Getting Information on Independently Prepared Quantum States — When Are Individual Measurements as Powerful as Joint Measurements?

Chi-Hang Fred Fung and H. F. Chau

Department of Physics and Center of Theoretical and Computational Physics,
University of Hong Kong, Pokfulam Road, Hong Kong

Given a composite quantum system in which the states of the subsystems are independently (but not necessarily identically) prepared, we construct separate measurements on the subsystems from any given joint measurement such that the former always give at least as large information as the latter. This construction offers new insights into the understanding of measurements on this type of composite systems. Moreover, this construction essentially proves the intuition that separate measurements on the subsystems are sufficient to extract the maximal information about the separately prepared subsystems, thus making a joint measurement unnecessary. Furthermore, our result implies that individual attacks are as powerful as collective attacks in obtaining information on the raw key in quantum key distribution.

PACS numbers: 03.67.-a, 03.67.Dd, 87.19.lo, 89.70.Cf

I. INTRODUCTION

Quantum states can be used to convey information. A sender, Alice, may prepare a few quantum particles, whose states depend on the message itself, and send them through a quantum channel to the receiver, Bob. To determine the message, Bob performs a quantum measurement on his received quantum states. When the quantum states live in multiple quantum subsystems, Bob may perform separate quantum measurements on the subsystems to learn about the message. Alternatively, he may perform a joint quantum measurement on all subsystems together. In general, performing separate quantum measurements on the subsystems is not powerful enough to extract maximal information on the input state. In fact, the capacities of certain quantum channels \[1, 2\] and the maximum information that can be extracted from certain unentangled but classically correlated states \[3\] can only be attained via joint measurements. But what if each subsystem is independently used to convey information? Perhaps measuring each subsystem separately is already good enough to extract maximum amount of information on the states in each subsystem. Here we prove this intuition by explicitly constructing an individual measurement from a given joint measurement such that information gain from former is at least as large as the latter. This construction offers new insights into the understanding of measurements on this type of composite systems. Furthermore, we explain the operational meaning of such construction and discuss its implication to quantum key distribution (QKD).

II. PRECISE DEFINITION OF OUR PROBLEM

Suppose that there are \(K\) subsystems. For each subsystem \(k\), Alice selects a state indexed by \(a_k\) from a set of normalized density matrices \(\{\phi_{a_k}^{(k)}\}\) with probability \(p_{a_k}^{(k)}\).

(In other words, \(\text{Tr}(\phi_{a_k}^{(k)}) = 1\) and \(\sum_{a_k} p_{a_k}^{(k)} = 1\). Furthermore, we do not limit the the number of subsystems \(K\), the Hilbert space dimension of each subsystem and the number of elements in the set \(\{\phi_{a_k}^{(k)}\}\) for each \(k\). These three numbers may well be infinite.) The state in each subsystem is selected independently but not necessarily identically. Suppose that Bob uses a particular joint positive operator-valued measure (POVM) to measure the \(K\) subsystems. In general, this POVM may contain elements that are entangled with the \(K\) subsystems. The purpose of this paper is to construct an individual measurement in the \(K\) subsystems that can extract no less Shannon mutual information about Alice’s states than the original joint POVM. Here individual measurement refers to the one composed of \(K\) independent POVM’s each operating on one subsystem. Thus, by showing that such an individual measurement exists, we confirm the intuition that maximal information on separately prepared subsystems can be extracted separately.

We present two methods for constructing such an individual measurement from the original joint measurement. Both methods draw on the observation that knowing the states of the other subsystems gives rise to a projected measurement on a subsystem. The first method is simpler to apply, while the second one admits an intuitive explanation for why it gives at least as large information as the original joint measurement. We also provide the operational meaning for the second method.

III. CONSTRUCTION 1

Without lost of generality, let us consider the case of having two subsystems (that is, \(K = 2\)) and denote the original joint POVM as \(\{M_b : \forall b\}\). The case of \(K > 2\) can be constructed and proven in a similar way. The key idea of constructing the individual measurement is to focus on measuring one particular subsystem and look at what effective measurement is performed on it. Thus, let
us focus, say, on subsystem 1 ($S_1$). If Alice always prepares subsystem 2 ($S_2$) in the state $\phi$, then the effective measurement on $S_1$ is

$$\{\text{Tr}_2(\langle I \otimes \phi \rangle M_b) : \forall b\}. \quad (1)$$

More generally, if Alice may prepare the state in $S_2$ in more than one way, then for each state sent by Alice in $S_2$, there corresponds a set of POVM elements similar to Eq. (1). The entire POVM is then composed of all these sets. One may regard the set of POVM elements corresponding to a state in $S_2$ as the effective measurement on $S_1$ when Alice sends that state. Therefore, how likely Bob uses this set of POVM elements should be weighted by the a priori probability of the corresponding state being sent. In summary, the effective POVM for $S_1$ is $\{M_{b_1}^{(1)} : \forall b_1\}$ where

$$M_{b_1}^{(1)} = \text{Tr}_2(\langle I \otimes \phi_a^{(2)} \rangle M_b)p_a^{(2)}, \quad (2)$$

and $b_1 \equiv (a_2, b)$ is the index of the POVM element specifying an input state in $S_2$ and an element of the original joint POVM. Using the same argument, the effective POVM $\{M_{b_2}^{(2)} : \forall b_2\}$ for $S_2$ is

$$M_{b_2}^{(2)} = \text{Tr}_1(\langle \phi_a^{(1)} \otimes I \rangle M_b)p_a^{(1)}, \quad (3)$$

where $b_2 \equiv (a_1, b)$ is similarly defined.

**Theorem 1.** Suppose Alice prepares the states of subsystems $k = 1, 2$ independently. Then, the amount of Shannon mutual information provided by an individual measurement on the two subsystem using the POVM's $\{M_{b_k}^{(k)}\}, k = 1, 2$ whose elements are given in Eq. (2) and Eq. (3) is at least as large as the Shannon mutual information provided by the original joint POVM $\{M_b\}$.

**Proof.** Let the capitalized symbols $A_1$, $A_2$, and $B$ denote the random variables for the input states $a_1$ and $a_2$ and the original joint POVM outcome $b$, respectively. Since the two subsystems are independent, the mutual information for the original joint POVM $\{M_b\}$ is given by

$$I_{1+2} \triangleq I(A_1, A_2; B) = H(A_1, A_2) - H(A_1, A_2|B) = H(A_1) + H(A_2) - H(A_1, A_2|B). \quad (4)$$

Here, the functions $I(\cdot; \cdot)$ and $H(\cdot)$ are the mutual information between its arguments and the entropy of its argument, respectively. The mutual information between the input and the output of subsystem $k = 1, 2$ is

$$I_k \triangleq I(A_k; B_k) = H(A_k) - H(A_k|B_k). \quad (5)$$

To prove this theorem, it suffices to show that $I_1 + I_2 \geq I_{1+2}$ which can be expressed as

$$H(B_1) - H(A_1, B_1) + H(B_2) - H(A_2, B_2) \geq H(B) - H(A_1, A_2, B). \quad (6)$$

We proceed by establishing a crucial relationship between the joint probability of the overall system and that of each subsystem. The former, with inputs $A_1$ and $A_2$ and output $B$, is given by

$$\Pr\{A_1 = a_1, A_2 = a_2, B = b\} = \text{Tr}[\phi_a^{(1)} \otimes \phi_a^{(2)} M_b]p_1^{(1)}p_2^{(2)}, \quad (7)$$

while the latter, with input $A_k$ and output $B_k$ for subsystem $k$, is given by

$$\Pr\{A_k = a_k, B_k = b_k\} = \text{Tr}[\phi^{(k)}_a M_{b_k}]p_a^{(k)}. \quad (8)$$

Here, the POVM element $M_{b_k}^{(k)}$ is given in Eq. (2) or Eq. (3). We relate these two probabilities for say $S_1$ by expanding the POVM element in Eq. (8) as follows:

$$\Pr\{A_1 = a_1, B_1 = (a_2, b)\} = \text{Tr}_1[\phi_a^{(1)} \text{Tr}_2(\langle I \otimes \phi_a^{(2)} \rangle M_b)p_a^{(2)}]p_a^{(1)} \equiv \Pr\{A_1 = a_1, A_2 = a_2, B = b\}. \quad (9)$$

This crucial relationship between the probabilities directly translates into a relationship between the entropies:

$$H(A_1, B_1) = \sum_{a_1, b_1} f(\Pr\{A_1 = a_1, B_1 = b_1\})$$

$$= \sum_{a_1, (a_2, b)} f(\Pr\{A_1 = a_1, B_1 = (a_2, b)\})$$

$$= \sum_{a_1, a_2, b} f(\Pr\{A_1 = a_1, A_2 = a_2, B = b\}) = H(A_1, A_2, B) \quad (10)$$

where $f(x) = -x \log_2 x$. Replacing $S_1$ by $S_2$, we have

$$H(A_1, B_1) = H(A_1, A_2, B) = H(A_2, B_2). \quad (11)$$

By the same token, we know that

$$H(B_1) = \sum_{a_2, b} f(\sum_{a_1} \Pr\{A_1 = a_1, A_2 = a_2, B = b\})$$

$$= H(A_2, B) \quad (12)$$

and

$$H(B_2) = H(A_1, B). \quad (13)$$

From Eqs. (11), (13), Eq. (6) is reduced to the well-known entropy inequality in (classical) information theory

$$H(A_1|B) \geq H(A_1|A_2, B). \quad (14)$$

Therefore, this theorem is proved.
IV. CONSTRUCTION 2

Recall that the effective POVM for each subsystem (given in Eq. (2) and Eq. (3)) is a mixture of sub-POVM’s each corresponding to a state sent in the other subsystem. Now the key observation is that Bob can use any of these sub-POVM’s on one subsystem irrespective of the actual state sent in the other. That is to say, Bob can use on $S_2$ the sub-POVM corresponding to one state in $S_1$ even though Alice has really sent another state in $S_1$. Therefore, among all sub-POVM’s for a particular subsystem, we can pick the one that provides the highest mutual information. This sub-POVM, alone, then constitutes the effective POVM for that subsystem. And this construction results in the effective POVM $\{M_b^{(1)} : \forall b\}$ for $S_1$ where

$$M_b^{(1)} = \text{Tr}_2[(I \otimes \phi_a^{(2)})M_b],$$

and $\phi_a^{(2)}$ is chosen to be one of the possible states of $S_2$ so that $\{M_b^{(1)}\}$ maximizes the mutual information for $S_1$. Similarly, the elements of the effective POVM $\{M_b^{(2)} : \forall b\}$ for $S_2$ are

$$M_b^{(2)} = \text{Tr}_1[(\phi_a^{(1)} \otimes I)M_b],$$

where $\phi_a^{(1)}$ is chosen to be one of the possible states of $S_1$ so that $\{M_b^{(2)}\}$ maximizes the mutual information for $S_2$.

Theorem 2. Suppose Alice prepares the states of subsystems $S_1,S_2$ independently. Then, the amount of Shannon mutual information provided by an individual measurement using the POVM’s whose elements are defined in Eq. (15) for $S_1$ and in Eq. (16) for $S_2$ is at least as large as the Shannon mutual information provided by the original joint POVM $\{M_b\}$.

Proof. We focus on $S_1$ as the case of $S_2$ is similar. It suffices to show that the mutual information for the POVM whose elements are defined in Eq. (15) is no less than that in Eq. (2), and invoke Theorem 1. Observe that

$$I(A_1;B_1) = H(A_1) - [H(A_1,B_1) - H(B_1)]$$
$$= H(A_1) - [H(A_1,A_2,B) - H(A_2,B)]$$
$$= H(A_1|A_2) - [H(A_1,B|A_2) - H(B|A_2)]$$
$$= \sum_{a_2} p_{a_2}^{(2)} I(A_1;B|A_2 = a_2)$$
$$\leq \max_{a_2} I(A_1;B|A_2 = a_2),$$

where the second line is due to Eqs. (11) and (12), and the third line is due to the fact that the states in the two subsystems are independent and that one can arbitrarily add and subtract $H(A_2)$. We proceed to verify that $I(A_1;B|A_2 = a_2)$ is indeed the mutual information for the sub-POVM consisting of elements given in Eq. (15). The probability of observing outcome $b$ with input $A_1 = a_1$ corresponding to Eq. (15) equals

$$\text{Tr}_1[\phi_a^{(1)}(\text{Tr}_2[(I \otimes \phi_b^{(2)})M_b])P_a^{(1)} = \text{Pr}(A_1 = a_1,B = b|A_2 = a_2).$$

This means that the corresponding mutual information for this POVM is $I(A_1;B|A_2 = a_2)$. Therefore, Eq. (17) shows that indeed the POVM of Eq. (2) can be broken down into sub-POVM’s each corresponding to one value of $a_2$. Thus, when Bob always uses the sub-POVM corresponding to the $a_2$ that maximizes $I(A_1;B|A_2 = a_2)$, the resulting mutual information is no less than that of using the weighted average of the sub-POVM’s.

V. INTUITIVE EXPLANATION OF CONSTRUCTION 2

Let us introduce two phantom subsystems (intended to be thrown away later) in addition to the two real subsystems. The phantom subsystem $P_k$ serves to replicate the phantom subsystem $P_k$ serves to replicate.
real subsystem $S_k$ for $k = 1, 2$ in the sense that they share
the same set of states in which Alice may send with the
same a priori probabilities. Nevertheless, they are inde-
pendent of each other and of other subsystems. Since all
the states sent by Alice in the four subsystems are inde-
pendent, the pair consisting of $S_1$ and $S_2$ and the pair
consisting of $S_1$ and $P_2$ appear to be identical to Bob.
Thus, the amount of information Bob can learn about
$S_1$ from measuring the first pair and that from measur-
ing the second pair using the same joint measurement
must be same. Because of this, we may consider that
Bob performs the joint measurement on $S_1$ and $P_2$ (see
Fig. 1(a)).

Now suppose that Alice tells Bob exactly which state
was sent in $P_2$ (and we will show that delaying this an-
nouncement indefinitely turns out to have no bearing on
Bob). Using this extra piece of information, Bob can pick
the corresponding POVM elements that are consist-
tent with the phantom state and project it onto $S_1$ as
a measurement operator (see Fig. 1(b)). Interestingly,
this projected measurement turns out to be the effective
measurement we have constructed in Eq. (15) for
various values of $a_2$. Essentially, for each state in $P_2$ an-
nounced by Alice, there corresponds an effective POVM
for $S_1$. Clearly, with the aid of the extra information in
the state of $P_2$, Bob’s information on $S_1$ in this case is
at least as large as that could be obtained with the original
joint measurement on the two systems when Alice did
not disclose the state of $P_2$.

Now the key point is that Bob can use any of these
effective POVM’s on $S_1$ irrespective of the actual state
sent in $P_2$. This is because Bob’s information on $S_1$ ob-
tained from using a particular effective POVM does not
depend on the state of $P_2$ as $P_2$ and $S_1$ are independent.
Therefore, we can regard that Bob always ignores Alice’s
announcement of the state in $P_2$ and uses the effective
POVM on $S_1$ that gives him the maximum amount of informa-
tion (see Fig. 1(c)). When Bob always uses only
one effective POVM on $S_1$, the existence of $P_2$ is irrele-
vant and thus we can completely discard $P_2$ along with
the announcement of its state (see Fig. 1(d)). Since Bob
always uses the best effective POVM on $S_1$, the amount
of information he gets on $S_1$ is at least as large as that
when he chooses the POVM based on Alice’s announce-
ment, which we have already argued is no worse than
that when he uses the original joint measurement.

We repeat the previous argument on the pair $S_2$ and
$P_1$ to obtain the best effective POVM for $S_2$. Finally,
the independence of $S_1$ and $S_2$ allows us to conclude that
using the best effective POVM for each of them gives no
less information on both as the original joint POVM.

VI. MULTIPARTITE SYSTEMS

Our results for the bipartite case given by Theorems 1
and 2 can easily be extended to the multipartite case (in-
cluding the case of an infinite number of subsystems). In
particular, the POVM elements corresponding to Theo-
rem 1 for a $K$-partite system are

$$M^{(k)}_{b_k} = \text{Tr}_{\ell \neq k} \left[ \bigotimes_{\ell=1}^{K} \phi_{a^{(\ell,k)}_a} M_\ell \right] \prod_{\ell \neq k} p_{a^{(\ell)}} \forall k, \quad (18)$$

where $\phi_{a^{(\ell,k)}_a} = \phi_{a^{(\ell)}}$ if $\ell \neq k$ and $\phi_{a^{(\ell,k)}_a} = I$ otherwise.

VII. IMPLICATION TO QUANTUM KEY
DISTRIBUTION

The result in this paper sheds some light on the var-
ious types of eavesdropping attacks in QKD [8, 9]. In
most QKD protocols such as the famous BB84 proto-
col [3], a legitimate party (Alice) sends a sequence of
quantum states each independently chosen from a set of
states to another legitimate party (Bob) through a hos-
tile channel controlled by an eavesdropper (Eve). The
goal of Alice and Bob is to derive a secret key from Al-
ice’s states and Bob’s states. Eve, on the other hand,
tries to steal their secret by launching an eavesdropp-
ing attack. Two types of keys can be distinguished:
the raw key and the final secret key. Alice’s raw key is
the bit string corresponding to the quantum states she
sends to Bob; whereas Bob’s raw key corresponds to his
measurement results on the received qubits [17]. Their
raw keys may not be secure and error-free; and they
derive their final keys from their raw keys via privacy
amplification. For QKD protocols in which Alice sends
out independent states (such as BB84 [3], SARG04 [7],
and Gaussian-modulated coherent states QKD [8]), Eve’s
probes become independent and our result in this paper
implies that individual attacks are as powerful as col-
lective attacks in obtaining information on Alice’s raw
key [13]. In contrast, Smith [3] shows that when the key
generation rate is concerned, collective attacks are
strictly more powerful. This makes sense since privacy
amplification correlates Alice’s raw keys in order to ob-
tain the final secret key.

VIII. CONCLUSIONS

We show that individual measurement is sufficient to
obtain optimal amount of information on the states in
which each subsystem is prepared independently but not
necessary identically based on the observation that know-
ing the state of the other subsystems gives rise to a pro-
jected measurement on a subsystem. Applying our result
to the QKD setting shows that individual and collective
attacks are equally powerful in obtaining information on
the raw key. Our work uses Shannon mutual information
as the information measure.

We note that Wootters has proved the same result as
ours that the accessible information is additive for in-

dependently prepared subsystems [10]. However, implementing his proof idea will result in an ensemble of individual measurements (each with a fixed probability of being drawn) for each subsystem. In contrast, both our construction methods lead to a single measurement for each subsystem.

Constructing individual measurements from a joint one giving at least as large information in terms of other information measures may be possible. For example, the Csiszár measure (see, e.g., [11, 12]), which is a generalization of Shannon information, allows such a construction in a special case [19].

Acknowledgments

We thank Debbie W. Leung, Hoi-Kwong Lo, Xiongfeng Ma, Marco Piani, and Ming-Yong Ye for enlightening discussions. This work is supported by the RGC grant No. HKU 701007P of the HKSAR Government and the Postdoctoral Fellowship program of NSERC of Canada.

[1] A. S. Holevo, IEEE Trans. Inf. Theory 44, 269 (1998).
[2] B. Schumacher and M. D. Westmoreland, Phys. Rev. A 56, 131 (1997).
[3] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A 59, 1070 (1999).
[4] T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley-Interscience, 2006), 2nd ed.
[5] C. H. Bennett and G. Brassard, in Proc. of IEEE Int. Conference on Computers, Systems, and Signal Processing (IEEE Press, New York, 1984), pp. 175–179.
[6] A. K. Ekert, Phys. Rev. Lett. 67, 661 (1991).
[7] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. 92, 057901 (2004).
[8] F. Grosshans, G. Van Assche, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier, Nature 421, 238 (2003).
[9] G. Smith (2007), arXiv:0705.3838 [quant-ph].
[10] Private communication between W. K. Wootters and D. P. DiVincenzo et al. as cited in Footnote 4 of D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, IEEE Trans. Inf. Theory 48, 580 (2002).
[11] K. Ferentinos and T. Papaioannou, Information and Control 51, 193 (1981).
[12] K. Zografos, K. Ferentinos, and T. Papaioannou, The Canadian Journal of Statistics 14, 355 (1986).
[13] I. Csiszár and J. Körner, IEEE Trans. Inf. Theory 24, 339 (1978).
[14] B. Kraus, N. Gisin, and R. Renner, Phys. Rev. Lett. 95, 080501 (2005).
[15] R. Renner, N. Gisin, and B. Kraus, Phys. Rev. A 72, 012332 (2005).
[16] M. Koashi, J. Phys. Conf. Ser. 36, 98 (2006).
[17] We do not consider noisy pre-processing here. See Refs. [13, 14, 15] for discussions on noisy pre-processing.
[18] We assume that Eve only uses her probes and Alice’s announce- ment of the basis information to learn about Alice’s raw key; in particular, she does not make use of the error correction information that Alice may reveal publicly. Note that error correction information may be transmitted by Alice in encrypted form, as is the case of Koashi’s security proof [16]. In this case, Eve is unable to use the error correction information.
[19] The Csiszár measure (see, e.g., [11, 12]) is defined as

\[ I_C(f, g) = \sum_x g(x)\Phi \left( \frac{f(x)}{g(x)} \right) \]

where \( \Phi \) is convex. Shannon mutual information \( I(X; Y) \) is recovered with \( I(X; Y) = I_C(p(x)p(y), p(x, y)) \) and \( \Phi(x) = -\log x \). Keeping the Csiszár measure in the form \( I_C(p(x)p(y), p(x, y)) =: I_C(X; Y) \), if \( \Phi \) satisfies the condition \( \Phi(x) + \Phi(y) \geq \Phi(xy) \) with \( x, y \geq 0 \), then we have \( I_C(A_1; E|A_2) + I_C(A_2; E) \geq I_C(A_1, A_2; E) \). By borrowing techniques in our paper and also that due to Wootters, it is not difficult to construct the desired individual measurements for the Csiszár measure from this inequality.