PROBABILITY LOADING VIA INVERSE TRANSFORMATION METHOD

DEKUAN DONG

Contents

Ref.[9] summarizes several properties that a good initialization method should have:

- being efficient, i.e., scaling polynomially in the number of qubits;
- being robust to noise;
- having controllable precision;
- being deterministic, i.e., not having any post-selection;
- having efficient circuit construction, i.e. that the compilation of the circuit is efficient.

Moreover, in the same reference [9], they summarized eight categories of initialization method but none of them fulfill all the requirements at once.

1. My Ideas

1.1. **Inverse Method to Prepare Distribution.** Suppose we want prepare a quantum state to approximately encode a distribution with cumulative function F. Inspired by the inverse transformation method, we may use the following steps to prepare such a quantum state.

$$|0^n\rangle|0^m\rangle \mapsto \sum_{x\in\{0,1\}^n} \frac{1}{\sqrt{2^n}}|x\rangle|0^m\rangle \mapsto \sum_{x\in\{0,1\}^n} \frac{1}{\sqrt{2^n}}|x\rangle|\widehat{F}_b^{-1}(x)\rangle,$$

where $\widehat{F}_b^{-1}:\{0,1\}^n \mapsto \{0,1\}^m$ is an proxy of F^{-1} in the sense that

$$\begin{cases} -\frac{1}{2^{d+1}} < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{\widehat{F}_b^{-1}(x)}{2^d} - M\right) \le \frac{1}{2^{d+1}}, & \forall x = 1,\dots,2^n-1, \\ \widehat{F}_b^{-1}(0) = 0. \end{cases} \tag{1.1}$$

Here, the pair (M,d) determines the way to interpret the m-length binary sequence stored in the second register. Then, we rearrange the final state in terms of a summation about the

Date: May 27, 2024.

second register and get

$$\sum_{x\in\{0,1\}^n}\frac{1}{\sqrt{2^n}}|x\rangle|\widehat{F}_b^{-1}(x)\rangle=\sum_{y\in\{0,1\}^m}\left(\sum_{x:\widehat{F}_b^{-1}(x)=y}\frac{1}{\sqrt{2^n}}|x\rangle\right)|y\rangle. \tag{1.2}$$

Note that

$$\begin{aligned}\left\{x\in\{0,1\}^n\setminus\{0^n\}:\widehat{F}_b^{-1}(x)=y\right\}&=\left\{x\in\{0,1\}^n\setminus\{0^n\}:-\frac{1}{2^{d+1}}<F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{y}{2^d}-M\right)\le\frac{1}{2^{d+1}}\right\}\\&=\left\{x\in\{0,1\}^n\setminus\{0^n\}:\frac{y-2^{-1}}{2^d}-M<F^{-1}\left(\frac{x}{2^n}\right)\le\frac{y+2^{-1}}{2^d}-M\right\}\\&=\left\{x\in\{0,1\}^n\setminus\{0^n\}:F\left(\frac{y-2^{-1}}{2^d}-M\right)<\frac{x}{2^n}\le F\left(\frac{y+2^{-1}}{2^d}-M\right)\right\}\end{aligned}$$

where the last equality due to the definition that $F^{-1}(u)=\inf\{x:F(x)\ge u\}$, and thus $\{F^{-1}(u)\le y\}=\{u\le F(y)\}$. Here, we give a proof to the first equality in the following way.

Proof. If $y=\widehat{F}_b^{-1}(x)$ and $x\neq 0$, by the definition of $\widehat{F}_b^{-1}$, we must have

$$-\frac{1}{2^{d+1}}<F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{y}{2^d}-M\right)\le\frac{1}{2^{d+1}}.$$

In another direction, if we have

$$-\frac{1}{2^{d+1}}<F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{y}{2^d}-M\right)\le\frac{1}{2^{d+1}},$$

but $y\neq\widehat{F}_b^{-1}(x)$. We must have

$$\left|\frac{\widehat{F}_b^{-1}(x)-y}{2^d}\right|\ge 2^{-d}\quad\text{and}\quad x\neq 0.$$

Thus, we have

$$\begin{aligned}&\left|F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{\widehat{F}_b^{-1}(x)}{2^d}-M\right)\right|\\\ge&\left|\frac{\widehat{F}_b^{-1}(x)-y}{2^d}\right|-\left|F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{y}{2^d}-M\right)\right|\\\ge&2^{-d}-2^{-d-1}=2^{-d-1}.\end{aligned}$$

When the equality does not achieve, we have derive a contradiction with the assumption of $\widehat{F}_b^{-1}(\cdot)$. In the case where we achieve equality in the above inequality, we must have

$$\begin{cases}F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{y}{2^d}-M\right)=2^{-d-1},\\\frac{\widehat{F}_b^{-1}(x)-y}{2^d}=2^{-d}.\end{cases}$$

Combining these two equalities, we must get

$$F^{-1}\left(\frac{x}{2^n}\right)-\left(\frac{\widehat{F}_b^{-1}(x)}{2^d}-M\right)=-2^{-d-1},$$

which also contradicts our assumption on $\widehat{F}_b^{-1}(\cdot)$. Thus, we must have $y = \widehat{F}_b^{-1}(x)$. Combine this two direction, we have proved the equailty

$$\left\{x \in \{0,1\}^n \setminus \{0^n\} : \widehat{F}_b^{-1}(x) = y\right\} = \left\{x \in \{0,1\}^n \setminus \{0^n\} : -\frac{1}{2^{d+1}} < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) \leq \frac{1}{2^{d+1}}\right\}.$$

□

Let N_y denote the number of elements in the set $\left\{x \in \{0,1\}^n : \widehat{F}_b^{-1}(x) = y\right\}$, then we have

$$\frac{F\left(\frac{y+2^{-1}}{2^d} - M\right) - F\left(\frac{y-2^{-1}}{2^d} - M\right)}{\frac{1}{2^n}} - 2 < N_y < \frac{F\left(\frac{y+2^{-1}}{2^d} - M\right) - F\left(\frac{y-2^{-1}}{2^d} - M\right)}{\frac{1}{2^n}} + 1. \tag{1.3}$$

When we measure the second register with computational basis, the probability to get an outcome y is

$$\frac{N_y}{2^n} \in \left(\widetilde{p}_y - 2^{-n+1}, \widetilde{p}_y + 2^{-n}\right),$$

where

$$\begin{aligned} \widetilde{p}_y &= F\left(\frac{y+2^{-1}}{2^d} - M\right) - F\left(\frac{y-2^{-1}}{2^d} - M\right) \\ &\equiv F\left(\overline{y}\right) - F\left(\underline{y}\right). \end{aligned}$$

Suppose $Y \sim F$, then we have

$$\widetilde{p}_y = \Pr\left[Y \in \left(\underline{y}, \overline{y}\right]\right] = \Pr\left[-2^{-(d+1)} < Y - B_y \leq 2^{-(d+1)}\right],$$

where $B_y = \frac{y}{2^d} - M$. We can see that $\{B_y : y \in \{0,1\}^m\}$ is a set of points distributed in the interval $[-M, 2^{m-d} - M - 2^{-d}]$ with equal space.

Question 1.1. *We can write the* (1.2) *as*

$$\sum_{y \in \{0,1\}^m} \left(\sum_{x : \widehat{F}_b^{-1}(x) = y} \frac{1}{\sqrt{2^n}} |x\rangle\right) |y\rangle \equiv \sum_{y \in \{0,1\}^m} \sqrt{p_y} |\psi_y\rangle |y\rangle,$$

where $p_y = \frac{N_y}{2^n}$. *Moreover, suppose the density corresponding to the distribution* F *is* f *and*

$$p(x) = \begin{cases} 2^d \cdot p_y & \exists y \in \{0,1,\ldots,2^m - 1\} \quad s.t. \quad x \in \left(\underline{y}, \overline{y}\right], \\ 0 & \textit{otherwise}. \end{cases}$$

Then we may regard $p(x)$ *as an approximation to* $f(x)$. *To achieve the error bound* $\|f - p\|_\infty \leq \epsilon$, *what values of* $M, m, n,$ *and* d *should be set?*

Solution. For $x \in (\underline{y}, \overline{y}]$, we have

$$
\begin{aligned}
|p(x) - f(x)| &= \left| \frac{2^d \cdot N_y}{2^n} - f(x) \right| \\
&\le \left| \frac{2^d \cdot N_y}{2^n} - 2^d \cdot \widetilde{p}_y \right| + \left| 2^d \cdot \widetilde{p}_y - f(x) \right| \\
&\le \frac{2^d}{2^{n-1}} + 2^d \int_{\underline{y}}^{\overline{y}} |f(z) - f(x)| \, dz \\
&\le \frac{2^d}{2^{n-1}} + \frac{1}{2^d} \cdot \max_{z \in [\underline{y}, \overline{y}]} |f'(z)| \\
&\le \frac{1}{2^{n-d-1}} + \frac{C}{2^d},
\end{aligned}
$$

with the assumption that $f'(z)$ exists and is uniformly bounded by a constant C in the real line. For $x \le -\frac{1}{2^{d+1}} - M$ or $x > 2^{m-d} - M - 2^{-(d+1)}$, we have $p(x) = 0$ and thus

$$
|p(x) - f(x)| = f(x).
$$

In summary, we should choose M, m, n, d such that

$$
\begin{cases} \frac{1}{2^{n-d-1}} + \frac{C}{2^d} \le \epsilon, \\ f(x) \le \epsilon, \quad \forall x \in (-\infty, -\frac{1}{2^{d+1}} - M] \cup (2^{m-d} - M - 2^{-(d+1)}, +\infty). \end{cases} \tag{1.4}
$$

□

For the case of standard Normal density

$$
f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},
$$

it is reasonable to construct a symmetric set of grids w.r.t. the original point, and hence we choose

$$
\frac{2M}{2^m} = \frac{1}{2^d} \Longrightarrow M = 2^{m-d-1}.
$$

Moreover, we can set $C = \frac{1}{\sqrt{\pi} e^{1/4}}$. Solving (1.4), we get the condition

$$
\begin{cases} \frac{1}{2^{n-d-1}} + \frac{C}{2^d} \le \epsilon, \\ 2^{m-d-1} = M \ge 2^{-(d+1)} + \sqrt{2 \log \frac{1}{\sqrt{2\pi}\epsilon}}. \end{cases}
$$

When ϵ is small, the term $2^{-(d+1)}$ is negligible compared to the term $\sqrt{2 \log \frac{1}{\sqrt{2\pi}\epsilon}}$. We can choose

$$
\begin{aligned}
d &= 1 + \log_2 \frac{C}{\epsilon}, \\
n &= d + 2 + \log_2 \frac{1}{\epsilon} = 3 + 2 \log_2 \frac{\sqrt{C}}{\epsilon}, \\
m &= d + \frac{5}{2} + \frac{1}{2} \log_2 \log \frac{1}{\sqrt{2\pi}\epsilon},
\end{aligned}
$$

where $C = \frac{1}{\sqrt{2\pi e}}$ for standard normal distribution.

Remark 1.2. *In the following applications, we use the state* $\sum_y \sqrt{p_y}|y\rangle$ *to approximate some expectation* $\int g(z)f(z)dz$ *by*

$$\sum_y p_y g(y).$$

The error is

$$\begin{aligned}
&\left|\sum_y p_y g(y) - \int g(z)f(z)dz\right| \\
\leq &\left|\sum_y p_y g(y) - \sum_y \widetilde{p}_y g(y)\right| + \left|\sum_y \widetilde{p}_y g(y) - \int g(z)f(z)dz\right| \\
\leq &\sum_y |g(y)| \cdot |p_y - \widetilde{p}_y| + \sum_y \int_{\underline{y}}^{\overline{y}} |g(y) - g(z)| f(z)dz \\
&\quad + \int_{-\infty}^{-M-2^{-d-1}} |g(z)|f(z)dz + \int_{2^{m-d}-2^{-d-1}-M}^{+\infty} |g(z)|f(z)dz.
\end{aligned}$$

We can see the last three terms are independent of the parameter n*. Therefore, to determine the parameter* n*, it is much more reasonable to consider the error term*

$$\sum_y |g(y)| \cdot |p_y - \widetilde{p}_y| \leq 2^{-n+1} \sum_y |g(y)|.$$

For bounded valued g*, we need to set* $n = 2m$ *approximately, which is consistent to the above result.*

In practice, the formula of the inverse CDF may not be given explicitly, and we can not compute the exact value with finite bits in the computer. Therefore, we need to choose suitable approximations. The next question is how to implement a function $\widehat{F}_b^{-1}$ satisfying (1.1) based on an approximation $\widehat{F}^{-1}$ to F^{-1}.

Question 1.3. *Suppose* $\widehat{F}^{-1} : \mathbb{R} \mapsto [0,1]$ *is an approximation to* F^{-1} *in the sense that*

$$\left|\widehat{F}^{-1}\left(\frac{x}{2^n}\right) - F^{-1}\left(\frac{x}{2^n}\right)\right| \leq \delta, \quad \forall x \in \{1, \ldots, 2^n - 1\},$$

and $\widehat{F}^{-1}(0) = -\infty$*. Can we implement* $\widehat{F}_b^{-1}$ *satisfying* (1.1) *based on* $\widehat{F}^{-1}$*, and what is the requirement to* δ*? If the condition* (1.1) *can not be satisfied, how to get a reasonable approximation?*

Solution. For a given $x \in \{1, \ldots, 2^n - 1\}$, we use B_x to denote the binary representation of $\widehat{F}^{-1}\left(\frac{x}{2^n}\right)$,

$$B_x = s_{m-d-1} \cdots s_1 s_0 . s_{-1} s_{-2} \cdots .$$

Here, we assume $\left|\widehat{F}^{-1}\left(\frac{x}{2^n}\right)\right| < 2^{m-d-1}$, and the most significant position s_{m-d-1} denote the sign of the number, 0 for negative and 1 for positive. Thus, the scheme of B_x can represent numbers in the interval $[-2^{m-d-1}, 2^{m-d-1}]$. Let $\mathtt{rd}_{m,d}(\cdot)$ denote the following operation:

$$\mathtt{rd}_{m,d}(B_x) = s_{m-d-1} \cdots s_1 s_0 . s_{-1} \cdots s_{-d+1}(s_{-d} + s_{-d-1}).$$

With this truncation, the numbers can be represented are in the interval $[-2^{m-d-1}+2^{-d}, 2^{m-d-1}-2^{-d}]$ and of the form $N \cdot 2^{-d}$ for some integer N. For convenience, we define the set

$$S := [-2^{m-d-1} + 2^{-d}, 2^{m-d-1} - 2^{-d}] \cap \left\{N \cdot 2^{-d} : N \in \mathbb{Z}\right\}.$$

Then, it can be shown that $|\mathtt{rd}_{m,d}(B_x) - B_x| \le 2^{-d-1}$:

- If $s_{-d-1} = 0$, then

$$\begin{aligned} B_x - \mathtt{rd}_{m,d}(B_x) &= 0.0\cdots 0 s_{-d-1} s_{-d-2}\cdots \\ &= 0.0\cdots 00 s_{-d-2}\cdots \\ &\le 0.0\cdots 0011\cdots = 2^{-d-1}. \end{aligned}$$

- If $s_{-d-1} = 1$, then

$$\begin{aligned} B_x - \mathtt{rd}_{m,d}(B_x) &= -2^{-d} + 0.0\cdots 0 s_{-d-1} s_{-d-2}\cdots \\ &= -2^{-d} + 0.0\cdots 01 s_{-d-2}\cdots \\ &\ge -2^{-d} + 0.0\cdots 010\cdots = -2^{-d-1}. \end{aligned}$$

In the above two cases, we assume $B_x \ge 0$. In the case $B_x < 0$, the discussion is the same. Then $\widehat{F}_b^{-1}$ can be defined by

$$\widehat{F}_b^{-1}(x) := 2^d \cdot (\mathtt{rd}_{m,d}(B_x) + M).$$

Thus, in the case $\mathtt{rd}_{m,d}(B_x) + M \in S$ we have

$$\frac{\widehat{F}_b^{-1}(x)}{2^d} - M = \mathtt{rd}_{m,d}(B_x).$$

With triangular inequality, for $x \in \{1, \ldots, 2^n - 1\}$, we have

$$\begin{aligned} \left|F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{\widehat{F}_b^{-1}(x)}{2^d} - M\right)\right| &\le \left|F^{-1}\left(\frac{x}{2^n}\right) - \widehat{F}^{-1}\left(\frac{x}{2^n}\right)\right| + \left|\widehat{F}^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{\widehat{F}_b^{-1}(x)}{2^d} - M\right)\right| \\ &= \left|F^{-1}\left(\frac{x}{2^n}\right) - \widehat{F}^{-1}\left(\frac{x}{2^n}\right)\right| + |B_x - \mathtt{rd}_{m,d}(B_x)| \\ &\le \delta + 2^{-d-1}. \end{aligned}$$

Thus, we can see (1.1) is not achievable in realistic since the round off to a d-accurate register causes an magnitude 2^{-d-1} error and the approximately computing of F^{-1} causes extra error δ. □

Remark 1.4. *The error dependence on δ can be eliminated if we can find an approximation $\widehat{F}^{-1}$ such that $\widehat{F}^{-1}(x)$ and $F^{-1}(x)$ have the same bits representation until the s_{-d-1} position.*

However, we should not expect that the error dependence on δ can be eliminated when δ is sufficiently small. Consider the case

$$F^{-1}(x) = s_{m-d-1}\cdots s_1 s_0.s_{-1}\cdots s_{-d+1}0100\cdots 01$$

but we get

$$\widehat{F}^{-1}(x) = s_{m-d-1}\cdots s_1 s_0.s_{-1}\cdots s_{-d+1}0011\cdots 11.$$

Suppose there are t bits behind the arithmetic point, the error between them is 2^{-t+1}, which can be arbitrary small. However,

$$\boldsymbol{\mathtt{rd}}_{m,d}\left(\widehat{F}^{-1}(x)\right) = s_{m-d-1}\cdots s_1 s_0.s_{-1}\cdots s_{-d+1}00.$$

Then, it achieves the error

$$2^{-d-1} + 2^{-t},$$

which still depend on the parameter t.

Remark 1.5. *From the above definition of* $\widehat{F}_b^{-1}$, *we see that*

$$\min_{x\in\{1,\ldots,2^n-1\}} \widehat{F}_b^{-1}(x) = \widehat{F}_b^{-1}(1) := 2^d\left(\boldsymbol{rd}_{m,d}(B_1) + M\right).$$

In order to take the full advantage of the register storing $\widehat{F}_b^{-1}$, *the value of* M *should be chosen such that* $\widehat{F}_b^{-1}(1) = 0$.

Question 1.6. *What happens if we relax the condition* (1.1) *to*

$$\begin{cases} -\gamma < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{\widehat{F}_b^{-1}(x)}{2^d} - M\right) \leq \gamma, & \forall x \in \{1,\ldots,2^n-1\}, \\ \widehat{F}_b^{-1}(0) = 0, \end{cases} \tag{1.5}$$

with $\gamma \in (2^{-d-1}, 2^{-d})$? *We assume that* F *is continuous over the real line, and has density function* f.

Sol. At first, we investigate the case $\gamma < 2^{-d-1}$, which will give us some insight to case $\gamma > 2^{-d-1}$.

- In the case $\gamma < 2^{-d-1}$, we have

$$\left\{x \in \{0,1\}^n \setminus \{0^n\} : \widehat{F}_b^{-1}(x) = y\right\} = \left\{x \in \{0,1\}^n \setminus \{0^n\} : -\gamma < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) \leq \gamma\right\}.$$

  For one direction, if x is in the first set, then $y = \widehat{F}_b^{-1}(x)$ satisfying the condition (1.5), and thus x belongs to the second set also. For another direction, if $x \in \{0,1\}^n \setminus \{0^n\}$ satisfies

$$-\gamma < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) \leq \gamma$$

  for a given y, but $\widehat{F}_b^{-1}(x) \neq y$. Then we have $\left|\widehat{F}_b^{-1}(x) - y\right| \geq 1$, and thus

$$\begin{aligned} \left|F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{\widehat{F}_b^{-1}(x)}{2^d} - M\right)\right| &\geq \left|\frac{y - \widehat{F}_b^{-1}(x)}{2^d}\right| - \left|F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right)\right| \\ &\geq 2^{-d} - \gamma \geq 2^{-d-1} > \gamma, \end{aligned}$$

  which is contradict to the condition (1.5).
- In the case $\gamma \geq 2^{-d-1}$ we can not argue $2^{-d} - \gamma \geq 2^{-d-1} > \gamma$ again for the second direction, but the first direction still holds true:

$$\left\{x \in \{0,1\}^n \setminus \{0^n\} : \widehat{F}_b^{-1}(x) = y\right\} \subset \left\{x \in \{0,1\}^n \setminus \{0^n\} : -\gamma < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) \leq \gamma\right\}.$$

  For the second direction, we define the quantity $\bar{\gamma} = 2^{-d} - \gamma$, if $x \in \{0,1\}^n$ satisfying

$$-\bar{\gamma} < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) < \bar{\gamma}$$

for a given y. Then we have $\widehat{F}_b^{-1}(x) = y$, or we have $\left|\widehat{F}_b^{-1}(x) - y\right| \geq 1$ and thus

$$\begin{aligned}\left|F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{\widehat{F}_b^{-1}(x)}{2^d} - M\right)\right| &\geq \left|\frac{y - \widehat{F}_b^{-1}(x)}{2^d}\right| - \left|F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right)\right| \\ &> 2^{-d} - \bar{\gamma} = \gamma,\end{aligned}$$

which contradicts the condition (1.5). In conclusion, we have

$$\left\{x \in \{0,1\}^n \setminus \{0^n\} : \widehat{F}_b^{-1}(x) = y\right\} \subset \left\{x \in \{0,1\}^n \setminus \{0^n\} : -\gamma < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) \leq \gamma\right\},$$
$$\left\{x \in \{0,1\}^n : \widehat{F}_b^{-1}(x) = y\right\} \supset \left\{x \in \{0,1\}^n : -\bar{\gamma} < F^{-1}\left(\frac{x}{2^n}\right) - \left(\frac{y}{2^d} - M\right) < \bar{\gamma}\right\}.$$

Let

$$\overline{p_y} = \int_{-\gamma + \frac{y}{2^d} - M}^{\gamma + \frac{y}{2^d} - M} f(x)dx, \quad \underline{p_y} = \int_{-\bar{\gamma} + \frac{y}{2^d} - M}^{\bar{\gamma} + \frac{y}{2^d} - M} f(x)dx,$$

and N_y be the number of elements in $\left\{x \in \{0,1\}^n : \widehat{F}_b^{-1}(x) = y\right\}$, then we have

$$\frac{\underline{p_y}}{2^{-n}} - 2 < N_y < \frac{\overline{p_y}}{2^{-n}} + 1.$$

Then, we get

$$\frac{N_y}{2^n} \in \left(\underline{p_y} - 2^{-n+1}, \overline{p_y} + 2^{-n}\right).$$

The desired quantity p_y is

$$p_y = \int_{-M + \frac{y}{2^d} - \frac{1}{2^{d+1}}}^{-M + \frac{y}{2^d} + \frac{1}{2^{d+1}}} f(x)dx = \int_{\underline{y}}^{\overline{y}} f(x)dx,$$

where we write $\underline{y} \equiv -M + \frac{y}{2^d} - \frac{1}{2^{d+1}}$ and $\overline{y} \equiv -M + \frac{y}{2^d} + \frac{1}{2^{d+1}}$ for simplicity. Then we have

$$\left|\frac{N_y}{2^n} - p_y\right| \leq \max\left\{\overline{p_y} - p_y, p_y + 2^{-n+1} - \underline{p_y}\right\},$$

where

$$\begin{aligned}\overline{p_y} - p_y &= \int_{-\delta + \underline{y}}^{\delta + \overline{y}} f(x)dx - \int_{\underline{y}}^{\overline{y}} f(x)dx = \int_{-\delta + \underline{y}}^{\underline{y}} f(x)dx + \int_{\overline{y}}^{\delta + \overline{y}} f(x)dx \\ &\leq 2\delta \cdot \max_{x \in \mathbb{R}} f(x), \\ p_y - \underline{p_y} &= \int_{\underline{y}}^{\overline{y}} f(x)dx - \int_{\delta + \underline{y}}^{-\delta + \overline{y}} f(x)dx = \int_{\underline{y}}^{\delta + \underline{y}} f(x)dx + \int_{-\delta + \overline{y}}^{\overline{y}} f(x)dx \\ &\leq 2\delta \cdot \max_{x \in \mathbb{R}} f(x).\end{aligned}$$

Here, we write $\gamma = 2^{-d-1} + \delta$ for some $\delta \in (0, 2^{-d-1})$. In conclusion, we have

$$\left|\frac{N_y}{2^n} - p_y\right| \leq 2^{-n+1} + 2\delta \cdot \max_{x \in \mathbb{R}} f(x).$$

□

1.2. **Implement the Inverse Transformation for Standard Normal Distribution.** Here we consider the gate complexity to implement the function Φ^{-1}, where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz.$$

To preserve the quadratic speedup of QMCI, we should use some approximations to Φ^{-1} with reasonable accuracy.

There is a plenty of works on the topic of approximating the inverse Normal CDF. For example in [8], there are some simple approximations

$$\Phi^{-1}(p) \approx 1.724 - 2.22t + 0.157t^2, \quad t = \sqrt{-\log p},$$
$$\Phi^{-1}(p) \approx 1.758 - 2.257t + 0.1661t^2, \quad t = \sqrt{-\log p}.$$

For more accurate approximations, there are

$$\begin{aligned} &y = -\log\left(2(1-p)\right), \\ &\Phi^{-1}(p) \approx \sqrt{\frac{\left(\left(4y+100\right)y+205\right)y^2}{\left(\left(2y+56\right)y+192\right)y+131}} = y \cdot \sqrt{\frac{\left(4y+100\right)y+205}{\left(\left(2y+56\right)y+192\right)y+131}} \end{aligned} \tag{1.6}$$

for $\frac{1}{2} < p < 1 - 10^{-7}$. This method is recorded in [13], and is claimed to achieve the approximate error within 1.3×10^{-4} in the interval $(0.5, 1-10^{-7})$. There is another interesting approximation in [13], say

$$\Phi^{-1}(p) \approx \frac{p^{0.135} - (1-p)^{0.135}}{0.1975}, \quad 0 < p < 1. \tag{1.7}$$

In the case $|\Phi^{-1}(p)| \leq 2$, this approximation achieve an error within 0.0093. These approximations are not that accurate, but I think they are enough for the present quantum hardware. For more accurate approximation, please refer to [8].

To implement such approximation in quantum circuit, I refered to [4] for the fixed point addition and multiplication. The quantum circuit implementation of $\log(\cdot)$ and $\sqrt{\cdot}$ should be considered further, or we use the approximation (1.7) with an implementation of $(\cdot)^{0.135}$. Since the more accurate approximation usually involves square root and logarithm, I think it is worth to talk about them.

Question 1.7. *What is the advantage of this method compared to the classical Grover-Rudolph method?*

1.3. **Box-Muller Method.** Suppose $U_1, U_2 \overset{\text{i.i.d.}}{\sim} \text{Unif}[0,1]$, and

$$\begin{cases} Z_1 = \sqrt{-2\log U_1} \cos\left(2\pi U_2\right), \\ Z_2 = \sqrt{-2\log U_1} \sin\left(2\pi U_2\right). \end{cases}$$

Then we have $Z_1, Z_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. To avoid the computation of sine and cosine, we can use the following procedure

$$\begin{aligned} &U_1 \leftarrow 2U_1 - 1, \quad U_2 \leftarrow 2U_2 - 1, \\ &X = U_1^2 + U_2^2, \quad Y = \sqrt{-2\log X / X}, \\ &Z_1 = U_1 Y, \quad Z_2 = U_2 Y. \end{aligned}$$

Moreover, to generate random vectors satisfying

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right),$$

we can compute

$$\begin{cases} X_1 = Z_1, \\ X_2 = \rho Z_1 + \sqrt{1-\rho^2} Z_2. \end{cases}$$

1.4. **Can we reduce the required qubits?** Figure 1 is a simple illustration to the inverse transformation method. Here we set $n = 5, m = 3, M = 1.75, d = 1$, and try to approximate the standard normal distribution. There are 2^n black points on the horizontal axis equally spaced by 2^{-n}, and 2^m red points on the vertical axis equally spaced by 2^{-d}. Then, we map the red points to the horizontal axis by the function $F(\cdot)$.

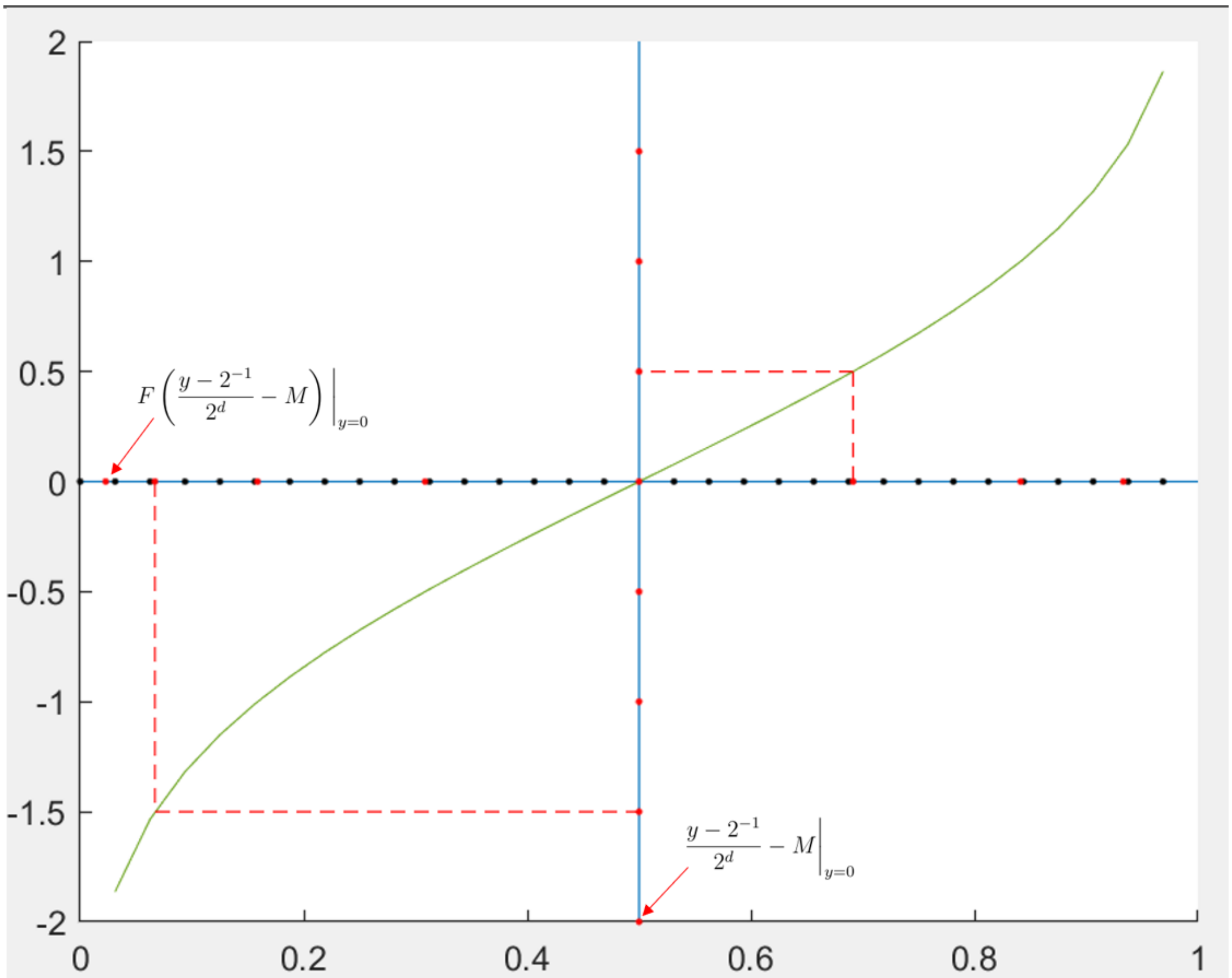

Figure 1. A Simple Illustration to The Inverse Method.

Finally, we can count the number of black points between two adjacent red points as the quantity N_y and get an approximate distribution to the standard normal distribution.

In this case, if we consider the corresponding quantum computation process, that is

$$\begin{aligned}
|0^5\rangle|0^3\rangle &\mapsto \sum_{x\in\{0,1\}^5} \frac{1}{\sqrt{2^5}}|x\rangle|0^3\rangle \\
&\mapsto (|00000\rangle + |00001\rangle + |00010\rangle)\,|000\rangle \\
&\quad + (|00011\rangle + |00100\rangle + |00101\rangle)\,|001\rangle \\
&\quad + (|00110\rangle + |00111\rangle + |01000\rangle + |01001\rangle)\,|010\rangle \\
&\quad + (|01010\rangle + |01011\rangle + |01100\rangle + |01101\rangle + |01110\rangle + |01111\rangle + |10000\rangle)\,|011\rangle \\
&\quad + (|10001\rangle + |10010\rangle + |10011\rangle + |10100\rangle + |10101\rangle + |10110\rangle)\,|100\rangle \\
&\quad + (|10111\rangle + |11000\rangle + |11001\rangle + |11010\rangle)\,|101\rangle \\
&\quad + (|11011\rangle + |11100\rangle + |11101\rangle)\,|110\rangle \\
&\quad + (|11110\rangle + |11111\rangle)\,|111\rangle.
\end{aligned}$$

Then, we can observe that if we only preserve the first three qubits in the first register, we will get the same distribution when we only measure the qubits in the second register. The reduced qubits state is

$$\begin{aligned}
&(|000\rangle + |001\rangle + |010\rangle)\,|000\rangle \\
&\quad + (|011\rangle + |100\rangle + |101\rangle)\,|001\rangle \\
&\quad + (|110\rangle + |111\rangle + |000\rangle + |001\rangle)\,|010\rangle \\
&\quad + (|010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle + |000\rangle)\,|011\rangle \\
&\quad + (|001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle)\,|100\rangle \\
&\quad + (|111\rangle + |000\rangle + |001\rangle + |010\rangle)\,|101\rangle \\
&\quad + (|011\rangle + |100\rangle + |101\rangle)\,|110\rangle \\
&\quad + (|110\rangle + |111\rangle)\,|111\rangle.
\end{aligned}$$

This means there is a space to reduce the required qubits.

Question 1.8. *How to reduce the required qubits? Find such methods as many as possible.*

Sol. To reduce qubits, we define the function $\widehat{F_b} : \{0,1\}^m \mapsto \{0,1\}^n$ as

$$\widehat{F_b}(y) = \begin{cases} 0, & y = 0 \\ \left\lceil 2^n F\left(\frac{y-2^{-1}}{2^d} - M\right)\right\rceil, & y \neq 0. \end{cases}$$

In the above illustration example, with the definition of $\widehat{F_b}^{-1}$ and $\widehat{F_b}$, we may have

$$\begin{aligned}
&\widehat{F_b}(000) = 00000, \quad \widehat{F_b}(001) = 00011, \quad \widehat{F_b}(010) = 00110, \quad \widehat{F_b}(011) = 01010, \\
&\widehat{F_b}(100) = 10001, \quad \widehat{F_b}(101) = 10111, \quad \widehat{F_b}(110) = 11011, \quad \widehat{F_b}(111) = 11110.
\end{aligned}$$

Suppose there is a quantum circuit to implement the transformation: $|y\rangle_m|0\rangle_n \mapsto |y\rangle|\widehat{F}_b(y)\rangle$, then we calim the following process can reduce the required qubits to represent the distribution corresponding to F.

$$\begin{aligned}
|0\rangle_n|0\rangle_m|0\rangle_n &\mapsto \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle_n|0\rangle_m|0\rangle_n \mapsto \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle_n|\widehat{F}_b^{-1}(x)\rangle_m|0\rangle_n \\
&\overset{\mathcal{U}}{\mapsto} \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle_n|\widehat{F}_b^{-1}(x)\rangle_m\left|\widehat{F}_b\left(\widehat{F}_b^{-1}(x)\right)\right\rangle_n \\
&\mapsto \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}\left|x-\widehat{F}_b\left(\widehat{F}_b^{-1}(x)\right)\right\rangle_n|\widehat{F}_b^{-1}(x)\rangle_m\left|\widehat{F}_b\left(\widehat{F}_b^{-1}(x)\right)\right\rangle_n \\
&\overset{\mathcal{U}^{-1}}{\mapsto} \frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}\left|x-\widehat{F}_b\left(\widehat{F}_b^{-1}(x)\right)\right\rangle_n|\widehat{F}_b^{-1}(x)\rangle_m\,|0\rangle_n\,.
\end{aligned}$$

Apply this process to the illustration example, we will finally get the state:

$$\begin{aligned}
&(|00000\rangle+|00001\rangle+|00010\rangle)\,|000\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle+|00010\rangle)\,|001\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle+|00010\rangle+|00011\rangle)\,|010\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle+|00010\rangle+|00011\rangle+|00100\rangle+|00101\rangle+|00110\rangle)\,|011\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle+|00010\rangle+|00011\rangle+|00100\rangle+|00101\rangle)\,|100\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle+|00010\rangle+|00011\rangle)\,|101\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle+|00010\rangle)\,|110\rangle|00000\rangle \\
&+(|00000\rangle+|00001\rangle)\,|111\rangle|00000\rangle \\
=\ &|00\rangle\big[\,(|000\rangle+|001\rangle+|010\rangle)\,|000\rangle \\
&+(|000\rangle+|001\rangle+|010\rangle)\,|001\rangle \\
&+(|000\rangle+|001\rangle+|010\rangle+|011\rangle)\,|010\rangle \\
&+(|000\rangle+|001\rangle+|010\rangle+|011\rangle+|100\rangle+|101\rangle+|110\rangle)\,|011\rangle \\
&+(|000\rangle+|001\rangle+|010\rangle+|011\rangle+|100\rangle+|101\rangle)\,|100\rangle \\
&+(|000\rangle+|001\rangle+|010\rangle+|011\rangle)\,|101\rangle \\
&+(|000\rangle+|001\rangle+|010\rangle)\,|110\rangle \\
&+(|000\rangle+|001\rangle)\,|111\rangle\big]|00000\rangle.
\end{aligned}$$

Thus, the first two and the last five qubits are released. □

Question 1.9. *How many qubits can we reduce?*

Sol. Let $f \equiv F'$ be the density function of the distribution F, and suppose $f_{max} = \max_{x\in\mathbb{R}} f(x)$. Then, we have

$$N_y = \frac{F\left(\frac{y+2^{-1}}{2^d}-M\right)-F\left(\frac{y-2^{-1}}{2^d}-M\right)}{\frac{1}{2^n}} \leq \frac{f_{max}\left(\frac{y+2^{-1}}{2^d}-M-\frac{y-2^{-1}}{2^d}+M\right)}{\frac{1}{2^n}} = 2^{n-d}f_{max},$$

where N_y is defined the same as in (1.3). Then, the number of qubits can be reduced is at least

$$n - \log_2(2^{n-d} f_{max}) = d - \log_2 f_{max}.$$

In the case of standard normal distribution, we have $f_{max} = \frac{1}{\sqrt{2\pi}}$ and thus the number of qubits can be reduced is at least $d + \log_2(\sqrt{2\pi}) \geq d + 1$. $\square$

References

[1] I. F. Araujo, D. K. Park, F. Petruccione, and A. J. da Silva. A divide-and-conquer algorithm for quantum state preparation. *Scientific reports*, 11(1):6329, 2021.

[2] S. Chakrabarti, R. Krishnakumar, G. Mazzola, N. Stamatopoulos, S. Woerner, and W. J. Zeng. A threshold for quantum advantage in derivative pricing. *Quantum*, 5:463, 2021.

[3] L. Grover and T. Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. *arXiv preprint quant-ph/0208112*, 2002.

[4] T. Häner, M. Roetteler, and K. M. Svore. Optimizing quantum circuits for arithmetic. *arXiv preprint arXiv:1805.12445*, 2018.

[5] S. Herbert. No quantum speedup with grover-rudolph state preparation for quantum monte carlo integration. *Physical Review E*, 103(6):063302, 2021.

[6] A. Holmes and A. Y. Matsuura. Efficient quantum circuits for accurate state preparation of smooth, differentiable functions. pages 169–179, 2020.

[7] J. Iaconis, S. Johri, and E. Y. Zhu. Quantum state preparation of normal distributions using matrix product states. *arXiv preprint arXiv:2303.01562*, 2023.

[8] S. A. Justin S.Dyer. Approximations to inverse error functions.

[9] A. Manzano, A. Gómez, and C. C. Vázquez. D5. 7: Update of review of state-of-the-art for pricing and computation of var. 2023.

[10] G. Marin-Sanchez, J. Gonzalez-Conde, and M. Sanz. Quantum algorithms for approximate function loading. *Physical Review Research*, 5(3):033114, 2023.

[11] S. McArdle, A. Gilyén, and M. Berta. Quantum state preparation without coherent arithmetic. *arXiv preprint arXiv:2210.14892*, 2022.

[12] A. Montanaro. Quantum speedup of monte carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150301, 2015.

[13] J. K. Patel and C. B. Read. *Handbook of the normal distribution*, volume 150. CRC Press, 1996.

[14] A. G. Rattew, Y. Sun, P. Minssen, and M. Pistoia. The efficient preparation of normal distributions in quantum registers. *Quantum*, 5:609, 2021.

[15] P. Rebentrost, B. Gupt, and T. R. Bromley. Quantum computational finance: Monte carlo pricing of financial derivatives. *Physical Review A*, 98(2):022321, 2018.

[16] A. C. Vazquez and S. Woerner. Efficient state preparation for quantum amplitude estimation. *Physical Review Applied*, 15(3):034027, 2021.