Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Experimental Test of Heisenberg's Measurement Uncertainty Relation Based on Statistical Distances

by

Wenchao Ma, Zhi-Hao Ma, Hengyan Wang, Zhi-Hua Chen, Ying Liu, Fei Kong, Zhaokai Li, Xinhua Peng, Mingjun Shi, Fazhan Shi, Shao-Ming Fei, and Jiangfeng Du

Preprint no.: 49 2016

Experimental Test of Heisenberg's Measurement Uncertainty Relation Based on Statistical Distances

Wenchao Ma,[1] Zhihao Ma,[2] Hengyan Wang,[1] Zhihua Chen,[3] Ying Liu,[1] Fei Kong,[1] Zhaokai Li,[1,4] Xinhua Peng,[1,4] Mingjun Shi,[1,4] Fazhan Shi,[1,4] Shao-Ming Fei,[5] and Jiangfeng Du[1,4,*]

[1] *Hefei National Laboratory for Physical Sciences at the Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*
[2] *Department of Mathematics, Shanghai Jiaotong University, Shanghai 200240, China*
[3] *Department of Applied Mathematics, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, China*
[4] *Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*
[5] *School of Mathematical Sciences, Capital Normal University, Beijing 100048, China*

Incompatible observables can be approximated by compatible observables in joint measurement or measured sequentially, with constrained accuracy as implied by Heisenberg's original formulation of the uncertainty principle. Recently, Busch, Lahti, and Werner proposed inaccuracy trade-off relations based on statistical distances between probability distributions of measurement outcomes [Phys. Rev. Lett. **111**, 160405 (2013); Phys. Rev. A **89**, 012129 (2014)]. Here we reformulate their theoretical framework, derive an improved relation for qubit measurement, and perform an experimental test on a spin system. The relation reveals that the worst-case inaccuracy is tightly bounded from below by the incompatibility of target observables, and is verified by the experiment employing joint measurement in which two compatible observables designed to approximate two incompatible observables on one qubit are measured simultaneously.

The uncertainty principle was first proposed by Heisenberg in the context of a measurement process [1]. He conceived of a γ-ray microscope and pointed out that the measurement of an electron's position Q disturbs the momentum P inevitably, and that the product of the error $\varepsilon(Q)$ and disturbance $\eta(P)$ cannot be arbitrarily small. A qualitative relation was written as $\varepsilon(Q)\eta(P) \sim h$, where h is Planck's constant. In the ensuing few years, Kennard [2], Weyl [3], Robertson [4], and Schrödinger [5] derived mathematically rigorous versions, including the famous relation $\sigma(A)\sigma(B) \geq \frac{1}{2}\left|\langle[A,B]\rangle\right|$, where the angle brackets represent expectations, the standard deviation $\sigma(A) = \sqrt{\langle A^2\rangle - \langle A\rangle^2}$ is called preparation uncertainty, and $[A,B] = AB - BA$. However, these formal inequalities, together with well-developed entropic uncertainty relations [6–12], do not handle the problem Heisenberg discussed, but refer to the uncertainties intrinsic to quantum states.

The attempts to quantify Heisenberg's original idea have a long history [13]. A dozen years ago, Ozawa proved a universally valid error-disturbance relation [14]. Shortly thereafter, Hall and Ozawa independently showed that the inaccuracies of any joint measurement estimating two incompatible observables satisfy similar relations [15–17]. Later, tighter inequalities were derived by Branciard [18, 19] and Weston *et al.* [20]. These relations have been experimentally verified using polarized neutrons [21, 22] and photons [20, 23–26]. Nevertheless, the physical validity of the definitions of error and disturbance in Ozawa's relation is in dispute [27–33]. Information-theoretic definitions for noise and disturbance were introduced and a state-independent trade-off relation was derived by Buscemi *et al* [34], and this relation has been verified by neutron spin qubits [35].

Recently, Busch, Lahti, and Werner (BLW) formulated uncertainty relations dealing with the imprecisions in joint measurements approximating incompatible observables [32, 33, 36]. The imprecisions or uncertainties, based on statistical distances between probability distributions of measurement outcomes, are processed into state independent and regarded as figures of merit of joint-measurement devices. Besides, the joint measurement scenario covers the successive measurement scenario with errors and disturbances. In this work, we reformulate BLW's theoretical framework, derive an improved relation for qubit measurement, and perform an experimental test on a spin system.

Consider a pair of incompatible observables A, B. Incompatible means A, B are not jointly measurable, i.e., cannot be measured simultaneously. For qubit measurements, A, B can be selected as two sharp observables with the spectral projections $A_\pm = (\mathbb{1} \pm \boldsymbol{a}\cdot\boldsymbol{\sigma})/2$ and $B_\pm = (\mathbb{1} \pm \boldsymbol{b}\cdot\boldsymbol{\sigma})/2$, where $\boldsymbol{a}, \boldsymbol{b}$ are unit vectors. If $\boldsymbol{a}$ and $\boldsymbol{b}$ are noncollinear, the observables A, B are incompatible. Approximate measurements of A, B can be performed simultaneously by a device composed of compatible observables C, D, each being used as an approximation of A, B, respectively. The observable C consists of positive operators $C_+ = (c_0\mathbb{1} + \boldsymbol{c}\cdot\boldsymbol{\sigma})/2$, $C_- = \mathbb{1} - C_+$, with $|\boldsymbol{c}| \leq \min\{c_0, 2 - c_0\} \leq 1$ for the positivity. Similarly, the observable D consists of $D_+ = (d_0\mathbb{1} + \boldsymbol{d}\cdot\boldsymbol{\sigma})/2$, $D_- = \mathbb{1} - D_+$, with $|\boldsymbol{d}| \leq \min\{d_0, 2 - d_0\} \leq 1$. Compatible or jointly measurable requires that there exists a positive-operator valued measure (POVM) M of which C, D are the marginals, namely, $C_i = \sum_j M_{ij}$ and

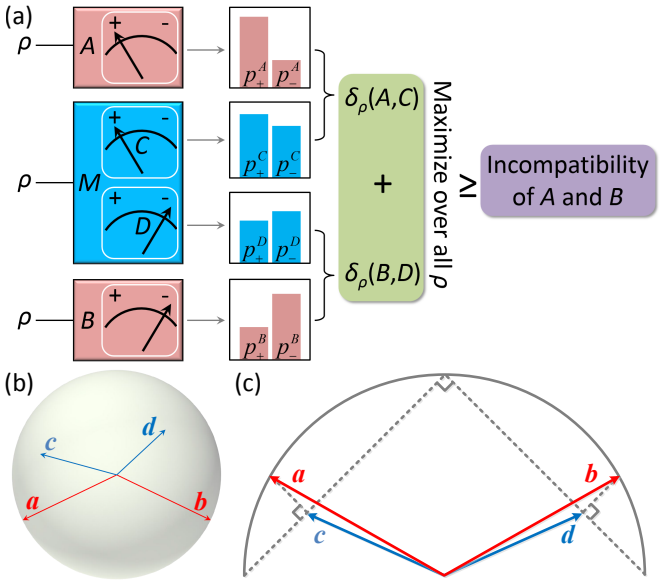


FIG. 1. (color online). Sketch of the theory. (a) Outline of the inaccuracy trade-off relation in Eq. (2). Here compatible observables C, D are jointly measured to approximate incompatible observables A, B. (b) An example of Bloch vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$. The lengths of $\boldsymbol{c}, \boldsymbol{d}$ are usually less than 1. (c) Bloch vectors $\boldsymbol{c}, \boldsymbol{d}$ of the optimal compatible observables C, D that approximate A, B. The dashed chords are legs of an inscribed isosceles right triangle. The vectors $\boldsymbol{a} - \boldsymbol{c}$ and $\boldsymbol{b} - \boldsymbol{d}$ are orthogonal to these dashed chords, and the end points of $\boldsymbol{c}, \boldsymbol{d}$ are the perpendicular feet. The whole pattern is coplanar and symmetric.

$D_j = \sum_i M_{ij}$, with $i, j = +$ or $-$ (hereinafter the same) and $M_{ij} = (C_i D_j + D_j C_i)/2$. When $c_0 = d_0 = 1$, the criterion of compatibility takes a simple form as [33, 37, 38]

$$|\boldsymbol{c} + \boldsymbol{d}| + |\boldsymbol{c} - \boldsymbol{d}| \leq 2. \tag{1}$$

Compatibility allows C, D to be performed simultaneously, but imposes a strong restriction that prevents C, D approaching A, B freely. The overall deviation of C, D from A, B, or the performance of the device, is rooted in probability distributions of measurement outcomes.

The operators $O_\pm$ $(O = A, B, C, D)$ are called effects, and the probabilities of measurement outcomes are $p_\pm^O = \mathrm{tr}\,(\rho O_\pm)$, where $\rho = (\mathbb{1} + \boldsymbol{r} \cdot \boldsymbol{\sigma})/2$ is the density operator of a qubit. We scale the outcomes of these measurements to be ± 1, and then the observable O is given as a map $\pm 1 \mapsto O_\pm$. The statistical difference between the measurements of A and C, namely, the inaccuracy or uncertainty, can be quantified as $\delta_\rho\,(A, C) := 2 \sum_i \left| p_i^A - p_i^C \right| = 2\,|1 - c_0 + \boldsymbol{r} \cdot (\boldsymbol{a} - \boldsymbol{c})|$ [33]. Likewise, the statistical difference between the measurements of B and D is written as $\delta_\rho\,(B, D) := 2 \sum_i \left| p_i^B - p_i^D \right| = 2\,|1 - d_0 + \boldsymbol{r} \cdot (\boldsymbol{b} - \boldsymbol{d})|$. Since the observables C, D are jointly measured, they face the same quantum state ρ and the combined difference $\Delta_\rho(A, B; C, D) := \delta_\rho\,(A, C) + \delta_\rho\,(B, D)$ is the inaccuracy of the joint measurement for a specific state. The state-dependent quantity $\Delta_\rho(A, B; C, D)$ has a vanishing lower bound for any given A, B, and ρ. In other words, $\Delta_\rho(A, B; C, D)$ vanishes when minimizing over all C, D for any given A, B, and ρ (see Sec. II.A in Ref. [39]). By maximizing $\Delta_\rho(A, B; C, D)$ over all ρ, one obtains $\Delta(A, B; C, D) := \max_\rho \Delta_\rho(A, B; C, D)$ as the worst-case inaccuracy of the joint measurement. The state-independent quantity $\Delta(A, B; C, D)$ characterizes the integrated deviation of C, D from A, B and thus is a figure of merit of the measurement device. The inaccuracy trade-off relation is

$$\begin{aligned} \Delta(A, B; C, D) &:= \max_\rho \left[\delta_\rho\,(A, C) + \delta_\rho\,(B, D)\right] \\ &\geq \Delta_{\mathrm{lb}}(A, B) := |\boldsymbol{a} + \boldsymbol{b}| + |\boldsymbol{a} - \boldsymbol{b}| - 2, \end{aligned} \tag{2}$$

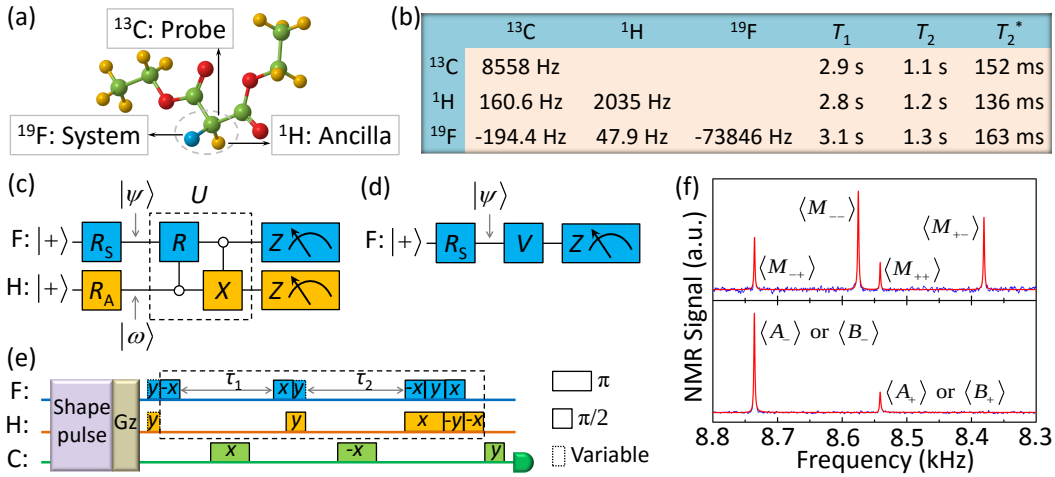
as outlined in Fig. 1(a). The lower bound $\Delta_{\mathrm{lb}}(A, B)$ is the amount of violation of the inequality (1), and is termed as the degree of incompatibility of A, B [33]. This lower bound is attained when C, D are the best approximations of A, B, and such C, D have $c_0 = d_0 = 1$ and $\boldsymbol{c}, \boldsymbol{d}$ depicted in Fig. 1(c). Compared with BLW's original relation, the characteristic quantity $\Delta(A, B; C, D)$ possesses an explicit physical meaning as the worst-case inaccuracy and has a delicate lower bound without an extra coefficient (see Sec. II.B in Ref. [39]). Additionally, as a modification of BLW's conceptual framework, the improved approach also applies to any pair of incompatible observables such as position and momentum although the combined inaccuracy may not be in the additive form.

The nuclear spin system, operated on a 400 MHz liquid-state nuclear magnetic resonance (NMR) spectrometer, is employed to experimentally demonstrate the relation in Eq. (2). We use diethyl fluoromalonate as the sample dissolved in ^{2}H-labeled chloroform at 304 K. As shown in Fig. 2(a), the ^{19}F, ^{1}H, and ^{13}C nuclear spins serve as the system, ancilla, and probe qubits, respectively. The Hamiltonian of the three-qubit system in the triple-resonance rotating frame is

$$H = 2\pi \sum_{1 \leq k < l \leq 3} J_{kl} I_z^k I_z^l, \tag{3}$$

with the scalar couplings J_{kl} listed in Fig. 2(b).

The experiment begins with preparing the pseudopure state (PPS) $\rho_{\mathrm{pps}} = (1 - \varepsilon)\mathbb{1}/8 + \varepsilon\,|{+}{+}{+}\rangle\langle{+}{+}{+}|$ from the thermal equilibrium state using the line-selective method [40, 41]. Here $\varepsilon \approx 10^{-5}$ denotes the polarization and $\mathbb{1}$ denotes the 8×8 identity matrix. In this paper, $|+\rangle$ and $|-\rangle$ represent the eigenvectors of the Pauli matrix Z. A shape pulse based on the gradient ascent pulse engineering (GRAPE) algorithm [42] and a pulsed field gradient are utilized for this step. After such initialization, a local operation R_{S} on the system qubit prepares the state $|\psi\rangle = R_{\mathrm{S}}\,|+\rangle$ that maximizes Δ_ρ. In our experiment, all involved C, D have $c_0 = d_0 = 1$, and the state



	^{13}C	^{1}H	^{19}F	T_1	T_2	T_2^*
^{13}C	8558 Hz			2.9 s	1.1 s	152 ms
^{1}H	160.6 Hz	2035 Hz		2.8 s	1.2 s	136 ms
^{19}F	-194.4 Hz	47.9 Hz	-73846 Hz	3.1 s	1.3 s	163 ms

FIG. 2. (color online). Experimental system and methods. (a),(b) Molecular structure and NMR parameters of diethyl fluoromalonate. The chemical shifts and scalar couplings are on and below the diagonal of the table. The relaxation times are in the right half of the table. (c) Quantum circuit for the joint measurement of C, D. (d) Quantum circuit for the measurement of A or B. (e) Experimental process for the joint measurement of C, D. All the square pulses are hard pulses with negligible duration compared to the free evolution time τ_1 and τ_2. The operation U is implemented by the pulse sequence in the dashed frame (see Sec. III in Ref. [39]). The π pulses applied on ^{13}C are for decoupling, and the final $\pi/2$ pulse is for readout. (f) ^{13}C spectra for the joint measurement of C, D via M (upper panel) and the measurement of A or B (lower panel). The jittering blue curves represent the experimental data and the smooth red curves are from Lorentzian fitting.

$|\psi\rangle$ is simply determined by $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ (see Section IV in Ref. [39]). If the angle between the vectors $\boldsymbol{a}-\boldsymbol{c}$ and $\boldsymbol{b}-\boldsymbol{d}$ is an acute or obtuse angle, the Bloch vector of $|\psi\rangle$ should be collinear with $\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}$ or $\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}$, respectively. If the angle is a right angle, the Bloch vector should be collinear with either.

For the measurement, different schemes are utilized to implement the observables A, B and C, D. The compatible observables C, D are measured in a joint way. In other words, we measure the joint observable M induced by C, D as previously mentioned. As a POVM with four components, M can be extended to orthogonal projective measurements on a certain two-qubit basis $|\chi_{ij}\rangle$ by encompassing the ancilla qubit. The logic circuit is illustrated in Fig. 2(c). A local operation R_{A} on the ancilla qubit prepares the state $|\omega\rangle = R_{\mathrm{A}}|+\rangle$. After that, a global operation U on the two qubits is performed with the controlled gates realized by hard pulses and free evolution. Finally, the probabilities on $|ij\rangle$ is measured. The combined effect of U and the measurement on $|ij\rangle$ amounts to the measurement on $|\chi_{ij}\rangle = U^{\dagger}|ij\rangle$, and the POVM on the system qubit is yielded as $M_{ij} = \mathrm{tr}_{\mathrm{A}}\left[|\chi_{ij}\rangle\langle\chi_{ij}|\left(\mathbb{1}\otimes|\omega\rangle\langle\omega|\right)\right]$, where $\mathbb{1}$ denotes the identity operator on the system qubit and tr_{A} the partial trace over the ancilla. The explicit form of M is

$$\begin{aligned} M_{\pm\pm} &= |\langle - \mid \omega\rangle|^2 |\pm\rangle\langle\pm|, \\ M_{\pm\mp} &= |\langle + \mid \omega\rangle|^2 R^{\dagger}|\pm\rangle\langle\pm| R. \end{aligned} \tag{4}$$

In this way, different M and thus C, D can be implemented by adjusting the operations R and R_{A}. As the upper panel of Fig. 2(f) shows, the statistics of M's four outcomes are displayed as four peaks on the NMR spectrum at once by introducing the ^{13}C nuclear spin as the probe qubit. The peak areas are proportional to the probabilities on $|ij\rangle$, and hence the probabilities relating to M. Therefore, the two compatible but typically non-commutative observables C, D are simultaneous measured with the outcome distributions derived from that of M.

The incompatible observables A and B are measured separately, and the ancilla qubit is not required. The logic circuit is illustrated in Fig. 2(d). The system qubit in the state $|\psi\rangle$ is measured on the basis $V^{\dagger}|i\rangle$ constructed as the eigenstates of $A_{\pm}$ or $B_{\pm}$. The statistics of measurement outcomes are displayed as two peaks on the spectrum as the lower panel of Fig. 2(f) shows. By comparing the outcome distributions of C, D with that of A, B, one obtains the state-dependent inaccuracy Δ_{ρ}. For the special preselected state $|\psi\rangle$, the value of Δ_{ρ} is identical to the state-independent inaccuracy Δ.

To experimentally validate the inaccuracy trade-off relation, we select several configurations of A, B and C, D with all the Bloch vectors in the xz plane. Every pair of $\boldsymbol{a}, \boldsymbol{b}$ is symmetric about the z axis, and is expressed as $\boldsymbol{a} = (-\sin(\theta/2), \cos(\theta/2))$, $\boldsymbol{b} = (\sin(\theta/2), \cos(\theta/2))$, where the former and latter components of a vector correspond to x and z, respectively (hereinafter the same). In the case where $\theta = 90^{\circ}$, the vectors $\boldsymbol{a}, \boldsymbol{b}$ are perpendicular to each other and the relevant observables A, B are maximally incompatible. Consider a series of compatible observables C, D with $\boldsymbol{c} =$

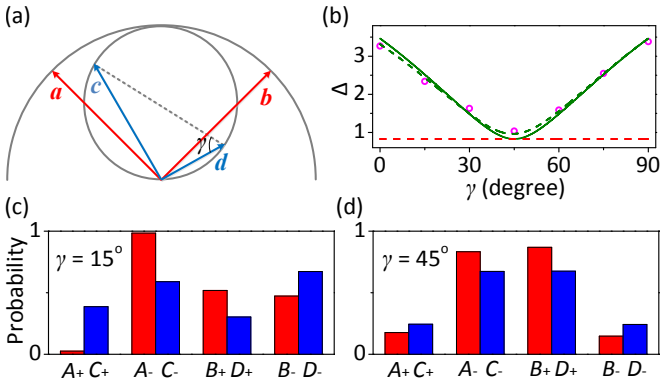


FIG. 3. (color online). Experimental demonstration of the inaccuracy trade-off relation with maximally incompatible A, B. (a) Bloch vectors of observables. The whole configuration is coplanar. The vectors $\boldsymbol{a}, \boldsymbol{b}$ are mutually perpendicular and incline symmetrically about the vertical direction, which is the z direction. The end points of vectors $\boldsymbol{c}, \boldsymbol{d}$ are the diameter ends of the small circle which is internally tangent to the semicircle and passes through its center. The center of the semicircle is the origin and the tangent point is in the z axis. (b) The inaccuracy Δ as a function of the angle $\gamma = \arctan\left(|\boldsymbol{c}| / |\boldsymbol{d}|\right)$. The circles, solid curves, and dashed curves represent the experimental, theoretical, and simulated values of Δ, respectively. The numerical simulation takes decoherence into account. The dashed straight line denotes the lower bound $\Delta_{\rm lb}$. (c),(d) Experimental statistics of measurement outcomes for $\gamma = 15°, 45°$.

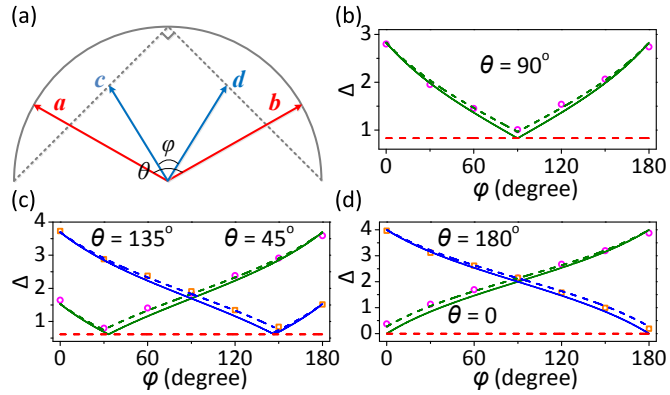


FIG. 4. (color online). Experimental demonstration of the inaccuracy trade-off relation with several pairs of A, B. (a) Bloch vectors of observables. The whole configuration is coplanar and symmetric. The dashed chords have the same meaning as that in Fig. 1(c). The angle between the vectors $\boldsymbol{a}, \boldsymbol{b}$ is θ and the angle between the vectors $\boldsymbol{c}, \boldsymbol{d}$ is φ. (b)-(d) The inaccuracy Δ as a function of the angle φ for several pairs of A, B with $\theta = 90°$(b), $\theta = 45°, 135°$(c), and $\theta = 0°, 180°$(d). The lower bound $\Delta_{\rm lb}$ varies with θ, and is the same for θ and $180° - \theta$. The symbols in these diagrams have the same meanings as that in Fig. 3(b).

$\left(-\sin\gamma\cos\gamma, \sin^2\gamma\right), \boldsymbol{d} = \left(\sin\gamma\cos\gamma, \cos^2\gamma\right)$ as illustrated in Fig. 3(a). The inaccuracy Δ varies with γ and is bounded by $\Delta_{\rm lb} = 2\left(\sqrt{2}-1\right) \approx 0.83$ as shown in Fig. 3(b). When $\gamma = 45°$, the corresponding C, D are the optimal approximations to A, B and the inaccuracy Δ reaches its lower bound $\Delta_{\rm lb}$. The statistics of measurement outcomes for two instances where $\gamma = 15°, 45°$ are listed in Figs. 3(c) and 3(d), which visually support the superiority of the optimal C, D. Another series of C, D with $\boldsymbol{c} = \left(-1/\left[1+\cot(\varphi/2)\right], 1/\left[1+\tan(\varphi/2)\right]\right), \boldsymbol{d} = \left(1/\left[1+\cot(\varphi/2)\right], 1/\left[1+\tan(\varphi/2)\right]\right)$ are illustrated in Fig. 4(a). For A, B with $\theta = 90°$, the optimal C, D are reached when $\varphi = 90°$ as shown in Fig. 4(b). This series of C, D are also used to approximate other A, B which are less incompatible. For $\theta = 45°$ or $135°$, the inaccuracy Δ is bounded by $\Delta_{\rm lb} = 2/\sqrt{2-\sqrt{2}} - 2 \approx 0.61$ as illustrated in Fig. 4(c). In the limit where $\theta = 0°$ or $180°$, A, B reduce to compatible with a vanishing $\Delta_{\rm lb}$ as shown in Fig. 4(d), and the optimal C, D are just A, B. All the experimental results confirm the inaccuracy trade-off relation in Eq. (2).

In conclusion, we have reformulated BLW's theoretical framework, derived an improved relation for qubit measurement, and performed an experimental test using the NMR technique. We show that as a figure of merit of the measurement device, the worst-case inaccuracy is tightly bounded by the incompatibility of target observables in the qubit case. In the experiment, the device is simulated by joint measurement which measures two compatible observables simultaneously. Our work represents an advance in quantitatively understanding and experimental verification of Heisenberg's uncertainty principle, and could have implications for the area of quantum information technology.

The authors thank Hui Zhou for helpful discussions. This work was supported by the 973 Program (Grant No. 2013CB921800), the National Natural Science Foundation of China (Grants No. 11227901, No. 31470835, No. 11371247, No. 11201427, No. 11571313, No. 11575173, No. 11575171, and No. 11275131), the China Postdoctoral Science Foundation, the CAS (Grant No. XDB01030400), and the Fundamental Research Funds for the Central Universities (WK2340000064). W. M., Z. M., and H. W. contributed equally to this work.

* djf@ustc.edu.cn

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927); *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton Univ. Press, Princeton, New Jersey, 1983), p. 62.
[2] E. H. Kennard, Z. Phys. **44**, 326 (1927).
[3] H. Weyl, *Gruppentheorie Und Quantenmechanik* (Hirzel, Leipzig, 1928).
[4] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[5] E. Schrödinger, Sitz. Preuss. Akad. Wiss. (Phys.-Math. Klasse) **19**, 296 (1930); A. Angelow and M.-C. Batoni,

arXiv: 9903100.
[6] I. I. Hirschman, Am. J. Math. **79**, 152 (1957).
[7] W. Beckner, Ann. Math. **102**, 159 (1975).
[8] I. Białynicki-Birula and J. Mycielski, Commun. Math. Phys. 44, 129 (1975).
[9] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
[10] K. Kraus, Phys. Rev. D **35**, 3070 (1987).
[11] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
[12] S. Wehner and A. Winter, New J. Phys. **12**, 025009 (2010).
[13] P. Busch, T. Heinonen, and P. Lahti, Phys. Rep. **452**, 155 (2007).
[14] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).
[15] M. Ozawa, Int. J. Quantum. Inform. **01**, 569 (2003).
[16] M. Ozawa, Phys. Lett. A **320**, 367 (2004).
[17] M. J. W. Hall, Phys. Rev. A **69**, 052113 (2004).
[18] C. Branciard, Proc. Natl. Acad. Sci. U.S.A. **110**, 6742 (2013).
[19] C. Branciard, Phys. Rev. A **89**, 022124 (2014).
[20] M. M. Weston, M. J. W. Hall, M. S. Palsson, H. M. Wiseman, and G. J. Pryde, Phys. Rev. Lett. **110**, 220402 (2013).
[21] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Nat. Phys. **8**, 185 (2012).
[22] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, Phys. Rev. A **88**, 022110 (2013).
[23] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg. Phys. Rev. Lett. **109**, 100404 (2012).
[24] S.-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, Sci. Rep. **3**, 2221 (2013).
[25] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, Phys. Rev. Lett. **112**, 020401 (2014).
[26] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, Phys. Rev. Lett. **112**, 020402 (2014).
[27] A. Di Lorenzo, Phys. Rev. Lett. **110**, 120403 (2013).
[28] J. Dressel and F. Nori, Phys. Rev. A **89**, 022106 (2014).
[29] K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. A **89**, 052108 (2014).
[30] P. Busch, T. Heinonen, and P. Lahti, Phys. Lett. A **320**, 261-270 (2004).
[31] R. F. Werner, Quantum Inf. Comput. **4**, 546 (2004).
[32] P. Busch, P. Lahti, and R. F. Werner, Phys. Rev. Lett. **111**, 160405 (2013).
[33] P. Busch, P. Lahti, and R. F. Werner, Phys. Rev. A **89**, 012129 (2014).
[34] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, Phys. Rev. Lett. **112**, 050401 (2014).
[35] G. Sulyok, S. Sponar, B. Demirel, F. Buscemi, M. J. W. Hall, M. Ozawa, and Y. Hasegawa, Phys. Rev. Lett. **115**, 030401 (2015).
[36] P. Busch, P. Lahti, and R. F. Werner, Rev. Mod. Phys. **86**, 1261 (2014).
[37] P. Busch, Phys. Rev. D **33**, 2253 (1986).
[38] S. Yu, N.-l. Liu, L. Li, and C. H. Oh, Phys. Rev. A **81**, 062116 (2010).
[39] See Supplemental Material for the derivation of Eq. (2), discussions on the two trade-off relations, detailed experimental methods, and additional data.
[40] N. A. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).
[41] X. Peng, X. Zhu, X. Fang, M. Feng, K. Gao, X. Yang, and M. Liu, Chem. Phys. Lett. **340**, 509 (2001).
[42] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

Supplementary Material for

Experimental test of Heisenberg's measurement uncertainty relation based on statistical distances

Wenchao Ma,[1] Zhihao Ma,[2] Hengyan Wang,[1] Zhihua Chen,[3] Ying Liu,[1] Fei Kong,[1] Zhaokai Li,[1,4] Xinhua Peng,[1,4] Mingjun Shi,[1,4] Fazhan Shi,[1,4] Shao-Ming Fei,[5] and Jiangfeng Du[1,4,*]

[1] *Hefei National Laboratory for Physical Sciences at the Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*
[2] *Department of Mathematics, Shanghai Jiaotong University, Shanghai 200240, China*
[3] *Department of applied mathematics, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, China*
[4] *Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*
[5] *School of Mathematical Sciences, Capital Normal University, Beijing 100048, China*

CONTENTS

I. Proof of the inaccuracy trade-off relation for qubit measurements 2

II. Discussion on the inaccuracy trade-off relation for qubit measurements 4
A. State-dependent inaccuracy trade-off relation 4
B. Comparison with BLW's original relation 5

III. Experimental methods 6
A. Quantum gates 6
B. Pulse sequences 7
C. Experimental errors 7

IV. Additional experimental data 7

References 9

* djf@ustc.edu.cn

I. PROOF OF THE INACCURACY TRADE-OFF RELATION FOR QUBIT MEASUREMENTS

The target observables A, B have the spectral projections $A_\pm = (\mathbb{1} \pm \boldsymbol{a}\cdot\boldsymbol{\sigma})/2$ and $B_\pm = (\mathbb{1} \pm \boldsymbol{b}\cdot\boldsymbol{\sigma})/2$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are unit vectors. If $\boldsymbol{a}$ and $\boldsymbol{b}$ are non-collinear, the observables A, B are incompatible. The compatible observables C, D to approximate A, B consist of positive operators $C_+ = (c_0\mathbb{1} + \boldsymbol{c}\cdot\boldsymbol{\sigma})/2$, $C_- = \mathbb{1} - C_+$, and $D_+ = (d_0\mathbb{1} + \boldsymbol{d}\cdot\boldsymbol{\sigma})/2$, $D_- = \mathbb{1} - D_+$, with $|\boldsymbol{c}| \le \min\{c_0, 2-c_0\} \le 1$ and $|\boldsymbol{d}| \le \min\{d_0, 2-d_0\} \le 1$. The operators $O_\pm$ $(O = A, B, C, D)$ are called effects, and the probabilities of measurement outcomes are $p^O_\pm = \mathrm{tr}\,(\rho O_\pm)$, where $\rho = (\mathbb{1} + \boldsymbol{r}\cdot\boldsymbol{\sigma})/2$ is the density operator of a qubit. We scale the outcomes of these measurements to be ± 1, and then the observable O is given as a map $\pm 1 \mapsto O_\pm$.

The statistical differences between the measurements of A and C and between the measurements of B and D, namely the inaccuracies, are [1]

$$\delta_\rho\,(A, C) := 2\sum_i \left|p^A_i - p^C_i\right| = 2\,|1 - c_0 + \boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{c})|\,, \quad \delta_\rho\,(B, D) := 2\sum_i \left|p^B_i - p^D_i\right| = 2\,|1 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{b}-\boldsymbol{d})|\,. \tag{1}$$

The combined inaccuracy is

$$\begin{aligned}\Delta_\rho(A, B; C, D) &:= \delta_\rho\,(A, C) + \delta_\rho\,(B, D) = 2\left(|1 - c_0 + \boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{c})| + |1 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{b}-\boldsymbol{d})|\right)\\ &= 2\cdot\begin{cases} |2 - c_0 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})|\,, & \text{if } [1 - c_0 + \boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{c})][1 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{b}-\boldsymbol{d})] \ge 0,\\ |c_0 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d})|\,, & \text{if } [1 - c_0 + \boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{c})][1 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{b}-\boldsymbol{d})] \le 0.\end{cases}\end{aligned} \tag{2}$$

By maximizing $\Delta_\rho(A, B; C, D)$ over all ρ, one obtains the worst-case combined inaccuracy $\Delta(A, B; C, D)$. Since

$$\begin{aligned}|2 - c_0 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})| &\le |2 - c_0 - d_0| + |\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,,\\ |c_0 - d_0 + \boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d})| &\le |c_0 - d_0| + |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\,,\end{aligned} \tag{3}$$

one has

$$\Delta(A, B; C, D) := \max_\rho\, \Delta_\rho(A, B; C, D) \le 2\max\left\{|2 - c_0 - d_0| + |\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,, |c_0 - d_0| + |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\right\}. \tag{4}$$

In fact, the equality in inequality (4) holds all the time. In other words, the right side of this inequality can always be reached by some Δ_{ρ_m} with a certain state ρ_m, and the Bloch vector of such state is

$$\boldsymbol{r}_\mathrm{m} = \begin{cases} (\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})/\,|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,, & \text{if } c_0 + d_0 \le 2 \text{ and } u \ge v,\\ -(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})/\,|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,, & \text{if } c_0 + d_0 \ge 2 \text{ and } u \ge v,\\ (\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d})/\,|\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\,, & \text{if } c_0 \le d_0 \text{ and } u \le v,\\ -(\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d})/\,|\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\,, & \text{if } c_0 \ge d_0 \text{ and } u \le v,\end{cases} \tag{5}$$

where $u = |2 - c_0 - d_0| + |\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|$ and $v = |c_0 - d_0| + |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|$. If the singular situation where $\boldsymbol{r}_\mathrm{m} = \boldsymbol{0}/0$ is encountered, it means that $\boldsymbol{r}_\mathrm{m}$ can be any Bloch vector of a qubit.

We only prove the first case where $c_0 + d_0 \le 2$ and $u \ge v$ with an additional assumption $c_0 \ge d_0$. Now $u \ge v$ turns into

$$2(1 - c_0) + |\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}| - |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}| \ge 0. \tag{6}$$

Since any two vectors $\boldsymbol{m}$ and $\boldsymbol{n}$ satisfy $|\boldsymbol{m}+\boldsymbol{n}|\,|\boldsymbol{m}-\boldsymbol{n}| \ge (\boldsymbol{m}+\boldsymbol{n})\cdot(\boldsymbol{m}-\boldsymbol{n}) = \boldsymbol{m}^2 - \boldsymbol{n}^2$, one has

$$|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,|\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}| \ge (\boldsymbol{b}-\boldsymbol{d})^2 - (\boldsymbol{a}-\boldsymbol{c})^2, \tag{7}$$

If $\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d} \ne \boldsymbol{0}$, inequality (7) can be rearranged as

$$\frac{2(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})\cdot(\boldsymbol{a}-\boldsymbol{c})}{|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|} \ge |\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}| - |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\,. \tag{8}$$

Inequalities (6) and (8) yield

$$1 - c_0 + \boldsymbol{r}_\mathrm{m}\cdot(\boldsymbol{a}-\boldsymbol{c}) \ge 0, \tag{9}$$

Inequality (6) and the condition $c_0 \ge d_0$ imply

$$2(1 - d_0) + |\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}| - |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}| \ge 0. \tag{10}$$

Similar to inequality (7), one has

$$|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,|\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}| \geq (\boldsymbol{a}-\boldsymbol{c})^2-(\boldsymbol{b}-\boldsymbol{d})^2, \tag{11}$$

Inequalities (10) and (11) yield

$$1-d_0+\boldsymbol{r}_{\mathrm{m}}\cdot(\boldsymbol{b}-\boldsymbol{d}) \geq 0. \tag{12}$$

From inequalities (9) and (12) one obtains

$$\Delta_{\rho_{\mathrm{m}}}(A,B;C,D)=2\,|2-c_0-d_0+\boldsymbol{r}_{\mathrm{m}}\cdot(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})|=2(|2-c_0-d_0|+|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|). \tag{13}$$

If $\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}=\boldsymbol{0}$, inequalities (7) and (10) entail

$$1-c_0+\boldsymbol{r}\cdot(\boldsymbol{a}-\boldsymbol{c}) \geq 1-c_0-|\boldsymbol{a}-\boldsymbol{c}| \geq 0, \qquad 1-d_0+\boldsymbol{r}\cdot(\boldsymbol{b}-\boldsymbol{d}) \geq 1-d_0-|\boldsymbol{b}-\boldsymbol{d}| \geq 0. \tag{14}$$

Then one can get

$$\Delta_{\rho}(A,B;C,D)=2\,|2-c_0-d_0+\boldsymbol{r}\cdot(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d})|=2\,|2-c_0-d_0|\,. \tag{15}$$

Hence in this situation, $\Delta_{\rho}(A,B;C,D)$ loses the state dependence and $\boldsymbol{r}_{\mathrm{m}}=\boldsymbol{0}/0$ can be any Bloch vector of a qubit. The reasoning is similar for other cases in Eq. (5). Therefore, the worst-case inaccuracy of the joint measurement is

$$\Delta(A,B;C,D)=2\max\left\{|2-c_0-d_0|+|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,,|c_0-d_0|+|\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\right\}. \tag{16}$$

The next task is to minimize $\Delta(A,B;C,D)$ over all possible C,D. The procedure is outlined as follows.

Firstly, $\Delta(A,B;C,D)$ is bounded by a piecewise function, that is,

$$\Delta(A,B;C,D) \geq f(\boldsymbol{c},\boldsymbol{d}) := 2\cdot\begin{cases}|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}|\,, & \text{if}\ \ (\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d}) \geq 0,\\ |\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}|\,, & \text{if}\ \ (\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d}) \leq 0.\end{cases} \tag{17}$$

Here $f(\boldsymbol{c},\boldsymbol{d})$ is a function of $\boldsymbol{c}$ and $\boldsymbol{d}$ for given $\boldsymbol{a}$ and $\boldsymbol{b}$, and the inequality holds iff $c_0=d_0=1$. As mentioned in the main text, the criterion of compatibility when $c_0=d_0=1$ is [1–3]

$$|\boldsymbol{c}+\boldsymbol{d}|+|\boldsymbol{c}-\boldsymbol{d}| \leq 2. \tag{18}$$

Secondly, the function $f(\boldsymbol{c},\boldsymbol{d})$ is bounded by another function, that is,

$$f(\boldsymbol{c},\boldsymbol{d}) \geq 2\sqrt{|\boldsymbol{a}-\boldsymbol{c}|^2+|\boldsymbol{b}-\boldsymbol{d}|^2}. \tag{19}$$

The equality holds iff $(\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d})=0$. We define $g(\boldsymbol{c},\boldsymbol{d}) := |\boldsymbol{a}-\boldsymbol{c}|^2+|\boldsymbol{b}-\boldsymbol{d}|^2$ for convenience.

Thirdly, the equality in the inequality (18), namely $|\boldsymbol{c}+\boldsymbol{d}|+|\boldsymbol{c}-\boldsymbol{d}|=2$, is a necessary condition for the function $g(\boldsymbol{c},\boldsymbol{d})$ to attain its minimum. This proposition is proved in a geometric way. The inequality (18) defines a region enclosed by an ellipsoid $E_{\boldsymbol{c}}$ or $E_{\boldsymbol{d}}$ for given $\boldsymbol{c}$ or $\boldsymbol{d}$, respectively. The interior of the ellipsoid corresponds to "<2" while the boundary corresponds to "$=2$". Without loss of generality, one can assume $\boldsymbol{c}$ is given and $E_{\boldsymbol{c}}$ is defined as shown in Fig. S1. Suppose some $\boldsymbol{d}$ with its terminal point lies inside $E_{\boldsymbol{c}}$. The line segment bounded by the terminal points of $\boldsymbol{b}$ and $\boldsymbol{d}$ intersects $E_{\boldsymbol{c}}$, and the intersection determines another vector $\boldsymbol{d}'$. Since $|\boldsymbol{b}-\boldsymbol{d}|>|\boldsymbol{b}-\boldsymbol{d}'|$, one gets $g(\boldsymbol{c},\boldsymbol{d})>g(\boldsymbol{c},\boldsymbol{d}')$ immediately.

Fourthly, the minimum of $g(\boldsymbol{c},\boldsymbol{d})$ is $(|\boldsymbol{a}+\boldsymbol{b}|+|\boldsymbol{a}-\boldsymbol{b}|-2)^2/4$. This proposition is proved below. Let $\boldsymbol{c}\cdot\boldsymbol{d}=x$, and one can get

$$|\boldsymbol{c}|^2+|\boldsymbol{d}|^2=1+x^2, \quad |\boldsymbol{c}+\boldsymbol{d}|=1+x, \quad |\boldsymbol{c}-\boldsymbol{d}|=1-x, \tag{20}$$

from $|\boldsymbol{c}+\boldsymbol{d}|+|\boldsymbol{c}-\boldsymbol{d}|=2$. Let α denote the angle between $\boldsymbol{a}+\boldsymbol{b}$ and $\boldsymbol{c}+\boldsymbol{d}$, and β denote the angle between $\boldsymbol{a}-\boldsymbol{b}$ and $\boldsymbol{c}-\boldsymbol{d}$, that is,

$$\begin{aligned}(\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{c}+\boldsymbol{d})&=|\boldsymbol{a}+\boldsymbol{b}|\,|\boldsymbol{c}+\boldsymbol{d}|\cos\alpha=|\boldsymbol{a}+\boldsymbol{b}|\,(1+x)\cos\alpha,\\ (\boldsymbol{a}-\boldsymbol{b})\cdot(\boldsymbol{c}-\boldsymbol{d})&=|\boldsymbol{a}-\boldsymbol{b}|\,|\boldsymbol{c}-\boldsymbol{d}|\cos\beta=|\boldsymbol{a}-\boldsymbol{b}|\,(1-x)\cos\beta.\end{aligned} \tag{21}$$

From the above equalities, one has

$$\begin{aligned}\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{d}&=\frac{1}{2}\left[|\boldsymbol{a}+\boldsymbol{b}|\,(1+x)\cos\alpha+|\boldsymbol{a}-\boldsymbol{b}|\,(1-x)\cos\beta\right],\\ \boldsymbol{a}\cdot\boldsymbol{d}+\boldsymbol{b}\cdot\boldsymbol{c}&=\frac{1}{2}\left[|\boldsymbol{a}+\boldsymbol{b}|\,(1+x)\cos\alpha-|\boldsymbol{a}-\boldsymbol{b}|\,(1-x)\cos\beta\right].\end{aligned} \tag{22}$$

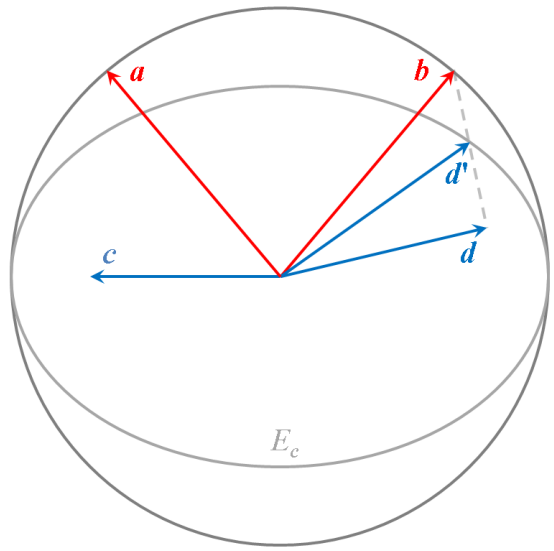


FIG. S1. (color online). The line segment bounded by the terminal points of $\boldsymbol{b}$ and $\boldsymbol{d}$ intersects $E_{\boldsymbol{c}}$, and the intersection determines another vector $\boldsymbol{d}'$. The original three-dimensional picture is illustrated in a two-dimensional manner here.

Now $g(\boldsymbol{c},\boldsymbol{d})$ can be rewritten in terms of three independent variables x,α,β as

$$\begin{aligned} g(\boldsymbol{c},\boldsymbol{d}) &= |\boldsymbol{a}-\boldsymbol{c}|^2+|\boldsymbol{b}-\boldsymbol{d}|^2=|\boldsymbol{a}|^2+|\boldsymbol{b}|^2+|\boldsymbol{c}|^2+|\boldsymbol{d}|^2-2(\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{d}) \\ &= x^2-|\boldsymbol{a}+\boldsymbol{b}|\,(1+x)\cos\alpha-|\boldsymbol{a}-\boldsymbol{b}|\,(1-x)\cos\beta+3 \\ &= x^2-(|\boldsymbol{a}+\boldsymbol{b}|\cos\alpha-|\boldsymbol{a}-\boldsymbol{b}|\cos\beta)\,x-|\boldsymbol{a}+\boldsymbol{b}|\cos\alpha-|\boldsymbol{a}-\boldsymbol{b}|\cos\beta+3. \end{aligned} \tag{23}$$

The above function gets its minimum $(|\boldsymbol{a}+\boldsymbol{b}|+|\boldsymbol{a}-\boldsymbol{b}|-2)^2/4$ when

$$\alpha=\beta=0,\qquad x=\frac{|\boldsymbol{a}+\boldsymbol{b}|-|\boldsymbol{a}-\boldsymbol{b}|}{2}, \tag{24}$$

which are equivalent to

$$\begin{aligned} \boldsymbol{c} &= \frac{1}{2}\left[\left(1+\frac{|\boldsymbol{a}+\boldsymbol{b}|-|\boldsymbol{a}-\boldsymbol{b}|}{2}\right)\frac{\boldsymbol{a}+\boldsymbol{b}}{|\boldsymbol{a}+\boldsymbol{b}|}+\left(1+\frac{|\boldsymbol{a}-\boldsymbol{b}|-|\boldsymbol{a}+\boldsymbol{b}|}{2}\right)\frac{\boldsymbol{a}-\boldsymbol{b}}{|\boldsymbol{a}-\boldsymbol{b}|}\right], \\ \boldsymbol{d} &= \frac{1}{2}\left[\left(1+\frac{|\boldsymbol{a}+\boldsymbol{b}|-|\boldsymbol{a}-\boldsymbol{b}|}{2}\right)\frac{\boldsymbol{a}+\boldsymbol{b}}{|\boldsymbol{a}+\boldsymbol{b}|}-\left(1+\frac{|\boldsymbol{a}-\boldsymbol{b}|-|\boldsymbol{a}+\boldsymbol{b}|}{2}\right)\frac{\boldsymbol{a}-\boldsymbol{b}}{|\boldsymbol{a}-\boldsymbol{b}|}\right]. \end{aligned} \tag{25}$$

for incompatible observables A,B with non-collinear $\boldsymbol{a}$ and $\boldsymbol{b}$. If A,B are compatible, that is, $\boldsymbol{a}$ and $\boldsymbol{b}$ are collinear, one has the minimum attained at $\boldsymbol{c}=\boldsymbol{a}$ and $\boldsymbol{d}=\boldsymbol{b}$, and such conditions agree with Eqs. (24) and (25). Note that Eq. (24) entails

$$(\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d})=\boldsymbol{a}\cdot\boldsymbol{b}+x-\frac{1}{2}\left[|\boldsymbol{a}+\boldsymbol{b}|\,(1+x)\cos\alpha-|\boldsymbol{a}-\boldsymbol{b}|\,(1-x)\cos\beta\right]=0. \tag{26}$$

Therefore, the two equalities in the inequality chain $f(\boldsymbol{c},\boldsymbol{d})\geq 2\sqrt{g(\boldsymbol{c},\boldsymbol{d})}\geq|\boldsymbol{a}+\boldsymbol{b}|+|\boldsymbol{a}-\boldsymbol{b}|-2$ achieve under the same condition. Overall we have $\Delta(A,B;C,D)\geq\Delta_{\rm lb}(A,B):=|\boldsymbol{a}+\boldsymbol{b}|+|\boldsymbol{a}-\boldsymbol{b}|-2$.

II. DISCUSSION ON THE INACCURACY TRADE-OFF RELATION FOR QUBIT MEASUREMENTS

A. State-dependent inaccuracy trade-off relation

Although one could use the state-dependent inaccuracy measure $\delta_\rho(A,C)$, $\delta_\rho(B,D)$, and $\Delta_\rho(A,B;C,D)$ to study the precision of measurements in a specific state, this framework does not support a nontrivial state-dependent trade-off relation in the form of $\Delta_\rho(A,B;C,D)\geq h_\rho(A,B)$, where $h_\rho(A,B)$ is a function or some combination of A, B, and

ρ. Since for arbitrary A, B, and ρ, one can construct compatible observables C, D as

$$c_0 = d_0 = 1, \qquad \boldsymbol{c} = \frac{\boldsymbol{r}\cdot\boldsymbol{a}}{|\boldsymbol{r}|^2}\boldsymbol{r}, \qquad \boldsymbol{d} = \frac{\boldsymbol{r}\cdot\boldsymbol{b}}{|\boldsymbol{r}|^2}\boldsymbol{r}, \tag{27}$$

such that $\delta_\rho(A,C) = \delta_\rho(B,D) = \Delta_\rho(A,B;C,D) = 0$, the bound $h_\rho(A,B)$ always vanishes and the state-dependent trade-off relation $\Delta_\rho(A,B;C,D) \geq h_\rho(A,B)$ is trivial. As shown in Fig. S2, the specific vectors $\boldsymbol{c}, \boldsymbol{d}$ are the projections of $\boldsymbol{a}, \boldsymbol{b}$ onto $\boldsymbol{r}$.

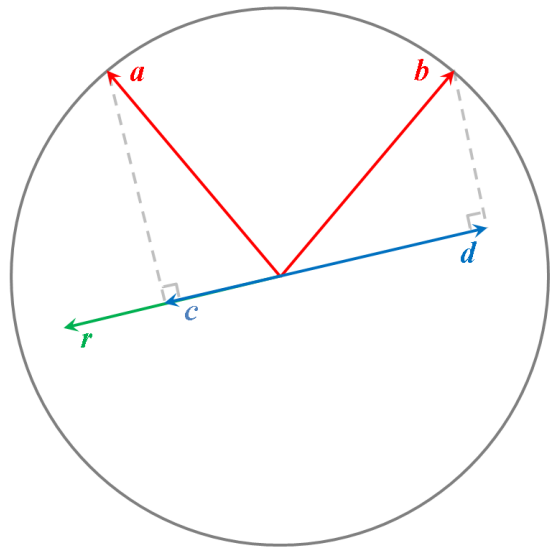


FIG. S2. (color online). The vectors $\boldsymbol{c}, \boldsymbol{d}$ are the projections of $\boldsymbol{a}, \boldsymbol{b}$ onto $\boldsymbol{r}$. The original three-dimensional picture is illustrated in a two-dimensional manner here.

B. Comparison with BLW's original relation

Here we repeat the meaning of observables A, B, C, D. The target observables A, B, which are typically incompatible, have the spectral projections $A_\pm = (\mathbb{1} \pm \boldsymbol{a}\cdot\boldsymbol{\sigma})/2$ and $B_\pm = (\mathbb{1} \pm \boldsymbol{b}\cdot\boldsymbol{\sigma})/2$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are unit vectors. The compatible observables C, D to approximate A, B consist of positive operators $C_+ = (c_0\mathbb{1} + \boldsymbol{c}\cdot\boldsymbol{\sigma})/2$, $C_- = \mathbb{1} - C_+$, and $D_+ = (d_0\mathbb{1} + \boldsymbol{d}\cdot\boldsymbol{\sigma})/2$, $D_- = \mathbb{1} - D_+$, with $|\boldsymbol{c}| \leq \min\{c_0, 2-c_0\} \leq 1$ and $|\boldsymbol{d}| \leq \min\{d_0, 2-d_0\} \leq 1$.

Since the observables C, D are jointly measured, they always face the same quantum state ρ, and then the state-independent inaccuracy in joint measurement is naturally defined as $\Delta(A,B;C,D) := \max_\rho[\delta_\rho(A,C) + \delta_\rho(B,D)]$. The resultant inaccuracy trade-off relation is

$$\Delta(A,B;C,D) := \max_\rho\left[\delta_\rho(A,C) + \delta_\rho(B,D)\right] \geq \Delta_{\rm lb}(A,B) := |\boldsymbol{a}+\boldsymbol{b}| + |\boldsymbol{a}-\boldsymbol{b}| - 2. \tag{28}$$

In BLW's original treatment, $\delta_\rho(A,C)$ and $\delta_\rho(B,D)$ are maximized separately, and the state-independent inaccuracy is $\widetilde{\Delta}(A,B;C,D) = \max_\rho \delta_\rho(A,C) + \max_\rho \delta_\rho(B,D)$. With the symbols adopted in this paper, their inaccuracy trade-off relation is rewritten as

$$\widetilde{\Delta}(A,B;C,D) := \max_\rho \delta_\rho(A,C) + \max_\rho \delta_\rho(B,D) \geq \sqrt{2}\Delta_{\rm lb}(A,B) = \sqrt{2}\left(|\boldsymbol{a}+\boldsymbol{b}| + |\boldsymbol{a}-\boldsymbol{b}| - 2\right). \tag{29}$$

But generally, $\delta_\rho(A,C)$ and $\delta_\rho(B,D)$ are maximized by different ρ. The definitions of $\delta_\rho(A,C)$ and $\delta_\rho(B,D)$ in Eq. (1) entail that $\delta_\rho(A,C)$ is maximized by $\boldsymbol{r}_1 = (\boldsymbol{a}-\boldsymbol{c})/|\boldsymbol{a}-\boldsymbol{c}|$ or $(\boldsymbol{c}-\boldsymbol{a})/|\boldsymbol{a}-\boldsymbol{c}|$ while $\delta_\rho(B,D)$ is maximized by $\boldsymbol{r}_2 = (\boldsymbol{b}-\boldsymbol{d})/|\boldsymbol{b}-\boldsymbol{d}|$ or $(\boldsymbol{d}-\boldsymbol{b})/|\boldsymbol{b}-\boldsymbol{d}|$. To be more specific, a concrete example is given as follows. For the observable A, B, C, D with

$$\boldsymbol{a} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad \boldsymbol{b} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad \boldsymbol{c} = \left(-\frac{1}{2}, \frac{1}{2}, 0\right), \quad \boldsymbol{d} = \left(\frac{1}{2}, \frac{1}{2}, 0\right), \quad c_0 = d_0 = 1, \tag{30}$$

$\delta_\rho(A,C)$ is maximized by $\rho_1 = (\mathbb{1} + \boldsymbol{r}_1 \cdot \boldsymbol{\sigma})/2$ with $\boldsymbol{r}_1 = \left(\pm 1/\sqrt{2}, \mp 1/\sqrt{2}, 0\right)$ and $\delta_\rho(B,D)$ is maximized by $\rho_2 = (\mathbb{1} + \boldsymbol{r}_2 \cdot \boldsymbol{\sigma})/2$ with $\boldsymbol{r}_2 = \left(\pm 1/\sqrt{2}, \pm 1/\sqrt{2}, 0\right)$, while $\delta_\rho(A,C) + \delta_\rho(B,D)$ is maximized by $\rho_{\mathrm{m}} = (\mathbb{1} + \boldsymbol{r}_{\mathrm{m}} \cdot \boldsymbol{\sigma})/2$ with $\boldsymbol{r}_{\mathrm{m}} = (\pm 1, 0, 0)$ or $(0, \pm 1, 0)$. As the argument here indicates, the treatment in (29) implies, in some sense, that C, D face different states at the same time. Such problem is naturally avoided by the treatment in (28), where the jointly measured observables C, D always face the same quantum state at the same time.

Both of the two state-independent quantities $\Delta(A,B;C,D)$ and $\widetilde{\Delta}(A,B;C,D)$ characterize the integrated deviation of C, D from A, B and act as figures of merit of the measurement device, but they have several differences. Firstly, from the definitions of the two quantities, one can see that $\Delta(A,B;C,D)$ is the worst-case inaccuracy in the joint measurement while $\widetilde{\Delta}(A,B;C,D)$ does not have an explicit physical meaning. Secondly, the inequality $\Delta(A,B;C,D) \leq \widetilde{\Delta}(A,B;C,D)$ always holds and $\Delta(A,B;C,D)$ has a delicate lower bound without the extra coefficient $\sqrt{2}$. Thirdly, the two quantities are not monotonic with each other, videlicet, sometimes they have conflict in comparing the performance of different pairs of C, D. That is to say, a device composed of C, D outperforms another device composed of C', D' under the evaluation by $\widetilde{\Delta}$ for given A, B, but the evaluation by Δ may be opposite sometimes. To be more specific, we give a concrete example, and the parameters of these observables are

$$\begin{aligned} &\boldsymbol{a} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad \boldsymbol{b} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad c_0 = d_0 = c_0' = d_0' = 1, \\ &\boldsymbol{c} = \left(\frac{1}{\sqrt{2}} - 1, \frac{1}{\sqrt{2}}, 0\right), \quad \boldsymbol{d} = \left(1 - \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad \boldsymbol{c}' = \left(-\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, 0\right), \quad \boldsymbol{d}' = \left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}, 0\right). \end{aligned} \tag{31}$$

The assessment of the quality of C, D yields $\widetilde{\Delta}(A,B;C,D) = \Delta(A,B;C,D) = 4\left(\sqrt{2} - 1\right)$, and the assessment of the quality of C', D' yields $\widetilde{\Delta}(A,B;C',D') = 2$ and $\Delta(A,B;C',D') = \sqrt{2}$. One can see that $\widetilde{\Delta}(A,B;C,D) < \widetilde{\Delta}(A,B;C',D')$ while $\Delta(A,B;C,D) > \Delta(A,B;C',D')$. In view of the physical meaning, Δ could be preferable in evaluating the performance of the joint-measurement device.

Despite these differences, $\Delta(A,B;C,D)$ and $\widetilde{\Delta}(A,B;C,D)$ attain their respective lower bounds $\Delta_{\mathrm{lb}}(A,B)$ and $\sqrt{2}\Delta_{\mathrm{lb}}(A,B)$ under the same conditions, which are $c_0 = d_0 = 1$ and $\boldsymbol{c}, \boldsymbol{d}$ in Eq. (25). This fact indicates that the two quantities are consistent in figuring out the optimal joint measurement.

III. EXPERIMENTAL METHODS

A. Quantum gates

The quantum gates illustrated in Fig. 2c,d of the main text can be represented by the rotation about the y axis, and the matrix form of such rotation is

$$\mathscr{R}(\zeta) = \begin{pmatrix} \cos\zeta & -\sin\zeta \\ \sin\zeta & \cos\zeta \end{pmatrix} \tag{32}$$

in the basis $|+\rangle = (1,0)^{\mathrm{T}}, |-\rangle = (0,1)^{\mathrm{T}}$. To construct the observables C, D with the Bloch vectors in Fig. 3a, the gates for measurement are selected as

$$R_A = \mathscr{R}\left(\frac{\pi}{2}\right), \qquad R = \mathscr{R}(\pi - 2\gamma). \tag{33}$$

For another series of C, D with the Bloch vectors in Fig. 4a, the gates are

$$R_A = \mathscr{R}\left(\arccos\frac{\sin\frac{\varphi}{2} - \cos\frac{\varphi}{2}}{\sin\frac{\varphi}{2} + \cos\frac{\varphi}{2}}\right), \qquad R = \mathscr{R}\left(\frac{\pi}{2}\right). \tag{34}$$

For the observables A, B with the Bloch vectors in Fig. 4a, the associated gates, in turn, are

$$V = \mathscr{R}\left(\frac{\theta}{2}\right), \qquad \mathscr{R}\left(-\frac{\theta}{2}\right). \tag{35}$$

B. Pulse sequences

The pulse sequence for the joint measurement of C, D is shown in Fig. 2e of the main text. A shape pulse based on the gradient ascent pulse engineering (GRAPE) algorithm and a pulsed field gradient are utilized to prepare the PPS. After such initialization, square pulses are applied. The controlled gates on the system and ancilla qubits are realized by hard pulses and free evolution. The free evolution time τ_1 or τ_2 depends on the flip angle of the related controlled gate. For the flip angle ζ (radian), the free evolution time is $\zeta/(2\pi J)$, where J is the scalar coupling between ^{19}F and ^{1}H. In our experiment, τ_1 varies with the R gate while τ_2 is fixed to $1/(2J)$ for the X gate. The flip angles of the three pulses with dotted borders also varies with their related quantum gates R_S, R_A, and R. The two pulses outside the dashed frame have the flip angles identical to that of the respective single-qubit gates R_S and R_A. But as a part of the controlled operation, the gate R with the flip angle ζ demands the flip angle of the pulse before τ_2 to be $\zeta/2$. The two π pulses applied on ^{13}C are for decoupling, and each of them sits on the midpoint of one free evolution period. The final $\pi/2$ pulse on ^{13}C is for readout. All the square pulses are hard pulses with negligible duration compared to the free evolution time. The pulse sequence for the measurement of A or B contains one pulse on ^{19}F to realize the combined effect of the two single-qubit gates and one $\pi/2$ pulse on ^{13}C for readout.

C. Experimental errors

Decoherence is negligible during hard pulses but effective during free evolution periods in the joint measurement of C, D (Fig. 2e). The numerical simulation takes such decoherence into account, and agrees with the experimental results. The contrast between dashed and solid curves in Figs. 3 and 4 indicates that the deviation of the experimental results from the theory primarily stems from decoherence. Since the joint measurement of C, D contains more pulses than the separate measurement of A, B, the former is more vulnerable to the radiofrequency (RF) inhomogeneity, which incurs a system error. The fluctuation of signal strength and fitting error lead to random readout error. Both the RF inhomogeneity and the readout error are responsible for the differences between experimental results and the simulation.

IV. ADDITIONAL EXPERIMENTAL DATA

Equation (5) indicates that the state $\rho_{\mathrm{m}} = (\mathbb{1} + \boldsymbol{r}_{\mathrm{m}} \cdot \boldsymbol{\sigma})/2$ maximizing $\Delta_\rho(A, B; C, D)$ depends on the vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ and c_0, d_0. When $c_0 = d_0 = 1$, Eq. (5) reduces to

$$\boldsymbol{r}_{\mathrm{m}} = \begin{cases} \pm\left(\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}\right)/\left|\boldsymbol{a}+\boldsymbol{b}-\boldsymbol{c}-\boldsymbol{d}\right|, & \text{if} \quad (\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d}) \geq 0, \\ \pm\left(\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}\right)/\left|\boldsymbol{a}-\boldsymbol{b}-\boldsymbol{c}+\boldsymbol{d}\right|, & \text{if} \quad (\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d}) \leq 0. \end{cases} \tag{36}$$

In the situation described by Fig. 3a of the main text, the Bloch vectors of the observables are

$$\boldsymbol{a} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \boldsymbol{b} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \boldsymbol{c} = \left(-\sin\gamma\cos\gamma, \sin^2\gamma\right), \quad \boldsymbol{d} = \left(\sin\gamma\cos\gamma, \cos^2\gamma\right), \tag{37}$$

where the former and latter components correspond to x and z, respectively (hereinafter the same). Since $(\boldsymbol{a}-\boldsymbol{c})\cdot(\boldsymbol{b}-\boldsymbol{d}) \leq 0$ always holds, we select $\boldsymbol{r}_{\mathrm{m}} = \left(\sqrt{2}-\sin 2\gamma, -\cos 2\gamma\right)/\sqrt{3-2\sqrt{2}\sin 2\gamma}$. The primitive experimental data for Fig. 3 of the main text are shown in Fig. S3.

In the situation described by Fig. 4a of the main text, the Bloch vectors of the observables are

$$\begin{aligned} &\boldsymbol{a} = \left(-\sin\frac{\theta}{2}, \cos\frac{\theta}{2}\right), \qquad \boldsymbol{b} = \left(\sin\frac{\theta}{2}, \cos\frac{\theta}{2}\right), \\ &\boldsymbol{c} = \left(-\left(1+\cot\frac{\varphi}{2}\right)^{-1}, \left(1+\tan\frac{\varphi}{2}\right)^{-1}\right), \qquad \boldsymbol{d} = \left(\left(1+\cot\frac{\varphi}{2}\right)^{-1}, \left(1+\tan\frac{\varphi}{2}\right)^{-1}\right). \end{aligned} \tag{38}$$

From Eq. (25) one can figure out that the angle between the Bloch vectors $\boldsymbol{c}, \boldsymbol{d}$ of the optimal C, D is

$$\varphi_{\mathrm{opt}} = \arccos\frac{\cos\frac{\theta}{2}-\sin\frac{\theta}{2}}{1-\cos\frac{\theta}{2}\sin\frac{\theta}{2}}. \tag{39}$$

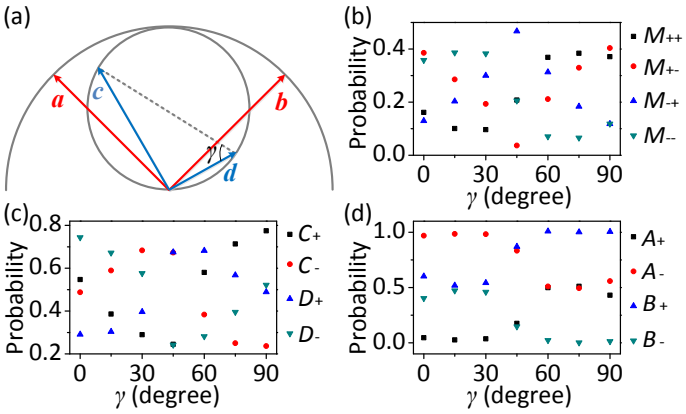


FIG. S3. (color online). Primitive data for Fig. 3 of the main text. (a) Bloch vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ of the observables A, B, C, D. (b) Probability distribution of measurement outcomes of the observable M. (c) Probability distributions of measurement outcomes of the observables C, D. These probabilities associated with C, D are combined from the probabilities associated with M. (d) Probability distributions of measurement outcomes of the observables A, B.

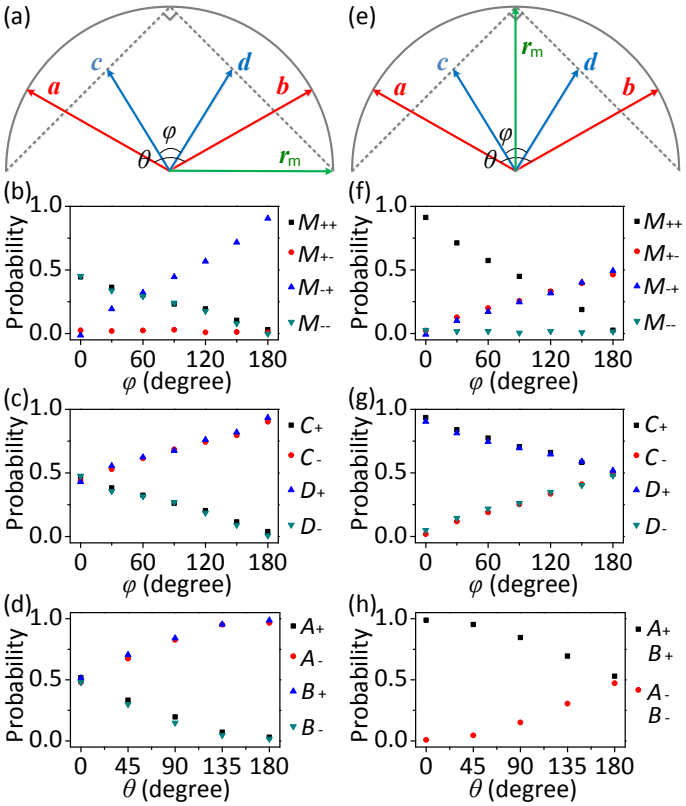


FIG. S4. (color online). Primitive data for Fig. 4 of the main text. (a) Bloch vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ and $\boldsymbol{r}_{\mathrm{m}} = (1, 0)$ of the observables A, B, C, D and the state ρ_{m}. (b-d) Probability distribution of measurement outcomes relating to the state with $\boldsymbol{r}_{\mathrm{m}} = (1, 0)$. (e) Bloch vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ and $\boldsymbol{r}_{\mathrm{m}} = (0, 1)$ of the observables A, B, C, D and the state ρ_{m}. (f-h) Probability distribution of measurement outcomes relating to the state with $\boldsymbol{r}_{\mathrm{m}} = (0, 1)$.

When $0 \leq \varphi < \varphi_{\mathrm{opt}}$, we have $(\boldsymbol{a} - \boldsymbol{c}) \cdot (\boldsymbol{b} - \boldsymbol{d}) < 0$ and select $\boldsymbol{r}_{\mathrm{m}} = (1, 0)$ in the experiment. When $\varphi_{\mathrm{opt}} \leq \varphi \leq 180°$, we have $(\boldsymbol{a} - \boldsymbol{c}) \cdot (\boldsymbol{b} - \boldsymbol{d}) \geq 0$ and select $\boldsymbol{r}_{\mathrm{m}} = (0, 1)$ in the experiment. The primitive experimental data for Fig. 4 of the main text are shown in Fig. S4.

[1] P. Busch, P. Lahti, and R. F. Werner, Phys. Rev. A **89**, 012129 (2014).
[2] S. Yu, N.-l. Liu, L. Li, and C. H. Oh, Phys. Rev. A **81**, 062116 (2010).
[3] P. Busch, Phys. Rev. D **33**, 2253 (1986).