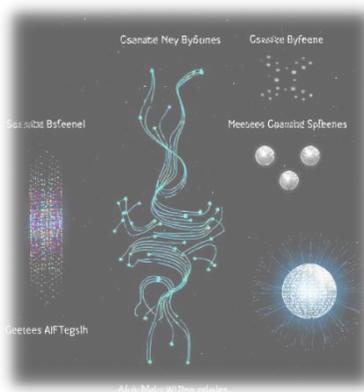




# 量子神经网络的演进

刘丁

2025. 1.10



- **哪些经典机器学习已经被“量子化”了？**
- **经典神经网络已经如此成功，我们为什么还要研究量子神经网络？我们真的需要量子神经网络吗？**
- **大模型可以被“量子化”吗？大模型究竟在做什么？**
- **量子神经网络是通向通用人工智能的有效路径吗？**
- **量子神经网络和参数化量子线路有何差异？**
- **现有的量子神经网络是研究经典神经网络量子加速的唯一途径吗？**
- **.....**

- **量子神经网络简史**
  - **几种经典的网络结构及案例**
  - **面临的问题**
  - **管窥未来**
- 

# 量子神经网络简史

## 量子神经网络的萌芽期

- 量子计算与神经网络的起源

1981年，理查德·费曼提出使用量子计算机模拟量子系统的设想，为量子计算埋下种子。

1985年，大卫·杜奇发表关于通用量子计算机的论文，奠定了量子算法研究的基础。

- 神经网络模型的提出

1943年，McCulloch和Pitts提出了第一个神经网络模型，为后续神经网络的发展奠定了基础。

1982年，霍普菲尔德网络模型提出，模仿人类大脑的记忆和学习机制，推动了神经网络研究。

- 量子神经计算的初步想法

1995年，Ron Chrisley和Subhash Kak独立发表了关于量子神经计算的想法，Subhash Kak讨论了量子神经计算机的性能优势。

1996年，Subhash Kak进一步探讨了量子神经计算在脑功能研究中的应用。

## 量子神经网络的探索期

- **量子M-P神经网络模型**

利用量子线性叠加原理，提出了量子M-P神经网络模型，描述了该网络在正交态和非正交态输入下的工作原理及权值更新算法。

这种模型展示了量子神经网络在处理复杂问题时的潜力，为后续研究提供了新的思路。

- **量子感知器网络模型**

结合量子计算和传统感知器网络，提出了量子感知器网络模型，通过实例分析和仿真实验，证明了其在实现经典感知器无法完成的XOR功能方面的优势。

该模型的提出，为量子神经网络在逻辑运算等领域的应用奠定了基础。

- **量子Hopfield神经网络模型**

提出了存储矩阵元素基于概率分布的量子Hopfield神经网络模型，其存储容量比传统模型提高了指数级。

该模型的工作过程符合量子演化过程，为量子神经网络在信息存储与检索方面的应用提供了新的方法。

- **基于Grover算法的权值训练方法**

在Grover量子搜索算法的基础上，提出了一种带权值的量子神经网络训练方法，该方法完全工作在量子机制下。这种训练方法能够以一定的概率学习训练样本，达到网络的目的，为量子神经网络的训练提供了新的途径。

- **量子竞争神经网络模型**

提出了量子竞争神经网络模型，通过量子竞争能够对模式进行识别和分类，其存储容量比传统竞争神经网络提高了指数级。该模型的提出，为量子神经网络在模式识别等领域的应用提供了新的思路。

## 量子神经网络的发展期

- 量子门控循环神经网络模型

量子门控循环神经网络（QGRNN），融合了变分量子算法和经典循环网络的优势。

该模型通过参数化量子线路实现了类似经典门控单元的功能机制，确保了QGRNN兼具高效学习性能和硬件可执行性。

- 量子卷积神经网络模型

量子卷积神经网络（QCNN）的出现，使得量子计算在处理量子数据或高维数据特征提取时展现出超强的能力。QCNN能够迅速识别出数据的风格和主题，帮助科学家们从海量实验数据中发现新的物理现象

- 量子GAN

- 量子图神经网络

- .....

# 几种经典的网络结构及案例

本节内容采用了《量子机器学习-基于Python  
的理论和实现》（姜楠等编著）配套课件，特致感谢

# 量子态编码

- 通过一个例子来说明，制备归一化数据  $\mathbf{x} = (0.4 \quad 0.4 \quad 0.8 \quad 0.2)^T$
- 因为有4个数据，所以位置信息用  $\log_4=2$  个比特表示，初始状态均为  $|0\rangle$
- 制备量子态的目的是得到  $|x\rangle = 0.4|00\rangle + 0.4|01\rangle + 0.8|10\rangle + 0.2|11\rangle$
- 方法是用旋转门修改振幅，并产生位置信息
- 如何获得旋转门的角度？

$$0.4^2 + 0.4^2 = (\sqrt{0.32})^2$$

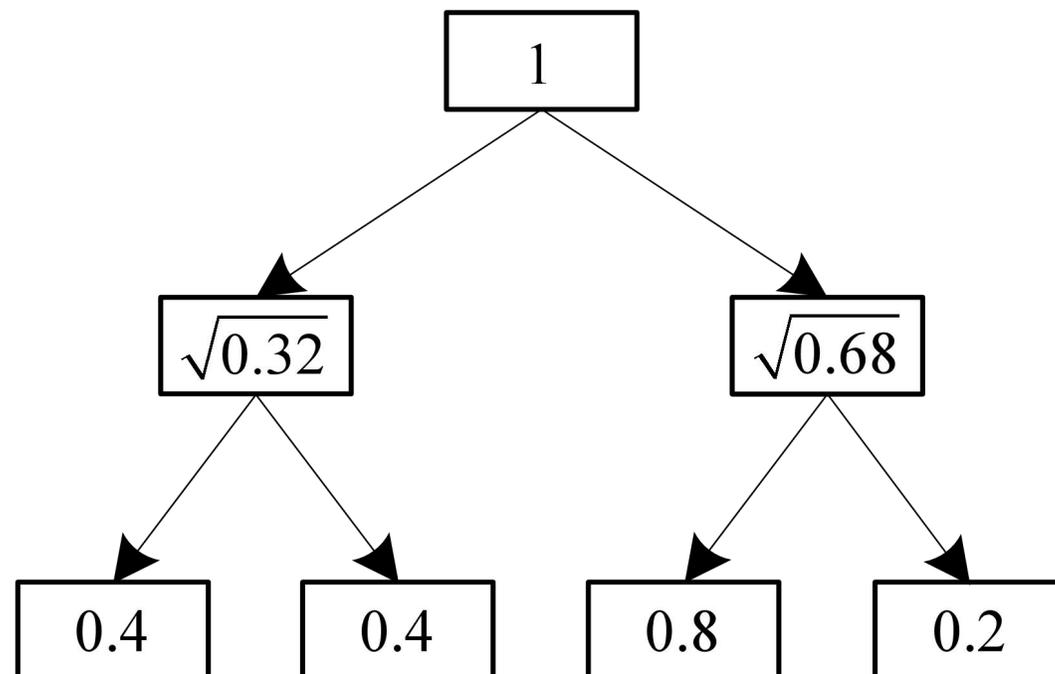
$$0.8^2 + 0.2^2 = (\sqrt{0.68})^2$$

$$(\sqrt{0.32})^2 + (\sqrt{0.68})^2 = (1)^2$$

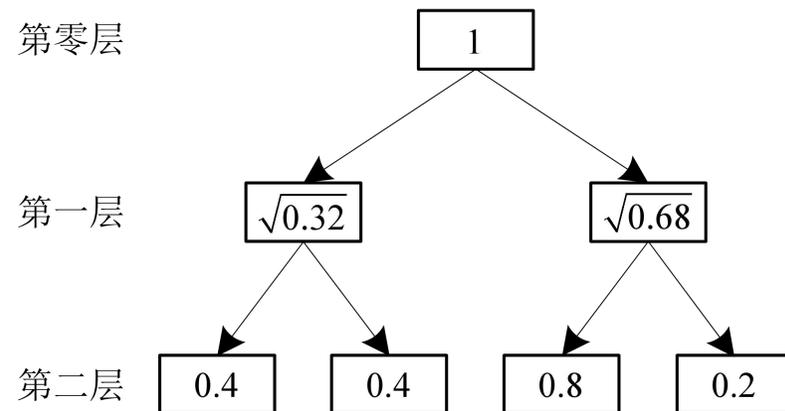
第零层

第一层

第二层



- 先实现第一层的  $\sqrt{0.32}$  和  $\sqrt{0.68}$
- 再分别实现第二层的0.4和0.4、0.8和0.2



- 用  $R_y(\theta_1^1)$  作用于第一个量子比特  $|0\rangle$  得到

$$\left( \sqrt{0.32} |0\rangle + \sqrt{0.68} |1\rangle \right) |0\rangle$$

• 由于  $R_y(\theta) |0\rangle = \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} |1\rangle$

- 因此  $\theta_1^1 = 2 \arccos \sqrt{0.32}$

## 量子线路

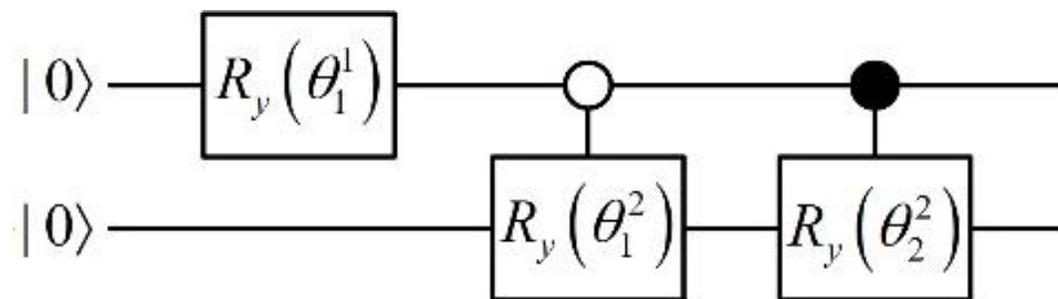
$$\left(\sqrt{0.32}|0\rangle + \sqrt{0.68}|1\rangle\right)|0\rangle = \sqrt{0.32}|00\rangle + \sqrt{0.68}|10\rangle$$

- 使用两个受控旋转门  $R_y(\theta_1^2)$  和  $R_y(\theta_2^2)$  都作用于第二个量子比特，得到

$$\begin{aligned} & \sqrt{0.32}|0\rangle \left[ \frac{1}{\sqrt{0.32}} (0.4|0\rangle + 0.4|1\rangle) \right] + \sqrt{0.68}|1\rangle \left[ \frac{1}{\sqrt{0.68}} (0.8|0\rangle + 0.2|1\rangle) \right] \\ & = 0.4|00\rangle + 0.4|01\rangle + 0.8|10\rangle + 0.2|11\rangle \end{aligned}$$

- 因此

$$\theta_1^2 = 2 \arccos \sqrt{\frac{1}{2}} \quad \theta_2^2 = 2 \arccos \sqrt{\frac{16}{17}}$$



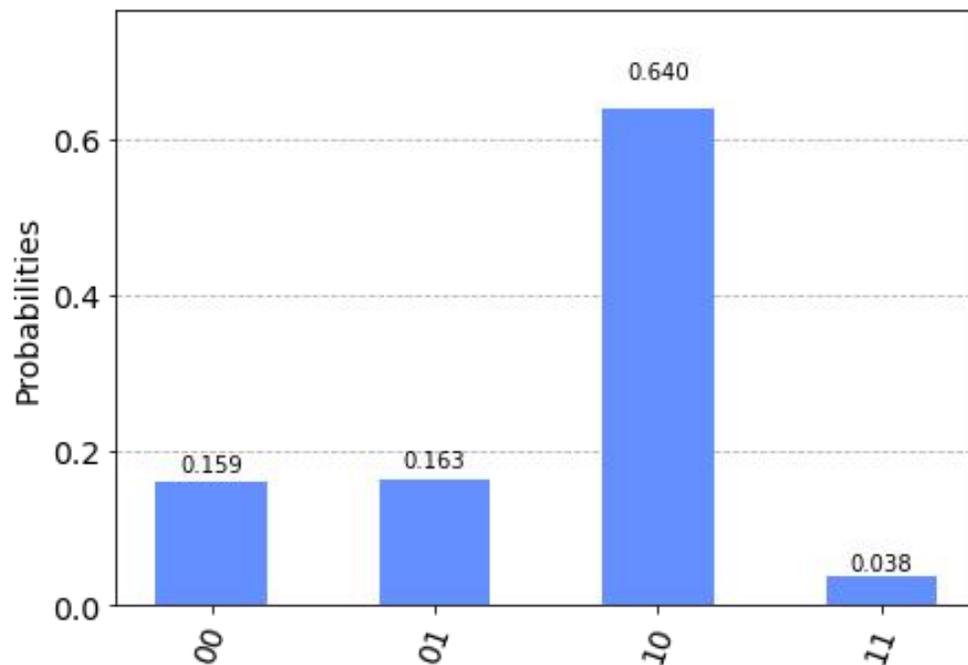
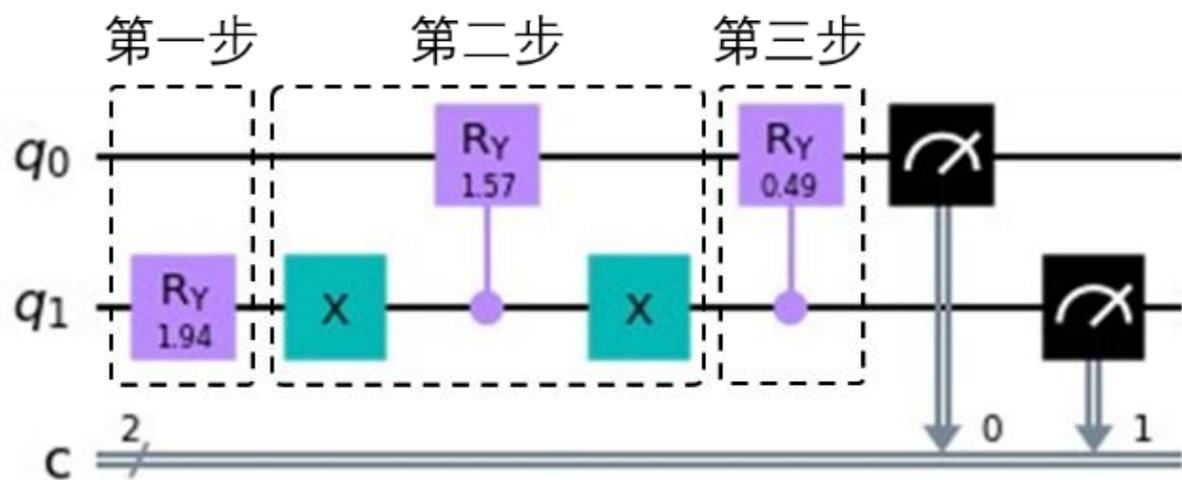
# 实现

$$\mathbf{x} = (0.4 \quad 0.4 \quad 0.8 \quad 0.2)^T$$

$$\theta_1^1 = 2 \arccos \sqrt{0.32} = 1.939$$

$$\theta_2^2 = 2 \arccos \sqrt{\frac{16}{17}} = 0.490$$

$$\theta_1^2 = 2 \arccos \sqrt{\frac{1}{2}} = 1.571$$



# 交换测试

- 计算  $a$  和  $b$  的内积 (的模)

$$|a\rangle = a_0|0\rangle + a_1|1\rangle, \quad |b\rangle = b_0|0\rangle + b_1|1\rangle$$

- 第一步: 使用  $U$  门和  $V$  门制备量子态

$$|\psi_1\rangle = |0\rangle|a\rangle|b\rangle$$

- 第二步: 交换测试

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle|a\rangle|b\rangle + |1\rangle|a\rangle|b\rangle)$$

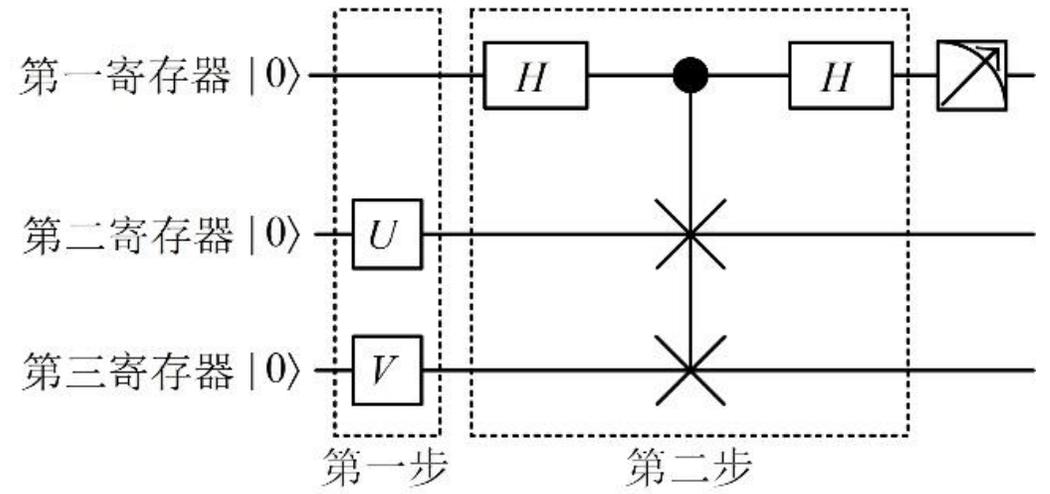
- 首先使用  $H$  门

- 然后应用受控交换门

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle|a\rangle|b\rangle + |1\rangle|b\rangle|a\rangle)$$

- 再使用  $H$  门

$$\begin{aligned} |\psi_4\rangle &= \frac{1}{\sqrt{2}} \left[ \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) |a\rangle |b\rangle + \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) |b\rangle |a\rangle \right] \\ &= \frac{1}{2} |0\rangle (|a\rangle |b\rangle + |b\rangle |a\rangle) + \frac{1}{2} |1\rangle (|a\rangle |b\rangle - |b\rangle |a\rangle) \end{aligned}$$



$$|\psi_4\rangle = \frac{1}{2}|0\rangle(|a\rangle|b\rangle + |b\rangle|a\rangle) + \frac{1}{2}|1\rangle(|a\rangle|b\rangle - |b\rangle|a\rangle)$$

- 对第一寄存器进行测量，得到  $|0\rangle$  的概率为

$$\begin{aligned} P(0) &= \left( \frac{1}{2} \langle a|\langle b| + \langle b|\langle a| \right) \left( \frac{1}{2} |a\rangle|b\rangle + |b\rangle|a\rangle \right) \\ &= \frac{1}{4} (\langle a|\langle b||a\rangle|b\rangle + \langle a|\langle b||b\rangle|a\rangle + \langle b|\langle a||a\rangle|b\rangle + \langle b|\langle a||b\rangle|a\rangle) \\ &= \frac{1}{4} (\langle b|a\rangle\langle a|b\rangle + 1 + 1 + \langle a|b\rangle\langle b|a\rangle) \\ &= \frac{1}{2} + \frac{1}{2} |\langle a|b\rangle|^2 \end{aligned}$$

- 同理，得到  $|1\rangle$  的概率为  $P(1) = \frac{1}{2} - \frac{1}{2} |\langle a|b\rangle|^2$

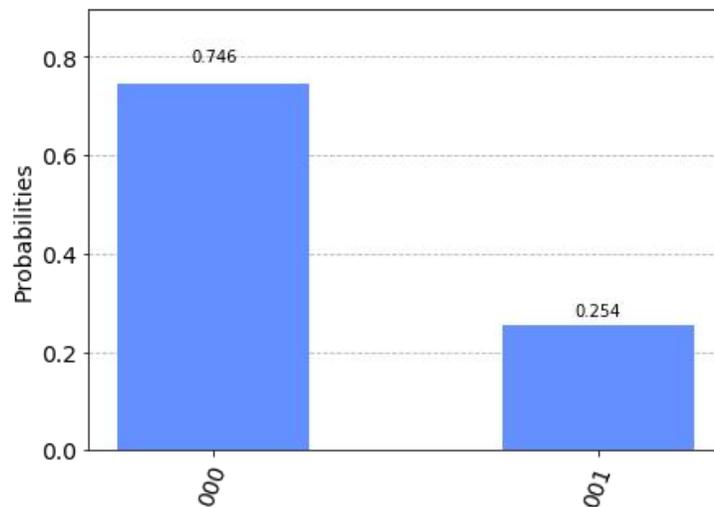
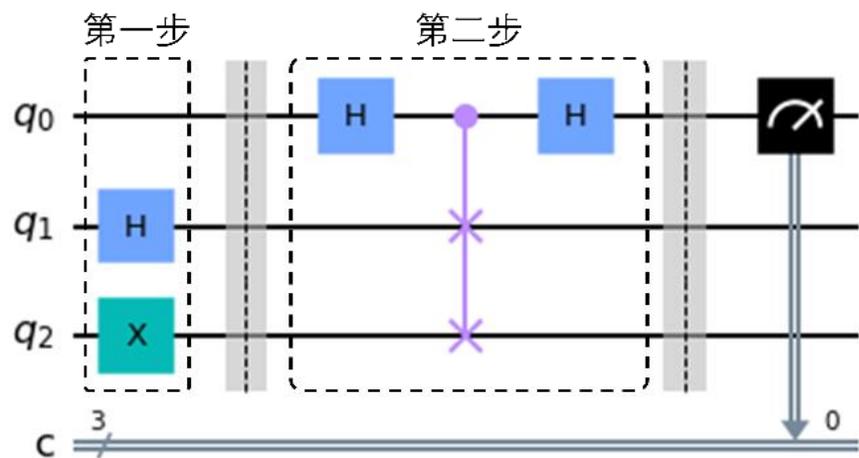
- 内积的模为

$$|\langle a|b\rangle| = \sqrt{2P(0) - 1} = \sqrt{1 - 2P(1)}$$

# 实现

- 计算  $|a\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \left(\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}}\right)^T$  和  $|b\rangle = |1\rangle = (0 \quad 1)^T$  的内积

- 理论值  $|\langle a|b\rangle|^2 = \left|\frac{1}{\sqrt{2}} \times 0 + \frac{1}{\sqrt{2}} \times 1\right|^2 = 0.5$



- 测量为0的概率0.746，测量为1的概率0.254

- 因此  $|\langle a|b\rangle|^2 = 2P(0) - 1 = 1 - 2P(1) = 0.492$

# 哈达玛测试

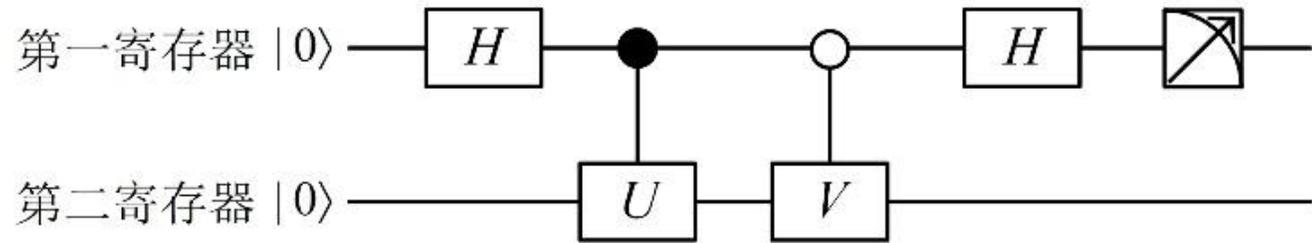
- 计算  $a$  和  $b$  的内积（非模）
- 内积计算分为实部和虚部
- 实部
- 第一步：使用  $H$  门
- 第二步：受第一寄存器的控制，制备  $a$  和  $b$

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|0\rangle)$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle|b\rangle + |1\rangle|a\rangle)$$

- 第三步：使用  $H$  门

$$|\psi_3\rangle = \frac{1}{2}[ (|0\rangle + |1\rangle)|b\rangle + (|0\rangle - |1\rangle)|a\rangle ] = \frac{1}{2}[ |0\rangle(|b\rangle + |a\rangle) + |1\rangle(|b\rangle - |a\rangle) ]$$



- 测量得到0的概率

$$\begin{aligned} P(0) &= \frac{1}{2}(\langle b| + \langle a|) \frac{1}{2}(|b\rangle + |a\rangle) \\ &= \frac{1}{4}(\langle b||b\rangle + \langle b||a\rangle + \langle a||b\rangle + \langle a||a\rangle) \\ &= \frac{1}{4}(2 + \langle a|b\rangle + \langle b|a\rangle) \\ &= \frac{1}{4}(2 + 2\operatorname{Re}\langle a|b\rangle) \\ &= \frac{1}{2} + \frac{\operatorname{Re}\langle a|b\rangle}{2} \end{aligned}$$

- 同理，测量得到1的概率

$$P(1) = \frac{1}{2} - \frac{\operatorname{Re}\langle a|b\rangle}{2}$$

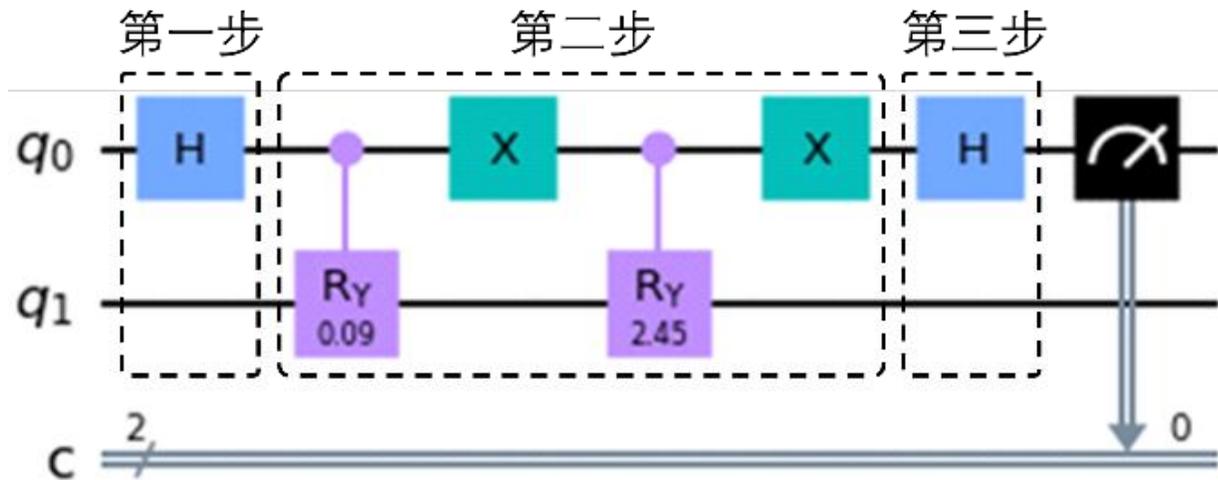
- 由此可得

$$\operatorname{Re}\langle a|b\rangle = 2P(0) - 1 = 1 - 2P(1)$$

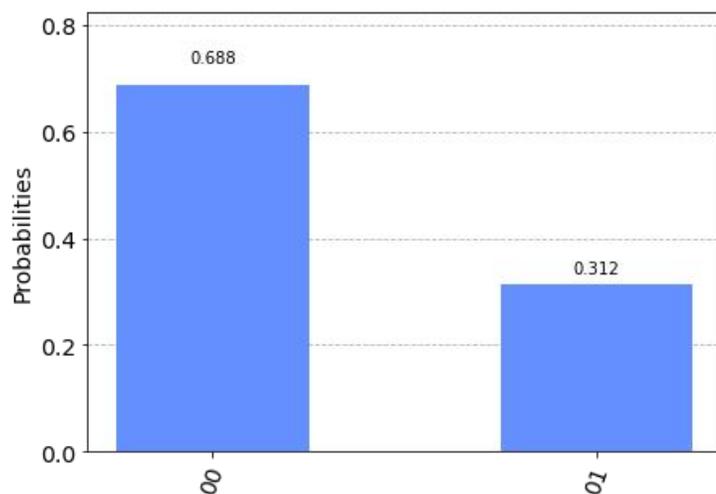
# 实现

- 求两个量子态  $|a\rangle = 0.999|0\rangle + 0.045|1\rangle$  和  $|b\rangle = 0.339|0\rangle + 0.941|1\rangle$
- 的内积  $0.999 \times 0.339 + 0.045 \times 0.941 = 0.381$

- 第一步，对  $q_0$  进行  $H$  门变换
- 第二步，用 1 控制和 0 控制的旋转门制备  $a$  和  $b$
- 第三步，对  $q_0$  进行  $H$  门变换
- 最后对  $q_0$  进行测量



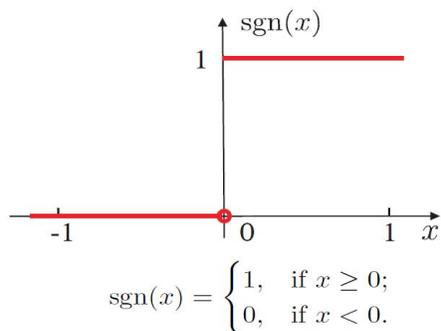
- 得到 1 的概率为 0.312
- 因此内积的实部为  $1 - 2 \times 0.312 = 0.376$



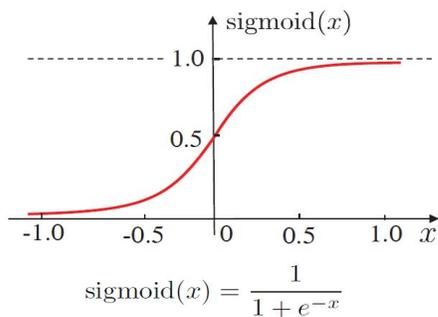
# 经典感知机

- 数据集  $\{(\mathbf{x}_i, y_i) : \mathbf{x}_i \in \mathbb{R}^M, y_i \in \{1, -1\}\}_{i=0, 1, \dots, N-1}$
- 感知机只有一个神经元，目标是找到一个超平面，能够将所有样本正确分类
- 感知机工作原理表示为

$$\hat{y}_i = \text{sign} \left( \sum_{j=0}^{M-1} w_j x_{ji} + b \right) = \begin{cases} +1, & \sum_{j=0}^{M-1} w_j x_{ji} + b \geq 0 \\ -1, & \sum_{j=0}^{M-1} w_j x_{ji} + b < 0 \end{cases}$$



(a) 阶跃函数



(b) Sigmoid 函数

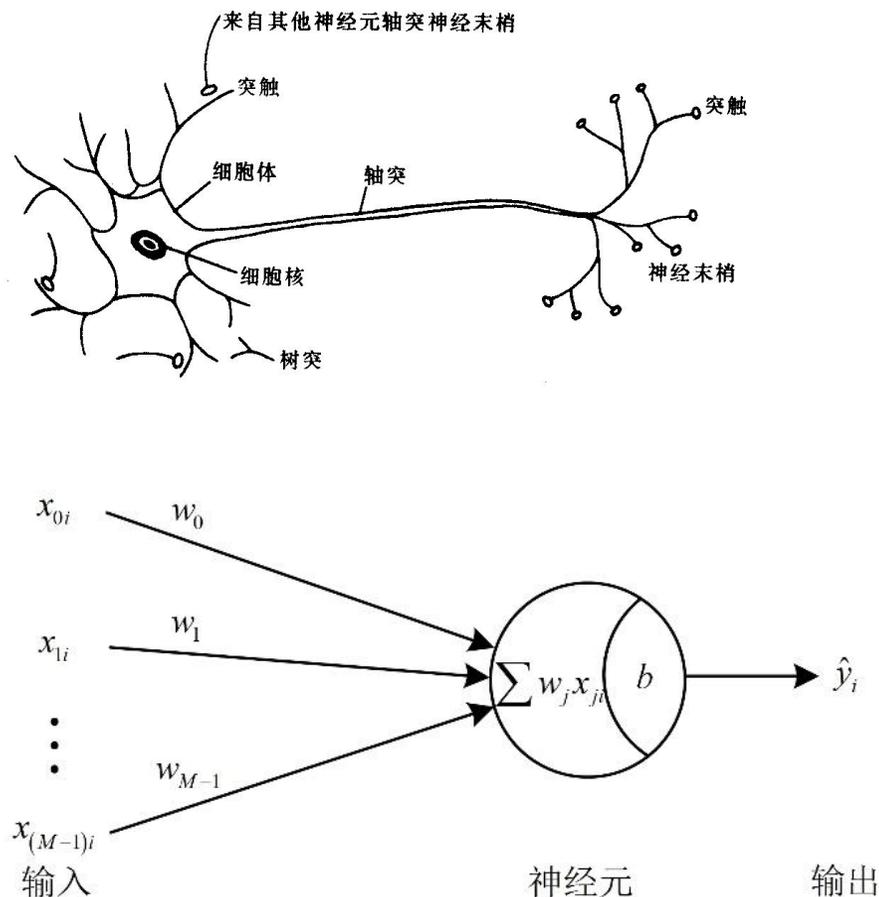


图 5.2 典型的神经元激活函数

通常使用梯度下降法更新权重系数 $w$ 和偏置 $b$

$$\begin{pmatrix} \mathbf{w} \\ b \end{pmatrix}^{t+1} = \begin{pmatrix} \mathbf{w} \\ b \end{pmatrix}^t + \eta y_i \begin{pmatrix} \mathbf{x}_i \\ 1 \end{pmatrix}$$

- $\eta \in (0,1)$  称为学习率， $t$  是训练次数
- 感知机的训练过程为按照一定的顺序取遍所有的样本 $(\mathbf{x}_i, y_i)$ ，判断其是否能被正确分类
  - 若能被正确分类，则跳过该样本
  - 否则，更新权重系数和偏置
- 直到所有样本被正确分类

# 量子感知机

- 量子感知机中，规定样本  $\mathbf{x}_i$  和权重  $w$  的元素只能为  $\pm 1$ ，**激活函数**为

$$\hat{y}_i = \begin{cases} 0, & \left| \sum_{j=0}^{M-1} w_j x_{ji} \right| \geq \theta \\ 1, & \left| \sum_{j=0}^{M-1} w_j x_{ji} \right| < \theta \end{cases}$$

## 内积计算

$$\left| \sum_{j=0}^{M-1} w_j x_{ji} \right|$$

- $x_j$  和  $w$  的量子态形式为

$$|x_i\rangle = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} x_{ji} |j\rangle \quad |w\rangle = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} w_j |j\rangle$$

- 计算内积

- 定义酉算子

$$U_{x_i} |0\rangle^{\otimes m} = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} x_{ji} |j\rangle = |\psi_1\rangle$$

$$U_w |1\rangle^{\otimes m} = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} w_j |j\rangle = |\psi_2\rangle$$

则

$$U_w^{-1} |\psi_2\rangle = |1\rangle^{\otimes m} = |M-1\rangle \quad \text{十进制形式}$$

第一步：使用酉算子  $U_{x_i}$  得到

$$|\psi_{1i}\rangle = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} x_{ji} |j\rangle$$

第二步：使用  $U_w^{-1}$  得到  $|\psi_3\rangle = U_w^{-1} |\psi_{1i}\rangle = \sum_{j=0}^{M-1} c_j |j\rangle$

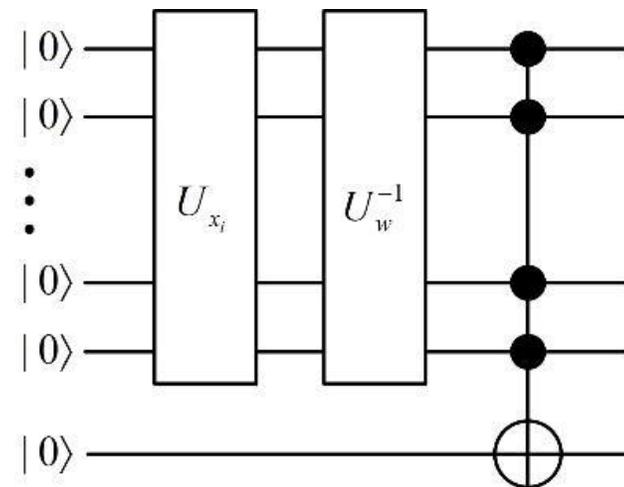
其中  $\sum_{j=0}^{M-1} c_j^2 = 1$

则  $\langle \psi_{1i} | \psi_2 \rangle = \langle \psi_{1i} | U_w U_w^{-1} | \psi_2 \rangle = \langle \psi_3 | M-1 \rangle = c_{M-1}$

而  $\langle x_i | w \rangle = \langle \psi_{1i} | \psi_2 \rangle = c_{M-1}$ ，因此  $c_{M-1}$  就是  $x_i$  和  $w$  的内积

第三步：得到  $|\psi_4\rangle = \sum_{j=0}^{M-2} c_j |j\rangle |0\rangle + c_{M-1} |M-1\rangle |1\rangle$

对最后一个量子比特测量，得到1的概率为  $|c_{M-1}|^2$ ，即  $\left| \sum_{j=0}^{M-1} w_j x_{ji} \right|^2$



# 分类

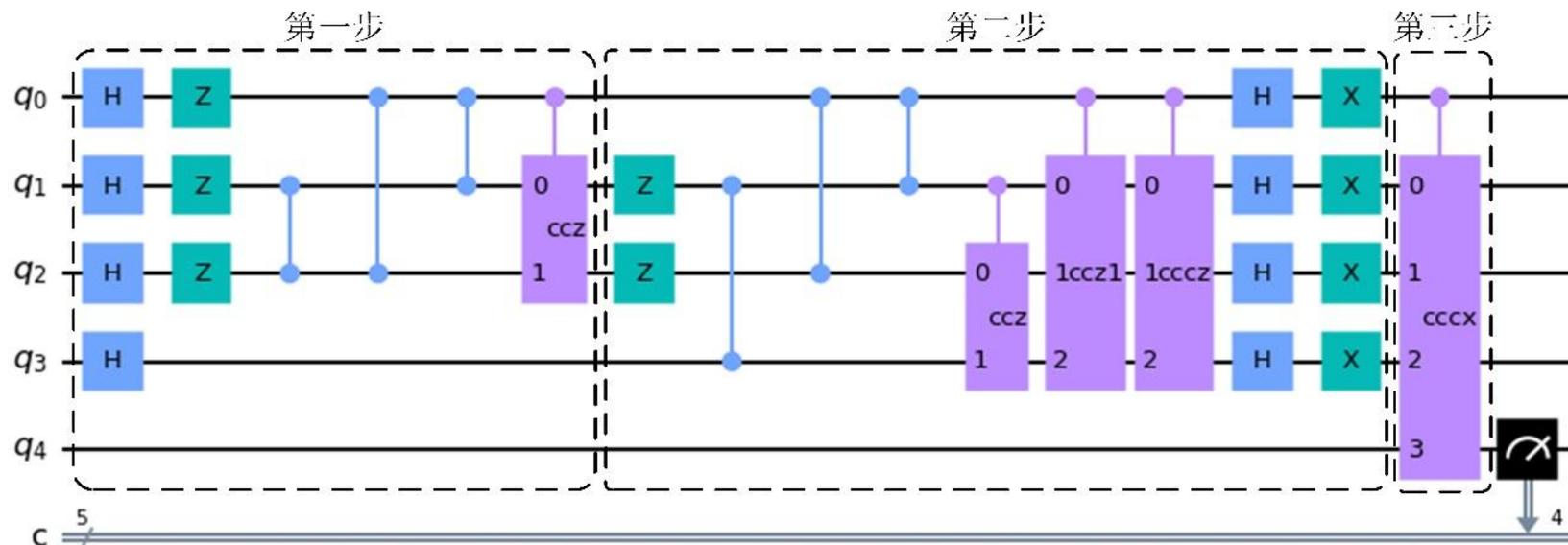
- 当输入向量被正确分类时，不改变权重系数  $w$ ；当输入向量被错误分类时，改变权重系数  $w$
- 错误分类有两种情况：
  - (1) 需要降低  $c_{M-1}$ ：翻转一部分相等位置处  $w$  元素的符号
  - (2) 需要增加  $c_{M-1}$ ：翻转一部分不相等位置处  $w$  元素的符号
- 翻转  $w$  元素的符号是经典算法

• 例如：

$$\mathbf{x}_i = \frac{1}{2\sqrt{2}}(-1 \ 1 \ 1 \ -1 \ 1 \ 1 \ -1 \ -1)$$
$$\mathbf{w} = \frac{1}{2\sqrt{2}}(-1 \ 1 \ 1 \ -1 \ -1 \ 1 \ -1 \ 1) \quad c_{M-1} = \langle \mathbf{x}_i | \mathbf{w} \rangle = \frac{1}{2}$$

- 如果错误属情况 (1)， $\mathbf{w}_1 = \frac{1}{2\sqrt{2}}(1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1 \ 1) \quad c_{M-1} = \langle \mathbf{x}_i | \mathbf{w}_1 \rangle = 0$
- 反转位置1和2
- 如果错误属情况 (2)， $\mathbf{w}_2 = \frac{1}{2\sqrt{2}}(-1 \ 1 \ 1 \ -1 \ -1 \ 1 \ -1 \ -1) \quad c_{M-1} = \langle \mathbf{x}_i | \mathbf{w}_2 \rangle = \frac{3}{4}$
- 反转位置8

# 线路实现



- 阈值  $\theta = 0.5$

$$\mathbf{x}_i^T = (-1 \quad -1 \quad 1 \quad 1)$$

$$\mathbf{w}^T = (1 \quad 1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1)$$

- 第一步得到  $|\psi_1\rangle = \frac{1}{\sqrt{16}} \sum_{j=0}^{15} x_{ji} |j\rangle$

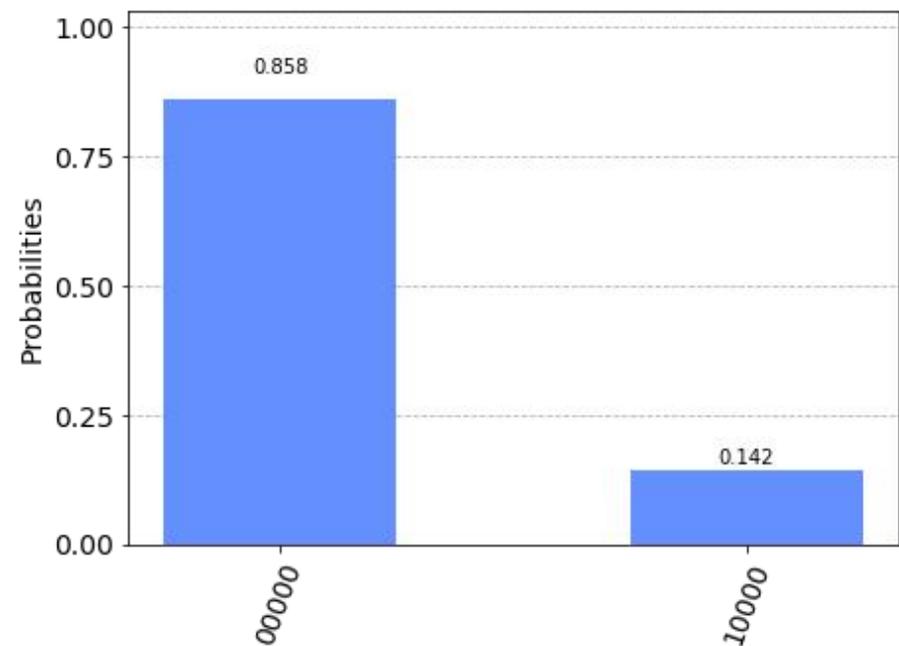
- 第二步执行能使  $U_w |1\rangle^{\otimes 4} = \frac{1}{\sqrt{16}} \sum_{j=0}^{15} w_j |j\rangle$  成立的  $U_w^{-1}$

- 第三步使用4受控非门将  $c_{15}$  提取到辅助量子比特  $q_4$  的振幅

- 最后通过测量辅助量子比特获取得到1的概率

- 1的概率为0.142，所以

$$|c_{M-1}| = \sqrt{0.142} = 0.3768 < 0.5$$

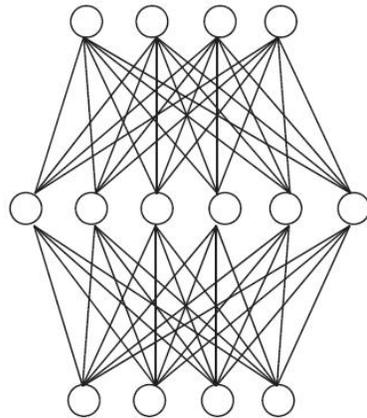


# 量子神经网络

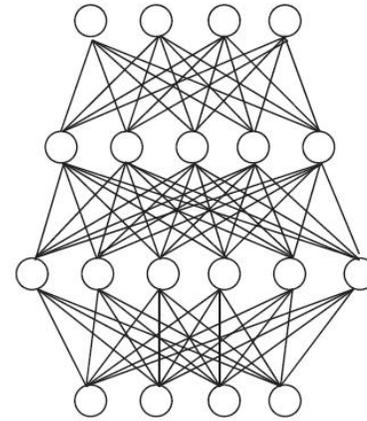
- 目前，最常见的量子神经网络模型为基于参数化量子线路的量子神经网络
- **使用参数化量子线路代替网络结构**，使用经典优化算法更新参数化量子线路中的参数

# 经典神经网络原理

- **定义**：每层神经元与下一层神经元全互联，神经元之间不存在同层连接也不存在跨层连接
- **前馈**：输入层接受外界输入，隐含层与输出层神经元对信号进行加工，最终结果由输出层神经元输出
- **学习**：根据训练数据来调整神经元之间的“**连接权**”以及每个功能神经元的“**阈值**”
- **多层网络**：包含隐层的网络



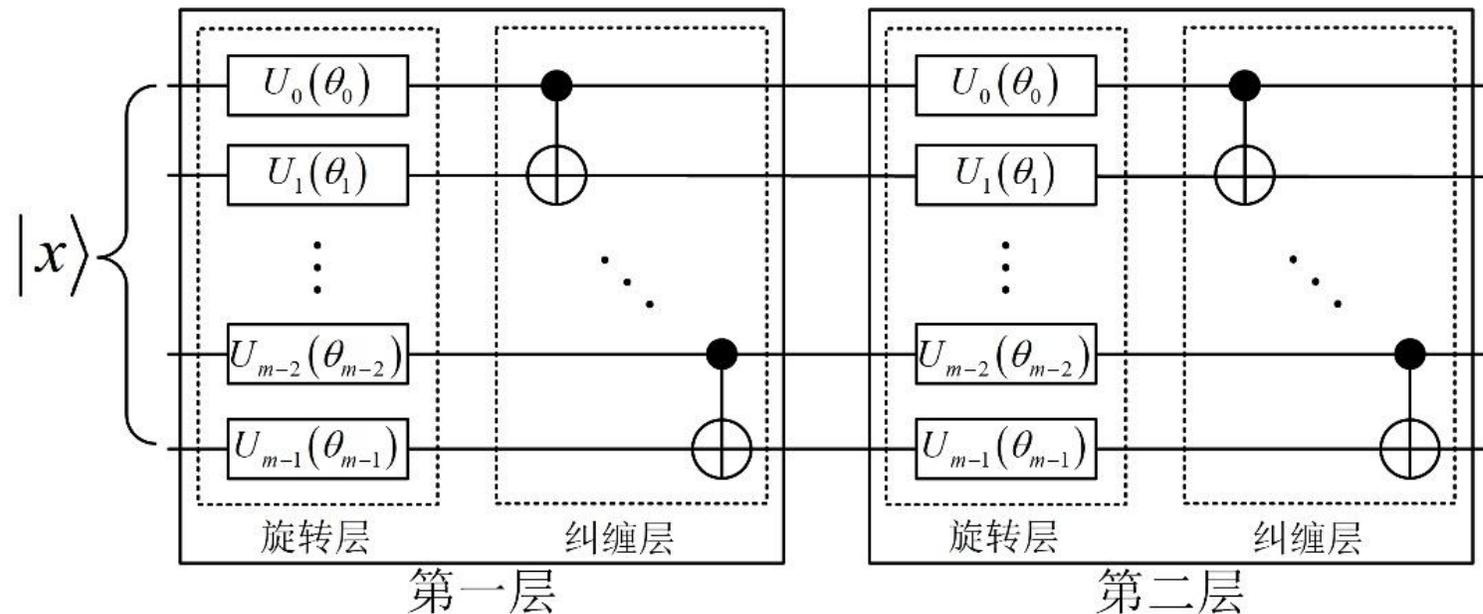
(a) 单隐层前馈网络



(b) 双隐层前馈网络

# 参数化量子线路

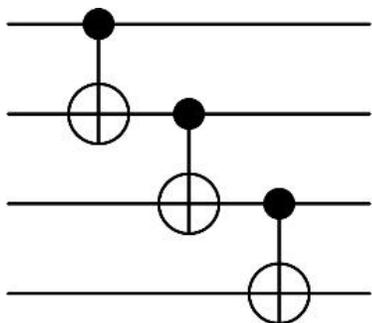
- 参数化量子线路通常由单量子比特旋转门和双量子比特纠缠门，如受控非门、受控旋转门等，按一定的规律交错排列组成
- 参数化指的是单量子比特旋转门或双量子比特纠缠门是含可训练参数的，例如 $R_x(\theta)$ 、 $R_y(\theta)$  和 $R_z(\theta)$  含有参数  $\theta$
- 这些含参数的量子门也被称为参数化量子门



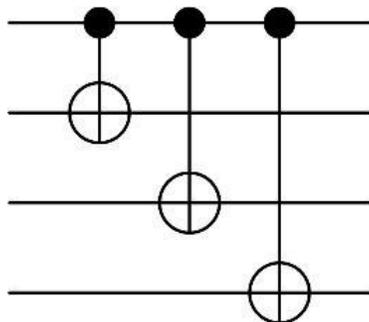
- 例如两层参数化量子线路的一般形式
- 每一层可以相同也可以不同，但是都由旋转层和纠缠层组成
- 旋转层由旋转门  $U_0(\theta_0)$ ,  $U_1(\theta_1)$ ,  $\dots$ ,  $U_{m-1}(\theta_{m-1})$  组成，纠缠层（记为  $U_m$ ）由CNOT门组成
- 该线路的作用表示为

$$U_m \left( U_{m-1}(\theta_{m-1}) \otimes \dots \otimes U_0(\theta_0) \right) U_m \left( U_{m-1}(\theta_{m-1}) \otimes \dots \otimes U_0(\theta_0) \right) |x\rangle$$

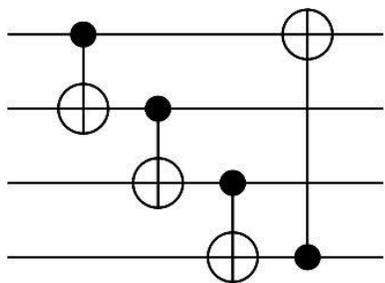
# 纠缠层的连接方法



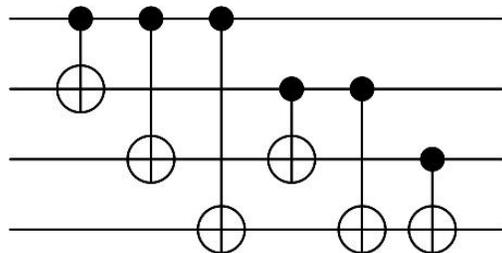
链式连接



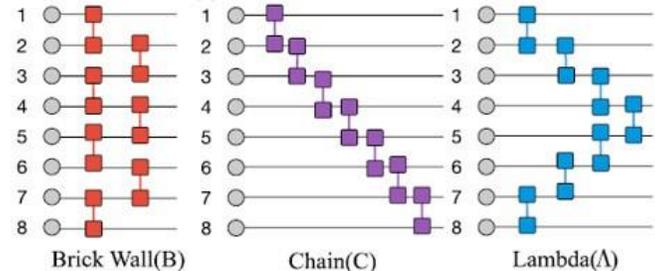
线性连接



环形连接



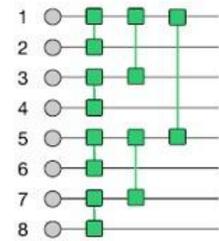
全连接



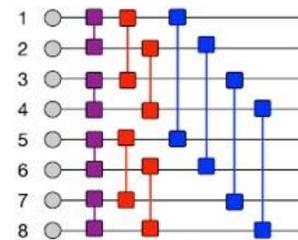
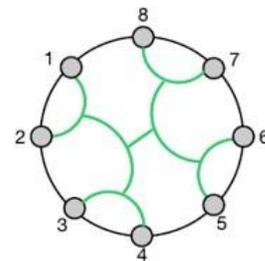
Brick Wall(B)

Chain(C)

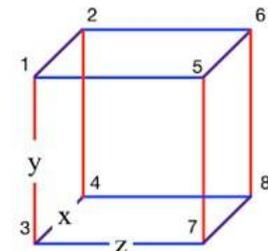
Lambda(A)



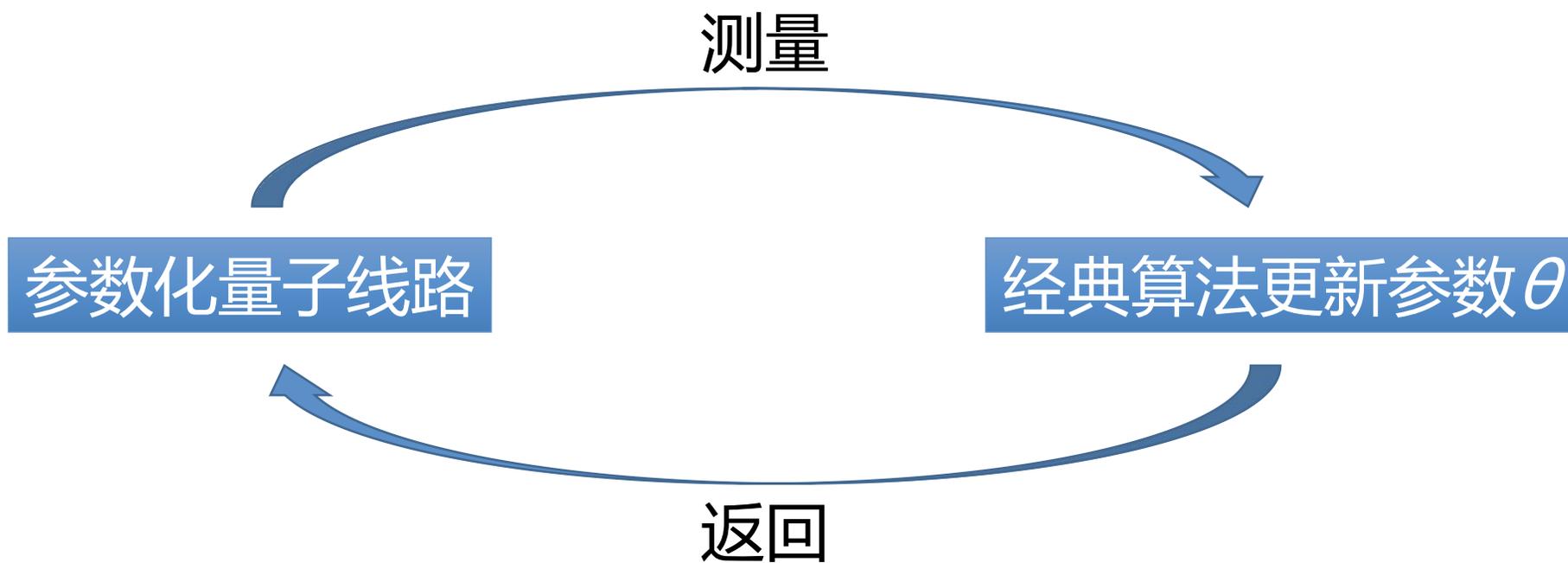
Hyperbolic(H)



Supercube(S)



- 参数化量子线路不单独运行，而是要与经典算法相结合



# 参数化量子线路所表示的函数

- 参数化量子线路所表示的函数与测量算子有关，此处用 $Z$ 算子
- 对只包含一个旋转层和一个纠缠层的单层参数化量子线路，记  $U_m(U_{m-1}(\theta_{m-1}) \otimes \dots \otimes U_0(\theta_0)) = U(\theta)$
- 经过酉变换 $U(\theta)$ 的量子态为  $|\psi(x, \theta)\rangle = U(\theta)|x\rangle$
- 用测量算子 $Z$ 测量  $|\psi(x, \theta)\rangle$ ，测量的平均值为  $\langle \psi(x, \theta) | Z | \psi(x, \theta) \rangle$
- 这就是参数化量子线路所表示的函数

- 由于
$$\begin{aligned} & \langle \psi(x, \theta) | Z | \psi(x, \theta) \rangle \\ &= \langle \psi(x, \theta) | (|0\rangle\langle 0| - |1\rangle\langle 1|) | \psi(x, \theta) \rangle \\ &= \langle \psi(x, \theta) | 0 \rangle \langle 0 | \psi(x, \theta) \rangle - \langle \psi(x, \theta) | 1 \rangle \langle 1 | \psi(x, \theta) \rangle \\ &= |\langle 0 | \psi(x, \theta) \rangle|^2 - |\langle 1 | \psi(x, \theta) \rangle|^2 \\ &= 2|\langle 0 | \psi(x, \theta) \rangle|^2 - 1 \end{aligned}$$

因此，可以通过测量得到0的概率算出  $\langle \psi(x, \theta) | Z | \psi(x, \theta) \rangle$

# 目标函数与优化

• 由前式： $\langle \psi(x, \theta) | Z | \psi(x, \theta) \rangle = 2 |\langle 0 | \psi(x, \theta) \rangle|^2 - 1$       $|\psi(x, \theta)\rangle = U(\theta) |x\rangle$

• 测量得到0的概率为  $\frac{1}{2} (1 + \langle x | U(\theta)^\dagger Z U(\theta) | x \rangle)$

• 定义样本预测标签的计算公式为

$$\pi(\theta, b) = \frac{1}{2} (1 + \langle x | U(\theta)^\dagger Z U(\theta) | x \rangle) + b$$

• 记样本  $\mathbf{x}$  的真实标签为  $y(\mathbf{x})$ ，量子神经网络的目标是找到合适的参数  $\theta$  和偏置  $b$  使得预测标签和真实标签尽量一致

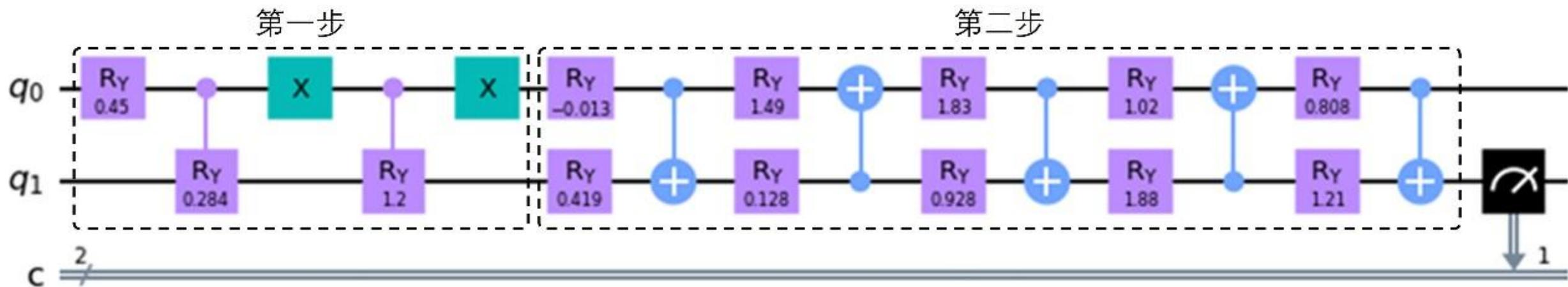
• 因此优化目标为最小化损失函数  $l(\theta, b) = \frac{1}{2} (\pi(\theta, b) - y(\mathbf{x}))^2$

• 该最小化过程为经典过程，得到参数  $\theta$  和偏置  $b$

# 实现

- 数据集为归一化的四维Iris数据集。该数据集中共有100个样本，选取75个样本作为训练集，25个样本为测试集

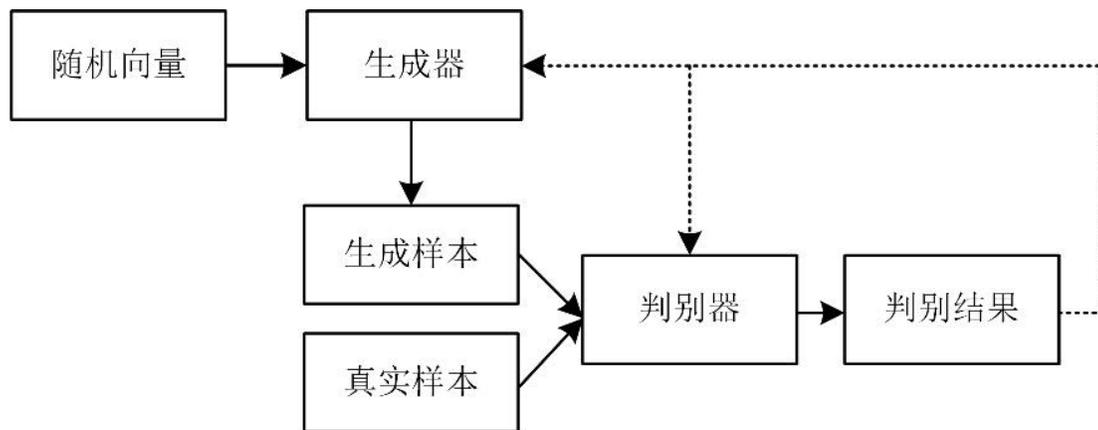
$$X = \begin{pmatrix} 0.803 & 0.714 & 0.776 & 0.860 & 0.690 \\ 0.551 & 0.266 & 0.549 & 0.440 & 0.321 \\ 0.220 & 0.618 & 0.307 & 0.248 & 0.607 \\ 0.031 & 0.191 & 0.032 & 0.057 & 0.226 \end{pmatrix}, y = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T$$



# 经典生成对抗网络

- 生成对抗网络由两部分组成：生成器和判别器，二者训练的目的都是为了打败彼此
- 生成器 $G$ 的输入为一个随机向量 $\mathbf{z}$ ，输出为一个生成样本 $G(\mathbf{z})$
- 判别器 $D$ 的输入为一个真实的或生成的样本 $\mathbf{x}$ ，输出为对该样本的判断
  - 当输入真实样本 $\mathbf{x}$ 时，判别器的输出为 $D(\mathbf{x})$ ，接近1
  - 当输入生成样本 $G(\mathbf{z})$ 时，判别器的输出为 $D(G(\mathbf{z}))$ ，接近0

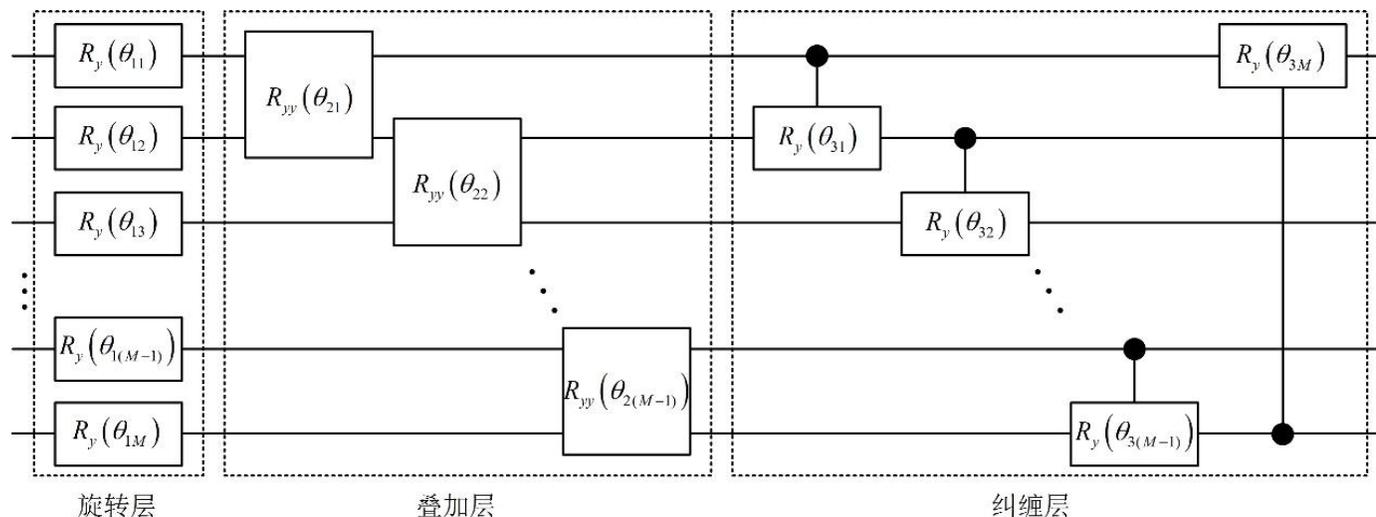
- GAN的优化目标函数可以表示为 
$$\min_G \max_D \left\{ E_{\mathbf{x} \sim p(\mathbf{x})} [\log D(\mathbf{x})] + E_{\mathbf{z} \sim p(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))] \right\}$$



训练过程中固定一方，更新另一方的网络参数，交替迭代

# 参数化量子线路

- 量子生成对抗网络所用的参数化量子线路主要由三部分组成：旋转层、叠加层和纠缠层
- 旋转层由旋转门 $R_y(\theta)$ 组成
- 叠加层使用的是双量子比特门，其作用是绕Y轴对两个量子比特进行耦合旋转



$$R_{yy}(\theta) = \begin{pmatrix} \cos\frac{\theta}{2} & 0 & 0 & i\sin\frac{\theta}{2} \\ 0 & \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} & 0 \\ 0 & -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} & 0 \\ i\sin\frac{\theta}{2} & 0 & 0 & \cos\frac{\theta}{2} \end{pmatrix}$$

- 纠缠层使用受控旋转门形成环形连接
- QGAN分别使用 $M$ 个量子比特表示 $M$ 个特征，每个特征用一个量子比特表示
- 用一个旋转门 $R_y(\theta_{ji})$ 制备

# 量子生成对抗网络算法

- 量子判别器和量子生成器都采用前述的参数化量子线路

- QGAN通过经典计算机和量子计算机之间的迭代切换，

找到参数化量子线路的最优参数

- QGAN将保真度作为目标函数，因此用交换测试得到保真度

- 假设量子判别器之后量子态为  $|\delta_1\rangle$ ，量子生成器之后量子态为  $|\gamma\rangle$

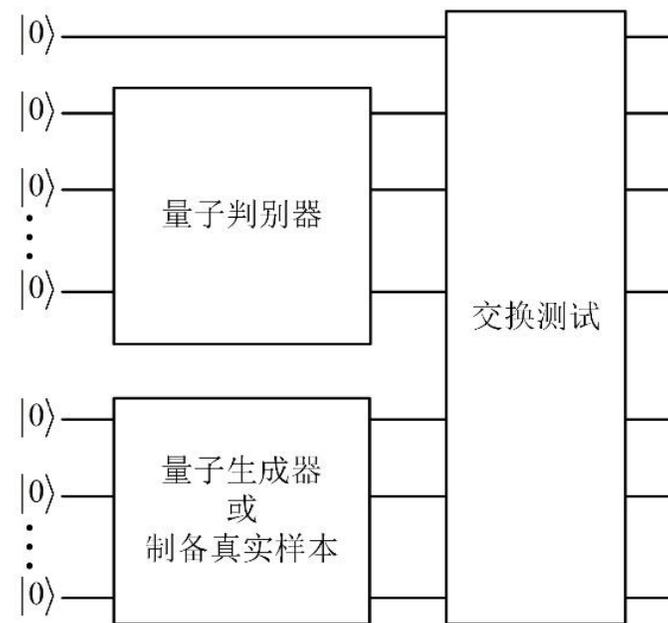
保真度  $|\langle\gamma|\delta_1\rangle|^2$  要尽量接近0

- 假设量子判别器之后量子态为  $|\delta_2\rangle$ ，真实样本为  $|\xi\rangle$ ，保真度  $|\langle\xi|\delta_2\rangle|^2$  要尽量接近1

- 因此训练量子判别器时要最小化  $-\log(1-|\langle\gamma|\delta_1\rangle|^2)$  和  $-\log(|\langle\xi|\delta_2\rangle|^2)$

- 训练量子生成器时要最小化

$$-\log(|\langle\xi|\delta_2\rangle|^2)$$



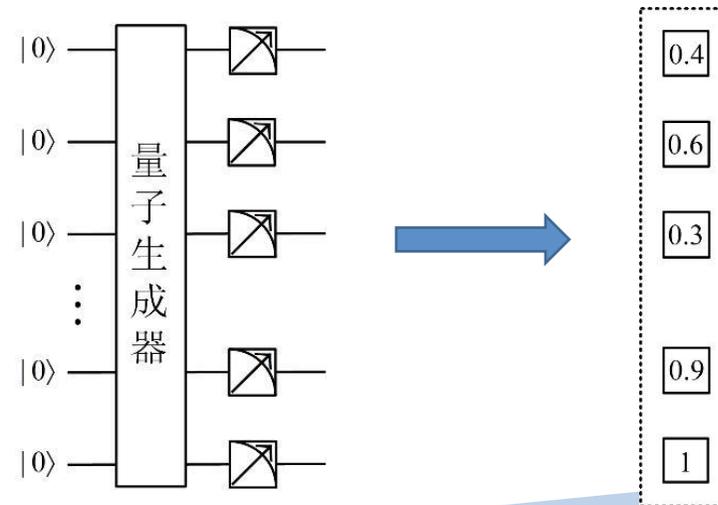
# 量子生成器后处理

- 量子生成器最后得到的是量子态，需要通过后处理才能得到近似于真实样本的生成样本
- 当制备一个样本时，它的每个特征都对应一个旋转门  $R_y(\theta_{ji})$ ，其中  $\theta_{ji} = 2 \times \arcsin(\sqrt{x_{ji}})$

- 即
$$R_y(\theta_{ji})|0\rangle = R_y(2 \times \arcsin(\sqrt{x_{ji}}))|0\rangle = \begin{pmatrix} \cos(\arcsin(\sqrt{x_{ji}})) & -\sin(\arcsin(\sqrt{x_{ji}})) \\ \sin(\arcsin(\sqrt{x_{ji}})) & \cos(\arcsin(\sqrt{x_{ji}})) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

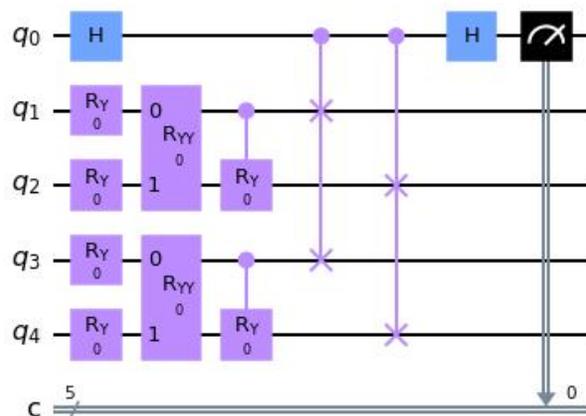
$$= \cos(\arcsin(\sqrt{x_{ji}}))|0\rangle + \sqrt{x_{ji}}|1\rangle$$

- 测量得到1的概率恰为特征  $x_{ji}$
- 因此，将量子生成器得到1的概率作为生成样本的特征

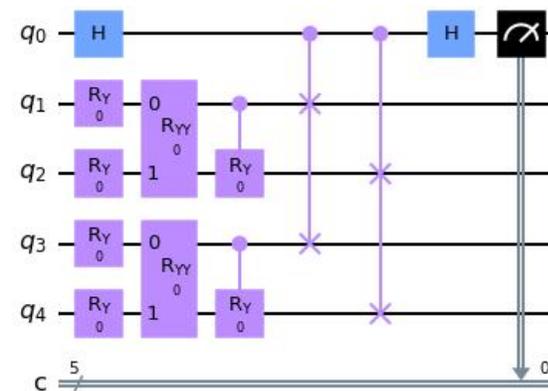


# 实现

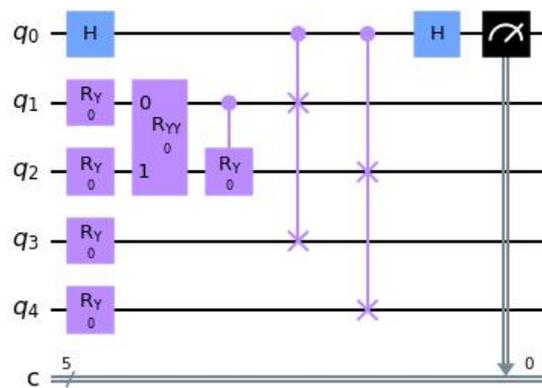
- 生成一个带有数字3或9的图片，真实图片由手写数字3和9组成，特征为2维



使用生成数据训练判别器



训练生成器



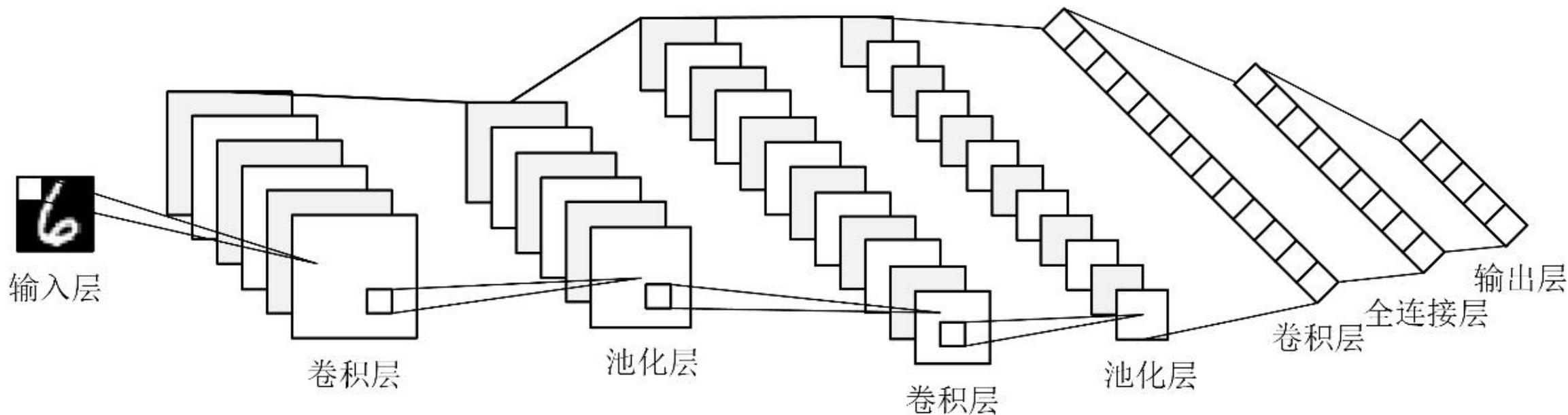
使用真实数据训练判别器



生成图片

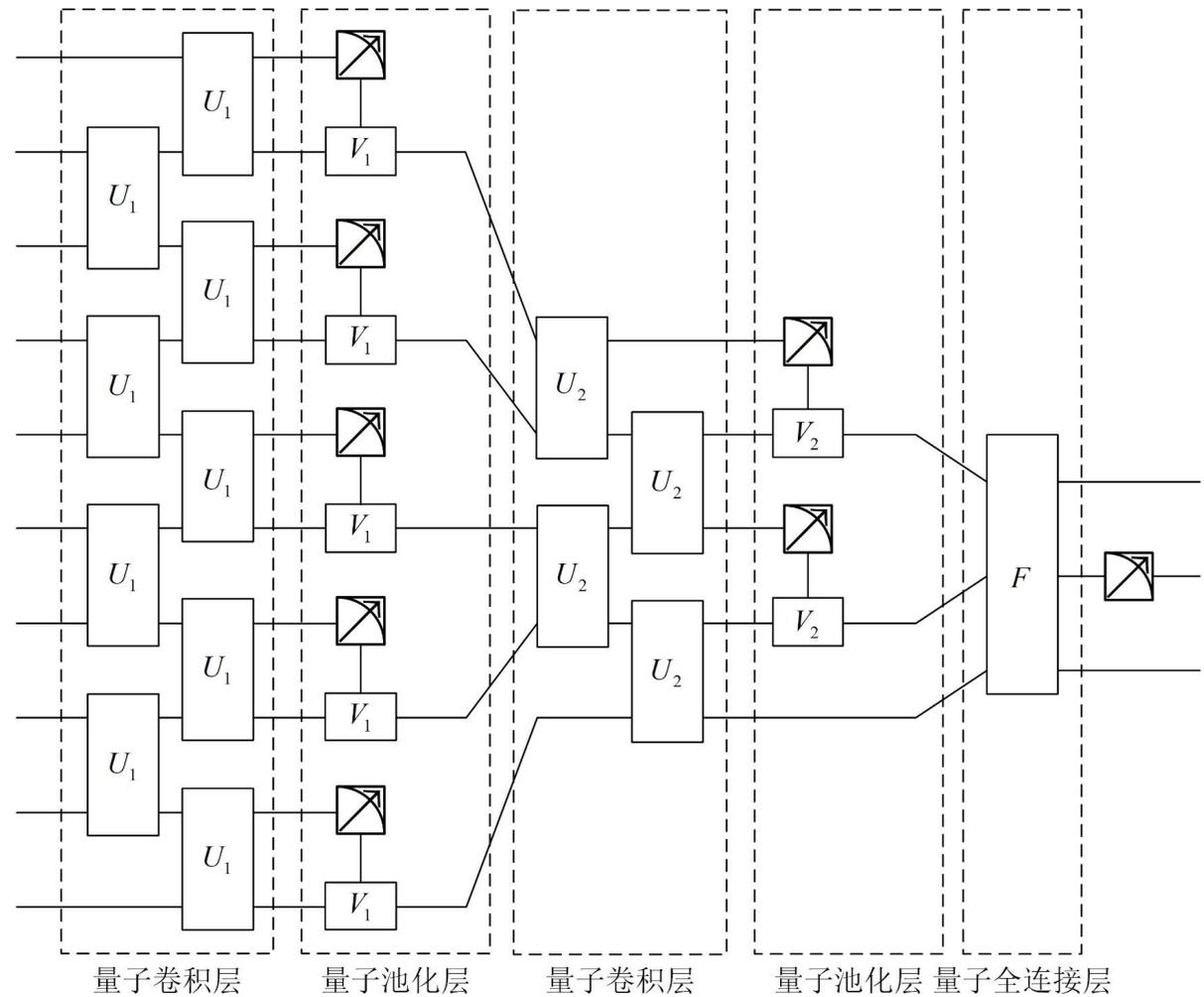
# 卷积神经网络

- 卷积神经网络（Convolutional Neural Network, CNN）是一种具有局部连接、权值共享等特点的深层前馈神经网络
- 它是一种自带特征提取和分类功能的深度学习模型



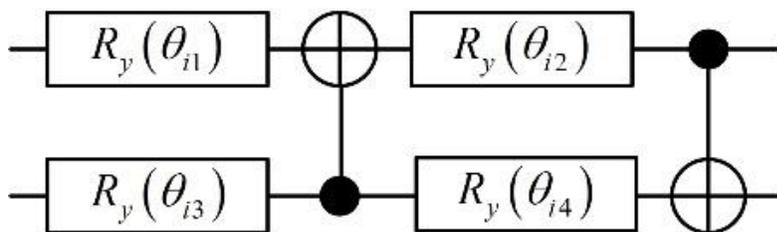
# 量子卷积神经网络

- 量子神经网络也由量子卷积层、量子池化层和量子全连接层组成
- 量子卷积层和量子池化层交替放置，分别实现特征提取和特征降维，之后通过量子全连接层进行特征综合
- 量子卷积层、量子池化层和量子全连接层分别由量子卷积单元、量子池化单元和量子全连接单元组成



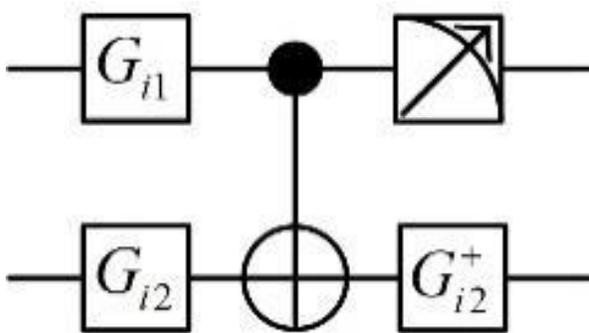
## 量子卷积层

- 量子卷积单元  $U_i$  类似于经典卷积神经网络中的卷积核
- 在量子卷积层内部，量子门作用于相邻的两个量子比特上，对量子态的部分振幅作变换，体现局部连接的特点
- 同一卷积层内，所有  $U_i$  具有相同的参数，则体现权值共享的特点



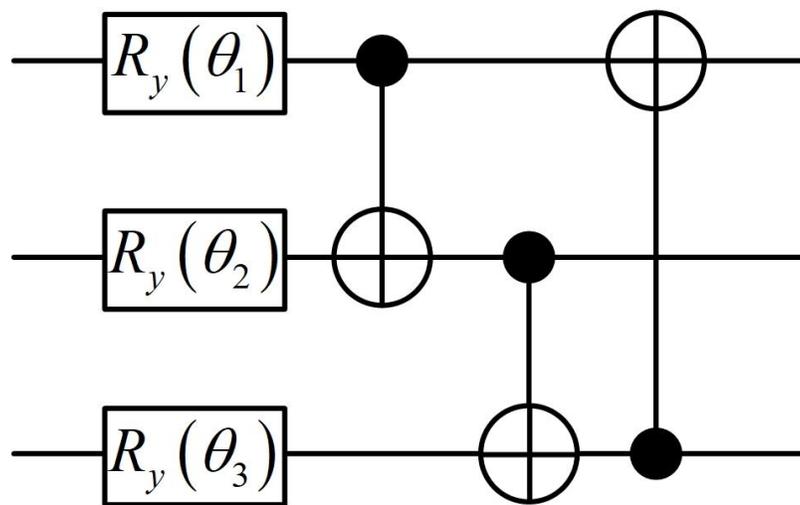
# 量子池化层

- 量子池化单元  $V_i$
- 量子池化层对量子卷积层所提取的特征进行降维
- 降维主要是通过测量完成的，测量使得被测量子比特坍塌，相干量子比特数量减少，量子态维度缩减



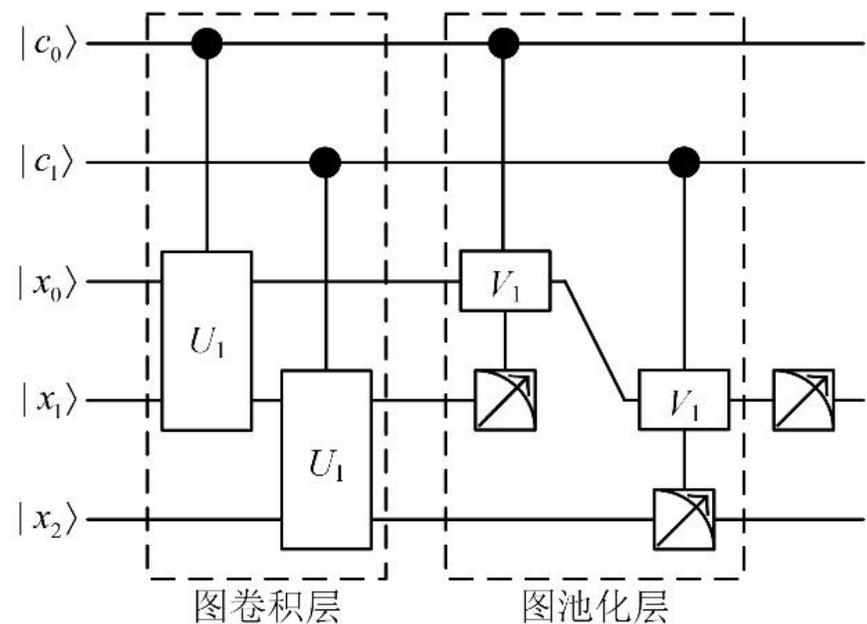
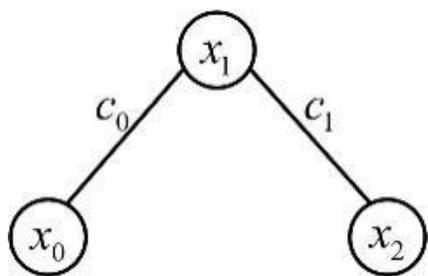
# 量子全连接层

- 量子全连接层起到了分类器的作用



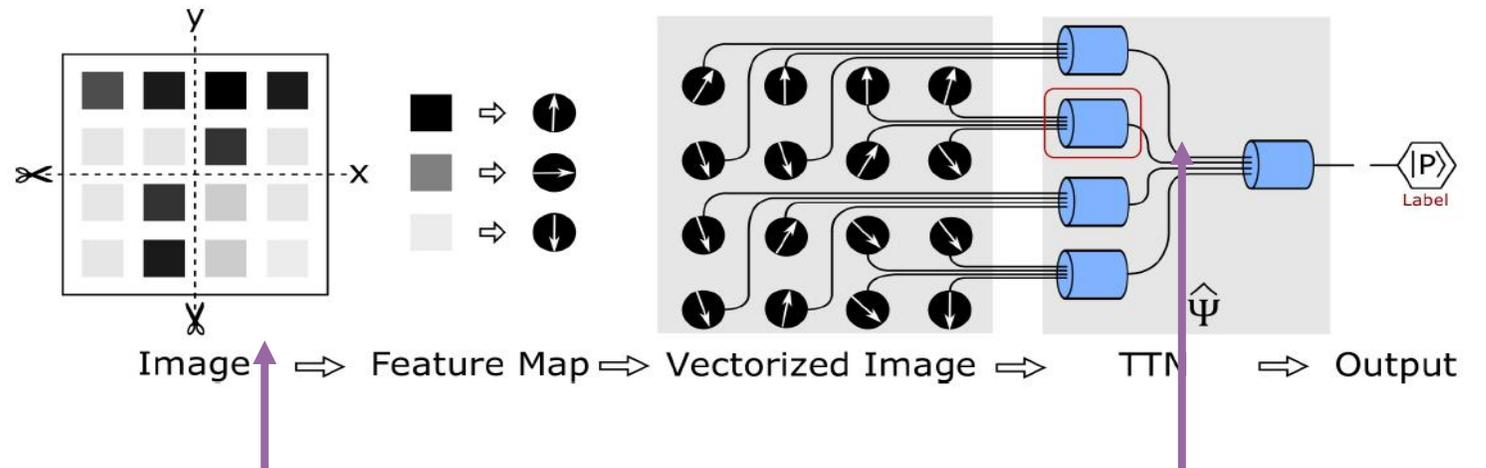
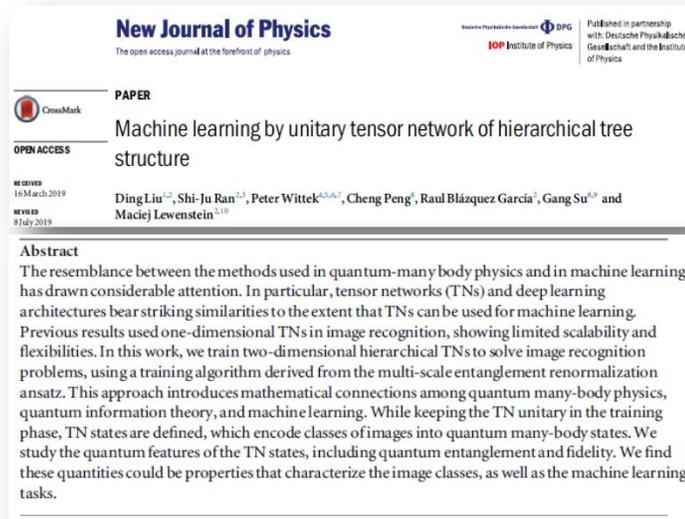
# 量子图神经网络

- 图结构，是一种典型的非欧几里得数据结构
- 可以用来描述社交网络、化学分子结构、知识图谱等
- 图中共有三个节点： $x_0$ 、 $x_1$ 和 $x_2$ ，两条边 $c_0$ 和 $c_1$
- 量子图卷积神经网络由量子图卷积层和量子图池化层组成
- 图中有几条边，在量子图卷积层中就有几个受控的量子门 $U_1$ ，在量子图池化层中就有几个受控量子门和测量操作



# 张量网络深度学习

Ding Liu, **Shi-Ju Ran\***, Peter Wittek\*, Cheng Peng, Rual Blázquez Garca, Gang Su, and Maciej Lewenstein.



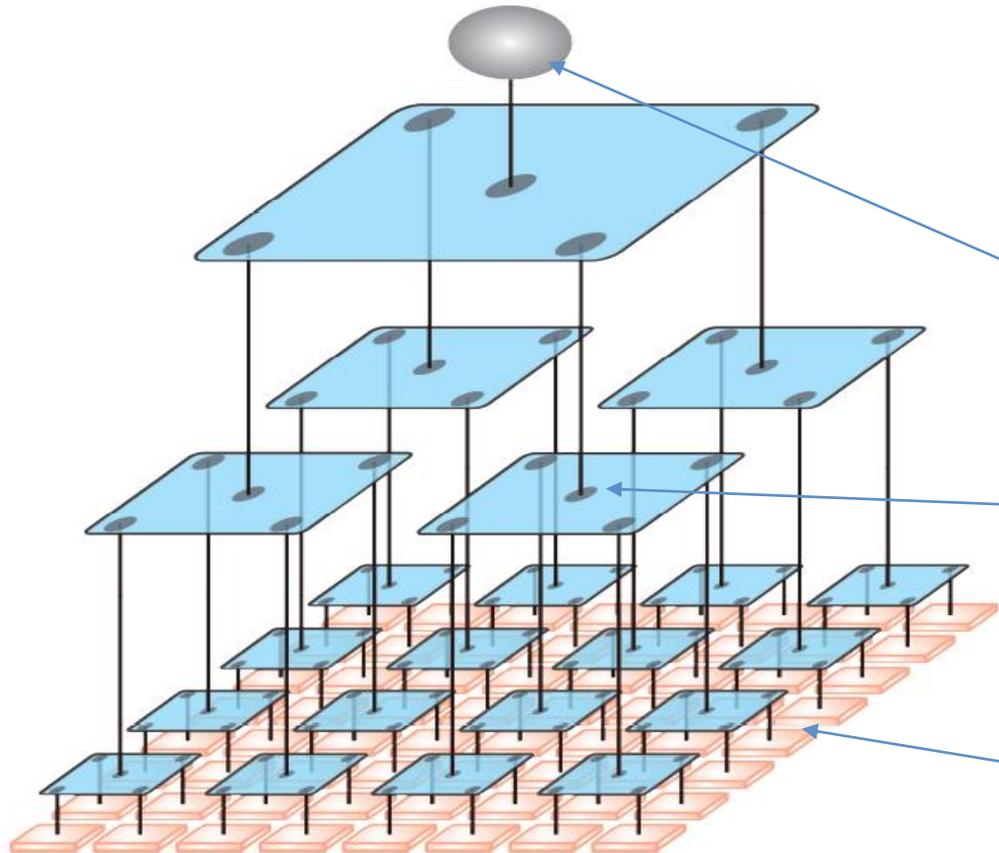
Map to the quantum many-body Hilbert space

Classifier in the quantum many-body Hilbert space

# Solving supervised machine learning problems by tree TN

Central clue: write the cost function (error) as a tree TN (TTN)

$$f = \sum_{j=1}^J |\tilde{\mathbf{L}}^{[j]} - \mathbf{L}^{[j]}|^2 \quad L_{\alpha}^{[j]} = \begin{cases} 1, & \text{if } \alpha = p \\ 0, & \text{otherwise} \end{cases}$$



**Label** of the input sample

**Orthogonal tensors:** mapping the vectorized images to the output vectors

An **input** sample mapped from scalars (e.g., grey pixels) to normalized vectors

$$\phi^{s_j}(x_j) = \sqrt{\binom{d-1}{s_j-1}} (\cos(\frac{\pi}{2}x_j))^{d-s_j} (\sin(\frac{\pi}{2}x_j))^{s_j-1}$$

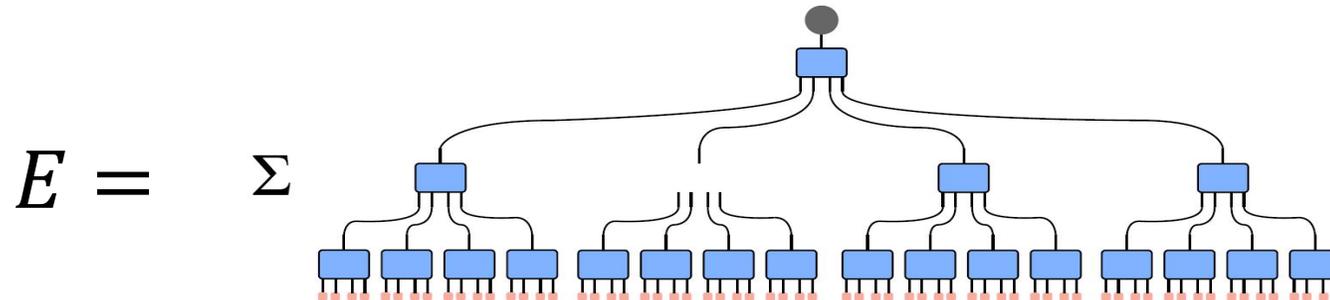


# Updating algorithm inspired by MERA

## Multi-scale entanglement renormalization ansatz

For updating one tensor, we define the environment tensor

Summing over the samples



Use SVD to reach the minimum while keeping the orthogonality

$$E = U\Lambda V^T \longrightarrow T = U^T V$$

At the minimum, we have  $f = -\sum_n \Lambda_n$

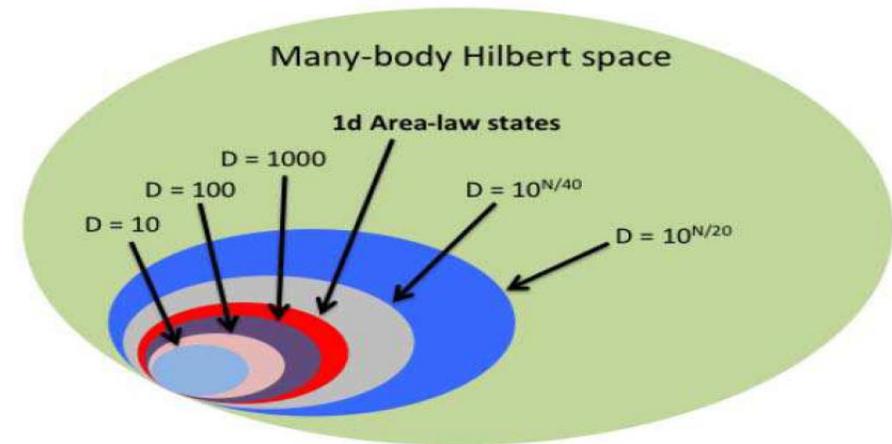
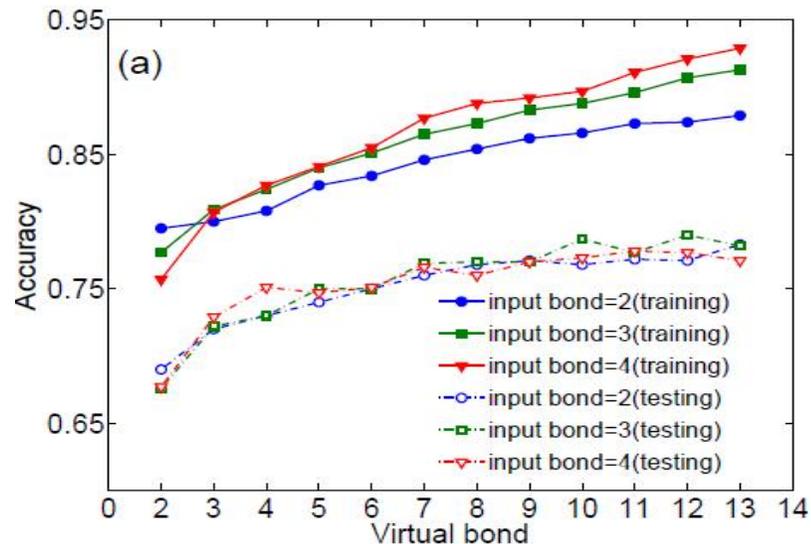
Iterate this procedure until all tensors converge

### Several tricks:

- Save intermediate vectors to avoid redundant computations
- Input part of the samples (randomly selected) in each training
- Keep the vectors normalized to ensure efficiency

# Benchmark on CIFAR database

Testing the *learning*: the accuracy of a TTN given certain fixed training samples (1000 training samples)

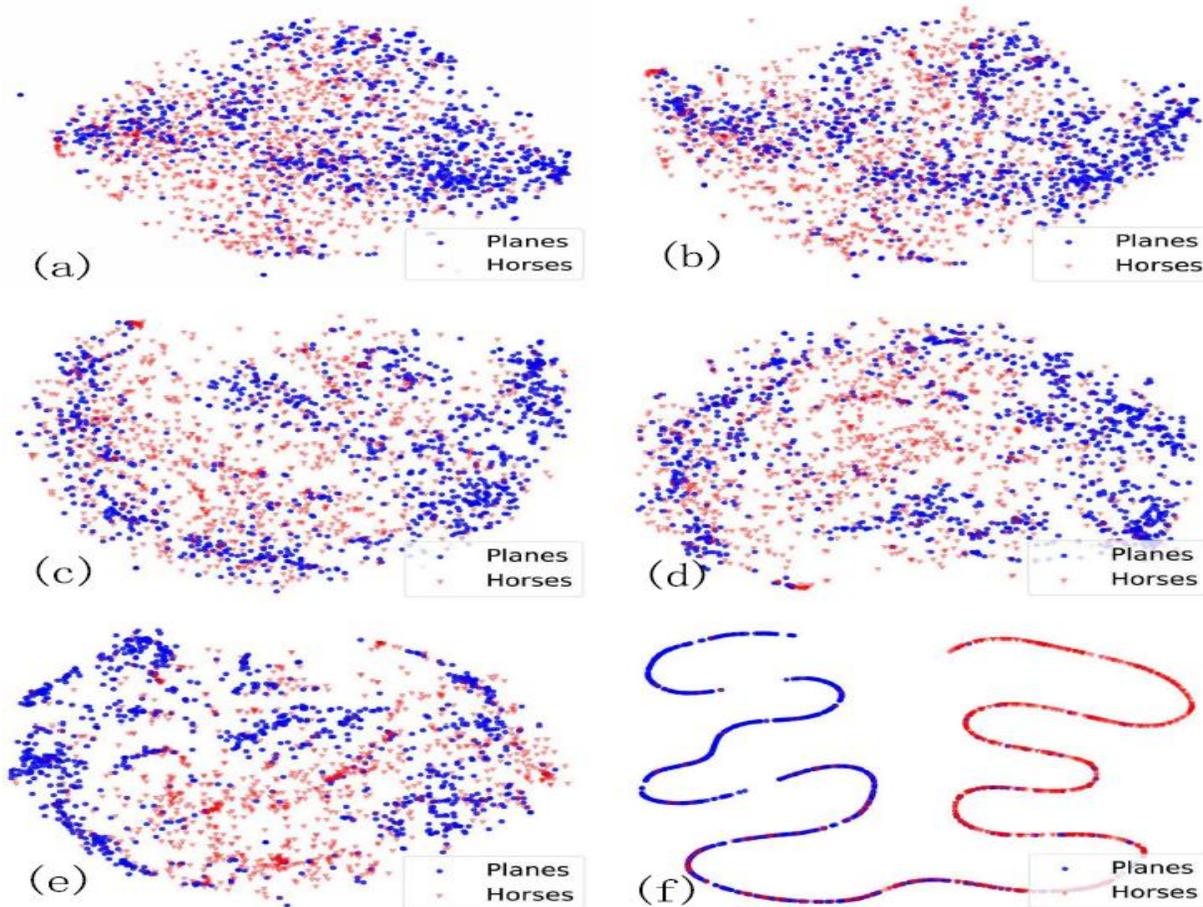


The upper limit of representation power (learnability) of a TTN depends on the input bond dimensions, and the virtual bond dimensions determine how well the TTN reaches this limitation.

# Checking how information is processed in the hierarchical structure by T-SNE

Input the vectors obtained on each layer to T-SNE

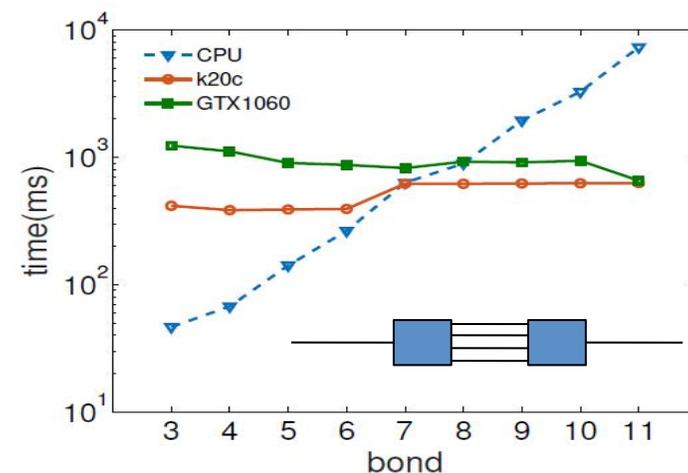
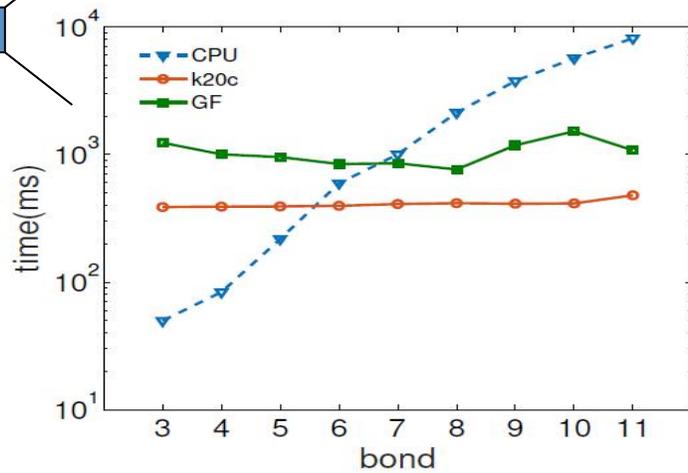
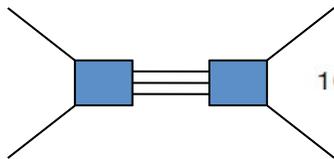
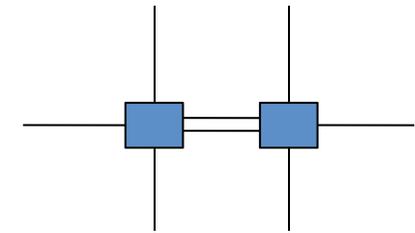
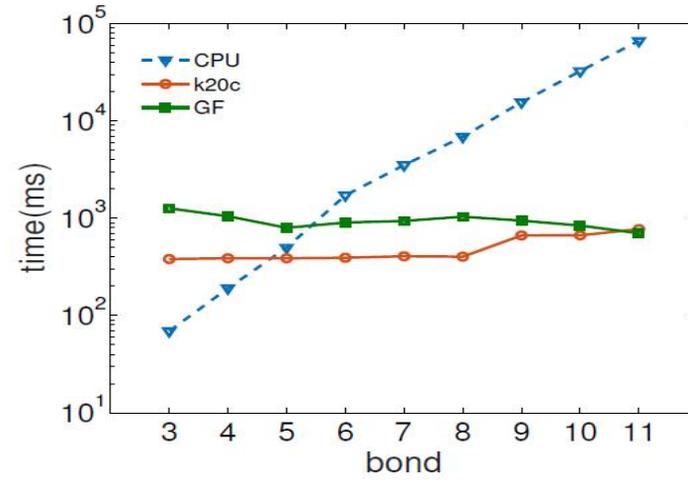
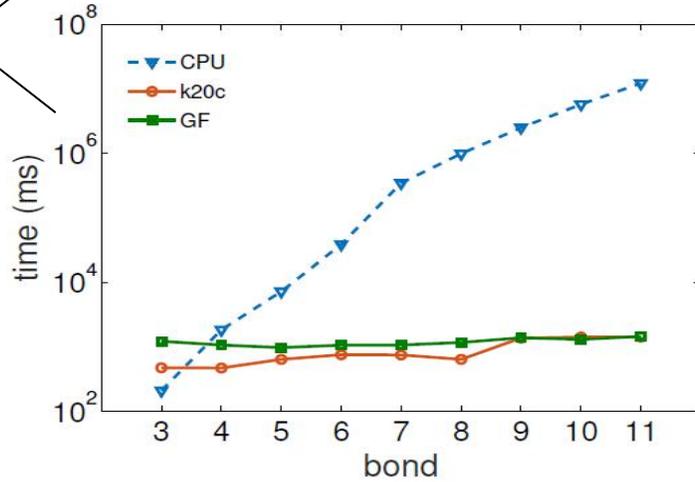
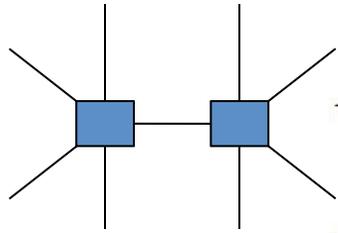
Mixed



Well separated

Further survey: consider the gauge degrees of freedom of TN!

# Related issue II: GPU acceleration for tensor network simulation



GPU is **10~10<sup>4</sup>** faster than CPU!

# 面临的问题

# 贫瘠高原 (barren plateau)

## • 网络架构改进

残差量子神经网络 (ResQNETs)：受经典残差神经网络的启发，将传统的量子神经网络架构拆分成多个量子节点，每个节点包含自己的参数化量子电路，并在这些节点之间引入残差连接。研究表明，残差连接的引入可以改善训练性能，从而有效缓解贫瘠高原问题。

量子卷积神经网络 (QCNNs)：通过引入卷积和池化层的序列来减少量子比特数量，同时保留数据特征的相关信息。研究发现，QCNN架构中参数的梯度方差消失速度最多是多项式级别的，这意味着QCNNs不会表现出贫瘠高原现象。

## • 训练策略优化

使用神经网络生成量子态：有研究提出利用神经网络生成量子态，可以有效地缓解随机变分量子电路中的贫瘠高原现象。不过，该方法的理论基础尚未完全验证，需要进一步研究。

自适应优化器：在量子物理信息神经网络 (QPINN) 的研究中发现，自适应优化器的性能不如基本的随机梯度下降 (SGD) 优化器。这提示在优化策略的选择上，可能需要针对量子神经网络的特点进行更深入的探索。

# 管窥未来

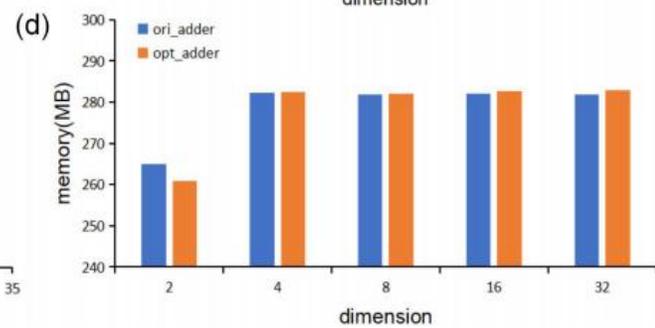
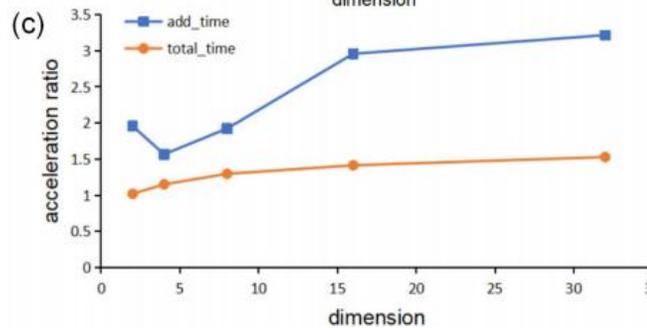
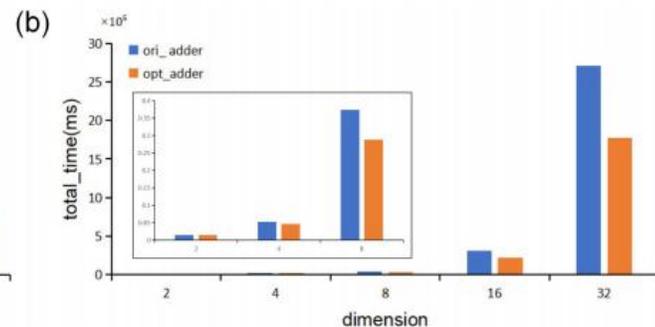
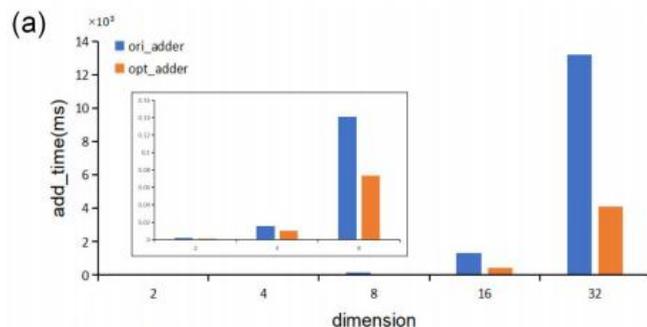
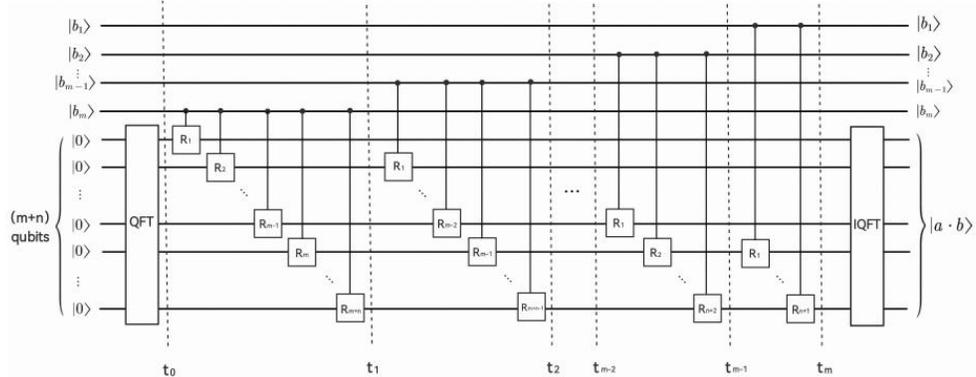
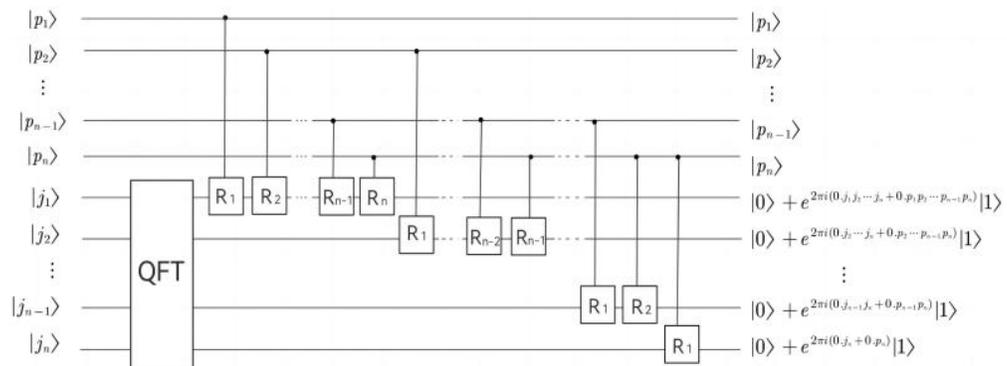
# 经典神经网络的量子加速

- **量子矩阵乘法**

- 基于交换测试、Hadamard测试的方案；
- 基于HHL算法的方案；
- 基于量子傅里叶变换的方案；
- 量子+经典的混合矩阵乘法（NISQ）

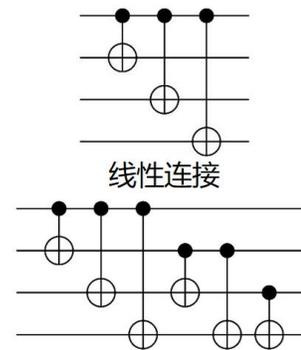
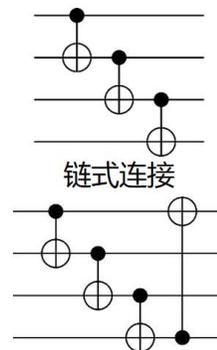
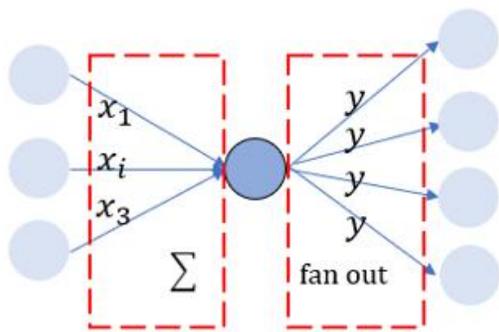
# 基于量子傅里叶变换的矩阵乘法

arXiv:2408.03085



# 一系列可讨论的问题

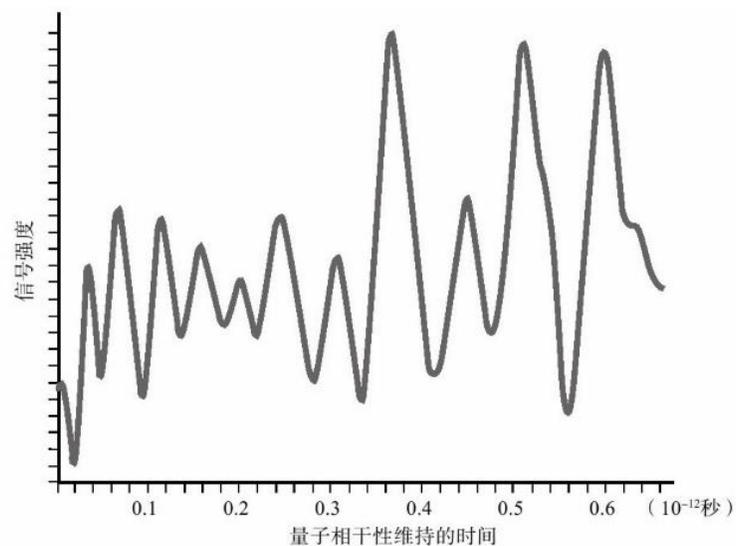
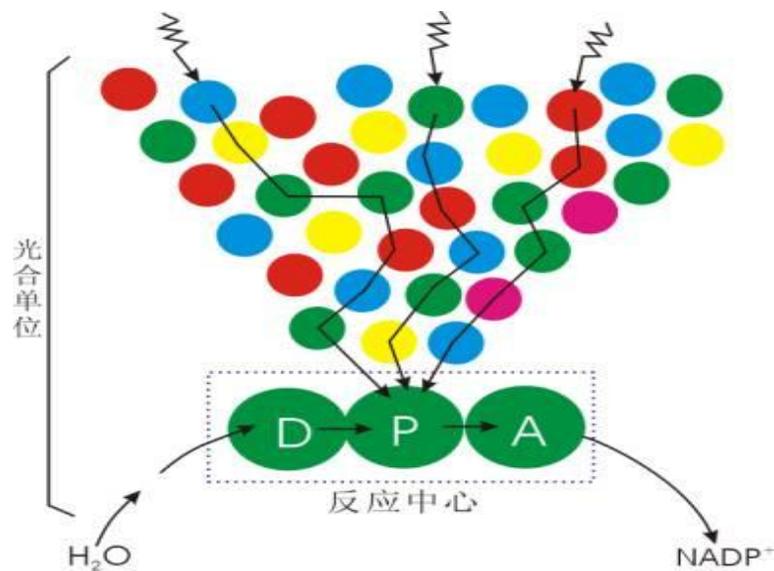
- 量子神经网络拟合连续非线性函数的能力如何？
- 量子神经网络就等于参数化量子线路吗？
- 网络架构与量子不可克隆定理的矛盾？
- 生物神经元处理信息是“量子”的吗？



# 光合作用中的量子节拍

- MIT的讨论会

- 2007年4月，麻省理工学院的一些数学家和物理学家正在举行例行的文献交流会
- Graham Fleming 2007的研究中通过飞秒激光脉冲研究FMO蛋白质光合作用复合体，得出如下数据图像，-196摄氏度



## 更多的证据

- 2009年，都柏林大学的Ian Mercer在另一种细菌的光合作用复合体LHC2中检测到了量子节拍，常温下完成。
- 2010年，安大略大学的Greg Scholes在海藻的光合作用系统中证实了量子节拍的存在。
- 其后，Fleming团队又在菠菜中发现了量子节拍。

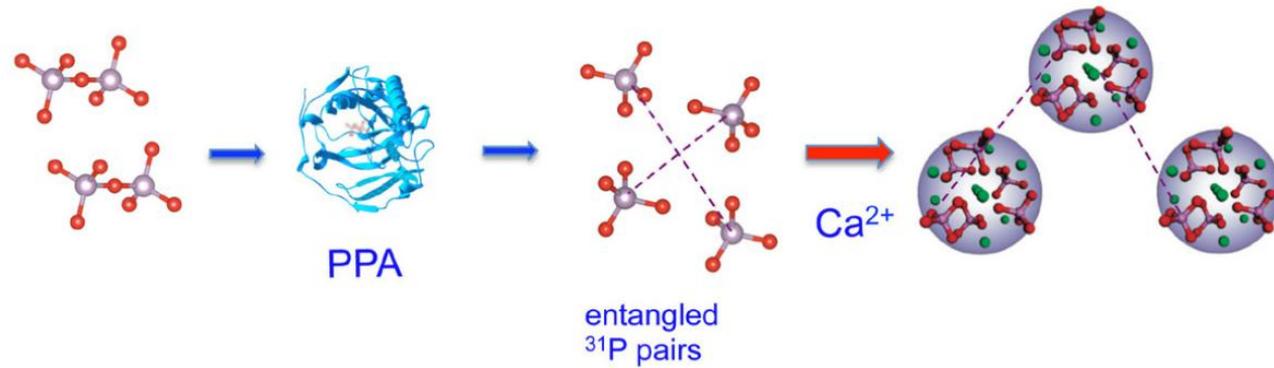
# 大脑是量子计算机吗？

- 这是一个大而无当的命题
  - 大脑中有量子效应吗？如果有，在哪儿？
  - 大脑如何利用高效的量子计算？
- 微管理论（一直被批判，《皇帝的新脑》）
- 神经元细胞膜上的离子通道可能是量子比特的载体
- 大脑中的核自旋

# 磷酸盐做为核自旋



Matthew Fisher



- PPA: 焦磷酸酶
- 波斯纳分子

## 结语

*Biologists and perhaps physicists will understand  
much better how the brain works. But why something that we call consciousness goes  
with those workings, I think that will remain mysterious*

—— E. Witten