Recycling Non-locality Using Minimal Resources

by Joel Johansen

Thesis submitted for the degree of
Masters of Physics
Project duration: 12 months (full-time) 60 hp

Supervisor: Armin Tavakoli
Co-supervisors: Elna Svegborn, Nicola D' Alessandro

Lund University

Department of Physics
Division of Mathematical Physics
Date of examination: June 4, 2026

Recycling Non-locality Using Minimal Resources

Joel Johansen

June 9, 2026

Abstract

The thesis aims to characterize the scenario in which the non-locality of an entangled state shared between sequential observers can be recycled utilizing as few resources as possible. This is accomplished by stochastically combining standard projective measurements using locally generated randomness. Using an entangled two-qubit state, we analytically minimize the number of parameters necessary to specify two Clauser-Horne-Shimony-Holt inequalities. Moreover, employing the Collins-Gisin-Linden-Massar-Popescu inequality, the protocol is then further extended to the higher-dimensional case in which the non-locality of pair of entangled qutrits is recycled. The trade-off in the violation magnitudes between Bell parameters is then investigated numerically, both for qubits and qutrits, and it is found that the optimal trade-off is enabled by non-maximally entangled states and that qutrits allows for larger sequential violations than does qubits.

Keywords: *Bell non-locality, CHSH, CGLMP.*

Contents

List of Abbreviations

- CGLMP – Collins-Gisin-Linden-Massar-Popescu.
- CHSH – Clauser-Horne-Shimony-Holt.
- CPTP – Completely Positive and Trace-Preserving.
- EPR – Einstein-Podolsky-Rosen
- LHVT – Local Hidden Variable Theory.
- POVM – Positive Operator-Valued Measure

I. INTRODUCTION

A central goal of quantum information science is to study how the quantum mechanical behavior of matter on the smallest scales can be exploited in order to gain an advantage in information processing tasks. The crucial resource utilized in all such quantum information schemes are quantum states, in which data can be encoded (Long 2025; Ramya et al. 2025). In this text we will investigate a method that permits certain properties of a single quantum state to be recycled, allowing it to be repurposed in various quantum information protocols.

In 1935, Einstein, Podolsky and Rosen (EPR) published a paper (Einstein et al. 1935) in which they argued for the incompleteness of quantum mechanics. In their efforts, they introduced the notion of *elements of reality*: a quantity that can be predicted with certainty without in any way disturbing the system. According to their reasoning, not all elements of reality have a counterpart in quantum mechanics and the description of objective reality as provided by this theory must therefore be incomplete. The thought experiment that allowed for this incorrect conclusion was later reformulated by (Bohm 1951), wherein a pair of spatially separated spin-1/2 particles prepared in an entangled state (Schrödinger 1936) serves as the centerpiece of the argument. The nature of this state makes it so that, if the spin of one particle is measured to be in one direction, the spin of other particle will, with complete certainty, be found to be in the opposite direction. Therefore, a measurement on one particle instantly affects the state of the other particle, regardless of how far apart they are. Further, according to quantum mechanics, measuring the spin of a particle along one direction will destroy information about its spin along any other direction, having complete simultaneous knowledge about the two is impossible (Heisenberg 1927). Consequently, since no signal can travel faster than the speed of light, and the direction in which the particles are measured is freely chosen, the EPR argument ruled that the spin of a particle must have a definite value prior to any measurement. That is to say, there are elements of reality which have no counterpart in quantum mechanics. For this reason, EPR deemed quantum mechanics to be incomplete and concluded that another theory, offering a complete description of objective reality, was in need (Harrigan and Spekkens 2010).

The publication of the EPR paper certainly sparked controversy, with mixed reactions from the physicists of the day. Most notably, (Bohr 1935) contended that quantum mechanics, in fact, was complete, an opinion not shared by (Schrödinger 1935) who asserted that the dominating interpretation of quantum mechanics, the so-called Copenhagen interpretation, was flawed.

The ideas of EPR was later formalized by (Bell 1964; Bell 1966), who devised a test with which the completeness of quantum mechanics could be experimentally investigated. In his paper, Bell showed that the dual assumptions of determinism and locality, meaning that measurements performed on one part of an entangled state cannot be used to communicate with the party in possession of the other part of the state, are incompatible with statistical predictions made by quantum mechanics. Mathematically, determinism is closely related with the idea of hidden variables, the hypothesized source of indeterminacy in quantum mechanics. Then, in particular, Bell provided a means, so-called Bell inequalities, by which it could be experimentally tested whether quantum mechanics could be described using hidden variables obeying the locality principle. In (Clauser et al. 1969) a Bell inequality more suitable for experimental considerations was presented, the so-called Clauser-Horne-Shimony-Holt

(CHSH) inequality. Using this inequality, it was then experimentally verified by (Freedman and Clauser 1972), and also (Aspect et al. 1982), that quantum mechanics indeed allows for the violation of Bell inequalities. Thus, one has to conclude that it is possible for matter to display non-local behavior. It should be stressed, however, that by no means do quantum mechanics permit information to be transmitted faster than the speed of light, as convincingly shown by (Ghirardi et al. 1980).

Any state violating a Bell inequality is said to be Bell non-local. In modern times, non-locality is considered an indispensable resource in many quantum information processing tasks, such as communication complexity (Buhrman et al. 2010), information theory (Cubitt et al. 2011) and quantum cryptography (Herbert 1975; Ekert 1991). Furthermore, the degree to which a Bell inequality is violated can be used to quantify how non-local a state is (Brunner et al. 2014). Therefore, studying the violations of Bell inequalities is of great importance in modern physics.

The analysis of non-locality is, however, not limited to two dimensional quantum systems. In fact, the correlations displayed by higher-dimensional quantum systems have proven to result in larger violations of local realism than possible using two qubits (Kaszlikowski et al. 2000). As such, using higher-dimensional quantum systems typically offers greater advantages in many quantum information protocols, such as higher noise tolerance and larger information capacity. One downside, however, of using higher-dimensional quantum states is that they are generally more challenging to realize experimentally than qubits (Cozzolino et al. 2019). Further, since the maximum violation of the CHSH inequality can be attained using qubits only (Cirel'son 1980), it is not a Bell inequality suitable for investigating higher-dimensional non-locality. In (Collins et al. 2002), a Bell inequality tailored for this purpose was introduced, the so-called Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality. To be precise, the CGLMP inequality constitutes a family of Bell inequalities, valid for quantum systems with an arbitrary number of dimensions.

In (Silva et al. 2015) it was shown that several independent and non-communicating observers can violate the CHSH inequality using a single two-qubit entangled state. This was accomplished using so-called unsharp measurements, which permits an observer to measure a quantum system without completely disturbing it, at the cost of less information gain (Zhou et al. 2025). Thus, measuring on a system using such measurements it is therefore possible to preserve some entanglement. It has since been demonstrated experimentally that it is possible to recycle non-locality according to this procedure (Foletto et al. 2020; Feng et al. 2020). Hence, non-locality is a resource that can be shared amongst multiple observers. Extensions of this protocol to higher-dimensional quantum systems have also been theoretically investigated by (Roy et al. 2024).

The sharing of non-local correlations is not only a scientific curiosity, it has far reaching implications for many different areas in quantum information science (Huang et al. 2023). From a resource perspective, being able to recycle a quantum state in order to perform various information processing tasks is clearly of great practical use. Further, applications of this scheme to a number of different quantum information protocols has also been demonstrated, such as quantum steering (Shenoy H. et al. 2019; Choi et al. 2020), quantum coherence (Datta and Majumdar 2018) and quantum networks (Mao et al. 2023).

The recycling of non-locality according to the protocol introduced in (Silva et al.2015) is enabled by the non-invasive nature of unsharp measurements. However, schemes based on sharp measurement, i.e. projective measurements, have also been developed by (Steffinlongo and Tavakoli 2022). By leveraging classical randomness, different projective measurement strategies can be stochastically combined in order to achieve sequential violations of the CHSH inequality. This protocol, too, has been experimentally realized by (Xiao et al. 2024), and projective measurements have therefore been proven to be sufficient for recycling non-locality. Admittedly, the double CHSH violations achievable using projective measurements are smaller than what is possible with unsharp measurements. However, performing unsharp measurements requires entangling the target system with an ancilla system, on which a projective measurement is subsequently performed. Such a procedure typically requires more complex experimental setups than what is needed for standard projective measurements (Choudhary et al. 2013; Adam and Diósi 2023). In this sense, recycling non-locality using projective measurement is not only easier in practice, it is also less resource demanding.

In (Steffinlongo and Tavakoli 2022) it was also demonstrated, via an explicit measurement strategy, that it is possible to recycle non-locality by stochastically combining projective measurements using only locally generated classical randomness. This way of recycling non-locality demands the least possible amount of resources. Indeed, locally generated classical randomness requires no quantum technology, not to mention any form of communication between the measuring parties. The investigation into this protocol by (Steffinlongo and Tavakoli 2022) is, however, quite limited, and it is not yet fully understood to what extent the resources can be minimized or what the optimal scheme is. Also, so far the scenario have not been treated in full generality.

The main objective of this thesis is to further investigate how projective measurements and local randomness can be utilized in order to share non-locality. More specifically, we will explore the performance of this method in the context of recycling a two-qubit entangled state using the CHSH inequality. Further, making use of the CGLMP inequality, we will then extend this protocol to the higher-dimensional case in which a two-qutrit state is recycled. In both cases, we find that the optimal state, leading to the largest double violation, is not maximally entangled. Moreover, the two-qutrit case leads to substantially larger double violations than does the two-qubit case.

The outline of this text is as follows. In Sec. II we provide a comprehensive overview of the physical concepts needed to fully appreciate the results. This section is further divided into several smaller sections, each devoted to a single central topic. In Sec. II.1 we briefly introduce quantum states, with particular emphasis on the density matrix formulation of quantum mechanics and higher-dimensional quantum systems. We then move on to quantum measurements, a topic that is treated in Sec. II.2. Afterwards, in Sec. II.3, the evolution of quantum states is addressed, especially, its modern formulation in terms of quantum channels is treated. Lastly, in Sec. II.5 we introduce the method of recycling non-locality using projective measurements and local randomness more formally. The results are then presented in Sec. III. This section is further divided into Secs. III.1 and III.2, wherein the results pertaining to the CHSH and CGLMP inequalities are treated separately. The text is concluded in Sec. IV where we briefly summarize our findings and discuss how the work can be further expanded.

II. THEORETICAL FRAMEWORK

II.1. *Quantum states*

Quite generally, a quantum state is a mathematical representation of a quantum system, and from it, all information about the system can be obtained. The purpose of the following sections is to give a comprehensive review of quantum states as a whole.

II.1.1. *Qubits*

A qubit represents a quantum system that can exist in two distinct states, or in any superposition of these. It is the simplest possible system one can think of in quantum mechanics. It should be emphasized, however, that a qubit does not have to represent any particular physical system, it is merely a mathematical construction that serves as a template for whatever one might study. Though, imagining the two states as the energy levels in a two-level atom, or the two spin states of a spin-1/2 particle, can sometimes be helpful.

Mathematically, a qubit can be written as

$$|\psi\rangle = a|0\rangle + b|1\rangle \quad a, b \in \mathbb{C}, \tag{1}$$

where $|a|^2 + |b|^2 = 1$ and the two states $|0\rangle = (0, 1)$ and $|1\rangle = (0, 1)$ are orthogonal basis vectors. Further, the two probability amplitudes a and b each requires two real numbers to be fully specified. However, the normalization condition together with the freedom in a global phase shift allows us to parametrize a qubit using only two real numbers. To see this, consider writing the probability amplitudes as $a = r_a e^{\mathrm{i}\alpha}$ and $b = r_b e^{\mathrm{i}\beta}$. One parameter can then be removed by simply multiplying $|\psi\rangle$ by $e^{-\mathrm{i}\alpha}$ (a global phase shift), and another by utilizing normalization, the end result being

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}e^{\mathrm{i}\phi}|1\rangle, \tag{2}$$

where we have defined $\phi = \beta - \alpha \in [0, 2\pi]$ and $\theta \in [0, \pi]$. Moreover, the two parameters θ and ϕ both allows for a simple physical interpretation, the former being related to the probability of finding the system in either $|0\rangle$ or $|1\rangle$, and the latter being the relative phase between the two orthogonal states.

II.1.2. *Density matrix*

The qubit state we saw above is an example of what is known as a *pure state*. In addition to pure states there also exist *mixed states*. From a physical standpoint, mixed states describes quantum systems that are not completely known to the observer, the information about the system is incomplete. In contrast, if everything is known about the system its state representation is said to be pure. In order to better understand the distinction between the two, let us consider the following scenario. Suppose we have a random number generator which generates the value i with probability p_i, where $\sum_i p_i = 1$. Depending on the outcome of the random number generator, a quantum system is prepared in the pure state $|\psi_i\rangle$. Since we do not know which pure state the system is in, we can, at best, describe the system by

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{3}$$

that is, as a convex combination of pure states. Any measurable quantity in nature can be represented by an *observable*: a Hermitian linear operator O whose real eigenvalues are the outcomes in experiments. Then, if a system is described by the state (3) and O is an observable property of this system, the mean value $\langle O\rangle$ can be expressed as a weighted sum of the mean values of O with respect to the ensemble $\{p_i, |\psi_i\rangle\}$, that is

$$\mathrm{Tr}\big(O\rho\big) = \sum_i p_i \langle\psi_i|O|\psi_i\rangle, \tag{4}$$

where the second equality follows from the cyclic and linear properties of the trace. Thus, the description of the system as given by (3) is well motivated from a physical standpoint since it leads to the correct statistics in experiments.

The quantity ρ is known as the *density matrix* and it represents the state of the system. In particular, if all but one of the probabilities are zero, that is, $p_i = 1$ for some i, the system is in a pure state. Conversely, if this is not the case, the system is in a mixed state. In fact, any operator ρ that is (i) Hermitian $\rho = \rho^\dagger$, (ii) positive semi-definite, and (iii) normalized $\mathrm{Tr}(\rho) = 1$ can be thought of as a state describing a quantum system (Nielsen and Chuang 2010). Moreover, the density matrix ρ generated by some ensemble $\{p_i, |\psi_i\rangle\}$, where $|\psi_i\rangle \in \mathcal{H}\ \forall\ i$, for some Hilbert space $\mathcal{H}$, is an element of the space of all positive linear operators. This space is complete in the sense of the *Hilbert Schmidt product* $\langle A, B\rangle = \mathrm{Tr}(A^\dagger B)$ and is therfore also a Hilbert space. Hence, we denote this space also by $\mathcal{H}$.

Further, we can determine the *purity* of any given state ρ by evaluating $\mathrm{Tr}(\rho^2)$. If $\mathrm{Tr}(\rho^2) = 1$ the state is pure, however, if $\mathrm{Tr}(\rho^2) < 1$ the state is mixed. If we know absolutely nothing about the system its state is said to be *maximally mixed* and its density matrix is given by $\rho = \mathbb{1}/d$, where d is the dimensionality of the system. On the other hand, if there is no uncertainty about which state the system is in, we can assign to it a pure state.

II.1.3. Bloch sphere

In order to fully characterize a qubit we only need to specify two real numbers. As a consequence, qubits can be visualized in three dimensions using a construction known as the *Bloch sphere*, originally devised by (Bloch 1946). The idea is quite simple, we associate with θ and ϕ in (2) the polar and azimuthal angles, respectively, defining the points on a unit sphere. A pure state is therefore represented by a *Bloch vector* pointing to the surface of the Bloch sphere. In this way, each point on the surface corresponds to a different pure state. Importantly, two states are orthogonal if their Bloch vectors are anti-podal. The reason for the half-angle in (2) can now also be understood. Had it not been included, states on the upper and lower hemispheres of the Bloch sphere would have been the same, differing only by a global phase.

Denoting by $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ the vector of Pauli matrices, where σ_1 is associated with the X-direction, σ_2 the Y-direction and σ_3 the Z-direction, the density matrix of a qubit can be written in a particularity simple form

$$\rho = \frac{1}{2}\big(\mathbb{1} + \vec{r}\cdot\vec{\sigma}\big), \tag{5}$$

where $\vec{r} = (r_1, r_2, r_3)$ is the Bloch vector, the components of which are given by $r_i = \mathrm{Tr}(\sigma_i \rho)$. Thus, any qubit, pure or mixed, can be expanded in the matrix basis $\{\mathbb{1}, \sigma_1, \sigma_2, \sigma_3\}$. Using the purity function $\mathrm{Tr}(\rho^2)$ one can show that a state is pure if the corresponding Bloch vector has unit length $|\vec{r}| = 1$, and mixed if its length is less than one $|\vec{r}| < 1$. As such, mixed states are represented by Bloch vectors directed towards points within the Bloch sphere. Notably, the maximally mixed state corresponds to the center of the Bloch sphere since its Bloch vector has length zero.

Moreover, one major advantage of working with qubits in the Bloch formalism is that the action of some arbitrary unitary operator U, satisfying $U^\dagger U = UU^\dagger = \mathbb{1}$, on (5) can be viewed as a rotation of the Bloch vector about some axis. Any unitary operator acting on qubits is an element of the *special unitary* group SU(2), that is, all 2×2 unitary matrices U with $\det(U) = 1$. The set of Pauli matrices $\{\sigma_1, \sigma_2, \sigma_3\}$ acts as *generators* for this group, meaning that any member of SU(2) can be written as a product of the elements in $\{\sigma_1, \sigma_2, \sigma_3\}$. Thus, an arbitrary unitary operator acting on a qubit can be written as

$$U(\vec{n}, \omega) = \exp(\mathrm{i}\frac{\omega}{2}\vec{n}\cdot\vec{\sigma}), \tag{6}$$

where $\vec{n}$ is a unit vector and ω the angle of rotation. Any unitary transformation on a qubit will result in a new state $U(\vec{n}, \omega)\rho U^\dagger(\vec{n}, \omega) = \rho'$. We can then relate the Bloch vector of ρ' to that of ρ via Rodrigues' formula (M. Barnett 2009)

$$\vec{r}' = \cos\omega\ \vec{r} + (1 - \cos\omega)\vec{n}(\vec{n}\cdot\vec{r}) + \sin\omega\ \vec{n}\times\vec{r}. \tag{7}$$

That is, rotating $\vec{r}$ about an axis $\vec{n}$ by an angle ω gives us $\vec{r}'$. To illustrate, in Fig. I we have drawn the Bloch vectors corresponding to the eigenstates of σ_1 and σ_3 with positive eigenvalues, respectively. Also depicted is the rotation of the state $|0\rangle$ into $(|0\rangle + |1\rangle)/\sqrt{2}$ by means of the unitary transformation $U = \exp(\mathrm{i}\frac{\pi}{4}\sigma_2)$.

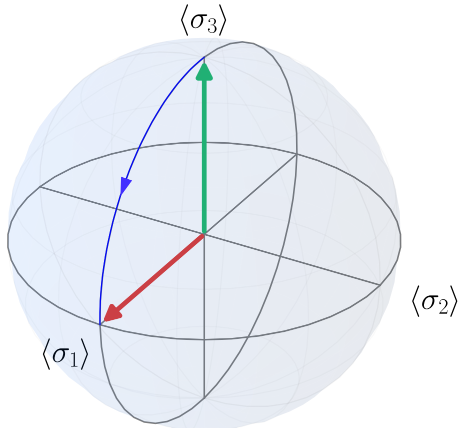


Figure I: The Bloch sphere. The red Bloch vector corresponds to the eigenstate of σ_1 with positive eigenvalue, and similarly for the green Bloch vector but instead for σ_3. We can rotate the latter into the former by means of the unitary transformation $U = \exp(\mathrm{i}\frac{\pi}{4}\sigma_2)$, as indicated by the blue arc connecting the two.

II.1.4. Higher-dimensional quantum states

Given a d-dimensional quantum system, we can represent its state, referred to as a *qudit*, using a density matrix of dimension $d \times d$. Similar to how a general qubit state can be written as a linear combination of $\{\mathbb{1}, \sigma_1, \sigma_2, \sigma_3\}$, a qudit can also be decomposed in terms of a suitable matrix basis. Though there are many different such bases (Kryszewski and Zachcial 2006), of particular interest to us is the *generalized Gell-Mann basis* (GGMB), which we can think of as an extension of the Pauli matrices to arbitrary dimensions (Bertlmann and Krammer 2008). Explicitly, the GGMB consists of $d^2 - 1$ matrices, of which there are three different kinds, namely, $d(d-1)/2$ symmetric

$$\Lambda_{\mathrm{s}}^{mn} = |m\rangle\langle n| + |n\rangle\langle m|, \quad 0 \leq m < n \leq d-1, \tag{8}$$

$d(d-1)/2$ anti-symmetric

$$\Lambda_{\mathrm{as}}^{mn} = -\mathrm{i}|m\rangle\langle n| + \mathrm{i}|n\rangle\langle m|, \quad 0 \leq m < n \leq d-1 \tag{9}$$

and $d-1$ diagonal

$$\Lambda_{\mathrm{d}}^{l} = \sqrt{\frac{2}{(l+1)(l+2)}}\Big(\sum_{m=0}^{l} |m\rangle\langle m| - (l+1)|l+1\rangle\langle l+1|\Big), \quad 0 \leq l \leq d-2. \tag{10}$$

Note that for $d = 2$ these indeed becomes the Pauli matrices. By construction, these matrices are all hermitian, traceless and mutually orthogonal. Specifically, they satisfy

$$\mathrm{Tr}(\Lambda_{\mathrm{s}}^{mn}\Lambda_{\mathrm{as}}^{op}) = \mathrm{Tr}(\Lambda_{\mathrm{s}}^{mn}\Lambda_{\mathrm{d}}^{l}) = \mathrm{Tr}(\Lambda_{\mathrm{as}}^{mn}\Lambda_{\mathrm{d}}^{l}) = 0 \tag{11}$$

as well as

$$\mathrm{Tr}(\Lambda_{\mathrm{s}}^{mn}\Lambda_{\mathrm{s}}^{op}) = \mathrm{Tr}(\Lambda_{\mathrm{as}}^{mn}\Lambda_{\mathrm{as}}^{op}) = 2\delta_{mo}\delta_{np} \tag{12}$$

and

$$\mathrm{Tr}(\Lambda_{\mathrm{d}}^{l}\Lambda_{\mathrm{d}}^{k}) = 2\delta_{lk}. \tag{13}$$

As such, together with the $d \times d$ identity operator they constitute a valid matrix basis in which we can expand any density matrix

$$\rho = \frac{1}{d}\Big(\mathbb{1} + \vec{r}\cdot\vec{\Lambda}\Big), \tag{14}$$

where $\vec{r} = (\{r_{\mathrm{s}}^{mn}\}, \{r_{\mathrm{as}}^{mn}\}, \{r_{\mathrm{d}}^{l}\})$ and the components are retrieved from $r_{\mathrm{s}}^{mn} = (d/2)\mathrm{Tr}(\rho\Lambda_{\mathrm{s}}^{mn})$, $r_{\mathrm{as}}^{mn} = (d/2)\mathrm{Tr}(\rho\Lambda_{\mathrm{as}}^{mn})$ and $r_{\mathrm{d}}^{l} = (d/2)\mathrm{Tr}(\rho\Lambda_{\mathrm{d}}^{l})$. Further, since any state must satisfy $\mathrm{Tr}(\rho^2) \leq 1$, the generalized Bloch vector $\vec{r}$ is restricted to a d^2-1 dimensional hyperball of radius $\sqrt{d(d-1)/2}$, i.e. $|\vec{r}| \leq \sqrt{d(d-1)/2}$. It should be emphasized, however, that this is not a sufficient condition for the density matrix to be positive, a valid state that is (Kimura 2003). Thus, unlike the qubit case, not all points on this Bloch hyperball corresponds to physical quantum states (Caves and Milburn 2000; Goyal et al. 2016).

Naturally, the GGMB acts as generators for the group $\mathrm{SU}(d)$, meaning that any unitary $d \times d$ matrix with determinant one can be written in the form

$$U(\vec{\varphi}) = \exp\Big(\mathrm{i}\vec{\varphi}\cdot\vec{\Lambda}\Big), \tag{15}$$

for some d^2-1 component vector $\vec{\varphi}$. In the case that $d=3$, qutrit systems that is, the GGMB becomes the so called *Gell-Mann matrices*, the matrix representation of which are given explicitly in Appendix E.

II.1.5. Bi-partite quantum states

In order to describe a system consisting of two interacting particles we need a way to relate the spaces of each separate particle with that of the larger system. To this end, say we have two particles interacting via their spin, each residing in Hilbert spaces denoted by $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Then, the joint state space of both particles is related to $\mathcal{H}_A$ and $\mathcal{H}_B$ via a tensor product structure $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ (Nielsen and Chuang 2010). The joint state of both particles ρ_{AB} is therefore an element of $\mathcal{H}_{AB}$. Although, in many situations we might only have access to one particle, i.e. a sub-system of a larger system, for which we assign the state ρ_A, the so-called *reduced state* of ρ_{AB}. We can relate ρ_A and ρ_{AB} via the *partial trace*

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \sum_n (\mathbb{1}\otimes\langle n|)\rho_{AB}(\mathbb{1}\otimes|n\rangle), \tag{16}$$

where a similar relation also holds for ρ_B. At first glance, this might seem rather arbitrary and overly complicated. It is, however, motivated by the fact that the measurement statistics retrieved from ρ_A and ρ_{AB} should be in agreement with one another. More specifically, if we have some observable in system A denoted by O_A, then, on purely physical grounds, we require that $\mathrm{Tr}(\rho_A O_A) = \mathrm{Tr}(\rho_{AB} O_A)$. The partial trace indeed ensures that this requirement is satisfied.

II.2. Measurements

In this section we introduce the mathematical description of quantum measurements. We focus on two different kinds of measurements and provide explicit examples for each in the context of qubits.

II.2.1. von Neumann measurements

The simplest type of measurements are so-called *von Neumann measurements*, also known as *projective measurements* (Holevo 2001). Given a d-dimensional quantum system, say we wish to measure some observable O having d possible outcomes labeled by $i = 0, \cdots, d-1$. In the von Neumann formalism, such a measurement is represented by a set of orthogonal projectors $\{P_i\}_{i=0}^{d-1}$, namely, operators satisfying $P_i P_j = P_i \delta_{ij}$. Then, given that the system is represented by the state ρ, the associated probability p_i of getting outcome i in an experiment is given by Born's rule $p_i = \mathrm{Tr}(P_i \rho)$. In order for the set $\{P_i\}_{i=0}^{d-1}$ to constitute a measurement, however, additional requirements must be satisfied. First, all members of the set must be non-negative $P_i \geq 0$. Second, the set must be *complete*, meaning that the members must satisfy the completeness relation $\sum_i P_i = \mathbb{1}$. These restrictions ensures that the probabilities are non-negative and sum to one. Further, if a measurement is performed on ρ, resulting in outcome i, the state is said to collapse to the projector associated with the outcome. Therefore, after observing outcome i we assign the following state to the system (Lüders 2006)

$$\rho' = \frac{P_i \rho P_i^\dagger}{\mathrm{Tr}(P_i \rho)}, \tag{17}$$

where we divide by $\mathrm{Tr}(P_i \rho)$ to make the post-measurement state normalized. Importantly, any subsequent measurement of the state will yield the same outcome.

For qubits, projective measurements takes a particularly simple form. Using the completeness relation it is possible to represent any measurement with two outcomes using a single Bloch vector. The observable relevant for qubits is $\vec{n} \cdot \vec{\sigma}$, that is, any qubit observable can be expanded in terms of the Pauli matrices. Then, the projectors associated with the two outcomes $i = 0$ and $i = 1$ are given by

$$P_i = \frac{1}{2}(\mathbb{1} + (-1)^i \vec{n} \cdot \vec{\sigma}). \tag{18}$$

Note that the two projectors P_0 and P_1 have Bloch vectors that are anti-podal. This is a consequence of the orthogonality constraint $P_i P_j = P_i \delta_{ij}$.

II.2.2. Generalized measurements

The operators constituting a measurement are not required to be projectors, since this limitation do not have a counterpart in the real world. In fact, any set of operators $\{E_n\}_n$ for which

$E_n \geq 0$ and $\sum_n E_n = \mathbb{1}$ are both satisfied is also a valid measurement and is called a *generalized measurements* or a *positive operator-valued measure* (POVM). In contrast to projective measurements, however, a POVM do not allow one to associate a unique post-measurement state to any given outcome (Brandt 1999).

To illustrate the difference between POVMs and projective measurements, consider a qubit measurement consisting of the following POVM elements

$$E_i = \frac{1}{2}(\mathbb{1} + (-1)^i \eta \vec{n} \cdot \vec{\sigma}) \quad 0 \leq \eta \leq 1, \tag{19}$$

known as an *unsharp measurement* or *weak measurement*. As the name suggests, such measurements do not fully disturb the state, rather, they only extract partial information from the state depending on the sharpness parameter η. Clearly, if $\eta = 1$ the measurement becomes projective, corresponding to maximal information gain and the largest possible disturbance to the state. On the other hand, if $\eta < 1$ the state is left with some coherence, at the cost of not gaining full information about the system.

II.3. Evolution of quantum states

We provide the necessary mathematical tools for describing the evolution of quantum states in the most general way possible. Additionally, we give some constructive examples that are relevant not only for the results but also aids the reader in their understanding of the concepts.

II.3.1. Quantum channels

The evolution of closed quantum systems is governed by the Schrödinger equation, as a consequence, closed systems evolves according to unitary transformations, ensuring that purity is preserved (Griffiths and Schroeter 2018). On the other hand, describing the evolution of open systems, i.e. the system together with its environment, as well as other operations through which a state can be transformed, for instance, state preparation, information loss, measurements etc., requires a more general mathematical construction. Collectively, all imaginable operations on states allowed by nature can be formulated in terms of *quantum channels* (Bertlmann et al. 2008). Simply put, a quantum channel maps states from one Hilbert space onto states in another Hilbert space $\rho \mapsto \Lambda[\rho]$, where $\rho \in \mathcal{H}_A$ and $\Lambda[\rho] \in \mathcal{H}_B$. However, in order to ensure that an arbitrary quantum state is mapped onto another valid state, a quantum channel has to fulfill a number of different conditions, namely, a quantum channel has to be

(i) **Linear**:

$$\Lambda\Big[\sum_i c_i \rho_i\Big] = \sum_i c_i \Lambda[\rho_i], \tag{20}$$

(ii) **Trace-preserving**:

$$\mathrm{Tr}(\Lambda[\rho]) = \mathrm{Tr}(\rho), \tag{21}$$

(iii) **Completely Positive**:

$$(\Lambda \otimes \mathbb{1})[\rho_{AB}] \geq 0 \quad \forall \rho_{AB} \geq 0. \tag{22}$$

Linearity is motivated by the fact that quantum mechanics itself is a linear theory and that the map must be well defined for both pure and mixed states. Further, the normalization of the evolved state is ensured by requiring the quantum channel to be trace-preserving. Lastly, we also have to take into consideration that the input state may be but a subsystem of a larger system, for which the output state must still remain positive. This is guaranteed by requiring the quantum channel to be completely positive. Thus, any quantum operation can be described in terms of a linear map that is completely positive and trace-preserving (CPTP).

The following theorem, due to (Kraus et al. 1983), provides an alternative, and equivalent, way in which quantum channels can be represented:

Theorem 1 (Kraus decomposition theorem)**.** *A linear map $\Lambda : \mathcal{H}_A \to \mathcal{H}_B$ is CPTP if and only if it admits a Kraus decomposition*

$$\Lambda[\rho] = \sum_{i=0}^{d-1} K_i \rho K_i^\dagger,$$

where $d \leq \dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$ and the Kraus operators K_i satisfy $\sum_{i=0}^{d-1} K_i^\dagger K_i = \mathbb{1}$.

II.3.2. The measurement and depolarization channels

To demonstrate the utility of the Kraus operator formalism, consider the act of measuring the state ρ using a POVM $\{E_i\}_i$. This process can mathematically be described using the *measurement channel*, which maps a quantum system into a classical one. It has the following mathematical form

$$\Lambda[\rho] = \sum_i \mathrm{Tr}(E_i \rho)|i\rangle\langle i|. \tag{23}$$

Classical states can be considered as a sub-class of quantum states with the distinction between the two being that the former is unable to exist in a coherent superposition. For density matrices, any coherence in the state will enter as non-diagonal entries, thus, the density matrix of a classical state will have diagonal entries only. This is indeed the case for the output state from the channel above, which takes the outcome probabilities $\mathrm{Tr}(E_i\rho)$ and organizes them into a classical state. To find the Kraus decomposition of the measurement channel we first write the elements E_i in spectral decomposition

$$E_i = \sum_j \lambda_{ij} |\phi_j^i\rangle\langle\phi_j^i|, \tag{24}$$

which when inserted into (23) leads to the following

$$\Lambda[\rho] = \sum_j \sum_i \sqrt{\lambda_{ij}}|i\rangle\langle\phi_j^i|\rho|\phi_j^i\rangle\langle i|\sqrt{\lambda_{ij}}. \tag{25}$$

Direct comparison with Thm. 1 suggests that the Kraus operators are $K_{ji} = \sqrt{\lambda_{ji}}|i\rangle\langle\phi_j^i|$, which certainly is the case since these indeed satisfy the condition $\sum_{ij} K_{ij}^\dagger K_{ij} = \mathbb{1}$ and also, by construction, implements the channel in (23).

Further, given some qubit state ρ, consider the following CPTP map

$$\rho_\nu = \nu\rho + (1-\nu)\mathrm{Tr}(\rho)\frac{\mathbb{1}}{2}, \tag{26}$$

known as the *depolarization channel*. Physically, this channel corresponds to preparing the state ρ with probability ν, the *visibility*, and the maximally mixed state $\mathbb{1}/2$ with probability $(1-\nu)$, effectively introducing white noise into the state ρ. Geometrically, the effect of this channel is to uniformly shrink the radius of the Bloch sphere, with $\nu = 0$ corresponding to the extreme case in which the entire Bloch sphere have been reduced to a single point, namely, the maximally mixed state. Furthermore, the Kraus operators corresponding to the depolarization channel are given by $K_0 = \sqrt{(1+3\nu)/4}\mathbb{1}$ and $K_i = \sqrt{(1-\nu)/4}\sigma_i$, for $i = 1, 2, 3$, as can be easily verified by direct substitution into the definition in Thm. 1 (Wilde 2013).

II.3.3. *Quantum instruments*

If we also wish to take into consideration the quantum aspects of the post-measurement state we need some other mathematical tool. This is made possible by *quantum instruments* (Davies and Lewis 1970). By definition, a quantum instrument consists of a set of completely positive maps $\{\mathcal{I}_i\}_i$ that are trace non-increasing with the property that their sum $\sum_i \mathcal{I}_i$ is a quantum channel, i.e. a CPTP map. The concept of quantum instruments allows for a more general description of measurements for which the quantum measurement channel and CPTP maps are but special cases. Given some input state ρ, the state after a measurement represented by the instrument $\{\mathcal{I}_i\}_i$ is

$$\rho \mapsto \mathcal{I}[\rho] = \sum_i |i\rangle\langle i| \otimes \mathcal{I}_i[\rho]. \tag{27}$$

We can think of $|i\rangle\langle i|$ as a classical register which records the outcome of the measurement and $\mathcal{I}_i[\rho]$ as the quantum post-measurement state conditioned on observing the outcome i. This post-measurement state is sub-normalized in the sense that $\mathcal{I}_i[\rho]$ is the post-measurement state times the probability of observing outcome i. Thus, if we take the partial trace of the second register above we will simply retrieve (23), the measurement channel that is.

A special class of quantum instruments that frequently occur in quantum information theory are so-called *Lüders instruments*. Consider a non-selective measurement, meaning that the outcome is not recorded, of the state ρ using a general measurement represented by the POVM $\{E_i\}_i$. Then, the state we assign to the quantum system after the measurement has been performed can then written as (Busch and Singh 1998)

$$\mathcal{I}_L[\rho] = \sum_i \mathcal{I}_{L,i}[\rho] = \sum_i \sqrt{E_i}\rho\sqrt{E_i}, \tag{28}$$

where the classical register is omitted since the outcome is never recorded. This instrument has the special property that $\mathrm{Tr}(\mathcal{I}_L[\rho]A) = \mathrm{Tr}(\rho A)$, if A is some observable that commutes with all E_i. This result is known as Lüders generalized theorem and it tells us that if we measure two commuting observables, then the resulting outcomes do not depend on the order in which we perform the two measurements (Łuczak and Podsdkowska 2015).

II.4. Non-classical correlations

This section is devoted to describing the correlations that arises in a typical measurement scenario featuring a bi-partite quantum state. We make the distinction between classical and non-classical correlations clear and also discuss the EPR paradox in depth. Based on this discussion we then provide a proper introduction to Bell inequalities.

II.4.1. Entanglement

Essentially, entanglement is a kind of correlation exhibited by some multipartite quantum systems. In a sense, systems that are entangled with one another can only be fully described as whole, information about the individual parts constituting a system do not allow one to characterize the system entirely. This holistic view is indeed very much different from our everyday experience and has no true classical analog.

To introduce the concept of entanglement, say we have two measuring parties, Alice and Bob, which have access to one system each with respective Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. The Hilbert space of the joint system comprised of both Alice's and Bob's respective system is then $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Any pure state which resides in $\mathcal{H}_{AB}$ can in the most general way be written as

$$|\psi\rangle_{AB} = \sum_{i=0}^{d_A-1} \sum_{j=0}^{d_B-1} c_{ij} |i\rangle_A \otimes |j\rangle_B \in \mathcal{H}_{AB}, \tag{29}$$

where $d_A = \dim(\mathcal{H}_A)$ and $d_B = \dim(\mathcal{H}_B)$, respectively, with $\dim(\mathcal{H}_A)$ not necessarily equal to $\dim(\mathcal{H}_B)$. By definition, if the coefficients c_{ij} are factorizable in the two indices i and j, i.e. $c_{ij} = a_i b_j$, the state is said to be a *product state*. In this case the state can be written in the form $|\psi\rangle_{AB} = |\phi\rangle \otimes |\varphi\rangle$, for some $|\phi\rangle \in \mathcal{H}_A$ and $|\varphi\rangle \in \mathcal{H}_B$. In contrast, if the coefficients do not factorize, the state is said to be entangled (Bertlmann et al. 2023). Note that entanglement in itself has no proper definition, it is simply defined as not being a product state.

This rather simple distinction between product and entangled states has profound consequences. To exemplify these, consider the state

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B). \tag{30}$$

If Alice measures her part of the state and receives the outcome corresponding to $|0\rangle \otimes |0\rangle$ (or $|1\rangle \otimes |1\rangle$) she immediately knows that Bob will get the same outcome. Thus, a local measurement of sub-system A permits the observer to make definitive predictions about the outcome of a measurement on sub-system B. Such correlations are not exhibited by systems in a product state, measuring on one part of a system do not allow the observer to make any predictions about a measurement performed on the other system.

For bi-partite systems there exists a theorem which is useful for determining whether a pure state is entangled or not. It is stated, without proof, as follows:

Theorem 2 (Schmidt decomposition theorem). *Given any pure bi-partite state* $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$*, we can write it as*

$$|\psi\rangle_{AB} = \sum_{i=0}^{k-1} \sqrt{\eta_i} |\phi_i\rangle_A \otimes |\varphi_i\rangle_B, \tag{31}$$

where $\{|\phi_i\rangle_A\}$ *and* $\{|\varphi_i\rangle_B\}$ *are orthonormal bases of* $\mathcal{H}_A$ *and* $\mathcal{H}_B$*, respectively, and* $\eta_i \geq 0$*, where* $\sum_i \eta_i = 1$*, are real numbers called Schmidt coefficients, of which there are at most* k*, the smallest of* $\dim(\mathcal{H}_A)$ *and* $\dim(\mathcal{H}_B)$*.*

Applications aside, this theorem has some quite interesting implications. A general bi-partite state (29) requires $\dim(\mathcal{H}_A)\dim(\mathcal{H}_A)$, potentially complex, numbers to fully specify its coefficient matrix. However, from the above theorem we see that a pure bi-partite state can be characterized using real numbers only, where we need at most the smallest of $\dim(\mathcal{H}_A)$ and $\dim(\mathcal{H}_B)$ such numbers.

Further, as mentioned above, any product state can be written on the form $|\psi\rangle_{AB} = |\phi\rangle \otimes |\varphi\rangle$. From Thm. 2 we see that this corresponds to all but a single Schmidt coefficient being zero, thus if a state has a Schmidt number k, the total number of non-zero Schmidt coefficients, larger than one, the state must be entangled. Further the state in (30) is said to be maximally entangled since its reduced states, the locally accessible states of each sub-system, are maximally mixed. This means that neither observer gets any information whatsoever about which state the system is in by performing a measurement on their respective sub-systems. In other words, there exist no local orthogonal measurement bases $\{|i\rangle_A\} \in \mathcal{H}_A$ and $\{|i\rangle_B\} \in \mathcal{H}_B$ that allow Alice and Bob to improve their probability in correctly guessing which state the system is in. This could only be accomplished by performing a measurement in some joint orthogonal basis $\{|i\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B\}$. In fact, the reduced states of any entangled state will always be mixed. In this way, when dealing with entangled systems, knowledge about the constituents of a system is not sufficient to characterize the system as a whole.

II.4.2. Entanglement of mixed states

The definitions of entangled and non-entangled states are easily extended to density matrices, allowing these concepts to be defined for mixed states as well. If a bi-partite state ρ_{AB} can be written as a convex combination of product states

$$\rho_{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B, \tag{32}$$

where $\sum_i p_i = 1$, then the state is said to be *separable*. In contrast, if the state do not admit a decomposition in the above form, then it is entangled. Further, taking mixed states into consideration, it is possible for quantum systems to be correlated without being entangled. These correlations, however, are of classical flavor and should be distinguished from those displayed by entangled systems.

With regards to correlations, states can be split into three distinct categories, namely, uncorrelated, classically correlated and quantum correlated states. Uncorrelated states represents systems that have been independently prepared and any observation of one system cannot be used to make any useful predictions about the measurement outcomes of the other system. These states are pure product states and corresponds to the case when $p_i = 1$ for some i in (32). Classically correlated states, on the other hand, we can think of as representing systems that have been prepared by some common, classical, influence. For instance, this influence may be a random number generator, the output of which determines which state Alice and Bob prepares their respective systems in. In this case, the joint state of Alice and Bob can be written in the form of (32). Note that this scenario do not necessarily require a physical interaction between the systems. The last kind of correlated states are then quantum correlated states. These are necessarily entangled and therefore require a coherent physical interaction between systems to be produced. Therefore, a quantum correlated state cannot be produced by local operations on a separable state.

Operationally, the distinction between classically and quantum correlated states is, however, a bit more nuanced than the above discussion suggests. Without going in too much detail, classically correlated states differ from the correlations arising due to entanglement in one important regard, namely, the former persists in several different measurement bases, as long as the basis is the same for both parties and the settings are unaltered. The same statement cannot be made about classically correlated states, however, which are typically only correlated in a specific local measurement basis. In the next section, we will see how this feature of quantum correlated states can be used to make predictions about the foundational nature of quantum mechanics.

The different types of correlations that can be displayed by quantum systems is well illustrated by the following example. Suppose we prepare a quantum system in $|\phi^+\rangle\langle\phi^+|$ with probability p or in $\mathbb{1}/4$ with probability $1-p$. The state describing the system is then

$$\rho_W = p|\phi^+\rangle\langle\phi^+| + (1-p)\frac{\mathbb{1}}{4}, \tag{33}$$

known as the two-qubit *Werner state* (Werner 1989), which can be shown to be entangled for $p > 1/3$ and separable for $p \leq 1/3$.

II.4.3. The EPR paradox

Consider two parties, Alice and Bob, each having access to one part of a larger system in an entangled state. For the sake of argument, assume that the shared state is the *singlet state*

$$|\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|\downarrow\rangle \otimes |\uparrow\rangle_B - |\uparrow\rangle_A \otimes |\downarrow\rangle_B), \tag{34}$$

which has total orbital angular momentum equal to zero. This means that the state is invariant with respect to rotations about an arbitrary axis. In other words, the state remains in the form of (34) regardless of the basis in which it is expressed. By virtue of the state being entangled, any measurement performed by Alice on her part of the state will inevitably affect the system of Bob. This is true even if the measurements of Alice and Bob takes place in such a way that no signal have time to reach one party before the other have had time to perform their measurement, if they are spacelike separated that is. Though, this might seem strange at first, we could easily explain this as the system always having been in the state observed by Alice, by which any measurement would then merely reveal this state. However, if we take into consideration that Alice can perform measurements in any basis of her choosing, the situation becomes a bit more peculiar. More specifically, Alice may choose to measure her part of the state in one of two bases corresponding to two non-commuting observables, for which there are no common eigenstates. For instance, if Alice chooses to measure the spin of her qubit along the axis $\vec{n}$ and find the spin of her particle to be in the positive direction, the spin of Bob's particle will inevitably be found to be in the opposite direction. Their measurements are perfectly anti-correlated. The same also holds if Alice were to measure in some other direction $\vec{n}'$ instead. Thus, Alice is effectively able to realize one of two incompatible properties of Bob's system. Consequently, the measurement performed by Alice seems to have instantaneously influenced Bob's system. It should be emphasized, however, that it is not possible to transmit information faster than light in this way. If Bob performs a measurement on his system he cannot tell if Alice already collapsed his state or if the state collapsed due to the act of his own measurement, the two are indistinguishable from his perspective (Uola et al. 2020).

A similar situation as outlined above was considered by Einstein, Podolsky and Rosen (EPR) in (Einstein et al. 1935) as a means to argue that quantum theory do not provide a complete description of objective reality. In their work, they define an *element of reality* to be a physical quantity that can be predicted with probability equal to unity without in any way disturbing the system. The incompleteness of quantum theory is then demonstrated with this definition in combination with one of the most venerable principles in physics, namely:

I **Locality**: The choice of measurement made by Alice do not affect the measurement outcome on Bob's side, and the other way around.

Via entanglement, Alice is able to establish elements of reality in Bob's system, i.e. quantities that can be predicted with certainty, without disturbing the system of Bob. However, these elements of reality are not contained in the quantum state description. Thus, quantum theory is incomplete in the sense that every element of reality do not have a counterpart in the theory. Therefore, EPR concluded that there must be some yet undiscovered theory that offers a more complete description of reality.

II.4.4. Bell inequalities

In (Bell 1964) the ideas formulated by EPR were put on more solid ground. In this paper the notion of *Bell inequalities* was also introduced, providing a means by which the foundational aspects of quantum theory could be experimentally tested. Central to the formulation of Bell inequalities is:

II **Hidden Variables**: Quantum systems are inherently deterministic, the inability to exactly predict the outcome in any single measurement is merely due to incomplete information about the system under observation.

The idea of hidden variables intends to put quantum mechanics on equal footing with theories such as classical statistical mechanics. The latter is governed by more fundamental mechanistic laws and, feasibility aside, any statistical description could be made completely deterministic if the degrees of freedom of an innumerable number of particles were known. In the same way, hidden variables tries to explain the probabilistic nature of quantum mechanics as being emergent from some more fundamental deterministic interactions beyond our control. The hypothetical source of the indeterminacy can be at the level of state preparation, unknown properties of measurement devices and so on.

Any theory in which the dual assumptions of locality and hidden variables are assumed we call a *local hidden variable theory* (LHVT). Without making any direct reference to what these hidden variables might physically correspond to, it is still possible to make experimentally testable predictions about such theories. With this intent, suppose that Alice and Bob each have access to one part of a classically correlated system, that is, a system describable by a LHVT. We collectively denote any source of indeterminacy, i.e. hidden variables, by λ. Then, the probability for Alice to observe outcome $a \in \{0,1\}$ given that she measured with setting $x \in \{0,1\}$ is denoted by $p(a|x,\lambda)$. In the same way, the probability of Bob getting outcome $b \in \{0,1\}$ using measurement setting $y \in \{0,1\}$ is $p(b|y,\lambda)$. Here we have already made the assumption of locality since $p(a|x,y,\lambda) = p(a|x,\lambda)$, and equivalently for Bob. Further, let λ be distributed according to $q(\lambda)$, where $\sum_\lambda q(\lambda) = 1$, the correlations between Alice and Bob as predicted by a LHVT can then be written on the form

$$p_L(a,b|x,y) = \sum_\lambda q(\lambda)p(a|x,\lambda)p(b|y,\lambda). \tag{35}$$

Importantly, the distributions $p(a|x,\lambda)$ and $p(b|y,\lambda)$ can always be assumed to be deterministic since all sources of indeterminacy can be absorbed into $q(\lambda)$ (Werner and Wolf 2001). On the other hand, if Alice and Bob each have access to one part of an entangled state ρ_{AB}, the correlations as predicted by quantum mechanics are

$$p_Q(a,b|x,y) = \mathrm{Tr}(A_{a|x} \otimes B_{b|y}\rho_{AB}), \tag{36}$$

where $\{A_{a|x}\}_a$ and $\{B_{b|y}\}_b$ are the measurements of Alice and Bob, respectively.

Quite generally, a Bell inequality can be written as

$$\mathcal{B} = \sum_{abxy} w_{abxy} p(a,b|x,y) \leq \mathcal{B}_L, \tag{37}$$

where w_{abxy} are real numbers and $\mathcal{B}_L$ is the so-called *local bound*. Any correlations that can be described by a LHVT, written as (35) that is, respects this inequality. Thus, if any $p(a,b|x,y)$ leads to a value of $\mathcal{B}$ larger than $\mathcal{B}_L$, these correlations cannot be explained by any LHVT.

II.4.5. The CHSH inequality

To test whether quantum mechanics is underpinned by a LHVT, consider the so-called CHSH scenario, originally devised by (Clauser et al. 1969). Without communicating with eachother, Alice and Bob are to partake in a game. Each have access to one part of a two-qubit system on which they measure. They get a positive score if the quantity $a + b + xy$ is even, or equal to zero, and a negative score otherwise, see Fig. II (a). The total score after many rounds can then be quantified by

$$\mathcal{S} = \sum_{abxy} (-1)^{a+b+xy} p(a,b|x,y). \tag{38}$$

We then compute the maximum possible score that can be achieved if the system on which they measure admits a LHVT. Inserting (35) into the above we get that

$$\begin{aligned}
\mathcal{S} &= \sum_{\lambda} \sum_{abxy} (-1)^{a+b+xy} q(\lambda) p(a|x,\lambda) p(b|y,\lambda) \\
&= \sum_{\lambda} \sum_{xy} (-1)^{xy} q(\lambda) \Big[\sum_{a} (-1)^{a} p(a|x,\lambda) \Big] \Big[\sum_{b} (-1)^{b} p(b|y,\lambda) \Big] \\
&= \sum_{\lambda} \sum_{xy} (-1)^{xy} q(\lambda) A(x,\lambda) B(y,\lambda),
\end{aligned}$$

where we introduced the quantities $A(x,\lambda) \in [-1,1]$ and $B(y,\lambda) \in [-1,1]$. If we then do the sums over x and y we arrive at

$$\mathcal{S} = \sum_{\lambda} q(\lambda) A(0,\lambda) \Big[B(0,\lambda) + B(1,\lambda) \Big] + \sum_{\lambda} q(\lambda) A(1,\lambda) \Big[B(0,\lambda) - B(1,\lambda) \Big], \tag{39}$$

whereupon using the Cauchy-Schwarz inequality separately for each term above yields

$$\mathcal{S} \leq \sum_{\lambda} q(\lambda) \Big| A(0,\lambda) \Big| \Big| B(0,\lambda) + B(1,\lambda) \Big| + \sum_{\lambda} q(\lambda) \Big| A(1,\lambda) \Big| \Big| B(0,\lambda) - B(1,\lambda) \Big|. \tag{40}$$

Clearly, the maximum occurs when $|A| = |B| = 1$, regardless of the measurement settings and λ. Thus, using the normalization of $q(\lambda)$ we then arrive at

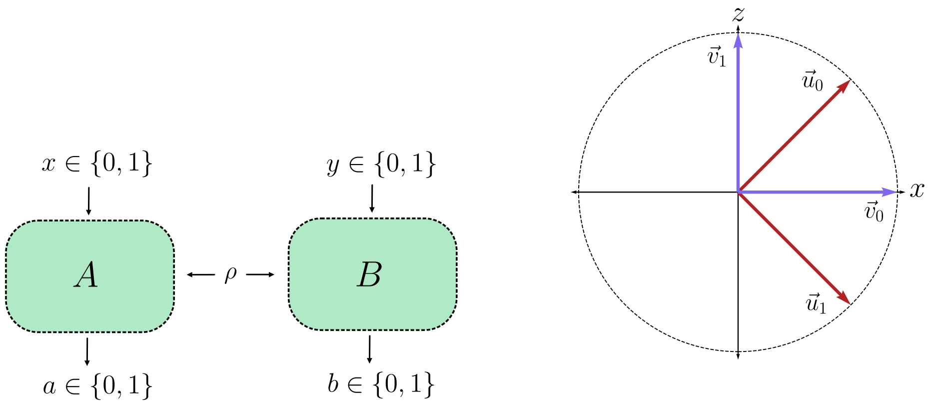
$$\mathcal{S} = \sum_{abxy} (-1)^{a+b+xy} p_L(a,b|x,y) \leq 2, \tag{41}$$

which is known as the Clauser-Horne-Shimony-Holt (CHSH) inequality.

Now, imagine that the shared system between Alice and Bob is represented by the maximally entangled state in (30). Suppose that Alice and Bob performs qubit measurements along the directions $\vec{u}_x$ and $\vec{v}_y$, respectively, which, without loss of generality, can be assumed to be limited to the XZ-plane of the Bloch sphere only. In this event, it can be shown that

$$\mathcal{S} = \vec{u}_0 \cdot (\vec{v}_0 + \vec{v}_1) + \vec{u}_1 \cdot (\vec{v}_0 - \vec{v}_1). \tag{42}$$

Then, if the shared state between Alice and Bob is maximally entangled, and Alice measures the observables $A_0 = (\sigma_1 + \sigma_3)/\sqrt{2}$ and $A_1 = (\sigma_1 - \sigma_3)/\sqrt{2}$, whereas Bob measures $B_0 = \sigma_3$ and $B_1 = \sigma_1$, we get that $\mathcal{S} = 2\sqrt{2}$, which indeed violates the CHSH inequality. The measurements that achieves this are illustrated in Fig. II (b). The violation of Bell inequalities by quantum mechanics is aptly named *Bell non-locality* and it forces us to accept that quantum mechanics is not underpinned by any LHVT and that non-local phenomena are an integral part of the world we inhabit. Further, the largest possible value of $\mathcal{S}$ achievable by quantum mechanics is in fact $2\sqrt{2}$, a result known as *Tsirelson's bound* (Cirel'son 1980).



(a) (b)

Figure II: (a) The CHSH scenario. A quantum source distributes one part of a two-qubit state ρ to two distant parties, Alice and Bob, respectively. With equal probability, Alice chooses one of two measurements, labeled by $x \in \{0, 1\}$, to perform on her part of the state and subsequently records the outcome $a \in \{0, 1\}$. Similarly, Bob also performs one of two measurements $y \in \{0, 1\}$ on his part of the state and records the outcome $b \in \{0, 1\}$. The aim is for Alice and Bob to violate the CHSH inequality (41). (b) The relative orientation of the observables of Alice, $A_0 = (\sigma_1 + \sigma_3)/\sqrt{2}$ and $A_1 = (\sigma_1 - \sigma_3)/\sqrt{2}$, and Bob, $B_0 = \sigma_3$ and $B_1 = \sigma_1$, respectively, that achieves Tsirelson's bound, i.e. the largest possible value of the CHSH parameter attainable by quantum mechanics.

The large difference between the local bound and Tsirelson's bound is indicative of the vast number of quantum states that are able to display non-local behavior. Further, entanglement is a necessary but not sufficient condition for the violation of Bell inequalities. For instance, the two-qubit Werner state in (33) only violates the CHSH inequality if $p > 1/\sqrt{2}$. Since entanglement is maintained only for $p > 1/3$, the state is entangled but incapable of violating the CHSH inequality if $1/3 < p \leq 1/\sqrt{2}$ (Li et al. 2021).

II.4.6. The CGLMP inequality

Since the maximum quantum bound of the CHSH inequality can be achieved using qubits only, it is not an appropriate Bell inequality for studying higher-dimensional non-locality. More suitable for this purpose is the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequality, which constitutes a family of Bell inequalities applicable to quantum systems of arbitrary dimension. This inequality is specifically devised for the Bell scenario in which Alice and Bob each have binary inputs, $x \in \{0, 1\}$ and $y \in \{0, 1\}$, and an arbitrary number of outputs $a \in \{0, \cdots, d-1\}$ and $b \in \{0, \cdots, d-1\}$, respectively. Optimally, the inequality is violated by a d-dimensional quantum system. Originally devised by (Collins et al. 2002), the CGLMP inequality was subsequently reformulated into a much simpler form in (Zohren and Gill 2008) which is more apt for our investigations. Using the same notation for the correlations as above, the CGLMP inequality can be written as

$$\begin{aligned}\mathcal{A}_d = &\sum_{b=0}^{d-1}\left(\sum_{a=b}^{d-1} p(a,b|x=1,y=0)\right) + \sum_{b=0}^{d-2}\left(\sum_{a=b+1}^{d-1} p(a,b|x=0,y=1)\right) \\ &+ \sum_{a=0}^{d-2}\left(\sum_{b=a+1}^{d-1} p(a,b|x=1,y=1)\right) + \sum_{a=0}^{d-2}\left(\sum_{b=a+1}^{d-1} p(a,b|x=0,y=0)\right) \geq 1.\end{aligned} \tag{43}$$

This form of the CGLMP inequality has the benefit of not depending on the measurement outcomes specifically, only the relative order of the outcomes is of significance. Further, note that the local bound is equal to unity, thus, non-locality of a d-dimensional system is certified whenever a value $\mathcal{A}_d < 1$ is measured.

In order to optimally violate this inequality, Alice and Bob can perform projective measurements $\{A_{a|x}\}_a$ and $\{B_{b|y}\}_b$, respectively, with eigenvectors on the form

$$|a\rangle_{A,x} = \frac{1}{\sqrt{d}}\sum_{l=0}^{d-1}\exp\left[\mathrm{i}\frac{2\pi}{d}l(a+\alpha_x)\right]|l\rangle_A \tag{44}$$

$$|b\rangle_{B,y} = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1}\exp\left[\mathrm{i}\frac{2\pi}{d}k(-b+\beta_y)\right]|k\rangle_B. \tag{45}$$

From a physical standpoint, these measurements are implemented in a series of different steps. Depending on which measurement they wish to perform, Alice and Bob first applies the variable phases α_x and β_y to their respective measurement bases $\{|l\rangle_A\}$ and $\{|k\rangle_B\}$. Afterwards, Alice performs a unitary in the form of a Fourier transform and similarly for Bob but with an inverse Fourier transform. It has been numerically shown that, whatever the value of d, the measurements that achieves the largest violation are $\alpha_0 = 0$, $\alpha_1 = 1/2$, $\beta_0 = 1/4$ and $\beta_1 = -1/4$ (Chen et al. 2006). However, for any $d > 2$, this inequality is not maximally violated by the maximally entangled state, as one would expect. Instead, the maximally violating states are partially entangled and changes for different d. Only in the case that $d = 2$ is the maximally violating state maximally entangled (Zohren et al. 2010).

II.5. Recycling Bell non-locality using projective measurements and local randomness

We are now in a suitable position to introduce the main topic of this thesis, namely, violating Bell inequalities sequentially using minimal recources. For this purpose, consider three spatially separated and non-communicating parties: Alice, Bob and Charlie. A quantum source distributes one part of an bi-partite entangled state ρ to Alice and Bob, respectively. On her part of the state, Alice performs one of two projective measurements labeled by $x \in \{0,1\}$, which she chooses from without bias. Resulting from these measurements are the outputs $a \in \{0,\cdots,d-1\}$. Bob also performs unbiased projective measurements, labeled by $y \in \{0,1\}$, on his part of the system, giving the output $b \in \{0,\cdots,d-1\}$. However, at his disposal he has several different measurement strategies to employ. Depending on some random variable λ, distributed according to $p(\lambda|y)$, Bob will then execute one of these measurement strategies. In particular, there are three different types of strategies Bob can apply.

First, both of Bob's measurements are non-trivial projective measurements. Second, one of his measurements are non-trivial and the other is a trivial identity projection. Third, both of his measurements are trivial identity projections. Importantly, any trivial measurement performed by Bob do not require him to interact with the system. He merely deterministically outputs $b = i$, corresponding to the measurement $\{B_{b|y}\}_b$ with $B_{i|y} = \mathbb{1}$ for some i and $B_{i'|y} = 0$ for $i' \neq i$. After Bob has performed his measurement, he subsequently sends his post-measurement state to Charlie who performs ordinary unbiased projective measurements, labeled by $z \in \{0, 1\}$, and produces the output $c \in \{0, \cdots, d-1\}$. The aim is for the observer pairs Alice-Bob and Alice-Charlie to both violate a Bell inequality, respectively. This scenario is illustrated in Fig. III.

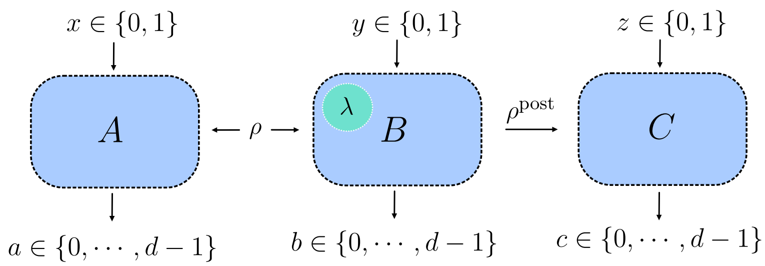


Figure III: Sequential Bell scenario using minimal resources. Two distant observers, Alice and Bob, share an entangled state ρ. Bob employs one of the measurement strategies outlined in the main text according to some rule $p(\lambda|y)$. The post measurement state is sequentially relayed to a third distant observer, Charlie. The aim is for Alice to violate a Bell inequality with both Bob and Charlie, respectively.

We denote by $\{A_{a|x}\}_a$ the measurements of Alice and by $\{B^{(\lambda)}_{b|y}\}_b$ the measurements of Bob corresponding to the strategy λ, respectively. Any general Bell parameter $\mathcal{B}$ between Alice and Bob can then be written as

$$\mathcal{B}_{AB} = \sum_{abxy} w_{abxy} \sum_{\lambda} p(\lambda|y) \mathrm{Tr}[(A_{a|x} \otimes B^{(\lambda)}_{b|y})\rho], \tag{46}$$

with w_{abxy} being real numbers. To get the corresponding expression for Alice and Charlie we first have to determine the post-measurement state resulting from Bob's measurements. To this end, let $K^{(\lambda)}_{b|y}$ be the Kraus operator representing the measurement operator $B^{(\lambda)}_{b|y} = (K^{(\lambda)}_{b|y})^\dagger K^{(\lambda)}_{b|y}$, where $\{K^{(\lambda)}_{b|y}\}_b$ is the quantum instrument implementing the measurement. Since all measurement operators are projective, the Kraus operators can be written as $K^{(\lambda)}_{b|y} = U^{(\lambda)}_{by}\sqrt{B^{(\lambda)}_{b|y}}$, where $U^{(\lambda)}_{by}$ is some unitary operation. Thus, depending on his outcome, Bob also has the freedom to perform a post-measurement unitary on the state before he sends it to Charlie. The average state shared by Alice and Charlie is then given by

$$\rho^{\text{post}} = \frac{1}{2} \sum_{\lambda} \sum_{by} p(\lambda|y) \Big(\mathbb{1} \otimes U^{(\lambda)}_{by} B^{(\lambda)}_{b|y}\Big) \rho \Big(\mathbb{1} \otimes B^{(\lambda)}_{b|y} (U^{(\lambda)}_{by})^\dagger\Big), \tag{47}$$

where we have employed Lüders instrument in (28) and used that $(B^{(\lambda)}_{b|y})^2 = B^{(\lambda)}_{b|y}$. The factor $1/2$ originates from the fact that Bob's measurements are unbiased. Also, whenever Bob does a trivial measurement, he only performs a unitary transformation on his state before he sends it to Charlie, in which case the unitary do not have to be conditioned on b. Denoting by $\{C_{c|z}\}_c$ the measurements performed by Charlie, the Bell parameter between Alice and Charlie can be expressed as

$$\mathcal{B}_{AC} = \sum_{acxz} w_{acxz} \mathrm{Tr}[(A_{a|x} \otimes C_{c|z})\rho^{\mathrm{post}}]. \tag{48}$$

To determine the optimal trade-off between the two CHSH parameters (46) and (48) we will therefore have to solve an optimization problem on the form

$$\begin{aligned} \max_{\{\rho, A_{a|x}, B_{b|y}^{(\lambda)}, U_{by}^{(\lambda)}, C_{c|z}, p(\lambda|y)\}} \quad & \mathcal{B}_{AC} \\ \text{given} \quad & \mathcal{B}_{AB} = C, \end{aligned} \tag{49}$$

where C is some fixed constant greater than or equal to the local bound of $\mathcal{B}$.

One could imagine a scenario in which Alice and Charlie also leverages local randomness. It turns out, however, that this would lead to a sub-optimal trade-off. Since their systems are never recycled, they can extract as much information from the state as possible during each round since none have to be saved for any subsequent observers. Thus, their optimal measurements are non-trivial projective measurements, which indeed will not be the case if they leverage local randomness. To see this, say Alice performs the measurement $\{A_{a|x}^{(\lambda)}\}_a$, where $A_{a|x}^{(\lambda)} A_{a'|x}^{(\lambda')} = \delta_{aa'}\delta_{\lambda\lambda'} A_{a|x}^{(\lambda)}$, with probability $p(\lambda|x)$. Then, the effective measurement she performs will be

$$A_{a|x} = \sum_{\lambda} p(\lambda|x) A_{a|x}^{(\lambda)}. \tag{50}$$

For this to be considered a projective measurement we require that $A_{a|x}A_{a'|x} = \delta_{aa'}A_{a|x}$, which, from

$$A_{a|x}A_{a'|x} = \sum_{\lambda,\lambda'} p(\lambda|x)p(\lambda'|x) A_{a|x}^{(\lambda)} A_{a'|x}^{(\lambda')} = \sum_{\lambda} [p(\lambda|x)]^2 A_{a|x}^{(\lambda)}, \tag{51}$$

is the case only if $p(\lambda|x) = 1$, if one strategy is chosen with certainty that is.

III. RESULTS

III.1. *Recycling non-locality using the CHSH inequality*

In the following we present the results related to recycling the non-locality in pairs of qubits. In Secs. III.1.1 and III.1.2 we first prove some simplifications that can be made regarding the measurements of Bob, after which we, in Sec. III.1.3, introduce the measurement protocol given that the shared state between Alice and Bob is maximally entangled. The numerical results for this case is then treated in Sec. III.1.4. Thereafter, in Sec. III.1.5, we highlight the dissimilarities in the measurement procedure when the shared state between Alice and Bob instead is partially entangled. The numerical findings in this event are then presented and quantitatively compared to the case of the maximally entangled state in Sec. III.1.6. Lastly, in Sec. III.1.7 we investigate the noise tolerance of the protocol by introducing white noise into the state.

III.1.1. *Reducing the number of measurement strategies*

Using an entangled pair of qubits, we wish to recycle non-locality via the procedure outlined in Sec. II.5. A suitable Bell inequality for this purpose is the CHSH inequality (41). In total, there are four distinct strategies that Bob can utilize. Although, two of these are of the same type, being strategies in which one measurement is trivial and the other non-trivial. More specifically, in one of these the trivial measurement is associated with setting $y = 0$ and the non-trivial with $y = 1$, whereas in the other it is the other way around. It is, however, redundant to include both of these. This is made possible by the fact that the CHSH parameter is invariant under the permutation $y \to \bar{y}$ if $x = 0$, and $y \to \bar{y}$ if $x = 1$ and $a \to \bar{a}$, where the bar denotes the bit-flip operation (Steffinlongo and Tavakoli 2022).

Furthermore, assuming the shared state between Alice and Bob to be the a pair of maximally entangled qubits $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, it is possible to prove that a strategy in which both Bobs measurements are trivial can be omitted since it do not contribute to any double violation. The proof of this is presented in Appendix A. However, there is strong numerical evidence suggesting that this strategy also can be disregarded when the shared state between Alice and Bob is partially entangled. As such, we will assume that Bob only have to consider the two strategies, labeled by λ, in Tab. I for all states.

Table I: The combination of trivial and non-trivial measurements that Bob have to consider in order to achieve a double CHSH violation. As argued in the main text, it can without loss of generality be assumed that the trivial measurement is associated with $y = 0$ in strategy $\lambda = 1$.

λ	$y = 0$	$y = 1$
0	$\{B^{(0)}_{b\|0}\}_b$	$\{B^{(0)}_{b\|1}\}_b$
1	$\{\mathbb{1}, 0\}$	$\{B^{(1)}_{b\|1}\}_b$

III.1.2. Reducing the number of measurement settings

Assuming that Bob only makes use of the measurement strategies $\lambda = 0$ and $\lambda = 1$ in Tab. I, we prove that the two non-trivial measurements corresponding to $y = 1$ are optimally the same, regardless of the state shared between Alice and Bob. To this end, assume ψ to be a pure two-qubit state, performing the sum over a, b and x in the CHSH parameter between Alice and Bob, that is, (46) using $w_{abxy} = (-1)^{a+b+xy}$, we arrive at

$$\mathcal{S}_{AB} = \sum_{\lambda}\sum_{y} p(\lambda|y) \underbrace{\left(\langle A_0, B_y^{(\lambda)}\rangle_\psi + (-1)^y \langle A_1, B_y^{(\lambda)}\rangle_\psi\right)}_{\tilde{\mathcal{S}}_{AB}(B_y^{(\lambda)})} = \sum_{\lambda}\sum_{y} p(\lambda|y) \tilde{\mathcal{S}}_{AB}(B_y^{(\lambda)}), \tag{52}$$

where $A_x = A_{0|x} - A_{1|x}$ and $B_y^{(\lambda)} = B_{0|y}^{(\lambda)} - B_{1|y}^{(\lambda)}$ are the observables of Alice and Bob, respectively, and $\langle A_x, B_y^{(\lambda)}\rangle_\psi$ is the joint expectation value of these given that the shared state is ψ.

We define $\sigma_{a|x} = \mathrm{Tr}_A[(A_{a|x} \otimes \mathbb{1})\psi]$ to be the sub-normalized state remotely prepared by Alice on Bob's side, meaning that $\mathrm{Tr}(\sigma_{a|x}) = p(a|x)$. With the shared state between Alice and Charlie being (47), the correlations between the two can then be written as

$$\begin{aligned} p(a,c|x,z) &= \frac{1}{2}\sum_{\lambda}\sum_{by} p(\lambda|y) \mathrm{Tr}\left[K_{b|y}^{(\lambda)} \sigma_{a|x} (K_{b|y}^{(\lambda)})^\dagger C_{c|z}\right] \\ &= \frac{1}{2}\sum_{\lambda}\sum_{y} p(\lambda|y) \mathrm{Tr}\left[\Gamma_y^{(\lambda)}[\sigma_{a|x}] C_{c|z}\right], \end{aligned} \tag{53}$$

where in the second line we defined the CPTP map $\Gamma_y^{(\lambda)}[\sigma_{a|x}] = \sum_b K_{b|y}^{(\lambda)} \sigma_{a|x} (K_{b|y}^{(\lambda)})^\dagger$ corresponding to the quantum instrument $\{K_{b|y}^{(\lambda)}\}_b$ employed by Bob. Inserting this into the CHSH parameter between Alice and Charlie (48) and summing over all the outcomes, introducing the notation $\sigma_x = \sum_a (-1)^a \sigma_{a|x}$, we end up with

$$\mathcal{S}_{AC} = \sum_{\lambda}\sum_{y} p(\lambda|y) \underbrace{\frac{1}{2}\sum_{xz} (-1)^{xz} \mathrm{Tr}\left[\Gamma_y^{(\lambda)}[\sigma_x] C_z\right]}_{\tilde{\mathcal{S}}_{AC}(\Gamma_y^{(\lambda)})} = \sum_{\lambda}\sum_{y} p(\lambda|y) \tilde{\mathcal{S}}_{AC}(\Gamma_y^{(\lambda)}). \tag{54}$$

We now expand both $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ in terms of the functionals $\tilde{\mathcal{S}}_{AB}$ and $\tilde{\mathcal{S}}_{AC}$, respectively. Using that $B_0^{(1)} = \mathbb{1}$ we arrive at

$$\mathcal{S}_{AB} = p(0|0)\tilde{\mathcal{S}}_{AB}(B_0^{(0)}) + p(1|0)\tilde{\mathcal{S}}_{AB}(\mathbb{1}) + p(0|1)\tilde{\mathcal{S}}_{AB}(B_1^{(0)}) + p(1|1)\tilde{\mathcal{S}}_{AB}(B_1^{(1)}), \tag{55}$$

and similarly

$$\mathcal{S}_{AC} = p(0|0)\tilde{\mathcal{S}}_{AC}(\Gamma_0^{(0)}) + p(1|0)\tilde{\mathcal{S}}_{AC}(\mathbb{1}) + p(0|1)\tilde{\mathcal{S}}_{AC}(\Gamma_1^{(0)}) + p(1|1)\tilde{\mathcal{S}}_{AC}(\Gamma_1^{(1)}), \tag{56}$$

where the notation $\tilde{\mathcal{S}}_{AC}(\mathbb{1})$ should be understood as Bob not using any quantum instrument. Note that the functional forms of $\tilde{\mathcal{S}}_{AB}(B_1^{(0)})$ and $\tilde{\mathcal{S}}_{AB}(B_1^{(1)})$ are the same, the only difference being that they are parametrized in terms of $B_1^{(0)}$ and $B_1^{(1)}$, respectively. Since $\tilde{\mathcal{S}}_{AB}(B_1^{(0)})$ and $\tilde{\mathcal{S}}_{AB}(B_1^{(1)})$ both enter into $\mathcal{S}_{AB}$ as a convex combination, maximizing $\mathcal{S}_{AB}$ separately over $B_1^{(0)}$ and $B_1^{(1)}$ it then follows that the optimal $B_1^{(0)}$ must be the same as the optimal $B_1^{(1)}$. That is,

$$\max_{\{B_1^{(0)},B_1^{(1)}\}} \mathcal{S}_{AB} \Rightarrow B_1^{(0)} = B_1^{(1)}. \tag{57}$$

By the same argument, selecting quantum instruments according to $\Gamma_1^{(0)} = \Gamma_1^{(1)}$ must also be optimal for $\mathcal{S}_{AC}$. Since this choice of Bob's measurements is optimal for both $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ separately, this must also be optimal for their trade-off. Moreover, the result that the two $y = 1$ measurement are optimally the same for both strategies further leads to the condition $p(\lambda|y) = p(\lambda)$. As such, the two CHSH parameters (55) and (56) are optimally be expressed as

$$\mathcal{S}_{AB} = p(0)\tilde{\mathcal{S}}_{AB}(B_0^{(0)}) + p(1)\tilde{\mathcal{S}}_{AB}(\mathbb{1}) + \tilde{\mathcal{S}}_{AB}(B_1^{(0)}), \tag{58}$$

and

$$\mathcal{S}_{AC} = p(0)\tilde{\mathcal{S}}_{AC}(\Gamma_0^{(0)}) + p(1)\tilde{\mathcal{S}}_{AC}(\mathbb{1}) + \tilde{\mathcal{S}}_{AC}(\Gamma_1^{(0)}). \tag{59}$$

This result is perhaps not so surprising. After all, the CHSH scenario is devised under the assumption that each measuring party only has two measurement settings at their disposal. If Bob had employed three different non-trivial measurements, Alice and Charlie would have to optimize their measurements with respect to one additional measurement setting of Bob, which is clearly disadvantageous.

Based on the above considerations, it therefore suffices to treat the simplified scenario in which Bob only makes use of two distinct measurement strategies, where the two measurements labeled by $y = 1$ are the same. Moreover, the distribution determining the frequency with which these strategies are employed do not have to be conditioned on the measurement setting of Bob.

III.1.3. Measurement protocol using maximally entangled qubits

Let the initial shared state between Alice and Bob be the maximally entangled state $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. We denote the observables of Alice by $A_x = \vec{u}_x \cdot \vec{\sigma}$, where $\vec{u}_x$ is a Bloch vector limited to the XZ-plane of the Bloch sphere. Since any local rotation performed by Alice on her part of the system can be absorbed into Bob's measurements using the action at a distance principle $R \otimes \mathbb{1}|\phi^+\rangle = \mathbb{1} \otimes R^T|\phi^+\rangle$, which holds for any linear operator R, we can parametrize Alice's Bloch vectors in terms of a single angle θ as $\vec{u}_x = (\cos\theta, 0, (-1)^x \sin\theta)$. The measurements $\{A_{a|x}\}_a$ of Alice are therefore rank-1 projective qubit measurements given by

$$A_{a|x} = \frac{1}{2}(\mathbb{1} + (-1)^a \vec{u}_x \cdot \vec{\sigma}). \tag{60}$$

For any pair of λ and y, Bob chooses his measurement strategy according to Tab. I. Since the state remotely prepared by Alice on Bob's side resides in the XZ-plane of the Bloch sphere, it is optimal for Bob's measurements to also be limited to this plane. Thus, the non-trivial observables of Bob can be written as $B_y^{(\lambda)} = \vec{v}_y^{(\lambda)} \cdot \vec{\sigma}$, with $\vec{v}_y^{(\lambda)} = (\cos\beta_y^{(\lambda)}, 0, \sin\beta_y^{(\lambda)})$, and the corresponding measurements $\{B^{(\lambda)}_{b|y}\}_b$ are

$$B^{(\lambda)}_{b|y} = \frac{1}{2}(\mathbb{1} + (-1)^b \vec{v}^{(\lambda)}_y \cdot \vec{\sigma}). \tag{61}$$

Furthermore, when Bob performs a trivial measurement he deterministically outputs $b = 0$. After his measurement, Bob performs a local unitary transformation of the state. Since we wish to remain in the same sub-space of the Bloch sphere, in which case the measurement outcomes of the parties are most strongly correlated, this unitary is optimally chosen to be a rotation about an axis orthogonal to the plane in which the parties measures. Thus, the post-measurement unitary transformations are taken to be rotations about the Y-axis in the Bloch sphere and are therefore given by

$$U^{(\lambda)}_{by} = \exp\Big(\mathrm{i}\frac{\mu^{(\lambda)}_{by}}{2}\sigma_2\Big). \tag{62}$$

Further, note that we have a global phase freedom in the unitaries, which allows us to set one of the post-measurement unitaries to be trivial. Without loss of generality, we therefore set the unitary performed after the trivial measurement $\{\mathbb{1}, 0\}$ to be $U^{(0)}_{00} = \mathbb{1}$. Lastly, Charlie measures the observables $C_z = \vec{w}_z \cdot \vec{\sigma}$, with corresponding measurements $\{C_{c|z}\}_c$ given by

$$C_{c|z} = \frac{1}{2}(\mathbb{1} + (-1)^c \vec{w}_z \cdot \vec{\sigma}), \tag{63}$$

where $\vec{w}_z = (\cos\gamma_z, 0, \sin\gamma_z)$.

Using (58) together with the above measurements we straightforwardly compute the CHSH parameter between Alice and Bob to be

$$\mathcal{S}_{AB} = 2p(0)\cos\theta\cos\beta^{(0)}_0 + 2\sin\theta\sin\beta^{(1)}_1, \tag{64}$$

where we have used that $B^{(0)}_{b|1} = B^{(1)}_{b|1}$. In the same way, the CHSH parameter between Alice and Charlie (59) can be parametrized as follows

$$\begin{aligned}
\mathcal{S}_{AC} = &\Big(\cos\theta\cos\beta^{(0)}_0\cos(\beta^{(0)}_0 - \gamma_0 + \mu^{(0)}_0) + \sin\theta\sin\beta^{(0)}_0\cos(\beta^{(0)}_0 - \gamma_1 + \mu^{(0)}_0)\Big)p(0) \\
&+ \Big(\cos\theta\cos\gamma_0 + \sin\theta\sin\gamma_1\Big)(1 - p(0)) \\
&+ \Big(\cos\theta\cos\beta^{(1)}_1\cos(\beta^{(1)}_1 - \gamma_0 + \mu^{(1)}_1) + \sin\theta\sin\beta^{(1)}_1\cos(\beta^{(1)}_1 - \gamma_1 + \mu^{(1)}_1)\Big).
\end{aligned} \tag{65}$$

Notice that the post-measurement unitaries do not depend on the outcome b, that is $\mu^{(\lambda)}_{by} = \mu^{(\lambda)}_y$. This follows straightforwardly when computing $\mathcal{S}_{AC}$, see Appendix B for a detailed derivation of (64) and (65). Though, we can understand this is a consequence of the initial shared state between Alice and Bob being maximally entangled, from which it follows that their reduced states have no local structure. This means that Bob gets no information whatsoever by performing a measurement on his local system, in other words, the distribution of his outcomes is completely uniform $p(b|y) = 1/2$. Consequently, there is no benefit in choosing the post-measurement unitaries based on the outcomes b, since no information can be exploited in order to rationalize such a choice.

III.1.4. Optimal double violation using maximally entangled qubits

To determine the optimal value of $\mathcal{S}_{AC}$ given some value of $\mathcal{S}_{AB} > 2$ we numerically solved the optimization problem given in (49) using (64) and (65). This we did using the Optimization Toolbox in Matlab, where we divided $\mathcal{S}_{AB}$ into a meshline consisting of 20 datapoints in the interval $\mathcal{S}_{AB} \in [2, 2.1]$. We then performed the optimization about 20-25 times for each single datapoint in order to ensure that the optimal value indeed was established. The largest $\mathcal{S}_{AC}$ for a given $\mathcal{S}_{AB}$ found through this iterative procedure is shown in Fig. IV below, where we present the optimal trade-off between $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ in the region in which both CHSH parameters surpasses the local bound $(\mathcal{S}_{AB}, \mathcal{S}_{AC}) > (2, 2)$.

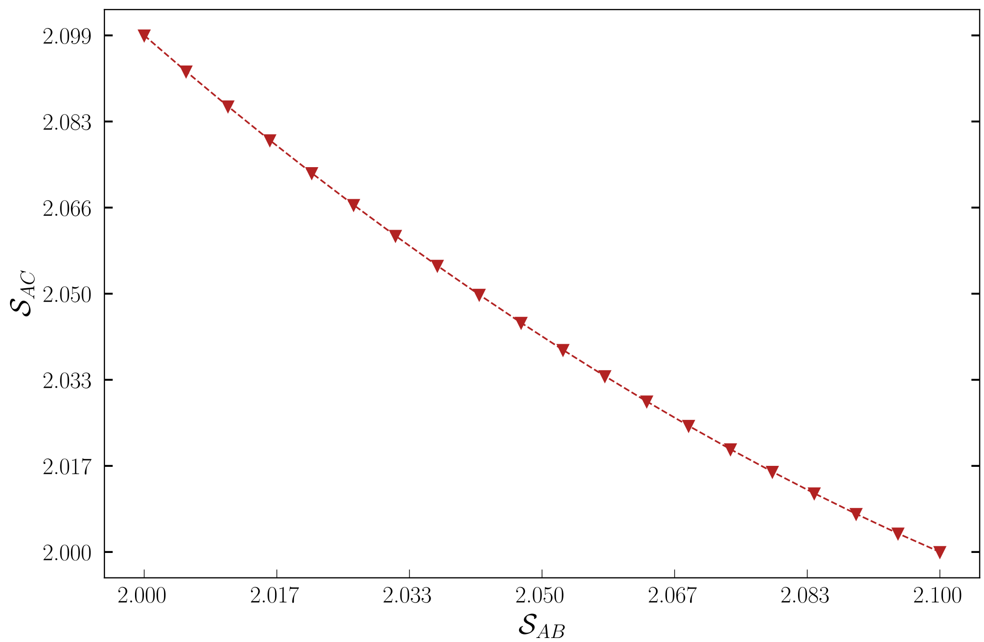


Figure IV: The optimal trade-off between two CHSH parameters $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$, in the double violation region $(\mathcal{S}_{AB}, \mathcal{S}_{AC}) > (2, 2)$, using only projective measurements and local randomness. The initial shared state between Alice and Bob is the two-qubit maximally entangled state. To achieve this, Bob need only consider two strategies, one in which both measurements are non-trivial projections and the other consisting of one trivial identity projection and one non-trivial projection.

Furthermore, we numerically found that the post-measurement unitaries are well approximated by the relation

$$\mu_y^{(\lambda)} = -\frac{e}{1+e}\Big(\frac{\gamma_0 + \gamma_1}{2} + \beta_y^{(\lambda)}\Big), \tag{66}$$

leading to deviations from the result shown in Fig. IV on the order of $10^{-4} - 10^{-5}$. Moreover, in Appendix D we present an explicit measurement strategy resulting in three sequential

CHSH violations using the maximally entangled state. The measurements presented therein yields the triple violation $(\mathcal{S}_{AB}, \mathcal{S}_{AC}, \mathcal{S}_{AD}) \approx (2.0001, 2.0001, 2.0002)$.

III.1.5. Measurement protocol using partially entangled qubits

We now wish to retrieve the optimal trade-off between $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ given that the initial shared state between Alice and Bob is the following partially entangled two-qubit state

$$|\psi_\phi\rangle = \cos\phi|00\rangle + \sin\phi|11\rangle, \tag{67}$$

where the restriction $\phi \in [0, \pi/4]$ ensures that the degree of entanglement of $|\psi_\phi\rangle$ increases monotonically with ϕ. Since the reduced states of Alice and Bob now features some local structure, for which a local rotation of Alice's system cannot be absorbed into the measurements of Bob, Alice's measurements can no longer be defined using a single parameter θ. Thus, Alice now performs qubit measurements represented by the Bloch vectors $\vec{u}_x = (\cos\alpha_x, 0, \sin\alpha_x)$, for $x \in \{0, 1\}$.

Furthermore, the post-measurement unitaries performed by Bob are found to optimally depend on the outcome of his measurements, i.e. $U_{0y}^{(\lambda)} \neq U_{1y}^{(\lambda)}$. In this scenario, the measurement outcomes of Bob are no longer equally distributed, that is $p(b|y) \neq 1/2$, and it is therefore advantageous for Bob to choose the post-measurement unitaries based on the outcome b. In Appendix C we present the explicit form of $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ when the initial state is partially entangled, together with the algebraic computation leading up to this result. From this derivation the condition $U_{0y}^{(\lambda)} \neq U_{1y}^{(\lambda)}$ also naturally emerge.

III.1.6. Optimal double violation using partially entangled qubits

Below in Fig. V we present the result of numerically solving the optimization problem in (49) using the state (67). This we did in two different ways. First we performed the optimization for a selection of fixed partially entangled states, corresponding to $\phi = \pi/5$ and $\phi = \pi/6$ in (67). Second, we left ϕ unspecified and included it as a variable to be optimized over. Either way, the trade-off between $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ is notably enhanced in comparison to the case when $\phi = \pi/4$, for the maximally entangled state that is. Clearly, the best trade-off is achieved when optimizing over ϕ. However, as $\mathcal{S}_{AB}$ increases, the curves corresponding to $\phi = \pi/5$ and the optimal ϕ begins to converge.

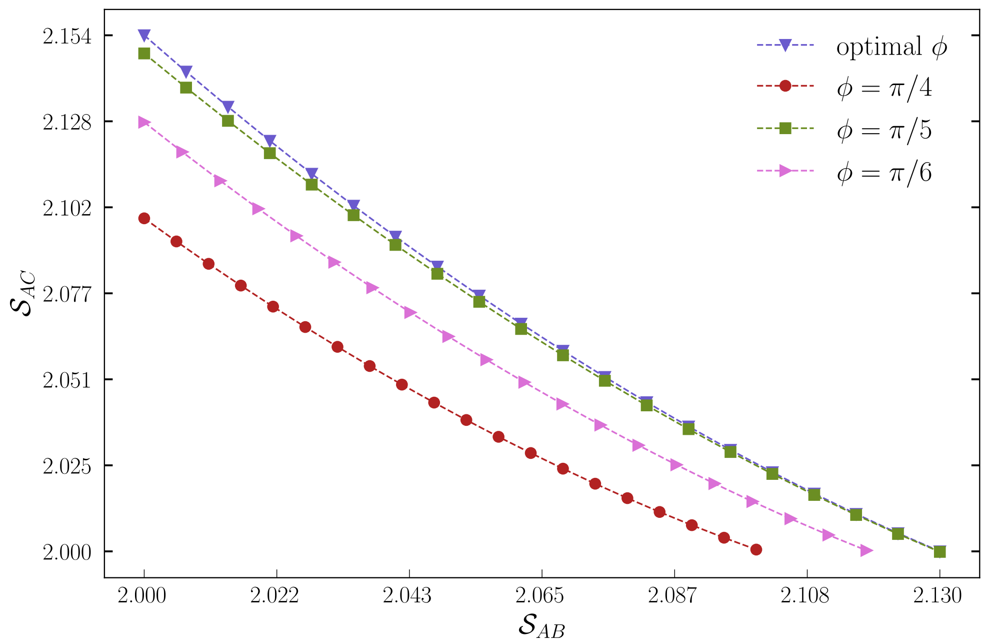


Figure V: Optimal trade-off between $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ in the double violation region $(\mathcal{S}_{AB}, \mathcal{S}_{AC}) > (2, 2)$ using the partially entangled state $|\psi_\phi\rangle = \cos\phi|00\rangle + \sin\phi|11\rangle$. Shown are the optimal trade-off curves resulting from performing the optimization using fixed $\phi \in \{\pi/4, \pi/5, \pi/6\}$ and including ϕ in the optimization procedure, respectively. In order to achieve this, Bob need only consider the same two measurement strategies as for the case in which the maximally entangled state is used.

It is instructive to compare the results shown in Fig. V with what is possible using other approaches. Using the maximally entangled state, a double violation with equal CHSH parameters is attained around $\mathcal{S}_{AB} = \mathcal{S}_{AC} \approx 2.046$. If we instead use the optimal state this value is somewhat increased to $\mathcal{S}_{AB} = \mathcal{S}_{AC} \approx 2.064$. In contrast, using projective measurements and shared global randomness, for which the measurement strategies of all observers can be collectively coordinated, it is possible to achieve even larger double violations $\mathcal{S}_{AB} = \mathcal{S}_{AC} \approx 2.108$ (Steffinlongo and Tavakoli 2022). The largest double violations, however, are made possible using weak measurements $\mathcal{S}_{AB} = \mathcal{S}_{AC} \approx 2.263$ (Silva et al. 2015).

Given that the maximally violating state in the usual CHSH scenario is maximally entangled, it is quite curious that the optimal state is found to be partially entangled when recycling non-locality using projective measurements. Why this is we can, in part, understand from Fig. VI, where we have plotted the frequency $p(0)$ with which Bob employs the measurement strategy $\lambda = 0$. This we have done both for the optimal state and for the maximally entangled state. Since the strategy $\lambda = 0$ is more invasive, in the sense that both measurements are non-trivial, the post-measurement state shared between Alice and Charlie must, on average, be more separable if this strategy is employed more frequently. Therefore, we can think of $p(0)$ as measure of the degree of disturbance induced to the state by the measurements of Bob. As such, smaller $p(0)$ must lead to larger attainable values of $\mathcal{S}_{AC}$, or, equivalently, smaller $\mathcal{S}_{AB}$. This behavior is clearly seen in both curves shown in Fig. VI. However, as can also be seen in Fig. VI, when the optimal state is used, Bob need not employ the strategy $\lambda = 0$ as frequently as for the case in which the maximally entangled state is used, if the same value of $\mathcal{S}_{AB}$ is to be achieved.

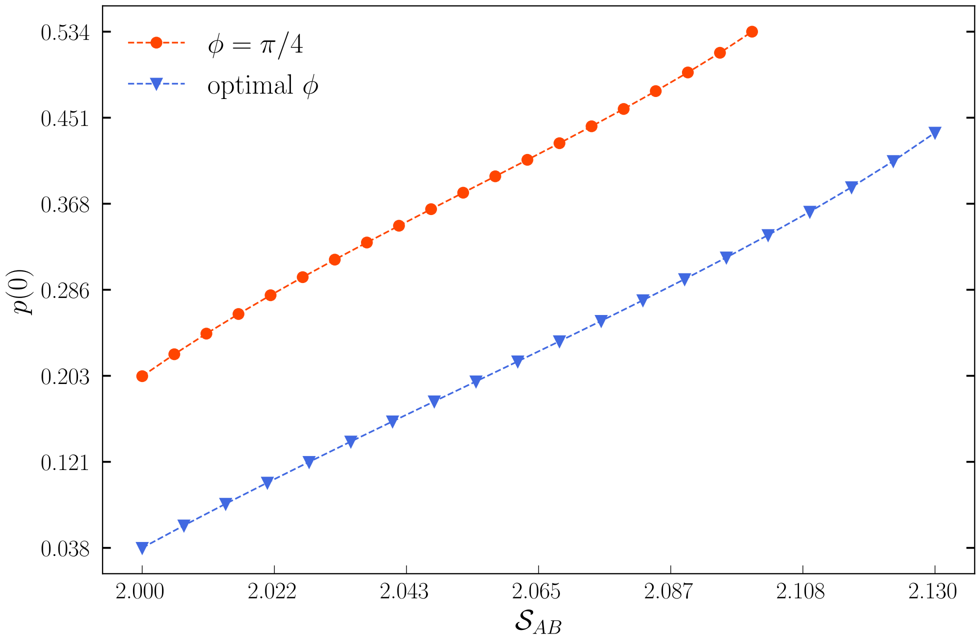


Figure VI: The probability $p(0)$ that Bob employs the measurement strategy $\lambda = 0$, i.e. the strategy in which both his measurements are non-trivial, when the optimal ϕ in the state shared between Alice and Bob $|\psi_\phi\rangle = \cos\phi|00\rangle + \sin\phi|11\rangle$ is used and for fixed $\phi = \pi/4$.

To fully appreciate why the optimal state is partially entangled we must therefore understand why Bob is able to employ $\lambda = 0$ less frequently and still achieve the same violation of $\mathcal{S}_{AB}$ with Alice, as for the maximally entangled state, when they share partially entangled qubits. An argument for why this is goes as follows. When Alice and Bob shares a partially entangled state, their respective reduced states features some local structure. Since no entanglement is required to achieve the local CHSH bound, they can leverage this locally accessible information in order to get closer to this bound without having to expend any entanglement. Thus, Bob no longer have to perform destructive non-trivial measurements as frequently. This is also evident from the explicit forms of $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$, as presented in Appendix C, since both of these features terms corresponding to the local states of Alice and Bob. On the other hand, when the shared state between Alice and Bob is maximally entangled, they have to utilize entanglement only in order to violate the CHSH inequality. In this case, there will not be as much entanglement left for Alice and Charlie to utilize in their violation as there would be for the partially entangled state.

In Fig. VII, we plot the value of the optimal ϕ against $\mathcal{S}_{AB}$. Evidently, the optimal ϕ decreases monotonically as $\mathcal{S}_{AB}$ increases. Since the degree of entanglement of $|\psi_\phi\rangle$ monotonically increases with ϕ, we conclude that the optimal state becomes less entangled as the violation of $\mathcal{S}_{AB}$ becomes larger. Given that the violation of $\mathcal{S}_{AC}$ becomes smaller as $\mathcal{S}_{AB}$ increases, the amount of entanglement required to achieve this violations must decrease together with $\mathcal{S}_{AC}$. The behavior displayed by ϕ in Fig. VII is therefore to be expected.

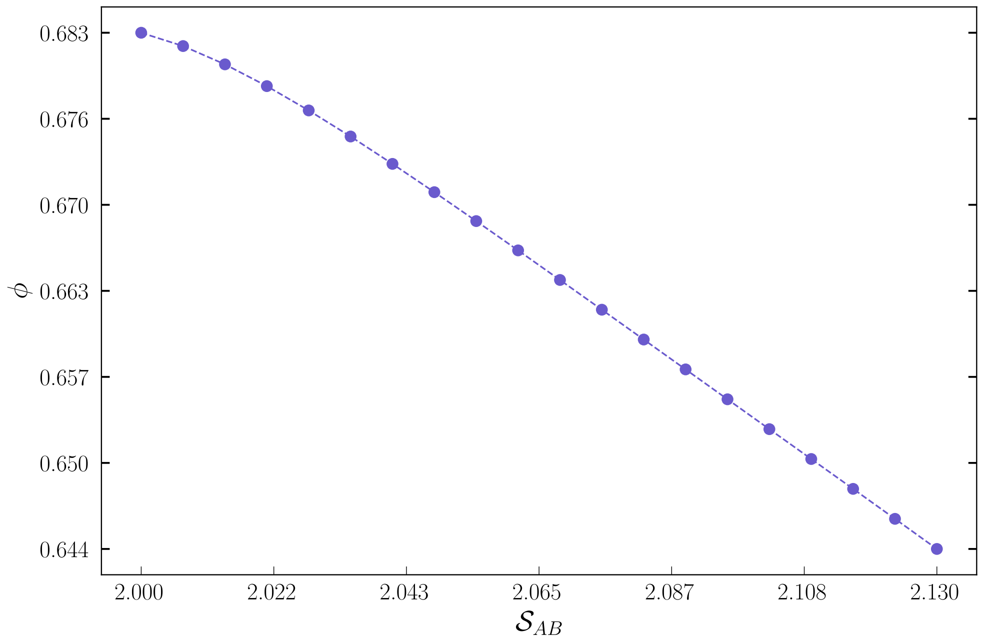


Figure VII: The optimal ϕ in the shared state between Alice and Bob $|\psi_\phi\rangle = \cos\phi|00\rangle + \sin\phi|11\rangle$ against $\mathcal{S}_{AB}$. This choice of ϕ given $\mathcal{S}_{AB} > 2$ leads to the largest possible violation of $\mathcal{S}_{AC}$ using projective measurements and local randomness.

III.1.7. Robustness to noise

Given that every observer measures optimally, we can appreciate the noise tolerance of the protocol by introducing white noise into the shared state by means of the depolarization channel (26). This we do by computing the smallest possible visibility $\nu^{\min}$ which still leads to a double violation. To this end, consider the CHSH parameter resulting from the initial state $\rho_\nu = \nu\rho + (1-\nu)\mathbb{1}/4$

$$\mathcal{S}(\rho_\nu) = \nu\mathcal{S}(\rho) + (1-\nu)\mathcal{S}(\mathbb{1}/4). \tag{68}$$

We can then simply set $\mathcal{S}(\rho_\nu) = 2$ and solve for ν to get $\nu^{\min}$, and since the CHSH parameters vanish for the maximally mixed state $\mathcal{S}(\mathbb{1}/4) = 0$ we get the simple relation

$$\nu^{\min} = \frac{2}{\mathcal{S}(\rho)}, \tag{69}$$

where $\mathcal{S}(\rho)$ is the CHSH value resulting from the optimal measurements. In Fig. VIII, we show the minimum visibility $\nu_{AB}^{\min}$ in the shared state between Alice and Bob against that of state shared between Alice and Charlie $\nu_{AC}^{\min}$. Furthermore, for the optimal state, $\mathcal{S}_{AC}$ is more tolerant to noise than $\mathcal{S}_{AB}$ in the sense that when $\nu_{AB}^{\min} = 1$ we find $\nu_{AC}^{\min} \approx 0.929$ whereas if $\nu_{AC}^{\min} = 1$ we instead have $\nu_{AB}^{\min} \approx 0.939$. Thus, with regards to any experimental considerations, it would be advantageous to measure a double violation for which $\mathcal{S}_{AC} > \mathcal{S}_{AB}$ rather than $\mathcal{S}_{AC} < \mathcal{S}_{AB}$.

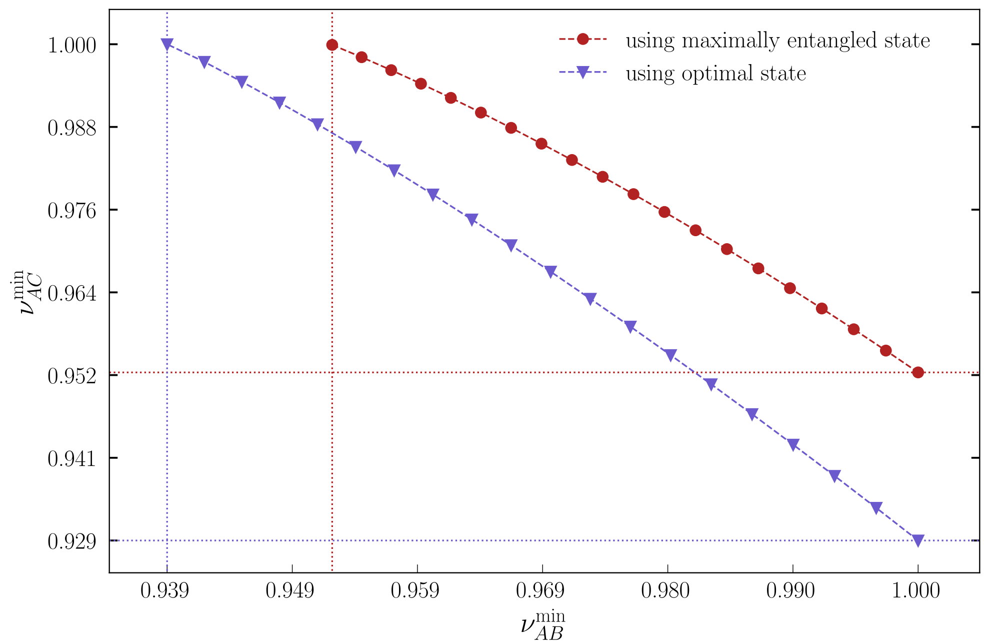


Figure VIII: The minimum visibility in the shared states between Alice and Bob against that of the shared state between Alice and Charlie. The curves traced by the points $(\nu_{AB}^{\min}, \nu_{AC}^{\min})$ corresponds to the critical noise for which both $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ reaches the local bound. All points above each respective curve leads to double violations whereas all points below do not.

To put things into perspective, using the maximally entangled state we have that $\mathcal{S}_{AB} = \mathcal{S}_{AC} \approx 2.046$. In order to measure this violation experimentally the setup would need to have a visibility around $\nu^{\min} \approx 0.978$. In contrast, measuring the violation $\mathcal{S}_{AB} = \mathcal{S}_{AC} \approx 2.064$ using the optimal state requires a visibility $\nu^{\min} \approx 0.969$. Considering the simplicity in the setup required to realize the quantum instrument used by Bob this is within current experimental capabilities. Further, producing non-maximally entangled states, in photonic experiments, do not require a more sophisticated setup than what is needed to produce maximally entangled states (White et al. 1999). Therefore, working with the optimal state in favor of the maximally entangled state do not come at the expense of additional resources.

III.2. Recycling non-locality using the CGLMP inequality

We present the results concerning the higher dimensional extension of the protocol in which the shared state between the observers is a pair of entangled qutrits. In Sec. III.2.1 we introduce the measurement protocol employed by the measuring parties with the aim of achieving two sequential CGLMP violations. It is then analytically proven in Sec. III.2.2 that the optimal post-measurement unitaries in this case also are outcome independent when the shared state between Alice and Bob is maximally entangled. The numerical results for both the maximally and partially entangled states are then presented in Sec. III.2.3. In Sec. III.2.4 we also conduct a quantitative comparison with the sequential protocol utilizing unsharp measurements.

III.2.1. Measurement protocol

Using the CGLMP inequality, we now wish to extend the protocol to three-dimensional systems, i.e. to the situation in which the shared state between Alice and Bob consists of two entangled qutrits. For this purpose, let Alice and Bob share a general pure two-qutrit state $|\psi\rangle$. On her part of the state, Alice will perform projective measurements according to the procedure outlined at the end of Sec. II.4.6. That is, based on her binary measurement choices $x \in \{0,1\}$ she applies a variable phase α_x to her measurement basis, on which she subsequently performs a Fourier transform. The resulting measurement operators $A_{a|x}$, where $a \in \{0,1,2\}$, will then have eigenvectors $|a\rangle_{A,x}$ given by (44). Using that $A_{a|x} = |a\rangle\langle a|_{A,x}$ together with the explicit form of the symmetric and anti-symmetric Gell-mann matrices, (8) and (9) with $d = 3$, respectively, we can express the measurement operators of Alice as

$$A_{a|x} = \frac{1}{3}\Big(\mathbb{1} + \sum_{i<j}^{2} \cos[(j-i)\phi^{A}_{ax}]\Lambda^{ij}_{\mathrm{s}} + \sum_{i<j}^{2} \sin[(j-i)\phi^{A}_{ax}]\Lambda^{ij}_{\mathrm{as}}\Big), \tag{70}$$

with $\phi^{A}_{ax} = 2\pi(a+\alpha_x)/3$. Further, based on numerical investigations, Bob only have to consider the same two strategies as in the qubit case. Then, whenever Bob performs a non-trivial measurement, he will apply the variable phase $\beta^{(\lambda)}_{y}$ to his measurement basis and thereafter perform an inverse Fourier transform. Thus, Bob's non-trivial measurements $\{B^{(\lambda)}_{b|y}\}_b$, with outcomes $b \in \{0,1,2\}$, consists of operators on the form

$$B^{(\lambda)}_{b|y} = \frac{1}{3}\Big(\mathbb{1} + \sum_{i<j}^{2} \cos[(j-i)\phi^{B}_{by}]\Lambda^{ij}_{\mathrm{s}} + \sum_{i<j}^{2} \sin[(j-i)\phi^{B}_{by}]\Lambda^{ij}_{\mathrm{as}}\Big), \tag{71}$$

where $\phi^{B,(\lambda)}_{by} = 2\pi(-b+\beta^{(\lambda)}_{y})/3$. On the other hand, whenever y and λ corresponds to a trivial measurement, Bob simply deterministically outputs $b = 0$. After his measurements, Bob performs a post-measurement unitary on the state before he sends it to Charlie. Due to the freedom in a global phase we may set the post-measurement unitary associated with the trivial measurement to be the identity operator. Otherwise, the unitary is optimally chosen to be expanded in a subspace orthogonal to the space in which the parties measures. This ensures that the measurement bases of all parties remains in the same subspace, for which the resulting outcomes will be most strongly correlated. Practically, what this means is that the action of the unitary on a measurement operator should serve as to add a constant shift to the measurement phase. Since the measurements are expressed in terms of the symmetric and anti-symmetric Gell-Mann matrices only, the unitaries must be expansions of the two remaining Gell-Mann matrices, i.e. the diagonal matrices in (10) for $d = 3$. Therefore, the post-measurement unitaries are in the most general way expressed as

$$U^{(\lambda)}_{by} = \exp\Big[\frac{\mathrm{i}}{2}(\mu^{(\lambda)}_{by}\Lambda^{0}_{\mathrm{d}} + \omega^{(\lambda)}_{by}\Lambda^{1}_{\mathrm{d}})\Big]. \tag{72}$$

However, as shown in Appendix E, it is optimal to set $\omega^{(\lambda)}_{by} = \sqrt{3}\mu^{(\lambda)}_{by}$, for which the above reduces to

$$U_{by}^{(\lambda)} = \exp\left[\mathrm{i}\mu_{by}^{(\lambda)} J_3\right], \tag{73}$$

where $J_3 = (\Lambda_{\mathrm{d}}^0 + \sqrt{3}\Lambda_{\mathrm{d}}^1)/2$ is the spin-1 operator associated with the Z-direction. After Bob has performed this unitary on his part of the state he subsequently sends it to Charlie, who performs the measurement $\{C_{c|z}\}_c$, resulting in the outcomes $c \in \{0, 1, 2\}$, with measurement operators given by

$$C_{c|z} = \frac{1}{3}\Big(\mathbb{1} + \sum_{i<j}^{2} \cos[(j-i)\phi_{cz}^C]\Lambda_{\mathrm{s}}^{ij} + \sum_{i<j}^{2} \sin[(j-i)\phi_{cz}^C]\Lambda_{\mathrm{as}}^{ij}\Big), \tag{74}$$

and $\phi_{cz}^C = 2\pi(-c + \gamma_z)/3$. Thus, using these measurements, Alice and Bob evaluates the CGLMP inequality (43) for $d = 3$

$$\mathcal{A}_{AB} = \sum_{abxy} w_{abxy} \sum_{\lambda} p(\lambda)\mathrm{Tr}[(A_{a|x} \otimes B_{b|y}^{(\lambda)})\rho], \tag{75}$$

where the components w_{abxy} can be written in matrix representation as

$$w_{ba01} = w_{ab11} = w_{ac00} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad w_{ac10} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}. \tag{76}$$

Similarly, with the shared state between Alice and Charlie being the average recycled state in (47), evaluated using (71) and (73), their CGLMP parameter is given by

$$\mathcal{A}_{AC} = \sum_{acxz} w_{acxz}\mathrm{Tr}[(A_{a|x} \otimes C_{c|z})\rho^{\mathrm{post}}]. \tag{77}$$

III.2.2. Post-measurement unitaries are optimally outcome independent when recycling maximally entangled qutrits

Just like in the qubit case, the post-measurement unitaries do not have to depend on the outcome b if the shared state between Alice and Bob is a pair of maximally entangled qutrits

$$|\phi^+\rangle = \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle). \tag{78}$$

Although this is expected, the formal proof is less straightforward compared to the qubit case. However, we only have to show that this holds for the correlations between Alice and Charlie originating from the non-trivial measurement of Bob, since no post-measurement unitary is performed after his trivial measurement. As shown explicitly in Appendix F, the correlations between Alice and Charlie whenever y and λ corresponds to non-trivial measurements are given by

$$
\begin{aligned}
p(a,c|x,z) = \frac{1}{9}\sum_{by}\left[\frac{1}{27}\sum_{k=1}^{2}\cos[k(\phi^A_{ax}+\phi^B_{by})]+\frac{1}{18}\right]\times \\
\left[3+2\sum_{k=1}^{2}\cos[k(\phi^C_{cz}-\phi^B_{by}+\mu_{by})]\right],
\end{aligned}
\tag{79}
$$

where the λ label have been omitted for notational convenience. We associate with the expressions in the first and second brackets in (79) with the functions $f(a,b)$ and $g(c,b,\mu_{by})$, respectively (omitting all other dependencies irrelevant for the proof). Thus we may write

$$
p(a,c|x,z) = \frac{1}{9}\sum_{by} f(a,b)g(c,b,\mu_{by}). \tag{80}
$$

We then define a new distribution according to

$$
\tilde{p}(a,c|x,z) = \frac{1}{3}\sum_{l=0}^{2} p(a\oplus_3 l, c\oplus_3 l|x,z), \tag{81}
$$

where $a\oplus_3 l$ denotes addition modulus three. Since $f(a,b)$ and $g(c,b,\mu_{by})$ only depends on the differences $a-b$ and $b-c$, respectively, these functions obey the following symmetries

$$
\begin{aligned}
f(a\oplus_3 l, b\oplus_3 l) &= f(a,b) \\
g(c\oplus_3 l, b\oplus_3 l, \mu_{by}) &= g(c,b,\mu_{by}),
\end{aligned}
\tag{82}
$$

for some arbitrary integer l. Using this we may rewrite (81) in terms of an auxiliary function $h(a,c,\mu_{by}) = \sum_{ly} f(a,l)g(c,l,\mu_{by})$ as

$$
\tilde{p}(a,c|x,z) = \frac{1}{27}\sum_{b} h(a,c,\mu_{by}) \tag{83}
$$

where the above mentioned symmetries in $f(a,b)$ and $g(c,b,\mu_{by})$ have been exploited in order to rearrange the terms appropriately. Since the only dependence on b is found to be in μ_{by}, evaluating the CGLMP parameter $\mathcal{A}_{AC}$ using this distribution we find that the optimal post-measurement unitaries are chosen as $\mu_{by}=\mu_y$. To complete the proof, however, we also have to show that the distribution (81) leads to the same value for $\mathcal{A}_{AC}$ as does (80). To this end, consider evaluating $\mathcal{A}_{AC}$ using the correlations (81), for which we get

$$
\mathcal{A}_{AC} = \sum_{acxz} w_{acxc}\tilde{p}(a,c|x,z) = \frac{1}{3}\sum_{l=0}^{2}\sum_{acxz} w_{acxz}p(a\oplus_3 l, c\oplus_3 l|x,z). \tag{84}
$$

Since $\mathcal{A}_{AC}$ is only concerned about the relative order in which a and c appears, we have the symmetries $w_{(a\oplus_3 l)(c\oplus_3 l)xz} = w_{acxz}$ and $w_{(a\ominus_3 l)(c\ominus_3 l)xz} = w_{acxz}$. This is also evident from the matrix representation of w_{acxz} in (76) since the elements along the diagonals are all the same. Thus, performing the variable substitutions $a' = a\oplus_3 l$ and $c' = c\oplus_3 l$, the above sum can be written as

$$\mathcal{A}_{AC} = \frac{1}{3}\sum_{l=0}^{2}\sum_{a'=l}^{2\oplus_3 l}\sum_{c'=l}^{2\oplus_3 l}\sum_{xz} w_{a'c'xz}\,p(a',c'|x,z). \tag{85}$$

Clearly, shifting the summation bounds of a' and c' by a constant l modulus three will only alter the order in which the terms are added and have no impact on the resulting CGLMP parameter value. Thus, this shift may be disregarded. Performing the sum over l and relabeling $a' \to a$ and $c' \to c$ then gives us the same value for for $\mathcal{A}_{AC}$ as does the original correlations (80), which concludes the proof. Consequently, when Alice and Bob share an entangled pair of maximally entangled qutrits, the optimal post-measurement unitaries performed by Bob are independent of Bob's outcome.

III.2.3. Optimal double violation using entangled pairs of qutrits

In order to find the optimal double violation, given that the shared state between Alice and Bob is maximally entangled $|\phi^+\rangle = (|00\rangle + |11\rangle + |22\rangle)/\sqrt{3}$, we performed the optimization procedure in (49) using the Bell expressions (75) and (77). The result of doing so is presented in Fig. IX where we also show the trade-off curve for $d = 2$ when the optimal state is used. Comparing the $d = 2$ and $d = 3$ curves, it is evident that a much larger double violation is possible using entangled qutrits rather than qubits. This is to be expected considering that greater violations of the local bound are successively shown as d increases in the usual CGLMP scenario (Acín et al. 2002).

We also performed the optimization procedure with non-maximally entangled states. Using the Schmidt decomposition theorem provided in Section II.4, we may write the shared state between Alice and Bob as $|\psi\rangle = \sum_{i=0}^{2}\sqrt{\eta_i}|ii\rangle$, where $\sum_{i=0}^{2}\eta_i = 1$. Doing so we numerically find that the optimal Schmidt coefficients can be expressed in terms of a single parameter ω according to

$$|\psi\rangle = \sqrt{\frac{1-\omega}{2}}|00\rangle + \sqrt{\omega}|11\rangle + \sqrt{\frac{1-\omega}{2}}|22\rangle. \tag{86}$$

This result is not very surprising, considering that the maximally violating state for $d = 3$ in the ordinary CGLMP scenario can be written in a similar way (Zohren and Gill 2008). The state (86), however, do not result in the post-measurement unitaries being dependent on the outcome of Bob's measurements. Why this is the case we can understand from the following. Computing the reduced state of Bob resulting from (86) we find that $\psi_B = \frac{1-\omega}{2}(|0\rangle\langle 0| + |2\rangle\langle 2|) + \omega|1\rangle\langle 1|$, for which we still get $p(b|y) = 1/3$. Therefore, even though the state (86) is not maximally entangled, the reduced state of Bob still features no local structure. As a consequence, no information on Bob's side can be exploited in order to warrant any outcome dependence in the post-measurement unitaries.

Since the optimal state should allow Alice and Bob to leverage local information in order to enhance the double violation, the above observation suggests that the state (86) is sub-optimal. We therefore also considered a more general state, expanded in an unspecified basis

$$|\psi\rangle = \sum_{i=0}^{2} \sqrt{\eta_i}(U \otimes V)|ii\rangle, \tag{87}$$

where U and V are two unitary transformations. We then performed the optimization over the basis by parameterizing U and V according to the procedure outlined in (Spengler et al. 2010). Not only do this state result in an enhanced double violation, we also get a dependence on Bob's outcome in the post-measurement unitaries. The optimal trade-off between $\mathcal{A}_{AB}$ and $\mathcal{A}_{AC}$, resulting from the states (86) and (87), is likewise presented in Fig. IX below.

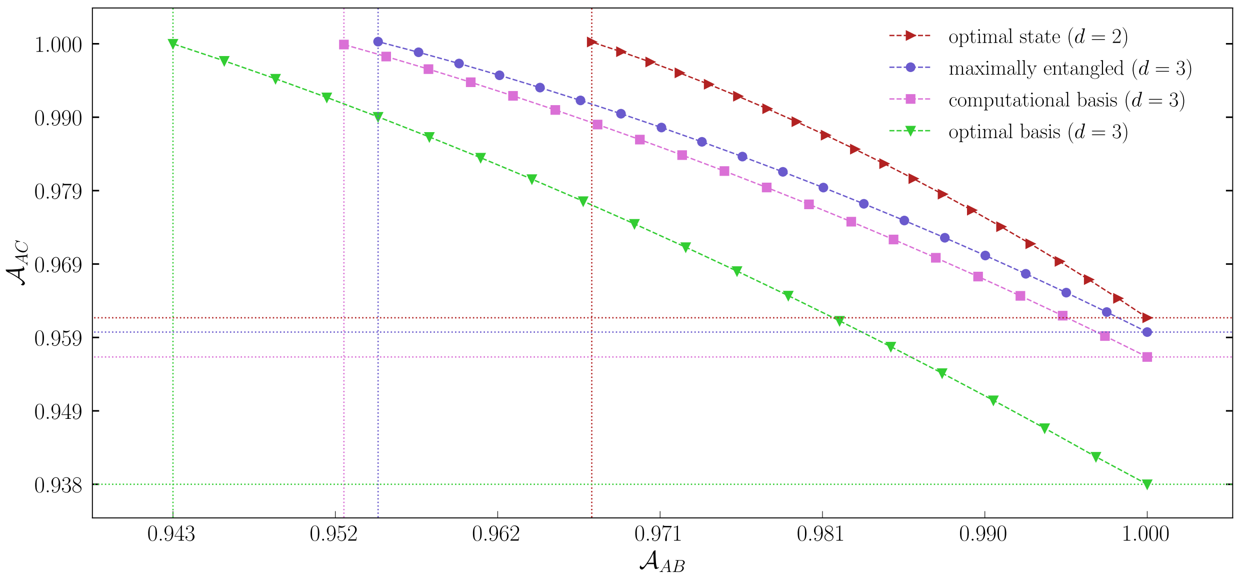


Figure IX: Optimal trade-off between $\mathcal{A}_{AB}$ and $\mathcal{A}_{AC}$ for $d = 3$ under projective measurements and local randomness when the shared state between Alice and Bob is given by (78) (maximally entangled), (86) (computational basis) and (87) (optimal basis), respectively, from the main text. For comparative reasons, the optimal violation for $d = 2$ has also been included.

Moreover, the variation of ω in (86) throughout the double violation is rather small, going from $\omega \approx 0.295$ to $\omega \approx 0.302$, corresponding to $\mathcal{A}_{AB} \approx 1$ and $\mathcal{A}_{AB} \approx 0.953$, respectively. We also find that the two unitaries U and V are optimally the same $U = V$. Introducing this new basis, however, makes it so that the Schmidt coefficients in (87) no longer can be written in terms of a single parameter.

III.2.4. Comparison with protocol using weak measurements

To gain some insights into how well the protocol performs we have made a quantitative comparison with what is possible using weak measurements. In Fig. X we have plotted the optimal trade-off using projective measurements and local randomness, with the shared state between Alice and Bob being (87), together with the trade-off resulting from weak measurements. The latter have been adapted from the the results retrieved by (Roy et al. 2024). Evidently, weak measurements offers greater non-locality sharing than does projective measurements and local randomness. However, as alluded to in the introduction, implementing weak measurements in practice requires more advanced quantum technology than what is needed for projective measurements. Therefore, recycling non-locality using projective measurements is more favorable from a resource point of view.

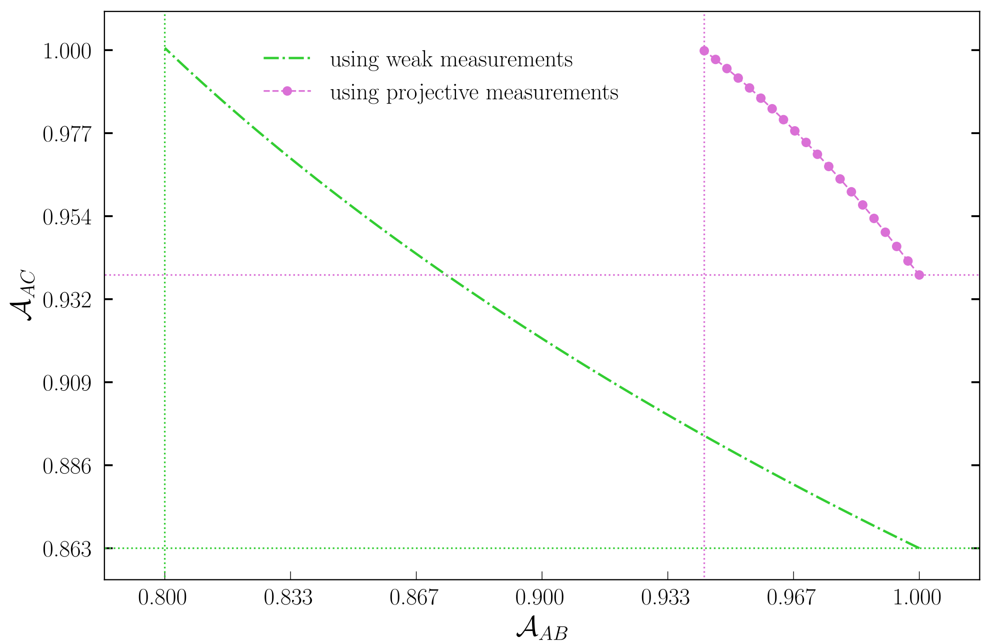


Figure X: The optimal trade-off between two CGLMP parameters $\mathcal{A}_{AB}$ and $\mathcal{A}_{AC}$ for $d = 3$ using projective measurements and local randomness as well as for weak measurements. The curve corresponding to weak measurements have been analytically characterized using the result presented in (Roy et al. 2024).

IV. OUTLOOK

We have explored a means with which one can recycle non-locality using nothing but projective measurements and local classical randomness. Using an explicit measurement strategy, it was initially demonstrated by (Steffinlongo and Tavakoli 2022) that it is possible to recycle non-locality in this way. We have advanced this work by numerically characterizing the optimal trade-off curve of two CHSH parameters in the double violation region. Moreover, we analytically proved some useful assumptions that can be made in the scenario, which enables one to reduce the number of measurement strategies and settings that have to be considered. Thus, we have successfully characterized the scenario that requires the least possible amount of resources necessary to experimentally recycle non-locality. It was also shown, by numerical means, that the states leading to the largest double violations are partially entangled.

Motivated by the advantages offered by higher-dimensional quantum systems, for instance, larger information capacity and higher noise tolerance, the protocol was then further extended to the case in which a pair of entangled qutrits is shared amongst the observers. This we did by explicitly introducing the measurement procedure employed by each measuring party as well as identifying the optimal definition of the post-measurement unitaries. We then applied this measurement procedure in order to numerically determine the optimal trade-off curve between two CGLMP parameters. It was found that qutrits indeed offer greater advantages in the sense that larger double violations are achievable. When the shared state between Alice and Bob is maximally entangled, we also analytically proved that the post-measurement unitaries are outcome independent. The largest possible double violations was also found to be enabled by partially entangled states.

Furthermore, we performed a quantitative comparison between the different protocols with which one can recycle non-locality. Despite the fact that using projective measurements and local randomness leads to the smallest double violations of them all, this method requires the least possible amount of resources and is also the simplest to implement experimentally.

Though we have attained the goals we initially set out to achieve, there are still many questions that remain open. In particular, is it possible, assuming a general two-qubit state, to analytically derive the optimal trade-off curve between two CHSH parameters? Considering the number of non-fixed variables in the CHSH parameters (e.g., see (64) and (65)) this task would very much be non-trivial. Moreover, in Appendix D we provide an explicit measurement strategy resulting in three sequential CHSH violations. It would be interesting to further investigate the entire triad of possible triple violations, essentially characterizing the entire triple violation region. Related to this, it is not yet known how many sequential violations one can actually achieve using projective measurements and local randomness. From a resource perspective, it is favorable to attain many sequential violations since this would allow the state to be even further repurposed in quantum information protocols relying on non-locality. In (Steffinlongo and Tavakoli 2022) it was proven that projective measurements and shared randomness also allows for three sequential CHSH violations, however, from numerical investigations no quadruple violation could be found. It would therefore be safe to assume that this is not possible using local randomness either. Moreover, since the generalized Gell-Mann matrix basis is applicable to quantum systems of arbitrary dimensionality, the measurement protocol provided for qutrits can easily be extended to also account for quantum states in any dimension. As such, it would be interesting to further explore the protocol in even higher dimensions. Especially how much the double violation grows as the number of dimensions are increased, but also what role the dimension plays for the number of possible sequential violations.

Acknowledgements

I have learned a lot during this past year and I am very thankful for all the support I have gotten throughout this experience. In particular, I would like to express my gratitude towards:

My co-supervisors Elna Svegborn and Nicola D' Alessandro. Our discussions have been a tremendous source of inspiration and I am extremely grateful for the assistance you have given me throughout the entire journey.

Every member of the Quantum Information Theory group at Lund who all have been extremely friendly and welcoming.

My family and friends whose emotional support have given me strength during times of difficulty.

References

Acín, A., Durt, T., Gisin, N., and Latorre, J. I. (2002). Quantum nonlocality in two three-level systems. *Physical Review A*, 65(5):052325.

Adam, P. and Diósi, L. (2023). Sequential unsharp measurement of photon polarization. *Physical Review A*, 107(6):063706.

Aspect, A., Grangier, P., and Roger, G. (1982). Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities. *Physical Review Letters*, 49(2):91–94.

Bell, J. S. (1964). On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195–200.

Bell, J. S. (1966). On the Problem of Hidden Variables in Quantum Mechanics. *Reviews of Modern Physics*, 38(3):447–452.

Bertlmann, R., Friis, N., Bertlmann, R., and Friis, N. (2023). *Modern Quantum Theory: From Quantum Mechanics to Entanglement and Quantum Information*. Oxford University Press, Oxford, New York.

Bertlmann, R. A. and Krammer, P. (2008). Bloch vectors for qudits. *Journal of Physics A: Mathematical and Theoretical*, 41(23):235303.

Bloch, F. (1946). Nuclear Induction. *Physical Review*, 70(7-8):460–474.

Bohm, D. (1951). *Quantum Theory*. Prentice-Hall, Inc.

Bohr, N. (1935). Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review*, 48(8):696–702.

Brandt, H. E. (1999). Positive operator valued measure in quantum information processing. *American Journal of Physics*, 67(5):434–439.

Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V., and Wehner, S. (2014). Bell nonlocality. *Reviews of Modern Physics*, 86(2):419–478.

Buhrman, H., Cleve, R., Massar, S., and de Wolf, R. (2010). Nonlocality and communication complexity. *Reviews of Modern Physics*, 82(1):665–698.

Busch, P. and Singh, J. (1998). Lüders theorem for unsharp quantum measurements. *Physics Letters A*, 249(1):10–12.

Caves, C. M. and Milburn, G. J. (2000). Qutrit entanglement. *Optics Communications*, 179(1):439–446.

Chen, J.-L., Wu, C., Kwek, L. C., Oh, C. H., and Ge, M.-L. (2006). Violating Bell Inequalities Maximally for Two \$d\$-Dimensional Systems. *Physical Review A*, 74(3):032106.

Choi, Y.-H., Hong, S., Pramanik, T., Lim, H.-T., Kim, Y.-S., Jung, H., Han, S.-W., Moon, S., and Cho, Y.-W. (2020). Demonstration of simultaneous quantum steering by multiple observers via sequential weak measurements. *Optica*, 7(6):675–679.

Choudhary, S. K., Konrad, T., and Uys, H. (2013). Implementation schemes for unsharp measurements with trapped ions. *Physical Review A*, 87(1):012131.

Cirel'son, B. S. (1980). Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100.

Clauser, J. F., Horne, M. A., Shimony, A., and Holt, R. A. (1969). Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23(15):880–884.

Collins, D., Gisin, N., Linden, N., Massar, S., and Popescu, S. (2002). Bell Inequalities for

Arbitrarily High-Dimensional Systems. *Physical Review Letters*, 88(4):040404.

Cozzolino, D., Da Lio, B., Bacco, D., and Oxenløwe, L. K. (2019). High-Dimensional Quantum Communication: Benefits, Progress, and Future Challenges. *Advanced Quantum Technologies*, 2(12):1900038.

Cubitt, T. S., Leung, D., Matthews, W., and Winter, A. (2011). Zero-Error Channel Capacity and Simulation Assisted by Non-Local Correlations. *IEEE Transactions on Information Theory*, 57(8):5509–5523.

Datta, S. and Majumdar, A. S. (2018). Sharing of nonlocal advantage of quantum coherence by sequential observers. *Physical Review A*, 98(4):042311.

Davies, E. B. and Lewis, J. T. (1970). An operational approach to quantum probability. *Communications in Mathematical Physics*, 17(3):239–260.

Einstein, A., Podolsky, B., and Rosen, N. (1935). Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10):777–780.

Ekert, A. K. (1991). Quantum cryptography based on Bell's theorem. *Physical Review Letters*, 67(6):661–663.

Feng, T., Ren, C., Tian, Y., Luo, M., Shi, H., Chen, J., and Zhou, X. (2020). Observation of nonlocality sharing via not-so-weak measurements. *Physical Review A*, 102(3):032220.

Foletto, G., Calderaro, L., Tavakoli, A., Schiavon, M., Picciariello, F., Cabello, A., Villoresi, P., and Vallone, G. (2020). Experimental Certification of Sustained Entanglement and Nonlocality after Sequential Measurements. *Physical Review Applied*, 13(4):044008.

Freedman, S. J. and Clauser, J. F. (1972). Experimental Test of Local Hidden-Variable Theories. *Physical Review Letters*, 28(14):938–941.

Ghirardi, G. C., Weber, T., and Rimini, A. (1980). A GENERAL ARGUMENT AGAINST SUPERLUMINAL TRANSMISSION THROUGH THE QUANTUM MECHANICAL MEASUREMENT PROCESS. *Lett. Nuovo Cim.*, 27:293–298.

Goyal, S. K., Simon, B. N., Singh, R., and Simon, S. (2016). Geometry of the generalized Bloch sphere for qutrits. *Journal of Physics A: Mathematical and Theoretical*, 49(16):165203.

Griffiths, D. J. and Schroeter, D. F. (2018). Introduction to Quantum Mechanics. ISBN: 9781316995433.

Harrigan, N. and Spekkens, R. W. (2010). Einstein, Incompleteness, and the Epistemic View of Quantum States. *Foundations of Physics*, 40(2):125–157.

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3):172–198.

Herbert, N. (1975). Cryptographic approach to hidden variables. *American Journal of Physics*, 43(4):315–316.

Holevo, A. S. (2001). Statistical Structure of Quantum Theory. volume 67 of *Lecture Notes in Physics Monographs*, Berlin, Heidelberg. Springer Berlin Heidelberg.

Huang, X., Zhan, L., Li, L., Bao, S., Tao, Z., and Ying, J. (2023). Quantum Correlation Resource Recycling via Sequential Measurements: Theoretical Models and Optical Experiments. *Photonics*, 10(12):1314.

Kaszlikowski, D., Gnaciński, P., Żukowski, M., Miklaszewski, W., and Zeilinger, A. (2000). Violations of Local Realism by Two Entangled N-Dimensional Systems Are Stronger than for Two Qubits. *Physical Review Letters*, 85(21):4418–4421.

Kimura, G. (2003). The Bloch Vector for N-Level Systems. *Physics Letters A*, 314(5-6):339–

349.
Kraus, K., Böhm, A., Dollard, J. D., and Wootters, W. H., editors (1983). *States, Effects, and Operations Fundamental Notions of Quantum Theory*, volume 190 of *Lecture Notes in Physics*. Springer, Berlin, Heidelberg.
Kryszewski, S. and Zachcial, M. (2006). Alternative representation of N \times N density matrix.
Li, N., Luo, S., and Sun, Y. (2021). Information-theoretic aspects of Werner states. *Annals of Physics*, 424:168371.
Long, G.-L. (2025). Progress in quantum information. *National Science Review*, 12(8):nwaf289.
Lüders, G. (2006). Concerning the state-change due to the measurement process * **. *Annalen der Physik*, 518:663–670.
M. Barnett, S. (2009). *Quantum Information*. Oxford Master Series in Physics. Oxford University Press, Oxford, New York.
Mao, Y.-L., Li, Z.-D., Steffinlongo, A., Guo, B., Liu, B., Xu, S., Gisin, N., Tavakoli, A., and Fan, J. (2023). Recycling nonlocality in quantum star networks. *Physical Review Research*, 5(1):013104.
Nielsen, M. A. and Chuang, I. L. (2010). Quantum Computation and Quantum Information: 10th Anniversary Edition. ISBN: 9780511976667.
Ramya, R., Kumar, P., Dhanasekaran, D., Kumar, R. S., and Sharavan, S. A. (2025). A review of quantum communication and information networks with advanced cryptographic applications using machine learning, deep learning techniques. *Franklin Open*, 10:100223.
Roy, S., Kumari, A., Mal, S., and Sen(De), A. (2024). Robustness of higher-dimensional nonlocality against dual noise and sequential measurements. *Physical Review A*, 109(6):062227.
Schrödinger, E. (1935). Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, 23(48):807–812.
Schrödinger, E. (1936). Probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(3):446–452.
Shenoy H., A., Designolle, S., Hirsch, F., Silva, R., Gisin, N., and Brunner, N. (2019). Unbounded sequence of observers exhibiting Einstein-Podolsky-Rosen steering. *Physical Review A*, 99(2):022317.
Silva, R., Gisin, N., Guryanova, Y., and Popescu, S. (2015). Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements. *Physical Review Letters*, 114(25):250401.
Spengler, C., Huber, M., and Hiesmayr, B. C. (2010). A composite parameterization of unitary groups, density matrices and subspaces. *Journal of Physics A: Mathematical and Theoretical*, 43(38):385306.
Steffinlongo, A. and Tavakoli, A. (2022). Projective Measurements Are Sufficient for Recycling Nonlocality. *Physical Review Letters*, 129(23):230402.
Uola, R., Costa, A. C., Nguyen, H. C., and Gühne, O. (2020). Quantum steering. *Reviews of Modern Physics*, 92(1):015001.
Werner, R. F. (1989). Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281.

Werner, R. F. and Wolf, M. M. (2001). Bell inequalities and entanglement. *Quant. Inf. Comput.*, 1(3):1–25.

White, A. G., James, D. F. V., Eberhard, P. H., and Kwiat, P. G. (1999). Nonmaximally Entangled States: Production, Characterization, and Utilization. *Physical Review Letters*, 83(16):3103–3107.

Wilde, M. M. (2013). *Quantum Information Theory.* Cambridge University Press, Cambridge.

Xiao, Y., Xin Rong, Y., Wang, S., Hong Han, X., Shi Xu, J., and Gu, Y. J. (2024). Experimental sharing of Bell nonlocality with projective measurements. *New Journal of Physics*, 26(5):053019.

Zhou, W., Li, X., Yang, R., Tao, L., Liu, Y.-C., Xie, Z., Gong, Y.-X., and Zhu, S.-N. (2025). Experimental device-independent efficient certification of unsharp measurements. *Physical Review Research*, 7(3):033106.

Zohren, S. and Gill, R. D. (2008). Maximal Violation of the Collins-Gisin-Linden-Massar-Popescu Inequality for Infinite Dimensional States. *Physical Review Letters*, 100(12):120406.

Zohren, S., Reska, P., Gill, R. D., and Westra, W. (2010). A tight Tsirelson inequality for infinitely many outcomes. *Europhysics Letters*, 90(1):10002.

Łuczak, A. and Podsdkowska, H. (2015). Lüders Instruments, Generalised Lüders Theorem, and Some Aspects of Sufficiency. *International Journal of Theoretical Physics*, 54(12):4283–4292.

A. Proving that the completely trivial measurement strategy can be disregarded when the maximally entangled state is recycled

Given that the shared state between Alice and Bob is a pair of maximally entangled qubits, we wish to prove that the completely trivial strategy, wherein both measurement settings corresponds to trivial measurements, can be omitted since it do not contribute to the simultaneous violation of $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ (see Tab. A1). In doing so we also find that the distribution of local randomness $p(\lambda|y)$ do not have to be conditioned on the measurement setting of Bob, i.e. that $p(\lambda|y) = p(\lambda)$.

Table A1: All the possible combinations of trivial and non-trivial measurements that Bob may consider.

λ	$y=0$	$y=1$
0	$\{B^{(0)}_{b\|0}\}_b$	$\{B^{(0)}_{b\|1}\}_b$
1	$\{\mathbb{1},0\}$	$\{B^{(1)}_{b1}\}_b$
2	$\{\mathbb{1},0\}$	$\{\mathbb{1},0\}$

Performing the sum over a, b and x in the CHSH parameter between Alice and Bob (46), with $w_{abxy} = (-1)^{a+b+xy}$, we arrive at

$$\mathcal{S}_{AB} = \sum_\lambda \sum_y p(\lambda|y)\Big(\langle A_0, B^{(\lambda)}_y\rangle_{\phi^+} + (-1)^y \langle A_1, B^{(\lambda)}_y\rangle_{\phi^+}\Big) = \sum_\lambda \sum_y p(\lambda|y)\tilde{\mathcal{S}}_{AB}(B^{(\lambda)}_y) \tag{A.1}$$

where $A_x = A_{0|x} - A_{1|x}$ and $B^{(\lambda)}_y = B^{(\lambda)}_{0|y} - B^{(\lambda)}_{1|y}$ are the observables of Alice and Bob, respectively, and $\langle A_x, B^{(\lambda)}_y\rangle_{\phi^+}$ is the joint expectation value of these given that the shared state is ϕ^+. Also, in the second line we defined the functional $\tilde{\mathcal{S}}_{AB}(B^{(\lambda)}_y)$ to be the expression within the parenthesis in the first equality. Further, note that $\tilde{\mathcal{S}}_{AB}(B^{(\lambda)}_y)$ will vanish whenever y and λ corresponds to Bob performing trivial measurements. This follows from the fact that any qubit observable can be expanded in terms of the Pauli matrices. Since these are traceless and Alice's local state is maximally mixed, $\tilde{\mathcal{S}}_{AB}(B^{(\lambda)}_y)$ must be zero whenever λ and y correspond to a trivial measurement. This is also physically reasonable, if Bob simply deterministically outputs $b=0$ (or $b=1$) this will clearly not be correlated with the measurement outcomes of Alice.

Before we compute $\mathcal{S}_{AC}$ we first have to retrieve the correlations between Alice and Charlie resulting from their average shared state in (47). For this purpose, we define $\sigma_{a|x} = \mathrm{Tr}_A[(A_{a|x}\otimes \mathbb{1})\phi^+]$ to be the sub-normalized state remotely prepared by Alice on Bob's side, meaning that $\mathrm{Tr}(\sigma_{a|x}) = p(a|x) = 1/2$, which is the probability of Alice observing outcome a. Then, using (47), with $\rho = \phi^+$, the correlations between Alice and Charlie can be written as

$$\begin{aligned} p(a,c|x,z) &= \frac{1}{2}\sum_\lambda \sum_{by} p(\lambda|y)\mathrm{Tr}\Big[C_{c|z}K^{(\lambda)}_{b|y}\sigma_{a|x}(K^{(\lambda)}_{b|y})^\dagger\Big] \\ &= \frac{1}{2}\sum_\lambda \sum_{y} p(\lambda|y)\mathrm{Tr}\Big[C_{c|z}\Gamma^{(\lambda)}_y[\sigma_{a|x}]\Big], \end{aligned} \tag{A.2}$$

where in the second line we defined the CPTP map $\Gamma^{(\lambda)}_y$ corresponding to the quantum instrument $\{K^{(\lambda)}_{b|y}\}_b$ employed by Bob. Inserting this into the CHSH parameter between Alice and Charlie (48) and summing over all the outcomes, using the notation $\sigma_x = \sum_a (-1)^a \sigma_{a|x}$, we end up with

$$\mathcal{S}_{AC} = \sum_\lambda \sum_y p(\lambda|y) \frac{1}{2} \sum_{xz} (-1)^{xz} \mathrm{Tr}\Big[C_z \Gamma_y^{(\lambda)}[\sigma_x]\Big] = \sum_\lambda \sum_y p(\lambda|y) \tilde{\mathcal{S}}_{AC}(\Gamma_y^{(\lambda)}), \tag{A.3}$$

where in the last equality we defined the functional $\tilde{\mathcal{S}}_{AC}(\Gamma_y^{(\lambda)})$.

We now expand both $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ in terms of $\tilde{\mathcal{S}}_{AB}(B_y^{(\lambda)})$ and $\tilde{\mathcal{S}}_{AC}(\Gamma_y^{(\lambda)})$, respectively, leading to

$$\mathcal{S}_{AB} = p(0|0)\tilde{\mathcal{S}}_{AB}(B_0^{(0)}) + p(0|1)\tilde{\mathcal{S}}_{AB}(B_1^{(0)}) + p(1|1)\tilde{\mathcal{S}}_{AB}(B_1^{(1)}) \tag{A.4}$$

and

$$\begin{aligned} \mathcal{S}_{AC} = p(0|0)\tilde{\mathcal{S}}_{AC}(\Gamma_0^{(0)}) + p(1|0)\tilde{\mathcal{S}}_{AC}(\mathbb{1}) + p(2|0)\tilde{\mathcal{S}}_{AC}(\mathbb{1}) + \\ p(0|1)\tilde{\mathcal{S}}_{AC}(\Gamma_1^{(0)}) + p(1|1)\tilde{\mathcal{S}}_{AC}(\Gamma_1^{(1)}) + p(2|1)\tilde{\mathcal{S}}_{AC}(\mathbb{1}), \end{aligned} \tag{A.5}$$

where the notation $\tilde{\mathcal{S}}_{AC}(\mathbb{1})$ implies that Bob do not make use of any quantum instrument. Note that the functional form of $\tilde{\mathcal{S}}_{AB}(B_1^{(0)})$ and $\tilde{\mathcal{S}}_{AB}(B_1^{(1)})$ are the same, the only difference being that they are functions of $B_1^{(0)}$ and $B_1^{(1)}$, respectively. Since $\tilde{\mathcal{S}}_{AB}(B_1^{(0)})$ and $\tilde{\mathcal{S}}_{AB}(B_1^{(1)})$ both enter into $\mathcal{S}_{AB}$ as a convex sum, maximizing $\mathcal{S}_{AB}$ separately over $B_1^{(0)}$ and $B_1^{(1)}$ it then follows that the optimal $B_1^{(0)}$ must be the same as the optimal $B_1^{(1)}$. That is,

$$\max_{\{B_1^{(0)}, B_1^{(1)}\}} \mathcal{S}_{AB} \Rightarrow B_1^{(0)} = B_1^{(1)}. \tag{A.6}$$

The same argument can be utilized to show that equating the quantum instruments $\Gamma_1^{(0)} = \Gamma_1^{(1)}$ must also be optimal for $\mathcal{S}_{AC}$. Since this choice of Bob's measurements is optimal for both $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ separately, this must also be optimal for their trade-off. This line of reasoning also allow us to set the post-measurement unitaries corresponding to the trivial measurements to be equal.

We can now write $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$ in a much simpler form. Let the distributions $p(\lambda|y)$ be parametrized according to $p(0|0) = s$, $p(1|0) + p(2|0) = 1 - s$, $p(0|1) + p(1|1) = t$ and $p(2|1) = 1 - t$, where $s, t \in [0, 1]$, which allows us to write

$$\mathcal{S}_{AB} = s\tilde{\mathcal{S}}_{AB}(B_0^{(0)}) + t\tilde{\mathcal{S}}_{AB}(B_1^{(1)}) \tag{A.7}$$

and

$$\mathcal{S}_{AC} = s\tilde{\mathcal{S}}_{AC}(\Gamma_0^{(0)}) + t\tilde{\mathcal{S}}_{AC}(\Gamma_1^{(1)}) + (2 - s - t)\tilde{\mathcal{S}}_{AC}(\mathbb{1}). \tag{A.8}$$

Our aim is to define a new distribution, not featuring the completely trivial strategy, in terms of s and t that preserves the relative frequency by which Bob employs trivial and non-trivial measurements, respectively, without altering the attainable values of $\mathcal{S}_{AB}$ and $\mathcal{S}_{AC}$. To this end, we have to show that the factor in front of $\tilde{\mathcal{S}}_{AC}(\mathbb{1})$ in (A.8) can be thought of as a probability, more specifically, that $0 \leq 2 - s - t \leq 1$. The lower bound is clearly respected since s and t themselves are probabilities. Thus, we have to show that $s + t > 1$. By assumption, we have

$$s\tilde{\mathcal{S}}_{AB}(B_0^{(0)}) + t\tilde{\mathcal{S}}_{AB}(B_1^{(1)}) > 2. \tag{A.9}$$

The smallest possible value of s and t are attained when Alice and Bob measures optimally, without any regards to the value of $\mathcal{S}_{AC}$. This is achieved if Alice measures the observables $A_0 = (\sigma_1 + \sigma_3)/\sqrt{2}$ and $A_1 = (\sigma_1 - \sigma_3)/\sqrt{2}$, and Bob $B_0^{(0)} = \sigma_3$ and $B_1^{(1)} = \sigma_1$, see Fig. II (b) in the main text, for which we get that $\tilde{\mathcal{S}}_{AB}(B_0^{(0)}) = \tilde{\mathcal{S}}_{AB}(B_1^{(1)}) = \sqrt{2}$. We therefore find $s + t > \sqrt{2}$, which in turn implies that $0 \leq 2 - s - t < 2 - \sqrt{2}$. In other words, we can interpret $2 - s - t$ as the total frequency with which Bob performs trivial measurements. As such, one may always construct some new distribution of measurements, according to Tab. A2 below, in which the relative frequency of trivial and non-trivial measurements is preserved. The difference being that the conditional probability distribution corresponding to $y = 1$ is deterministic. Then, if Bob performs his measurements in accordance with Tab. A2, the resulting CHSH parameters becomes

$$\mathcal{S}_{AB} = p(0)\tilde{\mathcal{S}}_{AB}(B_1^{(1)}) + \tilde{\mathcal{S}}_{AB}(B_1^{(1)}) \tag{A.10}$$

and

$$\mathcal{S}_{AC} = p(0)\tilde{\mathcal{S}}_{AC}(\Gamma_0^{(0)}) + \tilde{\mathcal{S}}_{AC}(\Gamma_1^{(1)}) + (1 - p(0))\tilde{\mathcal{S}}_{AC}(\mathbb{1}). \tag{A.11}$$

Note that we, in effect, can achieve this by just setting $t = 1$ in (A.7) and (A.8).

Table A2: The the frequency with which Bob performs trivial and non-trivial measurements remains invariant if one defines a new set of measurements and distributions according to the table below.

Probability Dist.	Measurement of Bob
$p(\lambda = 0\|y = 0) = (s + t) - 1$	$\{B_{b\|0}^{(0)}\}_b$
$p(\lambda = 1\|y = 0) = 2 - (s + t)$	$\{\mathbb{1}, 0\}$
$p(\lambda = 1\|y = 1) = 1$	$\{B_{b\|1}^{(1)}\}_b$

B. Derivation of the CHSH parameters when the maximally entangled state is recycled

We analytically derive the CHSH parameters between Alice-Bob and Alice-Charlie when the shared state between the former is maximally entangled. Furthermore, we assume that Bob only employs the strategies given in Tab. I in the main text.

To begin, we can parametrize the observables of A as $A_x = \vec{u}_x \cdot \vec{\sigma} = \cos\theta\sigma_1 + (-1)^x \sin\theta\sigma_3$, that is, we assume Alice to measure in the XZ-plane only and that the two measurement Bloch vectors are reflections about the X-axis in the Bloch sphere. This comes at no loss of generality since any global rotation of Alice's local system can be absorbed into the measurements of Bob via the so-called action at a distance relation $R \otimes \mathbb{1}|\phi^+\rangle = \mathbb{1} \otimes R^T|\phi^+\rangle$. Further, it is optimal for all parties to measure in the same plane, as such, the observables of Bob corresponding to non-trivial measurements can be parametrized as $B_y^{(\lambda)} = \vec{v}_y^{(\lambda)} \cdot \vec{\sigma} = \cos\beta_y^{(\lambda)}\sigma_1 + \sin\beta_y^{(\lambda)}\sigma_3$. Writing the state $|\phi^+\rangle$ in Bloch form

$$\phi^+ = \frac{1}{4}\Big(\mathbb{1} \otimes \mathbb{1} + \sum_i t_{ii}\sigma_i \otimes \sigma_i\Big), \tag{B.1}$$

where t_{ii} are elements of the corelation tensor $T = \mathrm{diag}(1, -1, 1)$, one can straightforwardly compute

$$\mathrm{Tr}[A_x \otimes B_y^{(\lambda)}\phi^+] = \cos\theta\cos\beta_y^{(\lambda)} + (-1)^x \sin\theta\sin\beta_y^{(\lambda)} = \cos(\beta_y^{(\lambda)} + (-1)^{x+1}\theta), \tag{B.2}$$

where we have used the identity $\mathrm{Tr}[(\vec{a}\cdot\vec{\sigma})\sigma_i] = 2a_i$. Since the Pauli matrices are traceless, the contribution from the trivial measurements performed by Bob will vanish. As such, performing the sum over x and y using

$$\begin{aligned} 2\cos\gamma\cos\delta &= \cos(\gamma + \delta) + \cos(\gamma - \delta) \\ 2\sin\gamma\sin\delta &= \cos(\gamma + \delta) - \cos(\gamma - \delta), \end{aligned} \tag{B.3}$$

we arrive at

$$\mathcal{S}_{AB} = 2p(0)\cos\theta\cos\beta_0^{(0)} + 2\sin\theta\sin\beta_1^{(1)}, \tag{B.4}$$

where we have used that $\beta_1^{(0)} = \beta_1^{(1)}$.

Moving on, it will be optimal for Charlie to also measure in the XZ-plane, as such, we can parametrize their observable as $C_z = \vec{w}_z \cdot \vec{\sigma} = \cos\gamma_z\sigma_1 + \sin\gamma_z\sigma_3$. Writing out the state ϕ^+ in Bloch form and using

$$\rho^{\text{post}} = \frac{1}{2}\sum_\lambda\sum_{by} p(\lambda)\Big(\mathbb{1}\otimes U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\Big)\phi^+\Big(\mathbb{1}\otimes B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger\Big), \tag{B.5}$$

as the shared state on which Alice and Charlie measures on, the joint expectation value between the two can be seen to be

$$\text{Tr}[(A_x\otimes C_z)\rho^{\text{post}}] = \frac{1}{8}\sum_{by}\sum_i t_{ii}p(\lambda)\text{Tr}[A_x\sigma_i\otimes C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_iB_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger]. \tag{B.6}$$

Then, inserting $A_x = \cos\theta\sigma_1 + (-1)^x\sin\theta\sigma_3$ and summing over i, the trace above can be written as

$$2\cos\theta\text{Tr}[(C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_1B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger] + (-1)^x2\sin\theta\text{Tr}[C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_3B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger].$$

Further, using the completeness relation we can write $C_z = 2C_{0|z} - \mathbb{1}$, for which the above instead becomes

$$4\cos\theta\text{Tr}[C_{0|z}U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_1B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger] + 4(-1)^x\sin\theta\text{Tr}[C_{0|z}U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_3B_{b|y}^{(\lambda)}(U_{by}^{\lambda})^\dagger]. \tag{B.7}$$

Here we have omitted the terms corresponding to $\mathbb{1}$ in $C_z = 2C_{0|z} - \mathbb{1}$, this since the sum over b would cause these terms to vanish, for instance

$$\sum_b \text{Tr}[U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_1B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger] = \sum_b \text{Tr}[B_{b|y}^{(\lambda)}\sigma_1] = \text{Tr}[\sigma_1] = 0. \tag{B.8}$$

We then end up with

$$\begin{aligned}\mathcal{S}_{AC} &= \frac{1}{2}\sum_{xz}\sum_{by\lambda}p(\lambda)(-1)^{xz}\cos\theta\text{Tr}[C_{0|z}U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_1B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger]\\ &= \frac{1}{2}\sum_{xz}\sum_{by\lambda}p(\lambda)(-1)^{x(z+1)}\sin\theta\text{Tr}[C_{0|z}U_{by}^{(\lambda)}B_{b|y}^{\lambda}\sigma_3B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger]\\ &= \sum_{by\lambda}p(\lambda)\cos\theta\text{Tr}[(U_{by}^{(\lambda)})^\dagger C_{0|0}U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_1B_{b|y}^{(\lambda)}] + \sum_{by\lambda}p(\lambda)\sin\theta\text{Tr}[(U_{by}^{(\lambda)})^\dagger C_{0|1}U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_3B_{b|y}^{(\lambda)}].\end{aligned} \tag{B.9}$$

Moreover, whenever $B_{b|y}^{(\lambda)}$ corresponds to a non-trivial measurement we have

$$B_{b|y}^{(\lambda)}\sigma_iB_{b|y}^{(\lambda)} = \frac{1}{2}v_{y,i}^{(\lambda)}\Big((-1)^b\mathbb{1} + (-1)^{2b}(\vec{v}_y^{(\lambda)}\cdot\vec{\sigma})\Big), \tag{B.10}$$

from which it follows

$$\text{Tr}[(U_{by}^{(\lambda)})^\dagger C_{0|z}U_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_iB_{b|y}^{(\lambda)}] = \frac{1}{2}v_{y,i}^{(\lambda)}\text{Tr}[(U_{by}^{(\lambda)})^\dagger C_{0|z}U_{by}^{(\lambda)}(\vec{v}_y^{(\lambda)}\cdot\vec{\sigma})]. \tag{B.11}$$

Since the optimal unitaries aims to align the Bloch vector $\vec{v}_y^{(\lambda)}$ with the Bloch vector corresponding to $C_{0|z}$, and neither of these depend on b, we may omit the dependence on b in the unitaries as well. From a physical standpoint this is reasonable since the local state of Bob is maximally mixed and no information about the state can therefore be extracted on Bob's side, that is, the distribution of outcomes is uniform. Thus, conditioning the post-measurement unitaries on this distribution will not improve the trade-off. Further, since every observer measures in the XZ-plane, the optimal unitaries must be rotations about the Y-axis, consequently

$$U_{by}^{(\lambda)} = U_y^{(\lambda)} = e^{i\frac{\mu_y^{(\lambda)}}{2}\sigma_2}. \tag{B.12}$$

Then, writing $C_{0|z} = (\mathbb{1} + \vec{w}_z \cdot \vec{\sigma})/2$ and using Rodrigues' rotation formula (7) we get that

$$\begin{aligned} e^{i\mu_y^{(\lambda)}\sigma_2/2}(\vec{w}_z \cdot \vec{\sigma})e^{-i\mu_y^{(\lambda)}\sigma_2/2} &= [\cos\mu_y^{(\lambda)}\vec{w}_z + (1-\cos\mu_y^{(\lambda)})\hat{n}_y(\hat{n}_y \cdot \vec{w}_z) - \sin\mu_y^{(\lambda)}\hat{n}_y \times \vec{w}_z] \cdot \vec{\sigma} \\ &= \vec{w}_{zy}^{(\lambda)} \cdot \vec{\sigma}, \end{aligned} \tag{B.13}$$

where $\bar{w}_{zy}^{(\lambda)} = (\cos[\gamma_z - \mu_y^{(\lambda)}], 0, \sin[\gamma_z - \mu_y^{(\lambda)}])$. In the event that Bob performs an identity projection, we may always associate the relevant outcome with $b = 0$, that is, if $B_y^{(\lambda)} = \mathbb{1}$ we can choose $B_{0|y}^{(\lambda)} = \mathbb{1}$ and $B_{1|y}^{(\lambda)} = 0$. This is a consequence of the CHSH parameter being invariant under the coordinate permutation $b \to \bar{b}$ if $y = 0$, $x \to \bar{x}$ and $a \to \bar{a}$.

We are now in a position to expand our expression for $\mathcal{S}_{AC}$ in (E.9). The contribution from the terms corresponding to non-trivial projective measurements takes the following form

$$p(\lambda)\big(\cos\theta\cos\beta_y^{(\lambda)}\ \vec{w}_{0y}^{(\lambda)} \cdot \vec{v}_y^{(\lambda)} + \sin\theta\sin\beta_y^{(\lambda)}\ \vec{w}_{1y}^{(\lambda)} \cdot \vec{v}_y^{(\lambda)}\big), \tag{B.14}$$

where

$$\vec{w}_{zy}^{(\lambda)} \cdot \vec{v}_y^{(\lambda)} = \cos(\gamma_z - \mu_y^{(\lambda)})\cos\beta_y^{(\lambda)} + \sin(\gamma_z - \mu_y^{(\lambda)})\sin\beta_y^{(\lambda)} = \cos(\beta_y^{(\lambda)} - \gamma_z + \mu_y^{(\lambda)}). \tag{B.15}$$

On the other hand, the measurements corresponding to identity projections are effectively reduced to a unitary acting on the measurement Bloch vectors of Charlie, accordingly, the contributions from these are given as

$$p(\lambda)\big(\cos\theta\cos(\gamma_0 - \mu_y^{(\lambda)}) + \sin\theta\sin(\gamma_1 - \mu_y^{(\lambda)})\big). \tag{B.16}$$

Further, we can fix the unitary $U_0^{(1)}$ as a reference due to the freedom of a global shift in the unitaries, i.e, $U_0^{(1)} = \mathbb{1}$. With the above considerations in mind the final expression for $\mathcal{S}_{AC}$ can be seen to be

$$\begin{aligned} \mathcal{S}_{AC} = &\Big(\cos\theta\cos\beta_0^{(0)}\cos(\beta_0^{(0)} - \gamma_0 + \mu_0^{(0)}) + \sin\theta\sin\beta_0^{(0)}\cos(\beta_0^{(0)} - \gamma_1 + \mu_0^{(0)})\Big)p(0) \\ &+ \Big(\cos\theta\cos\gamma_0 + \sin\theta\sin\gamma_1\Big)(1 - p(0)) \\ &+ \Big(\cos\theta\cos\beta_1^{(1)}\cos(\beta_1^{(1)} - \gamma_0 + \mu_1^{(1)}) + \sin\theta\sin\beta_1^{(1)}\cos(\beta_1^{(1)} - \gamma_1 + \mu_1^{(1)})\Big). \end{aligned} \tag{B.17}$$

C. Derivation of the CHSH parameters when a partially entangled state is recycled

We now wish to characterize the CHSH parameters of Alice-Bob and Alice-Charlie for the case in which the initial shared state between the former is the partially entangled state $|\psi_\phi\rangle = \cos\phi|00\rangle + \sin\phi|11\rangle$,

$\phi \in [0, \pi/4]$. We can write this state in Bloch form as

$$\psi_\phi = \frac{1}{4}\Big(\mathbb{1}\otimes\mathbb{1} + \cos 2\phi(\sigma_3\otimes\mathbb{1} + \mathbb{1}\otimes\sigma_3) + \sum_i t_{ii}\sigma_i\otimes\sigma_i\Big), \tag{C.1}$$

where the elements of the correlation tensor now are given by $T = \mathrm{diag}(\sin 2\phi, -\sin 2\phi, 1)$. Since the local qubit states no longer are maximally mixed we cannot parametrize the measurements of Alice in terms of a single angle θ. We therefore denote the observables of Alice by $A_x = \cos\alpha_x\sigma_1 + \sin\alpha_x\sigma_3$, while otherwise using the same notation and assumptions as for the maximally entangled state. Whenever $B_y^{(\lambda)}$ is a non-trivial projector we get

$$\mathrm{Tr}[(A_x\otimes B_y^{(\lambda)})\psi_\phi] = \frac{1}{4}\sum_i t_{ii}\Big[(\bar{u}_x\cdot\bar{\sigma})\sigma_i\otimes(\bar{v}_y^{(\lambda)}\cdot\bar{\sigma})\sigma_i\Big] = \sin 2\phi\cos\theta_x\cos\beta_y^{(\lambda)} + \sin\theta_x\sin\beta_y^{(\lambda)}. \tag{C.2}$$

Whereas if $B_y^{(\lambda)}$ is an identity projection

$$\mathrm{Tr}[(A_x\otimes\mathbb{1})\psi_\phi] = \cos 2\phi\sin\alpha_x. \tag{C.3}$$

Introducing the shorthand notation $\eta = (\alpha_0+\alpha_1)/2$ and $\delta = (\alpha_0-\alpha_1)/2$ we then straightforwardly compute

$$\begin{aligned}\mathcal{S}_{AB} &= (1-p(0))2\cos 2\phi\sin\eta\cos\delta\\ &+ 2p(0)\cos\delta\Big(\sin 2\phi\cos\eta\cos\beta_0^{(0)} + \sin\eta\sin\beta_0^{(0)}\Big)\\ &+ 2\sin\delta\Big(\cos\eta\sin\beta_1^{(1)} - \sin 2\phi\sin\eta\cos\beta_1^{(1)}\Big)\end{aligned} \tag{C.4}$$

This indeed reduces to (B.4) for $\phi = \pi/4$, $\alpha_0 = \theta$ and $\alpha_1 = 2\pi - \theta$, as one would expect. Note that we also get a contribution from the local term in ψ_ϕ.

Using the same formula for the average state between Alice and Charlie as previously, the joint expectation value between Alice and Charlie will be

$$\begin{aligned}\mathrm{Tr}[(A_x\otimes C_z)\psi_\phi^{\mathrm{post}}] &= \frac{1}{8}\sum_{by\lambda}p(\lambda)\cos 2\phi\,\mathrm{Tr}[A_x\sigma_3\otimes C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger]\\ &+\frac{1}{8}\sum_{by\lambda}p(\lambda)\sum_i t_{ii}\mathrm{Tr}[A_x\sigma_i\otimes C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_iB_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger].\end{aligned} \tag{C.5}$$

The first trace evaluates into

$$\mathrm{Tr}[A_x\sigma_3\otimes C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger] = 2(-1)^b u_{x,3}\vec{w}_{zby}^{(\lambda)}\cdot\vec{v}_y^{(\lambda)}, \tag{C.6}$$

where $\bar{w}_{zby}^{(\lambda)}$ is defined in the same way as before, and the second trace becomes

$$\mathrm{Tr}[A_x\sigma_i\otimes C_zU_{by}^{(\lambda)}B_{b|y}^{(\lambda)}\sigma_iB_{b|y}^{(\lambda)}(U_{by}^{(\lambda)})^\dagger] = 2u_{x,i}v_{y,i}^{(\lambda)}\vec{w}_{zby}^{(\lambda)}\cdot\vec{v}_y^{(\lambda)}. \tag{C.7}$$

At this point we have

$$\mathrm{Tr}[(A_x\otimes C_z)\psi_\phi^{\mathrm{post}}] = \frac{1}{4}\sum_{by\lambda}p(\lambda)\Big(\cos 2\phi(-1)^b u_{x,3}\vec{w}_{zby}^{(\lambda)}\cdot\vec{v}_y^{(\lambda)} + \sum_i t_{ii}u_{x,i}v_{y,i}^{(\lambda)}\vec{w}_{zby}^{(\lambda)}\cdot\vec{v}_y^{(\lambda)}\Big). \tag{C.8}$$

We then perform the sum over b using the relations

$$\begin{aligned}
\sin\theta + \sin\varphi &= 2\sin\frac{1}{2}(\theta+\varphi)\cos\frac{1}{2}(\theta-\varphi)\\
\sin\theta - \sin\varphi &= 2\cos\frac{1}{2}(\theta+\varphi)\sin\frac{1}{2}(\theta-\varphi)\\
\cos\theta + \cos\varphi &= 2\cos\frac{1}{2}(\theta+\varphi)\cos\frac{1}{2}(\theta-\varphi)\\
\cos\theta - \cos\varphi &= -2\sin\frac{1}{2}(\theta+\varphi)\sin\frac{1}{2}(\theta-\varphi),
\end{aligned} \tag{C.9}$$

where the relevant parts are

$$\begin{aligned}
\sum_b \vec{w}^{(\lambda)}_{zby} &= (\cos[\gamma_z - 2\mu^{(\lambda)}_{0y}] + \cos[\gamma_z - 2\mu^{(\lambda)}_{1y}], 0, \sin[\gamma_z - 2\mu^{(\lambda)}_{0y}] + \sin[\gamma_z - 2\mu^{(\lambda)}_{1y}])\\
&= 2\cos[\mu^{(\lambda)}_{1y} - \mu^{(\lambda)}_{0y}](\cos[\gamma_z - \mu^{(\lambda)}_{0y} - \mu^{(\lambda)}_{1y}], 0, \sin[\gamma_z - \mu^{(\lambda)}_{0y} - \mu^{(\lambda)}_{1y}]),
\end{aligned} \tag{C.10}$$

and

$$\sum_b (-1)^b \vec{w}^{\lambda}_{zby} = 2\sin[\mu^{(\lambda)}_{1y} - \mu^{(\lambda)}_{0y}](-\sin[\gamma_z - \mu^{(\lambda)}_{0y} - \mu^{(\lambda)}_{1y}], 0, \cos[\gamma_z - \mu^{(\lambda)}_{0y} - \mu^{(\lambda)}_{1y}]). \tag{C.11}$$

In order to abbreviate the notation somewhat we will define $\sigma^{(\lambda)}_y = (\mu^{(\lambda)}_{0y} + \mu^{(\lambda)}_{1y})$, and $\omega^{(\lambda)}_y = (\mu^{(\lambda)}_{0y} - \mu^{(\lambda)}_{1y})$. Further, summing over i we have

$$\sum_i t_{ii} u_{x,i} v^{(\lambda)}_{y,i} = \sin 2\phi \cos\alpha_x \cos\beta^{(\lambda)}_y + \sin\alpha_x \sin\beta^{(\lambda)}_y, \tag{C.12}$$

and, consequently, using the trigonometric angle sum formulas the expectation value $\mathrm{Tr}[(A_x \otimes C_z)\psi^{\mathrm{post}}_\phi]$ can be written as (whenever $B^{(\lambda)}_{b|y}$ is a non-trivial projector)

$$\begin{aligned}
&\frac{1}{2}\sum_y p(\lambda)\Big(\cos 2\phi \sin\alpha_x \sin\omega^{(\lambda)}_y \sin[\beta^{(\lambda)}_y - \gamma_z + \sigma^{(\lambda)}_y] +\\
&\qquad \cos\omega^{(\lambda)}_y \cos[\beta^{(\lambda)}_y - \gamma_z + \sigma^{(\lambda)}_y](\sin 2\phi \cos\alpha_x \cos\beta^{(\lambda)}_y + \sin\alpha_x \sin\beta^{(\lambda)}_y)\Big).
\end{aligned} \tag{C.13}$$

In the case that $B^{(\lambda)}_{b|y}$ is an identity projection, one instead get

$$\frac{1}{2}\sum_\lambda p(\lambda)\Big(\sin 2\phi \cos\alpha_x \cos\gamma_z + \sin\alpha_x \sin\gamma_z\Big). \tag{C.14}$$

Where we have set the corresponding unitary to be trivial. In the end, we can therefore write the joint expectation value between Alice and Charlie $\mathrm{Tr}[(A_x \otimes C_z)\psi^{\mathrm{post}}_\phi]$ as

$$
\begin{aligned}
&\frac{1}{2}\cos 2\phi \sin\alpha_x \sin\omega_0^{(0)} \sin[\beta_0^{(0)} - \gamma_z + \sigma_0^{(0)}]p(0)+\\
&\frac{1}{2}\cos\omega_0^{(0)} \cos[\beta_0^{(0)} - \gamma_z + \sigma_0^{(0)}](\sin 2\phi \cos\alpha_x \cos\beta_0^{(0)} + \sin\alpha_x \sin\beta_0^{(0)})p(0)+\\
&\frac{1}{2}\cos 2\varphi \sin\alpha_x \sin\omega_1^{(1)} \sin[\beta_1^{(1)} - \gamma_z + \sigma_1^{(1)}]+\\
&\frac{1}{2}\cos\omega_1^{(1)} \cos[\beta_1^{(1)} - \gamma_z + \sigma_1^{(1)}](\sin 2\phi \cos\alpha_x \cos\beta_1^{(1)} + \sin\alpha_x \sin\beta_1^{(1)})+\\
&\frac{1}{2}(\sin 2\phi \cos\alpha_x \cos\gamma_z + \sin\alpha_x \sin\gamma_z)(1 - p(0)).
\end{aligned} \tag{C.15}
$$

We can then retrieve S_{AC} from

$$
S_{AC} = \sum_{xz} (-1)^{xz} \mathrm{Tr}[(A_x \otimes C_z)\psi_\phi^{\mathrm{post}}]. \tag{C.16}
$$

which results in

$$
\begin{aligned}
\mathcal{S}_{AC} = \Big\{ &\sin\omega_0^{(0)} \cos 2\phi \Big(\sin\eta \cos\delta \sin\left[\beta_0^{(0)} - \gamma_0 + \sigma_0^{(0)}\right] + \cos\eta \sin\delta \sin\left[\beta_0^{(0)} - \gamma_1 + \sigma_0^{(0)}\right]\Big)+\\
&\cos\omega_0^{(0)} \sin\beta_0^{(0)} \Big(\sin\eta \cos\delta \cos\left[\beta_0^{(0)} - \gamma_0 + \sigma_0^{(0)}\right] + \cos\eta \sin\delta \cos\left[\beta_0^{(0)} - \gamma_1 + \sigma_0^{(0)}\right]\Big)+\\
&\cos\omega_0^{(0)} \cos\beta_0^{(0)} \sin 2\phi \Big(\cos\eta \cos\delta \cos\left[\beta_0^{(0)} - \gamma_0 + \sigma_0^{(0)}\right] - \sin\eta \sin\delta \cos\left[\beta_0^{(0)} - \gamma_1 + \sigma_0^{(0)}\right]\Big)\Big\}p(0)+\\
&\sin\omega_1^{(1)} \cos 2\phi \Big(\sin\eta \cos\delta \sin\left[\beta_1^{(1)} - \gamma_0 + \sigma_1^{(1)}\right] + \cos\eta \sin\delta \sin\left[\beta_1^{(1)} - \gamma_1 + \sigma_1^{(1)}\right]\Big)+\\
&\cos\omega_1^{(1)} \sin\beta_1^{(1)} \Big(\sin\eta \cos\delta \cos\left[\beta_1^{(1)} - \gamma_0 + \sigma_1^{(1)}\right] + \cos\eta \sin\delta \cos\left[\beta_1^{(1)} - \gamma_1 + \sigma_1^{(1)}\right]\Big)+\\
&\cos\omega_1^{(1)} \cos\beta_1^{(1)} \sin 2\phi \Big(\cos\eta \cos\delta \cos\left[\beta_1^{(1)} - \gamma_0 + \sigma_1^{(1)}\right] - \sin\eta \sin\delta \cos\left[\beta_1^{(1)} - \gamma_1 + \sigma_1^{(1)}\right]\Big)+\\
&\Big(\sin 2\phi \Big(\cos\eta \cos\delta \cos\gamma_0 - \sin\eta \sin\delta \cos\gamma_1\Big) + \sin\eta \cos\delta \sin\gamma_0 + \cos\eta \sin\delta \sin\gamma_1\Big)\Big(1 - p(0)\Big).
\end{aligned} \tag{C.17}
$$

D. Explicit measurement strategy resulting in three sequential CHSH violations

As it turns out, it is possible to have three sequential CHSH violations using projective measurements and local randomness. In this extended sequential scenario, Charlie further relays her post-measurement state to yet another observer, Dave, whom also performs a CHSH test with Alice, resulting in the CHSH parameter $\mathcal{S}_{AD}$. The aim is then for all pairs of measurement parties to simultaneously violate the CHSH inequality $(\mathcal{S}_{AB}, \mathcal{S}_{AC}, \mathcal{S}_{AD}) > (2, 2, 2)$. We prove that this is possible using an explicit measurement strategy.

Let Alice and Bob share a pair of maximally entangled qubits. On her part of the state, Alice measures the observables $A_x = \cos[(2e(3+e)/(9e-5)]\sigma_1 + (-1)^x \sin[(2e(3+e)/(9e-5)]$. Just as before, Bob chooses between two measurement strategies, labeled by $\lambda_B \in \{0,1\}$. Then, with probability $p(\lambda_B = 0) = 0.2575$, Bob measures the observables $B_0^{(0)} = \cos(7999\pi/15\,422)\sigma_1 + \cos(7999\pi/15\,422)\sigma_3$ and $B_1^{(0)} = \cos(9959\pi/19\,868)\sigma_1 + \cos(9959\pi/19\,868)\sigma_3$, after which he performs the unitary transformations $U_0^{(0)} = \mathbb{1}$ and $U_1^{(0)} = \exp[i(9e/3850)^{3/4})\sigma_2]$, respectively, on the post-measurement state. On the other hand, if $\lambda_B = 1$, he instead measures $B_0^{(1)} = \mathbb{1}$ and $B_1^{(1)} = B_1^{(0)}$, with probability $p(\lambda_B = 1) = 1 - 0.2575$, and subsequently performs the unitaries $U_0^{(1)} = \exp[i(209e/40\,083)\sigma_2]$ and $U_1^{(1)} = U_1^{(0)}$. In the same way, Charlie also leverages local randomness and chooses between two sets of measurements, labeled by $\lambda_C \in \{0,1\}$, according to $p(\lambda_C)$. Then, with probability $p(\lambda_C = 0) = 0.4155$, she measures the observables $C_0^{(0)} = \cos(20\,569\pi/37\,125)\sigma_1 + \sin(20\,569\pi/37\,125)\sigma_3$ and $C_1^{(0)} = \cos(46\,979\pi/91\,826)\sigma_1 + \sin(46\,979\pi/91\,826)\sigma_3$, and likewise performs unitary transformations $V_0^{(0)} = \mathbb{1}$ and $V_1^{(0)} = \exp[i(9833\pi/459\ 698)\sigma_2]$. Further, with probability $p(\lambda_C = 1) = 1 - 0.4155$, she instead measures $C_0^{(1)} = \mathbb{1}$ and $C_1^{(1)} = C_1^{(0)}$, with the corresponding post-measurement unitaries being $V_0^{(1)} = \exp[i(5813\pi/485\ 823)\sigma_2]$ and $V_1^{(1)} = V_1^{(0)}$. Lastly, Dave measures the two observables $D_0 = \cos(33\ 793\pi/34\ 132)\sigma_1 + \sin(33\ 793\pi/34\ 132)\sigma_3$ and $D_1 = \cos(33\ 243\pi/62\ 008)\sigma_1 + \sin(33\ 243\pi/62\ 008)\sigma_3$. These measurements yields the triple violation $(\mathcal{S}_{AB}, \mathcal{S}_{AC}, \mathcal{S}_{AD}) \approx (2.0001, 2.0001, 2.0002)$.

E. Optimal definition of the post-measurement unitary when recycling entangled pairs of qutrits

As argued in the main text, the optimal post-measurement unitary should but add a constant shift to the measurement phase when acting on a measurement operator. This ensures that we remain in the same measurement sub-space. Since the measurement operators are all expanded in the symmetric and anti-symmetric Gell-Mann matrices, the unitaries should be expressed in terms of the diagonal Gell-Mann matrices. Thus, the most general post-measurement unitary should be on the form

$$U(\vec{\varphi}) = \exp\left[\frac{i}{2}(\varphi_0 \Lambda_{\mathrm{d}}^0 + \varphi_1 \Lambda_{\mathrm{d}}^1)\right], \tag{E.1}$$

for some $\varphi_0, \varphi_1 \in \mathbb{R}$. Note that Λ_{d}^0 and Λ_{d}^1 commute since they are both diagonal. We then set out to compute the action of this unitary on the measurement operators. For this we need the explicit form of the Gell-Mann matrices. We have three symmetric

$$\Lambda_{\mathrm{s}}^{01} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_{\mathrm{s}}^{02} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \Lambda_{\mathrm{s}}^{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{E.2}$$

three anti-symmetric

$$\Lambda_{\mathrm{as}}^{01} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_{\mathrm{as}}^{02} = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \Lambda_{\mathrm{as}}^{12} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \tag{E.3}$$

and two diagonal

$$\Lambda_{\mathrm{d}}^{0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Lambda_{\mathrm{d}}^{1} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{E.4}$$

The action of (E.1) on the symmetric and anti-symmetric Gell-Mann matrices is straightforwardly computed using the above matrix representations together with the Baker-Hausdorff Lemma

$$e^{i\lambda A} B e^{-i\lambda A} = B + i\lambda[A,B] + \frac{(i\lambda)^2}{2!}[A,[A,B]] + \frac{(i\lambda)^3}{3!}[A,[A,[A,B]]] + \cdots, \tag{E.5}$$

which holds for any two linear operators A and B. We can combine this Lemma together with the Taylor expansions of $\sin x$ and $\cos x$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \tag{E.6}$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \tag{E.7}$$

to arrive at the rotations generated by Λ_{d}^0 on the symmetric and anti-symmetric Gell-Mann matrices, respectively, leading to

$$\begin{aligned}
e^{i\varphi\Lambda_{\mathrm{d}}^0/2}\Lambda_{\mathrm{s}}^{01}e^{-i\varphi\Lambda_{\mathrm{d}}^0/2} &= \Lambda_{\mathrm{s}}^{01}\cos\varphi - \Lambda_{\mathrm{as}}^{01}\sin\varphi \\
e^{i\varphi\Lambda_{\mathrm{d}}^0/2}\Lambda_{\mathrm{as}}^{01}e^{-i\varphi\Lambda_{\mathrm{d}}^0/2} &= \Lambda_{\mathrm{as}}^{01}\cos\varphi + \Lambda_{\mathrm{s}}^{01}\sin\varphi \\
e^{i\varphi\Lambda_{\mathrm{d}}^0/2}\Lambda_{\mathrm{s}}^{02}e^{-i\varphi\Lambda_{\mathrm{d}}^0/2} &= \Lambda_{\mathrm{s}}^{02}\cos\frac{\varphi}{2} - \Lambda_{\mathrm{as}}^{02}\sin\frac{\varphi}{2} \\
e^{i\varphi\Lambda_{\mathrm{d}}^0/2}\Lambda_{\mathrm{as}}^{02}e^{-i\varphi\Lambda_{\mathrm{d}}^0/2} &= \Lambda_{\mathrm{as}}^{02}\cos\frac{\varphi}{2} + \Lambda_{\mathrm{s}}^{02}\sin\frac{\varphi}{2} \\
e^{i\varphi\Lambda_{\mathrm{d}}^0/2}\Lambda_{\mathrm{s}}^{12}e^{-i\varphi\Lambda_{\mathrm{d}}^0/2} &= \Lambda_{\mathrm{s}}^{12}\cos\frac{\varphi}{2} + \Lambda_{\mathrm{as}}^{12}\sin\frac{\varphi}{2} \\
e^{i\varphi\Lambda_{\mathrm{d}}^0/2}\Lambda_{a}^{12}e^{-i\varphi\Lambda_{\mathrm{d}}^0/2} &= \Lambda_{\mathrm{as}}^{12}\cos\frac{\varphi}{2} - \Lambda_{\mathrm{s}}^{12}\sin\frac{\varphi}{2},
\end{aligned} \tag{E.8}$$

and similarly for Λ_{d}^1

$$\begin{aligned}
e^{i\varphi\Lambda_{\mathrm{d}}^1/2}\Lambda_{\mathrm{s}}^{12}e^{-i\varphi\Lambda_{\mathrm{d}}^1/2} &= \Lambda_{\mathrm{s}}^{01} \\
e^{i\varphi\Lambda_{\mathrm{d}}^1/2}\Lambda_{\mathrm{as}}^{12}e^{-i\varphi\Lambda_{\mathrm{d}}^1/2} &= \Lambda_{a}^{01} \\
e^{i\varphi\Lambda_{\mathrm{d}}^1/2}\Lambda_{\mathrm{s}}^{02}e^{-i\varphi\Lambda_{\mathrm{d}}^1/2} &= \Lambda_{\mathrm{s}}^{02}\cos\frac{\sqrt{3}\varphi}{2} - \Lambda_{\mathrm{as}}^{02}\sin\frac{\sqrt{3}\varphi}{2} \\
e^{i\varphi\Lambda_{\mathrm{d}}^1/2}\Lambda_{\mathrm{as}}^{13}e^{-i\varphi\Lambda_{\mathrm{d}}^1/2} &= \Lambda_{\mathrm{as}}^{02}\cos\frac{\sqrt{3}\varphi}{2} + \Lambda_{\mathrm{s}}^{02}\sin\frac{\sqrt{3}\varphi}{2} \\
e^{i\varphi\Lambda_{\mathrm{d}}^1/2}\Lambda_{\mathrm{s}}^{23}e^{-i\varphi\Lambda_{\mathrm{d}}^1/2} &= \Lambda_{\mathrm{s}}^{12}\cos\frac{\sqrt{3}\varphi}{2} - \Lambda_{\mathrm{as}}^{12}\sin\frac{\sqrt{3}\varphi}{2} \\
e^{i\varphi\Lambda_{\mathrm{d}}^1/2}\Lambda_{\mathrm{as}}^{12}e^{-i\varphi\Lambda_{\mathrm{d}}^1/2} &= \Lambda_{\mathrm{as}}^{12}\cos\frac{\sqrt{3}\varphi}{2} + \Lambda_{s}^{12}\sin\frac{\sqrt{3}\varphi}{2}.
\end{aligned} \tag{E.9}$$

Since Λ_{d}^0 and Λ_{d}^1 commute we can combine (E.8) and (E.9) in order to compute the action of (E.1) on the symmetric and anti-symmetric Gell-Mann matrices. This then leads to the following relations

$$\begin{aligned}
e^{i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2}\Lambda_{\mathrm{s}}^{01}e^{-i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2} &= \Lambda_{\mathrm{s}}^{01}\cos\varphi_0 - \Lambda_{\mathrm{as}}^{01}\sin\varphi_0 \\
e^{i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2}\Lambda_{\mathrm{as}}^{01}e^{-i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2} &= \Lambda_{\mathrm{as}}^{01}\cos\varphi_0 + \Lambda_{\mathrm{s}}^{01}\sin\varphi_0 \\
e^{i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2}\Lambda_{\mathrm{s}}^{02}e^{-i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2} &= \Lambda_{\mathrm{s}}^{02}\cos\left(\frac{\varphi_0}{2} + \frac{\sqrt{3}\varphi_1}{2}\right) - \Lambda_{\mathrm{as}}^{02}\sin\left(\frac{\varphi_0}{2} + \frac{\sqrt{3}\varphi_1}{2}\right) \\
e^{i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2}\Lambda_{\mathrm{as}}^{02}e^{-i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2} &= \Lambda_{\mathrm{as}}^{02}\cos\left(\frac{\varphi_0}{2} + \frac{\sqrt{3}\varphi_1}{2}\right) + \Lambda_{\mathrm{s}}^{02}\sin\left(\frac{\varphi_0}{2} + \frac{\sqrt{3}\varphi_1}{2}\right) \\
e^{i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2}\Lambda_{\mathrm{s}}^{12}e^{-i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2} &= \Lambda_{\mathrm{s}}^{12}\cos\left(\frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\right) + \Lambda_{\mathrm{as}}^{12}\sin\left(\frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\right) \\
e^{i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2}\Lambda_{\mathrm{as}}^{12}e^{-i\vec{\varphi}\cdot\vec{\Lambda}_{\mathrm{d}}/2} &= \Lambda_{\mathrm{as}}^{12}\cos\left(\frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\right) - \Lambda_{\mathrm{s}}^{12}\sin\left(\frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\right),
\end{aligned} \tag{E.10}$$

from which we then get the result

$$
\begin{aligned}
U^\dagger(\vec{\varphi})C_{c|z}U(\vec{\varphi}) = \frac{1}{d}\Big(&\mathbb{1} + \cos\phi^C_{cz}\Big[U^\dagger\Lambda^{12}_{\rm s}U + U^\dagger\Lambda^{23}_{\rm s}U\Big] + \sin\phi^C_{cz}\Big[U^\dagger\Lambda^{12}_{\rm as}U + U^\dagger\Lambda^{23}_{\rm as}U\Big] \\
&+ \cos 2\phi^C_{cz}\Big[U^\dagger\Lambda^{13}_{\rm s}U\Big] + \sin 2\phi^C_{cz}\Big[U^\dagger\Lambda^{13}_{\rm as}U\Big]\Big) \\
=& \frac{1}{d}\Big(\mathbb{1} + \Lambda^{12}_{\rm s}\cos[\phi^C_{cz} - \varphi_0] + \Lambda^{12}_{\rm as}\sin[\phi^C_{cz} - \varphi_0] \\
&+ \Lambda^{23}_{\rm s}\cos\Big[\phi^C_{cz} + \frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\Big] + \Lambda^{23}_{\rm as}\sin\Big[\phi^C_{cz} + \frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\Big] \\
&+ \Lambda^{13}_{\rm s}\cos\Big[2\phi^C_{cz} - \frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\Big] + \Lambda^{13}_{\rm as}\sin\Big[2\phi^C_{cz} - \frac{\varphi_0}{2} - \frac{\sqrt{3}\varphi_1}{2}\Big]\Big).
\end{aligned} \tag{E.11}
$$

If we then set $\varphi_1 = \sqrt{3}\varphi_0 = \varphi$ we find that the action of $U(\vec{\varphi})$ on $C_{c|z}$ results in $\phi^C_{cz} \to \phi^C_{cz} - \varphi$, as desired. Furthermore, since

$$
J_3 = \frac{1}{2}\Lambda^0_{\rm d} + \frac{\sqrt{3}}{2}\Lambda^1_{\rm d}, \tag{E.12}
$$

the optimal definition of the post measurement unitary is on the form $U = \exp[i\varphi J_3]$, which was to be shown.

F. Derivation of correlations when recycling pairs of maximally entangled qutrits

We wish to compute the correlations between Alice-Bob and Alice-Charlie when y and λ corresponds to non-trivial measurements. For notational simplicity we will omit the λ label in the measurements of Bob since it is always assumed that his measurements are non-trivial. We can express the maximally entangled qutrit state in terms of the GGMB (with $d = 3$) as

$$
\phi^+ = \frac{1}{9}\mathbb{1}\otimes\mathbb{1} + \frac{1}{6}\Lambda, \tag{F.1}
$$

where

$$
\Lambda := \sum_{m<n}^{2}\Lambda^{mn}_{\rm s}\otimes\Lambda^{mn}_{\rm s} - \sum_{m<n}^{2}\Lambda^{mn}_{\rm as}\otimes\Lambda^{mn}_{\rm as} + \sum_{m=0}^{1}\Lambda^{m}_{\rm d}\otimes\Lambda^{m}_{\rm d}. \tag{F.2}
$$

In order to compute the correlations $p(a,b|x,y) = \mathrm{Tr}[(A_{a|x}\otimes B_{b|y})\phi^+]$ we first expand $A_{a|x}\otimes B_{b|y}$ using the measurements presented in the main text, this leads to the following expression

$$
\begin{aligned}
&\frac{1}{9}\Big(\mathbb{1}\otimes\mathbb{1} + \sum_{k<l}\cos[(k-l)\phi^B_{by}]\mathbb{1}\otimes\Lambda^{kl}_{\rm s} + \sum_{k<l}\sin[(k-l)\phi^B_{by}]\mathbb{1}\otimes\Lambda^{kl}_{\rm as} \\
&+ \sum_{i<j}\cos[(j-i)\phi^A_{ax}]\Lambda^{ij}_{\rm s}\otimes\mathbb{1} + \sum_{i<j}\sum_{k<l}\cos[(j-i)\phi^A_{ax}]\cos[(l-k)\phi^B_{by}]\Lambda^{ij}_{\rm s}\otimes\Lambda^{kl}_{\rm s} \\
&+ \sum_{i<j}\sum_{k<l}\cos[(j-i)\phi^A_{ax}]\sin[(l-k)\phi^B_{by}]\Lambda^{ij}_{\rm s}\otimes\Lambda^{kl}_{\rm as} \\
&+ \sum_{i<j}\sin[(j-i)\phi^A_{ax}]\Lambda^{ij}_{\rm as}\otimes\mathbb{1} + \sum_{i<j}\sum_{k<l}\sin[(j-i)\phi^A_{ax}]\cos[(l-k)\phi^B_{by}]\Lambda^{ij}_{\rm as}\otimes\Lambda^{kl}_{\rm s} \\
&+ \sum_{i<j}\sum_{k<l}\sin[(j-i)\phi^A_{ax}]\sin[(l-k)\phi^B_{by}]\Lambda^{ij}_{\rm as}\otimes\Lambda^{kl}_{\rm as}\Big).
\end{aligned} \tag{F.3}
$$

We then treat the terms separately. The $\mathbb{1}\otimes\mathbb{1}$ in ϕ^+ will only contribute with a constant term equal to 1/9. Furthermore, we have that

$$\begin{aligned}\sum_{m<n}\mathrm{Tr}([A_{a|x}\otimes B_{b|y}]\Lambda_{\mathrm{s}}^{mn}\otimes\Lambda_{\mathrm{s}}^{mn}) &= \sum_{m<n}\sum_{i<j}\sum_{k<l}\cos[(j-i)\phi_{ax}^A]\cos[(l-k)\phi_{by}^B]\mathrm{Tr}\Big(\Lambda_{\mathrm{s}}^{mn}\Lambda_{\mathrm{s}}^{ij}\otimes\Lambda_{\mathrm{s}}^{mn}\Lambda_{\mathrm{s}}^{kl}\Big)\\ &= 4\sum_{m<n}\sum_{i<j}\sum_{k<l}\cos[(j-i)\phi_{ax}^A]\cos[(l-k)\phi_{by}^B]\delta^{mi}\delta^{nj}\delta^{mk}\delta^{nl}\\ &= 4\sum_{m<n}\cos[(n-m)\phi_{ax}^A]\cos[(n-m)\phi_{by}^B],\end{aligned} \tag{F.4}$$

where we have made use of the orthogonality relations provided in the main text. In the same way we have that

$$\sum_{m<n}\mathrm{Tr}([A_{a|x}\otimes B_{b|y}]\Lambda_{\mathrm{as}}^{mn}\otimes\Lambda_{\mathrm{as}}^{mn}) = 4\sum_{m<n}\sin[(n-m)\phi_{ax}^A]\sin[(n-m)\phi_{by}^B]. \tag{F.5}$$

The diagonal terms in ϕ^+ will not contribute. Then, using the angle sum relations $\cos(\alpha\pm\beta)=\cos\alpha\cos\beta\mp\sin\beta\sin\alpha$ and $\sin(\alpha\pm\beta)=\cos\alpha\sin\beta\pm\cos\beta\sin\alpha$ we can add (F.4) and (F.5), leading to the final result

$$p(a,b|x,y)=\frac{1}{9}+\frac{2}{27}\Big(\cos[\phi_{ax}^A+\phi_{by}^B]+\cos[2(\phi_{ax}^A+\phi_{by}^B)]\Big). \tag{F.6}$$

Moving on to the correlations between Alice and Charlie. Their shared state is on the form

$$\begin{aligned}\rho^{\mathrm{post}} =& \frac{1}{18}\sum_{by}\mathbb{1}\otimes U_{by}B_{b|y}U_{by}^\dagger+\frac{1}{12}\sum_{by}\Big(\sum_{i<j}\Lambda_{\mathrm{s}}^{ij}\otimes U_{by}B_{b|y}\Lambda_{\mathrm{s}}^{ij}B_{b|y}U_{by}^\dagger\\ &-\sum_{i<j}\Lambda_{\mathrm{as}}^{ij}\otimes U_{by}B_{b|y}\Lambda_{\mathrm{as}}^{ij}B_{b|y}U_{by}^\dagger+\sum_{m=0}^{1}\Lambda_{\mathrm{d}}^{m}\otimes U_{by}B_{b|y}\Lambda_{\mathrm{d}}^{m}B_{b|y}U_{by}^\dagger\Big).\end{aligned} \tag{F.7}$$

The contribution from the first term above is

$$\begin{aligned}\frac{1}{18}\sum_{by}\mathrm{Tr}[(A_{a|x}\otimes C_{c|z})\mathbb{1}\otimes U_{by}B_{b|y}U_{by}^\dagger] &= \frac{1}{18}\sum_{by}\mathrm{Tr}[U_{by}^\dagger C_{c|z}U_{by}B_{b|y}]\\ &= \frac{1}{81}\sum_{by}\Big(\frac{3}{2}+\cos[\phi_{by}^B-\phi_{cz}^C+\mu_{by}]\\ &\quad+\cos[2(\phi_{by}^B-\phi_{cz}^C+\mu_{by})]\Big),\end{aligned} \tag{F.8}$$

where we have used the definition of U_{by} arrived at in Appendix E. The contribution from the second term in ρ^{post} is

$$\frac{1}{12}\mathrm{Tr}[(A_{a|x}\otimes C_{c|z})\sum_{i<j}\Lambda_{\mathrm{s}}^{ij}\otimes U_{by}B_{b|y}\Lambda_{\mathrm{s}}^{ij}B_{b|y}U_{by}^\dagger]=\frac{1}{18}\sum_{i<j}\cos[(j-i)\phi_{ax}^A]\mathrm{Tr}[U_{by}^\dagger C_{c|z}U_{by}B_{b|y}\Lambda_{\mathrm{s}}^{ij}B_{b|y}]. \tag{F.9}$$

To continue, observe that

$$B_{b|y}\Lambda_{\mathrm{s}}^{ij}B_{b|y}=\mathrm{Tr}(B_{b|y}\Lambda_{\mathrm{s}}^{ij})B_{b|y}=\frac{2}{3}\sum_{i<j}\cos[(j-i)\phi_{by}^B]B_{b|y}, \tag{F.10}$$

therefore, we can write (F.9) as

$$\frac{1}{27}\sum_{i<j}\cos[(j-i)\phi^A_{ax}]\cos[(j-i)\phi^B_{by}]\mathrm{Tr}\Big(U^\dagger_{by}C_{c|z}U_{by}B_{b|y}\Big), \tag{F.11}$$

where $\mathrm{Tr}[U^\dagger_{by}C_{c|z}U_{by}B_{b|y}]$ was computed previously in (F.8). In the same way we compute the contribution from the third term in ρ^{post} to be

$$-\frac{1}{27}\sum_{i<j}\sin[(j-i)\phi^A_{ax}]\sin[(j-i)\phi^B_{by}]\mathrm{Tr}\Big(U^\dagger_{by}C_{c|z}U_{by}B_{b|y}\Big). \tag{F.12}$$

The joint contribution from the second and third term in ρ^{post} can then compactly be written as

$$\frac{1}{27}\sum_{k=1}^{2}\cos[k(\phi^A_{ax}+\phi^B_{by})]\mathrm{Tr}[U^\dagger_{by}C_{c|z}U_{by}B_{b|y}]. \tag{F.13}$$

The last term in ρ^{post} will not enter into $p(a,c|x,z)$ since it features the diagonal Gell-Mann matrices. Hence, in the end, we therefore get that

$$p(a,c|x,z)=\sum_{by}\left(\frac{1}{27}\sum_{k=1}^{2}\cos[k(\phi^A_{ax}+\phi^B_{by})]\mathrm{Tr}[U^\dagger_{by}C_{c|z}U_{by}B_{b|y}]+\frac{1}{18}\mathrm{Tr}[U^\dagger_{by}C_{c|z}U_{by}B_{b|y}]\right), \tag{F.14}$$

where

$$\mathrm{Tr}[U^\dagger_{by}C_{c|z}U_{by}B_{b|y}]=\frac{1}{9}\Big(3+2\sum_{k=1}^{2}\cos[k(\phi^B_{by}-\phi^C_{cz}+\mu_{by})]\Big), \tag{F.15}$$

which is what we wanted to arrive at.