

## Using Entanglement Improves the Precision of Quantum Measurements

G. Mauro D'Ariano,<sup>1,2,3,\*</sup> Paoloplacido Lo Presti,<sup>1,†</sup> and Matteo G. A. Paris<sup>1,‡</sup>

<sup>1</sup>*Quantum Optics & Information Group,<sup>§</sup> Istituto Nazionale di Fisica della Materia, Dipartimento di Fisica "A. Volta,"  
Università di Pavia, via Bassi 6, I-27100 Pavia, Italy*

<sup>2</sup>*Istituto Nazionale di Fisica Nucleare, Sezione di Pavia, Pavia, Italy*

<sup>3</sup>*Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois 60208*  
(Received 8 September 2001; published 12 December 2001)

We show how entanglement can be used to improve the estimation of an unknown transformation. Using entanglement is always of benefit in improving either the precision or the stability of the measurement. Examples relevant for applications are illustrated, for either qubits or continuous variables.

DOI: 10.1103/PhysRevLett.87.270404

PACS numbers: 03.65.Ud, 03.65.Ta

Entanglement is certainly the most distinctive feature of quantum mechanics. The quantum nonlocality due to entanglement, which has puzzled generations of theoreticians since the work of Einstein, Podolsky, and Rosen [1], in the last decade eventually has been harnessed for practical use in the new quantum information technology [2,3]. Entanglement has become the essential resource for quantum computing, quantum teleportation, and secure cryptographic protocols [3]. Recently, entanglement has been proved as a valuable resource for improving optical resolution [4], spectroscopy [5], quantum lithography [6], and has shown to be a crucial ingredient for making the tomography of a quantum device [7], with a single input entangled state playing the role of all possible states at the input of the device—another manifestation of the *quantum parallelism*, the feature of entanglement that is the core of quantum computing algorithms [8,9].

In this Letter we will show how in general entanglement can be used to improve quantum measurements, for either precision or stability. The measurement scheme will be considered in the general framework of quantum estimation theory [10], in which one needs to estimate the parameter  $\theta$  of the density operator  $\rho_\theta$  on the Hilbert space  $\mathcal{H}$  as the result of a unitary transformation  $\rho \rightarrow \rho_\theta = U_\theta \rho U_\theta^\dagger$ —more generally a quantum operation  $Q_\theta$  could be considered, with  $\rho_\theta = Q_\theta(\rho)$ , corresponding to a parameter of any physical (amplifying, measuring, etc.) device. This situation for known input state  $\rho$  is very common in practice, e.g., in interferometry [11], and more generally whenever the measurement is *indirect*, resorting to the detection of a change in an ancillary part of the measuring apparatus. In this scenario we will consider the use of an entangled input state  $R$  in place of  $\rho$ , with the unknown transformation  $U_\theta$  acting locally only on one side of the entangled state. In tensor notation,  $R \rightarrow R_\theta = U_\theta \otimes I R U_\theta^\dagger \otimes I$ . The situation is depicted in Fig. 1. As we will see in this Letter, the entangled configuration is better than the conventional one, for either precision or stability of the measurement. This is due to the fact that, in some sense, the input entangled state is

equivalent to many input states in “quantum parallel.” In the following we will examine different measurement situations separately, and we will draw general conclusions at the end.

*Covariant measurements.*—In a covariant measurement the parameter  $\theta$  is the element  $g \in \mathbf{G}$  of a group  $\mathbf{G}$  of transformations. This kind of measurement has been thoroughly analyzed in Ref. [12].

Let us first illustrate the mechanism of entanglement on a simple example. We want to discriminate among the four unitary transformations represented by the Pauli matrices  $\sigma_0 \equiv I$ ,  $\sigma_1 \equiv \sigma_x$ ,  $\sigma_2 \equiv \sigma_y$ ,  $\sigma_3 \equiv \sigma_z$ . As is well known, they form a unitary discrete group [13]. By applying the four transformations to any single-qubit input state  $|\psi\rangle \in \mathbb{C}^2$  we always obtain four linearly dependent states, which makes the conventional scheme in Fig. 1 useless for a reliable discrimination. On the contrary, if we apply the four matrices to the maximally entangled input state  $\frac{1}{\sqrt{2}}|I\rangle\rangle$  we obtain the four Bell states  $\sigma_j \otimes I \frac{1}{\sqrt{2}}|I\rangle\rangle \equiv \frac{1}{\sqrt{2}}|\sigma_j\rangle\rangle$ , which are mutually orthogonal. Here we use the notation  $|A\rangle\rangle \doteq \sum_{ij} A_{ij}|i\rangle|j\rangle \equiv A \otimes |I\rangle\rangle$ , which puts vectors  $|A\rangle\rangle \in \mathcal{H} \otimes \mathcal{H}$  into correspondence with operators  $A$  on  $\mathcal{H}$ ,  $A_{ij}$  denoting the matrix elements of  $A$  on the fixed basis  $\{|i\rangle\}$  for  $\mathcal{H}$ , and  $I$  being the

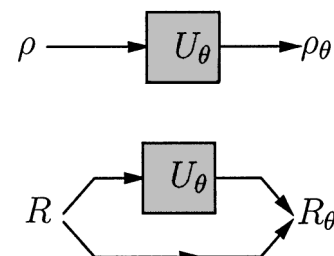


FIG. 1. Measurement schemes considered in the present Letter. The parameter  $\theta$  of the density operator  $\rho_\theta$  is estimated as the result of a unitary transformation  $\rho \rightarrow \rho_\theta = U_\theta \rho U_\theta^\dagger$  (up figure). In this scenario the use of an entangled input  $R$  in place of  $\rho$  is considered, with the unknown transformation  $U_\theta$  acting locally on one Hilbert space only (down figure).

identity operator. This simple example is very instructive: the discrimination among the four Pauli transformations  $\sigma_j$ , which is impossible with a single qubit input state, becomes possible and exact when applying  $\sigma_j$  to a maximally entangled state. The mechanism is clear: using an entangled state instead of a single qubit doubles the dimension of the Hilbert space  $\mathcal{H}_{\text{out}}$  spanned by the output states, allowing perfect discrimination of the four  $\sigma_j$ . This example can be generalized easily to any dimension  $d$ , when discriminating among the  $d^2$  unitary transformations  $U(m, n) = \sum_{k=0}^{d-1} e^{2\pi i km/d} |k\rangle \langle k \oplus n|$ ,  $n$  and  $m$  ranging in  $0-(d-1)$ , and  $\oplus$  denoting addition modulo  $d$  [14]. Now, using the maximally entangled state  $\frac{1}{\sqrt{d}}|I\rangle\rangle$  at the input will produce the  $d^2$  orthogonal output states  $\frac{1}{\sqrt{d}}|U(m, n)\rangle\rangle$ , which allows perfect discrimination among all  $U(m, n)$ , whereas a nonentangled input  $|\psi\rangle \in \mathcal{H}$  would output  $d^2$  linearly dependent states in the  $d$ -dimensional  $\mathcal{H}$ . More generally, let us consider a set of unitary transformations  $\{U_g\}$ ,  $g \in \mathbf{G}$  that form a (projective) representation of the group  $\mathbf{G}$ , i.e.,  $U_g U_h = \omega(g, h) U_{gh}$ , where  $\omega(g, h)$  is a suitable phase [15]. For simplicity let us consider the case of an irreducible representation (the reducible case is technically more complicated, and needs the knowledge of all irreducible components on invariant subspaces). For every operator  $O$  on  $\mathcal{H}$ , from the Schur's lemma one has the trace identity

$$[U_g O U_g^\dagger]_{\mathbf{G}} = \text{Tr}[O], \quad (1)$$

where  $[f(g)]_{\mathbf{G}}$  denotes the group averaging  $[f(g)]_{\mathbf{G}} \doteq \frac{1}{|\mathbf{G}|} \sum_{g \in \mathbf{G}} f(g)$  with suitable normalization, with  $|\mathbf{G}|$  the cardinality of  $\mathbf{G}$ . Equation (1) generalizes to the continuous case for group averaging  $[f(g)]_{\mathbf{G}} \doteq \int_{\mathbf{G}} dg f(g)$ ,  $dg$  being a (normalized) invariant measure on  $\mathbf{G}$ . For a general input state  $|E\rangle\rangle \in \mathcal{H} \otimes \mathcal{H}$ , the Hilbert space  $\mathcal{H}_{\text{out}}$  spanned by the output states is the support of the operator  $O = [|\Psi_g\rangle\rangle \langle\langle \Psi_g|]_{\mathbf{G}}$ , with  $\Psi_g = U_g E$ . One has  $O = I \otimes \text{Tr}_1[|E\rangle\rangle \langle\langle E|] = I \otimes (E^\dagger E)^T$ ,  $\text{Tr}_1$  representing the partial trace over the first Hilbert space, and  $T$  denoting transposition with respect to the basis  $\{|i\rangle\rangle$  for  $\mathcal{H}$ . Therefore,  $\dim(\mathcal{H}_{\text{out}}) = d \times \text{rank}(E)$ , and since  $\text{rank}(E)$  is equal to the Schmidt number of  $|E\rangle\rangle$  [3], we conclude that an entangled input always increases the dimension of  $\mathcal{H}_{\text{out}}$ ; i.e., it improves the precision of the measurement.

Since the Schmidt number does not depend on the actual amount of entanglement of  $|E\rangle\rangle$ , a more refined goodness criterion can be given in terms of the Holevo bound [3]  $\chi = S(\frac{1}{d}[|\Psi_g\rangle\rangle \langle\langle \Psi_g|]_{\mathbf{G}}) - \frac{1}{d}[S(|\Psi_g\rangle\rangle \langle\langle \Psi_g|)]_{\mathbf{G}}$  for the information accessible from the measurement,  $S$  denoting the von Neumann entropy. Equation (1) gives  $\chi = \log d + S[E^\dagger E]$ ; i.e., the bound is increased exactly of the amount of entanglement  $S[E^\dagger E]$  of the input state.

With the measurement problem addressed in a maximum likelihood strategy, it is easy to see that the optimal POVM  $d\Pi_g$  is of the form

$$d\Pi_g = dg(U_g \otimes I)P(U_g^\dagger \otimes I), \quad (2)$$

with  $P \geq 0$  a positive operator on  $\mathcal{H} \otimes \mathcal{H}$  normalized as  $\text{Tr}_1[P] = I$ . By covariance, the maximum average likelihood is equal to  $\langle\langle E|P|E\rangle\rangle \leq d$ , since normalization limits the maximum eigenvalue of  $P$  below  $d$ . The bound is saturated for  $E = d^{-1/2}U$ , with  $U$  unitary, i.e., for maximally entangled input, and  $P = |U\rangle\rangle \langle\langle U|$ .

Another way to see the optimality of a maximally entangled input is to notice that the average overlap  $\Omega(E) \doteq [|\langle\langle \Psi_g|E\rangle\rangle|^2]_{\mathbf{G}} \equiv \text{Tr}[(E^\dagger E)^2]$  is a Schur convex function of the reduced density operator  $E^\dagger E$ . Following Ref. [16], this implies that if  $|A\rangle\rangle \prec |B\rangle\rangle$  ( $|A\rangle\rangle$  is "majorized" by  $|B\rangle\rangle$ ), then  $\Omega(A) \leq \Omega(B)$ , whence the minimum averaged overlap between output states comes from a maximally entangled input, since this is majorized by any other state. That the optimal estimation strategy to discriminate among unitaries needs entangled inputs has also been noticed in Ref. [17].

As an example in infinite dimensions, consider the problem of estimating the displacement of a harmonic oscillator in the phase space, i.e., the parameter  $\alpha \in \mathbb{C}$  of the transformation  $\rho \rightarrow \rho_\alpha = D(\alpha)\rho D^\dagger(\alpha)$ , where  $D(\alpha) = \exp(\alpha a^\dagger - \bar{\alpha}a)$  is the displacement operator for annihilation and creation operators  $a$  and  $a^\dagger$ , respectively (in this case  $\mathbf{G}$  is the Weyl-Heisenberg group). For unentangled  $\rho$ , an estimation of  $\alpha$  isotropic on  $\mathbb{C}$  is equivalent to an optimal joint measurement of position and momentum, which, as well known, is affected by an unavoidable minimum noise of 3 dB [18]. Here the optimal state (for fixed minimum energy) is the vacuum, and the corresponding conditional probability of measuring  $z$  given  $\alpha$  is  $p(z|\alpha) = \pi^{-1} \exp[-|z - \alpha|^2]$ . Now, consider the case in which the estimation is made with  $D(\alpha)$  acting on the entangled state

$$|E\rangle\rangle = \sqrt{1 - |x|^2} \sum_{n=0}^{\infty} x^n |n\rangle |n\rangle, \quad (3)$$

with  $|x| \leq 1$  [the state (3) can be achieved by parametric down-conversion of vacuum]. Here we can use the orthonormal resolution of the identity  $|D(z)\rangle\rangle \langle\langle D(z)|$  of eigenvectors  $|D(z)\rangle\rangle$  of  $Z = a \otimes I - I \otimes a^\dagger$  with eigenvalue  $z$  (this is just a heterodyne measurement [19]), now achieving  $p(z|\alpha) = (\pi\Delta^2)^{-1} \exp[-\Delta^{-2}|z - \alpha|^2]$ , with variance  $\Delta^2 = \frac{1-|x|}{1+|x|}$  that, in principle, can be decreased at will with the state (3) approaching an eigenstate of  $Z$  (by increasing the gain of the down-converter).

*Measurement in the presence of noise.*—What happens if the estimation is performed in the presence of noise, namely the channel before and after the unknown transformation is affected by noise? Here it is instructive to reconsider the problem of estimating the displacement of a harmonic oscillator in the phase space in the presence of Gaussian displacement noise, which maps states as follows:

$$\rho \rightarrow \Gamma_{\bar{n}}(\rho) \doteq \int_{\mathbb{C}} \frac{d^2\gamma}{\pi\bar{n}} \exp[-|\gamma|^2/\bar{n}] D(\gamma)\rho D^\dagger(\gamma). \quad (4)$$

The variance  $\bar{n}$  of the noise is usually referred to as “mean thermal photon number.” The case of Gaussian displacement noise is particularly simple, since one has the composition law  $\Gamma_{\bar{n}} \circ \Gamma_{\bar{m}} = \Gamma_{\bar{n}+\bar{m}}$ , and, moreover  $\Gamma_{\bar{n}}[D(\alpha)\rho D^\dagger(\alpha)] = D(\alpha)\Gamma_{\bar{n}}(\rho)D^\dagger(\alpha)$ . Therefore, if the measurement is made on the entangled state (3), one can easily derive a Gaussian conditional probability distribution with variance  $\delta^2 = \Delta^2 + 2\bar{n}_T$ , where  $\bar{n}_T$  is the total Gaussian displacement noise before and after the displacement  $D(\alpha)$ , and the noise is doubled since it is supposed equal on the two entangled Hilbert spaces. On the other hand, in the measurement scheme with unentangled input (note that the optimal is the vacuum), one has  $\delta^2 = 1 + \bar{n}_T$ . One concludes that the entangled input is no longer convenient above one thermal photon  $\bar{n}_T = 1$  of noise. This is exactly the threshold of noise above which the entanglement is totally degraded to a separable state [20], and therefore the quantum capacity of the noisy channel vanishes [21].

*Discrimination between two unitaries.*—What about when one discriminates between only two unitary transformations? It is clear that here using an entangled input as in Fig. 1 would be of no benefit, since it is useless to increase the dimensionality of  $\mathcal{H}_{\text{out}}$ . However, we will see that in this case a multipartite entanglement would allow *perfect discrimination for a finite number of copies* of the transformation to be determined.

In an optimized strategy [10] the minimum error probability in the discrimination of the two output states  $U_1|\psi\rangle$  and  $U_2|\psi\rangle$  for any (also entangled) input state  $|\psi\rangle$  is

$$P_E = \frac{1}{2} [1 - \sqrt{1 - 4p_1p_2|\langle\psi|U_2^\dagger U_1|\psi\rangle|^2}], \quad (5)$$

$p_1$  and  $p_2$  being the *a priori* probability of the two transformations. For simplicity, in the following we set  $p_1 = p_2 = \frac{1}{2}$ . Clearly, the optimum input states  $|\psi\rangle$  are those minimizing the overlap  $|\langle\psi|U_2^\dagger U_1|\psi\rangle|$ . It is easy to show that the minimum overlap is given by [22]

$$\min_{\|\psi\|=1} |\langle\psi|U_2^\dagger U_1|\psi\rangle| = r(U_2^\dagger U_1), \quad (6)$$

where  $r(W)$  denotes the distance between the origin of the complex plane and the polygon whose vertices are the eigenvalues of the unitary operator  $W$ . Moreover, optimizing the overlap over entangled  $|\psi\rangle$  gives again Eq. (6) [23]. From the rule of the minimum overlap (6) we conclude that the discrimination is perfect if and only if  $z = 0 \in K(U_2^\dagger U_1)$ , namely the polygon of the eigenvalues of  $W = U_2^\dagger U_1$  encircles the origin. Then, it is obvious that an entangled input as in Fig. 1 would be of no use, since  $W$  and  $W \otimes I$  have the same spectrum. However, the situation changes dramatically if one has  $N$  copies

of the unitary transformation  $U = U_{1,2}$  to be determined, and a  $N$ -partite entangled state is available for a measurement scheme as in Fig. 2. Now the spectrum of  $W^{\otimes N}$  must be considered, and the angular spread  $\Delta(W)$  of the eigenvalues is increased as  $\Delta(W^{\otimes N}) = \min(N\Delta(W), 2\pi)$  [ $\Delta(W)$  is the angle subtended at the origin by the polygon of eigenvalues of  $W$ ]. There, *the discrimination is always exact for sufficiently many uses  $N$* . This result should be compared to the case of *state* discrimination. There, for nonorthogonal states the probability of failure is always nonvanishing for any  $N$ . Here, instead, for nonorthogonal transformations the discrimination among unitaries is always exact for  $N$  sufficiently large. It is clear that the above arguments could be extended to the case of multiple testing, whenever the strategy leads to an overlap criterion (as, for example, in Ref. [24]). That exact discrimination between unitaries is virtually possible, for a finite number  $N$  of uses has also been noticed in Ref. [25].

*Improving the stability of the measurement.*—In the instances in which the optimal discrimination between transformations is already optimized by an unentangled input, an entangled state can still be better in achieving a more stable sensitivity. We have seen that an unentangled input is already optimal in the discrimination of (one use of) two unitaries. An unentangled input is also optimal in the covariant measurement for Abelian  $\mathbb{G}$ , since the irreducible representations are one dimensional. Consider, for example, the problem of distinguishing among displacements on a fixed direction of the phase space, say  $D(x)$ , with  $x \in \mathbb{R}$ . In this case one could use a squeezed state  $|x_0\rangle_s \doteq \exp\{\frac{s}{2}[(a^\dagger)^2 - a^2]\}D(x_0)|0\rangle$ , with  $s > 0$ , i.e., squeezed in the direction of the “quadrature”  $X = \frac{1}{2}(a^\dagger + a)$ . Then, a conditional Gaussian probability with variance  $\langle\Delta X^2\rangle = \frac{1}{4}e^{-2s}$  is obtained, which can be narrowed at will by using  $n_s = \sinh^2 s$  squeezing photons. However, if the phase of the quadrature is slightly mismatched, and the quadrature  $X_\phi = \frac{1}{2}(a^\dagger e^{i\phi} + a e^{-i\phi})$  is measured instead, then the variance becomes

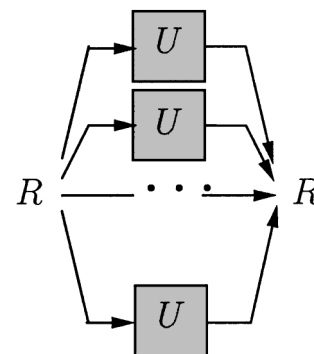


FIG. 2. When testing between two unitaries  $U = U_{1,2}$  it is possible to achieve perfect discrimination even for nonorthogonal  $U_1$  and  $U_2$  for sufficiently large number  $N$  of copies of the unitary transformation, if a  $N$ -partite entangled state is available for a measurement scheme as figure (see text).

$\langle \Delta X_\phi^2 \rangle = \frac{1}{4}(e^{2s} \sin^2 \phi + e^{-2s} \cos^2 \phi)$ , and the sensitivity is exponentially unstable. Using the entangled input in Eq. (3), instead, gives the same Gaussian noise  $\Delta^2 = \frac{1-|x|}{1+|x|}$ , independently of  $\phi$ , by using  $n = 2|x|^2/(1 - |x|^2)$  down-converted photons.

*Further generalizations and conclusions.*—Up to now we have focused our analysis only on discrimination among unitaries; however, we could have considered more generally nonunitary quantum operations, to see that entanglement is still a useful resource for improving the measurement. For the case of two operations  $Q_1$  and  $Q_2$  the distinguishability is related to the completely bounded (cb) norm [21]  $\|p_1 Q_1 - p_2 Q_2\|_{cb}$  which is the supremum over all possible entangled input states of the trace distance between the output states. Since the cb norm is equivalent to the usual trace norm for completely positive maps, it follows that an unentangled state already achieves optimality in the special case that the difference  $p_1 Q_1 - p_2 Q_2$  is completely positive.

In conclusion, we have seen that entanglement is a useful resource for upgrading the quantum measurements which are based on the estimation of a quantum transformation. It is always of benefit, in improving either precision or stability. In many cases the measurement precision becomes in principle unbounded, even when the conventional measurement is noise limited. The upgrading is effective in the presence of noise, below the threshold of total entanglement degradation.

This work has been funded by the EC program ATESIT, Contract No. IST-2000-29681. G. M. D. acknowledges support by DARPA Grant No. F30602-01-2-0528.

\*Email address: [dariano@unipv.it](mailto:dariano@unipv.it)

†Email address: [lopresti@unipv.it](mailto:lopresti@unipv.it)

‡Email address: [paris@unipv.it](mailto:paris@unipv.it)

§Electronic addresses: [www.qubit.it](http://www.qubit.it),  
[www.quantummechanics.it](http://www.quantummechanics.it),  
[www.quantumoptics.it](http://www.quantumoptics.it)

- [1] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
- [2] *Introduction to Quantum Computation and Information*, edited by H.-K. Lo, S. Popescu, and T. Spiller (World Scientific, Singapore, 1998).
- [3] I. L. Chuang and M. A. Nielsen, *Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, U.K., 2000).
- [4] M. I. Kolobov and C. Fabre, *Phys. Rev. Lett.* **85**, 3789 (2000).
- [5] B. E. A. Saleh, B. M. Jost, H.-B. Fei, and M. C. Teich, *Phys. Rev. Lett.* **80**, 3483 (1998).
- [6] M. D'Angelo, M. V. Chekhova, and Y. Shih, LANL arXive quant-ph/0103035.

- [7] G. M. D'Ariano and P. Lo Presti, *Phys. Rev. Lett.* **86**, 4195 (2001).
- [8] P. W. Show, in *Proceedings of the 35th Annual Symposium of the Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, Los Alamitos, CA, 1994), p. 124.
- [9] L. K. Grover, in *Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (STOC)*, 1996, p. 212 (e-print quant-ph/9605043); *Phys. Rev. Lett.* **79**, 325 (1997).
- [10] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [11] J. H. Shapiro and S. R. Shepard, *Phys. Rev. A* **43**, 3795 (1991).
- [12] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
- [13] The Pauli matrix forms a projective non-Abelian irreducible representation of the (Abelian) dihedral group  $D_2$  of  $\pi$  rotations in three dimensions.
- [14] The unitary operators  $U(m, n)$  form a projective non-Abelian irreducible representation of the (Abelian) group  $\mathbb{Z}_d \times \mathbb{Z}_d$  describing translations on a two-dimensional lattice embedded in a torus. The dihedral group  $D_2$  corresponds to the particular case  $d = 2$ .
- [15] The projective representations are unitary group representations with an additional phase  $\omega(g, h)$  in the composition law called "cocycle." The cocycle must satisfy the Jacobi associativity constraints, namely that  $\omega(gh, l)\omega(g, h) = \omega(g, hl)\omega(h, l)$  and  $\omega(g, g^{-1}) = \omega(g, e) = 1$ , for  $g, h, l \in \mathbf{G}$ ,  $e$  being the identity element.
- [16] M. A. Nielsen, *Phys. Rev. Lett.* **83**, 436 (1999).
- [17] A. Acín, E. Jané, and G. Vidal, *Phys. Rev. A* **64**, 050302 (2001).
- [18] E. Arthurs and M. S. Goodman, *Phys. Rev. Lett.* **60**, 2447 (1988).
- [19] H. P. Yuen and J. H. Shapiro, *IEEE Trans. Inf. Theory* **IT-26**, 78 (1980).
- [20] R. Simon, *Phys. Rev. Lett.* **84**, 2726 (2000).
- [21] A. S. Holevo and R. F. Werner, *Phys. Rev. A* **63**, 032312 (2001).
- [22] In fact, by expanding  $|\psi\rangle = \sum_j \psi_j |w_j\rangle$  over the eigenvectors  $|w_j\rangle$  of  $W = U_2^\dagger U_1$  with eigenvalue  $e^{i\gamma_j}$ , consider the convex set  $K(W) \doteq \{z = \sum_j |\psi_j|^2 e^{i\gamma_j}, \|\psi\|^2 = 1\}$ , namely the polygon whose vertices are the eigenvalues of  $W$ . Then, the minimum of  $|\langle \psi | U_2^\dagger U_1 | \psi \rangle|$  is just the minimum  $|z|$  with  $z \in K(W)$ , namely the distance of the polygon from the origin of the complex plane.
- [23] Using entangled states  $|\psi\rangle \equiv |E\rangle$  corresponds to finding the minimum of  $|\text{Tr}[W(E^\dagger E)^T]|$ , which is the average  $\sum_j \rho_j |\langle \lambda_j | U_2^\dagger U_1 | \lambda_j \rangle|$  for  $(E^\dagger E)^T = \sum_j \rho_j |\lambda_j\rangle \langle \lambda_j|$ , whence the minimum distance is always obtained for a  $E^\dagger E$  pure, i.e., for an unentangled  $|\psi\rangle$ .
- [24] H. P. Yuen, R. S. Kennedy, and M. Lax, *IEEE Trans. Inf. Theory* **IT-21**, 125 (1975).
- [25] A. Acín, *Phys. Rev. Lett.* **87**, 177901 (2001).