



量子态制备及其在量子机器学习中的前景

赵健 陈昭昀 庄希宁 薛程 吴玉椿 郭国平

Quantum state preparation and its prospects in quantum machine learning

Zhao Jian Chen Zhao-Yun Zhuang Xi-Ning Xue Cheng Wu Yu-Chun Guo Guo-Ping

引用信息 Citation: *Acta Physica Sinica*, 70, 140307 (2021) DOI: 10.7498/aps.70.20210958

在线阅读 View online: <https://doi.org/10.7498/aps.70.20210958>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于混合光模式阵列的自由空间编码通信

Free-space optical communication based on hybrid optical mode array encoding

物理学报. 2017, 66(14): 144102 <https://doi.org/10.7498/aps.66.144102>

基于机器学习构建的环三亚甲基三硝胺晶体势

Energetic potential of hexogen constructed by machine learning

物理学报. 2020, 69(23): 238702 <https://doi.org/10.7498/aps.69.20200690>

铅基钙钛矿铁电晶体高临界转变温度的机器学习研究

High critical transition temperature of lead-based perovskite ferroelectric crystals: A machine learning study

物理学报. 2019, 68(21): 210502 <https://doi.org/10.7498/aps.68.20190942>

基于量子算法的量子态层析新方案

A novel scheme of quantum state tomography based on quantum algorithms

物理学报. 2019, 68(14): 140301 <https://doi.org/10.7498/aps.68.20190157>

一种编码式宽带多功能反射屏

A wideband coding reflective metasurface with multiple functionalities

物理学报. 2017, 66(6): 064203 <https://doi.org/10.7498/aps.66.064203>

基于鲁棒极端学习机的混沌时间序列建模预测

Chaotic time series prediction based on robust extreme learning machine

物理学报. 2018, 67(3): 030501 <https://doi.org/10.7498/aps.67.20171887>

专题: 机器学习与物理

量子态制备及其在量子机器学习中的前景*

赵健¹⁾ 陈昭昀¹⁾ 庄希宁¹⁾³⁾ 薛程¹⁾ 吴玉椿^{1)2)†} 郭国平¹⁾²⁾³⁾

1) (中国科学技术大学, 中国科学院量子信息重点实验室, 合肥 230026)

2) (合肥综合性国家科学中心人工智能研究院, 合肥 230088)

3) (合肥本源量子计算科技有限责任公司, 合肥 230026)

(2021 年 5 月 21 日收到)

经典计算机的运算能力依赖于芯片单位面积上晶体管的数量, 其发展符合摩尔定律. 未来随着晶体管的间距接近工艺制造的物理极限, 经典计算机的运算能力将面临发展瓶颈. 另一方面, 机器学习的发展对计算机的运算能力的需求却快速增长, 计算机的运算能力和需求之间的矛盾日益突出. 量子计算作为一种新的计算模式, 比起经典计算, 在一些特定算法上有着指数加速的能力, 有望给机器学习提供足够的计算能力. 用量子计算来处理机器学习任务时, 首要的一个基本问题就是如何将经典数据有效地在量子体系中表示出来. 这个问题称为态制备问题. 本文回顾态制备的相关工作, 介绍目前提出的多种态制备方案, 描述这些方案的实现过程, 总结并分析了这些方案的复杂度. 最后对态制备这个方向的研究工作做了一些展望.

关键词: 态制备, 量子机器学习, 编码**PACS:** 03.67.Ac, 03.67.Lx, 03.67.-a**DOI:** 10.7498/aps.70.20210958

1 引言

机器学习是一门人工智能领域的科学, 其通过计算机学习训练已知的数据, 并利用训练好的数据模式预测未知数据的信息. 随着计算机性能的不断增强, 机器学习对数据的处理能力也不断提升, 被广泛应用到各个领域^[1]. 这包括图像识别^[2,3]、人脸识别^[4-6]等分类问题, 也包括最优决策问题, 如 Alpha Go^[7], Alpha Zero^[8] 的围棋对弈等. 经典数据有许多处理和训练方式, 如神经网络、聚类等方法. 为了准确提取未知数据的特征, 训练方式的选择需要参考相应的数据类型. 当处理大规模的数据时, 为了获取数据特征, 往往采取深度学习的学习方式, 如包含数十亿权重的神经网络^[9], 这充分展示了深度学习在处理大数据时的效果.

当今的机器学习发展, 特别是在大数据的处理方面, 对经典计算机的运算能力有很高的需求. 1965 年戈登·摩尔提出摩尔定律, 指集成电路上可容纳的元器件数目约每两年增加一倍. 一方面, 在不久的将来随着晶体管在芯片上的间距接近 1 nm, 接近传统工艺制造的物理极限; 另一方面数据的爆炸式增长, 对算力需求越来越高. 于是为了应对大数据的处理, 需要一个创新的计算体系结构. 量子计算作为一种新的计算模型, 比起经典计算, 在一些特定算法上有着指数加速的能力, 有望为大数据的处理提供足够的计算需求. 如果一个量子信息计算处理器能够产生经典计算机难以模拟的统计模式, 那么量子计算与机器学习结合便可能识别经典机器学习难以识别的特征. 为此, 人们将量子运算和经典的机器学习相结合, 提出了机器学习的量子版本, 称为量子机器学习, 并将这种寄希望于量子

* 国家重点研究发展计划 (批准号: 2016YFA0301700)、国家自然科学基金 (批准号: 11625419)、安徽省量子信息技术倡议 (批准号: AHY080000) 和中国科学院战略重点研究计划 (批准号: XDB24030600) 资助的课题.

† 通信作者. E-mail: wuyuchun@ustc.edu.cn

机器学习的优势称为量子计算在经典机器学习中的潜在的加速能力^[10]. 量子机器学习包括用经典机器学习的方法处理量子物理中的问题和用量子计算的方式解决经典机器学习的问题. 前者需要将量子物理中的量子态转换为经典数据, 再用经典机器学习的方法来提取数据信息, 如构造经典神经网络训练这些经典数据后, 得到某些量子态的特征. 后者在处理经典数据时, 某些步骤中的计算过程可以通过量子态的西变换来辅助实现, 这其中不可避免地需要将经典数据对应成量子态.

量子机器学习中, 需要运用量子计算机处理经典数据, 这涉及经典数据的在量子体系中的表示问题. 这种将经典数据映射到量子计算机中的过程, 称为态制备问题^[11,12]. 态制备的种类有很多, 大部分是将经典数据转换为了量子态, 也存在一些将经典数据映射到哈密顿量的方式. 态制备种类的选择直接影响了执行机器学习算法的选择, 这意味着不同的态制备方法决定了提取经典数据信息的差异, 影响了后续在量子系统里的操作, 影响整个机器学习算法的计算复杂度. 同时, 态制备作为量子机器学习的其中一部分, 其制备精度和成功率会影响整个机器学习算法的有效性.

态制备问题不受限于机器学习的应用, 它同样是一些算法的基础, 如解线性方程组的 HHL 量子算法^[13]. 基于解线性方程组的量子算法, 有量子主成分分析算法^[14], 可用于聚类 and 特征识别; 也有支持向量机算法^[15], 用于对大规模的数据分类问题. 这类量子算法的共同点都是为了解决实际的经典问题, 需要以经典数据为输入和输出. 这可以分为三个步骤: 首先运用态制备将经典数据转为量子态, 再用量子计算机对量子态进行酉变换, 最后多次地量子测量概率性得到一个经典结果. 整个算法的复杂度受各个步骤的影响, 本文仅列出不同态制备方式的复杂度. 如果考虑量子算法的复杂度, 可通过量子线路的语言, 对所需的基本量子操作, 即基本量子门计数得到所有门的个数. 类比于经典算法的分类方式, 量子算法分为不含黑箱 (oracle) 的显式算法和含黑箱的算法. 前者的复杂度指的是所有基本量子门个数, 后者往往忽略黑箱的执行时间而考虑黑箱的执行次数, 称为质询复杂度. 一般地, 若数据规模是 $O(N)$ 的, 量子基本门的时间层数是 $O(\text{Poly}(\log N))$ 的, 称量子算法的执行时间是有效

的. 在每个时间层, 允许多个量子比特同时执行一次量子基本门. 同样的数据规模, 若用到的量子比特是 $O(\text{Poly}(\log N))$ 的, 称量子算法的比特数是有效的. 量子基本门的个数受量子比特数和时间层数的影响, 在一个时间层至多有量子比特数的量子门同时执行, 故显式算法的复杂度上界为量子基本门的时间层数与比特数的乘积.

从经典计算机到量子系统态制备的方式叫作编码. 编码的种类大体上可以分为三种, 分别是基底编码、振幅编码和量子抽样编码^[12]. 基底编码用于处理二值数据向量, 将数据编码到量子态的基底上; 振幅编码是最为常见的态制备方式, 将数据编码到量子态的振幅上, 数据向量可以是连续变量, 数据特征信息体现到量子态的振幅大小; 量子抽样编码可以看成前两种编码的结合, 是对在整个计算基基底的经典概率分布进行振幅编码, 对于某个给定的经典概率分布, 量子抽样编码退化为了振幅编码. 上述由经典数据编码到量子态的过程, 在量子系统中也可以视为从初态到目标量子态的一种酉变换. 广义上讲, 可以称从经典数据到酉变换的过程为编码, 如由经典数据决定量子系统演化哈密顿量的方式也可以看成一种编码, 这种编码称为哈密顿量编码.

态制备中振幅编码的相关工作最为丰富, 除了平凡的编码方式, 振幅编码可以从 2002 年 Grover 和 Rudolph^[16] 的工作谈起, 其将满足条件可积的一种数据分布制备成了量子态, 制备过程依赖于经典函数的有效计算, 且没有给出量子线路语言, 编码的有效性需进一步探讨. Kaye 等^[17] 以类似的方式得到了任意量子态的制备, 给出了可称之为含黑箱的量子线路. Soklakov 和 Schack^[11] 于 2005 年用其他形式的黑箱给出了在一定限制条件下的有效的概率性算法. 振幅编码中不得不提到量子随机存取存储器 (quantum random access memory, QRAM) 的方法, 这是一种从已知量子态出发, 由经典数据直接得到新的量子态的过程.

本综述简要叙述各种态制备的编码方式, 并给出一些简单的例子. 根据各个编码方式的适用情景, 对不同编码进行比较, 列出态制备的复杂度, 表明应谨慎乐观对待量子态制备问题.

符号说明 文中希尔伯特空间用 \mathcal{H} 表示, 任意的单位化向量 $|\psi\rangle \in \mathcal{H}$ 表示量子态, 其中 $|0\rangle = (1, 0)^T$,

$|1\rangle = (0, 1)^T$. 泡利算符用 σ_x, σ_y 和 σ_z 表示, $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. 在 Bloch 球上, 单比特的绕 A 轴旋转门 $R_A(\theta) = e^{-i\theta\sigma_A/2} = \cos(\theta/2)I - i\sin(\theta/2)\sigma_A$, 其中 $A = x, y, z$.

2 编码方式

这里给出态制备的问题模型. 对给定的经典数据, 不妨假定数据集 $\mathcal{D} \subset \mathcal{R}^n$ 是有限集, $|\mathcal{D}| = l$, 每个数据 $x = (x_1, \dots, x_n) \in \mathcal{D}$, 用一个单射 f 将 \mathcal{D} 的所有子集构成的集合, 记为 $2^{\mathcal{D}}$, 映射到某个希尔伯特空间 \mathcal{H}^m , 使得对 $\mathcal{D}' \subset \mathcal{D}$, $f(\mathcal{D}') \in \mathcal{H}^m$. 称 f 为态制备, 其中 \mathcal{H}^m 中的元素都视为单位向量, 对应量子态. 例如, $\mathcal{D} = \{x_1 = (1, 0), x_2 = (0, 1)\}$, 可以找到一种态制备的映射, 使得 $f(\{x_1\}) = |10\rangle$, $f(\{x_2\}) = |01\rangle$, $f(\mathcal{D}) = 1/\sqrt{2}(|10\rangle + |01\rangle)$.

2.1 基底编码

这类编码中, 限定所有数据是二元向量, 或二值化处理后的经典数据是二元向量, 即 $\mathcal{D} \subset \{0, 1\}^n$. 对任意数据集的子集 $\mathcal{D}' \subset \mathcal{D}$, $\mathcal{D}' = \{x^j\}_{j=1}^{l'}$, $x^j = (x_1^j, \dots, x_n^j)$,

$$f(\mathcal{D}') = \frac{1}{\sqrt{l'}} \sum_{j=1}^{l'} |x^j\rangle = \frac{1}{\sqrt{l'}} \sum_{j=1}^{l'} |x_1^j, \dots, x_n^j\rangle.$$

这种编码方式将数据集中的所有数据, 编码到量子态的计算基上, 等权叠加. 制备过程中用到的量子比特数为 $\mathcal{O}(n)$. 制备的思路是运用迭代法.

1) 制备 $|\psi_1\rangle = \frac{1}{\sqrt{l'}} |x^1\rangle |0\rangle_1^{\otimes n} |0\rangle_2 + \sqrt{\frac{l'-1}{l'}} |0\rangle_1^{\otimes n} |1\rangle_2$. 注意到这里增加了两个寄存器, 其中第一个寄存器为了存放 $|x^2\rangle$, 第二个寄存器放置辅助比特标记位置.

2) 制备 $|\psi'_1\rangle = \frac{1}{\sqrt{l'}} |x^1\rangle |0\rangle_1^{\otimes n} |0\rangle_2 + \sqrt{\frac{l'-1}{l'}} |x^2\rangle^{\otimes n} |x^2\rangle_1^{\otimes n} |1\rangle_2$. 这里的操作是一系列并行的 X 门和单比特受控门.

3) 得到 $|\psi_{12}\rangle = \frac{1}{\sqrt{l'}} (|x^1\rangle + |x^2\rangle) |0\rangle_1^{\otimes n} |0\rangle_2 + \sqrt{\frac{l'-2}{l'}} |0\rangle_1^{\otimes n} |0\rangle_1^{\otimes n} |1\rangle_2$. 这里的操作涉及对上一步含 $|x^2\rangle$ 的态进行受控 Y 操作, 含 $|x^2\rangle$ 态的振幅也成为 $1/\sqrt{l'}$, 再进行退计算的操作还原即可得到.

以同样的迭代方法, 可以得到 $|\psi_{1\dots l'}\rangle$, 即得到 $f(\mathcal{D}')$.

2.2 振幅编码

这类编码要求的数据不再是二元向量, 可以是任意实数. 对于任何 $\mathcal{D}' = \{x^j\}_{j=1}^{l'}$, $x^j = (x_1^j, \dots, x_n^j)$,

$$f(\mathcal{D}') = \frac{1}{C} \sum_{j=1}^{l'} \sum_{i=1}^n x_i^j |i, j\rangle,$$

这里 C 为归一化常数, $C = \sqrt{\sum_{j=1}^{l'} \sum_{i=1}^n (x_i^j)^2}$. 可以看出, 如果对数据集中所有数据振幅编码, 当 ln 是 2 的幂次时, 只需要 $\log(ln)$ 个量子比特便可以编码 ln 个振幅. 例如, $ln = 4$ 时, 只需制备一个 2 bit 的量子态, 使得在四个计算基 $|00\rangle, |01\rangle, |10\rangle$ 和 $|11\rangle$ 上的振幅为数据大小即可. 振幅编码问题可以简化为, 给定一个单点集合 $X = \{x = (x_0, \dots, x_{N-1})\} \subset \mathcal{R}^N$, N 为 2 的幂次, 使得在忽略归一化常数的条件下

$$f(X) = \sum x_i |i\rangle.$$

2.2.1 显式的编码

1) 用 $\log N$ 个量子比特编码.

基本的想法是利用迭代法, 用部分量子态对新粒子多重控制操作, 直到全部粒子完成态制备. 这个算法的执行时间是 $\mathcal{O}(N)$. 假定制备出的量子态的每个振幅的大小已知, 即每个计算基上测量得到相应的结果的概率 $p_i^{\log N} = x_i^2 / \|x\|^2$ 已知, 并且我们定义边际概率 $p_i^k = \sum_{i'=2i, 2i+1} p_{i'}^{k+1}$, $k = 1, \dots, \log N - 1$, 如图 1 所示.

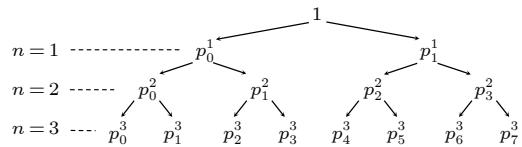
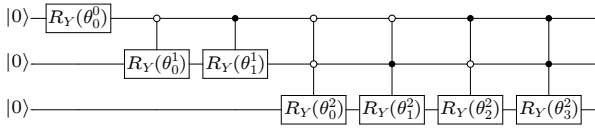


图 1 当 $N = 8$ 时所有的边际概率
Fig. 1. The marginal probabilities for $N = 8$.

假定 $\theta_i^k = 2 \arccos \sqrt{p_{i/2}^{k+1} / p_i^k}$ ($k = 0, \dots, n-1$; $i = 0, \dots, 2^k - 1$), 则态制备的整个迭代流程图可以参看图 2.

$$f(X) = \sum_{i=0}^{N-1} x_i |i\rangle.$$

图 2 $N = 8$ 时用 $\mathcal{O}(N)$ 的时间制备 $f(X)$ Fig. 2. Preparation for $f(X)$ in $\mathcal{O}(N)$ time for $N = 8$.

在用 $\log N$ 个量子比特编码时, 每个基底前的振幅都不能并行运算, 导致了这个方法的运行时间为 $\mathcal{O}(N)$. 如果这些多重受控操作可以并行操作, 运行时间将大大降低.

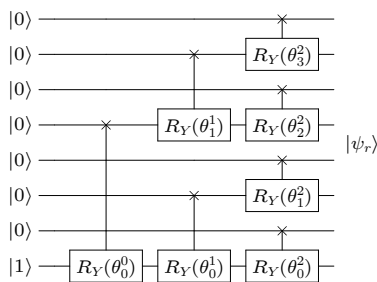
2) 用 N 个量子比特编码. 基于减少运行时间的考量, 可以增加量子比特, 使得编码振幅的基底选择性更多, 从而增加并行运算的可行性. 这里选取 W 态的基底, 得到态制备的映射为

$$f(X) = x_0|0 \cdots 01\rangle + \cdots + x_{N-1}|10 \cdots 0\rangle \\ = \sum_{i=1}^N x_i |2^i\rangle.$$

令

$$Y(\theta) = \begin{bmatrix} 1 & & & \\ & \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} & \\ & \sin \frac{\theta}{2} & \cos \frac{\theta}{2} & \\ & & & 1 \end{bmatrix} \triangleq \begin{array}{c} \text{---} \times \text{---} \\ \text{---} R_Y(\theta) \text{---} \end{array},$$

可以注意到 $Y(\theta)|01\rangle = \cos(\theta/2)|01\rangle + \sin(\theta/2)|10\rangle$, $Y(\theta)|10\rangle = -\sin(\theta/2)|01\rangle + \cos(\theta/2)|10\rangle$, 这类似于对单比特量子门 $R_Y(\theta)$ 在 $|0\rangle$ 和 $|1\rangle$ 上的操作, 故将 $Y(\theta)$ 定义为由符号“ \times ”控制的 $R_Y(\theta)$ 量子门. 这里直接给出 $N = 8$ 时的整个迭代过程, 详见图 3.

图 3 $N = 8$ 时用 $\mathcal{O}(\log N)$ 的时间获取振幅Fig. 3. Acquiring the amplitudes in $\mathcal{O}(\log N)$ time for $N = 8$.

2.2.2 含黑箱的编码

这类编码不考虑黑箱构造的问题, 有两大类制备方案.

1) Grover 等^[16] 和 Kaye 等^[17] 的工作

Grover 在 2002 年提出将满足条件可积的经典数据制备成量子态的方法. 给定一个离散概率分布 $\{p_i\}_{i=0}^{N-1}$, 目标制备 $|\psi\rangle = \sqrt{p_i}|i\rangle$. 等价于给定单点集 $\{x = (x_0, \cdots, x_{N-1}), x_i \geq 0\}$, 满足归一化条件 $\sum_i x_i^2 = 1$. 制备的思路与显式编码相同, 都是运用迭代法, 为了描述方便, 仍然采用边际概率的记号 $p_i^k = \sum_{i'=2i, 2i+1} p_{i'}^{k+1}$, $k = 1, \cdots, \log N - 1$.

1) 制备 $|\psi_1\rangle = \sqrt{p_0^1}|0\rangle + \sqrt{p_1^1}|1\rangle$. 这一步操作只需要一个单比特门即可.

2) 制备 $|\psi_2\rangle = \sqrt{p_0^2}|00\rangle + \sqrt{p_1^2}|01\rangle + \sqrt{p_2^2}|10\rangle + \sqrt{p_3^2}|11\rangle$. 该步骤的思路是运用迭代法, 具体来讲, 首先, 令 $f(0) = \frac{p_0^2}{p_0^2 + p_1^2}$, $f(1) = \frac{p_2^2}{p_2^2 + p_3^2}$ 表示条件概率, 在 1) 的基础上加上一个寄存器, 用来存储条件概率, 即

$$\sqrt{p_i^1}|i\rangle|0\rangle \rightarrow \sqrt{p_i^1}|i\rangle|\theta_i\rangle, \quad (1)$$

这里 $\theta_i = \arccos \sqrt{f(i)}$. 这一步操作并未指明寄存器量子比特的数量与执行时间. 可以视为以 f 为黑箱的一步操作. 原文中该步骤的有效实现需要 f 为条件可积函数, 并要求在经典计算中对 f 这步黑箱操作是有效计算的. 那么从经典计算机到量子计算机, 通过辅助量子比特将经典计算转变为可逆计算并退计算等一系列操作, 一定存在时间复杂度与经典运算的时间复杂度相同的量子线路, 使得该步骤可以有效实现^[18]. 接着, 再增加一个寄存器, 受控于 θ_i 的操作, 得到 $|\psi_2\rangle$ 的振幅信息, 即

$$\sqrt{p_i^1}|i\rangle|\theta_i\rangle|0\rangle \rightarrow \sqrt{p_i^1}|i\rangle|\theta_i\rangle(\cos \theta_i|0\rangle + \sin \theta_i|1\rangle).$$

之后进行退计算操作, 将 θ_i 擦除, 这一步操作的量子比特数量与执行时间及存储 θ_i 相同, 同样是含 f 的黑箱操作, 得到 $|\psi_2\rangle$.

运用迭代法, 由 $|\psi_k\rangle$ 得到 $|\psi_{k+1}\rangle$, 最终得到 $|\psi_{\log_2 N}\rangle$, 即目标量子态 $|\psi\rangle$.

Kaye 等^[17] 的态制备方法与 Grover 类似, 给出了量子线路的语言. 其中对存储 θ_i 的步骤进行细化, 在已知 $|\psi\rangle$ 的前提下, 将 (1) 式改写为

$$|\psi\rangle \sqrt{p_i^1}|i\rangle|0\rangle \rightarrow |\psi\rangle \sqrt{p_i^1}|i\rangle|\theta_i\rangle, \quad (2)$$

这步黑箱操作表明 θ_i 的获得需要整个态 $|\psi\rangle$ 的各个分量的值, 并未给出黑箱操作的具体构造, 在这一点上与 Grover 的算法没有本质区别.

评价含黑箱的算法复杂度, 通常不考虑黑箱以外的线路, 这是由于黑箱的结构相比于显式的量子线路更为复杂. 如果同时考虑黑箱内部的执行时间和黑箱外的量子门执行时间, 对于任意 N 规模的量子态, 是不可能用 $\mathcal{O}(n)$ 的量子比特在有效时间完成的. 因此, 我们往往考虑黑箱的执行次数, 称为质询复杂度, 以此来衡量含黑箱算法的计算复杂度. Kaye 的算法对于任意的量子态都可以制备, 并且从含黑箱的角度看出是以 (2) 式为黑箱的振幅编码, 该编码方式具有有效的质询复杂度. 不过, 值得说明的是 Grover 和 Kaye 的算法原文中并没有指明是含黑箱的算法. 给定数据集 X , 制备过程可以视为含黑箱的量子算法. 若是未指明某个数据集上的经典数据, 对数据集中的元素随机化处理, 如数据集中的元素满足某种概率分布函数 g , 对这种分布的态制备问题可能是有效的, 因为 g 的参数可能不依赖于 n . 这种含黑箱的编码比较广泛, 将在 2.3 节的量子抽样编码中再次提及.

II) Soklakov 和 Schack^[11] 的工作

真正意义上经典的含黑箱的振幅编码可参看 Soklakov 等的工作. 这类算法属于概率性的量子算法, 态制备给出了理想态的近似量子态. 数据向量不局限于实空间, 即 $X = \{x = (x_0, \dots, x_{N-1})\} \subset \mathcal{C}^N$, N 为 2 的幂次, 这里 $x_i = |x_i|e^{i\alpha_i}$, $\alpha_i \in [0, 2\pi)$, 但 $|x_i|$ 不可以全相等. 理想的量子态为

$$f(X) = \frac{1}{\|x\|} \sum_{i=0}^{N-1} x_i e^{i\alpha_i} |i\rangle \triangleq |\psi\rangle. \quad (3)$$

该编码的执行时间受限于两个因素. 一方面是数据集本身 x 各个分量实部的差异, 如果各个分量 $|x_i|$ 大小都比较接近, 那么编码执行时间会很快. 另一方面是对制备量子态结果保真度和成功率的要求, 如果对态制备的结果要求严苛, 会导致执行时间变慢. 令 $\gamma, \lambda, \eta \in (0, 1)$, 如果对任意的数据分量 $x_i^2 < 1/\eta N$, 以不小于 $1 - \gamma$ 的成功率制备的近似态为 $|\tilde{\psi}\rangle$, 满足 $|\langle \tilde{\psi} | \psi \rangle| = 1 - \lambda$, 所需要的计算复杂度为 $\mathcal{O}(P(\log_2 N \cdot \gamma^{-1} \lambda^{-1} \eta^{-1}))$.

算法的核心内容是选择合适的黑箱, 对获取目标量子态所有分量振幅的大小做分割, 并从振幅大分量向振幅小的分量标记, 最终用 Grover 搜索算法, 将目标态的近似态以一定的成功率找到.

2.2.3 QRAM

量子随机存取存储器 (QRAM) 是类比于经典

内存存取数据的一种装置, 可以将经典数据存储到相干的量子态各个分量地址中. 在读取量子态的任意一个分量时, 每个分量地址上都需要附带经典数据的信息:

$$\sum_i \psi_i |i\rangle \rightarrow \sum_i \psi_i |i\rangle |D_i\rangle. \quad (4)$$

QRAM 存取数据的过程中, 第一个寄存器存储经典数据作为指标, 要求对任意量子态 $\sum_i \psi_i |i\rangle$, 分量都需要存储经典数据地址信息. 第二个寄存器是数据寄存器, 用于存储经典数据 X . 这种装置类似于 (1) 式, (2) 式的操作方式, 故一定程度上, QRAM 的模型包含了 Grover 和 Kaye 等的态制备工作. 例如在量子推荐系统算法中^[19], 概率分布被提前储存到 QRAM 中, Grover 的算法也可以实现. 理论上, QRAM 的模型可以通过增加大量的比特数来减小执行时间. QRAM 通过二分的树状图和桶队结构 (bucket-brigade) 来实现, 这种实现方式可以做到 $\mathcal{O}(n)$ 的时间复杂度, 但量子比特数是 $\mathcal{O}(N)$ 的. QRAM 的量子线路语言实现方式种类较多^[20,21]. 人们在后续的工作中更关心哪一种 QRAM 的实现方式更具有噪声的抗性和可拓展性, Hann^[22] 给出了一种关于噪声抗性的论证.

(1) 式, (2) 式以及 QRAM 的直接形式 (4) 式都是将经典数据存储到辅助比特上, 每个分量对应的辅助比特上都有经典数据的信息. 特别地, 如果辅助比特可以写成二进制数, 这种变换称为数字编码. 与之对应的, Mitarai 称振幅编码为模拟编码^[23], 并介绍了数字编码和模拟编码的转换关系. 利用这种数模转换的关系, 可以得到振幅编码的具体形式.

具体来讲, 如果 QRAM 的操作完成后, 振幅编码可以通过条件受控和后选择的方式得到振幅编码的概率性量子算法. 给定数据集 X , QRAM 可以将 $x = (x_0, \dots, x_{N-1})$ 存储到等权叠加的量子态上, 忽略归一化, 得到

$$\sum_{i=0}^{N-1} |i\rangle |0\rangle \rightarrow \sum_{i=0}^{N-1} |i\rangle |x_i\rangle. \quad (5)$$

进行条件受控操作, 通过增加辅助比特和受控操作实现,

$$\sum_{i=0}^{N-1} |i\rangle |x_i\rangle \rightarrow \sum_{i=0}^{N-1} |i\rangle |x_i\rangle (x_i |0\rangle + \sqrt{1 - |x_i|^2} |1\rangle), \quad (6)$$

(6) 式中的受控旋转操作可以将 x_i 表示为 t 比特的二进制数, 分别对辅助比特做控制 $R_Y(\pi/2^t)$ 类似的操作来实现. 最后一步进行后选择的操作, 对辅助比特进行测量, 当测量到 $|1\rangle$ 态时, 制备失败. 需要重复这个算法的流程, 直到测量值 $|0\rangle$. 当测量值为 $|0\rangle$ 态时, 成功制备, 成功率可以计算, 得到

$$\sum_{i=0}^{N-1} x_i |i\rangle |x_i\rangle. \quad (7)$$

(7) 式得到的量子态与目标态还多了数据寄存器的数据, 需要擦除. 这步擦除数据是退计算的过程, 也是 QRAM 里 (5) 式的逆操作, 即

$$\sum_{i=0}^{N-1} x_i |i\rangle |x_i\rangle \rightarrow \sum_{i=0}^{N-1} x_i |i\rangle. \quad (8)$$

2.3 量子抽样编码

本节介绍基底编码和振幅编码的一种混合编码方式——量子抽样编码. 振幅编码的数据集 X 是单点集, $X = \{x = (x_0, \dots, x_{N-1})\}$, 如果用 $\log N$ 个量子比特, 时间复杂度为 $\mathcal{O}(N)$. 在量子抽样编码中, 给定一种概率分布, 不妨假定为 $g(x')$, $x' \in [0, N]$, 表示量子态的在每个基底的概率为 $p_i = \int_i^{i+1} g(x') dx'$, $p_i \geq 0$, $\sum p_i = 1$. 数据集仍为单点集 $X = \{x = (x_0, \dots, x_{N-1})\}$, $x_i = \sqrt{p_i}$. 目标量子态是 X 对应的振幅编码

$$f(x) = \sum_{i=0}^{N-1} x_i |i\rangle. \quad (9)$$

这类编码的编码技术是从已有的量子态出发, 根据经典数据的分布 $g(x')$ 得到新的量子态, 与一般的振幅编码相比, 这类编码的 $g(x)$ 参数可能与 N 无关. 可以通过对分布函数 $g(x')$ 做一些限制, 得到关于某些函数性质有关的态制备方法.

在 Grover 等 [16] 的工作中, 作者进一步提出对于很大一类被称为“对数凸”的函数, 都可以通过这种编码方式进行制备, 这其中包括常见的正态分布和指数分布. 除了 Grover 的工作, Kitaev 和 Webb [24] 也分析了高斯分布的量子态制备. 文献 [25] 给出了一种基于矩阵积态 (matrix product state) 方法, 得到了当 $g(x')$ 为光滑可微且导函数有界时的编码方式. 该算法需要 $\mathcal{O}(n)$ 个量子比特, 执行时间为 $\mathcal{O}(n)$. 实验方面 Vazquez 和 Woerner [26] 给出

了基于量子振幅放大算法简化态制备的方法, 并在 IBM 的物理比特上展示了实现了该算法.

以机器学习的方式研究这类情形的态制备问题有大量的工作, 如生成对抗网络 [27], 给定经典数据的分布, 利用含参数的量子线路生成一种量子分布, 再由对抗识别器对量子分布采样, 反复调整, 直到识别不出经典数据分布与生成的分布, 训练完毕. 数值模拟过程中所需要的量子门数量控制在 $\mathcal{O}(P(n))$. 也有其他机器学习相关的工作, 如 Arrazola 等 [28] 用含参线路在光量子计算机模拟器演示了许多量子态的生成, 如 ON 态和 GKP 态. 值得一提的是, 此类统计分布的态制备问题, 在量子蒙特卡罗模拟算法中处于非常核心的地位 [29], 而后者已经被证明在很多金融和其他模拟问题中显示出量子优越性 [30–32].

2.4 哈密顿量编码

经典数据转为量子态的另一种方案是将经典数据的信息编码到某个量子系统的哈密顿量中, 运用哈密顿模拟的方式代替将经典数据转为量子态的方法 [33]. 记经典数据集为 $X = \{x = (x_0, \dots, x_{N-1})\}$, 记对应的哈密顿量为 H_x , 表明哈密顿量依赖于经典数据的选择. 则对于量子系统的初态 $|\psi\rangle$, 演化时间为 t , 得到演化后的量子态为

$$|\psi'\rangle = e^{-iH_x t} |\psi\rangle.$$

哈密顿量的演化过程可以由量子线路语言实现 [18]. 考虑一个 n 量子比特的量子系统, 哈密顿量可以分解为一些哈密顿量的和, 即 $H = \sum_{i=0}^{L-1} H_i$, 其中 H_i 为较易模拟的哈密顿量, $L = \mathcal{O}(P(n))$. 由 Trotter 公式 [33],

$$e^{-iHt} = (e^{-iHt/n})^n = \left(\prod_{i=0}^{L-1} e^{-iH_i t/n} + \mathcal{O}((t/n)^2) \right)^n.$$

当 n 充分大时, 可以用多次的演化 $e^{-iH_i t/n}$ 来实现量子模拟. H_i 常见的选择是泡利算符.

哈密顿量编码的步骤分为两大类. 一类是从哈密顿量 H 出发. 1) 由经典数据确定哈密顿量 H , 如经典数据为动能、势能函数决定的参量; 2) 在某个量子系统中选定基底, 确定哈密顿量的矩阵元; 3) 哈密顿量模拟. 模拟过程包含哈密顿量的分解, 需要确定 H_i . Matto 等 [34] 用格雷码序的方式对经典数据进行哈密顿量编码, 是一种比特数有效的编

码方案. 另一类是从分解后的哈密顿量 H_i 出发. 1) 选定 H_i , 由经典数据得到每个 H_i 的系数; 2) 在某个量子系统中选定基底, 确定哈密顿量 H_i 的矩阵元; 3) 得到总的哈密顿量, 即为经典数据的哈密顿量编码. 例如用第二类的方法, 假定经典数据 $x = (x_0, \dots, x_{4^n-1})$, 我们将经典数据 x 编码到 n 粒子哈密顿量中, 以泡利算符和单位算符在计算基上的表示为基底, 基底个数为 4^n , 总的哈密顿量为

$$H = x_0(I_1 \otimes \dots \otimes I_n) + x_1(\sigma_{1x} \otimes I_2 \otimes \dots \otimes I_n) + x_{4^n-1}(\sigma_{1z} \otimes \sigma_{2z} \otimes \dots \otimes \sigma_{nz}).$$

3 复杂度

当数据规模为 $\mathcal{O}(N)$ 时, 基底编码的执行时间为 $\mathcal{O}(N)$ 次, 需要的量子比特数为 $\mathcal{O}(N)$, 复杂度为 $\mathcal{O}(N^2)$. 振幅编码适用范围广, 其中显式振幅编码的复杂度为 $\mathcal{O}(N) \log N$, 对于含黑箱的振幅编码, 仅考虑质询复杂度, 可以做到有效制备, 即在一定条件下质询复杂度可以达到 $\mathcal{O}(\log(N))$, 但黑箱的执行时间在实际操作中需要考量. QRAM 的编码方式从已知量子态获取经典数据得到新的量子态, 这个过程中需要 $\mathcal{O}(N)$ 个量子比特, 但执行时间可以做到 $\mathcal{O}(n)$. 而量子抽样编码也是直接从量子态出发, 比较目标量子态与量子初态的差异得到新的量子态, 比特数和时间有效性都可以实现, 这是由于给出的分布函数可以不依赖于数据量规模. 在哈密顿量编码中, 哈密顿量的合理选择可使得编码方式中的比特数有效, 执行时间取决于哈密顿量的矩阵形式和哈密顿量模拟的精度. 进一步的, 考虑到量子比特数和时间执行次数的平衡和取舍, 噪声抗

性等因素, 态制备问题应该被仔细斟酌. 以上的复杂度分析可参看表 1.

4 研究前景和展望

在未来, 经典计算机芯片的工艺制程接近摩尔定律极限, 经典计算机的算力发展达到瓶颈期. 而大数据的处理使得算力需求呈快速增长趋势. 这之间算力的供需矛盾关系使得人们迫切地寻找新的计算模式. 研究量子机器学习的出发点是解决这种矛盾关系. 具体来说是在处理某一类问题时, 量子机器学习的方法能够大大缩短传统经典机器学习需要的时间, 继而在更广泛的问题中表现出加速能力. 量子计算机的实用化受限于量子比特、量子门的质量和量子操作系统等诸多实验因素, 故量子机器学习的研究大多停留在数值模拟或是构建理论模型的阶段. 在这个大背景下, 人们不过多关注量子机器学习中的物理实现.

目前态制备问题里更受关注的是量子抽样编码, 其中涌现出了许多利用量子机器学习研究态制备的工作. 这种编码方式通过经典神经网络与含参量子线路的结合, 以监督学习的方式训练参数, 不断优化量子线路得到近似的目标量子态. 复杂度的分析通常考虑参数的数量, 但含参量子线路的表示能力与学习方式的选择都会影响其编码的有效性. 这类工作比较丰富, 例如生成对抗网络^[27], 利用含参数的量子线路生成一种分布并由对抗识别器采样, 机器学习的方式训练参数, 直到对抗识别器识别不出目标分布与生成的分布; Arrazola 等^[28]采用的是光量子计算机模拟器, 利用自动微分法优化得到目标量子态; 最近 Zhou 等^[35]提出了一种自

表 1 态制备的不同编码方式的复杂度分析
Table 1. Complexity analysis of kinds of encoding methods for state preparations.

编码方式	数据类型	比特数	执行时间
基底编码	二值数据	$\mathcal{O}(N)$	$\mathcal{O}(N)$
显式振幅编码	计算基 $ i\rangle$ 上编码	连续变量	$\mathcal{O}(\log N)$
	$ 2^i\rangle$ 上编码	连续变量	$\mathcal{O}(N)$
含黑箱振幅编码	Grover和Kaye	连续变量	$\mathcal{O}(\log N)$
	Soklakov和Schack	连续变量	$\mathcal{O}(\log N)$
	QRAM	连续变量	$\mathcal{O}(\log N)$
量子抽样编码	分布函数	$\mathcal{O}(\log N)$	$\mathcal{O}(\log N)$
哈密顿量编码	连续变量	$\mathcal{O}(N)/\mathcal{O}(\log N)$	$\mathcal{O}(\text{Poly}(\log N))^*$

注: *同时受保真度和成功率的影响.

动微分的量子含参线路, 可以优化得到任意的量子态.

另一方面, 态制备问题作为经典数据和量子态的桥梁, 在量子机器学习中的使用不可避免. 相较于经典计算机编码数据的方式, 量子计算机在态制备时编码数据的指数级加速能力是没有问题的, 但这是以大量量子比特数为前提的实现方式. 研究量子机器学习的初衷是实用化解决经典问题, 更应该考虑其中的态制备方案的时间计算资源和空间计算资源. 对于复杂度的分析, 态制备的算法复杂度至少是数据自由度的量级, 既要分析时间复杂度, 也要考虑量子比特数的规模. 单看时间复杂度, 得出具有加速能力的结论还不足以体现量子机器学习的能力, 分析时应该谨慎. 但同时也要乐观对待量子机器学习的能力, 至少以发展的眼光去看待. 例如, 大数分解的量子算法复杂度比已知最优的经典算法有指数级的提高, 而人们在大数分解算法提出前也不清楚量子计算的加速能力. 总的来说, 随着量子计算机的发展特别是硬件水平的提升, 相信会有更多的人关注态制备问题.

参考文献

- [1] Jordan M I, Mitchell T M 2015 *Science* **349** 255
- [2] Lay K T, Katsaggelos A K 1990 *Opt. Eng.* **29** 436
- [3] Lu D, Weng Q 2007 *Int. J. Remote Sens.* **28** 823
- [4] Samaria F S, Harter A C 2002 *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision Sarasota, December 5–7, 1994* p138
- [5] Guillaumin M, Verbeek J, Schmid C 2009 *In 2009 IEEE 12th international conference on computer vision Kyoto, Japan, September 29–October 2, 2009* p498
- [6] Sun Y 2015 *Deep Learning Face Representation by Joint Identification-verification* (Ann Arbor: ProQuest LLC) pp40–57
- [7] Silver D, Huang A, Maddison C J, Guez A, Sifre L, Driessche G V, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D 2016 *Nature* **529** 484
- [8] Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, Driessche G V, Graepel T, Hassabis D 2017 *Nat. Nature* **550** 354
- [9] Le Q V 2013 *IEEE International Conference on Acoustics, Speech and Signal Processing Vancouver, Canada, May 26–31, 2013* p8595
- [10] Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S 2017 *Nature* **549** 195
- [11] Soklakov A N, Schack R 2006 *Phys. Rev. A* **73** 012307
- [12] Schuld M, Petruccione F 2018 *Supervised Learning with Quantum Computers* (Vol. 17) (Berlin: Springer) pp139–171
- [13] Harrow A W, Hassidim A, Lloyd S 2009 *Phys. Rev. Lett.* **103** 150502
- [14] Lloyd S, Mohseni M, Rebentrost P 2014 *Nat. Phys.* **10** 631
- [15] Rebentrost P, Mohseni M, Lloyd S 2014 *Phys. Rev. Lett.* **113** 130503
- [16] Grover L, Rudolph T 2002 arXiv: 0208112 v1 [quant-ph].
- [17] Kaye P, Mosca M 2001 *International Conference on Quantum Information New York, USA, June 13, 2001* p28
- [18] Nielsen M A, Chuang I 2002 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press) pp120–215
- [19] Kerenidis I, Prakash A 2016 arXiv: 1603.08675 v3 [quant-ph].
- [20] Matteo O D, Gheorghiu V, Mosca M 2020 *IEEE Trans. Quantum Eng.* **1** 4500213
- [21] Paler A, Oumarou O, Basmadjian R 2020 *Phys. Rev. A* **102** 032608
- [22] Hann C T, Lee G, Girvin S M, Jiang L 2021 *PRX Quantum* **2** 020311
- [23] Mitarai K, Kitagawa M, Fujii K 2019 *Phys. Rev. A* **99** 012301
- [24] Kitaev A, Webb W A 2008 arXiv: 0801.0342 [quant-ph].
- [25] Holmes A, Matsuura A Y 2020 *In 2020 IEEE International Conference on Quantum Computing and Engineering (QCE) Denver, CO, USA, October 12–16, 2020* p169
- [26] Vazquez A C, Woerner S 2021 *Phys. Rev. A* **15** 034027
- [27] Zoufal C, Lucchi A, Woerner S 2019 *npj Quantum Inf.* **5** 103
- [28] Arrazola J M, Bromley T R, Izaac J, Myers C R, Brádler K, Killoran N 2019 *Quantum Sci. Technol.* **4** 024004
- [29] Montanaro A 2015 *Proc. R. Soc. A* **471** 20150301
- [30] Orus R, Mugel S, Lizaso E 2019 *Rev. Phys.* **4** 100028
- [31] Stamatopoulos N, Egger D J, Sun Y, Zoufal C, Iten R, Shen N, Woerner S 2020 *Quantum* **4** 291
- [32] Woerner S, Egger D J 2019 *npj Quantum Inf.* **5** 15
- [33] Lloyd S 1996 *Science* **273** 1073
- [34] Matteo O D, McCoy A, Gysbers P, Miyagi T, Woloshyn R M, Navrátil P 2021 *Phys. Rev. A* **103** 042405
- [35] Zhou P F, Hong R, Ran S J 2021 arXiv: 2104.14949[quant-ph].

SPECIAL TOPIC—Machine learning and physics

Quantum state preparation and its prospects in quantum machine learning^{*}

Zhao Jian¹⁾ Chen Zhao-Yun¹⁾ Zhuang Xi-Ning¹⁾³⁾ Xue Cheng¹⁾
Wu Yu-Chun^{1)2)†} Guo Guo-Ping¹⁾²⁾³⁾

1) (*CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China*)

2) (*Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China*)

3) (*Origin Quantum Computing Company Limited, Hefei 230026, China*)

(Received 21 May 2021)

Abstract

The development of traditional classic computers relies on the transistor structure of microchips, which develops in accordance with Moore's Law. In the future, as the distance between transistors approaches to the physical limit of manufacturing process, the development of computation capability of classical computers will encounter a bottleneck. On the other hand, with the development of machine learning, the demand for computation capability of computer is growing rapidly, and the contradiction between computation capability and demand for computers is becoming increasingly prominent. As a new computing model, quantum computing is significantly faster than classical computing for some specific problems, so, sufficient computation capability for machine learning is expected. When using quantum computing to deal with machine learning tasks, the first basic problem is how to represent the classical data effectively in the quantum system. This problem is called the state preparation problem. In this paper, the relevant researches of state preparation are reviewed, various state preparation schemes proposed at present are introduced, the processes of realizing these schemes are described, and the complexities of these schemes are summarized and analyzed. Finally, some prospects of the research work in the direction of state preparation are also presented.

Keywords: state preparation, quantum machine learning, encoding

PACS: 03.67.Ac, 03.67.Lx, 03.67.-a

DOI: 10.7498/aps.70.20210958

^{*} Project supported by the National Key Research and Development Program of China (Grant No. 2016YFA0301700), the National Natural Science Foundation of China (Grant No. 11625419), the Anhui Initiative in Quantum Information Technologies, China (Grant No. AHY080000), and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB24030600).

[†] Corresponding author. E-mail: wuyuchun@ustc.edu.cn