Information gain versus coupling strength in quantum measurements

Xuanmin Zhu\(^1\), Yuxiang Zhang\(^3,4\), Quanhui Liu\(^2\), and Shengjun Wu\(^5\)

\(^1\) School of Science, Xidian University, Xi’an 710071, China
\(^2\) School for Theoretical Physics, and Department of Applied Physics, Hunan University, Changsha 410082, China
\(^3\) Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China
\(^4\) School for the Gifted Young, University of Science and Technology of China, Hefei, Anhui 230026, China

(Dated: May 5, 2012)

We investigate the relationship between the information gain and the interaction strength between the quantum system and the measuring device. A strategy is proposed to calculate the information gain of the measuring device as a function of the coupling strength. For qubit systems, we prove that the information gain increases monotonically with the coupling strength. It is obtained that the information gain of the projective measurement along the \(x\) direction decreases with increasing measurement strength along the \(z\) direction, and a complementarity of information gain in the measurements along those two directions is presented.

PACS numbers: 03.67.-a, 03.65.Ta

I. INTRODUCTION

In quantum information theory, quantum systems can be used to transmit classical information. But quantum systems can be neither unambiguously distinguished\(^1,2\), nor perfectly cloned\(^3\) in general. Usually the information encoded in quantum systems cannot be transmitted without any distortion. Even if the transmitted quantum states are not disturbed during the transmission, there is a Holevo bound that limits the accessible information of the receivers\(^4,5\).

The information transmission process can be described as follows: there is a classical information source which produces symbols \(i = 1, ..., n\) according to a probability distribution \(p_1, ..., p_n\). The classical information is quantified by the Shannon entropy \(H(p_i) = -\sum p_i \log p_i\), where the base of the logarithm function is 2 in this paper. The message sender Alice encodes the information into the quantum state \(\rho_i\) with the probability \(p_i\) where \(i = 1, ..., n\). The receiver Bob performs a measurement described by the positive operator valued measure (POVM), \(\{E_j\} = \{E_1, ..., E_m\}\), to gain the information\(^6\). If the measured state is \(\rho_j\), the probability of obtaining output \(j\) is \(p_{ij} = \text{Tr}(E_j \rho_i)\), and \(p_{ij}\) is \(p_i \text{Tr}(E_j \rho_i)\). The accessible information on Bob is \(I_{\text{acc}} = H(p_i) + H(p_j) - H(p_{ij})\). The Holevo bound is \(\chi = S(\rho) - \sum p_i S(\rho_i)\), where \(S(\rho) = \sum p_i S(\rho_i)\), and \(S(\rho)\) is the von Neumann entropy of the state \(\rho\). From the properties of the von Neumann entropy\(^5,6\), we get that \(I_{\text{acc}} \leq S(\rho) - \sum p_i S(\rho_i) \leq H(p_i)\) which means that the accessible information on the receiver is less than the original information.

From another perspective, the Holevo bound \(\chi\) is equal to the mutual information \(S(A : B)\) of the bipartite state \(\rho_{AB} = \sum p_i|i\rangle_A \langle i| \otimes \rho_{iB}\), where \(\{i\rangle_A\) is an orthonormal basis, and \(\rho_{iB} = \rho_i\) is the transmitted state. From the aspect of the quantum correlation theory\(^7,8,11\), the mutual information \(S(A : B)\) is considered to be the total correlation of subsystems \(A\) and \(B\). The maximal classical information that Bob can gain from states \(\{\rho_{iB}\}\) is the classical correlation \(I_{\text{max}}(A : B)\) of state \(\rho_{A,B}\) which is defined as \(I_{\text{max}}(A : B) = \max_{\{E_j\}} \{H(p_i) + H(p_j) - H(p_{ij})\}\). It has been proven that \(I_{\text{max}}(A : B) \leq S(A : B)\)

II. THE INFORMATION GAIN

The quantum states sent by Alice constitute an ensemble \(\{p_i, \rho_{iB}\}\), specified by Alice sending state \(\rho_i\) with probability \(p_i\), where \(i = 1, ..., n\). The ensemble can be described by the density matrix \(\rho_B = \sum p_i \rho_{iB}\). The state of the measuring device is \(|\Phi_B\rangle\); the interaction between the quantum systems and the measuring device is...
where $B$ is an observable operator of the quantum systems, $D$ is an operator of the measuring device, and $g$ is the coupling strength with the assumption that $g \geq 0$. We introduce a fictitious auxiliary system $A$ which can be thought of as the "preparation" system. The auxiliary system has an orthonormal basis $\{|i\}_A$ whose elements correspond to the labels $1, 2, \ldots, n$ on the possible preparations for the transmitted system, $B$. The states of $A$ can be considered as the memory of the original information source. Before the interaction, the overall state of $A$, $B$, and the measuring device $D$ is

$$\rho_{ABD} = \sum_i p_i|\psi_i\rangle_A \langle \psi_i| \otimes \rho_B \otimes |\Phi\rangle_D\langle \Phi|.$$

(2)

After the interaction the overall state evolves into

$$\rho'_{ABD} = \sum_i p_i|\psi_i\rangle_A \langle \psi_i| \otimes U\rho_B \otimes |\Phi\rangle_D\langle \Phi|U^\dagger,$$

(3)

where $U = e^{-iH_{int}dt} = e^{-igB\otimes D}$, with $\hbar = 1$ throughout this paper. It is assumed that the complete orthonormal eigenstates of the observable $B$ are $\{|b_m\}$, and the corresponding eigenvalues are $\{b_m\}$. States $\rho_B$ and $\rho_D$ can be written as

$$\rho_B = \sum_{mn} \rho_{mn}|b_m\rangle \langle b_n|,$$

$$\rho_D = \sum_{mn} \rho_{mn}|b_m\rangle \langle b_n|.$$

(4)

After the interaction, state $\rho_{BD} = \rho_{BD} \otimes |\Phi\rangle_D\langle \Phi|$ evolves into

$$\rho'_{BD} = \sum_{mn} \rho_{mn} e^{-igb_nD} |b_m\rangle \langle b_n| \otimes |\Phi\rangle_D\langle \Phi|e^{igb_nD}.$$

(5)

We get the measuring device’s state

$$\rho_D' = \text{tr}_B(\rho'_{BD}) = \sum_{mn} \rho_{mn} e^{-igb_nD} |\Phi\rangle_D\langle \Phi|e^{igb_nD}.$$

(6)

Similarly, we can obtain the final overall state of the measuring device and the information source,

$$\rho'_{AD} = \text{tr}_B(\rho'_{ABD}) = \sum_i p_i |\psi_i\rangle_A \langle \psi_i| \otimes \rho_D'.$$

(7)

The mutual information $S(A : D)$ of the measuring device and system $A$ represents the correlation of the measuring device and the information source [17], so we define the information gain of the measuring device,

$$I_a = S(A : D) = S(\rho_A') + S(\rho_D') - S(\rho_{AD}')$$

$$= S(\rho_D') - \sum_i p_i S(\rho'_{iD}),$$

(8)

where $\rho_A' = \sum_i p_i |\psi_i\rangle_A \langle \psi_i|$ is the total density matrix of the information source, and $\rho_D' = \sum p_i \rho'_{iD}$ is the total density matrix of the measuring device. From another perspective, the information gain $I_a$ is the Holevo bound for the case that the classical information is encoded in the measuring device’s states $\{|\psi_i\rangle\}$ with the probabilities $\{p_i\}$.

Now, we prove that the information gain $I_a$ is less than the Holevo bound $\chi = S(\rho_B) - \sum_i p_i S(\rho_B)$.

From the theory of relative entropy [5, 18], we have

$$\chi = S(\rho_{ABD} \otimes |\Phi\rangle_D\langle \Phi|)$$

$$= S(\rho_{ABD}' \otimes |\Phi\rangle_D\langle \Phi|U^\dagger_{ABD})$$

(9)

where $U_{ABD} = U_A \otimes e^{-iH_{int}dt}$ is a unitary operator and $\rho_{ABD}' = U_{ABD}\rho_{ABD} \otimes |\Phi\rangle_D\langle \Phi|U_{ABD}^\dagger$. Based on the monotonicity of relative entropy [5, 18], we obtain

$$\chi \geq S(\text{tr}_B(\rho'_{ABD} \otimes |\Phi\rangle_D\langle \Phi|U_{ABD}^\dagger))$$

$$= S(\rho_{AD}' \otimes \rho_D')$$

$$= S(\rho_D') - \sum_i p_i S(\rho'_{iD}) = I_a.$$  

(10)

Thus we obtain that the information gain $I_a$ is less than the Holevo bound $\chi$.

Without loss of generality, the initial state of the measuring device is assumed to be a Gaussian wave function centered on $q = 0$

$$\Phi(q) = \frac{1}{(2\pi\Delta^2)^{\frac{1}{4}}} \exp\left(-\frac{q^2}{4\Delta^2}\right),$$

(11)

where the standard deviation $\Delta = \Delta$. The original density matrix of the measuring device is

$$\rho_D = \frac{1}{(2\pi\Delta^2)^{\frac{1}{4}}} \int e^{-\frac{(q - q')^2}{4\Delta^2}} \Phi(q) \Phi(q') dq dq'.$$

(12)

The interaction Hamiltonian considered is $H_{int} = g\delta(t - t_0)B \otimes p$. From Eq. (10), we obtain the density matrix $\rho_D'$ is

$$\rho_D' = \sum_m \rho_{mm} e^{-igb_mD} \rho_D e^{igb_mD}$$

$$= \sum_m \frac{\rho_{mm} e^{-igb_mD} \rho_D e^{igb_mD}}{(2\pi\Delta^2)^{\frac{1}{4}}} \int e^{-\frac{(q - q')^2}{4\Delta^2}} \Phi(q) \Phi(q') dq dq'.$$

(13)

Since the $\rho_D'$ is a continuous variable density matrix, it is not easy to calculate its von Neumann entropy directly. We can introduce an auxiliary system $R$ to purify the state of the measuring device, and the state of the combined system is

$$|\Psi_i\rangle_{DR} = \sum_m \sqrt{\rho_{mm}} \int e^{-\frac{(q - q')^2}{4\Delta^2}} |m\rangle_R |q\rangle dq,$$

(14)
where $\{|m\rangle_R\}$ is an orthonormal basis of the auxiliary system, and $\rho'_{ID} = \text{tr}_R(\rho'_i D_R |\Psi_i\rangle \langle \Psi_i| D_R)$. As $|\Psi_i\rangle D_R$ is a pure state, we have

$$S(\rho'_{ID}) = S(\rho_{IR}),$$

and the density matrix $\rho_{IR}$ is

$$\rho_{IR} = \text{tr}_D(|\Psi_i\rangle D_R |\Psi_i\rangle) = \sum_{m n} \rho_{mn}^R |m\rangle \langle n|.$$  

We can obtain the matrix elements of $\rho_{IR}$

$$\rho_{mn}^R = \frac{\sqrt{\rho_{mm}^R \rho_{nn}^R}}{(2\pi \Delta^2)^{2}} \int \int e^{-\frac{(s - b_m - b_n)^2}{4\Delta^2}} e^{-\frac{(s' - b_{m'} - b_{n'})^2}{4\Delta^2}} \langle q'| q \rangle dq dq'.$$

It can be seen that the dimension of the matrix $\rho_{IR}$ is the same with the observable $B$. By a similar derivation, we obtain

$$S(\rho'_{D}) = S(\rho_{R}) = S(\sum_{m n} \rho_{mn}^R |m\rangle \langle n|),$$

and the matrix element $\rho_{mn}^R = \sqrt{\rho_{mm}^R \rho_{nn}^R} e^{-\frac{g^2 (b_m - b_n)^2}{8\Delta^2}}$.

So we can get the von Neumann entropy of the measuring device by calculating the entropy of the auxiliary system $R$. From Eqs. (3), (13), and (18), the information gain of the measuring device is

$$I_a = S(\rho_{R}) - \sum_i p_i S(\rho_{IR}).$$  

When the coupling strength is strong (i.e., $g \gg \Delta$), and the eigenvalues of $B$ are nondegenerate, we will prove that the information gain is equal to the information $I_p$ extracted by the projective measurement along the orthonormal eigenstates $\{|b_m\rangle\}$ of $B$. The information $I_p$ obtained in the projective measurement along the basis $\{|b_m\rangle\}$ is

$$I_p = H(p_m) + H(p_i) - H(p_{im}).$$

where $H(p_m) = -\sum m p_m \log p_m$ is the Shannon entropy, $p_m = \text{tr}(|b_m\rangle \langle b_m| \rho_B)$, and the joint probability $p_{im} = \text{tr}(|b_m\rangle \langle b_m| \rho_B)$. When $g \gg \Delta$, and $m \neq n$, we have $e^{-g^2 (b_m - b_n)^2/8\Delta^2} \rightarrow 0$, the nondiagonal elements of matrices $\rho_{IR}$ and $\rho_R$ are approximatively equal to 0, and we have

$$\rho_{IR} \approx \sum m \rho_{mn}^R |m\rangle \langle m|, \rho_R \approx \sum m \rho_{mn} |m\rangle \langle m|.$$  

From Eq. (4), we have $\rho_{IR} = \rho_{IR}^I = \sum m |b_m\rangle \langle b_m| b_m\rangle \langle b_m|$, and $\rho_R = \rho_{IR}^B = \sum m |b_m\rangle \langle b_m| b_m\rangle \langle b_m|$, which are the states after the projective measurements on states $\rho_{IR}$ and $\rho_B$ along the basis $\{|b_m\rangle\}$, respectively. From Eq. (19), the information gain is

$$I_a = S(\rho_B) - \sum_i p_i S(\rho'_B) = H(p_m) - \sum_i p_i H(p_{im})$$

Thus we have proved that the information gain $I_a$ equals the information obtained by measuring the states transmitted $\{\rho_B\}$ along the basis $\{|b_m\rangle\}$. It can be seen that when the coupling strength is large, the information gain of the measuring device is equal to the information obtained in the ideal projective measurement, which is consistent with our expectation.

### III. The Monotonicity of the Information Gain

Now we study the monotonicity of the information gain $I_a$ and the coupling strength $g$ when the transmitted quantum systems are qubits. For two-dimensional systems, the orthonormal eigenstates of observable $B$ can be denoted as $\{|0\rangle, |1\rangle\}$, and without loss of generality, the corresponding eigenvalues are assumed to be $1, -1$. The state of a qubit can be represented as a point in the Bloch sphere. We can use three parameters, $r$ (radius), $\theta$ (polar angle), and $\phi$ (phase angle), to define a qubit state, where $0 \leq r \leq 1$, $0 \leq \theta < \pi$, and $0 \leq \phi < 2\pi$. In the representation $\{|0\rangle, |1\rangle\}$, the transmitted state $\rho_B$ can be written as

$$\rho_{IB} = \left( \begin{array}{cc} \frac{1 + r \cos \theta}{2} & \frac{r \sin \theta e^{-i\phi}}{2} \\ \frac{r \sin \theta e^{i\phi}}{2} & \frac{1 - r \cos \theta}{2} \end{array} \right).$$  

Then $\rho_{I1} = \frac{1 + r \cos \theta}{2}$ and $\rho_{I2} = \frac{1 - r \cos \theta}{2}$, and from Eqs. (16) and (17), we have

$$\rho_{IR} = \left( \begin{array}{cc} \frac{1 + r \cos \theta}{4} & \frac{\sqrt{1 - r^2 \cos^2 \theta}}{4} e^{-\frac{g^2}{2\Delta^2}} \\ \frac{-\sqrt{1 - r^2 \cos^2 \theta}}{4} e^{-\frac{g^2}{2\Delta^2}} & \frac{1 - r \cos \theta}{4} \end{array} \right).$$

Then we obtain the entropy of the state $\rho_{IR}$ as

$$S(\rho_{IR}) = H_B(\lambda_i),$$

where $\lambda_i = (1 + (r_i^2 \cos^2 \theta_i + (1 - r_i^2 \cos^2 \theta_i) e^{-\frac{g^2}{2\Delta^2}})^{1/2})/2$, and $H_B(\lambda_i) = -\lambda_i \log \lambda_i - (1 - \lambda_i) \log (1 - \lambda_i)$ is the binary Shannon entropy. By similar calculations, we have

$$S(\rho_{R}) = H_B \left( \frac{1 + s^{1/2}}{2} \right),$$

where $s = (\sum_i p_i r_i \cos \theta_i)^2 + (1 - (\sum_i p_i r_i \cos \theta_i)^2) e^{-\frac{g^2}{2\Delta^2}}$. From Eqs. (19), (23), and (29), the information gain of
the measuring device is

$$I_a = H_B \left( \frac{1 + s^{1/2}}{2} \right) - \sum_i p_i H_B \left( \lambda_i \right). \tag{27}$$

In the following theorem, we present that the information gain \( I_a \) increases with the coupling strength.

**Theorem 1.** When the transmitted systems are qubits, the information gain \( I_a \) monotonically increases with the coupling strength \( g \).

The proof of this theorem is given in the Appendix.

Without loss of generality, we have assumed that this measurement is along the \( z \) direction. The information gain \( I_{a,z} \) of Eve is given by Eq. (27). After the measurement performed by Eve, the state \( \rho_{jB} \) given in Eq. (27) is changed into

$$\rho_{jB} = \text{tr}_D(\rho_{jBD}) = \left( \begin{array}{cc} \frac{1 + r_j \cos \theta_j}{2} & \frac{r_j \sin \theta_j e^{-i \phi_j} e^{-s^2/2}}{2} \\ \frac{r_j \sin \theta_j e^{i \phi_j} e^{-s^2/2}}{2} & \frac{1 - r_j \cos \theta_j}{2} \end{array} \right), \tag{29}$$

and the total density matrix of the ensemble evolves into

$$\rho_B = \sum_i p_i \rho_i B = \left( \begin{array}{cc} \frac{1 + \sum_j p_j r_j \cos \theta_j}{2} & \frac{\sum_j p_j r_j \sin \theta_j e^{-i \phi_j} e^{-s^2/2}}{2} \\ \frac{\sum_j p_j r_j \sin \theta_j e^{i \phi_j} e^{-s^2/2}}{2} & \frac{1 - \sum_j p_j r_j \cos \theta_j}{2} \end{array} \right). \tag{30}$$

Finally, the legitimate receiver Bob performs a projective measurement on his received quantum system along the \( x \) direction to gain the information from Alice. The projective measurement operators are \( \{ |+\rangle \langle + |, |−\rangle \langle − | \} \), where \( |+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \) and \( |−\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \). The information gain \( I_{a,x} \) of Bob is

$$I_{a,x} = H(p_i) + H(p_{j,x}) - H(p_{ij,x}), \tag{31}$$

where \( j = +, - \), \( p_+ = \text{tr}(|+\rangle \langle + | \rho_B) \), \( p_- = \text{tr}(|−\rangle \langle − | \rho_B) \), \( p_i = \text{tr}(|+\rangle \langle + | \rho_{ij,B}) \), and \( p_{ij} = \text{tr}(|−\rangle \langle − | \rho_{ij,B}) \). From Eqs. (27), (30), and (31), we have

$$I_{a,x} = H_B \left( \frac{1 + \sum_j p_j r_j \sin \theta_j e^{-i \phi_j} e^{-s^2/2}}{2} \right) \tag{32}$$

$$- \sum_i p_i H_B \left( \frac{1 + r_i \sin \theta_i e^{-i \phi_i} e^{-s^2/2}}{2} \right).$$

Now, we give a theorem to show that the information gain \( I_{a,x} \) decreases with an increase in the coupling strength \( g \).

**Theorem 2.** For qubit systems, the information gain \( I_{a,x} \) of the projective measurement along the \( x \) direction monotonically decreases with the measurement coupling strength \( g \) along the \( z \) direction.

This monotonicity is consistent with the widely studied information-disturbance trade-off relation \[19, 22\]. The proof of this theorem is given in the Appendix.

**IV. Complementarity of the Information Gain**

From Theorem 1, we know that the information gain \( I_{a,z} \) of Eve increases with \( g \). For \( g \to +\infty \), from Eq. (27), the information gain is equal to the information gain of the projective measurement along the basis \( \{|0\rangle, |1\rangle \} \), which is

$$I_z = H(p_i) + H(p_{j,z}) - H(p_{ij,z}), \tag{33}$$

where \( p_{1,z} = \text{tr}(|0\rangle \langle 0 | \rho_B) \), \( p_{2,z} = \text{tr}(|1\rangle \langle 1 | \rho_B) \), \( p_{1,z} = p_i \text{tr}(|0\rangle \langle 0 | \rho_{ij,B}) \), \( p_{2,z} = p_i \text{tr}(|1\rangle \langle 1 | \rho_{ij,B}) \), and we have \( I_{a,z} \leq I_z \). In Theorem 2, it is shown that the information gain \( I_{a,x} \) decreases with the coupling strength \( g \), when \( g = 0 \), we have

$$I_x = H(p_i) + H(p_{j,x}) - H(p_{ij,x}), \tag{34}$$

where \( p_{1,x} = \text{tr}(|+\rangle \langle + | \rho_B) \), \( p_{2,x} = \text{tr}(|−\rangle \langle − | \rho_B) \), \( p_{1,x} = p_i \text{tr}(|+\rangle \langle + | \rho_{ij,B}) \), \( p_{2,x} = p_i \text{tr}(|−\rangle \langle − | \rho_{ij,B}) \), and we have \( I_{a,x} \leq I_x \). Then we obtain

$$I_{a,z} + I_{a,x} \leq I_z + I_x \tag{35}$$

$$= H(p_{j,z}) + H(p_{j,x}) - \sum_i p_i (H(p_{j,z|i}) + H(p_{j,x|i})),$$

where the conditional probabilities \( p_{1,z|i} = \text{tr}(|0\rangle \langle 0 | \rho_{ij,B}) \), \( p_{2,z|i} = \text{tr}(|1\rangle \langle 1 | \rho_{ij,B}) \), \( p_{1,x|i} = \text{tr}(|+\rangle \langle + | \rho_{ij,B}) \), and \( p_{2,x|i} = \text{tr}(|−\rangle \langle − | \rho_{ij,B}) \). From the entropic uncertainty relation given in \[23\] and \[24\], we have

$$H(p_{j,z|i}) + H(p_{j,x|i}) \geq 1 + S(\rho_B). \tag{36}$$

As \( \{ p_{j,z} \} \) and \( \{ p_{j,x} \} \) are the two-outcome probability distributions, we have \( H(p_{j,z}) + H(p_{j,x}) \leq 2 \), and from Eqs. (35) and (36), we obtain

$$I_{a,z} + I_{a,x} \leq 1 - \sum_i p_i S(\rho_{ij,B}) \tag{37}.$$
which is a complementarity of the information gain of the measurements along two mutually unbiased bases [23–27]. By complementarity, we mean that the more information eavesdropper Eve extracted from the measurement along the z direction, the less information Bob could gain by the projective measurement along the x direction. Numerical calculations indicate that there is a complementarity of the information gain of the measurements along two mutually unbiased bases [25–27]. This means that the measurement performed by Eve along the z direction destroys the information that Bob could gain in the projective measurement along the x direction.

V. CONCLUSIONS

In conclusion, we have studied the relationship between information gain and measurement coupling strength.

For a finite interaction, the information gain of the measuring device is calculated when the measuring device’s states are of the Gaussian type. When the coupling strength is high, we have shown that the information gain of the measuring device is equal to the information obtained in the projective measurement. It has been proved that the information gain increases with the coupling strength g monotonously for qubit systems. Complementarity of the information obtained in the measurements along two different mutually unbiased bases is given. The research in this paper is useful for evaluating the information gain in finite-interaction measurements.

ACKNOWLEDGMENTS

This work was financially supported by the National Natural Science Foundation of China (Grants No. 11075148, and No. 11175063).

VI. APPENDIX

A. Proof of Theorem 1

Proof. Here, we show that the information gain $I_x$ is a monotonic function of $g$. As $g \geq 0$, let $t = \frac{g}{2\sqrt{\Delta}}$, $s = (\sum_i p_i r_i \cos \theta_i)^2 + (1 - (\sum_i p_i r_i \cos \theta_i)^2) e^{-t}$, and $s_i = (r_i \cos \theta_i)^2 + (1 - (r_i \cos \theta_i)^2) e^{-t}$, we have

$$F = \frac{dI_a}{dt} = \frac{(1 - (\sum_i p_i r_i \cos \theta_i)^2) e^{-t}}{4s^{1/2}} \log \frac{1 + s^{1/2}}{1 - s^{1/2}}$$

$$- \sum_i p_i (1 - (r_i \cos \theta_i)^2) e^{-t} \log \frac{1 + s_i^{1/2}}{1 - s_i^{1/2}}$$

To show the monotonicity of $I_x$, we only need to prove that $F \geq 0$. Let $a = e^{-t}$, and we define a function

$$h(x) = \frac{(1 - x^2) a}{4(x^2 + (1 - x^2) a)^{1/2}} \log \frac{1 + (x^2 + (1 - x^2) a)^{1/2}}{1 - (x^2 + (1 - x^2) a)^{1/2}},$$

where $x \in [-1, 1]$, we get

$$F = h(\sum_i p_i r_i \cos \theta_i) - \sum_i p_i h(r_i \cos \theta_i).$$

Since $\sum_i p_i = 1$, if we could prove that $h(x)$ is a concave function, we will get $F \geq 0$. The second derivative of $h(x)$ is

$$\frac{d^2 h(x)}{dx^2} = \frac{(1 - a)^2 aC(x, a)}{4(x^2 - 1)(a + x^2 - ax^2)^{1/2} D(a, x) \ln 2},$$

where $C(x, a) = 2(a + x^2 - ax^2)^{1/2}(2x^2 + a(x^2 - 1)) + (x^2 - 1)(2x^2 + a^2(x^2 - 1) - a(1 + 3x^2)) \ln \frac{1 + (x^2 + (1 - x^2) a)^{1/2}}{1 - (x^2 + (1 - x^2) a)^{1/2}}$.
and \( D(a, x) = (a + x^2 - 2ax^2 + a^2(-1 + x^2))^2 \); Let \( w = (a + x^2 - ax^2)^{1/2} \), we have \( 0 \leq w \leq 1 \) and \( x^2 = \frac{w^2 - a}{1 - a} \). Let

\[
G(w, a) = \frac{C(x, a)}{x^2 - 1} = \frac{2(2 + a)w^3 - 6aw}{w^2 - 1} + ((2 - a)w^2 - 3a) \ln \frac{1 + w}{1 - w} \quad (43)
\]

For a fixed value of \( w \), we search the extreme value of \( G(w, a) \); from

\[
\frac{\partial G(w, a)}{\partial a} = \frac{2w^3 - 6w}{w^2 - 1} - (3 + w^2) \ln \frac{1 + w}{1 - w} = 0, \quad (44)
\]

we obtain

\[
\ln \frac{1 + w}{1 - w} = \frac{2w^3 - 6w}{(w^2 - 1)(3 + w^2)}. \quad (45)
\]

Substituting this solution into Eq. (43), the extreme value of this function is

\[
G(w, a)_{\text{ext}} = \frac{8w^5}{(w^2 - 1)(3 + w^2)}. \quad (46)
\]

Since \( 0 \leq w \leq 1 \), we have \( G(w, a)_{\text{ext}} \leq 0 \). As \( 0 < a = e^{-t} \leq 1 \), for \( a = 0 \), we have

\[
G(w, 0) = \frac{2w^2}{w^2 - 1}(2w - (1 - w^2) \ln \frac{1 + w}{1 - w}). \quad (47)
\]

Let \( K(w) = 2w - (1 - w^2) \ln \frac{1 + w}{1 - w} \), the first derivative of \( K \) is

\[
\frac{dK}{dw} = 2w \ln \frac{1 + w}{1 - w}. \quad (48)
\]

Since \( 0 \leq w \leq 1 \), we have \( \frac{dK}{dw} \geq 0 \), as \( K(0) = 0 \), so \( K(w) \geq 0 \). From Eq. (47), we have \( G(w, 0) \leq 0 \). When \( a = 1 \), we obtain

\[
G(w, 1) = 6w + (w^2 - 3) \ln \frac{1 + w}{1 - w}. \quad (49)
\]

The partial derivative of \( G(w, 1) \) is

\[
\frac{\partial G(w, 1)}{\partial w} = \frac{2w}{w^2 - 1}(2w - (1 - w^2) \ln \frac{1 + w}{1 - w}) = \frac{2w}{w^2 - 1} K. \quad (50)
\]

Since \( K \geq 0 \) and \( 0 \leq w \leq 1 \), so \( \frac{\partial G(w, 1)}{\partial w} \leq 0 \), and as \( G(0, 1) = 0 \), we have \( G(w, 1) \leq 0 \). Now we have proved that \( G(w, 0) \leq 0, \ G(w, 1) \leq 0 \), and the extreme value \( G(w, a)_{\text{ext}} \leq 0 \), and as the function \( G(w, a) \) is a continuum function of \( a, 0 \leq w \leq 1, \) and \( 0 < a \leq 1 \), we have \( G(w, a) \leq 0 \). From Eqs. (42) and (43), we have

\[
\frac{d^2 h(x)}{dx^2} \leq 0, \quad (51)
\]

and \( h(x) \) is a concave function. From Eq. (41), we have \( F \leq 0 \). Thus we have proved Theorem 1. \( \square \)

**B. Proof of Theorem 2**

Proof. As \( g \geq 0 \), let \( t = \frac{s^2}{2}, s = \sum_i p_i r_i \sin \theta_i \cos \phi_i \), and \( s_i = r_i \sin \theta_i \cos \phi_i \), we have

\[
W = \frac{dI_{a,x}}{dt} = \frac{se^{-t/2}}{4} \log 1 + se^{-t/2} - \sum_i p_i \frac{s_i e^{-t/2}}{4} \log 1 + s_i e^{-t/2} - \sum_i s_i e^{-t/2}. \quad (52)
\]

To show the monotonicity of \( I_{a,x} \), we only need to prove that \( W \leq 0 \). We define a function \( f(x) = \frac{x e^{-t/2}}{4} \log 1 + x e^{-t/2} \), where \( x \in [-1, 1] \). We have

\[
W = f(\sum_i p_i r_i \sin \theta_i \cos \phi_i) - \sum_i p_i f( r_i \sin \theta_i \cos \phi_i). \quad (53)
\]

The second derivative of \( f(x) \) is

\[
\frac{d^2 f(x)}{dx^2} = \frac{e^{-t}}{(1 - e^{-t}x^2)^2} \ln 2 \quad (54)
\]

Since \( 0 < e^{-t} \leq 1 \) and \(-1 \leq x \leq 1 \), we have \( \frac{d^2 f(x)}{dx^2} > 0 \), so \( f(x) \) is a convex function. Since \( \sum_i p_i = 1 \), from Eq. (53), we obtain that \( W \leq 0 \). Thus, it has been proven that the information gain \( I_{a,x} \) monotonically decreases with the coupling strength \( g \). \( \square \)

---

[1] A. S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory (North-Holland, Amsterdam, 1982).
[2] S. Pang and S. Wu, Phys. Rev. A 80, 052320 (2009).
[3] W. K. Wootters and W. H. Zurek, Nature (London) 299, 802(1982).
[4] A. S. Holevo, Probl. Peredachi Inf. 9, 3 (1973); Probl. Inf. Transm. 9, 177 (1973).
[5] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, England, 2000).
[6] A. Wehrl, Rev. Mod. Phys. 50 221 (1978).
