)

all relevant experiments are contained in a bounded region of this probability space. Since we want to restrict to no-signaling probability distributions, we impose also the no-signaling conditions

\[
\sum_{r_A=0,1} P_{ij}(r_A, r_B) = P_j(r_B) \quad \forall i \\
\sum_{r_B=0,1} P_{ij}(r_A, r_B) = P_i(r_A) \quad \forall j .
\]  

(5.1)
Conditions (5.1) mean that Alice’s output cannot depend on Bob’s setting, and *vice versa*. This shrinks further the region of possible experiments, and the dimension of the probability space is now reduced to \( d = N(N + 2) \). So each no-signaling experiment is represented by a point in a \( d \)-dimensional probability space. In fact the region containing all relevant probability distributions (strategies), i.e. satisfying positivity, normalization and no-signaling, forms a polytope, i.e. a convex set with a finite number of vertices. This is the no-signaling polytope (NSP).

One can restrict even further the probability distributions, by requiring that these are built only by local means, such as shared randomness. We then obtain a smaller polytope: the local polytope (LP). The facets of this polytope are Bell inequalities, in the sense that a probability distribution lying inside (outside) LP, satisfies (violates) a Bell inequality. The vertices (extremal points) of LP are deterministic strategies obtained by setting the outputs \( r_A \) and \( r_B \) always to 0 or always to 1. Finding the facets of some polytope knowing only its vertices is a computationally difficult task, as shown by Pitowsky [59]. That’s why all Bell inequalities have been listed for the case of two or three settings, whereas not much is known for a larger number of settings.

In the simple case of two settings on each side, the situation is well known [60]. Fine [61] showed that all non-trivial facets of the LP are equivalent to the famous CHSH inequality

\[
\text{CHSH} \equiv \begin{array}{ccc}
-1 & 0 \\
1 & 1 \\
0 & 1 & -1
\end{array} \leq 0.
\] (5.2)

Here the notation represents the coefficients that are put in front of the probabilities, according to

\[
\frac{P_j(r_B = 0)}{P_i(r_A = 0)} = P_{ij}(r_A = r_B = 0).
\] (5.3)

Inequality (5.2) can indeed be violated by quantum mechanics, and the maximal violation is \( 1/\sqrt{2} - 1/2 \approx 0.2071 \), obtained by suitable measurements on the singlet state. Of course the quantum set is included in the NSP, but the converse is not true. There are no-signaling correlations that are more non-local than those of quantum mechanics. Among these figures indeed the PR-box [7].

The PR-box (see Fig. 5.1) is a two-input, two-output NLM. Alice inputs \( x \) into the machine and gets outcome \( a \), while Bob inputs \( y \) and gets output \( b \). The outcomes are correlated such that \( a \oplus b = xy \). The local marginals are however completely random, i.e. \( P(a = 0) = P(b = 0) = \frac{1}{2} \), which ensures no-signaling. All non-local vertices of the NSP are represented by some PR-box [60].
5.2 Entanglement and non-locality are different resources

In this section we study Bell-type inequalities which allow the partners to use one PR-box. This means that all strategies satisfying such inequality can be reproduced by local means (i.e. shared randomness, etc) together with a single use of a PR-box. Thus any strategy violating the inequality needs a more powerful non-local resource in order to be simulated. In the case of two settings, described in the previous section, such inequalities cannot exist, because the most elementary non-local resource, the PR-box, suffices already to generate the whole NSP.

Therefore we have to consider more settings on each side. We start with the case of three settings. Then we briefly present results when more settings are added.

5.2.1 Three settings on each side

For this case, all facets of the LP have been listed [62]. While there are still some facets of the CHSH form, there is only one new inequality in this case, which is

$$I_{3322} = \begin{vmatrix} -1 & 0 & 0 \\ -2 & 1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{vmatrix} \leq 0$$ (5.4)

Quantum mechanics indeed violates inequality $I_{3322}$ up to $\frac{1}{4}$ for the maximally-entangled state. Furthermore this inequality is relevant compared to CHSH, in the sense that it is violated by some quantum states which do not violate the CHSH inequality [62].

The no-signaling strategy that achieves the larger violation of $I_{3322}$ is described by the relation $[xy/2] = a + b(\mod 2)$, where $x, y \in \{0, 1, 2\}$ are the inputs and $a, b \in \{0, 1\}$ the outputs. This machine, which will be referred to as $PR_3$, is more powerful than a PR-box; in fact it needs two PR-boxes to be simulated. This means that there are no-signaling correlations which cannot be reproduced with a single PR-box. Now we go on showing that some of these correlations can be achieved by appropriate measurements on a family of non-maximally entangled states of two
Figure 5.2: Sketch of the no-signaling polytope above the $I_{3322}$ facet. The new inequality $M_{3322}$ is a facet of the 1-PR-box polytope (see text). Remarkably a family of partially entangled states violate this inequality, and thus cannot be simulated with a single PR-box.

qubits. In order to do this, we present an inequality, which cannot be violated if Alice and Bob share at most one PR-box.

Let us first gain some geometrical intuition! Recently Jones and Masanes [63] characterized all extremal points of the NSP in the case of three settings on each side. They showed that there are two type of these extremal points: first those who can be obtained with a PR-box and, second, those who require a stronger non-local resource, namely the $PR_3$ box. Thus there is a new polytope, sandwiched between the local and the no-signaling polytopes. It is formed by all strategies that can be simulated using at most one PR-box (see Fig. 5.5). By characterizing all the extremal points of this 1-PR-box polytope, we found one of its facets, which corresponds to the inequality

$$M_{3322} \equiv \begin{array}{ccc} -2 & 0 & 0 \\ -2 & 1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{array} \leq 0 \quad (5.5)$$

Remarkably $M_{3322}$ is violated by a family of non-maximally entangled states of two qubits of the form $|\psi(\alpha)\rangle = \cos \alpha |00\rangle + \sin \alpha |11\rangle$ with $0 \leq \alpha \leq \frac{2}{15} \pi$. Indeed the maximally entangled state does not violate this inequality, since its correlations can be simulated using a single PR-box [57]. Note that the structure of $M_{3322}$ is similar to $I_{3322}$, the only difference being the coefficient of Alice’s first marginal.

Indeed $M_{3322}$ can be violated with a $PR_3$ box, thus also with two PR-boxes. Surprisingly it can also be violated with a single bit of communication; an example of such a strategy was found by S. Pironio and can be found in [E]. This proves that, even considering only no-signaling strategies, a PR-box is a strictly weaker resource than a single bit of communication.
5.2.2 Extension to more settings

The most interesting case is when Alice chooses among four settings and Bob among three settings. Here all Bell inequalities are also known [62], among which

\[
I_{4322}^{(2)} = \begin{vmatrix}
-2 & -1 & 0 & 0 \\
1 & 1 & 1 & 0 \\
-1 & 1 & 0 & -1 \\
0 & 1 & -1 & 0 \\
\end{vmatrix} \leq 0. \quad (5.6)
\]

Above \(I_{4322}^{(2)}\) we found a single new inequality with one PR-box, given by

\[
M_{4322} = \begin{vmatrix}
-2 & -1 & 0 & 0 \\
1 & 1 & 1 & 0 \\
-1 & 1 & 0 & -1 \\
0 & 1 & -1 & 0 \\
\end{vmatrix} \leq 0. \quad (5.7)
\]

Quite similarly to the case of three settings on each side, the difference between the original inequality and the new one is just a larger penalty on one marginal.

Again the interest of \(M_{4322}\) comes from the quantum violation, which is larger than the violation of \(M_{3322}\), thus allowing to extend the range of states for which one PR-box is not enough. Specifically, one finds that \(M_{4322}\) is violated for \(0 < \alpha \lesssim \frac{\pi}{78}\).

Beyond the 3322 and 4322 cases, the facets of the deterministic polytope have not been listed exhaustively, but several examples of facets are available [62, 64]. On these, we searched for possible extensions of our results by increasing the penalties on the marginals. 4422 facets, as well as some 5522 and 6622 ones, did not appear to be worth a closer study after our survey. We have considered neither inequalities with larger numbers of outcomes, nor multi-partite scenarios.

5.3 Bell-type inequalities for stronger non-local resources

Next we extended the work of the previous section to more powerful non-local resources [F]. First, we introduced N-inputs NLM, which appear as a natural extension of the two-inputs PR-box. These machines, denoted \(PR_N\), have a nice connection to a family of N-settings Bell inequalities known as \(I_{NN22}\) [62], similar to the one that relates the PR-box to the CHSH inequality (see section 5.1).

In fact, the structure of the N-inputs NLM can be directly deduced from the corresponding \(I_{NN22}\) inequality. Then we derived a family of N-settings inequalities, \(M_{NN22}\), which allow one use of \(PR_{N-1}\) machine. Again, the structure of these new inequalities is easily deduced from the corresponding inequality \(I_{NN22}\).

Thus a nice construction appears: for any number of settings \(N\), we have a Bell inequality \(I_{NN22}\) and the related NLM, \(PR_N\), which reaches the upper (no-signaling) bound of the inequality. Adding one setting we find another inequality,
\(M_{(N+1)(N+1)22}\), that cannot be violated by strategies which use at most one \(PR_N\) and the next NLM \(PR_{N+1}\), i.e.

\[
(I_{NN22}, PR_N) \rightarrow (M_{(N+1)(N+1)22}, PR_{N+1})
\]  \hspace{1cm} (5.8)

Unfortunately, for \(N \geq 4\), inequalities \(M_{NN22}\) are not violated by non-maximally entangled state of two qubits.

### 5.4 Outlook

The main result of this chapter is that some non-maximally entangled states of two qubits require a strictly larger amount of non-local resources than the maximally entangled state in order to be simulated. This result strongly suggests that entanglement and non-locality are different resources. This intuition also appears in some previous results. For example it is known that some Bell inequalities are maximally violated by states that are partially entangled [65]. Another example comes from the detection loophole, where non-maximally entangled states can tolerate smaller detector efficiencies than the maximally-entangled one [66]. We will come back to this in the next chapter.

Finally we point out some interesting open questions on the simulation of quantum correlations. In the light of our result, a natural question is whether two PR-boxes are enough to reproduce the correlations of all two-qubits states. Two different approaches may bring an answer. First try to construct explicitly a model that gives the correct correlations. Second find an inequality valid for any strategy with two PR-boxes, that would be violated by some non-maximally entangled states. Both approaches seem difficult. Another problem is whether one bit of communication is enough to simulate any entangled state of two-qubit state. Generalizing the Toner-Bacon model [51] to non-maximally entangled states turns out to be very difficult. The best model known to date needs one bit of communication and a PR-box to simulate non-maximally entangled states [67]. Whether or not this bit of communication can be removed is an interesting open problem.
Chapter 6

Detection loophole in asymmetric Bell experiments

In 1964 John Bell came out with his famous Bell inequality [2]. It is remarkable that all experiments realized since Bell’s breakthrough have confirmed quantum non-locality [3, 4, 68, 69, 70, 71]. However, from a strictly logical point of view, none of these experiments definitely rules out the possibility of some local realistic model; some loopholes exist in all these experiments that make it is still possible to reproduce their results with a local hidden variable (lhv) model. To summarize, there is strong experimental evidence that Nature is non-local, but, considering the importance of such a statement, it is crucial to perform an experiment free of any loopholes. Another motivation comes from quantum information science, where the security of some quantum communication protocols is based on the loophole-free violation of Bell inequalities [54, 55].

Performing a loophole-free Bell test is quite challenging. One first has to ensure that no signal can be transmitted from one particle to the other during the measurement process. Thus the measurement choice on one side and the measurement result on the other side should be space-like separated. If this is not the case, one particle could send some information about the measurement setting it experiences to the other particle. This is the locality loophole [72]. Secondly, the particles must be detected with a high enough probability. If the detection efficiency is too low, a lhv model can reproduce the quantum correlations. In this picture a hidden variable affects the probability that the particle is detected depending on the measurement setting chosen by the observer. This is the detection loophole [73].

In practice, photon experiments have been able to close the locality loophole [4, 68, 69]. However the optical detection efficiencies are still too low to close the detection loophole. For the CHSH [52] inequality, an efficiency larger than 82.8% is required to close the detection loophole with maximally entangled states. Surprisingly, Eberhard [66] showed that this threshold efficiency can be lowered to 66.7% by using non-maximally entangled states. Threshold efficiencies for other Bell inequalities have also been studied [74]. On the other hand an experiment carried out on trapped ions [71] closed the detection loophole, but the ions were only a few micrometers apart which leaves the locality wide open. It would already be a
significant step forward to close the detection loophole for well separated systems.

In this chapter we focus on asymmetric Bell tests [G], where the two particles are detected with different probabilities. This is the case e.g. in an atom-photon system: the atom is measured with an efficiency close to one while the probability to detect the photon is smaller. Intuition suggests that if one party can do very efficient measurements, then the minimal detection efficiency on the other side should be considerably lowered compared to the case where both detectors have the same efficiency. Experimentally this approach might be quite promising, since recent experiments have demonstrated atom-photon entanglement [75, 76] and violation of the CHSH inequality [77] with such systems. In section 6.1 we present the general approach to the study of the detection loophole in asymmetric systems. We focus on the case where one of the systems is detected with efficiency \( \eta_A = 1 \) and we compute the threshold efficiency \( \eta_B^{th} \) for the detection of the other system. Section 6.2 reviews the results, namely the threshold efficiencies as well as the influence background noise and noisy detectors. Finally a discussion about the experimental feasibility and a brief outlook are presented in section 6.3.

### 6.1 Detection loophole in asymmetric Bell experiments

We consider a standard Bell experiment scenario. Alice and Bob test a Bell inequality \( I \leq L \) on a state \( \rho_{AB} \) having two different detection efficiencies, \( \eta_A \) and \( \eta_B \). Since they must always produce an outcome, they have to decide on a strategy in case of non-detection; they could for example agree to output “0” when no particle is detected.

So, four different kind of events can occur. They are listed below. For each event, we give the corresponding value of the Bell inequality.

1. Both Alice’s and Bob’s detectors fire \( \rightarrow I^{(d,d)} \equiv Q \)
2. Only Alice’s detector fires \( \rightarrow I^{(d,\emptyset)} \equiv M_A \)
3. Only Bob’s detector fires \( \rightarrow I^{(\emptyset,d)} \equiv M_B \)
4. No detector fires \( \rightarrow I^{(\emptyset,\emptyset)} \equiv X \)

Event 1, is the only one which can violated the inequality; \( Q \) is the amount of violation. From a mathematical point of view \( Q = Tr(I\rho_{AB}) \) is the mean value of the Bell operator \( I \). Indeed, whenever one, or both, of the partners does not detect its particle, he will give a predetermined result. Thus the inequality cannot be violated in those cases, i.e. \( M_A, M_B, X \leq L \).

Considering all four events, inequality \( I \) is violated when

\[
I_{\eta_A,\eta_B} \equiv \eta_A \eta_B Q + \eta_A (1 - \eta_B)M_A + (1 - \eta_A)\eta_B M_B + (1 - \eta_A)(1 - \eta_B)X > L .
\]
So Alice and Bob must choose the measurement settings \( \{A_i, B_j\} \) and the value they output in the case of no detection, in order to maximize \( I_{\eta_A, \eta_B} \). We stress that the measurement settings that maximize \( I_{\eta_A, \eta_B} \) are not those that maximize the violation of the inequality \( Q \) for the same quantum state, except for the maximally entangled state. Concerning the values assigned to the outputs by Alice and Bob in the case of no detection, we limited ourselves to dichotomic outcomes, i.e. \( a, b = \{0, 1\} \), whether the particle is detected or not. Note that they could also use a third outcome for no detection [74].

6.2 Threshold detection efficiencies

The general approach above can be carried out for any specific values of the efficiencies; now we consider the limit where Alice’s detector is perfect, \( \eta_A = 1 \). We focus on two inequalities: CHSH [52] and \( I_{3322} \) [62]. For those inequalities \( L = 0 \), which implies from (6.1) that the efficiency of Bob’s detector must be above the threshold

\[
\eta_B > \eta_B^{th} = \frac{1}{1 - Q/M_A} \tag{6.1}
\]

in order to close the detection loophole. For any given state, the measurement settings and Bob’s output in case of no detection must be chosen as to maximize \( |Q/M_A| \).

Let us first focus on pure states of two qubits of the form \( |\psi(\theta)\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle \). For the maximally entangled state, \( \theta = \frac{\pi}{4} \), the optimal settings are those which maximize the violation of the inequality. So the threshold efficiency is easily computed from equation (6.1). We obtain \( \eta_B^{th} = \frac{1}{\sqrt{2}} \approx 70.7\% \) for the CHSH inequality and \( \eta_B^{th} = \frac{2}{3} \approx 66.7\% \) for the \( I_{3322} \) inequality. Note that a lhv model is known, which reproduces the correlations of the maximally entangled state under the assumption \( \eta_A = 1 \) and \( \eta_B = 50\% \) [50]; it is an interesting open question to close this gap by finding either a better Bell-type inequality, or a better lhv model.

For non-maximally entangled states, we performed a numerical minimization of \( \eta_B^{th} \): we find that \( \eta_B^{th} \) decreases with decreasing \( \theta \) both for CHSH and \( I_{3322} \), as shown in Fig. 6.2 (thick lines), in analogy with Eberhard’s result [66]. In the limit of weakly entangled states (\( \theta \to 0 \)), one finds \( \eta_B^{th} \to 50\% \) for CHSH and \( \eta_B^{th} \to \sim 43\% \) for \( I_{3322} \) (see Fig. 6.2). It is remarkable that the detection loophole can in principle be closed with \( \eta_B < 50\% \).

We have just seen that \( \eta_B^{th} \) decreases with the degree of entanglement for pure states. But the amount of violation of the inequality decreases as well, so it is important to study the effect of noise. We considered two models of noise: background noise and detection errors. The threshold efficiencies for \( I_{3322} \) as a function of \( \theta \) are shown in Fig. 6.2. The behaviour is similar for both noise models. As expected, when \( \theta \) decreases, the threshold efficiency reaches a minimum: for less entangled states the violation of the inequality is rapidly overcome by the noise.

For background noise, one sees (see Fig. 6.2) that inequality \( I_{3322} \) can tolerate lower efficiencies than the CHSH inequality for \( p \lesssim 6\% \), where \( p \) is the amount of noise in the state, i.e. the global state reads \( \rho = (1 - p)|\psi(\theta)\rangle\langle\psi(\theta)| + p1/4. \)
Figure 6.1: Numerical optimization of the threshold efficiency $\eta_{th}^B$ as a function of $\theta$. Thick lines: results for pure states, for CHSH (dashed-dotted line) and $I_{3322}$ (full line). For $I_{3322}$ the threshold efficiency $\eta_{th}^B$ goes down to $\sim 43\%$ in the limit of weakly entangled states. Thin full lines: $I_{3322}$ with background noise; thin dashed lines: $I_{3322}$ with detection errors; with error value for both.

Figure 6.2: Minimal detection efficiency $\eta_B$ required for a given noise power. The curves for the symmetric case, $\eta_A = \eta_B$, are also plotted, the curve for CHSH being Eberhard’s result. Though CHSH provides a smaller threshold efficiency for any noise power than $I_{3322}$ in the symmetric case, $I_{3322}$ can tolerate smaller efficiencies than CHSH when $p < 6\%$ and $\eta_A = 1$. This is probably due to the fact that inequality $I_{3322}$ is asymmetric, contrary to CHSH.
6.3 Experimental feasibility and outlook

In this chapter we discussed the detection loophole in asymmetric Bell tests. In particular we showed that, for the inequality $I_{3322}$, a minimal detection efficiency of $\eta_B \approx 43\%$ can be tolerated (for $\eta_A = 1$), considering non-maximally entangled states. For maximally entangled states, the threshold efficiency is $\eta_B = 66.7\%$. For these states the lhv model of Ref. [50], based on the detection loophole, provides a lower bound for the threshold efficiency $\eta_B > 50\%$. It is an interesting question whether this bound can be reached by considering other Bell inequalities, or if on the contrary one can find a better model than the one of Ref. [50].

From the experimental point of view, the situation can be summarized as follows. Atom-photon entanglement has been recently demonstrated both with Cd ions [77, 76] and with Rb atoms [75]. Non-maximally entangled atom-photon states were already created in Ref. [76]. An important drawback of these experiments is the very low overall photon detection efficiency, mostly due to inefficient photon collection. However there is good hope that the collection efficiency could be brought to the required level by placing the atom inside a high-finesse cavity [78, 79]. So the perspective for closing the detection loophole for two well-separated systems seems excellent using atom-photon implementations.

Performing a loophole-free Bell experiment is far more demanding. In atom-photon systems, the main inconvenient is the slowness of the atomic state measurement, which is typically based on detecting fluorescence from a cycling transition. As a consequence, enforcing locality in such an experiment would require a large separation between the two detection stations for the atom and the photon. Nevertheless a loophole-free Bell experiment might be possible with asymmetric atom-photon systems, with a photon wavelength that minimizes propagation losses, and highly efficient photon detection.
Chapter 7

Hilbert space dimension witnesses

A central issue in quantum information science is to quantify the computational power of a quantum state, or some quantum correlations. As we have seen in Chapter 5, there is a very well defined hierarchy of correlations; local correlations are strictly weaker than quantum correlations, which themselves are strictly weaker than no-signaling correlations.

In this chapter we address the following problem: given some quantum correlations, i.e. provided by measurements on some unknown quantum state, how does one efficiently bound the dimension of the Hilbert space necessary to fully describe this state?

It is clear that one can always extend the Hilbert space dimension which describes a quantum system; a $d$-dimensional quantum state may as well be described in a Hilbert space of dimension $d'$, where $d' > d$, by adding $(d' - d)$ useless extra dimensions. Thus the relevant question here is to find the dimension $d$ of a state, such that a $(d - 1)$-dimensional Hilbert space does not allow a complete description of this state.

While the problem of witnessing the Hilbert space dimension is of fundamental interest, it is also motivated by quantum cryptography. In standard security proofs of QKD, the correlations shared by the authorized partners, Alice and Bob, are supposed to come from measurements on a quantum state of a given dimension. There are cases where this assumption turns out be crucial for the security of the protocol (see Section 8.2 for an example); in fact one can prove the protocol to be insecure whenever this condition is not satisfied. From a practical point of view it is very difficult to guarantee the dimensionality of a quantum system. There may be so-called side-channels, i.e. other degrees of freedom that contain information about the logical value of the quantum state. If the spy gets access to these side-channel, the security is compromised. We will come back to these issues in the next Chapter, in Section 8.2.

In this Chapter we will focus on bipartite correlations. Basically two cases should be considered independently. First the correlations shared by Alice and Bob are local, in the sense that there exists a lhv model that reproduces perfectly this set of correlations. Then one may ask what is the minimal dimension for a quantum state to reproduce the desired correlations. Here one should also distinguish the
case where the state is separable from the case of an entangled state. Second, the correlations are non-local, in the sense that they violate a Bell inequality. Then one wants to find a constraint that indicates that these correlations cannot be produced by measurements on a quantum state of a dimension less than a given $d$.

In this Chapter we will present results for the second case, i.e. non-local correlations. In section 7.1 we review the Collins-Gisin-Linden-Massar-Popescu (CGLMP) [80] Bell inequality. Then we present a family of linear dimension witnesses in section 7.2. Section 7.3 gives a short outlook.

### 7.1 CGLMP

In this section we consider a typical Bell-test scenario where Alice and Bob choose among two settings, $x, y \in \{0, 1\}$. Each of these settings provides a ternary result $a, b \in \{0, 1, 2\}$. The experiment is described by a set of joint probabilities $p(a, b|x, y)$. If Alice and Bob answer according to some local strategy, their set of joint probabilities will satisfy the Collins-Gisin-Linden-Massar-Popescu (CGLMP) [80] Bell inequality

\[
\begin{array}{cccc}
-1 & -1 & 0 & 0 \\
-1 & 1 & 0 & 1 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & -1 \\
\end{array}
\]

which is the only Bell inequality for such configuration [62], i.e. two ternary settings on each side.

Now, if Alice and Bob are allowed to do measurements on pairs of entangled particles, they will be able to violate the CGLMP inequality. The largest violation of the inequality, 0.3050, is obtained for some an entangled state of two qutrits, which is surprisingly non-maximally entangled [65]. For qubits, there is strong numerical evidence that the largest violation is only $\frac{1}{\sqrt{2}} - \frac{1}{2} \simeq 0.2071$, obtained by suitable measurements on the singlet state.

So the CGLMP inequality provides a first example of a Hilbert space dimension witness, since a violation larger than $\frac{1}{\sqrt{2}} - \frac{1}{2}$ excludes (according to numerical evidences) that the probability distribution can be obtained from measurements on entangled qubits. However this witness provides only a poor vision of the situation; to understand this we use the geometrical description of polytopes.

### 7.2 Linear dimension witnesses

For Bell-tests with two ternary settings on each side, the polytope of all no-signaling correlations lives in a 24-dimensional probability space. Indeed it is very difficult to gain geometrical intuition in such high dimensions. However, in [I] a two-dimensional
slice this no-signaling polytope has been characterized. All points belonging to this slice can be written as the following table of probabilities

\[
\begin{array}{cccc}
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & p_0 & p_2 & p_0 \\
\frac{1}{3} & p_1 & p_0 & p_2 \\
\frac{1}{3} & p_2 & p_0 & p_0 \\
\end{array}
\] 

where \( p_0 + p_1 + p_2 = 1 \).

(7.2)

So any point on this slice is described by two parameters \( p_0 \) and \( p_1 \). Note also that there is a depolarization protocol which allows one to map any point from the complete 24-dimensional NS polytope onto this slice \([I]\). The slice is an equilateral triangle. Its top right corner corresponds to a non-local machine called \( PR_{23} \) defined by the relation \((a + b = \bar{x}y + 1)\) \([60]\). Below this machine, which is an extremal point of the NS polytope, we find the CGLMP inequality, a facet of the local polytope.

The quantum curve above the CGLMP, as plotted in Fig. 7.1, is shown to be the quantum boundary, using the technique of \([81]\). The curve is obtained by using always the same family of settings (known as the Fourier transform settings) on a family of two-qutrits states of the form \( |\psi\rangle = \frac{1}{\sqrt{2+\gamma^2}} (|00\rangle + \gamma |11\rangle + |22\rangle) \). The largest violation (0.3050) is obtained for a non-maximally entangled state \((\gamma = [\sqrt{11} - \sqrt{3}] / 2)\), whereas the maximally entangled state violates only up to 0.2910.

\[\text{Figure 7.1: Above the CGLMP facet.}\]

Indeed the best dimension witness would be to know the analytic form of the qubit curve as plotted in Fig. 7.1. This is a very difficult problem since the quantum region
is not a polytope; it has an infinite number of extremal points — in fact very few is known on the quantum boundaries in such geometrical representations [82].

What we do here is that we construct a family of linear dimension witnesses which are represented by family of Bell-type inequalities

\[
I_b \equiv \begin{vmatrix}
-1 & -1 & -b & 0 \\
-1 & 1 & b & 1 - b \\
-1 & 1 - b & 1 & b \\
0 & 1 & b - 1 & -1 \\
-b & b & 1 & b - 1
\end{vmatrix} \leq 0.
\]  (7.3)

From a geometrical point of view, these inequalities are generated by tilting the CGLMP inequality around the M₃ local point (see Fig. 7.1). Indeed \(I_{b=0}\) is CGLMP, and \(I_{b=2}\) is the vertical facet of the NS polytope. For each inequality the maximal qubit and qutrit violations can be computed numerically. For \(b \geq 1\) qubits no longer violate inequality \(I_b\). It can be shown that \(I_{b=1}\) cannot be violated if A and B dismiss one output for each setting, i.e. if they use binary outputs only. Thus neither a PR-box, nor Von Neumann measurements on any two-qubit state can violate \(I_{b=1}\); so \(I_{b=1}\) is a witness for binary non-locality. Numerical studies also show that three-outputs POVM on qubits cannot violate \(I_{b=1}\), which strongly suggests that \(I_{b=1}\) is a true dimension witness; quantum mechanically it can only be violated with (at least) qutrits.

### 7.3 Outlook

We have presented the problem of detecting the dimension of the Hilbert for quantum correlations. We have presented some first results for the case of non-local correlations. In particular we have shown a family of linear dimension witnesses derived from geometrical considerations of the CGLMP inequality. A drawback of our results is that most of them are numerical. Deriving analytical proofs is very difficult here, in particular for the case of POVM’s. There is nevertheless hope, through the use of new techniques such as the one developed in Barcelona for bounding the set of quantum correlations [81]. It might be possible to strengthen this criteria in order to restrict it to states of a given dimension.
Chapter 8

Other contributions

This chapter is dedicated to contributions which were done during this PhD thesis. They concern quantum cryptography, both applied and theoretical.

In section 8.1, a novel protocol for practical quantum key distribution (QKD) is presented. This protocol is known as coherent-one-way (COW) [H]. It has the nice feature of being both fast and simple.

Section 8.2 contains two contributions in theoretical QKD. The first one is an analysis of the security of a quantum cryptography protocol against a post-quantum Eve [I]. The second one is a security proof of device-independent quantum cryptography [J].

8.1 Fast and simple one-way quantum key distribution

Quantum key distribution (QKD) is the only method to distribute a secret key between two distant partners, Alice and Bob, whose security is based on the laws of physics [83]. While QKD is certainly amongst the most mature fields of quantum information science, there is still much work ahead in order to build a system being reliable and fast at the same time. The COW protocol is significant step in this direction [H]. It is tailored for an implementation with weak coherent pulses to obtain a high key generation rate. Reliability is achieved by using standard telecom components. An important feature of the protocol is its robustness against photon-number-splitting (PNS) attacks which are devastating for other protocols such as BB84 for example.

We now describe briefly the protocol and its practical implementation. Alice holds a source which consists of a continuous wave (cw) laser followed by an intensity modulator (IM), which either prepares a pulse with a mean number of photon $\mu$ or an empty pulse. A logical bit is encoded on a sequence of two successive pulses. Only one of them is empty, such that the k-th logical bit is

$$|0_k\rangle = |\sqrt{\mu}\rangle_{2k-1}|0\rangle_{2k}, \quad |1_k\rangle = |0\rangle_{2k-1}|\sqrt{\mu}\rangle_{2k}. \quad (8.1)$$
Figure 8.1: The COW protocol.

Note that $|0_k⟩$ and $|1_k⟩$ are not orthogonal due to their non-zero vacuum component. However when the time-of-arrival measurement is conclusive, i.e. it gives a click, it provides the optimal unambiguous discrimination of the bit value. With a probability $f << 1$, Alice also sends decoy sequences: $|√μ⟩_{2k−1}|√μ⟩_{2k}$. To check for the presence of an eavesdropper a monitoring line with an interferometer controls the coherence between two successive non-empty pulses. Since the phase remains constant between the pulses — the coherence length being much longer than the interval between the pulses — the interferometer can be aligned such that the interference of two successive pulses is constructive in one of the outputs and destructive in the other one. Thus Alice and Bob can quantify coherence of the $1−0$ bit sequence and the decoy sequences, through the visibility of the interference.

The performance of a QKD protocol is quantified by the achievable secret key rate $R_{sk}$. After sifting, Alice and Bob are left with a fraction $R_S(μ)$ of the key. Then they have to perform some error correction and privacy amplification, which will remove a fraction $h(Q) + I_{EVE}$, where $Q$ is the quantum-bit-error-rate (QBER), $I_{EVE}$ is the fraction of the raw key known to Eve. Finally Alice and Bob have extracted a secret and errorless key. The rate is given by

$$R_{sk} = R_S(μ)[1 - h(Q) - I_{EVE}] . \quad (8.2)$$

With this figure of merit, COW can be compared to other protocols, such as BB84. The main advantage of COW is its robustness to PNS attacks, which allows using a much larger mean number of photons per pulse, typically $μ ≈ 0.5$ — while for BB84 $μ ≈ 0.1$.

A proof-of-principle experiment of the protocol has been realized. It demonstrated that a reasonably low QBER ($\sim 5\%$) and good visibilities ($\sim 98\%$) can be obtained using standard telecom components in an implementation with fiber optics.

To conclude let us mention that a real-world prototype implementing the COW protocol is now under test. On the other hand, the security of this protocol against very general attacks is also a hot topic of research [84].
8.2 Device-independent quantum cryptography

Quantum cryptography shows that one can guarantee the secrecy of correlation on the sole basis of the laws of physics, without limiting the computational power of the eavesdropper. The usual security proofs of QKD [16] suppose that the authorized partners, Alice and Bob, have perfect knowledge and control of their quantum systems and devices; for instance they must be sure that the logical bits of the key have been encoded in true qubits, and not in a higher dimensional system.

An example by Acin, Masanes and Gisin [55] clearly shows how crucial this assumption is. They considered the BB84 protocol [85]. It is well known result from Shor and Preskyll [16] that this protocol is secure as long as the QBER is inferior to 11%. However this is true only for two-dimensional systems, i.e. qubits. Ref. [55] provides an example of a four-dimensional state, which achieves a zero error rate while giving full information about the key to Eve.

In practical implementations it is very difficult to guarantee the dimensionality of a quantum system. There may be some side-channels, unnoticed from Alice and Bob, which contain some information about the encoded quantum state. For example, the four states of the BB84 protocol might be encoded in optical pulses with slightly different frequencies. Indeed if Eve can access this side-channels, which is the case when Eve’s power is only limited by the laws of physics, the security of the protocol is entirely compromised.

Recently a novel approach has been proposed to circumvent this problem. It is based on the idea of device-independent cryptography. The central idea here is that the security is based on violation of a Bell inequality. Note that this was already Ekert’s intuition back in 1991 [5], though he could not formalize it in this way. In this scenario one can even assume that it is Eve who is distributing the correlations, in other words she is selling the apparatus to Alice and Bob.

A first step towards device-independent cryptography was done by Acin, Masanes and Gisin [55], who demonstrated that it is possible to prove the security of QKD without the heavy mathematic tools of QM, but only assuming the no-signaling condition. Note that a previous result in the same spirit was derived Barrett, Hardy and Kent [54]. In Ref. [55], the authors presented a new QKD protocol, the CHSH protocol, whose security is proved against an adversary limited only by the no-signaling condition. Indeed such an eavesdropper is more powerful than any quantum eavesdropper since she is allowed to distribute correlations which are post-quantum. The intuition behind their proof is the following: comparing their list of bits, Alice and Bob can compute the violation of the CHSH Bell inequality of their data. This violation of the CHSH inequality then bounds Eve’s information on the key. Thus whenever the violation of the Bell inequality is large enough, Alice and Bob can, via some post-processing (error correction and privacy amplification) factorize Eve out, such that they share a list secret bits. Remarkably, in the standard scenario where the eavesdropper is quantum and Alice and Bob really have qubits, the CHSH protocol turns out to be equivalent to the BB84.

This study was then extended to a protocol with ternary outputs [I]. Here the security is checked through the violation of the CGLMP [80] Bell inequality. In
this case bounds for security could also be derived. Furthermore the work provided useful results on the topology of the no-signaling polytope for this configuration.

Finally a last work demonstrated the device-independent security of quantum cryptography [J] against so-called collective attacks, which are believed to be the most general attack possible for Eve. Here the main improvement over the previous work is that Eve is now restricted to quantum physics. Indeed, as intuition suggests, limiting Eve’s power improves the key rate that Alice and Bob can achieve. The paper also explicitly gives the optimal attack for Eve in an untrusted-device scenario. The protocol can be shown to be secure up to a maximal QBER of 7.1%; in the usual scenario, i.e. when Alice and Bob use trusted devices, this bound goes up to 11%, the bound of Shor and Preskill [16].
Deuxième partie
Version française
Chapitre 9

Introduction

Quelques années après la découverte de la théorie quantique moderne, il fut réalisé que cette théorie, malgré son impressionnante capacité prédictive, contenait des aspects défiant toute explication rationnelle. Ces étranges propriétés firent à la base du célèbre dialogue entre Albert Einstein et Niels Bohr. Dans un premier, Einstein tenta de démontrer l’inconsistance de la physique quantique, puis il reformula ses critiques dans le but de prouver que la théorie était incomplète. Les objections d’Einstein devinrent célèbres grâce au célèbre article d’Einstein, Podolsky et Rosen (EPR) [1]. EPR argumentait que la théorie quantique et l’hypothèse de localité étaient incompatibles.

Pendant longtemps le débat EPR se poursuivit. Toutefois la majorité de la communauté scientifique le percevait plus comme un problème philosophique que scientifique. Cependant en 1964, le débat fut ravivé par John Bell, qui découvrit le moyen de tester la complétude de la physique quantique [2]. Cette découverte, connue sous le nom d’inégalité de Bell, ne se limite en réalité pas à la simple physique quantique. Il s’agit plutôt d’une preuve générale qu’il existe une limite supérieure pour la corrélation d’événements distants, sous l’hypothèse de localité.

La découverte de Bell créa un nouvel intérêt pour la non-localité. En particulier, une poignée de physiciens chevronnés entreprirent de démontrer expérimentalement la non-localité quantique, par le biais du résultat de Bell [3]. En 1982 cette recherche atteignit son apogée, lorsque Alain Aspect et ses collègues [4] réalisèrent le premier test de Bell avec des mesures séparées spatialement, au sens de la relativité restreinte. Leur résultat confirmèrent les prédictions quantiques!

Ainsi, au cours d’un demi-siècle, la non-localité quantique passa du statut de vague problème philosophique, à celui de fait expérimentalement démontré. Malgré son aspect pour le moins fascinant, la non-localité quantique est restée pendant longtemps un sujet dédaigné, voir même méprisé. Ce désintérêt général était dû à un consensus tacite parmi les physiciens, qui faisait de la non-localité une curiosité de laboratoire, un phénomène qui trouverait sans doute un jour une explication rationnelle.

Durant les années 90, la situation changea radicalement. Grâce à une poignée de physiciens et informaticiens visionnaires, il fut réalisé que l’intrication n’était pas uniquement un concept fascinant, mais une ressource extrêmement puissante.
Cette ressource permet même d’accomplir des tâches impossibles classiquement ! Ici il faut mentionner la découverte d’Arthur Ekert [5], qui en 1991 démontra que la non-localité quantique peut servir à distribuer des clés cryptographiques secrètes entre deux partenaires éloignés l’un de l’autre. Qui plus est, la sécurité de la clé est vérifiée par une violation de l’inégalité de Bell, et ainsi garantie contre tout espion limité uniquement par les lois de la physique. Cette découverte et un poignée d’autres marquent la naissance de la science de l’information quantique, qui depuis vingt ans, n’a cessé de se développer [6].

Il y a encore quelques années, les concepts d’intrication et de non-localité étaient considérés comme deux aspects différents du même phénomène physique. La vérité est bien plus excitante ! Il fut réalisé récemment que la non-localité est une notion fondamentale, qui peut être étudiée per se, sans la moindre référence à la physique quantique. Cette découverte fondamentale repose sur un travail précurseur de Popescu et Rohrlich [7], qui démontrèrent en 1994 que les corrélations quantiques ne forment qu’une petite partie du puzzle ; il existe des correlations contenant plus de non-localité que celles de la physique quantique et qui néanmoins, ne permettent pas de communiquer instantanément. Ces corrélations sont dites no-signaling. Etonnamment, il semble que nombre de propriétés fascinantes de la physique quantique, comme le théorème de non-clonage, la monogamie ou encore les relations d’incertitudes, sont en fait des propriétés générales de toutes les théories no-signaling [9]. Par ailleurs, les nouveau outils développés pour l’étude de la non-localité permettent de quantifier la puissance des corrélations quantiques, ce qui représente en théorie de l’information quantique un problème central.

Cette thèse comprend différents travaux, allant des aspects fondamentaux de la théorie quantique jusqu’à l’optique quantique moderne. À priori certains de ces travaux peuvent sembler très éloignés les uns des autres. Cependant, ils peuvent tous être perçus comme des manifestations diverses du même phénomène : la non-localité quantique. La philosophie de cette recherche est la suivante : la découverte d’applications intéressante passe par une meilleure compréhension des concepts fondamentaux. D’un point de vue fondamental, les différents travaux présentés dans cette thèse abordent deux sujets essentiels de la physique quantique moderne : la théorie de la mesure et les corrélations quantiques.

Les mesures en physique quantique

Le processus de la mesure est fondamental en physique quantique car c’est par lui que l’espace de Hilbert est traduit en langage classique [11]. Dans cette thèse, nous évoquerons deux concepts modernes en théorie de la mesure : les mesures faibles et les mesures généralisées (POVM).

Mesures faibles. Tout d’abord nous mettrons en évidence une belle analogie entre deux sujets de la physique moderne à priori sans relation : la théorie abstraite des mesures faibles [10]et les effets de polarisation dans les réseaux de fibres optiques [17]. En particulier nous montrerons comment le formalisme quantique des mesures faibles simplifie le calcul des effets de polarisation dans les fibres [A]. De plus cette analogie donne une meilleure compréhension du concept de mesure faible. A noter ici que l’intrication joue un rôle clé. C’est le degré d’intrication entre le pointeur et le système physique à mesurer qui caractérise la mesure effectuée ; dans le cas d’une
mesure forte, l'intrication est maximale, alors que pour une mesure faible, l'état est presque séparable. Ensuite, sur la base de cette analogie, nous présenterons une expérience dite de slow et fast light dans une fibre optique [B]. Les résultats incluent des mesures directes des vitesses de groupe supralumineuses, ainsi que des mesures de vitesses de signal. Nous insisterons sur le fait que même dans le cas d'une propagation supraluminineuse, la vitesse de signal, la vitesse à laquelle se propage l'information, reste toujours strictement inférieure à la vitesse de la lumière c.

POVM. Un nouveau schéma pour la mesure de Bell [6], une opération cruciale dans une multitude de protocoles en information quantique, sera présenté [C,D]. Ce schéma est une implémentation en optique linéaire d’un POVM [11] à 21 sorties. Ce POVM permet de distinguer trois des quatre états de Bell. L’implémentation en optique linéaire est optimale dans le sens où l’efficacité maximale de 50% est atteinte [15]. Une expérience de téléportation quantique implémentant avec succès ce nouveau schéma sera décrite.

Corrélations quantiques et inégalités de Bell
Au travers d’une étude de la non-localité, en particulier à l’aide des inégalités de Bell, nous discuterons quatre sujets principaux, étroitement reliés les uns aux autres.

Intrication et non-localité. Par l’exploration des propriétés géométriques des inégalités de Bell [8], nous présentons des résultats illustrant la différence entre les concepts d’intrication et de non-localité. Nos outils principaux seront les polytopes et les machines non-locales. Tout d’abord nous démontrerons que certains états non-maximalement intriqués requièrent une ressource non-locale plus puissante que les états maximalement intriqués, pour être simulés [E,F]. Ceci suggère fortement que l’intrication et la non-localité sont des ressources différentes, une hypothèse déjà sous-jacente dans certains articles récemment publiés.

Detection loophole. Nous dériverons les seuils d’efficacité de détection pour clore le detection loophole dans des expériences de Bell asymétriques [G], comme par exemple pour un système composé d’un atome et d’un photon intriqués. En particulier nous montrerons que lorsque l’une des deux particules est toujours détectée, alors une efficacité de 43% peut être tolérée pour l’autre particule. Par analogie au cas symétrique, nous montrerons que les états non-maximalement intriqués tolèrent des efficacités plus faibles que les états maximalement intriqués [66]. Donc ces états contiennent plus de non-localités que les états maximalement intriqués, dans le sens où une plus petite efficacité de détection permet de mettre en évidence leur non-localité. Ceci confirme l’intuition des résultats du travail précédent [E]. Finalement, nous discutons aussi les implémentations possibles de ces tests de Bell asymétriques. Au vu de nos résultats, il semble tout à fait prometteur de fermer le detection loophole pour des systèmes asymétriques avec la technologie actuelle.


Cryptographie quantique. Bien que l’intrication ne soit pas essentielle pour la
cryptographie quantique, elle reste néanmoins un ingrédient fondamental pour les preuves théoriques de sécurité. Des contributions de cette thèse vont à des nouvelles preuves de sécurité, notamment contre des espions dit post-quantique [I], ainsi que des preuves de sécurité indépendante de l'appareillage [J]. Dans chacune de ces deux preuves, la sécurité du protocole est garantie par la violation d'une inégalité de Bell. Finalement une dernière contribution de cette thèse présente un nouveau protocole de cryptographie quantique expérimentale [H], qui possède deux avantages importants : être à la fois rapide et simple.

Ce manuscrit se compose de plusieurs chapitres, chacun résumant une partie du travail effectué au cours de cette thèse. Chaque chapitre peut se lire indépendamment. La plupart des résultats présentés ici ont fait l’objet de publication dans des revues scientifiques. Ces articles, figurant à la fin de ce manuscrit, sont référencés alphabétiquement. Les autres articles sont référencés numériquement. Bien entendu le lecteur intéressé par plus de détails se tournera directement vers les articles, ou alors vers la version anglaise de cette thèse.
Chapitre 10

Réseau télécom vus comme des mesures faibles avec post-sélection


Avec l’apparition des réseaux optiques à grande distance, les systèmes de télécommunication deviennent de plus en plus complexes. Typiquement un système moderne inclut une multitude de composants, tel que des coupleurs, des circulateurs ou encore des diviseurs de longueurs d’ondes (WDM). Tout ces éléments sont placés entre les tronçons de fibre optique. Un des aspect les plus problématique de ces réseaux est la distortion des impulsions optiques, dues à la combinaison de deux effets de polarisation dans les réseaux : la dispersion de polarisation (PMD) et les pertes dépendantes de la polarisation (PDL) [17]. La PMD est importante dans les fibres, mais plutôt négligeable dans les éléments optiques placés entre les fibres (circulateurs, WDM, etc). La PMD provient de la biréfringence de la fibre optique. Son effet se traduit par une dépendance de la vitesse de groupe d’une impulsion lumineuse en fonction de son état de polarisation. En particulier les modes les plus lent et rapide correspondent à des modes de polarisation orthogonaux. Le second effet, la PDL, est présent dans les éléments optiques placés entre les tronçons de fibre. Ainsi on peut représenter un réseau optique comme une concatenation d’éléments optiques, avec en alternance PMD ou PDL.

Le formalisme standard pour calculer l’effet combiné de la PMD et de la PDL est plutôt lourd. Basés sur la théorie des vecteurs de Jones [18], les calculs sont souvent
longs et pénibles, et doivent souvent être effectués numériquement pour des réseaux de taille réaliste. Il est aussi connu que l’effet combiné de la PMD et de la PDL peut entraîner des phénomènes surprenants, comme une dispersion anormale [17], et ceci même pour des réseaux très simples.

Ici nous montrons qu’un réseau optique peut être décrit comme des mesures faibles avec post-sélection [A]. Cette analogie fonctionne en deux parties. Dans la première nous démontrons que la PMD réalise des mesures de la polarisation. En particulier il est démontré que le régime télécom de la PMD correspond exactement à ce que les théoriciens quantiques appellent une mesure faible. Ensuite nous montrons que la PDL introduit la post-sélection requise pour compléter l’analogie. Nous traitons deux cas distincts : tout d’abord le cas d’une PDL infinie correspondant à une post-sélection sur un état quantique pur, puis une PDL finie correspondant à une post-sélection sur un état quantique mixte. Il faut souligner ici le rôle naturel de la post-sélection, parfois considéré comme un concept quelque peu arbitraire. Dans un réseau optique, on ne considère que les photons présents à la sortie du réseau ; ceux qui ont été perdus dans la fibre ne nous intéressent plus ! À noter encore que ceci est de la physique non-triviale, puisque la post-sélection se fait justement sur le degré de liberté mesuré par la PMD, à savoir la polarisation. Notre résultat principal est que le temps moyen d’arrivée d’une impulsion optique s’exprime mathématiquement comme une valeur faible [10]. En particulier, pour le cas d’une fibre PMD coincée entre deux polariseurs (PDL infinie), nous avons que le temps moyen d’arrivée est donnée par

\[ \langle t \rangle = \frac{\delta \tau}{2} \frac{\langle \phi | \sigma | \psi \rangle}{\langle \phi | \psi \rangle}, \]

où \( | \psi \rangle \) et \( | \phi \rangle \) sont respectivement les états de polarisation pré et post-sélectionnés, et \( \hat{b} \) est l’axe de biréfringence de la fibre PMD et \( \delta \tau \) le délai de groupe (DGD) [17]. Comme une valeur faible a la propriété de pouvoir prendre n’importe quelle valeur, une impulsion peut parcourir le réseau arbitrairement vite ou lentement. Ceci rejoint le fait, déjà connu, qu’une dispersion anormale peut survenir pour ce genre de réseau. Finalement nous dérivons encore des formules générales pour le temps moyen d’arrivée, dans le cas de réseau arbitrairement compliqué. Par la compréhension intuitive des réseaux optiques qu’il fournit, le formalisme des mesures faibles permet de simplifier considérablement les calculs.
Chapitre 11

Slow- et fast-light dans une fibre optique

Dans sa fameuse théorie de la relativité restreinte, Albert Einstein a postulé l'existence d'une vitesse maximale à laquelle tout objet physique peut se déplacer. Cette limite est donnée par la vitesse de la lumière $c$. Un siècle après les célèbres travaux d'Einstein, bon nombre d'articles ont été publiés dans des revues prestigieuses, annonçant des observations expérimentales de vitesses supra-luminiques [23]. Toutefois, aucune de ces expériences ne se trouve être en conflit avec la causalité d'Einstein. Cette contradiction apparente est résolue lorsque le concept de vitesse de groupe est correctement défini. La vitesse de groupe d'une impulsion lumineuse est la vitesse à laquelle se déplace son centre de masse [26]. Remarquablement, ni l'information, ni l'énergie ne se déplace en général à la vitesse de groupe. Par conséquent il n'y a aucun paradoxe dans l'observation de vitesses de groupe dépassant la vitesse de la lumière ! Depuis les célèbres travaux de Sommerfeld, ensuite poursuivis par Brillouin [24], il est connu que l'information se déplace à la vitesse dite de signal, définie comme la vitesse à laquelle se déplace le front montant d'une impulsion. Cette vitesse de signal ne peut en aucun cas dépasser $c$.

Dans ce chapitre nous présentons une expérience de slow- et fast-light dans une fibre optique standard [B]. D'un point de vue théorique cette expérience trouve ses bases dans les résultats présentés au chapitre précédent [A]. En d'autres termes, nous exploitons ici l'analogie entre les réseaux optiques et les mesures faibles, pour réaliser une expérience dans laquelle la vitesse de groupe d'une impulsion lumineuse est considérablement modifiée.

Les précédentes expériences de slow- et fast-light exploitaient en général les propriétés dispersives de certains milieux bien choisis, comme des gaz froids par exemple [29]. Bon nombre de ces expériences utilisent le principe de la transparence induite électromagnétiquement (EIT) [32, 33]. Dans la présente nous allons au contraire créer une dispersion artificielle grâce à l'effet combiné de la PMD et de la PDL [B]. Nous pourrons donc observer des propagations ultra-rapides (plus rapide que la lumière) ainsi que des propagations ultra-lent. De plus notre expérience ne nécessite aucun équipement complexe, mais uniquement des composants télécom standards.
Le schéma expérimental comprend une fibre PMD à maintien de polarisation placée entre deux cubes polarisants. Nous utilisons un OTDR à comptage de photons [34], à la fois comme source et comme détecteur. Cet appareil fréquemment utiliser dans les télécoms sert à créer des profils de fibre optique, en envoyant une quantité bien précise de lumière, puis en analysant la quantité de lumière retro-diffusée. Ici l’OTDR nous permet de faire des photographies instantanées des impulsions lumineuses, grâce auxquelles nous mesurons directement les vitesses de groupe et de signal. Lorsque l’état de post-sélection devient presque orthogonal à celui de pré sélection, on observe une augmentation de la vitesse de groupe. Celle-ci peut même dépasser la vitesse de la lumière. Finalement nos résultats montrent clairement que pour toute vitesse de groupe, la vitesse de signal reste inchangée, et bien entendu strictement inférieure à $c$. 
Chapitre 12
Mesure de Bell

La mesure de Bell est un élément clé d’une multitude de protocole de l’information quantique, comme la téléportation [12, 36, 38], le dense coding [13, 41] et la téléportation d’intrication [39, 40]. La mesure de Bell ont été implémentées avec succès pour des systèmes à variables continues [42] et de l’optique non-linéaire [44]. En optique linéaire la situation est un peu plus compliquée. Pour des qubits il est démontré qu’une mesure de Bell parfaite ne peut être effectuée [15]. Il faut cependant distinguer deux cas. Si l’on tolère l’utilisation d’un grand nombre de photons auxiliaires, alors l’efficacité de la mesure de Bell tend vers un. Ceci est l’essence du fameux résultat de Knill-Laflamme et Milburn (KLM) [45]. Lorsqu’aucun photon auxiliaire n’est toléré, l’efficacité est limitée à 50% [15]. A noter que cette preuve limite uniquement l’efficacité totale de la mesure de Bell, mais pas le nombre d’états de Bell qu’il est possible de discriminer. Jusqu’à aujourd’hui les mesures de Bell proposée en optique linéaire ne permettent de distinguer que deux parmi les quatre états de Bell avec une efficacité globale optimale.

Ici nous présentons un nouveau schéma pour la mesure de Bell permettant de discriminer trois parmi quatre états de Bell, avec une efficacité globale de 50% [C,D]. Mathématiquement cette mesure de Bell est représentée par une mesure généralisée (POVM) [11] à 21 sorties. En pratique elle peut être implémentée pour tous les types de codage. Ici nous présenterons deux implémentations possibles en optique linéaires : en time-bins [47] et en polarisation. L’implémentation en time-bins présentent l’avantage de ne nécessiter que deux détecteurs, alors que l’implémentation en polarisation permet d’obtenir l’efficacité optimale de 50% avec des détecteurs actuels, c’est-à-dire en considérant le temps mort du détecteur.

Finalement une expérience de téléportation démontre expérimentalement la faisabilité de l’implémentation time-bins de la nouvelle mesure de Bell [C,D]. Les résultats expérimentaux reproduisent bien les prédictions théoriques. Il faut néanmoins reconnaître que cette nouvelle mesure de Bell est d’un intérêt limité pour la pratique. En effet les pertes supplémentaires introduites par ce nouveau schéma (environ 3 dB dans l’expérience mentionnée) réduisent bien sur le nombre d’événements conclusifs, ce qui est un sérieux inconvénient pour les applications. Néanmoins l’implémentation d’un tel POVM reste très intéressante au niveau fondamental.
Chapitre 13

Inégalité de Bell pour des ressources non-locales

L’intrication est l’une des plus étonnantes propriétés de la physique quantique. Il est maintenant bien connu que lorsque deux observateurs, partageant d’un état intriqué, et effectuant les mesures appropriées, établissent des corrélations dites non-locales, dans le sens où celles-ci violent une inégalité de Bell [2]. Une caractéristique clé de l’intrication est qu’elle ne permet pas aux deux observateurs de communiquer instantanément, respectant ainsi le principe de causalité. On dit alors que les corrélations quantiques sont no-signaling [9].

Un aspect frustrant de la physique quantique est qu’elle ne donne pas la moindre indication sur le mécanisme qu’elle utilise pour produire cette incroyable non-localité ! Il est dès lors naturel d’essayer de reproduire, ou de simuler, les corrélations quantiques par d’autres moyens. Pour cela, il faut un ingrédient essentiel : une ressource non-locale. L’idée est que cette ressource non-locale, bien que contenant suffisamment de non-localité pour violer les inégalités de Bell autant que les corrélations quantiques, admette une description simple hors du cadre de la théorie quantique. Cette approche permet également de quantifier la puissance des corrélations quantiques, ce qui est fondamental en théorie de l’information quantique.

La communication classique est un premier choix très naturel pour une telle ressource non-locale. Cependant reproduire des corrélations no-signaling avec de la communication n’est pas trivial. Il s’agit en effet de cacher habilement la communication afin d’éviter le signaling. Plusieurs travaux [48, 49] entreprirent la tâche ardue d’estimer la quantité de communication nécessaire pour simuler les corrélations d’un état maximalement intriqué de deux qubits. En 2003, ces résultats partiels furent dépasser par le modèle de Toner et Bacon [51], capable de reproduire les corrélations du singlet avec seulement un bit de communication.

Récemment une autre ressource non-locale s’est imposée pour l’étude de ce problème : la boîte PR [7]. En 1994, Popescu et Rohrlich découvrent qu’il existe des corrélations no-signaling plus non-locales que celles de la physique quantique. En particulier ils présentent un ensemble de corrélations, représenté par une boîte non-locale, capable de violer l’inégalité de Clauser-Horne-Shimony-Holt (CHSH) [52] jusqu’à sa borne algébrique. Cette boîte non-locale est maintenant connue sous le nom de boîte PR.
Bien sûr la boîte PR est parfaitement adaptée pour la simulation des corrélations quantiques. En effet elle est no-signaling par définition et contient plus de non-localité que les états quantiques. Entre temps il fut aussi démontré que la boîte PR est une ressource très puissance pour aborder des problèmes comme la complexité de la communication [53] ou la cryptographie [54, 55]. Récemment Cerf, Gisin, Massar et Popescu [57] ont présenté un modèle capable de simuler les corrélations du singlet avec une boîte PR. Un lien étroit entre ce modèle et celui de Toner et Bacon a été mis en évidence par Degorre, Laplante et Roland [58].

Au vu de ces résultats, il semble naturel d’étendre cette étude à d’autres états. Or il s’avère que ceci est très difficile, même dans le cas d’un état de deux qubits partiellement intriqués. Dans ce chapitre nous démontrons que le modèle de Cerf et al. ne peut pas être étendu à une famille d’états partiellement intriqués. Plus précisément nous dérivons une inégalité de Bell ne pouvant être violée à l’aide d’une boîte PR, puis nous montrons que certains partiellement intriqués violent cette inégalité [E,F]. En d’autres termes, ces états quantiques requièrent une quantité plus importante de ressources non- locales pour être simulés. Ce résultat suggère fortement que non-localité et intrication sont deux ressources différentes.

Dans ce travail, la géométrie des inégalités de Bell [8] est notre principal outil. Sans entrer dans les détails, nous travaillons dans un espace de probabilité, dans lequel chaque point représente un ensemble de corrélations. Dans cet espace, toutes les corrélations pouvant être établie par une stratégie prédéfinie forme un polytope [F]. Les faces de ce polytope représentent des inégalités de Bell, dans le sens où un point situé à l’intérieur (à l’extérieur) du polytope satisfait (viole) l’inégalité. Notre démarche est la suivante. Partant d’une inégalité de Bell récemment découverte, I_{3322} [62], nous listons toutes les stratégies extrêmes no-signaling, au-dessus de cette face, utilisant au plus une boîte PR. Comme le nombre de ces stratégies est fini, on obtient un nouveau polytope. Celui-ci contient toutes les stratégies réalisables avec une boîte PR. Ensuite nous caractérisons une des faces de ce polytope, puis montrons qu’une famille d’états partiellement intriqués violes cette inégalité. Ceci illustre bien la différence entre l’intrication et la non-localité, déjà mise en évidence dans de précédents travaux. On sait par exemple que certaines inégalités de Bell sont maximalement violées par des états non-maximalement intriqués [65]. De plus les états partiellement intriqués tolèrent de plus petites efficacités de détection pour clore le detection loophole [66]. Cet aspect sera le sujet du prochain chapitre de cette thèse.

Finalement nous terminons par quelques questions ouvertes. Notre résultat n’a aucune conséquence sur la simulabilité des états partiellement intriqués avec un bit de communication, puisqu’une boîte est une ressource strictement plus faible qu’un bit de communication, même restreint au cas des stratégies no-signaling [E]. Par conséquent savoir si tout les états de qubits sont simulables avec un seul bit de communication est un problème toujours ouvert et très intéressant. Au niveau de boîte PR, en considérant notre résultat, on peut se demander combien de boîtes PR il faut pour simuler tout les états partiellement intriqués. Ceci est un problème difficile car il s’agit ici, soit de construire un modèle explicite reproduisant les bonnes corrélations, ou alors de trouver une inégalité valable pour toute stratégie avec plusieurs boîte PR violée par la physique quantique.
Chapitre 14

Loophole de détection pour les tests de Bell asymétriques

En 1964 John Bell dérive sa fameuse inégalité [2]. Depuis, toutes les expériences sur la non-localité quantique ont confirmé la théorie [3, 4, 69, 70], à savoir une violation de l’inégalité de Bell. Toutefois, d’un point de vue strictement logique, aucune de ces expériences ne rejette définitivement une explication par des théories locales. En effet il existe dans toutes ces expériences des loopholes, laissant la porte ouverte à des modèles à variable cachées (lhv). En résumé, il y a donc de très fortes évidences expérimentales que la Nature est non-locale, mais compte tenu de l’importance d’une pareille affirmation il faut impérativement réaliser une expérience sans loophole! D’autre part la science de l’information quantique donne une motivation supplémentaire, puisque la sécurité de certains protocoles de communication quantique est garantie par une violation ”loophole-free” des inégalités de Bell [54, 55].

Réalisé un test de Bell sans loophole est très difficile. Il s’agit dans un premier temps de s’assurer qu’aucun signal ne peut être transmis entre les deux particules au cours de la mesure. Par conséquent la mesure d’un côté doit se faire hors du cône de lumière de la mesure sur l’autre particule. Si cette condition n’est pas satisfaite, il devient alors possible qu’une particule envoie une information à l’autre, lui communiquant par exemple la mesure qu’elle a subi. Ceci est l’essence du loophole de localité [72]. Ensuite il faut encore que les particules soient détectées avec une efficacité suffisamment élevée. En effet si ce n’est pas le cas, il existe des modèles lhv qui reproduisent exactement les prédictions quantiques. Par exemple on peut imaginer qu’une variable cachée portée par la particule décide, en fonction du choix de la mesure, si la particule est détectée ou pas. Ceci est le loophole de détection [73].

En pratique, les expériences conduites sur des photons ont pu clore le loophole de localité [4, 69]. En effet les photons se déplaçant très vite, avec peu de pertes, il est relativement facile distribuer des paires de photons intriqués sur des grandes distances. Par contre, les efficacités de détection optique sont toujours trop faibles pour clore le loophole de détection. Pour l’inégalité CHSH, une efficacité d’au moins 82% est requise, pour l’état singlet. Etonnamment, cette efficacité limite peut être diminuée jusqu’à 67% en utilisant des états non maximalement intriqués [66]. Ce
résultat est dû à Eberhard. Part ailleurs une expérience sur des ions intriqués [71], distants de quelques microns, a fermé le loophole de détection. Par conséquent, il serait déjà intéressant de clore le loophole de détection sur une distance appréciable.

Dans ce chapitre nous présentons une étude sur les tests de Bell asymétriques, où chacune des deux particules est détectée avec une efficacité différente. En effet, comme le suggère l’intuition, lorsque l’une des deux particules est détectée très efficacement, on peut tolérer une efficacité plus faible pour la seconde particule. C’est le cas par exemple dans les systèmes d’atome et photon intriqués, dans lesquels l’atome est toujours détecté. Ici nous nous concentrons sur ce cas et démontrons qu’une efficacité de 43% peut être tolérée pour le photon lorsque l’atome est détecté avec probabilité un [G]. En analogie avec le résultat d’Eberhard, nous montrons que les états partiellement intriqués aident également.

Finalement nous discutons les implémentations expérimentales, et montrons que, au vue de la technologie actuelle [75], fermer le loophole de détection dans des systèmes atome-photon semble tout à fait réaliste.
Chapitre 15

Détecter la dimension de l’espace de Hilbert

En théorie de l’information quantique il est fondamental de caractériser la puissance des corrélations quantiques. Comme on l’a vu précédemment il existe maintenant toute une classification des corrélations : les corrélations locales sont strictement plus faibles que les corrélations quantiques, qui sont elles-mêmes strictement plus faibles que les corrélations no-signaling. Dans ce chapitre, nous essayons de caractériser la puissance des corrélations quantiques en fonction de la dimensionnalité de l’état mesuré. Le problème peut s’enoncer comme suit : étant donné des corrélations quantiques, obtenues par des mesures sur un état quantique, comment borne-t-on efficacement la dimension de l’espace de Hilbert dans lequel vit cet état ?

Tout d’abord, il est clair que la dimension d’un état quantique peut toujours être agrandie. En d’autres termes, un système de dimension $d$ peut toujours être décrit comme un système de dimension supérieure $d’ > d$. Pour cela, il suffit simplement d’ajouter des dimensions supplémentaires inutiles. Ainsi la question pertinente est de trouver la dimension $d$ d’un état, telle qu’un espace de Hilbert de dimension $(d – 1)$ ne donne pas de description complète de l’état.

Ici nous nous concentrerons sur les corrélations bipartites non-locales. Nous étudions particulièrement l’inégalité Collins-Gisin-Linden-Massar-Popescu (CGLMP) [80]. Cette inégalité est déjà en quelque sorte un témoin de dimensionnalité, puisque les états de quqrits permettent d’obtenir des violations plus grandes que les états de qubit. Grâce à une approche géométrique nous trouvons une famille de témoins de dimensionnalité, qui ne sont autres que des inégalités à la Bell. En particulier nous trouvons une inégalité ne pouvant être violée par la non-localité binaire. Pour en conclure que cette inégalité ne peut effectivement pas être violée par des qubits, il faut encore montrer que les POVM à trois sorties ne permettent pas de violer cette inégalité. Ceci est une recherche en cours.
Chapitre 16

Contributions en cryptographie quantique

Ce chapitre est dédié à plusieurs contributions de cette thèse à la cryptographie quantique, appliquée et théorique. La cryptographie quantique [83] figure sans aucun doute parmi les branches les plus mûres de l’information quantique. Il reste toutefois énormément de progrès à faire, au niveau théorique ainsi qu’au niveau des systèmes expérimentaux. Dans la pratique ce sont la rapidité, la fiabilité et la simplicité qui sont les aspects clés d’un protocole. Du point de vue théorique, les preuves de sécurité très générales sont devenues une priorité. Dans ce type de preuves mathématiques, le but est de garantir la sécurité d’un protocole contre un espion limité aux seules lois de la physique.

La première contribution de cette thèse est l’étude d’un nouveau protocole de cryptographie quantique expérimentale [H]. Ce protocole connu sous le nom de Coherent-One-Way (COW), se distingue par sa capacité à être à la fois simple et rapide. Il est basé sur une implémentation en pulses faibles, obtenus par un laser atténué, et permet d’obtenir des taux de clé élevés. Un avantage du protocole COW est sa robustesse aux attaques dites de photon-number-splitting (PNS), qui sont dévastatrices pour beaucoup de protocole.

Dans les preuves standards de sécurité de cryptographie quantique, il est toujours supposé qu’Alice et Bob partagent un état quantique d’une dimension parfaitement connue. Donc ils doivent s’assurer par exemple que les états qu’ils mesurent sont bien des qubits, et pas des systèmes de dimension plus élevée. Un exemple récent présenté par Acin, Masanes et Gisin [55] illustrent parfaitement l’importance de cette hypothèse. Les auteurs considèrent le célèbre protocole BB84 [85], dont la sécurité est démontrée dans un célèbre papier de Shor et Preskyll [16]. Acin et collaborateurs [55] donne un état de dimension quatre, n’introduisant aucune erreur dans le protocole, mais donnant néanmoins toute l’information à l’espion ! En pratique il est extrêmement difficile de déterminer la dimensionnalité d’un système quantique. En effet il peut toujours exister des degrés de liberté auxiliaires, inaperçus d’Alice et de Bob, qui contiennent toute l’information sur le codage de l’état quantique. Bien entendu lorsque l’espion peut accéder à ces degrés de liberté, la sécurité du protocole est entièrement compromise. Pour contourner ce problème, une nouvelle approche a
été récemment proposée. L'idée est d'étudier la sécurité dite "device-independent" de certains protocoles, c'est-à-dire qu'Alice et Bob ne font aucune confiance au canal quantique qu'il mesure, ni à l'appareil de mesure lui-même. C'est en quelque sorte comme si Alice et Bob avaient acheté leur système de cryptographie chez l'espion directement. Etonnament il est tout de même possible de garantir la sécurité dans ce cas. Celle-ci repose ici sur la violation d'une inégalité de Bell. A noter que cette intuition était déjà à la base de l'idée d'Ekert en 1991 [5], même s'il n'a pu la formuler dans ces termes.

Une première étape vers la cryptographie "device independent" fut réalisée récemment par Acín, Masanes et Gisin [55]. Ils ont démontré la sécurité d'un nouveau protocole, appelé protocole-CHSH, contre un espion post-quantique. En clair, ce nouvel espion n'est plus limité par les lois de la physique quantique, mais seulement par le principe de no-signaling. Dans cette situation, on suppose que c'est l'espion qui distribue les corrélations. La sécurité est garantie par une violation suffisamment grande de l'inégalité CHSH [52]. Dans une contribution de cette thèse, cette étude est étendue à l'inégalité CGLMP [I]. La sécurité de la distribution de clé y est démontrée. Le papier comprend également une étude de la géométrie du polytope no-signaling au dessus de la face CGLMP.

Finalement, un papier plus récent, démontre la sécurité de la cryptographie dite device-independent contre les attaques collectives [J], supposées être les plus générales. Ici le résultat précédent est amélioré dans le sens où l'espion est à présent limité à la physique quantique. Comme le suggère l'intuition, le taux de clé secrète augmente dans ce cas, puisque le pouvoir de l'espion est réduit. La sécurité est assurée jusqu'à un taux d'erreur (QBER) de 7.1% ; dans la limite du scénario standard, dans lequel les appareils sont parfaitement connus, la borne de sécurité remonte à 11%, qui n’est autre que la célèbre borne de Shor-Preskill [16].
Bibliography / Bibliographie


Publication list / Liste des publications


Remerciements

Je tiens à remercier tout d’abord le Prof. Nicolas Gisin pour m’avoir donné l’opportunité d’effectuer cette thèse au sein du Groupe de Physique Appliquée, ainsi que pour sa confiance, son soutien et sa disponibilité.

Ensuite je remercie chaleureusement le Prof. Valerio Scarani pour son soutien, sa confiance, ainsi que pour les (innombrables) discussions passionantes que nous avons eues au cours de cette thèse (et que nous aurons encore, j’espère !).

Je remercie également le Prof. Sandu Popescu, membre du jury de cette thèse.

Bien entendu je remercie encore tout les (ex-)membres du groupe de théorie au GAP. En particulier le Prof. Antonio Acin et le Dr. Christoph Simon, qui outre le fait d’être membres du jury de cette thèse, m’ont appris énormément. Ma gratitude va également à Cyril Branciard, André Méthot, Daniel Collins, Sofyan Iblisdir et Barbara Kraus.

Je dois également saluer la performance de Matthieu Legré et Mark Wegmüller, qui m’ont magistralement accueilli dans leur laboratoire, le temps d’une expérience. Toujours dans le registre expérimental, je remercie Jeroen van Houwelingen et Alexios Beveratos pour leur acharnement à réaliser la mesure de Bell proposée dans cette thèse.

Je remercie encore Matthäus Halder, Sylvain Fasel, Rob Thew, Mikael Afzelius, Hugues de Riedmatten, Hugo Zbinden, et Sébastien Tanzili pour les discussions scientifiques et l’ambiance.

Merci encore à mes parents pour leur soutien et leurs encouragements.

Finalement merci à Louise.
Part III

Published articles /
Articles publiés
Optical Telecom Networks as Weak Quantum Measurements with Postselection

Nicolas Brunner, Antonio Acín,* Daniel Collins, Nicolas Gisin, and Valerio Scarani
Group of Applied Physics, University of Geneva, 20, rue de l’Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland
(Received 17 June 2003; published 28 October 2003)

We show that weak measurements with postselection, proposed in the context of the quantum theory of measurement, naturally appear in the everyday physics of fiber optics telecom networks through polarization-mode dispersion (PMD) and polarization-dependent losses (PDL). Specifically, the PMD leads to a time-resolved discrimination of polarization; the postselection is done in the most natural way: one postselects those photons that have not been lost because of the PDL. The quantum formalism is shown to simplify the calculation of optical networks in the telecom limit of weak PMD.

DOI: 10.1103/PhysRevLett.91.180402 PACS numbers: 03.65.Ta, 42.81.Gs

Several times in the history of science, different people working on different fields and with different motivations happened to discover the same thing, or to introduce the same concepts. Think of the connection between differential geometry and general relativity: physics received a convenient mathematical tool for its predictions, mathematics gained in popularity and interest because, apart from its intrinsic beauty, it proved useful. In this paper, we point out a connection which should help to bring together two very different communities: quantum theorists and telecom engineers. The physical degree of freedom that supports this connection is the polarization of light; we show that the quantum formalism of weak measurements and postselection [1–3] applies to the description of polarization effects in optical networks [4]. The structure of this Letter is as follows: we give first a qualitative description of the announced connection. Then, we introduce the mathematical formalism, and show that the connection does indeed hold down to the detailed formulas; in particular, the knowledge of the “quantum” formalism can simplify some “telecom” calculations.

A modern optical network is composed of different devices connected through optical fibers. With respect to polarization, two main physical effects are present. The first one is polarization-mode dispersion (PMD); due to birefringency, different polarization modes (P modes in the following) propagate with different velocities; in particular, the fastest and the slowest polarization modes are orthogonal. PMD is the most important polarization effect in the fibers. The second effect is polarization-dependent loss (PDL), that is, different P modes are differently attenuated. PDL is negligible in fibers, but is important in devices such as amplifiers, wavelength-division multiplexing couplers, isolators, circulators, etc. In particular, a perfect polarizer is an element with infinite PDL, since it attenuates completely a P mode. Thus, an optical network can be described by a concatenation of trunks, alternating PMD and PDL elements. Combined effects of PMD and PDL elements have been studied in Refs. [5,6]; in particular, interesting phenomena such as anomalous dispersion have been shown to arise even in simple concatenations; namely, a PDL element sandwiched between two PMD elements.

The first piece of the connection we want to point out is the following: a PMD element performs a measurement of polarization on light pulses (Fig. 1). In fact, PMD leads to the separation of two orthogonal P modes in time; this separation is called differential group delay (DGD), noted $\delta \tau$. If $\delta \tau$ is larger than the pulse width, the measurement of the time of arrival is equivalent to the measurement of polarization —PMD acts then as a “temporal polarizing beam splitter.” However, in the usual telecom regime $\delta \tau$ is much smaller than the pulse width. In this case, the time of arrival does not achieve a complete discrimination between two orthogonal P modes anymore; but still, some information about the polarization of the input pulse is encoded in the modified temporal shape of the output pulse. We are in a regime of weak measurement of the polarization; we are going to show later that we recover indeed the notion of weak measurement of the quantum theorists, by measuring the mean time of arrival (that is, the “center of mass” of the output pulse).

The second piece of the connection defines the role of PDL: a PDL element performs a postselection of some polarization modes. Far from being an artificial ingredient, postselection of some modes is the most natural situation in the presence of losses: one does always postselect those photons that have not been lost. This would be trivial physics if the losses were independent of any degree of freedom, just like random scattering; but in the

FIG. 1. When a polarized pulse passing through a PMD fiber, the P mode $H$ (parallel to the birefringency axis in the Poincaré sphere) and its orthogonal $V$ are separated in time. A measurement of the time of arrival (TOA) is a measurement, strong or weak, of the polarization.
case of PDL, the amount of losses depends on the meaningful degree of freedom, polarization. An infinite PDL, as we said above, would correspond to the postselection of a precise \( P \) mode (a pure state, in the quantum language); a finite PDL corresponds to postselecting different \( P \) modes with different probabilities (a mixed quantum state).

In summary: by tuning the PMD, we can move from weak to strong measurements of polarization; by tuning the PDL, we can study the postselection of a pure or of a mixed state of polarization. This is the main result of this Letter, that we are now going to present in mathematical terms.

It is convenient to use the formalism of two-dimensional Jones vectors, in which the description of classical polarization is identical to the quantum description of the spin \( \frac{1}{2} \) [7]. Thus, e.g., the three typical pairs of orthogonal polarizations — horizontal-vertical linear, diagonal linear, left-right circular — are described, respectively, by the eigenvectors of the Pauli matrices \( \sigma_x \), \( \sigma_y \), and \( \sigma_z \). In this Letter, we shall only need to define the eigenvectors of \( \sigma_z \): \( \sigma_z | H \rangle = | H \rangle \), \( \sigma_z | V \rangle = | V \rangle \). Any pure polarization state can be described as a superposition of these vectors, with complex coefficients, the state corresponding to the point \( \hat{n} = (\theta, \phi) \) on the Poincaré sphere being | + \( \hat{n} \rangle = \cos \frac{\theta}{2} | H \rangle + \sin \frac{\theta}{2} e^{i\phi} | V \rangle \).

On a monochromatic wave of frequency \( \omega \), a PDL that separates the eigenvectors of \( \sigma_z \) for a birefringency \( b \) is represented by the operator [5]

\[
\text{PMD: } U(b\omega, \hat{z}) = e^{i b\omega \sigma_z / 2} = \cos \frac{b\omega}{2} I + i \sin \frac{b\omega}{2} \sigma_z. \tag{1}
\]

This is a unitary operation that describes a global rotation of the state of polarization around the \( z \) axis of the Poincaré sphere. As for PDL: since the most and least attenuated states are always orthogonal, they can be written as the eigensates of \( \sigma_z = \hat{n} \cdot \hat{\sigma} \), where the direction \( \hat{n} \) has a priori no link with the direction \( \hat{z} \) of the birefringency axis. Neglecting a global attenuation, the PDL is represented by the operator [5]

\[
\text{PDL: } F(\mu, \hat{n}) = e^{\mu \sigma_z / 2} = \cos \frac{\mu}{2} I + i \sin \frac{\mu}{2} \sigma_z. \tag{2}
\]

This is a nonunitary operation, sometimes called a filter; in the quantum theory, it appears also in the unambiguous discrimination of nonorthogonal quantum states [8]. It has been shown in Ref. [5] that any optical network can be modeled by an effective PDL followed by an effective PDL, that is, by an operator of the form \( F(\mu, \hat{n})U(\eta, \hat{m}) \). However, the study of the general case is involved because the effective parameters \( \mu, \hat{n}, \eta, b \), and \( \hat{m} \) depend on the optical frequency \( \omega \) in a nontrivial way, leading to deformations in the shape of the light pulse. Thus, we focus initially on the simplest optical network, namely a PDL fiber followed by a PDL element.

The input state is a Gaussian (Fourier-transform limited) light pulse of coherence time \( t_c \), of central frequency \( \omega_0 \), prepared in a pure polarization state \( | \psi_0 \rangle \):

\[
| \Psi_{\text{in}} \rangle = \mathcal{A} \ e^{-(1/4)(t/t_c)^2} e^{-i\omega_0 t} \otimes (\alpha | H \rangle + \beta | V \rangle) = g(t) \otimes | \psi_0 \rangle, \tag{3}
\]

with \( \mathcal{A} = (\sqrt{2\pi} t_c)^{-1/2} \) so that \( g(t) = |g(t)|^2 \) is a probability distribution [9]. To compute the state of the light at the output of the PMD fiber, we must Fourier transform \( | \Psi_{\text{in}} \rangle \) into the frequency domain, apply (1) to any monochromatic component, and integrate back to the time domain. This gives [10]

\[
| \Psi_{\text{PMD}} \rangle = \int d\omega \ e^{-i\omega\hat{g}(\omega - \omega_0) U(b\omega, \hat{z}) | \psi_0 \rangle} = \alpha g_- (| H \rangle + \hat{\beta} g_+ (| V \rangle), \tag{4}
\]

where \( g_\pm(t) = g[t \pm (\delta \tau/2)] \) with \( \delta \tau = b, \alpha = e^{ib\omega_0/2} \) and \( \hat{\beta} = e^{-ib\omega_0/2} \). We see that, in addition to the global rotation around the birefringency axis at the frequency \( \omega_0 \), the PMD has delayed the \( V \) polarization with respect to the \( H \) polarization, as announced. According to whether the delay \( \delta \tau \) is much larger or much smaller than the width \( t_c \) of the input pulse, the recording of the time of arrival will provide us with a strong or a weak measurement [11]. For further reference, let us define the polarization state

\[
| \psi \rangle = U(b\omega_0, \hat{z}) | \psi_0 \rangle = \alpha | H \rangle + \hat{\beta} | V \rangle \tag{5}
\]

obtained by retaining only the global rotation, that is, in the limit of continuous light \( \delta \tau/t_c \approx 0 \).

Now, we should apply the PDL operator (2) to \( | \Psi_{\text{PMD}} \rangle \). Before presenting the general case, to become familiar with the concepts, we study the case of postselection of a pure state: the PDL element is then a polarizer that projects onto a polarization state \( | \psi_1 \rangle = \mu | H \rangle + \nu | V \rangle \).

Thus, at the output of the optical network we have

\[
| \Psi_{\text{out}} \rangle = [\alpha \hat{\mu} g_- (t) + \hat{\beta} \hat{\nu} g_+ (t)] | \psi_1 \rangle = f(t) | \psi_1 \rangle, \tag{6}
\]

where \( \hat{z} \) is the conjugate of a complex number \( z \). Clearly \( f(t) \) is the temporal shape of the selected component of the field. Now we measure the intensity \( I(t) = |f(t)|^2 \), with \( A = \alpha \hat{\mu} \) and \( B = \hat{\beta} \hat{\nu} \), we have

\[
I(t) = |A|^2 G_- (t) + |B|^2 G_+ (t) + 2 \text{Re}(\bar{A}B) g_- (t) g_+ (t). \tag{7}
\]

In the limit of strong measurement, \( \delta \tau \gg t_c \), the overlap \( \bar{g} \cdot g \) is essentially 0, so the detected intensity corresponds to two well-separated Gaussians: \( I(t) = |A|^2 G_- (t) + |B|^2 G_+ (t) \). A detection in \( G_- \) corresponds to the \( H \) polarization, so the probability that the polarization was \( | H \rangle \) given the preparation and postselection is simply the integral of the Gaussian \( G_- \), normalized to the total intensity: \( P(H) = \int_0^\infty I(t) dt / \left( \int_0^\infty I(t) dt \right) = \frac{|A|^2}{(|A|^2 + |B|^2)} \) but \( |A|^2 \) is the probability \( P(H) | \psi_0 \rangle \) of finding a photon polarized...
along $|H\rangle$ given that the state is $|\psi_0\rangle$; using similar notations for $|\beta|^2$, $|\mu|^2$, and $|\nu|^2$, we have found

$$P(H) = \frac{P(|\psi_1\rangle|H\rangle P(H)|\psi_0\rangle}{\sum_{K=H,V} P(|\psi_1\rangle|K\rangle P(K)|\psi_0\rangle).} \quad (8)$$

This is the Aharonov-Bergmann-Lebowitz (ABL) rule [12], which corresponds to the classical rule for the probability of sequential events. Since we have access to both $P(H)$ and $P(V)$, we can compute $\langle \sigma_\tau \rangle = P(H) - P(V)$, Moreover, the mean time of arrival, defined as usual by $\langle t \rangle = (\int t|I(t)|dt)/\int |I(t)|dt$, is here $P(H)(\delta\tau/2) + P(V)(-\delta\tau/2)$. So, for the case of a strong measurement, we have derived the relation

$$\langle t \rangle = \frac{\delta\tau}{2} \langle \sigma_\tau \rangle. \quad (9)$$

This is the relation that appears in any measurement theory between the pointer or meter (here, the mean time of arrival) and physical quantity to be measured (here, $\sigma_\tau$). Even though it has been derived from more intuitive grounds in the regime of strong measurements, the relation (9) is the fundamental relation of a measurement process in which the coupling between the pointer and the observable quantity is made by the PMD [11]. In particular, contrary to $P(H)$ and $P(V)$, $\langle t \rangle$ can be defined and measured for any $I(t)$. We shall then take (9) as the definition of the mean value of $\sigma_\tau$ when measured by the PMD. With this, we can remove the assumption of strong measurement.

Starting with $I(t)$ given by (7), $\langle t \rangle$ can be calculated analytically in a straightforward way, and the relation (9) yield

$$\langle \sigma_\tau \rangle = \frac{|A|^2 - |B|^2}{|A|^2 + |B|^2 + 2\text{Re}(AB) e^{-(1/2)(\delta\tau/2)t^2}}. \quad (10)$$

Note that the dependence in the strength of the measurement (i.e., in $\delta\tau/2t_\tau$) is very explicit in (10). In the limit of strong measurement, $\delta\tau/2t_\tau \to \infty$, we recover the above results. In the opposite limit, $e^{-(1/2)(\delta\tau/2t_\tau)^2} \to 1 - O(\delta\tau/2t_\tau)$, corresponding to a weak measurement, we have $\langle \sigma_\tau \rangle_w = \text{Re}(A - B)(A + B)$. Noticing that

$$A \pm B = \bar{\alpha} \bar{\mu} \pm \bar{\nu} \bar{\nu} = \langle |\psi_1\rangle |\psi\rangle, \quad \langle |\nu\rangle |\bar{\sigma}_\tau \rangle |\psi_0\rangle,$$

with $|\psi\rangle$ given by (5), we find

$$\langle \sigma_\tau \rangle_w = \text{Re}\left(\langle |\psi_1\rangle |\sigma_\tau \rangle |\psi\rangle, \quad \langle |\nu\rangle |\bar{\sigma}_\tau \rangle |\psi_0\rangle\right). \quad (12)$$

This is exactly the formula for the weak value of $\sigma_\tau$ when the postselection is done on a pure state $|\psi_1\rangle$ as given by the quantum theorists [1,2]. Note, in particular, that $\langle \sigma_\tau \rangle_w$ can reach arbitrarily large values, leading to an apparently paradoxical situation since the eigenvalues of $\sigma_\tau$ are $\pm 1$. But there is no paradox at all: $\langle \sigma_\tau \rangle_w > 1$ simply means $\langle t \rangle > (\delta\tau/2)$, and this situation is reached by postselecting a state $|\psi_1\rangle$ that is almost orthogonal to $|\psi\rangle$; these are very rare events, the shape $f(t)$ of the pulse is strongly distorted, and it is not astonishing that its center of mass could be found far away from its expected position in the absence of postselection.

We can now examine the case of a finite value of the PDL after the PMD fiber. For conciseness, we write $F(\mu, \hat{n}) = F$ for the PDL operator (2). At the output of the PMD-PDL truncated, the state is

$$|\psi_\text{out}\rangle = F|\psi_\text{PMD}\rangle = A(t)|H\rangle + B(t)|V\rangle, \quad (13)$$

where $A(t) = \langle H|F|H\rangle \bar{\sigma}_g(t) + \langle H|F|V\rangle \bar{\sigma}_g(t)$,

$$B(t) = \langle V|F|V\rangle \bar{\sigma}_g(t) + \langle V|F|H\rangle \bar{\sigma}_g(t), \quad (14)$$

with $C = \cosh^2 \gamma, S = \sinh^2 \gamma, n_c = n_n \pm \sinh \gamma$. We can then calculate the detected intensity $I(t) = |A(t)|^2 + |B(t)|^2 = |a|^2 (\cosh \mu + n_c \sinh \mu)G_+ + |\beta|^2 (\cosh \mu - n_c \sinh \mu)G_+ + 2 \sinh \mu \Re(\bar{\sigma}_g \bar{\sigma}_g(t))$. The mean time of arrival is then calculated; with $\gamma = \tanh \mu$, the result is

$$\langle t \rangle = \frac{\delta\tau}{2} + \frac{1}{1 + \gamma n_c (|a|^2 - |\beta|^2) + 2 \text{Re}^{-1/2}(\delta\tau/2t_\tau)^2}. \quad (16)$$

Again, in the limit of weak measurement and using (9), we find

$$\langle \sigma_\tau \rangle_w = \frac{\langle \sigma_\tau \rangle^\phi + \gamma n_c}{1 + \gamma n_c \cdot \langle \bar{\sigma}_\tau \rangle^\phi} = \text{Re}\left(\langle F^\dagger F \sigma_\tau \rangle^\phi \langle F^\dagger F \rangle^\phi\right). \quad (17)$$

with $|\psi\rangle$ given by (5) as before. The right-hand side is the weak value obtained by postselection on the mixed state $\rho = [1/\text{Tr}(F^\dagger F)] F^\dagger F$ [2,3]. The limiting case $\gamma = 0$ means $\mu = 0$, thence $\rho = \frac{1}{2} (I + \sigma_\tau)$, and we recover the formula (12) for the postselection of the pure state $|\psi_1\rangle = |\pm \hat{n}\rangle$. Finally, we stress that the principal states of polarization of the PMD-PDL network, as defined, e.g., in Ref. [5], are $F(H)|H\rangle$ and $F(V)|V\rangle$ [13].

We have then demonstrated our claims: an optical PMD-PDL network is an everyday realization of the abstract notions of weak measurement and postselection introduced in the theory of quantum measurement. We had also said that telecom engineers would benefit by learning some quantum formalism, were it only because it could simplify their calculations. Indeed, consider a more complicated optical network, composed of three trunks: PMD-PDL-PMD, represented by the operator $T = U(b_2, \omega, \hat{m}) F(\mu, \hat{n}) U(b_1, \omega, \hat{z})$. As we noticed above, this simple network is sufficiently complex to yield anomalous dispersion. The calculation can of course be done following the same steps as above, but it is heavy and not really instructive. Another approach, that is
moreover scalable to any network consisting of $2N + 1$ trunks alternating PMD and PDL, is possible if the two PMD’s are weak, that is, in the telecom limit where the DGD’s $\delta \tau_x = b_k$ are much smaller than the width $t_f$ of the pulse; for conciseness, we write $\epsilon = \tau_x / t_f$. This means that $\bar{g}(\omega) = \bar{g}(\omega_0 + x)$ is significantly different from zero only for $|x| \leq \frac{1}{t_f}$, that is, $b_k x = O(\epsilon)$. So we can expand all the PMD operators (1) as

$$U(b \omega_0, \bar{m}) = \left[1 + i(b x/2)\sigma_m + O(\epsilon^2)\right]U(b \omega_0, \bar{m}).$$

Let us then calculate the three-trunk network:

$$T(x) = \mathcal{F} + i x \left(\frac{b_1}{2} \mathcal{F} \sigma_x + \frac{b_2}{2} \sigma_m \mathcal{F}\right) + O(\epsilon^2),$$

with $\mathcal{F} = U(b_2 \omega_0, \bar{m})FU(b_1 \omega_0, \bar{z})$. In what follows, we define the two orthogonal states of polarization $|\psi_F\rangle = \mathcal{F}|\psi\rangle/\sqrt{(\mathcal{F}^\dagger \mathcal{F})_{\phi_0}}$ and $|\psi_{\bar{F}}\rangle$, and we systematically omit global attentations. We have:

$$T(x)|\psi_0\rangle = (\psi_F[T]|\psi_0\rangle|\psi_F\rangle + (\psi_{\bar{F}}[T]|\psi_0\rangle|\psi_{\bar{F}}\rangle$$

$$= (1 + i x W)|\psi_F\rangle + x D|\psi_{\bar{F}}\rangle + O(\epsilon^2),$$

where $W = \langle \psi_0|\mathcal{F}^\dagger (\frac{b_1}{2} \mathcal{F} \sigma_x + \frac{b_2}{2} \sigma_m \mathcal{F})|\psi_0\rangle/(\mathcal{F}^\dagger \mathcal{F})_{\phi_0}$ and $x D \sim O(\epsilon)$. The passage from the Fourier to the time domain yields

$$|\Psi_3\rangle = \int dx e^{-i(x + \omega_0) t} \bar{g}(x) \otimes T(x)|\psi_0\rangle$$

$$= g(t - \text{Re}(W))e^{-i\omega_0 t} \otimes |\psi_F\rangle + h(t) \otimes |\psi_{\bar{F}}\rangle,$$

where we used $1 + i x W = e^{i x W} + O(\epsilon^2)$ and where $h(t) \sim O(\epsilon)$. The measurement of the intensity of the light pulse $|\Psi_3\rangle$ gives $I(t) \approx G(t - \text{Re}(W)) + O(\epsilon^2)$: the center of the pulse is now in

$$\langle \hat{t} \rangle = \text{Re}(W) = \frac{b_1}{2} w_1 + \frac{b_2}{2} w_2,$$

with $w_1$ given by (17) and $w_2 = \langle \psi_F|\sigma_m|\psi_F\rangle$. This result is intuitively clear: the first term is the weak value obtained by forgetting the second PMD element; the second term is just the mean value of $\sigma_m$ on the filtered state obtained by forgetting the first PMD element. For the case of any network composed of $2N + 1$ trunks alternating PMD and PDL elements, the result generalizes immediately as $\langle \hat{t} \rangle = \sum_k (\delta \tau_k/2) w_k$, with $w_k$ the suitable weak values [13]. This example shows how the formalism of weak measurements simplifies some calculations of networks combining PMD and PDL, adding an intuitive meaning to the formulas.

In conclusion, we have shown that the quantum theoretical formalism of weak measurements and postselection, often thought of as a weirdness of theorists, describes important effects in the physics of telecom fibers. In particular, the notion of postselection appears naturally, since the telecom engineers select only those photons that are not lost in the fiber.

Just a final remark, to say that, with this investigation, we close a loop of analogies. On the one hand, in Ref. [14], Gisin and Go stressed the analogy between the PMD-PDL effects in optical networks and the mixing and decay of kaons. On the other hand, in Ref. [15] it was shown that adiabatic measurements in metastable systems are a kind of weak measurement, and point out that kaons provide experimental examples of this. By showing the link between PMD-PDL and weak measurements with postselection, this work closes the loop.

Note added in proof.—For a related independent work, see [16].

*Present address: Institut de Ciències Fotòniques, Jordi Girona 29, 08034 Barcelona, Spain.


[9] The “time” $t$ in our equations is not the evolution parameter alone, but should rather be $t - n\zeta/c$, where $\zeta$ is the position in the fiber and $c/n$ is the average light speed in the fiber. That is why quantum theorists can well consider $t$ as a “position” and $\omega$ as its conjugate “momentum.”

[10] Note that $U(b(\omega_0 + \delta \omega), \bar{z}) = U(b(\omega_0), \bar{z})U(b \delta \omega, \bar{z})$.

[11] Note that the PMD alone is a reversible operation: the simple fact of delaying one polarization mode is not a complete measurement—it is sometimes called a premeasurement.


Optical Telecom Networks as Weak Quantum Measurements with Postselection

Nicolas Brunner, Antonio Acín,* Daniel Collins, Nicolas Gisin, and Valerio Scarani

Group of Applied Physics, University of Geneva, 20, rue de l’Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland

(Received 17 June 2003; published 28 October 2003)

We show that weak measurements with postselection, proposed in the context of the quantum theory of measurement, naturally appear in the everyday physics of fiber optics telecom networks through polarization-mode dispersion (PMD) and polarization-dependent losses (PDL). Specifically, the PMD leads to a time-resolved discrimination of polarization; the postselection is done in the most natural way: one postselects those photons that have not been lost because of the PDL. The quantum formalism is shown to simplify the calculation of optical networks in the telecom limit of weak PMD.

DOI: 10.1103/PhysRevLett.91.180402 PACS numbers: 03.65.Ta, 42.81.Gs

Several times in the history of science, different people working on different fields and with different motivations happened to discover the same thing, or to introduce the same concepts. Think of the connection between differential geometry and general relativity: physics received a convenient mathematical tool for its predictions, mathematics gained in popularity and interest because, apart from its intrinsic beauty, it proved useful.

In this paper, we point out a connection which should help to bring together two very different communities: quantum theorists and telecom engineers. The physical degree of freedom that supports this connection is the polarization of light; we show that the quantum formalism of weak measurements and postselection [1–3] applies to the description of polarization effects in optical networks [4]. The structure of this Letter is as follows: we give first a qualitative description of the announced connection. Then, we introduce the mathematical formalism, and show that the connection does indeed hold down to the detailed formulas; in particular, the knowledge of the “quantum” formalism can simplify some “telecom” calculations.

A modern optical network is composed of different devices connected through optical fibers. With respect to polarization, two main physical effects are present. The first one is polarization-mode dispersion (PMD): due to birefringency, different polarization modes (P modes in the following) propagate with different velocities; in particular, the fastest and the slowest polarization modes are orthogonal. PMD is the most important polarization effect in the fibers. The second effect is polarization-dependent loss (PDL), that is, different P modes are differently attenuated. PDL is negligible in fibers, but is important in devices such as amplifiers, wavelength-division multiplexing couplers, isolators, circulators, etc. In particular, a perfect polarizer is an element with infinite PDL, since it attenuates completely a P mode. Thus, an optical network can be described by a concatenation of trunks, alternating PMD and PDL elements. Combined effects of PMD and PDL elements have been studied in Refs. [5,6]; in particular, interesting phenomena such as anomalous dispersion have been shown to arise even in simple concatenations; namely, a PDL element sandwiched between two PMD elements.

The first piece of the connection we want to point out is the following: a PMD element performs a measurement of polarization on light pulses (Fig. 1). In fact, PMD leads to the separation of two orthogonal P modes in time; this separation is called differential group delay (DGD), noted $\delta \tau$. If $\delta \tau$ is larger than the pulse width, the measurement of the time of arrival is equivalent to the measurement of polarization—PMD acts then as a “temporal polarizing beam splitter.” However, in the usual telecom regime $\delta \tau$ is much smaller than the pulse width. In this case, the time of arrival does not achieve a complete discrimination between two orthogonal P modes anymore; but still, some information about the polarization of the input pulse is encoded in the modified temporal shape of the output pulse. We are in a regime of weak measurement of the polarization; we are going to show later that we recover indeed the notion of weak measurement of the quantum theorists, by measuring the mean time of arrival (that is, the “center of mass” of the output pulse).

The second piece of the connection defines the role of PDL: a PDL element performs a postselection of some polarization modes. Far from being an artificial ingredient, postselection of some modes is the most natural situation in the presence of losses: one does always postselect those photons that have not been lost. This would be trivial physics if the losses were independent of any degree of freedom, just like random scattering; but in the
case of PDL, the amount of losses depends on the meaningful degree of freedom, polarization. An infinite PDL, as we said above, would correspond to the postselection of a precise $P$ mode (a pure state, in the quantum language); a finite PDL corresponds to postselecting different $P$ modes with different probabilities (a mixed quantum state).

In summary: by tuning the PMD, we can move from weak to strong measurements of polarization; by tuning the PDL, we can study the postselection of a pure or of a mixed state of polarization. This is the main result of this Letter, that we are now going to present in mathematical terms.

It is convenient to use the formalism of two-dimensional Jones vectors, in which the description of classical polarization is identical to the quantum description of the spin $\frac{1}{2}$ [7]. Thus, e.g., the three typical pairs of orthogonal polarization — horizontal-vertical linear, diagonal linear, left-right circular — are described, respectively, by the eigenvectors of the Pauli matrices $\sigma_x$, $\sigma_y$, and $\sigma_z$. In this Letter, we shall only need to define the eigenvectors of $\sigma_z$: $|H\rangle = |H\rangle$, $|V\rangle = |V\rangle$. Any pure polarization state can be described as a superposition of these vectors, with complex coefficients, the state corresponding to the point $\hat{n} = (\theta, \phi)$ on the Poincaré sphere being $| + \hat{n}\rangle = \cos \frac{\theta}{2} |H\rangle + \sin \frac{\theta}{2} e^{i\phi} |V\rangle$.

On a monochromatic wave of frequency $\omega$, a PMD that separates the eigenvectors of $\sigma_z$ for a birefringency axis $b$ is represented by the operator [5]

$$\text{PMD:} U(b\omega, \hat{z}) = e^{i b \omega \sigma_z / 2} = \cos \frac{b \omega}{2} + i \sin \frac{b \omega}{2} \sigma_z.$$

(1)

This is a unitary operation that describes a global rotation of the state of polarization around the $z$ axis of the Poincaré sphere. As for PDL: since the most and least attenuated states are always orthogonal, they can be written as the eigenstates of $\sigma_x = \hat{n} \cdot \hat{\sigma}$, where the direction $\hat{n}$ has a priori no link with the direction $\hat{z}$ of the birefringency axis. Neglecting a global attenuation, the PDL is represented by the operator [5]

$$\text{PDL:} F(\mu, \hat{n}) = e^{i \mu \sigma_x / 2} = \cos \frac{\mu}{2} + i \sin \frac{\mu}{2} \sigma_x.$$

(2)

This is a nonunitary operator, sometimes called a filter; in the quantum theory, it appears also in the unambiguous discrimination of nonorthogonal quantum states [8]. It has been shown in Ref. [5] that any optical network can be modeled by an effective PDL followed by an effective PDL, that is, by an operator of the form $F(\mu, \hat{n}) U(b, \hat{m})$. However, the study of the general case is involved because the effective parameters $\mu$, $\hat{n}$, $b$, and $\hat{m}$ depend of the optical frequency $\omega$ in a nontrivial way, leading to deformations in the shape of the light pulse. Thus, we focus initially on the simplest optical network, namely a PDL fiber followed by a PDL element.

The input state is a Gaussian (Fourier-transform limited) light pulse of coherence time $t_c$, of central frequency $\omega_0$, prepared in a pure polarization state $|\psi_0\rangle$:

$$|\Psi_{\text{in}}\rangle = A e^{-(i/4)(t/t_c)^2} e^{-i \omega_0 t} \otimes (\alpha |H\rangle + \beta |V\rangle)$$

$$= g(t) \otimes |\psi_0\rangle,$$

(3)

with $A = (\sqrt{2\pi} t_c)^{-1/2}$ so that $G(t) = |g(t)|^2$ is a probability distribution [9]. To compute the state of the light at the output of the PMD fiber, we must Fourier transform $|\Psi_{\text{in}}\rangle$ into the frequency domain, apply (1) to any monochromatic component, and integrate back to the time domain. This gives [10]

$$|\Psi_{\text{PMD}}\rangle = \int d\omega e^{-i \omega t} \tilde{g}(\omega - \omega_0) U(b\omega, \hat{z}) |\psi_0\rangle$$

$$= \tilde{\alpha} g_-(t) |H\rangle + \tilde{\beta} g_+(t) |V\rangle,$$

(4)

where $g_{\pm}(t) = g[t \pm (\delta \tau / 2)]$ with $\delta \tau = b, \tilde{\alpha} = \alpha e^{ib\omega_0 / 2}$ and $\tilde{\beta} = \beta e^{-ib\omega_0 / 2}$. We see that, in addition to the global rotation around the birefringency axis at the frequency $\omega_0$, the PMD has delayed the $V$ polarization with respect to the $H$ polarization, as announced. According to whether the delay $\delta \tau$ is much larger or much smaller than the width $t_c$ of the input pulse, the recording of the time of arrival will provide us with a strong or a weak measurement [11]. For further reference, let us define the polarization state

$$|\psi\rangle = U(b\omega_0, \hat{z}) |\psi_0\rangle = \tilde{\alpha} |H\rangle + \tilde{\beta} |V\rangle$$

(5)

obtained by retaining only the global rotation, that is, in the limit of continuous light $\delta \tau / t_c = 0$.

Now, we should apply the PDL operator (2) to $|\Psi_{\text{PMD}}\rangle$. Before presenting the general case, to become familiar with the concepts, we study the case of postselection of a pure state: the PDL element is then a polarizer that projects onto a polarization state $|\psi_1\rangle = \mu |H\rangle + \nu |V\rangle$. Thus, at the output of the optical network we have

$$|\Psi_{\text{out}}\rangle = [\tilde{\alpha} \tilde{\mu} g_-(t) + \tilde{\beta} \tilde{\nu} g_+(t)] |\psi_1\rangle = f(t) |\psi_1\rangle,$$

(6)

where $\tilde{\nu}$ is the conjugate of a complex number $\nu$. Clearly $f(t)$ is the temporal shape of the selected component of the field. Now we measure the intensity $I(t) = |f(t)|^2$; with $A = \tilde{\alpha} \tilde{\mu}$ and $B = \tilde{\beta} \tilde{\nu}$, we have

$$I(t) = |A|^2 G_- (t) + |B|^2 G_+ (t) + 2 \text{Re}(\bar{A}B) \tilde{g}_-(t) g_+(t).$$

(7)

In the limit of strong measurement, $\delta \tau \gg t_c$, the overlap $\tilde{g}_- \cdot g_+$ is essentially 0, so the detected intensity corresponds to two well-separated Gaussians: $I(t) = |\alpha\mu|^2 G_- (t) + |\beta\nu|^2 G_+ (t)$. A detection in $G_-$ corresponds to the $H$ polarization, so the probability that the polarization was $|H\rangle$ given the preparation and postselection is simply the integral of the Gaussian $G_-$, normalized to the total intensity: $P(H) = \int_0^\infty I(t) dt / \int_{-\infty}^\infty I(t) dt = \left[ |\alpha\mu|^2 + |\beta\nu|^2 \right][t|\alpha|^2 + |\beta|^2]]$. But $|\alpha|^2$ is the probability $P(H)|\psi_0\rangle$ of finding a photon polarized
along $|H\rangle$ given that the state is $|\psi_0\rangle$; using similar notations for $|\beta|^2$, $|\mu|^2$, and $|\nu|^2$, we have found
\[ P(H) = \frac{P(\psi_1|H)P(H|\psi_0)}{\sum_{K=H,N} P(\psi_1|K)P(K|\psi_0)}. \tag{8} \]

This is the Aharonov-Bergmann-Lebowitz (ABL) rule [12], which corresponds to the classical rule for the probability of sequential events.

Since we have access to both $P(H)$ and $P(V)$, we can compute $\langle \sigma_\tau \rangle = P(H) - P(V)$. Moreover, the mean time of arrival, defined as usual by $\langle t \rangle = \int t P(t) dt$, is here $P(H)(\delta \tau/2) + P(V)(-\delta \tau/2)$. So, for the case of strong measurement, we have derived the relation
\[ \langle t \rangle = \frac{\delta \tau}{2} \langle \sigma_\tau \rangle. \tag{9} \]

This is the relation that appears in any measurement theory between the pointer or meter (here, the mean time of arrival) and physical quantity to be measured (here, $\sigma_\tau$). Even though it has been derived from more intuitive grounds in the regime of strong measurements, the relation (9) is the fundamental relation of a measurement process in which the coupling between the pointer and the observable quantity is made by the PMD [11]. In particular, contrary to $P(H)$ and $P(V)$, $\langle t \rangle$ can be defined and measured for any $P(t)$. We shall then take (9) as the definition of the mean value of $\sigma_\tau$ when measured by the PMD. With this, we can remove the assumption of strong measurement.

Starting with $P(t)$ given by (7), $\langle t \rangle$ can be calculated analytically in a straightforward way, and the relation (9) yield
\[ \langle \sigma_\tau \rangle = \frac{|A|^2 - |B|^2}{|A|^2 + |B|^2 + 2Re(AB) e^{-(1/2)(\delta \tau/2t)\gamma}}. \tag{10} \]

Note that the dependence in the strength of the measurement (i.e., in $\delta \tau/2t_\tau$) is very explicit in (10). In the limit of strong measurement, $\delta \tau/2t_\tau \to \infty$, we recover the above results. In the opposite limit, $e^{-(1/2)(\delta \tau/2t_\tau)} = 1 - O(\delta \tau/2t_\tau)$, corresponding to a weak measurement, we have $\langle \sigma_\tau \rangle = Re(A - B)(A + B)$. Noticing that
\[ A \pm B = \tilde{\alpha} \tilde{\mu} \pm \tilde{\nu} \tilde{\nu} = \langle \psi_1 \rangle T \langle \psi_1 \rangle, \tag{11} \]
with $|\psi_1\rangle$ given in (5), we find
\[ \langle \sigma_\tau \rangle = Re\left(\frac{\langle \psi_1 \rangle T \langle \sigma_\tau \rangle}{\langle \psi_1 \rangle T \langle \psi_1 \rangle}\right). \tag{12} \]

This is exactly the formula for the weak value of $\sigma_\tau$ when the postselection is done on a pure state $|\psi_1\rangle$ as given by the quantum theorists [1,2]. Note, in particular, that $\langle \sigma_\tau \rangle$ can reach arbitrarily large values, leading to an apparently paradoxical situation since the eigenvalues of $\sigma_\tau$ are $\pm 1$. But there is no paradox at all: $\langle \sigma_\tau \rangle > 1$ simply means $\langle t \rangle > (\delta \tau/2)$, and this situation is reached by postselecting a state $|\psi_1\rangle$ that is almost orthogonal to $|\psi\rangle$; these are very rare events, the shape $f(t)$ of the pulse is strongly distorted, and it is not astonishing that its center of mass could be found far away from its expected position in the absence of postselection.

We can now examine the case of a finite value of the PDL after the PMD fiber. For conciseness, we write $F(\mu, \hat{\alpha}) = F$ for the PDL operator (2). At the output of the PMD-PDL trunk, the state is
\[ |\Psi_{out}\rangle = F|\Psi_{PMD}\rangle = A(t)|H\rangle + B(t)|V\rangle, \tag{13} \]
where
\[ A(t) = \langle H|F|H\rangle \tilde{a} g_- - \langle H|F|V\rangle \tilde{b} g_+ + (C + n_z) \tilde{a} g_- - n_z \tilde{b} g_+, \tag{14} \]
\[ B(t) = \langle V|F|V\rangle \tilde{b} g_+ + \langle V|F|H\rangle \tilde{a} g_- - (C - n_z) \tilde{b} g_+ + n_z \tilde{a} g_. \tag{15} \]

Again, in the limit of weak measurement and using (9), we find
\[ \langle \sigma_\tau \rangle = \frac{\langle \sigma_\tau \rangle + \gamma n_z}{1 + \gamma n_z}, \tag{17} \]
with $|\psi\rangle$ given by (5) as before. The right-hand side is the weak value obtained by postselection on the mixed state $\rho = [1/Tr(F^\dagger F)] F^\dagger F$ [2,3]. The limiting case $\gamma = 0$ means $\mu = 0$, thence $\rho = \frac{1}{2} (I + \sigma_\tau)$; if there is no PDL, $\langle \sigma_\tau \rangle$ is as it should. At the other extreme, $\gamma = 1$ means $\mu \to \infty$ thence $\rho = \frac{1}{2} (I + \sigma_\tau)$, and we recover the formula (12) for the postselection of the pure state $|\psi_1\rangle$. Finally, we stress that the principal states of polarization of the PMD-PDL network, as defined, e.g., in Ref. [5], are $F(H)$ and $F(V)$ [13].

We have then demonstrated our claims: an optical PMD-PDL network is an everyday realization of the abstract notions of weak measurement and postselection introduced in the theory of quantum measurement. We had also said that telecom engineers would benefit by learning some quantum formalism, were it only because it could simplify their calculations. Indeed, consider a more complicated optical network, composed of three trunks: PMD-PDL-PMD, represented by the operator $T = U(b_2, \omega, \hat{m}) F(\mu, \hat{\alpha}) U(b_1, \omega, \hat{\beta})$. As we noticed above, this simple network is sufficiently complex to yield anomalous dispersion. The calculation can of course be done following the same steps as above, but it is heavy and not really instructive. Another approach, that is
Let us then calculate the three-trunk network:

\[ U(b, \omega, \dot{m}) = [1 + i(b/2 \sigma_m + O(e^2))] U(b, \omega_0, \dot{m}). \]  

(18)

Let us then calculate the three-trunk network:

\[ T(x) \approx F + i x \left( \frac{b_1}{2} \mathcal{F} \sigma_x + \frac{b_2}{2} \sigma_m \mathcal{F} \right) + O(e^2), \]  

(19)

with \( \mathcal{F} = U(b_2, \omega_0, \dot{m}) F U(b_1, \omega_0, \dot{z}) \). In what follows, we define the two orthogonal states of polarization \( |\psi_F \rangle = \mathcal{F}^\dagger |\psi \rangle / \sqrt{\langle \mathcal{F}^\dagger \mathcal{F} \rangle_{\theta_0}} \), and \( |\psi_F^\perp \rangle = \mathcal{F}^\dagger |\psi \rangle / \sqrt{\langle \mathcal{F}^\dagger \mathcal{F} \rangle_{\theta_0}} \), and we systematically omit global attenuations. We have:

\[ T(x)|\psi_0 \rangle = (\langle \psi_F | T |\psi_0 \rangle |\psi_F \rangle + \langle \psi_F^\perp | T |\psi_0 \rangle |\psi_F^\perp \rangle) \]  

(20)

where \( W = \langle \psi_0 | \mathcal{F}^\dagger (\frac{b_1}{2} \mathcal{F} \sigma_x + \frac{b_2}{2} \sigma_m \mathcal{F}) |\psi_0 \rangle / \langle \mathcal{F}^\dagger \mathcal{F} \rangle_{\theta_0} \) and \( x D \sim O(e) \). The passage from the Fourier to the time domain yields

\[ |\Psi_3 \rangle = \int dx e^{-i(x + \omega t)} \tilde{g}(x) \otimes T(x)|\psi_0 \rangle \]  

(21)

where we used \( 1 + i x W = e^{ixW} + O(e^2) \) and where \( h(t) \sim O(e) \). The measurement of the intensity of the light pulse \( |\Psi_3 \rangle \) gives \( I(t) \propto G(t - Re(W)) + O(e^2) \): the center of the pulse is now in

\[ \langle t \rangle = \text{Re}(W) = \frac{b_1}{2} w_1 + \frac{b_2}{2} w_2. \]  

(22)

with \( w_1 \) given by (17) and \( w_2 = \langle \psi_F | \sigma_m |\psi_F \rangle \). This result is intuitively clear: the first term is the weak value obtained by forgetting the second PMD element; the second term is just the mean value of \( \sigma_m \) on the filtered state obtained by forgetting the first PMD element. For the case of any network composed of \( 2N + 1 \) trunks alternating PMD and PDL elements, the result generalizes immediately as \( \langle t \rangle = \sum_k (\delta \tau_k/2) w_k \), with \( w_k \) the suitable weak values [13]. This example shows how the formalism of weak measurements simplifies some calculations of networks combining PMD and PDL, adding an intuitive meaning to the formulas.

In conclusion, we have shown that the quantum theoretical formalism of weak measurements and post-selection, often thought of as a weirdness of theorists, describes important effects in the physics of telecom fibers. In particular, the notion of postselection appears naturally, since the telecom engineers select only those photons that are not lost in the fiber.

Just a final remark, to say that, with this investigation, we close a loop of analogies. On the one hand, in Ref. [14], Gisin and Go stressed the analogy between the PMD-PDL effects in optical networks and the mixing and decay of kaons. On the other hand, in Ref. [15] it was shown that adiabatic measurements in metastable systems are a kind of weak measurement, and point out that kaons provide experimental examples of this. By showing the link between PMD-PDL and weak measurements with post-selection, this work closes the loop.

*Note added in proof.—For a related independent work, see [16].

[9] The “time” \( t \) in our equations is not the evolution parameter alone, but should rather be \( t - nz/c \), where \( z \) is the position in the fiber and \( c/n \) is the average light speed in the fiber. That is why quantum theorists can well consider \( t \) as a “position” and \( \omega \) as its conjugate “momentum.”
[10] Note that \( U(b, \omega_0 + \delta \omega, \dot{z}) = U(b, \omega_0, \dot{z}) U(b \delta \omega, \dot{z}) \).
[11] Note that the PMD alone is a reversible operation: the simple fact of delaying one polarization mode is not a complete measurement—it is sometimes called a premeasurement.
Quantum Teleportation with a Three-Bell-State Analyzer

J. A. W. van Houwelingen,* N. Brunner, A. Beveratos, H. Zbinden, and N. Gisin

Group of Applied Physics, University of Geneva, Rue de l’École-de-Médecine 20 CH-1211 Genève 4, Switzerland

(Received 5 October 2005; published 7 April 2006)

We present a novel Bell-state analyzer (BSA) for time-bin qubits allowing the detection of three out of four Bell states with linear optics, two detectors, and no auxiliary photons. The theoretical success rate of this scheme is 50%. Our new BSA demonstrates the power of generalized quantum measurements, known as positive operator valued measurements. A teleportation experiment was performed to demonstrate its functionality. We also present a teleportation experiment with a fidelity larger than the cloning limit.

DOI: 10.1103/PhysRevLett.96.130502 PACS numbers: 03.67.Hk, 03.65.Ud, 42.50.Dv, 42.81.–i

A Bell-state analyzer (BSA) is an essential part of quantum communications protocols such as a quantum relay based on quantum teleportation [1–7], entanglement swapping [8,9], or quantum dense coding [10,11]. It has been shown that, using only linear optics, a BSA for qubits has a maximal success rate of 50% when no auxiliary photons are used [12,13]. This, however, does not limit the number of Bell states one can measure, but only the overall success rate. A complete BSA could be achieved using either nonlinear optics [14] or using continuous variable encoding [15,16]. However, each of these two alternatives carry some significant drawback. The nonlinear optics approach has exceedingly low efficiency; while continuous variable encoding has the disadvantage that postselection is not possible. Note that postselection is a very useful technique that allows one to use only “good” measurement results and straightforwardly eliminate all others without the need for a lot of computing power [17].

Today’s optimal BSA schemes based on linear optics for qubits are only able to detect two out of four Bell states [2,6,8,9], or are able to detect more than two states but not optimally [18]. Here we present a novel scheme for a BSA which achieves the 50% upper bound of success rate, but can distinguish three out of the four Bell states. We demonstrate this scheme in a quantum teleportation experiment at telecom wavelengths. The new BSA is inspired by, although not limited to, the time-bin implementation of qubits, and is thus fully compatible with the field of quantum communications [19].

At first, one may think that detecting 3 out of 4 Bell states provides a full BSA. Indeed, if the BSA would consist of a standard von Neumann projective measurement, then the fourth Bell state would merely correspond to the nondetection of the 3 others. But our BSA is a new example of the power of generalized quantum measurements, it uses a positive operator valued measurement (POVM) with 21 possible outcomes. Some outcomes of this POVM (see Fig. 1) correspond to one of the 3 Bell states that can be distinguished unambiguously and thus detect this state. The others correspond to inconclusive results. More specifically, the Bell state $|\psi_+\rangle$ is always detected, $|\phi_-\rangle$ is never detected, while $|\psi_-\rangle$ and $|\phi_+\rangle$ are detected with a 50% success rate.

In previous BSAs, the main method was to use a beam splitter followed by detectors to determine the input Bell state. We replace this standard approach by a time-bin interferometer equivalent to the ones used to encode and decode time-bin qubits (Fig. 1) [19]. The two photons of the Bell state enter in port $a$ and $b$, respectively. In a BSA using only a beam splitter, one is able to distinguish between $\psi_\pm$ but not $\phi_\pm$, since in the later case the two photons will experience photon bunching but this interference does not contain the phase information that distinguishes $\phi_-$ from $\phi_+$. In our case, the first beam splitter acts like above, but we introduce a second interference possibility between the two photons on the second beam splitter. This second interference allows one to distinguish more than just two Bell states. The input modes $a$ and $b$ evolve in the interferometer as follows (see Fig. 1):

\begin{align}
\hat{a}_i^\dagger &= \frac{1}{\sqrt{4}} (-\hat{e}_i^\dagger + \hat{e}_i^\dagger \hat{f}^{\dagger}_{t+\Delta t} + i \hat{f}_i^\dagger + i e^{i\delta} \hat{f}^{\dagger}_{t+\Delta t}), \\
\hat{b}_i^\dagger &= \frac{1}{\sqrt{4}} (\hat{f}_i^\dagger - \hat{f}_i^\dagger \hat{f}^{\dagger}_{t+\Delta t} + i \hat{f}_i^\dagger + i e^{i\delta} \hat{f}^{\dagger}_{t+\Delta t}),
\end{align}

where $\hat{f}_j$ is a photon at time $j$ in mode $i$. Using these

FIG. 1. A schematic representation of the new type of Bell-state measurement. When two qubit states are sent into a time-bin interferometer the output state is a mixture of photons in two directional modes and three temporal modes. By looking at certain combinations of these photons a Bell-state measurement can be performed for three different Bell states.
TABLE I. The table shows the probability to find any of the 21 possible coincidences as a function of the input Bell-State. A 0 in row D1 means that a photon was found at detector “D1” and at a time corresponding to the photon having taken the short path in the interferometer and it was originally a photon in time-bin $t_0$, a 1 corresponds to $t_0 + 1 \times \tau$ with $\tau$ corresponding to the difference between the time-bins, etc., Note that several combinations of detection are possible for only one Bell-state (the bold entries), therefore when such a combination is found a Bell-state measurement was performed. The theoretical probability of a successful measurement is 0.5 which is the optimal value using only linear optics [12].

| $|\phi_+\rangle$ | $1/16$ | $1/16$ | $1/16$ | $1/16$ | $1/8$ | $1/8$ | $1/2$ |
| $|\phi_+\rangle$ | $1/16$ | $1/16$ | $1/4$ | $1/4$ | $1/4$ | $1/16$ | $1/8$ | $1/8$ |
| $|\psi_-\rangle$ | $1/8$ | $1/8$ | $1/8$ |
| $|\psi_-\rangle$ | $1/4$ | $1/4$ | $1/8$ | $1/8$ | $1/8$ |

Formulas it is possible to calculate all possible outputs of the interferometer as a function of any input state.

These output coincidences in ports e and f, i.e., on detectors D1 and D2, are summarized in Table I. By convention, a photon detected at time “0” means that the photon did not accumulate any delay with regards to a fixed reference. This is only possible if the photon took the short path in the BSA and it was originally a photon in time-bin $t_0$ (Fig. 1). A photon detected at time “1” signifies that the photon was originally in $t_1$ and took the short path of the BSA interferometer or it was in $t_0$ and took the long path. A detection at time “2” then means the photon was in $t_1$ and took the long path.

In Table I one can distinguish two cases. Either the result unambiguously distinguishes one Bell state. Or the result could have been caused by two specific Bell states, i.e., the result is ambiguous and hence inconclusive.

The above described approach is correct in the case were the phase $\delta$ of the BSA interferometer is set to 0. Let us thus briefly analyze in this paragraph the situation of an arbitrary phase $\delta$. In such a case, our BSA still distinguishes 3 Bell states, but these are no longer the Bell states of the computational basis. For a teleportation experiment this means the basis for the measured Bell states is not the same as the basis for the entangled states shared between Bob and Charlie. Still, perfect teleportation is possible, but with the difference that the unitary transformations that Bob has to apply after receiving the classical information about the result of the BSA have changed and no longer include the identity: all unitary transformations are non-trivial, but they remain experimentally feasible. More specifically, the analyzed Bell states are:

$$\phi'_\pm = |00\rangle \pm e^{i\delta}|11\rangle, \quad (3)$$

$$\psi'_\pm = e^{i\delta}(|01\rangle \pm |10\rangle). \quad (4)$$

These Bell states are equivalent to the standard states except that the $|1\rangle$ is replaced by $e^{i\delta}|1\rangle$ for each of the input modes. Therefore the unitary transformations that have to be applied to retrieve the original state of the teleported photon also have to be modified from $[I, \sigma_x, \sigma_y, \sigma_z]$ to $[\sigma_{2\delta}, \sigma_z, \sigma_{2\delta}, \sigma_x, \sigma_z, \sigma_{2\delta}]$. Here $\sigma_{2\delta} = e^{-i2\delta}P_{|1\rangle} + P_{|0\rangle}$ is a phase shift of $2\delta$ to be applied to the time-bin $|1\rangle$.

In a realistic experimental environment the success probabilities of the BSA will be affected by detector limitations, because existing photon detectors are not fast enough to distinguish photons which follow each other closely (in our case two photons separated by $\tau = 1.2 \text{ ns}$) in a single measurement cycle. Hence a coincidence, for example, “02” on D2, cannot be detected with our detectors. This limitation rises from the dead time of the photodetectors. When including this limitation we find that the maximal attainable probabilities of success in our experimental setup are reduced to 1/2, 1/4, and 1/2 for $\psi'_+, \psi'_-$, and $\phi'_+$, respectively. This leads to an overall probability of success of 5/16, which is greater by 25% than the success rate of 1/4 with a BSA consisting only of one beam splitter and two identical detectors.

In order to demonstrate successful Bell-state analysis we performed a teleportation experiment. A schematic of the experimental setup is shown in Fig. 2. Alice prepares a photon in the state $|\psi_A\rangle = |0\rangle + e^{i\alpha}|1\rangle$. Bob analyzes the teleported photon and measures interference fringes for each successful BSA announced by Charlie. The setup consist of a mode-locked Ti:sapphire laser creating 150 fs pulses with a spectral width of 4 nm, a central

FIG. 2 (color online). A rough overview of the experimental setup. The fiber interferometers shown here are in reality Michelson interferometers; for the interferometer in the BSA two circulators are used to have two separate inputs and outputs. Not shown in the figure is the method used for stabilizing the interferometers.
wavelength of 711 nm and a mean power of 400 mW. This beam is split in two beams using a variable coupler (\(\lambda/2\) and a PBS). The reflected light (Alice) is sent to a Lithium tri-Borate crystal (LBO, Crystal Laser) were by parametric down-conversion a pair of photons is created at 1.31 and 1.55 \(\mu m\). Pump light is suppressed with a Si filter, and the created photons are collected by a single mode optical fiber and separated with a wavelength-division multiplexer (WDM). The 1.55 \(\mu m\) photon is ignored whereas the 1.31 \(\mu m\) is send to a fiber interferometer which encodes the qubit on the equator of the Bloch sphere. In the same way, the transmitted beam (Bob) is send onto another LBO crystal after having passed through an unbalanced Michelson bulk optics interferometer, the phase of this interferometer is considered as the reference phase. The nondegenerate entangled photons produced in this way corresponds to the \(\phi_+\) state. The photons at 1.31 \(\mu m\) are send to Charlie in order to perform the Bell-state measurement. In order to assure temporal indistinguishability, Charlie filters the received photons down to a spectral width of 5 nm. In this way the coherence time of the generated photons is greater than that of the photons in the pump beam, and as such we can consider the photons to be emitted at the same time. Bob filters his 1.55 \(\mu m\) photon to 15 nm in order to avoid multiphoton events [20]. A liquid-nitrogen-cooled Ge avalanche photon detector (APD) D1 with passive quenching detects one of the two photons in the BSA and triggers the commercial infrared APDs (id Quantique) D2 and D3. Events are analyzed with a time to digital converter (TDC, Acam) in order to check the fidelity of the teleported state. The photons at \(\frac{3}{2}\) and \(\frac{5}{2}\) of the pump beam, and as such we can consider the photons to be emitted at the same time. Bob sends the teleported photon through an analyzing interferometer and measures the interference fringes conditioned on a successful BSA. When Bob scans the phase of his interferometer we obtained a raw visibility of \(V = 57%\) \((F = 79%)\) and a net visibility of \(V = 83\% \pm 4\) \((F = 91\% \pm 2)\) clearly higher than the cloning limit of \(F = 5/6\) (Fig. 3). We then switched to the new BSA. This new setup introduces about 3 dB of excess loss, due to added optical elements including the interferometer and its stabilization optics. These losses result in a lower count rate. For experimental reasons we now scan the interferometer of Alice instead of Bob. The experiments were performed for approximately 4.4 hours per point in order to accumulate enough data to have low statistical noise. The expected interference fringes after Bob’s interferometer are of the form \(1 \pm \cos(\alpha + \beta)\) for a projection on \(\psi_\pm\) and \(1 + \cos(\alpha - \beta - 2\delta)\) for a projection on \(\phi_+\). Hence one would

\[ \psi = \frac{1}{\sqrt{2}} (\psi_+ + \psi_-) \]

\[ \phi_+ = \cos(\theta) \psi_+ \sin(\phi) \]

\[ \phi_- = \cos(\theta) \psi_- \sin(\phi) \]

\[ \psi_\pm = \frac{1}{\sqrt{2}} (\psi_0 \pm \psi_2) \]

\[ \phi_\pm = \frac{1}{\sqrt{2}} (\phi_0 \pm \phi_2) \]

\[ \psi_\pm = \frac{1}{\sqrt{2}} (|\psi_0\rangle \pm |\psi_2\rangle) \]

\[ \phi_\pm = \frac{1}{\sqrt{2}} (|\phi_0\rangle \pm |\phi_2\rangle) \]

\[ |\psi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\psi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

\[ |\phi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\phi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

The temporal indistinguishability of the photons arriving at the BSA is usually tested by measuring a Hong-Ou-Mandel dip. In our BSA this is not directly possible. Photons that have bunched on the first interferometer will be split up by the second interferometer in a nondeterministic manner and as such there will be no decrease in the number of coincidences when looking at the same time of arrival. There will actually be an increase for these coincidences because the amount of photons that took a different path in the interferometer will decrease. In our experimental setup this means one has to look at an increase in the rate of detecting a “0” on both D1 and D2 and the same is true for the rate of “2” on D1 and D2. A typical result from this alignment procedure can be found in Fig. 3.

First we performed a quantum teleportation with independent units using a beam splitter for the Bell-state measurement. This enables us to test our setup in terms of fidelity. In order to check the fidelity of the teleported state, Bob sends the teleported photon through an analyzing interferometer and measures the interference fringes conditioned on a successful BSA. When Bob scans the phase of his interferometer we obtained a raw visibility of \(V = 57%\) \((F = 79%)\) and a net visibility of \(V = 83\% \pm 4\) \((F = 91\% \pm 2)\) clearly higher than the cloning limit of \(F = 5/6\) (Fig. 3). We then switched to the new BSA. This new setup introduces about 3 dB of excess loss, due to added optical elements including the interferometer and its stabilization optics. These losses result in a lower count rate. For experimental reasons we now scan the interferometer of Alice instead of Bob. The experiments were performed for approximately 4.4 hours per point in order to accumulate enough data to have low statistical noise. The expected interference fringes after Bob’s interferometer are of the form \(1 \pm \cos(\alpha + \beta)\) for a projection on \(\psi_\pm\) and \(1 + \cos(\alpha - \beta - 2\delta)\) for a projection on \(\phi_+\). Hence one would

\[ \psi = \frac{1}{\sqrt{2}} (\psi_+ + \psi_-) \]

\[ \phi_+ = \cos(\theta) \psi_+ \sin(\phi) \]

\[ \phi_- = \cos(\theta) \psi_- \sin(\phi) \]

\[ \psi_\pm = \frac{1}{\sqrt{2}} (|\psi_0\rangle \pm |\psi_2\rangle) \]

\[ \phi_\pm = \frac{1}{\sqrt{2}} (|\phi_0\rangle \pm |\phi_2\rangle) \]

\[ |\psi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\psi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

\[ |\phi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\phi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

\[ |\psi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\psi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

\[ |\phi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\phi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

\[ |\psi_\pm\rangle = \frac{1}{\sqrt{2}} (|\psi_0\rangle \pm |\psi_2\rangle) \]

\[ |\phi_\pm\rangle = \frac{1}{\sqrt{2}} (|\phi_0\rangle \pm |\phi_2\rangle) \]

\[ |\psi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\psi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]

\[ |\phi_0\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |2\rangle) \]

\[ |\phi_2\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |2\rangle) \]
expect to find three distinct curves, two with a phase difference of $\pi$, and the third dephased by $-2(\beta + \delta)$, where $\alpha$, $\beta$, and $\delta$ are the phases of Alice, Bob, and the BSA interferometer, respectively. Note that Bob is able to derive $\delta$ using the phase difference between the fringes made by $\psi_+$ and $\phi_+$ since this difference corresponds to $2(\delta + \beta)$, where $\beta$ is Bob’s local phase.

In Fig. 4 we show the coincidence interference fringes between Bob and a successful BSA. As expected fringes for $|\psi_-\rangle$ and $|\psi_+\rangle$ have a $\pi$ phase difference due to the phase flip caused by the teleportation. On the other hand, fringes for $|\psi_-\rangle$ and $|\phi_+\rangle$ are dephased by $-2(\beta + \delta)$, which in this case we had arranged to be approximately 0.

The count rates for the different Bell states differ because we correspond to different numbers of detection combinations. The raw (Fig. 4) visibilities obtained for the projection on each Bell state are $V_{\phi_-} = 0.38$, $V_{\phi_+} = 0.22$, $V_{\psi_-} = 0.43$, which leads to an overall value of $V = 0.34$ ($F = 0.67$). If we subtract the noise we find net visibilities of 0.51, 0.69, and 0.55 for $|\psi_-\rangle$, $|\phi_+\rangle$ and $|\phi_-\rangle$ which leads to an average of $V = 0.58$ ($F = 0.79$). Note that the maximal value that can be attained without the use of entanglement is $V_{\text{max}} = 1/3$ [21]. Two of the raw visibilities and all of the net visibilities break this limit. In order to check the dependence of $|\phi_-\rangle$ on $\delta$ we also performed a teleportation with a different value and we clearly observe the expected shift in the fringe (Fig. 5) while measuring similar visibilities.

Finally the authors would like to stress two points about this novel BSA. First it is possible to implement this BSA for polarization encoded photons by creating a second interference possibility for H and V polarization. This would require 4 detectors, but an overall efficiency of 50% can be achieved with current day detectors. The limit of 50% can also be achieved with time-bin encoded qubits but would require the use of ultra fast optical switches and two more detectors. We did not implement this due to the losses associated with introducing current day high speed integrated modulators. Second, even though three out of four Bell states can be distinguished, one cannot use this scheme in order to increase the limit of $\log_2 3$ bits per symbol for quantum dense coding.

In conclusion, we have shown experimentally that it is possible to perform a three-state Bell analysis while using only linear optics and without any actively controlled local operations on a single qubit. In principle this measurement can obtain a success rate of 50%. We have used this BSA to perform a teleportation experiment, and obtained a non-corrected overall fidelity of 67%, after noise substraction we find $F = 76\%$. Also, we performed a teleportation experiment with a one state BSA which exceeded the cloning limit.

The authors acknowledge financial support from the European IST-FET Project “RamboQ” and from the Swiss NCCR Project “Quantum Photonics” and thank C. Barreiro and J.-D. Gautier for technical support.
Experimental quantum teleportation with a three-Bell-state analyzer

J. A. W. van Houwelingen,* A. Beveratos, N. Brunner, N. Gisin, and H. Zbinden
Group of Applied Physics, University of Geneva, CH-1211 Geneve 4, Switzerland
(Received 28 April 2006; published 3 August 2006)

We present a Bell-state analyzer for time-bin qubits allowing the detection of three out of four Bell states with linear optics, two detectors, and no auxiliary photons. The theoretical success rate of this scheme is 50%. A teleportation experiment was performed to demonstrate its functionality. We also present a teleportation experiment with a fidelity larger than the cloning limit of $F = \frac{3}{5}$.

DOI: 10.1103/PhysRevA.74.022303

PACS number(s): 03.67.Hk, 42.50.Dv, 42.81.–i

I. INTRODUCTION

Bell-state analyzers (BSA) form an essential part of quantum communications protocols. Their uses range from quantum relays based on teleportation [1-7] or entanglement swapping [8,9] to quantum dense coding [10,11]. An important restriction for BSAs is that a system based on linear optics, without using auxiliary photons, is limited to a 50% overall success rate [12,13]. This important result does not restrict the number of Bell states that can be measured, but only the overall efficiency of a measurement. Nevertheless, a complete BSA is possible for at least two different cases: the first approach uses nonlinear optics [14] but this has the drawback of an exceedingly low efficiency and is, therefore, not well adapted for quantum communication protocols. Another possibility is the use of continuous variable encoding [15,16], however, this technique has the disadvantage that postselection is not possible. Note that postselection is a very useful technique that allows one to use only “good” measurement results and straightforwardly eliminate all others without the need for great computational analysis.

Many experiments have been done up to date that use BSAs. In this paper a different BSA is introduced [17]. It has the maximum possible efficiency that can be obtained when using only linear optics without ancilla photons. It is different with respect to other BSAs since it can distinguish three out of the four Bell states. All of the used BSAs up to date that can reach the maximum efficiency, without the use of ancilla photons, are limited to two (or less) Bell states [2,6,8,9,18]. There have also been experiments of a BSA that detects all four Bell states but its overall efficiency does not reach 50% and it requires the use of an entangled ancilla photon pair [19].

II. THEORY

A. Time-bin encoding

In our experiments a qubit is encoded on photons using time bins [20]. This means that a photon is created that exists in a superposition of two well-defined instants in time (time bins) that have a fixed temporal separation of $\tau$. By convention the Fock state with $N=1$ corresponding to a photon in the early time of existence $t_0$ is written as $|0\rangle$ and for the later time $t_1 = t_0 + \tau$ as $|1\rangle$. Photons in such a state can be created in several ways. The simplest method is to pass a single photon through an unbalanced interferometer with a path length difference of $n c \tau$, where $n$ is the refractive index. After the interferometer the photon will be in the qubit state $A|0\rangle + e^{i\alpha}B|1\rangle$. Here $A$ and $B$ are amplitudes that depend on the characteristics of the interferometer and $\alpha$ is the phase difference between the interferometer paths which is directly determined by $\alpha = (2 \pi n c \tau / \lambda) \mod 2 \pi$. For the sake of readability we will use the word qubit when talking about a “photon that is in a qubit state.”

B. Bell-state analyzer

In a large part of all experiments using BSAs that have been performed up to date, the BSA consists essentially of a beamsplitter and single-photon detectors (SPDs). In such a beamsplitter-BSA(BS-BSA) the “clicks” of the SPDs are analyzed and, depending on their results, the input state will be projected onto a particular Bell state. With time-bin qubits as described above a simple BS-BSA works as follows: two qubits arrive at the same time on a beamsplitter but at different entry ports. Since the four standard Bell states

\[ |\phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \]
\[ |\phi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \]

form a complete basis we can write our two-qubit input state as a superposition of these four states. One can calculate for each Bell input state the possible output states. These states can then be detected using SPDs. The different detection patterns and their probabilities are shown in Table I. By convention, a detection click at time “0” (“1”) means that the photon was detected in the time bin $t_0$ ($t_1$). The output combinations show that, if one detects two photons in the same path but in a different time bin, the input state could only have been caused by the Bell state $|\psi_\pm\rangle$ and therefore the overall state of the system is projected onto this state. When the photons arrive at different detectors with a time-bin difference the input state is projected onto the state $|\psi_\pm\rangle$. However, when one measures two photons in the same time bin in the same detector the state could either be $|\phi_\pm\rangle$ or $|\phi_\mp\rangle$, and therefore the state has not been projected onto a single Bell

*Email address: jeroen.vanhouwelingen@physics.unige.ch
TABLE I. The table shows the probability of finding specific coincidences as a function of the input Bell state in the case of a single beamsplitter as a BSA. A “0” (“1”) in row D1 means that a photon was found at detector “D1” at time $t_0$ ($t_1$), etc. Note that only half of the combinations of detection are possible for only one Bell state (the bold entries), therefore when such a combination is found a projection onto this Bell state was performed. The theoretical success probability is 50%.

<table>
<thead>
<tr>
<th></th>
<th>D1</th>
<th></th>
<th>D2</th>
<th></th>
<th>D2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>00</td>
<td>11</td>
<td>01</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>$</td>
<td>\phi_0\rangle$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>$</td>
<td>\phi_-\rangle$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>$</td>
<td>\psi_+\rangle$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>

The two remaining Bell states

state but onto a superposition of two Bell States. This method has a success rate of 50% which corresponds to the maximal possible success rate that can be obtained while using only linear optics and no auxiliary photons [12].

Here we propose a BSA which is capable of distinguishing more than two Bell states while still having the maximum success rate of 50%. This is possible by replacing the beamsplitter with a time-bin interferometer equivalent to the ones used to encode and decode time-bin qubits (Fig. 1). This BSA will be capable of distinguishing three out of four Bell states, but $|\phi_0\rangle$ and $|\psi_\pm\rangle$ will only be discriminated 50% of the time as will be explained shortly. Two qubits enter in port $a$ and $b$, respectively. The first beamsplitter acts like above, allowing the distinction of two Bell states ($|\phi_0\rangle$ and $|\psi_\pm\rangle$). A second possibility for interference is added by another BS for which the inputs are the outputs of the first BS, with one path having a delay corresponding to the time-bin separation $\tau$. The two-photon effects on this beamsplitter lead to fully distinguishable photon combinations of one of the two remaining Bell states ($|\phi_\pm\rangle$) while still allowing a partial distinction of the first two.

One might expect that when it is possible to measure three out of four states that the fourth, nonmeasured, state can simply be inferred from a negative measurement result of the three measurable states. This is, however, not the case. The above described measurement is a positive operator valued measure (POVM) with 21 possible outcomes, some of these outcomes are only possible for one of the four input Bell states. Therefore, when such an outcome is detected it unambiguously discriminates the corresponding input Bell state. The rest of the 21 outcomes correspond to outcomes which can result from more than one input Bell state. In other words, their results are ambiguous and the input state is not projected onto a single Bell state but onto a superposition of Bell states.

The state after the interferometer can be calculated for any input state using

$$\hat{a}^\dagger(t) \Rightarrow \frac{1}{\sqrt{4}}[-\hat{e}^\dagger(t) + e^{i\delta}\hat{e}^\dagger(t + \tau) + i\hat{f}^\dagger(t) + i\hat{e}^{i\delta}\hat{f}^\dagger(t + \tau)],$$

(3)

$$\hat{b}^\dagger(t) \Rightarrow \frac{1}{\sqrt{4}}[\hat{f}^\dagger(t) - e^{i\delta}\hat{f}^\dagger(t + \tau) + i\hat{e}^\dagger(t) + i\hat{e}^{i\delta}\hat{e}^\dagger(t + \tau)],$$

(4)

where $\hat{a}^\dagger(j)$ is the creation operator of a photon at time $j$ in mode $i$. When the input states are qubits and the photons are detected after the interferometers the detection patterns are readily calculated and are shown in Table II. The output coincidences on detectors $D1$ (port $e$) and $D2$ (port $f$) are shown as a function of a Bell state as input. By convention, a detection at time “0” means that the photon was in time $t_0$ after the BSA interferometer. This is only possible if it took the short path in the BSA and it was originally a photon in time bin $t_0$ (Fig. 1). Similarly a detection at time “1” means that either the photon was originally in $t_1$ and took the short path of the BSA interferometer or it was in $t_0$ and took the long path. A detection at time “2” means the photon was in $t_1$ and took the long path. In Table II we see that some of the patterns corresponds to a single Bell state and therefore the measurement is unambiguous. For the other cases the result could have been caused by two Bell states, i.e., the result is ambiguous and hence inconclusive. More specifically, the Bell state $|\psi_-\rangle$ is detected with probability 1, $|\phi_0\rangle$ is never detected, and both $|\psi_\pm\rangle$ and $|\phi_\pm\rangle$ are detected with probability $\frac{1}{2}$.

The above described approach is correct in the case where the separation $\tau$ of the incoming qubits is equal to the time-bin separation caused by the interferometer. If this is not the case and the interferometer creates a time-bin separation of $\tau + n\delta/(2\pi c)$, where $\delta$ is a phase, the situation is slightly more complicated In such a case, our BSA still distinguishes three Bell States, but these are no longer the standard Bell States but are the following:

$$|\phi_0\rangle = |00\rangle \pm e^{i\delta}|11\rangle = (\sigma_\delta \otimes \sigma_\delta)|\phi_0\rangle,$$

(5)

$$|\psi_-\rangle = e^{i\delta}|01\rangle \pm |10\rangle) = e^{i\delta}|\psi_-\rangle.$$  

(6)

Here $\sigma_\delta = P_{\{00\}} + e^{i\delta}P_{\{11\}}$ is a phase shift of $\delta$ to be applied to the time bin $|1\rangle$. These new Bell States are equivalent to the standard states except that the $|1\rangle$ is replaced by $e^{i\delta}|1\rangle$ for each of the input modes.

In a realistic experimental environment the success probabilities of the BSA are affected by detector limitations. This

FIG. 1. A schematic representation of the new type of Bell-state measurements. When two qubit states are sent into a time-bin interferometer the output state is a mixture of photons in two directional modes and three temporal modes. By looking at certain combinations of these photons a Bell-state measurement can be performed for three different Bell States.
is because existing photon detectors are not fast enough to distinguish photons which follow each other closely (in our case two photons separated by $\tau=1.2$ ns) in a single measurement cycle. This limitation rises from the dead time of the photodetectors. When including this limitation we find that the maximal probabilities of success in our experimental setup are reduced to $1/4$ respectively. This leads to an overall probability of success of $5\%$, which is greater by $25\%$ than the success rate of $1/32$ for a BSA consisting only of one beamsplitter and two detectors with the same limitation. This limitation could be partially eliminated by using a beamsplitter and two detectors in order to detect the state $50\%$ of the time, or it could be completely eliminated by using an ultrafast optical switch (sending each time bin to a different detector). Both of these methods are associated with a decrease in the signal-to-noise ratio. This is caused by the additional noise from the added detector and by additional losses from the optical switch, respectively.

C. Bell-state analyzer for polarization qubits

So far the discussion about this BSA only considered time-bin qubits. The authors would like to note at this point that it is also possible to implement a similar BSA for polarization encoded photons. This can be done by the equivalent polarization setup as shown in Fig. 2. This setup would require four detectors but there will never be two photons on one detector and therefore dead times do not hinder the measurement of all the detection patterns and the overall efficiency can reach $50\%$ with today’s technology.

D. Four-Bell-state analyzer?

This paper discusses our results testing a three-Bell-state analyzer. It is obviously interesting to also consider the possibility of a linear optics four-Bell-state analyzer with $50\%$ efficiency and no ancilla photons. Such a system was not used for the simple reason that there is no known method to make such a measurement. Is there a fundamental reason to suspect that such a BSA cannot be performed? No such reason is known to the authors, therefore this paper will be limited to the three-Bell-state analyzer.

E. Teleportation

One of the most stunning applications of a BSA is its use in the teleportation protocol. In order to perform a teleportation experiment an entangled qubit photon pair is created (EPR) as well as a qubit to be teleported (Alice). One photon of the entangled pair is made to interact with Alice’s qubit in a BSA (Charlie). This interaction followed by detection projects the overall state onto a Bell state (if the BSA is successful). The remaining photon (Bob) now carries the same information as the photon from Alice up to a unitary transformation. The situation for the new BSA is slightly different since the entangled pair is not a member of the detected Bell basis [Eqs. (5) and (6)]. However this has no major influence on the theory. After a successful measurement of the BSA the remaining photon at Bob is equal to the original qubit up to a unitary transformation. This transformation, however, has to be adapted with regards to the standard case from $[1, \sigma_+, \sigma_-, \sigma_z]$ to $[\sigma_z, \sigma_+ \sigma_z^{-1}, \sigma_+, \sigma_z \sigma_+]$, as can be seen from the following calculation:

\[
\lambda/2
\]

FIG. 2. A schematic representation of the new type of a Bell-state analyzer for polarization qubits. The gray cubes represent non-polarizing beamsplitters and the white cubes are polarizing beamsplitters.
Recall Fig. 3; the experiment is an adaptation of a setup used previously for long distance teleportation and discussed. Finally, the results of the experiment are presented in this section is the experimental setup that is the method used for stabilizing the interferometers.

\[
|\psi_{abc}\rangle = |\psi_1\rangle \otimes |\phi_4\rangle
\]

\[
= \frac{1}{2} (|\phi_1\rangle_{ab} \otimes \sigma_{z0}\sigma_z|\psi_1\rangle_c + |\phi_2\rangle_{ab} \otimes \sigma_z|\psi_1\rangle_c + |\psi_1\rangle_{ab} \otimes e^{-i\delta}\sigma_z|\psi_1\rangle_c + |\psi_2\rangle_{ab} \otimes e^{-i\delta}\sigma_z|\psi_1\rangle_c).
\]

Recall \(\sigma_{z0}^{-1} = P_{00} + e^{2i\delta} P_{11}\) is a phase shift of the bit 1).

III. EXPERIMENTAL TELEPORTATION

The BSA was tested in a quantum teleportation experiment. Presented in this section is the experimental setup that was used as well as some of the required preliminary alignment experiments. Finally, the results of the experiment are given and discussed.

A. Experimental setup

A rough schematic of the experimental setup is shown in Fig. 3; the experiment is an adaptation of a setup used previously for long distance teleportation and for entanglement swapping. Alice prepares a photon in the state \(|\psi_1\rangle = (1/\sqrt{2})(|0\rangle + e^{i\theta}|1\rangle\). A BSA is used by Charlie on Alice’s qubit combined with a part of an entangled qubit pair. Bob analyzes the other half of the pair (the teleported qubit) and measures interference fringes for each successful BSA announced by Charlie.

The setup consists of a mode-locked Ti:sapphire laser (MIRA with 8W VERDI pump-laser, Coherent) creating 150 fs pulses with a spectral width of 4 nm, a central wavelength of 711 nm, a mean power of 400 mW and a repetition rate of 80 MHz. This beam is split in two beams using a variable coupler (\(\lambda/2\) and a PBS). The reflected light (Alice) is sent to a scannable delay and afterwards to a lithium triborate crystal (LBO, crystal laser) where by parametric downconversion a pair of photons is created at 1.31 and 1.55 \(\mu\)m. Pump light is suppressed with a Si filter, and the created photons are collected by a single mode optical fiber and separated with a wavelength-division multiplexer (WDM). The 1.55 \(\mu\)m photon is ignored, whereas the photon at 1.31 \(\mu\)m is send to a fiber interferometer which encodes the qubit state \(|\psi_1\rangle\) onto the photon. The transmitted beam (Bob) is passed through an unbalanced Michelson-type bulk interferometer. The separation between the two time bins after this interferometer is considered as the reference for all the other separations. The phase of the interferometer can, therefore, be considered as a reference phase and can be defined as 0. After the interferometer the beam passes a different LBO crystal. The nondegenerate photon pairs created in this crystal are entangled and their state corresponds to \(|\phi_4\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)\).

The photons at 1.31 \(\mu\)m are send to Charlie in order to perform the Bell-state measurement. To assure temporal indistinguishability, Charlie filters the received photons down to a spectral width of 5 nm using a bulk interference filter. Because of this the coherence time of the generated photons is greater than that of the photons in the pump beam, and as such no distinguishability between photons can be caused by jitter in their creation time. Bob filters his 1.55 \(\mu\)m photon to 15 nm in order to avoid multi photon-pair events, this filtering is done by the WDM that separates the photons at 1.31 and 1.5 \(\mu\)m. This filter is larger than Charlie’s filter for experimental reasons. A liquid nitrogen cooled Ge avalanche-photon detector (APD) \(D_1\) with passive quenching detects one of the two photons in the BSA and triggers the InGaAs APDs (id Quantique) \(D_2\) and \(D_3\). Events are analyzed with a time to digital converter (TDC, Acam) and coincidences are recorded on a computer.

Each interferometer is stabilized in temperature and for greater stability an active feedback system adjusts the phase every 100 s using reference lasers. The reference for Bob’s interferometer is a laser (Dicos) stabilized on an atomic transition at 1531 nm and for both Alice’s and Charlie’s
interferometer a stabilized distributed-feedback (DFB) laser (Dicosa) at 1552 nm is used. It is possible to use different lasers for Alice and Charlie if one wants to create two independent units. By using independent interferometer units using different stabilization lasers it was assured that this experiment is ready for use “in the field.” A more detailed description of the active stabilization of the interferometers is given in Ref. [9]. For sake of clarity the interferometers shown in the setup (Fig. 3) are Mach-Zender-type interferometers but in reality they are Michelson interferometers which use Faraday mirrors in order to avoid distinguishability due to polarization differences [25].

B. Alignment experiments

There are two important, nontrivial alignments that have to be made before one can perform a quantum teleportation experiment with time-bin encoded qubits. First, one would have to assure that all the time-bin interferometers have the same difference in length between the two paths. Second, it is required that there is temporal indistinguishability between qubits coming from Alice and Bob on the BSA. The equalization of the interferometers is needed in order to assure that all the interferometers have a difference in length of \( cT/n \) with a precision higher than the coherence length of the photons \((\approx 150 \mu m)\). We have two mechanisms to actively change the optical path lengths: the first is changing the temperature of the interferometers and thus allowing the long arm to increase or decrease its length more than the short arm and the second is to directly change the length of only one arm by means of a cylindric piezoelectric element. When changing the voltage over the piezo we change the diameter of the cylinder and thus the length of the fiber changes. This is used for the active feedback stabilization system. In order to align the interferometers with each other we perform two different experiments: First, we optimize the visibility of single-photon interference fringes for photons from Alice detected in D1. This aligns Alice’s interferometer with the BSA interferometer. Next, we optimize a Fransson-type Bell test of the entangled photon pair [26]. While optimizing this experiment we do not change the BSA interferometer. This optimization aligns the bulk interferometer and Bob’s analysis interferometer to the other two. Using this method we found visibilities of 97% \( \pm 1\% \) for the single photon interference and 94% \( \pm 1\% \) for the Bell test (Fig. 4).

The second alignment procedure is necessary in order to assure temporal indistinguishability between the photons arriving at the BSA. In the case of a BS-BSA this can be assured by performing a Hong-Ou-Mandel dip type experiment [27], which is to say, make a scan in a delay for one of the incoming photons and look at a decrease in the number of coincidences as a result of photon bunching (Fig. 5). The position where the minimum is obtained corresponds to the point with maximal temporal overlap of the two photons.

In the case of an interferometer-BSA (IF-BSA) this procedure becomes more complicated. We can no longer look at a mandel dip because the second beamsplitter will probabilistically split up the photons that bunched on the first beamsplitter. However, the photon bunching remains and it can still be seen by a different method. Consider the situation where two single photons, both in the state \( |0\rangle\), are send to the different inputs of an IF-BSA (Fig. 6). If the photons are not temporally indistinguishable there are three possible output differences between detection times, corresponding to “10”, “00&11”, and “01”. If the photons are indistinguishable they bunch at the first BS and therefore the difference in arrival time between the photons has to be zero. This means that “10” and “01” are not possible anymore and the possibility for “00&11” is larger. If the inputs are arbitrary qubits instead of the simple example above there will be more coincidence possibilities and some of them will be subject to single-photon interference and/or photon bunching. It is possible to see an increase in the coincidences for “00” and “22”, which is not affected by a single-photon interference, for similar reasons as the increase that was explained above. These coincidences can be measured in a straightforward way with our setup. A more rigorous calculation and explanation of this alignment procedure is given in the Appendix. A typical result of an experiment in which the count rate is measured while changing a delay is shown in Fig. 5 and clearly shows the expected increase in count rate.

The measured antidips have a net visibility of 32\( \pm 3\% \) and 26\( \pm 2\% \) after noise substraction. The maximal attainable value is 1 due to undesired but unavoidable double-pair events (see appendix). The large visibilities mean that the temporal indistinguishability is very good, this will thus not
be limiting for our experiments. The noise substraction for this estimation is justified because in a teleportation experiment the noise will be reduced since one will consider only three-photon events. The difference in height of the two coincidences is related to an electronic loss of signal in an electrical delay line.

C. Experimental results

Two different types of teleportation experiments were performed. Namely a standard BS-BSA teleportation in order to benchmark our equipment followed by the IF-BSA experiment. For the BS-BSA, the main difference with regards to previous experiments [9,23] was that the interferometers now all had an active stabilization. This allows for large stability and long measurement times. The experiment consisted of Bob scanning of his interferometer phase while the other interferometers where kept constant, we therefore expect to find an interference curve of the form \(1 + V \sin(\beta + \alpha)\) where \(\alpha\) is kept constant. The results of the experiment (Fig. 7) clearly shows the expected behavior. The visibility measured was \(V=0.57\pm0.03\) \(F=0.79\pm0.02\). After conservative noise substraction we find \(V=0.83\pm0.04\) \(F=0.91\pm0.02\). This clearly is higher than the strictest threshold that has been associated with quantum teleportation of \(F=5/6\) [28,29]. The limiting factors of this experiment are the detectors and the fiber coupling after the LBO crystals.

After this experiment the setup was changed to the IF-BSA. The count rates in this experiment with regards to the previous one is reduced due to two reasons. The introduction of the BSA interferometer and its stabilization optics means an additional 3 dB of loss which reduces count rates. Another difference is that now the counts are distributed over three different Bell states, whereas before there was only one. Therefore an overall reduction of counts per state will occur. Combined these effects result in a large reduction of the count rate per Bell state. This problem was overcome by, on the one hand, an overall increase of the BSA efficiency by \(1/2\) (from 25% for the BS-BSA to 31.25% for the IF-BSA) and, on the other hand, by integrating data over longer time periods. During the teleportation experiment scans were made in the interferometer of Alice rather than Bob. This was done since the most important noise is dependent of the phase of Bob’s interferometer but not of Alice’s (more details are given in the next subsection). The experiments were performed with approximately 4.4 h per phase setting in order to have low statistical noise.

For this IF-BSA all the different unambiguous results (Table II) were analyzed both separately (for example, “02”) and combined as a Bell state (for example, \(|\psi_+\rangle=|02\rangle+|20\rangle\)). For the separate results it is expected that each BSA outcome will have count rates depending on the phases of the

FIG. 5. (Color online) Top, IF-BSA: Graph showing the number of measured coincidences as a function of a change in the delay line. Both “00” and “22” clearly show an antidip at the same location. The net visibilities are \(V_{00}=32\pm3\%\) and \(V_{22}=26\pm2\%\). Bottom, BS-BSA: Graph showing the decrease in the number of measured coincidences 00 as a function of a change in the delay line [27]. The net visibility is \(V=29\pm3\%\).

FIG. 6. The simple experiment on the left (one photon in each input of an IF-BSA) will have the following property. If the photons are not temporally indistinguishable one will find three different coincidence peaks: “10”, “00& 11”, and “01” (dotted curve), however, if the photons are indistinguishable there will be only one peak: 00 and 11 (plain curve). This is caused by photon bunching.

FIG. 7. (Color online) The result of the one-Bell state teleportation experiment (a beamsplitter instead of the interferometer) with \(F_{raw}=0.79\pm0.02\) and \(F_{net}=0.91\pm0.02\).
interferometers as $R[1 + V \cos(\alpha + \rho)]$. Here $R$ is dependent on the overall efficiency of the experiment and is different for each BSA outcome and $\rho$ is a combination of the constant phases of the interferometers of Bob and Charlie and is different for different BSA results:

$$|\psi_\alpha\rangle, |01\rangle, |10\rangle, |12\rangle, |21\rangle \rightarrow \rho = \beta, \quad (9)$$

$$|\phi_\alpha\rangle, |11\rangle \rightarrow \rho = -\beta - 2\delta, \quad (10)$$

$$|\psi_\gamma\rangle, |02\rangle, |20\rangle \rightarrow \rho = \beta + \pi. \quad (11)$$

As is evident from the differences in $\rho$ we expect that fringes corresponding to one particular Bell state are in phase with each other, but have a well-determined phase difference with fringes corresponding to another Bell state.

The measured count rates as a function of the phase of Alice’s interferometer are shown in (Fig. 8). Note that, due to experimental restrictions, the absolute phases of the interferometers are not known and therefore all phase values have an unknown offset. The results clearly show that each of the outcomes has the expected interference behavior. Furthermore, the fringe corresponding to “01” is in phase with the fringe “21”. The same is true for the fringe “10” with “12” and for “02” with “20”. It is expected that all four of the fringes corresponding to $|\psi_\alpha\rangle$ (“01”, “10”, “12”, and “21”) are all in phase with each other, but there is a clear phase shift between the first two and the last two. The average phase of these four fringes is different by $180^\circ$ from the fringes corresponding to “02” and “20” as expected. The fringe corresponding to “11” is in phase with the fringes of “02” and “20” as was expected since for this measurement we had arranged $2(\beta + \delta) = \pi (\text{mod} 2\pi)$. The results of the fits to these fringes is shown in Table III. The differences in phase and visibility are in part due to noise (see next subsection).

The results corresponding to each of the three possible Bell states can be found by adding the measurements of the constituent parts. When doing this one would expect coincidence fringes of the form $R[1 + V \cos(\alpha + \rho)]$ with $R$ and $\rho$ as above. This corresponds to three distinct interference fringes, with $|\phi_\alpha\rangle$ and $|\psi_\alpha\rangle$ in phase and $|\psi_\gamma\rangle$ with a $180^\circ$ phase difference with respect to the other two.

In Fig. 9 we show the raw coincidence interference fringes between the detection rate at Bob and a successful BSA as a function of a change of phases in Alice’s interferometer. As expected fringes for $|\psi_\alpha\rangle$ and $|\psi_\gamma\rangle$ have a $180^\circ$ phase difference due to the phase flip caused by the
teleportation. On the other hand, the fringe for $|\psi_+\rangle$ is in phase with the $|\phi_+\rangle$ as expected. The raw visibilities obtained for the projection on each Bell state are $V_{\psi}=0.38\pm0.05$, $V_{\phi}=0.22\pm0.01$, $V_{\phi}=0.43\pm0.03$ which leads to an overall value of $V=0.34\pm0.06$ ($F=0.67\pm3$). In order to check the dependence of $|\phi_+\rangle$ on $\delta$ we also performed a teleportation with a different value for $\delta$ and we clearly observe the expected shift in the fringe (Fig. 10) while measuring similar visibilities.

Note that Bob is able to derive the phase value $\delta$ of the BSA interferometer just by looking at the phase differences between the fringes made by $\psi_+$ and $\phi_+$ and his knowledge about $\beta$. It is important to know $\delta$ since this allows Bob to perform the unitary transformation $\sigma_{2,8}$ on the teleported photon. Since the count rates were quite low we expected to have an important noise factor, an analysis of the noise follows in the next subsection.

### D. Noise analysis and discussion

In the case of the BS-BSA the noise analysis is straightforward. All of the important noise counts are completely independent of the phase, since they concern situations in which there is no single-photon interference possible. The most important sources of noise were estimated and then measured. The estimated signal-to-noise ratio was 2.6, measurements find a SNR of approximately 2.2$\pm$0.5. The largest source of noise are darkcounts at one detector combined with two real detections.

<table>
<thead>
<tr>
<th>Result 3BSA</th>
<th>$V_{raw}$</th>
<th>$V_{net}$</th>
<th>$\rho_{raw}$</th>
<th>$\rho_{net}$</th>
<th>$P_{raw}$</th>
<th>$P_{net}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>\psi_+\rangle$</td>
<td>$</td>
<td>01\rangle$</td>
<td>35±3</td>
<td>61±6</td>
<td>278±4</td>
</tr>
<tr>
<td>$</td>
<td>\psi_+\rangle$</td>
<td>$</td>
<td>10\rangle$</td>
<td>43±3</td>
<td>72±13</td>
<td>339±4</td>
</tr>
<tr>
<td>$</td>
<td>\psi_+\rangle$</td>
<td>$</td>
<td>12\rangle$</td>
<td>18±3</td>
<td>64±7</td>
<td>340±7</td>
</tr>
<tr>
<td>$</td>
<td>\psi_+\rangle$</td>
<td>$</td>
<td>21\rangle$</td>
<td>13±2</td>
<td>36±2</td>
<td>227±9</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$</td>
<td>11\rangle$</td>
<td>43±3</td>
<td>55±2</td>
<td>136±3</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$</td>
<td>02\rangle$</td>
<td>40±5</td>
<td>83±13</td>
<td>126±12</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$</td>
<td>20\rangle$</td>
<td>39±4</td>
<td>62±10</td>
<td>153±4</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$</td>
<td>12\rangle$</td>
<td>22±1</td>
<td>51±3</td>
<td>311±3</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$</td>
<td>11\rangle$</td>
<td>43±3</td>
<td>55±2</td>
<td>136±10</td>
</tr>
<tr>
<td>$</td>
<td>\phi_+\rangle$</td>
<td>$</td>
<td>11\rangle$</td>
<td>38±5</td>
<td>69±10</td>
<td>140±6</td>
</tr>
</tbody>
</table>

FIG. 9. (Color online) Uncorrected teleportation fringes found when scanning the interferometer at Bob. The fitted curves have visibilities of 0.22, 0.43, and 0.38 for $|\phi_+\rangle$, $|\phi_+\rangle$, and $|\psi_+\rangle$. The average visibility of the BSA is $V_{avg}=0.34$ ($F=0.67$).

FIG. 10. (Color online) Teleportation fringes measured in two distinct measurements with a $\delta$ which had changed by $70^\circ\pm10^\circ$. In the measurement a clear shift is visible of the fringe $|\phi_+\rangle$ by $74^\circ$ with regards to the other fringes.
The situation for the IF-BSA is more complicated. The additional interferometer has an unfortunate side effect. There are now possibilities for noise to depend on the phases of the interferometers. In other words, while measuring interference fringes there are also noise fringes. It is obviously important to be able to distinguish between the two. The most important fluctuating noise is caused by false coincidences that involve one or more photons coming from Alice and no photons from the EPR source at the BSA. These noise sources depend on the phases $\beta$ and $\delta$. This corresponds to a combination of a Franson-type Bell test with a darkcount. The fluctuation of this noise was avoided in our experiment since we only changed the phase of Alice’s interferometer ($\alpha$).

Not all possible sources of noise depend on the phases of the interferometers, there are also stable sources of noise, which are different for each of the BSA possibilities. The average value of the most important noise sources are shown in Table IV, which shows that by choosing to scan Alice instead of Bob a large fluctuating noise was avoided. It also shows that the fluctuating noise from Alice is only a small part of the total noise and therefore its effect will only be limited. Another source of errors that is different for each coincidence combination is electronical loss. These losses are caused by long (up to 100 ns) electronical delays that are required in the treatment of the coincidence signals.

The results, after noise substraction and correction for electronical transmission differences for the individual coincidence combinations, are shown in Table III. There is a clear agreement with theory, for example, the probability of finding a “11” is approximately 4 times larger than the probability for any of the other possibilities (Fig. 12).

There are a few differences worth noting between the results and theory. First of all there are small differences in visibility, these are probably caused by several small unmeasured noise sources and partially they are real physical differences which are caused by imperfect interferometers, an indication of these imperfections is given by the quality of the alignment experiments. Second, the phases of the curves

| TABLE IV. The average noise counts of several noise possibilities. Note that each measured value concerns a combination of different sources of noise. The most important noise (source Alice blocked) did not fluctuate during the experiment because the scan in the phase was done by Alice. |
|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 01  | 02  | 10  | 11  | 12  | 20  | 21  |
| Source Alice blocked | 70±4 | 60±4 | 60±4 | 88±6 | 147±6 | 46±3 | 157±5 |
| EPR to Bob blocked    | 2.7±0.1 | 1.0±0.1 | 3.5±0.1 | 5.3±0.1 | 4.3±0.1 | 1.7±0.1 | 4.2±0.1 |
| EPR to BSM blocked    | 13±1 | 7±1 | 12±1 | 15±1 | 11±1 | 7.8±1 | 13.07±1 |

FIG. 11. (Color online) Measurements of the noise for a interferometer-BSA teleportation experiment. Top: “01” and “10” are in antiphases as expected and have visibilities of $V_{01} = 0.77±0.12$, $V_{10} = 0.65±0.12$. Bottom: “12” and “21” have a $\pi$ phasesshift as expected and have visibilities of $V_{12} = 0.66±0.14$, $V_{21} = 0.91±0.13$. 

EXPERIMENTAL QUANTUM TELEPORTATION WITH A photoemission interferometer-BSA ¼ PHYSICAL REVIEW A 74, 022303 (2006)
show an interesting phase difference between “01” and “10”. The reason for this shift is unknown, but the average value of the two phases corresponds with the phase that is expected from the curve for “02” and “20”. This suggests that this effect is caused by a fluctuating noise that is out of phase with the teleportation fringe.

When the different possibilities of the BSA are summed, in order to have the Bell states, the noise will no longer have any fluctuations. This is because the different noise possibilities had a $\pi$ phase difference. After summing the different parts of the noise of a Bell state the result will be constant. For example, the noise of “01” combined with “10” is approximately constant. The overall resulting noise is in practice independent of the phase. The results after noise subtraction and correction for electronic transmission differences are shown in Fig 13 and Table III. The results show excellent correspondence between theory and experiment. The visibilities are similar within their errors. The difference in phase between $|\psi_+\rangle$ and $|\psi_-\rangle$ ($189^\circ \pm 9^\circ$) corresponds with theory ($180^\circ$). Also, since the phases were arranged in such a way that $\beta=-\delta\mod 2\pi$, the fringe of $|\psi_+\rangle$ is in phase with $|\psi_-\rangle$ (phase difference of $4^\circ \pm 9^\circ$). The normalized probabilities of a measurement (Fig. 12) show that $|\psi_+\rangle$ and $|\phi_\pi\rangle$ have the same probability (43%, respectively, 41%) and these values correspond with the theoretical value of 40%. The probability of $|\psi_-\rangle$ is 15% with a theoretical value of 20%. These excellent agreements with theory suggest that the discrepancies as seen for the individual results are caused by differences in noise that cancel out when they are added to each other.

IV. CONCLUSIONS

In conclusion we have shown experimentally that it is possible to perform a three-state Bell analysis while using
The maximum visibility when measuring the difference between aligned and nonaligned can be calculated by taking into account the probability of creating two photons in Alice \(P(00 \mid (a^t)^2\rangle\) or at the EPR source \(P(00 \mid (b^t)^2\rangle\).

\[
P(00 \mid (a^t)^2\rangle = 1, \quad P(00 \mid (b^t)^2\rangle = \frac{1}{3}, \quad V = \frac{P_{\text{out}} - P_{\text{in}}}{P_{\text{out}}},
\]

\[
P_{\text{out}} = \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}, \quad P_{\text{in}} = \frac{1}{2} + 2 \times \frac{1}{2} = \frac{5}{12}, \quad V_{\text{max}} = \frac{1}{3}.
\]

Note that when making measurements of “antidips” the photons at Bob are completely ignored.

The antidips discussed above are not the only method of aligning the setup. It is also possible to look at a dip. For example, there will be a decrease in the number of “01” depending on whether there is photon bunching or not. During measurements of such a decrease the interferometers are not stabilized for experimental reasons. Since the coincidence rate is dependent on single-photon interference it is very difficult to clearly see the decrease in counts (Fig. 14).

One way to avoid this problem is to use a baby peak as a normalization. Baby peaks are coincidences with one (or more) laser pulses of difference between the creation time of the detected photons. For example, laser pulse \(n\) creates a photon in Alice and this photon goes to detector \(D_0\), while laser-pulse \(n+1\) creates a photon in the EPR source which goes to \(D_1\). The amount of coincidences measured for these

**APPENDIX: TEMPORAL ALIGNMENT**

For a BSA to work it is important to have complete indistinguishability of the incoming qubits. This includes a indistinguishability in time. In order to align the path lengths in an experiment it is useful to perform photon bunching experiments, since photon bunching only occurs for indistinguishable photons. In the case of a teleportation experiment using a BS-BSA it is possible to perform a Mandel-dip experiment [27] by looking at the coincidence rate “00” or “11”. A decrease in the number of coincidences between the BSA detectors is observed when the photons from Alice’s source and the EPR source are indistinguishable. When an IF-BSA is used it is not trivial to directly measure such an effect, without having to make significant changes to the optical setup (such as replacing the interferometer by a beamsplitter) in between two teleportation experiments. In order to avoid any changes to the setup another method of checking indistinguishability was used. An increase in the number of coincidences “00” or “22” is dependent on indistinguishability, as was explained in the main text of the paper. The difference is clearly seen by calculating the probability to find “0” in both \(D_0\) and \(D_1\) for indistinguishable photons \(P(00 \mid \text{aligned})\) and distinguishable \(P(00 \mid \text{nonaligned})\) photons:

\[
P(00 \mid \text{aligned}) = \frac{1}{4}, \quad P(00 \mid \text{nonaligned}) = \frac{1}{8},
\]

only linear optics without the use of ancilla photons. In principle, this measurement can reach a success rate of 50%.

The authors acknowledge financial support from the European IST-FET project “RamboQ,” the European project “QAP,” and from the Swiss NCCR project “Quantum Photonics” and we thank C. Barreiro and J.-D. Gautier for technical support. Also, we would like to express our gratitude to Dr. Keller of the ETH Zurich for the lending of materials.

**ACKNOWLEDGMENTS**

The case in our experiment. The dashed curves are for the case of distinguishable photons at the BSA and the plain curves are for indistinguishable photons (teleportation).

**FIG. 15.** Simulation of result from a teleportation experiment in the case that the interferometers have been aligned to have \(\delta = \beta\) as was the case in our experiment. The dashed curves are for the case of distinguishable photons at the BSA and the plain curves are for indistinguishable photons (teleportation).

The maximum visibility when measuring the difference between aligned and nonaligned can be calculated by taking into account the probability of creating two photons in Alice \(P(00 \mid (a^t)^2\rangle\) or at the EPR source \(P(00 \mid (b^t)^2\rangle\).

\[
P(00 \mid (a^t)^2\rangle = 1, \quad P(00 \mid (b^t)^2\rangle = \frac{1}{3}, \quad V = \frac{P_{\text{out}} - P_{\text{in}}}{P_{\text{out}}},
\]

\[
P_{\text{out}} = \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}, \quad P_{\text{in}} = \frac{1}{2} + 2 \times \frac{1}{2} = \frac{5}{12}, \quad V_{\text{max}} = \frac{1}{3}.
\]
pulses will depend on the single-photon interference but there will clearly not be any photon bunching. Since such coincidences have the same interference effects as for the real coincidences it can be used to normalize a measurement and in this way a dip can be found (Fig. 14). Since this normalization method is much more complicated and less accurate it was not used for alignment, only antidip alignment was used.

If temporal alignment is not accomplished in a BS-BSA teleportation experiment the resulting coincidence rates will not depend on the phases of the interferometers and therefore a fringe with V=0% is found. When using an IF-BSA this is not the case since the presence of the extra interferometer leads to a single photon interference when changing the phase α. It is clearly important to be able to distinguish between these interferences and the interference fringes caused by teleportation. The behavior of the nonaligned setup can be readily calculated and the fringes that will be found are shown in Table V. One important fact clearly stands out straight away: there is no interference for “02” and “20” if the photons are distinguishable but there is when the photons are indistinguishable. The visibility of these fringes are an important indication whether or not there was temporal alignment during the experiment. In the experiment presented here a visibility of V=55%±3% was found which indicates that there was temporal indistinguishability.

Other indications whether there is good temporal alignment can be found when simulating the result of an unaligned experiment. Such a simulation is shown (Fig. 15) for the case of δ=−β as was used during our experiments. The simulation clearly shows differences between the two cases which are readily identifiable in an experiment, such as the phaseshift of ς between the fringes for “01” and “10”. These differences make it possible to see after an experiment whether or not the alignment was good and remained good.

Entanglement and non-locality are different resources

Nicolas Brunner, Nicolas Gisin and Valerio Scarani
Group of Applied Physics, University of Geneva, 20, rue de l’Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland
E-mail: nicolas.brunner@physics.unige.ch

New Journal of Physics 7 (2005) 88
Received 14 February 2005
Published 5 April 2005
Online at http://www.njp.org/
doi:10.1088/1367-2630/7/1/088

Abstract. Bell’s theorem states that, to simulate the correlations created by measurement on pure entangled quantum states, shared randomness is not enough: some ‘non-local’ resources are required. It has been demonstrated recently that all projective measurements on the maximally entangled state of two qubits can be simulated with a single use of a ‘non-local machine’. We prove that a strictly larger amount of this non-local resource is required for the simulation of pure non-maximally entangled states of two qubits \( |\psi(\alpha)\rangle = \cos \alpha |00\rangle + \sin \alpha |11\rangle \) with \( 0 < \alpha \lesssim \frac{\pi}{1.8} \).
1. Introduction

There exists in Nature a channel that allows one to distribute correlations between distant observers, such that (i) the correlations are not already established at the source, and (ii) the correlated random variables can be created in a configuration of space-like separation, i.e. no normal signal can be the cause of the correlations [1]. This intriguing phenomenon, often called quantum non-locality, has been repeatedly observed, and it is natural to look for a description of it. A convenient description is already known: quantum mechanics (QM) describes the channel as a pair of entangled particles. In recent years, there has been a growing interest in providing other descriptions of this channel, mainly assuming a form of communication. Usually, the interest in these descriptions does not come from a rejection of QM and the desire to replace it with something else; rather the opposite: the goal is to quantify how powerful QM is by comparing its achievements to those of other resources.

For instance, one may naturally ask how much information should be sent from one party (Alice) to the other (Bob) in order to reproduce the correlations that are obtained by performing projective measurements on entangled pairs (to ‘simulate entanglement’). The amount of communication is something that we are able to quantify, thus the answer to this question provides a measure of the non-locality of the channel. Bell’s theorem implies that some communication is required, but does not quantify this amount. Several works [2, 3] underwent the task of estimating the amount of communication required to simulate the maximally entangled state of two qubits (singlet). These partial results were superseded in 2003, when Toner and Bacon [4] proved that the singlet can be simulated exactly using local variables plus just one bit of communication per pair. This amount of communication is tight, in the absence of block-coding—which is indeed the way Nature does it: in an experiment, each pair of entangled particles is ‘processed’ independently of those that preceded and those which will follow it.

More recently, something other than communication has been proposed as a tool to study non-locality: the non-local machine (NLM) described by Popescu and Rohrlich [5], sometimes called the PR box—actually, the first appearance of this ‘machine’ is equation (1.11) of [6]. This hypothetical machine was constructed to violate the Clauser–Horne–Shimony–Holt (CHSH) inequality \[ B \leq 2 \] up to its algebraic bound of \( B = 4 \) (while it is known that QM reaches up only to \( B = 2\sqrt{2} \)) without violating the no-signalling constraint; and it would also provide a very powerful primitive for information-theoretical tasks [8, 9]. Cerf et al [10] have shown that the singlet can be simulated by local variables plus just a single use of the NLM per pair: this is the analogue of the Toner–Bacon result for communication. While the NLM is by far a less familiar object than bits of communication, in the context of simulation of entanglement it has a very pleasant feature: it automatically ensures that the no-signalling condition is respected. In contrast, bits of communication imply signalling: to reproduce quantum correlations, as in the Toner–Bacon model, one must cleverly mix different communication strategies in order to hide the existence of communication. The idea itself of hidden communication between quantum particles has several drawbacks [11] and is hard to reconcile with the persistency of correlations in experiments with moving devices [12].

Thus, to date, the simulation of quantum non-locality has been studied for two resources (communication and the NLM) and the results are similar: the basic unit of the resource (one bit, or a single use of the NLM) is sufficient for the simulation of the singlet. Very little is known beyond the case of the singlet. Even staying with just two qubits, the only known result is that two bits of communication are enough to simulate all states [4], but this is not claimed to be tight. In this paper, we study the analogue problem using the NLM and demonstrate that, in order to simulate the correlations of some pure non-maximally entangled state of two qubits, a single use of the NLM is not sufficient: an amount of non-local resources strictly larger than amount for the simulation of the maximally entangled state is needed. Curious as it may seem, this is not the first example in the literature where entanglement and non-locality do not behave monotonically with one another: Eberhard proved that non-maximally entangled states require lower detection efficiencies than maximally entangled ones, in order to close the detection loophole [13]; Bell inequalities have been found whose largest violation is given by a non-maximally entangled state [14] and this has some consequences on the communication cost as well [15]. It is also known that some mixed entangled states admit a local variable model, even for the most general measurements [16].

The present paper is structured as follows. As a necessary introduction, we start by recalling the meaningful mathematical tools for this investigation (section 2). In section 3, we demonstrate the main claim, by showing that there exists a unique Bell-type inequality using three settings for both Alice and Bob which is not violated by any strategy using the NLM at most once, and which is violated by all the states of the form \( |\psi(\alpha)\rangle = \cos \alpha |00\rangle + \sin \alpha |11\rangle \) for \( 0 \leq \alpha \leq \frac{\pi}{7} \) (the sign of approximate inequality means that these are numerical, not analytical results). Thus, these states cannot be simulated by a single use of the NLM. In the same section, we show how this new inequality can be violated by two uses of the NLM or by one bit of communication, and comment on these features. In section 4, we consider extension to more settings on Alice’s and/or on Bob’s side. The case of four settings for Alice and three settings for Bob allows us to extend the result to the range \( 0 < \alpha \leq \frac{\pi}{7} \) (in particular, we prove that at least two uses of the NLM are required to simulate pure states arbitrary close to the product state \( \alpha = 0 \)). The (admittedly incomplete) survey of other cases did not provide further improvements. The concluding remarks are given in section 5.
Figure 1. Representation of the regions in probability space. The thick line represents the polytope of shared randomness; its vertices are deterministic strategies (D). Above it lies the new polytope obtained with a single use of the NLM. The curved line encloses the points achievable with measurements on quantum states. The black dot represents the measurement on the singlet that gives the maximal violation of the inequality corresponding to the facet (CHSH or I_{3322}). The existence of the grey regions above the I_{3322}-facet is the result of section 3.

2. Tools: polytopes and the no-signalling condition

Instead of tackling the issue of simulating all possible measurements done on an entangled state, we consider a restricted protocol, as typical in Bell’s inequalities. Obviously, if this restricted protocol cannot be simulated, a fortiori it will be impossible to simulate all the correlations. We allow then each of the two physicists, called Alice and Bob, to choose between a finite set of possible measurements \{A_i\}_{i=1}^{mA}; \{B_j\}_{j=1}^{mB}. As a result of each measurement on a pair, they get an outcome noted \(r_A, r_B\). We focus here on dichotomic observables (like von Neumann measurements on qubits), with the convention \(r_A, r_B \in \{0, 1\}\). An ‘experiment’ is fully characterized by the family of probabilities \(P(r_A, r_B | A_i, B_j) \equiv P_{ij}(r_A, r_B)\). There are \(d = 4m_A m_B\) such probabilities, so each experiment can be seen as a point in a region of a \(d\)-dimensional space, bounded by the conditions that probabilities must be positive and sum up to one. By imposing restrictions on the possible distributions, the region of possible experiment shrinks, thus adding non-trivial boundaries [6, 17, 18]. For instance, one may require that the probability distribution must be built without communication, only with shared randomness. In this case, the bounded region actually forms a polytope, that is a convex set bounded by hyperplanes (‘facets’) which are Bell’s inequalities in the usual sense. The vertices of this polytope are deterministic strategies, that is, probability distributions obtained by setting \(r_A, r_B\) always at 0 or always at 1 for each setting. The vertices are thus easily listed, but to find the facets given the vertices is a computationally hard task. The probability distributions obtained with a single use of the NLM also form a polytope, obviously larger than the one of shared randomness—actually, the vertices associated to deterministic strategies remain vertices of this new polytope, but more vertices are added. Finally, the probability distributions that can be obtained from measurements on quantum states form a convex set which is not a polytope.\(^1\) A sketch of this structure is given in figure 1, which will be commented in more detail in the following.

\(^1\) This statement is valid for the set of all possible measurements on all possible quantum states, see [17], as well as paragraph V.C. of [19]. If one restricts to the measurements on a given state, or even to the von Neumann measurements on a Hilbert space with given dimension, convexity is not proved (although no counterexample is known, to our knowledge).
Since our goal is to simulate QM, we impose from the beginning the constraints of no-signalling; that is, we focus only on those probability distributions which fulfil

$$\sum_{r_A} P_{ij}(r_A, r_B) = P_j(r_B) \quad \text{for all } i$$

(1)

and a similar condition for the marginal of A. Under no-signalling, the full probability distribution is entirely characterized by $d_{ns} = m_A m_B + m_A + m_B$ probabilities, which we choose conventionally to be the $P_i(r_A = 0)$, $P_j(r_B = 0)$ and $P_{ij}(r_A = r_B = 0)$ as in [20].

3. Main result

3.1. Basic notations

Let us focus more specifically on the first case of interest for this paper, $m_A = m_B = 3$. The no-signalling probability space is 15-dimensional. All the facets of the deterministic polytope are known [20]: up to relabelling of the settings and/or of the outcomes, they are equivalent either to the usual two-settings CHSH inequality, or to the truly three-settings inequality $I_{3322}$ that reads

$$I_{3322} = \begin{bmatrix} -2 & 1 & -1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} \leq 0. \quad (2)$$

Here, the notation represents the coefficients that are put in front of the probabilities, according to

$$P_i(r_A = 0) \quad P_j(r_B = 0) \quad P_{ij}(r_A = r_B = 0).$$

(3)

The maximal violation allowed by QM is obtained for the singlet and is $\langle I_{3322} \rangle = \frac{1}{2}$. To become familiar with the notations, the deterministic strategy in which Alice outcomes $r_A = 0$ for $A_0$, $A_1$ and $r_A = 1$ for $A_2$, and Bob outcomes always $r_B = 0$, corresponds to the probability point

$$[0_d0_d1_d; 0_d0_d0_d] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \quad (4)$$

To see the result of $I_{3322}$ on this strategy, one has simply to multiply the arrays term-by-term: here we find $I_{3322} = 0$. There are obviously $2^6 = 64$ deterministic strategies; among these, 20 saturate the inequality $I_{3322} = 0$ (i.e., they lie on the facet) while the others give $I_{3322} < 0$. To verify that this is indeed a facet, it is enough to show that the rank of the matrix containing the 20 points that saturate the inequality is $d_{ns} - 1 = 14$, so that the condition $I_{3322} = 0$ really defines a hyperplane [21].
3.2. The polytope of a single use of the NLM

The NLM is defined as a two-input and two-output channel. Alice inputs $x$ and gets the outcome $a$, Bob inputs $y$ and gets the outcome $b$; all these numbers take the values 0 or 1, the marginal distribution on each side is completely random, $P(a = 0) = P(b = 0) = \frac{1}{2}$, but the outcomes are correlated as

$$a + b = xy. \quad (5)$$

Explicitly, if either $x = 0$ or $y = 0$, then $(a, b) = (0, 0)$ or $(1, 1)$ with equal probability; if $x = y = 1$, then $(a, b) = (0, 1)$ or $(1, 0)$ with equal probability.

The most general strategy allowing a single use of the NLM is sketched in figure 2. Alice and Bob share some random variable $\lambda$. Alice inputs in the machine the bit $x = x(A_i, \lambda)$; the machine gives the output $a$, and Alice outputs $r_A = r_A(A_i, \lambda, a)$. The extremal strategies are such that a given $x$ is associated to each $A_i$, and the outcome is either $r_A = a$ or $r_A = 1 - a$. Similarly for Bob. Note that it is also possible that for a given pair $(A_i, B_j)$ Alice uses the machine while Bob outputs a deterministic bit; in this case, we can suppose that Bob inputs $y = 0$ in the machine but does not use the output $b$. We shall come back below to the listing of extremal strategies. Let us now see how the polytope of possible probability distributions is enlarged by allowing a single use of the NLM.

By construction, with a single use of the machine, one can violate the CHSH inequality more than is possible in QM. In the polytope picture, one new point appears above any face corresponding to CHSH: the facets of the enlarged polytope should now pass through this points. They must pass through the deterministic points as well, because these points are still extremal. All this is sketched in figure 1. We will not study this example further, however, since there is no hope of finding something interesting above a CHSH-like facet: all the probability distributions involving two settings which are no-signalling (in particular all distributions arising from measurements on a quantum state) can be simulated by a single use of the NLM [18].

So let us consider the facet defined by $I_{3322} = 0$. Again, a single use of the NLM allows to violate $I_{3322}$, which is expected since the NLM can in particular simulate the singlet. For instance, consider the strategy in which Alice inputs $x = 0$ in the NLM for $A_0$ and $A_1$, and $x = 1$ for $A_2$, while Bob inputs $y = 0$ for $A_0$ and $B_2$, and $y = 1$ for $B_1$; in each case, the outputs are $r_A = a$, $r_B = b$. Figure 2. Schematics of a strategy allowing a single use of the NLM between Alice and Bob. See text for details.
This strategy gives the probability point

\[
[0_m 0_m 1_m; 0_m 1_m 0_m] = \frac{1}{2} \times \begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 0 \\
1 & 1 & 1
\end{pmatrix}
\]

yielding \( I_{3322} = \frac{1}{2} \): the machine can violate \( I_{3322} \) more than QM. However, here we are going to show what is graphically represented in figure 1: for non-maximally entangled states, some points achievable with quantum states lie outside the enlarged polytope. In other words, the facets of this polytope define generalized Bell’s inequalities that can still be violated by QM.

To find the facets of the polytope allowing a single use of the machine, one must first list all the vertices (extremal strategies). This can be done systematically on a computer, once having noticed that for each setting, Alice and Bob have six choices: deterministically output 0 or 1 (noted 0\(_d\), 1\(_d\)), input 0 or 1 in the machine and keep the output of the machine (noted 0\(_m\), 1\(_m\)), input 0 or 1 in the machine and flip the output of the machine (noted 0\(_f\), 1\(_f\)).\(^2\) This listing gives \textit{a priori} \(6^6 = 46656\) strategies, although many of them are equal\(^3\) and only 3088 different strategies are left after inspection. Note that some of these are not even extremal points of the polytope: for instance, the strategy \([0_m, 0_m, 0_m; 0_m, 0_m, 0_m]\), in which both Alice and Bob input 0 in the NLM for all settings, yields the same probability point as the equiprobable mixture of the deterministic strategies \([0_d, 0_d, 0_d; 0_d, 0_d, 0_d]\) and \([1_d, 1_d, 1_d; 1_d, 1_d, 1_d]\). Certainly not equivalent to a mixture of deterministic strategies, however, are those strategies like (6) which violate \(I_{3322}\): upon counting, there are 28 of these, all giving the same violation \(I_{3322} = \frac{1}{2}\). To find the facets of the new polytope, we use available computer programs.\(^4\) We find some trivial facets, plus a single non-trivial one\(^5\) which reads

\[
M_{3322} = \begin{pmatrix}
-2 & 0 & 0 \\
1 & 1 & 1 \\
-1 & 1 & -1 \\
0 & 1 & -1 & 0
\end{pmatrix} \leq 0. \tag{7}
\]

This new inequality is extremely similar to \(I_{3322}\), equation (2): only the coefficient of \(P_0(r_A = 0)\) is now \(-2\) instead of \(-1\). The origin of this difference can be appreciated, at least to some extent: one of the biggest difficulties found in adapting the Toner–Bacon model [4] to non-maximally entangled states lies in the need of simulating not only the correlations, but also the non-trivial marginal distributions. It is thus a good idea, for our purpose, to add penalties on the marginal distributions.

Summarizing, we have found a tight inequality \(M_{3322} \leq 0\) which is satisfied by all the 3088 extremal points of the polytope of probabilities achievable by shared randomness plus a single use of the NLM. Now, we move on to show that QM violates this inequality.

\(^2\) This takes automatically into account the possibility that Alice and Bob use the alternative version of the NLM defined by \(a + b = xy + 1\).

\(^3\) For instance, \([0_f, 1_m, 1_f; 0_f, 0_f, 0_f]\) and \([0_f, 1_f, 1_m; 1_f, 0_f, 1_f]\) define the same probability point.

\(^4\) We used the function \texttt{convhulln} of \textsc{Matlab} on the 28 points which yield \(I_{3322} = \frac{1}{2}\) alone, or on these points plus some of the 20 deterministic points which yield \(I_{3322} = 0\). The result is the same.

\(^5\) Actually, the program outputs two facets: inequality (7), and a similar one where the first line \((2, 0, 0)\) is replaced by \((1, 1, 0)\). If one writes down the inequality explicitly, it becomes evident that one is transformed into the other by exchanging \(A_0\) and \(A_1\) and by flipping the bit \(r_{B2}\).
Figure 3. Value of the quantum mechanical expectation $M(\alpha)$ as a function of $\alpha$ for the optimized settings of Alice and Bob, for the inequalities $M_{3322}$, equation (7), and $M_{4322}$, equation (14). The region where $M(\alpha) > 0$ corresponds to the grey regions in figure 1: the corresponding quantum states cannot be simulated by a single use of the NLM.

3.3. Violation of the inequality with pure non-maximally entangled states

We consider states of the form

$$|\psi(\alpha)\rangle = \cos \alpha |00\rangle + \sin \alpha |11\rangle$$

with $\cos \alpha \geq \sin \alpha \geq 0$. Up to local operations, this is the most general pure state of two qubits (Schmidt decomposition). We form the Bell operator $\mathcal{M}$ as usual, by replacing the probabilities in (7) by the corresponding one-dimensional projectors, like

$$P_0(r_A = 0) \longrightarrow \frac{1}{2} (1 + \vec{a}_0 \cdot \vec{\sigma}).$$

For each $\alpha$, we have to find the settings which maximize $M(\alpha) = \langle \psi(\alpha) | \mathcal{M} | \psi(\alpha) \rangle$. We have not found a closed analytical formula, but it is easy for a computer to optimize over twelve real parameters. The result is shown in figure 3: one finds $M(\alpha) > 0$ for $0 < \alpha < \frac{2}{19} \pi$, with a maximal violation $M(\bar{\alpha}) \approx 0.0061$ at $\bar{\alpha} \approx 0.0712 \pi$. It appears that all the optimal settings are of the form

$$\hat{a}_i = \cos \theta_i \hat{\alpha} \hat{\zeta} + \sin \theta_i \hat{\alpha} \hat{\chi}, \quad \hat{b}_j = \cos \theta_j \hat{\beta} \hat{\zeta} + \sin \theta_j \hat{\beta} \hat{\chi}.$$  

The curve of figure 3 is the exact version of the pictorial argument of figure 1. Note in particular the following features: (i) as expected, there is no violation for the singlet ($\alpha = \frac{\pi}{4}$), because this state can be simulated with the NLM; even more, $M(\frac{\pi}{4}) = -\frac{1}{4}$ which is the difference between $\langle I_{3322} \rangle = \frac{1}{4}$ on the singlet and the maximal value $I_{3322} = \frac{1}{2}$ achievable with the NLM, the picture of figure 1 yields in this case even a quantitative prediction; and (ii) as mentioned [19], it is not obvious that the set of probabilities obtained from quantum measurements on
$|\psi(\alpha)\rangle$ is convex; so at this point it is not proved that states arbitrarily close to the product state $|00\rangle$ cannot be simulated by a single use of the NLM—the proof will be provided in section 4.

3.4. Violation of the inequality with two NLMs

As we mentioned in section 1, two bits of communications are a sufficient resource to simulate any state of two qubits. The analogue simulation by twice using the NLM is still missing, and may even not exist. While waiting for more clarification, we have found a way of violating inequality (7) by using the NLM twice. The settings are coded as $A \rightarrow (x', x'')$, $B \rightarrow (y', y'')$ according to: $A_0 \rightarrow (0, 0), A_1 \rightarrow (0, 1), A_2 \rightarrow (1, 0); B_0 \rightarrow (0, 0), B_1 \rightarrow (1, 0)$ and $B_2 \rightarrow (0, 1)$. Then, $x'$ and $y'$ are used as inputs in the first use of the machine, whose outcomes are denoted $a'$ and $b'$; $x''$ and $y''$ are used as inputs in the second use of the machine, whose outcomes are denoted $a''$ and $b''$. Finally, Alice outputs $r_A = a' + a''$, Bob outputs $r_B = b' + b''$ (sum modulo 2). This strategy gives the probability point

$$I_{3322} = \frac{1}{2} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

yielding $I_{3322} = 1$ and $M_{3322} = \frac{1}{2}$.

We can go a step further. Consider a mixture of two strategies: with probability $p$, the strategy just described which uses two NLMs; with probability $1 - p$, the deterministic strategy in which Alice and Bob output always $r_A = r_B = 1$ for all settings (all entries of the table are zeros), and which obviously does not require any use of the NLM. Such a mixed strategy yields a violation $M_{3322} = \frac{p}{2}$ of the inequality. Now, if $p < \frac{1}{2}$, this strategy uses less than one NLM on average. This is not a contradiction with our main result: at least two NLMs must be available to simulate non-maximally entangled states, albeit possibly this resource is not used for all items.

3.5. Violation of the inequality with one bit of communication

It is natural to ask whether inequality (7) provides also a Bell inequality for one bit of communication. Bacon and Toner [22] had studied such inequalities for three settings, but they restricted themselves to correlation inequalities, whereas inequality (7) is a probability inequality. The problem is complex because, as we mentioned, pure strategies with one bit of communication imply signalling; we must find a mixture of such strategies which is no-signalling and which violates our inequality. It turns out that such mixtures do exist, so that our inequality is not an inequality for one bit of communication. In other words, the polytope of the probability distributions obtained with one bit of communication plus the no-signalling constraint is larger than the one associated to a single use of the NLM.

As an explicit example [23], it can be verified that the no-signalling strategy

$$[\text{One bit, no-signalling}] = \frac{1}{5} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

This was pointed out to us by [23].
yields the violation \( M_{3322} = \frac{1}{5} \) and can be obtained as the equiprobable mixture of the following five one-bit strategies:

\[
\{(r_{A_i}; r_{B_j})|c_0, c_1, c_2\} = \begin{cases} 
[1, 1, 0; c, 1, c|1, 1, 0] \\
[1, 0, 1; c, c, 1|1, 0, 1] \\
[0, 1, 1; c, c, c|0, 1, 1] \\
[1, 1, 1; 1, 1, c|0, 1, 1] \\
[1, 1, 1; 1, c, 1|1, 1, 0] 
\end{cases}.
\] (12)

In these notations, \( c_i \) is the value of the bit that Alice sends to Bob when she has used the setting \( A_i \); \( r_{B_j} = c \) means that, upon choosing the setting \( B_j \), Bob outputs the value of the bit received from Alice.

The fact that the inequality (7) can be violated by one bit of communication shows that a single use of the NLM does not correspond to a single bit of communication plus no-signalling: the NLM is a resource strictly weaker than communication, as argued in [10], and grasps finer details of the structure of quantum non-locality. The question whether one bit of communication is sufficient to simulate non-maximally entangled states is obviously still open.

4. Extensions to more settings

In the previous section, we have provided a complete study of the case \( m_A = m_B = 3 \): there cannot be any inequality other than (7) which has the desired properties. In this section, we explore other cases, starting from the next easiest, namely \( m_A = 4 \) and \( m_B = 3 \).

4.1. The case \( m_A = 4 \) and \( m_B = 3 \)

In the case \( m_A = 4 \) and \( m_B = 3 \), the no-signalling probability space is 19-dimensional. All the facets of the deterministic polytope have been listed in appendix A of [20]: one finds of course CHSH, \( I_{3322} \), plus three new inequalities. The one which turns out to be of interest is (A.2) of that\(^7\); using the properties of no-signalling distributions, and providing Alice instead of Bob with the four settings, this inequality can be rewritten in the form

\[
I_{4322}^{(2)} = \begin{pmatrix}
-2 & -1 & 0 & 0 \\
1 & 1 & 1 & 0 \\
-1 & 0 & -1 & 0 \\
0 & 1 & -1 & 0 \\
\end{pmatrix} \leq 0.
\] (13)

The polytope of a single use of the NLM can be found as in section 3. After listing, one finds that the \( 6^7 \) possible extremal strategies produce 17272 different points, 63 of which violate (13).

\(^7\) Actually, we did not perform a systematic search because the polytope becomes rather big. Based on the result of section 3, we have taken inequalities (A.1)–(A.3) of [20], verified that all of them can be violated by a single use of the NLM, then modified the marginals. It turns out that, even by modifying the marginals by 1, the new inequalities derived from (A.1) and (A.3) are not violated by QM.

by $I_{4322}^{(2)} = \frac{1}{2}$. By numerical inspection, we found that these points define a single non-trivial new facet

$$M_{4322} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix} \leq 0. \quad (14)$$

Quite similarly to the case $m_A = m_B = 3$, the difference between the original inequality and the new one is just a larger penalty on one marginal. Also similar is the fact that (14) can be violated by strategies which use twice the NLM or one bit of communication, as can be easily verified. In fact, simply by taking the corresponding strategies of section 3 and adding the condition $A_3 = A_1$, we can produce the probability point

$$[\text{two NLMs, one bit}] = \lambda \times \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad (15)$$

which gives $M_{4322} = \lambda$, with $\lambda = \frac{1}{2}$ in the case of two NLMs and $\lambda = \frac{1}{3}$ in the case of one bit of communication.

The interest of $M_{4322}$ comes from the quantum violation, which (i) is larger than the violation of $M_{3322}$, thus allowing one to extend the range of $\alpha$ for which one NLM is not enough, and (ii) is obtained for a family of settings which can be easily parametrized (see appendix A). Specifically, one finds $M(\alpha) > 0$ for $0 < \alpha \lesssim \frac{\pi}{7}$, with a maximal violation $M(\tilde{\alpha}) \approx 0.0102$ at $\tilde{\alpha} \approx \frac{\pi}{12}$ (figure 3). For small values of $\alpha$, moreover, one can prove

$$M(\alpha) \geq \frac{1}{4} \alpha^2 + O(\alpha^4). \quad (16)$$

Thus $M(\alpha) > 0$ as soon as $\alpha > 0$: the simulation of pure states with arbitrarily weak entanglement requires more than one NLM—again, as we noticed at the end of section 3.4, it may be the case that correlations can be reproduced by using this resource only on a subset of the particles.

### 4.2. Other inequalities

Beyond the 3322 and 4322 cases, the facets of the deterministic polytope have not been listed exhaustively, but several examples of facets are available [20, 24]. On these, we searched for possible extensions of our results by increasing the penalties in some marginals.

Starting from the inequality $I_{4422} \leq 0$ given in [20], the corresponding inequality $M_{4422} \leq 0$ is obtained exactly as above, just replacing $-1$ with $-2$ as the coefficient of $P_0(r_A = 0)$. The result is similar: $M_{4422} \leq 0$ indeed holds for all strategies allowing a single use of the non-local machine, and QM violates it. If all four settings are used, the range of values of $\alpha$ in which...
we found a violation is however smaller than for $M_{3322}$, only up to $\sim \frac{\pi}{17}$. Note that $I_{4422}$ is less violated than $I_{3322}$ by the singlet [20], while the violation achievable by the NLM is $\frac{\pi}{2}$ for both; by looking at figure 1, it becomes intuitive that the range of violation should decrease. Interestingly, one can recover the (better) result of figure 3 by setting $a_3$ and $b_0$ to the value $1_d$, thus reducing $M_{4422}$ to $M_{3322}$. This assignment reads $P_1(r_A = 0)$, $P_0(r_B = 0) \rightarrow 0$ and is thus not of the form (9): it describes a degenerate measurement.

The other 4422 deterministic facets, as well as some 5522 and 6622 ones, did not appear to be worth a closer study after our survey. We have considered neither inequalities with larger number of outcomes, nor multi-partite scenarios.

5. Conclusion and perspectives

In conclusion, we have shown that the simulation of non-maximally entangled states of qubits requires a strictly larger amount of resources (use of the NLM) than the simulation of the singlet. We have completely solved the problem in the case of three settings and two outcomes for both Alice and Bob, and found an extension in the case where Alice chooses among four settings. At present, therefore, we know that the singlet can be simulated by a single use of the NLM, while the simulation of states with $0 < \alpha \lesssim \frac{\pi}{7.8}$ requires that more than one NLM be available (even though possibly this resource is seldom used). It will be of great interest to fill the gap, and to see whether a similar result holds when the resource used to simulate correlation are bits of communications instead of the NLM.

From a very fundamental point of view, we have discovered a new surprising feature of the quantum world, which shows once more how far this world lies from our intuition. But a precise understanding of the incommensurability between entanglement and non-locality would be of interest for applications as well. For instance, it would allow us to study whether in a given quantum information protocol (cryptography, teleportation, an algorithm . . . ) it is better to look for the largest amount of entanglement or for the largest amount of non-locality.

Acknowledgments

We acknowledge financial support from the EU Project RESQ and from the Swiss NCCR ‘Quantum photonics’.

Appendix A. Optimal settings for $M_{4322}$

To compute the maximal violation of $M_{4322}$, equation (14), on qubit states, one must perform an optimization over 14 parameters. This we first performed numerically; by looking at the result, however, an analytical form for the settings has been guessed. We give the settings by indicating the azimuthal and polar angle of the vector in the Bloch sphere $\hat{n} \equiv (\theta, \phi)$.

For $0 \leq \alpha \lesssim \frac{\pi}{10.6}$, the optimal settings lie in the $(x, z)$ plane and only two parameters depend on $\alpha$; specifically

\begin{align*}
A_0 &= (\pi, 0) = -\hat{z}, & A_1 &= (\theta_A, \pi), & A_2 &= \left(\frac{\pi}{2}, \pi\right) = -\hat{x}, & A_3 &= (\theta_A, 0), \\
B_0 &= (\theta_B, \pi), & B_1 &= (\theta_B, 0), & B_2 &= (\pi, 0) = -\hat{z}.
\end{align*}
This gives
\[ M(\alpha) = \frac{1}{4} [ -3 + \cos 2\alpha + \cos \theta_A - \cos \theta_B (1 + \cos 2\alpha) + \cos \theta_A \cos \theta_B + \sin \theta_B(1 + \sin \theta_A) \sin 2\alpha ] . \]

We have not been able to find a closed formula for \( \theta_{A,B} \) as a function of \( \alpha \).

For \( \frac{\pi}{10} \leq \alpha \leq \frac{\pi}{4} \), the optimal settings do not lie in the \((x, z)\) plane any longer, and only one parameter depends on \( \alpha \); specifically
\[ A_0 = (\pi, 0) = -\hat{z}, \quad A_1 = \left( \frac{\pi}{2}, \frac{5\pi}{6} \right), \quad A_2 = \left( \frac{\pi}{2}, \frac{\pi}{2} \right) = +\hat{y}, \quad A_3 = \left( \frac{\pi}{2}, \frac{7\pi}{6} \right), \]
\[ B_0 = \left( \theta_B, \frac{4\pi}{3} \right), \quad B_1 = \left( \theta_B, \frac{2\pi}{3} \right), \quad B_2 = (\theta_B, 0). \]

This gives
\[ M(\alpha) = \frac{1}{4} [ -7 + \cos 2\alpha - 3 \cos \theta_B (1 + \cos 2\alpha) + 3\sqrt{3} \sin \theta_B \sin 2\alpha ] \]
which can easily be maximized to find
\[ \theta_B(\alpha) = \pi - \arctan \left( \frac{\sqrt{3} \sin 2\alpha}{1 + \cos 2\alpha} \right) . \]

Even though these are not the best settings for small values of \( \alpha \), we can study the limit \( \alpha \to 0 \) and we find \( M(\alpha) = \frac{1}{4} \alpha^2 + O(\alpha^4) \) which means a violation of the inequality for arbitrary small values of \( \alpha \). Thus, we prove analytically that at least this value can be reached, as written in the main text, equation (16).

References

Bell J S 1964 Physics 1 195
Steiner M 2000 Phys. Lett. A 270 239
Scarani V, Tittel W, Zbinden H and Gisin N 2000 Phys. Lett. A 276 1
[16] Pitowski I 1989 Quantum Probability, Quantum Logic (Lecture Notes in Physics vol 321) (Heidelberg: Springer)
[23] Pironio S 2004 private communication
Bell-type inequalities for nonlocal resources

Nicolas Brunner, Valerio Scarani, and Nicolas Gisin

Group of Applied Physics, University of Geneva, 20 rue de l’Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland

(Received 6 April 2006; accepted 16 August 2006; published online 1 November 2006)

We present bipartite Bell-type inequalities which allow the two partners to use some nonlocal resource. Such inequalities can only be violated if the parties use a resource which is more nonlocal than the one permitted by the inequality. We introduce a family of $N$-input nonlocal machines, which are generalizations of the well-known PR (Popescu-Rohrlich) box. Then we construct Bell-type inequalities that cannot be violated by strategies that use one of these new machines. Finally we discuss implications for the simulation of quantum states. © 2006 American Institute of Physics. [DOI: 10.1063/1.2352857]

I. INTRODUCTION

One of the most striking properties of quantum mechanics is nonlocality. It is well known that two separated observers, each holding half of an entangled quantum state and performing appropriate measurements, share correlations which are nonlocal, in the sense that they violate a Bell inequality. Indeed, this has been demonstrated in many laboratory experiments. A key feature of entanglement is that it does not allow the two distant observers to send information to each other faster than light, i.e., correlation from measurements on quantum states is no-signaling.

It is an interesting problem to quantify how powerful the nonlocal correlations of quantum mechanics are. In order to do that, one has to use some nonlocal resource. A quite natural choice is indeed classical communication. In 2003, Toner and Bacon showed that a single bit of communication is enough to reproduce the correlations of the singlet state. In the last years another nonlocal resource, the PR (Popescu-Rohrlich) box, has also been proposed to study this problem. Introduced in 1994 by Popescu and Rohrlich, the PR box was then proven to be a powerful resource for information theoretic tasks, such as communication complexity and cryptography. It was also recently suggested that the PR box is a unit of nonlocality. The PR box has the appealing feature that it is intrinsically nonsignaling, which is of course not the case of classical communication. Note that a PR box is a strictly weaker resource than a bit of communication. Recently, Cerf et al. presented a model using a single PR box which simulates correlations from any projective measurement on the singlet. It appears very natural to extend this study to other quantum states, but this turns out to be quite difficult, even for nonmaximally entangled pure states of two qubits. In a recent paper we showed a family of nonmaximally entangled states, whose correlations cannot be reproduced by a single PR box. In other words, some nonmaximally entangled states require a strictly larger amount of nonlocal resources than the maximally entangled state to be simulated. This suggests that entanglement and nonlocality are different resources. To demonstrate this result we found a Bell-type inequality allowing some nonlocal resource; in this case a single use of a PR box. Then, it was proven that this inequality is violated by some nonmaximally entangled state.

In the present paper, we introduce $N$-input bipartite nonlocal machines (NLM), which appear as a natural extension of the two-input PR box. These machines, denoted $PR_N$, have a nice connection to a family of $N$-setting Bell inequalities known as $I_{N_2}$, similar to the one that relates the PR box to the Clauser-Horne-Shimony-Holt (CHSH) inequality.

$$\text{CHSH} \Rightarrow \text{PR box},$$

(1)
In fact, the structure of the \( N \)-input NLM can be directly deduced from the corresponding \( I_{NN2} \) inequality. Then, we present a family of \( N \)-setting inequalities, \( M_{NN2} \), which allows one use of the \( PR_{N-1} \) machine. Again, the structure of these new inequalities is easily deduced from the structure of the \( I_{NN2} \) inequalities, i.e.,

\[
I_{NN2} \Rightarrow PR_N.
\]

Thus, a nice construction appears: for any number of settings \( N \), we have a Bell inequality \( I_{NN2} \) and the related NLM, \( PR_N \), which reaches the upper (no-signaling) bound of the inequality. Adding one setting we find another inequality, \( M_{(N+1)(N+1)22} \), which cannot be violated by strategies which require a single use of \( PR_N \).

The organization of the paper is as follows. In Sec. II we present the mathematical tools and introduce the notations by reviewing the simplest case of two settings on each side. The link between the PR box and the CHSH inequality is pointed out. Section III is devoted to the case of three settings: we introduce a three-setting NLM and study an inequality for a single use of a PR box. In Sec. IV, the construction of Sec. III is extended to the case \( N \) settings. Section V concludes the paper by reviewing the main results about Bell inequalities with and without resources. Our present work is then clearly situated in this context.

II. TOOLS

Let us consider a typical Bell test scenario. Two distant observers, Alice and Bob, share some quantum state. Each of them chooses between a set of measurements (settings) \( \{A_i\}_{i=1...N_A} \), \( \{B_j\}_{j=1...N_B} \). The result of the measurement is noted \( r_A, r_B \). Here we will focus on dichotomic observables and we will restrict Alice and Bob to use the same number of settings, i.e., \( r_{A,B} \in \{0,1\} \) and \( N_A = N_B = N \). An “experiment” is fully characterized by the family of \( 4N^2 \) probabilities \( P(r_A, r_B | A_i, B_j) = P_{ij}(r_A, r_B) \) and can be seen as a point in a \( 4N^2 \)-dimensional probability space. As probabilities must satisfy

(i) Positivity:

\[
P_{ij}(r_A, r_B) \geq 0 \quad \forall \ i,j, r_A, r_B
\]

(ii) Normalization:

\[
\sum_{r_A, r_B = 0, 1} P_{ij}(r_A, r_B) = 1 \quad \forall \ i, j,
\]

all relevant experiments are contained in a bounded region of this probability space. Since we want to restrict ourselves to no-signaling probability distributions, we impose also the no-signaling conditions

\[
\sum_{r_A = 0, 1} P_{ij}(r_A, r_B) = P_j(r_B) \quad \forall \ i,
\]

\[
\sum_{r_B = 0, 1} P_{ij}(r_A, r_B) = P_i(r_A) \quad \forall \ j.
\]

Conditions (3) mean that Alice’s output cannot depend on Bob’s setting, and vice versa. This shrinks further the region of possible experiments, and the dimension of the probability space is now reduced to \( d = N(N+2) \). So, each no-signaling experiment is represented by a point in a \( d \)-dimensional probability space. In fact, the region containing all relevant probability distributions (strategies), i.e., satisfying positivity, normalization, and no-signaling, forms a polytope, i.e., a convex set with a finite number of vertices. It is the no-signaling polytope.

One can restrict the probability distributions even further by requiring that these are built only by local means, such as shared randomness. We then obtain a smaller polytope: the local polytope. The facets of this polytope are Bell inequalities, in the sense that a probability distribution lying inside (outside) the local polytope satisfies (violates) a Bell inequality. The vertices (extremal
points) of this polytope are deterministic strategies obtained by setting the outputs \( r_A \) and \( r_B \) always to 0 or always to 1. Finding the facets of a polytope knowing its vertices is a computationally difficult task. In fact, Pitowsky has shown this problem to be NP-complete.\(^{15}\) That is why all Bell inequalities have been listed for the case of two or three settings, whereas not much is known for a larger number of settings.

Let us start with a brief review of the simplest situation: two settings on each side. This case has been largely studied, and both the local and the no-signaling polytope have been completely characterized.\(^{16}\) The probability space has eight dimensions. We choose the eight probabilities \( P_i(r_A=0) \), \( P_j(r_B=0) \), and \( P_{ij}(r_A=r_B=0) \) to characterize the space.

**The local polytope.** The local polytope has 16 vertices. Fine\(^{17}\) showed that all nontrivial facets are equivalent to the CHSH inequality

\[
\text{CHSH} = -1 \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \leq 0. \tag{4}
\]

Here, the notation represents the coefficients that are put in front of the probabilities, according to

\[
P_i(r_A=0) \\ P_j(r_B=0) \\ P_{ij}(r_A=r_B=0).
\]

The extremal points (vertices) of the local polytope are deterministic strategies, i.e., for each setting Alice and Bob always output 0 or always output 1. Let us do an example: Alice outputs bit 0 for the first setting \( A_0 \) and outputs 1 for the second setting \( A_1 \); Bob always outputs 0, for both settings. This strategy corresponds to the point in probability space

\[
\begin{bmatrix} 0_{A0}1_{A1} \\ 0_{B0}0_{B1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}.
\tag{5}
\]

All probability distributions lying outside this polytope are nonlocal.

The *quantum set* is the set of correlations that can be obtained by local measurements on quantum states. Inequality (4) can indeed be violated by quantum mechanics, and the maximal violation is \( 1/\sqrt{2}-1/2 \approx 0.2071 \), obtained by suitable measurements on the singlet state. Of course the quantum set is included in the no-signaling polytope, but the converse is not true. There are no-signaling correlations that are more nonlocal than those of quantum mechanics. The PR box is indeed among these correlations.

**The no-signaling polytope.** The no-signaling polytope has 24 vertices: 16 of them are the local vertices seen before and the eight others are the nonlocal vertices. Each one of these points corresponds to a PR box. Let us make this clear. The PR box is a two-input, two-output NLM. Alice inputs \( x \) into the machine and gets outcome \( a \), while Bob inputs \( y \) and gets output \( b \). The outcomes are correlated such that \( a \otimes b = xy \). The local marginals are however completely random, i.e., \( P(a=0)=P(b=0)=1/2 \), which ensures no-signaling. In probability space, the PR box corresponds to the point

\[
\text{PR} = \frac{1}{2} \times \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.
\tag{6}
\]

According to the symmetries \( x \rightarrow x+1, \ y \rightarrow y+1, \ a \rightarrow a+1 \), there are eight ”equivalent” PR boxes. As pointed out in Ref. 16 there is a strong correspondence between the eight CHSH facets of the local polytope and the eight PR boxes. Above each CHSH inequality lies one of the PR boxes. Each PR box violates its corresponding inequality up to 0.5, which is the maximal value for
a no-signaling strategy. Formally, this correspondence is also pretty obvious by looking at Eqs. (4) and (7). To get the PR box from the CHSH inequality, proceed as follows:

Recipe. When the coefficient of a joint probability is +1 or 0 in the inequality, replace it with 0.5; when a coefficient is equal to −1, replace it with 0 in the machine.

In other words, when a joint probability appears with a coefficient +1 or 0, the outputs of the machine are correlated, and when the coefficient is −1, the outputs are anticorrelated. This simple recipe can be straightforwardly extended to Bell inequalities with more settings. For a Bell inequality with \( N \) settings, we then get a new NLM, denoted \( \text{PR}_N \). This machine has \( N \) inputs and binary outcomes (see Fig. 1).

III. MAIN RESULT—THREE SETTINGS

In this paper we present Bell-type inequalities allowing the use of some nonlocal resource. This means that all strategies satisfying such inequalities can be simulated by local means (i.e., shared randomness, etc.) together with some nonlocal resource—for example, one NLM. In other words, any strategy violating such inequalities would require a strictly larger amount of nonlocal resource than is allowed by the inequality. In the case of two settings, described in the previous section, such inequalities cannot exist. This is because the most elementary nonlocal resource, the PR box, suffices already to generate all the nonlocal vertices of the no-signaling polytope.

Therefore we switch to the next case, i.e., three settings (with two outputs) on each side. Here, the situation becomes much more complicated but remains tractable. All facets of the local polytope have been listed.\(^{13}\) No-signaling strategies are now living in a 15-dimensional space.

The local polytope. The local polytope has 64 vertices. Surprisingly, it turns out that each of the 648 nontrivial facets is equivalent to one of the two following Bell inequalities:

\[
\begin{array}{ccc}
-1 & 0 & 0 \\
0 & 0 & 0 \\
-1 & 1 & 1 \\
0 & 1 & -1 \\
\end{array} = 0,
\]

\[
\begin{array}{ccc}
-2 & -1 & 0 \\
1 & 1 & 0 \\
-1 & 1 & -1 \\
0 & 1 & 0 \\
\end{array} = 0.
\]

The CHSH inequality is still a facet of the local polytope. This is a general property of Bell inequalities, known as “lifting”\(^{18}\): a facet Bell inequality, defined in a given configuration, remains a facet when the number of settings, outcomes, or parties is augmented.

Quantum mechanics indeed violates the three-setting CHSH inequality. The second inequality, \( I_{3322} \), is also violated by quantum mechanics. Furthermore, this inequality is relevant, since it is violated by some quantum states which do not violate the CHSH inequality.\(^ {13}\)

The no-signaling polytope. The local polytope has 72 CHSH-type facets. Above each of these facets lies a PR box. This is clear since the CHSH inequality, while still being a facet of any local polytope with more settings, is a true two-input Bell inequality. Now, it is interesting to see that
above each $I_{3322}$ inequality (which is a true three-input Bell inequality) we find a no-signaling strategy which is more nonlocal than a PR box. This strategy is represented by a three-input NLM, defined by the relation $\frac{x+y}{2} = a + b \mod 2$, where $x,y \in \{0,1,2\}$ and $a,b \in \{0,1\}$. This machine will be referred to as PR$_3$. In probability space this new machine corresponds to the point

$$\frac{1}{2} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$ (10)

Note that $\langle I_{3322} | PR_3 \rangle = 1$, while $\langle I_{3322} | PR \rangle = 0.5$ (see Fig. 2). Here, we have used a scalar product-type notation $\langle I | S \rangle = z$, which means that testing inequality $I$ with strategy $S$ gives a value $z$. The machine $PR_3$ can be simply obtained from the inequality $I_{3322}$ using the Recipe mentioned at the end of Sec. II. One needs two PR boxes to simulate $PR_3$, as shown in Appendix A. $PR_3$ can also be rewritten in the elegant manner $x = y \Leftrightarrow a = b$, which corresponds to the probability point

$$\frac{1}{2} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$ (11)

The distribution (11) is indeed equivalent to (10) up to local symmetries: here, both Alice and Bob flip their outputs for their first setting.

In a recent paper Jones and Masanes$^{19}$ gave a complete characterization of all the vertices of the no-signaling polytope for any number of settings and two outcomes—note that Barrett et al. studied the reversed case: two settings and any number of outcomes.$^{16}$ From their result it is clear that all vertices of the no-signaling polytope for three settings and two outputs can be constructed with a $PR_3$.

Numerically we find all the vertices of the no-signaling polytope. We proceed as follows. First we generate all strategies that use at most one $PR_3$. These are all the strategies where Alice and Bob can choose each of their three inputs in the set $\{0d,1d,0m,1m,2m,0mf,1mf,2mf\}$. Here $0d,1d$ means that they deterministically output the value 0 or 1; $0m,1m,2m$ means that they input 0,1,2 in the machine $PR_3$; $0mf,1mf,2mf$ means that they input 0,1,2 in $PR_3$ and flip the output of the machine. Second, we remove those strategies which are inside the local polytope by testing all the 648 Bell inequalities. Finally there are 1344 strategies left which are the nonlocal vertices.
of the (three-input, two-outcome) no-signaling polytope. We find four different classes of those vertices—given in Appendix B. A curious feature of those points is that each of them violates several inequalities of the local polytope. For example, the strategy

$$PR = \frac{1}{2} \times \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

violates the CHSH inequality (8). But, it clearly also violates eight $I_{3322}$-type inequalities, among which

$$\begin{bmatrix} -1 & 0 & 0 \\ -2 & 1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \leq 0, \quad \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \leq 0. \quad (13)$$

Formally this is clear, since each of these eight $I_{3322}$ inequalities [for example, (13)] reduces to the CHSH inequality (8) once Alice’s third setting and Bob’s first setting are discarded. Figure 3 gives some geometrical intuition of the situation.

**Inequality with a PR box.** We have just seen that, in the case of three settings on each side, there are two types of NLM, the PR box and the $PR_3$, generating different types of nonlocal vertices of the no-signaling polytope. As mentioned, the $PR_3$ is a stronger nonlocal resource than the PR box—it needs two PR boxes to be simulated. Thus there is a new polytope, sandwiched between the local and the no-signaling polytopes. It is formed by all strategies that can be simulated using at most one PR box (see Fig. 2). A facet of this polytope was recently found.\(^1\) It corresponds to the inequality

$$M_{3322} = \begin{bmatrix} -2 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \leq 0. \quad (14)$$

Although $M_{3322}$ is not violated by the maximally entangled state, it is violated by a family of nonmaximally entangled states of two qubits.\(^1\) Indeed, the maximally entangled state does not
violate this inequality, since its correlations can be simulated using a single PR box. Note that
the structure of $M_{3322}$ is similar to $I_{3322}$, the only difference being the coefficient of Alice\'s first
marginal.

We prove now that $M_{3322}$ is a facet of the polytope of all strategies using at most one PR box.
This result will be extended to the case of $N$ settings in the next section.

The proof consists of two parts: first we show that no strategy with a single use of a PR box
violates $M_{3322}$; then we show that there are (at least) $N(N+2)=15$ linearly independent strategies
using at most one PR box which saturate $M_{3322}$. Here, we just sketch the idea of the proof; see
Appendix C for details.

To prove the first part, we state a Lemma. Any no-signaling strategy $S$ violating $M_{3322}$, also
violates the two following inequalities:

$$
C_1 = \begin{array}{ccc}
-1 & 0 & 0 \\
0 & 0 & 0 \\
-1 & 1 & 0 \\
0 & 1 & -1 \\
\end{array} \leq 0, \quad C_2 = \begin{array}{ccc}
-1 & 0 & 0 \\
1 & 0 & 1 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
\end{array} \leq 0.
$$

The proof is straightforward. One needs only to note that, for no-signaling strategies, joint prob-
obabilities are smaller (or equal) than their respective marginals. Then by inverting the Lemma, we
get the following proposition: if $S$ does not violate both inequalities $C_1$ and $C_2$, then $S$ does not
violate $M_{3322}$. Finally, it is obvious that with a single PR box one can violate either $C_1$ or $C_2$, but
not both at the same time.

For the second part of the proof, we find numerically eight local deterministic distributions
which saturate $M_{3322}$. Then we find 57 other strategies with one PR box saturating $M_{3322}$. Alto-
together these strategies form a hyperplane of dimension 14. This completes the proof that $M_{3322}$ is
a facet of the polytope.

IV. N SETTINGS

In this section, the results of Sec. III are extended to the case of an arbitrary number of
settings $N$. We use a family of Bell inequalities, known as $I_{NN22}$, which were proven to be facets
of the local polytope. These inequalities are generalization of the $I_{3322}$ seen before. For $N$
settings, the inequality reads

$$
I_{NN22} = -(N-3) \begin{array}{ccccccc}
-1 & 1 & 1 & \cdots & 1 & 0 \\
1 & 1 & 1 & \cdots & 1 & -1 \\
: & : & : & \cdots & : & : \\
0 & 1 & -1 & \cdots & 0 & 0 \\
\end{array} \leq 0.
$$

Using the Recipe of Sec. II, we construct a family of $N$-settings NLM

$$
PR_N = \frac{1}{2} \begin{array}{ccccccc}
1 & 1 & 1 & \cdots & 1 & 1 \\
1 & 1 & 1 & \cdots & 1 & 0 \\
1 & 1 & 1 & \cdots & 1 & 1 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
1 & 1 & 0 & \cdots & 1 & 1 \\
1 & 1 & 0 & \cdots & 1 & 1 \\
\end{array}
$$

(15)
One can simulate $PR_N$ with $N - 1$ PR-boxes. This is easily shown using a straightforward generalization of Appendix A. We think that $PR_N$ cannot be simulated with fewer than $(N - 1)$ PR-boxes, but we do not have a proof. The inequality

$$M_{(N + 1)22} \equiv - (N - 2) 1 1 \cdots - 1 0 \leq 0$$

is an $(N + 1)$-setting Bell inequality that cannot be violated by strategies which require a single use of $PR_N$, as proven in Appendix C. In (16) we have omitted a factor $(N + 1)$ in the name of the inequality for practical reasons. Again the structure of $M_{NN22}$ is similar to $I_{NN22}$, up to Alice’s first marginal: in order to get $M_{NN22}$ from $I_{NN22}$, one simply changes Alice’s first marginal to $-(N - 1)$.

So, finally we get the following nice construction. For any number of settings $N$ we have a Bell inequality $I_{NN22}$ and an $N$-input NLM ($PR_N$) which reaches the upper no-signaling bound of $I_{NN22}$. From there, we construct an $(N + 1)$-setting inequality ($M_{(N + 1)(N + 1)22}$), which cannot be violated with one use of $PR_N$, i.e.,

$$I_{NN22} \rightarrow (M_{(N + 1)(N + 1)22}, PR_{N+1}).$$

V. CONCLUSION

To conclude, we review briefly the main results concerning polytopes and Bell inequalities with and without nonlocal resources. We focus on two-outcome settings. Table I summarizes the situation. The oldest result is due to Fine, who showed that all (nontrivial) facets of the two-input, two-outcome local polytope are equivalent to the CHSH inequality. Then Collins and Gisin completely characterized the case of three settings. In particular, they showed that there is a single new inequality ($I_{3322}$) which is inequivalent to CHSH. They also found a family of facet inequalities $I_{NN22}$ of the $N$ setting local polytope, but for $N > 3$ it is not known if there are other inequalities. The vertices of the no-signaling polytope for two settings and any number of outcomes have been characterized by Barrett et al., while Jones and Masanes studied the reversed case: an arbitrary number of settings with two outcomes.

Not much is known about inequalities allowing nonlocal resources. In 2003 Toner and Bacon found inequalities allowing one bit of communication for the case of two and three settings. They showed that the correlations from measurements on any quantum state satisfy those inequalities.
ties. In the present paper we introduced a family of $N$-input NLMs $(PR_N)_{N \geq 3}$, which are a generalization of the well-known PR box. These NLMs can be derived from Bell inequalities in the same way than the PR box is derived from the CHSH inequality. Then, we presented a new family of inequalities $(M_{NN22})_{N \geq 3}$ allowing one use of $PR_N$.

For $N=3$, we get an inequality which cannot be violated with a single PR box. This inequality, presented in a previous work,\textsuperscript{11} is however violated by some nonmaximally entangled state of two qubits. Here, we checked numerically that no states of two qubits violates $M_{4422}$ and $M_{5522}$, which suggests that these states could be simulated with two PR boxes, or even a $PR_3$ box. However such model has still not been found.

**ACKNOWLEDGMENT**

We acknowledge support from the project QAP (IST-FET FP6-015848).

**APPENDIX A**

Since $(M_{3322})_{PR_3}=0.5$, $PR_3$ cannot be simulated with a single PR box. In this Appendix, we show how to construct a $PR_3$ with 2 PR boxes.

Alice and Bob each receive a trit. For each value of the trit they input one bit in each PR box. The strategy is the following:

\begin{align*}
\begin{array}{c|ccc|c|ccc}
& x_1 & x_2 & y & y_1 & y_2 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 1 & 0 \\
2 & 1 & 0 & 2 & 0 & 1 \\
\end{array}
\end{align*}

where $x, y$ denote the settings, and $x_i, y_i$ are the binary inputs into PR-box number $i$. Finally, Alice and Bob output the sum (modulo 2) of both outputs of the PR boxes. See Fig. 4. Intuitively the strategy works as follows. The first machine introduces an anticorrelation of the outputs for the pair of settings $x=1, y=2$. The second PR box does the same for $x=2, y=1$. A nice way to show that this strategy works is by computing the parity of the outputs for each pair of settings. So, we compute a parity matrix $P$ by multiplying Alice strategy by the transpose of Bob’s strategy

\begin{align*}
P = S_A S_B^T &= 
\begin{pmatrix}
0 & 0 \\
0 & 1 \\
1 & 0 \\
\end{pmatrix}
\begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0 \\
\end{pmatrix}
= 
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0 \\
\end{pmatrix}.
\end{align*}

Note that matrix $P$ has the same structure as the correlation terms of $I_{3322}$. So, Alice and Bob’s outputs are identical when a 1 appears in the inequality and different when −1 is in the inequality.
This construction is easily generalized to $N$ settings. Since $I_{NN22}^N$ has $N-1$ correlation terms equal to $-1$, one simply uses a PR box to anticorrelate the outcomes for each of those terms. Thus it can be shown that a $PR_N$ NLM is constructed with $N-1$ PR-boxes.

**APPENDIX B**

We find four classes of nonlocal vertices of the three-setting, two-outcome no-signaling polytope

$$S_1 = \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} x & x & 0 \\ 0 & 0 & 0 \\ x & x & 0 \end{pmatrix},$$

$$S_3 = \begin{pmatrix} x & x & 0 \\ x & x & 0 \\ x & x & 0 \end{pmatrix}, \quad S_4 = \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \end{pmatrix},$$

where $x = 1/2$. Class $S_1$ corresponds to strategies with a $PR_3$. They violate maximally $I_{3322}$, i.e., up to 1. Classes $S_2 - S_4$ are strategies which can be obtained with a PR box. In $S_2$, Alice and Bob have a deterministic output for one of their setting; in $S_3$, only Alice (or Bob) has a deterministic setting; in $S_4$, no one outputs deterministic values.

There are 192 vertices in class 1, 288 in class 2, 576 in class 3, and 288 in class 4. All strategies in the same class violate the same number of CHSH inequalities and the same number of $I_{3322}$ inequalities. These numbers are summarized in the table below. For each class of vertices, the number of CHSH and $I_{3322}$ inequalities violated is given:

<table>
<thead>
<tr>
<th>Class</th>
<th>CHSH</th>
<th>$I_{3322}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>$S_2$</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>$S_3$</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>$S_4$</td>
<td>4</td>
<td>24</td>
</tr>
</tbody>
</table>

**APPENDIX C**

This Appendix is divided in two parts. In the first part, it is shown that inequality

$$M_{NN22} \equiv - (N-3) \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & 1 & -1 \\ 0 & 1 & 1 & \cdots & 1 & 0 \\ 0 & 1 & -1 & \cdots & 1 & 0 \end{pmatrix}$$

cannot be violated by strategies using (at most) one $PR_{N-1}$. In the second part, we show that, for $N \leq 5$, inequality $M_{NN22}$ is tight, i.e., it is a facet of the polytope containing all strategies that use (at most) one $PR_{N-1}$. We conjecture that this is also true for any number of settings $N$. To motivate our conjecture we show that there are $2^N$ deterministic strategies lying on the
\((d-1)\)-dimensional hyperplane, where \(d=N(N+2)\) is the dimension of the probability space.

**Part 1.** We start with a Lemma.

**Lemma 1.** Let us define the two inequalities

\[
C_1^N = -(N - 3) \quad \begin{array}{cccccc}
-1 & 1 & 1 & \cdots & -1 & 0 \\
0 & 1 & -1 & \cdots & 0 & 0
\end{array}
\]

\[
C_2^N = -(N - 4) \quad \begin{array}{cccccc}
1 & 1 & 1 & \cdots & 1 & -1 \\
0 & 0 & 0 & \cdots & 0 & 0
\end{array}
\]

\[
\begin{array}{cccccc}
0 & 0 & 0 & \cdots & 0 & 0 \\
- (N - 2) & 1 & 1 & 1 & \cdots & 1 & 0 \\
- (N - 2) & 1 & 0 & 1 & \cdots & 1 & 1 \\
- (N - 3) & 1 & 0 & 1 & \cdots & 1 & -1
\end{array}
\]

Let \(S\) be a strategy with \(N\) settings for each of the two partners. \(S\) is in a probability space of dimension \(N(N+2)\). If \(S\) violates inequality \(M_{NN22}\), then \(S\) also violates both inequalities \(C_1^N\) and \(C_2^N\).

**Proof.** \(S\) violates \(M_{NN22}\), i.e.,

\[
-(N - 1)P(A_0) - \sum_{k=0}^{N-2} (N - k - 1)P(B_k) + \sum_{k=0}^{N-1} P(A_kB_0) + \sum_{m=1}^{N-1} \left( \sum_{k=0}^{N-m-1} P(A_kB_m) - P(A_{N-m}B_m) \right) \\
> 0.
\]

(C1)

According to the no-signaling condition, we have

\[
P(A_0) \geq P(A_0B_0),
\]

(C2)

\[
(N - 1)P(B_0) \equiv \sum_{k=1}^{N-1} P(A_kB_0),
\]

(C3)

\[
P(A_{N-1}B_1) \geq 0.
\]

(C4)

Inserting these relations into (C1) we get

\[
-(N - 2)P(A_0) - \sum_{k=1}^{N-2} (N - k - 1)P(B_k) + \sum_{k=0}^{N-2} P(A_kB_1) + \sum_{m=2}^{N-1} \left( \sum_{k=0}^{N-m-1} P(A_kB_m) - P(A_{N-m}B_m) \right) \\
> 0,
\]

which means \(S\) violates inequality \(C_1^N\).

Again from the no-signaling condition, we have

\[
P(A_0) \equiv P(A_0B_{N-1}),
\]

(C5)
\[ P(B_j) \equiv P(A_1B_j) \quad \forall j \in \{0, N - 1\}, \]  
\[ P(A_1B_{N-1}) \geq 0, \]  
which inserted into (C1) gives

\[ -(N - 2)P(A_0) - \sum_{k=0}^{N-3} (N - k - 2)P(B_k) + \sum_{k=0}^{N-2} P(A_0B_k) + \sum_{k=2}^{N-1} P(A_kB_0) \]

\[ + \sum_{m=1}^{N-3} \left[ \sum_{k=2}^{N-m-1} P(A_kB_m) \right] - P(A_{N-m}B_m) \]

\[ - P(A_{N-2}B_{N-2}) > 0, \]

which means \( S \) violates inequality \( C_N^2 \).

This completes the proof.

**Part 2.** For \( N = 5 \), we checked numerically that all strategies using at most one \( PR_{N-1} \) form a subspace of dimension \( d-1 \), where \( d = N(N+2) \) is the dimension of the probability space. This shows that inequality \( M_{NN22} \) is tight. We conjecture that this is also the case for any number of settings \( N \). Below we show that there are \( 2^N \) deterministic strategies on the hyperplane \( M_{NN22} = 0 \), which strongly supports our conjecture.

First we note that there are eight local strategies on the \( M_{3322} \) facet.

<table>
<thead>
<tr>
<th>0 1 1</th>
<th>0 1 1</th>
<th>0 1 0</th>
<th>0 1 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 1 1</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>0 0 0 0</td>
<td>0 0 0 1</td>
<td>0 1 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>0 0 1 0</td>
<td>0 0 1 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
<td>0 0 0 0</td>
</tr>
<tr>
<td>1 0 0 1</td>
<td>0 0 0 0</td>
<td>1 0 0 0</td>
<td>0 0 0 0</td>
</tr>
</tbody>
</table>

Obviously the marginals fix entirely a deterministic strategy. Then it is clear that if a three-setting strategy

\[
\begin{pmatrix}
A_0 & A_1 & A_2 \\
B_0 & & \\
B_1 & & \\
B_2 & & \\
\end{pmatrix}
\]

is on the facet \( M_{3322} \), then both (four settings) strategies

\[
\begin{pmatrix}
A_0 & A_1 & A_2 & 0 \\
0 & 0 & 0 & 0 \\
B_0 & & \\
B_1 & & \\
B_2 & & \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
A_0 & A_1 & A_2 & 0 \\
0 & 0 & 0 & 0 \\
B_0 & & \\
B_1 & & \\
B_2 & & \\
\end{pmatrix}
\]
are on the facet $M_{4422}$. The notation $\beta_j$ for some correlation coefficients means that their value depends on Bob’s marginal. Indeed, all these strategies are extremal since they are deterministic.

Then the argument is extended to the next case: for each of the 16 strategies (constructed above) which lie on $M_{4422}$, there are two strategies on $M_{5522}$. Thus there are $2^N$ deterministic strategies on $M_{NN22}$. Note that $2^N > N(N+2)$ for $N \geq 6$. In this case the number of local strategies on the hyperplane $M_{NN22} = 0$ is larger than the dimension of the probability space. This suggests that $M_{NN22}$ is a facet of the polytope of all strategies using at most one PR$_{N-1}$.

1. J. S. Bell, Physics 1, 195 (1964).
10. In order to construct a no-signaling strategy with classical communication, one has to cleverly hide the communication, to avoid signaling. See, for example, J. Degorre, S. Laplante, and J. Roland, Phys. Rev. A 72, 062314 (2005) (or also Ref. 11).
Detection Loophole in Asymmetric Bell Experiments

Nicolas Brunner,* Nicolas Gisin, Valerio Scarani, and Christoph Simon

Group of Applied Physics, University of Geneva, CH-1211 Geneva 4, Switzerland

(Received 13 February 2007; published 31 May 2007)

The problem of closing the detection loophole with asymmetric systems, such as entangled atom-photon pairs, is addressed. We show that, for the Bell inequality $I_{322}$, a minimal detection efficiency of 43% can be tolerated for one of the particles, if the other one is always detected. We also study the influence of noise and discuss the prospects of experimental implementation.

DOI: 10.1103/PhysRevLett.98.220403

PACS numbers: 03.65.Ud

Nonlocality is one of the most striking properties of quantum mechanics. Two distant observers, each holding half of an entangled quantum state and performing appropriate measurements, share correlations which are non-local, in the sense that they violate a Bell inequality [1].

In other words, those correlations cannot be reproduced by any local hidden variable (LHV) model. All laboratory experiments to date have confirmed quantum nonlocality [2–8]. There is thus strong evidence that nature is non-local. However, considering the importance of such a statement, it is crucial to perform an experiment free of any loopholes, which is still missing today. Another motivation comes from quantum information science, where the security of some quantum communication protocols is based on the loophole-free violation of Bell inequalities [9].

Performing a loophole-free Bell test is quite challenging. One first has to ensure that no signal can be transmitted from one particle to the other during the measurement process. Thus the measurement choice on one side and the measurement result on the other side should be spacelike separated. If this is not the case, one particle could send some information about the measurement setting it experiences to the other particle. This is the locality loophole [10]. Secondly the particles must be detected with a high enough probability. If the detection efficiency is too low, a LHV model can reproduce the quantum correlations. In this picture a hidden variable affects the probability that the particle is detected depending on the measurement setting chosen by the observer. This is the detection loophole [11,12].

In practice, photon experiments have been able to close the locality loophole [3–7]. However the optical detection efficiencies are still too low to close the detection loophole. For the Clauser-Horne-Shimony-Holt (CHSH) [13] inequality, an efficiency larger than 82.8% is required to close the detection loophole with maximally entangled states. Surprisingly, Eberhard [14] showed that this threshold efficiency can be lowered to 66.7% by using nonmaximally entangled states. Threshold efficiencies for other Bell inequalities have also been studied [15–17]. On the other hand, an experiment carried out on trapped ions [8] closed the detection loophole, but the ions were only a few micrometers apart. It would already be a significant step forward to close the detection loophole for well-separated systems. Recently, new proposals for closing both loopholes in a single experiment were reported [18,19].

In this Letter we focus on asymmetric setups, where the two particles are detected with different probabilities. This is the case, e.g., in an atom-photon system: the atom is measured with an efficiency close to 1 while the probability to detect the photon is smaller. Intuition suggests that if one party can do very efficient measurements, then the minimal detection efficiency on the other side should be considerably lowered compared to the case where both detectors have the same efficiency. Experimentally this approach might be quite promising, since recent experiments have demonstrated atom-photon entanglement [20,21] and violation of the CHSH inequality [22]. In the following, after presenting the general approach to the study of the detection loophole in asymmetric systems, we focus on the case where one of the systems is detected with efficiency $\eta_A = 1$ and we compute the threshold efficiency $\eta_B^\theta$ for the detection of the other system. The best results are obtained for the three-setting $I_{322}$ inequality [23]. In analogy to Eberhard’s result [14], we show that nonmaximally entangled states require a lower efficiency; moreover, here, the threshold goes down to ~43%. Then we study two noise models: background noise and noisy detectors. Finally, we discuss the feasibility of experiments in the light of these results.

General approach.—Let us consider a typical Bell test scenario. Two distant observers, Alice and Bob, share some quantum state $\rho_{AB}$. Each of them chooses randomly between a set of measurements (settings) $\{A_i\}_{i=1,\ldots,N_A}$ for Alice, $\{B_j\}_{j=1,\ldots,N_B}$ for Bob. The result of the measurement is noted $a, b$. Here we will focus on dichotomic observables (corresponding to von Neumann measurements on qubits) and Alice and Bob will use the same number of settings, i.e., $a, b \in \{0, 1\}$ and $N_A = N_B = N$. Repeating the experiment many times, the two parties can determine the joint probabilities $p(a, b|A_i, B_j)$ for any pairs of settings, as well as marginal probabilities $p(a|A_i)$ and $p(b|B_j)$. A Bell inequality is a constraint on those probabilities, which is satisfied for all LHV models. We say that...
a quantum state is nonlocal if and only if there are measurement settings such that a Bell inequality is violated. Mathematically speaking a Bell inequality is a polynomial of joint and marginal probabilities. In the case \( N = 2 \) the only relevant Bell inequality is the CHSH inequality, which is defined here using the Clauser-Horne polynomial \([24]\),

\[
I_{\text{CHSH}} = P(A_1B_1) + P(A_1B_2) + P(A_2B_1) - P(A_2B_2)
- P(A_1) - P(B_1),
\]

where \( P(A_1B_1) \) is a shortcut for \( P(00|A_1B_1) \), the probability that \( a = b = 0 \). The bound for LHV models is \( I_{\text{CHSH}} \leq 0 \), while quantum mechanics can reach up to \( I_{\text{CHSH}} = \frac{1}{\sqrt{2}} - \frac{1}{2} \).

We also introduce the Bell polynomial

\[
I_{3322} = P(A_1B_1) + P(A_1B_2) + P(A_1B_3)
+ P(A_2B_1) + P(A_2B_2) + P(A_2B_3)
- P(A_1B_2) - 2P(A_1) - P(A_2) - P(B_1),
\]

which is the only relevant Bell inequality for the case \( N = 3 \) \([23]\). The local limit is \( I_{3322} \leq 0 \) and quantum mechanics violates \( I_{3322} \) up to \( \frac{1}{2} \).

As an introductory example, consider the case where Alice and Bob share maximally entangled states and detect their particles with the same limited efficiency \( \eta \); since they must always produce an outcome, they agree to output “0” in case of no detection. When both detectors fire, the CHSH inequality is maximally violated, i.e., \( I_{\text{CHSH}} = Q = \frac{1}{\sqrt{2}} - \frac{1}{2} \). Here \( I_{\text{CHSH}}(d,d) \) is the value obtained for the CHSH polynomial when both parties detect their particles. When only Alice’s detector fires, \( P(A_1) = \frac{1}{2}, \) \( P(B_1) = 1, \) and \( P(A_1B_1) = \frac{1}{2} \), therefore \( I_{\text{CHSH}} = M_A = -\frac{1}{2} \); similarly, when only Bob’s detector fires, \( I_{\text{CHSH}}(d,d) = M_B = -\frac{1}{2} \).

When no detector fires, the LHV bound is reached, \( I_{\text{CHSH}} = 0 \), since \( P(A_1) = P(B_1) = P(A_1B_1) = 1 \).

Consequently, the whole set of data violates the CHSH inequality if and only if

\[
\eta^2 Q + \eta(1-\eta)(M_A + M_B) > 0,
\]

yielding the well-known threshold efficiency \( \eta > 82.84\% \).

In general, Alice and Bob test a Bell inequality \( I \leq L \) on a state \( \rho_{AB} \) having two different detection efficiencies, \( \eta_A \) and \( \eta_B \). In analogy to the previous example, Alice and Bob must choose the measurement settings \( \{A_i, B_j\} \) and the value they output in the case of no detection, in order to maximize

\[
I_{\eta_A, \eta_B} = \eta_A \eta_B Q + \eta_A (1 - \eta_B) M_A + (1 - \eta_A) \eta_B M_B + (1 - \eta_A)(1 - \eta_B)X,
\]

where \( Q = \text{Tr}(I \rho_{AB}) \) is the mean value of the Bell operator \( I \) associated to the inequality, \( M_A, B \) and \( X \) are the values of \( I \) obtained when one or both detectors do not fire. We stress that the measurement settings that maximize \( I_{\eta_A, \eta_B} \) are not those that maximize \( Q \) for the same quantum state, except for the maximally entangled state. Concerning the values assigned to the outputs by Alice and Bob in the case of no detection, we limit ourselves to dichotomic outcomes, i.e., \( a, b \in \{0, 1\} \). Note that they could also use a third outcome for no detection \([15]\).

**Case study:** \( \eta_A = 1 \).—The general approach above can be carried out for any specific values of the efficiencies; now we consider the limit where Alice’s detector is perfect, \( \eta_A = 1 \). Moreover, we consider inequalities such that \( L = 0 \). From (4) one obtains immediately that the efficiency of Bob’s detector must be above the threshold

\[
\eta_B > \eta_B^\text{th} = \frac{1}{1 - Q/M_A}
\]

in order to close the detection loophole. For any given state, the measurement settings and Bob’s output in case of no detection must be chosen as to maximize \( Q/M_A \). Note that \( M_A \leq L = 0 \), because the events where only Alice’s detector clicks never violate the Bell inequality.

Consider first pure states. For the maximally entangled state, one obtains \( \eta_B^\text{th} = 70.7\% \) for the CHSH inequality and \( \eta_B^\text{th} = 66.7\% \) for the \( I_{3322} \) inequality (the settings are those that achieve \( Q = \frac{1}{\sqrt{2}} \) \([23]\), and in the absence of detection Bob outputs 0 leading to \( M_A = -\frac{1}{2} \)). Note that a LHV model is known, which reproduces the correlations of the maximally entangled state under the assumption \( \eta_A = 1 \) and \( \eta_B = 50\% \) \([25]\); it is an interesting open question to close this gap by finding either a better Bell-type inequality, or a better LHV model.

For pure nonmaximally entangled states \( |\psi_\theta\rangle = \cos \theta |00\rangle + \sin \theta |11\rangle \), we performed a numerical minimization of \( \eta_B^\text{th} \): we find that \( \eta_B^\text{th} \) decreases with decreasing \( \theta \) both for CHSH and \( I_{3322} \), as shown in Fig. 1 (thick lines), in analogy with Eberhard’s result \([14]\). The optimal settings can always be found to lie in the \((x, z)\) plane of the Bloch sphere, i.e., \( A_1 = \cos(\alpha_i)\sigma_3 + \sin(\alpha_i)\sigma_x \) and \( B_1 = \cos(\beta_i)\sigma_z + \sin(\beta_i)\sigma_x \). In the case of no detection, we found that it is optimal for Bob to output always 0; note that in this case, \( M_A = P(A_1) = \frac{1}{4}(|\psi_\theta\rangle[A_1 \otimes 1]|\psi_\theta\rangle - 1) \) for both inequalities we consider here, CHSH and \( I_{3322} \).

In the limit of weakly entangled states \( \theta \to 0 \), one finds \( \eta_B^\text{th} \approx 50\% \) for CHSH and \( \eta_B^\text{th} \approx 43\% \) for \( I_{3322} \) (see Fig. 1). This is our main result. It is remarkable that the detection loophole can in principle be closed with \( \eta_B < 50\% \). Though we could not find an analytical expression for the optimal settings as a function of \( \theta \), we provide a numerical example: for \( \theta = 0^\circ \), \( I_{3322} \) gives \( \eta_B^\text{th} \approx 43.3\% \) \((Q \approx 0.0013 \) and \( M_A \approx -0.001) \) for the optimal settings \( \alpha_0 = -0.0012, \alpha_1 = 0.1331 \pi, \alpha_2 = 0.5494, \beta_0 = 0.0101 \pi, \beta_1 = -0.0038 \pi, \) and \( \beta_2 = -0.0924 \pi \).

We have seen that \( \eta_B^\text{th} \) decreases with the degree of entanglement for pure states. However, the violation of the inequality decreases as well. It is therefore important to study the effect of noise. We consider two models of noise. The first is background noise as in Ref. \([14]\): Alice and Bob share a state of the form

\[Q = Q \rho_{AB} + (1 - Q) \rho_n\]
function of noise, but the threshold efficiencies are lower. We have also shown in Fig. 1 (thin full lines). As expected, when $\theta$ decreases, the threshold efficiency reaches a minimum: for less entangled states the violation of the inequality is rapidly overcome by the noise. In Fig. 2, one sees that the $I_{3322}$ inefficiency can tolerate lower efficiencies than the CHSH inequality for $p \approx 6\%$.

Another noise model, probably more relevant for experiments, supposes that Alice’s and Bob’s detectors have a certain probability of error, $\epsilon_A^j$ and $\epsilon_B^j$, e.g., due to dark counts in the case of photon detection. The statistics are then described by the state

$$\rho_{AB} = (1 - p)|\psi_\theta\rangle\langle\psi_\theta| + p \frac{I}{4}.$$  

For $\theta = \frac{\pi}{2}$, the state (6) is the Werner state. The threshold efficiency for $I_{3322}$ as a function of $\theta$ is shown in Fig. 1 (thin full lines). As expected, when $\theta$ decreases, the threshold efficiency reaches a minimum: for less entangled states the violation of the inequality is rapidly overcome by the noise. In Fig. 2, one sees that the $I_{3322}$ inefficiency can tolerate lower efficiencies than the CHSH inequality for $p \approx 6\%$.

Another noise model, probably more relevant for experiments, supposes that Alice’s and Bob’s detectors have a certain probability of error, $\epsilon_A^j$ and $\epsilon_B^j$, e.g., due to dark counts in the case of photon detection. The statistics are then described by the state

$$\rho_{AB} = (1 - \epsilon_A^j)|(1 - \epsilon_B^j)|\langle\psi_\theta\rangle\langle\psi_\theta| + \epsilon_A^j \left( \frac{I}{2} \otimes \rho_B \right) + \epsilon_B^j \left( \frac{I}{2} \otimes \rho_A \right)$$

where $\rho_A = \text{Tr}_B|\psi_\theta\rangle\langle\psi_\theta|$ and $\rho_B = \text{Tr}_A|\psi_\theta\rangle\langle\psi_\theta|$ are the reduced states of Alice and Bob. In the recent atom-photon experiment done in Munich [20], the atom measurement has $\eta_A = 1$ and $\epsilon_A^j = 5\%$, whereas the photon measurement is much less efficient but also less noisy. In the light of this, we focus for definiteness on the case $\eta_A = 1$ and $\epsilon_A^j = 0$. Again, the computed threshold efficiency as a function of $\theta$ is shown in Fig. 1 (thin dashed lines). The behavior is qualitatively the same as for the background noise, but the threshold efficiencies are lower. We have also found that $I_{3322}$ can tolerate higher error rates than CHSH as soon as $\eta_B < 75\%$. Note that for both noise models, the optimal settings can be found to lie in the $(x, z)$ plane of the Bloch sphere and that the optimal strategy for Bob in the case of no detection is to output always 0.

Experimental feasibility.—Atom-photon entanglement has been demonstrated both with Cd ions in an asymmetric quadrupole trap [21,22] and with Rb atoms in an optical dipole trap [20]. Nonmaximally entangled atom-photon states were already created in Ref. [21]. The overall photon detection efficiency is very low in these experiments, mostly due to inefficient photon collection. The collection efficiency could be brought to the required level by placing the atom inside a high-finesse cavity. For example, Ref. [26] demonstrated coupling of a trapped ion to a high-finesse cavity and achieved $\beta = 0.51$, where $\beta$ is the fraction of spontaneously emitted photons that are emitted into the cavity mode. The experimental conditions in Ref. [27] correspond to $\beta$ very close to 1. In real experiments there are other sources of loss, such as propagation losses and detector inefficiency. However, detection efficiencies of order 90% have already been achieved [28], and propagation losses can be kept small for moderate distances (see below). Overall, the perspective for closing the detection loophole for two well-separated systems seems excellent using atom-photon implementations.

Performing a loophole-free Bell experiment requires enforcing locality of the measurements [4,5] in addition to closing the detection loophole. The measurement of the atomic state, which is typically based on detecting fluorescence from a cycling transition, is relatively slow. As a
consequence, enforcing locality in an experiment with atom-photon pairs requires a large separation between the two detection stations for the atom and the photon. For example, Ref. [18] estimated that for trapped Ca ions the atomic measurement could be performed in 30 μs, assuming that 2% of the photons from the cycling transition are collected, leading to a required separation of order 5 km for an asymmetric configuration [29]. For distances of this order propagation losses for the photon become significant. For example, 5 km of telecom fiber have a transmission of order 80% for the optimal wavelength range around 1.5 μm, but only of order 30% for wave-lengths around 850 nm [30]. Provided that one can achieve that, for the inequality derived by Cabello and Larsson [33].

Note added in proof.—We discussed the detection loophole in asymmetric Bell tests. In particular, we showed that, for the inequality $I_{3322}$, a minimal detection efficiency of $\eta_B = 43\%$ can be tolerated (for $\eta_A = 1$), considering nonmaximally entangled states. For maximally entangled states, the threshold efficiency is $\eta_B = 66.7\%$. For these states the LHV model of Ref. [25], based on the detection loophole, provides a lower bound for the threshold efficiency $\eta_B > 50\%$. It is an interesting question whether this bound can be reached by considering other Bell inequalities. We have found no improvement using the following inequalities: $I_{4422}$ and $I_{3322}^{2,5}$ from Ref. [23], $A_5$ from Ref. [31], and $A S_{1,2}$ from Ref. [32]. From an experimental point of view, we have argued that atom-photon entanglement seems promising for closing the detection loophole for well-separated systems. We also briefly discussed the experimental requirements for realizing a loophole-free Bell experiment using an asymmetric approach.

The authors thank A.A. Méthot, J. Volz, H. de Riedmatten, and M.Legré for discussions, and S. Pironio for pointing out a mistake in a previous version of this paper. We acknowledge financial support from the project QAP (No. IST-FET FP6-015848) and from the Swiss NCCR “Quantum Photonics” project.

Note added in proof.—While finishing the writing of this manuscript, we became aware that the results presented here about the CHSH inequality were independently derived by Cabello and Larsson [33].

*Electronic address: nicolas.brunner@physics.unige.ch

[29] In Ref. [20] the authors suggest that it may be enough for the quantum state of the atom to have collapsed, rather than for the result of the measurement to have been recorded, which leads them to a smaller required separation of 150 m. However, from the strict perspective of testing the most general hidden variable models this assumption is somewhat unsatisfactory.
Fast and simple one-way quantum key distribution

Damien Stucki, Nicolas Brunner, Nicolas Gisin, Valerio Scarani, and Hugo Zbinden

Group of Applied Physics, University of Geneva, 20, rue de l’Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland

(Received 17 June 2005; accepted 15 September 2005; published online 2 November 2005)

We present and demonstrate a new protocol for practical quantum cryptography, tailored for an implementation with weak coherent pulses to obtain a high key generation rate. The key is obtained by a simple time-of-arrival measurement on the dataline; the presence of an eavesdropper is checked by an interferometer on an additional monitoring line. The setup is experimentally simple; moreover, it is tolerant to reduced interference visibility and to photon number splitting attacks, thus featuring a high efficiency in terms of distilled secret bit per qubit. © 2005 American Institute of Physics. [DOI: 10.1063/1.2126792]

Quantum key distribution (QKD) is the only method to distribute a secret key between two distant authorized partners, Alice and Bob, whose security is based on the laws of physics. QKD is the most mature field in quantum information; nevertheless, there is still some work ahead in order to build a practical system that is reliable and at the same time fast and provably secure. In this paper we present an important improvement in this direction. The quest for rapidity is the inspiring motivation of this system: the idea is to obtain the secret bits from the simplest possible measurement (here, the time of arrival of a pulse) without introducing lossy optical elements at Bob’s. Security is obtained by occasionally checking quantum coherence: in QKD, a decrease of coherence is attributed to the presence of the eavesdropper Eve, who has attacked the line and obtained some information on the bit values, at the price of introducing errors. Reliability is achieved by using standard telecom components; in particular, the source is an attenuated laser, and bits are encoded in time bins, robust against polarization effects in fibers. In this paper, we first define the protocol and demonstrate its advantages: simplicity, and robustness against both reduced interference visibility and photon number splitting (PNS) attacks. Then, we present a first proof-of-principle experiment.

To date, the most developed setups for practical QKD implement the Bennett-Brassard 1984 (BB84) protocol using phase encoding between two time bins, as sketched in the top of Fig. 1 (see Ref. 1 for a detailed description). The four states belonging to two mutually orthogonal bases are the bits 0 and 1 in the Y basis

\[ |0\rangle = \sqrt{\mu} |2k-1\rangle |0\rangle_{2k}, \]

\[ |1\rangle = |0\rangle_{2k-1} |\sqrt{\mu} \rangle_{2k}. \]

Note that $|0\rangle$ and $|1\rangle$ are not orthogonal, due to their vacuum component; however, a time-of-arrival measurement, whenever conclusive, provides the optimal unambiguous determination of the bit value. To check coherence, we produce a fraction $f < 1$ of decoy sequences $|\sqrt{\mu} \rangle_{2k-1} |\sqrt{\mu} \rangle_{2k}$; while for BB84, one should produce the two states $|\sqrt{\mu} \rangle_{2k-1} |\sqrt{\mu} \rangle_{2k}$. Now, due to the coherence of the laser, there is a well-defined phase between any two nonempty pulses: within each decoy sequence, but also across the bit separation in the case where bit number $k$ is 1 and bit number $k+1$ is 0 (a “1–0 bit sequence”). Since we produce equally spaced pulses, the coherence of both decay and 1–0 bit sequences can be checked with a single interferometer (see Fig. 1, bottom). And there is a further benefit: coherence being distributed both within and across the bit separations, Eve cannot count the number of photons in any finite number of pulses without introducing errors. In our scheme the PNS attacks can be detected. To detect PNS attacks in BB84, one needs to complicate the protocol by the technique of decoy states, which consists of varying $\mu$.

The pulses propagate to Bob on a quantum channel characterized by a transmission $t$, and are split at a nonequilibrated beamsplitter with transmission coefficient $t_B \leq 1$. The pulses that are transmitted (dataline) are used to establish the
raw key by measuring the arrival times of the photons. The counting rate is \(R = 1 - e^{-\mu t_\eta} = \mu t_\eta\), where \(\eta\) is the quantum efficiency of the photon counter. The pulses that are reflected at Bob’s beamsplitter go to the interferometer that is used to check quantum coherence (monitoring line). Indeed, when both pulses \(j\) and \(j+1\) are nonempty, then only detector \(D_{Mj}\) can fire at time \(j+1\). Coherence can be quantified by Alice and Bob through the visibility of the interference,

\[
V = \frac{p(D_{M1}) - p(D_{M2})}{p(D_{M1}) + p(D_{M2})},
\]

where \(p(D_{Mj})\) is the probability that detector \(D_{Mj}\) fired at a time where only \(D_{M1}\) should have fired. These probabilities are small, the average detection rate on the monitoring line being \(\frac{1}{2} \mu t_\eta \eta\) per pulse. Still, if the bit rate is high, meaningful estimates can be done in a reasonable time.

Let’s summarize the protocol.

1. Alice sends a large number of times “bit 0” with probability \((1-f)/2\), “bit 1” with probability \((1-f)/2\) and the decoy sequence with probability \(f\).
2. At the end of the exchange, Bob reveals for which bits he obtained detections in the dataline and when detector \(D_{2M}\) has fired.
3. Alice tells Bob which bits he has to remove from his raw key, since they are due to detections of decoy sequences (sifting).
4. Analyzing the detections in \(D_{2M}\), Alice estimates the break of coherence through the visibilities \(V_{1-d}\) and \(V_d\) associated, respectively, with 1–0 bit sequences and to decoy sequences, and computes Eve’s information.
5. Finally, Alice and Bob run an error correction and a privacy amplification and end up with a secret key.

The performance of a QKD protocol is quantified by the achievable secret key rate \(R_{sk}\). To compute this quantity, we need to introduce several parameters. The fraction of bits kept after sifting (sifted key rate) is \(R_{sk}[\mu] = \frac{1}{2} + 2p_x(1 - f)\), with \(R = \mu t_\eta\) the counting rate due to photons defined above, \(p_d\) the probability of a dark count, and \(p_x = 1 - f\) the amount of errors in the sifted key is called the quantum bit error rate (QBER, \(Q\)). Moreover, this key is not secret: Eve knows a fraction \(I_{Eve}\) of it. Some classical post-processing (error correction and privacy amplification) allows us to extract a key that is errorless and secret, while removing a fraction \(h(Q) + I_{Eve}\), where \(h\) is binary entropy. Then,

\[
R_{sk} = R_{sk}[\mu][1 - h(Q) - I_{Eve}].
\]

With this figure of merit, we can compare our scheme to BB84 implemented using the interferometric bases \(X\) and \(Y\), as it is done today, with an asymmetric use of the bases such that \(p_x = 1 - f\) (BB84\(_{XY}\)). We require that all the visibilities are equal: \(V_X = V_Y = 0.9\) in BB84\(_{XY}\), \(V_{1-d} = V_d\) in our scheme—otherwise, Alice and Bob abort the protocol. Under this assumption, the QBER of BB84 is \(Q(\mu) = \frac{1}{2}(1 - V)/2\) and \(I_{Eve} = I_{Eve} + I_{Eve}\), where in our scheme \(Q(\mu) = I_{Eve}\), independent of \(V\).

In order to estimate \(I_{Eve}\), we restrict the class of Eve’s attacks, waiting for a full security analysis. Because of losses and the existence of multiphoton pulses, Eve can gain full information on a fraction of the bits without introducing any errors. This fraction is either \(r = \mu(1-t)\) or \(r = \mu/2t\), according to whether PNS attacks do not or do introduce errors. Then Eve performs the intercept-resend attack on a fraction \(p_{IR}\) of the remaining pulses. In BB84\(_{XY}\), she introduces the error \((1-r)p_{IR} = (1-V)/2\) and gains the information \(I = (1-r)p_{IR} = V\). On the present protocol, the IR will be performed in the time basis, so \(I = (1-r)p_{IR}\). However, since we use only one decoy sequence, if Eve detects a photon in two successive pulses she knows what sequence to prepare; the introduced error is then \(1-V = I\xi\) with \(\xi = 2e^{-\mu t}/(1 + e^{-\mu})\) the probability that Eve detects something in one pulse and nothing in the other. Plugging \(Q(\mu)\) and \(I_{Eve} = r + I\) into Eq. (4), we have \(R_{sk}\) as an explicit function of \(\mu\); Alice and Bob must choose \(\mu\) in order to maximize it. The result of numerical optimization is shown in Fig. 2.

As expected, the present protocol is more robust than BB84\(_{XY}\) against the decrease of visibility.

We show that a reasonably low QBER and good visibility can be obtained using standard telecom components in an implementation with optical fibers. The experimental setup is sketched in Fig. 3. The light of a cw laser (wavelength 1550 nm) passes through an intensity modulator (IM), which prepares the chosen pulse sequence. For simplicity, we send always the same eight-pulse sequence as shown in the figure, namely the string \(DD10\), where \(D\) stands for a decoy sequence. The frequency of 434 MHz of clock \(C_1\) defines the time \(\tau\) between two successive pulses. The frequency of logical bits in a sequence is half this frequency. The clock \(C_2\) at
The visibility of the interfering pulses on detector $D_M$ is measured by varying the phase (i.e., the temperature) of the interferometer. The raw visibility is $V_{\text{raw}} \approx 92\%$, if we consider 1.7 ns time windows. The net visibility, obtained deducing the dark counts and afterpulses, is $V = 98\%$. We attribute the slight reduction of the visibility to a nonperfect overlap of the interfering pulses due to timing jitter and fluctuations in the intensity modulation. However, this reduced visibility has no significant consequence on the secret key rate (Fig. 2). This tolerance in visibility simplifies the adjustment of the interferometers. With our basic thermal stabilization the interferometer needed to be readjusted only about every 30 min. Indeed, for our pathlength difference, a temperature stability of 0.01 K guarantees $V \geqslant 80\%$. Note that, as the clock frequency of $C_1$ increases, the stabilization of the interferometer becomes easier.

We have introduced a scheme for QKD and presented the experimental results. The scheme features several advantages: The dateline is very simple, with low losses at Bob’s side and small optical QBER. The scheme is tolerant against reduced interference visibility and is robust against PNS attacks (thus allowing the mean photon number to be large, typically $\mu = 0.5$). Finally, it is polarization insensitive. The existence of such a scheme shows that the main limiting parameter for practical quantum cryptography are the imperfections of the detectors.

The authors acknowledge financial support from the Swiss NCCR “Quantum photonics” and the European Project SECOQC, and thank Avanex for the loan of an intensity modulator.

---


4Our dateline is that of a classical communication channel, but with a photon counter. The same line is used in a different QKD protocol: T. Debuisschert and W. Boucher, Phys. Rev. A 70, 042306 (2004).


6N. Gisin et al., quant-ph/0411022.


9Alternatively, the source could be a pulsed mode-locked laser followed by a pulse picker.

10If dark counts can be neglected ($Q_{\text{dark}} = 0$), the optimization can be done analytically: for the present protocol, $\mu_{\text{opt}} = f(V)/[2(1-V)-f(V)]$; for BB84, $\mu_{\text{opt}} = V/[2(1-V)-f(V)]$ with, and $\mu_{\text{opt}} = f(V)$ without decoy states, where $f(V) = V - h[(1-V)/2]$). For simplicity, in the optimization we have taken $t_\text{opt} = 1$ for both protocols, although this may be technically harder to achieve in BB84 because there are more optical components.

R. Thew et al. (in preparation).
Secrecy extraction from no-signaling correlations

Valerio Scarani, Nicolas Gisin, and Nicolas Brunner
Group of Applied Physics, University of Geneva, 20, rue de l’Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland

Lluis Masanes
School of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom

Sergi Pino and Antonio Acín
ICFO—Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels, Barcelona, Spain

(Received 27 June 2006; published 30 October 2006)

Quantum cryptography shows that one can guarantee the secrecy of correlation on the sole basis of the laws of physics, that is, without limiting the computational power of the eavesdropper. The usual security proofs suppose that the authorized partners, Alice and Bob, have a perfect knowledge and control of their quantum systems and devices; for instance, they must be sure that the logical bits have been encoded in true qubits and not in higher dimensional systems. In this paper, we present an approach that circumvents this strong assumption. We define protocols, both for the case of bits and for generic $d$-dimensional outcomes, in which the security is guaranteed by the very structure of the Alice-Bob correlations, under the no-signaling condition. The idea is that if the correlations cannot be produced by shared randomness, then Eve has poor knowledge of Alice’s and Bob’s symbols. The present study assumes on the one hand that the eavesdropper Eve performs only individual attacks (this is a limitation to be removed in further work), and on the other hand that Eve can distribute any correlation compatible with the no-signaling condition (in this sense her power is greater than what quantum physics allows). Under these assumptions, we prove that the protocols defined here allow extracting secrecy from noisy correlations, when these correlations violate a Bell-type inequality by a sufficiently large amount. The region in which secrecy extraction is possible extends within the region of correlations achievable by measurements on entangled quantum states.

DOI: 10.1103/PhysRevA.74.042339 PACS number(s): 03.67.Dd, 03.65.Ud

I. INTRODUCTION

Quantum physics has been shown to provide a means to distribute correlations at a distance whose secrecy can be guaranteed by the laws of physics, without any assumption on the computational power of the eavesdropper. This is the nowadays largely studied field of quantum cryptography (or quantum key distribution, QKD), the most mature development of quantum information science [1]. The fact itself, that quantum physics can be used to distribute secrecy, is safe: if the authorized partners share a maximally entangled state, then secrecy is definitely guaranteed. But of course, one must verify that secrecy is not immediately spoiled by any small departure from this ideal case; this is why much theoretical research has been devoted to the derivation of rigorous bounds for the security of quantum cryptography [2]. Still, a lot of questions remain unsolved: for instance, the theorists, who find security proofs, and the experimentalists, who realize devices, tend to make different and often incompatible assumptions when figuring out their schemes.

In particular, an assumption in theoretical proofs has gone unnoticed until recently [3]: one assumes that the logical bits are encoded in quantum systems whose dimension is under perfect control (generally, qubits). Why do we question this assumption? First, because it is interesting in itself to ask whether one can remove an assumption, that is, whether one can base the studies of security on weaker constraints [4]. Second, because side channels are a serious issue in practical quantum cryptography. Experimentalists have to be careful that, when they encode (say) polarization, they encode only polarization, and that the device does not change the spectral line, or the spatial mode, or the temporal mode of the photon as well. Third, because it is important for practical reasons: quantum cryptography is becoming a commercial product. If a security expert recommends a quantum cryptography device, he should be able to assess that the device acts as it should with “reasonable” means. After all, the eavesdropper Eve could be herself the provider of the device!

Anyone faced with this scenario feels at first that if Eve is allowed to sell you the devices and you cannot know them in detail, there is no hope for security. Surprisingly, recent advances in quantum information suggest that this despair, reasonable as it is, may be too pessimistic. Let us see where the hope lies and which assumptions are really crucial.

The scheme to distribute correlations we have in mind is represented in Fig. 1. In Alice’s and Bob’s laboratories, the dark gray square represents the device possibly provided by Eve. The distribution of correlations is made in three steps. In the first step, both laboratories are open to the signal that correlate them. This signal comes either from outside, or is emitted by Alice’s device to Bob’s, or vice versa: in any case, it must be assumed to be under Eve’s full control. In the second step, the laboratories are completely sealed, an obviously necessary condition as we are going to see. On the device that reads the signal, Alice and Bob must have a knob which allows them to choose among at least two alternatives (in usual QKD, this is, for instance, the choice of the basis). It is obviously necessary to assume that no information about the position of the knob leaks out of Alice’s and Bob’s laboratories (in QKD, if Eve would know the basis, she can
measure the state without introducing errors). Now, conditioned on the choice of an input (a position of the knob labeled for Alice and for Bob), an output is produced ( for Alice, for Bob). The lists of and constitute the raw key. How can there be some secrecy in this raw key? The insight from quantum physics is that the outputs may not be under the provider’s control: if the probability distribution of the outputs violates some kind of Bell inequality, then by definition those outputs have not been produced by shared randomness—in other words, the correlations have been produced by the measurements themselves and did not pre-exist to them. They could have been produced by communication, if information about the inputs and/or would have propagated between Alice and Bob; but we have insisted on the no-signaling assumption: no information about and should leak out of Alice’s and Bob’s laboratories, respectively [5]. The third step is usual: Alice and Bob can make classical data processing in order to distill a fully secret key.

The reasoning above is exactly the intuition that led Ekert to discover (independently of previous works) quantum cryptography in 1991 [6]. Ekert’s work contains in nuce the idea of a device-independent security proof: it should be possible to demonstrate that a probability distribution, which violates some Bell inequality, is secure by this very fact, without any reference to the formalism of quantum physics. Of course, in physics as we know it today, a Bell inequality can only be violated with entanglement: that is why people immediately used the quantum formalism to study Ekert’s intuition [7]. But recently, tools have been developed that allow one to study no-signaling distributions in themselves, without the formalism of Hilbert spaces. It is then possible to come back to the original intuition by Ekert and try to prove security only through the violation of a Bell-type inequality. This is the main theme of the present paper.

Since we need to introduce in more detail the tools used in this work, we do this in Sec. II and postpone the outline of the paper to Sec. II D.

II. CRYPTOGRAPHY IN THE NO-SIGNALING POLYTOPE

This first section introduces the language and the tools which are needed in a general framework. We focus from the very beginning on bipartite correlations, i.e., correlations involving two partners, traditionally called Alice and Bob.

A. Formalization of Bell-type experiments

The physical situation one must keep in mind is a Bell-type experiment. Alice and Bob receive several pairs of entangled quantum particles. On each particle, Alice performs the measurement randomly drawn from a finite set of possibilities; as a result, she obtains the output out of a discrete set containing symbols. Independently from Alice, Bob performs the measurement randomly drawn from a finite set of possibilities; as a result, he obtains the output out of a discrete set containing symbols. Such an experiment is characterized by the family of probabilities

\[ P(a, b, x, y) = P(a, b|x, y)P(x)P(y). \] (1)

There are such numbers, so each experiment can be described by a point in a -dimensional space; more precisely, in a region of such a space, bounded by the conditions that probabilities must be positive and sum up to one. By imposing further restrictions on the possible probability distributions, the region of possible experiment shrinks, thus adding nontrivial boundaries [8–10]. For our study, three restrictions are meaningful.

The first restriction is the requirement that the probability distribution must be built without communication, only with shared randomness. In the literature, this has been known as the hypothesis of local hidden variables. In our context, these variables are not hidden “in nature” (as in the original interpretational debates about quantum physics); they may rather be hidden in Alice’s and Bob’s laboratories, in the devices that Eve has provided to them. The bounded region, which contains all probability distributions that can be obtained by shared randomness, forms a polytope, that is a convex set bounded by a finite number of hyperplanes (“facets”); therefore we refer to it as to the local polytope. The vertices of the local polytope are the points corresponding to deterministic strategies. The vertices are thus easily listed, but to find the facets given the vertices is a computationally hard task. The importance of finding the facets is pretty clear. If a point, representing an experiment, lies within the polytope, then there exists a strategy with shared randomness (a local variable model) that produces the same probability distribution. On the contrary, if a point lies outside the local polytope, then the experiment cannot be reproduced with shared randomness only. The interpretation of the

FIG. 1. A pictorial description of the no-signaling assumption in our context. The dark gray boxes in Alice’s and Bob’s laboratories are the devices provided by Eve. In a first step, the laboratories are open for the signal that correlates them (gray spheres). The arrows on the channel indicate that it is not important whether this signal comes from outside, or is emitted by Alice’s device to Bob’s, or vice versa: in any case, it is under Eve’s control. What is important, is that the inputs have not been chosen yet. In a second step, the laboratories are absolutely closed: no leakage of information about the inputs or the outputs is allowed. In a third step (not shown), Alice and Bob can carry out the usual procedures of error correction and privacy amplification by communicating on an authenticated channel.
facets of the local polytope is therefore obvious: they correspond to Bell’s inequalities. We shall call the region which lies outside the local polytope a nonlocal region.

The second restriction is the requirement that the probability distribution must be obtained from measurements on quantum bipartite systems. The bounded region thus obtained shall be called the quantum region. It is not a polytope, since there is not a finite set of extremal points. It is a convex set if one really allows all possible measurements on all possible states in arbitrary-dimensional Hilbert space [9,11]; if one restricts to the measurements on a given state, or even to von Neumann measurements on a Hilbert space with given dimension, convexity is not proved in general (although no counterexample is known, to our knowledge).

Needless to recall, the quantum region contains the local polytope but is larger than it; measurement on quantum states can give rise to nonlocal correlations (Bell inequalities are violated).

The third restriction is the requirement that the probability distribution must not allow signaling from Alice to Bob or vice versa. The no-signaling requirement is fulfilled if and only if Alice’s marginal distribution does not depend on Bob’s choice of input, and vice versa: that is, the probability distributions must fulfill

\[ \sum_b P(a,b|x,y) = P(a|x), \]

and

\[ \sum_a P(a,b|x,y) = P(b|y). \]

These conditions define again a polytope, the no-signaling polytope, which contains the quantum region. The deterministic strategies are still vertices for this polytope: to these, one must add other vertices which represent, loosely speaking, purely nonlocal no-signaling strategies. These additional points, sometimes called nonlocal machines or nonlocal boxes, have been fully characterized only in a few cases.

B. Secrecy of probability distributions

Here is the question that we are going to address in this paper. Alice and Bob have repeated many times the “measurement” procedure and share an arbitrary large number of realizations of the random variables distributed according to \( P(a,b|x,y) \). By revealing a fraction of their lists, they can estimate whether their probability distribution lies in the local polytope or in the nonlocal region. The goal is to study whether Alice and Bob can extract secrecy out of their data with this knowledge only.

To motivate the question, let us consider the best-known quantum cryptography protocol, the one invented by Bennett and Brassard in 1984 (BB84) [12]. In this protocol, \( a, b \in \{0,1\} \) and \( x, y \in \{X,Z\} \) are both binary. In the absence of any error, the BB84 protocol distributes perfect correlations when \( x = y \) and no correlations when \( x \neq y \), that is, \( P(0,0|X,X) = P(1,1|X,X) = \frac{1}{2} \), \( P(0,0|Z,Z) = P(1,1|Z,Z) = \frac{1}{2} \), and \( P(a,b|X,Z) = P(a,b|Z,X) = \frac{1}{4} \). If Alice and Bob have obtained their results by measuring two-dimensional quantum systems (qubits), such correlations provide secrecy under the usual assumption that the eavesdropper is limited only by the laws of quantum physics [13]. However, this distribution can also be obtained with shared randomness: if Alice and Bob would share randomly distributed pairs of classical bits \( (r_X, r_Z) \), they simply have to output \( r_Z \) (respectively \( r_X \)) if they are asked to measure \( Z \) (respectively \( X \)). Thus, we see the importance of the additional assumption on the physical realization, namely, that both Alice and Bob are measuring a qubit, and therefore the pair \( (r_X, r_Z) \) is not available because \( [X,Z] \neq 0 \).

In other words, the correlations of BB84, even in the absence of errors, are not secure by themselves. They are secure only provided the quantum degrees of freedom are under good control. The question we raised can now be put in its true perspective: Are there correlations that are secure by themselves, by the very fact of being what they are, without having even to describe how Alice and Bob managed to obtain them from a real channel?

It turns out that it is easier to tackle this question by considering that the eavesdropper Eve is not even limited by quantum physics but only by the no-signaling constraint. This means that Eve can distribute any many-instances probability distribution \( P(\tilde{a}, \tilde{b} | \tilde{x}, \tilde{y}) \) that lies within the no-signaling polytope; Alice and Bob have the freedom of choosing their sequence of measurements (\( \tilde{a} \) and \( \tilde{y} \), respectively) and will obtain the corresponding outcomes. By making this assumption, we stand clearly on the conservative side: if we can demonstrate that a nonvanishing secret key can be extracted against such a powerful eavesdropper, then the secret key achievable against a “realistic” (i.e., quantum) eavesdropper will be at least as long.

In quantum cryptography, secrecy relies on entanglement. On which physical quantity can such a strong security, as the one we are asking for, rely? The answer is: On the nonlocality of the correlations, that is, on the fact that the correlations cannot be obtained by shared randomness [5]. No secrecy can be extracted if Alice and Bob share a probability distribution which lies within the local polytope, just as no secrecy against a quantum Eve can be extracted out of separable states [14,15].

C. Individual eavesdropping strategies

Barrett, Hardy, and Kent [4] have shown an example of a protocol in which quantum correlations can provide secrecy against the most powerful attack by a no-signaling Eve. This is the first example that one can achieve security even against a supraquantum Eve, showing that security in key distribution arises from general features of no-signaling distributions rather than from the specificities of the Hilbert space structure. However, their example has important limitations: actually, it provides a protocol to distribute a single secret bit (hence, zero key rate) in the case when Alice and Bob share correlations that can be ascribed to noiseless quantum states.

In this paper, we tackle the problem from the other side: we do not go straight for security against the most powerful adversary, but we follow the same path that was followed historically by quantum cryptography, namely, we limit the eavesdropper to adopt an individual strategy. This means the
following: Eve follows the same procedure for each instance of measurement—that is, she is not allowed to correlate different instances. Moreover, Eve is asked to put her input $z$ before any error correction and privacy amplification. Consequently, any individual attack is described of a three-partite probability distribution $P(a,b,e|x,y,z)$ such that

$$P(a,b|x,y) = \sum_e P(e|z)P(a,b|x,y,e,z).$$

(4)

Note that Eve is also limited by no-signaling; that is why the left-hand side does not depend on $z$. One can see that this is an individual attack by looking at it as follows: when Eve gets outcome $e$ out of her input $z$, she sends out the point $P(a,b|x,y,e,z)$.

Now we demonstrate two similar, important results about individual eavesdropping strategies:

Theorem 1. Eve can limit herself in sending out extremal points of the no-signaling polytope.

Proof. Suppose that an attack is defined in which one of the $P(a,b|x,y,e,z)$ is not an extremal point. Then, this point can be itself decomposed on extremal points:

$$P(a,b|x,y,e,z) = \sum_\lambda P(\lambda)P(a,b|x,y,e,z,\lambda)$$

where the $P(a,b|x,y,e,z,\lambda)$ are all extremal. But the knowledge of $\lambda$ must be given to Eve. By redefining Eve’s symbol as $(e,\lambda) \rightarrow e$, we have an attack which is as powerful as the one we started from and is of the form (4) while having only extreme points in the decomposition.

Theorem 2. Suppose that Alice and Bob can transform $P(a,b|x,y)$ into $\tilde{P}(a,b|x,y)$ by using only local operations and public communication independent of $a,b,x,y$. Then there exists a purification of $\tilde{P}(a,b|x,y)$ that gives Eve as much information as the best purification of $P(a,b|x,y)$.

Proof. Suppose (4) is the best purification of $P$ from Eve’s point of view; for clarity, let us use Theorem 1 to say that the $P(a,b|x,y,e,z)$ are extremal points. The procedure of Alice and Bob can be described as follows: for each realization of the variables $(a,b,x,y)$, Alice draws a random number and reveals publicly its value $j$; then, she and Bob apply the local transformation $T_j$ on which they have previously agreed, transforming $x \rightarrow X_j, a \rightarrow A_j, \text{ etc.}$ Since there is no correlation between $j$ and $(e,z)$, each extremal point $P(e|z,a,b|x,y,e,z)$ is transformed into

$$\tilde{P}(a,b|x,y,e,z) = \sum_j P(j)P(A_j,B_j|X_j,Y_j,e,z)$$

$$= \sum_{j,e} P(j)P(e|j)P(a,b|x,y,e,z).$$

(5)

Consequently, $\tilde{P}(a,b|x,y)$ is a mixture of the extremal points $P(a,b|x,y,e,z)$ with weight $P(e|z) = \sum_j P(e|z)P(j) \times P(e|j)$. To conclude the proof, just notice that Eve has been able to follow the full procedure, because she has learned $j$ and the list of the $T_j$ is publicly known. Thus, there exists a decomposition of $\tilde{P}(a,b|x,y)$ onto extremal points that gives Eve as much information as the best decomposition of $P(a,b|x,y)$.

D. Outline of the paper

This is all that could be said in full generality. In what follows, we study mainly scenarios in which $x,y \in \{0,1\}$. Alice and Bob choose between two possible measurements. In Sec. III, we address the case where also the outcomes $a$ and $b$ are both binary; apart from Sec. III D, all the results of this section have been announced in Ref. [3]. In Sec. IV, we explore the case where both the outcomes are $d$-valued, in particular for $d=3$. In both situations, we shall consider an explicit protocol for Alice and Bob, without claim of optimality. Conclusions and perspectives are listed in Sec. V.

III. BINARY OUTCOMES

In this section, we consider $m_A=m_B=n_A=n_B=2$, that is, $a,b,x,y \in \{0,1\}$ are all binary. Below, all the sums involving bits are to be computed modulo 2.

A. The polytopes and the quantum region

In the case of binary inputs and outputs, the local and the no-signaling polytopes have been fully characterized and their structure is rather simple. A lot (but not all) is known about the quantum region too. Under no-signaling, the full probability distribution is entirely characterized by eight probabilities; therefore, all these objects live in an eight-dimensional space.

The local polytope [16,17] has eight nontrivial facets. Up to symmetries like relabeling of the inputs and of the outputs, they are all equivalent to the Clauser-Horne-Shimony-Holt (CHSH) inequality [18]. The representative of this inequality reads

$$I_{CHSH} = \sum_{x,y=0}^{1} P(a+b=xy|xy) \leq 3.$$  

(6)

On each facet lie eight out of the sixteen deterministic strategies; these are said to saturate the inequality, because by definition they give $I_{CHSH}=3$. Note that the eight points on a facet are linearly independent from one another [19]. The deterministic strategies that saturate our representative (6) are readily seen to be the following ones:

$$L'_1 = \{a(x)=r, b(y)=r\},$$

$$L'_2 = \{a(x)=x+r, b(y)=r\},$$

$$L'_3 = \{a(x)=r, b(y)=y+r\},$$

$$L'_4 = \{a(x)=x+r, b(y)=y+r+1\},$$

(7)

where $r=0,1$.

The no-signaling polytope [10] is obtained from the local polytope by adding a single extremal nonlocal point on top of each CHSH facet. The nonlocal point on top of our representative is defined by
PLY the entries of the table by pretty clear. For instance, one finds that points are summarized in Table I. The reading of this table is the CHSH inequality. It follows that in the quantum region

In other words, the set of correlations that are producible by measuring quantum states is known and corresponds to what can be produced by measurement on two-qubit states [21]. It is an open question whether the analysis of the marginals can reveal further features of the quantum region. The violation of CHSH is bounded by

\[ I_{\text{CHSH}}(QM) \leq 2 + \sqrt{2}, \]  

where the maximum is reached with the probability distribution

\[ P(a + b = xy|xy) = \frac{1 + \sqrt{2}}{2}, \]

obtained by measuring suitable observables on a maximally entangled state.

**B. The nonlocal raw probability distribution**

To study the possibility of secret key extraction, we can restrict our attention to the sector of the nonlocal region that lies above a given facet of the local polytope, say the representative one for which we have collected the tools above. Any point in this sector, by definition, can be decomposed as a convex combination of the PR box (8) and of the eight deterministic strategies on the facet (7). As shown above, we can assume without loss of generality that Eve distributes these nine strategies. We shall write \( p_{NL} \) (for nonlocal) the probability that Eve sends the PR box to Alice and Bob, and \( p'_{j} \) the probability that Eve sends the deterministic strategy \( L'_{j} \). We shall also write \( p_{j} = \sum_{j} p'_{j} = 1 - p_{NL} \).

The statistics generated by Eve sending the extremal points are summarized in Table I. The reading of this table is pretty clear. For instance, one finds that \( P(a + b = 0|xy = 0) = p_{NL}/2 + p'_{0} + p'_{3} + p'_{5} \). To obtain \( P(a, b, x, y) \), one must multiply the entries of the table by \( P(x)P(y) \). Since we are supposing that the extremal points are sent to Alice and Bob by Eve, the label of each point can be considered also as Eve’s symbol.

Finally, we note that using Table I in (6), one finds

\[ p_{NL} = I_{\text{CHSH}} - 3. \]  

In other words, \( p_{NL} \) measures directly the violation of the CHSH inequality. It follows that in the quantum region

\[ p_{NL}(QM) \leq \sqrt{2} - 1 = 0.414. \]  

**C. The CHSH protocol for cryptography**

Whenever Eve distributes the PR box, she has no information at all about the bits received by Alice and Bob because of the monogamy of those correlations [10]. On the contrary, when she distributes a deterministic strategy she has some information, depending on the actual cryptographic protocol. The question is thus: Which is the best procedure to extract a secret key out of the raw distribution of Table I? We have no answer in full generality, but we can notice a few things and propose a protocol which is a reasonable candidate for optimality. A good cryptography protocol should (i) present high correlations between Alice and Bob and (ii) reduce Eve’s information as much as possible. Now, in the raw data, we see that Alice and Bob are highly anticorrelated when \( x = y = 1 \). It is thus natural to devise a procedure that allows them to transform these anticorrelations in correlations. A good procedure reveals as little as possible information on the public channel.

The protocol we propose, and that we call CHSH protocol for obvious reasons, is the following:

1. **Distribution.** Alice and Bob repeat the measurement procedure on arbitrarily many instances and collect their data.
2. **Parameter estimation.** By revealing publicly some of their results, they estimate the parameters of their distribution, in particular the fraction \( p_{NL} \) of intrinsically nonlocal correlations.
3. **Pseudosifting.** For each instance, Alice reveals the measurement she has performed \( (x = 0 \) or \( x = 1 \)). Whenever Alice declares \( x = 1 \) and Bob has chosen \( y = 1 \), Bob flips his bit. Bob does not reveal the measurement he has performed. This is the procedure which transforms anticorrelations into correlations while revealing the smallest amount of information on the public channel. We call it pseudosifting, because it enters in the protocol at the same place as sifting occurs in

<table>
<thead>
<tr>
<th>( A \backslash B )</th>
<th>( y = 0, b = 0 )</th>
<th>( y = 0, b = 1 )</th>
<th>( y = 1, b = 0 )</th>
<th>( y = 1, b = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x = 0 )</td>
<td>( p_{NL}/2 )</td>
<td>( p_{NL}/2 )</td>
<td>( p_{NL}/2 )</td>
<td>( p_{NL}/2 )</td>
</tr>
<tr>
<td>( a = 0 )</td>
<td>( p_{0}^{0} (L_{0}) )</td>
<td>( p_{0}^{0} (L_{0}) )</td>
<td>( p_{0}^{0} (L_{0}) )</td>
<td>( p_{0}^{0} (L_{0}) )</td>
</tr>
<tr>
<td>( a = 1 )</td>
<td>( p_{1}^{0} (L_{0}) )</td>
<td>( p_{1}^{0} (L_{0}) )</td>
<td>( p_{1}^{0} (L_{0}) )</td>
<td>( p_{1}^{0} (L_{0}) )</td>
</tr>
<tr>
<td>( x = 1 )</td>
<td>( p_{NL}/2 )</td>
<td>( p_{NL}/2 )</td>
<td>( p_{NL}/2 )</td>
<td>( p_{NL}/2 )</td>
</tr>
<tr>
<td>( a = 0 )</td>
<td>( p_{0}^{1} (L_{1}) )</td>
<td>( p_{0}^{1} (L_{1}) )</td>
<td>( p_{0}^{1} (L_{1}) )</td>
<td>( p_{0}^{1} (L_{1}) )</td>
</tr>
<tr>
<td>( a = 1 )</td>
<td>( p_{1}^{1} (L_{1}) )</td>
<td>( p_{1}^{1} (L_{1}) )</td>
<td>( p_{1}^{1} (L_{1}) )</td>
<td>( p_{1}^{1} (L_{1}) )</td>
</tr>
</tbody>
</table>
Remarkably, the protocol we have defined exhibits a feature which is also present in quantum cryptography, namely, the fact that Eve gains information on a “basis” at the expense of introducing errors in the complementary one. Here it is precisely.

Refer to Table II, recalling that $\xi_0 = \xi = \frac{1}{2}$. The probabilities $p(a \neq b | x)$ of error between Alice and Bob when $x = 0$ and $x = 1$ are, respectively

$$e_{AB|0} = \frac{1}{2} (p_0^0 + p_0^1 + p_0^2 + p_0^3),$$

$$e_{AB|1} = \frac{1}{2} (p_1^0 + p_1^1 + p_1^2 + p_1^3).$$

Eve’s uncertainty on Bob’s symbol, measured by conditional Shannon entropy, is

$$H(B | E, x = 0) = 1 - (p_0^0 + p_1^1 + p_2^0 + p_3^1),$$

$$H(B | E, x = 1) = 1 - (p_0^0 + p_0^1 + p_2^0 + p_1^1).$$

Thus, there appears in our protocol a cryptographic uncertainty relation in the form
\[ H(B|E,x) = 1 - 2\epsilon_{AB|x+1}. \] (18)

The origin of this relation is rather clear. The pseudosifting phase of the protocol is optimized to extract correlations from the nonlocal strategy (PR box), but on deterministic strategies, the pseudosifting has another action. Specifically: for \( L_1 \) and \( L_2 \), after pseudosifting we have \( b(y=0)=b(y=1) \) \( \equiv a \) when \( x=0 \) (no error, and Eve knows \( b \)), and \( b(y=0) \neq b(y=1) \) when \( x=1 \) (error in half cases, and Eve does not know \( b \)); for \( L_1^\prime \) and \( L_2^\prime \), it is just the opposite. In summary, for each local strategy, Eve learns everything only for one Alice’s setting, and for the other an error between Alice and Bob occurs half of the time.

This is the first evidence of an analog of quantum mechanical uncertainty relations in a generic no-signaling theory. We can now move to the main issue, the extraction of a secret key.

E. One-way classical postprocessing

1. Generalities

For one-way classical postprocessing, the bound for the length of the achievable secret key rate under the assumption of individual attacks is the Csiszár–Körner (CK) bound [23,24]. In the case where Eve’s knows more about Alice’s symbol than about Bob’s, as is the case here, the CK bound reads

\[ R_{CK} = \sup_{(B',T) \to B} \left[ H(B'|E,T) - H(B'|A,T) \right], \] (19)

where \( B \to (B',T) \) is called preprocessing. From his initial data \( B \), Bob obtains some processed data \( B' \) that he does not reveal and some other processed data \( T \) that are broadcast on a public channel. For classical distributions, bitwise preprocessing is already optimal [23,24]. In this paper, we have not explored the possible use of \( T \). In this case, the preprocessing reduces to flipping each bit with some probability \( q \). Consequently, we will have an estimate \( r_{CK} \leq R_{CK} \) for the achievable secret key rate. Recalling the link \( I(X:Y) = H(X) - H(X|Y) \) between Shannon entropies and mutual information, we write our estimate for the CK bound as

\[ r_{CK} = \max_{B' \to B} \left[ I(A:B') - I(B':E) \right] \]

\[ = \frac{1}{2} \sum_{x=0,1} \max_{B' \to B} \left[ I(A:B'|x) - I(B'|E|x) \right]. \] (20)

Let us sketch the computation explicitly for \( x=0 \) (for conciseness, we omit this condition in the formula below). In Table II, one reads for \( p(a,b) \),

\[ p(0,0) = \frac{p_{NL} + p_1 + p_2 + p_3 + p_4}{2}, \]

\[ p(0,1) = \frac{p_1 + p_3}{2}, \]

\[ p(1,0) = \frac{p_1 + p_4}{2}. \]

\[
\begin{array}{c|cc}
\text{Isotropic} & b=0 & b=1 \\
\hline
a=0 & p_{NL}/2 (0,?) & p_L/4 (0,0) & p_L/8 (0,?) \\
a=1 & p_{NL}/2 (0,?) & p_L/4 (1,1) & p_L/8 (1,?) \\
\end{array}
\]

If we denote by \( q \) the probability that Bob flips his bit in the preprocessing, then

\[ p(a,b') = (1-q)p(a,b=b') + qp(a,b=b'+1). \] (22)

These four probabilities allow us to compute the mutual information \( I(A:B') = H(A) - H(A|B') \). Turning to Eve, before preprocessing she has full knowledge of Bob’s symbol for \( L_1^\prime \) and \( L_2^\prime \), and no knowledge for \( L_1^\prime \) and \( L_2^\prime \). As a consequence of the fact that Eve knows exactly on which items she has full information and on which she has no information at all, one has simply

\[ H(B'|E) = H(B|E) + \left[ 1 - H(B|E) \right] h(q), \] (23)

where \( h \) is binary entropy. The calculation is, of course, identical for \( x=1 \) and this allows computation of \( r_{CK} \) for any probability distribution. We focus explicitly on two cases.

2. Isotropic distribution

Let us consider an isotropic probability distribution, that is, a distribution of the form

\[ p(a,b|x,y) = \frac{1}{4} p_{NL} \delta(a+b-xy) + p_k. \] (24)

This necessarily implies \( p_j = p_k/8 \) for all \( j, r \), since recall that the \( L_j^\prime \) are linearly independent. Note that the point of highest violation in the quantum region (10) is of this form, with \( p_{NL} = \sqrt{2} - 1 \).

Remarkably, Alice and Bob can transform any distribution with a given \( p_{NL} \) to the isotropic distribution defined by the same \( p_{NL} \) with local operations and public communication, a procedure called depolarization [25]. This implies that the results of this paragraph are in some sense generic. In fact, by theorem 2 of Sec. II C, Eve’s best individual eavesdropping strategy for a fixed value of \( p_{NL} \) consists in preparing an isotropic distribution. Alternatively, we can modify the protocol to add the fact that Alice and Bob apply systematically the depolarization procedure.

For isotropic distributions, the two tables for \( x=0 \) and \( x=1 \) become identical, and we can rewrite them as Table III. In this table, we have changed the notation for Eve’s knowledge and have written \((a,b)\) when Eve knows both out-
comes, (a,?) when she knows only Alice’s, and (?,?,?) when she knows none.

This distribution has $p(a=0)=p(a=1)=\frac{1}{2}$. Before preprocessing, the error between Alice and Bob is $e_{AB}=p_L/4$; after preprocessing, the quantity to be corrected in error correction is $e'_{AB}=(1-q)e_{AB}+q(1-e_{AB})$. Eve’s information is $p_{NL} \left[1-h(q)\right]$. Thus,

$$r_{CK} = \max_{q \in [0,1/2]} \left[1-h(e'_{AB})-\frac{p_{NL}}{2} \left[1-h(q)\right]\right].$$

This quantity is plotted in Fig. 2 as a function of the disturbance $D$ defined by $p_{NL,\text{opt}}=\sqrt{2(1-2D)}-1$. This parameter characterizes the properties of the channel linking Alice and Bob; it is therefore useful for comparison with a quantum realization of the CHSH protocol and with BB84 (see Sec. III G and Appendix A). We see that $r_{CK}>0$ for $D \approx 6.3\%$ that is $p_{NL,\text{opt}} \geq 0.236$ for the optimal preprocessing. Without preprocessing, the bound becomes $p_{NL,\text{opt}} \geq 0.318$. The important remark is that both these values are within the quantum region (12). This means that using quantum physics, one can distribute correlations which allow (at least against individual attacks) the extraction of a secret key without any further assumption about the details of the physical realization.

3. Reaching the Bell limit

Another interesting example deals with the following question: Can one find one-parameter families of probability distributions for which $R_{CK}>0$ as soon as $p_{NL,\text{opt}}>0$, that is, distributions for which one can extract a secret key out of one-way processing, down to the limit of the local polytope? The answer is yes, and this can be achieved even without preprocessing. Here is an example: set $p_1=\frac{1}{2}, p_2=p_L$, and $p_{3,4}'=0$. For both $x=0$ and $x=1$ we have $p(a=0)=p(a=1)=\frac{1}{2}$. For $x=0$, Alice and Bob make no errors ($e_{AB}^{(0)}=0$), and Eve’s information is $I(B:E|x=0)=p_L$; for $x=1$, the errors of Alice and Bob are $e_{AB}^{(1)}=\frac{p_L}{2}$ and Eve has no information. In summary, even neglecting preprocessing,

$$r_{CK} = 1 - \frac{1}{2} h(p_L/2) - \frac{p_L}{2},$$

which is strictly positive in the whole region $p_L<1$.

Note that the distributions described here cannot be broadcasted using quantum states. The reason is that the quantum interaction with the nonlocal region is strictly inside this region, where inside means that as soon as $p_{NL,\text{opt}}>0$, all the $p_j'$ must be nonzero, because the $L_j'$ are linearly independent. On the contrary, here we have set $p_{3,4}'=0$. Anyway, in spite of the fact that we are not able to broadcast this distribution with known physical means, it is interesting to notice that there exists a family of probability distributions that can lead to a secret key under one-way postprocessing, for any amount of nonlocality.

F. Two-way classical postprocessing

1. Advantage distillation (AD)

Contrary to the one-way case, no tight bound like the Csiszár-Körner bound is known when two-way classical postprocessing is allowed, nor is the optimal procedure known. The best-known two-way postprocessing is the so-called advantage distillation (AD). Forgetting about preprocessing, one can see the effect of AD as follows: starting from a situation where $I(A:B)<I(B:E)$, one makes a processing at the end of which the new variables satisfy $I(A:B)>I(B:E)$ and at this point, one applies the one-way postprocessing.

In AD, Alice reveals $N$ instances such that her $N$ bits are equal: $a_i=\cdots=a_{i+N}$. Bob looks at the same instances and announces whether his bits are also all equal. If indeed $b_{i+N}=\cdots=b_{i+N+\beta}$, which happens with probability $1-\left(1-e_{AB}\right)^N + e_{AB}^N$, Alice and Bob keep one instance; otherwise, they discard all the $N$ bits. Bob’s error on Alice’s symbols becomes

$$\bar{e}_{AB} = \frac{e_{AB}^N}{(1-e_{AB})^N + e_{AB}^N} = \left(\frac{e_{AB}}{1-e_{AB}}\right)^N.$$  

Notice that $\bar{e}_{AB} \to 0$ in the limit $N \to \infty$. This means that $\alpha=\beta$ almost always, for $N$ sufficiently large. This remark is used to estimate Eve’s probability of error (see below for concrete applications). Typically, one finds that Eve’s error on Bob’s symbols goes as

$$\bar{e}_E \approx C[f(e_{AB})]^N,$$

with $f(\cdot)$ some function which depends on the probability distribution under study. Now, as long as the condition

$$f(e_{AB}) > \frac{e_{AB}}{1-e_{AB}}$$

is fulfilled, Eve’s error at the end of AD is exponentially larger than Bob’s for increasing $N$. There exists always a finite value of $N$ such that Eve’s error becomes larger than Bob’s. The bound on the tolerable error after AD is then computed by solving Eq. (29).
TABLE IV. Probability distribution Alice-Bob-Eve, in the case of isotropic distribution, after Alice’s and Bob’s preprocessing.

<table>
<thead>
<tr>
<th>a = 0</th>
<th>b = 0</th>
<th>b = 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNL(\tilde{q}_Aq_B + q_Aq_B)/2 (\text{?},\text{?})</td>
<td>pNL(\tilde{q}_Aq_B + q_Aq_B)/2 (\text{?},\text{?})</td>
<td>pNL(\tilde{q}_Aq_B + q_Aq_B)/2 (\text{?},\text{?})</td>
</tr>
<tr>
<td>p\tilde{q}_A\tilde{q}_B/4 (0,0)</td>
<td>p\tilde{q}_A\tilde{q}_B/4 (0,0)</td>
<td>p\tilde{q}_A\tilde{q}_B/4 (0,0)</td>
</tr>
<tr>
<td>p\tilde{q}_Bq_B/4 (1,1)</td>
<td>p\tilde{q}_Bq_B/4 (1,1)</td>
<td>p\tilde{q}_Bq_B/4 (1,1)</td>
</tr>
<tr>
<td>p\tilde{q}_A\tilde{q}_B/8 (0,2)</td>
<td>p\tilde{q}_A\tilde{q}_B/8 (0,2)</td>
<td>p\tilde{q}_A\tilde{q}_B/8 (0,2)</td>
</tr>
<tr>
<td>p\tilde{q}_Aq_B/8 (1,?)</td>
<td>p\tilde{q}_Aq_B/8 (1,?)</td>
<td>p\tilde{q}_Aq_B/8 (1,?)</td>
</tr>
<tr>
<td>a = 1</td>
<td>b = 0</td>
<td>b = 1</td>
</tr>
<tr>
<td>pNL(\tilde{q}_Aq_B + \tilde{q}_Aq_B)/2 (\text{?},\text{?})</td>
<td>pNL(\tilde{q}_Aq_B + q_Aq_B)/2 (\text{?},\text{?})</td>
<td>pNL(\tilde{q}_Aq_B + q_Aq_B)/2 (\text{?},\text{?})</td>
</tr>
<tr>
<td>p\tilde{q}_A\tilde{q}_B/4 (0,0)</td>
<td>p\tilde{q}_A\tilde{q}_B/4 (0,0)</td>
<td>p\tilde{q}_A\tilde{q}_B/4 (0,0)</td>
</tr>
<tr>
<td>p\tilde{q}_Bq_B/4 (1,1)</td>
<td>p\tilde{q}_Bq_B/4 (1,1)</td>
<td>p\tilde{q}_Bq_B/4 (1,1)</td>
</tr>
<tr>
<td>p\tilde{q}_A\tilde{q}_B/8 (0,2)</td>
<td>p\tilde{q}_A\tilde{q}_B/8 (0,2)</td>
<td>p\tilde{q}_A\tilde{q}_B/8 (0,2)</td>
</tr>
<tr>
<td>p\tilde{q}_Aq_B/8 (1,?)</td>
<td>p\tilde{q}_Aq_B/8 (1,?)</td>
<td>p\tilde{q}_Aq_B/8 (1,?)</td>
</tr>
</tbody>
</table>

We apply this procedure to the isotropic correlations described above (Sec. III E 2), first without preprocessing and then by allowing Alice and Bob to perform some bit flip before starting AD. We anticipate the result. We find that a key can be extracted for \(p_{NL} \approx 0.09\), that is, even with two-way postprocessing we are not able to reach the Bell limit for isotropic correlations. It is an open question whether the Bell limit can be reached by a better two-way postprocessing for the isotropic distribution.

2. AD without preprocessing

We refer to Table III. We have, as above, \(e_{AB} = \frac{p_{L}}{2}\). We must now estimate Eve’s error on Bob’s symbol after AD. Eve knows \(\alpha\) as soon as she knows one of Alice’s symbols \(a_i\), and recall that asymptotically the guess \(\beta = \alpha\) is correct. The only situation in which Eve is obliged to make a random guess is therefore the case in which all the \(N\) instances correspond to Eve’s symbol \((?,?)\). The probability that Eve’s guess of Bob’s symbol is wrong is therefore

\[\tilde{e}_E \geq \frac{1}{2} \left( \frac{p_{NL}}{1 - e_{AB}} \right)^N, \tag{30}\]

where the denominator comes from the fact that we must condition on the bit’s acceptance. Using (29), we obtain that secrecy can be extracted as long as \(p_{NL} > p_{L}/4\) that \(p_{NL} > 1/5\). This is lower than the bound obtained for one-way postprocessing, as expected.

3. AD with preprocessing

The previous bound can be further improved by allowing Alice and Bob to preprocess their lists before starting AD. For two-way postprocessing, it is not known whether bitwise preprocessing is already optimal, but we restrict to it in this work. Specifically, we suppose that Alice flips her bit with probability \(q_A\), Bob with probability \(q_B\). By inspection, one finds that the probability distribution obtained from Table III after this preprocessing is the one of Table IV, where we have written \(\tilde{q} = 1 - q\). Just by looking at the table, one can guess the interest of preprocessing. The five possible symbols for Eve are now spread in all the four cells of the table. For instance, Eve’s symbol as \((0,0)\) was present only in the case \(a = b = 0\) in Table III, that is, whenever she had this symbol Eve had full information; this is no longer the case in Table IV. Note also that the roles of \(q_A\) and \(q_B\) are not symmetric, because only \(q_A\) mixes the strategies for which Eve does not know Bob’s symbol.

The distribution of Table IV is such that

\[e_{AB}' = \left( \frac{p_{NL} + \frac{p_{L}}{2}}{2} \right)(q_A\tilde{q}_B + \tilde{q}_Aq_B) + \frac{p_{L}}{4}. \tag{31}\]

The estimate of Eve’s error requires some attention. As before, we assume that as soon as Eve guesses correctly Alice’s symbol \(\alpha\), she automatically guesses also \(\beta\). So the question is, when is Eve uncertain about \(\alpha\), in the asymptotic regime of large \(N\)? Of course, inequality (30) still holds with \(e_{AB}'\) replacing \(e_{AB}\), but this condition is too weak here. It does not make any use of the uncertainty introduced on Eve’s knowledge by the preprocessing.

Eve’s situation now is such that even if she has a symbol \((a,b)\) or \((a,\bar{a})\), she cannot be completely sure whether \(a = a\) or not. Suppose that among her \(N\) symbols, Eve has \(n_0\) times the symbol \((\text{?},\text{?})\), \(n_1^0\) times the symbol \((0,\text{?})\), \(n_1^1\) times the symbol \((1,\text{?})\), \(n_0^1\) times the symbol \((\text{?},0)\), and \(n_1^2\) times the symbol \((1,1)\). Eve cannot avoid errors when \(p(a = 0 | e) = p(a = 1 | e)\), that is, when \(n_1^0 = n_1^1 = n_1^2 = n_2\). We have therefore the bound

\[\tilde{e}_E' \geq \frac{1}{2} \sum_{n_0,n_1,n_2} \frac{N!}{n_0!(n_1^0+n_1^2)!2(n_2)!} q_0^{n_0} q_1^{n_1} \gamma_{1,1}^{2n_1} \gamma_{1,0}^{2n_2}, \tag{32}\]

where the sum is taken under the constraint \(n_0 + 2n_1 + 2n_2 = N\). \(\gamma_e\) is the probability that Eve has symbol \(e\) conditioned on the bit’s acceptance, and \(\gamma_{1,1} = \sqrt{\gamma(0,1)\gamma(1,1)}\). Using \((n!)^2 \sim (2n)! / 2^{2n}\) and summing the multinomial expansion, we obtain

\[\tilde{e}_E' \geq \frac{1}{8} \gamma_{1,1} + 2 \gamma_1 + 2 \gamma_2 N. \tag{33}\]

Now we must find the expressions for the \(\gamma_e\) in Table IV. Suppose for definiteness that Alice and Bob have accepted the bit \(a = b = 0\). This happens with probability \(1/\gamma_{1,0}\). The probability that this happens and that Eve has got the symbol \((\text{?},?)\) is \(\frac{p_{NL}}{2}(q_A\tilde{q}_B + \tilde{q}_Aq_B)\); whence \(\gamma(\text{?},?) = \frac{1}{1 - e_{AB}'}\). Similarly, the probability that Alice and Bob accept the bit 0 and that Eve has got \((0,?)\), respectively \((1,?)\), is \(\frac{p_{NL}}{2} q_A\); respectively \(\frac{p_{NL}}{2} \tilde{q}_A\); whence \(\gamma_1 = \frac{p_{NL} q_A}{4(1 - e_{AB}')}\) and \(\gamma_2 = \frac{p_{NL} \tilde{q}_A}{4(1 - e_{AB}')}\). In a similar way, one computes \(\gamma_{2,\gamma}\). By writing \(\delta_e = (1 - e_{AB}') \gamma_e\), we have then

\[\delta_{e(1,1)} = \frac{p_{NL}(\tilde{q}_Aq_B + q_Aq_B)}, \quad \delta_1 = \frac{p_{L}}{4} \sqrt{q_Aq_B}.\]
The optimization over $q_A$ and $q_B$ can be done numerically. The result is that a secret key can be extracted at least down to $p_{NL} \approx 0.09$.

4. Positivity of intrinsic information

Given a tripartite probability distribution, $P(a,b,e)$, an upper bound to the secret-key rate $R$ is given by the so-called intrinsic information $I(A:B|\bar{E})$, denoted more briefly in what follows by $I_1$. This function, introduced in [26], reads

$$I(A:B|\bar{E}) = \min_{E \rightarrow \bar{E}} I(A:B|\bar{E}),$$

with the minimization running over all the channels $E \rightarrow \bar{E}$. Here, $I(A:B|\bar{E})$ denotes the mutual information between Alice and Bob conditioned on Eve. That is, for each value of Eve’s variable $e$, the correlations between Alice and Bob are described by the conditioned probability distribution $P(a,b|e)$. The conditioned mutual information $I(A:B|\bar{E})$ is equal to the mutual information of these probability distributions averaged over $P(e)$. The exact computation of the intrinsic information is, in general, difficult. However, a huge simplification was obtained in [27], where it was shown that the minimization in Eq. (36) can be restricted to variables $\bar{E}$ of the same size as the original one, $E$. This allows a numerical approach to this problem.

The intrinsic information can be understood as a witness of secret correlations in $P(a,b,e)$. Indeed, a probability distribution can be established by local operations and public communication if, and only if, its intrinsic information is zero [28]. It is then clear why the positivity of the intrinsic information is a necessary condition for a positive secret-key rate. Whether it is sufficient is at present unknown. Strong support has been given to the existence of probability distributions such that $R=0$ and $I_1>0$. These would constitute examples of probability distributions containing bound information [29], that is, nondistillable secret correlations. The existence of bound information has been proven in a multipartite scenario consisting of $N>2$ honest parties and the eavesdropper [30]. However, it remains as an open problem for the more standard bipartite scenario.

Using these tools, it is possible to study the secrecy properties of the probability distribution $P(a,b,e)$ derived from the previous CHSH protocol. A first computation of its conditioned mutual information gives $I(A:B|\bar{E})=p_{NL}$. This result easily follows from Table II. When $e=??$, which happens with probability $p_{NL}$, Alice and Bob are perfectly correlated, so their mutual information is equal to one. In all the remaining cases, e.g., $e=(0,0)$, Alice and Bob have no correlations. Using this observation, one can guess the optimal map $E \rightarrow \bar{E}$. In order to minimize the conditioned mutual information, this map should deteriorate the perfect correlations between Alice and Bob when $e=(?,?)$. A way of doing this is by mapping $(0,?)$ and $(1,?)$ into $(?,?,?)$, leaving the other symbols unchanged [31]. We conjecture that this defines the optimal map for the computation of the intrinsic information. Actually, all our numerical evidence supports this conjecture. Thus, the conjectured value for the intrinsic information is

$$I_1 = \left(1 - \frac{p_{NL}}{2}\right) \left[1 - h\left(\frac{p_{NL}}{4 - 2p_{NL}}\right)\right].$$

Interestingly, this quantity is positive whenever $p_{NL}>0$. If the conjecture is true, it implies that either (i) it is possible to have a positive secret-key rate for the whole region of Bell violation, using a new key-distillation protocol, or (ii) the probability distribution of Table II represents an example of bipartite bound information for sufficiently small values of $p_{NL}$.

G. Quantum cryptographic analysis of the CHSH protocol

It is interesting to analyze the CHSH protocol with the standard approach of quantum cryptography. Alice and Bob share a quantum state of two qubits and have agreed on the physical measurements corresponding to each value of $x$ and $y$. Eve is constrained to distribute quantum states, of which she keeps a purification. Recent advances have provided a systematic recipe to find a lower bound on the secret key rate, that is, to discuss security when Eve is allowed to perform the most general strategy compatible with quantum physics (such bounds have been called “unconditional security proofs,” but it should be clear by all that precedes that this wording is unfortunate).

The resulting bound on the achievable secret key rate is plotted in Fig. 2. Since the formalism used to compute this bound is entirely different from the tools used in the present study, we give this calculation in Appendix A. It turns out that the CHSH protocol is equivalent to the BB84 protocol plus some classical preprocessing. In particular, the robustness to noise is the same for both protocols. For low error rate, BB84 provides higher secret-key rate; however, BB84 cannot be used for a device-independent proof, since (as we noticed above) its correlations become intrinsically insecure if the dimensionality of the Hilbert space is not known.

IV. LARGER-DIMENSIONAL OUTCOMES

In this section, we explore the generalization of the previous results to the case of binary inputs and $d$-nary outputs: $m_A=m_B=2$, $n_A=n_B=d$, that is, $x,y \in \{0,1\}$ and $a,b \in \{0,1,\ldots,d-1\}$. Below all the sums involving $d$s are to be computed modulo $d$.

For this study, it is useful to introduce a notation for probability distributions and inequalities [32]. While the full probability space is $4d^2$-dimensional, one can verify that only $4d(d-1)$ parameters are needed to characterize completely a no-signaling probability distribution—in other words, $D=4d(d-1)$ is the dimension of the space in which the no-signaling and the local polytopes are embedded. We choose the $\{P(a|x),a=0,1,\ldots,d-2;x=0,1\}$ ($d-1$ numbers for each value of $x$), the $\{P(b|y),b=0,1,\ldots,d-2;y=0,1\}$
(d−1 numbers for each value of y), and the
\{P(a,b|x,y),a,b=0,1,\ldots,d−2;x,y=0,1\} \quad [(d−1)^2 numbers
for each value of x,y]. This we arrange in arrays as follows:

\[
P = \begin{bmatrix}
P(a|0) & P(a|0,0) & P(a|0,1) \\
P(a|1) & P(a|1,0) & P(a|1,1)
\end{bmatrix}
\] (38)

Note that this array has 2(d−1) lines and as many columns. Information on the values a,b=d−1 is redundant for all inputs because of no-signaling. Of course, there is no problem in working with the “full” array with 2d×2d if one finds it more convenient, provided the additional entries are filled consistently, because these parameters are not free.

This notation will also be used for inequalities. In this case, the numbers in the arrays are the coefficients which multiply each probability in the expression of the inequality. Examples will be provided below.

A. Polytopes and the quantum region

1. Known characterization

As one might expect, characterization of the local and the no-signaling polytopes is an increasingly hard task, as the dimensions of the output increase.

Numerical studies [17] have provided an unexpectedly simple structure for the local polytope for small d: as it happened for d=2, all the nontrivial facets appear to be equivalent to the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality [33,34,17,35]

\[
I_d = \begin{bmatrix}
-1 & -1 & \cdots & -1 & 0 & 0 & \cdots & 0 \\
-1 & 0 & 1 & \cdots & 1 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
-1 & 0 & 0 & \cdots & 1 & 1 & \cdots & 1 \\
0 & 1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\
0 & 1 & 1 & \cdots & 0 & -1 & -1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1
\end{bmatrix}
\leq 0.
\] (39)

It is conjectured that all nontrivial facets are equivalent to the CGLMP inequality for all d. Anyway, our work is independent of the truth of this conjecture. We are going to study the possibility of secret key extraction for nonlocal distributions which lie above a CGLMP facet, irrespective of whether inequivalent facets exist or not.

The no-signaling polytope appears to have a richer structure than in the case d=2. All the extremal nonlocal points are generalizations of the PR box [10]. We are interested in those that lie above our representative CGLMP facet. The highest violation of CGLMP is provided by the extremal point

\[
P_{PR_{2,d}} = \frac{1}{d} \delta(b - a = xy),
\] (40)

whose corresponding array is

\[
\begin{array}{cccccccc}
A & B & 1 & 1 & \ldots & 1 & 1 & 0 & \ldots & 0 \\
1 & 0 & 0 & \ldots & 1 & 0 & \ldots & 0 & \ldots & 0 \\
1 & 0 & 1 & \ldots & 0 & 1 & \ldots & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
1 & 0 & 0 & \ldots & 1 & 0 & \ldots & 0 & \ldots & 0
\end{array}
\]

Its violation of the inequality can be rapidly calculated by a term-by-term multiplication (a formal scalar product) of the two arrays (39) and (41), yielding

\[
\langle I_d, P_{PR_{2,d}} \rangle = \frac{d−1}{d}.
\] (42)

However, P_{PR_{2,d}} is not the only nonlocal extremal point which lies above a CGLMP facet. In fact, for all d'<d, there is at least one P_{PR_{2,d}} above the facet. For instance, a possible version of P_{PR_{2,2}}=P_{PR} reads (boldface 0 standing for matrices filled with zeros)

\[
P_{PR} = \frac{1}{2}
\]

whence a violation \langle I_d, P_{PR} \rangle=\frac{1}{2}. For d=3, we shall give below (Sec. IV C) some additional elements on the structure of the no-signaling polytope.

The boundaries of the quantum region are basically unknown to date; it is not even clear whether they coincide with all possible results of measurements on two-qutrits states.

2. A slice in the nonlocal region

As we said, a no-signaling probability distribution is characterized by 4d(d−1) parameters. However, when one reviews the results obtained for the CGLMP inequality in the context of quantum physics (see Appendix B for all details), one finds that the probability distributions associated to the optimal settings belong to a very symmetric family. Specifically, these distributions are such that (i) for fixed inputs x
and $y$, $P(a,b|x,y)$ depends only on $\Delta = a-b$, and (ii) the probabilities for the different inputs are related as $P(\Delta|0,0) = P(-\Delta|0,1) = P(\Delta|1,0) = P(-\Delta|1,1)$. Compactly,

$$P(a,b = a-\Delta|x,y) = \frac{1}{d} p_f,$$  \hspace{1cm} (44)

with $f = (-1)^{xy} \Delta + xy$ and $p_f = \Sigma_a P(a,b = a-f|0,0)$. The corresponding array is

$$P = \frac{1}{d} \begin{bmatrix} A \setminus B & 1 & 1 & \ldots & 1 & 1 & 1 & \ldots & 1 \\ 1 & p_0 & p_{-1} & \ldots & p_2 & p_0 & p_1 & \ldots & p_{-2} \\ p_1 & p_0 & \ldots & p_3 & p_{-1} & p_0 & \ldots & p_{-3} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{-2} & p_{-3} & \ldots & p_0 & p_2 & p_3 & \ldots & p_0 \\ p_0 & p_1 & \ldots & p_{-2} & p_1 & p_0 & \ldots & p_3 \\ p_{-1} & p_0 & \ldots & p_{-3} & p_2 & p_1 & \ldots & p_4 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_2 & p_3 & \ldots & p_0 & p_{-1} & p_{-2} & \ldots & p_1 \end{bmatrix}.$$  \hspace{1cm} (45)

This family defines a slice in the no-signaling polytope. Note that all the marginals are equal, that is, $P(a|x) = P(b|y) = \frac{1}{d}$. Moreover, the $d$ numbers $p_f$ define uniquely and completely a point $P$ in the slice; thus, given the constraint $\Sigma_f p_f = 1$, the slice defined by (44) is $(d-1)$-dimensional. A single extremal nonlocal point belongs to the slice, namely, $P_{PB_{2,d}}$ obtained by setting $p_0 = 1$ (40); in fact, none of the $P_{PB_{2,d}}$ with $d' < d$ has the correct marginals.

As it happened for the isotropic distributions for $d=2$, there exists a depolarization procedure that maps any probability distribution onto this slice by local operations and public communication, while keeping the violation $\langle I_d, P \rangle$ constant. The procedure is given in Appendix C. As a consequence, Eve’s optimal individual eavesdropping, for a fixed value of the violation of the inequality, consists in distributing a point in the slice.

**B. Cryptography**

1. **The protocol**

We suppose from the beginning $p(x=i)=p(y=j)=\frac{1}{2}$. The protocol is the analog of the CHSH protocol described in Sec. III C above. When Alice announces $x=1$ and Bob has measured $y=1$, Bob corrects his dit according to $b \rightarrow b-1$. In other words, the pseudosifting implements $\Delta \rightarrow \Delta - xy$. The Alice-Bob distribution after pseudosifting, averaged on Bob’s settings, becomes independent of $x$ (as in the case of isotropic distribution for $d=2$):

$$P(a,a-\Delta|x) = \frac{1}{d} \Sigma_{y=0,1} p_{(-1)^{xy} \Delta} = \frac{p_{\Delta} + p_{-\Delta}}{2d}. \hspace{1cm} (46)$$

In a protocol with $d$-dimensional outcomes, Alice and Bob can estimate not just one; but several error rates, one for each value of $\Delta$. We have just found that these error rates exhibit the symmetry

$$e_{AB}(\Delta) = e_{AB}(-\Delta) = \frac{p_{\Delta} + p_{-\Delta}}{2}. \hspace{1cm} (47)$$

As in the case of bits, we think of Eve as sending either a local or a nonlocal probability distribution. Let us discuss in some detail the points which lie on and above a CGLMP facet.

2. **Eve’s strategy: Local points**

To understand what follows, we do not need a full characterization of the deterministic strategies that saturate the CGLMP inequality. Some facts are, however, worth noting. The proof of these statements and some other features are given in Appendix D.

The first fact is that for $d > 2$, the number of deterministic points on the CGLMP facet is strictly larger than $D = 4d(d-1)$, the dimension of the local and the no-signaling polytope. This implies that for some points on the facet, several decompositions as a convex combination of extremal points are possible.

The second fact is that no extremal deterministic strategy belongs to the slice (44). To see it, just recall that the marginals in the slice are completely random. Since we require the final distribution to belong to the slice, Eve must manage to send deterministic strategies with the suitable probabilities. As a consequence of the previous remark, at least one local point on the slice can be obtained by several different decompositions on extremal points. We’ll have to choose the decomposition that optimizes Eve’s information.

As a third fact, we elaborate on the same idea that led to the uncertainty relations in Sec. III D. We know that all deterministic strategies are not equally interesting for Eve. In fact, two kinds of local points are of special interest for her: (i) those for which $b(0) = b(1)$, because Eve knows Bob’s symbol when Alice announces $x=0$, and which we denote by the set $L_0$; and (ii) those for which $b(0) = b(1)-1$, because Eve knows Bob’s symbol when Alice announces $x=1$, and which we denote by the set $L_1$. In all the other cases, Eve does not learn Bob’s symbol with certainty. Now, in the complexity of the list of deterministic points on the CGLMP facet, a remarkable feature appears:

(i) There are exactly $d^2$ points in $L_0$, namely, those for which $a(0) = b(0) = b(1)$ and $a(1)$ can take any value. In other words, there are no points on the CGLMP facet such that $b(0) = b(1)$ but $a(0)$ is different from this value. Whenever Eve learns Bob’s symbol for $x=0$, Alice and Bob make no error for $x=0$.

(ii) There are exactly $d^2$ points in $L_1$, namely, those for which $a(1) = b(0) = b(1)-1$ and $a(0)$ can take any value. This has a similar interpretation as the statement above, in the case $x=1$.

Now, since the error rate Alice-Bob depends only on $P(a,b|x,y)$ and not on the particular decomposition chosen by Eve to realize this distribution, it is obvious that Eve’s interest lies in distributing local points that belong to $L = L_0 \cup L_1$ as often as possible. For $d=3$, we shall prove that she can prepare any point in the slice by distributing only these kinds of local points. Finally, we want to introduce a further distinction within $L$, which appears explicitly in the
study of $d=3$ but may play a more general role. We shall call $\mathcal{L}^3$ the subset of $\mathcal{L}$, whose points satisfy three out of the four relations $a(0) = b(0)$, $a(0) = b(1)$, $a(1) = b(0)$, and $a(1) = b(1) = 1$. The complementary set, containing the points that satisfy two of the relations $a(0) = b(0)$ or $a(1) = b(0) = b(1) = 1$, is written $\mathcal{L}^2$.

3. Eve’s strategy: Nonlocal point

We have said that, among the extremal nonlocal points which lie above the CGLMP facet, the only one on the slice (44) is $P_{PR_{2d}}$. However, it may be the case that mixtures of other extremal nonlocal points lie as well in the slice. For $d=3$, this is not the case (see Appendix E), but we have not been able to generalize this statement. In this study, we suppose tentatively that Eve sends a unique nonlocal strategy, namely $P_{PR_{2d}}$. Under this assumption, we can define $P_{NL}$ as the probability that Eve sends $P_{PR_{2d}}$. To find the expression of $P_{NL}$, we notice that $\langle I_0, P_{NL} \rangle = \frac{\Delta}{d}$ should correspond to $p_{NL} = 1$, and that $\langle I_j, L \rangle = 0$ for all local points on the CGLMP facet should correspond to $p_{NL} = 0$. Moreover, $P_{NL}$ measures the geometrical distance from the facet and is therefore an affine function of the violation of CGLMP. Thus, for a generic distribution $P$ of the form (45) we have

$$p_{NL} = \frac{d}{d-1} \langle I_0, P \rangle = -2 + \sum_{\Delta=0}^{d-1} \left(1 - \frac{\Delta}{d-1}\right) \left|3p_{-\Delta} - p_{\Delta+1}\right|. \tag{48}\$$

Now we can present the results for the possibility of extracting a secret key, starting from a detailed study of the case $d=3$ (Sec. IV C), then generalizing some results for arbitrary $d$ (Sec. IV D).

C. Secret key extraction: $d=3$

1. The slice of the polytope

The slice (44) is two-dimensional for $d=3$, and we choose $p_0$ and $p_1$ as free parameters. This gives $p_2 = 1 - p_0 - p_1$ and

$$p_{NL} = 2(p_0 - p_1) - 1. \tag{49}\$$

The full slice has a form of an equilateral triangle (Fig. 3), whose vertices $V_\Delta$ are defined by $p_\Delta = 1$. As mentioned, $V_0 = P_{PR_{23}}$. The vertex $V_2$ is also a $P_{PR_{23}}$, the one defined by $b-a = x+y+1$ with $x = 1 - z$. On the contrary, $V_1$ is a mixture of deterministic strategies. The middle of the triangle, $p_0 = p_1 = p_2 = \frac{1}{3}$, is the completely random strategy (obtained, e.g., when measuring the maximally mixed quantum state, the “identity”).

We are going to focus on the nonlocal region close to $V_0$ (Fig. 4). The intersection with the CGLMP facet is the segment $p_0 - p_1 = \frac{1}{2}$, whose ends are the points labeled $M_2$ ($p_0 = \frac{1}{2}, p_1 = 0$) and $M_3$ ($p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$). The decompositions of these mixtures on the extremal deterministic strategies are

$$M_2 = \sum_{L \in \mathcal{L}^2} \frac{1}{6} L, \quad M_3 = \sum_{L \in \mathcal{L}^3} \frac{1}{12} L, \tag{50}\$$

where the sets of local points $\mathcal{L}^2$ and $\mathcal{L}^3$ have been defined above. In fact, the decomposition of $M_3$ is unique.

Conversely, $M_2$ can be decomposed in an infinity of ways (see Appendix E), but all the others involve also the points that do not belong to $\mathcal{L}$ and are therefore suboptimal for Eve.

The quantum-mechanical studies (see Appendix B for more details) have singled out two nonlocal probability distributions in this region. The first one corresponds to the maximal violation of CGLMP using two qutrits, $p_{NL} = 0.4574$, noted $Q_{mm}$ and defined by (B6) with $\gamma = \frac{1}{2}$. The second one corresponds to the highest violation achievable with the maximally entangled state of two qutrits, $p_{NL} = 0.4365$, noted $Q_{mc}$ and defined by (B6) with $\gamma = 1$.

2. One-way, classical postprocessing

To write down the table for the correlations Alice-Bob-Eve, one needs to list explicitly the deterministic points that saturate the CGLMP and the corresponding information Eve can extract. This is done in Appendix E. The result is Table V. It can be verified easily that all the probabilities in a row/column sum up to $\frac{1}{3}$; moreover.
TABLE V. Probability distribution Alice-Bob-Eve for $d=3$, after pseudotyping, assuming decomposition (50) for $M_2$. We indicate by $?_2$ the case where Eve hesitates among two values of Bob’s symbol (instead of three).

<table>
<thead>
<tr>
<th>$b=0$</th>
<th>$b=1$</th>
<th>$b=2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=0$</td>
<td>$p_{NL}/3$ (2,?)</td>
<td>$p_L/6$ (0,0)</td>
</tr>
<tr>
<td></td>
<td>$p_L/12-p_2/6$ (0,?2)</td>
<td>$(1-p_0)/6$ (0,?2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(1-p_0)/6$ (0,?2)</td>
</tr>
<tr>
<td>$a=1$</td>
<td>$p_{NL}/3$ (?2)</td>
<td>$p_L/6$ (1,1)</td>
</tr>
<tr>
<td></td>
<td>$(1-p_0)/6$ (1,?2)</td>
<td>$p_L/12-p_2/6$ (1,?2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(1-p_0)/6$ (1,?2)</td>
</tr>
<tr>
<td>$a=2$</td>
<td>$p_{NL}/3$ (?)</td>
<td>$p_L/6$ (2,2)</td>
</tr>
<tr>
<td></td>
<td>$(1-p_0)/6$ (2,?2)</td>
<td>$(1-p_0)/6$ (2,?2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_L/12-p_2/6$ (2,?2)</td>
</tr>
</tbody>
</table>

$$e_{AB}(+1) = e_{AB}(-1) = \frac{1-p_0}{2},$$

(51)
as expected from (47). We have introduced the symbol $?_2$ to describe the situation where Eve is uncertain of Bob’s symbol, but only among two possibilities: this is clearly the case whenever the uncertainty derives from a deterministic strategy. In all that follows, information is quantified in trits, and we write $h([v_1, v_2, v_3]) = \sum v_j \log_3 v_j$.

In the absence of preprocessing, Eve has no information with probability $p_{NL}$, full information with probability $p_L/3$, and information $1-h([1/2, 1/2, 0])=1-\log_3 2$ with probability $p_L/2$. Therefore the estimate for the CK bound is

$$R_{CK}(q = 0) = r_{CK} = 1-h\left(\frac{1-p_0}{2}, \frac{1-p_0}{2}\right) - p_L\left(1-\frac{1}{2} \log_3 2\right).$$

(52)
The curve $r_{CK}(q=0)=0$ is shown in Fig. 4; it clearly cuts the quantum region.

A natural question is, which is the point that maximizes $r_{CK}(q=0)$ under the requirement that the correlations should belong to the quantum region? In the slice under consideration, we find a rate $r_{max}(q=0)=0.09$ trits =0.144 bits for the correlations defined by $p_0=0.8286$, $p_1=0.1093$. These correlations can be obtained by measuring the quantum state

$$|\psi(\gamma)\rangle = \frac{1}{\sqrt{2+\gamma^2}}(|00\rangle + \gamma |11\rangle + |22\rangle)$$

(53)
for $\gamma=0.9875$. This state is close to, but certainly different from, the maximally entangled state. Thus, the secret key rate exhibits the same form of anomaly as all the other measures of nonlocality known to date [45]; maximal nonlocality is obtained with nonmaximally entangled states.

We consider now Bob’s preprocessing. For one-way postprocessing, dit-wise preprocessing is already optimal. A priori, one can define two different flipping probabilities $q_{+1}$ and $q_{-1}$, associated respectively to $b \rightarrow b+1$ and $b \rightarrow b+2$. But it turns out by inspection that the optimal is always obtained for $q_{+1}=q_{-1}=q$, so we write down directly this case. From (51) it is clear that after preprocessing

$$e_{AB}(+1) = e_{AB}(-1) = \frac{1-p_0}{2} + q\frac{3p_0-1}{2},$$

(54)
whence

$$I(A:B') = 1-h\left(1 - e'_{AB} - e'_{AB}\right).$$

(55)

Eve’s information is computed by recalling that for any local point she sends out, before preprocessing (i) for one value of $x$, she knows perfectly Bob’s symbol $b$; (ii) for the other value of $x$, she hesitates between two values of $b$. Preprocessing leaves $b$ unchanged with probability $1-q$, and sends it to $b+1$ with probability $q$ each. Therefore, in case (i), Eve’s information is lowered from 1 to $1-h(1-2q, q, q) = h_1(q)$; in case (ii), Eve’s information is lowered from $1-h\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ to $1-h\left(\frac{1}{2}, \frac{1}{2}, q\right)$, since each case is equiprobable,

$$I(E:B') = 1 - \frac{1}{2}[h_1(q) + h_2(q)].$$

(56)

From (55) and (56), we can compute $r_{CK}$ by optimizing the value of $q$. We did the optimization numerically. The improvement due to preprocessing is clear in Fig. 4.

3. Two-way, classical postprocessing

We have also studied the possibility of extracting a secret key from the correlations of Table V using AD (without preprocessing). Alice selects $N$ of her symbols that are identical, and Bob accepts if and only if his corresponding symbols are also identical. The probability that Bob accepts is $p_0^N + [e_{AB}(+1)]^N + [e_{AB}(-1)]^N$, and consequently,

$$\bar{e}_{AB}(\pm 1) = \frac{\left(1-p_0\right)^{N}}{p_0^N + 2\left(1-p_0\right)} = \left(\frac{1-p_0}{2p_0}\right)^{N}.$$  

(57)

As in the case $d=2$, Eve has to make a random guess if and only if she has sent $P_{PR_{(d)\setminus}}$ for all the $N$ instances,

$$\bar{e}_E(\pm 1) \approx \frac{1}{3}\left(p_{NL}\right)^N.$$  

(58)

Thus, a secret key can be extracted using AD as long as $p_{NL}>\frac{1-p_0}{3}$, that is as long as

$$5p_0 > 4p_1 + 3.$$

(59)

The limiting curve is also plotted in Fig. 4. Its extremal points are $p_{NL}>\frac{1}{2}$ for $p_1=0$ (the same value as obtained for $d=2$) and $p_{NL}>\frac{5}{3}$ for $p_2=0$. 

042339-14
4. Intrinsic information

It is straightforward to generalize the map used above in the computation of the intrinsic information to the $d=3$ case. Looking at Table V, one has to map all the symbols $(i,j,k)$ into $(?,?)$, where $i=0,1,2$. The obtained conditional mutual information reads

$$I(A;B|E) = P(?,?) \left[ 1 - h \left( \frac{2p_0 - p_L}{2 - p_L} \frac{1 - p_0}{2 - p_L} \frac{1 - p_0}{2 - p_L} \right) \right].$$

(60)

This is, of course, an upper bound to the intrinsic information, since the employed map may not be the optimal one. Contrary to what happens in the $d=2$ case, this quantity vanishes for some points inside the region of Bell violation. Indeed, Eq. (60) is zero on the line $5p_0 - 2p_1 - 3 = 0$; by changing slightly Eve’s map (specifically, she applies the map above only with a suitable probability and makes nothing in the other cases), it can be verified that the intrinsic information is zero also below the line, that is, for

$$5p_0 - 2p_1 - 3 \leq 0.$$  

(61)

This region overlaps with the nonlocal region (Fig. 4).

D. Secret key extraction: Generic $d$

For generic $d$, we want to prove that secrecy can be generated using quantum states. The statistics Alice-Bob can be computed using quantum mechanics, in particular, the error rates $e_{AB}(\Delta)$ of Eq. (47). The question is, how to estimate Eve’s information. To compute this quantity exactly, one must describe the points in the CGLMP facet in some detail. However, some interesting bound can be derived from what we have already said and the intuition developed in the study of $d=3$.

Consider first one-way postprocessing. The discussion of Sec. IV B 2 implies the bound

$$I(B:E) \leq I_E = \frac{P_L}{2} + \frac{P_L}{2} (1 - \log_d 2).$$

(62)

The bound is reached if and only if Eve distributes strategies that belong to $\mathcal{L}$, as it happened to be always possible for $d=3$. Moreover, this bound can also be computed from the Alice-Bob distribution only assuming (48). Consequently, we can estimate

$$R_{CK}(q=0) \geq r = 1 - h(e_{AB}(\Delta))_E - I_E$$

(63)

with $h$ the Shannon entropy measured in bits. We have studied the right-hand side numerically for $d=10$, for correlations in the quantum region obtained from states that are Schmidt-diagonal in the computational basis, $|\psi\rangle = \sum_{k=0}^{d-1} c_k |kk\rangle$. The general features that emerge are as follows.

(i) The maximal value of $R$ achievable in the quantum region increases with $d$, reaching up to $R = 0.692$ bits for $d=10$.

(ii) The quantum state corresponding to the maximal value of $R$ is always such that $c_k = c_{d-1-k}$. It seems that the overlap $\eta$ of this state with the maximally entangled one decreases with $d$, but the decrease is very slow (we have $\eta=1$ for $d=2$, and for $d=10$ we still have $\eta \approx 0.998$).

A similar simple approach can be found to explore the possibilities of two-way postprocessing. We have

$$\tilde{e}_{AB}(\Delta) \leq [e_{AB}(\Delta)]_{\Delta = \max_{\Delta=1,\ldots,d-1}}$$

(64)

and Eve’s error is $\tilde{e}_{E} \leq \frac{1}{2} p_{NL}$. Consequently, AD will certainly work for

$$p_{NL} > \max_{\Delta=1,\ldots,d-1} e_{AB}(\Delta).$$

(65)

All the quantities in this relation can be computed from the Alice-Bob correlations alone. As before, we have studied this condition numerically for $d=100$. This time, we have concentrated on correlations of the form $P = w P_{me} + \frac{1-w}{d}$, where $P_{me}$ are the correlations obtained when measuring the maximally entangled state (this is, of course, a completely arbitrary choice, but seems interesting from the point of view of quantum physics). One observes that, as expected, the use of two-way postprocessing significantly decreases the value $p_{NL}(0)$ of $p_{NL}$ for which no secrecy can be extracted. Moreover, $p_{NL}(0)$ decreases when $d$ increases, but very slowly; so slowly, in fact, that it cannot be guessed from the numerical results, whether ultimately $p_{NL}(0) \to 0$ for $d \to \infty$.

In summary, we have obtained a few results for generic $d$. In spite of a large number of assumptions and approximations (not least the choice of the protocol), we can conjecture that secrecy can be extracted from quantum nonlocal correlations for any $d$, and more precisely, that the amount of extractable secrecy increases with increasing $d$.

V. CONCLUSIONS AND PERSPECTIVES

In conclusion, we have presented a first approach to a device-independent security proof for cryptography, expanding and generalizing the work of Ref. [3]. Under the assumption of individual attacks, we have proved that a secret key can be extracted from some no-signaling probability distributions, using only the very fact that they violate a Bell-type inequality and cannot therefore originate from shared randomness. In particular, noisy quantum states can be used to distribute correlations that are nonlocal enough to contain distillable secrecy: so our result is also of practical interest.

We would like to finish by raising some of the questions and perspectives that are opened by this work.

(i) A first objective is to extend our analysis beyond the assumption of individual attacks, proving ultimately the security against the most general attacks by an eavesdropper limited by no-signaling. A first step in this direction has been recently derived [36].

(ii) One can make a step further: Can one make a device-independent proof of security against an eavesdropper which would be limited by quantum physics? On the side of Alice and Bob, nonlocality should still be the physical basis for security, because there exists no other entanglement witness which works independently of the dimension of the Hilbert space. On the side of Eve, the requirement that she must
We thank Stefano Pironio, Sandu Popescu, and Renato Renner for discussion. This work has been supported by the European Commission, under the Integrated Project Qubit Applications (QAP) funded by the IST directorate as Contract Number 015848, and the Spanish MEC, under a Ramón y Cajal grant. We acknowledge also financial support from the Swiss NCCR “Quantum Photonics.”

ACKNOWLEDGMENTS

APPENDIX A: LOWER BOUND FOR A QUANTUM IMPLEMENTATION OF THE CHSH PROTOCOL

In this appendix, we study the security of the CHSH protocol in the standard scenario where Eve is limited by the quantum formalism, and Alice and Bob have a perfect knowledge on their quantum devices. More precisely, Alice and Bob know their Hilbert spaces are two-dimensional and they apply the spin measurements that produce the largest Bell violation for the noiseless state \(|\Phi^+\rangle\). For instance, Alice and Bob measure in the \(xz\) plane, their spin measurement being defined by the angle \(\theta\) with the \(z\) axis on the Poincaré sphere. Alice measures in the \(\theta = \pi/2\) and \(\theta = 0\) bases, corresponding to \(x = 0, 1\), respectively, while Bob does it in the \(\theta = \pi/4, 3\pi/4\) directions, corresponding to \(y = 0, 1\).

As shown in Refs. [38,39], the bound for security against the most general attacks (unconditional security) can be computed by focusing on collective attacks, where Eve prepares the same two-qubit state \(\rho_{AB}\) on all instances, but is allowed to make a coherent measurement of her ancillae after error correction and privacy amplification.

By inspection, or by using the formalism developed in Ref. [38], it can be proved that Eve’s optimal strategy uses a Bell-diagonal state of the form

\[
\rho_{AB} = \lambda_1 |\Phi^+\rangle\langle\Phi^+| + \lambda_2 (|\Phi^-\rangle\langle\Phi^-| + |\Psi^-\rangle\langle\Psi^-|) + \lambda_4 |\Psi^-\rangle\langle\Psi^-|,
\]

where \(|\Phi^\pm\rangle\) and \(|\Psi^\pm\rangle\) denote the projectors onto the Bell basis

\[
|\Phi^\pm\rangle = \frac{1}{\sqrt{2}} (|00\rangle \pm |11\rangle), \quad |\Psi^\pm\rangle = \frac{1}{\sqrt{2}} (|01\rangle \pm |10\rangle).
\]

By assumption, Eve holds a purification of each pair: before error correction and privacy amplification, and after error correction and privacy amplification.

(ii) In this paper, we have defined protocols which look as “natural” for the CHSH and the CGLMP inequalities. But there is no claim of optimality. In fact, it is not even proved that the pseudosifting that we have used is the best way of extracting secrecy from the raw correlations of CHSH-like measurements (Table I). Other protocols may be better suited for cryptographic tasks, as discussed in Ref. [37].

(iii) A particular consequence of the previous item is worth mentioning in itself. On the one hand, it has been proved that all nonlocal probability distributions have positive intrinsic information [25]. On the other hand, as mentioned several times in this paper, we have not been able to find an explicit procedure for extracting a secret key in the whole nonlocal region. This means either that a better procedure does exist, or that nonlocal distributions close to the local limit provide examples of bipartite bound information [29,30].

(iv) A technical open point, which we mentioned and would be very meaningful for the present studies, is the characterization of the quantum region in probability space for a given number of inputs and outcomes.

(v) A technical open point, which we mentioned and would be very meaningful for the present studies, is the characterization of the quantum region in probability space for a given number of inputs and outcomes.

\[
R^- \geqslant R_{DW} = I(A:B) - \chi(B:E). \quad (A3)
\]

Here, \(I(A:B)\) denotes the standard mutual information between Alice and Bob’s classical outcomes, while \(\chi(B:E)\) is the Holevo quantity for the effective channel between Bob and Eve. Indeed, Bob’s measurement outcome prepares a quantum state on Eve’s side (see [40] for more details). Contrary to the more standard situation, Eve does not know which measurement Bob has applied, so she has to sum over the two possibilities. Her states read, up to normalization,

\[
\rho_E^i = \text{tr}_{AB} [\rho_{AB} |i⟩⟨i|_{z=\pi/4} + |i⟩⟨i|_{z=3\pi/4} \otimes |ψ⟩⟨ψ|_{AB}], \quad (A4)
\]

where \(i = 0, 1\) and \(|i⟩\theta\) denote the basis elements in the direction specified by \(\theta\), as above. It is straightforward to see that for the CHSH protocol

\[
I(A:B) = 1 - h\left(\frac{1 + (\lambda_1 - \lambda_2)/\sqrt{2}}{2}\right). \quad (A5)
\]

The computation of \(\chi(B:E) = S(\rho_E^0) - [S(\rho_E^0) + S(\rho_E^1)]/2\), where \(S(\rho)\) denotes the von Neumann entropy for a state \(\rho\), is slightly more involved. However, after some patient algebra one can see that the maximum of this quantity is obtained, for fixed disturbance, when
\[ \lambda_1 = (1 - D)^2, \quad \lambda_2 = D(1 - D), \quad \lambda_3 = D^2, \]  \hspace{1cm} (A6)

which defines Eve’s optimal attack. Not surprisingly, this attack corresponds to a phase covariant cloning machine (see for instance [41]) that optimally clones all the states in the \( \chi \) plane. This attack is also optimal for the standard BB84 protocol.

The obtained critical disturbance for CHSH is \( D \approx 12\% \). This is larger than the well-known Shor-Preskill bound \( D = 11\% \) for security of BB84 [42]. This bound, however, has recently been improved by allowing any of the parties, say Alice, to introduce some preprocessing of her outcome before the reconciliation [38]. Alice then flips her bit with probability \( q \). This local noise worsens the correlations between Alice and Bob, but it deteriorates in a stronger way the correlation between Alice and Eve. For any value of the disturbance there exists an optimal preprocessing \( g(D) \), depending on the protocol, which maximizes the key rate. This explains the improvement on the critical disturbance that moves up to \( D \approx 12.4\% \) both for BB84 and for the CHSH protocol described here.

Actually, the close relation between the CHSH protocol and the BB84 protocol is made clear by this preprocessing. Let \( Q_b = D \) and \( q_b \) be, respectively, the QBER and the preprocessing rate for BB84, and \( Q_C \) and \( q_C \) denote the same quantities for the CHSH protocol. Note first that the channel defined in (A6) induces a QBER \( Q_b = D \) in BB84, and a QBER \( Q_C = Q_0 + \frac{D}{2} \) for the CHSH protocol; whence,

\[ Q_b = \sqrt{2}(Q_C - Q_0). \]  \hspace{1cm} (A7)

It can then be shown that \( R_{CHSH}^{CHSH}(Q_C, q_C) = R_{BB84}^{BB84}(Q_b, q_b) \) when the QBERs are related as (A7) and when

\[ q_b = Q_0 + \frac{q_C}{\sqrt{2}}. \]  \hspace{1cm} (A8)

These two relations imply

\[ Q' = Q_C(1 - q_C) + (1 - Q_C)q_C = Q_0(1 - q_b) + (1 - Q_b)q_b. \]  \hspace{1cm} (A9)

Consider first for clarity the case \( q_C = 0 \). The error in CHSH due to the nonperfect overlap of the bases can be attributed to the application of a preprocessing \( q_b = Q_0 \) onto the correlations obtained with perfectly overlapping bases (indeed, the errors \( Q_0 \) are intrinsic to the protocol, and Eve cannot gain anything from them). In general, the rates obtained for the CHSH protocol in the standard quantum scenario coincide with those derived for the BB84 protocol when the preprocessing is optimized under the constraint \( q_b = Q_0 \). If we now compare \( R_{CHSH}^{CHSH}(Q_C) \) and \( R_{BB84}^{BB84}(Q_b) \) for a fixed value of \( D \) [that is, (A7) holds] and choosing the optimal preprocessing in each case, we find the following. For small error rates, the optimal preprocessing on BB84 is smaller than \( Q_0 \); in other words, even for \( q_C = 0 \) CHSH corresponds to BB84 with an excessive preprocessing, whence \( R_{BB84}^{BB84}(Q_b) < R_{BB84}^{BB84}(Q_b) \). The optimal preprocessing on BB84 becomes equal to \( Q_0 \) for \( D \approx 11.7\% \); from this point on, the optimal \( q_C \) is larger than zero, and the rates for the two protocols become identical: \( R_{CHSH}^{CHSH}(Q_C) = R_{BB84}^{BB84}(Q_b) \). In particular, as announced, both become zero for \( D \approx 12.4\% \).

**APPENDIX B: THE CGLMP INEQUALITY IN QUANTUM PHYSICS**

The CGLMP inequalities [33] have been the object of several studies in the context of quantum physics. Here we summarize the results without any proof.

One unexpected feature of CGLMP is the fact that the maximal violation is not reached by measurements on the maximally entangled state [43]. Also unexpected is the fact that the settings that maximize the violation are the same for a wide class of states (including the maximal entangled one and the one which gives the maximal violation). These are the settings we consider here. We label them \( A_0 \) and \( A_1 \) for Alice, \( B_0 \) and \( B_1 \) for Bob:

\[ A_x = \{\Psi_x(a)\}_{a=0}^{d-1}, \quad \Psi_x(a) = \sum_{k=0}^{d-1} e^{i(2\pi/d)ak} (e^{i2\theta_k} |k\rangle), \]  \hspace{1cm} (B1)

\[ B_y = \{\Phi_y(b)\}_{b=0}^{d-1}, \quad \Phi_y(b) = \sum_{k=0}^{d-1} e^{-i(2\pi/d)bk} (e^{i2\theta_k} |k\rangle). \]  \hspace{1cm} (B2)

In operational terms, both Alice and Bob apply first global phases in the computational basis, then make a quantum Fourier transform (Bob makes the inverse as Alice), and finally measure in the new basis and outcome the value \( a \) or \( b \).

Consider quantum states that are Schmidt-diagonal in the computational basis,

\[ |\psi\rangle = \sum_{k=0}^{d-1} c_k |kk\rangle, \]  \hspace{1cm} (B3)

with \( c_k \in \mathbb{R} \). On this family, one finds

\[ P(a,b|x,y) = \frac{1}{d^2} \sum_{k,k'=0}^{d-1} c_k c_{k'} \cos \left( \frac{2\pi}{d} \Delta + \phi_x + \theta_x (k - k') \right), \]  \hspace{1cm} (B4)

with \( \Delta = a - b \). The only freedom left is the choice of the four angles \( \phi_x \) and \( \theta_x \). The settings we are interested in are defined by

\[ \phi_0 = 0, \quad \phi_1 = \frac{\pi}{d}, \quad \theta_0 = -\frac{\pi}{2d}, \quad \theta_1 = \frac{\pi}{2d}. \]  \hspace{1cm} (B5)

With these settings, (44) holds.

For the case \( d = 3 \), all the interesting states found to date are of the form (53). For instance, \( \gamma = 1 \) is the maximally entangled state; the maximal violation is obtained for \( \gamma = 1 - \frac{2}{3} = 0.7923 [43] \); the largest Kullback-Leibler distance from the set of local distributions is obtained for \( \gamma = 0.6529 [44] \); and we have shown above (Sec. IV C) that the maximal amount of secret key rate under one-way processing is found for \( \gamma = 0.9875 \). For the states \( |\psi(\gamma)\rangle \), the \( p_\Delta = P(a, \Delta - a | 0, 0) \) are
\[ p_0 = \frac{1}{3} \left( 1 + \frac{1 + 2\sqrt{3}y}{2 + y^2} \right), \]
\[ p_1 = \frac{1}{3} \left( 1 + \frac{2}{2 + y^2} \right), \]
\[ p_2 = \frac{1}{3} \left( 1 + \frac{1 - 2\sqrt{3}y}{2 + y^2} \right). \]  

(B6)

APPENDIX C: DEPOLARIZATION FOR ARBITRARY D

1. The procedure

Suppose Alice and Bob share initially an arbitrary no-signaling probability distribution \( P(a,b|x,y) \). The depolarization procedure that brings \( P \) in the slice defined by (44) is very similar to the one described in Ref. [25] for \( d=2 \). It consists of two steps.

Step 1. Alice chooses \( k \in \{0, \ldots, d-1\} \) with probability \( \frac{1}{d} \) and communicates it to Bob on a public channel. Both Alice and Bob perform
\[
a \rightarrow a + k, \\
b \rightarrow b + k. 
\]  

(C1)

This implements \( P \rightarrow P_1 \), which is such that \( P_1(a,b|x,y) = \frac{1}{d} \sum_k P(a+k,b+k|x,y) \) and is consequently a function only of \( \Delta = a-b \).

Step 2. With probability \( \frac{1}{d} \), Alice chooses one of the following four procedures and asks Bob on the public channel to act accordingly:

\[
\begin{array}{c|c|c}
A \backslash B & 1 & 0 \\
\hline
1 & 1_{d'} & 0 \cdot_{d'} \cdot 0 \\
\cdot & 0 & 0 \cdot 0 \cdot 0 \\
1 & 1_{d'} & 0 \cdot_{d'} \cdot 0 \\
\hline
\end{array}
\]  

where boldface numbers indicate arrays containing all ones or all zeros, \( 1_{d'} \) is the identity matrix of dimension \( d' \times d' \), and where \( U_{d'} \) is the \( d' \times d' \) matrix
\[
U_{d'} = \begin{pmatrix}
0 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & 0 & \cdots & 0
\end{pmatrix}.
\]  

(C4)

The arrays (39) and (C3) allow us to compute immediately the “scalar product”
\[
\langle I_{d'}, \hat{P}(d') \rangle = 1 - \frac{1}{d'} 
\]  

generalizing the results we gave in the main text for \( d'=2 \) and \( d'=d \).

By following the steps of the depolarization protocol, one finds that \( \hat{P}(d') \) goes to the distribution in the slice which is given by
\[
\hat{P}(d') \rightarrow \hat{P}_2(d') = \begin{pmatrix}
p_0 & \frac{1}{4d'} \\
p_0 & \frac{1}{4d'} \\
p_0 & \frac{1}{4d'} \
\end{pmatrix},
\]  

(C6)

and obviously all the other \( p_j \) are zero. Using (48), one can verify that \( \langle I_{d'}, \hat{P}_2(d') \rangle = 1 - \frac{1}{d'} \); the violation is preserved. As we said in the main text, this is not peculiar to this example but is a general feature, as we show in the next section.

3. Preservation of the violation of CGLMP

We want to prove that this depolarization preserves the violation of the CGLMP inequality, that is

\[ 4P_2(a,b|1,1) = P_1(a,b|1,1) + P_1(-a,-b+1|0,1) + P_1(-a-1,-b|1,0) + P_1(a+1,b|0,0), \]
\[ 4P_2(a,b|0,0) = P_1(a,b|0,0) + P_1(-a,-b|0,1) + P_1(-a-1,-b|1,1) + P_1(a+1,b+1|0,1), \]
\[ 4P_2(a,b|0,1) = 4P_2(a,b|1,0) = P_1(a,b|0,1) + P_1(-a-1,-b|1,1) + P_1(-a,-b|0,0) + P_1(a+1,b+1|0,1), \]
The easiest way is to write down $I_d$ as it appeared in the original paper [33], namely, $I_d \leq 2$ with

$$I_d = \sum_{k=0}^{[d/2]-1} \left( 1 - \frac{2k}{d-1} \right) \left[ P(-k|0,0) + P(k|0,1) + P(k|1,0) \right] + P(-k-1|1,1) - [P(k+1|0,0) + P(-k-1|0,1) + P(k|1,1)],$$

where $P(\Delta|x,y) = P(a-b=\Delta|x,y)$. The link between $\bar{I}_d$ and our definition of $I_d$ is provided by

$$I_d = \frac{d-1}{2d}(-2 + \bar{I}_d).$$

Using the expression of $\bar{I}_d$, the proof is straightforward. In fact, Step 1 keeps by definition all the $P(\Delta|x,y)$ constant, while Step 2 keeps both sums in $\lfloor \ldots \rfloor$ constant.

**APPENDIX D: DETERMINISTIC STRATEGIES THAT SATURATE CGLMP**

We present here a more detailed study of the extremal points that lie on the CGLMP facet, completing what has been written in Sec. IV B 2.

Consider the array which represents the CGLMP inequality $I_d \leq 0$, Eq. (39); here, it is more convenient to look at it as having $2d \times 2d$ entries [35]. Let $[i,j]$ denote an entry of this array. For the deterministic strategy $\{a(0), a(1); b(0), b(1)\}$, the value of CGLMP is simply

$$I_d = -2 + \sum_{x,y=0}^{1} I[a(x),b(y)] = -2 + \delta[b(0) \geq a(0)] + \delta[a(0) \geq b(1)] + \delta[a(1) \geq b(0)] - \delta[a(1) \geq b(1)],$$

where the $-2$ comes from the marginals of $a(0)$ and $b(0)$, and where $\delta[C]$ is equal to 1 if condition $C$ is satisfied and to 0 otherwise. The inequality is saturated by all the strategies such that $I_d=0$.

Consider the points such that $b(0)=b(1)$; the last two conditions become equal and the $\delta$'s compensate each other for all $a(1)$, so the only way to saturate the inequality is to fulfill both $b(0) \geq a(0)$ and $a(0) \geq b(1)$; whence, $a(0)=b(0) = b(1)$ as announced in the main proof. The proof of the analog statement in the case $b(0)=b(1)-1$ is similarly done by inspection. One first considers $b(0)<d-1$. In this case, $b(1) > b(0)$; therefore, the first two conditions can be both fulfilled, whatever $a(0)$ is. One can then easily verify that only the choice $a(1)=b(0)$ leads to a saturation of the inequality. The last remaining case is $b(0)=d-1, b(1)=0$. It can be read directly from the array and leads to the same conclusion.

So, we have proved the properties of sets $\mathcal{L}_0$ and $\mathcal{L}_1$, which consist of $d^2$ points each. We still have to prove that the number of points on the facet is larger than $D=4d(d-1)$. This is easily done by noticing the following: the four “natural” relations associated to the CGLMP inequality, those that are simultaneously fulfilled by $P_{PR_{2,d'}}$ are

$$R_{00}: a(0) = b(0),$$
$$R_{01}: a(0) = b(1),$$
$$R_{10}: a(1) = b(0),$$
$$R_{11}: a(1) = b(1) - 1.$$  

Because of the specific pseudosifting of our cryptographic protocol, we grouped them by pairs according to Alice’s input. But from the standpoint of the inequality, any pairwise grouping is equally meaningful. It can indeed be easily verified using (D1) that all the points that fulfill at least two among these relations saturate the inequality. There are therefore $4d$ strategies that fulfill three relations and $6d(d-2)$ strategies that fulfill exactly two relations. In conclusion, by looking only at the points that fulfill at least two among the four relations (D2), we have already $6d^2-8d$ deterministic points on the CGLMP facet, and this number is larger than $D$ for $d \geq 2$. We note that the list is exhaustive for $d=3$ (see Appendix E), but not in general. For instance, for $d=5$, the strategy $\{a(0)=4, a(1)=1; b(0)=5, b(1)=3\}$ fulfills none of the relations (D2) but achieves, nevertheless, $I_d=0$.

**APPENDIX E: EXPLICIT ANALYSIS FOR D=3**

1. **Deterministic strategies on the facet**

We give here the explicit list of the 30 deterministic strategies that saturate CGLMP. We note $r=0,1,2$.

The 12 strategies in $\mathcal{L}^3_0$ are

$$L_{0}^{3}; \begin{cases} L_{1,1}^{3} = \{a(x) = r, b(y) = r\}, \\ L_{1,2}^{3} = \{a(x) = r-x, b(y) = r\}, \end{cases}$$

The six strategies in $\mathcal{L}^2_0$ are

$$L_{0}^{2}; \begin{cases} L_{1,1}^{2} = \{a(x) = r+x, b(y) = r\}, \\ L_{1,2}^{2} = \{a(x) = r+y, b(y) = r+y+1\}, \end{cases}$$

The 12 strategies outside $\mathcal{L}$ are

$$L_{r}^{c}; \begin{cases} L_{1,1}^{c} = \{a(x) = r, b(y) = r-y\}, \\ L_{1,2}^{c} = \{a(x) = r+x, b(y) = r-y\}, \\ L_{c,3}^{c} = \{a(x) = r+x, b(y) = r-y+1\}, \\ L_{c,4}^{c} = \{a(x) = r-x, b(y) = r+y\}, \end{cases}$$

As we said in the main text, the decomposition of $M_2$ given in (50) is only one possible decomposition, the one
TABLE VI. Probability distribution Alice-Bob-Eve after pseudosigning for $d=3$ and Alice’s setting $x$, for the general decomposition (E6) of $M_2$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$b=0$</th>
<th>$b=1$</th>
<th>$b=2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=0$</td>
<td>$p_{NL}/3$ (??,?) [ f(p^{0-x}_2) (0,0) ] [ g(p^{0-x}_2) (0,0) ] [ (1-p_0)/6 (0,0,?) ] [ (1-p_0)/6 (0,0,?) ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a=1$</td>
<td></td>
<td>$p_{NL}/3$ (??,?) [ f(p^{1-x}_2) (1,1) ] [ g(p^{1-x}_2) (1,1) ] [ (1-p_0)/6 (1,?) ] [ (1-p_0)/6 (1,?) ]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$a=2$</td>
<td></td>
<td>$p_{NL}/3$ (??,?) [ f(p^{2-x}_2) (2,2) ] [ g(p^{2-x}_2) (2,2) ] [ (1-p_0)/6 (2,?) ] [ (1-p_0)/6 (2,?) ]</td>
</tr>
</tbody>
</table>

which optimizes Eve’s information on Bob’s symbol. It can be checked that the general decomposition is defined by

\[
p_{e,2} = \frac{1}{6} - (p_e^f + p_{e,1}^f),
\]

\[
p_{e,3}^f = p_{e,1}^f,
\]

\[
p_{e,4}^f = \frac{1}{6} - (p_e^f + p_{e,1}^f).
\]

There are thus six free parameters \{$p_{e,2}^f, p_{e,1}^f$\}, constrained of course by the positivity of probabilities (in particular, none of these parameters can exceed $\frac{1}{2}$). A possible realization of $M_2$ is the equiprobable mixture of the eighteen points which are not in $L^3$. The choice leading to (50) is the equiprobable mixture of the six points in $L^2$ ($p_{e,2}^f = \frac{1}{6}$, implying automatically $p_{e,1}^f = 0$).

2. Alice-Bob-Eve correlations

Having the explicit deterministic strategies, it is a matter of patience to derive the tables for the correlations Alice-Bob-Eve. The result is given in Table VI, in which we have introduced the notations

\[
f(p) = \frac{2p_1}{3} + 2p_2p, \quad g(p) = \frac{1-p_0}{3} - 2p_2p.
\]

Note that in each of the nine cells, the sum of the probabilities does not depend on the $p_{e,2}^f$ as it should: the decomposition of $M_2$ is known only to Eve. Eve is obviously interested in maximizing the probability of knowing both symbols, measured by $f(p)$; whence, the choice $p_{e,2}^f = \frac{1}{6}$ made in the main text.

3. About nonlocal points that violate CGLMP

Here, we want to list some nonlocal points other than $P_{PR_{2,3}}$ that violate CGLMP, and study their relation with the slice (44).

Consider first the nonlocal points equivalent to $P_{PR_{2,2}} = P_{PR}$. There are 24 such points in the no-signaling polytope: in fact, there are three choices for the two outcomes \{$(0,1), (0,2)$ or $(1,2)$\} and for each choice there are eight PR-like points, obtained as usual by relabeling inputs and/or outputs. By inspection, it can be seen that $I_3 > 0$ (in fact, $I_3 = \frac{1}{2}$) for our representative (39) of $I_3$ is achieved only by three PR-like points defined by

$$\begin{cases}
a = b & \text{if } xy = 0 \quad \text{for } (a,b) \in (0,1), \ (0,2), \ (1,2).
\end{cases}$$

It is readily seen that no mixture of these three strategies can belong to the slice (44). To obtain all the marginals equal to $\frac{1}{4}$, the only possible mixture is the equiprobable one. This one reads

\[
\begin{array}{cccc}
A & V & B & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 \\
\end{array}
\]

\[
M_{PR} = \frac{1}{3}
\]

\[
\begin{array}{cccc}
1 & 1 & 0 & 0 & 1 & 2 \\
1 & 0 & 1 & 1 & 2 & 0 \\
\end{array}
\]

and is clearly not of the form (45). This negative result is important for our study: had such a mixture belonged to the slice, Eve would have sent these nonlocal points, for which she would have gained some information (because in each case one result is impossible).

Actually, there is a mixture of nonlocal points on the slice. It is a mixture of other $P_{PR_{2,3}}$-like strategies, which optimize the violation of different representatives of CGLMP, and violate our representative by $I_3 = \frac{1}{2}$. The strategies are those in which $b-a$ is equal to $-xy, x(2-y), y(2-x)$, and $(x+y + 1) \bmod 2$. (Note that this last one is indeed a $P_{R_{2,3}}$. The nonlocality is embedded on the fact that the right-hand side is computed modulo 2 instead of modulo 3, as is the case for the others.) The equiprobable mixture of these four strategies is the point $p_0 = \frac{1}{2}$ and $p_1 = 0$ in the slice. Obviously, Eve has no interest in sending these strategies instead of the $P_{PR_{2,3}}$ which gives the maximal violation. In all cases she is going to learn nothing about the outcomes.
[5] We explicitly stress an important point. There are several ways of guaranteeing that no information leakage has taken place about \(x\) and \(y\). One of these ways is based on physics: if the measurement-and-detection events of Alice and Bob are space-like separated events, then relativity guarantees that no signal has propagated from one location to the other. This is a sufficient but not necessary condition for independence. Our work only assumes that indeed there is no information leakage, independently of the way this is achieved.
[19] If the polytope is embedded in a \(D\)-dimensional space, a facet is a hyperplane of dimension \(D-1\). To define such a hyperplane, one needs \(D\) linearly independent points, just as one needs two different points to define a line, three nonaligned points to define a plane, etc.
[31] This map looks counterintuitive from Eve’s point of view. Why should she map \((0, ?)\) and \((1, ?)\), where she knows Alice’s variable, into \((? , ?)\)? Recall, however, that this map is only a mathematical tool for the computation of an upper bound on the secret-key rate.
[32] We could have used the same formalism for the case \(d=2\); we did not, because we found it preferable to tackle that case without the introduction of a lot of notations. All that we write in Sec. IV for general \(d\) applies in particular for \(d=2\).
[35] The array (39) that defines \(I_d\) has exactly the same form, whether one considers it as having \(2(d-1)\times 2(d-1)\) entries, or in the “full” form, which has \(2d\times 2d\) entries. It is in fact easy to verify that all the terms that are added cancel out.
Device-Independent Security of Quantum Cryptography against Collective Attacks

Antonio Acín,1,2 Nicolas Brunner,3 Nicolas Gisin,3 Serge Massar,4 Stefano Pironio,1 and Valerio Scarani3
1ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain
2ICREA-Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain
3Group of Applied Physics, University of Geneva, CH-1211 Geneva 4, Switzerland
4Laboratoire d’information quantique, CP 225, Université Libre de Bruxelles, 1050 Brussels, Belgium

(Received 20 February 2007; published 4 June 2007)

We present the optimal collective attack on a quantum key distribution protocol in the “device-independent” security scenario, where no assumptions are made about the way the quantum key distribution devices work or on what quantum system they operate. Our main result is a tight bound on the Holevo information between one of the authorized parties and the eavesdropper, as a function of the amount of violation of a Bell-type inequality.

DOI: 10.1103/PhysRevLett.98.230501
PACS numbers: 03.67.Dd, 03.65.Ud

Quantum key distribution (QKD) allows two parties, Alice and Bob, to generate a secret key in the presence of an eavesdropper, Eve [1]. All QKD schemes rely for security on several assumptions. The basic one is that any eavesdropper, however powerful, must obey the laws of quantum physics. In addition to it, there are two other requirements, without which no shared secret key can be established. The first one is the freedom and secrecy of measurement settings: on each particle, both Alice and Bob should be allowed to choose freely among at least two measurement settings [e.g., the two bases of the Bennett-Brassard 1984 (BB84) protocol [2]] and this choice should not be known to Eve, at least as long as she can act on the incoming quantum states (in BB84, the bases are revealed, but only after the measurements are performed). The second requirement, even more obvious, is the secrecy of outcomes: at no stage should there be a leakage of information about the final key. These two requirements can be summarized by saying that no unwanted classical information must leak out of Alice’s and Bob’s laboratories. If an implementation has a default in this point (e.g., if a Trojan Horse attack is possible, or if Eve can access Bob’s computer), no security can be guaranteed.

In addition to these essential requirements, existing security proofs [3–5] assume that Alice and Bob have (almost) perfect control of the state preparation and of the measurement devices. This assumption is often critical: for instance, the security of the BB84 protocol is entirely compromised if Alice and Bob, instead of sharing qubits as usually assumed, share four-dimensional systems [6,7].

At first sight, control of the apparatuses seems to be an inescapable requirement. Remarkably, this is not the case: we present here a device-independent security proof against collective attacks by a quantum Eve for the protocol described in Ref. [8]. Our proof holds under no other requirements than the essential ones listed above. It is therefore “device independent” in the sense that it needs no knowledge of the way the QKD devices work, provided quantum physics is correct and provided Alice and Bob do not allow any unwanted signal to escape from their laboratories.

In a collective attack, Eve applies the same attack on each particle of Alice and Bob, but no other limitations are imposed to her. In particular, she can keep her systems in a quantum memory and perform a (coherent) measurement on them at any time. Collective attacks are very meaningful in QKD because a bound on the key rate for these attacks becomes automatically a bound for the most general attacks if a de Finetti theorem can be applied, as is the case in the usual security scenario [9].

The physical basis for our device-independent security proof is the fact that measurements on entangled particles can provide Alice and Bob with nonlocal correlations, i.e., correlations that cannot be reproduced by shared randomness (local variables), as detected by the violation of Bell-type inequalities. Considered in the perspective of QKD, the fact that Alice’s and Bob’s symbols are correlated in a nonlocal way, whatever be the underlying physical details of the apparatuses that produced those symbols, implies that Eve cannot have full information about them, otherwise her own symbol would be a local variable able to reproduce the correlations.

This intuition was at the origin of Ekert’s 1991 proposal [10] and implicit in subsequent works [11,12]. Quantitative progress has been possible, however, only recently, thanks to the pioneering work of Barrett, Hardy, and Kent [13] and to further extensions [6,8,14]. For conceptual interest and mathematical simplicity, all these works studied security against a supra-quantum Eve, who could perform any operation compatible with the no-signaling principle. The proof of Ref. [13] applies only to the zero-error case; those in Refs. [6,8] allow for errors but restrict Eve to perform individual attacks; Masanes and Winter [14] proved non-universally composable security under the assumption that Eve’s attack is arbitrary but is not correlated with the classical post-processing of the raw key. In this Letter, we focus on the more realistic situation in which Eve is constrained by quantum physics, and we prove universally composable security against collective attacks.
The protocol.—The protocol that we study is a modification of the Ekert 1992 protocol [10] proposed in Ref. [8]. Alice and Bob share a quantum channel consisting of a source that emits pairs of entangled particles. On each of her particles, Alice chooses between three possible measurements $A_0$, $A_1$, and $A_2$, and Bob between two possible measurements $B_1$ and $B_2$. All measurements have binary outcomes labeled by $a_i$, $b_j \in \{+1, -1\}$ (note, however, that the quantum systems may be of dimension larger than 2). The raw key is extracted from the pair $(A_0, B_1)$. In particular, the quantum bit error rate (QBER) is $Q = \text{prob} (a_0 \neq b_1)$. As mentioned in the introduction, Eve’s information is bounded by evaluating Bell-type inequalities, since these are the only entanglement witnesses which are independent of the details of the system. In our case, Alice and Bob use the measurements $A_1$, $A_2$, $B_1$, and $B_2$ on a subset of their particles to compute the Clau泽-Horne-Shimony-Holt (CHSH) polynomial [15]

$$S = \langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle,$$

which defines the CHSH inequality $S \leq 2$. We note that there is no a priori relation between the value of $S$ and the value of $Q$: these are the two parameters which are available to estimate Eve’s information. Without loss of generality, we suppose that the marginals are random for each measurement, i.e., $\langle a_i \rangle = \langle b_j \rangle = 0$ for all $i$ and $j$. Were this not the case, Alice and Bob could achieve it a posteriori through public one-way communication by agreeing on flipping a chosen half of their bits. This operation would not change the value of $Q$ and $S$ and would be known to Eve.

Eavesdropping.—In the device-independent scenario, Eve is assumed not only to control the source (as in usual entanglement-based QKD), but also to have fabricated Alice’s and Bob’s measuring devices. The only data available to Alice and Bob to bound Eve’s knowledge are the observed relation between the measurement settings and outcomes, without any assumption on how the measurements are actually carried out or on what system they operate. In complete generality, we may describe this situation as follows. Alice, Bob, and Eve share a state $|\Psi_{ABE}\rangle$ in $H_A^2 \otimes H_B^n \otimes H_E$, where $n$ is the number of bits of the raw key. The dimension $d$ of Alice and Bob Hilbert spaces $H_A = H_B = \mathbb{C}^d$ is unknown to them and fixed by Eve. The measurement $M_k$ yielding the $k$th outcome of Alice is defined on the $k$th subspace of Alice and chosen by Eve. This measurement depends on the $k$th setting $A_{k_{-1}}$ chosen by Alice, but possibly also on all previous settings and outcomes: $M_k = M(A_{k_{-1}}, A_{k_{-1}})$, where $A_{k_{-1}} = (A_{j_1}, \ldots, A_{j_{k-1}})$ and $a_{k_{-1}} = (a_{j_1}, \ldots, a_{j_{k-1}})$. The situation is similar for Bob.

Collective attacks.—In this Letter, we focus on collective attacks where Eve applies the same attack to each system of Alice and Bob. Specifically, we assume that the total state shared by the three parties has the product form $|\Psi_{ABE}\rangle = |\Psi_{ABE}\rangle^\otimes n$ and that the measurements are a function of the current setting only, e.g., for Alice $M_k = M(A_j) = A_{j_{k_{-1}}}$. For collective attacks, the secret-key rate $r$ under one-way classical post-processing from Bob to Alice is lower bounded by the Devetak-Winter rate [16],

$$r \geq r_{\text{DW}} = I(A_0; B_1) - \chi(B_1; E),$$

which is the difference between the mutual information between Alice and Bob, $I(A_0; B_1) = 1 - h(Q)$ ( $h$ is the binary entropy), and the Holevo quantity between Eve and Bob, $\chi(B_1; E) = S(\rho_E) - \frac{1}{2} \sum_{b_{k_{-1}}} S(\rho_{E b_{k_{-1}}})$. Note that the rate is given by (2) because $\chi(A_0; E) \geq \chi(B_1; E)$ holds for our protocol [8]; it is therefore advantageous for Alice and Bob to do the classical post-processing with public communication from Bob to Alice.

Upper bound on the Holevo quantity.—To find Eve’s optimal collective attack, we must find the largest value of $\chi(B_1; E)$ compatible with the observed parameters without assuming anything about the physical systems and the measurements that are performed. Our main result is the following.

Theorem.—Let $|\psi_{ABE}\rangle$ be a quantum state and $\{A_1, A_2, B_1, B_2\}$ a set of measurements yielding a violation $S$ of the CHSH inequality. Then after Alice and Bob have symmetrized their marginals,

$$\chi(B_1; E) \leq h \left( \frac{1 + \sqrt{(S/2)^2 - 1}}{2} \right).$$

Before presenting the proof of this bound, we give an explicit attack which saturates it; this example clarifies why the bound (3) is independent of $Q$. Eve sends to Alice and Bob the two-qubit Bell-diagonal state

$$\rho_{AB} = \frac{1 + C}{2} \rho_{\Phi^+} + \frac{1 - C}{2} \rho_{\Phi^-},$$

where $\rho_{\Phi^\pm}$ are the projectors on the Bell states $|\Phi^{\pm}\rangle = ((00 \pm 11))\sqrt{2}$ and $C = \sqrt{(S/2)^2 - 1}$. She defines the measurements to be $B_1 = \sigma_2$, $B_2 = \sigma_3$, and $A_{1,2} = \frac{1}{\sqrt{2}} \sigma_3 \pm \frac{i}{\sqrt{2}} \sigma_2$, choosing $A_0$ to be $\sigma_r$ with probability $1 - 2Q$ and to be a randomly chosen bit with probability $2Q$. This attack is impossible within the usual assumptions because here not only the state $\rho_{AB}$, but also the measurements taking place in Alice’s apparatus depend explicitly on the observed values of $S$ and $Q$. The state (4) has a nice interpretation: it is the two-qubit state which gives the highest violation $S$ of the CHSH inequality for a given value of the entanglement, measured by the concurrence $C$ [17].

We now present the proof of the theorem stated above, in four steps; more details will be given in a forthcoming paper.

Proof, Step 1.—It is not restrictive to suppose that Eve sends to Alice and Bob a mixture $\rho_{AB} = \sum_i P_i \rho_{AB}^i$ of two-qubit states, together with a classical ancilla (known to her) that carries the value $c$ and determines which measurements $A_i$ and $B_j$ are to be used on $\rho_{AB}^i$. 

230501-2
The proof of this first statement relies critically on the simplicity of the CHSH inequality (two binary settings on each side). We present the argument for Alice; the same holds for Bob. First, we may assume that the two measurements $A_{1,2}$ of Alice are von Neumann measurements, if necessary by including ancillas in the state $\rho_{AB}$ shared by Alice and Bob. Thus $A_1$ and $A_2$ are Hermitian operators on $\mathbb{C}^{d}$ with eigenvalues $\pm 1$. It follows from this that $A_1 A_2$ is a unitary, hence diagonalizable, operator. In the basis of $\mathbb{C}^{d}$ formed by the eigenvectors of $A_1 A_2$, one can show that $A_1$ and $A_2$ are block diagonal, with blocks of size $1 \times 1$ or $2 \times 2$ [18]. In other words, $A_j = \sum_i P_i A_i P_c$, where the $P_i$s are projectors of rank 1 or 2. From Alice’s standpoint, the measurement of $A_1$ thus amounts at projecting in one of the (at most) two-dimensional subspaces defined by the projectors $P_c$, followed by a measurement of the reduced observable $P_c A_i P_c = \hat{a}_i^* \cdot \hat{a}_i$. Clearly, it cannot be worse for Eve to perform the projection herself before sending the state to Alice and learning the value of $c$. The same holds for Bob. We conclude that in each run of the experiment Alice and Bob receive a two-qubit state. The deviation from usual proofs lies in the fact that the measurements to be applied can depend explicitly on the state.

Proof, Step 2.—Each state $I_{AB}^\rho$ can be taken to be a Bell-diagonal state and the measurements $A_j^\rho$ and $B_j^\rho$ to be measurements in the $(x, z)$ plane.

To reduce the problem further in this way, we use some freedom in the labeling together with two applications of a suitable symmetry. The reason is as follows. First note that since the (classical) randomization protocol that ensures $\langle a_i \rangle = \langle b_j \rangle = 0$ is done by Alice and Bob through public communication, we can as well assume that it is Eve who does it; i.e., she flips the value of each outcome bit with probability one half. But because the measurements of Alice and Bob are in the $(x, z)$ plane, we can equivalently, i.e., without changing Eve’s information, view the classical flipping of the outcomes as the quantum operation $\rho \rightarrow \tilde{\rho} = (\sigma_y \otimes \sigma_y) \rho (\sigma_y \otimes \sigma_y)$ on the state $\rho$. We conclude that it is not restrictive to assume that Eve is in fact sending the mixture $\tilde{\rho} = \frac{1}{2} (\rho + \tilde{\rho})$, i.e., that she is sending a state invariant under $\sigma_y \otimes \sigma_y$. Now, through an appropriate choice of basis that leaves invariant the $(x, z)$ plane, and corresponding to the freedom to define the orientation of $\tilde{x}$ and the direction of $\tilde{y}$ for both Alice and Bob, every $\sigma_y \otimes \sigma_y$ invariant two-qubit state can be written in the Bell basis, ordered as $\{|\Phi^+\rangle, |\Psi^-\rangle, |\Phi^-\rangle, |\Psi^+\rangle\}$, in the canonical form

$$\tilde{\rho} = \begin{pmatrix} \lambda_{\Phi^+} & i r_1 \\ -i r_1 & \lambda_{\Psi^-} \\ \lambda_{\Phi^-} & -i r_2 \\ i r_2 & \lambda_{\Psi^+} \end{pmatrix},$$

with $\lambda_{\Phi^+} \equiv \lambda_{\Psi^-}, \lambda_{\Phi^-} \equiv \lambda_{\Psi^+}$ and $r_1, r_2$ real.

Finally, we repeat an argument similar to the one given above: since $\tilde{\rho}$ and its conjugate $\tilde{\rho}^*$ produce the same statistics for Alice and Bob’s measurements and provide Eve with the same information, we can suppose without loss of generality that Alice and Bob rather receive the mixture $\frac{1}{2}(\rho + \tilde{\rho}^*)$, which is Bell diagonal.

Proof, Step 3.—For a Bell-diagonal state $\rho_\lambda$ with eigenvalues $\lambda$ ordered as above and for measurements in the $(x, z)$ plane,

$$\chi_\lambda(B_1:E) \leq F(S_\lambda) = h\left(1 + \sqrt{(S_\lambda/2)^2 - 1/2}\right),$$

where $S_\lambda = 2 \sqrt{2} (\lambda_{\Phi^+} - \lambda_{\Psi^-})^2 + (\lambda_{\Phi^-} - \lambda_{\Psi^+})^2$ is the largest violation of the CHSH inequality by the state $\rho_\lambda$.

This step is mainly computational; we sketch it here and refer to a forthcoming paper for details. For Bell-diagonal states, for any choice of $B_1 = \cos \varphi \sigma_x + \sin \varphi \sigma_y$, one has $S(\rho_{B_1|0}) = S(\rho_{B_1|1}) \geq h(\lambda_{\Phi^+} + \lambda_{\Psi^-})$ with equality if and only if $B_1 = \sigma_x$. It follows that $\chi_\lambda(B_1:E) \leq H(\lambda) - h(\lambda_{\Phi^+} + \lambda_{\Psi^-})$. The right-hand side of this expression is in turn bounded by the function $F(S_\lambda)$ appearing in (6). It now suffices to notice that $\chi_\lambda = 2 \sqrt{2} (\lambda_{\Phi^+} - \lambda_{\Psi^-})^2 + (\lambda_{\Phi^-} - \lambda_{\Psi^+})^2$ is the maximal violation of the CHSH inequality by the state $\rho_\lambda$ [17,19]; it is achieved for $B_1 = \sigma_x$, $B_2 = \sigma_y$, and $A_1$ and $A_2$ depending explicitly on the $\lambda$’s.

Proof, Step 4.—To conclude the proof, note that if Eve sends a mixture of Bell-diagonal states $\sum_\lambda \rho_\lambda$ and chooses the measurements to be in the $(x, z)$ plane, then $\chi(B_1:E) = \sum_\lambda \rho_\lambda \chi_\lambda(B_1:E)$. Using (6), we then find $\chi(B_1:E) \leq \sum_\lambda \rho_\lambda F(S_\lambda) \leq F(\sum_\lambda \rho_\lambda S_\lambda)$, where the last inequality holds because $F$ is concave. But since the observed violation $S$ of CHSH is necessarily such that $S \leq \sum_\lambda \rho_\lambda S_\lambda$ and since $F$ is a monotonically decreasing function, we find $\chi(B_1:E) \leq F(S)$.

Key rate.—Given the bound (3), the key rate (2) can be computed for any values of $Q$ and $S$. As an illustration, we study correlations satisfying $S = 2 \sqrt{2} (1 - 2Q)$, and which arise from the state $|\Phi^+\rangle$ after going through a depolarizing channel, or through a phase-covariant cloner, or more generally from any Bell-diagonal $\rho_{AB}$ such that $\lambda_{\Phi^+} = \lambda_{\Psi^-}$ and $\lambda_{\Phi^-} = \lambda_{\Psi^+}$, when doing the measurements $A_0 = B_1 = \sigma_x$, $B_2 = \sigma_y$, $A_1 = (\sigma_z + \sigma_y)/\sqrt{2}$, and $A_2 = (\sigma_z - \sigma_y)/\sqrt{2}$. We consider these correlations because of their experimental significance, but it is important to stress that Alice and Bob do not need to assume that they perform the above qubit measurements. The corresponding key rate is plotted in Fig. 1 as a function of $Q$. For the sake of comparison, we have also plotted the key rate under the usual assumptions of QKD for the same set of correlations.
In this case, Alice and Bob have a perfect control of their apparatuses, which we have assumed to faithfully perform the qubit measurements given above. The protocol is then equivalent to Ekert’s, which in turn is equivalent to the entanglement-based version of BB84, and one finds

\[ \chi(B_1;E) \leq h(Q + S/2\sqrt{2}). \]  

(7)

If \( S = 2\sqrt{2}(1 - 2Q) \), this expression yields the well-known critical QBER of 11% \[3\], to be compared to 7.1% in the device-independent scenario (Fig. 1). [Note that the key rate given by Eq. (3) is much higher than the one against a no-signaling eavesdropper obtained by applying the security proof of [14].]

Final remarks.—Through its remarkable generality, our device-independent security proof allows us to ignore the detailed implementation of the QKD protocol and therefore applies in a simple way to situations where the quantum apparatuses are noisy or where uncontrolled side channels are present. It also applies to the situation where the apparatuses are entirely untrusted and provided by the eavesdropper herself. In this latter case, the proof cannot be applied to any existing device yet, because of the detection loophole which arises due to inefficient detectors and photon absorption. These processes imply that sometimes Alice’s and Bob’s detectors will not fire. A possible strategy to apply our proof to this new situation is for Alice and Bob to replace the absence of a click by a chosen outcome, in effect replacing detection inefficiency by noise. However, the amount of detection inefficiency that can be tolerated in this way is much lower than the one present in current quantum communication experiments. In Bell tests, this problem is often circumvented by invoking additional assumptions such as the fair sampling hypothesis—a very reasonable one if the aim is to constrain possible models of Nature, but hardly justified if the device is provided by an untrusted Eve. In the light of the present work, the “detection loophole” thus becomes a meaningful issue in applied physics.

In conclusion, we have found the optimal collective attack on a QKD protocol in the device-independent scenario, in which no other assumptions are made than the validity of quantum physics and the absence of any leakage of classical information from Alice’s and Bob’s laboratories. If a suitable de Finetti-like theorem can be demonstrated in this scenario, the bound that we have presented here will in fact be the bound against the most general attacks.

We are grateful to C. Branciard, I. Cirac, A. Ekert, A. Kent, R. Renner, and C. Simon for fruitful discussions. We acknowledge financial support from the Swiss NCCR “Quantum Photonics”, the EU Qubit Applications Project (QAP) Contract No. 015848, the Spanish Projects No. FIS2004-05639-C02-02 and Consolider QOIT, the Spanish MEC for a “Juan de la Cierva” Grant, and the IAP project Photonics@be of the Belgian Science Policy.

---