Circuit Design and Efficient Simulation of Quantum Inner Product and Empirical Studies of Its Effect on Near-Term Hybrid Quantum-Classic Machine Learning

Supplementary Material

Contents

G. Preliminaries of Quantum Computing

In this section, we first go over the basic concepts of QC in Sec. G.1, and then introduce two important quantum algorithms for this paper, Hadamard test and Quantum Phase Estimation, in Sec. G.2 and Sec. G.3 respectively.

G.1. Quantum State, Gate, and Measurement

Analogue to the classical computing carried out in classical computers where the state of binary bits are changed via logical gates, quantum computing is physically realized by the quantum computer comprised of quantum bits and quantum gates. The quantum bit named as ***qubit*** is the basic unit of quantum computing, which is written in the conventional *Dirac notation*, i.e., the qubit $|0\rangle$ and qubit $|1\rangle$ are quantum versions of the classic bit 0 and bit 1 respectively. Mathematically, one qubit is denoted as a 2-dimensional unit vector, i.e., $|0\rangle = [1\ 0]^T, |1\rangle = [0\ 1]^T$. Accordingly, we call the $|0\rangle$ and $|1\rangle$ are *computational basis states* since they are orthonormal and can be spanned into $\mathbb{R}^2$. In a general case, the qubit can be represented by a linear superposition of its two computational basis states, which is given as $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle = [\alpha\ \beta]^\top$ where α and β are two complex numbers namely *amplitudes* and $|\alpha|^2 + |\beta|^2 = 1$, which can be also written as $\langle\phi|\phi\rangle = 1$ where $\langle\phi|$ denotes the conjugate transposition of $|\phi\rangle$. For a quantum state comprised of n qubits, it can be represented by a 2^n-dimensional vector whose l_2-norm is 1 such as $|01\rangle = |0\rangle \otimes |1\rangle = [0\ 1\ 0\ 0]^T$ where $\otimes$ denotes tensor product operation. We also use the superscript notation $|\phi\rangle^{\otimes n}$ to indicate a quantum state comprised of n qubits in the same state $|\phi\rangle$. There also exits some *entangled states* that can not be decomposed by tensor product operation. For example, the state comprised of two qubits $|\psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ can not be decomposed into a tensor product of two individual qubits. However, the entangled state would have the ability of storing much more information since an entangled state containing n qubits hold up to 2^n non-zero amplitudes to encode the classic datum.

Given the initial qubits encoded the classical information, ***quantum gates*** are implemented to transform the state of qubits to achieve a given computing task. According to the quantum mechanics, quantum gates are *unitary* operations to ensure the normalization property of quantum states before and after the transformation. For example, the Hadamard gate H whose mathematical formulation is written as

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{4}$$

By applying such gate on state $|0\rangle$, the state is transformed to another state given by $H|0\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$. The quantum gate establishes the basic building block for ***quantum circuit***.

After a sequence of unitary transformations, the result would be generated and encoded in the final quantum states. The result can not be observed until ***quantum measurement*** is implemented to interact with the quantum system. The measurement is described by an observable, M, a *Hermitian* operator on the state space of the system being observed. The information would be restored from the quantum state to the classical representation by calculating average values for measurements $E(M) = \langle\phi|M|\phi\rangle$. For example, after doing

a measurement on state $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ using an observable $|0\rangle\langle 0| - |1\rangle\langle 1|$, the state will collapse to either $|0\rangle$ or $|1\rangle$ with the same probability 50%.

G.2. Hadamard Test

In quantum mechanism, the ***fidelity*** is a measure of the overlap between two quantum states, which formally is defined in an inner-product form as $f_{\text{fid}} = |\langle\phi|\psi\rangle|^2$ for two quantum states $|\phi\rangle$ and $|\psi\rangle$. Let $|\phi\rangle$ and $|\psi\rangle$ be two n-qubit quantum states that are prepared by unitary operators U_ϕ and U_ψ respectively. That is $|\phi\rangle = U_\phi|0\rangle^{\otimes n}$, $|\psi\rangle = U_\psi|0\rangle^{\otimes n}$. The ***Hadamard test*** is an algorithm to estimate the quantum fidelity. The circuit of the Hadamard test for two single qubits $|\phi\rangle$ and $|\psi\rangle$ is given in Fig. 6 where the qubit on the top is the auxiliary qubit used for measurement. By the circuit, we can get output quantum state $|\rho\rangle$ step by step as follows:

$$\begin{aligned} &\text{Input} = |0\rangle|0\rangle^{\otimes n} \\ &\xrightarrow{1} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle^{\otimes n} \\ &\xrightarrow{2} \frac{1}{\sqrt{2}}|0\rangle|\phi\rangle + \frac{1}{\sqrt{2}}|1\rangle|\psi\rangle \\ &\xrightarrow{3} \frac{1}{2}|0\rangle(|\phi\rangle + |\psi\rangle) + \frac{1}{2}|1\rangle(|\phi\rangle - |\psi\rangle). \end{aligned} \tag{5}$$

Then, the probability of measuring auxiliary qubit as $|0\rangle$ is:

$$\Pr(0) = \frac{1}{4}(\langle\phi| + \langle\psi|)(|\phi\rangle + |\psi\rangle) = \frac{1 + \textbf{Re}(\langle\psi|\phi\rangle)}{2}. \tag{6}$$

When all of the amplitudes of $|\phi\rangle$ and $|\psi\rangle$ are real numbers, Eq. 6 is reduced to

$$\Pr(0) = \frac{1 + \langle\psi|\phi\rangle}{2}, \tag{7}$$

by which we have $\langle\phi|\psi\rangle = 2\Pr(0) - 1$. The score of $\langle\psi|\phi\rangle$ can also be estimated by the same circuit but swapping the positions of the two unitaries U_ϕ and U_ψ. Thus we can compute the fidelity by $f_{\text{fid}} = |\langle\phi|\psi\rangle|^2 = \langle\phi|\psi\rangle \cdot \langle\psi|\phi\rangle$.

It's necessary to mention that there are various definitions and formulations with respect to the fidelity of quantum states. In some literature, the type of fidelity we take into account in this paper is referred to as the "Hilbert-Schmidt distance". Also a number of quantum algorithms have been proposed to estimate such kind of fidelity, among which the Hadamard test is very elegant and can be efficiently implemented (i.e., the number of gates does not grow with the dimension of the Hilbert space in which the data is embedded), and is therefore feasible for near-term quantum computers.

G.3. Quantum Phase Estimation

Quantum phase estimation (QPE) algorithm, also known as the quantum eigenvalue estimation algorithm, is the key component of many quantum algorithms. Suppose a unitary

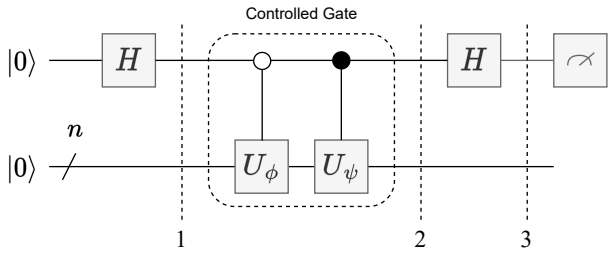

Figure 6. Quantum circuit of the Hadamard test for two quantum states $|\phi\rangle$ and $|\psi\rangle$.

operator U has an eigenvector $|\psi\rangle$ with eigenvalue $e^{2\pi i\varphi}$, where the value of φ is unknown. The goal of the phase estimation algorithm is to estimate φ. As illustrated in Figure 7, the quantum phase estimation procedure uses two registers. The first register contains t auxiliary qubits initially in the state $|0\rangle$. The value of t should be carefully considered since it determines the number of digits of accuracy we wish to have in our estimate for φ. The second register begins in the state $|\psi\rangle$, and contains as many qubits as is necessary to store $|\psi\rangle$. QPE is performed in two stages. The circuit begins by applying a Hadamard gate to the first register, followed by the application of controlled-U operations on the second register, with U raised to successive powers of two. The value of φ would be encoded into the amplitudes of a superposition state in the first register after controlled-U operations. The second stage of QPE is to apply the inverse quantum Fourier transform (inverse QFT) on the first register to change the Fourier basis to the computational basis. The third and final stage of QPE is to read out the state of the first register by doing a measurement in the computational basis. We give the mathematical definition of QPE in Def. 1. And the following theorem shows that QPE provides a pretty good estimate of φ with an appropriate number of auxiliary qubits.

Definition 1 (Quantum Phase Estimation).

$$|u\rangle|0\rangle^{\otimes t} \xrightarrow{QPE} \frac{1}{2^t}\sum_{x=0}^{2^t-1}\sum_{k=0}^{2^t-1} e^{-\frac{2\pi \mathbf{i} k}{2^t}(x - 2^t\varphi)}|u\rangle|x\rangle \tag{8}$$

Theorem 4 (Performance of QPE [24]). *Suppose a unitary operator U has an eigenvector $|u\rangle$ with eigenvalue $e^{2\mathbf{i}\varphi}$ for $\varphi \in [0, \pi)$ with $\mathbf{i}$ being the imaginary unit, i.e. we have $U|u\rangle = e^{2\mathbf{i}\varphi}|u\rangle$. The quantum phase estimation algorithm can map a state $|u\rangle|0\rangle^{\otimes t}$ to the state $|u\rangle|2^t\overline{\varphi}/\pi\rangle$ such that $|\varphi - \overline{\varphi}| \leq \epsilon$ in time $O(\frac{T(U)}{\epsilon})$, where $T(U)$ is the time to implement U.*

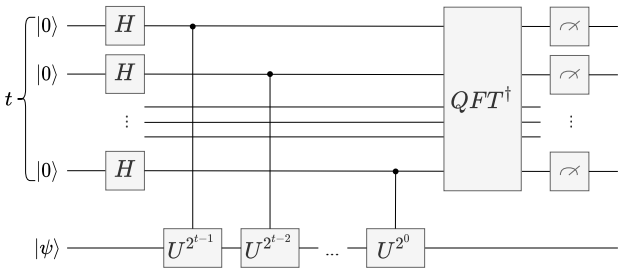

Figure 7. Quantum circuit of Quantum Phase Estimation.

H. Quantum Circuits Construction

H.1. Loading Data onto Quantum Devices by Amplitude Encoding

Generating the quantum state with classical information is the primary task of using quantum computing techniques to solve classical problems. Since our goal is to do arithmetic operations w.r.t. estimating the inner product in quantum devices, we adopt the famous amplitude encoding method, which has been implemented in many quantum algorithms [12, 13, 16, 27] to transform the classical vector-like or matrix-like data into their quantum representations. Now we give the formal definition of the quantum data loading procedure.

Definition 2 (Quantum Data Loading). Given access to classical data $\mathbf{x} \in \mathbb{R}^d$ whose i-th element is written as $\mathbf{x}[i]$, and assume d is a power of 2, there exists a quantum algorithm to generate quantum states of the form

$$|\mathbf{x}\rangle = U_x|0\rangle^{\otimes \log d} = \frac{1}{||\mathbf{x}||}\sum_{i=0}^{d-1}\mathbf{x}[i]|i\rangle, \tag{9}$$

where U_x is a unitary transformation also named as amplitude encoding in some literature. As described in [16], the complexity to prepare and readout an element scales logarithmically with respect to d.

H.2. Proofs of the 1-to-1 Quantum Inner Product Theorem

In our model, we treat the process executed on quantum processors as a dedicated procedure for accelerating the calculation of inner product scores. Remarkably, the quantum inner product estimation can achieve exponential speedup and no classical implements can match it as far. Previous works [2, 19, 33] study the quantum algorithm used to estimate the inner product score given one pair of vectors. After loading the vectors into two quantum states respectively, they employ Hadamard test to generate a state containing the information of the inner product score between such two states, followed by the QPE used to recover the score in some computational basis states. We adopt their idea and name such quantum algorithm as the 1-to-1 QIP, which is introduced in details as follows.

It is straightforward to achieve the 1-to-1 QIP by the quantum circuit which is an integration of Hadamard test and quantum phase estimation. As illustrated in Fig. 8, at the beginning of the circuit, two classical vectors $\mathbf{z}$ and $\mathbf{x}$ are loaded into a quantum state in the input register, and then transform the inner product into quantum phases through Hadamard test. Only consider the quantum state in the input register, the transformation is operated step by step as follows:

$$\begin{aligned}
&\text{Input} = |0\rangle|0\rangle^{\otimes k} \\
&\xrightarrow{1} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle^{\otimes k} \\
&\xrightarrow{2} \frac{1}{\sqrt{2}}(|0\rangle|\mathbf{z}\rangle + |1\rangle|0\rangle^{\otimes k}) \\
&\xrightarrow{3} \frac{1}{\sqrt{2}}(|0\rangle|\mathbf{z}\rangle + |1\rangle|\mathbf{x}\rangle) \\
&\xrightarrow{4} \frac{1}{2}|0\rangle(|\mathbf{z}\rangle + |\mathbf{x}\rangle) + \frac{1}{2}|1\rangle|(|\mathbf{z}\rangle - |\mathbf{x}\rangle).
\end{aligned} \tag{10}$$

Define the final state of Eq. 10 as $|g\rangle$ and the unitary transformation that generates such state, i.e. the modified Hadamard test as U_g. Let $|p\rangle$ and $|\tilde{p}\rangle$ be the normalized state of $\frac{|\mathbf{z}\rangle+|\mathbf{x}\rangle}{2}$ and $\frac{|\mathbf{z}\rangle-|\mathbf{x}\rangle}{2}$, respectively. Then the quantum state of Eq. 10 is written as

$$|g\rangle = \sin\theta|0\rangle|p\rangle + \cos\theta|1\rangle|\tilde{p}\rangle, \tag{11}$$

where $\cos\theta = \sqrt{\frac{1-\mathbf{Re}\langle\mathbf{z}|\mathbf{x}\rangle}{2}}$ contains the inner product score. Thus the inner product between $\mathbf{z}$ and $\mathbf{x}$ is given by

$$s = \mathbf{Re}\langle\mathbf{z}|\mathbf{x}\rangle = 1 - 2\cos^2\theta = -\cos 2\theta. \tag{12}$$

The quantum state $|g\rangle$ can be reformulated by applying Schmidt decomposition:

$$|g\rangle = \frac{-\mathbf{i}}{\sqrt{2}}(e^{\mathbf{i}\theta}|w_+\rangle - e^{-\mathbf{i}\theta}|w_-\rangle), \tag{13}$$

where $|w_\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle|p\rangle \pm \mathbf{i}|1\rangle|\tilde{p}\rangle)$ are two orthonormal quantum states. Then we estimate the phase θ through quantum phase estimation. The approximated inner product can be obtained from the measured output quantum states of the circuits. Define the unitary operating on the input register

$$G = U_g(I^{\otimes(1+k)} - 2|0\rangle^{\otimes(1+k)}\langle 0|^{\otimes(1+k)})U_g^\dagger(Z \otimes I^{\otimes k}). \tag{14}$$

It can be checked that

$$G|w_\pm\rangle = e^{\pm 2\mathbf{i}\theta}|w_\pm\rangle, \tag{15}$$

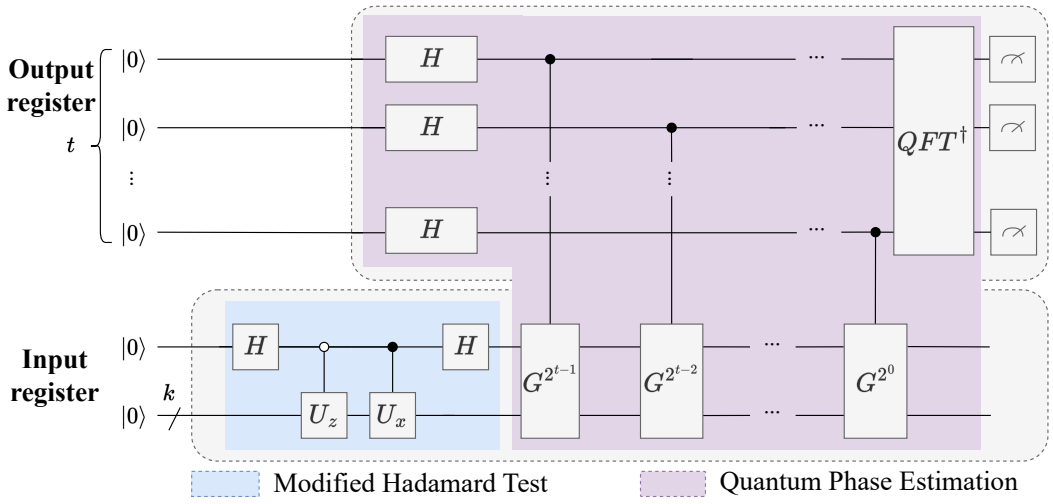

Figure 8. Quantum circuit of the 1-to-1 quantum inner product estimation.

where $|w_\pm\rangle$ are the eigenvectors of G and the corresponding eigenvalues are $e^{\pm 2\mathrm{i}\theta}$. After performing successive exponentialized unitary transformation G^{2^l} controlled by the l-th qubit in output register for $l \in [0, t-1]$, the quantum state of the output register is

$$\frac{1}{\sqrt{2}}(e^{\mathrm{i}\theta}|\boldsymbol{R}_+\rangle|w_+\rangle - e^{-\mathrm{i}\theta}|2^{t-1} - \boldsymbol{R}_+\rangle|w_-\rangle), \tag{16}$$

where $\boldsymbol{R}_+ = r^0 r^1 \dots r^{t-1} \in [0, 2^{t-1}]$ (written as the binary code) for all $r^l \in \{0, 1\}$. Since $|w_\pm\rangle$ are two orthonormal quantum states, we can employ amplitude estimation to make purification and the purified state is given as $e^{\mathrm{i}\theta}|\overline{\boldsymbol{R}_+}\rangle|w_+\rangle$. Finally, the approximation of the phase θ can be obtained by measuring all the qubits of the output register in computational basis:

$$\overline{\theta} = \pi(r^0 2^{-1} + r^1 2^{-2} + \dots + r^{t-1}2^{-t}) = \frac{\pi\overline{\boldsymbol{R}_+}}{2^t}. \tag{17}$$

Thus the inner product score is approximated by

$$s \approx \overline{s} = -\cos\frac{\pi\overline{\boldsymbol{R}_+}}{2^{t-1}}. \tag{18}$$

Note that in practice, the classical data must be normalized before the quantum data loading procedure to ensure the unit l_2 norm, such that the quantum estimation of the inner product is $\frac{\langle \mathbf{z}|\mathbf{x}\rangle}{||\mathbf{z}||||\mathbf{x}||}$. We assume that the norm of the vectors is already calculated and stored in classic memory before loading them onto the quantum device. We can retrieve the norm and recover the actual inner product $\langle \mathbf{z}|\mathbf{x}\rangle$ by scalar multiplication without additional overheads.

H.2.1 Complexity Analysis

Now we investigate the time complexity of the 1-to-1 QIP algorithm. The most time-consuming part of the modified Hadamard test is loading the classical data into their quantum representation by $U_\mathbf{x}$ and $U_\mathbf{z}$ both of which cost $O(\log d)$ running time, where d is the feature dimension of $\mathbf{x}$ and $\mathbf{z}$. Such loading procedure is also known as the ampltitude encoding. General results show that the amplitude encoding can be implemented in $O(\log nd)$ running time for n datapoints either by a dedicated quantum random access memory (QRAM) data structure [16] or by a parallel unary loader constructed from parameterized gates [13]. For the QPE, there have been various improvements [1] which try to soften the dependence on the inverse QFT, while retaining the accuracy guarantees offered by the QFT in estimating the phase value. We conclude that the complexity of inverse QFT is independent on the feature dimension, such that most of the time is spent on implementing the gate G which is equivalent to the unitary U_g multiplied by some element quantum gates. Thus the most time-consuming part of QPE is to implement the unitary U_g which costs $O(\frac{\log d}{\epsilon})$ time, where $\epsilon = |\theta - \overline{\theta}|$ is the precision indicating the difference between the theoretical phase θ and the estimated one $\overline{\theta}$. For the final stage and before the measurement, the amplitude estimation is employed and the time complexity is also logarithmically with respect to d [14]. In conclusion, the overall running time of the 1-to-1 QIP is $O(\frac{\log d}{\epsilon})$.

H.3. Proofs of the 1-to-N Quantum Inner Product Theorem

We go one step further and demonstrate how quantum parallelism can be used to estimate the inner product scores of a collection of paired embedding vectors, including one positive pair and several negative pairs. This allows us to introduce the quantum batch inner product estimation algorithm, which is essential for batch network embedding training.

For batch training, there are two main problems which can

not be circumvented by the quantum inner product estimation between two vectors without any modification. First, the vanilla Hadamard test can only be employed to estimate the fidelity (in this context, equivalent to inner product) between two quantum states, which is not suitable for both loading a collection of paired embedding vectors into a superposition state and evaluating the inner product score of many pairs of quantum states. Second, the output of the quantum circuit, i.e., the final quantum state containing the information of inner product score would be also a superposition state. The inner product scores stored in the superposition would be entangled and inseparable. A general way to retrieve the precise inner product score is quantum tomography [16]. However, the number of measurements required for the quantum tomography scales exponentially with respect to the number of qubits [9], such that the speedup benefited from the quantum parallelism may be offset. As a result, we suggest two solutions: the **modified Hadamard test** and **multiple phase estimation**, both of which serve as the foundation for the quantum batch inner product estimation.

The circuit of the proposed quantum batch inner product estimation is shown in Fig. 10. There are three quantum registers initiated by state $|0\rangle$, from top to bottom are index register, output register and input register. The index register has $\log N$ qubits where N is the batch size. Each qubit is transformed via a Hadamard gate and such operation produces a uniform superposition state $\frac{1}{\sqrt{N}}\sum_{i=0}^{N-1}|i\rangle$. Each component $|i\rangle$, that is the computational basis of the N-dimensional quantum state, is used to turn into an index aligned with the state containing the information of the inner product score and control the unitary operation in the other registers. One of the responsibilities of the input register is to perform modified Hadamard test. Different from the vanilla Hadamard test that can only load and estimate the inner product of two individual vectors, a number of classical vectors can be loaded and a batch of inner product scores will be generated into a superposition by the modified Hadamard test. Then we can retrieve the estimation of the scores by applying multiple phase estimation in the input and the output register. Compared to the QPE illustrated in Fig. 7, the most distinct part of multiple phase estimation is that N inner product scores are estimated and engraved at the same time in N mini registers, each of which has t qubits similar to the QPE. We describe modified Hadamard test and multiple phase estimation in more detail below.

H.3.1 Modified Hadamard Test

We zoom up the circuit layout of the modified Hadamard test and exhibit it in Fig 9. The Hadamard test circuit is modified by encoding a batch of embedding vectors into an entangled superposition state by introducing a small number of auxiliary qubits. There are $1+k$ qubits employed to perform the modified Hadamard test. The amplitude encoding unitary is

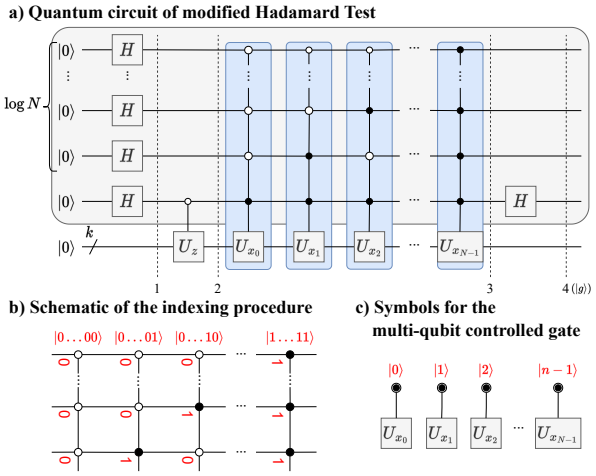

Figure 9. a) Quantum circuit of the modified Hadamard test for N pairs of quantum states. b) It depicts the layout of control operations which can be interpreted as an indexing procedure determined by the binary choice of each of the control qubits. c) Simplified symbol for the multi-qubit controlled gate.

controlled by the $1+\log N$ auxiliary qubits including all the qubits in the index register and one qubit in the input register. Concretely, we define $U_{\mathbf{x}_i}$ as the amplitude encoding unitary that transforms the zero states to the input quantum state by $U_{\mathbf{x}_i}|0\rangle^{\otimes k} = |\mathbf{x}_i\rangle$, where $k = \lceil \log d \rceil$ is the minimum number of qubits needed to load a d-dimensional vector. Thus the inner product score can be also written as $s_i = \langle \mathbf{z}|\mathbf{x}_i\rangle$. The resulting quantum state includes $2^k - d$ amplitudes that are redundant and have values of 0. We assume in the remaining of this papers that d is the power of 2 for simplicity, such that $k = \log d$ and no redundant amplitude exists. Note that $U_{\mathbf{x}_i}$ operates on the last k qubits when the first $\log N$ auxiliary qubits are in the state $|i\rangle$ and the last auxiliary qubit is in the state $|1\rangle$. While the amplitude encoding unitary $U_{\mathbf{z}}$ operates on the last k qubits when the last auxiliary qubit is in the state $|0\rangle$ regardless of the state of the other qubits. We give the output quantum state step by step as follows:

$$
\begin{aligned}
&\text{Input} = |0\rangle^{\otimes \log N}|0\rangle|0\rangle^{\otimes k} \\
&\xrightarrow{1} \frac{1}{\sqrt{N}}\sum_{i=0}^{N-1}|i\rangle\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle^{\otimes k} \\
&\xrightarrow{2} \frac{1}{\sqrt{N}}\sum_{i=0}^{N-1}|i\rangle\frac{1}{\sqrt{2}}(|0\rangle|\mathbf{z}\rangle + |1\rangle|0\rangle^{\otimes k}) \\
&\xrightarrow{3} \frac{1}{\sqrt{2N}}\sum_{i=0}^{N-1}(|i\rangle|0\rangle|\mathbf{z}\rangle + |i\rangle|1\rangle|\mathbf{x}_i\rangle) \\
&\xrightarrow{4} \frac{1}{\sqrt{N}}\sum_{i=0}^{N-1}|i\rangle\left[\frac{1}{2}|0\rangle(|\mathbf{z}\rangle + |\mathbf{x}_i\rangle) + \frac{1}{2}|1\rangle|(|\mathbf{z}\rangle - |\mathbf{x}_i\rangle)\right].
\end{aligned}
\tag{19}
$$

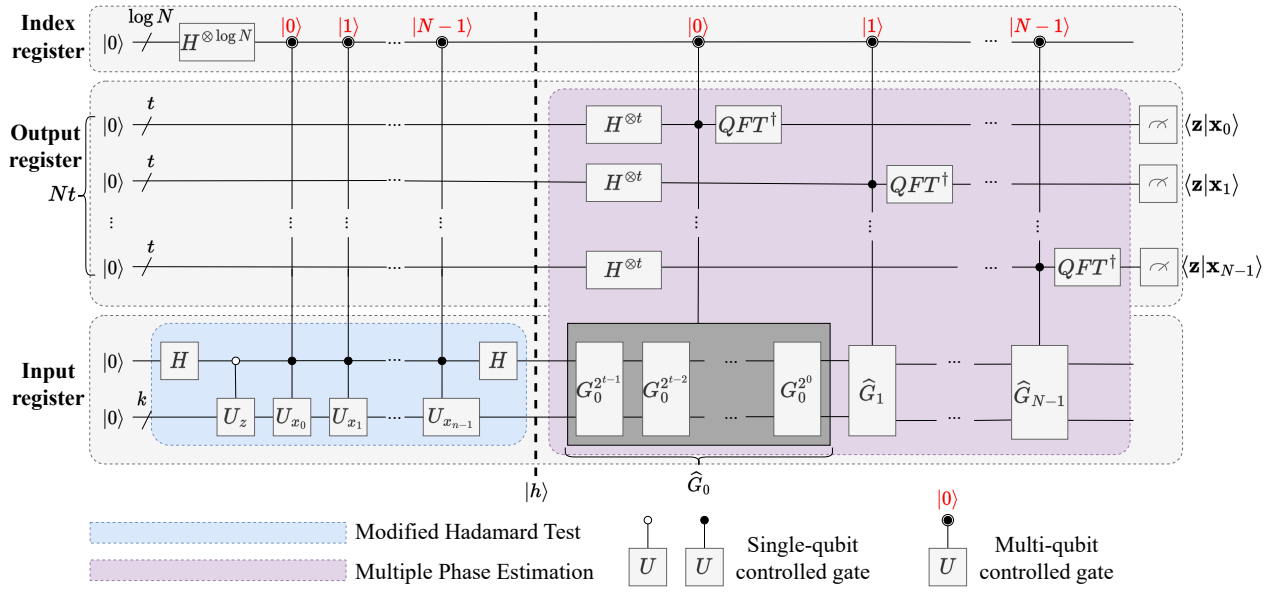

Figure 10. Quantum circuit of the quantum batch inner product estimation.

Let $|p_i\rangle$ is the normalized states of $\frac{|\mathbf{z}\rangle+|\mathbf{x}_i\rangle}{2}$, *i.e.*,

$$\begin{aligned}|p_i\rangle &= (\frac{|\mathbf{z}\rangle+|\mathbf{x}_i\rangle}{2})/\|\frac{|\mathbf{z}\rangle+|\mathbf{x}_i\rangle}{2}\|_2 \\ &= (\frac{|\mathbf{z}\rangle+|\mathbf{x}_i\rangle}{2})/\sqrt{\frac{1+\mathbf{Re}\langle\mathbf{z}|\mathbf{x}_i\rangle}{2}} \\ &= (\frac{|\mathbf{z}\rangle+|\mathbf{x}_i\rangle}{2})/\sin\theta_i,\end{aligned} \tag{20}$$

where $\sin\theta_i = \sqrt{\frac{1+\mathbf{Re}\langle\mathbf{z}|\mathbf{x}_i\rangle}{2}}$, thereby

$$\cos\theta_i = \sqrt{1-\sin^2\theta_i} = \sqrt{\frac{1-\mathbf{Re}\langle\mathbf{z}|\mathbf{x}_i\rangle}{2}}. \tag{21}$$

The inner product between $\mathbf{z}$ and $\mathbf{x}_i$ can be obtained by:

$$s_i = \mathbf{Re}\langle\mathbf{z}|\mathbf{x}_i\rangle = 1-2\cos^2\theta_i = -\cos 2\theta_i, \tag{22}$$

where $\theta_i \in [0, \frac{\pi}{2}]$. In addition, $|\tilde{p}_i\rangle$ is defined as the normalized states of $\frac{|\mathbf{z}\rangle-|\mathbf{x}_i\rangle}{2}$, *i.e.*,

$$|\tilde{p}_i\rangle = (\frac{|\mathbf{z}\rangle-|\mathbf{x}_i\rangle}{2})/\cos\theta_i. \tag{23}$$

The top $1+\log N$ qubits of Fig. 9 can be regarded as the index register, and the other qubits is the value register. Without loss of generality, we ignore the index register in the following analysis for simplicity, and let U_{g_i} be a unified unitary transformation of the modified Hadamard test indexed by the state $|i\rangle$, and then the quantum state of the value register can be written as:

$$\begin{aligned}|g_i\rangle &= U_{g_i}|0\rangle^{\otimes(1+k)} \\ &= \frac{1}{2}|0\rangle(|\mathbf{z}\rangle+|\mathbf{x}_i\rangle)+\frac{1}{2}|1\rangle|(|\mathbf{z}\rangle-|\mathbf{x}_i\rangle) \\ &= \sin\theta_i|0\rangle|p_i\rangle+\cos\theta_i|1\rangle|\tilde{p}_i\rangle.\end{aligned} \tag{24}$$

Applying the Schmidt decomposition method to the quantum state $|g_i\rangle$, the quantum state can be further written as

$$|g_i\rangle = \frac{-\mathbf{i}}{\sqrt{2}}(e^{\mathbf{i}\theta_i}|w_{i+}\rangle - e^{-\mathbf{i}\theta_i}|w_{i-}\rangle), \tag{25}$$

where $|w_{i\pm}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|p_i\rangle \pm \mathbf{i}|1\rangle|\tilde{p}_i\rangle)$. Now we can think about the index register again, the overall state is

$$|h\rangle = \frac{1}{\sqrt{N}}\sum_{i=0}^{N-1}|i\rangle(e^{\mathbf{i}\theta_i}|0\rangle^{\otimes Nt}|w_{i+}\rangle - e^{-\mathbf{i}\theta_i}|0\rangle^{\otimes Nt}|w_{i-}\rangle), \tag{26}$$

which is depicted in Fig. 10. So far, we have encoded the inner product results into a supposition state. The next stage is retrieving the results from quantum states. Rather than quantum tomography, we prefer to use QPE to get all of the inner product results concurrently. However, the applicable requirements of the QPE are violated by the entangled and interdependent nature of the superposition. We resort to the multiple quantum phase estimation introduced in the following Sec. H.3.2.

H.3.2 Multiple Quantum Phase Estimation

In the last Sec. H.3.1, we have introduced how to encode the N inner product scores of a batch of vectors into the

quantum state by the modified Hadamard test. Now we develop a novel quantum algorithm to output these scores in parallel.

We define the following unitary operating on the input register:

$$G_i = U_{g_i}(I^{\otimes(1+k)} - 2|0\rangle^{\otimes(1+k)}\langle 0|^{\otimes(1+k)})U_{g_i}^{\dagger}(Z \otimes I^{\otimes k}) \tag{27}$$

where I denotes the identity operator and $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$ is the Pauli-Z operator. Applying the unitary operator G_i defined above to the quantum state $|g_i\rangle$, we can get

$$G_i|w_{i\pm}\rangle = e^{\pm 2\mathbf{i}\theta_i}|w_{i\pm}\rangle. \tag{28}$$

It can be seen that $|w_{i\pm}\rangle$ are the eigenvectors of G_i and the corresponding eigenvalues are $e^{\pm 2\mathbf{i}\theta_i}$.

In the second stage, we use multiple phase estimation to estimate $\theta_0, \theta_1, \ldots, \theta_{N-1}$ in parallel. The first register contains nt auxiliary qubits initially in the state $|0\rangle$. It can be seen as a register stacked with N identical mini registers, each of which has t qubits on which a series of Hadamard gates and inverse QFT will be operate. The second register is initiated by state $|0\rangle$ comprised of $k+1$ qubits followed by the Hadamard test U_{g_i} and the resulting quantum state is $|g_i\rangle$. The control gate $\widehat{G}_i$ should be regarded as a composition of exponentialized controlled gates $G_i^{2^l}$ by viewing both the qubits in the index register and the l-th qubit in the i-th mini register as the control qubits, where $l \in [0, t-1]$. Consequently, the effect of the multiple phase estimation is equivalent to take the QPE on the state $|g_i\rangle$ conditioned on the index $|i\rangle$, and the phase estimation result $|\overline{\theta}_i\rangle$ will be encoded in the i-th mini register. By applying multiple phase estimation on the state $|h\rangle$ given in Eq. 26, and we discard the other mini registers except for the i-th mini register, the quantum state of the output register is transformed to

$$\frac{1}{\sqrt{N}}\sum_{i=0}^{N-1}|i\rangle(e^{\mathbf{i}\theta_i}|\boldsymbol{R}_{i+}\rangle|w_{i+}\rangle - e^{-\mathbf{i}\theta_i}|2^{t-1} - \boldsymbol{R}_{i+}\rangle|w_{i-}\rangle), \tag{29}$$

where $\boldsymbol{R}_{i+} = r_i^0 r_i^1 \ldots r_i^{t-1} \in [0, 2^{t-1}]$ (written as the binary code) for all $r_i^j \in \{0, 1\}$. Now we denote the notation $\mathcal{U}$ as a unified unitary transformation to generate the state given in Eq. 29. Then by using amplitude estimation algorithm [4], we can generate a state Δ-close to $\frac{1}{\sqrt{N}}\sum_{i=0}^{N-1} e^{\mathbf{i}\theta_i}|i\rangle|\overline{\boldsymbol{R}}_{i+}\rangle|w_{i+}\rangle$ in time $O(T(\mathcal{U})\ln(1/\Delta))$. The output register will output the state $|\overline{\boldsymbol{R}}_{0+}\overline{\boldsymbol{R}}_{1+}\overline{\boldsymbol{R}}_{2+}\ldots\overline{\boldsymbol{R}}_{(N-1)+}\rangle$ and the approximation of the phase θ_i is given by

$$\overline{\theta}_i = \pi(r_i^0 2^{-1} + r_i^1 2^{-2} + \cdots + r_i^{t-1}2^{-t}) = \frac{\pi\overline{\boldsymbol{R}}_{i+}}{2^t}, \tag{30}$$

which is supposed to be in $[0, \pi/2]$. The third and final stage is to read out the state of the output register by doing a measurement in the computational basis. All of the auxiliary qubits with respect to the output register of the multiple QPE will collapse to state $|0\rangle$ or $|1\rangle$. According to Eq. 22 and Eq. 30, by using $t + \lceil \log\left(2 + \frac{1}{2\epsilon}\right)\rceil$ auxiliary qubits to implement the inverse QFT, we can recover the inner product score of each of the node pairs from the measurement results on the first t qubits:

$$s_i \approx \overline{s}_i = -\cos\frac{\pi\overline{\boldsymbol{R}}_{i+}}{2^{t-1}}. \tag{31}$$

H.3.3 Complexity Analysis

We now investigate the complexity of the three main stages of the proposed algorithm. At the first stage where the modified Hadamard test is applied, the time complexity is highly dominated by the amplitude encoding procedure, which costs $O(\log Nd)$ time. For the second stage when employing multiple phase estimation, it costs $T(G_i^{2^l}) = O(\log Nd)$ time to implement exponentialized controlled gates which is defined by Eq. 27 where the most time-consuming part is the U_{g_i} operation. By the Theorem 4, the time complexity of multiple phase estimation is $O(\frac{\log Nd}{\epsilon})$ where ϵ is the precision denoting the difference between the theoretical value and the estimated value of the phase.

H.4. Proofs of the M-to-N Quantum Inner Product

Proof (A Sketch). The idea is to stack the 1-to-N circuits in the similar way as by adding an index register marked in orange, shown as following Eq. 32:

$$\begin{aligned}
&\text{Input} = |0\rangle^{\otimes\lceil\log N\rceil}|0\rangle^{\otimes\lceil\log M\rceil}|0\rangle^{\otimes MNt}|0\rangle|0\rangle^{\otimes\lceil\log d\rceil} \\
&\xrightarrow[\text{Hardamard Test}]{E(\{\mathbf{x}_i\},\{\mathbf{y}_j\})} \frac{1}{\sqrt{M}}\sum_{i=0}^{M-1}\bigotimes_{i=0}^{M-1}|i\rangle\left(\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle(e^{\mathbf{i}\theta_{ij}}|0\rangle^{\otimes MNt}|w_{ij+}\rangle\right. \\
&\left.- e^{-\mathbf{i}\theta_{ij}}|0\rangle^{\otimes MNt}|w_{ij-}\rangle)\right) \\
&\xrightarrow[QPE]{D(\{\mathbf{x}_i\},\{\mathbf{y}_j\})} \frac{1}{\sqrt{M}}\sum_{i=0}^{M-1}\bigotimes_{i=0}^{M-1}|i\rangle\left(\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle(e^{\mathbf{i}\theta_{ij}}|\boldsymbol{R}_{ij+}\rangle|w_{ij+}\rangle\right. \\
&\left.- e^{-\mathbf{i}\theta_{ij}}|\boldsymbol{R}_{ij+}\rangle|w_{ij-}\rangle)\right),
\end{aligned} \tag{32}$$

where $|\boldsymbol{R}_{ij+}\rangle$ is similar to $|\boldsymbol{R}_{i+}\rangle$ in Theorem 2 and the inner products can be recovered. □

The detailed proof: Given a bunch of classical vectors $\{\mathbf{z}_i\}_{i=0}^{M-1}$ and another bunch of classical vectors $\{\mathbf{x}_j\}_{j=0}^{N-1}$, where $\mathbf{z}_i \in \mathbb{R}^d$ and $\mathbf{x}_j \in \mathbb{R}^d$. We show that the quantum algorithm can output the inner product scores $s_{ij} = \mathbf{z}_i^{\top}\mathbf{x}_j$ at the same time with the complexity $O(\frac{\log MNd}{\epsilon})$. The quantum circuit of the M-to-N quantum batch inner product estimation is given in Fig. 11. Rather than using only one index register as the 1-to-N QIP does, we use two index registers in which the qubits serve as index states aligned with

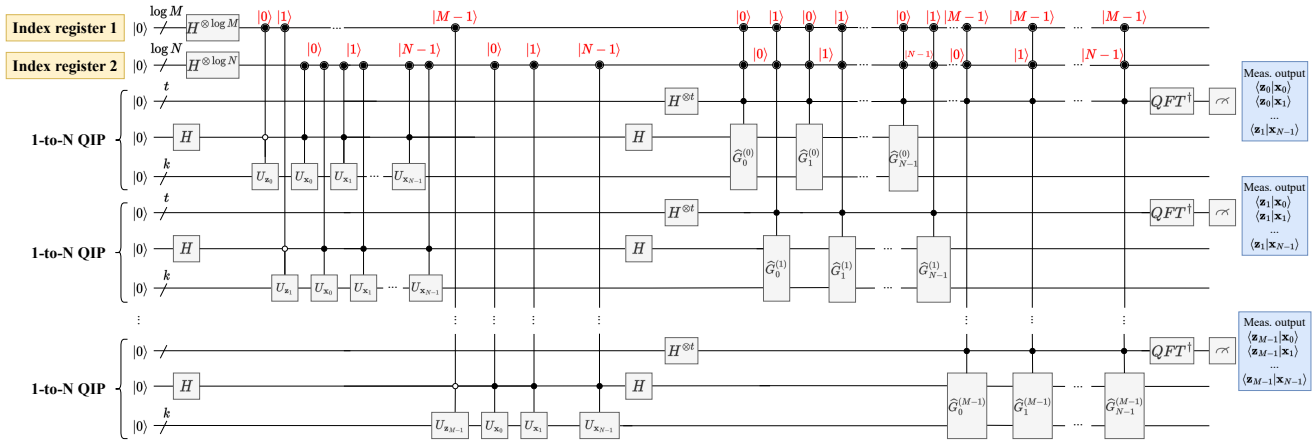

Figure 11. Quantum circuit of the M-to-N quantum batch inner product estimation.

the states containing the information of the inner product score and control the unitary operation in the other registers. Concretely, $\{\mathbf{z}_i\}_{i=0}^{M-1}$ are indexed by a superposition state $\frac{1}{\sqrt{M}}\sum_{i=0}^{M-1}|i\rangle$ in the first index register, while $\{\mathbf{x}_j\}_{j=0}^{N-1}$ are indexed by another superposition state $\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle$ in the second index register. Apart from the two index registers, the other quantum registers can be divided into M components, each of which is exactly the quantum circuit excluding the index register of 1-to-N QIP circuit. In other words, the i-th 1-to-N QIP circuit for $i \in [0, M-1]$ is used to estimate inner product scores between $\mathbf{z}_i$ and $\{\mathbf{x}_j\}_{j=0}^{N-1}$ by using modified Hadamard test and multiple phase estimation. By performing the modified Hadamard test conditioned on the two index registers, the transformation of the quantum state step by step is given as follows:

$$
\begin{aligned}
&\text{Input} = |0\rangle^{\otimes \log M}|0\rangle^{\otimes \log N}\left(|0\rangle|0\rangle^{\otimes k}\right)^{\otimes M} \\
&\xrightarrow{1} \frac{1}{\sqrt{MN}}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}|ij\rangle\left[\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)|0\rangle^{\otimes k}\right]^{\otimes M} \\
&\xrightarrow{2} \frac{1}{\sqrt{MN}}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\bigotimes_{i=0}^{M-1}\left[\frac{1}{\sqrt{2}}(|ij\rangle|0\rangle|\mathbf{z}_i\rangle+|ij\rangle|1\rangle|0\rangle^{\otimes k})\right] \\
&\xrightarrow{3} \frac{1}{\sqrt{M}}\sum_{i=0}^{M-1}\bigotimes_{i=0}^{M-1}|i\rangle\left\{\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle\left[\frac{1}{\sqrt{2}}(|0\rangle|\mathbf{z}_i\rangle+|1\rangle|\mathbf{x}_j\rangle)\right]\right\} \\
&\xrightarrow{4} \frac{1}{\sqrt{M}}\sum_{i=0}^{M-1}\bigotimes_{i=0}^{M-1}|i\rangle\left\{\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle\left[\frac{1}{2}|0\rangle|(|\mathbf{z}_i\rangle+|\mathbf{x}_j\rangle)\right.\right. \\
&\left.\left.+\frac{1}{2}|1\rangle(|\mathbf{z}_i\rangle-|\mathbf{x}_j\rangle)\right]\right\},
\end{aligned}
\tag{33}
$$

where $|ij\rangle$ is the abbreviation of $|i\rangle\otimes|j\rangle$. Notice that the final quantum state formulated in $\{\cdots\}$ is similar to the transformed state in Eq. 19. Akin to the method which has been introduced in Sec. H.3.2, we now implement the multiple phase estimation to retrieve inner product scores. The different part is that the number of control qubits of exponentialized controlled gates is increased (qubits in the first index register).

In the following we show that how to implement multiple phase estimation to estimate MN inner product scores. For convenience, we ignore the two index registers in the following analysis and only consider the quantum state of one of the 1-to-N QIP circuits. Let $|p_{ij}\rangle$ is the normalized states of $\frac{|\mathbf{z}_i\rangle+|\mathbf{x}_j\rangle}{2}$, and $|\tilde{p}_{ij}\rangle$ is the normalized states of $\frac{|\mathbf{z}_i\rangle-|\mathbf{x}_j\rangle}{2}$. Let $U_{g_{ij}}$ be a unified unitary transformation of the modified Hadamard test indexed by the state $|ij\rangle$, and then the quantum state of the value register in the i-th 1-to-N QIP circuit can be given by:

$$
\begin{aligned}
|g_{ij}\rangle &= U_{g_{ij}}|0\rangle^{\otimes(1+k)} \\
&= \frac{1}{2}|0\rangle(|\mathbf{z}_i\rangle+|\mathbf{x}_j\rangle)+\frac{1}{2}|1\rangle|(|\mathbf{z}_i\rangle-|\mathbf{x}_j\rangle) \\
&= \sin\theta_{ij}|0\rangle|p_{ij}\rangle+\cos\theta_{ij}|1\rangle|\tilde{p}_{ij}\rangle.
\end{aligned}
\tag{34}
$$

This quantum state also can be written as

$$
|g_{ij}\rangle = \frac{-\mathbf{i}}{\sqrt{2}}(e^{\mathbf{i}\theta_{ij}}|w_{ij+}\rangle - e^{-\mathbf{i}\theta_{ij}}|w_{ij-}\rangle) \tag{35}
$$

after applying Schmidt decomposition, where $|w_{ij\pm}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|p_{ij}\rangle \pm \mathbf{i}|1\rangle|\tilde{p}_{ij}\rangle)$. The value of θ_{ij} will be approximated by QIP. Define the following unitary operating on the input register

$$
G_j^{(i)} = U_{g_{ij}}(I^{\otimes(1+k)} - 2|0\rangle^{\otimes(1+k)}\langle 0|^{\otimes(1+k)})U_{g_{ij}}^{\dagger}(Z\otimes I^{\otimes k}). \tag{36}
$$

Next, we define the gate $\widehat{G}_j^{(i)}$ be a composition of exponentialized controlled gates $\left(G_j^{(i)}\right)^{2^l}$ where $l \in [0, t-1]$. The

three indices indicate that such kind of gates is controlled by the qubits in the index state $|i\rangle$ in the first register, the index state $|j\rangle$ in the second register, and the l-th qubit in the j-th mini register of the i-th 1-to-N QIP circuit .

Reconsider the state in the index registers, the quantum state of the i-th 1-to-N QIP circuit after performing multiple phase estimation is written as

$$|i\rangle\otimes\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle(e^{\mathbf{i}\theta_{ij}}|\boldsymbol{R}_{ij+}\rangle|w_{ij+}\rangle-e^{-\mathbf{i}\theta_{ij}}|2^{t-1}-\boldsymbol{R}_{ij+}\rangle|w_{ij-}\rangle), \tag{37}$$

where $\boldsymbol{R}_{ij}=r_{ij}^0r_{ij}^1\dots r_{ij}^{t-1}\in[0,2^{t-1}]$ for all $r_{ij}^l\in\{0,1\}$. Such quantum state is purified by amplitude estimation and the state is transformed to

$$|i\rangle\otimes\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}e^{\mathbf{i}\theta_{ij}}|j\rangle|\overline{\boldsymbol{R}}_{ij+}\rangle|w_{ij+}\rangle, \tag{38}$$

where $|\overline{\boldsymbol{R}}_{ij+}\rangle$ is an approximation of $|\boldsymbol{R}_{ij+}\rangle$. The output register will output the state $|\overline{\boldsymbol{R}}_{i0+}\overline{\boldsymbol{R}}_{i1+}\overline{\boldsymbol{R}}_{i2+}\dots\overline{\boldsymbol{R}}_{i(N-1)+}\rangle$ and the approximation of the phase θ_{ij} can be retrieved by measuring the output register in computational basis:

$$\overline{\theta}_{ij}=\pi(r_{ij}^02^{-1}+r_{ij}^12^{-2}+\dots+r_{ij}^{t-1}2^{-t})=\frac{\pi\overline{\boldsymbol{R}}_{ij+}}{2^t}, \tag{39}$$

which is supposed to be in $[0,\pi/2]$. Finally, we can recover the MN inner product scores between $\{\mathbf{z}_i\}_{i=0}^{M-1}$ and $\{\mathbf{x}_j\}_{j=0}^{N-1}$ from the measurement results on the first t qubits in every 1-to-N QIP circuit:

$$s_{ij}\approx\overline{s}_{ij}=-\cos\frac{\pi\overline{\boldsymbol{R}}_{ij+}}{2^{t-1}}. \tag{40}$$

H.4.1 Complexity Analysis

Here we theoretically analyze the complexity of the M-to-N quantum inner product algorithm. Retrospecting the circuit of 1-to-N quantum inner product in Fig. 10 and the circuit of M-to-N QIP in Fig. 11, the M-to-N quantum inner product can be viewed as stacking M 1-to-N QIP. In other words, the i-th ($i\in[0,M-1]$) stacked 1-to-N QIP is used to encode the inner product scores between vector $\mathbf{z}_i$ and $\{\mathbf{x}_j\}_{j=0}^{N-1}$ into an entangled state, and decode the scores by multiple phase estimation. For the modified Hadamard test which loading the inner product scores into an entangled state, it costs $O(\log MNd)$ time to encode MN vectors by amplitude encoding. For the multiple phase estimation, the time complexity of each stacked 1-to-N QIP is $O(\frac{\log Nd+\log M}{\epsilon})$ where $O(\log M)$ is the additional time for conducting unitary controlled by the $\log M$ qubits in the first index register, and ϵ is the precision parameter. Given that our quantum algorithm is designed to estimate the MN inner product scores in parallel, the time complexity of implementing the M-to-N QIP is $O(\frac{\log MNd}{\epsilon})$.

```python
import torch, torch_qip
from torch.autograd.function import Function

class QIPMatMul(Function):
    @staticmethod
    def forward(ctx, input, weight, num_qubits,
                sample_times, out_strategy):
        ctx.save_for_backward(input, weight)
        output = input.mm(weight)
        output = torch_qip.qip(output, num_qubits,
                               sample_times, out_strategy)
        return output

    @staticmethod
    def backward(ctx, grad_output):
        input, weight = ctx.saved_tensors
        grad_input = grad_weight = None
        if ctx.needs_input_grad[0]:
            grad_input = grad_output.mm(weight.t())
        if ctx.needs_input_grad[1]:
            grad_weight = input.t().mm(grad_output)
        return grad_input, grad_weight, None, None, None
```

Figure 12. An example of the code for differentiable quantum matrix multiplication operator based on `torch_qip`.

Table 9. The base implementation of machine learning models.

Model	Implementation
ProjUNN [17]	https://github.com/facebookresearch/projUNN. The repository includes implementations of ProjUNN on several ML models. The optimizer with two variants ProjUNN-D and ProjUNN-T for mapping matrices to unitaries is the key technique where QIP can be introduced.
node2vec [11]	https://github.com/xgfs/node2vec-c. The repository contains the C++ implementation of node2vec. In the embedding training, the computation of loss is mainly composed of inner product calculations.
K-Means	Implemented by the authors with numpy.

I. Enable Automatic Differentiation for QIP in Pytorch

We give the example of enabling automatic differentiation for QIP in deep learning in Fig. 12.

J. Experimental Details

In our numerical experiments, we substitute the classical inner product computations with the developed QIP operators. The base implementations of the including ML models and brief introductions are listed in Table 9.

J.1. Accuracy Study

Metrics. The used metrics mean squared error (MSE) and mean absolute error (MAE) are defined as follows:

$$\begin{aligned}MSE&=\frac{1}{|\mathcal{D}|}\sum_{i=1}^{|\mathcal{D}|}\left((\mathbf{x},\mathbf{y})_i-\widetilde{(\mathbf{x},\mathbf{y})}_i\right)^2,\\ MAE&=\frac{1}{|\mathcal{D}|}\sum_{i=1}^{|\mathcal{D}|}\left|(\mathbf{x},\mathbf{y})_i-\widetilde{(\mathbf{x},\mathbf{y})}_i\right|,\end{aligned} \tag{41}$$

where $\mathcal{D}$ is the dataset, i is the index of inner product $(\mathbf{x},\mathbf{y})$ in $\mathcal{D}$.

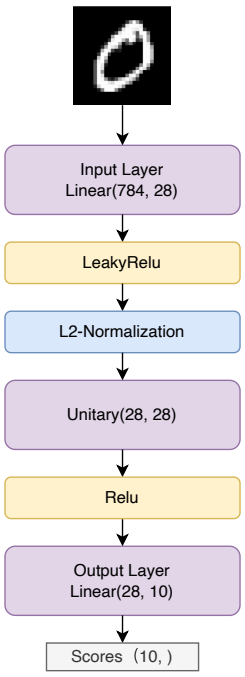

Figure 13. The MLP model for image classification on MNIST.

J.2. Training Unitary Neural Networks

Approach and Model Architecture. Considering that multiplying of a unitary matrix and a normalized vector can be achieved by QIP circuits without further normalization, we embed the unitary layer proposed in ProjUNN [17] in a typical multi-layer perceptron (MLP) model, as shown in Fig. 13. The input is a figure in array of size $(28 * 28,)$. Among the layers the second one adopts LeakyRelu to avoid an all-zero output; the third layer, an l_2-normalization layer, is adopted to normalize input vectors to length 1, for quantum state preparation; and the fourth layer is the unitary mapping. The output is the scores of different classes which is an array of size $(10,)$.

Dataset. MNIST is a dataset of handwritten digits, each has 28x28 pixels. It contains 60,000 training images and 10,000 testing images.

J.3. Embedding Learning by node2vec

J.3.1 Approach of Node2vec

Biased Random Walk. We define a network $G = (\mathcal{V}, \mathcal{E})$ by node set $\mathcal{V}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. First, node2vec defines a biased random walk method to generate random walk sequences by a combination of BFS and DFS. The second-order transition weight is defined as:

$$\alpha_{pq}(t,x) = \begin{cases} 1/p & d_{tx} = 0 \\ 1 & d_{tx} = 1 \\ 1/q & d_{tx} = 2 \\ 0 & \text{otherwise} \end{cases}, \tag{42}$$

where α_{pq} is the transition weight from node x to t, p and q are two parameters controlling the BFS/DFS preference, d_{tx} is the distance of the two nodes. According to the transition weight, a random walk sequence can be obtained by starting from an arbitrary node and walk for a certain number of steps.

Objective of Negative Sampling. For each node i in the sequence, we define i as the 'context' node, and nodes that appear within a slide window is called 'content' nodes. For such a pair of context node i and content node j, the objective based on negative sampling is defined as:

$$\log \sigma(\mathbf{x}_i^\top \mathbf{h}_j) + \sum_{b=1}^{B} \mathbb{E}_{j' \sim \mathbb{P}_{neg}(\cdot)} \log \sigma(-\mathbf{x}_i^\top \mathbf{h}_{j'}), \tag{43}$$

where $\mathbf{x}_i$ is the context embedding of node i and $\mathbf{h}_j$ is the content embedding of node j, $\sigma(\cdot)$ is the sigmoid function, $\mathbb{P}_{neg}$ is the distribution of negative samples (usually empirically set proportional to node degrees), and B is the number of negative samples for each positive sample.

So far, we can see that the main component of the computation of the objective Eq. 43 is the computation of inner products, which can be solved by our 1-to-many QIP circuit.

J.3.2 Experimental Setting

Parameter settings. We set the embedding dimension as 128, $p = 1$, $q = 1$ for the biased random walk, the number of negative samples for each positive sample as 5, and perform 80 times of random walk for each nodes with window size 10.

Datasets. BlogCatalog [30] is a social network of 10,312 bloggers who have social connections with each other. There are 39 different labels of the nodes; **PPI** [5] is a subgraph of the PPI network for Homo Sapiens which has 3890 nodes. There are 50 different labels of the nodes.

J.4. K-Means Clustering

Approach. K-Means is an unsupervised clustering algorithm, which has the following four steps: 1) Randomly initialize centroids, 2) calculate pairwise Euclidean distances between the data points, 3) find the closest centroid to a given data point, 4) create clusters, and 5) update the centroids as the means of each cluster. The algorithm repeats step 2) to step 5) until convergence or reaching maximum iterations. In the step 2), the Euclidean distance computation is time-consuming and will be intractable for a big dataset. It is also where the QIP operators be adopted. Considering two data points $\mathbf{x}_i$ and $\mathbf{x}_i$, the Euclidean distance can be obtained by several inner product terms:

$$\|\mathbf{x}_i - \mathbf{x}_j\|_2 = \Big(\langle \mathbf{x}_i, \mathbf{x}_i \rangle + \langle \mathbf{x}, \mathbf{x}_j \rangle - 2\langle \mathbf{x}_i, \mathbf{x}_j \rangle\Big)^{1/2}. \tag{44}$$

In this way, we use inner product operators to compute the Euclidean distance of data points, whose complexity only

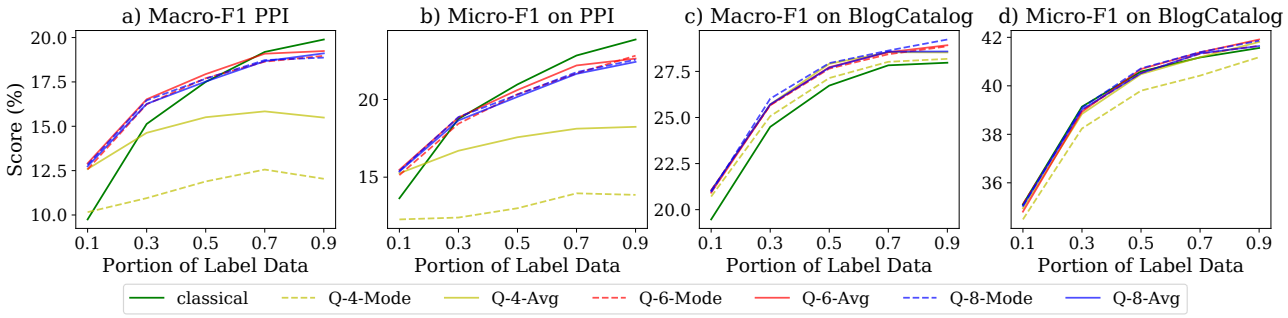

Figure 14. (Full version) Node classification on PPI and BlogCatalog.

depends on the complexity of inner products hence QIP works.

Metrics. The metrics, rand index (RI), normalized mutual information (NMI), and adjusted mutual information (AMI) are used to measure the similarity of two data clusterings.

By regarding the clustering problem as classification problem, RI can be defined as follows:

$$RI := \frac{TP + TN}{TP + FP + TN + FN}, \tag{45}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

Given the clustering results X and ground-truth clustering Y, NMI is defined as follows:

$$NMI := \frac{2I(X;Y)}{H(X) + H(Y)}, \tag{46}$$

where $I(\cdot;\cdot)$ is the mutual information, and $H(\cdot)$ represents marginal entropy. Then, AMI is defined as follows:

$$AMI := \frac{I(X;Y) - \mathbb{E}(I(X;Y))}{\frac{1}{2}(H(X) + H(Y)) - \mathbb{E}(I(X;Y))}. \tag{47}$$

K. Additional Experiments

K.1. Training Unitary Neural Networks by ProjUNN-T

We use ProjUNN-T [17] to optimize the untary layer. The model convergence curve is plotted in Fig. 15 and the test accuracy after 20 epochs of training is given in Table 10. We find that under ProjUNN-T, the performance of classical method keeps outperforming QIP-based methods on both model convergence speed and test accuracy, which is slightly different from the results of ProjUNN-D where QIP-based methods have chance to outperform the classical method. It indicates that, developing a effective optimizer for the normalized vectors or unitary matrices is also very important for QIP-based ML models.

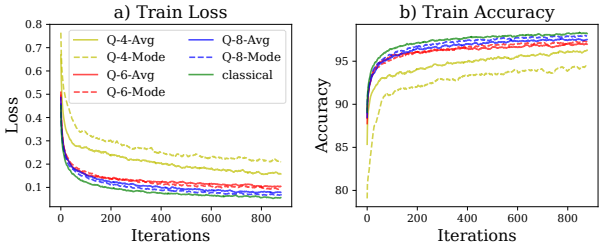

Figure 15. Convergence of using ProjUNN-T.

Table 10. Test precision of adopting ProjUNN-T on MNIST.

	Accuracy (%)
Q-4-Avg	94.76
Q-4-Mode	93.26
Q-6-Avg	95.40
Q-6-Mode	95.65
Q-8-Avg	95.99
Q-8-Mode	95.55
classical	**96.18**