Quantum machine learning

T. Garipov, P. Izmailov

May 15, 2015
Outline

1. A brief introduction to quantum mechanics
2. A brief introduction to quantum computations
3. Quantum algorithms for supervised and unsupervised machine learning
4. Algorithm implementation and experiments
Outline

1. A brief introduction to quantum mechanics
2. A brief introduction to quantum computations
3. Quantum algorithms for supervised and unsupervised machine learning
4. Algorithm implementation and experiments
Outline

1. A brief introduction to quantum mechanics
2. A brief introduction to quantum computations
3. Quantum algorithms for supervised and unsupervised machine learning
4. Algorithm implementation and experiments
1. A brief introduction to quantum mechanics
2. A brief introduction to quantum computations
3. Quantum algorithms for supervised and unsupervised machine learning
4. Algorithm implementation and experiments
Section 1

A brief introduction to quantum mechanics
Motivation

Consider the following experiment

Experiment

- $S$ is a source of electrons;
- $B$ is a screen with two holes in it;
- $C$ is another screen, with detectors covering its surface.
Now consider the plot of the probability of the electron to reach the screen $C$ at the height $x$ as a function of $x$.

Experiment shows, that

$$P \neq P_1 + P_2,$$

where $P$ is the probability distribution with both holes open (subplot a), and $P_i$ is the probability distribution with only $i$-th hole open (subplots b and c).
Now consider the plot of the probability of the electron to reach the screen \( C \) at the height \( x \) as a function of \( x \).

Experiment shows, that

\[ P \neq P_1 + P_2, \]

where \( P \) is the probability distribution with both holes open (subplot a), and \( P_i \) is the probability distribution with only \( i \)-th hole open (subplots b and c).
Postulate 1

Associated to any isolated physical system is a complex Hilbert space known as the state space of the system. The system is completely described by its state vector, which is a unit vector in the system’s state space.

We will denote the column vector $a$ from the state space as $|a\rangle$ and the corresponding row vector as $\langle a|$. We will also denote the inner product of $a$ and $b$ as $\langle a||b\rangle = \langle a|b\rangle$. 
Postulate 1

Associated to any isolated physical system is a complex Hilbert space known as the state space of the system. The system is completely described by its state vector, which is a unit vector in the system’s state space.

We will denote the column vector \( a \) from the state space as \( |a\rangle \) and the corresponding row vector as \( \langle a| \). We will also denote the inner product of \( a \) and \( b \) as \( \langle a||b\rangle = \langle a|b\rangle \).
The simplest quantum system is a *qubit*. It has a 2-dimensional state space $\mathcal{H}$ with orthonormal basis vectors $|0\rangle$ and $|1\rangle$.

The state of a qubit can be written in the form

$$|\psi\rangle = a|0\rangle + b|1\rangle,$$

where $a, b \in \mathbb{C}$.

As a state vector is a unit vector, we have

$$\langle \psi | \psi \rangle = |a|^2 + |b|^2 = 1.$$
The simplest quantum system is a *qubit*. It has a 2-dimensional state space $\mathcal{H}$ with orthonormal basis vectors $|0\rangle$ and $|1\rangle$.

The state of a qubit can be written in the form

$$|\psi\rangle = a|0\rangle + b|1\rangle,$$

where $a, b \in \mathbb{C}$.

As a state vector is a unit vector, we have

$$\langle \psi | \psi \rangle = |a|^2 + |b|^2 = 1.$$
Postulate 2
The evolution of a closed quantum system is described by a unitary transformation. That is, the state $|\psi\rangle$ of the system at time $t_1$ is related to the state $|\psi'\rangle$ of the system at time $t_2$ by a unitary operator $U$ which depends only on the times $t_1$ and $t_2$.

Postulate 2'
The evolution of the state of a closed quantum system is described by the Schrödinger equation,

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle,$$

where $\hbar$ is the Planks constant and $H$ is a hermitian operator (that might depend on time) known as the *Hamiltonian* of a closed system.
Postulate 2
The evolution of a closed quantum system is described by a unitary transformation. That is, the state \( |\psi\rangle \) of the system at time \( t_1 \) is related to the state \( |\psi'\rangle \) of the system at time \( t_2 \) by a unitary operator \( U \) which depends only on the times \( t_1 \) and \( t_2 \).

Postulate 2’
The evolution of the state of a closed quantum system is described by the Schrödinger equation,

\[
i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t)|\psi(t)\rangle,
\]

where \( \hbar \) is the Planks constant and \( H \) is a hermitian operator (that might depend on time) known as the Hamiltonian of a closed system.
Evolution

Consider the following example of the quantum system evolution.

**Hadamard gate**

The matrix representation of the Hadamard gate is the following

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

It’s easy to see, that

$$H|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2},$$

$$H|1\rangle = (|0\rangle - |1\rangle)/\sqrt{2}.$$
Postulate 3

Quantum measurements are described by a collection \( \{ M_m \} \) of measurement operators, acting on the state space of the system being measured. The index \( m \) refers to the measurement outcomes. If the system is in the state \( |\psi\rangle \) immediately before the measurement then the probability that \( m \) occurs is

\[
p(m) = \langle \psi | M_m^* M_m | \psi \rangle,
\]

and the state after the measurement is

\[
\frac{M_m | \psi \rangle}{\sqrt{\langle \psi | M_m^* M_m | \psi \rangle}}.
\]
As an example, consider the measurement of a qubit in the computational basis.

This is a measurement on a single qubit with two outcomes, defined by the operators

\[ M_0 = |0\rangle \langle 0|, \]
\[ M_1 = |1\rangle \langle 1|. \]

If the state of the system before the measurement is \( |\psi\rangle = a|0\rangle + b|1\rangle \), then the probability to obtain a 0 outcome, for example, is

\[ \langle \psi | M_0^* M_0 | \psi \rangle = \langle \psi | M_0 | \psi \rangle = |a|^2. \]
**Postulate 4**

The state space of a composite system is a tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 to \( n \), and \( i \)-th system is in the state \(|\psi_i\rangle\), then the joint state of the total system is \(|\psi_1\rangle \otimes \ldots \otimes |\psi_n\rangle\).

State space of a 2 qubit system

For example, for two qubits the state vector is of the form

\[ |\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle, \]

where \( |a|^2 + |b|^2 + |c|^2 + |d|^2 = 1 \).
Postulate 4

The state space of a composite system is a tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 to $n$, and $i$-th system is in the state $|\psi_i\rangle$, then the joint state of the total system is $|\psi_1\rangle \otimes \ldots \otimes |\psi_n\rangle$.

State space of a 2 qubit system

For example, for two qubits the state vector is of the form

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle,$$

where $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. 
Entagled states

*Entagled state* is a quantum state $|\psi\rangle$ of $S_1, S_2$ system that can’t be represented as a tensor product of any states $|\psi_{S_1}\rangle$ and $|\psi_{S_2}\rangle$ from $S_1$ and $S_2$ state spaces respectively.

$$|\psi_{S_1}\rangle \otimes |\psi_{S_2}\rangle \neq |\psi\rangle$$

**Entaglement example**

Consider a two-qubit state $|\phi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$. It is easy to show that this is an entagled state.
Section 2

A brief introduction to quantum computations
Main steps of quantum computation:

1. Initialization quantum states.
2. Performance of unitary operations on some subsets of qubits.
3. State measurement of the interesting set of qubits.

Almost always correct outcome of computations is obtained with some probability or algorithm aims to estimate some probability.
Quantum computer includes two parts.

- **Quantum part:**
  - Initialize quantum memory.
  - Perform an unitary operation on some subset of qubits.
  - Make a measurement of a quantum state.

- **Classical part:**
  - Choose the current unitary operation and a subset of qubits to perform it on.
  - Control the computation process.
Motivation of using quantum computers

- State of a set of $N$ qubits is a superposition of $2^N$ possible states.
- The result of a measurement is one of $2^N$ states, which is obtained with some probability.
- So the quantum memory is principally more powerful than classical memory.
- One quantum operation makes much more work than one classical computer operation.
- Main source of speed-up is the existence of entangled states.
Example of quantum algorithm

SWAP-test

**Task:** determine are two quantum states $|\phi\rangle$ and $|\psi\rangle$ different or equal.

**Solution:**

1. Add an ancillary qubit in the state $|0\rangle$.
2. Apply the Hadamard transform to the ancillary qubit.
   
   \[
   H = \begin{bmatrix}
   \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
   \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}
   \end{bmatrix}
   \]
3. Perform the controlled swap procedure (CSWAP).
4. Apply another Hadamard transform to the ancillary qubit.
5. Measure state of the ancillary qubit. If $|\phi\rangle = |\psi\rangle$ then $|0\rangle$ is obtained with probability 1.
Example of quantum algorithm

SWAP-test

Task: determine are two quantum states $|\phi\rangle$ and $|\psi\rangle$ different or equal.

Solution:

1. Add an ancillary qubit in the state $|0\rangle$.
2. Apply the Hadamard transform to the ancillary qubit.

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

3. Perform the controlled swap procedure (CSWAP).
4. Apply another Hadamard transform to the ancillary qubit.
5. Measure state of the ancillary qubit. If $|\phi\rangle = |\psi\rangle$ then $|0\rangle$ is obtained with probability 1.
Example of quantum algorithm

SWAP-test

Task: determine are two quantum states $|\phi\rangle$ and $|\psi\rangle$ different or equal.

Solution:

1. Add an ancillary qubit in the state $|0\rangle$.
2. Apply the Hadamard transform to the ancillary qubit.

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

3. Perform the controlled swap procedure (CSWAP).
4. Apply another Hadamard transform to the ancillary qubit.
5. Measure state of the ancillary qubit. If $|\phi\rangle = |\psi\rangle$ then $|0\rangle$ is obtained with probability 1.
Example of quantum algorithm

SWAP-test

**Task:** determine are two quantum states $|\phi\rangle$ and $|\psi\rangle$ different or equal.

**Solution:**

1. Add an ancillary qubit in the state $|0\rangle$.
2. Apply the Hadamard transform to the ancillary qubit.
   \[
   H = \begin{bmatrix}
   \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
   \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}
   \end{bmatrix}
   \]
3. Perform the controlled swap procedure (CSWAP).
4. Apply another Hadamard transform to the ancillary qubit.
5. Measure state of the ancillary qubit. If $|\phi\rangle = |\psi\rangle$ then $|0\rangle$ is obtained with probability 1.
Task: determine if two quantum states $|\phi\rangle$ and $|\psi\rangle$ are different or equal.

Solution:

1. Add an ancillary qubit in the state $|0\rangle$.
2. Apply the Hadamard transform to the ancillary qubit.

\[ H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \]

3. Perform the controlled swap procedure (CSWAP).
4. Apply another Hadamard transform to the ancillary qubit.
5. Measure state of the ancillary qubit. If $|\phi\rangle = |\psi\rangle$ then $|0\rangle$ is obtained with probability 1.
Example of quantum algorithm

SWAP-test

The evolution through algorithm steps:

\[ |0\rangle|φ\rangle|ψ\rangle \xrightarrow{H} \frac{|0\rangle + |1\rangle}{\sqrt{2}} |φ\rangle|ψ\rangle \xrightarrow{CSWAP} \frac{|0\rangle|φ\rangle|ψ\rangle + |1\rangle|ψ\rangle|φ\rangle}{\sqrt{2}} \xrightarrow{H} \]

\[ |0\rangle \left[ |φ\rangle|ψ\rangle + |ψ\rangle|φ\rangle \right] + |1\rangle \left[ |φ\rangle|ψ\rangle - |ψ\rangle|φ\rangle \right] \]

\[ \frac{2}{2} \]

The probability of passing the test:

\[ P = \frac{1}{4} (\langle φ | ψ \rangle + \langle ψ | φ \rangle)(|φ\rangle|ψ\rangle + |ψ\rangle|φ\rangle) = \frac{1 + |\langle φ | ψ \rangle|^2}{2} \]
### Famous quantum algorithms

#### Grover's search algorithm

**Task:** Find a solution of an equation \( f(x) = 1 \), where \( x \in \{0, 1\}^n \), and \( f(x) \) is a boolean function.

**Time:** \( O(\sqrt{2^N}) \)

**Memory:** \( O(N) \)

#### Shor's factorization algorithm

**Task:** Factorization of number \( N \).

**Time:** \( O([\log N]^3) \)  
**Goodbye RSA!**

**Memory:** \( O(\log N) \)

#### Quantum algorithm for linear systems of equations

**Task:** Solving a linear system of equations with \( N \) variables.

**Assumption:** The system is sparse and has a low conditional number \( \kappa \).

**Time:** \( O(\kappa \log N) \) (versus classical \( O(\kappa N) \))
Why don’t we have a quantum computer

Theory of quantum computations is quite well developed and it have shown possible huge potential of quantum computers compared to classical computers.

But modern implementations of quantum computer are strongly restricted and are used only for specialized tasks. Building a full quantum computer as a real physical device is a fundamental problem of 21th century physics.
Motivation of quantum machine learning:

- There is a challenge that machine learning with rapidly growing “big data” could become intractable for classical computers.
- Manipulation with high-dimensional vector is a core routine for machine learning algorithms.
- Quantum computers appear to be good at such manipulations.

Vector storing

Storing a representation of a $2^N$-dimensional unit vector requires only $N$ qubits. And such representations could be constructed in $O(N)$ time.

- Quantum machine learning may provide an exponential speed-up over known algorithms for problems involving evaluating distances and inner products between large vectors.
Section 3

Quantum algorithms for supervised and unsupervised machine learning
We assume, that the data sets that consist of vectors and collections of vectors are originally stored in qRAM.

Then constructing the $\log_2 N$ qubit quantum state $|\psi\rangle = |\psi\rangle - \frac{1}{2}\psi$ takes $O(\log_2 N)$ steps.
Supervised cluster assignment problem

Consider the task of assigning a post-processed vector $u \in \mathbb{R}^n$ to one of two sets $V, W$ given $M$ representatives of each classes.

A common method for such an assignment is evaluating the distances

$$
\left| u - \frac{1}{M} \sum_{j=1}^{M} v_j \right|
$$

and

$$
\left| u - \frac{1}{M} \sum_{j=1}^{M} w_j \right|
$$

and assign the vector to the class, for which this distance is smaller.
Supervised cluster assignment

distance determining algorithm

1. Create an ancillary variable with $M + 1$ states.
2. Construct the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle|u\rangle + \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |j\rangle|v_j\rangle).$$

3. Use the swap-test to determine whether the ancillary variable is in the state

$$|\varphi\rangle = \frac{1}{\sqrt{Z}} (|u\rangle|0\rangle - \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |v_j\rangle|j\rangle),$$

where $Z = |u|^2 + (1/M) \sum_{j=1}^{M} |v_j|^2$.

The probability of success in this measurement is equal to the distance

$$|u - (1/M) \sum_{j=1}^{M} v_j|^2$$

divided by $Z$. Thus, performing this measurement several times we can estimate the distance.
Supervised cluster assignment

distance determining algorithm

1. Create an ancillary variable with $M + 1$ states.
2. Construct the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle |u\rangle + \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |j\rangle |v_j\rangle).$$

3. Use the swap-test to determine, whether the ancillary variable is in the state

$$|\varphi\rangle = \frac{1}{\sqrt{Z}} (|u\rangle |0\rangle - \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |v_j\rangle |j\rangle),$$

where $Z = |u|^2 + (1/M) \sum_{j=1}^{M} |v_j|^2$.

The probability of success in this measurement is equal to the distance $|u - (1/M) \sum_{j=1}^{M} v_j|^2$ divided by $Z$. Thus, performing this measurement several times we can estimate the distance.
Supervised cluster assignment

distance determining algorithm

1. Create an ancillary variable with $M + 1$ states.
2. Construct the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |j\rangle|v_j\rangle).$$

3. Use the swap-test to determine, whether the ancillary variable is in the state

$$|\varphi\rangle = \frac{1}{\sqrt{Z}}(|u|0\rangle - \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |v_j|j\rangle),$$

where $Z = |u|^2 + \left(\frac{1}{M}\right) \sum_{j=1}^{M} |v_j|^2$.

The probability of success in this measurement is equal to the distance $|u - (1/M) \sum_{j=1}^{M} v_j|^2$ divided by $Z$. Thus, performing this measurement several times we can estimate the distance.
Supervised cluster assignment

distance determining algorithm

1. Create an ancillary variable with $M + 1$ states.
2. Construct the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle|u\rangle + \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |j\rangle|v_j\rangle).$$

3. Use the swap-test to determine, whether the ancillary variable is in the state

$$|\varphi\rangle = \frac{1}{\sqrt{Z}} (|u\rangle|0\rangle - \frac{1}{\sqrt{M}} \sum_{j=1}^{M} |v_j\rangle|j\rangle),$$

where $Z = |u|^2 + \frac{1}{M} \sum_{j=1}^{M} |v_j|^2$.

The probability of success in this measurement is equal to the distance

$$|u - \frac{1}{M} \sum_{j=1}^{M} v_j|^2$$

divided by $Z$. Thus, performing this measurement several times we can estimate the distance.
Adiabatic quantum computations

As already been mentioned, the time evolution of a quantum system is described by the Schrödinger equation:

\[ i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle, \]

where \( H(t) \) is the Hamiltonian of the system. A Hamiltonian is described by a hermitian matrix, whose eigenvectors represent the eigenstates of the system. The corresponding eigenvalues refer to the different energies of the eigenstates. The state with the lowest energy is called the ground state of the system.

Quantum adiabatic theorem

A physical system that is initially in its ground state, tends to stay in this lowest energy state, provided that the Hamiltonian of the system is changed ’slowly enough’.
As already been mentioned, the time evolution of a quantum system is described by the Schrödinger equation:

\[
\frac{i\hbar}{dt} \langle \psi(t) \rangle = H(t) \langle \psi(t) \rangle,
\]

where \( H(t) \) is the Hamiltonian of the system. A Hamiltonian is described by a hermitian matrix, whose eigenvectors represent the eigenstates of the system. The corresponding eigenvalues refer to the different energies of the eigenstates. The state with the lowest energy is called the ground state of the system.

**Quantum adiabatic theorem**

A physical system that is initially in its ground state, tends to stay in this lowest energy state, provided that the Hamiltonian of the system is changed 'slowly enough'.
Consider the following algorithm for minimization an objective function $f$.

**Adiabatic quantum computations framework**

1. Define the Hamiltonian $H_0$ so that it’s eigenstates are easy to compute. Set the system to the ground state of $H_0$.

2. Define the final Hamiltonian as

   $$H_f = \sum_{z \in \{0, 1\}^n} f(z)|z\rangle\langle z|.$$ 

3. Gradually transform the Hamiltonian of the system from $H_0$ to $H_f$.

4. The final state of the system is the minimizer of the objective function.
Adiabatic quantum computations

Consider the following algorithm for minimization an objective function $f$.

### Adiabatic quantum computations framework

1. Define the Hamiltonian $H_0$ so that its eigenstates are easy to compute. Set the system to the ground state of $H_0$.
2. Define the final Hamiltonian as
   \[
   H_f = \sum_{z \in \{0,1\}^n} f(z) |z\rangle \langle z|.
   \]
3. Gradually transform the Hamiltonian of the system from $H_0$ to $H_f$.
4. The final state of the system is the minimizer of the objective function.
Consider the following algorithm for minimization an objective function $f$.

### Adiabatic quantum computations framework

1. Define the Hamiltonian $H_0$ so that it’s eigenstates are easy to compute. Set the system to the ground state of $H_0$.

2. Define the final Hamiltonian as

   $$H_f = \sum_{z \in \{0,1\}^n} f(z) |z\rangle\langle z|.$$

3. Gradually transform the Hamiltonian of the system from $H_0$ to $H_f$.

4. The final state of the system is the minimizer of the objective function.
Consider the following algorithm for minimization of an objective function $f$.

**Adiabatic quantum computations framework**

1. Define the Hamiltonian $H_0$ so that its eigenstates are easy to compute. Set the system to the ground state of $H_0$.
2. Define the final Hamiltonian as

   $$H_f = \sum_{z \in \{0,1\}^n} f(z) |z\rangle \langle z|.$$ 

3. Gradually transform the Hamiltonian of the system from $H_0$ to $H_f$.
4. The final state of the system is the minimizer of the objective function.
Consider the task of assigning $M$ vectors to $k$ clusters in a way, that minimizes the average distance to the centroid of the cluster.

The standard classical algorithm for doing this is the Lloyd’s algorithm:

**Lloyd algorithm**

1. Choose $k$ initial centroids.
2. Assign each vector to the cluster with the closest mean.
3. Recalculate the centroids of the clusters.
4. Repeat steps (2–3) until a stationary assignment is attained.
Consider the task of assigning $M$ vectors to $k$ clusters in a way, that minimizes the average distance to the centroid of the cluster.

The standard classical algorithm for doing this is the Lloyd’s algorithm:

**Lloyd algorithm**

1. Choose $k$ initial centroids.
2. Assign each vector to the cluster with the closest mean.
3. Recalculate the centroids of the clusters.
4. Repeat steps (2–3) until a stationary assignment is attained.
Consider the task of assigning $M$ vectors to $k$ clusters in a way, that minimizes the average distance to the centroid of the cluster.

The standard classical algorithm for doing this is the Lloyd’s algorithm:

**Lloyd algorithm**

1. Choose $k$ initial centroids.
2. Assign each vector to the cluster with the closest mean.
3. Recalculate the centroids of the clusters.
4. Repeat steps (2–3) until a stationary assignment is attained.
Unsupervised quantum learning

Consider the task of assigning $M$ vectors to $k$ clusters in a way, that minimizes the average distance to the centroid of the cluster.

The standard classical algorithm for doing this is the Lloyd’s algorithm:

**Lloyd algorithm**

1. Choose $k$ initial centroids.
2. Assign each vector to the cluster with the closest mean.
3. Recalculate the centroids of the clusters.
4. Repeat steps (2–3) until a stationary assignment is attained.
Unsupervised quantum learning

The quantum Lloyd algorithm

1. Choose \( k \) initial centroids \( i_1, \ldots, i_k \).

2. Using the adiabatic quantum computations find the re-clustering that minimizes the distances from each point to the corresponding centroid. The result is a state \( |\psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{c,j \in c} |c\rangle |j\rangle \).

3. Construct multiple copies of this state and perform projective measurements of the \( |c\rangle \). Thus we obtain individual cluster states \( |\varphi_1^c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle \).

4. Use the adiabatic quantum computations to find the re-clustering that minimizes the distances from each point to the corresponding centroid of the clusters, obtained on the previous iteration. The result is a state \( |\psi_i\rangle \).

5. Construct multiple copies of the state \( |\psi_i\rangle \) and perform projective measurements of the \( |c\rangle \). Thus we obtain individual cluster states \( |\varphi_i^c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle \).

6. Repeat the steps (4 – 5) until two successive states \( |\psi_i\rangle \) and \( |\psi_{i+1}\rangle \) coincide, which can be verified using the swap test.
Unsupervised quantum learning

The quantum Lloyd algorithm

1. Choose $k$ initial centroids $i_1, \ldots, i_k$.
2. Using the adiabatic quantum computations find the re-clustering that minimizes the distances from each point to the corresponding centroid. The result is a state $|\psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{c,j \in c} |c\rangle |j\rangle$.
3. Construct multiple copies of this state and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi_c^1\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle$.
4. Use the adiabatic quantum computations to find the re-clustering that minimizes the distances from each point to the corresponding centroid of the clusters, obtained on the previous iteration. The result is a state $|\psi_i\rangle$.
5. Construct multiple copies of the state $|\psi_i\rangle$ and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi^i_c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle$.
6. Repeat the steps (4–5) until two successive states $|\psi_i\rangle$ and $|\psi_{i+1}\rangle$ coincide, which can be verified using the swap test.
Unsupervised quantum learning

The quantum Lloyd algorithm

1. Choose $k$ initial centroids $i_1, \ldots, i_k$.

2. Using the adiabatic quantum computations find the re-clustering that minimizes the distances from each point to the corresponding centroid. The result is a state $|\psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{c,j \in c} |c\rangle |j\rangle$.

3. Construct multiple copies of this state and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi_1^c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle$.

4. Use the adiabatic quantum computations to find the re-clustering that minimizes the distances from each point to the corresponding centroid of the clusters, obtained on the previous iteration. The result is a state $|\psi_i\rangle$.

5. Construct multiple copies of the state $|\psi_i\rangle$ and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi_i^c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle$.

6. Repeat the steps (4–5) until two successive states $|\psi_i\rangle$ and $|\psi_{i+1}\rangle$ coincide, which can be verified using the swap test.
The quantum Lloyd algorithm

1. Choose $k$ initial centroids $i_1, \ldots, i_k$.
2. Using the adiabatic quantum computations find the re-clustering that minimizes the distances from each point to the corresponding centroid. The result is a state $|\psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{c,j \in c} |c\rangle |j\rangle$.
3. Construct multiple copies of this state and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi_1^c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle$.
4. Use the adiabatic quantum computations to find the re-clustering that minimizes the distances from each point to the corresponding centroid of the clusters, obtained on the previous iteration. The result is a state $|\psi_i\rangle$.
5. Construct multiple copies of the state $|\psi_i\rangle$ and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi_i^c\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle$.
6. Repeat the steps (4–5) until two successive states $|\psi_i\rangle$ and $|\psi_{i+1}\rangle$ coincide, which can be verified using the swap test.
Unsupervised quantum learning

The quantum Lloyd algorithm

1. Choose $k$ initial centroids $i_1, \ldots, i_k$.
2. Using the adiabatic quantum computations find the re-clustering that minimizes the distances from each point to the corresponding centroid. The result is a state $|\psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{c,j\in c} |c\rangle |j\rangle$
3. Construct multiple copies of this state and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi^c_1\rangle = \frac{1}{\sqrt{M}} \sum_{j\in c} |j\rangle$.
4. Use the adiabatic quantum computations to find the re-clustering that minimizes the distances from each point to the corresponding centroid of the clusters, obtained on the previous iteration. The result is a state $|\psi_i\rangle$.
5. Construct multiple copies of the state $|\psi_i\rangle$ and perform projective measurements of the $|c\rangle$. Thus we obtain individual cluster states $|\varphi^c_i\rangle = \frac{1}{\sqrt{M}} \sum_{j\in c} |j\rangle$.
6. Repeat the steps (4–5) until two successive states $|\psi_i\rangle$ and $|\psi_{i+1}\rangle$ coincide, which can be verified using the swap test.
The quantum Lloyd algorithm

1. Choose \( k \) initial centroids \( i_1, \ldots, i_k \).
2. Using the adiabatic quantum computations find the re-clustering that minimizes the distances from each point to the corresponding centroid. The result is a state \( |\psi_1\rangle = \frac{1}{\sqrt{M}} \sum_{c,j \in c} |c\rangle |j\rangle \).
3. Construct multiple copies of this state and perform projective measurements of the \( |c\rangle \). Thus we obtain individual cluster states \( |\varphi^c_1\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle \).
4. Use the adiabatic quantum computations to find the re-clustering that minimizes the distances from each point to the corresponding centroid of the clusters, obtained on the previous iteration. The result is a state \( |\psi_i\rangle \).
5. Construct multiple copies of the state \( |\psi_i\rangle \) and perform projective measurements of the \( |c\rangle \). Thus we obtain individual cluster states \( |\varphi^c_i\rangle = \frac{1}{\sqrt{M}} \sum_{j \in c} |j\rangle \).
6. Repeat the steps (4 – 5) until two successive states \( |\psi_i\rangle \) and \( |\psi_{i+1}\rangle \) coincide, which can be verified using the swap test.
Unsupervised quantum learning

The output of the algorithm is a state

$$|\chi\rangle = \frac{1}{\sqrt{M}} \sum_j |c_j\rangle |j\rangle$$

that contains the labels $j$ of vectors, correlated with their cluster assignments $c_j$. Sampling from this state we obtain a statistical picture of the clustering.
For the vector assignment algorithm the computational complexity estimate is $O(\log MN)$.

For the quantum Lloyd algorithm the computational complexity of constructing the state $|\chi\rangle$ estimation is $O(k \log kMN)$. 
For the vector assignment algorithm the computational complexity estimate is $O(\log MN)$.

For the quantum Lloyd algorithm the computational complexity of constructing the state $|\chi\rangle$ estimation is $O(k \log kMN)$. 
Section 4

Algorithm implementation and experiments
Task: Classification of $N$-dimensional vectors (where $N = 2, 4, 8$) to two classes with one reference vector in each class.

Notation:
- Classes are labeled as $A$ and $B$.
- $\vec{v}_A$ and $\vec{v}_B$ are reference vectors.
- $\vec{u}$ is the new sample vector.

Classification of the new sample is done by comparing distances:

$$D_A = |\vec{u} - \vec{v}_A|, \quad D_B = |\vec{u} - \vec{v}_B|$$
**Experiments: algorithm**

**Computing distance** $|\vec{u} - \vec{v}|$:

1. Represent vectors with quantum states:
   $$\vec{u} = |u|u\rangle, \quad \vec{v} = |v|v\rangle$$

2. Add an ancillary qubit and create an entangled state:
   $$|\phi\rangle = (|0\rangle_{anc} |u\rangle_{new} + |1\rangle_{anc} |v\rangle_{ref})/\sqrt{2}$$

3. Make a measurement on the ancillary qubit, projecting it onto the state:
   $$|\psi\rangle = (|u\rangle|0\rangle - |v\rangle|1\rangle)/\sqrt{|u|^2 + |v|^2}$$

4. Estimate the success probability of measurement $p$ by repeated measurements.

5. Calculate $|\vec{u} - \vec{v}|$ using $p$. 
Experiments: algorithm

**Computing distance** $|\vec{u} - \vec{v}|$:

1. **Represent vectors with quantum states:**

$$\vec{u} = |u||u\rangle, \quad \vec{v} = |v||v\rangle$$

2. **Add an ancillary qubit and create an entangled state:**

$$|\phi\rangle = (|0\rangle_{anc}|u\rangle_{new} + |1\rangle_{anc}|v\rangle_{ref})/\sqrt{2}$$

3. **Make a measurement on the ancillary qubit, projecting it onto the state:**

$$|\psi\rangle = (|u||0\rangle - |v||1\rangle)/\sqrt{|u|^2 + |v|^2}$$

4. **Estimate the success probability of measurement $p$ by repeated measurements.**

5. **Calculate $|\vec{u} - \vec{v}|$ using $p$.**
**Experiments: algorithm**

**Computing distance** $|\vec{u} - \vec{v}|$:

1. Represent vectors with quantum states:

   \[ \vec{u} = |u| |u\rangle, \quad \vec{v} = |v| |v\rangle \]

2. Add an ancillary qubit and create an entangled state:

   \[ |\phi\rangle = (|0\rangle_{anc} |u\rangle_{new} + |1\rangle_{anc} |v\rangle_{ref})/\sqrt{2} \]

3. Make a measurement on the ancillary qubit, projecting it onto the state:

   \[ |\psi\rangle = (|u\rangle |0\rangle - |v\rangle |1\rangle)/\sqrt{|u|^2 + |v|^2} \]

4. Estimate the success probability of measurement $p$ by repeated measurements.

5. Calculate $|\vec{u} - \vec{v}|$ using $p$. 
Experiments: algorithm

**Computing distance** $|\vec{u} - \vec{v}|$:

1. Represent vectors with quantum states:

   $$\vec{u} = |u||u\rangle, \quad \vec{v} = |v||v\rangle$$

2. Add an ancillary qubit and create an entangled state:

   $$|\phi\rangle = (|0\rangle_{anc}|u\rangle_{new} + |1\rangle_{anc}|v\rangle_{ref})/\sqrt{2}$$

3. Make a measurement on the ancillary qubit, projecting it onto the state:

   $$|\psi\rangle = (|u||0\rangle - |v||1\rangle)/\sqrt{|u|^2 + |v|^2}$$

4. Estimate the success probability of measurement $p$ by repeated measurements.

5. Calculate $|\vec{u} - \vec{v}|$ using $p$.  

T. Garipov, P. Izmailov
Quantum machine learning
May 15, 2015
Experiments: algorithm

Computing distance $|\vec{u} - \vec{v}|$:

1. Represent vectors with quantum states:
   
   $$\vec{u} = |u\rangle u, \quad \vec{v} = |v\rangle v$$

2. Add an ancillary qubit and create an entangled state:
   
   $$|\phi\rangle = (|0\rangle_{anc} |u\rangle_{new} + |1\rangle_{anc} |v\rangle_{ref})/\sqrt{2}$$

3. Make a measurement on the ancillary qubit, projecting it onto the state:
   
   $$|\psi\rangle = (|u\rangle |0\rangle - |v\rangle |1\rangle)/\sqrt{|u|^2 + |v|^2}$$

4. Estimate the success probability of measurement $p$ by repeated measurements.

5. Calculate $|\vec{u} - \vec{v}|$ using $p$. 
The distance between $\vec{u}$ and $\vec{v}$ can be directly calculated from $p$:

$$|\vec{u} - \vec{v}| = \sqrt{2p(|u|^2 + |v|^2)}$$

Also the inner product between $|u\rangle$ and $|v\rangle$ can be obtained:

$$\langle u|v \rangle = (0.5 - p)(|u|^2 + |v|^2)/(|u||v|)$$
Experiments: setup

Figure: Experimental setup with 4 photonic qubits
Experiment for 2D vectors:

- Only 2 of 100 samples are misclassified.
- Errors occurred at vectors that are close to boundary where absolute error of probability estimation is close to $|D_A - D_B|$.
- Estimation of probability done by 10,000 repeated measurements for each vector.
- Time needed for processing one vector is 1 sec.
### Experiments: results

#### 4D classification

<table>
<thead>
<tr>
<th></th>
<th>( \vec{v}_A )</th>
<th>( \vec{v}_B )</th>
<th>Group</th>
<th>Correct?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(2.00, 0.00, 0.00, 0.00)</td>
<td>-1.45, -0.93</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>(0.00, 0.00, 0.00, 2.00)</td>
<td>0.82, 0.50</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>(0.35, 0.20, 0.00, 0.00)</td>
<td>-0.79, -0.71</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>(0.23, 0.19, 0.08, 0.07)</td>
<td>-0.54, -0.51</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>(1.32, 3.62, 1.57, 4.32)</td>
<td>0.74, 0.48</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>(0.15, 0.17, 0.82, 0.98)</td>
<td>1.26, 0.72</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>(0.18, 0.10, 1.02, 0.59)</td>
<td>0.98, 0.76</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>(0.97, 0.17, 0.17, 0.03)</td>
<td>-1.37, -0.93</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>9</td>
<td>(0.68, 0.25, 0.00, 0.00)</td>
<td>-1.18, -0.79</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>10</td>
<td>(0.83, 0.48, 1.44, 0.83)</td>
<td>0.67, 0.17</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>11</td>
<td>(1.27, 1.06, 3.48, 2.92)</td>
<td>1.13, 0.76</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>12</td>
<td>(0.40, 0.40, 0.40, 0.40)</td>
<td>-0.10, -0.26</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>13</td>
<td>(0.09, 0.15, 0.49, 0.85)</td>
<td>0.80, 0.55</td>
<td>B</td>
<td>✓</td>
</tr>
<tr>
<td>14</td>
<td>(0.10, 0.55, 0.06, 0.32)</td>
<td>-0.19, -0.28</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>15</td>
<td>(1.94, 0.34, 0.34, 0.06)</td>
<td>-1.22, -1.10</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>16</td>
<td>(3.42, 1.24, 1.97, 0.72)</td>
<td>-0.34, -0.39</td>
<td>A</td>
<td>✓</td>
</tr>
<tr>
<td>17</td>
<td>(0.66, 0.00, 1.80, 0.00)</td>
<td>0.40, -0.02</td>
<td>A</td>
<td>✗</td>
</tr>
</tbody>
</table>

#### Experiment for 4D vectors:

- \( \vec{v}_A = (1, 0, 0, 0) \).
- \( \vec{v}_B = (0, 0, 1, 1) \).

500 measurements per estimation.

Data acquisition time for one vector is 2 min.
### Experiment for 8D vectors:

- \( \vec{v}_A = (1, 0, 0, 0, 0, 0, 0, 0) \).
- \( \vec{v}_B = (0, 0, 0, 0, 0, 0, 1) \).
- 500 measurements per estimation.
- Data acquisition time for one vector is 4 min.
Experiments summary:

- First experimental demonstration of machine learning on a photonic quantum computer.
- Experimental prove of suitability and potential power of quantum machine learning.
- Future studies are planned to design circuits that will be able to achieve speed-up not only with respect to \( N \) but also with respect to number of training samples \( M \).
References
