University of Glasgow

Blume-Kohout, R., Croke, S., and Zwolak, M. (2013) Quantum data gathering. Scientific Reports, 3 (1800). ISSN 2045-2322

http://eprints.gla.ac.uk/89032

Deposited on: 07 January 2014

Quantum data gathering

Robin Blume-Kohout[1], Sarah Croke[2] & Michael Zwolak[3]

[1]Sandia National Laboratories, MS-1322, Albuquerque, NM 87185, [2]Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada, [3]Department of Physics, Oregon State University, Corvallis, OR 97331.

Measurement of a quantum system – the process by which an observer gathers information about it – provides a link between the quantum and classical worlds. The nature of this process is the central issue for attempts to reconcile quantum and classical descriptions of physical processes. Here, we show that the conventional paradigm of quantum measurement is directly responsible for a well-known disparity between the resources required to extract information from quantum and classical systems. We introduce a simple form of *quantum* data gathering, "coherent measurement", that eliminates this disparity and restores a pleasing symmetry between classical and quantum statistical inference. To illustrate the power of quantum data gathering, we demonstrate that coherent measurements are optimal and strictly more powerful than conventional one-at-a-time measurements for the task of discriminating quantum states, including certain entangled many-body states (e.g., matrix product states).

O bservation is at the heart of understanding physical phenomena. More broadly, it serves as the first step in information processing. In the canonical example, a classical observer gathers data about a quantum system by measuring it. This process yields classical information about the state, but at a high price – the observer's information gain is incomplete (it doesn't uniquely identify the state), and accompanied by an irreversible loss of information. It is typically further restricted by locality; most joint measurements on an N-component quantum system are practically impossible when the components are spatially delocalized. This has profound consequences for extracting information from multiple samples of a quantum state. Measuring them individually is strictly less powerful than a joint measurement on the whole ensemble. By contrast, in classical theories, measurement is a passive procedure with no accompanying information loss, and complete information can be gathered from independent measurements of the samples. This implies a fundamental asymmetry between data-gathering in quantum and classical theories.

The conventional paradigm of quantum measurement comprises: (i) a controlled unitary interaction between a system $\mathcal{S}$ and an apparatus $\mathcal{A}$; (ii) decoherence on $\mathcal{A}$, which forces its state into a mixture of "pointer basis" states[1]; and (iii) experimental readout of the classical result from $\mathcal{A}$ (arguably accompanied by "collapse" of $\mathcal{A}$'s state). But this prescription breaks down if any part of the observing apparatus can sustain coherence. A *coherent apparatus* has other, more powerful information-gathering strategies available to it. To demonstrate this in the context of traditional information processing, we allow $\mathcal{A}$ to be a quantum information processor (QIP) – basically a very small (perhaps just 1 qubit) non-scalable quantum computer. We focus on the case where $\mathcal{A}$ is protected from decoherence until the very end of the protocol. What remains is a *coherent measurement*, a unitary interaction between $\mathcal{S}$ and $\mathcal{A}$ that transfers information from $\mathcal{S}$ to $\mathcal{A}$.

To explore and demonstrate the power of coherent measurement, let us consider the specific problem of *quantum state discrimination*[2,3]: Given N quantum systems that were all prepared in one of K distinct states $|\psi_1\rangle,\dots,|\psi_K\rangle$, decide in which state they were prepared. In principle, finding the optimal measurement is a straightforward convex program. But when $N > 1$ copies of $|\psi_k\rangle$ are available, this task (famously) requires joint measurement on all N copies, which is prohibitively difficult. Observing each of the N copies independently yields a strictly lower probability of success[4,5]. This contrasts starkly with the corresponding classical problem of distinguishing K distinct probability distributions, where one-at-a-time observations are completely sufficient. Discrimination thus provides an ideal scenario to test the information-processing utility of more general data gathering paradigms.

Results

Simple discrimination with N copies. Suppose we are given N quantum systems ($\mathcal{S}_1 \dots \mathcal{S}_N$) with d-dimensional Hilbert spaces $\mathcal{H}_n$, and a promise that they were all identically prepared in one of K nonorthogonal states $\{|\psi_1\rangle \dots |\psi_K\rangle\}$. Their joint state is $|\psi_k\rangle^{\otimes N} \in \mathcal{H}^{\otimes N}$, with k unknown. Identifying k with maximum success probability requires a joint measurement on all N samples. Non-adaptive, one-at-a-time measurement cannot

1

achieve the optimal success probability. For $K = 2$ candidate states, there is an adaptive local measurement scheme that achieves the optimal success probability[6], but no such protocol has been found for $K > 2$. It has recently been shown that for K = 3, there exist cases for which local measurements cannot achieve the optimal success probability, even with multiple rounds of classical communication[7].

All the information about k is contained in a K-dimensional subspace

$$\mathcal{K}_N = \mathrm{Span}\left(\left\{|\psi_k\rangle^{\otimes N}\right\}\right). \qquad (1)$$

So while the optimal measurement is a joint measurement, it does not need to explore the majority of $\mathcal{H}$. We will implement it by rotating the entire subspace $\mathcal{K}_N$ into the state space of our K-dimensional QIP $\mathcal{A}$ (the coherent measurement apparatus). We do so via sequential independent interactions between $\mathcal{A}$ and each of the N samples $\mathcal{S}_n$, "rolling up" all information about k into $\mathcal{A}$.

$\mathcal{A}$ is initially prepared in the $|0\rangle$ state. We bring it into contact with $\mathcal{S}_1$, and execute a SWAP gate between $\mathcal{S}_1$ and the $\{|0\rangle,...,|d-1\rangle\}$ subspace of $\mathcal{A}$. This transfers all information from the first sample into $\mathcal{A}$, leaving $\mathcal{S}_1$ in the $|0\rangle$ state.

Now we bring $\mathcal{A}$ into contact with $\mathcal{S}_2$. Their joint state is $|\psi_k\rangle^{\otimes 2}$, although we do not know k. But we do know that the state lies within $\mathcal{K}_2 = \mathrm{Span}\left(\left\{|\psi_k\rangle^{\otimes 2}\right\}\right)$, (see Eq. 1), whose dimension is at most K. A basis $\{|\phi_j\rangle : j = 1...K\}$ for this space can be obtained by Gram-Schmidt orthogonalization. We apply a unitary interaction between $\mathcal{A}$ and $\mathcal{S}_2$,

$$U_2 = \sum_j |0_{\mathcal{S}_2} j_{\mathcal{A}}\rangle\langle\phi_j|. \qquad (2)$$

where we have defined U_2 only on the subspace of interest (for completeness, it can be extended to the complement, $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_2/\mathcal{K}_2$, in any convenient manner). It rotates $\mathcal{K}_2$ into $\{|0\rangle\}_{\mathcal{S}_2} \otimes \mathcal{H}_{\mathcal{A}}$, which places all the information about k in $\mathcal{A}$ and decouples $\mathcal{S}_2$ ($\mathcal{S}_2$ is left with *no* information about k if and only if $\mathcal{A}$ is left with *all* the information about k.). $\mathcal{A}$ is now in one of K possible states $|\psi_k^{(2)}\rangle$, which (as a set) are unitarily equivalent to $\{|\psi_k\rangle^{\otimes 2}\}$ – e.g., $\langle\psi_j^{(2)} | \psi_k^{(2)}\rangle = \langle\psi_j | \psi_k\rangle^2$.

The rest of the algorithm is now fairly obvious; we move on and interact $\mathcal{A}$ with $\mathcal{S}_3$ in the same way, etc, etc. At each step, when $\mathcal{A}$ comes into contact with $\mathcal{S}_n$, their joint state is $|\psi_k^{(n-1)}\rangle \otimes |\psi_k\rangle$. These K alternatives reside in a K-dimensional space $\mathcal{K}_n$ (see Eq. 1), spanned by a basis $\left\{|\phi_j^{(n)}\rangle\right\}$, which is then rotated into $\{|0\rangle\}_{\mathcal{S}_n} \otimes \mathcal{H}_A$ by applying

$$U_n = \sum_j |0_{\mathcal{S}_n} j_{\mathcal{A}}\rangle\langle\phi_j^{(n)}|. \qquad$$

Each sample system is left in the $|0\rangle$ state, indicating that all its information has been extracted. After every sample has been sucked dry, we simply measure $\mathcal{A}$ to extract k. This final measurement can be efficiently computed via convex programming, since $\mathcal{A}$ is only K-dimensional.

The sequence of coherent measurement interactions is independent of what sort of discrimination we want to do – e.g., minimum-error[4], unambiguous discrimination[8–11], maximum-confidence[12], etc. – because $\mathcal{K}_N$ is a *sufficient statistic* for any inference about k, and our protocol simply extracts it whole, leaving the decision rule up to the final measurement on $\mathcal{A}$. As in the classical case, data gathering can now be separated from data analysis.

As an example of the data gathering protocol, we consider now the $K = 2$ case, and take the two possible initial states to be single qubit states, parametrized as follows

$$|\psi_1\rangle = \cos\theta\,|0\rangle - \sin\theta\,|1\rangle, \qquad (3)$$

$$|\psi_2\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle. \qquad (4)$$

The joint state of the first n samples is given by

$$|\psi_k\rangle^{\otimes n} = \sum_{j=0}^{n} (-1)^{jk}\sqrt{\binom{n}{j}}(\cos\theta)^{n-j}(\sin\theta)^j|n,j\rangle \qquad (5)$$

where $|n,j\rangle$ is a normalized equal superposition of all tensor product states with $n - j$ systems in state $|0\rangle$ and j systems in state $|1\rangle$. A natural basis for the two-dimensional subspace $\mathcal{K}_n = \mathrm{Span}\left(\left\{|\psi_k\rangle^{\otimes n}\right\}\right)$ is then given by

$$|0\rangle_n = \sum_{j\,\mathrm{even}} \frac{\sqrt{\binom{n}{j}}(\cos\theta)^{n-j}(\sin\theta)^j}{\sqrt{\frac{1}{2}(1+(\cos 2\theta)^n)}} \,|n,j\rangle, \qquad (6)$$

$$|1\rangle_n = \sum_{j\,\mathrm{odd}} \frac{\sqrt{\binom{n}{j}}(\cos\theta)^{n-j}(\sin\theta)^j}{\sqrt{\frac{1}{2}(1-(\cos 2\theta)^n)}} \,|n,j\rangle, \qquad (7)$$

and the goal of the data gathering algorithm is to transfer the subspace spanned by these states into the processor $\mathcal{A}$ via sequential interactions with each of the samples $\mathcal{S}_1 \ldots \mathcal{S}_n$.

At the first step, as indicated above, we perform a SWAP gate between $\mathcal{A}$ and $\mathcal{S}_1$. At the next step, the joint state of $\mathcal{A}$ and $\mathcal{S}_2$ is

$$|\psi_k\rangle_{\mathcal{S}_2}|\psi_k\rangle_{\mathcal{A}} = \cos^2\theta|0_{\mathcal{S}_2}0_{\mathcal{A}}\rangle + \sin^2\theta|1_{\mathcal{S}_2}1_{\mathcal{A}}\rangle$$

$$+ (-1)^k\cos\theta\,\sin\theta(|0_{\mathcal{S}_2}1_{\mathcal{A}}\rangle + |1_{\mathcal{S}_2}0_{\mathcal{A}}\rangle)$$

$$= \sqrt{\frac{1}{2}(1+\cos^2 2\theta)}|0\rangle_2$$

$$+ (-1)^k\sqrt{\frac{1}{2}(1-\cos^2 2\theta)}|1\rangle_2,$$

where $|0\rangle_2$, $|1\rangle_2$ are joint $\mathcal{S}_2\mathcal{A}$ states, given by the $n = 2$ case of Eqs. 6 and 7. A two-qubit unitary acting as follows

$$U_2|0\rangle_2 = |0_{\mathcal{S}_2}0_{\mathcal{A}}\rangle,$$

$$U_2|1\rangle_2 = |0_{\mathcal{S}_2}1_{\mathcal{A}}\rangle,$$

thus leaves $\mathcal{S}_2$ in state $|0\rangle$ while the processor, in state

$$|\psi_k^{(2)}\rangle = \sqrt{\frac{1}{2}(1+\cos^2 2\theta)}\,|0\rangle + (-1)^k\sqrt{\frac{1}{2}(1-\cos^2 2\theta)}\,|1\rangle$$

now holds all the information previously contained in the two samples read so far. Now suppose the processor has interacted with and processed the first $n - 1$ samples. Then we have performed a unitary which acts as follows on the relevant subspace:

$$U_{n-1}\ldots U_2 U_1\left(|0\rangle_{n-1}|0\rangle_{\mathcal{A}}\right) = |0\rangle^{\otimes(n-1)}|0\rangle_{\mathcal{A}}$$

$$U_{n-1}\ldots U_2 U_1\left(|1\rangle_{n-1}|0\rangle_{\mathcal{A}}\right) = |0\rangle^{\otimes(n-1)}|1\rangle_{\mathcal{A}}$$

Defining $\cos\theta_n = \sqrt{\frac{1}{2}(1+(\cos 2\theta)^n)}$, and using Eqs. 5–7, the state of the processor is given by

$$|\psi_k^{(n-1)}\rangle = \cos\theta_{n-1}|0\rangle + (-1)^k\sin\theta_{n-1}|1\rangle$$

and we wish to update this state through interaction with the nth sample. The basis states $|0\rangle_n$, $|1\rangle_n$ spanning $\mathcal{K}_n$ are related to the corresponding basis states spanning $\mathcal{K}_{n-1}$, and the basis $|0\rangle$, $|1\rangle$ of

2

$\mathcal{S}_n$ as follows

$$|0\rangle_n = \frac{\cos\theta\,\cos\theta_{n-1}}{\cos\theta_n}|0\rangle_{n-1}|0\rangle + \frac{\sin\theta\,\sin\theta_{n-1}}{\cos\theta_n}|1\rangle_{n-1}|1\rangle,$$

$$|1\rangle_n = \frac{\sin\theta\,\cos\theta_{n-1}}{\sin\theta_n}|0\rangle_{n-1}|1\rangle + \frac{\cos\theta\,\sin\theta_{n-1}}{\sin\theta_n}|1\rangle_{n-1}|0\rangle.$$

Thus we wish to implement a 2-qubit unitary on the processor $\mathcal{A}$ and the nth sample, which acts as follows:

$$U_n\left(\frac{\cos\theta\,\cos\theta_{n-1}}{\cos\theta_n}|0_{\mathcal{S}_n}0_{\mathcal{A}}\rangle + \frac{\sin\theta\,\sin\theta_{n-1}}{\cos\theta_n}|1_{\mathcal{S}_n}1_{\mathcal{A}}\rangle\right) = |0_{\mathcal{S}_n}0_{\mathcal{A}}\rangle,$$

$$U_n\left(\frac{\sin\theta\,\cos\theta_{n-1}}{\sin\theta_n}|1_{\mathcal{S}_n}0_{\mathcal{A}}\rangle + \frac{\cos\theta\,\sin\theta_{n-1}}{\sin\theta_n}|0_{\mathcal{S}_n}1_{\mathcal{A}}\rangle\right) = |0_{\mathcal{S}_n}1_{\mathcal{A}}\rangle.$$

This can be achieved by, e.g., performing a CNOT gate with $\mathcal{S}_n$ as the control and $\mathcal{A}$ as the target, followed by a controlled unitary on $\mathcal{S}_n$ to rotate the residual state to $|0\rangle$, controlled on the state of $\mathcal{A}$. Continuing in this way we can roll up all the information contained in the provided samples $\mathcal{S}_1\ldots\mathcal{S}_N$ into the processor $\mathcal{A}$. Each step requires only a 2 qubit unitary, acting on the nth sample and the processor. During the course of this process all the information contained in the joint N qubit state is transferred to the processor, where a single qubit measurement is sufficient to perform any allowed data analysis operation.

This protocol can be modified to discriminate non-symmetric product states, e.g., $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots |\psi_N\rangle$ vs. $|\phi_1\rangle \otimes |\phi_2\rangle \otimes \ldots |\phi_N\rangle$. A more general extension, however, is to discriminating many-body states of the matrix product form.

Matrix product state discrimination. The information about k can be "rolled up" using sequential interactions because it is contained in a subspace $\mathcal{K}_N$ with Schmidt rank at most K across any division of the N systems. By this we mean that, given any pure or mixed state ρ on $\mathcal{K}_N$, if we trace out $\{\mathcal{S}_1\ldots\mathcal{S}_n\}$, the reduced state for $\{\mathcal{S}_{n+1}\ldots\mathcal{S}_N\}$ has rank at most K. Low Schmidt rank is critical. Consider distinguishing between two states that are each maximally entangled between the first $N/2$ samples and the last $N/2$ samples. They lie in a 2-dimensional subspace, but it is *not* accessible through our protocol with a low-dimensional QIP. The first $N/2$ samples are maximally entangled with the rest, so their reduced state has rank $d^{N/2}$. At least $N\,\log_2 d/2$ qubits would be needed to store the information extracted from the first $N/2$ samples.

But whenever the Schmidt rank condition is satisfied, a variation of our algorithm will work. For product states (above), each state has Schmidt rank 1, and the span of K such states has Schmidt rank at most K.

This property of low Schmidt rank is generalized by *matrix product states* (MPS)[13,14]. An N-particle MPS with *bond dimension* D is guaranteed to have Schmidt rank at most D across any division of the 1D lattice. Thus, the span of K such MPS, each with bond dimension $\le D$, has Schmidt rank at most DK. We denote such a subspace,

$$\mathcal{K} = \mathrm{Span}(\{|\psi_k\rangle\}),$$

a matrix product subspace with bond dimension DK. Such a set of MPS can be distinguished optimally with coherent measurements and $(\log_2 D + \log_2 K)$ qubits.

Our algorithm is a straightforward generalization of the one for product states, and proceeds as shown in Figure 1. First, we represent the MP subspace $\mathcal{K}$ by its *purification* – a single MPS $|\Psi\rangle$ for $\mathcal{S}_1\ldots\mathcal{S}_N$ *and* a fictitious reference system R,

$$|\Psi\rangle \propto \sum_k |k\rangle_R |\psi_k\rangle_{\mathcal{S}_1\ldots\mathcal{S}_N}, \tag{8}$$

Information about k, which is contained in $\mathcal{K}$, equates to correlation with the imaginary R.

Now, we initialize $\mathcal{A}$ in the $|0\rangle$ state, then SWAP its state with that of $\mathcal{S}_1$ (the first lattice site). This decouples $\mathcal{S}_1$, leaving $\mathcal{A}\otimes\mathcal{S}_2\ldots\mathcal{S}_N\otimes R$ in a matrix product state,

$$|0\rangle_{\mathcal{A}}|\Psi\rangle_{\mathcal{S}_1\ldots\mathcal{S}_N,R} \to |0\rangle_{\mathcal{S}_1}|\Psi\rangle_{\mathcal{A},\mathcal{S}_2\ldots\mathcal{S}_N,R}, \tag{9}$$

with Schmidt rank no greater than DK.

Now, to roll up each successive site $\mathcal{S}_n(n=2,\ldots N)$, we find the Schmidt decomposition of the current state between $\mathcal{A}\otimes\mathcal{S}_n$ and the remainder of the lattice, write it (generically) as

$$\sum_{j=1}^{DK} c_j|\mu_j\rangle_{\mathcal{A}\mathcal{S}_n}|\nu_j\rangle_{\mathcal{S}_{n+1}\ldots\mathcal{S}_N,R}, \tag{10}$$

and apply a unitary operation to $\mathcal{A}\otimes\mathcal{S}_n$,

$$U_n = \sum_j |0_{\mathcal{S}_n}j_{\mathcal{A}}\rangle\langle\mu_j|, \tag{11}$$

which decouples $\mathcal{S}_n$ and leaves all the information previously in $\mathcal{S}_n\otimes\mathcal{A}$ in $\mathcal{A}$. As in Eq. 2, this unitary is specified only on the subspace of interest, and can be completed in any convenient fashion. Doing this successively at each site decorrelates all the $\mathcal{S}_n$, and we are left in the state

$$|\Psi'\rangle \propto \sum_k |k\rangle_R |\psi_k\rangle_{\mathcal{A}},$$

with all information about k in $\mathcal{A}$, where it can be extracted by a simple measurement.

Recent proposals for local tomography[15] are also based on sequential interactions. Our protocol, with coherent measurements, offers a tremendous efficiency advantage (at the cost of requiring a small QIP!). It can distinguish near-orthogonal MPS states with a single copy, whereas local tomography requires $O(N)$ copies. Distinguishing non-orthogonal states requires multiple (M) copies. To apply our algorithm, we simply line up the copies (they do not have to exist simultaneously) and treat them as a single NM-particle MPS of bond dimension D.

Mixed state discrimination and tomography. In the context of N-copy states, one may ask:

1. Can coherent measurement be used to discriminate mixed states, i.e. $\rho_k^{\otimes N}$?
2. Can coherent measurement be used for full state tomography (rather than discrimination)?

The answer to both is "Yes, but it seems to require an $O(\log N)$-qubit apparatus." This is a very favorable scaling, but less remarkable (and less immediately useful) than the $O(1)$ scaling for pure state discrimination.

This is possible because the order of the samples is completely irrelevant. As we scan through the samples, we can discard ordering information, keeping only a sufficient statistic for inference about ρ. The *quantum Schur transform* does this[16]. It is based on Schur-Weyl duality[17], which states that because permutations of the N samples commute with collective rotations on them, the Hilbert space $\mathcal{H}_d^{\otimes N}$ can be refactored as

$$\mathcal{H}^{\otimes N} = \bigoplus_\lambda \mathcal{U}_\lambda \otimes \mathcal{P}_\lambda.$$

The $\mathcal{U}_\lambda$ factors are *irreducible representation spaces* (irreps) of $SU(d)$, the $\mathcal{P}_\lambda$ factors are irreps of S_N, and λ labels the various irreps. The Schur transform can be implemented by a unitary circuit that acts sequentially on the samples, mapping N qudit registers into three quantum registers containing (respectively) λ, $\mathcal{U}$, and $\mathcal{P}$:
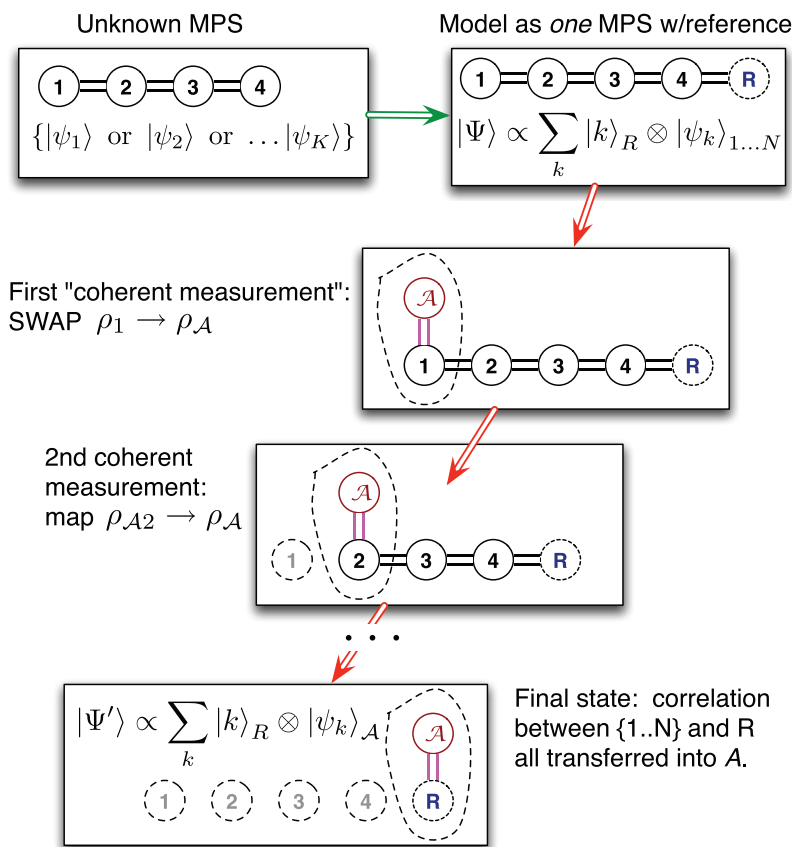
Figure 1 | Discrimination of MPS. Our protocol represents an unknown MPS $|\psi_k\rangle$ from a set $\{|\psi_1\rangle \ldots |\psi_K\rangle\}$ by its *purification* – a single MPS $|\Psi\rangle$ involving a fictitious reference R. The algorithm then successively decorrelates each sample $\mathcal{S}_n$ from the rest, storing $\mathcal{S}_n$'s correlations with the remainder of the lattice in the "apparatus" $\mathcal{A}$. Ultimately, all existing information about k (i.e., its correlation with R) is contained in $\mathcal{A}$, which can be measured.

$$\mathcal{H}^{\otimes N} \rightarrow \mathcal{H}_\lambda \otimes \mathcal{H}_{\mathcal{U}} \otimes \mathcal{H}_{\mathcal{P}}.$$

The "ordering" register $\mathcal{H}_{\mathcal{P}}$ accounts for nearly all the Hilbert space dimension of $\mathcal{H}_d^{\otimes N}$, and since it is irrelevant to inference it can be discarded as rapidly as it is produced. What remains to be stored in $\mathcal{A}$ is:

1. a "label" register λ (a sufficient statistic for the spectrum of ρ),
2. a $SU(d)$ register $\mathcal{U}$ (a sufficient statistic for the eigenbasis of ρ).

The λ register requires a Hilbert space with dimension $\geq$ the number of Young diagrams with N boxes in at most d rows, which is approximately

$$\dim(\lambda) \approx \frac{1}{d!} \binom{N}{d}.$$

The $\mathcal{U}$ register must hold the largest N-copy irrep of $SU(d)$, whose size can be calculated using hook-length formulae[18], and upper bounded by

$$\dim_{\max}(\mathcal{U}) = (N+d-1)^{\frac{1}{2}d(d-1)}.$$

Together, these registers require $O(d^2 \log N)$ qubits of memory (although for pure state tomography, $O(d \log N)$ qubits of memory are sufficient).

$O(\log N)$ memory appears to be *necessary* for optimal accuracy. Consider the simplest possible case – discrimination of two *classical* 1-bit probability distributions

$$\binom{p}{1-p} \quad \text{vs.} \quad \binom{q}{1-q}.$$

The sufficient statistic is frequencies of "0" and "1", $\{n, N-n\}$. For any given problem, there is a threshold value n_c such that the answer

depends only on whether $n < n_c$, so only one bit of information is required. However, extracting that bit via sequential queries requires storing n exactly at every step (using $O(\log N)$ bits of memory). Any loss of precision could cause a ± 1 error at the final step, and thus a wrong decision. In this example, classical storage is sufficient. But in the general case, where the candidate ρ_k do not commute, no method is known to compress the intermediate data into classical memory without loss (previous work suggests it is probably impossible[19]).

Discussion

Quantum information science is rife with gaps between what is theoretically achievable and what is practically achievable. Our algorithm eliminates performance gaps for pure state discrimination with local measurements – but it requires a new kind of measurement apparatus with at least 1 controllable qubit of quantum storage. Its utility depends on its applications, and on the difficulty of implementation.

One *immediate* application of our protocol is detection of weak forces and transient effects. A simple force detector (e.g., for magnetic fields) might comprise a large array of identical systems (e.g., $|\uparrow\downarrow\rangle$ spins). Each system is only weakly perturbed by the force, so information about the force is distributed across the entire array. Our algorithm efficiently gathers up that information with no loss – whereas local measurements with classical processing waste much information.

A more sensitive N-particle "antenna" would incorporate entanglement between the N particles[20]. High sensitivity can be achieved by simple MPS states with $D = 2$, like N00N states[21],

4

$$|\psi\rangle = \frac{|N,0\rangle + |0,N\rangle}{\sqrt{2}} = \frac{|\uparrow\rangle^{\otimes N} + |\downarrow\rangle^{\otimes N}}{\sqrt{2}}.$$

Collective forces do not change D, so the final states to be discriminated are also MPS. Our approach can discriminate such states *and* it can be used to prepare them, by running the "rolling up" process in reverse[22].

More ambitious applications include direct probing of many-body states, to test a particular MPS ansatz for a lattice system, or to characterize results of quantum simulations in optical lattices or ion traps. Without fully scalable quantum computers that can couple directly to many-body systems, coherent measurements may be the only way to efficiently probe complex N-particle states. Our protocol does not obviously scale to *PEPS*, the higher-dimensional analogues of MPS[14]. Like MPS, PEPS obey an *area law* – entanglement across a cut scales not with the volume of the lattice (N), but with the area of the cut. For a 1-dimensional MPS on N systems, any cut has area 1, so the Schmidt rank scales as $O(1)$, and our algorithm requires an $O(1)$ qubit QIP. Rolling up a general PEPS on an n-dimensional lattice would require $O\left(N^{\frac{n-1}{n}}\right)$ bits of quantum memory. However, some PEPS can be sequentially generated[23], and are likely amenable to our protocol.

We note that the engineering requirements for a coherent measurement apparatus are achievable with near-future technology. $\mathcal{A}$ should be a clean K-dimensional quantum system with:

1. Universal local control,
2. Long coherence time relative to the gate timescale,
3. Controllable interaction with an external d-dimensional "sample" system,
4. Sequential coupling to each of N samples,
5. Strong measurements (which may be destructive).

Although adaptive local measurements can discriminate 2 states optimally, $K = 2$ is nevertheless sufficient for a proof of principle experiment demonstrating quantum data gathering. That a sufficient statistic for any desired measurement is stored as quantum data in the processor could be demonstrated by e.g. performing minimum error or unambiguous discrimination through eventual measurement on the processor in different runs of the same experiment. However, $K \geq 3$ would be more exciting as optimal discrimination of 3 states via local, adaptive measurements is not possible in all cases[7].

These requirements are much weaker than those for scalable quantum computing. Coherent measurement could be an early application for embryonic quantum architectures. Furthermore, scalability is not required, just a single K-dimensional system. Only local control has to be universal, since the interaction with external systems is limited. Error correction is not mandatory, for coherence need only persist long enough to interact with each of the N systems of interest. Since measurements are postponed until the end, they can be destructive.

We do require $\mathcal{A}$ to be portable – i.e., sequentially coupled to each of the N samples – whereas a quantum computer can be built using only nearest neighbor interactions. Fortunately, most proposed architectures have selective coupling either through frequency space (NMR, ion traps with a phonon bus) or physical motion of the qubits (some ion traps) or flying qubits (photonic architectures). Devices that are not viable for full-scale quantum computing may be even better for coherent measurement. For example, an STM might pick up and transport a single coherent atomic spin along an array of sample atoms, interacting sequentially with each of them.

There are broad implications stemming from the results above. Coherent measurements are a genuinely new way to gather information. We have not just removed collapse from standard quantum measurements! That kind of coherent measurement is used already in quantum error correction, where it's common to replace a measurement of X with a controlled unitary of the form

$$U_{\mathcal{S}\mathcal{A}} = \sum_x |x\rangle\langle x|_{\mathcal{S}} \otimes U_{\mathcal{A}}^{(x)}. \tag{12}$$

Such unitaries transfer information about a specific observable X from $\mathcal{S}$ to $\mathcal{A}$. For appropriate $|\psi_0\rangle_{\mathcal{A}}$ and $U_{\mathcal{A}}^{(x)}$, later measurements of $\mathcal{A}$ produce exactly the same result as if $\mathcal{S}$ had been measured directly. The coherent measurements in our discrimination protocols are *not* of this form. They do not measure (i.e., transfer information to $\mathcal{A}$ about) a specific basis. For example, in N-copy state discrimination, $\mathcal{A}$ interacts with the first sample by a SWAP operation, which has no preferred basis. Later interactions are also not of controlled-U form (Eq. 12).

One might ask where the "measurement" occurs, since the interaction between $\mathcal{S}$ and $\mathcal{A}$ is purely unitary. The essence of measurement is that an observer or apparatus gains information. Quantum measurements are usually construed as mysterious processes that consume quantum states and excrete specific, definite measurement outcomes. Quantum theories of measurements usually represent them as (i) unitary interaction, (ii) decoherence and superselection, and finally (iii) wavefunction collapse or splitting of the universe. Our results suggest that unitary interaction (the only part of this sequence that is really understood) can stand alone as an information-gathering "measurement." And by avoiding decoherence, we can gather information strictly *more* effectively.

Decoherence is ubiquitous in human experience. But in its absence, there is no compelling reason why gathering information must be accompanied by collapse or definite outcomes. The whole point of quantum information science is to produce devices that do *not* decohere, and that can *process* information more efficiently than classical computers. The central message of this paper is that they can also *gather* information more efficiently. Unfettered by decoherence, they may still be constrained by locality. For such devices, coherent measurements are the natural way to gather information.

1. Zurek, W. H. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **75**, 715–75 (2003).
2. Chefles, A. Quantum state discrimination. *Contemp. Phys.* **41**, 401–424 (2000).
3. Barnett, S. M. & Croke, S. Quantum state discrimination. *Adv. Opt. Photon.* **1**, 238–278 (2009).
4. Helstrom, C. W. *Quantum detection and estimation theory* (Academic Press Inc. (New York), 1976).
5. Peres, A. & Wootters, W. K. Optimal detection of quantum information. *Phys. Rev. Lett.* **66**, 1119–1122 (1991).
6. Acín, A., Bagan, E., Baig, M., Masanes, L. & Muñoz Tapia, R. Multiple-copy two-state discrimination with individual measurements. *Phys. Rev. A* **71**, 032338 (2005).
7. Chitambar, E. & Hsieh, M.-H. A return to the Optimal Detection of Quantum Information. arxiv:1304.1555 (2013).
8. Ivanovic, I. D. How to differentiate between non-orthogonal states. *Phys. Lett. A* **123**, 257–259 (1987).
9. Dieks, D. Overlap and distinguishability of quantum states. *Phys. Lett. A* **126**, 303–306 (1988).
10. Peres, A. How to differentiate between non-orthogonal states. *Phys. Lett. A* **128**, 19–19 (1988).
11. Jaeger, G. & Shimony, A. Optimal distinction between two non-orthogonal quantum states. *Phys. Lett. A* **197**, 83–87 (1995).
12. Croke, S., Andersson, E., Barnett, S. M., Gilson, C. R. & Jeffers, J. Maximum confidence quantum measurements. *Phys. Rev. Lett.* **96**, 070401 (2006).
13. Fannes, M., Nachtergaele, B. & Werner, R. Finitely correlated states on quantum spin chains. *Commun. Math. Phys.* **144**, 443–490 (1992).
14. Verstraete, F., Murg, V. & Cirac, J. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.* **57**, 143–224 (2008).
15. Cramer, M. *et al.* Efficient quantum state tomography. *Nat. Commun.* **1**, 149 (2010).
16. Bacon, D., Chuang, I. L. & Harrow, A. W. Efficient Quantum Circuits for Schur and Clebsch-Gordan Transforms. *Phys. Rev. Lett.* **97**, 170502 (2006).
17. Harrow, A. W. *Applications of coherent classical communication and the Schur transform to quantum information theory.* Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2005 (2005).
18. Georgi, H. *Lie algebras in particle physics* (Westview Press, 1999).
19. Koashi, M. & Imoto, N. Quantum information is incompressible without errors. *Phys. Rev. Lett.* **89**, 097904 (2002).

20. Bartlett, S. D., Rudolph, T. & Spekkens, R. W. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.* **79**, 555–609 (2007).
21. Dowling, J. P. Quantum optical metrology – The lowdown on high-N00N states. *Contemp. Phys.* **49**, 125–143 (2008).
22. Schön, C., Solano, E., Verstraete, F., Cirac, J. I. & Wolf, M. M. Sequential generation of entangled multiqubit states. *Phys. Rev. Lett.* **95**, 110503 (2005).
23. Bañuls, M. C., Pérez-García, D., Wolf, M. M., Verstraete, F. & Cirac, J. I. Sequentially generated states for the study of two-dimensional systems. *Phys. Rev. A* **77**, 052306 (2008).

Acknowledgments

Author contributions

R.B.K., S.C. and M.Z. contributed equally to this work.

Additional information