Fidelity Balance in Quantum Operations

Konrad Banaszek

Center for Quantum Information and Rochester Theory Center for Optical Science and Engineering, University of Rochester, Rochester, New York 14627

and Instytut Fizyki Teoretycznej, Uniwersytet Warszawski, Hoża 69, PL-00-681 Warszawa, Poland

(Received 1 May 2000; revised manuscript received 30 June 2000)

I derive a tight bound between the quality of estimating the state of a single copy of a d-level system, and the degree the initial state has to be altered in the course of this procedure. This result provides a complete analytical description of the quantum mechanical trade-off between the information gain and the quantum state disturbance expressed in terms of mean fidelities. I also discuss consequences of this bound for quantum teleportation using nonmaximally entangled states.

DOI: 10.1103/PhysRevLett.86.1366 PACS numbers: 03.67.–a, 03.65.Ta

As a general rule, the more information is obtained from an operation on a quantum system, the more its state has to be altered. This heuristic statement was first exemplified by the Heisenberg microscope *gedankenexperiment* [1], where the spatial resolution of the apparatus was shown to scale inversely with the uncertainty of the momentum transferred during the observation. Presently, the disturbance caused by the information gain has become an issue of practical significance, as it underlies the security of quantum key distribution [2].

The balance between the information gain and the state disturbance attracts currently a lot of interest, particularly in the context of quantum cryptography [3]. Information theory provides a selection of concepts to quantify both the information gain and the state disturbance. The choice of measures for these two effects is usually dictated by the relevance to a specific application. In most cases, however, derivation of the actual balance represents a highly nontrivial task, especially if one is tempted to resign from numerical means. The purpose of this Letter is to present a formulation of the information gain versus state disturbance trade-off which is completely solvable using elementary analytical techniques. This formulation is motivated by recent works on quantum state estimation [4], where the information obtained from the operation is converted into an estimate for the initial state of the system.

The problem considered in this Letter can be formulated as follows. Suppose we are given a single d-level particle in a completely unknown pure state $|\psi\rangle$. We want to make a guess about the quantum state of this particle, but at the same time we would like to alter the state as little as possible. One can associate two fidelities with such a procedure. The first one, which we will denote by F, describes how much the state after the operation resembles the original one. The second fidelity, denoted by G, characterizes the average quality of our guess. It is natural to expect a trade-off between these two quantities: the more information is extracted from the system, i.e., the larger G, the less the final state should resemble the initial one, hence the smaller F should be. What is the actual quantitative bound between F and G?

Two extreme cases are well known: if nothing is done to the particle we have $F = 1$, but then our guess about the state of the particle has to be random, which yields $G = 1/d$. On the other hand, the optimal estimation strategy for a single copy [5] yields $G = 2/(d + 1)$, but then the particle after the operation cannot provide any more information on the initial state; thus also $F = 2/(d + 1)$. I prove here that quantum mechanics imposes a general constraint between F and G in the form of the following inequality:

$$\sqrt{F - \frac{1}{d+1}} \leq \sqrt{G - \frac{1}{d+1}} + \sqrt{(d-1)\left(\frac{2}{d+1} - G\right)}. \tag{1}$$

I also show that this inequality cannot be further improved; i.e., there exist quantum operations saturating the equality sign.

The most general strategy that can be applied to the particle has the form of a trace-preserving operation described by a set of operators $\hat{A}_r$, where $r = 1, \ldots, N$. These operators satisfy the completeness relation

$$\sum_{r=1}^{N} \hat{A}_r^\dagger \hat{A}_r = \hat{\mathbb{1}}. \tag{2}$$

The classical information gained from this operation is given by the index r, which is subsequently used to estimate the initial state of the particle. The outcome r of the operation performed on a state $|\psi\rangle$ is obtained with the probability $\langle\psi|\hat{A}_r^\dagger \hat{A}_r|\psi\rangle$. This corresponds to the following conditional transformation of the quantum state [6]:

$$|\psi\rangle \rightarrow \frac{\hat{A}_r|\psi\rangle}{\sqrt{\langle\psi|\hat{A}_r^\dagger \hat{A}_r|\psi\rangle}}. \tag{3}$$

We shall measure the resemblance of the transformed state to the original one using the squared modulus of the scalar

 0031-9007/01/86(7)/1366(4)$15.00

product, equal to $|\langle\psi|\hat{A}_r|\psi\rangle|^2/\langle\psi|\hat{A}_r^\dagger\hat{A}_r|\psi\rangle$. Summation of this expression over r with the weights $\langle\psi|\hat{A}_r^\dagger\hat{A}_r|\psi\rangle$, and integration over all possible input states $|\psi\rangle$, yields the complete expression for the mean operation fidelity F:

$$F = \int d\psi \sum_{r=1}^{N} |\langle\psi|\hat{A}_r|\psi\rangle|^2. \tag{4}$$

Here the integral $\int d\psi$ over the space of pure states is performed using the canonical measure invariant with respect to the group of unitary transformations on the state vectors of the particle.

Given the outcome r of the operation, we can make a guess $|\psi_r\rangle$ what the state originally was. The quality of this guess, assuming that the initial state was $|\psi\rangle$, can be quantified with the help of the overlap $|\langle\psi_r \mid \psi\rangle|^2$. The mean estimation fidelity G is given by the average of this expression over all outcomes r with the probability distribution $\langle\psi|\hat{A}_r^\dagger\hat{A}_r|\psi\rangle$, and by integration over states $|\psi\rangle$:

$$G = \int d\psi \sum_{r=1}^{N} \langle\psi|\hat{A}_r^\dagger\hat{A}_r|\psi\rangle |\langle\psi_r \mid \psi\rangle|^2. \tag{5}$$

We will start derivation of the trade-off between the fidelities F and G by evaluating the integrals over $|\psi\rangle$. For this purpose, let us introduce in Eq. (4) two decompositions of unity in a certain orthonormal basis $|i\rangle$:

$$\begin{aligned} F &= \sum_{r=1}^{N} \sum_{i,j=0}^{d-1} \int d\psi \langle\psi \mid i\rangle\langle i|\hat{A}_r^\dagger|\psi\rangle\langle\psi|\hat{A}_r| j\rangle\langle j \mid \psi\rangle \\ &= \sum_{r=1}^{N} \sum_{i,j=0}^{d-1} \langle i|\hat{A}_r^\dagger \hat{M}_{ij}\hat{A}_r| j\rangle, \end{aligned} \tag{6}$$

where by $\hat{M}_{ij}$ we have denoted the following integrals of projectors on the states $|\psi\rangle\langle\psi|$:

$$\begin{aligned} \hat{M}_{ij} &= \int d\psi \langle\psi \mid i\rangle\langle j \mid \psi\rangle |\psi\rangle\langle\psi| \\ &= \frac{1}{d(d+1)} (\delta_{ij}\hat{\mathbb{1}} + |i\rangle\langle j|). \end{aligned} \tag{7}$$

The second explicit form of the operators $\hat{M}_{ij}$ has been derived in Ref. [7]. This formula allows us to simplify the expression for the mean operation fidelity F to the form

$$\begin{aligned} F &= \frac{1}{d(d+1)} \left(\sum_{i=0}^{d-1} \sum_{r=1}^{N} \langle i|\hat{A}_r^\dagger\hat{A}_r|i\rangle + \sum_{r=1}^{N} \left| \sum_{i=0}^{d-1} \langle i|\hat{A}_r|i\rangle \right|^2 \right) \\ &= \frac{1}{d(d+1)} \left(d + \sum_{r=1}^{N} |\mathrm{Tr}\hat{A}_r|^2 \right). \end{aligned} \tag{8}$$

Let us now consider the estimation fidelity G. The guess $|\psi_r\rangle$ can be represented as a result of a certain unitary transformation $\hat{U}_r$ acting on a reference state, which we will take for concreteness to be $|0\rangle$,

$$|\psi_r\rangle = \hat{U}_r|0\rangle. \tag{9}$$

Using this representation, and changing the integration measure in Eq. (5) according to $|\psi\rangle \rightarrow \hat{U}_r|\psi\rangle$, we can evaluate the integral over $|\psi\rangle$,

$$\begin{aligned} G &= \sum_{r=1}^{N} \int d\psi |\langle 0 \mid \psi\rangle|^2 \langle\psi|\hat{U}_r^\dagger\hat{A}_r^\dagger\hat{A}_r\hat{U}_r|\psi\rangle \\ &= \sum_{r=1}^{N} \mathrm{Tr}(\hat{U}_r^\dagger\hat{A}_r^\dagger\hat{A}_r\hat{U}_r\hat{M}_{00}). \end{aligned} \tag{10}$$

Inserting the explicit form of the operator $\hat{M}_{00} = (\hat{\mathbb{1}} + |0\rangle\langle 0|)/[d(d+1)]$ yields

$$\begin{aligned} G &= \frac{1}{d(d+1)} \left(\sum_{r=1}^{N} \mathrm{Tr}(\hat{U}_r^\dagger\hat{A}_r^\dagger\hat{A}_r\hat{U}_r) + \sum_{r=1}^{N} \langle 0|\hat{U}_r^\dagger\hat{A}_r^\dagger\hat{A}_r\hat{U}_r|0\rangle \right) \\ &= \frac{1}{d(d+1)} \left(d + \sum_{r=1}^{N} \langle\psi_r|\hat{A}_r^\dagger\hat{A}_r|\psi_r\rangle \right). \end{aligned} \tag{11}$$

This expression provides directly a recipe for optimal assignment of guesses $|\psi_r\rangle$ to outcomes of the operation: each of the components $\langle\psi_r|\hat{A}_r^\dagger\hat{A}_r|\psi_r\rangle$ in the sum over r is maximized if $|\psi_r\rangle$ is the eigenvector of $\hat{A}_r^\dagger\hat{A}_r$ corresponding to its maximum eigenvalue. Consequently, the maximum value of the mean estimation fidelity G for a given operation $\{\hat{A}_r\}$ can be written as

$$G = \frac{1}{d(d+1)} \left(d + \sum_{r=1}^{N} \|\hat{A}_r\|^2 \right), \tag{12}$$

where the operator norm is defined in the standard way,

$$\|\hat{A}_r\| = \sup_{\langle\varphi \mid \varphi\rangle = 1} \sqrt{\langle\varphi|\hat{A}_r^\dagger\hat{A}_r|\varphi\rangle}. \tag{13}$$

In order to relate the fidelities F and G to each other, let us consider the singular-value decomposition [8] of the operators $\hat{A}_r$,

$$\hat{A}_r = \hat{V}_r\hat{D}_r\hat{W}_r, \tag{14}$$

where $\hat{V}_r$ and $\hat{W}_r$ are unitary, and $\hat{D}_r$ is a semipositive definite diagonal matrix,

$$\hat{D}_r = \sum_{i=0}^{d-1} \lambda_i^r |i\rangle\langle i|, \tag{15}$$

with the diagonal elements put in a decreasing order: $\lambda_0^r \geq \cdots \geq \lambda_{d-1}^r \geq 0$. We will first show that only the diagonal matrices $\hat{D}_r$ are relevant to the trade-off. Indeed, the modulus of the trace of the matrix $\hat{A}_r$ appearing in Eq. (8) is bounded by

$$\begin{aligned} |\mathrm{Tr}\hat{A}_r| &= \left| \sum_{i=0}^{d-1} \langle i|\hat{W}_r\hat{V}_r\hat{D}_r|i\rangle \right| \\ &\leq \sum_{i=0}^{d-1} \lambda_i^r |\langle i|\hat{W}_r\hat{V}_r|i\rangle| \leq \sum_{i=0}^{d-1} \lambda_i^r, \end{aligned} \tag{16}$$

and moreover any quantum operation can easily be modified in such a way that the equality sign is reached. What needs to be done is to follow the operation $\{\hat{A}_r\}$ with an extra unitary transformation $\hat{W}_r^\dagger \hat{V}_r^\dagger$ depending on the outcome r. Let us note that this corresponds to the modification of the operation according to $\hat{A}_r \rightarrow \hat{W}_r^\dagger \hat{V}_r^\dagger \hat{A}_r$, which makes each element of the operation a semipositive Hermitian operator. As we are interested in the maximum value of F, we can further assume with no loss of generality that

$$F = \frac{1}{d(d+1)}\left[d + \sum_{r=1}^{N}\left(\sum_{i=0}^{d-1}\lambda_i^r\right)^2\right]. \tag{17}$$

The expression for the estimation fidelity written in terms of λ_i^r takes the form

$$G = \frac{1}{d(d+1)}\left(d + \sum_{r=1}^{N}(\lambda_0^r)^2\right). \tag{18}$$

In addition, the trace of the completeness condition given in Eq. (2) yields the following constraint on λ_i^r:

$$\sum_{r=1}^{N}\sum_{i=0}^{d-1}(\lambda_i^r)^2 = d\,. \tag{19}$$

To complete the proof of the inequality (1), it is convenient to introduce vector notation. Let us define d real vectors $\mathbf{v}_i = (\lambda_i^1, \ldots, \lambda_i^N)$, where the index i runs from 0 to $d-1$. Sums over r appearing in Eqs. (17) and (18) can be written as

$$f = \sum_{r=1}^{N}\left(\sum_{i=0}^{d-1}\lambda_i^r\right)^2 = \sum_{i,j=0}^{d-1}\mathbf{v}_i\cdot\mathbf{v}_j\,, \tag{20}$$

$$g = \sum_{r=1}^{N}(\lambda_0^r)^2 = |\mathbf{v}_0|^2, \tag{21}$$

where the dot denotes the scalar product, and $|\cdot|$ is the standard quadratic norm. The completeness condition (19) for the operation $\{\hat{A}_r\}$ written in the vector notation takes the form

$$\sum_{i=0}^{d-1}|\mathbf{v}_i|^2 = d\,. \tag{22}$$

Let us now suppose that the vector $\mathbf{v}_0$ is fixed. The estimation fidelity is then given by $G = (d + |\mathbf{v}_0|^2)/[d(d+1)]$. What is the maximum operation fidelity F that can be achieved with this constraint? The answer to this question is provided by an application of the Schwarz inequality to Eq. (20):

$$f \le \sum_{i,j=0}^{d-1}|\mathbf{v}_i||\mathbf{v}_j| = \left(\sum_{i=0}^{d-1}|\mathbf{v}_i|\right)^2 = \left(\sqrt{g} + \sum_{i=1}^{d-1}|\mathbf{v}_i|\right)^2. \tag{23}$$

We have excluded here from the sum over i the norm of the vector $\mathbf{v}_0$ which is fixed and equal to $\sqrt{g}$. The sum of the norms of the remaining vectors can be estimated using the inequality between the arithmetic and quadratic means,

$$\frac{1}{d-1}\sum_{i=1}^{d-1}|\mathbf{v}_i| \le \sqrt{\frac{1}{d-1}\sum_{i=1}^{d-1}|\mathbf{v}_i|^2} = \sqrt{\frac{d-g}{d-1}}\,, \tag{24}$$

where we have evaluated the sum $\sum_{i=1}^{d-1}|\mathbf{v}_i|^2$ using Eq. (22). Inserting this bound into Eq. (23) we finally obtain the inequality

$$f \le [\sqrt{g} + \sqrt{(d-1)(d-g)}]^2, \tag{25}$$

which expressed in terms of the fidelities F and G takes the form of Eq. (1).

The necessary and sufficient conditions for a quantum operation to reach the equality sign can most easily be formulated in the vector notation. The Schwarz inequality (23) becomes equality if all the vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{d-1}$ are collinear. Furthermore, the equality sign in Eq. (24) holds if and only if $|\mathbf{v}_1| = \cdots = |\mathbf{v}_{d-1}|$. It is straightforward to see that an exemplary operation satisfying these conditions for a given estimation fidelity $G = (1 + g/d)/(d+1)$ is defined by

$$\begin{aligned}\hat{A}_r = {} & \sqrt{\frac{g}{d}}\,|r-1\rangle\langle r-1| \\ & + \sqrt{\frac{d-g}{d(d-1)}}\,(\hat{\mathbb{1}} - |r-1\rangle\langle r-1|)\,,\end{aligned} \tag{26}$$

where the index r runs from 1 to d, and the projectors $|r-1\rangle\langle r-1|$ are constructed using any orthonormal basis. This confirms the inequality (1) is indeed a tight one and cannot be further improved.

A simple transformation of Eq. (1) shows that the quantum mechanically allowed region for the fidelities F and G is bounded by a quadratic curve, which turns out to be a fragment of an ellipse given by the equation

$$\begin{aligned}&(F-F_0)^2 + d^2(G-G_0)^2 + \\ &\quad 2(d-2)(F-F_0)(G-G_0) = \frac{d-1}{(d+1)^2}\end{aligned} \tag{27}$$

with $F_0 = (d+2)/(2d+2)$ and $G_0 = 3/(2d+2)$. The shape of the region for several values of d is depicted in Fig. 1.

The balance between the operation and estimation fidelities derived in this Letter has interesting consequences in quantum teleportation based on nonmaximally entangled states. If two parties share a pure bipartite state of the Schmidt form $|\mathrm{tele}\rangle = \sum_{k=0}^{d-1}\mu_k|k\rangle\otimes|k\rangle$, then the maximum teleportation fidelity attainable using this state is given by [7,9]

$$F_{\mathrm{tele}} = \frac{1 + (\sum_{k=0}^{d-1}\mu_k)^2}{d+1}\,. \tag{28}$$

Furthermore, for a nonmaximally entangled state the measurement performed during the teleportation protocol reveals some information on the teleported state. This information can be converted into an estimate for the initial state, whose maximum average fidelity has been shown to equal [7]

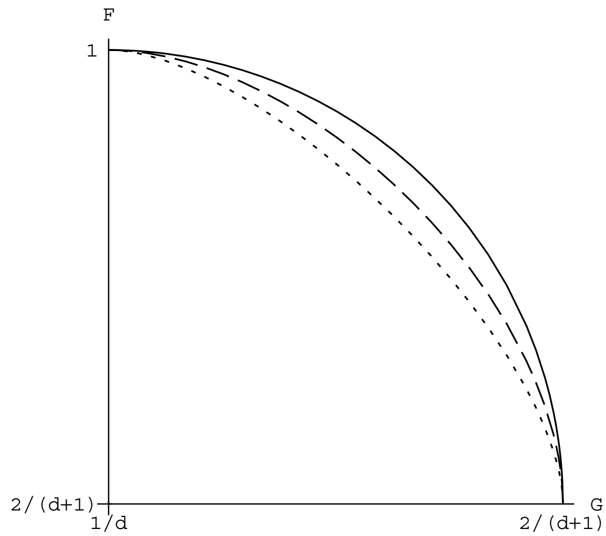

FIG. 1. Rescaled bound for the operation fidelity F versus the estimation fidelity G, plotted for $d = 2$ (solid line), $d = 4$ (dashed line), and $d = 8$ (dotted line).

$$G_{\mathrm{tele}} = \frac{1 + \mu_0^2}{d + 1}, \tag{29}$$

where μ_0 denotes the largest Schmidt coefficient for the state $|\mathrm{tele}\rangle$. As the procedure of teleportation can be viewed as a special case of a quantum operation [10], the bound (1) applies as well to the pair of fidelities F_{tele} and G_{tele}. Consequently, for a given teleportation fidelity F_{tele}, the maximum value of the estimation fidelity is achieved for the state $|\mathrm{tele}\rangle$ satisfying the condition $\mu_1 = \cdots = \mu_{d-1} = \sqrt{(1 - \mu_0^2)/(d - 1)}$. This condition defines a class of pure bipartite states which are optimal from the point of view of the trade-off between the teleportation fidelity and the estimation fidelity.

In conclusion, I have obtained a tight bound for the fidelities describing the quality of estimating the state of a single copy of a d-level particle, and the degree the initial state has to be changed during this operation. This result seems to be one of very few cases, when the trade-off between the information gain and the state disturbance can be derived in a closed analytical form.

I acknowledge useful discussions with J.H. Eberly, C.A. Fuchs, N. Lütkenhaus, V. Vedral, and I.A. Walmsley. This research was partially supported by ARO-administered MURI Grant No. DAAG-19-99-1-0125, NSF Grant No. PHY-9415583, and KBN Grant No. 2 P03B 089 16.

[1] W. Heisenberg, *The Physical Principles of the Quantum Theory* (Dover, New York, 1930). Let us note that the Heisenberg uncertainty relation can be assigned two meanings. The first one refers to variances of the position and the momentum when these quantities are determined from disjoint subsets of an ensemble of identically prepared particles. The second meaning, illustrated by the microscope experiment and relevant in the context of information gain versus disturbance, describes a bound on the resolution of an apparatus measuring *simultaneously* both the conjugate variables. For a detailed discussion, see K. Wódkiewicz, Phys. Lett. A **124**, 207 (1987); S. Stenholm, Ann. Phys. (N.Y.) **218**, 233 (1992) and references therein.

[2] C.H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.

[3] A.K. Ekert, B. Huttner, G.M. Palma, and A. Peres, Phys. Rev. A **50**, 1047 (1994); C.A. Fuchs and A. Peres, *ibid.* **53**, 2038 (1996); E. Biham and T. Mor, Phys. Rev. Lett. **78**, 2256 (1997); N. Lütkenhaus, Phys. Rev. A **59**, 3301 (1999); H. Bechmann-Pasquinucci and N. Gisin, *ibid.* **59**, 4238 (1999).

[4] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995); R. Derka, V. Bužek, and A.K. Ekert, *ibid.* **80**, 1571 (1998); G. Vidal, J.I. Latorre, P. Pascual, and R. Tarrach, Phys. Rev. A **60**, 126 (1999).

[5] D. Bruss and C. Macchiavello, Phys. Lett. A **253**, 149 (1999); A. Acín, J.I. Latorre, and P. Pascual, Phys. Rev. A **61**, 022113 (2000).

[6] We restrict our attention to the class of ideal operations, which preserve the purity of the state for each conditional transformation. This introduces no loss of generality, as for the purpose of further considerations all nonideal operations can be represented as ideal ones with an additional constraint on the assignment of the guesses to the classical outcomes.

[7] K. Banaszek, Phys. Rev. A **62**, 024301 (2000).

[8] K.E. Atkinson, *An Introduction to Numerical Analysis* (Wiley, New York, 1978), Sec. 7.2.

[9] G. Vidal, D. Jonathan, and M.A. Nielsen, Phys. Rev. A **62**, 012304 (2000).

[10] M.A Nielsen and C.M. Caves, Phys. Rev. A **55**, 2547 (1997).