The power of quantum sampling

Aram Harrow

Bristol

NEC-Rutgers
25 March, 2010

Classical random sampling

"Useful with high probability"

- Possibly helpful in computational complexity

  e.g. polynomial identity checking, approximating the permanent and other hard-to-compute quantities

- Provably helpful in communication complexity

  e.g. testing of two n-bit strings requires n bits deterministically, $O(\log n)$ bits with randomness or $O(1)$ bits with shared randomness.

- Provably helpful in query complexity

  e.g. estimating averages or the volume of convex bodies. The key tool is always sampling.

- Absolutely necessary for cryptography

  Without (subjective) randomness there are no secrets.

What it means to sample on a quantum computer

There is no canonical answer. Let $p \in \mathbb{R}^N$ be a probability distribution. Here are three possibilities, in order of increasing strength.

	Model	Cost of uniformity testing
1	The ability to prepare an $N \times N$ density matrix ρ with $\operatorname{spec} \rho = p$	$\Theta(N)$
2	The existence of an efficient classical circuit that can sample from p.	$\sqrt{N}$ classically, $N^{1/3}$ quantumly
3	The ability to prepare $\sum_{i=1}^{N} \sqrt{p_i} \, \|i\rangle$.	$O(1)$

When are these possible?

Scenario 1: Preparing ρ with $\operatorname{spec}\rho = p$.

- Can be obtained via "Quantum Metropolis Sampling" [Temme *et al.*, 0911.3635], which prepares $\rho \propto e^{-\beta H}$ for $\beta \geq 0$ and H a sum of terms each involving $O(1)$ qubits.
- Information-theoretic tasks often work in this setting; e.g. compressing $\rho^{\otimes n}$ for unknown ρ requires only knowing the Shannon entropy of $p = \operatorname{spec}\rho$.
- This ability is weaker than the ability to sample from p.

Scenario 2: p is efficiently classically samplable

- By definition, quantum criteria are same as classical.
- Being able to sample from p using a quantum circuit is *not* necessarily as strong, since classical circuits can be written as deterministic functions of a random seed.

When are these possible?

Scenario 3: preparing $\sum_i \sqrt{p_i}\,|i\rangle$.

- Also called "q-sampling" p.
- Of the three scenarios, this is the most powerful.
- Can be prepared using an oracle that maps $j \rightarrow \sum_{i\leq j} p_i$.
- Using [Aharonov and Ta-Shma; STOC '03], we can generate this state if
  1. There is a sequence of Markov chains $M^{(0)}, \ldots, M^{(t)}$ with stationary states $p^{(0)}, \ldots, p^{(t)} = p$.
  2. $p^{(0)}$ is easy to q-sample.
  3. Each $M^{(t)}$ has gap $\geq 1/\operatorname{poly}(n)$ and each $p^{(t)}$ has $\geq 1/\operatorname{poly}(n)$ overlap with $p^{(t+1)}$.
  4. For each t, i, j we can efficiently compute $p_i^{(t)}/p_j^{(t)}$.
- If $f : [N] \rightarrow [M]$ is efficiently computable, then we can efficiently create

$$\frac{\sum_{i\in f^{(-1)}(x)} |i\rangle}{\sqrt{|f^{(-1)}(x)|}},$$

for a random choice of x.

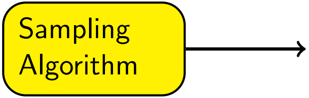
Scenario 1: Can create ρ

Measuring N^2 observables can estimate ρ.
Unfortunately, this requires $\Omega(N^2)$ copies in general.
Quantum birthday problem: In [Childs, Harrow, Wocjan; STACS '07], we considered the problem of distinguishing the case $\rho = I/N$ from the case $\operatorname{rank} \rho \leq N/2$.

- If the eigenbasis of ρ is known, then this is the classical birthday problem and $\Theta(\sqrt{N})$ copies are sufficient.
- When the basis is not known, $\Theta(N)$ copies are necessary and sufficient to distinguish these two cases.
- The intuition is that the quantum problem is harder because of the unknown basis. This raises the effective number of degrees of freedom to $\sim N^2$, so that N copies are needed for a "collision."
- The proof uses the Schur basis, a quantum generalization of the classical method of types.

Scenario 2: Defining sampling oracles

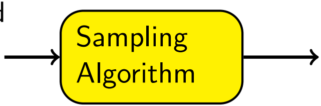
We want to treat the classical algorithm creating samples of p as a black box.



Sampling Algorithm $\longrightarrow i$

with probability p_i

However, this model is too restrictive.



Random seed

$r \in \{0,1\}^m \longrightarrow$ Sampling Algorithm $\longrightarrow i = f(r)$ where $p_i = \frac{|f^{-1}(i)|}{2^m}$

Our quantum algorithm will make use of oracle access to f.

Scenario 2: Classical results

The symmetric case was mostly solved by [Valiant, STOC '08].

Canonical tester

1. Draw M samples according to p.
2. Suppose that item i appears $s(i)$ times.
3. If $s(i) \geq \theta$, then estimate $\hat{p}_i = s(i)/M$. Otherwise, consider the range $\hat{p}_i \in [0, \frac{\theta}{M}]$.
4. Hope that $\hat{p}$ gives an unambiguous answer.

Applications

- Estimating trace distance in general requires $N^{1-o(1)}$ samples.
- Determining whether $p = q$ or $\frac{1}{2}\|p - q\|_1 \geq \epsilon$ requires $\Theta(N^{2/3})$ samples.

Scenario 2: Previous quantum results

- The first example of quantum advantage is Grover's 1996 search algorithm. (Proved optimal by BBBV in 1994!)
- For any subset $S \subset [N]$, let $\pi = \sum_{i \in S} p_i$. Grover's algorithm can determine whether $\pi = 0$ or $\pi \geq \theta$ in time $O(1/\sqrt{\theta})$.
- This can be used distinguish 1-1 functions from 2-1 functions in time $O(N^{1/3})$. [Brassard, Høyer, Tapp; quant-ph/9705002]
- More generally, we can output $\pi \pm O(\epsilon)$ in time $O(\sqrt{\pi}/\epsilon)$.
  [Brassard, Høyer, Mosca, Tapp; quant-ph/0005055]
  Compare with $O(\pi/\epsilon^2)$ for classical sampling.
- [Aaronson and Ambainis; arXiv:0911.0996] prove that for symmetric problems, any Q-query quantum algorithm can be turned into an $O(Q^9)$-query randomized algorithm.

Scenario 2: our results

[Bravyi, Harrow and Hassidim; STACS '10] Given two distributions p, q and constants $0 < \epsilon \leq \theta \leq 1$ we consider three problems.

Goal	to distinguish	
Uniformity testing	$p = u$	$\frac{1}{2}\|p - u\|_1 \geq \epsilon$
Statistical distance	$\frac{1}{2}\|p - q\|_1 \leq \theta - \epsilon$	$\frac{1}{2}\|p - q\|_1 \geq \theta$
Orthogonality	$\frac{1}{2}\|p - q\|_1 \leq 1 - \epsilon$	$\frac{1}{2}\|p - q\|_1 = 1$

(u denotes the uniform distribution on $[N]$.)

Results:

Goal	Classical	Quantum
Uniformity testing	$\Theta(N^{1/2})$	$\Theta(N^{1/3})$
Statistical distance	$N^{1-o(1)}$	$O(N^{1/2})$
Orthogonality	$\Theta(N^{1/2})$	$\Theta(N^{1/3})$

(Uniformity lower bound from [Chakraborty *et. al*, unpublished].)

Distribution testing protocols

Algorithm for statistical distance

Consider the r.v. X which equals $\frac{|p_i - q_i|}{p_i + q_i}$ with probability $\frac{1}{2}(p_i + q_i)$.

- $\mathbb{E}(X) = \frac{1}{2}\|p - q\|_1$
- $\text{Var}(X) \leq \mathbb{E}(X^2) \leq 1$
- Estimating X to constant multiplicative accuracy requires $O(\sqrt{N/\delta})$ queries when $\max(p_i, q_i) \geq \delta/2N$.
  This happens with probability $\geq 1 - \delta$.
- Therefore $O(\sqrt{N})$ queries suffice.

Algorithm for fidelity: $\sum_{i=1}^N \sqrt{p_i q_i}$

Now let X equal $\sqrt{p_i/q_i}$ with probability q_i.

- $\mathbb{E}(X) = \sum_{i=1}^N \sqrt{p_i q_i}$ $\qquad$ $\text{Var}(X) \leq \mathbb{E}(X^2) = \sum_{i=1}^N p_i = 1$
- Again $O(\sqrt{N})$ queries suffice.

Distribution testing protocols

Algorithm for uniformity testing

- Take $M \sim N^{1/3}$.
- Sample $S = \{i_1, \ldots, i_M\}$ according to p.
- If there is a collision, then output "not uniform."
- Let $p_S = p_{i_1} + \ldots + p_{i_M}$.
- Estimate p_S to constant accuracy and output "uniform" iff $p_S \approx M/N$.

Algorithm for orthogonality testing

- Take $M \sim N^{1/3}$.
- Sample $S = (i_1, \ldots, i_M)$ according to p, ignoring duplicates.
- Estimate $q_S = q_{i_1} + \ldots + q_{i_M}$ and output "orthogonal" iff the estimate is 0.

Scenario 2: Discussion

Unlike the classical “canonical tester,” there are many different quantum approaches.

It is unknown whether there is a general framework that encompasses all optimal quantum algorithms for testing symmetric properties of distributions.

The only general purpose lower bound is the Aaronson-Ambainis result. It is probably not tight, and does not suggest a canonical algorithm.

One other subtlety: what if we have a quantum algorithm that can generate samples according to p? Now there is no seed, but the subroutine to estimate probabilities still works. Is this ever a weaker primitive?

Scenario 3: *q*-sampling

The ability to prepare $\sum_{i=1}^{N} \sqrt{p_i} \left| i \right\rangle$ is a potentially much more powerful form of sampling.

Swap test

Given $\sum_{i=1}^{N} \sqrt{p_i} \left| i \right\rangle$ and $\sum_{i=1}^{N} \sqrt{q_i} \left| i \right\rangle$, the swap test accepts with probability

$$\frac{1 + \left(\sum_{i=1}^{N} \sqrt{p_i q_i} \right)^2}{2}.$$

Quantities such as fidelity and statistical distance can be estimated with $O(1)$ samples.

q-samples and Markov chains

If M is a gapped Markov chain with stationary distribution p, then we can efficiently distinguish $\sum_i \sqrt{p_i}\,|i\rangle$ from any orthogonal state.

This is used by [Lutomirski *et al.*, 0912.3825] to verify a particular candidate for quantum money.

Similarly, quantum expanders can be used to verify arbitrarily large maximally entangled states. Define $|\Phi_D\rangle = \frac{1}{\sqrt{D}}\sum_{i=1}^{D}|i\rangle \otimes |i\rangle$.

Let $U_1, \ldots, U_k$ be a D-dimensional quantum expander. Then Alice and Bob can verify that they share $|\Phi_D\rangle$ by

- Alice prepares the control register $\frac{1}{\sqrt{k}}\sum_{i=1}^{k}|i\rangle$.
- Conditioned on the control register she applies U_i to her data.
- She sends the control register to Bob.
- Conditioned on the control register Bob applies $\bar{U}_i$ to his data.
- Bob accepts if the control register is still in the $\frac{1}{\sqrt{k}}\sum_{i=1}^{k}|i\rangle$ state.

Achieves accuracy ϵ using $O(\log 1/\epsilon)$ communication.

Testing productness

Problem: p is a distribution on $\overbrace{[d] \times \ldots \times [d]}^{n}$. Determine whether p is a product distribution or is far from any product distribution.

Classical version: Can be achieved with $O(d^{n/2})$ samples by a relation to uniformity testing.

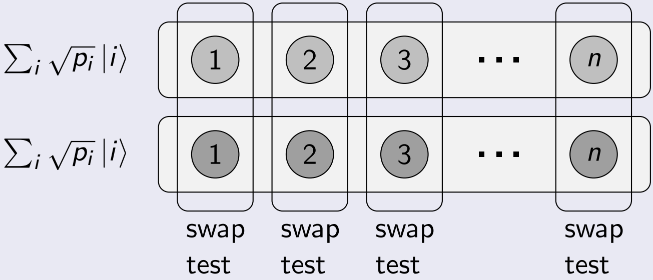
Using q-samples:

- One sample gives no information.
- Two samples are enough for constant accuracy [Harrow and Montanaro; 1001.0017].

Let $1 - \epsilon$ be the maximum fidelity of $\sum_{i=1}^{N} \sqrt{p_i} \, |i\rangle$ with any product state. Then our test outputs ACCEPT with probability $1 - \Theta(\epsilon)$.

Testing productness

Product test algorithm



$\sum_i \sqrt{p_i}\,|i\rangle$

1 2 3 $\cdots$ n

$\sum_i \sqrt{p_i}\,|i\rangle$

1 2 3 $\cdots$ n

swap test swap test swap test swap test

Accept iff all n swap tests pass.

Why it works. Applied to $\rho \otimes \rho$, the swap test accepts with probability $(1 + \operatorname{tr} \rho^2)/2$. If $\sum_i \sqrt{p_i}\,|i\rangle$ is entangled, some of its subsystems must be mixed and so some swap tests are likely to fail.

Application of product test

QMA(*k*)

$QMA(k)_{c,s}$ is the class of Quantum Merlin-Arthur proofs with k unentangled Merlins. For $L \in QMA(k)_{c,s}$, there is a quantum poly-time verifier V such that

- If $x \in L$, then there exists a proof $|\psi_1\rangle \otimes \ldots \otimes |\psi_k\rangle$ such that V accepts $x, |\psi_1\rangle \otimes \ldots \otimes |\psi_k\rangle$ with probability $\geq c$.
- If $x \notin L$, then *for any* purported proof $|\psi_1\rangle \otimes \ldots \otimes |\psi_k\rangle$, V accepts $x, |\psi_1\rangle \otimes \ldots \otimes |\psi_k\rangle$ with probability $\leq s$.

Amplification

- $QMA(1)_{c,s}$ can be amplified, so $c - s > 1/\operatorname{poly}(n)$ is just as strong as $c = 1 - 2^{-\operatorname{poly}(n)}$, $s = 2^{-\operatorname{poly}(n)}$.
- Amplification was not known to work for $QMA(k)_{c,s}$ when $k \geq 2$. Indeed, perfect parallel repetition fails because $S_\infty^{\min}$ is not additive.

Application of product test

Our result implies:

- Two provers can simulate k provers: $QMA(k)_{c,1-\epsilon} \subset QMA(2)_{c,1-\Theta(\epsilon)}$.
- $QMA(2)$ can be amplified: If $c - s > 1/\operatorname{poly}(n)$ then $QMA(2)_{c,s} \subset QMA(2)_{1-2^{-\operatorname{poly}(n)},0.98}$.
- Short $QMA(2)$ proofs for 3-SAT: [based on Aaronson *et al.*, 0804.0802]
  - Two provers.
  - Perfect completeness ($c = 1$).
  - Constant soundness ($s = 0.98$).
  - Quantum proofs each using $\sqrt{n} \cdot \operatorname{poly}\log(n)$ qubits.

  Scaling down, we have $NP_{\log^2} \subset QMA_{\log}(2)_{0.98,1}$.
- Hardness of approximation This implies that the following problems are $NP_{\log^2}$-hard to approximate to within a constant factor: separability, $S_\alpha^{\min}$ and ground-state energy density of mean-field Hamiltonians.

Open questions

Scenario 1: spec $\rho = p$

- Optimal state tomography (i.e. how to estimate ρ) is still not known, although the Schur basis helps.
- Are there algorithmic applications of e.g. the quantum birthday problem?
- Determine performance of the quantum Metropolis algorithm.

Scenario 2: classical sampling possible

- Is there a quantum "canonical tester"?
- Is the ability to sample with a quantum algorithm ever weaker than the ability to sample with a classical algorithm?

Open questions

Scenario 3: q-sampling

- We can create a quantum state quickly with the assistance of an oracle. Can we do the same for a unitary? [Berry and Childs, 0910.4157] presents one difficulty.
- Analyze the product state tester for states that are very far from any product state. This may require new ideas that are more interesting than the original tester, such as an exponential de Finetti theorem for product states, or a proof that weak additivity holds.
- Can q-sampling be of use in other quantum proof systems?
- How strong is $QMA_{\text{log}}(2)_{\text{const},1}$?
- In general there should be more scope for exponential speedups using tools associated with q-sampling.