Quantum Information and Computation
Homework 2

(Report Due: May 14, 2022)

1. **(15 points) Period-finding algorithm.**
Consider the function $f(x) = 5^x \bmod 39$ on the domain $x \in \mathbb{Z}_{2^t}$ with say $t = 11$. The reason of choosing such $a = 5$ and $N = 39$ is that we might want to factor the number $N = 39$ as in Shor's algorithm. We randomly picked up an number $1 < a < N$ and find $a = 5$. Moreover, using an *extended Euclidean algorithm* would efficiently (taking time $O(t^2)$) to see that $a = 5$ is coprime to $N = 39$.

   (a) **(2 points)** Show that f is periodic and determine its period r (e.g. by using a calculator.)

   (b) **(2 points)** Suppose we construct the equal superposition state $|f\rangle$ of $(x, f(x))$ values over the domain $\mathbb{Z}_{2^t}$, measure the second register. What are the possible resulting measurement values z and their probabilities?

   (c) **(2 points)** Suppose the measurement outcome was $z = 8$. Find the resulting state $|\alpha\rangle$ of the first register after the measurement.

   (d) **(2 points)** Perform the quantum Fourier transform mod 2^t on $|\alpha\rangle$, and finally measure it (in the computational basis). What is the probability for each possible outcome $0 \leq y < 2^t$?

   (e) **(5 points)** How to use the above measurement outcome to determine the period r of the function f? What is the probability that we successfully determine r from this measurement result, using the standard process of the quantum period finding algorithm?

   (f) **(2 points)** Use the obtained period r to find a factor of the target integer $N = 39$.

2. **(15 points) Shor's factoring algorithm, continued fractions.**
Suppose we wish to factor $N = 21$ using Shor's algorithm and we have chosen $a = 2$ so we aim to determine the period of $f(x) = 2^x \bmod 21$. We proceed through the quantum algorithm and finally measure the x registers obtaining measurement result $y = 427$.

   (a) **(0 point)** What is the number t of qubits that is used for the x register?

   (b) **(10 points)** Use the continued fraction method to find a fraction ℓ/r with denominator less than 21, that is within $1/2^{t+1}$ of the ratio $y/2^t$.

   (c) **(5 points)** We hope that the denominator of ℓ/r (when the fraction is canceled down to lowest terms) is the period of $f(x)$. Check to see that it is indeed the period in this example. Use your value of r to find factors of 21 (following the method used in Shor's algorithm).

3. **(30 points) Shor's factoring algorithm for discrete logarithms.**
For any prime p consider the set $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\} \subset \mathbb{Z}_p$ of nonzero integers modulo p, with the operation of *multiplication* mod p. A *generator* for $\mathbb{Z}_p^*$ is an element g whose powers generate

all elements of $\mathbb{Z}_p^*$ i.e. for all $x \in \mathbb{Z}_p^*$, there is a $y \in \mathbb{Z}_{p-1}$ with $x = g^y \mod p$. Then, such y is called the *discrete logarithm* of x (to base g). You may assume that $\mathbb{Z}_p^*$ always has a generator g and that it satisfies $g^{p-1} = 1 \mod p$.

Suppose we are given a generator g and element $x \in \mathbb{Z}_p^*$, and we wish to compute its discrete logarithm y.

(a) (**5 points**) Consider the function $f : \mathbb{Z}_{p-1} \times \mathbb{Z}_{p-1} \to \mathbb{Z}_p^*$ given by

$$f(a,b) = g^a x^{-b} \operatorname{mod} p.$$

For each fixed $c \in \mathbb{Z}_p^*$, show that there is a corresponding fixed $k \in \mathbb{Z}_{p-1}$ such that

$$f(a,b) = c \quad \text{iff} \quad a = by + k \mod (p-1).$$

(b) (**5 points**) Suppose we have constructed the state

$$|\phi\rangle = \frac{1}{(p-1)} \sum_{a,b \in \mathbb{Z}_{p-1}} |a\rangle |b\rangle |f(a,b)\rangle$$

(in Hilbert space $\mathbb{C}^{p-1} \otimes \mathbb{C}^{p-1} \otimes \mathbb{C}^p$, with standard computational basis), and we measure the third register obtaining a result c_0. Find the post-measurement state of the first two registers.

(c) (**10 points**) If we then apply the quantum Fourier transform mod $(p-1)$ to each of these two registers and measure both registers, which output pairs $(c_1, c_2) \in \mathbb{Z}_{p-1} \times \mathbb{Z}_{p-1}$ can be obtained with non-zero probability?

Okay, so far so good, now can y be determined from any such pair? Not exactly. To solve a linear equation in a finite group, we would need some mathematical preliminaries to do the job. We call a *modular multiplicative inverse* of an integer $z \in \mathbb{Z}_p^*$ with respect to the modulus p if x is a solution to

$$xz = zx = 1 \operatorname{mod} p. \tag{1}$$

It is known that

$$z \text{ is coprime to } p \iff z^{-1} := x \text{ satisfying (1) is unique.} \tag{2}$$

Moreover, one can find the multiplicative inverse z^{-1} efficiently by a classical *extended Euclidean algorithm* (with running time in $O(n^2)$ if $z, p < 2^n$).

(d) (**10 points**) Using the results you got from above to outline a quantum algorithm for computing discrete logarithms, that runs in time $O(\text{poly}(\log p))$ for large p, and succeeds with probability $1 - \epsilon$ for any chosen constant $\epsilon > 0$. You may assume that calculating f and implementing $\mathbf{QFT}_{p-1}$ may be implemented in $O(\text{poly}(\log p))$ time.

Remark. While you are doing this problem, you may have sensed why finding a discrete logarithm is hard. Some cryptography systems in use today do not rely on the difficulty of integer factorization, but on finding discrete logarithms in groups (e.g. the ElGamal encryption and the Diffie–Hellman key exchange). groups.

4. **(20 points) POVMs versus PVMs.** Consider a quantum binary hypothesis testing problem: the null hypothesis H_0 is that the quantum system is in the state $|\psi_0\rangle = |0\rangle \in \mathbb{C}^2$, and the alternative hypothesis H_1 is that the quantum system is in the state $|\psi_1\rangle = |+\rangle := (|0\rangle + |1\rangle)/\sqrt{2} \in \mathbb{C}^2$, and we assume that both hypotheses happen equally likely. The goal is to perform a measurement to tell which hypothesis is true. Unfortunately, the states $|0\rangle$ and $|+\rangle$ are not orthogonal, so you cannot perfectly distinguish them (as taught in class).

   Now as in Homework 1, you are allowed to do the *unambiguous discrimination* as follows. You can report one of three possible outcomes, which can be modeled as a quantum measurement $\{\Pi_i\}_{i\in\{0,1,*\}}$: the true state is either $|\psi_0\rangle$, or $|\psi_1\rangle$, or that the measurement outcome is inconclusive (denoted as '$*$'), but no wrong answers is allowed, i.e.

$$\Pr(\text{outcome is } j \,|\, \mathsf{H}_i = |\psi_i\rangle) = 0, \forall i \in \{0,1\}, \forall j \neq i. \tag{3}$$

   We define the success probability of such a measurement scheme to be the probability that you identify the true state:

$$P_{\mathrm{s}} := \sum_{i=0,1} \frac{1}{2} \Pr(\text{outcome is } i \,|\, \mathsf{H}_i = |\psi_i\rangle). \tag{4}$$

   (a) **(5 points)** Show that the success probability using projection-valued measure (PVM) measurements on $\mathbb{C}^2$ is at most $1/4$.

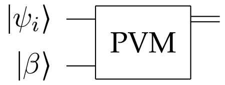
   Here, a PVM $\{\Pi_i\}_{i\in\{0,1,*\}}$ means that Π_i is a projection for each $i \in \{0,1,*\}$ and the completeness relation also holds, i.e. $\sum_{i\in\{0,1,*\}} \Pi_i = I$. Moreover, an important feature of PVMs is that the PVM elements are mutually orthogonal.

   Hint. In this problem, you only access to a qubit system. Then, one of the PVM element Π_i has to be zero projection unfortunately due to the requirement of the mutual orthogonality. Now you may see why the PVM measurements impose stringent conditions on the state discrimination problems.

   (b) **(15 points)** Find a positive operator-valued measure (POVM) measurement on $\mathbb{C}^2$ that achieves a success probability strictly larger than $1/4$.

   (c) **(Bonus 5 points)** How large the success probability using POVM measurements can be (in unambiguous discrimination)?

   (d) **(Bonus 5 points)** Suppose that you have resources to enlarge your working space by appending the state $|\psi_i\rangle$ with some ancilla $|\beta\rangle$ who may live in a very high Hilbert space, and then you may perform PVM measurements (on a larger Hilbert space), i.e.



Can the success probability using PVM measurements outperform 1/4 this time?

How is the PVM measurement on a larger Hilbert space compared with the POVM measurement on the original space? This is closely related to the *Naimark theorem* as we have seen in Stinespring's dilation for purifying a noisy quantum operation.

Now you may have a sense why those corporations such as Google and IBM eager to build large-scale quantum computers (aside from quantum computing tasks).

5. **(20 points) Bloch-vector representation and state tomography.**

Any 2×2 Hermitian matrix lives in a 4-dimensional Euclidean space. Hence, in principle we can use 4 parameters to completely describe a 2×2 Hermitian matrix. If we further require that its trace equal 1, then 3 parameters are sufficient. Now, for any Hermitian matrix ρ with $\mathrm{Tr}[\rho] = 1$, we write

$$\rho = \frac{1}{2}\left(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z\right).$$

Here, $\vec{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$ is called the *Bloch vector* for ρ, and $\sigma_x = X$, $\sigma_y = Y$, and $\sigma_z = Z$ are the usual qubit Pauli matrices.

(a) **(2 points)** Show that $r_x = \mathrm{Tr}[\rho\sigma_x]$, $r_y = \mathrm{Tr}[\rho\sigma_y]$, and $r_z = \mathrm{Tr}[\rho\sigma_z]$.

(b) **(2 points)** Show that ρ is a quantum state (i.e. a density operator) if and only if $\|\vec{r}\|_2 \le 1$. Here $\|\cdot\|_2$ denotes the 2-norm or the Euclidean norm. (Note that ρ here could be a general mixed state not only a pure state.)

(c) **(2 points)** Let σ be another qubit state, with Bloch vector $\vec{s}$. Verify that $\mathrm{Tr}[\rho\sigma] = \frac{1}{2}(1+\vec{r}\cdot\vec{s})$. If ρ is orthogonal to σ (under the Hilbert–Schmidt inner product), what can you visualize on the Bloch sphere?

(d) **(2 points)** Let $\{|\psi_i\rangle\}_{i=0,1}$ denote an orthonormal basis of $\mathbb{C}^2$, and let $\vec{r}_i$ be the Bloch vector of $|\psi_i\rangle\langle\psi_i|$ for each $i \in \{0,1\}$. Show that the probability of obtaining outcome $i \in \{0,1\}$ when measuring ρ using measurement $\{|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|\}$ is given by $\frac{1}{2}(1+\vec{r}\cdot\vec{r}_i)$.

(e) **(4 points)** Now imagine that ρ is an unknown qubit state whose Bloch vector $\vec{r}$ you would like to characterize completely (since if you know $\vec{r}$ you would know ρ mathematically). Consider the following six operators:

$$\Pi_{(a,b)} := \frac{I + (-1)^b\sigma_a}{6},$$
$$a \in \{x,y,z\},$$
$$b \in \{0,1\}.$$

Show that $\{\Pi_{(a,b)}\}$ forms a valid POVM.

(f) **(4 points)** Using the measurement given in 5e, show that the probabilities of measurement outcomes are given by

$$\Pr(a,b) = \frac{1+(-1)^b r_a}{6}.$$

(g) **(4 points)** How can you visualize this formula on the Bloch sphere? Describe how measuring (possibly infinitely) many copies of ρ using the above-mentioned measurement allows for estimating the entries of $\vec{r}$ to arbitrary accuracy. Here, recalling that whenever we measure a state, it collapses, and by many copies we meant the quantum system is prepared independently and identically distributed (i.i.d) in the state ρ. Namely, we are given $\rho \otimes \rho \otimes \cdots$. When we measure, say, the state ρ at the first register, we obtain a measurement outcome. Then, we measure the (identical) state ρ at the second register, and so on so forth to collect enough statistics.

Hint. You may wonder why we would require so many copies of quantum states. We actually require this even in the classical case. Let us imagine we want to determine the state of a coin. That is, to know the probability distribution of flipping to bottom (say outcome 0) and to head (say outcome 1). How would you do? Just flip it *many times*, and calculate the *sample average*, i.e. the number of outcomes 1 (the realizations of the random variable X taking values in 0 and 1), say N_1, dividing by the total number of trials, say N. Then, according to the weak law of large numbers, the fraction $^{N_1}/_N$ converges to the value $\Pr(1)$ in probability. So you would need *many trials* as well. To quantify how fast the convergence is, you need the so-called *concentration inequalities* such as the Chebyshev bound or the Chernoff bound. (Under mild conditions it converges exponentially fast.)

Remark. What if the coin is actually *deterministic* (say just only head), i.e. a pure state? (In the quantum case if you are unlucky to choose the wrong basis, you will never get right answer since you can never use a X basis measurement to tell whether it is $|0\rangle$ or the maximally mixed state.) Well, you will always get 1 in this case. However, even if you get 1 a million times, can you be so sure that it is a deterministic coin? Why can't the tail be a *black swam*? Well, you never know when you only make trials in finite times. But, what's the odd? Even if you are not so sure, the probability of hitting the jackpot is so small (10^{-6}). Ok, you may wonder 你到底看了什麼. I just wanna say, in the world full of uncertainty, you never know everything for sure; but it's good to know that something is unknown. What needs to be worried about is the *unknown unknowns*—if things are *known unknowns*, that's not too bad (非戰之罪). It would be fine if you are aware of the risks (and the chances and your confidence levels, and so on) and get the risks (or errors) in control. That is the very reason why theoretical analysis beforehand matters and hence why this course matters (神展開？).

6. **(Optional) Relating error probability to the trace-norm distance.**
Suppose we have M equiprobable random messages, say $1, 2, \ldots, M$, and our goal is to make an observation to see which message it is. (You can imagine that we are rolling a fair dice with M faces and to see which side it turns out.) We can model this equiprobable random messages by using a random variable, say M, governed by a *uniform distribution* on the set $\{1, 2, \ldots, M\}$. Unfortunately, before observation, the random message is corrupted by some quantum noise, i.e. for each $m \in \{1, 2, \ldots, M\}$, we can only measure on a quantum state, described by a density

operator ρ_B^m on some Hilbert space $\mathcal{H}_B$ which may be viewed as a channel output state, i.e.

$$m \mapsto \mathcal{N}_{\mathsf{M}\to B}(m) =: \rho_B^m. \tag{5}$$

Here, the superscript 'm' means which message m is sent, and the channel $\mathcal{N}_{\mathsf{M}\to B}$ characterizes the quantum noise.

So now we use a POVM, $\{\Pi_B^m\}_{m=1}^M$, to decide which message was sent. Note that here the outcomes of the quantum measurement can be modeled as a random variable $\hat{\mathsf{M}}$. Then, the average probability of erroneous decision of the message M, which can be calculated as follows:

$$\begin{aligned}
\Pr\left\{\hat{\mathsf{M}} \neq \mathsf{M}\right\} &= \sum_{m=1}^M \Pr\{\mathsf{M} = m\} \cdot \Pr\left\{\hat{\mathsf{M}} \neq m \mid \mathsf{M} = m\right\} &(6)\\
&\overset{\text{(a)}}{=} \sum_{m=1}^M \frac{1}{M} \Pr\left\{\hat{\mathsf{M}} \neq m \mid \mathsf{M} = m\right\} &(7)\\
&= \sum_{m=1}^M \frac{1}{M} \left(1 - \Pr\left\{\hat{\mathsf{M}} = m \mid \mathsf{M} = m\right\}\right) &(8)\\
&\overset{\text{(b)}}{=} \sum_{m=1}^M \frac{1}{M} \left(1 - \operatorname{Tr}\left[\rho_B^m \Pi_B^m\right]\right), &(9)
\end{aligned}$$

where (a) follows from the uniform distribution, and (b) follows from Born's rule.

Since we have the freedom to choose the POVM, minimizing (9) will give us the *minimum error probability* of deciding M. We have seen such a research problem before, and it is termed **quantum state discrimination** or **minimum quantum error detection**. We also remark that in the classical scenario (when $\{\rho_B^m\}_m$ mutually commute), the optimal measurement minimizing (9) is given by the *maximum a posteriori* (MAP) decision rule, or the *maximum likelihood* (ML) decision rule with equiprobable priors.

The goal of this exercise is not to derive the minimum error, but to rewrite the (conditional) error probability $\Pr\left\{\hat{\mathsf{M}} \neq m \mid \mathsf{M} = m\right\}$. Let us rewrite the POVM $\{\Pi_B^m\}_{m=1}^M$ as the following quantum measurement procedure on the Hilbert space $\mathcal{H}_B$:

$$\mathcal{M}_{B\to\hat{\mathsf{M}}}(\sigma_B) := \sum_{\hat{m}=1}^M \operatorname{Tr}\left[\sigma_B \Pi_B^{\hat{m}}\right] |\hat{m}\rangle\langle\hat{m}|_{\hat{M}}, \quad \forall \text{ density operator } \sigma_B \text{ on } \mathcal{H}_B. \tag{10}$$

Prove that for each $m \in \{1, 2, \dots, M\}$,

$$\Pr\left\{\hat{\mathsf{M}} \neq m \mid \mathsf{M} = m\right\} = \frac{1}{2} \left\| \mathcal{M}_{B\to\hat{M}}(\rho_B^m) - |m\rangle\langle m|_{\hat{\mathsf{M}}} \right\|_1. \tag{11}$$

Here, $\frac{1}{2}\|\rho - \sigma\|_1 := \frac{1}{2}\operatorname{Tr}\left[|\rho - \sigma|\right]$ stands for the *trace-norm distance* between density operators ρ and σ. (Here, we use the notation $|M| := \sqrt{M^\dagger M}$ for any operator M; moreover, if $M = \sum_i \lambda_i P_i$ admits a spectral decomposition, then $M = \sum_i |\lambda_i| P_i$.)

Remark. This exercise manifests the fact that the conditional error probability is given by the trace distance between the post-measurement state $\mathcal{M}_{B\to\hat{M}}(\rho_B^m)$ an the register $|m\rangle\langle m|_{\hat{\mathsf{M}}}$. This thus gives the trace distance an *operational meaning*.

Let us denote by the joint state $\rho_{\mathsf{M}B} := \sum_{m=1}^{M}\frac{1}{M}|m\rangle\langle m|_{\mathsf{M}} \otimes \rho_B^m$, and $\rho_{\mathsf{M}\hat{\mathsf{M}}} := \sum_{m=1}^{M}\frac{1}{M}|m\rangle\langle m|_{\mathsf{M}} \otimes |m\rangle\langle m|_{\hat{\mathsf{M}}}$. Note here that not only $\rho_{M\hat{\mathsf{M}}}$ is a pure state, but also an *purification* of the equiprobable state ρ_{M}, i.e.

$$\mathrm{Tr}_{\hat{\mathsf{M}}}\left[\rho_{\mathsf{M}\hat{\mathsf{M}}}\right] = \rho_{\mathsf{M}} = \sum_{m=1}^{M}\frac{1}{M}|m\rangle\langle m|_{\mathsf{M}}. \tag{12}$$

What does this mean? The system M (i.e. ρ_{M}) means the original source, while the system $\hat{\mathsf{M}}$ "records the correct answer" at the environment since now systems M and $\hat{\mathsf{M}}$ are perfectly correlated (e.g. if m at system M is sent, the system $\hat{\mathsf{M}}$ will be in the state $|m\rangle\langle m|_{\hat{\mathsf{M}}}$).

Equipped with the above notation and (11), we can rewrite the trace distance by using the direct sum property of the trace-norm as

$$\begin{aligned}
\Pr\left\{\hat{\mathsf{M}} \neq \mathsf{M}\right\} &= \sum_{m=1}^{M} \Pr\left\{\mathsf{M} = m\right\} \Pr\left\{\hat{\mathsf{M}} \neq m \mid \mathsf{M} = m\right\} && (13)\\
&= \sum_{m=1}^{M} \frac{1}{M}\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}(\rho_B^m) - |m\rangle\langle m|_{\hat{\mathsf{M}}}\right\|_1 && (14)\\
&= \frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}}(\rho_{\mathsf{M}B}) - \rho_{\mathsf{M}\hat{\mathsf{M}}}\right\|_1 && (15)\\
&= \frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}} \circ \mathcal{N}_{\mathsf{M}\to B}\left(\rho_{\mathsf{M}\hat{\mathsf{M}}}\right) - \rho_{\mathsf{M}\hat{\mathsf{M}}}\right\|_1. && (16)
\end{aligned}$$

The last equation (16) illustrates the central idea of the **error criterion** used in quantum information theory and quantum error correction.

For example, in addition to the noise/corruption operation $\mathcal{N}_{\mathsf{M}\to B}$ and the decision/measurement operation $\mathcal{M}_{B\to\hat{M}}$, one may introduce an *encoding operation*, say $\mathcal{E}_{\mathsf{M}\to A}$, such that the error criterion becomes

$$\frac{1}{2}\left\|\mathcal{M}_{B\to\hat{M}} \circ \mathcal{N}_{\mathsf{M}\to B} \circ \mathcal{E}_{\mathsf{M}\to A}\left(\rho_{\mathsf{M}\hat{\mathsf{M}}}\right) - \rho_{\mathsf{M}\hat{\mathsf{M}}}\right\|_1. \tag{17}$$

Since we have freedoms to control/design encoding $\mathcal{E}_{\mathsf{M}\to A}$ and decoding $\mathcal{M}_{B\to\hat{M}}$. This is the general aim of quantum error correction and quantum channel coding.

One may further consider "quantum source" instead of classical source (e.g. a classical probability distribution or a diagonal operator $\rho_M = \sum_{m=1}^{M}\frac{1}{M}|m\rangle\langle m|_{\mathsf{M}}$). Now, the source may be a general quantum state (a density operator) ρ_S on some Hilbert space $\mathcal{H}_S$. Denote by $\rho_{SS'}$ to be a purification of ρ_S. Then, the error criterion is now

$$\frac{1}{2}\left\|\mathcal{M}_{B\to S'} \circ \mathcal{N}_{A\to B} \circ \mathcal{E}_{S\to A}\left(\rho_{SS'}\right) - \rho_{SS'}\right\|_1, \tag{18}$$

where the decoding operation $\mathcal{M}_{B\to S'}$ could be a general quantum operation (i.e. completely positive and trace-preserving map), not just a quantum measurement, since now the system S may be quantum.

Hope this exercise provides you a flavor of the general paradigm of the field *quantum information processing*. In your final project, please go ahead to explore more interesting topics and research directions therein.