Quantum union bounds for sequential projective measurements

Jingliang Gao*

State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China

(Received 12 September 2015; published 24 November 2015)

We present two quantum union bounds for sequential projective measurements. These bounds estimate the disturbance accumulation and probability of outcomes when the measurements are performed sequentially. These results are based on a trigonometric representation of quantum states and should have wide application in quantum information theory for information-processing tasks such as communication and state discrimination, and perhaps even in the analysis of quantum algorithms.

DOI: 10.1103/PhysRevA.92.052331 PACS number(s): 03.67.−a, 03.65.Ta

In order to acquire information from a quantum system, we must perform a quantum measurement on it. According to quantum theory, a von Neumann measurement yields an eigenvalue of the measured observable with a probability given by the Born rule, and simultaneously this measurement could disturb the measured system. However, by weakly coupling the measuring device to the system, it is possible to read out certain information while limiting the disturbance to the system [1].

In some cases, one can perform a sequence of measurements in order to acquire the desired information, but the situation becomes more complex as the number of measurements increases. Although a single measurement does not necessarily disturb the system in some cases, the disturbance could potentially accumulate gradually when the measurements are performed in a sequential fashion. So some natural questions for sequential measurements are as follows: *Can we bound the accumulated disturbance in a meaningful way or, related to this, understand how many measurements can be performed until the final state is no longer close to the initial state?* Moreover, performing a larger number of measurements results in a variety of possible sequences. *Then how can we estimate the probability of occurrence of the resulting sequences?*

Having sharp answers to these questions would be very helpful in analyzing many situations, such as quantum property testing [2], quantum sequential decoding [3–6], sequential state discrimination [7,8], quantum tomography [9], or any other task which requires a large number of measurements. In former work, Aaronson presented a union bound for general measurements [10]. Thereafter, Sen proposed a significantly improved bound for projective measurements [5], his bound now being known as the "noncommutative union bound." Wilde then generalized Sen's bound to apply to general measurements and analyzed classical communication over a single instance of a quantum channel with this approach [6].

In this paper, we present some useful bounds for sequential projective measurements which can be used to estimate the disturbance and the probability of occurrence separately. Our results given here strengthen previously known results from Refs. [5] and [6], and we establish them by employing a trigonometric representation of quantum states. As an example of the application, we provide general formulas for the sequential decoding strategy [3–5].

*gaojingliang@stu.xidian.edu.cn

We begin by clarifying what we mean by a sequential measurement. Suppose that the initial state of a quantum system is given by the density operator ρ. Now we perform a sequence of measurements on the system. Specifically, we first perform a two-outcome measurement $\mathcal{M}_1$ on ρ and obtain a postmeasurement state ρ_1. Then we perform another two-outcome measurement $\mathcal{M}_2$ on ρ_1 and obtain the postmeasurement state ρ_2. Next, we perform a third two-outcome measurement $\mathcal{M}_3$ on ρ_2 and obtain ρ_3. And so it carries on, with each measurement being performed on the state resulting from the previous measurement. After N measurements, we obtain the state ρ_N. It should be emphasized that the final state ρ_N can take many forms because each step has several possible results. Without loss of generality, we suppose that each measurement is given by $\mathcal{M}_i = \{P_i, I - P_i\}$ for $i = 1,\dots,N$, where P_i are projectors. (The generality of this approach follows from [6, Lemma 3.1].) Now, suppose we are only interested in the case in which each measurement gives the outcome corresponding to P_i rather than $I - P_i$. In other words, the desired postmeasurement state sequence is as follows:

$$\rho_1 = \frac{P_1 \rho P_1}{\mathrm{tr}\,(P_1 \rho)},$$
$$\rho_2 = \frac{P_2 P_1 \rho P_1 P_2}{\mathrm{tr}\,(P_2 P_1 \rho P_1 P_2)},$$
$$\vdots$$
$$\rho_N = \frac{P_N \cdots P_2 P_1 \rho P_1 P_2 \cdots P_N}{\mathrm{tr}\,(P_N \cdots P_2 P_1 \rho P_1 P_2 \cdots P_N)}.$$

We can now present our main result. The disturbance and the probability of ρ_N can be estimated as stated in the following theorem:

Theorem 1. Given a density operator ρ and projectors $P_1, P_2, \dots, P_N$ such that

$$\mathrm{tr}\,(P_i \rho) = 1 - \varepsilon_i, \quad i = 1,2,\dots,N,$$

then we have the following bounds:

1-a The trace distance between ρ and ρ_N obeys

$$D(\rho,\rho_N) \leqslant 2\sqrt{\sum \varepsilon_i},$$

where $D(\rho,\rho_N) = \mathrm{tr}\,\sqrt{(\rho - \rho_N)(\rho - \rho_N)^\dagger}$.

1-b The probability of the occurrence of ρ_N obeys

$$\mathrm{tr}\,(P_N\cdots P_2P_1\rho P_1P_2\cdots P_N)\geqslant 1-4\sum\varepsilon_i.$$

The equality holds if and only if all ε_i are equal to 0.

The bound (1-a) reveals how the disturbance increases when the measurements are sequentially performed. In prior work, Wilde proposed a method [6] to guarantee that the postmeasurement state is close to the original one. He showed that one can perform the projectors P_1 through P_m and then perform them again in the opposite order. The distance between the postmeasurement state and the original one can be upper-bounded by $(\sum\varepsilon_i)^{1/4}$. The bound (1-a) improves Wilde's result and reveals that the measurements in opposite order are not necessary.

The bound (1-a) implies that the probability of occurrence of possible results may change by as much as $O((\sum\varepsilon_i)^{1/2})$. However, the bound (1-b) provides an even better estimate and controls the change to $O(\sum\varepsilon_i)$. It can be thought as a noncommutative analog of union bound from classical probability theory:

$$\Pr\{\overline{A_1\cap\cdots\cap A_N}\}=\Pr\{\overline{A}_1\cup\cdots\cup\overline{A}_N\}\leqslant\sum_{i=1}^{N}\Pr\{\overline{A}_i\},$$

where $A_1,\ldots,A_N$ are events. If we think of $P_1\cdots P_N\cdots P_1$ as the intersection of P_is, then the best analogous bound for projector logic would be

$$1-\mathrm{tr}\,(P_1\cdots P_N\cdots P_1\rho)\leqslant\sum_{i=1}^{N}\mathrm{tr}\,[(I-P_i)\rho],$$

although the above bound only holds if the projectors are commuting. For the noncommutative case, the bound (1-b) turns out to be the next best thing.

The bound (1-b) can be further generalized as follows:

Corollary 1. For projectors $P_1,P_2,\ldots,P_N$, let $\overline{P}_i=I-P_i$, then we have

$$P_1\cdots P_N\cdots P_1\geqslant I-4\sum_{i=1}^{N}\overline{P}_i.$$

Proof. This corollary is equivalent to the following: for any vector $|\nu\rangle$, it holds that

$$\langle\nu|P_1\cdots P_N\cdots P_1|\nu\rangle\geqslant\langle\nu|\nu\rangle-4\sum_{i=1}^{N}\langle\nu|\overline{P}_i|\nu\rangle.$$

Let $\rho=\frac{|\nu\rangle\langle\nu|}{\langle\nu|\nu\rangle}$; then ρ is a density operator. Applying the bound (1-b), the above inequality follows. ■

Remark 1. In prior work, Sen [5] proved that, for any positive operator ρ such that $\mathrm{tr}\,\rho\leqslant 1$, it holds that

$$\mathrm{tr}\,(P_N\cdots P_2P_1\rho P_1P_2\cdots P_N)\geqslant\mathrm{tr}\,\rho-2\sqrt{\sum\mathrm{tr}(\overline{P}_i\rho)}.$$

Corollary 1 shows that the above inequality can be enhanced to the following version:

$$\mathrm{tr}\,(P_N\cdots P_2P_1\rho P_1P_2\cdots P_N)\geqslant\mathrm{tr}\,\rho-4\sum\mathrm{tr}(\overline{P}_i\rho).$$

The new bound improves Sen's result, particularly in the "Zeno" regime where each measurement succeeds with high probability.

In the following, we detail the proof of Theorem 1. It will be first shown that the bounds hold if ρ is a pure state, and then extended to the mixed state. Our proof is based on the trigonometric representation of quantum states.

Suppose that $\rho=|\psi\rangle\langle\psi|$ is a pure state and the final state is $\rho_N=|\psi_N\rangle\langle\psi_N|$, then we have

$$|\psi_1\rangle=\frac{P_1|\psi\rangle}{\sqrt{\langle\psi|P_1|\psi\rangle}},$$
$$|\psi_2\rangle=\frac{P_2|\psi_1\rangle}{\sqrt{\langle\psi_1|P_2|\psi_1\rangle}},$$
$$\vdots$$
$$|\psi_N\rangle=\frac{P_N|\psi_{N-1}\rangle}{\sqrt{\langle\psi_{N-1}|P_N|\psi_{N-1}\rangle}}.$$

Consider the ith measurement,

$$|\psi_i\rangle=\frac{P_i|\psi_{i-1}\rangle}{\sqrt{\langle\psi_{i-1}|P_i|\psi_{i-1}\rangle}}.$$

If we let

$$|\psi_i^{\perp}\rangle=\frac{(I-P_i)|\psi_{i-1}\rangle}{\sqrt{\langle\psi_{i-1}|I-P|\psi_{i-1}\rangle}},$$

we can write $|\psi_{i-1}\rangle$ in terms of $|\psi_i\rangle$ and $|\psi_i^{\perp}\rangle$ as follows:

$$|\psi_{i-1}\rangle=\cos\theta_i|\psi_i\rangle+\sin\theta_i|\psi_i^{\perp}\rangle,\tag{1}$$

where $\theta_i=\arccos|\langle\psi_i|\psi_{i-1}\rangle|$. θ_i can be regarded as the angle between $|\psi_{i-1}\rangle$ and $|\psi_i\rangle$. The advantage of this representation is that the trace distance and probability can be expressed in a simple form [11,12]:

$$D(\psi_{i-1},\psi_i)=2\sin\theta_i,\tag{2}$$

$$\mathrm{tr}\,(P_i|\psi_{i-1}\rangle\langle\psi_{i-1}|)=|\langle\psi_i|\psi_{i-1}\rangle|^2=\cos^2\theta_i.\tag{3}$$

If we perform the measurement $\{P_i,I-P_i\}$ on ρ directly, then the resulting state would be

$$|\psi_i'\rangle=\frac{P_i|\psi\rangle}{\sqrt{\langle\psi|P_i|\psi\rangle}}$$

or

$$|\psi_i'^{\perp}\rangle=\frac{(I-P_i)|\psi\rangle}{\sqrt{\langle\psi|I-P_i|\psi\rangle}}.$$

Likewise, $|\psi\rangle$ can be written as

$$|\psi\rangle=\cos\alpha_i|\psi_i'\rangle+\sin\alpha_i|\psi_i'^{\perp}\rangle,\tag{4}$$

where $\alpha_i=\arccos|\langle\psi_i'|\psi\rangle|$. α_i is the angle between $|\psi\rangle$ and $|\psi_i'\rangle$, and it holds that

$$D(\psi,\psi_i')=2\sin\alpha_i,\tag{5}$$

$$\mathrm{tr}\,(P_i|\psi\rangle\langle\psi|)=\cos^2\alpha_i=1-\varepsilon_i.\tag{6}$$

Thus, we have

$$\sin^2\alpha_i=\varepsilon_i.\tag{7}$$

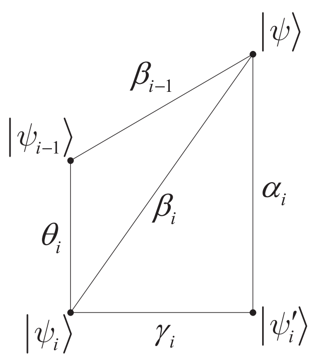

FIG. 1. The relationship between the states in the ith measurement.

We can also write $|\psi\rangle$ in terms of $|\psi_i\rangle$ and its orthogonal complement $|\psi_i^c\rangle$,

$$|\psi\rangle = \cos\beta_i|\psi_i\rangle + \sin\beta_i|\psi_i^c\rangle,$$

where $\beta_i = \arccos|\langle\psi_i|\psi\rangle|$. β_i is the angle between $|\psi\rangle$ and $|\psi_i\rangle$, and it holds that

$$D(\psi,\psi_i) = 2\sin\beta_i. \tag{8}$$

Likewise, let γ_i be the angle between $|\psi_i\rangle$ and $|\psi_i'\rangle$, then $\gamma_i = \arccos|\langle\psi_i|\psi_i'\rangle|$.

For convenience, the states and angles are shown in Fig. 1. Every vertex in the figure represents a state and the edges indicate the angles.

From the trigonometric representation of the states, we can get two important points. First, from the definition of β_i,

$$\begin{aligned}\cos\beta_i &= |\langle\psi_i|\psi\rangle| \\ &= |\cos\alpha_i\langle\psi_i|\psi_i'\rangle + \sin\alpha_i\langle\psi_i|\psi_i'^{\perp}\rangle| \\ &= \cos\alpha_i|\langle\psi_i|\psi_i'\rangle| \\ &= \cos\alpha_i\cos\gamma_i.\end{aligned} \tag{9}$$

The equality uses the fact that $P_i(I-P_i)=0$.

Second, by Eqs. (1) and (4) we have

$$\begin{aligned}\cos\beta_{i-1} &= |\langle\psi_{i-1}|\psi\rangle| \\ &= |\cos\theta_i\cos\alpha_i\langle\psi_i|\psi_i'\rangle + \sin\theta_i\sin\alpha_i\langle\psi_i^{\perp}|\psi_i'^{\perp}\rangle| \\ &\leqslant \cos\theta_i\cos\alpha_i|\langle\psi_i|\psi_i'\rangle| + \sin\theta_i\sin\alpha_i \\ &= \cos\theta_i\cos\alpha_i\cos\gamma_i + \sin\theta_i\sin\alpha_i.\end{aligned} \tag{10}$$

Equations (9) and (10) are crucial for our proof of the bounds. We will use them repeatedly in the following.

We now prove the following lemma which allows us to lower bound the disturbance in a simple way.

Lemma 1. For the ith measurement, we have

$$D^2(\psi,\psi_i) \leqslant D^2(\psi,\psi_{i-1}) + D^2(\psi,\psi_i').$$

Proof. From the trigonometric representation of the trace distance, this lemma can be equivalently stated as

$$\sin^2\beta_i \leqslant \sin^2\beta_{i-1} + \sin^2\alpha_i.$$

Furthermore, by Eq. (9), it is easy to find that

$$\sin^2\beta_i = \cos^2\alpha_i\sin^2\gamma_i + \sin^2\alpha_i.$$

Therefore, to prove the lemma, we only need to show that $\sin^2\beta_{i-1} \geqslant \cos^2\alpha_i\sin^2\gamma_i$.

Squaring Eq. (10), we have

$$\begin{aligned}\sin^2\beta_{i-1} &\geqslant 1 - (\cos\theta_i\cos\alpha_i\cos\gamma_i + \sin\theta_i\sin\alpha_i)^2 \\ &= (\sin\theta_i\cos\alpha_i\cos\gamma_i + \cos\theta_i\sin\alpha_i)^2 \\ &\quad + \cos^2\alpha_i\sin^2\gamma_i \\ &\geqslant \cos^2\alpha_i\sin^2\gamma_i.\end{aligned}$$

This completes the proof. ■

Applying Lemma 1, we can obtain that

$$\begin{aligned}D^2(\psi,\psi_N) &\leqslant D^2(\psi,\psi_{N-1}) + D^2(\psi,\psi_N') \\ &\leqslant D^2(\psi,\psi_{N-2}) + D^2(\psi,\psi_{N-1}') + D^2(\psi,\psi_N') \\ &\vdots \\ &\leqslant \sum_{i=1}^{N} D^2(\psi,\psi_i') = 4\sum_{i=1}^{N}\varepsilon_i.\end{aligned}$$

Thus, the bound (1-a) is true for a pure state.

Now let us consider the case for which ρ is a mixed state. Suppose that $|\psi\rangle^{RA}$ and $|\psi_N\rangle^{RA}$ are purifications of ρ and ρ_N, where R denotes the reference system. Let $Q_i = I^R \otimes P_i$, then the state $|\psi_N\rangle^{RA}$ is generated by performing the projective measurements $\{Q_i, I - Q_i\}$ sequentially on $|\psi\rangle^{RA}$. Moreover, the probability of each step obeys

$$\mathrm{tr}(Q_i|\psi\rangle\langle\psi|^{RA}) = \mathrm{tr}(P_i\rho) = 1 - \varepsilon_i.$$

Applying the bound for the pure state and the monotonicity of trace distance [11,12], we can obtain

$$D(\rho,\rho_N) \leqslant D\big(\psi^{RA},\psi_N^{RA}\big) \leqslant 2\sqrt{\sum\varepsilon_i}.$$

This completes the proof of the bound (1-a).

The bound (1-b) obviously holds if $\sum\varepsilon_i > \frac{1}{2}$ because the right side would be negative. In the following, we show that it still holds if $\sum\varepsilon_i \leqslant \frac{1}{2}$.

For the pure states, the condition $\sum\varepsilon_i \leqslant \frac{1}{2}$ implies that

$$0 \leqslant \alpha_i, \quad \beta_i \leqslant \frac{\pi}{4}, \quad i = 1,\ldots,N. \tag{11}$$

The probability that $|\psi_N\rangle$ occurs is

$$\begin{aligned}&\mathrm{tr}\,(P_N\cdots P_1|\psi\rangle\langle\psi|P_1\cdots P_N) \\ &\quad = \mathrm{tr}\,(P_1|\psi\rangle\langle\psi|)\cdots\mathrm{tr}\,(P_N|\psi_{N-1}\rangle\langle\psi_{N-1}|) \\ &\quad = \cos^2\theta_1\cos^2\theta_2\cdots\cos^2\theta_N.\end{aligned} \tag{12}$$

From Eq. (10), we can see that

$$\begin{aligned}\cos\beta_{N-1} &\leqslant \cos\theta_N\cos\alpha_N + \sin\theta_N\sin\alpha_N \\ &= \cos\,(\theta_N - \alpha_N),\end{aligned}$$

so it holds that $\theta_N \leqslant \beta_{N-1} + \alpha_N$. Then we have

$$\cos\theta_1\cdots\cos\theta_N \geqslant \cos\theta_1\cdots\cos\theta_{N-1}\cos\,(\beta_{N-1} + \alpha_N). \tag{13}$$

To continue, we need the following lemma:

Lemma 2. Define $\{a_k\}$ by

$$a_k = \frac{\cos\alpha_N \cos\beta_k - \sqrt{\sum_{i=k+1}^{N}\sin^2\alpha_i}\sqrt{\sin^2\beta_k + \sum_{i=k+1}^{N-1}\sin^2\alpha_i}}{1+\sum_{i=k+1}^{N-1}\sin^2\alpha_i}.$$

Then we have

$$(\cos\theta_k)a_k \geqslant a_{k-1}$$

The proof is given in Appendix A.

Note that $\cos(\beta_{N-1}+\alpha_N) = a_{N-1}$. Applying Lemma 2 repeatedly, we can get

$$\begin{aligned}
&\cos\theta_1\cdots\cos\theta_{N-2}\cos\theta_{N-1}\cos(\beta_{N-1}+\alpha_N)\\
&\quad = \cos\theta_1\cdots\cos\theta_{N-2}\cos\theta_{N-1}(a_{N-1})\\
&\quad \geqslant \cos\theta_1\cdots\cos\theta_{N-2}(a_{N-2})\\
&\quad \vdots\\
&\quad \geqslant a_0. \qquad (14)
\end{aligned}$$

Continuing, from the fact that $\beta_0 = 0$, we have

$$\begin{aligned}
a_0 &= \frac{\cos\alpha_N - \sqrt{\sum_{i=1}^{N}\sin^2\alpha_i\sum_{i=1}^{N-1}\sin^2\alpha_i}}{1+\sum_{i=1}^{N-1}\sin^2\alpha_i}\\
&\geqslant \frac{1-\sum_{i=1}^{N}\sin^2\alpha_i}{1+\sum_{i=1}^{N}\sin^2\alpha_i} \qquad (15)\\
&= \frac{1-\sum\varepsilon_i}{1+\sum\varepsilon_i} \geqslant 1-2\sum\varepsilon_i. \qquad (16)
\end{aligned}$$

The inequality (15) is proven in Appendix B.

Combining Eqs. (12)–(14) and (16), we get

$$\mathrm{tr}\,(P_N\cdots P_1|\psi\rangle\langle\psi|P_1\cdots P_N) \geqslant 1-4\sum\varepsilon_i.$$

Thus, the bound (1-b) is true for the pure states.

If ρ is a mixed state, then

$$\begin{aligned}
&\mathrm{tr}\,(P_N\cdots P_1\rho P_1\cdots P_N)\\
&\quad = \mathrm{tr}(Q_N\cdots Q_1|\psi\rangle\langle\psi|^{RA}Q_1\cdots Q_N)\\
&\quad \geqslant 1-4\sum\varepsilon_i
\end{aligned}$$

This completes the proof of bound (1-b). ■

Theorem 1 reveals how the disturbance accumulates when the measurements are performed sequentially. The generality and simplicity of the bounds imply that they should be nice tools for analyzing many situations. As an example, we show how to achieve the Holevo bound via a sequential decoding strategy. The sequential decoding scheme was first proposed by Lloyd, Giovannetti, and Maccone (LGM)[3,4]. They showed that it is possible to achieve the Holevo bound by performing sequential measurements. After the work of LGM, Sen presented a simplification of the error analysis by establishing the noncommutative bound [5]. The new bounds presented in this paper provide more general formulas for the sequential decoding strategy.

The basic sets of this problem are as follows: $\{j\}$ is a set of possible inputs to the quantum channel and $\{\sigma_j\}$ are the corresponding outputs. Let $\{p_j\}$ be a probability distribution over the indices $\{j\}$ and $\sigma \equiv \sum p_j\sigma_j$. Alice wants to send a message chosen from the set $\{1,\ldots,2^{nR}\}$ to Bob by using the quantum channel n times. The Holevo bound sets a limit on the rate R that can be achieved when the messages are transferred. We are going to outline a proof that there exists an error-correcting code that accomplishes this task with low probability of error in the limit of large n and provided $R < S(\sigma) - \sum_j p_j S(\sigma_j)$. This proof is based on the random coding and sequential decoding scheme. The transmission of messages can be decomposed into three stages: the encoding, the transmission, and the decoding. In the encoding stage, we adopt the standard random coding scheme. Alice associates with the ith message a codeword $\vec{c}_i = c_1c_2\cdots c_n$, where $c_1,c_2,\cdots c_n$ are chosen from the index set $\{j\}$ according to the distribution $\{p_j\}$. She repeats this procedure for 2^{nR} times, creating a codebook $\mathcal{C}$ of 2^{nR} entries. The corresponding output of the channel is denoted by $\sigma_{\vec{c}_i}$. When Bob receives a particular state $\sigma_{\vec{c}_m}$ he tries to determine what the message was. To do this, he has two tools: the projector P onto the δ-typical subspace of $\sigma^{\otimes n}$ and the projectors $\{P_{\vec{c}_i}\}$ onto the δ-typical subspace of the corresponding $\sigma_{\vec{c}_i}$. They have the following properties [12]: for any $\varepsilon > 0$ and sufficiently large n,

$$\mathrm{tr}\left(P\sigma^{\otimes n}\right) \geqslant 1-\varepsilon, \qquad (17)$$

$$\mathrm{tr}\left(P_{\vec{c}_i}\sigma_{\vec{c}_i}\right) \geqslant 1-\varepsilon, \qquad (18)$$

$$\mathrm{tr}\left(P_{\vec{c}_i}\right) \leqslant 2^{n[\sum p_j S(\sigma_j)+\delta]}, \qquad (19)$$

$$P\sigma^{\otimes n}P \leqslant 2^{-n[S(\sigma)-\delta]}I. \qquad (20)$$

To decode the message, Bob first performs the measurement $\{P, I-P\}$ to detect whether the received state is in the typical subspace of $\sigma^{\otimes n}$. If yes, he then asks in sequential order "Is the received codeword $\vec{c}_i$?" by performing the measurements $\{P_{\vec{c}_i}, I-P_{\vec{c}_i}\}$.

The probability of detecting $\vec{c}_m$ correctly under this sequential decoding scheme is

$$p_c = \mathrm{tr}\left(P_{\vec{c}_m}\overline{P}_{\vec{c}_{m-1}}\cdots\overline{P}_{\vec{c}_1}P\sigma_{\vec{c}_m}P\overline{P}_{\vec{c}_1}\cdots\overline{P}_{\vec{c}_{m-1}}P_{\vec{c}_m}\right).$$

Consider the expectation of p_c over all possible codes $\mathcal{C}$,

$$\begin{aligned}
\mathbb{E}_{\mathcal{C}}\{p_c\} &= \mathbb{E}_{\mathcal{C}}\left\{\mathrm{tr}\left(P_{\vec{c}_m}\overline{P}_{\vec{c}_{m-1}}\cdots\overline{P}_{\vec{c}_1}P\sigma_{\vec{c}_m}P\overline{P}_{\vec{c}_1}\cdots\overline{P}_{\vec{c}_{m-1}}P_{\vec{c}_m}\right)\right\}\\
&\geqslant \mathbb{E}_{\mathcal{C}}\left\{\mathrm{tr}\left(P\sigma_{\vec{c}_m}\right) - 4\,\mathrm{tr}\left(\overline{P}_{\vec{c}_m}P\sigma_{\vec{c}_m}P\right)\right.\\
&\quad \left. -4\sum_{i=1}^{m-1}\mathrm{tr}\left(P_{\vec{c}_i}P\sigma_{\vec{c}_m}P\right)\right\}.
\end{aligned}$$

The inequality follows from Corollary 1.

For the first term of the right side,

$$\mathbb{E}_{\mathcal{C}}\{\operatorname{tr}(P\sigma_{\vec{c}_m})\} = \operatorname{tr}(P\mathbb{E}_{\mathcal{C}}\{\sigma_{\vec{c}_m}\}) = \operatorname{tr}(P\sigma^{\otimes n}) \geqslant 1-\varepsilon.$$

The second equality is due to the fact that $\mathbb{E}_{\mathcal{C}}\{\sigma_{\vec{c}_m}\} = \sigma^{\otimes n}$. The inequality follows from Eq. (17).

For the second term, we have

$$\begin{aligned}\mathbb{E}_{\mathcal{C}}\{\operatorname{tr}(\overline{P}_{\vec{c}_m}P\sigma_{\vec{c}_m}P)\} &= \mathbb{E}_{\mathcal{C}}\{\operatorname{tr}(P\sigma_{\vec{c}_m}) - \operatorname{tr}(P_{\vec{c}_m}P\sigma_{\vec{c}_m}P)\}\\ &\leqslant 1-\mathbb{E}_{\mathcal{C}}\{\operatorname{tr}(P_{\vec{c}_m}P\sigma_{\vec{c}_m}P)\}\\ &\leqslant 4\mathbb{E}_{\mathcal{C}}\{\operatorname{tr}(\overline{P}\sigma_{\vec{c}_m}) + \operatorname{tr}(\overline{P}_{\vec{c}_m}\sigma_{\vec{c}_m})\}\\ &\leqslant 8\varepsilon.\end{aligned}$$

The first inequality uses the fact $\operatorname{tr}(P\sigma_{\vec{c}_m}) \leqslant 1$. The second inequality is due to the bound (1-b). The last inequality follows from Eqs. (17) and (18).

For the third term, we have

$$\begin{aligned}\mathbb{E}_{\mathcal{C}}\left\{\sum_{i=1}^{m-1}\operatorname{tr}(P_{\vec{c}_i}P\sigma_{\vec{c}_m}P)\right\} &\leqslant \mathbb{E}_{\mathcal{C}}\left\{\sum_{i\neq m}\operatorname{tr}(P_{\vec{c}_i}P\sigma_{\vec{c}_m}P)\right\}\\ &= \sum_{i\neq m}\operatorname{tr}(\mathbb{E}_{\mathcal{C}}\{P_{\vec{c}_i}\}P\sigma^{\otimes n}P)\\ &\leqslant 2^{-n[S(\sigma)-\delta]}\sum_{i\neq m}\mathbb{E}_{\mathcal{C}}\{\operatorname{tr}(P_{\vec{c}_i})\}\\ &\leqslant 2^{-n[S(\sigma)-\delta]}(2^{nR}-1)2^{n[\sum p_i S(\sigma_i)+\delta]}\\ &< 2^{n[R-(\chi-2\delta)]},\end{aligned}$$

where $\chi = S(\sigma) - \sum p_j S(\sigma_j)$ is the Holevo quality. The first inequality follows from summing all of the codewords not equal to $\vec{c}_m$(this sum can only be larger). The second inequality is due to Eq. (20). The third inequality follows from Eq. (19).

Thus, the average probability to get the correct result turns to be

$$\mathbb{E}_{\mathcal{C}}\{p_c\} > 1 - 33\varepsilon - 4\times 2^{n[R-(\chi-2\delta)]}.$$

The error probability $p_e = 1 - p_c$, so

$$\mathbb{E}_{\mathcal{C}}\{p_e\} < 33\varepsilon + 4\times 2^{n[R-(\chi-2\delta)]},$$

which means that there exists at least one code such that

$$p_e < 33\varepsilon + 4\times 2^{n[R-(\chi-2\delta)]}.$$

ε and δ can be arbitrary small, so for any R such that $R<\chi$, $p_e \to 0$ when $n\to\infty$. This completes our proof. ■

Remark 2. Sen also provided a similar decoding procedure in Ref. [5]. In his proof, the expected error probability is

$$p_e < 2\sqrt{4\times 2^{n[R-(\chi-2\delta)]} + 13\sqrt{\varepsilon}}. \tag{21}$$

We can see that the error analysis that we have shown above is significantly better than Sen's result.

Remark 3. It would be interesting to compare Corollary 1 with the Hayashi–Nagaoka inequality [13] which plays the key role in the "pretty good measurement." In the pretty good measurement, the detecting operator of $\vec{c}_m$ is defined by

$$\Lambda_m^p = \left(\sum_i PP_{\vec{c}_i}P\right)^{-\frac{1}{2}} PP_{\vec{c}_m}P\left(\sum_i PP_{\vec{c}_i}P\right)^{-\frac{1}{2}}. \tag{22}$$

The error probability can be bounded by applying the Hayashi–Nagaoka inequality

$$(S+T)^{-\frac{1}{2}}S(S+T)^{-\frac{1}{2}} \geqslant I - 2(I-S) - 4T. \tag{23}$$

Let $S = PP_{\vec{c}_m}P$, $T = \sum_{i\neq m}PP_{\vec{c}_i}P$, then

$$\Lambda_m^p \geqslant P - 2P\overline{P}_{\vec{c}_m}P - 4\sum_{i\neq m}PP_{\vec{c}_i}P. \tag{24}$$

In our sequential decoding scheme, the detecting operator of $\vec{c}_m$ is

$$\Lambda_m^s = P\overline{P}_{\vec{c}_1}\cdots\overline{P}_{\vec{c}_{m-1}}P_{\vec{c}_m}\overline{P}_{\vec{c}_{m-1}}\cdots\overline{P}_{\vec{c}_1}P. \tag{25}$$

Applying Corollary 1, we have

$$\Lambda_m^s \geqslant P - 4P\overline{P}_{\vec{c}_m}P - 4\sum_{i=1}^{m-1}PP_{\vec{c}_i}P. \tag{26}$$

We see that Corollary 1 actually plays a similar role as the Hayashi–Nagaoka inequality plays in pretty good measurement, and they give a very similar error analysis.

Conclusion. With the aid of the trigonometric representation of quantum states, we find two union bounds for estimating the disturbance and probability of the sequential projective measurements. Our result provides a powerful tool for analyzing many situations. As an example, we provide a new proof of achieving the Holevo bound via sequential measurements.

It is not clear to us whether the bounds still hold for sequential positive-operator-value measures (POVMs), or stronger, for sequential general measurements. It would be an interesting open problem for further study. What we have known so far is that the bound (1-b) holds when we perform the same POVM repeatedly, i.e., if $\operatorname{tr}(E\rho) = 1-\varepsilon$, then $\operatorname{tr}(E^m\rho) > 1 - m\varepsilon$. This is a simple consequence of the quantum Jensen inequality.

ACKNOWLEDGMENTS

I thank Mark. M. Wilde and N. Cai for useful discussions. This work is supported by the National Natural Science Foundation of China under Grants No. 61271174, No. 61372076, No. 61301178, No. 61301172, and No. 61502376 and the 111 Program of China under Grant No. B08038.

APPENDIX A

In this appendix, the proof of Lemma 2 is specified. From Eqs. (9)–(11), we have

$$\cos\beta_k\cos\theta_k \geqslant \cos\beta_{k-1} - \sin\theta_k\sin\alpha_k \geqslant 0.$$

Let $x = \sin\theta_k$, then from the definition of a_k, we have

$$\cos\theta_k(a_k) = \frac{\cos\alpha_N\cos\beta_k\cos\theta_k - \sqrt{\sum_{i=k+1}^{N}\sin^2\alpha_i}\sqrt{\cos^2\theta_k - \cos^2\beta_k\cos^2\theta_k + \cos^2\theta_k\sum_{i=k+1}^{N-1}\sin^2\alpha_i}}{1+\sum_{i=k+1}^{N-1}\sin^2\alpha_i}$$

$$\geqslant \frac{\cos\alpha_N(\cos\beta_{k-1} - \sin\theta_k\sin\alpha_k) - \sqrt{\sum_{i=k+1}^{N}\sin^2\alpha_i}\sqrt{\cos^2\theta_k - (\cos\beta_{k-1} - \sin\theta_k\sin\alpha_k)^2 + \cos^2\theta_k\sum_{i=k+1}^{N-1}\sin^2\alpha_i}}{1+\sum_{i=k+1}^{N-1}\sin^2\alpha_i}$$

$$= \frac{\cos\alpha_N\cos\beta_{k-1} - x\cos\alpha_N\sin\alpha_k - \sqrt{\sum_{i=k+1}^{N}\sin^2\alpha_i}\sqrt{-\left(1+\sum_{i=k}^{N-1}\sin^2\alpha_i\right)x^2 + 2x\cos\beta_{k-1}\sin\alpha_k + \sum_{i=k+1}^{N-1}\sin^2\alpha_i + \sin^2\beta_{k-1}}}{1+\sum_{i=k+1}^{N-1}\sin^2\alpha_i}.$$

Denote the right side by $g(x)$. From $g'(x) = 0$, we can obtain the minimum value of $g(x)$. It can be verified that $g_{\min}(x) = a_{k-1}$ if and only if

$$x = \frac{\cos\beta_{k-1}\sin\alpha_k\left(\sum_{i=k}^{N}\sin^2\alpha_i\right) + \sin\alpha_k\cos\alpha_N\sqrt{\sum_{i=k}^{N}\sin^2\alpha_i}\sqrt{\left(\sin^2\beta_{k-1} + \sum_{i=k}^{N-1}\sin^2\alpha_i\right)}}{\left(1+\sum_{i=k}^{N-1}\sin^2\alpha_i\right)\left(\sum_{i=k}^{N}\sin^2\alpha_i\right)}.$$

APPENDIX B

To prove the inequality (15), we first define W by

$$W = \cos\alpha_N - \sqrt{\sum_{i=1}^{N}\sin^2\alpha_i\sum_{i=1}^{N-1}\sin^2\alpha_i} - \left(1 - \sum_{i=1}^{N}\sin^2\alpha_i\right).$$

Clearly, if $W \geqslant 0$, then the inequality holds. It can be verified that

$$W = \frac{\sin^2\alpha_N}{1+\sqrt{1-\frac{\sin^2\alpha_N}{\sum\sin^2\alpha_i}}} - \frac{\sin^2\alpha_N}{1+\sqrt{1-\sin^2\alpha_N}}.$$

Since $\sum\sin^2\alpha_i = \sum\varepsilon_i \leqslant \frac{1}{2}$, we have $W \geqslant 0$. ■

[1] G. Mitchison, R. Jozsa, and S. Popescu, Phys. Rev. A **76**, 062105 (2007).
[2] A. Montanaro and R. Wolf, A survey of quantum property testing, arXiv:1310.2035.
[3] S. Lloyd, V. Giovannetti, and L. Maccone, Phys. Rev. Lett. **106**, 250501 (2011).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. A **85**, 012302 (2012).
[5] P. Sen, Achieving the Han–Kobayashi inner bound for the quantum interference channel by sequential decoding, arXiv:1109.0802.
[6] M. M. Wilde, Proc. R. Soc. London, Ser. A **469**, 2157 (2013).
[7] J. Bergou, E. Feldman, and M. Hillery, Phys. Rev. Lett. **111**, 100501 (2013).
[8] C. Q. Pang, F. L. Zhang, L. F. Xu, M. L. Liang, and J. L. Chen, Phys. Rev. A **88**, 052331 (2013).
[9] C. H. Bennett, A. W. Harrow, and S. Lloyd, Phys. Rev. A **73**, 032336 (2006).
[10] S. Aaronson, in *Proceedings of 21st Annual IEEE Conference on Computational Complexity (CCC)* (IEEE, Piscataway, NJ, 2006), p. 273.
[11] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[12] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
[13] M. Hayashi and H. Nagaoka, IEEE Trans. Inf. Theory **49**, 1753 (2003).