Quantum enhancement to information speed

Sebastian Horvat*
Faculty of Physics, University of Vienna, Austria and
Department of Physics, Faculty of Science, University of Zagreb, Croatia

Borivoje Dakic†
Vienna Center for Quantum Science and Technology (VCQ),
Faculty of Physics, University of Vienna, Austria and
Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, Vienna, Austria
(Dated: November 28, 2019)

The speed of the transmission of a physical signal from a sender to a receiver is limited by the speed of light, regardless of the physical system being classical or quantum. In this sense, quantum mechanics can not provide any enhancement of the speed of information. If instead we consider that the information needing to be transmitted is not localized at the sender’s location, but dispersed throughout space, spatial coherence might provide an enhancement of the information speed. In this work, we demonstrate a quantum mechanical advantage in the speed of acquirement and transmission of information globally encoded in space. Our findings can naturally be applied in situations where the information source has limited power, i.e. bounded number of signals that can be sent per unit time. We conclude by showing a significant advantage when using single-photon (or Fock) states with respect to coherent states for communication purposes.

INTRODUCTION

Usually, when one talks about information speed, one envisions two parties A and B, where A aims to communicate some message to B; A then sends the message encoded in a physical system (signal or information carrier) to B and the information speed is simply the speed of the signal which is limited by the speed of light. In this sense, quantum mechanics cannot provide any enhancement to the information speed. However, what if the information that needs to be transmitted is not localized at the sender’s station, as it was at A in the given example? What if the information of interest is encoded in a global property of dispersed pieces of information, each localized at a different location? In this case, if we define information speed as a quantity inversely proportional to the time needed to acquire and transmit some generally global information, quantum mechanics may provide some advantage with respect to classical theory.

In this paper we show that preparing information carriers in spatial superposition provides an arbitrarily high speed up of an information theoretic task involving the acquisition and transmission of globally encoded information. In order to formally address the subject matter we first describe the scenario of interest and introduce the auxiliary notion of k-way signaling behaviors within a device-independent formalism [1]. We then proceed by proving that a single quantum particle in spatial superposition drastically outperforms classical particles at collecting and transmitting delocalized information. This is shown by the violation of a specific inequality which poses sharp bounds on the performance of k-way signaling processes. Our findings have a natural application in scenarios where information sources have limited power. Finally, we address our protocol in the context of bounded energy and communication with light, and compare the information speed achievable by using a source that creates coherent light versus one that creates single photons. Specifically, we show a significant advantage of using single photons for a task involving information globally encoded in two locations. The obtained quantum enhancement is based solely on the quantum superposition principle. Our result can be seen in the light of recent developments that put forward quantum superposition as a genuine resource for information processing, such as in two-way communication with one particle [2], quantum acausal processes [3], superposition of orders [4] and directions [5], quantum combs [6], quantum switch [7] and quantum causal models [8]. Some of these novel phenomena have been demonstrated in recent experiments [9–11].

ACQUIREMENT AND TRANSMISSION OF DELOCALIZED INFORMATION

The scenario of interest consists of one party, whom we’ll refer to as Alice, and N pieces of information \( \{x_1, x_2, ..., x_N\} \) dispersed at N different locations, as pictured in Figure 1. Alice is connected to each of the N locations with communication channels which enable a bidirectional transmission of information. Her goal is to learn some global property \( \alpha = \alpha(x_1, ..., x_N) \) as a function of spatially dispersed information pieces \( x_i \). For simplicity we assume that the N locations containing the local information are not mutually connected by any communication channel, i.e. the information does not flow in between the different regions. This restriction can be understood as one forcing the pieces of information to be truly isolated/localized and removing the dependence on the geometry of the problem.

One way for Alice to learn \( \alpha(x_1, ..., x_N) \) is to send \( N \) signals which shall encode the information pieces \( x_i \), retrieve them and calculate the function \( \alpha \). We denote the time she needs to complete this action (acquisition and transmission of information pieces) as \( \tau \) and call it unit time. Suppose

---

* sebastian.horvat@univie.ac.at
† borivoje.dakic@univie.ac.at
now that Alice has limited resources to complete the task, i.e. she has power to send at most $k$ signals per unit time, where $k < N$. This restriction can come from some natural assumption, such as the limited power of the source of information carriers. For example, Alice sends light signals via a source of limited power (average energy per unit time). Upon receiving back the signals, she decodes the message and produces a classical output $a$.

In general, the information pieces are randomly sampled from some distribution; the process is thus mathematically fully characterized by the following set of conditional probabilities, or behavior

$$
\{ P(a|x_1, ..., x_N) ; \forall a \in O ; x_1, ..., x_N \in I \},
$$

where $I$ and $O$ denote the input and output alphabets pertaining respectively to $\{x_1, ..., x_N\}$ and to $a$. $P(a|x_1, ..., x_N)$ thus denotes the probability that Alice produces the output $a$ conditioned on the dispersed information being $\{x_1, ..., x_N\}$. For a moment, we shall forget that $A$ is using $k$ classical signals, and define the more general notion of $k$-way signaling, which we introduce in a device-independent manner in what follows.

**Definition.** A behavior $\{ P(a|x_1, ..., x_N) ; \forall a, x_i \}$ is said to be $k$-way signaling iff there exists a set of weights $(q_{j_1, ..., j_k}, \forall j_1, ..., j_k)$ and a set of probability distributions $\{ P(a|x_{j_1}, ..., x_{j_k}) ; \forall j_1, ..., j_k \}$ such that the following is satisfied:

\[
P(a|x_1, ..., x_N) = \sum_{j_1, ..., j_k} q_{j_1, ..., j_k} P(a|x_{j_1}, ..., x_{j_k});
\]

\[
\sum_{j_1, ..., j_k} q_{j_1, ..., j_k} = 1; \quad q_{j_1, ..., j_k} \geq 0, \quad \forall j_1, ..., j_k,
\]

where the domain of the indices $\{j_1, ..., j_k\}$ ranges over all $\binom{N}{k}$ subsets of the $N$ locations.

The intuition behind the latter definition is the following: if the system exhibits $k$-way signaling, it means that its behavior can be modeled by Alice choosing to communicate with locations pertaining to $\{x_{j_1}, ..., x_{j_k}\}$ with probability $q_{j_1, ..., j_k}$. For example, for $N = 3$, a two-way signaling distribution can be decomposed as

\[
P(a|x_1, x_2, x_3) = q_{12} P(a|x_1, x_2) + q_{13} P(a|x_1, x_3) + q_{23} P(a|x_2, x_3);
\]

\[
\sum_{i < j} q_{ij} = 1; \quad q_{ij} \geq 0, \forall i < j,
\]

where $q_{ij}$ denotes the probability of Alice communicating with locations pertaining to $x_i$ and $x_j$. The definition above will help us to put strict bounds on Alice’s performance when using $k$ signals per unit time.

**GENUINE N-WAY SIGNALING**

The set of all $k$-way signaling behaviors forms a polytope when embedded in a real vector space, since it is closed under probabilistic mixing (a simple proof is given in Appendix A). Thus, one can characterize this set via necessary and sufficient conditions in form of facet inequalities. These and similar methods have been applied e.g. in the investigation of Bell’s inequalities [12] and in causal modeling [13, 14].

Providing the full characterization of a polytope is hard in general. Instead, here we focus on a particular inequality that can be used to bound the performance of any $k$-way signaling distribution. We focus our attention to the case of binary inputs and outputs, i.e. $a, x_i \in \{0, 1\}$.

Any $k$-way signaling behavior, where $k < N$, satisfies the following inequality:

\[
B \equiv -P(1|0, 0, ..., 0) + P(1|1, 0, ..., 0) + P(1|0, 1, ..., 0) + ... + P(1|0, 0, ..., 1) \leq N - 1,
\]

(3)

To see that this is the case, notice that any $(N-1)$-way signaling behavior can be expressed as a convex sum of processes which leave out one location from the communication. If the $i$-th location is left out, then

\[
P(a|0, 0, ..., x_i = 0, ..., 0) = P(a|0, 0, ..., x_i = 1, ..., 0),
\]

(4)

so the first negative term in $B$ cancels at least one of the positive terms and leaves the maximum achievable value equal to $N - 1$. The analogous reasoning holds for $k < N - 1$. Thus, the violation of this inequality necessarily implies $N$-way signaling.

Note that our inequality represents the probability of successfully accomplishing a task involving information genuinely globally encoded in space. Namely, Alice is supposed to compute the function

\[
a(x_1, ..., x_N) = \begin{cases} 
0, & \text{if } x_i = 0, \quad \forall i, \\
1, & \text{if } x_j = 1 \text{ and } x_i = 0, \forall i \neq j, 
\end{cases}
\]

(5)

where the $N + 1$ settings are randomly sampled from a uniform distribution. Inequality (3) then implies that Alice’s probability to successfully compute the latter function is bounded by $P_{\text{bound}} = \frac{N}{N-1}$, unless her performance exhibits $N$-way signaling.

Suppose now that Alice possesses limited resources and sends $k$ classical signals per unit time $\tau$ thus achieving $k$-way signaling. In this case, it is clear that she can achieve at best $P_{\text{bound}}$ in a single shot experiment. In order to surpass this
threshold she needs to send her \( k \) signals at least \( \left\lceil \frac{N}{2} \right\rceil \) times before producing an output, thereby effectively exhibiting \( N \)-way signaling behavior.

**QUANTUM ENHANCED INFORMATION SPEED**

In the following we show the possibility of achieving the violation of inequality (3) for arbitrary \( N \) by using a single quantum particle (one signal per unit time) prepared in spatial superposition. We assume a simple model where at each location \( x_i \) the information is encoded into a crystal that applies a local phase shift \( e^{ix_i\phi} \) to the state of the particle, where \( \{\phi_i\} \) are fixed angles known to Alice. The protocol can be summed up as follows. Alice prepares her signal in a spatial superposition of trajectories directed towards the \( N \) locations; after interacting with the crystals, the wave packets are bouned back to Alice who performs a binary measurement thereby producing an output \( a \).

The initial wave function is:

\[
|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_n |n\rangle,
\]

where \( \{|n\rangle\} \) is the basis of spatial modes corresponding to the \( N \) trajectories directed towards their pertaining locations. After encoding, the wave function is transformed to

\[
|\psi\rangle_{x_1,\ldots,x_N} = \frac{1}{\sqrt{N}} \sum_n e^{i\phi_n x_n} |n\rangle.
\]

Upon getting back her signals, Alice performs a binary measurement defined by a general POVM \( \Pi \equiv \{\Pi_0, \Pi_1\} \), thereby producing an outcome \( a \in \{0, 1\} \).

Let’s denote the quantum state that arises via encoding when \( \{x_1 = 0, x_2 = 0, \ldots, x_N = 0\} \) with \( \rho_0 \), and the one that arises from encoding when \( \{x_1 = 0, x_2 = 0, \ldots, x_N = 0\} \) with \( \rho_1 \). Then, if we introduce the following averaged state

\[
\rho = \frac{1}{N} \sum_i \rho_i,
\]

the left hand side of inequality (3) can be rewritten as

\[
B = -1 + (N + 1) \left[ \frac{1}{N + 1} \text{Tr}(\Pi_0 \rho_0) + \frac{N}{N + 1} \text{Tr}(\Pi_1 \rho_1) \right]
\]

\[
= -1 + (N + 1) P_W.
\]

The expression \( P_W = \left[ \frac{1}{N + 1} \text{Tr}(\Pi_0 \rho_0) + \frac{N}{N + 1} \text{Tr}(\Pi_1 \rho_1) \right] \) is the probability of successfully distinguishing the quantum states \( \rho_0 \) and \( \rho_1 \) given their respective prior probabilities \( p_0 = \frac{1}{N + 1} \) and \( p_1 = \frac{N}{N + 1} \). It is known [15] that this probability is bounded by

\[
\max_{\Pi} P_W = \frac{1}{2} (1 + \|p_1 \rho_1 - p_0 \rho_0\|_1),
\]

where \( \|A\|_1 \) denotes the trace norm of \( A \).

The maximum achievable value of \( B \) for a given encoding scheme is then given by

\[
\max_B = -1 + \frac{N + 1}{2} (1 + \|p_1 \rho_1 - p_0 \rho_0\|_1) \equiv N - 1 + \delta,
\]

where \( \delta = \frac{1}{2} - \frac{N}{2} + \frac{N - 1}{2} \|p_1 \rho_1 - p_0 \rho_0\|_1 \) is the amount of violation.

Let us first analyze the case \( N = 2 \). We set the two crystals’ phases to \( \phi_{1,2} = \pi \). Alice is supposed to output \( a = 0 \) if both settings are equal to 0 and \( a = 1 \) if one of the settings is equal to 1. Clearly, states \( \rho_0 \) and \( \rho_1 \) are mutually orthogonal and thus perfectly distinguishable, thereby enabling Alice to saturate the logical bound of our inequality, i.e. \( \delta = 1 \) (the details are presented in Appendix B). In contrast, if she uses one classical signal per unit time, she needs double the time to achieve a violation. Therefore, spatial coherence doubles the information speed involved in completing the task.

The \( N \geq 3 \) case is much more complicated and the detailed analysis is presented in Appendix C. In order to optimize Alice’s performance we set \( \frac{\delta}{2} \) of the crystals’ phases to some angle \( \phi \) and the rest to \(-\phi\) and show that a clear violation \( \delta > 0 \) is achieved for \( \cos(\phi) > \frac{N(N-2)}{2(N-1)^2} \) (with a small correction for odd \( N \)). The numerical results of the violation are shown in Figure 2.

\[\text{Figure 2. Violation of the inequality. The left graph represents the quantum violation of inequality (3) as function of the number of information locations \( N \). The table on the right compares the value of the inequality achievable using classical particles (B_{CL}) and the bound achievable with quantum particles in spatial superposition (B_{QM}).}\]

We have therefore proved the possibility of genuine multi-way signaling with an arbitrary number of locations using one particle in spatial superposition within one unit of time \( \tau \). On the other hand, suppose that Alice possesses a source producing one particle with a defined trajectory per unit time: in this case, she cannot achieve the quantum performance for this task even in \((N - 1) \tau\) time, since she can communicate with maximally one location per unit time. Hence, spatial coherence as a resource provides an arbitrarily large enhancement of the information speed as defined by our task and inequality (3). Note that there is no conflict with special relativity, since the information carriers’ speed is limited by the speed of light, which is reflected in the time \( \tau \) that light needs to travel back and forth from Alice to the
crystals.

**INFORMATION SPEED WITH BOUNDED ENERGY**

In practice, the most common situation deals with sources of limited power (energy per unit time). When speaking of communication with light sources, typically we model (semi)classical resources as coherent pulses defined by coherent states $|\alpha\rangle$. The associated mean energy is $E = \hbar \omega |\alpha|^2$, where $\omega$ is the frequency of the used light source and $|\alpha|^2$ is the mean photon number. In contrast, a single-photon state is a described by a Fock state $|1\rangle$ which is usually associated with genuine quantum behavior. Nevertheless, for most practical purposes, a single photon state is approximated by an attenuated coherent state of $|\alpha|^2 \sim 1$ [16, 17]. This approximation is ubiquitous in almost all applications and quantum communication protocols. From this point of view, Fock states are as good as coherent states for all practical purposes. Our goal here is to show that this correspondence is not valid for our task, and that Fock states provide a significant enhancement of the information speed with respect to coherent states.

Suppose Alice possesses a light source of bounded power and wants to violate inequality (3) for the $N = 2$ case. As already mentioned, she can achieve the maximum violation $\delta = 1$ by using a Fock state. On the other hand, she could use coherent states as resources. In that case, the protocol proceeds as follows. Alice prepares a coherent state $|\alpha\rangle$ in one spatial mode and sends it on a beam splitter directing the state $|\alpha\rangle_1 |0\rangle_2 \rightarrow |\alpha/\sqrt{2}\rangle_1 |\alpha/\sqrt{2}\rangle_2$, (12) where the subscripts 1,2 denote the Fock spaces corresponding to the trajectories directed towards the two crystals. The state that returns to Alice after the interaction with the crystals is given by

$$|\psi_{x_1 x_2}\rangle = |\alpha e^{i\pi x_1}/\sqrt{2}\rangle_1 |\alpha e^{i\pi x_2}/\sqrt{2}\rangle_2.$$  (13)

The states $|\psi_{00}\rangle$ and $|\psi_{01,10}\rangle$ are not orthogonal and consequently not perfectly distinguishable. Their overlap is given by $|\langle \psi_{00}|\psi_{01,10}\rangle| = e^{-|\alpha|^2}$ which tends to zero for $|\alpha| \rightarrow \infty$. Therefore, one needs infinite energy per coherent pulse in order to achieve the maximal violation of our inequality, while using Fock states, it is possible to do it with single photons (single energy quanta).

However, the maximal violation using Fock states requires an overly idealized experimental procedure, i.e. perfect beam splitters and phase shifters set exactly to $\pi$. One of the most common practical challenges for distant communication is the phase stabilization of interferometers. Therefore, in order to make a fair comparison between the two schemes, let us assume the crystals’ phases to be perturbed by a small $\epsilon$, i.e. $\phi = \pi - \epsilon$, and see by how much one needs to increase the coherent source’s power in order to achieve the same performance as when using Fock states.

In this case, the state that returns to Alice after the interaction with the crystals is given by

$$|\psi_{x_1 x_2}\rangle = |\alpha e^{i(\pi-\epsilon)x_1}/\sqrt{2}\rangle_1 |\alpha e^{i(\pi-\epsilon)x_2}/\sqrt{2}\rangle_2.$$  (14)

The maximum violation can then be calculated using expression (11). The obtained bound is given as function of $|\alpha|$ and of the perturbation $\epsilon$, i.e. $\delta_{\text{Cod}}(|\alpha|, \epsilon)$. It is interesting to note that the perturbation $\epsilon$ effectively rescales $\alpha$ as $\delta_{\text{Cod}}(|\alpha|, \epsilon) = \delta_{\text{Cod}}(|\alpha| \cos(\epsilon/2), 0)$. The details of the calculation and the analytic form are presented in Appendix D.

On the other hand, as shown in Appendix B, using single photons with the crystals’ phases fixed at $\pi - \epsilon$ achieves

$$\delta_{\text{Fock}}(\epsilon) = 1 - 2 \sin^2 \left( \frac{\epsilon}{2} \right),$$  (15)

which reduces to the logical bound $\delta = 1$ for $\epsilon = 0$.

By solving numerically the equation $\delta_{\text{Cod}}(|\alpha(\epsilon)|, \epsilon) = \delta_{\text{Fock}}(\epsilon)$ for various $\epsilon$ we obtained the value of the coherent state’s $|\alpha(\epsilon)|$ needed to achieve the same violation of the inequality for various crystal perturbations. The power ratio between the two methods is then equal to $R = |\alpha(\epsilon)|^2$ and is shown in Figure 3.

The phase stabilization of interferometers highly depends on the length of its arms: 6 km long fiber interferometers can reduce phase fluctuations up to $\epsilon \approx 0.06$ [29], while tabletop setups with 1 meter long arms can reduce them up to $\epsilon \approx 10^{-4}$. In the former case, using single-photon states requires 2.5 times less power than by using coherent states for achieving the same violation. Alternatively, we expect that the average time needed to achieve the violation using coherent states is 2.5 times longer than by using single photons with the same power consumption. The difference is even more significant for tabletop experiments, since in this case the coherent states scheme requires 4.5 times more power. Therefore, Fock states provide a significant enhancement to the information speed with respect to coherent states when the source’s power is bounded.

**DISCUSSION**

An interesting remark about our scheme is the fact that the only resource used is coherence of paths and all the information is encoded via local phases. It still remains open whether internal degrees of freedom could increase the performance:
one may expect enhancement coming from entangling internal and spatial degrees of freedom since this might increase the correlation between spatial modes. It would also be interesting to analyze the information speed involved in accomplishing more generic tasks: specifically, one could classify global functions \( f(x_1, \ldots, x_N) \) that Alice is supposed to learn with respect to the advantage provided by spatial coherence. This discussion would then be strongly related to generic quantum algorithmic speedups (for instance hidden subgroup problems \([18]\)).

Another interesting point analyzed here is the practical aspect of communication with light. We have shown a clear advantage of using Fock states over coherent states in the context of information speed enhancement. While our practical analysis was mainly limited to the phase instability and coherent errors, the complete framework shall incorporate all the losses and energy consumption for creating single-photon states (for example via spontaneous down-conversion processes \([19]\)). This will provide a more realistic assessment on energy-time trade offs in communication tasks. These studies are relevant for practical device-independent information processing \([20, 21]\), quantum key distribution \([22–26]\) and direct quantum communication \([27, 28]\) protocols.

**ACKNOWLEDGMENTS**

B.D. acknowledges support from an ESQ Discovery Grant of the Austrian Academy of Sciences (OAW) and the Austrian Science Fund (FWF) through BeyondC (F71). S.H. acknowledges support from the Erasmus Programme. Both authors thank Joshua Morris and Francesco Massa for helpful discussions.

[1] Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. and Wehner, S., 2014. Bell nonlocality. Reviews of Modern Physics, 86(2), p.419.
[2] Del Santo, F. and Dakić, B., 2018. Two-way communication with a single quantum particle. Physical review letters, 120(6), p.060503.
[3] Oreshkov, O., Costa, F. and Brukner, Č., 2012. Quantum correlations with no causal order. Nature communications, 3, p.1092.
[4] Feix, A., Araújo, M. and Brukner, Č., 2016. Causally non-separable processes admitting a causal model. New Journal of Physics, 18(8), p.083040.
[5] Guérin, P.A., Feix, A., Araújo, M. and Brukner, Č., 2016. Exponential communication complexity advantage from quantum superposition of the direction of communication. Physical review letters, 117(10), p.110052.
[6] Chiribella, G., D’Ariano, G.M. and Perinotti, P., 2008. Quantum circuit architecture. Physical review letters, 101(6), p.060401.
[7] Chiribella, G., D’Ariano, G.M., Perinotti, P. and Valiron, B., 2013. Quantum computations without definite causal structure. Physical Review A, 88(2), p.022318.
[8] Allen, J.M.A., Barrett, J., Horsman, D.C., Lee, C.M. and Spekkens, R.W., 2017. Quantum common causes and quantum causal models. Physical Review X, 7(3), p.031021.
[9] Procopio, L.M., Moqanaki, A., Araújo, M., Costa, F., Calafell, I.A., Dowd, E.G., Hamel, D.R., Rozema, L.A., Brukner, Č. and Walther, P., 2015. Experimental superposition of orders of quantum gates. Nature communications, 6, p.7913.
[10] Rubino, G., Rozema, L.A., Feix, A., Araújo, M., Zeuner, J.M., Procopio, L.M., Brukner, Č. and Walther, P., 2017. Experimental verification of an indefinite causal order. Science advances, 3(3), p.e1602589.
[11] Massa, F., Moqanaki, A., Baumeler, A., Del Santo, F., Kettlewell, J.A., Dakić, B. and Walther, P. (2019), Experimental Two-Way Communication with One Photon. Adv. Quantum Technol., 2: 1900050. doi:10.1002/qute.201900050
[12] Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. and Wehner, S., 2014. Bell nonlocality. Reviews of Modern Physics, 86(2), p.419.
[13] Branciard, C., Araújo, M., Feix, A., Costa, F. and Brukner, Č., 2015. The simplest causal inequalities and their violation. New Journal of Physics, 18(1), p.013008.
A behavior \( P (a|x_1, x_2, ..., x_N) : \forall a \in O; x_1, ..., x_N \in I \) can be regarded as an element of a real vector space of dimension \( D = L K^N \), where \( K \) and \( L \) are cardinalities of the input and output alphabets. In what follows we prove that the subset \( S_k \) of all \( k \)-way signaling behaviors forms a polytope.

A general conditional probability can be written as a convex sum of deterministic distributions

\[
P(a|b) = \sum_f \mu_f \delta_{a, f(b)}
\]  

where the sum runs over all functions \( f : I \rightarrow O \), where \( I \) and \( O \) are respectively the input and output alphabets to which \( b \) and \( a \) pertain. A general \( k \)-way signaling correlation can thus be expressed as

\[
P(a|x_1, x_2, ..., x_N) = \sum_{j_1, j_2, ..., j_k} q_{j_1, j_2, ..., j_k} P(a|x_{j_1}, x_{j_2}, ..., x_{j_k})
\]  

\[
= \sum_{j_1, j_2, ..., j_k} q_{j_1, j_2, ..., j_k} \sum_f \mu_f \delta_{a, f(x_{j_1}, x_{j_2}, ..., x_{j_k})}
\]  

\[
= \sum_{f, j_1, ..., j_k} \lambda_{f, j_1, ..., j_k} \delta_{a, f(x_{j_1}, x_{j_2}, ..., x_{j_k})},
\]

where we defined a new set of weights

\[
\lambda_{f, j_1, ..., j_k} = \mu_f q_{j_1, j_2, ..., j_k}; \quad \lambda_{f, j_1, ..., j_k} \geq 0, \quad \sum_{f, j_1, ..., j_k} \lambda_{f, j_1, ..., j_k} = 1.
\]

Therefore, any \( k \)-way signaling behavior can be written as a convex combination of a finite number of deterministic behaviors, which implies that \( S_k \) is a polytope.

**APPENDIX B: QUANTUM VIOLATION FOR \( N = 2 \)**

Suppose we have only two crystals which provide phase shifts \( e^{i \phi_i} x_i \) depending on their local bits \( x_i \) and let both phases be set to \( \pi \).

Alice sends her photon in a homogeneous superposition

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

where \( |1/2\rangle \) represent states of defined trajectories directed towards crystals 1 and 2 respectively. Since the crystals’ phases are fixed to \( \pi \), the state Alice receives after the interaction with the crystal is

\[
|\psi_{x_1, x_2}\rangle = \frac{1}{\sqrt{2}} [(-1)^{x_1} |1\rangle + (-1)^{x_2} |2\rangle].
\]

Plugging \( \rho_{x_1, x_2} = |\psi_{x_1, x_2}\rangle \langle \psi_{x_1, x_2}| \) into expression (11), we obtain \( \max B = 2 \), which saturates the logical bound of the inequality. The latter follows directly from the orthogonality of states \( |\psi_{00}\rangle \) and \( |\psi_{01}/10\rangle \) which enables a perfect state discrimination. Specifically, the required measurement is given by a projection on vectors \( |\pm\rangle = \frac{1}{\sqrt{2}} [\pm 1 \pm 2]. \)

If the crystals’ phases are perturbed by \( \epsilon \), the state Alice receives back is

\[
|\psi_{x_1, x_2}\rangle (\epsilon) = \frac{1}{\sqrt{2}} \left[ e^{i(\pi-\epsilon)x_1} |1\rangle + e^{i(\pi-\epsilon)x_2} |2\rangle \right].
\]

Plugging these states into expression (11) we obtain

\[
\delta(\epsilon) = -\frac{1}{2} + ||\rho_{01} + \rho_{10} - \rho_{00}||_1 = 1 - 2 \sin^2 \left( \frac{\epsilon}{2} \right).
\]
APPENDIX C: PROOF OF THE QUANTUM VIOLATION FOR ARBITRARY N

In this appendix we provide a step by step proof of the violation of inequality (3) for \( N \geq 2 \) by using a single particle. According to the encoding scheme portrayed in section Quantum enhanced information speed, the density operators of interest are

\[
\rho_0 = \frac{1}{N} \sum_{n,m} |n\rangle \langle m|
\]

and

\[
\rho_1 = \frac{1}{N^2} \sum_{n,m} |n\rangle \langle m| \sum_{k} e^{i(\phi_n \delta_{n,k} - \phi_m \delta_{m,k})} = \frac{1}{N^2} \sum_{n,m} |n\rangle \langle m| \left[ N + (1 - \delta_{n,m}) \left( e^{i\phi_n} + e^{-i\phi_m} - 2 \right) \right].
\]

(24)

The goal is to calculate the trace norm of the following operator

\[
p_1 \rho_1 - p_0 \rho_0 = \frac{1}{N(N+1)} \sum_{n,m} |n\rangle \langle m| \left[ (N - 3) + e^{i\phi_n} + e^{-i\phi_m} + 2\delta_{n,m} - \delta_{n,m} \left( e^{i\phi_n} + e^{-i\phi_m} \right) \right].
\]

(25)

The trace norm of an operator \( M \) can be expressed succinctly as

\[
||M||_1 = \sum_i |\lambda_i|,
\]

(26)

where \( \lambda_i \) are the eigenvalues of the given operator. The calculation has thus been reduced to an eigenvalue problem.

Let’s further specify the encoded phases by setting half of them equal to an arbitrary phase \( \phi \) and the other half to \( -\phi \). More specifically, if \( N = 2K \) for some \( K \in \mathbb{N} \), set \( K \) of them to \( \phi \) and \( K \) of them to \( -\phi \), while if \( N = 2K + 1 \), set \( K + 1 \) of them to \( \phi \) and \( K \) of them to \( -\phi \). The operator (25) is then equal to

\[
p_1 \rho_1 - p_0 \rho_0 = \frac{1}{N+1} \left[ \frac{2}{N} (1 - \cos(\phi)) \mathbb{1} + (N - 3) |\psi_0\rangle \langle \psi_0| + |\phi\rangle \langle \psi_0| + |\psi_0\rangle \langle \phi| \right],
\]

(27)

where we introduced an auxiliary phase vector

\[
|\phi\rangle = \frac{1}{\sqrt{N}} \sum_n e^{i\phi_n} |n\rangle.
\]

(28)

Let’s define

\[
M \equiv \langle N - 3 \rangle |\psi_0\rangle \langle \psi_0| + |\phi\rangle \langle \psi_0| + |\psi_0\rangle \langle \phi|
\]

(29)

and diagonalize it in the two-dimensional subspace. Two orthogonal vectors in the given subspace are

\[
|0\rangle \equiv |\psi_0\rangle,
|1\rangle \equiv \frac{1}{\sqrt{1 - |\langle \psi_0| \phi\rangle|^2}} \left( |\phi\rangle - \langle \psi_0| \phi\rangle |\psi_0\rangle \right).
\]

(30)

We’re going to treat even and odd \( N \) cases separately. Starting with even \( N = 2K \), the following holds:

\[
|\psi_0\rangle = |0\rangle,
|\phi\rangle = \cos(\phi) |0\rangle + \sin(\phi) |1\rangle,
\]

(31)

where we used

\[
\langle \psi_0| \phi\rangle = \frac{1}{N} \left( e^{i\phi} + e^{-i\phi} \right) = \cos(\phi).
\]

(32)

Substituting the latter into \( M \) we obtain:

\[
M = \langle N - 3 + 2 \cos(\phi) \rangle |0\rangle \langle 0| + \sin(\phi) |0\rangle \langle 1| + \sin(\phi) |1\rangle \langle 0|.
\]

(33)
The eigenvalues are
\[
\lambda_{\pm} = \frac{1}{2} \left[ A \pm \sqrt{A^2 + (2 \sin(\phi))^2} \right],
\]
\[A \equiv N - 3 + 2 \cos(\phi) .\] (34)

Now we have to return to the full operator (27); since the identity matrix is diagonal in any basis, the eigenvalues trivially follow: two of them are equal to
\[
\mu_{\pm} = \frac{1}{N + 1} \left( \frac{2}{N} (1 - \cos \phi) + \lambda_{\pm} \right),
\]
\[\mu_0 = \frac{1}{N + 1} \left( \frac{2}{N} (1 - \cos \phi) + \lambda_0 \right),\] (35)
and \((N-2)\) of them which correspond to eigenvectors orthogonal to our two-dimensional subspace are equal to
\[
\mu_i = \frac{1}{N + 1} \frac{2}{N} (1 - \cos \phi).
\] (36)

The trace norm is then
\[
||p_1 p_1 - p_0 p_0||_1 = \sum_j |\mu_j| = \frac{1}{N + 1} \left( (N-2) |\frac{2}{N} (1 - \cos(\phi))| + \frac{2}{N} (1 - \cos(\phi)) + \lambda_+ | + \frac{2}{N} (1 - \cos(\phi)) + \lambda_- | \right).
\] (37)

If \(|\frac{2}{N} (1 - \cos(\phi))| > |\lambda_-|\) holds, then \(\delta\) turns out to be 0 and independent of \(\phi\), hence not violating the \((N-1)\)-way signaling bound.

On the contrary, if \(|\frac{2}{N} (1 - \cos(\phi))| < |\lambda_-|:\)
\[
\delta = \frac{3}{2} - \frac{N}{2} - \frac{2}{N} + \frac{2}{N} \cos(\phi) - \cos(\phi) + \frac{1}{2} \sqrt{A^2 + (2 \sin(\phi))^2}.
\] (38)

Inserting the assumed inequality in the previous expression we get
\[
\delta > \frac{3}{2} - \frac{N}{2} - \cos(\phi) + \frac{1}{2} \left[ A - \sqrt{A^2 + (2 \sin(\phi))^2} \right] + \frac{1}{2} \sqrt{A^2 + (2 \sin(\phi))^2} = 0,
\] (39)
which means that the inequality is violated. Now it only remains to be shown that for any even \(N\) there exists \(\phi\) such that \(|\frac{2}{N} (1 - \cos(\phi))| < |\lambda_-|\) is satisfied.

Rearranging and squaring the inequality, we obtain
\[
\left[ \frac{4}{N} (1 - \cos(\phi)) + A \right]^2 < A^2 + (2 \sin(\phi))^2.
\] (40)

A few trigonometric manipulations lead to
\[
\frac{8}{N^2} \sin^2 \left( \frac{\phi}{2} \right) \left\{ (N-2)^2 \cos(\phi) - N(N-6) - 4 \right\} > 0,
\] (41)
which is satisfied if
\[
\cos(\phi) > \frac{N(N-6) + 4}{(N-2)^2}.
\] (42)

This means that for any even \(N\) it is possible to find \(\phi\) such that our communication scheme violates the \((N-1)\)-way signaling bound. In particular, \(\phi\) has to be chosen such that \(\frac{N(N-6) + 4}{(N-2)^2} < \cos(\phi) < 1\).

The previous analysis holds for even \(N\); for odd \(N = 2K + 1\) we get
\[
\langle \psi | \phi \rangle = \frac{1}{N} \left( (N-1) \cos(\phi) + e^{i\phi} \right) = \cos(\phi) + \frac{i}{N} \sin(\phi),
\] (43)
and
\[
M' = (N - 3 + 2 \cos(\phi)) |0\rangle \langle 0| + \sin(\phi) \sqrt{1 - \frac{1}{N^2}} |0\rangle \langle 1| + \sin(\phi) \sqrt{1 - \frac{1}{N^2}} |1\rangle \langle 0|.
\] (44)
which is equivalent to the even $N$ case up to the factor $\sqrt{1 - \frac{1}{N^2}}$ in the off-diagonal elements. Following the analogous procedure, one obtains a clear violation $\delta > 0$ if $\left|\frac{3}{N} (1 - \cos(\phi))\right| < |\lambda_-|$, where $\lambda_-$ is the negative eigenvalue of the operator $M'$. The assumed inequality can be cast in a simpler form using trigonometric relations and can be shown to be equivalent to the condition

$$\cos(\phi) > \frac{N(N - 6) + 5}{N^2 - 2N + 3}. \tag{45}$$

Therefore, we showed the possibility of achieving multi-way signaling with an arbitrary number of parties.

**APPENDIX D: QUANTUM VIOLATION USING COHERENT STATES FOR $N = 2$**

As described in *Information speed with bounded energy*, if Alice uses a source of coherent photons $|\alpha\rangle$, the states she receives back are

$$|\psi_{x1x2}\rangle = |\alpha e^{i(\pi - \epsilon)/\sqrt{2}}\rangle \quad \text{and} \quad |\psi_{x1x2}\rangle = |\alpha e^{i(\pi - \epsilon)/\sqrt{2}}\rangle. \tag{46}$$

In order to simplify the calculation, we apply a displacement operator $D_1 \left( \alpha \sqrt{2} \right) D_2 \left( \alpha \sqrt{2} \right)$ to the latter states, since their distinguishability doesn’t change under this operation. We thus obtain

$$|\psi_00\rangle = |0\rangle_1 |0\rangle_2$$
$$|\psi_10\rangle = |0\rangle_1 \left( 1 + e^{-i\epsilon} \right)/\sqrt{2} |0\rangle_2$$
$$|\psi_01\rangle = |0\rangle_1 \left( 1 + e^{-i\epsilon} \right)/\sqrt{2} |0\rangle_2. \tag{47}$$

Computing expression (11) is now rather easy since the problem is reduced to two qubits, where each of them is spanned by the vacuum $|0\rangle$ and by the state $|x\rangle \equiv -\frac{\alpha}{\sqrt{2}} \left( 1 + e^{-i\epsilon} \right)$. Their overlap is given by

$$\langle 0|x\rangle = e^{-|\alpha \cos(\frac{\epsilon}{2})|^2}, \tag{48}$$

which shows that the effect of the perturbation $\epsilon$ can be absorbed in a rescalement of $|\alpha|$, i.e. $|\alpha| \to |\alpha \cos(\frac{\epsilon}{2})|$. Plugging into (11) we obtain the violation $\delta$ as function of $\alpha$ and $\epsilon$:

$$\delta_{\text{Coh}} (\alpha, \epsilon) = \frac{1}{4A^2} \left[ -1 - A + \sqrt{9 + A (-14 + A (17 + 8A (-3 + 2A)))} \right] \tag{49}$$

where $A \equiv e^{i|\alpha \cos(\frac{\epsilon}{2})|^2}$.

As expected, the latter expression reduces to $\delta = 0$ for $|\alpha| = 0$ and converges to $\delta = 1$ for $|\alpha| \to \infty$. 

