Can communication power of separable correlations exceed that of entanglement resource?

Paweł Horodecki,1,3 Jan Tuziemski,1,3 Paweł Mazurek,2,3 and Ryszard Horodecki2,3

1Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-952 Gdańsk, Poland
2Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland
3National Quantum Information Centre of Gdańsk, 81-824 Sopot, Poland

(Dated: May 7, 2014)

The scenario of remote state preparation with shared correlated quantum state and one bit of forward communication [B. Dakić et al. Nature Physics 8, 666-670 (2012)] is considered. Optimisation of the transmission efficiency is extended to include general encoding and decoding strategies. The importance of use of linear fidelity is recognised. It is shown that separable states cannot exceed the efficiency of entangled states in standard “local operations plus classical communication” paradigm. It is proven however that such a surprising phenomena may naturally occur when the decoding agent has limited resources in the sense that either (i) has to use decoding which is insensitive to change of coordinate system in the plane being in question (which is the natural choice if the receive does not know the latter) or (ii) is forced to use bistochastic operations which may be imposed by physically inconvenient local thermodynamical conditions.

PACS numbers: 03.67.-a, 03.67.Hk, 03.65.Ud
Keywords:

Introduction. - It was recognized that quantum correlations provide a resource for special tasks such as computing, teleportation, dense coding. Originally the quantum advantage in realization of those task was due to quantum entanglement. However there is a more general phenomenon of quantum correlations involving the correlations beyond entanglement (see [1] and references therein). It turned out that efficiency of the protocols involving the latter may also exceed the efficiency of any classical solution of Deutsch Jozsa problem [2], or Knill-Laflamme scheme [4]. Quite recently two proposals of application of quantum correlations beyond entanglement (QCBE) have been provided. One of them have theoretically and experimentally supported the significance of their role in the analogue of quantum dense coding [2] on the level of continuous variables. The other [6] has addressed the issue of the importance of QCBE for the remote state preparation (RSP). Since RSP is one of the significant building blocks in quantum communication the question is very important. While it has already been adapted to weak entanglement scenarios including bound entanglement [5] it has never been proven to be possible in QCBE which involves separable states. The paper [6] announces the surprising possibility of the fact that in some cases the communication power of QCBE represented by separable states may exceed that of some entangled ones. The authors have also provided the direct connection of the transfer fidelity they have chosen to measure (called geometrical discord) of quantum correlations of the resource state. However this conclusion - unlike the experimental results of the paper fully transparent and impressive - seems to be not fully justified.

First of all the (quadratic) measure of transfer fidelity based on average of Bloch vector used in the description may be highly misleading. For instance it is completely confusing in fully classical protocol in which Alice and Bob preagreed on a fixed (say standard) basis. Then completely random choice of the Bob state in that basis gives already maximal figure of merit for the protocol. Even more intriguing fact is that in RSP protocol any classical deterministic strategy of Bob preparing pure state along arbitrary fixed axis exceeds efficiencies of the originally proposed schemes in Ref. [6]. This makes the presence of shared quantum state irrelevant not speaking about need of classical communication resource [17].

This shows that the correct figure of merit of the state transfer in RSP protocol should be the standard linear one (see [2]).

Before detailed analysis it is always important to make consistent assumptions about the resources. In the standard LOCC paradigm the local Bloch coordinates (reference frame) are assumed to be known and the observers are allowed to use that knowledge. From that perspective the classical bit of information is only the result of the measurement on Alice side.

As we shall prove in this paper, for linear fidelity and fully optimised encoding and decoding strategies, working under assumption of standard LOCC paradigm, there is no chance for separable state to beat entanglement efficiency as a resource in any RSP protocol. Then the basic question arises if there is any other natural scenario in which separable QCBE can exceed the efficiency offered by some entangled state in RSP protocols. The result of [6] give the strong evidence that it may be so in cases, when decoding is insensitive to change of coordinates in the sender’s plane. However it is based on misleading fidelity and, for instance, not naturally restricted use of the classical communication channel: Al-
ice only uses von Neumann measurements and Bob uses only unitary operations. In fact it was suggested in [10] that the link between optimal fidelity of RSP and quantum discord presented in [8] may be caused by the usage of non-optimised protocol on Bob’s side. This suggestion was based on hybrid version of RSP protocol.

Here we state the problem in a natural perspective. We use the correct fidelity and consider the other class of the protocols i.e. the ones in which the Bob action should be invariant under the rotation in the plane form which the qubit state is to be prepared. This is a very natural choice if Bob does not know the Bloch reference frame of Alice and has no way to infer it from his data. We shall call this class restricted-LOCC (R-LOCC). We also consider the other scenario i.e. standard LOCC but with extra restriction on Bob side who is forced to apply bistochastic operations only, which may be caused by infinite temperature of its working environment. We prove by explicit analysis that both in R-LOCC and in case of standard one with bistochastic restriction the QCBE may work better.

It is intriguing that in second scenario the final fidelity for Bell diagonal states depends on the same parameters as geometrical discord. This makes in some way the intuition of the paper [9] correct.

\[\rho = \rho(\tilde{x}, \tilde{y}, T) = \frac{1}{4}[I \otimes I + \tilde{x}\tilde{\sigma} \otimes I + I \otimes \tilde{y}\tilde{\sigma} + \sum_{ij} T_{ij}\sigma_i \otimes \sigma_j].\] (1)

The most general form of Alice binary POVM must be a function of the following family of parameters \(A = \{\tilde{a}, a_+, a_-\}\) and is defined by the formula \(M_{\pm} = a_{\pm}I \pm \tilde{a}\tilde{\sigma}\) with the probability-like parameters \(a_{\pm}\) and vector \(\tilde{a}\) satisfying the conditions:

\[a_+ + a_- = 1, \quad 0 \leq a_{\pm} \leq 1, \quad ||\tilde{a}|| \leq \min[a_+, a_-] \leq \frac{1}{2},\] (2)

where in general both \(a_{\pm}\) and \(\tilde{a}\) are functions of the unit vector \(\hat{s}\) perpendicular to \(\tilde{\beta}\) which has a fixed orientation during the protocol. Finally the payoff function of the protocol is minimised over \(\tilde{\beta}\).

The resulting probabilities of the Alice outcomes on the state \(\tilde{\rho}\) and the resulting states \(\rho_{\pm}\) on Bob side are defined by the relations:

\[p_{\pm} = Tr_{AB}[M_{\pm} \otimes I\rho] = (a_{\pm} \pm \tilde{a}\tilde{x}),\] (3)

\[p_{\pm}p_{\pm} = Tr_{A}[M_{\pm} \otimes I\rho] = \frac{1}{2}[(a_{\pm} \pm \tilde{a}\tilde{x})I + (\pm T\tilde{a} + a_{\pm}\tilde{x})\tilde{\sigma}]\] (4)

Bob is allowed to perform channels \(\Lambda_{\pm}\) which depend upon the result \(\pm\) of the Alice measurement and act on any qubit state \(\rho(\tilde{u}) = \frac{1}{2}(I + \tilde{u}\tilde{\sigma})\) as:

\[\Lambda_{\pm}[\rho(\tilde{u})] = \frac{1}{2}[I + (T_{\pm}\tilde{u} + \tilde{u}_{\pm})\tilde{\sigma}].\] (5)

Note that by standard convexity arguments the decoding channels may always be chosen to be extremal.

After the action of \(\Lambda_{\pm}\) the final Bob state is

\[\tilde{\rho}_B = \sum_{r=\pm} p_r\Lambda_r(\tilde{\rho}_B^{(r)}) = \frac{1}{2}(I + \tilde{r}\tilde{\sigma}),\] (6)

with the final Bob Bloch vector

\[\tilde{r} = \sum_{r=\pm} T_r(a_r\tilde{y} + rT\tilde{a}) + (a_r + r\tilde{a}\tilde{y})\tilde{v}_r.\] (7)

**Probabilistic fidelity**. For fixed \(\hat{s}\) the probabilistic fidelity of the success in remote state preparation of a pure state \(\rho(\hat{a}) = \frac{1}{2}(I + \hat{s}\hat{\sigma})\) is defined as (see [3]):

\[F(\hat{s}) = \frac{1}{2}(1 + \hat{r}\hat{s}).\] Averaging over \(\hat{s}\) gives

\[\bar{F} = \frac{1}{2}(1 + G),\] (8)

where the fidelity parameter is

\[G = G(\rho; \hat{\beta}, A, T) = \int d\hat{s}(\hat{r}\hat{s}) = \int d\hat{s}[(T_+ - T_-)T\tilde{a} + (\tilde{v}_+ - \tilde{v}_-)]||\tilde{a}\| + a_+(T_+\tilde{y} + \tilde{v}_+) + a_-(T_-\tilde{y} + \tilde{v}_-)]\hat{s}.\] (10)
Here \( \rho = \rho(x, y, T) \), \( \hat{\beta} \) defines the plane to which the vectors \( \hat{s} \) belongs and the explicit dependence on the encoding \( A = \{ a, a+, a- \} \) and decoding strategy \( T = \{ T_+, T_+, T_- \} \) is written. The full range of parameters describing the encoding \( A \) is written explicitly in (2). The range of parameters \( T \) is determined by the structure of the extremal one-qubit channels [18].

Optimal RSP protocol for separable states. Advantage of entangled states in standard LOCC paradigm.- Here we keep the assumption of LOCC paradigm in which Alice and Bob naturally share the reference frame on the Bloch sphere. We may choose the coordinates as \( \{ \hat{\beta}, \hat{e}, \hat{e}' \} \) where \( \hat{\beta} \times \hat{e} = \hat{e}' \) and \( \{ \hat{e}, \hat{e}' \} \) represent the coordinates system in the \( \hat{s} \) plane.

Because fidelity is convex, for separable states it is sufficient to consider pure states. For pure states \( p_s = a_s(\hat{s}) + \bar{a}(\hat{s})\hat{\xi} \), where \( \hat{\xi} \) is Alice Bloch vector. The reduced state of Bob is \( \rho_+ = \frac{1}{2}(I + \bar{n}_\pm \hat{\sigma}) \), where \( \bar{n}_\pm \) is Bob Bloch vector transformed by respective channel. Then (10) is of a form \( \hat{F} = \frac{1}{2} \int d\rho_+(\hat{s}) \bar{n}_+(\hat{s}) - \bar{n}_-(\hat{s}) \hat{s} \).

As a result using separable states cannot lead to better fidelity than using entangled states \( (\rho_{ent}) \) and inserting it into (12), we obtain optimised formula for \( \max_A G(\rho(x, y, T); \hat{\beta}; A, T) \). In the case of an isotropic correlations \( T = -\lambda I \) we have following facts:

Fact 1.- We can always decompose \( \hat{y} \) as \( \hat{y} = \| \hat{y} \| \left( \alpha \hat{u} + (1 - \alpha)\hat{\beta} \right) \). Then the formula \( \max_A G(\rho(x, y, T); \hat{\beta}; A, T) \) is monotonic function of parameter \( \alpha = \| \hat{y} \| \).

Fact 2.- The optimisation over the R-LOCC class (which naturally corresponds to the situation with an unknown coordinates system, as in case b) of Fig. 2) yields \( \min_{\lambda, \max_A} G(\rho(x, y, -\lambda I); \hat{\beta}; A, T) = \lambda \) and then consequently

\[
F_{R-LOCC}(\rho(x, y, -\lambda I)) = \frac{1}{2} (1 + \lambda).
\]

For details of the proofs of the above see [10]. Now following the Ref. (16) consider the following class: \( \rho(tz, t\hat{z}, -\lambda I) \); or in other words the states with the parameters: \( T = -\lambda I, \hat{x} = \hat{y} = t\hat{z} \) where the positivity condition determines the following range of parameter \( t \):

\[
|t| \leq \frac{1}{2} \sqrt{1 - 2\lambda - 3\lambda^2}.
\]

All of those entangled states \( \rho(x, y, -\lambda I) \) with \( \lambda\leq\frac{1}{2} \) will lead to the fact that the RSP fidelity (under the restriction of the extremal one-qubit channels) is independent of the setting on Alice side. This means that Alice, after establishing decoding strategy with Bob, can e.g. change type of input state by choosing different angle \( \varphi \) or her coordinates system in the plane orthogonal to \( \hat{\beta} \) and Bob’s strategy should remain optimal. As a consequence, this strategy cannot depend on the parametrisation of the input Bloch vector. Technically in this case decoding operations should be restricted to the class, that is invariant under averaging in the plane orthogonal to \( \hat{\beta} \) or always look the same after any rotation in that plane. We will denote this class as \( \Gamma_{R-LOCC} \).

For operations belonging to \( \Gamma_{R-LOCC} \) we have that \( T \{ T_\pm, \{ v_\pm \} \} = \hat{T} \{ \{ \hat{T}_\pm \}, \{ \hat{v}_\pm \} \} \), where \( \hat{T}_\pm = \frac{1}{2\pi} \int_0^{2\pi} d\varphi O_\varphi(\varphi) T_\varphi O_\varphi^T(\varphi), \hat{v}_\pm = \frac{1}{2\pi} \int_0^{2\pi} d\varphi O_\varphi(\varphi) v_\pm \), where \( O_\varphi(\varphi) \) denotes rotation in the \( \hat{s} \) plane. As a result of averaging \( \hat{T}_\pm = \text{diag}[t_\pm, \hat{T}_\pm(1)] \), where \( |t_\pm| \leq 1 \), \( \hat{T}_\pm(1) \) are 2 by 2 matrices acting in the \( \hat{s} \) plane. The use of this class is natural also from game-like perspective: let us allow Bob to use arbitrary channel \( T \). Since he does not know the coordinates he must average his decoding strategy over all possible orientations of reference frame on the \( \hat{s} \) plane. This results in decoding from \( \Gamma_{R-LOCC} \). As a consequence we get that \( \hat{v}_\pm \) have no components parallel to \( \hat{s} \):

\[
\hat{v}_s = \sum_{r=\pm} [\hat{T}_r(1)(a_r \hat{y} + rT\hat{a})] \hat{s}.
\]

By setting \( M = (\hat{T}_+ - \hat{T}_-)^{-1} \) and \( \hat{V} = \hat{T}_+^T(1) \hat{y} \) and inserting it into (12), we obtain optimised formula for \( \max_A G(\rho(x, y, T); \hat{\beta}; A, T) \). In the case of an isotropic correlations \( T = -\lambda I \) we have following facts:

Fact 1.- We can always decompose \( \hat{y} \) as \( \hat{y} = \| \hat{y} \| \left( \alpha \hat{u} + (1 - \alpha)\hat{\beta} \right) \). Then the formula \( \max_A G(\rho(x, y, T); \hat{\beta}; A, T) \) is monotonic function of parameter \( \alpha = \| \hat{y} \| \).

Fact 2.- The optimisation over the R-LOCC class (which naturally corresponds to the situation with an unknown coordinates system, as in case b) of Fig. 2) yields \( \min_{\lambda, \max_A} G(\rho(x, y, -\lambda I); \hat{\beta}; A, T) = \lambda \) and then consequently

\[
F_{R-LOCC}(\rho(x, y, -\lambda I)) = \frac{1}{2} (1 + \lambda).
\]
unknown coordinate system) than the separable states
\( \rho(0, 0, -\lambda t') \) with \( \lambda' = \frac{1}{2} \). (This comprises as special cases considered in Ref. 10: the separable case \( \lambda' = \frac{1}{2} \), 
\( t = 0 \) and an entangled one with \( \lambda = \frac{1}{2}, \ t = \frac{1}{2} \). The overall conclusion is that whenever Bob does not know
the coordinates of Alice in the \( s \) plane then entanglement may be less useful than quantum correlations beyond it, i.e. the ones contained in separable states.

**Optimal RSP with help of Bell diagonal states and one bit of communication in the case of bistochastic channels**. Let us consider a situation when temperature of Bob’s environment is infinite. Then he is restricted to use bistochastic channels. Detailed analysis shows that for Bell diagonal states the formula (12) can be fully optimised. Here by Bell diagonal states we mean all the states that are local unitary (i.e. \( U_1 \otimes U_2 \) type) rotations of the states diagonal in the standard Bell basis
\( \Psi_\pm = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \) and \( \Phi_\pm = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \). It is known that all such states can be represented by
\( \rho(0, \bar{\theta}, T) \). Then after some algebra (see 14) one gets
\[
\min_{\beta} \max_{A,T} G(\rho(0, \bar{\theta}, T); \beta; A, T) = \frac{2|t_2|}{\pi} E\left(\sqrt{1 - \frac{t_1^2}{t_2^2}}\right)
\]
and
\[
\bar{F}_{\text{bistochastic}} = \frac{1}{2} \left[1 + \frac{2|t_2|}{\pi} E\left(\sqrt{1 - \frac{t_1^2}{t_2^2}}\right)\right],
\]
where \( t_1^2, t_2^2 \) are two lowest eigenvalues of \( T^T T \) and \( E(x) \) is complete elliptic integral of the second kind [15].

In this case it is also possible to show that there exists separable states leading to higher fidelity of RSP protocol than entangled states. Let us consider two Bell diagonal states with following correlation tensors:
\[
T_1 = \left[-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right], \ T_2 = \left[-\frac{1}{2} - 2\epsilon, -\frac{1}{2} + \frac{\bar{\theta}}{2}, -\frac{1}{2} - \frac{\bar{\theta}}{2}, -\frac{1}{2} + \frac{\bar{\theta}}{2}\right]
\]
with \( \epsilon > 0 \). The set of separable Bell diagonal states is specified by condition \( |t_1| + |t_2| + |t_3| \leq 1 \). Clearly the state corresponding to \( T_1 \) is separable whereas that corresponding to \( T_2 \) is not. Using (13) one immediately obtains that
\[
\bar{F}_{\text{bistochastic}}(\rho(0, \bar{\theta}, T_1)) = \frac{2 \bar{\theta}}{3} > \frac{2 \bar{\theta}}{3} - \frac{2 \bar{\theta}}{3} = F_{\text{bistochastic}}(\rho(0, \bar{\theta}, T_2)).
\]
Interestingly (13) depends only on the two smallest
eigenvalues of \( T^T T \). Since the geometric discord is in this case of the form \( D(\rho(0, \bar{\theta}, T)) = \frac{2}{3}(t_1^2 + t_2^2) \) the optimised fidelity depends on the same parameters like the one used in (3), because the latter was just proportional to discord. This shows that in the case of bistochastic decoding the presented result based on standard fidelity and the one based on quadratic fidelity are consistent.

**Conclusions**. It is known that quantum correlations without entanglement, contained in separable states may be useful in quantum information processing. The basic issue is whether they may outperform entanglement in any case. Our analysis shown that one should be careful in comparison of the two resources. In fact two-qubit separable states cannot outperform two-qubit entanglement in the process of remote state preparation of quantum bit under most general assumptions i.e. standard LOCC paradigm. This lies in the heart of the balance of quantum resources within so called LOCC scenario: whenever initial entanglement is too weak, Bob may remove entanglement and prepare the optimal state on his own still achieving the best efficiency provided by all separable correlations based protocols. Thus any protocol with initial entanglement cannot be worse than the one with separable state.

The apparent contradictions to the above may only take place if one use nonstandard figure of merit. To be correct the latter must make a difference between the orthogonal states in any basis and the standard linear fidelity works well from that perspective.

However it turns out that all the above does not prevent quantum information from „quantum separability advantage” in the cases when the users have extra restrictions, somehow very natural, on the operations they are allowed to apply. Here we have shown that if in the RSP protocol of a qubit state with two-qubit correlations the receiver is restricted to the decoding class, which reflects his ignorance about the coordinates in the plane the origin state comes from, then separability can work better than entanglement. The second scenario when the latter may happen is the one when the receiver is forced to use bistochastic decodings. Then, whenever Alice and Bob share Bell diagonal quantum state, the final fidelity of the protocol depends on the same set of parameters as geometric quantum discord.

The latter result is even more intriguing, when one realises that the restriction of bistochastic character of the decoding may be interpreted as a presence of „thermodynamically unsuitable” ancillas namely those of infinite temperature. Note that in this case quantum correlations beyond entanglement help better than entanglement itself and that one option to discriminate those correlations form classical ones is just the thermodynamical picture of local engines (see 13, 14). This suggests that possible thermodynamical perspective of the discussed protocol (and also practical aspects of other protocols aimed in using quantum correlations beyond entanglement) should be examined more in future.

**Acknowledgements**. PH and RH thank A. Miranowicz.
and R. Chhajlany for discussions. This paper is supported by Polish Ministry of Science and Higher Education Grant no. IdP2011 000361. PM is supported by the Foundation for Polish Science International PhD Projects Programme co-financed by the EU European Regional Development Fund. JT is supported by EU Project QESSENCE.

[1] K. Modi et al., Rev. Mod. Phys. 84, 1655 (2012).
[2] D. Kenigsgberg et al., Quantum Inform. Compu. 6(7), 606 (2006).
[3] E. Biham et al., Theor. Comput. Sci. 320(1), 15 (2004).
[4] A. Datta et al., Phys. Rev. Lett. 100, 050502 (2008).
[5] S. Popescu, Phys. Rev. Lett. 72, 797 (1994).
[6] B. Dakić et al., Nature Phys. 8, 666 (2012).
[7] M. Gu et al. Nature Phys. 8, 671 (2012).
[8] M. Murao et al., Phys. Rev. Lett. 86, 352 (2001).
[9] M. B. Ruskai et al., Lin. Alg. Appl. 347, 159 (2002).
[10] T. Tufarelli et al., Phys. Rev. A 86, 052326 (2012)
[11] R. Horodecki et al., Phys. Rev. A 54, 1838 (1996).
[12] B. C. Carlson, Elliptic Integrals in NIST Handbook of Mathematical Functions ed. by F. W. J. Olver et al. (Cambridge University Press, Cambridge, 2010).
[13] J. Oppenheim, et al., Phys. Rev. Lett. 89, 180402 (2002).
[14] W. H. Zurek, Phys. Rev. A 67, 012320 (2003).
[15] B. C. Carlson, Elliptic Integrals in NIST Handbook of Mathematical Functions ed. by F. W. J. Olver et al. (Cambridge University Press, Cambridge, 2010).
[16] Supplemental Material
[17] In fact Alice and Bob may share no quantum state at all, and Bob may produce the pure state in question without requiring any information from Alice. In such a case he obtains the fidelity that is much better than any offered by the protocol in [12].
[18] $T_\pm = O^{(1)}_\pm T^{(2)}_\pm O^{(2)}_\pm$, $\vec{v}_\pm = O^{(1)}_\pm \vec{v}_0^{(2)}$, where $O^{(1)}_\pm, O^{(2)}_\pm$ are arbitrary rotations combined with representation of completely positive trace preserving maps. Here we use a family given by $\vec{v}_\pm = [0, 0, \sin u_{\pm} \sin w_{\pm}]$, $T_\pm^{(2)} = [\cos u_{\pm}, \cos w_{\pm}, \cos u_{\pm} \cos w_{\pm}]$, $T_\pm^{(1)} = [\cos w_{\pm}, \cos u_{\pm} \cos w_{\pm}, \cos w_{\pm}]$ with $u_{\pm} \in [0, 2\pi]$, $w_{\pm} \in [0, \pi]$. This is the most general form of any one-qubit channel belonging to the closure of the set of extreme one-qubit channels. By simple convexity argument it is enough to consider only these channels.

**SUPPLEMENTAL MATERIAL**

**Optimisation over Alice POVMs**.- Here we shall optimise the quantity

$$G = \int d\bar{s}(T_+ - T_-)^T \bar{s} + (\bar{v}_+ - \bar{v}_-) \bar{s} + a_+(T\bar{y}_+ + \bar{v}_+) + a_-(T\bar{y}_- + \bar{v}_-))\bar{s}$$

with respect to $A$ (to be concise we shall omit all the arguments in its notation). Define the matrix $M = (T_+ - T_-)T + (|\bar{v}_+| - |\bar{v}_-|) \langle \bar{s} \rangle$, and $\tilde{V}_+ = (T\bar{y}_+ + \bar{v}_+)$, $\tilde{V}_- = (T\bar{y}_- + \bar{v}_-)$. Then the above function is of the form

$$G = \int d\bar{s}[M\bar{a} + a_+ \tilde{V}_+ + a_- \tilde{V}_-] \bar{s},$$

where the vector $\bar{a}$ and the scalars $a_{\pm}$ depend in general on $\bar{s}$ and satisfy the conditions (2) (main text). Let us put $\bar{a} = a\bar{a}$, where $0 \leq a = ||\bar{a}|| \leq a_\pm$. Clearly the best choice to maximise the value of $G$ is to put $\bar{a}$ parallel to the vector $MT\bar{s}$ or, in other words, $\bar{a} = \frac{M^T \bar{s}}{||M^T \bar{s}||}$. Then the value of the integral becomes $G = \int d\bar{s}a[||M^T \bar{s}|| + a_+ \tilde{V}_+ + a_- \tilde{V}_-] \bar{s}$, which may be further optimised with respect to $a$ by taking its maximal allowed value $a = \min(a_+, a_-)$.

Eventually, this gives the function optimised over $\bar{a}$ for fixed $a_{\pm}$ and all the other parameters:

$$G = \int d\bar{s}[||M^T \bar{s}||\min(a_+, a_-) + a_+ \tilde{V}_+ + a_- \tilde{V}_-] \bar{s}.$$  

Using notation $M' = ||[M^T \bar{s}]|| \geq 0$, $A_{\pm} = \tilde{V}_{\pm} \bar{s}$, we may carefully consider the maximum of

$$f(p) = M'\min[p, 1 - p] + pA_+ + (1 - p)A_-$$

over the interval $p \in [0, 1]$, where we put $p = a_+$ and $1 - p = a_-$ for conciseness. The above function has the following maxima:

(a) if $M' \geq |A_+ - A_-|$, then $\max_{p \in [0, 1]} f(p) = \frac{M + A_+ + A_-}{2}$ achieved at $p = \frac{1}{2}$;

(b) if $M' < |A_+ - A_-|$, then either (i) $\max_{p \in [0, 1]} f(p) = A_+$ for $A_+ - A_- > 0$ (achieved at $p = 0$) or (ii) $\max_{p \in [0, 1]} f(p) = A_-$ for $A_- - A_+ > 0$ (achieved at $p = 1$).

In case (a) the strategy of Alice is naturally the one of Ref. (13); she performs the von Neumann measurement with the projections $P_{\pm} = \frac{1}{2}(I \pm a\bar{a}^T)$. An intriguing strategy of Alice in case (b) is that she just does nothing (since then the POVM is the identity) and puts the message $r$ to Bob depending on the sign of $(A_+ - A_-) = (\tilde{V}_+ - \tilde{V}_-) \bar{s}$. Quite remarkably this strategy gives always nonnegative contribution form the part of the integral involving the vectors $\tilde{V}_\pm$.

We have then the three sets in the unit circle on the $\bar{s}$ plane: $\Omega_0$, $\Omega_\pm$ defined as
(i) $\Omega_0 = \{ \hat{s} : \hat{s}\tilde{\beta} = 0, ||M^T\hat{s}|| \geq (\tilde{V}_+ - \tilde{V}_-)\hat{s} \}$; (ii) $\Omega_+ = \{ \hat{s} : \hat{s}\tilde{\beta} = 0, ||M^T\hat{s}|| < (\tilde{V}_+ - \tilde{V}_-)\hat{s} \}$; (iii) $\Omega_- = \{ \hat{s} : \hat{s}\tilde{\beta} = 0, ||M^T\hat{s}|| < (\tilde{V}_+ - \tilde{V}_-)\hat{s} \}$.

The final formula optimised over $A$ is of the form

$$\max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; A, T)) = \int_{\Omega_0^+} d\hat{s} ||M^T\hat{s}|| + \int_{\Omega_-} d\hat{s} (\tilde{V}_+ - \tilde{V}_-)\hat{s}.$$  \hfill (17)

We have also a following Observation .- i) $\Omega_0 = -\Omega_0$, $\Omega_0$ is symmetrical, ii) $\Omega_+ = -\Omega_-$, i.e. after the reflection the sets are equal. From i) we get that $\int_{\Omega_0} d\hat{s} ||M^T\hat{s}|| + \tilde{V}_+\hat{s} + \tilde{V}_-\hat{s} = \int_{\Omega_0} d\hat{s} ||M^T\hat{s}||$, and from ii) $\int_{\Omega_-} d\hat{s} \tilde{V}_+\hat{s} + \int_{\Omega_+} d\hat{s} \tilde{V}_-\hat{s} = \int_{\Omega_+} d\hat{s} (\tilde{V}_+ - \tilde{V}_-)\hat{s}$. Let us define $\Omega_0^+ = -\Omega_0$ as any of two subsets of original $\Omega_0$ such that $\Omega_0 = \Omega_0^+ \cup \Omega_0^-$. The final formula optimised over $A$ is of the form

$$\max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; A, T)) = \max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; A^*, T))$$

$$\int_{\Omega_0^+} d\hat{s} ||M^T\hat{s}|| + \int_{\Omega_+} d\hat{s} (\tilde{V}_+ - \tilde{V}_-)\hat{s}.$$  \hfill (18)

**Proof of Fact 1 .-** Let us consider $\alpha' > \alpha$, where $\alpha = |\tilde{y}|\nu$. Parameters $\alpha, \alpha'$ correspond to two different orientations of $\tilde{y}$ with respect to $\hat{s}$ plane i.e. $\tilde{y} = ||\tilde{y}|| [\alpha\hat{u} + (1 - \alpha)\tilde{\beta}]$ and $\tilde{y}' = ||\tilde{y}'|| [\alpha'\hat{u} + (1 - \alpha')\tilde{\beta}]$. For $\alpha'$ (\alpha) we will denote solutions of inequalities defining the sets as $\Omega_0^{\alpha'}$, $\Omega_+^{\alpha'}$, $A^{\alpha}$, $\|M^T\hat{s}\| = f'(\tilde{V}_+ - \tilde{V}_-)\hat{s} = g'(\|M^T\hat{s}\|) = f(\tilde{V}_+ - \tilde{V}_-)\hat{s} = g$. Let us recall that here we consider only the restricted class of the invariant (equivalently averaged) decodings $\tilde{T}$. It follows from the definition of the sets that $\Omega_0^{\alpha'} < \Omega_0^\alpha$ and $\Omega_+^{\alpha'} > \Omega_+^\alpha$ as well as $f' = f = \lambda(\|\tilde{T}^{(1)}_+ - \tilde{T}^{(1)}_-\|\|T\|)$. We can rewrite $g'$ as $g' = (\tilde{T}^{(1)}_+ - \tilde{T}^{(1)}_-)T\hat{s} = g'(\|M^T\hat{s}\|) = f'(\tilde{V}_+ - \tilde{V}_-)\hat{s} = g$. As a consequence the following relation holds $g' = g\tilde{w}(\hat{s}) = \|g\| [\alpha'\hat{u} + (1 - \alpha')\tilde{\beta}] > ||\tilde{y}|| \alpha'\hat{u}(\hat{s})$. Thus we can write

$$\max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; \tilde{T})) = \int_{\Omega_0^{\alpha'}} d\hat{s} f' + \int_{\Omega_+^{\alpha'}} d\hat{s} g' = \int_{\Omega_0^\alpha} d\hat{s} f + \int_{\Omega_+^\alpha} d\hat{s} g' \geq \int_{\Omega_0^{\alpha'}} d\hat{s} f + \int_{\Omega_+^{\alpha'}} d\hat{s} f = \int_{\Omega_0^\alpha} d\hat{s} = \max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; \tilde{T})).$$.  \hfill (19)

As a result, for $\alpha' > \alpha$ it holds that $\max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; \tilde{T})) > \max_{\Delta} A(G(\rho; \hat{\beta}; \hat{\alpha}; \tilde{T})).$

**Proof of Fact 2 .-** It follows from the Fact 1 that $\max_{\Delta} A(G(\rho(\tilde{x}, \tilde{y}, -\lambda\tilde{I}); \hat{\beta}; \hat{\alpha}; \tilde{T}))$ is monotonic in $\alpha$, where $\alpha = |\tilde{y}|\nu$. Let us consider a simple

**Lemma .-** Let $f(a, x)$ be a function with $a \in [a_0, a_1]$ and $x \in \Omega \subset \mathbb{R}^n$ where $\Omega$ is compact. Suppose that (i) for any $a \leq a'$ and for any $x$ one has $f(a, x) \leq f(a', x)$; (ii) the $x(a)$ is some (may be not unique) point realising maximum of $f(a, x)$ over $x$ for fixed $a$, i.e. $f(a, x(a)) = \max_x f(a, x)$. Then the function $f(a, x(a))$ is monotonic in $a$. As a result $\min_{a\in[a_0, a_1]} \max_x f(a, x) = \max_{a\in[a_0, a_1]} \min_x f(a, x)$.

**Proof of Lemma .-** Consider any $a \leq a'$. Then we have $f(a, x(a)) \leq f(a', x(a)) \leq f(a', x(a'))$, where the first inequality follows from (i) and the second one from (ii).

Coming back to the proof of the Fact 2 we may put in place of $a \in [a_0, a_1]$ in the Lemma above the parameter $\alpha \in [0, 1]$ and in place of $x$ all the other parameters contained in the sets $A, \tilde{T}$ getting the desired monotonicity in Fact 2.

As a result of the Fact 2 we can set $\alpha = 0$ which implies $\hat{\beta} = \tilde{y} \Omega_0^+ = (0, x), \Omega_+ = \emptyset$.

$$\min_{\Delta} \max_{\Delta} A(\tilde{T}(\tilde{x}, \tilde{y}, -\lambda\tilde{I}); \hat{\beta}; \hat{\alpha}; \tilde{T}) = \max_{\Delta} A(\tilde{T}(\tilde{x}, \tilde{y}, -\lambda\tilde{I}); \hat{\beta}; \hat{\alpha}; \tilde{T}) = \max_{T_+} \mathbb{A}^{\alpha} \mathbb{I}^{\lambda} \int_{\Omega_0^+} d\lambda \|\tilde{T}^{(1)}_+ - \tilde{T}^{(1)}_-\| T\hat{s} = \lambda$$

since (i) $\alpha = 0$ implies $\tilde{y}|\tilde{\beta}$, (ii) the triangle inequality $\|\tilde{T}^{(1)}_+ - \tilde{T}^{(1)}_-\| T\hat{s} \leq \|\tilde{T}^{(1)}_-\| T\hat{s} \leq 2$ is saturated for the $T^{(1)}_+ = \pm I$ choice and (iii) $d\lambda$ represents the measure on the plane $d\tilde{\lambda} = \frac{d\tilde{\lambda}}{2\pi}$.

**Optimisation of formula (12) (main text) for Bell diagonal states and bistochastic decodings.-** Maximisation $\|M^T\hat{s}\|$ over decoding strategies $\tilde{T}$ gives $\tilde{T}_+ = I$ and $\tilde{T}_- = R_\beta(\pi)$ - rotation about $\beta$ direction. For Bell states $TT^T = \text{diag}[t_1^2, t_2^2, t_3^2]$, $t_3^2 \geq t_2^2 \geq t_1^2$. Consider the situation when $\hat{s}$ is not orthogonal to the largest eigenvalue of $TT^T$. Let us denote $\max_{\Delta} A(\tilde{T}(\tilde{x}, 0, 0); \hat{\beta}; \hat{\alpha}; \tilde{T}) = \int_{0}^{2\pi} d\phi \tilde{g}(\hat{s}(\theta(\phi), \phi))$, where effectively $\phi$ parametrises $\hat{s}$. We can decompose any rotation into rotation about $y$ axis in a plane perpendicular to $\tilde{z}$ followed by rotation about $\tilde{\beta}$. Then $\hat{s}$ is transformed into $\hat{s}'$, what corresponds to the change of parametrization $(\phi, \theta(\phi)) \rightarrow (\phi'(\phi), \theta'(\phi))$. In the rotated frame $\max_{\Delta} A(\tilde{T}(\tilde{x}, 0, 0); \hat{\beta}; \hat{\alpha}; \tilde{T}) = \int_{0}^{2\pi} d\phi' \tilde{g}^\prime(\hat{s}'(\phi', \phi')) = \int_{0}^{2\pi} d\phi \tilde{g}(\hat{s}(\theta'(\phi'), \phi'))$.\hfill \hfill
We will need the following

**Lemma.** - The function $f(x) = \int_0^{2\pi} \cos^2 \varphi \, d\varphi$ is a decreasing function of $x$.

**Proof.** - We have

$$\frac{\partial}{\partial x} f(x) = \int_0^{2\pi} \frac{\sin \varphi \cos \varphi}{\sqrt{1 - C \sin^2 \varphi + x \sin 2\varphi}} = 0.$$

To optimise the formula (12) (main text), we consider the function $g(\hat{s}(\theta'(\phi), \phi'(\phi)))$ which can be explicitly written as

$$g(\hat{s}(\theta'(\phi), \phi'(\phi))) = \sqrt{(R_{\theta}(\theta) R_{\varphi}(\varphi) \hat{s}, TT^T R_{\theta}(\theta) R_{\varphi}(\varphi) \hat{s})}.$$

Now the following relation holds

$$\int_0^{2\pi} \cos^2 \varphi \left[ \cos^2 \psi(t_1^2 \cos^2 \theta + t_2^2 \sin^2 \theta) + \sin^2 \psi t_3^2 \right] + \frac{1}{2} \sin 2\varphi \sin 2\theta \cos \psi(t_2^2 - t_1^2) + \sin^2 \varphi (\sin^2 \theta t_1^2 + \cos^2 \theta t_2^2) \geq 0.$$