Algorithms for a Scalable Quantum Computer

Optimal Pulse Sequences, Montgomery Factoring, and Bayesian Inference

Guang Hao Low, Rich Rines, Theodore Yoder, Isaac Chuang

Introduction

Quantum computers promise speedups to a whole host of classical algorithms, from exponential advantage in factoring to a plethora of polynomial advantages in machine learning tasks. However, as the size of the problem grows, the quantum circuit required to solve it also grows, and small errors in single gates add to have a substantial effect on the output. Scaling up a quantum computer therefore requires either simplifying the circuits or reducing errors. To that end, we present open-loop error correction for single-qubit gates in the form of pulse sequences of optimal length. We also show how to reduce the circuit for Shor's algorithm using the Montgomery product from number theory, bringing the factoring of 35 within experimental grasp. Finally, we demonstrate how a Bayesian inference task can be performed with polynomial advantage on a scalable quantum computer.

Optimal Pulse Sequences

Say that we wish to perform a single-qubit rotation $R_\phi[\theta]$ by angle θ around the axis $\sigma_\phi = \hat{X}\cos\phi + \hat{Y}\sin\phi$ but we only have access to a faulty operation $M_\phi[\theta] = R_\phi[\theta(1+\epsilon)]$ which over- or under-rotates by a factor ϵ. In the presence of these amplitude errors, how can we still perform $R_\phi[\theta]$ accurately?

The answer comes from the non-commutativity of single-qubit operations which allows us to chain faulty pulses in sequence such that

$$U_0[\theta] = \underbrace{\prod_{j=1}^{L} M_{\varphi_j}[\theta_0]}_{\equiv \mathcal{S}(\vec{\varphi})} \cdot M_0[\theta] = R_0[\theta] + \mathcal{O}(\epsilon^{n+1}). \quad (1)$$

The dependence of L on n that is implicit in this equation is crucial. We prove that $L > n$ and argue that $L = 2n$ is achievable. This is in contrast to the next best sequences for arbitrary $\gamma \equiv \theta/2\pi$, which scale as $L = \mathcal{O}(n^{3.09})$[1].

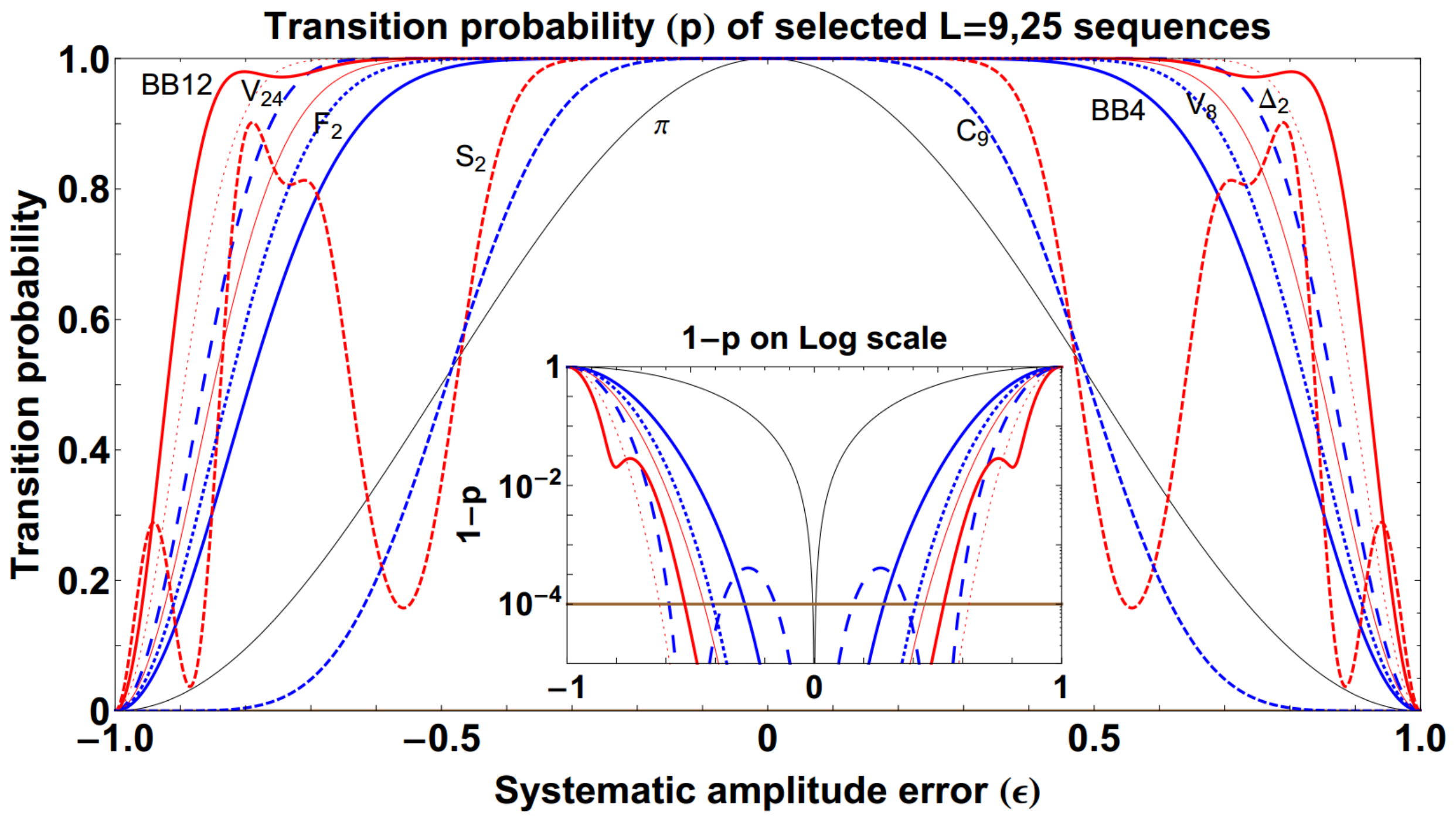


Figure 1: We compare the transition probability $p = |\langle 0|U_0[\pi]|1\rangle|^2$ effected by pulse sequences with amplitude error ϵ in the cases of length $L = 9$ (blue) and also $L = 25$ (red). Included are $\mathrm{BB}4{,}12$ of this work, Vitanov's $V_{8,24}$, Shaka's Δ_2, Cho's C_9, Husain's F_2, and Tycko's S_2.

Our sequences are obtained algebraically by expanding (1). We make the assumption $\theta_0 = 2\pi$ to dramatically simplify the result. The resulting n complex equations to be solved for $\vec{\varphi}$ are

$$\Phi_L^j(\vec{\varphi}) = f_L^j(\gamma), \quad j = 1, \ldots, n \quad (2)$$

where

$$\Phi_L^j(\vec{\varphi}) = \sum_{\{h_k\}}{}' \exp\left(-i\sum_{k=1}^{j}(-1)^k \varphi_{h_k}\right), \; f_L^j(\gamma) = \sum_{k=0}^{j}(-1)^k \binom{T}{k}\binom{L-T}{j-k}.$$

The primed sum enforces $1 \le h_1 < \cdots < h_j \le L$ and $T = (\gamma + L)/2$.

Bayesian Inference

Say that we have a probability distribution $P(\vec{x})$ on n binary variables. After observing some set of evidence variables $\mathcal{E} = e$ how can we sample from $P(\mathcal{Q}|e)$ for a set of query variables? Such sampling allows us to perform approximate inference, inferring the most likely values of $\mathcal{Q}$ given e by evaluating $\mathrm{argmax}_{\mathcal{Q}} P(\mathcal{Q}|e)$.

$$P(\vec{x}) = P(x_1)P(x_2|x_1)P(x_3|x_1)P(x_4|x_1,x_3) \quad (3)$$

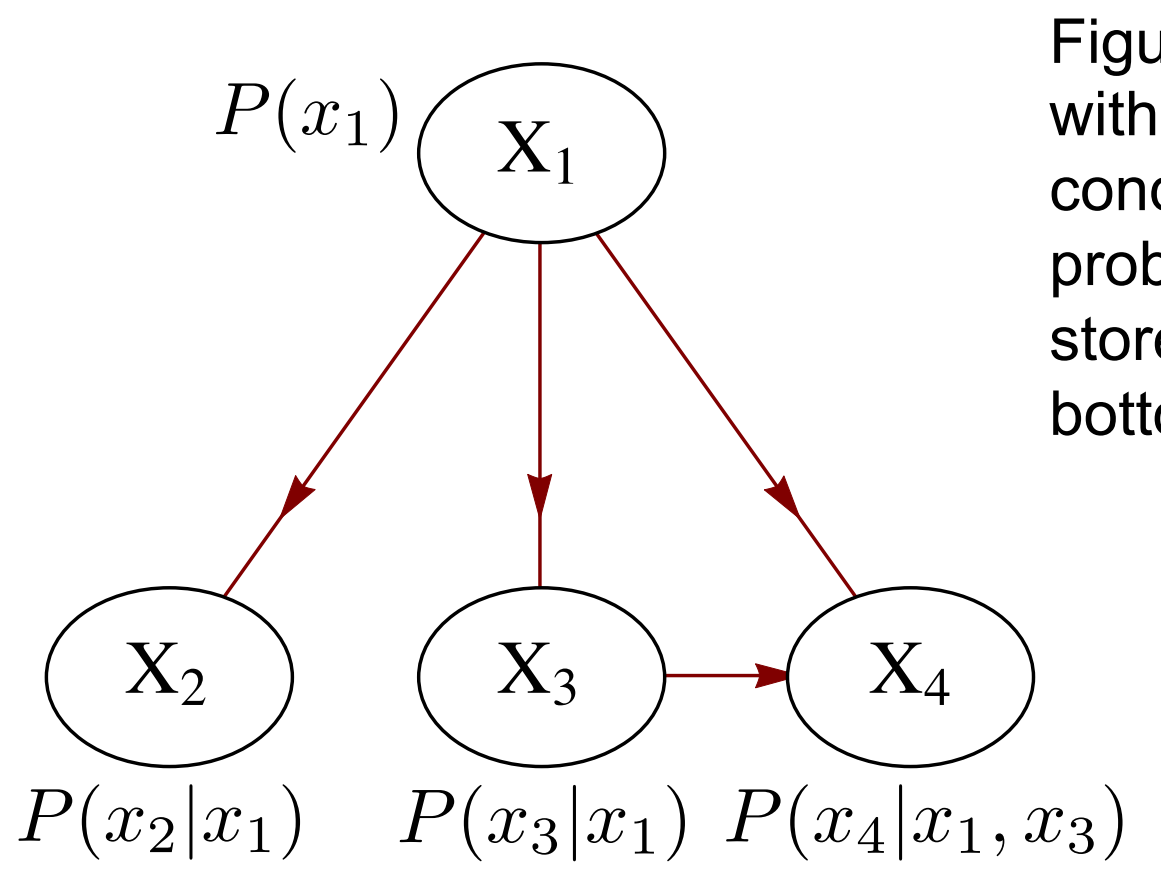
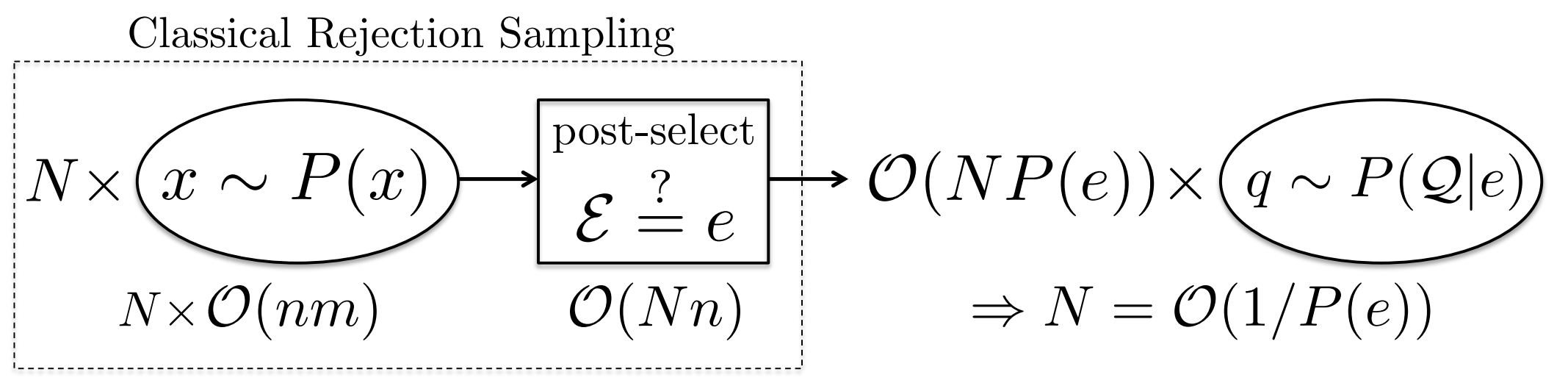
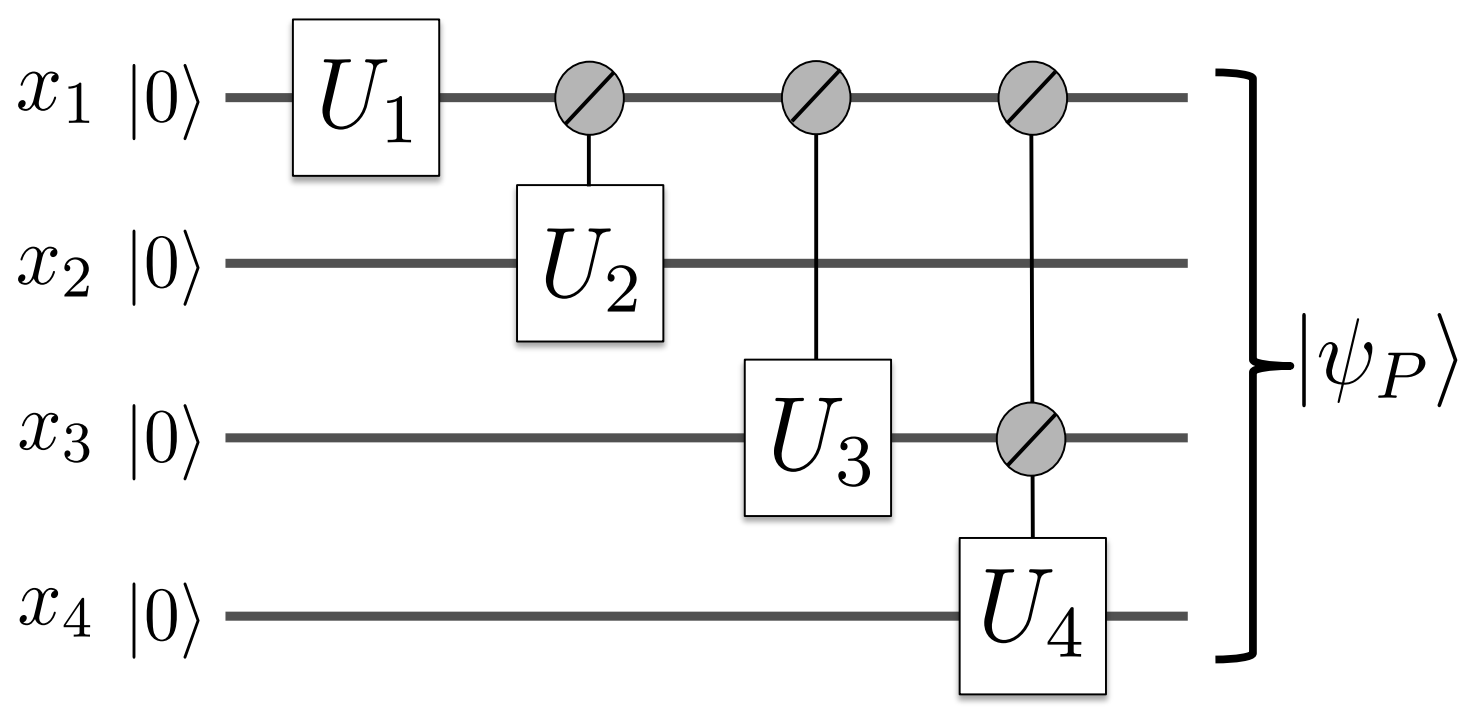


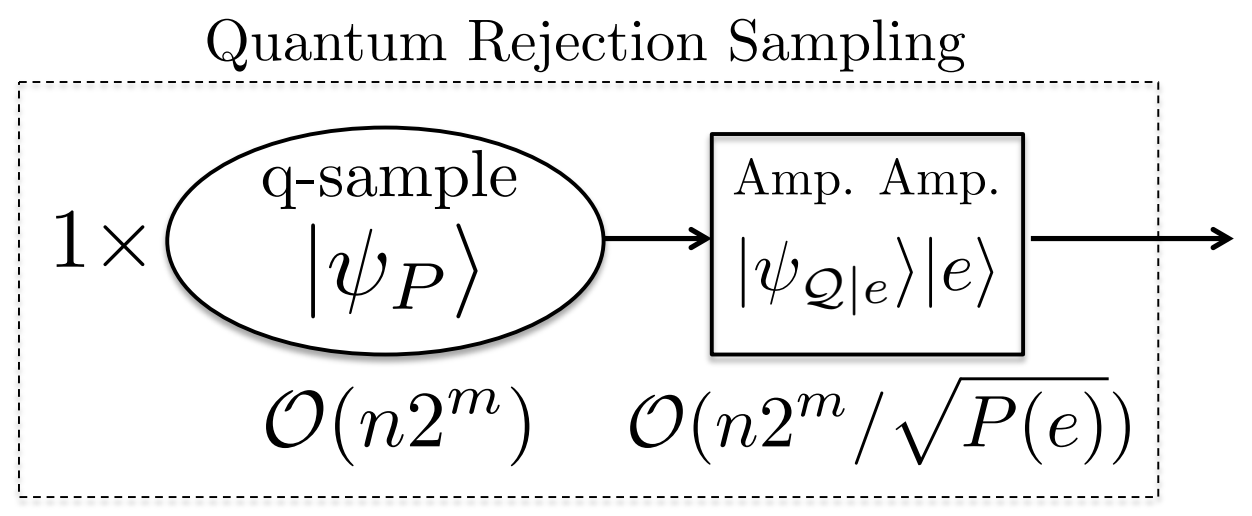
Figure 2: (Left) A Bayes net is a directed acyclic graph with n nodes and maximum indegree m showing the conditional dependences of random variables of a probability distribution $P(\vec{x})$ (Eq.(3)). With each node is stored a table of probabilities. Sampling nodes top to bottom, gives a sample of $P(\vec{x})$ in time $\mathcal{O}(nm)$.
(Below) Classical rejection sampling takes N samples from $P(\vec{x})$ and rejects those with incorrect evidence $\mathcal{E} \neq e$. Since the correct evidence only appears with probability $P(e)$, if we desire one sample from the posterior $P(\mathcal{Q}|e)$ we must set $N = \mathcal{O}(1/P(e))$ and take $\mathcal{O}(nm/P(e))$ time.


Classical Rejection Sampling
$N\times$
$x \sim P(x)$
post-select
$\mathcal{E} \overset{?}{=} e$
$\mathcal{O}(NP(e))\times$
$q \sim P(\mathcal{Q}|e)$
$N\times\mathcal{O}(nm)$
$\mathcal{O}(Nn)$
$\Rightarrow N = \mathcal{O}(1/P(e))$

q-sample state [3]:
$$|\psi_P\rangle = \sum_{x_1,\ldots,x_n} \sqrt{P(x_1,\ldots,x_n)}|x_1,\ldots,x_n\rangle \quad (4)$$

Figure 3: (Right) This circuit prepares the q-sample $|\psi_P\rangle$ based on the structure of the Bayes net in Fig. 2. Once compiled, the circuit complexity is $\mathcal{O}(n2^m)$. (Below) The quantum rejection sampling algorithm produces one sample of $P(\mathcal{Q}|e)$ per q-sample prepared. It does so by enhancing the component of $|\psi_P\rangle$ that has the correct evidence $|e\rangle$ using amplitude amplification. This effectively speeds up the classical post-selection by a square-root.




Quantum Rejection Sampling
$1\times$
q-sample
$|\psi_P\rangle$
Amp. Amp.
$|\psi_{\mathcal{Q}|e}\rangle|e\rangle$
$1\times$
$q \sim P(\mathcal{Q}|e)$
$\mathcal{O}(n2^m)$
$\mathcal{O}(n2^m/\sqrt{P(e)})$

Solving for Sequences

The symmetries $\vec{\varphi} = \vec{\varphi}_R$ ($\vec{\varphi} = -\vec{\varphi}_R$) result in $\mathrm{Im}\Phi_L^j(\vec{\varphi}) = 0$ for j even (odd) and the resulting sequences are called initialed PD (AP). All such sequences have $L = 2n$.

We introduce three techniques for solving (2):

- Analytical: The substitution $t_k = \tan(\phi_k/2)$ converts (2) into a system of polynomial equations, which can be solved by Groebner bases to yield $\mathrm{AP}1,2,3$ and $\mathrm{PD}2,4$.
- Perturbative: Given a solution at γ_0 and nonzero Jacobian $\frac{\partial}{\partial\varphi_k}\Phi_L^j|_{\gamma_0}$ solutions $\varphi(\gamma)$ for $\gamma \approx \gamma_0$ can be obtained. For a certain AP sequence ToP, $\varphi_k^{\mathrm{ToP}}(0) = \pi$ and we prove a nonzero Jacobian at 0.
- Numerical: We use Mathematica [2] to find roots to equation (2) up to $n = 12$.

Name	Length	Notes
SCROFULOUS	3	$n=1$, non-uniform θ_j [9]
Pn, Bn	$\mathcal{O}(e^{n^2})$	Closed-form [5]
SKn	$\mathcal{O}(n^3)$	$n \le 30$, numerical [5] $n > 30$, conjectured
APn (PDn)	$2n$	$n \le 3(4)$, closed-form $n \le 12$, analytic continuation $n > 12$, conjectured
ToPn	$2n$	arbitrary n, perturbative

Table I: Summary of existing pulse sequences and our optimal sequences AP, PD, and ToP.

Montgomery Factoring

Repeated application of the controlled modular product operation,

$$|x\rangle \longrightarrow |a^{2^k} x \bmod N\rangle, \quad (5)$$

represents the computational bottleneck of Shor's algorithm. Montgomery reduction is a commonly used classical computational technique for efficient modular exponentiation under a large modulus, reducing the algorithmic complexity of modular reduction by division to that of multiplication:

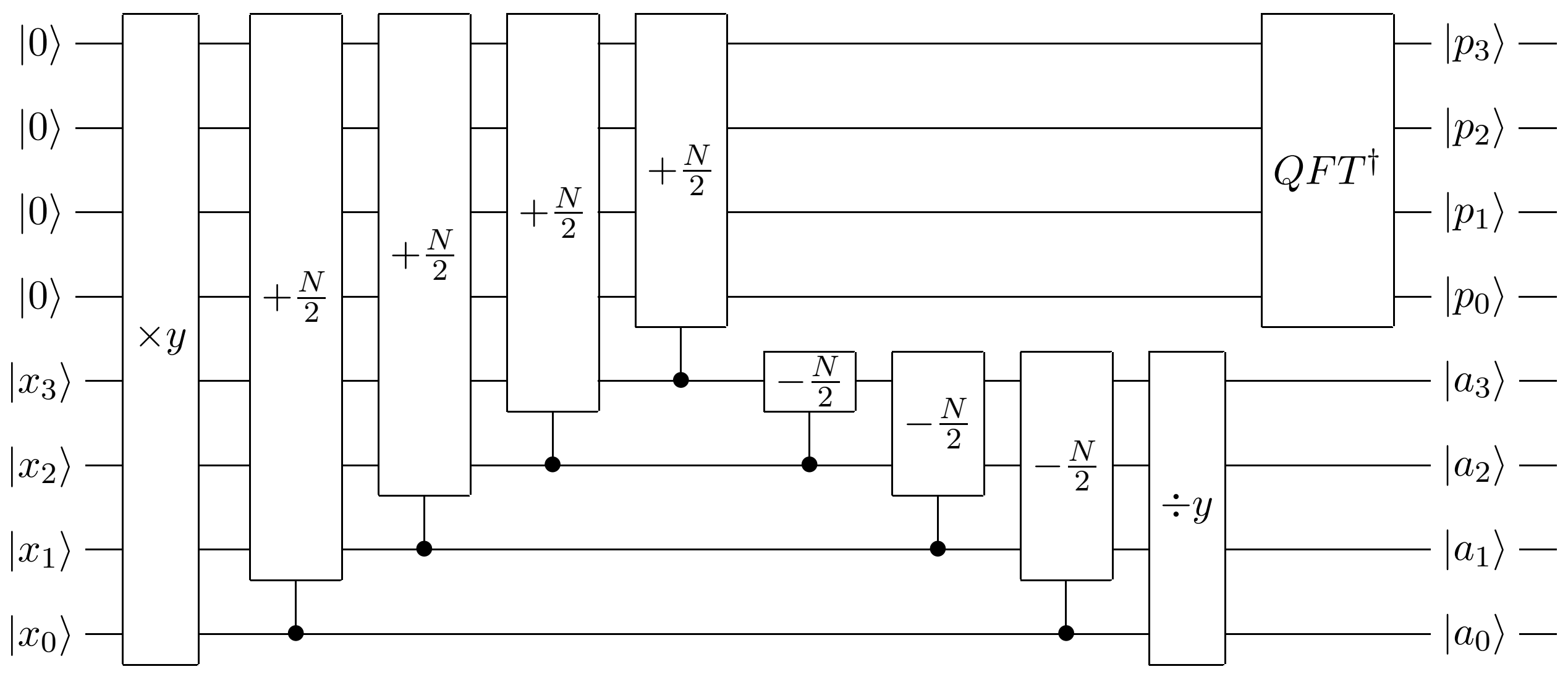
$$\begin{gathered}\text{Given: } T = xy < NR \\ N' := -N^{-1} \bmod R \\ U := TN' \bmod R \\ (T+UN)/R \equiv TR^{-1} \pmod{N} \\ (T+UN)/R < 2N\end{gathered}$$

The classical advantage of Montgomery reduction results from the ability to implicitly calculate U with no computational overhead. Unfortunately, this implicit calculation is not immediately reversible, mitigating the advantage of Montgomery reduction when naively adapted to quantum modular exponentiation.

Remarkably, we can reclaim the algorithmic advantage by adapting to a uniquely quantum mechanical arithmetic model acting in quantum Fourier space. The resulting multiplier provides a unique scalable building block for constructing Shor's circuits:

Multiplication Method	Depth	Gates	Qubits
Fourier Montgomery	$\mathcal{O}(n)$	$12n^2+3n+6$	$2n+2$
Shift-and-Add [4]	$\mathcal{O}(n^2)$	$24n^2+\mathcal{O}(n)$	$2n+\mathcal{O}(1)$
Fourier Shift-and-Add, Draper [5]	$\mathcal{O}(n^2)$	$2n^3+\mathcal{O}(n^2)$	$3n$
Fourier Shift-and-Add, Beauregard [6]	$\mathcal{O}(n^3)$	$2n^3+\mathcal{O}(n^2)$	$2n+2$

Small circuit width and small lower-order terms in the circuit depth and gate count make the Fourier Montgomery method immediately applicable to foreseeable factoring experiments, such as N = 35.



Bibliography

1. Ken Brown, Aram Harrow, Isaac Chuang, Phys. Rev. A **70**, 052318 (2004).
2. Wolfram Research Inc., *Mathematica*, Version 9.0, Champaign, IL (2012).
3. Dorit Aharonov, Amnon Ta-Shma, STOC '03, ACM, pp. 20-29 (2003).
4. Christof Zalka, arXiv:quant-ph/9806084 (1998)
5. Thomas Draper, arXiv:quant-ph/0008033 (2000)
6. Stephane Beauregard, arXiv:quant-ph/0205095 (2003)

The authors acknowledge the support of the NSF CUA, NSF iQuISE IGERT, IARPA QCS ORAQL, and NSF CCF-RQCC projects.