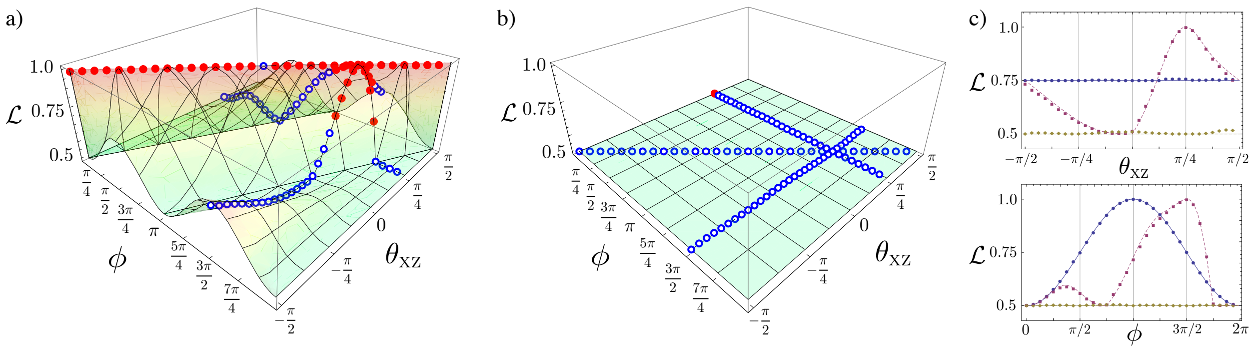
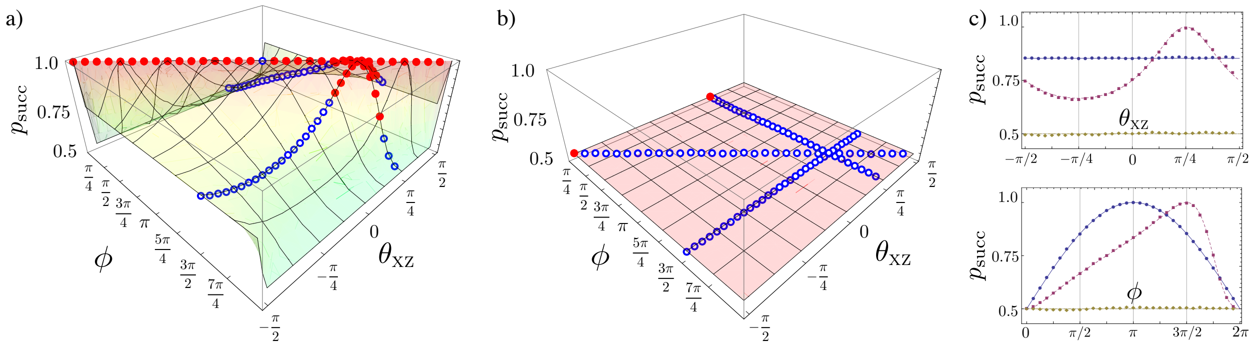


Supplementary Figure 1. **Experimental results for the state discrimination scenario with an optimal CTC implementation.** Probability of state discrimination for **a)** locally prepared and **b)** non-locally prepared states $|\psi_0\rangle{=}|H\rangle$ and $|\psi_1\rangle = \cos(\frac{\phi}{2})|H\rangle + \sin(\frac{\phi}{2})|V\rangle$ as measured by $\mathcal{L}$. The surface represents the theoretically predicted probability for the optimal CTC implementation, depending on the state and gate parameters ϕ and θ_{XZ}, respectively. Solid, red (open, blue) data-points indicate better (worse) performance than standard quantum mechanics, which also implements the optimal measurement, making use of all the available information. **c)** Cross-sectional views of the combined plots a) and b) as in Fig. 5. Error bars obtained from a Monte Carlo routine simulating the Poissonian counting statistics are too small to be visible on the scale of this plot.



Supplementary Figure 2. **Experimental results for the state identification scenario.** Probability of state identification p_{succ} for **a)** locally prepared and **b)** non-locally prepared states $|\psi_0\rangle = |H\rangle$ and $|\psi_1\rangle = \cos(\frac{\phi}{2})|H\rangle + \sin(\frac{\phi}{2})|V\rangle$. The surface represents the theoretically predicted probability depending on the state and gate parameters ϕ and θ_{XZ}, respectively. Solid, red (open, blue) data-points indicate better (worse) performance than standard quantum mechanics, which implements the optimal measurement, making use of all the available information. **c)** Cross-sectional views of the combined plots a) and b) as in Fig. 5. Error bars obtained from a Monte Carlo routine simulating the Poissonian counting statistics are too small to be visible on the scale of this plot.

Supplementary Note 1. Distinguishing (mixed) quantum states

The measure $\mathcal{L}$ introduced in eq. (4) of the main text has an operational interpretation as the probability of obtaining the outcome "different" when comparing two quantum states by means of a single projective measurement on each system. Notably, the minimum-error measurement for discriminating two quantum states is indeed a projective measurement in a direction that depends on the two states [1–3]. Hence, considering only projective measurements is not a restriction and the measure is optimal with the right choice of measurement direction. This result in particular also holds for mixed quantum states, which will become very relevant in the next section.

The situation in the main text can be recast as a game where Alice prepares two quantum systems, one in state $|\psi_0\rangle$ and one in state $|\psi_1\rangle$ and sends them to Bob, whose task is to determine whether they are different or not. If Alice indeed sends two different states, then the measure $\mathcal{L}$ is understood as Bob's probability of either guessing both states correctly or both incorrectly, which are the two cases where he successfully distinguishes the states. Hence, $\mathcal{L}$ is a natural measure for this task and—given that Bob uses the knowledge about the states to be distinguished—also optimal. In fact, given an optimal choice of measurement direction, $\mathcal{L}$ is directly related to the trace-distance metric $\mathcal{D}$ and therefore a similarly suitable measure of distinguishability:

$$\mathcal{L} = p_{\text{correct}}^2 + p_{\text{error}}^2 = 1 - \frac{1}{2}|\langle\psi_0|\psi_1\rangle|^2 = \frac{1}{2}\big(1 + \mathcal{D}^2\big)$$

Supplementary Note 2. Optimal CTC implementation

In the main text we investigated the case where the controlled unitary CU_{xz} is chosen non-optimally for the state $|\psi_1\rangle$. Notably, this is considered a conscious choice of the experimenter, in contrast to decoherence of the gate, which is beyond their control. Hence, the knowledge about θ_{xz} is available and can be used to optimize the measurement direction of the final projective measurement as is done in the case of standard quantum mechanics. Although in the non-optimal CTC-case the output states are mixed, this does not change the fact that the optimal measurement is projective. Hence, the quantity $\mathcal{L}$ can be optimized depending on the states $|\psi_0\rangle, |\psi_1\rangle$ and the gate CU_{xz}. The corresponding results are shown in Supplementary Figure 1, which differs from Fig. 5 in the main text in that the CTC case is now optimal for the chosen gate.

Note that now the CTC-circuit always achieves the best distinguishability of 1/2 for non-local state preparation. In the local case the advantage over standard quantum mechanics is extended to a wider range of non-optimal combinations. Furthermore, we observe recovery of distinguishability for combinations far from optimal, see Supplementary Figure 1.

Supplementary Note 3. State identification

As an alternative approach we consider a scenario where Alice prepares two known quantum states $|\psi_0\rangle$ and $|\psi_1\rangle$ at random and sends them—one at a time—to Bob, who is given the task of identifying each of the states. Similarly to the state-discrimination case, the optimal measurement is a projective measurement in a direction that depends on the two states. The figure of merit that is intrinsically related to this task is the probability of success,

$$p_{\mathrm{succ}} = p_{|\psi_0\rangle} p(\psi_0 \mid |\psi_0\rangle) + p_{|\psi_1\rangle} p(\psi_1 \mid |\psi_1\rangle),$$

where $p_{|\psi\rangle}$ is the probability for the state $|\psi\rangle$ to be sent and $p(\phi \mid |\psi\rangle)$ is Bob's probability for guessing $|\phi\rangle$ in the case where he received the state $|\psi\rangle$. The optimal measurement direction can again be chosen based on knowledge of the two states to be identified. In the scenario considered here, this information is available to Bob and the states are prepared with equal probability. The optimal probability of success is then given by

$$p_{\mathrm{succ}} = \frac{1}{2}\big(1 + \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}\big) = \frac{1}{2}\big(1 + \mathcal{D}\big).$$

Hence, in this case p_{succ} is also directly related to $\mathcal{D}$, making it an equivalent measure of distinguishability. We have analyzed our experiment from this point of view and find the same qualitative behavior: the CTC circuit outperforms standard quantum mechanics for the optimally chosen unitary interaction, as well as for a range of non-optimal choices. The results are shown in Supplementary Figure 2, which parallels Fig. 5 in the main text.

Note that in contrast to Fig. 5 of the main text, the CTC-circuit always achieves, but never surpasses 1/2 in the case of non-locally prepared states.

Supplementary References

[1] C. W. Helstrom, Quantum detection and estimation theory, Journal of Statistical Physics, **1**, 231-252, (1969).

[2] U. Herzog and J. Bergou, Distinguishing mixed quantum states: Minimum-error discrimination versus optimum unambiguous discrimination, Physical Review A, **70**, 022302, (2004).

[3] G. Jaeger and A. Shimony, Optimal distinction between two non-orthogonal quantum states, Physics Letters A, **197**, 83-87, (1995).