Experimental Test of Entropic Noise-Disturbance Uncertainty Relations for Spin-1/2 Measurements

Author

Sulyok, Georg, Sponar, Stephan, Demirel, Bülent, Buscemi, Francesco, Hall, Michael JW, Ozawa, Masanao, Hasegawa, Yuji

Published

Journal Title

Version

DOI

Copyright Statement

Downloaded from

Griffith Research Online

Experimental test of entropic noise-disturbance uncertainty relations for spin-$1/2$ measurements

Georg Sulyok[1], Stephan Sponar[1], Bülent Demirel[1], Francesco Buscemi[2], Michael J. W. Hall[3], Masanao Ozawa[2], and Yuji Hasegawa[1]

[1]*Atominstitut, Vienna University of Technology, 1020 Vienna, Austria*

[2]*Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan and*

[3]*Centre for Quantum Computation and Communication Technology (Australian Research Council),*
Centre for Quantum Dynamics, Griffith University, Brisbane, QLD 4111, Australia

(Dated: June 18, 2015)

Information-theoretic definitions for noise and disturbance in quantum measurements were given in [Phys. Rev. Lett. **112**, 050401 (2014)] and a state-independent noise-disturbance uncertainty relation was obtained. Here, we derive a tight noise-disturbance uncertainty relation for complementary qubit observables and carry out an experimental test. Successive projective measurements on the neutron's spin-$1/2$ system, together with a correction procedure which reduces the disturbance, are performed. Our experimental results saturate the tight noise-disturbance uncertainty relation for qubits when an optimal correction procedure is applied.

Introduction - The uncertainty principle, first formulated by Heisenberg in 1927 [1], expresses an intuitive understanding of the physical consequences of noncommutativity. Heisenberg argued that it is impossible to simultaneously measure noncommuting observables with arbitrary precision, and used the famous γ-ray microscope thought-experiment to obtain the 'noise-disturbance' uncertainty relation

$$q_1 \, p_1 \sim h \tag{1}$$

for the product of the mean error (noise) q_1 of a position measurement and the discontinuous change (disturbance) p_1 of the particle's momentum. In the subsequent mathematical derivation of Eq. (1), he showed that the product of the position and momentum standard deviations, $\Delta q \, \Delta p$, was equal to $\hbar/2$ for a class of Gaussian wavefunctions, which was generalised by Kennard to

$$\Delta q \, \Delta p \geq \frac{\hbar}{2} \tag{2}$$

for all states [2].

Note that relation (2) sets a limitation as to how one can precisely prepare both the position and momentum of a quantum system, independently of whether these observables are actually measured. Hence, such preparation relations, whether formulated in terms of standard deviations [2, 3] or Shannon entropies [4], do not place any restrictions *per se* on joint or successive measurements of noncommuting observables.

In contrast, Heisenberg's formulation of the uncertainty relation in Eq. (1) is all about the unavoidable influence of measuring instruments on quantum systems: the more precisely one observable such as position is measured, the greater is the disturbance to another observable such as momentum [1, 5]. As an aside, we notice that Eq. (1) can in fact be derived from Eq. (2) under the repeatability hypothesis [6], which was implicitly assumed in most of arguments on quantum measurements until the 1970's [7].

A rigorous error-disturbance uncertainty relation, generalising Eq. (1) to arbitrary pairs of observables and measurements without assuming the repeatability hypothesis, was derived by Ozawa [8–10] and has recently received considerable attention. The validity of Ozawa's relation, as well as of a stronger version of this relation [11], were experimentally tested with neutrons [12–14] and with photons [15–19]. Other approaches generalising Heisenberg's original relation can be found, for example, in [20–22].

It is very natural to also seek a formulation of the uncertainty principle in terms of the information gained and lost due to measurement influences. Such a formulation was recently introduced by Buscemi *et al.* [1], leading to a state-independent uncertainty relation. Here, noise and disturbance are quantified not by a difference between a system observable and the quantity actually measured, but by the correlations between input states and measurement outcomes, independently of how they are labelled. In this letter we derive a tight uncertainty relation for information-theoretic noise and disturbance, in the qubit case, and demonstrate its validity in a neutron polarimeter experiment.

Theoretical framework - Consider an observable A, acting on a finite-dimensional Hilbert space, with eigenvalues α belonging to the non-degenerate eigenstates $|a\rangle$ and a measurement apparatus $\mathcal{M}$ representing a quantum instrument [24–26] with possible outcomes μ. All eigenstates $|a\rangle$ of A are now fed with equal probability into the apparatus, which is schematically illustrated in Fig. 1. The conditional probability $p(\alpha|\mu)$ that the eigenstate $|a\rangle$ was sent, given a specific measurement outcome
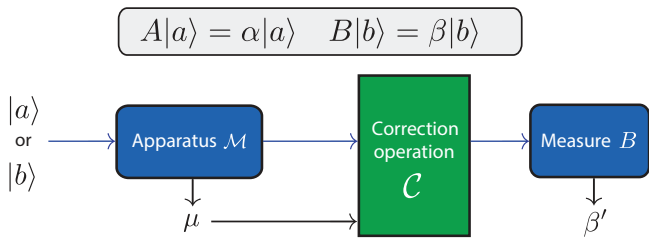
Figure 1: (color online) Experimental concept for determination of noise and disturbance. Randomly selected eigenstates of A and B are sent into a measurement apparatus $\mathcal{M}$. After a correction operation $\mathcal{C}$, and a precise measurement of B, the information-theoretic noise $N(\mathcal{M}, A)$ and disturbance $D(\mathcal{M}, B)$ are calculated using the conditional probabilities $p(\alpha|\mu)$ and $p(\beta|\beta')$, respectively.

μ, and the marginal probability $p(\mu)$ for occurrence of the specific outcome are used to define the information-theoretic noise $N(\mathcal{M}, A)$ as

$$N(\mathcal{M}, A) := -\sum_{\alpha,\mu} p(\mu)p(\alpha|\mu) \log p(\alpha|\mu) = H(\mathbb{A}|\mathbb{M}). \tag{3}$$

Equation (22) is just the conditional entropy $H(\mathbb{A}|\mathbb{M})$, where $\mathbb{A}$ and $\mathbb{M}$ denote the classical random variables associated with input α and output μ. The information-theoretic noise thus quantifies how well the value of A can be inferred from the measurement outcome and only vanishes if an absolutely correct guess is possible.

The information-theoretic disturbance is defined in a similar manner as

$$D(\mathcal{M}, B) := -\sum_{\beta,\beta'} p(\beta')p(\beta|\beta') \log p(\beta|\beta') = H(\mathbb{B}|\mathbb{B}'). \tag{4}$$

Here uniformly distributed eigenstates $|b\rangle$ of an observable B are input to the apparatus $\mathcal{M}$, and a subsequent measurement of B is performed, with outcomes labeled by β' (Fig. 1). The disturbance $D(\mathcal{M}, B)$ thus quantifies the correlation between the initial and final values of B, and is a measure of how much information about B is lost through the measurement $\mathcal{M}$.

In order to determine the *irreversible* loss of information about B, a correction operation $\mathcal{C}$ can be performed before the B-measurement to decrease the disturbance (Fig. 1), and consists of any completely positive, trace preserving map. We deal with two cases here; one is the uncorrected disturbance which we write as D_0. The other is the optimally corrected disturbance denoted as D_{opt} corresponding to the correction operation that minimizes the disturbance. For any correction procedure the information theoretic noise and disturbance fulfil the following uncertainty relation [1]

$$N(\mathcal{M}, A) + D(\mathcal{M}, B) \geq c_{AB} := -\log \max |\langle a|b\rangle|^2, \tag{5}$$

where $|a\rangle$ and $|b\rangle$ denote the eigenstates of the observables A and B.

For maximally incompatible qubit observables, represented by the Pauli matrices σ_z and σ_y, we have been able to significantly strengthen this relation (see Sec. I of the Supplemental Material [27]) to the tight relation

$$g[N(\mathcal{M}, \sigma_z)]^2 + g[D(\mathcal{M}, \sigma_y)]^2 \leq 1. \tag{6}$$

Here $g[x]$ denotes the inverse of the function $h(x)$ on the interval $x \in [0, 1]$ given by

$$h(x) := -\frac{1+x}{2} \log \frac{1+x}{2} - \frac{1-x}{2} \log \frac{1-x}{2}. \tag{7}$$

Experimental procedure - In our experiment projective measurements on neutron spin qubits are utilized. The observables are chosen to be Pauli spin matrices $A = \sigma_z$ and $B = \sigma_y$, having the eigenvalues $\alpha = \pm 1$ and $\beta = \pm 1$. We denote the eigenstates of A and B as $|\alpha z\rangle \equiv |\pm z\rangle$ for $\alpha = \pm 1$ and $|\beta y\rangle \equiv |\pm y\rangle$ for $\beta = \pm 1$, respectively. For projective measurements the measurement apparatus $\mathcal{M}$ is simply characterized by a measurement operator

$$M = \vec{m} \cdot \vec{\sigma} = \sigma_y \sin \theta + \sigma_z \cos \theta, \tag{8}$$

representing spin along the axis $\vec{m} = (0, \sin \theta, \cos \theta)$. It has the eigenvalues/outcomes $\mu = \pm 1$ and projects the system onto its eigenstates denoted as $|\mu m\rangle \equiv |\pm m\rangle$ for $\mu = \pm 1$ after the measurement.

The experiment is performed on the neutron's spin-$1/2$ qubit system using the polarimeter beam line of the tangential beam port at the research reactor facility TRIGA Mark II of the Vienna University of Technology [12, 13]. The setup is depicted in Fig. 2 and illustrates the generic
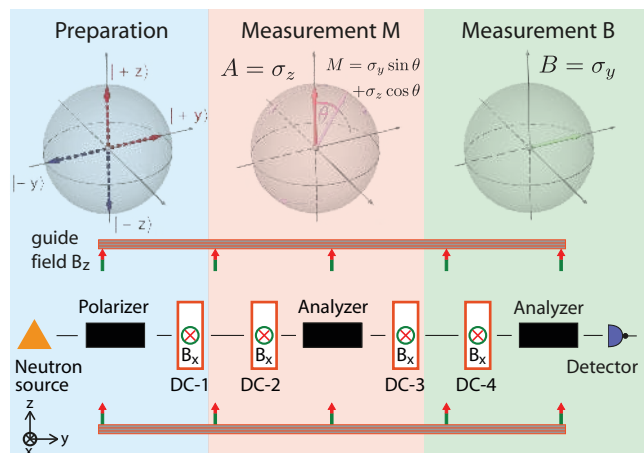

Figure 2: (color online) Neutron polarimetric setup for the demonstration of information-theoretic uncertainty relations for noise and disturbance. Exploiting Larmor precession of the Bloch vector around magnetic fields (B_x, B_z) and using supermirror arrays (polarizer, analyzers) as projectors all required spin states can be prepared and measured.

experimental procedure: An unpolarized thermal neutron beam, incident from a pyrolytic graphite crystal, with a mean wavelength of 2.02 Å and spectral width $\Delta\lambda/\lambda = 0.015$, is spin polarized up to $\sim 99\%$ via reflection from a bent Co-Ti supermirror array, with polarization in the $+z$-direction. To prevent depolarization by stray fields a 13 Gauss guide field B_z pointing in the $+z$-direction is applied along the entire setup.

For the generation of the desired initial states $|\pm z\rangle$ and $|\pm y\rangle$ the first spin turner coil DC-1 is used. Within the coil region a field B_x, pointing in $+x$-direction, is effectively applied. Larmor precession around the x-axis is induced and the strength of B_x is tuned such that it causes a spin rotation by an angle of 0, π, or $\pm\pi/2$ radians within the coil DC-1. In order to achieve the uniform distribution of the eigenstates as required for the determination of noise and disturbance all four input states are sent one after another.

For the measurement of M another spin turner coil (DC-2) is used. It is placed such that within the distance to DC-1 integer multiples of the full rotation period around the z-axis are performed in the guide field. Then, by correctly adjusting the strength of B_x in DC-2, the $|\mu m\rangle$-component of the spinor is rotated to $|+z\rangle$. After the projection onto $|+z\rangle$ in the second supermirror (first analyzer in Fig. 2) spin turner coil DC-3 rotates the analyzer's output state $|+z\rangle$ to $|\mu m\rangle$ thus completing the projective measurement of M. In an analogous way, DC-4 and the third supermirror perform the B-measurement. The recovering of the eigenstates of B can be omitted since the neutron detector is not sensitive to spin (for more details of the experimental procedure see Sec. II of [27]).

The two successively performed projective spin measurements result in four output intensities for each input eigenstate. We label the intensities as $I^A_{\alpha\mu\beta'}$ and $I^B_{\beta\mu\beta'}$ where all lower indices can take the values ± 1. The different upper indices A and B discriminate between the input states. For example, $I^A_{+\cdot\cdot}$ indicates that the eigenstate $|+z\rangle$ of $A = \sigma_z$ has been sent and $I^B_{+\cdot\cdot}$ stands for output intensities when $|+y\rangle$ has been fed to the measurement apparatus. From $I^A_{\alpha\mu\beta'}$, the probabilities required for the determination of the information theoretic noise can be deduced, and $I^B_{\beta\mu\beta'}$ yields the probabilities for the information theoretic disturbance:

$$p(\alpha) = \frac{\sum_{\mu,\beta'} I^A_{\alpha\mu\beta'}}{\sum_{\alpha,\mu,\beta'} I^A_{\alpha\mu\beta'}} \qquad p(\mu|\alpha) = \frac{\sum_{\beta'} I^A_{\alpha\mu\beta'}}{\sum_{\mu,\beta'} I^A_{\alpha\mu\beta'}} \quad (9)$$

$$p(\beta) = \frac{\sum_{\mu,\beta'} I^B_{\beta\mu\beta'}}{\sum_{\beta,\mu,\beta'} I^B_{\beta\mu\beta'}} \qquad p(\beta'|\beta) = \frac{\sum_{\mu} I^B_{\beta\mu\beta'}}{\sum_{\mu,\beta'} I^B_{\beta\mu\beta'}} (10)$$

It is important to note here that we first determine which eigenstate has been sent and then record the probability for a specific outcome μ. We thus obtain the conditioned probabilities $p(\mu|\alpha)$ and $p(\beta'|\beta)$ respectively instead of
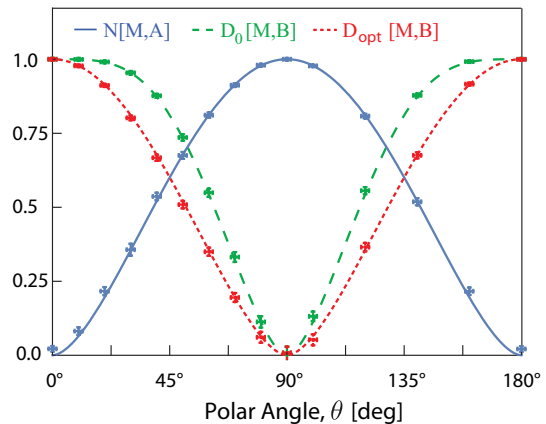

Figure 3: (color online) Noise $N(\mathcal{M}, A)$ (straight blue line), uncorrected disturbance $D_0(\mathcal{M}, B)$ (dashed green line), and optimally corrected disturbance $D_{\mathrm{opt}}(\mathcal{M}, B)$ (dotted red line) vs. polar angle θ of M.

$p(\alpha|\mu)$ and $p(\beta|\beta')$ for which noise and disturbance are defined. But, by using Bayes' theorem for conditioned probabilities they can be converted into each other (see Sec. III of [27]).

No correction procedure- In the first experiment, no additional correction is applied ($\mathcal{C} = \mathbb{1}$), we just successively measure M and B. The setup then consists of the three stages depicted in Fig. 2: i) state preparation - the corresponding eigenstates of the observables σ_z and σ_y are generated. ii) measurement of $M = \sigma_y \sin\theta + \sigma_z \cos\theta$, and iii) measurement of $B = \sigma_y$.

We record the intensities while varying the polar angle θ of M in the interval $\theta \in [0°, 90°]$ with increment $\Delta\theta = 10°$ and with a smaller step width of $\Delta\theta = 20°$ in interval $\theta \in [100°, 180°]$ since noise and disturbance are mirror-symmetric around $\theta = 90°$. Fig. 9 shows the measured data points and their theory curves, with the latter given in terms of h from Eq. (27) by

$$N(\mathcal{M}, A) = h(\cos\theta), \quad D_0(\mathcal{M}, B) = h(\sin^2\theta). \quad (11)$$

An intuitive understanding for the information-theoretic meaning of noise and disturbance can be reached by looking at special values of θ. For $\theta = 0°$, the measurement operator $M = \sigma_z = A$. The measurement result μ is numerically identical to the "value" α of the observable A in the eigenstate $|\alpha z\rangle$ and thus obviously perfectly correlated to it leading to vanishing noise. With increasing polar angle θ of M the correlation is lost. For $\theta = 90°$ ($M = \sigma_y$) the measurement outcome does not allow any inference as to which eigenstate of A was sent, $p(\alpha|\mu) = \frac{1}{2}$, $\forall \alpha, \mu$. Thus, the information-theoretic noise is maximal. For $\theta = 180°$ ($M = -\sigma_z$) the measurement result is perfectly anti-correlated with the observable's value. A guessing function of type $\alpha = f(\mu) = -\mu$ allows a flawless determination of the observable's value from the measurement outcome and thus, this setup is

noiseless as well, although the numerical deviation between the measurement outcome and the value of A is maximal. While in the standard "noise operator" approach [9, 10] the noise then becomes maximal (see for example error (dashed blue line) in Fig. 8 of [13]), in the information-theoretic approach the degree of correlation rather than its sign is relevant. For the same reason, the behaviour of noise for $\theta = 180°$ to $360°$ is identical to $\theta = 0°$ to $180°$ (see also Sec. III of [27]).

The disturbance induced on $B = \sigma_y$ by the measurement of M behaves reciprocally to the noise in Fig. 9, exemplifying the trade-off between noise and disturbance in Eq. (5). If $M = \sigma_z$ the result of a subsequent B-measurement allows no inference of the initially sent eigenstate $|\beta y\rangle$. The outcomes $\beta' = \pm 1$ are equally probable for both eigenstates and the information-theoretic disturbance becomes maximal. If $M = \sigma_y$ (for $\theta = 90°$) the outcomes of the B-measurement are perfectly correlated with the input eigenstates. The prior measurement of M leads to no loss of correlation between the β and the measurement outcome β' and the information theoretic disturbance vanishes. For $\theta = 180°$ ($M = -\sigma_z$), the correlation is lost entirely, we again have $p(\beta|\beta') = \frac{1}{2}$, $\forall \beta, \beta'$ as for $\theta = 0°$ and therefore maximal disturbance.

Optimal correction procedure- For the special cases $M = \pm A$ and $M = \pm B$ the disturbance is fixed to be either 1 or 0, but for the intermediate values, it can be reduced by performing a correction operation $\mathcal{C}$ after the M-measurement. An important class of correction operations are unitary transformations which could be experimentally realized by an additional spin turner device. However, we can concatenate state preparation after the M-measurement and correction $\mathcal{C}$ and immediately prepare the rotated state with DC-3 alone. In
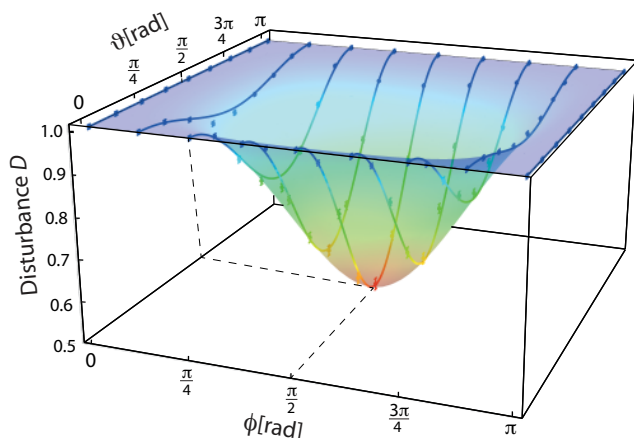
Fig .4, we show experimental results if, after the projective measurement of M with $\theta = 50°$, the eigenstates $|\mu m\rangle$ are rotated along arbitrary directions, that is, onto states $|\psi(\vartheta, \phi)\rangle = \cos\frac{\vartheta}{2}|+z\rangle + e^{i\phi}\sin\frac{\vartheta}{2}|-z\rangle$ for $|+m\rangle$ and $|-\psi\rangle = |\psi(\pi - \vartheta, \phi + \pi)\rangle$ for $|-m\rangle$. The directions of the output states are varied over the region $[\vartheta \times \phi] = [[0°, 180°] \times [0°, 180°]]$ with step widths $\Delta\vartheta = \Delta\phi = 22.5°$. The minimal disturbance is obtained if the eigenstates $|\pm m\rangle$ are rotated exactly onto the eigenstates $|\pm y\rangle$ of B, that is for $\vartheta = \phi = 90°$.

This experimental result can be generalized for the projective measurement of $M = \vec{m}\cdot\vec{\sigma}$ on observable $B = \vec{b}\cdot\vec{\sigma}$ yielding the optimal error correction $\mathcal{C}_{opt}$

$$\mathcal{C}_{opt}(|\mu m\rangle) := \begin{cases} |\mu b\rangle, & \vec{b}\cdot\vec{m} \geq 0 \\ |-\mu b\rangle, & \vec{b}\cdot\vec{m} < 0 \end{cases} \quad (12)$$

with $\mu = \pm 1$ and $|\pm m\rangle, |\pm b\rangle$ being the respective eigenstates of M and B (see Sec. IV of [27] for a detailed explanation and proof). The results for the optimally corrected disturbance are depicted in Fig. 9, together with its theoretically expected curve given, using $h(x)$ from Eq. (27), by

$$D_{opt}(\mathcal{M}, B) = h(\sin\theta). \quad (13)$$

In order to investigate the uncertainty relations Eqs. (5) and (6) we plot disturbance and noise data pairs from Fig. 9 against each other in Fig. 5. We immediately see that the noise disturbance uncertainty relation Eq. (5) is always fulfilled, but not saturable apart from extremal values, that is when either N or D vanishes. In contrast, the improved qubit relation Eq. (6) provides a tight bound and can be saturated if the optimal correction procedure is applied, as in our experiment.

Figure 4: (color online) Disturbance data points and their respective theory curves if, after the projective measurement of M (with θ=50°) the eigenstate $|+m\rangle$ is rotated onto states $|\psi(\vartheta, \phi)\rangle = \cos\frac{\vartheta}{2}|+z\rangle + e^{i\phi}\sin\frac{\vartheta}{2}|-z\rangle$ and $|-m\rangle$ onto $|-\psi\rangle = |\psi(\pi - \vartheta, \phi + \pi)\rangle$.
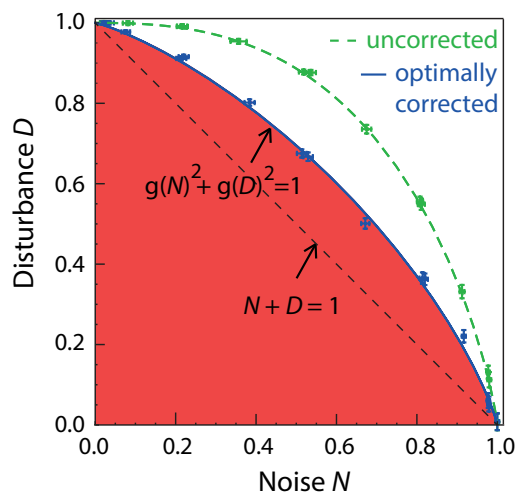

Figure 5: (color online) Disturbance vs. Noise with and without optimal correction procedure. The red shaded area marks the region which are prohibited according to Eq. (6).

Conclusions - We have shown the experimental validity of the information-theoretic formulation of Heisenberg's noise-disturbance uncertainty principle in qubit measurements. As soon as we obtain knowledge about the value of a certain spin observable by applying a suitable measurement the information which can be extracted about another, incompatible observable is reduced. For maximally incompatible spin observables we observe a completely reciprocal trade-off. In order to characterize the irreversible loss of correlations, correction operations are performed which minimize the disturbance, but the sum of noise and disturbance is always bounded from below. We mathematically characterized and experimentally confirmed the optimal correction procedure for qubits, leading to a tight noise-disturbance uncertainty relation. This result should stimulate the search for improved entropic uncertainty relations for observables of higher dimensional Hilbert spaces as well.

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927); English translation in Quantum Theory and Measurement, edited by J. A. Wheeler and W. H. Zurek (Princeton Univ. Press, Princeton, NJ, 1984), p. 62.
[2] E. H. Kennard, Z. Phys. **44**, 326 (1927).
[3] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[4] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
[5] W. Heisenberg, *The Physical Principles of Quantum Mechanics*, (University of Chicago Press, Chicago, IL, 1930).
[6] M. Ozawa, J. Phys. Conf. Ser. **504**, 012024 (2014).
[7] E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[8] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).
[9] M. Ozawa, Phys. Lett. A **318**, 21 (2003).
[10] M. Ozawa, Ann. Phys. (N.Y.) **311**, 350 (2004).
[11] C. Branciard, Proc. Natl. Acad. Sci. USA **110**, 6742 (2013).
[12] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, Nat. Phys. **8**, 185 (2012).
[13] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, Phys. Rev. A **88**, 022110 (2013).
[14] S. Sponar, G. Sulyok, J. Erhart, and Y. Hasegawa, Advances in High Energy Physics **2014**, 735398 (2014).
[15] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Phys. Rev. Lett. **109**, 100404 (2012).
[16] S.-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, Scientific Reports **3**, 2221 (2013).
[17] M. M. Weston, M. J. W. Hall, M. S. Palsson, H. M. Wiseman, and G. J. Pryde, Phys. Rev. Lett. **110**, 220402 (2013).
[18] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, Phys. Rev. Lett. **112**, 020401 (2014).
[19] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, Phys. Rev. Lett. **112**, 020402 (2014).
[20] P. Busch, P. Lahti, and R. F. Werner, Phys. Rev. Lett. **111**, 160405 (2013).
[21] P. Busch, P. Lahti, and R. F. Werner, Rev. Mod. Phys. **86**, 1261 (2014).
[22] X.-M. Lu, S. Yu, K. Fujikawa, and C. H. Oh, Phys. Rev. A **90**, 042113 (2014).
[1] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, Phys. Rev. Lett. **112**, 050401 (2014).
[24] E. B. Davies and J. T. Lewis, Communications in Mathematical Physics **17**, 239 (1970).
[25] E. B. Davies, *Quantum theory of open systems* (Academic Press London, New York).
[26] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
[27] See Supplemental Material at

Supplementary Material

Improved information-theoretic uncertainty relation for qubits

A generally valid uncertainty relation between noise $N(\mathcal{M}, A)$ and disturbance $D(\mathcal{M}, B)$, for any measurement apparatus M, is given by Eq. (4) of the main text—which for our investigated qubit scenario reduces to

$$N(\mathcal{M}, \sigma_z) + D(\mathcal{M}, \sigma_y) \geq \log 2 = 1 \text{ bit.} \tag{14}$$

For our experiment, this inequality is only saturated for extremal values, that is, if either N or D vanish. By applying the optimal correction operation for rank-one projective measurements, the measured values come closer to the straight line $N + D = 1$ in the N-D plane, but do not reach it (see blue data points in Fig. 5 of the main text). Thus, the question arises as to whether the above inequality can be saturated by a different class of measurements, or, conversely whether an improved, saturable inequality exists.

Here we show that the above inequality can in fact be substantially improved, to

$$g[N(\mathcal{M}, \sigma_z)]^2 + g[D(\mathcal{M}, \sigma_y)]^2 \geq 1, \tag{15}$$

where $g[x]$ is the inverse of the function $h(x) := -\frac{1+x}{2} \log_2 \frac{1+x}{2} - \frac{1-x}{2} \log_2 \frac{1-x}{2}$ on the interval $x \in [0, 1]$. This inequality is in fact *optimal*, i.e., it is the tightest possible inequality for the noise and disturbance of σ_z and σ_y, for arbitrary measurement apparatuses. Moreover, this optimal inequality is saturated in our experimental scenario, as depicted in Fig. 5 of the main text.

To prove Eq. (15), let $R := \{(N, D)\}$ denote the region of possible values of $N = N(\mathcal{M}, \sigma_z)$ and $D = D(\mathcal{M}, \sigma_y)$, over all possible measurements $\mathcal{M}$. Hence, R has some lower boundary, C, that in general prevents the noise and disturbance from both being arbitrarily small. Indeed, from Eq. (14) above, all points in C must lie on or above the line $N + D = 1$ bit. Our aim is to show that C is given by Eq. (15).

Now, as shown in the Supplemental Material of [1], the noise and disturbance of two system observables A and B, for any measurement M, satisfy

$$N(\mathcal{M}, A) \geq \sum_u p_u H(A^T | \rho_u) = \sum_u p_u H(A | \rho_u^T), \qquad D(\mathcal{M}, B) \geq \sum_u p_u H(B^T | \rho_u) = \sum_u p_u H(B | \rho_u^T),$$

where $\mathcal{E} = \{\rho_u; p_u\}$ is an ensemble of reduced states describing the system S, following measurement of some POVM $\{\Pi_u\}$ on a reference copy R of the system that is maximally entangled with S. Here $H(C|\rho)$ denotes the entropy of C for state ρ, and the transpose C^T of operator C is defined with respect to the Schmidt basis of the maximally entangled state of R and S [1].

It immediately follows that C, the lower boundary of R, lies on or above the lower boundary C^* of the region

$$R^* := \left\{ \left(\sum_m p_m H(\sigma_z | \rho_m), \sum_m p_m H(\sigma_y | \rho_m) \right) \right\}, \tag{16}$$

where $\mathcal{E} = \{\rho_m; p_m\}$ ranges over all possible ensembles of qubit states (thus including the ensembles $\{\rho_u^T; p_u\}$ in particular). Remarkably, it turns out that $C \equiv C^*$, i.e., C^* specifies the optimal uncertainty relation for $N(\mathcal{M}, \sigma_z)$ and $D(\mathcal{M}, \sigma_y)$.

The explicit form of the curve C^* may be determined by showing that attention can be restricted to the subset of pure-state ensembles, and performing a suitable variational calculation. First, for a given ensemble $\mathcal{E} = \{\rho_m; p_m\}$, let $r^{(m)}$ denote the Bloch vector corresponding to the qubit state ρ_m, i.e., $\rho_m = \frac{1}{2}(1 + \sigma \cdot r^{(m)})$. Now define a corresponding *pure*-state ensemble, $\mathcal{E}' := \{\rho_m'; p_m\}$, by taking $r^{(m)'}$ to be the unit vector in the y-z plane which has (i) the same component as $r^{(m)}$ in the y-direction; (ii) no component in the x-direction; and (iii) a remaining component in the $\pm z$-direction, with the $+$ sign ($-$ sign) chosen if $r^{(m)} \cdot z \geq 0$ (< 0). Thus, $|r^{(m)'}| = 1$, $r^{(m)'} \cdot y = r^{(m)} \cdot y$, and $|r^{(m)'} \cdot z| \geq |r^{(m)} \cdot z|$ (since, by construction, $r^{(m)'}$ has a longer component than $r^{(m)}$ in the z-direction). It follows immediately that $H(\sigma_z | \rho_m') \leq H(\sigma_z | \rho_m)$ and $H(\sigma_y | \rho_m') = H(\sigma_y | \rho_m)$. Hence,

$$\sum_m p_m H(\sigma_z | \rho_m') \leq \sum_m p_m H(\sigma_z | \rho_m), \qquad \sum_m p_m H(\sigma_y | \rho_m') = \sum_m p_m H(\sigma_y | \rho_m).$$

Thus, for any point $(N, D) \in R^*$ generated by some ensemble $\mathcal{E}$, there is a point $(N', D) \in R^*$ generated by a
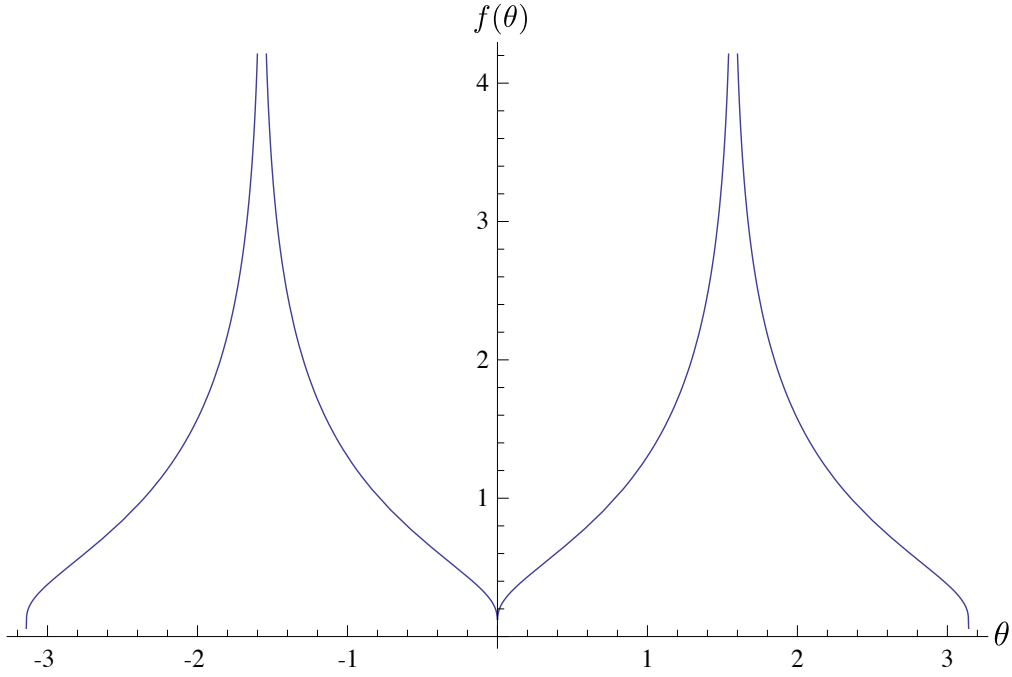
Figure 6: The function $f(\theta)$ in Eq. (19).

corresponding pure-state ensemble $\mathcal{E}'$, with $N' \leq N$.

Hence, to find the point (N, D) on the lower boundary C^* of R^*, for any given value of D, one only has to minimise the variational quantity

$$
\begin{aligned}
J &= \sum_m p_m H(\sigma_z | \rho'_m) + \kappa \left[\sum_m p_m H(\sigma_y | \rho'_m) - D \right] + \lambda \left[\sum_m p_m - 1 \right] \\
&= \sum_m p_m \, h(\sin \theta_m) + \kappa \left[\sum_m p_m \, h(\cos \theta_m) - D \right] + \lambda \left[\sum_m p_m - 1 \right]
\end{aligned}
\tag{17}
$$

over all pure-state ensembles $\mathcal{E}' = \{\rho'_m; p_m\}$ in the yz-plane. Here κ and λ are Lagrange multipliers, the Bloch vector is parameterised as $r^{(m)'} = (0, \cos \theta_m, \sin \theta_m)$, and $h(x)$ is defined as above.

There are two sets of variational equations, for $\{p_m\}$ and $\{\theta_m\}$ respectively, which fix the Lagrange multipliers λ and κ. In particular, $\partial J / p_m = 0$ yields

$$
h(\sin \theta_m) + \kappa \, h(\cos \theta_m) = -\lambda,
\tag{18}
$$

while $\partial J / \theta_m = 0$ yields

$$
p_m \left[\cos \theta_m h'(\sin \theta_m) - \kappa \sin \theta_m h'(\cos \theta_m) \right] = 0.
$$

The latter reduces to, for all $p_m \neq 0$ (i.e, for those p_m which actually contribute to J)

$$
f(\theta_m) := \frac{h'(\sin \theta_m)/ \sin \theta_m}{h'(\cos \theta_m)/ \cos \theta_m} = \kappa.
\tag{19}
$$

It may be checked (see Fig. 6) that the function $f(\theta)$ in Eq. (19) is symmetric about $\theta = 0$ and $\theta = \pi/2$, and monotonic on $[0, \pi/2]$. Hence, one has up to four possible solutions of $f(\theta_m) = \kappa$, for a given value of κ, of the form $\theta_m = \pm \pi/2 \pm \theta$ for some $\theta \in [0, \pi/2]$. Thus,

$$
\sin \theta_m = \pm \cos \theta, \qquad \cos \theta_m = \mp \sin \theta, \qquad \forall m : p_m \neq 0,
$$

for some fixed $\theta \in [0, \pi/2]$, and $\kappa = f(\pi/2 - \theta)$. Moreover, noting $h(x) = h(-x)$, Eq. (18) is automatically satisfied, with $\lambda = -h(\cos \theta) - \kappa \, h(\sin(\theta))$.

The corresponding extremal value of J follows from Eq. (17) as $J_{\min} = h(\cos \theta)$, independently of p_m, yielding the lower boundary of the region R^* to be

$$C^* = \{ \, (h(\cos \theta), h(\sin \theta)) : \theta \in [0, \pi/2] \, \}. \tag{20}$$

Noting that $\cos^2 \theta + \sin^2 \theta = 1$, the equation of this curve corresponds to equality in Eq. (15).

Finally, to show that $C^* \equiv C$, note that all the points $(N, D) \in C^*$ correspond to the values of noise and disturbance for the measurements in our experimental scenario (with optimal correction), as per Eqs. (10) and (11) of the main text. Hence, $C^* \subset R$, implying C lies on or below C^*. But, by construction, C^* lies on or below C, and it immediately follows that $C = C^*$ as desired.

Similar techniques to the above may be used to derived improved noise-disturbance relations for other observables, as will be explored elsewhere. We note here that, by restricting attention to single-member ensembles in Eq. (17), it further follows that C^* is the lower boundary of the region $\{H(\sigma_z|\rho), H(\sigma_y|\rho)\}$, where ρ ranges over all possible qubit states ρ. Hence, C^* also gives the optimal trade-off between the entropies of the mutually complementary qubit observables σ_y and σ_z, improving on the standard Maassen-Uffink relation $H(\sigma_y) + H(\sigma_z) \geq \log 2$.

State preparation and measurement in the neutron polarimetric setup

Here, the experimentally interested reader can find more detailed informations on how the neutron's spin state is manipulated along our polarimeter beam line in order to prepare the desired initial states ($|\pm z\rangle$, $|\pm y\rangle$) and successively measure the observables of interest $M = \sigma_y \sin \theta + \sigma_z \cos \theta$ and $B = \sigma_y$.

State preparation

In order to prepare the eigenstates $|\alpha z\rangle$ of A, that is $|+z\rangle$ and $|-z\rangle$ since $\alpha = \pm 1$, the current through the spin turner coil DC-1 is simply turned off for the former leaving the the spin in the state it possesses after the first supermirror, i.e. $|+z\rangle$, and set to the predetermined flip current generating B_x^{π} for the latter. B_x^{π} is just the field strength that causes a rotation of π of the Bloch vector and thus converts $|+z\rangle$ to $|-z\rangle$. To prepare the eigenstates $|\beta y\rangle$ of B, the respective currents generating the fields $B_x^{\pm \pi/2}$, i.e., that cause $\pm \pi/2$-roations of the Bloch vector, are applied in DC-1. Each of the two eigenstates of σ_z and σ_y is sent with equal probability, experimentally realized by sending each input state for the same, sufficiently long time period.

Measurement of M

The projective measurement of M consists of two steps. At first, we have to project the initially prepared state onto the eigenstates of M, then, in order to complete the measurement, we have to prepare the neutron spin in the eigenstates of M. Since the input eigenstates $|\pm z\rangle$ and $|\pm y\rangle$ and the $|\pm m\rangle$ all lie in the zy-plane of the Bloch sphere, the distance between DC-1 and DC-2 has to be chosen such that the Bloch vector undergoes integer multiples of the full rotation period in their intermediate guide field. Then, spin turner DC-2 rotates the spin component to be measured, which depends on the polar angle θ ($M = \sigma_y \sin \theta + \sigma_z \cos \theta$), towards the z-direction. For the eigenstate belonging to eigenvalue $\mu = +1$ the component along $+\vec{m}$ and for eigenvalue $\mu = -1$ the spin component along $-\vec{m}$ is rotated in the $+z$ direction. The second supermirror (first analyzer) then selects only the $|+z\rangle$ part of the spinor wave function. The projective measurement is completed by the preparation of the measured spin component with spin turner DC-3. In analogous manner to the preparation of the initial state, this is accomplished by properly setting the respective currents in DC-3 required for the fields B_x^{θ} and $B_x^{\pi+\theta}$. Thus when leaving DC-3 the system is in the appropriate eigenstate of M.

Measurement of B

Here we apply the same procedure as for the M measurement, that is to rotate the $\pm y$ component towards the $+z$-direction with DC-4, followed by another supermirror (second analyzer). A further DC coil for preparing the measured spin state can be omitted, since the neutron detection is insensitive to the spin.

From intensities to noise and disturbance

In this section, we want to explain in detail how the probabilities needed for the calculation of noise and disturbance are obtained from the intensities measured in the experiment.

Noise

In order to determine the information-theoretic noise, the eigenstates $|\pm z\rangle$ of A are sent onto the measurement apparatus which then projectively measures M and B resulting in four different output intensities for each input eigenstate. We have schematically depicted the measurement process in Fig. 7. The polarimeter setup is adjusted such that it realizes one of the eight possible "arms" of Fig. 7 after the other. The output intensities get labeled with
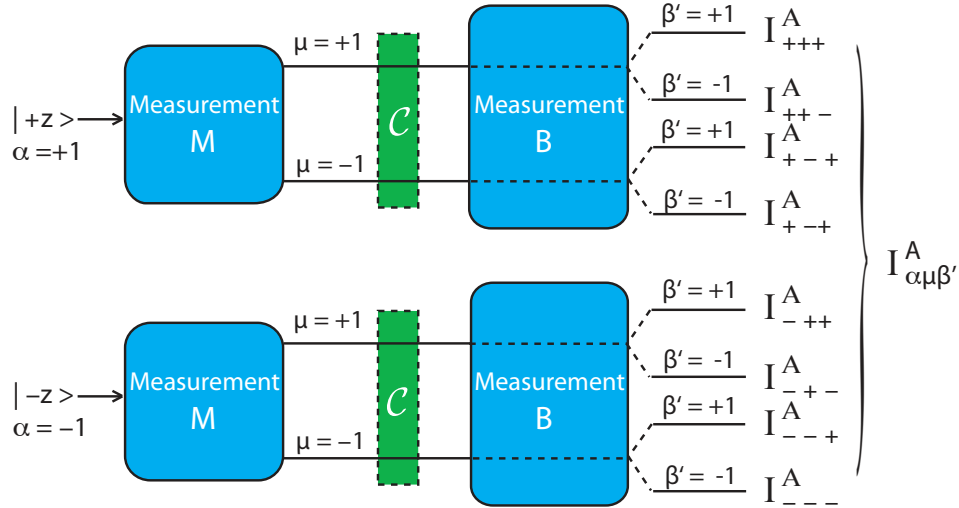

Figure 7: Schematic illustration of the complete noise measurement procedure. All probabilities required for the determination of the information-theoretic noise can be obtained from the output intensities $I^A_{\alpha\mu\beta'}$.

three lower indices $I^A_{\alpha\mu\beta'}$ having the values ± 1 where α gives the sign of $|\pm z\rangle$, μ indicates which projection operator of M has been realized, and β' does the same for the projection operator of B. The probabilities are connected to the intensities via

$$p(\alpha) = \frac{\sum_{\mu,\beta'} I^A_{\alpha\mu\beta'}}{\sum_{\alpha,\mu,\beta'} I^A_{\alpha\mu\beta'}} \qquad p(\mu|\alpha) = \frac{\sum_{\beta'} I^A_{\alpha\mu\beta'}}{\sum_{\mu,\beta'} I^A_{\alpha\mu\beta'}}. \tag{21}$$

However, the definition of the information theoretic noise is not in terms of the conditioned probability $p(\mu|\alpha)$ but $p(\alpha|\mu)$, since it quantifies how well the observable's value can be guessed from the outcome and not contrariwise:

$$N(\mathcal{M}, A) := -\sum_{\alpha,\mu} p(\mu)p(\alpha|\mu) \log p(\alpha|\mu) = H(\mathbb{A}|\mathbb{M}) \tag{22}$$

Here $H(\mathbb{A}|\mathbb{M})$ is the conditional entropy and $\mathbb{A}$ and $\mathbb{M}$ denote the classical random variables associated with input α and output μ. The information-theoretic noise thus quantifies how well the value of A can be inferred from the measurement outcome and only vanishes if an absolutely correct guess is possible. Thus, we have to use Bayes'

theorem to connect these different conditional probabilities

$$p(\alpha|\mu) = \frac{p(\alpha)p(\mu|\alpha)}{p(\mu)}. \tag{23}$$

The marginal probability $p(\mu)$ is given by summation over α of the joint probability distribution $p(\alpha, \mu)$

$$p(\mu) = \sum_\alpha p(\alpha, \mu) = \sum_\alpha p(\alpha)p(\mu|\alpha). \tag{24}$$

Now, the information theoretic noise as defined in Eq. (22) can be calculated.

Disturbance

For the determination of the information-theoretic disturbance, the eigenstates $|\beta y\rangle$, that is $|\pm y\rangle$, of the disturbed observable $B = \sigma_y$ are sent onto the apparatus. By labeling the output intensities with $I^B_{\beta\mu\beta'}$ (see Fig. 8) we get the required probabilities from

$$p(\beta) = \frac{\sum_{\mu,\beta'} I^B_{\beta\mu\beta'}}{\sum_{\beta,\mu,\beta'} I^B_{\beta\mu\beta'}} \qquad p(\beta'|\beta) = \frac{\sum_\mu I^B_{\beta\mu\beta'}}{\sum_{\mu,\beta'} I^B_{\beta\mu\beta'}}. \tag{25}$$

By again using Bayes theorem we obtain the probabilities as they occur in the definition of the information-theoretic disturbance.

$$D(\mathcal{M}, B) := -\sum_{\beta,\beta'} p(\beta')p(\beta|\beta') \log p(\beta|\beta') = H(\mathbb{B}|\mathbb{B}'). \tag{26}$$

as given in the main text, with

$$h(x) := -\frac{1+x}{2} \log \frac{1+x}{2} - \frac{1-x}{2} \log \frac{1-x}{2}. \tag{27}$$

In our scenario, the measurement operator M is varied over the zy-plane spanned by A and B

$$A = \sigma_z, \quad B = \sigma_y, \quad M = \vec{m} \cdot \vec{\sigma} = \sigma_y \sin\theta + \sigma_z \cos\theta, \tag{28}$$
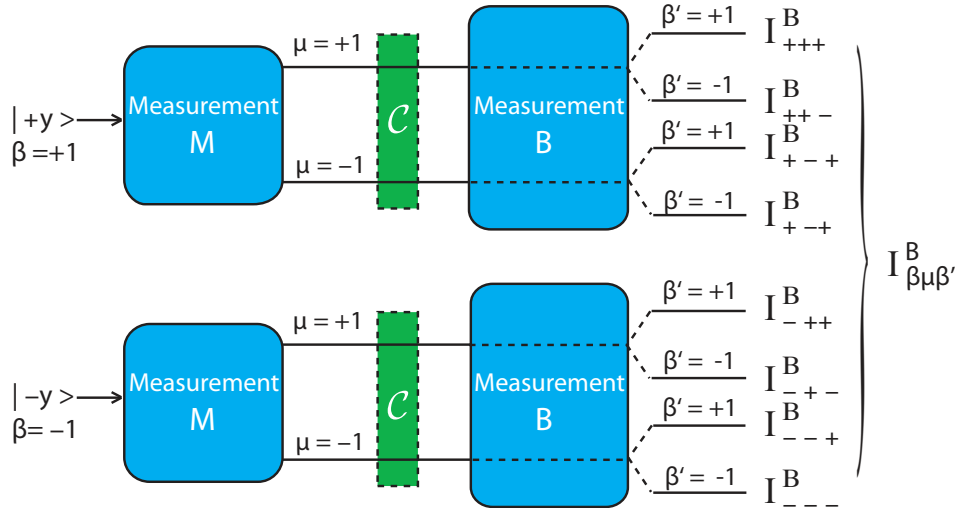

Figure 8: Schematic illustration of the complete disturbance measurement procedure. All probabilities required for the determination of the information-theoretic disturbance can be obtained from the output intensities $I^B_{\beta\mu\beta'}$.

and the theoretically expected expressions for the probabilities are

$$p(\alpha) = p(\beta) = p_{\text{opt}}(\beta) = \frac{1}{2}, \qquad p(\mu|\alpha) = \frac{1 + \mu\alpha\cos\theta}{2},$$

$$p(\beta'|\beta) = p(\beta|\beta') = \frac{1 + \beta'\beta\sin^2\theta}{2}, \quad p_{\text{opt}}(\beta'|\beta) = p_{\text{opt}}(\beta|\beta') = \frac{1 + \beta'\beta|\sin\theta|}{2}. \tag{29}$$

Here p_{opt} is used to denote the probabilities in the case that an optimal correction (see section) is applied after the M-measurement. These expressions yield

$$N(\mathcal{M}, A) = h\left(\cos\theta\right), \quad D_0(\mathcal{M}, B) = h\left(\sin^2\theta\right), \quad D_{\text{opt}}(\mathcal{M}, B) = h\left(\sin\theta\right), \tag{30}$$

These theoretically expected value of noise and disturbance are depicted here in Fig. 9, along with the corresponding experimentally measured values.
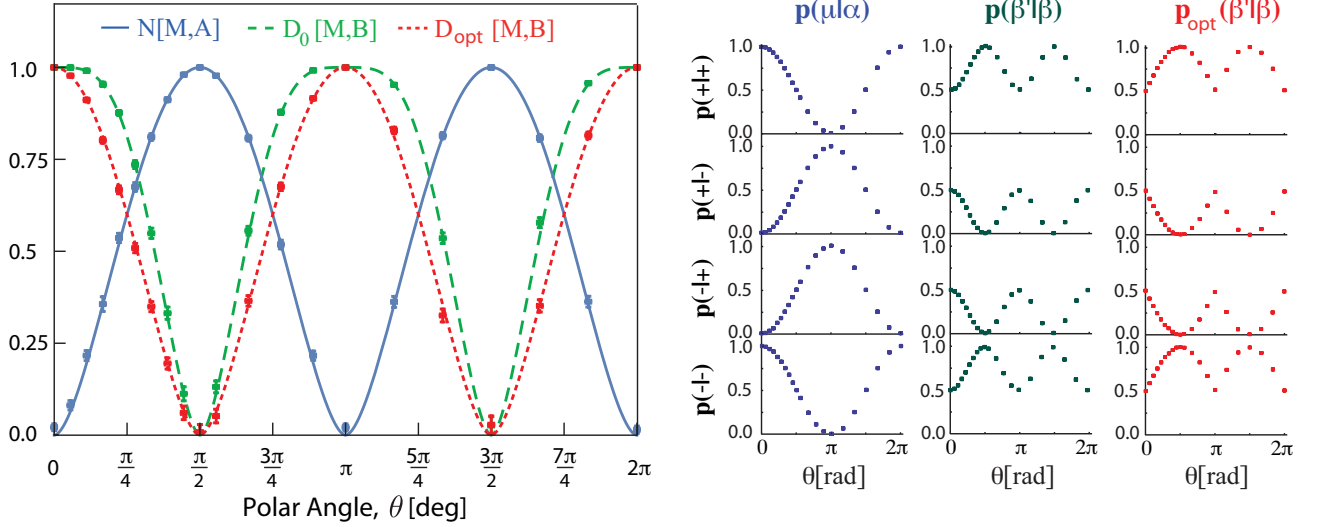

Figure 9: Noise $N(M, A)$ (straight blue line), uncorrected disturbance $D_0(M, B)$ (dashed green line), and optimally corrected disturbance $D_{\text{opt}}(M, B)$ (dotted red line) normalized with the apparatus' efficiency vs. polar angle θ of M. The respective conditional probabilities are plotted on the right hand side.

Optimal correction procedure for projective qubit measurements

For a spin measurement operator $M = \vec{m} \cdot \vec{\sigma}$ and an observable $B = \vec{b} \cdot \vec{\sigma}$ the optimal correction minimizing the disturbance after the projective measurement of M on B is given by

$$\mathcal{C}_{\text{opt}}(|\mu m\rangle) := \begin{cases} |\mu b\rangle, & \vec{b} \cdot \vec{m} \geq 0 \\ |-\mu b\rangle, & \vec{b} \cdot \vec{m} < 0 \end{cases} \tag{31}$$

with $\mu = \pm 1$ and $|\pm m\rangle$ being the respective eigenstates of M. The above formula can be intuitively understood as a sort of *maximum likelihood* correction procedure: the output $|\mu m\rangle$ of the apparatus is rotated onto the closest eigenvector of B, which is $|\mu b\rangle$, if M and B are more correlated than anti-correlated (i.e., $\vec{b} \cdot \vec{m} \geq 0$), or $|-\mu b\rangle$ otherwise (i.e., $\vec{b} \cdot \vec{m} < 0$). This guarantees that the outcome obtained from the final measurement of B is perfectly correlated with the outcome of M, if $\vec{b} \cdot \vec{m} \geq 0$, or perfectly anti-correlated otherwise—in either cases, correlations are kept maximal, by avoiding the occurrence of extra random noise.

The fact that such a simple correction procedure is indeed the optimal one, is a consequence of having an apparatus performing a sharp measurement (i.e., measurement operators are all rank-one). In such a case, in fact, the apparatus is necessarily of the form 'measure-and-prepare,' in the sense that the apparatus has to completely absorb the input quantum system in order to produce the outcome. Therefore, the state of the output quantum system emerging from

the apparatus only depends on the outcome, and all the information about the state of the input quantum system is encoded on the apparatus outcome only.

This in turns implies that, from an information-theoretic viewpoint, in the case of an apparatus performing a sharp measurement (as we have here), the best thing to do, in order to minimize the disturbance in Eq. (3) of the main text, is to have the random output variable $\mathbb{B}'$ perfectly correlated (or perfectly anti-correlated) with the apparatus' random variable $\mathbb{M}$, so that all available information about the input system is encoded in $\mathbb{B}'$. Thus, the minimum possible disturbance is

$$D_{\text{opt}}(\mathcal{M}, B) = H(\mathbb{B}|\mathbb{B}') = H(\mathbb{B}|\mathbb{M}), \tag{32}$$

for such measurements, corresponding to the optimal correction operation in the main text.

———————

[1] F. Buscemi, M. J. W. Hall, M. Ozawa, and M. M. Wilde, Phys. Rev. Lett. **112**, 050401 (2014).