Adaptive versus non-adaptive quantum measurements for estimation and discrimination

H. Wiseman*, D. Berry*, S. Bartlett, A. Doherty, B. Booth, B. Higgins and G. Pryde

Centre for Quantum Dynamics, Griffith University, Brisbane

QISC, Macquarie University, Sydney → QIC, Waterloo, Canada

School of Physics, University of Sydney

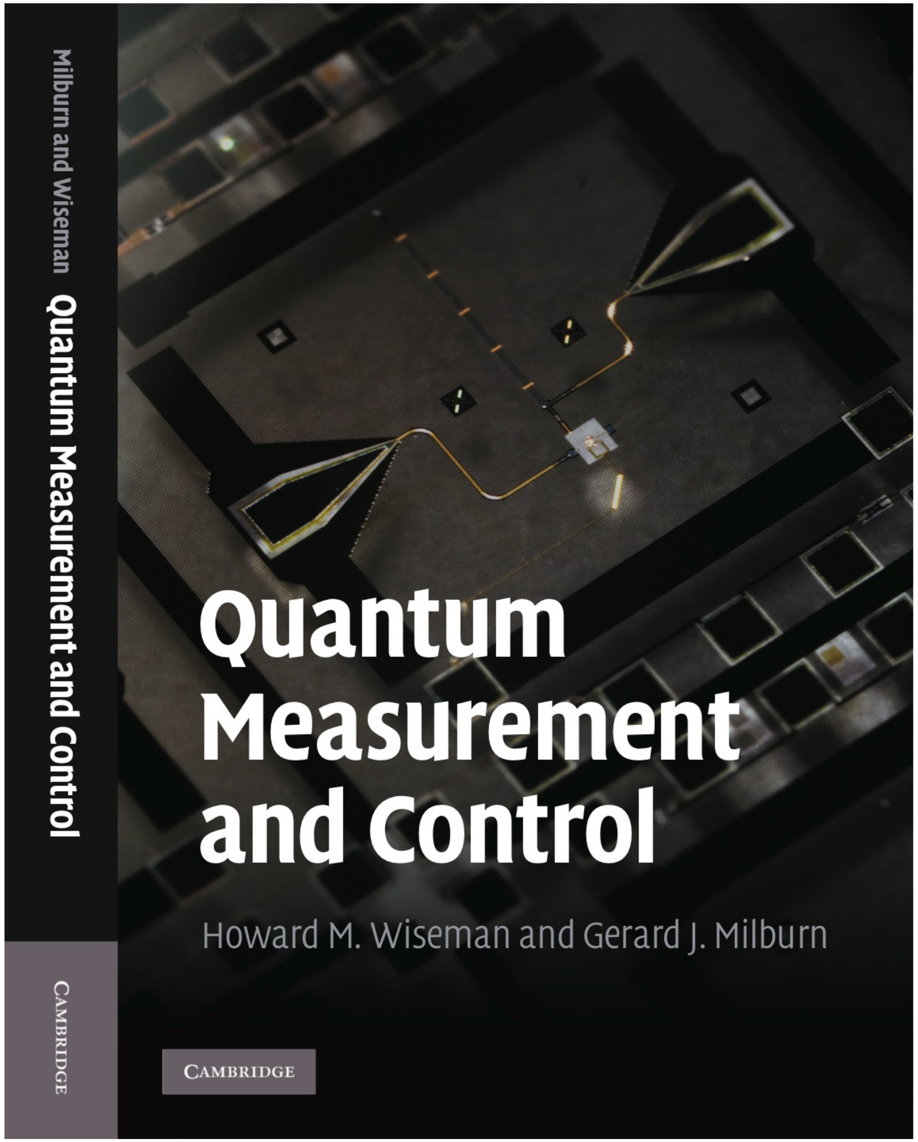
School of Mathematics and Physics, University of Queensland

*Centre for Quantum Computer Technology, Australia

Griffith: Centre for Quantum Dynamics

- Theory: Quantum information, measurement, control and foundations (HMW, David Pegg, Joan Vaccaro).
- Ion trap quantum computer laboratory (Dave Kielpinski)
- Quantum optical information laboratory (Geoff Pryde)
- Laser cooling and trapping of atoms (Robert Sang & DK)
- The Australian Attosecond Science Facility (DK & RS & Igor Litvinyuk)

Quantum Control



- Control is intervening in the world to (try to) optimize something, under given constraints.
- Quantum control is when working out how to do that requires some knowledge of quantum physics.
- e.g. Maximizing the creation of some molecular product, subject to a bound on laser intensity and modulation bandwidth.
- e.g. Minimizing the uncertainty in the estimate of a unitary-gate parameter, subject to a bound on the number of applications of the gate.

Part I — Phase Estimation

- The Rules of the Game
- The Standard Quantum Limit
- The Heisenberg Limit
- The Quantum Phase Estimation Algorithm
- Our 1st algorithm: Generalized QPEA [Nature **450**, 393-6 (2007)]
- Our new algorithm: Non-Adaptive Multi-Pass [arXiv:0809.3308v2]
- Experiment [Nature **450**, 393-6 (2007) and arXiv:0809.3308v2]
- Conclusion

The Rules of the Game

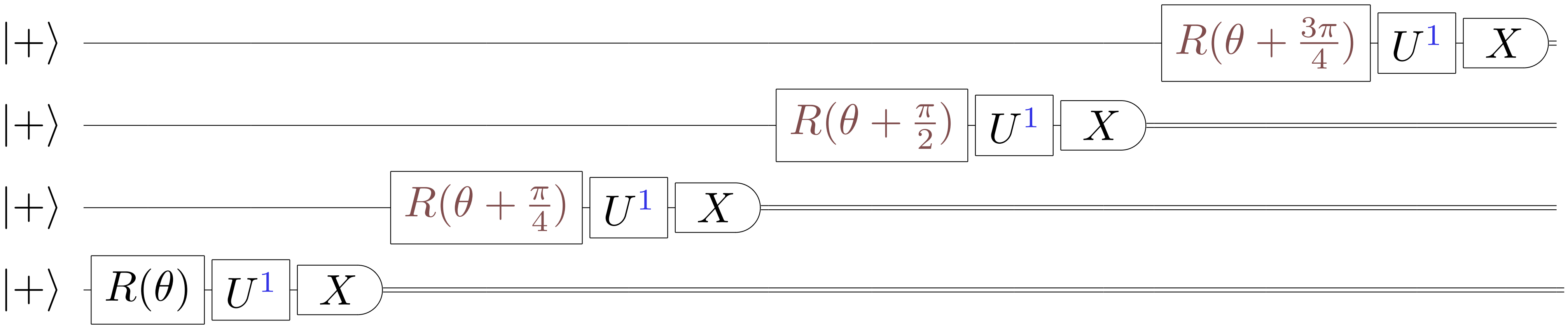
1. We have a gate that performs the unitary operation $U = \exp(i\phi |1\rangle\langle 1|)$ on a specific sort of qubit, and an auxilliary gate $R(\theta) \equiv \exp(i\theta |0\rangle\langle 0|)$.
   e.g. (as in our experiment) the qubit could be a photon-polarization qubit, and an equivalent gate implemented by passing the photon through a HWP at angle $\phi/4$.
2. We have an indefinite supply of these qubits.
3. The parameter ϕ is initially **completely unknown**.
4. We are allowed *at most* N **applications of the gate** U.
5. We aim to minimize the **variance** in our best estimate $\phi_{\rm est}$ of ϕ.
   Technically, we use a cyclic variance measure, $V_{\rm Holevo} = \langle \exp[i(\phi - \phi_{\rm est})] \rangle^{-2} - 1$.

We do not impose temporal or "spatial" (number of qubits) constraints.

The Standard Quantum Limit

N qubits, independently prepared in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, independently measured in the X basis ($|\pm\rangle$), and with $\exp(i\phi|1\rangle\langle 1|)$ applied once on each. $\phi_{\rm est}$ is inferred from the results of the measurement.

For even sampling, $\theta_{\rm init}$ is random, and θ is incremented by π/N between one qubit and the next. Here $N = 4$:



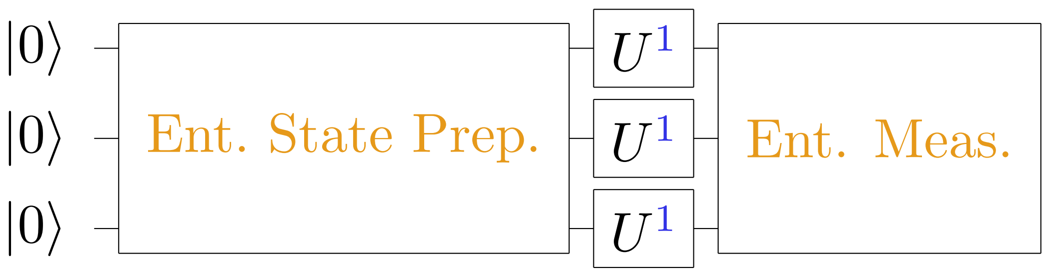
$$\mathrm{SQL} = V[\phi_{\rm est}] \sim 1/N \text{ for } N \gg 1.$$

The Heisenberg Limit (i)

Theoretically, the ultimate limit allowed by QM [1] is much better:

$$\mathrm{HL} = V[\phi_{\mathrm{est}}] \sim \pi^2/N^2 \text{ for } N \gg 1.$$

This requires creating the optimal entangled state [Berry & HMW, PRL (2000)] and a measurement in the phase basis. Here $N = 3$:



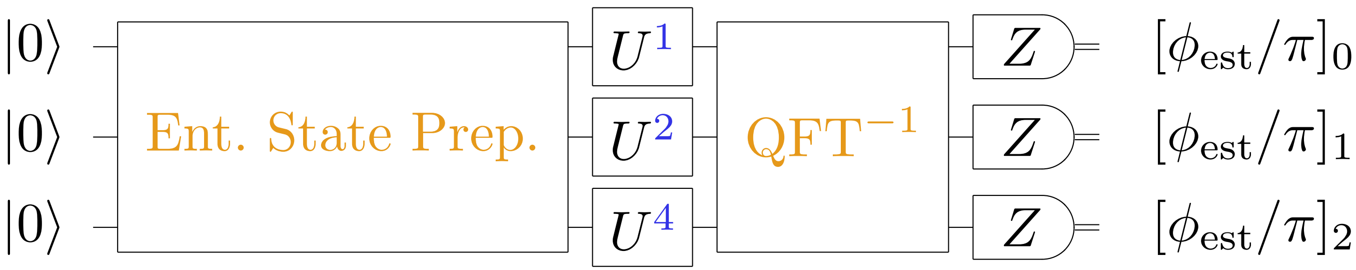
This requires "spatial" resources $O(N)$ but only constant time.

[1] This is called the Heisenberg Limit because the scaling can be derived from the H.U.P. $V[\phi]V[\hat{n}] \geq 1/2$, where $0 \leq \hat{n} \leq N$ is the operator such that the full unitary $U_{\mathrm{total}} = \exp(i\phi\hat{n})$.

The Heisenberg Limit (ii)

Alternatively, we can use **binary encoding** where U acts on the kth qubit ($k = 0, 1, \cdots, K$) $P = 2^k$ times, which we represent by U^P.

Here $N = 2^{K+1} - 1 = 4 + 2 + 1 = 7$:

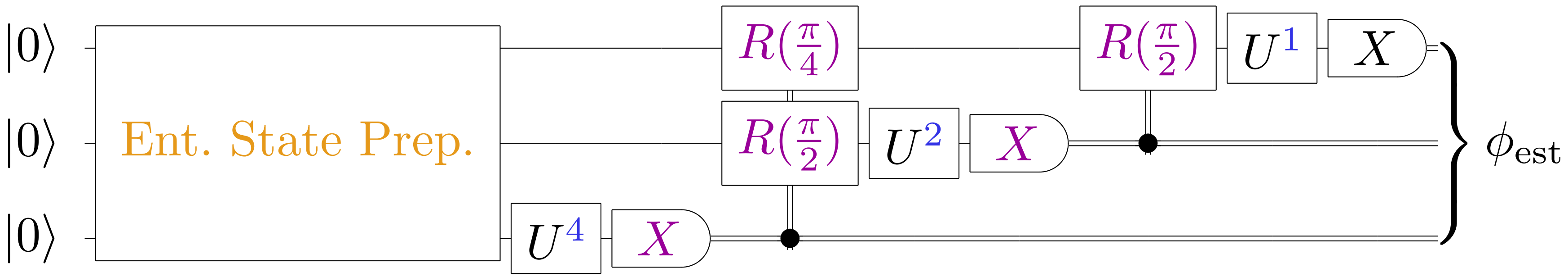


The QFT^{-1} [Shor, 1994] takes the phase basis to the number (logical) basis so that $\phi_{\rm est}$ is read-out from Z measurements ($r = [r]_0.[r]_1[r]_2\ldots$).

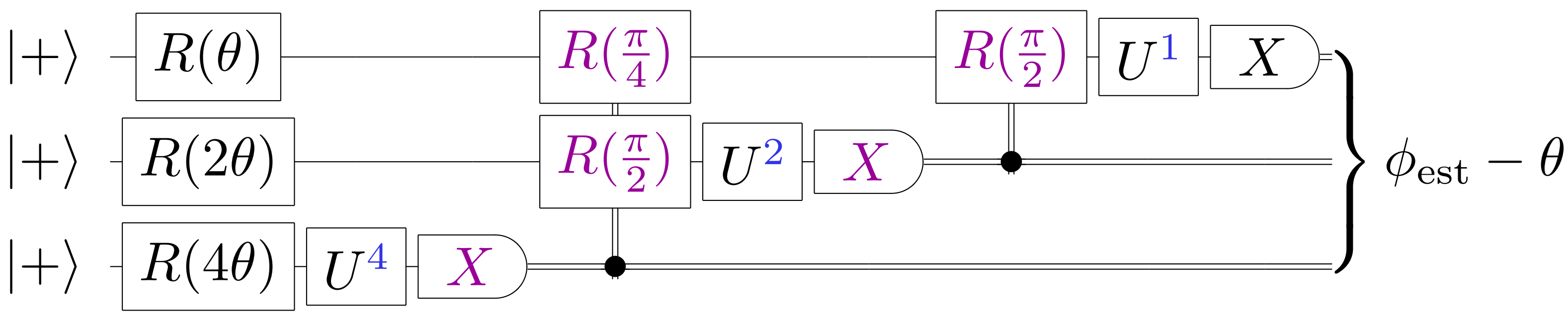
This uses only $O(\log N)$ spatial resources, but a time $O(N)$.

The Quantum Phase Estimation Algorithm (i)

As shown by Griffiths and Niu (PRL, 1996), the QFT^{-1} can be achieved by local (single-qubit) measurement and feedback:



Entangling operation on many qubits is hard. So we can try replacing the entangled state by independent qubits as in the SQL, yielding the QPEA:



The Quantum Phase Estimation Algorithm (ii)

Since the QPEA gives $K + 1$ bits of $\phi_{\rm est}/\pi$, and $N \sim 2^{K+1}$ we would expect

$$\text{QPEA } V[\phi_{\rm est}] \propto (\pi/2^{K+1})^2 \propto \pi^2/N^2 = \text{HL}.$$

But an exact calculation gives

$$\text{QPEA } V[\phi_{\rm est}] \sim 2/N \propto \text{SQL}.$$

What went wrong?

Outliers. The distribution $P(\phi_{\rm est})$ is sharply peaked around at ϕ, with

$$\text{QPEA (HWHM)}^2 \simeq 2.81^2/N^2 \propto HL.$$

But it has high wings, giving SQL scaling for the variance.

The Quantum Phase Estimation Algorithm (ii)

Since the QPEA gives $K+1$ bits of $\phi_{\rm est}/\pi$, and $N \sim 2^{K+1}$ we would expect

$$\text{QPEA } V[\phi_{\rm est}] \propto (\pi/2^{K+1})^2 \propto \pi^2/N^2 = \text{HL}.$$

But an exact calculation gives

$$\text{QPEA } V[\phi_{\rm est}] \sim 2/N \propto \text{SQL}.$$

What went wrong?

Outliers. The distribution $P(\phi_{\rm est})$ is sharply peaked around at ϕ, with

$$\text{QPEA (HWHM)}^2 \simeq 2.81^2/N^2 \propto HL.$$

But it has high wings, giving SQL scaling for the variance.

The Quantum Phase Estimation Algorithm (ii)

Since the QPEA gives $K+1$ bits of $\phi_{\rm est}/\pi$, and $N \sim 2^{K+1}$ we would expect

$$\text{QPEA } V[\phi_{\rm est}] \propto (\pi/2^{K+1})^2 \propto \pi^2/N^2 = \text{HL}.$$

But an exact calculation gives

$$\text{QPEA } V[\phi_{\rm est}] \sim 2/N \propto \text{SQL}.$$

What went wrong?

Outliers. The distribution $P(\phi_{\rm est})$ is sharply peaked around at ϕ, with

$$\text{QPEA } (\text{HWHM})^2 \simeq 2.81^2/N^2 \text{ as in the HL}.$$

But it has high wings, giving SQL scaling for the variance.

Our 1st algorithm: Generalized QPEA

QPEA: the kth qubit $(k = 0, 1, \cdots K)$ passes the phase gate 2^k times.

We generalize this by having, for each k, M independent qubits which pass the gate 2^k times, so that the total number of passes through the phase gate is

$$N = M \times (2^{K+1} - 1).$$

We use the algorithm of Berry and HMW (PRL 2000) to make the *locally optimal* adaptive measurement.

- For $M = 1$, this **exactly** reproduces the optimal QFT^{-1} of the QPEA.
- Numerically we find [Nature **450**, 393-6 (2007)] $M = 5$ is best:

$$M = 5 \text{ GQPEA } V[\phi_{\text{est}}] \simeq (4.8/N)^2 \propto (\pi/N)^2 = \text{HL}.$$

Our new algorithm: Non-Adaptive Multi-Pass

Previous work [Giovannetti, Lloyd, and Maccone, PRL '06] has claimed one can more simply attain the Heisenberg Limit by using **non-adaptive measurements** and "large" M.

Actually this is **impossible** *even if* M *is chosen depending on* K.

Can we get to the HL with **no feedback** with a more general algorithm, with a function $M(K,k)$ that assigns more qubits to smaller k-values (which use exponentially fewer resources)?

Yes, for some functions of the form $M(K,k) = M_K + \mu(K-k)$.

Numerically we find the best results are for $M_K = 2$ and $\mu = 3$ [arXiv:0809.3308v2]

$$\text{NAMP } V[\phi_{\text{est}}] \simeq (6.4/N)^2 \propto (\pi/N)^2 = \text{HL}.$$

Our new algorithm: Non-Adaptive Multi-Pass

Previous work [Giovannetti, Lloyd, and Maccone, PRL '06] has claimed one can more simply attain the Heisenberg Limit by using **non-adaptive measurements** and "large" M.

Actually this is **impossible** *even if* M *is chosen depending on* K.

Can we get to the HL with **no feedback** with a more general algorithm, with a function $M(K, k)$ that assigns more qubits to smaller k-values (which use exponentially fewer resources)?

Yes, for some functions of the form $M(K, k) = M_K + \mu(K - k)$.

Numerically we find the best results are for $M_K = 2$ and $\mu = 3$ [arXiv:0809.3308v2]

$$\text{NAMP } V[\phi_{\text{est}}] \simeq (6.4/N)^2 \propto (\pi/N)^2 = \text{HL}.$$

Our new algorithm: Non-Adaptive Multi-Pass

Previous work [Giovannetti, Lloyd, and Maccone, PRL '06] has claimed one can more simply attain the Heisenberg Limit by using **non-adaptive measurements** and "large" M.

Actually this is **impossible** *even if* M *is chosen depending on* K.

Can we get to the HL with **no feedback** with a more general algorithm, with a function $M(K,k)$ that assigns more qubits to smaller k-values (which use exponentially fewer resources)?

Yes, for some functions of the form $M(K,k) = M_K + \mu(K-k)$.

Numerically we find the best results are for $M_K = 2$ and $\mu = 3$ [arXiv:0809.3308v2]

$$\text{NAMP } V[\phi_{\text{est}}] \simeq (6.4/N)^2 \propto (\pi/N)^2 = \text{HL}.$$

Our new algorithm: Non-Adaptive Multi-Pass

Previous work [Giovannetti, Lloyd, and Maccone, PRL '06] has claimed one can more simply attain the Heisenberg Limit by using **non-adaptive measurements** and "large" M.

Actually this is **impossible** *even if* M *is chosen depending on* K.

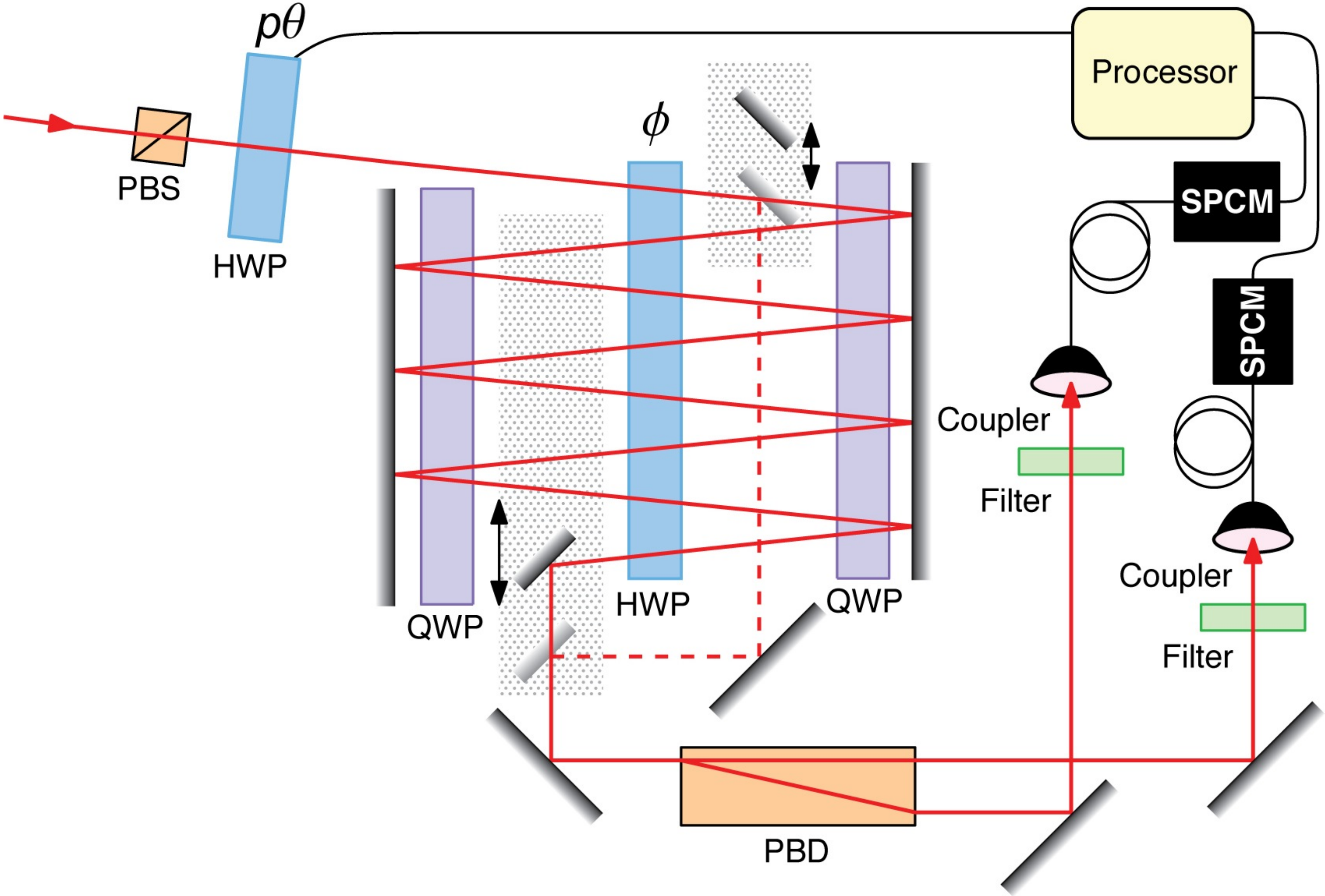
Can we get to the HL with **no feedback** with a more general algorithm, with a function $M(K,k)$ that assigns more qubits to smaller k-values (which use exponentially fewer resources)?

Yes, for some functions of the form $M(K,k) = M_K + \mu(K-k)$.

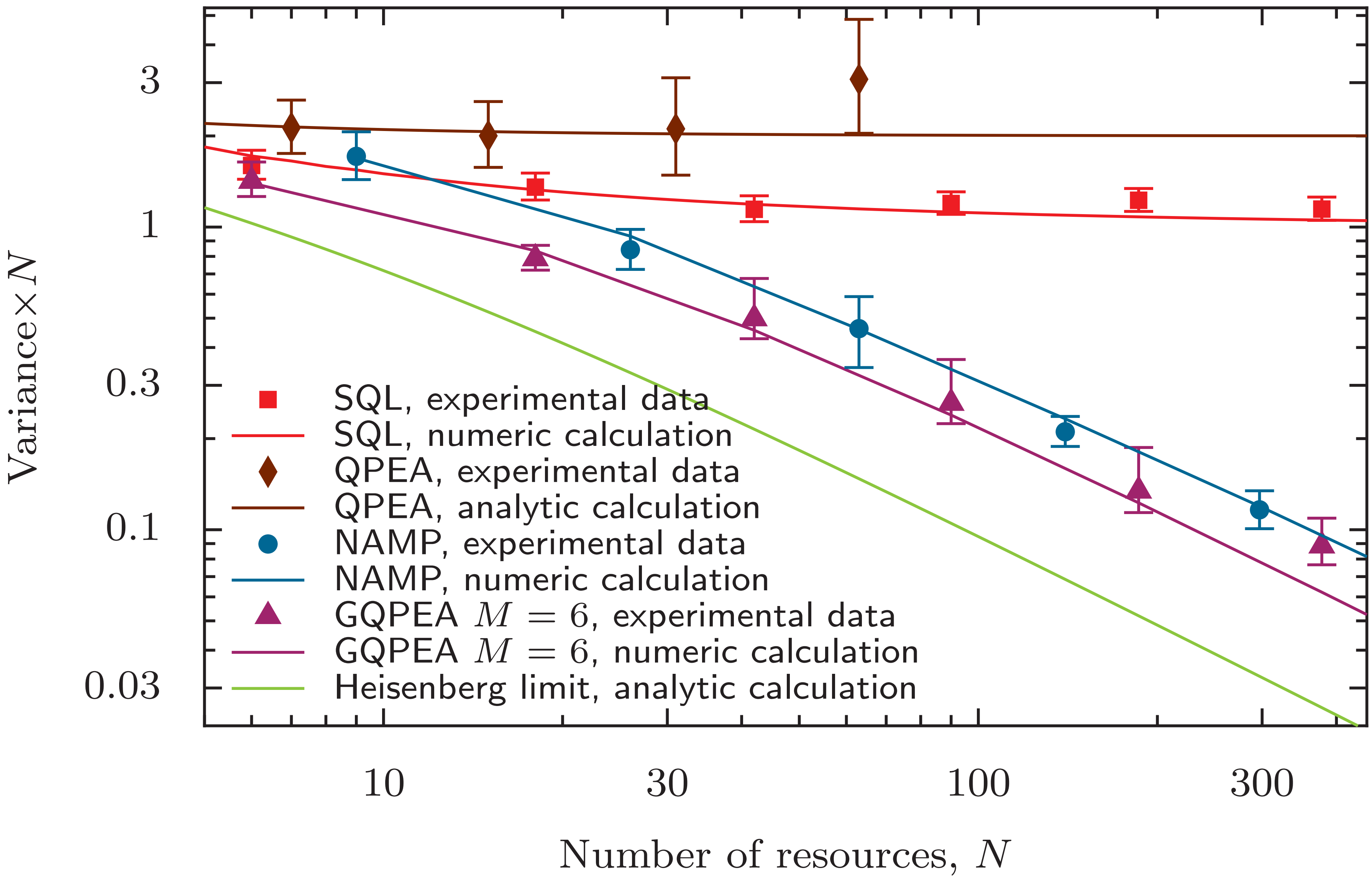
Numerically we find the best results are for $M_K = 2$ and $\mu = 3$ [arXiv:0809.3308v2]

$$\text{NAMP } V[\phi_{\text{est}}] \simeq (6.4/N)^2 \propto (\pi/N)^2 = \text{HL}.$$

The Experimental Apparatus



Experiment [Nature (2007), arXiv:0809.3308v2]



Conclusions (Part I)

- The absolute quantum limit to estimating the phase ϕ of a qubit gate $\exp(i\,|1\rangle\,\langle 1|\,\phi)$, with N gate applications, is

$$V[\phi_{\rm est}] \sim (\pi/N)^2 = {\rm HL},$$

To attain this **exactly**, while not using *exponential "space"*, requires

  - preparing an entangled state of $O(\log N)$ qubits.
  - multiple passes through the gate of any given qubit.
  - control of individual qubits based on prior results.

- We have shown **analytically**, **numerically**, and **experimentally** that HL-**scaling** can be attained with **only**

  - multiple passes through the gate of any given qubit.

- Future directions: not using *exponential time*.

Conclusions (Part I)

- The absolute quantum limit to estimating the phase ϕ of a qubit gate $\exp(i\,|1\rangle\langle 1|\,\phi)$, with N gate applications, is

$$V[\phi_{\rm est}] \sim (\pi/N)^2 = \mathrm{HL},$$

To attain this **exactly**, while not using *exponential "space"*, requires

- preparing an entangled state of $O(\log N)$ qubits.
- multiple passes through the gate of any given qubit.
- control of individual qubits based on prior results.

- We have shown **analytically**, **numerically**, and **experimentally** that HL-**scaling** can be attained with **only**
  - multiple passes through the gate of any given qubit.
- Future directions: not using *exponential time*.

Conclusions (Part I)

- The absolute quantum limit to estimating the phase ϕ of a qubit gate $\exp(i\,|1\rangle\langle 1|\,\phi)$, with N gate applications, is

$$V[\phi_{\rm est}] \sim (\pi/N)^2 = \mathrm{HL},$$

To attain this **exactly**, while not using *exponential "space"*, requires

  - preparing an entangled state of $O(\log N)$ qubits.
  - multiple passes through the gate of any given qubit.
  - control of individual qubits based on prior results.

- We have shown **analytically**, **numerically**, and **experimentally** that HL-**scaling** can be attained with **only**

  - multiple passes through the gate of any given qubit.

- Future directions: not using *exponential time*.

Conclusions (Part I)

- The absolute quantum limit to estimating the phase ϕ of a qubit gate $\exp(i\,|1\rangle\langle 1|\,\phi)$, with N gate applications, is

$$V[\phi_{\rm est}] \sim (\pi/N)^2 = \mathrm{HL},$$

To attain this **exactly**, while not using *exponential "space"*, requires

  - preparing an entangled state of $O(\log N)$ qubits.
  - multiple passes through the gate of any given qubit.
  - control of individual qubits based on prior results.

- We have shown **analytically**, **numerically**, and **experimentally** that HL-**scaling** can be attained with **only**

  - multiple passes through the gate of any given qubit.

- Future directions: entangled states to avoid *exponential time*.

Part II — State Discrimination

- The Rules of the Game (and a Primer)
- Potential Strategies, including SQL and Helstrom Limit
- Pure State Case: Theory (Acín *et al.*) and Experiment (us)
- Mixed State Case: Theory and Experiment (us)
- Conclusion

The Rules of the Game (and a Primer)

1. We are given N qubits either in state $\rho_+^{\otimes N}$ or in state $\rho_-^{\otimes N}$, where

$$\rho_\pm = \tfrac{1}{2}(I + r\cos\theta\ \hat{\sigma}_x \pm r\sin\theta\ \hat{\sigma}_z)\,,$$

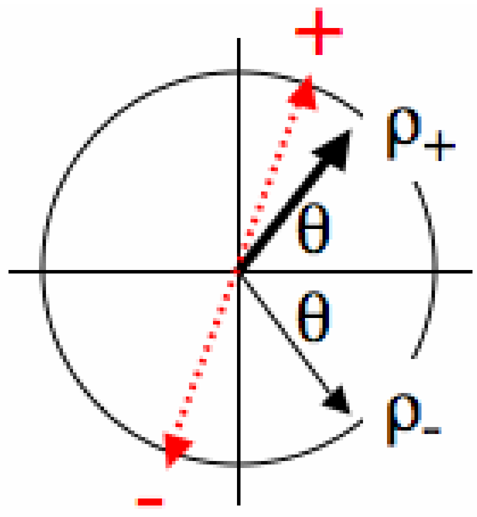
   with prior probabilities $\wp_0$ and $1-\wp_0$ (we always assume $\wp_0 = 0.5$).

2. We have to decide which state it is, and the cost function (to be minimized) is the *probability of error* $C(N)$.

For the case $N = 1$, the optimal strategy is to make the Helstrom measurement (1976) by measuring

$$\hat{H}(1,\wp_0) \equiv \wp_0\rho_+ - (1-\wp_0)\rho_-$$

and depending on whether the outcome is positive or negative, declare $+$ or $-$.



Potential Strategies

1. Majority Vote (SQL): Measure $\hat{H}(1,\frac{1}{2})$ on each qubit and declare $\pm$ depending on which outcome occurs more often.

2. Globally Optimal Meas$^{\underline{t}}$ (Helstrom L.): Measure $\hat{H}(N,\frac{1}{2}) \propto \rho_+^{\otimes N} - \rho_-^{\otimes N}$ and declare on the basis of the sign of the outcome.

3. Globally Optimal Local Meas$^{\underline{t}}$: Use *Dynamic Programming* to determine the optimal observable $\hat{O}_n(N)$ for the nth qubit, based on prior results.

4. Locally Optimal Local Meas$^{\underline{t}}$: Measure $\hat{H}(1,\frac{1}{2})$ on the first qubit, update prior to $\wp_1$ using Bayes' theorem, then measure $\hat{H}(1,\wp_1)$ on the second qubit, update prior to $\wp_2$ and so on ....

5. Fully Biased Meas$^{\underline{t}}$: Measure $\hat{H}(1,1)$ $[\hat{H}(1,0)]$ on every qubit, and update the prior using Bayes' theorem. For the pure state case $(r=1)$ this means a "$+$" ["$-$"] is declared if and only if the 'vote' is unanimous.

Pure State Case (Theory)

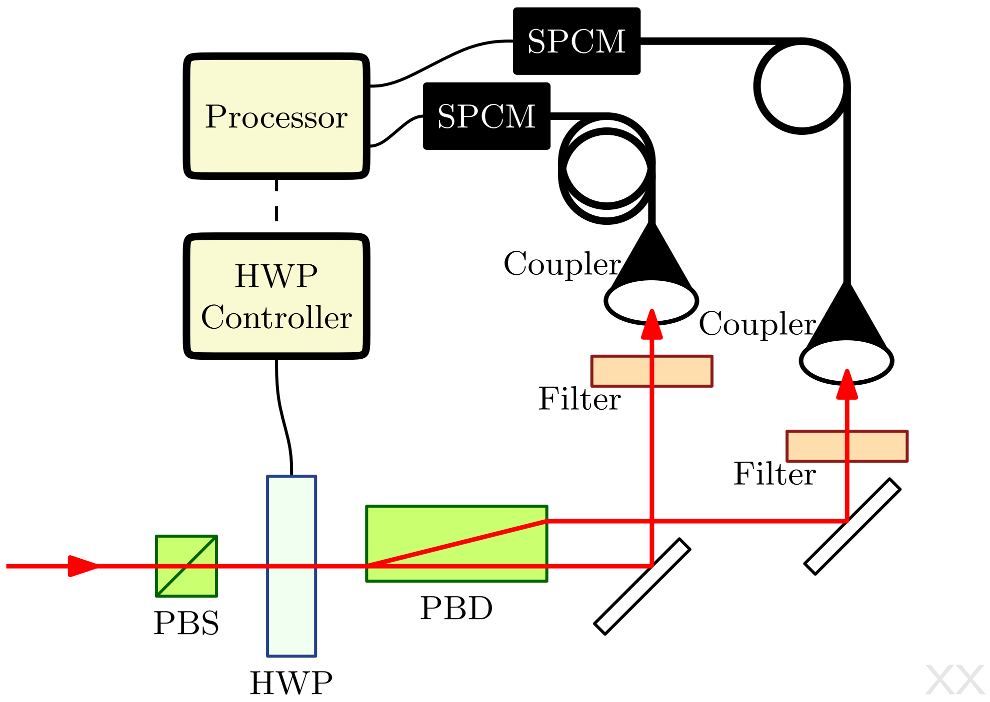
If $\rho_\pm \to |\phi_\pm\rangle$, we have a simple problem. **Theory** by Acín *et al.*, 2005:

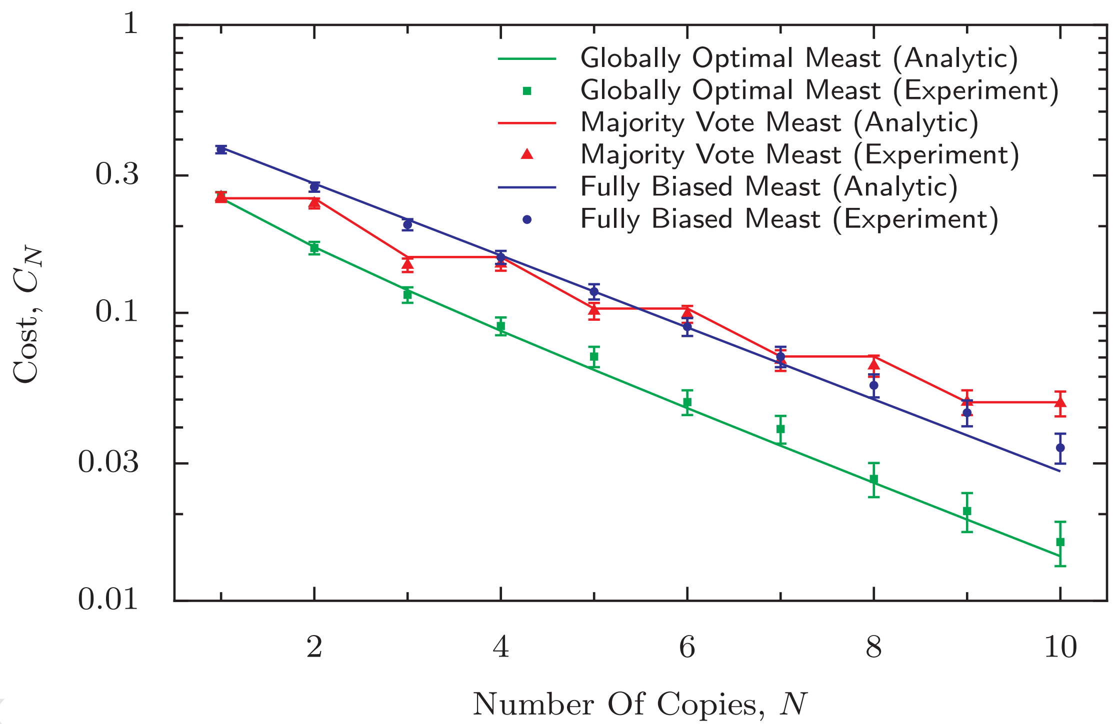
- Majority Vote (SQL): $C(N) = c^N$, where $c \equiv \cos 2\theta = |\langle\phi_+|\phi_-\rangle|$.
- Globally Opt. = Glob. Opt. Local = Locally Opt. Loc.: $C(N) = c^{2N}$.
- Fully Biased (Unanimity Vote): $C(1) > c$, but $\lim_{N\to\infty} C(N) \propto c^{2N}$.

Pure State Case (Experiment)

Higgins, Booth, Doherty, Bartlett, HMW, Pryde (unpub.).

Parameters: $\theta = 15°$, $r > 0.9999$.



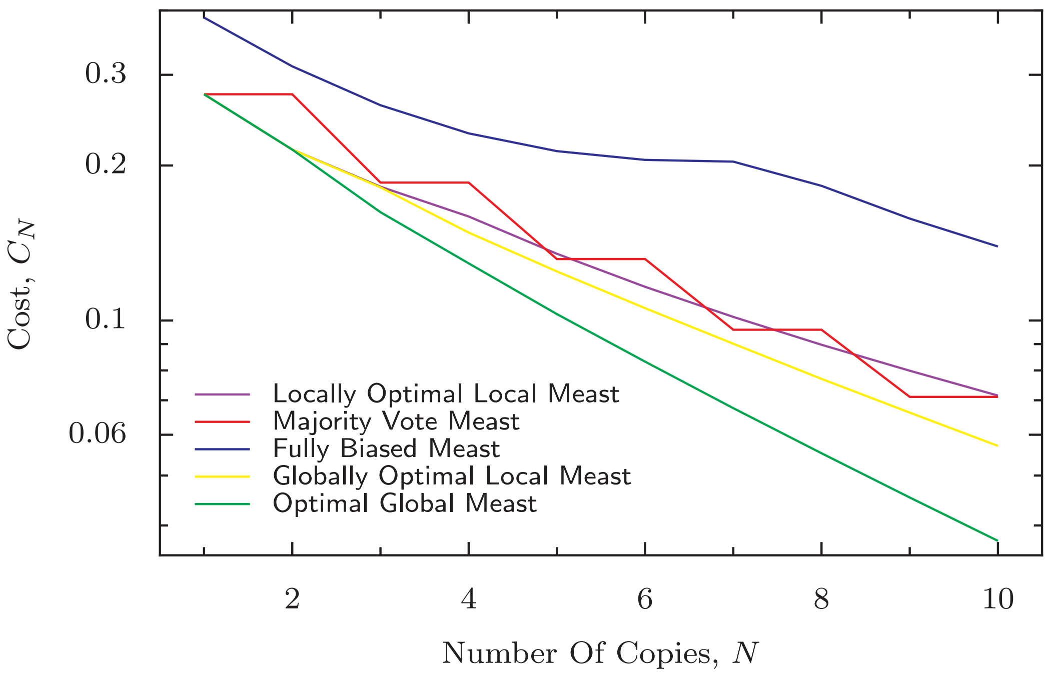


Mixed State Case (Theory & Experiment)

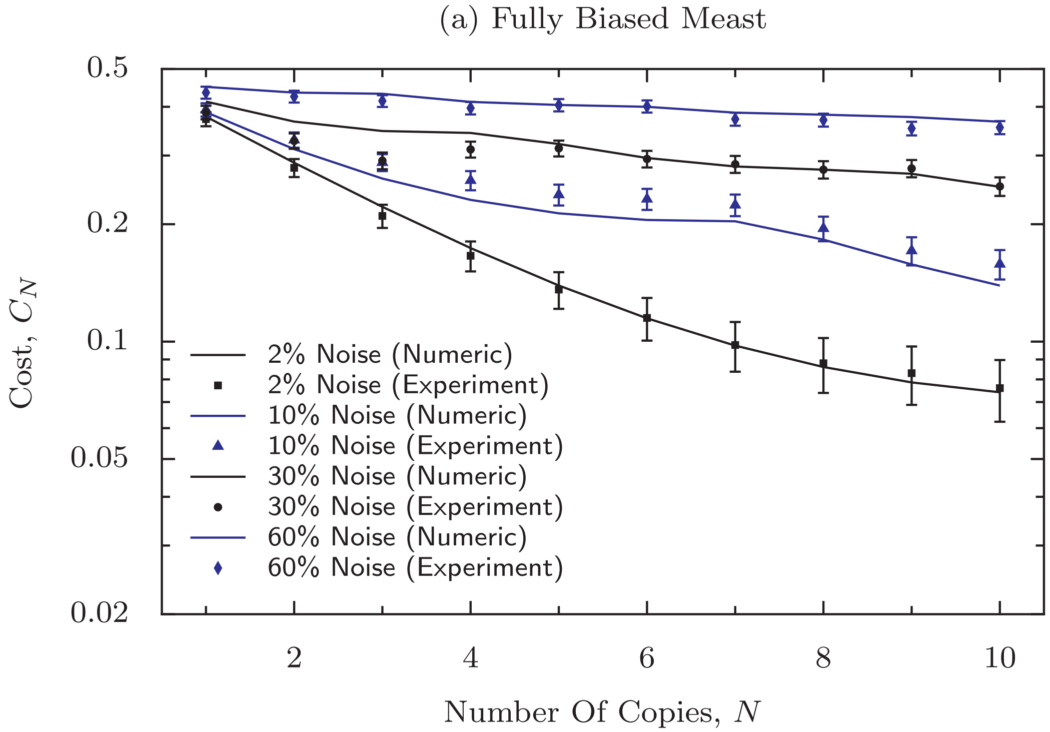
For systems with non-zero noise ($= 1 - r$), the problem is much more complicated — analytical results possible only for MV and GO.

All schemes are now different, and FB and LOL can be worse than SQL.

Theory for 10% noise:



Experiment:

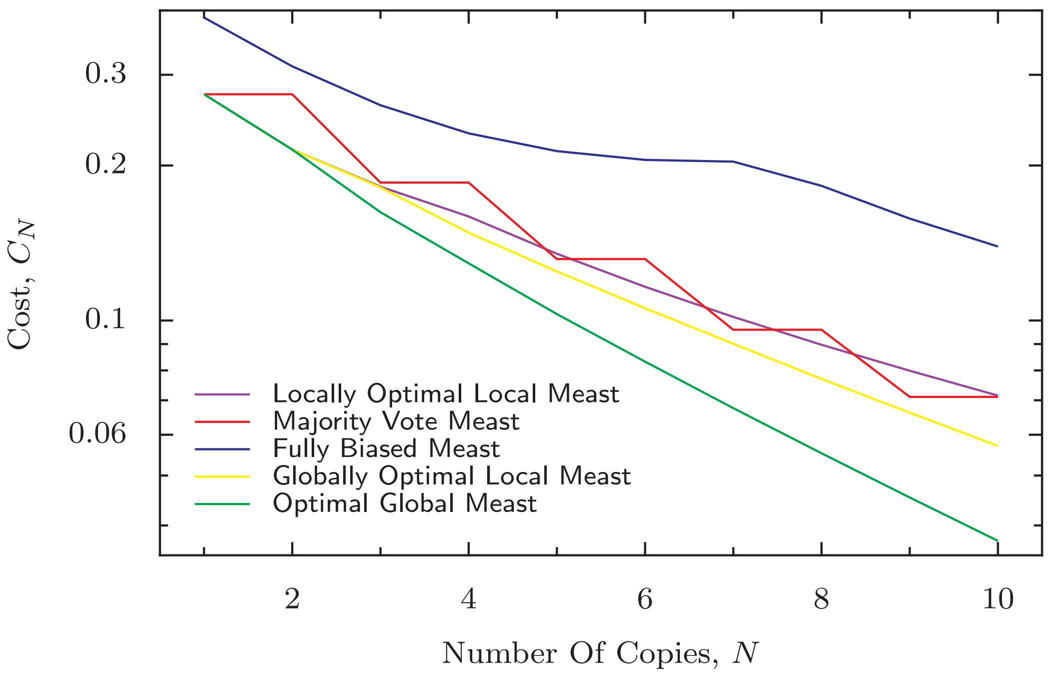


Mixed State Case (Theory & Experiment)

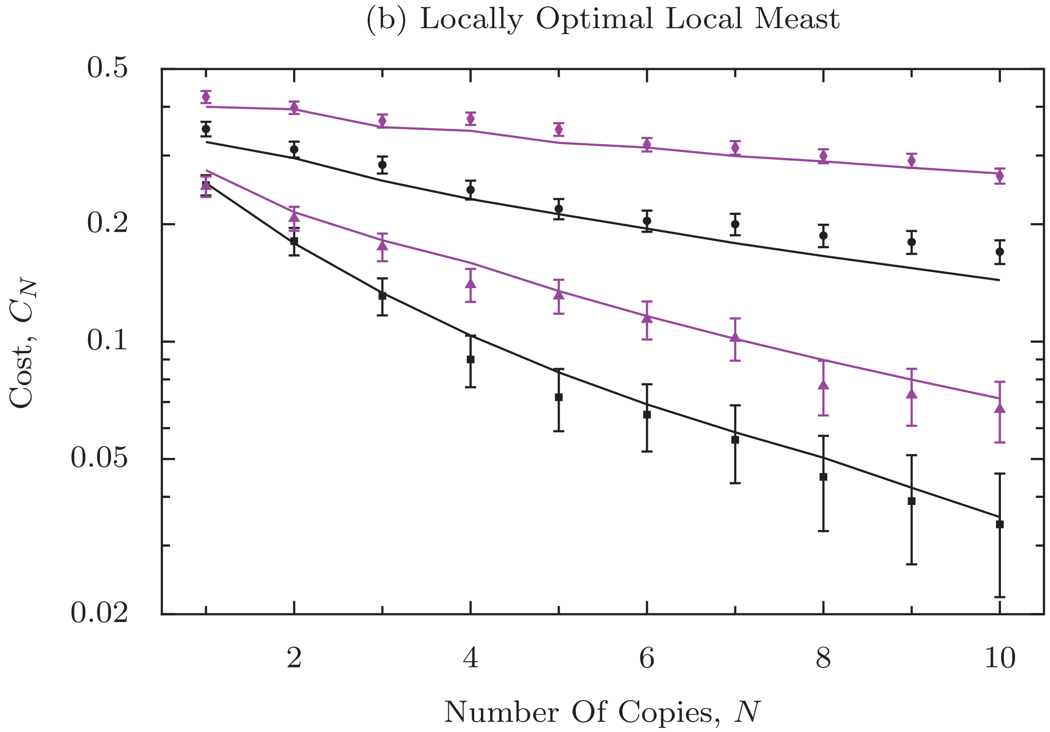
For systems with non-zero noise ($= 1 - r$), the problem is much more complicated — analytical results possible only for MV and GO.

All schemes are now different, and FB and LOL can be worse than SQL.

Theory for 10% noise:



Experiment:

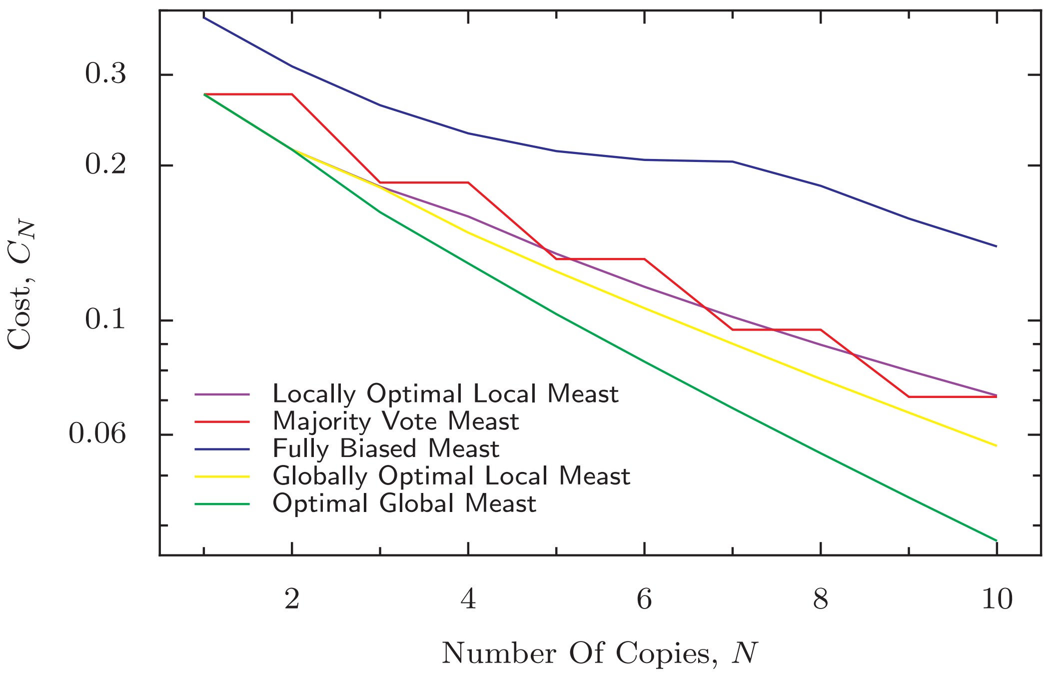


Mixed State Case (Theory & Experiment)

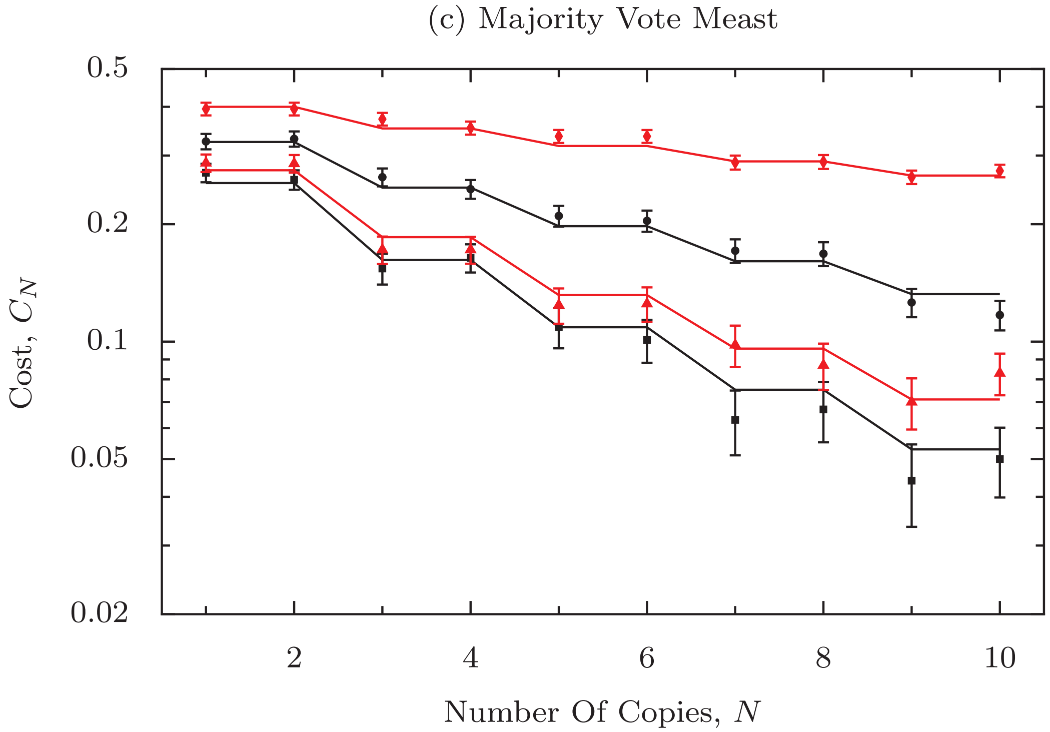
For systems with non-zero noise ($= 1 - r$), the problem is much more complicated — analytical results possible only for MV and GO.

All schemes are now different, and FB and LOL can be worse than SQL.

Theory for 10% noise:



Experiment:

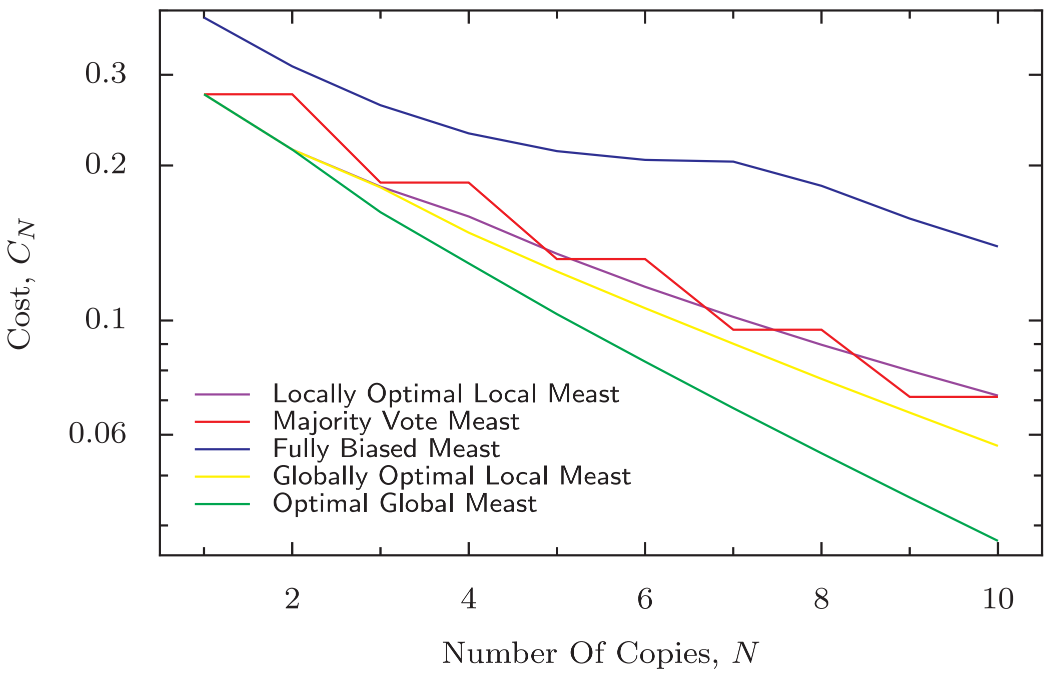


Mixed State Case (Theory & Experiment)

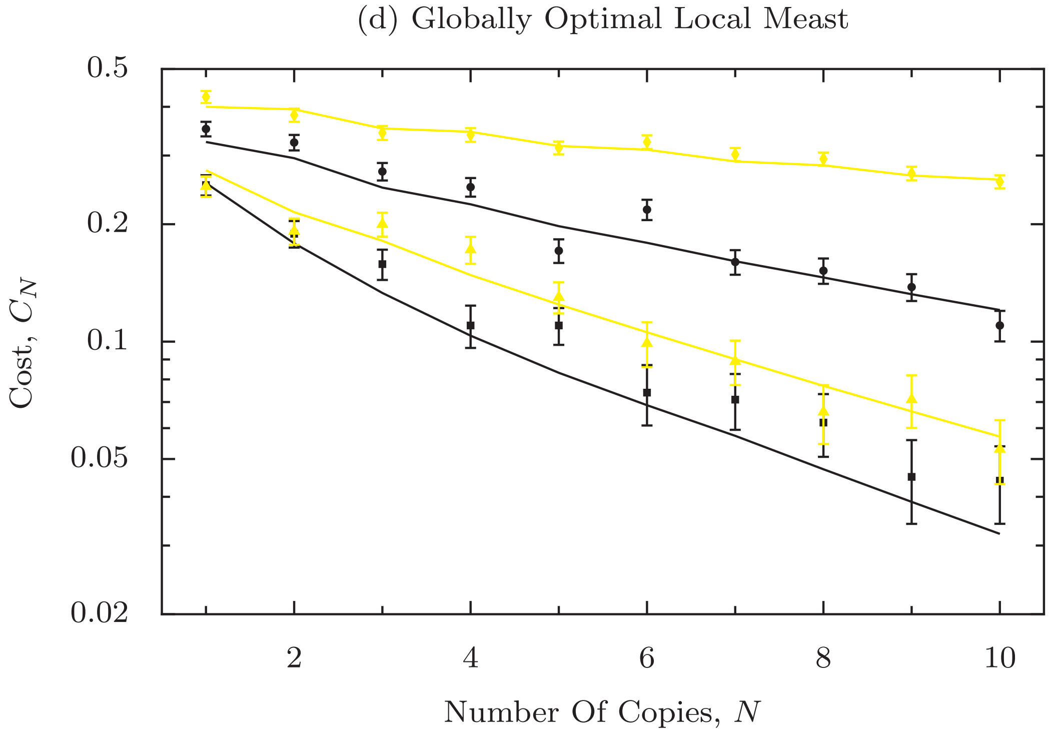
For systems with non-zero noise $(= 1 - r)$, the problem is much more complicated — analytical results possible only for MV and GO.

All schemes are now different, and FB and LOL can be worse than SQL.

Theory for 10% noise:



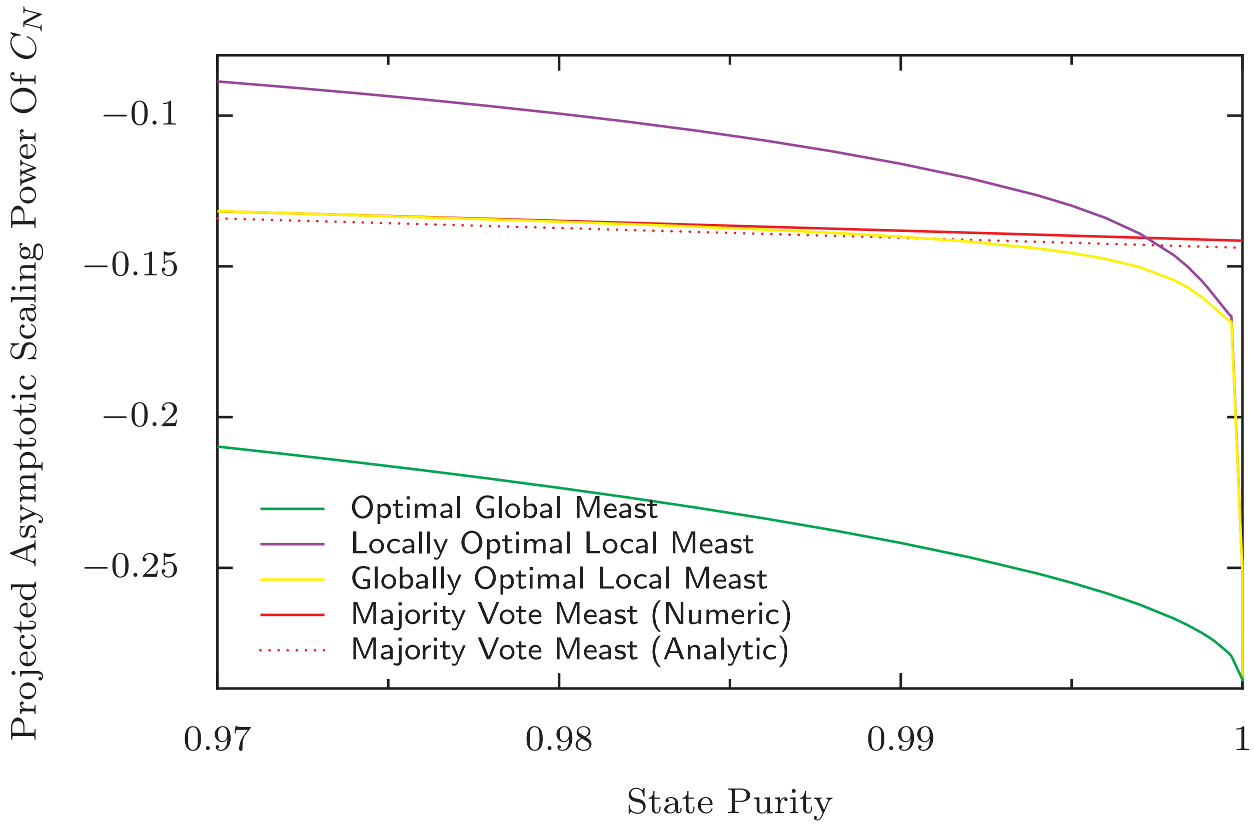
Experiment:



Mixed State Case — Asymptotic Theory

Look at $L = \lim_{N\to\infty}(\partial/\partial N)\log C(N)$. In practice $N \sim 200$ is sufficient.

To calculate accurately with DP, we need small *grid spacing* S for $\{\wp\}$.
We fit the data to $L(S) = a - b|\log S|^{-1.22}$, then extrapolate to $L(0) = a$.



Conclusions (Part II)

1. For mixed states, the optimal local (single qubit) state discrimination scheme can only be achieved by applying dynamic programming, a technique from optimal stochastic control theory.

2. In $N \gg 1$ limit, the different schemes behave very differently in different regimes of purity:

	How Pure are the States?		
Measurement Scheme	100% Pure	Almost ($\gtrsim 99.9\%$)	Not Very ($\lesssim 99\%$)
Majority Vote Meas.t	SQL	SQL	SQL
Fully Biased Meas.t	$\sim$ Helstrom Limit	Bad!	Bad!
Locally Optimal Local Meas.t	Helstrom Limit	sub-SQL	Bad!
Globally Optimal Local Meas.t	Helstrom Limit	**more** sub-SQL	$\approx$ SQL
Optimal Global Meas.t	Helstrom Limit	Helstrom Limit	Helstrom Limit

Conclusions (Global)

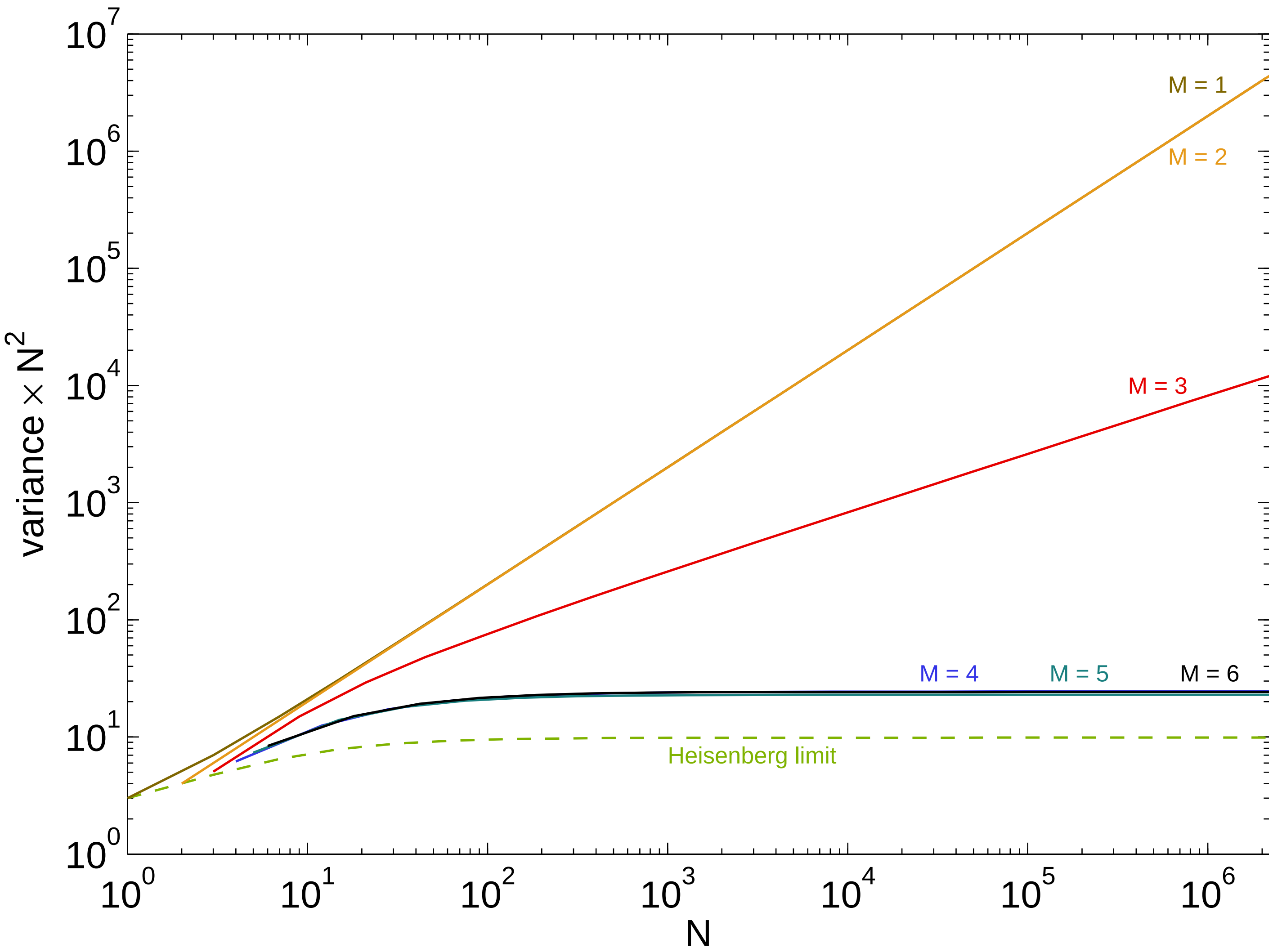
Adaptive local measurements always give better performance than non-adaptive local measurements.

However, in terms of **asymptotic** $(N \gg 1)$ **scaling** of the performance:

1. in phase estimation and **pure** state discrimination,
   - adaptation **is sufficient** to achieve the Heisenberg/Helstrom Limit.
   - adaptation is **not necessary** for the Heisenberg/Helstrom Limit.

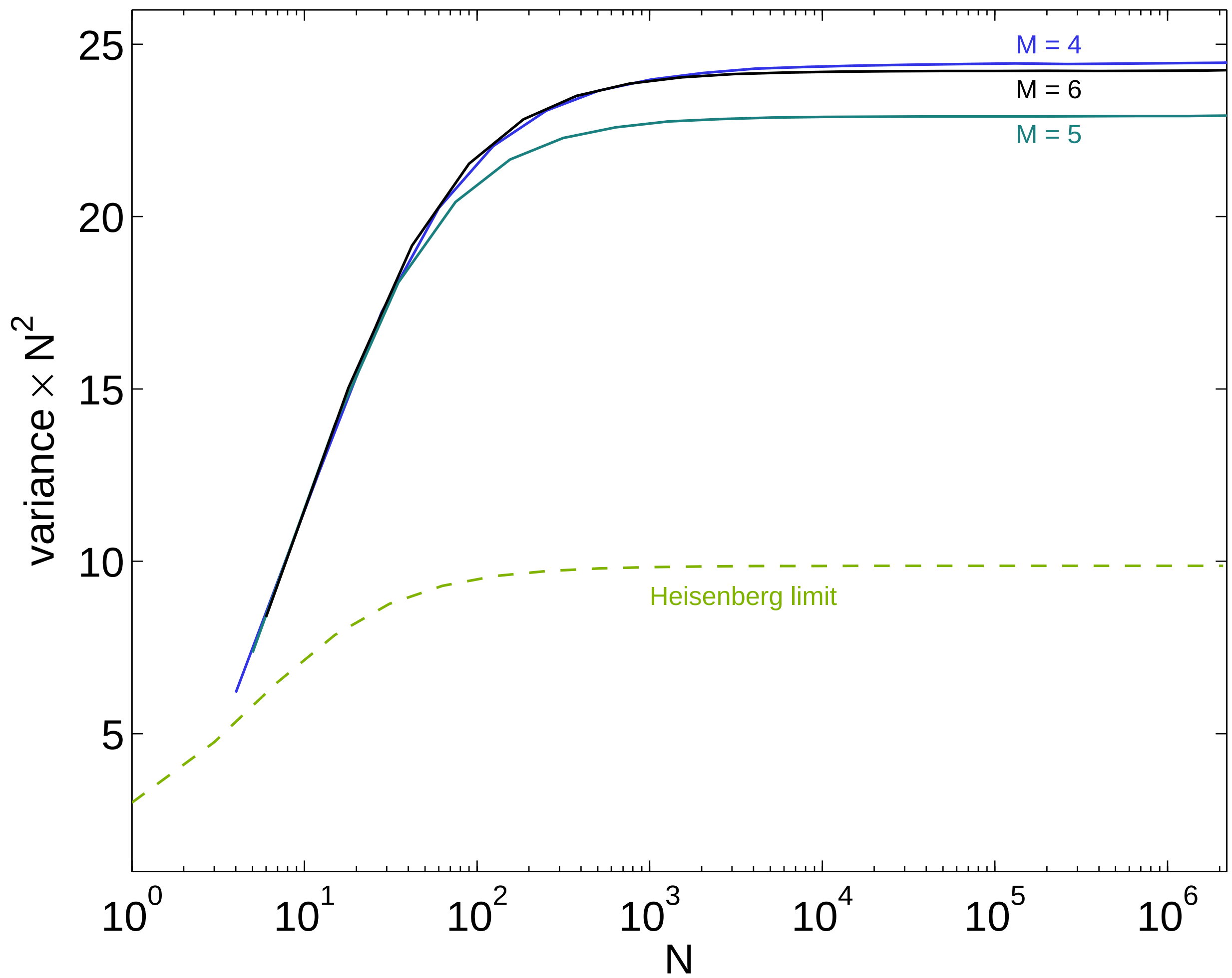
2. in **almost-pure** state discrimination
   - adaptation is **not sufficient** to achieve the Helstrom Limit.
   - adaptation **is sufficient** (and perhaps **necessary**) to **beat** the SQL.

Numerical Results: Variances for all M



Our adaptive scheme acheives HL scaling for $M \geq 4$ ...

Numerical Results: Selected Variances



… with an overhead as small as ≈ 2.3 for $M = 5$.