One of the characteristic features of quantum mechanics is that every measurement that extracts information about a general quantum system necessarily causes an unavoidable disturbance to the state of this system. A plethora of different approaches has been developed to characterize and optimize this tradeoff. Here, we apply the framework of quantum instruments to investigate the optimal tradeoff and to derive a class of procedures that is optimal with respect to most meaningful measures. We focus our analysis on binary measurements on qubits as commonly used in communication and computation protocols and demonstrate theoretically and in an experiment that the optimal universal asymmetric quantum cloner, albeit ideal for cloning, is not an optimal procedure for measurements and can be outperformed with high significance.

**Introduction.**—The work of Heisenberg, best visualized by the Heisenberg microscope [1], teaches us that every measurement is accompanied by a fundamental disturbance of a quantum system. The question about the precise relation between the information gained about the quantum system and the resulting disturbance has since inspired numerous studies [2–19]. A central problem is to find a tight, quantitative tradeoff relation, e.g., for the maximally achievable information for a given disturbance or, vice versa, for the minimal disturbance for a certain amount of extracted information. Obviously, this is not only relevant for quantum foundations, but also for many applications in quantum communication [20, 21] and quantum computation [22–24]. Initially studied in the context of which-path information and loss of visibility in interferometers [2, 3], quantifying the information-disturbance tradeoff was based on various measures such as the traditional root mean squared distance [4, 5], the distance of probability distributions [6], operation and estimation fidelities [7–9], reversibility [13–15], stabilized operator norms [16, 17], state discrimination probability [13], probability distribution fidelity [18], and Fisher information [19]. In spite of all these distinct approaches, no clear candidate for a most fundamental framework for the analysis of the information-disturbance tradeoff in quantum mechanics has yet emerged.

Here we build upon a novel, comprehensive information-disturbance relation introduced recently by two of us [25]. There, optimal measurement devices have been proven to be independent of the chosen quality measures, as long as these fulfill some reasonable assumptions, such as convexity and basis-independence. This approach is unique with respect to the employment of reference observables. On one hand, since information eventually is obtained via measurements of observables, we base the quantification of the measurement error on a reference observable. On the other hand, the measurement induced disturbance is defined without relying on any reference observable in order not to restrict the further usage of the post-measurement state. For a finite-dimensional von Neumann measurement, the optimal tradeoff can be achieved with quantum instruments described by at most two parameters.

![FIG. 1. The optimal quantum instruments in terms of measurement error and disturbance clearly outperform the optimal asymmetric cloner (red curve) and the coherent swap operation (green line). Our measurements (blue crosses) come close to the theoretical curve (blue curve). The violet marked instrument is discussed in Fig. 5 in more detail. The error bars are too small to be visible; for a detailed discussion see [26].](image-url)
form all (asymmetric) quantum cloners \[26\]. We test the tradeoff relation experimentally using a tunable Mach-Zehnder-Interferometer and implement a large range of quantum instruments. We apply these instruments to a two-dimensional quantum system encoded in the photon polarization and investigate the relation between the error of the measurement and the disturbance of the qubit state. As distance measures we consider exemplarily some of the measures recommended in \[16\], i.e., the worst-case total variational distance and the worst-case trace norm. For other measures see supplemental material (SM) \[26\]. The experiment clearly shows that the optimal universal asymmetric cloner as well as the coherent swap scheme are suboptimal (Fig. 1).

**Measurements as quantum instruments.**—To generally quantify both the measurement error and the measurement induced disturbance, we describe the measurement of observables on a quantum system by means of quantum instruments \[27, 28\] as illustrated in Fig. 2. Formally, a quantum instrument \(I\) is defined as a set of completely positive linear maps \(I := \{I_j\}_{j=1}^m\) that fulfills the normalization condition \(\sum_{j=1}^m I_j(1) = 1\), where \(I'_j\) denotes the dual map to \(I_j\) with respect to the Hilbert-Schmidt inner product. This description naturally encompasses the connection between the observable given by a positive operator valued measure (POVM) \(E' := \{E'_j\}_{j=1}^m\) and the quantum channel \(T_s\), which describes the measurement induced change of the state.

In general, a quantum channel is a completely positive trace preserving linear map. In the context of quantum instruments, the channel is given by the sum of the linear maps with \(T_s := \sum_{j=1}^m I_j\), where each map corresponds to one measurement operator \(E_j\) of the POVM. The normalization condition of the quantum instrument ensures that the corresponding quantum channel is trace-preserving. Expressing the channel in terms of \(I\) as above reflects the decohering effect of the measurement on the quantum state of the measured system.

The measurement operators \(\{E'_j\}_{j=1}^m\) themselves are fully determined by \(I\) via \(E'_j := I'_j(1)\), where the probability distribution for outcomes \(\{j\}_{j=1}^m\) on state \(\rho\) is given by \(\text{tr} (I_j(\rho)) = \text{tr} (I'_j(1)) = \text{tr} (\rho I'_j)\). From this point of view, the normalization condition of the quantum instrument ensures that the distribution \(\{\text{tr} (E'_j(\rho))\}_{j=1}^m\) is normalized. The instrument description based on the normalized set of maps \(I\), which implies the pair \((E', T_s)\), is sufficient to exhaustively describe all possible quantum measurement processes.

**Distance measures.**—From the notion of quantum instruments it becomes immediately clear that \(E'\) and \(T_s\) are not independent, i.e. the change of the state has a fundamental dependence on the information gained and vice versa. To enable a thorough quantitative analysis of this measurement-disturbance tradeoff, we use distance measures to assess the quality of the approximate measurement and to quantify the disturbance. We quantify the disturbance \(\Delta\) caused to the system by the deviation of the channel \(T_s\) from the identity channel \(T_{id}\) \((\rho) := \rho\).

The measurement error \(\delta\) quantifies the deviation of the measurement \(E'\) from a reference measurement \(E\). This approach utilizes a reference POVM \(E\) to quantify the measurement error, but not the disturbance, in contrast to all other approaches found in the literature, where either a reference system is used for both, measurement error and disturbance, or none is used at all.

The measurement error \(\delta\) can be quantified by defining a worst-case total variational distance based on the \(l_1\)-distance between probability distributions. The \(l_1\)-distance, also called total variational distance, displays the largest possible difference between the probabilities that two probability distributions assign to the same event and therefore is the relevant distance measure for hypothesis testing \[28, 29\]. In our case, these two probability distributions stem from the target measurement \(E\) and the actual measurement \(E'\) for some quantum state. To generalize the measure for the measurement error to take into account all possible quantum states \(\rho\) of the system we additionally take the worst case w.r.t. all states, which is natural when considering the maximal difference, i.e., worst-case characteristic of the \(l_1\)-distance itself. Thus our worst-case total variational distance is defined as

\[
\delta(E') := \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\text{tr} (E'_i(\rho)) - \text{tr} (E_i(\rho))|.
\]

(1)

The quantum analogue of the worst-case total variational distance is the worst-case trace norm distance, which we thus use to quantify the distance between the quantum channel \(T_s\) and the identity channel \(T_{id}\),

\[
\Delta(T_s) := \frac{1}{2} \sup_{\rho} \|T_s(\rho) - \rho\|_1.
\]

(2)

This disturbance measure quantifies how well the quantum channel \(T_s\) can be distinguished from the identity channel \(T_{id}\) in a statistical experiment, if no auxiliary systems are allowed \[30\].

**Optimal instruments and tradeoff.**—As reference measurement, we choose the ideal projective measurement of the qubit with \(E = \{(j),j\}_{j=1}^m\). As proven in \[25\] for the optimal quantum instruments each element \(I_j\) can be expressed by a single Kraus operator, agreeing with the intuition that additional Kraus operators introduce

![FIG. 2. General description of a measurement using a quantum instrument \(I\). Obtaining information about the quantum state via the POVM \(E'\) (dashed line, classical output) induces a change of the quantum state described by the quantum channel \(T_s\) (solid line, quantum output).](image-url)
noise to the system. In the case of a qubit this leads to
\[ T_s(\rho) = \sum_{j=1}^{2} K_j \rho K_j^\dagger \quad \text{and} \quad \{ E_j' = K_j^\dagger K_j \}_{j=1}^{2}. \]  

The Kraus operators of an optimal instrument can be chosen diagonal in the basis \( \{|j\rangle\}^2_{j=1} \) given by the target measurement \[ \text{[24].} \] Since for a qubit there are only two of them and they must satisfy the normalization condition, in general their form is
\[ K_1 = \sqrt{1 - b_1^2} |1\rangle\langle 1| + e^{i\beta_1} b_1 |2\rangle\langle 2|, \quad \text{(4a)} \]
\[ K_2 = b_2 |1\rangle\langle 1| + e^{i\beta_2} \sqrt{1 - b_1^2} |2\rangle\langle 2|, \quad \text{(4b)} \]
with \( 0 \leq b_1^2, b_2^2 \leq 1 \) and two arbitrary phases \( \beta_1 \) and \( \beta_2 \).

As proven in \[ \text{[26],} \] for such an instrument, the worst-case total variational distance \( \delta \) and its trace-norm analogue \( \Delta \), Eqs. \[ \text{[12],} \] quantifying measurement error and disturbance respectively, satisfy
\[ \Delta \geq \left\{ \begin{array}{ll} \frac{1}{2} \left( \sqrt{1 - \delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{array} \right. \]  

The inequality is tight and cannot be exceeded by any quantum measurement procedure. Equality in Eq. \[ \text{(5),} \] is attained for the family of optimal instruments defined by
\[ K_1 = \frac{1}{\sqrt{2}} \left( \sqrt{1 - \gamma} |1\rangle\langle 1| + \sqrt{1 + \gamma} |2\rangle\langle 2| \right), \quad \text{(6a)} \]
\[ K_2 = \frac{1}{\sqrt{2}} \left( \sqrt{1 + \gamma} |1\rangle\langle 1| + \sqrt{1 - \gamma} |2\rangle\langle 2| \right), \quad \text{(6b)} \]
with \( \gamma \in [0, 1] \), leading to \( \delta(\gamma) = (1 - \gamma)/2 \).

**Other known measurement schemes.**—Let us evaluate common quantum measurement procedures in terms of their measurement-disturbance tradeoff. For perfect quantum cloning, there would be no measurement-disturbance tradeoff, as one of the perfect clones could be measured without error with the other clone staying undisturbed. Although perfect cloning is impossible \[ \text{[31],} \] one can derive a protocol that is optimal for approximate quantum cloning. Hence, it is a manifest intuition that the optimal universal asymmetric quantum cloner provides a promising measurement protocol that naturally leads simultaneously to a small disturbance and a small measurement error. It is illustrated in Fig. \[ \text{[5,]} \] The quantum channel \( T_s(\rho) = \text{tr}_\nu (T_{\text{clo}}(\rho)) \), a marginal of the cloning channel \( T_{\text{clo}} \), corresponds to the evolution of the system state, obtained when tracing out the second (primed) clone. The corresponding channel of the second clone, \( T_{\nu'}(\rho) = \text{tr}_s (T_{\text{clo}}(\rho)) \), provides an approximate copy to which the reference POVM \( E \) is applied. Asymmetry within the quality of the clones determines the tradeoff between the measurement error and the disturbance.

The optimal universal asymmetric quantum cloning channel \( T_{\text{clo}} \) for any initial quantum state \( \rho \) reads \[ \text{[32]} \]
\[ T_{\text{clo}} (\rho) = (a_2 \mathbb{1} + a_1 \mathbb{F}) \left( \rho \otimes \frac{\mathbb{1}}{2} \right) (a_2 \mathbb{1} + a_1 \mathbb{F}), \quad \text{(7)} \]
with \( a_1^2 + a_2^2 + a_1 a_2 = 1, \ a_1, a_2 \in \mathbb{R}, \) and the flip (or swap) operator \( \mathbb{F} := \sum_{j=1}^{2} |j\rangle\langle j| \). The parameter \( a_1 \) determines the amplitude of a swap operation between both qubits.

With our measures, the measurement-disturbance tradeoff for the asymmetric quantum cloning channel satisfies
\[ \Delta = \left\{ \begin{array}{ll} \frac{1}{4} \left( \sqrt{2 - 3\delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{array} \right. \]  

with \( \delta(a_2) = a_2^2/2 \) \[ \text{[26].} \]

As the cloning operation cannot be realized by a unitary two-qubit transformation, any real implementation of the protocol is embedded in a larger system. Let us thus consider an obvious analogue to the cloning operation, which can be realized by a unitary two-qubit operation. For the swapping channel \( T_{\text{sw}} \), the system interacts with the auxiliary system via a Heisenberg Hamiltonian as
\[ T_{\text{sw}} (\rho) = e^{itF} (\rho \otimes \tilde{\rho}) e^{-itF} = (a_2 \mathbb{1} + ia_1 \mathbb{F}) (\rho \otimes \tilde{\rho}) (a_2 \mathbb{1} - ia_1 \mathbb{F}), \quad \text{(9)} \]
with \( t \in [0, \pi/2] \) or using a parametrization analogous to the cloning scheme with \( a_1^2 + a_2^2 = 1, \ a_1, a_2 \in \mathbb{R}. \) The extreme cases are no swap \( (t = 0, a_2 = 1) \) and full swap \( (t = \pi/2, a_1 = 1). \)

The \( \delta-\Delta \)-tradeoff for the target measurement \( E = \{|j\rangle\langle j|\}_{j=1}^{2} \) performed on one of the outputs satisfies
\[ \Delta = \frac{1}{2} - \delta, \quad \text{(10)} \]
with \( \delta(t) = (1 - a_1^2)/2, \) for the coherent swap \[ \text{[26],} \] evidently also inferior to our optimal instruments, Eq. \[ \text{(6),} \] with the tradeoff given in Eq. \[ \text{(5).} \]

**Experimental implementation.**—For our experimental evaluation of the measurement-disturbance tradeoff we want to realize a broad range of quantum instruments including the optimal ones. For that purpose we consider

\[
\begin{array}{ccc}
\rho & \rightarrow & T_{\text{clo}} \\
& \rightarrow & T_{\text{sw}} (\rho) \\
& \rightarrow & T_{\nu'} (\rho) \leftrightarrow E
\end{array}
\]
the polarization degree of freedom of photons to encode $\rho$, with $|1\rangle \leftrightarrow |H\rangle$ and $|2\rangle \leftrightarrow |V\rangle$, where $|H\rangle$ ($|V\rangle$) denotes horizontally (vertically) polarized light. The Kraus operators describing the chosen set of instruments are thus given by

$$K_{1,2} = \frac{1}{\sqrt{2}} \left[ \sqrt{1 \pm \gamma} |H\rangle\langle H| + e^{i\beta} \sqrt{1 \mp \gamma} |V\rangle\langle V| \right]$$

with an arbitrary phase $\beta$. The optimal cases Eqs. (9) are achieved for $\beta = 0$.

To experimentally realize a quantum instrument and to enable analysis of the two outputs $T_s$ and $E'$, it is necessary to employ an additional auxiliary quantum system, which is not yet explicitly present in the instrument description of Fig. 2. For the measurement of photon polarization a natural candidate is the path degree of freedom of the photons. Since in our case a two-dimensional auxiliary system is sufficient, we employ a Mach-Zehnder interferometer, which provides the two path states $|A\rangle$ and $|B\rangle$, see Fig. 4. The properties of the instrument are then determined by the initial state of this auxiliary system, $|\phi_0\rangle = \cos \alpha |A\rangle + e^{i\varphi} \sin \alpha |B\rangle$, the measurement performed on it, i.e., the detection in the output path states $|C\rangle$ and $|D\rangle$, as well as by an intermediate interaction between path and polarization. The interaction is given by a unitary evolution $U$, which exchanges information between the systems. We use $U = \sigma_2 \otimes |A\rangle\langle A| + \mathbb{1} \otimes |B\rangle\langle B|$, which introduces a polarization dependent phase shift in arm $|A\rangle$.

For an initial path state $|\phi_0\rangle$ the Kraus operators, which act on the polarization, can then be obtained as

$$K_1 = \text{tr}_{\text{path}} \left[ (\mathbb{1} \otimes |C\rangle\langle C|) U \left( \mathbb{1} \otimes |\phi_0\rangle\langle \phi_0| \right) \right],$$

$$K_2 = \text{tr}_{\text{path}} \left[ (\mathbb{1} \otimes |D\rangle\langle D|) U \left( \mathbb{1} \otimes |\phi_0\rangle\langle \phi_0| \right) \right].$$

Relating these expressions with Eq. (11), the parameters $\gamma$ and $\beta$ are given by the experimental parameters $\alpha$ and $\varphi$ by $\gamma = \sin (2\alpha) \sin \varphi$ and $\beta = \arctan [\tan (2\alpha) \cos \varphi]$. The outcome of the measurement $E'$ is then obtained by determining the total intensity in the output $C$ ($E'_1$) and $D$ ($E'_2$), respectively, the action of the quantum channel $T_s$ by state tomography of the polarization degree of freedom.

Measurements and results.—According to Eqs. (1) and (2), the measures $\delta$ and $\Delta$ use the supremum over different input states $\rho$. We thus prepare for each quantum instrument different linearly polarized states $\rho$, which are analyzed after the interaction. The prepared polarization state $\rho = |\psi\rangle\langle \psi|$ in both arms is given by $|\psi\rangle = \cos \frac{\delta}{2} |H\rangle + \sin \frac{\delta}{2} |V\rangle$, where $|H\rangle$ and $|V\rangle$ as the eigenstates of the Pauli matrix $\sigma_z$ with eigenvalues $+1$ and $-1$, respectively, denote horizontal and vertical po-
polarization. We use 16 different values for \( \theta \), including those where extremal behavior for the disturbance or the measurement error is expected. The set of pure, linearly polarized states is sufficient as the suprema in Eqs. [1] and [2] are attained in our experimental implementation, see SM [26].

An intuitive strategy consists of setting a specific instrument and then varying the polarization state \( \rho \), which however requires to keep the instrument parameters (\( \alpha \) and \( \varphi \)) stable. It turns out to be experimentally more favorable to prepare different polarization states \( \rho \) and then vary the phase \( \varphi \) for fixed \( \alpha \) and \( \rho \). One thus associates measurements which correspond to the same state \( |\varphi_0\rangle \) of the auxiliary system to the same instrument.

The evaluation of the measurement error and the disturbance for one instrument of Fig. [1] is shown in Fig. [5] a) and b), respectively. The supremum over a great circle of the Bloch sphere, described by \( |\psi\rangle \), has been used for the analysis. The measurement error is given by the maximal deviation of the measurement (red crosses) to the best fitting target measurement (blue solid line), see Eq. [1]. While some states as eigenstates of the transformation (theoretically) do not show any disturbance, for the disturbance, the largest trace distance has to be taken into account, see Eq. [2].

The obtained values for measurement error and state disturbance are shown in Fig. [1] for the set of experimentally prepared quantum instruments. Each data point here identifies one quantum instrument, for which the supremum of the prepared quantum states in terms of measurement error and disturbance is determined. The horizontal structure is explained when considering that for a fixed \( \alpha \), various measurements with different \( \varphi \) have been taken, see Eq. [11]. We could show that there exist quantum instruments, also experimentally accessible, which significantly outperform the optimal universal asymmetric cloner (red curve) and the coherent swap operation (green line) in terms of the considered distances.

**Conclusion.**—We applied the novel approach derived in [25] to the setting of binary qubit measurements achieving an optimal measurement-disturbance tradeoff. In this setting a reference measurement is used to quantitatively obtain the measurement error. The disturbance, on the other hand, does not depend on any reference measurement, but solely on comparing the state before and after the measurement. Our protocol is tailored for applications based on a specific measurement without restricting subsequent use of the post-measurement state.

Furthermore, we have demonstrated that the strategies of optimal universal asymmetric quantum cloning and coherent swap do not perform optimally when considering the tradeoff relation between measurement error and disturbance. Those protocols are optimal for their respective purposes such as approximate quantum cloning, but cannot compete with the optimal quantum instruments in the measurement scenario as in general they result in worse measurement-disturbance tradeoff relations. We have shown that the advantage of optimal instruments over other schemes is experimentally accessible and not only a mere theoretical improvement. In future applications our findings allow to identify these procedures which retrieve information at the physically lowest cost in terms of state disturbance.

**Acknowledgments.**—We thank Jonas Goeser for stimulating discussions. This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915 and by the German excellence initiative Nanosystems Initiative Munich. LK and AKH are supported by the PhD program Exploring Quantum Matter of the Elite Network of Bavaria. JD acknowledges support by the International Max-Planck Research Program for Quantum Science and Technology (IMPRS-QST). JDMA is supported by an LMU research fellowship.

---

[1] Werner Heisenberg, *The Physical Principles of the Quantum Theory* (University of Chicago Press, 1930).

[2] Gregg Jaeger, Abner Shimony, and Lev Vaidman, “Two interferometric complementarities,” *Phys. Rev. A* **51**, 54-67 (1995).

[3] Berthold-Georg Englert, “Fringe Visibility and Which-Way Information: An Inequality,” *Phys. Rev. Lett.* **77**, 2154-2157 (1996).

[4] Masanao Ozawa, “Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement,” *Phys. Rev. A* **67**, 042105 (2003).

[5] Cyril Branciard, “Error-tradeoff and error-disturbance relations for incompatible quantum measurements,” *Proceedings of the National Academy of Sciences* **110**, 6742-6747 (2013).

[6] Paul Busch, Pekka Lahti, and Reinhard F. Werner, “Proof of Heisenberg’s Error-Disturbance Relation,” *Phys. Rev. Lett.* **111**, 160405 (2013).

[7] Konrad Banaszek, “Fidelity Balance in Quantum Operations,” *Phys. Rev. Lett.* **86**, 1366-1369 (2001).

[8] Christopher A. Fuchs and Asher Peres, “Quantum-state disturbance versus information gain: Uncertainty relations for quantum information,” *Phys. Rev. A* **53**, 2038-2045 (1996).

[9] Lorenzo Maccone, “Information-disturbance tradeoff in quantum measurements,” *Phys. Rev. A* **73**, 042307 (2006).

[10] Francesco Buscemi, Masahito Hayashi, and Michal Horodecki, “Global Information Balance in Quantum Measurements,” *Phys. Rev. Lett.* **100**, 210504 (2008).

[11] Francesco Buscemi, Michael J. W. Hall, Masanao Ozawa, and Mark M. Wilde, “Noise and Disturbance in Quantum Measurements: An Information-Theoretic Approach,” *Phys. Rev. Lett.* **112**, 050401 (2014).

[12] Jun Zhang, Yang Zhang, and Chang-Shui Yu, “The Measurement-Disturbance Relation and the Disturbance Trade-off Relation in Terms of Relative Entropy,” *International Journal of Theoretical Physics*, **55** (2016).
Giacomo M. D’Ariano, “On the Heisenberg principle, namely on the information-disturbance trade-off in a quantum measurement,” Fortschritte der Physik 51, 318–330 (2005).

Andrew N. Jordan and Alexander N. Korotkov, “Uncollapsing the wavefunction by undoing quantum measurements,” Contemporary Physics 51, 125–147 (2010).

Yong Wook Cheong and Seung-Woo Lee, “Balance Between Information Gain and Reversibility in Weak Measurement,” Phys. Rev. Lett. 109, 150402 (2012).

Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen, “Distance measures to compare real and ideal quantum processes,” Phys. Rev. A 71, 062310 (2005).

Dennis Kretschmann, Dirk Schlingemann, and Reinhard F. Werner, “The Information-Disturbance Trade-off and the Continuity of Stinespring’s Representation,” IEEE Transactions on Information Theory 54, 1708–1717 (2008).

Longfei Fan, Wenchao Ge, Hyunchul Nha, and M. S. Zubairy, “Trade-off between information gain and fidelity under weak measurements,” Phys. Rev. A 92, 022114 (2015).

Tomohiro Shitara, Yui Kuramochi, and Masahito Ueda, “Trade-off relation between information and disturbance in quantum measurement,” Phys. Rev. A 93, 032134 (2016).

Nicolas Gisin, Gréggoire Ribordy, Wolfgang Tittel, and Hugo Zbinden, “Quantum cryptography,” Reviews of Modern Physics 74, 145–195 (2002).

Jian-Wei Pan, Zeng-Bing Chen, Chao-Yang Lu, Harald Weinfurter, Anton Zeilinger, and Marek Žukowski, “Multiphoton entanglement and interferometry,” Reviews of Modern Physics 84, 777–838 (2012).

Artur Ekert and Richard Jozsa, “Quantum computation and Shor’s factoring algorithm,” Reviews of Modern Physics 68, 733–753 (1996).

Viatche Vedral and Martin B. Plenio, “Basics of quantum computation,” Progress in Quantum Electronics 22, 1–39 (1998).

Andrew Steane, “Quantum computing,” Reports on Progress in Physics 61, 117–173 (1998).

Anna-Lena K. Haszgen and Michael M. Wolf, “Universality and Optimality in the Information-Disturbance Tradeoff,” ArXiv e-prints (2018), arXiv:1802.09893 [quant-ph].

Supplemental Material.

E. Brian Davies and John T. Lewis, “An operational approach to quantum probability,” Communications in Mathematical Physics 17, 239–260 (1970).

John Watrous, The Theory of Quantum Information (Cambridge University Press, 2018).

Jerzy Neyman and Egon S. Pearson, “On the Problem of the Most Efficient Tests of Statistical Hypotheses,” Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 231, 289–337 (1933).

Allowing auxiliary systems, the relevant disturbance measure is the diamond norm, \( \Delta_c(T) := \frac{1}{2} \sup \| (\rho - T_{a,d}) \otimes T_{b,d}(\xi) \|_1 \), where the state \( \xi \) includes auxiliary systems. Here, for the optimal tradeoff curve, the trace norm turns out to be equal to the diamond norm distance [25].

William K. Wootters and Wojciech H. Zurek, “A single quantum cannot be cloned,” Nature 299, 802–803 (1982).

Anna-Lena K. Haszgen, “Universal Asymmetric Quantum Cloning Revisited,” Quant. Inf. Comp. 17, 0747–0778 (2017).
SUPPLEMENTAL MATERIAL

SM 1: OPTIMAL TRADEOFF RELATION

Theorem 1 (Total variation - trace norm tradeoff). Consider a von Neumann target measurement given by an orthonormal basis \( \{ |i \rangle \in \mathbb{C}^2 \}^2_{i=1} \), and an instrument with two corresponding outcomes. Then the worst-case total variational distance \( \Delta \) and its trace-norm analogue \( \delta \), defined as in Eqs. (1,2), quantifying measurement error and disturbance respectively, satisfy

\[
\Delta \geq \begin{cases} \frac{1}{2} \left( \sqrt{1 - \delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{S1}
\]

The inequality is tight and equality is attained within the family of instruments defined by

\[
I_j(\rho) := K_j \rho K_j, \quad j = 1, 2, \tag{S2}
\]

with

\[
K_{1,2} = \frac{1}{\sqrt{2}} \left( \sqrt{1 + \gamma} |1\rangle\langle 1| + \sqrt{1 - \gamma} |2\rangle\langle 2| \right) \tag{S3}
\]

with \( \gamma \in [0, 1] \).

Proof. In order to derive the information-disturbance tradeoff, we need to solve the following optimization problem:

For \( \gamma \in [0, 1] \)

\[
\begin{aligned}
\text{minimize} & \quad \Delta \left( T_N = \sum_{j=1}^{2} I_j \right) \\
\text{subject to} & \quad \delta \left( E' = \{ I_j^\gamma(1) \}^2_{j=1} \right) \leq \gamma, \\
& \quad I_j \text{ is c.p. and } \\
& \quad \sum_{j=1}^{2} I_j^\gamma(1) = \mathbb{1},
\end{aligned} \tag{S4}
\]

where the last two constraints ensure that \( I \) is an instrument. As discussed before, we assume that every element of the instrument can be expressed using a single Kraus operator. This agrees well with intuition, because more Kraus operators introduce more noise to the system. Furthermore, we assume that these Kraus operators can be chosen diagonal in the basis of the target measurement, \( E = \{ |j\rangle\langle j| \}^2_{j=1} \), to reflect the symmetry of the optimization problem. These assumptions simplify the optimization problem significantly. The Kraus operators given in Eq. (3) then yield the following POVM elements of the approximate measurement

\[
E_j = (1 - b_j^2) |j\rangle\langle j| + b_j^2 (\mathbb{1} - |j\rangle\langle j|), \tag{S5}
\]

for \( j = 1, 2 \), where \( j = 2 \) if \( j = 1 \) and \( j = 1 \) if \( j = 2 \) with \( 0 \leq b_1^2, b_2^2 \leq 1 \). The measurement error is thus given as

\[
\delta(E') = \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr} \left( E'_j \rho - (j| \rho |j) \right) \right| \\
= \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr} \left( (b_j^2 \mathbb{1} - (b_j^2 + b_j^2)|j\rangle\langle j|) \rho \right) \right| \\
= \sup_{\|\psi\|=1} \frac{1}{2} \sum_{j=1}^{2} \left| \langle \psi | (b_j^2 \mathbb{1} - (b_j^2 + b_j^2)|j\rangle\langle j| |\psi\rangle \right| \\
= \frac{1}{2} (b_1^2 + b_2^2),
\]

where the convexity of the \( l_1 \)-norm was used. The disturbance follows from direct calculations,

\[
\Delta(T_1) = \frac{1}{2} \sup_{\rho} \| T_1(\rho) - \rho \|_1 \\
= \frac{1}{2} \sup_{\rho} \left\| \sum_{j=1}^{2} K_j \rho K_j^\dagger - \rho \right\|_1 \\
= \frac{1}{2} \left| 1 - e^{i\beta_1} b_1 \sqrt{1 - b_1^2} - e^{i\beta_2} b_2 \sqrt{1 - b_2^2} \right|.
\]

Without loss of generality, we may assume that \( b_1, b_2 \geq 0 \) in the optimization problem, such that an optimum is attained for \( \beta_1 = \beta_2 = 0 \). The optimization problem given in Eq. (S4) therefore simplifies:

For \( \gamma \in [0, 1] \)

\[
\begin{aligned}
\text{minimize} & \quad \frac{1}{2} \left( 1 - b_1 \sqrt{1 - b_1^2} - b_2 \sqrt{1 - b_2^2} \right) \tag{S6} \\
\text{subject to} & \quad \frac{1}{2} (b_1^2 + b_2^2) \leq \frac{1}{2} (1 - \gamma), \\
& \quad 0 \leq b_1, b_2 \leq 1.
\end{aligned}
\]

The global minimum is achieved at

\[
b_1 = b_2 = \begin{cases} \sqrt{\frac{1}{\gamma}} & \gamma \in [-1, 0] \\ \frac{\sqrt{1 - \gamma}}{2} & \gamma \in [0, 1] \end{cases}
\]

and as stated in Eq. (S1).

\( \square \)

SM 2: TRADEOFF RELATION FOR OPTIMAL UNIVERSAL ASYMMETRIC CLONING

Theorem 2 (Total variation - trace norm tradeoff using optimal universal asymmetric cloning). Consider a von Neumann measurement given by an orthonormal basis \( \{ |i\rangle \}^2_{i=1} \) on one of the outputs of the optimal universal \( 1 \rightarrow 2 \) asymmetric quantum cloning channel. Then the worst-case total variational distance \( \Delta \) and its trace-norm analogue \( \delta \) satisfy

\[
\Delta = \begin{cases} \frac{1}{2} \left( \sqrt{2 - 3\delta} - \sqrt{\delta} \right)^2 & \text{if } \delta \leq \frac{1}{2}, \\ 0 & \text{if } \delta \geq \frac{1}{2}. \end{cases} \tag{S7}
\]
Proof. The marginals of the optimal cloning channel are given by

\[ T_{\text{clo},i}(\rho) = a_i^2 \frac{\mathbb{I}}{2} \text{tr}(\rho) + (1 - a_i^2)\rho, \quad i = 1, 2, \]

with \( T_{\text{clo},1} = T_s \) and \( T_{\text{clo},2} = T_{s'} \). The marginal quantum channel \( T_s \) describes the evolution of the quantum state and its distance to the identity channel \( T_{\text{id}} \) then quantifies the disturbance. Similarly, the marginal \( T_{s'} \), whose output is measured by the target measurement \( E'_j \), describes the measurement itself through \( E'_j = T_{s'}(E_j) \). This is illustrated in Fig. 3. This yields for the disturbance

\[
\Delta(T_s) := \frac{1}{2} \sup_{\rho} \| \text{tr}(T_s(\rho)) - \rho \|_1 \\
= \frac{1}{2} \sup_{\rho} \left\| a_1^2 \frac{\mathbb{I}}{2} - a_1^2 \rho \right\|_1 \\
= \frac{a_1^2}{2}.
\]

The measurement error turns out to be

\[
\delta(E') := \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr}(E'_j \rho) - \langle j|\rho|j \rangle \right| \\
= \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr}(T_{s'}(\rho) - \langle j|\rho|j \rangle \right| \\
= \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr}((\rho - \rho) T_{s'}(\rho)) - \langle j|\rho|j \rangle \right| \\
= \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \langle j|a_2^2 \frac{\mathbb{I}}{2} - a_2^2 \rho|j \rangle \right| \\
= \frac{a_2^2}{2}.
\]

Substituting this into the trace-preserving condition of the optimal universal asymmetric quantum cloning channel, we obtain the theorem [2].

**SM 3: TRADEOFF RELATION FOR COHERENT SWAP**

**Theorem 3** (Total variation - trace norm tradeoff using the coherent swap). Consider a von Neumann measurement given by an orthonormal basis in \( \mathbb{C}^2 \) on one of the outputs of a coherent swap channel. Then the worst-case total variational distance \( \delta \) and its trace-norm analogue \( \Delta \) satisfy

\[
\Delta = \frac{1}{2} - \delta.
\]

**Proof.** Using the substitution \( a_1 = a \) and \( a_2 = \sqrt{1 - a^2} \) with \( a \in [0, 1] \) yields the two marginals of the coherent swap quantum channel,

\[
T_s(\rho) = a^2 \rho + (1 - a^2)\rho
\]

and

\[
T_s(\rho) = (1 - a^2) \rho + a^2 \rho.
\]

The disturbance is therefore

\[
\Delta(T_s) := \frac{1}{2} \sup_{\rho} \| T_s(\rho) - \rho \|_1 \\
= \frac{1}{2} a^2 \sup_{\rho} \| \rho - \rho \|_1.
\]

The optimal choice for \( \tilde{\rho} \) should clearly satisfy the points \( (\Delta(T_s) = 0, \delta(E') = 1/2) \) and \( (\Delta(T_s) = 1/2, \delta(E') = 0) \), where again \( E' = T_{s'}(E) \). For any such choice of \( \tilde{\rho} \) the disturbance thus satisfies \( \Delta(T_s) \geq a^2/2 \). The measurement error turns out to be

\[
\delta(E') := \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr}(E'_j \rho) - \langle j|\rho|j \rangle \right| \\
= \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr}(T_{s'}(\rho) - \langle j|\rho|j \rangle \right| \\
= \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \text{tr}(\rho - \rho T_{s'}(\rho)) - \langle j|\rho|j \rangle \right| \\
= (1 - a^2) \sup_{\rho} \frac{1}{2} \sum_{j=1}^{2} \left| \langle j|\rho|j \rangle - \langle j|\rho|j \rangle \right|.
\]

Thus, an optimal choice for \( \tilde{\rho} \) that minimizes the disturbance and the measurement error is \( \tilde{\rho} = \frac{1}{2} \). A pure state with the same diagonal entries yields the same measurement error; it would, however, increase the disturbance caused to the system.

The disturbance is then

\[
\Delta(T_s) = \frac{a^2}{2},
\]

and the measurement error is

\[
\delta(E') = \frac{1}{2} (1 - a^2).
\]

This gives the linear tradeoff curve given in theorem [3].

**SM 4: PROPERTIES OF DISTANCE MEASURES**

The distance measures used throughout this manuscript to quantify the measurement error and the disturbance, denoted by \( \delta \) and \( \Delta \), satisfy Assumption 1 and Assumption 2 of [25] respectively.
Lemma 4. \( \delta \) as defined in Eq. 7 satisfies the following properties:

(a) \( \delta(\{|i\rangle\langle i|\}_{i=1}^2) = 0 \),

(b) \( \delta \) is convex,

(c) \( \delta \) is permutation invariant, i.e., for every permutation \( \pi \) and any measurement \( M \)

\[
\delta \left( \{U_\pi^\dagger M_{\pi(i)}U_\pi\}_{i=1}^2 \right) = \delta \left( \{M_i\}_{i=1}^2 \right),
\]

where \( U_\pi \) is the permutation matrix that acts as \( U_\pi \langle i | = | \pi(i) \rangle \), and

(d) \( \delta \) is invariant under diagonal unitaries, i.e., for every diagonal unitary \( D \) and any measurement \( M \)

\[
\delta \left( \{D^\dagger M_i D\}_{i=1}^2 \right) = \delta \left( \{M_i\}_{i=1}^2 \right).
\]

Proof. Let \( \delta(M) := \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | \rho | i \rangle - \langle i | \rho | i \rangle| \).

Then

(a) \( \delta(\{|i\rangle\langle i|\}_{i=1}^2) = 0 \), since

\[
\delta(\{|i\rangle\langle i|\}_{i=1}^2) = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | \rho | i \rangle - \langle i | \rho | i \rangle| = 0,
\]

(b) \( \delta \) is convex, since for any measurements \( M, M' \) and for all \( \lambda \in [0, 1] \),

\[
\delta((\lambda M + (1 - \lambda)M')) = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | (\lambda M_i + (1 - \lambda)M_i') | i \rangle - \langle i | \rho | i \rangle| \leq \lambda \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | M_i | i \rangle - \langle i | \rho | i \rangle| + (1 - \lambda) \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | M_i' | i \rangle - \langle i | \rho | i \rangle| = \lambda \delta(M) + (1 - \lambda)\delta(M'),
\]

(c) \( \delta \) is permutation invariant, since for every permutation \( \pi \) and any measurement \( M \)

\[
\delta \left( \{U_\pi^\dagger M_{\pi(i)}U_\pi\}_{i=1}^2 \right) = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle i | U_\pi^\dagger M_{\pi(i)}U_\pi \rho | i \rangle - \langle i | \rho | i \rangle| = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle \pi(i) | \rho | \pi(i) \rangle - \langle \pi(i) | \rho | \pi(i) \rangle| = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle M_i \rho | i \rangle - \langle i | \rho | i \rangle| = \delta \left( \{M_i\}_{i=1}^2 \right),
\]

where \( U_\pi \) is the permutation matrix that acts as \( U_\pi \langle i | = | \pi(i) \rangle \), and

(d) \( \delta \) is invariant under diagonal unitaries, since for every diagonal unitary \( D \) and any measurement \( M \)

\[
\delta \left( \{D^\dagger M_i D\}_{i=1}^2 \right) = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle D^\dagger M_i D \rho | i \rangle - \langle i | \rho | i \rangle| = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle M_i \rho | i \rangle - \langle i | D^\dagger \rho D | i \rangle| = \sup_{\rho} \frac{1}{2} \sum_{i=1}^2 |\langle M_i \rho | i \rangle - \langle i | \rho | i \rangle| = \delta \left( \{M_i\}_{i=1}^2 \right).
\]

\( \Box \)

Lemma 5. \( \Delta \) as defined in Eq. 8 satisfies the following properties:

(a) \( \Delta(T_{id}) = 0 \),

(b) \( \Delta \) is convex,

(c) \( \Delta \) is basis-independent, i.e., for every unitary \( U \) and every quantum channel \( \Phi \)

\[
\Delta \left( U \Phi (U^\dagger \cdot U) U^\dagger \right) = \Delta(\Phi).
\]

Proof. Let \( \Delta(\Phi) := \frac{1}{2} \sup_{\rho} \|\Phi(\rho) - \rho\|_1 \).

Then

(a) \( \Delta(T_{id}) = 0 \), since \( \Delta(T_{id}) = \frac{1}{2} \sup_{\rho} \|\rho - \rho\|_1 = 0 \),

(b) \( \Delta \) is convex, since for any quantum channels \( \Phi, \Phi' \) and for all \( \lambda \in [0, 1] \),

\[
\Delta((\lambda \Phi + (1 - \lambda)\Phi')) = \frac{1}{2} \sup_{\rho} \|\lambda \Phi + (1 - \lambda)\Phi' - \rho\|_1 \geq \frac{1}{2} \sup_{\rho} \|\lambda(\Phi(\rho) - \rho) + (1 - \lambda)(\Phi'(\rho) - \rho)\|_1 \leq \lambda \Delta(\Phi) + (1 - \lambda)\Delta(\Phi'),
\]

where we have used properties of a norm and properties of a supremum of a convex functional over a convex set,

(c) \( \Delta \) is basis-independent, i.e., for every unitary \( U \) and every quantum channel \( \Phi \)

\[
\Delta \left( U \Phi (U^\dagger \rho U) U^\dagger \right) = \frac{1}{2} \sup_{\rho} \|U \Phi (U^\dagger \rho U) U^\dagger - \rho\|_1 \geq \frac{1}{2} \sup_{\rho} \|U \Phi(\rho) U^\dagger - U \rho U^\dagger\|_1 \geq \frac{1}{2} \sup_{\rho} \|\Phi(\rho) - \rho\|_1 = \Delta(\Phi),
\]
where we have used the fact that the trace norm is unitarily invariant.

SM 5: DIFFERENT MEASURES

The optimal instruments as explained in the main text and derived in Sec. result in optimal measurement-disturbance relations for all distance measures which satisfy the assumptions of [25]. For more details on the distance measure used in the main text see Sec. .

We here show the tradeoff relations for different choices of disturbance measures, while the measurement error is always quantified as in Eq. (1). For various meaningful measures, we observe that the optimal instruments outperform the cloner, see Fig. S1.

SM 6: EXPERIMENTAL SETUP

Due to experimental and practical limitations, the actual experimental setup has been slightly different than described in the main text. However, the actual implementation is fully equivalent to the description there. In order to be able to fully tune the attenuation in one of the interferometer arms, we use a half wave plate (HWP) sandwiched between two polarizers. Therefore, the polarization state $\rho$ cannot be set before. Hence, we decided to first create the spatial superposition state $|\phi_0\rangle$ using waveplates and polarizers and subsequently set $\rho$ in both interferometer arms separately. With this approach, we still achieve at this stage a separable state $\rho \otimes |\phi_0\rangle\langle\phi_0|$ within the interferometer before the interaction. As we set the polarization state directly in front of the second beam splitter of the interferometer, the reflection of beam $A$ on the beam splitter already provides the interaction between system and auxiliary system. This reflection induces the unitary transformation $U$ as described in the main text, enabling us to obtain the Kraus operators given in Eq. (11).

Since for a perfect beam splitter the output ports are interchanged for $\varphi_0 \leftrightarrow \varphi_0 + \pi$, we use only output port $C$ to obtain data for both projections, considering the phases $\varphi_1 = \varphi_0$ and $\varphi_2 = \varphi_0 + \pi$. This way, both projections are carried out with exactly the same equipment, reducing possible experimental errors.

SM 7: CHOICE OF POLARIZATION STATES

According to the parametrization $|\psi\rangle = \cos \frac{\theta}{2} |H\rangle + \sin \frac{\theta}{2} |V\rangle$, the experimentally prepared values for $\theta$ were $\{-20^\circ, -10^\circ, 0^\circ, 10^\circ, 20^\circ, 70^\circ, 80^\circ, 90^\circ, 100^\circ, 110^\circ, 160^\circ, 170^\circ, 180^\circ, 190^\circ, 200^\circ, 270^\circ\}$. For $\theta = 0^\circ$ and $\theta = 180^\circ$, the prepared state corresponds to horizontal polarization $|H\rangle$ and vertical polarization $|V\rangle$, respec-
E\text{xpected to give the largest disturbance }\Delta. \{\text{characterized by}

\begin{align*}
|H\rangle \otimes (\cos \alpha |A\rangle + \sin \alpha e^{i\varphi} |B\rangle) &\rightarrow \\
|H\rangle \otimes (i \cos \alpha |A\rangle + \sin \alpha e^{i\varphi} |B\rangle),
\end{align*}

which does not change the state of the polarization. The disturbance therefore (ideally) vanishes. In contrast, for \(\theta = 90^\circ\), we expect

\begin{align*}
(|H\rangle + |V\rangle) \otimes (\cos \alpha |A\rangle + \sin \alpha e^{i\varphi} |B\rangle) &\rightarrow \\
i (|H\rangle - |V\rangle) \otimes \cos \alpha |A\rangle + (|H\rangle + |V\rangle) \otimes \sin \alpha e^{i\varphi} |B\rangle,
\end{align*}

where normalization is omitted. For a given instrument characterized by \{\alpha, \varphi\}, this polarization state is expected to give the largest disturbance \Delta.

For the Kraus operators given in Eq. (11), we find for \(E_j^\prime = K_j^i K_j\) for \(j = 1, 2\),

\begin{equation}
E_{1,2}^\prime = \frac{1}{2} \begin{pmatrix} 1 \pm \sin 2\alpha \cos \varphi & 0 \\ 0 & 1 \mp \sin 2\alpha \cos \varphi \end{pmatrix}.
\end{equation}

Therefore, the distance of the outcome probabilities, used to obtain \(\delta\), becomes

\begin{equation}
\frac{1}{2} \sum_i \left| \text{tr} \left( E_j^\prime |\psi\rangle\langle\psi| \right) - |\langle i|\psi\rangle|^2 \right| = \\
|\cos \theta (1 - \cos \varphi \sin 2\alpha)|,
\end{equation}

which vanishes for \(\theta = 90^\circ\) (and \(\theta = 270^\circ\)) and can be maximal for \(\theta = 0^\circ\) (and \(\theta = 180^\circ\)).

**SM 8: ERROR ANALYSIS OF EXPERIMENTAL DATA**

The statistical error of the data shown in Fig. 1 is estimated by comparing the results obtained in redundant measurements. The standard deviation of the measurement error is estimated to be around \(8.3 \cdot 10^{-5}\), whereas the 1σ-error bar for the estimated disturbance is approximately \(7.0 \cdot 10^{-5}\). Those values are thus too small to be visible in Fig. 1.

Additionally to statistical errors, two different sources of systematic errors have been identified. First, the state preparation as well as the interaction are not perfectly implemented. The imperfect preparation of the initial polarization state and of the state analysis are the main reasons that the identity channel with no disturbance at all (but high measurement error) cannot be implemented perfectly, leading to a residual disturbance, which appears as an increase of the minimal disturbance \(\Delta\) of the data in the plot. In any case, this type of error only reduces the quality of the prepared quantum instruments and does not lead to faulty conclusions.

However, as a second type of systematic error one has to ensure that the prepared polarization states are describing a great circle on the Bloch sphere and contain the states with extremal results sufficiently well. This error can be approximated by considering the data as shown in Fig. 5. By applying a parabolic model for the data points around the extrema of the probability graphs and the maxima of the trace distance graphs, the deviation of the extrema from the measured points can be estimated. This effect might cause a quantum instrument to look better than it actually is, i.e., less disturbing together with smaller measurement error. Yet, for the dataset shown in Fig. 5(b), the parabolic fit results in a maximum at \(\theta \approx 89.95^\circ\) with a trace distance larger by only 0.02% compared to the trace distance at \(\theta = 90^\circ\). The probabilities in Fig. 5(a) around \(\theta = 0^\circ\) and \(\theta = 180^\circ\) can nicely be described by parabolas, where the extrema coincide with our measured points. Thus, the systematic effect of underestimating the measurement error or the disturbance due to badly chosen measurement states is negligibly small.

In conclusion, the different sources of errors overall reduce the quality of the implemented quantum instruments and do not lead to an underestimation of disturbance and measurement error, respectively. We can thus show the implementation of instruments better than the optimal quantum cloner with high significance.