

## 量子生成模型

孙太平 吴玉椿 郭国平

### Quantum generative models for data generation

Sun Tai-Ping Wu Yu-Chun Guo Guo-Ping

引用信息 Citation: *Acta Physica Sinica*, 70, 140304 (2021) DOI: 10.7498/aps.70.20210930

在线阅读 View online: <https://doi.org/10.7498/aps.70.20210930>

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

周期驱动量子伊辛模型中非热统计的形成与抑制

Formation and suppression of nonthermal statistics in periodically driven quantum Ising models

物理学报. 2020, 69(14): 140501 <https://doi.org/10.7498/aps.69.20191657>

量子计算与量子模拟

Quantum computation and quantum simulation

物理学报. 2018, 67(12): 120301 <https://doi.org/10.7498/aps.67.20180710>

量子相干

Quantum coherence

物理学报. 2019, 68(3): 030304 <https://doi.org/10.7498/aps.68.20181779>

横场中具有周期性各向异性的一维XY模型的量子相变

Quantum phase transitions of one-dimensional period-two anisotropic XY models in a transverse field

物理学报. 2017, 66(18): 180302 <https://doi.org/10.7498/aps.66.180302>

基于量子隐形传态的量子保密通信方案

Quantum communication scheme based on quantum teleportation

物理学报. 2017, 66(23): 230303 <https://doi.org/10.7498/aps.66.230303>

Tavis-Cummings模型中的几何量子失协特性

Geometric quantum discord in Tavis-Cummings model

物理学报. 2018, 67(11): 110301 <https://doi.org/10.7498/aps.67.20172699>

## 专题: 机器学习与物理

## 量子生成模型\*

孙太平<sup>1)</sup> 吴玉椿<sup>1)2)†</sup> 郭国平<sup>1)2)3)</sup>

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 17 日收到; 2021 年 6 月 30 日收到修改稿)

近年来, 很多基于生成模型的机器学习算法, 如生成对抗网络、玻尔兹曼机、自编码器等, 在数据生成、概率模拟等方面有广泛的应用. 另一方面, 融合量子计算和经典机器学习的量子机器学习算法也不断被提出. 特别地, 量子生成模型作为量子机器学习的分支, 目前已有很多进展. 量子生成算法是一类量子-经典混合算法, 算法中引入参数量子线路, 通过执行参数线路得到损失函数及其梯度, 然后通过经典的优化算法来优化求解, 从而完成对应的生成任务. 与经典生成模型相比, 量子生成模型通过参数线路将数据流映射到高维希尔伯特空间, 在高维空间中学习数据的特征, 从而在一些特定问题上超越经典生成模型. 在中等规模含噪声的量子体系下, 量子生成模型是一类有潜力实现量子优势的量子机器学习算法.

关键词: 量子机器学习, 量子生成模型, 量子优势

PACS: 03.67.-a, 03.67.Ac

DOI: 10.7498/aps.70.20210930

## 1 引言

在经典计算领域, 生成模型被广泛应用于各个领域, 如图像生成<sup>[1]</sup>、语音合成<sup>[2]</sup>、药物设计<sup>[3]</sup>、图像转换<sup>[4]</sup>等. 近来已有对经典生成模型的量子版本的研究, 如量子生成对抗网络<sup>[5,6]</sup>、量子玻尔兹曼机<sup>[7,8]</sup>、量子玻尔兹曼机<sup>[9–12]</sup>、量子自编码器<sup>[13–15]</sup>以及量子贝叶斯网络<sup>[16]</sup>等.

量子生成模型是受经典机器学习模型启发, 使用量子计算机来改进经典模型中的生成步骤, 以完成数据生成、概率模拟任务的一类算法.

这类算法的主要步骤是, 通过量子线路生成量子态, 用量子态观测结果构建损失函数, 再通过经典的优化算法进行迭代优化, 从而得到最优的参数变量集合以完成相应的生成任务. 因此, 这些算法

被统称为量子-经典混合算法. 这种量子-经典混合的闭环优化思想始于变分量子算法<sup>[17]</sup>, 即通过量子计算机和参数化量子门操作集合获得分子的量子态表示, 测量分子模型对应的哈密顿量的期望值, 以此作为目标函数进行变分优化.

量子相位估计<sup>[18]</sup>在大数分解<sup>[18]</sup>、HHL<sup>[19]</sup>等被证明有量子优势的算法上具有广泛的应用, 但基于以下原因相位估计算法无法在当前的计算体系下有效实现: 1) 当前硬件体系多集中于单量子门和两比特纠缠门的实现, 相位估计用到的受控幺正门 ( $c-U^j$ ) 必须分解为基本门的组合, 此过程会显著降低受控幺正门的保真度; 2) 相位估计用到的量子傅里叶逆变换需要  $O(n^2)$  ( $n$  为量子比特个数) 数量级的量子门, 使得深层量子线路的态保真度无法达到量子计算要求. 量子计算体系的噪声不可避免, 当前只有中等规模的量子比特、无法实现

\* 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)、中国科学院战略重点研究计划 (批准号: XDB24030600) 和安徽省量子信息技术倡议 (批准号: AHY080000) 资助的课题.

† 通信作者. E-mail: wuyuchun@ustc.edu.cn

有效的量子纠错<sup>[20]</sup>技术, 这使量子-经典混合算法受到广泛关注.

尽管已有多种量子生成任务在真实量子芯片上完成, 如利用量子生成对抗网络进行图像生成与识别<sup>[21]</sup>, 但对于参数量子线路 (Parameterized Quantum Circuit, PQC, 见附录) 的表示能力与经典网络相比是否具有优势还缺少理论的结果. Du 等<sup>[22]</sup>尝试从内在的线路结构出发证明其结论, 而经典深度学习神经网络和量子线路的比较还没有结论. 同样的, 量子生成网络作为量子机器学习的一部分, 如何将经典数据转化为量子态、如何从测量结果获取经典信息仍需要更直观更普适的方法.

本文主要分析一些常见量子生成模型的优势和劣势, 并对其扩展性展开讨论. 量子生成模型的分类与经典生成模型类似, 是以任务为导向的. 量子生成模型基本沿袭与之对应的经典生成模型的命名, 如量子生成对抗网络是使用量子线路来完成经典生成对抗任务的一种量子生成模型. 对于利用量子体系特有的性质以完成生成任务的, 如量子玻恩机, 则是用量子专有的玻恩测量命名. 还有一些模型保留了原作者的命名方式, 归并在“其他生成模型”这一节进行介绍. 一些量子生成模型如量子生成对抗网络、量子玻恩机、量子自编码器等用到的参数量子线路, 在附录里会有详细的介绍. 通过对这些模型比较, 可以从不同角度理解量子生成模型的构建思路, 以此为不同的计算任务构建与之适应的计算模型提供启发.

## 2 量子生成模型

### 2.1 量子生成对抗网络

经典的生成对抗网络<sup>[23]</sup>(generative adversarial Network, GAN) 被视为深度学习中的前沿进展. GAN 在生成图像<sup>[1]</sup>、语音<sup>[2]</sup>、文本<sup>[24]</sup>等方面具有非常广泛的应用. 经典生成对抗网络是一类有两个玩家的博弈游戏, 一个是生成器 (generator), 记为  $G$ , 另一个为鉴别器 (discriminator), 记为  $D$ . 对于已存在的源数据  $\vec{x} \sim p_{\text{data}}$ , 生成器  $G$  生成新的数据以逼近真实的源数据分布, 使  $D$  无法区分数据的来源; 而鉴别器  $D$  对每个输入数据给出真或假的决策概率, 以最大程度地区分这两类数据.

如果  $G$  产生的数据足够逼近目标数据,  $D$  无

论对于何种输入数据, 都有 50% 的几率误判, 这时生成对抗网络就达到了 Nash 平衡点. 设真实的源数据标签为 1, 由  $G$  产生的数据标签为 0,  $D$  输出  $(0, 1)$  间的概率值, 那么生成对抗网络的损失函数可以被定义为:

$$\mathcal{L}(D, G) = \mathbb{E}_{\vec{x} \sim p_{\text{data}}} [\log D_{\theta_D}(\vec{x})] + \mathbb{E}_{\vec{z} \sim p(\vec{z})} [\log(1 - D_{\theta_D}(G_{\theta_G}(\vec{z})))] \quad (1)$$

生成对抗网络的训练策略是, 更新其中一方的参数时, 另一方的参数保持不变. 对于上述的损失函数模型, 优化任务为  $\min_{\theta_G} \max_{\theta_D} \mathcal{L}(D, G)$ . 采用经典神经网络构建的 GAN 模型, 可以用反向传播算法来实现梯度计算.

量子的生成对抗网络<sup>[5,21,25,26]</sup>(quantum GAN, QGAN) 基本保留了经典生成对抗模型的特征, 不同之处在于量子生成对抗网络用到的数据集既可以是量子的也可以是经典的, 而且可选择地, 生成器与鉴别器也可以为经典神经网络或参数量子线路.

对于数据是量子的, 生成器与鉴别器也都是量子的情况, Dallaire-Demers 等<sup>[5]</sup>提出了这样的 QGAN 模型: 源数据的概率分布模型为  $p_R(x)$ , 由某个未知 (有可能很复杂) 的过程  $R$  生成. 由参数  $\theta_G$  确定的生成器  $G$  生成新的数据样本, 由参数  $\theta_D$  确定的鉴别器  $D$  识别这两类数据. 用标签  $\lambda$  指定源数据中的类别,  $R = R(|\lambda\rangle) = \rho_{\lambda}^R$ , 考虑噪声  $|z\rangle$ , 可以按照 (1) 式定义与参数有关的损失函数  $\mathcal{L}(\theta_D, \theta_G)$ , 这里  $D$  的输入是  $R(|\lambda\rangle)$  或  $G(\theta_G, |\lambda, z\rangle)$ :

$$\mathcal{L}(\theta_D, \theta_G) = \frac{1}{2} + \frac{1}{2\Lambda} \sum_{\lambda=1}^{\Lambda} [\cos^2(\phi) \text{tr}(Z \rho_{\lambda}^{DR}(\theta_D)) - \sin^2(\phi) \text{tr}(Z \rho_{\lambda}^{DG}(\theta_D, \theta_G, z))], \quad (2)$$

其中含  $\phi$  的项是调整源数据和生成部分占整体的比例. 一般取  $\phi = \frac{\pi}{4}$ . Dallaire-Demers 等构建的量子生成对抗网络线路图如图 1 所示.

量子生成对抗网络的任务就是通过参数优化来实现  $\min_{\theta_G} \max_{\theta_D} \mathcal{L}(\theta_D, \theta_G)$ , 从而得到最接近于源数据概率分布的分布模型与鉴别模型. 量子线路的梯度可以用数值差分<sup>[27]</sup>、参数位移法则<sup>[27]</sup>、以及 Hadamard test 线路<sup>[5]</sup>来求. Dallaire-Demers 在数值模拟上实现了源数据为  $|0\rangle, |1\rangle$  两个量子态的生成对抗任务.

Lloyd 等<sup>[6]</sup>讨论了当数据是量子的, 鉴别器是

量子或经典的, 生成器是经典的量子生成对抗模型. Lloyd 所述的量子数据是通过对一个量子体系进行固定的测量得到的数据分布. 根据量子优越性<sup>[19]</sup>的表述, 经典的生成器不能有效地匹配量子数据分布, 或者更准确地, 除非有指数的资源可用, 经典的生成器无法模拟量子数据的分布  $p_R(x)$ . 对于数据是经典的, 生成器和鉴别器都是量子的情况, Lloyd 通过对比得出, 经典的生成器和鉴别器在计算梯度时需要  $O(N^2)$  个比特和  $O(N^2)$  步时间 ( $N$  是经典数据的维度), 而量子的生成器和鉴别器在计算协方差矩阵时, 如果矩阵是低秩和稀疏的, 需要  $O(\log N)$  个量子比特和  $O(\text{poly}(\log N))$  个逻辑门操作.

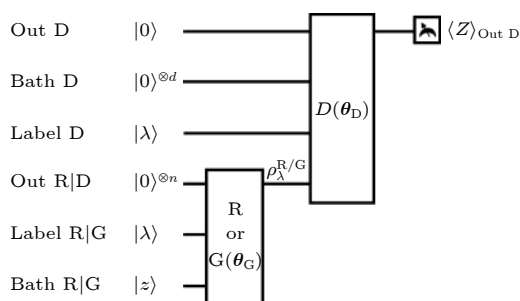


图 1 量子生成对抗网络结构<sup>[6]</sup>

Fig. 1. The general structure of QGAN<sup>[6]</sup>.

Zeng 等<sup>[26]</sup>结合量子玻恩机构建了经典数据-量子生成器-经典鉴别器的量子生成对抗模型, 实现条带图 (bars and stripes, BAS)<sup>[8,16,26,28]</sup>的概率生成与识别 (见第 2 节). 类似的, Situ 等<sup>[29]</sup>将 BAS 图的每个点都编码至量子比特上, 通过参数量子线路构建的生成器生成量子态, 测量结果输入至经典神经网络构建的鉴别器, 也实现了 BAS 图的生成. Zouful 等<sup>[25]</sup>通过测量量子态在计算基上的投影, 将其转化为数据对应的概率, 以此为基础构建生成对抗网络, 实现了对数正态分布、双峰高斯分布等概率分布的学习和生成. 值得注意的是, 连续数据被截断并舍入到临近整数值, 所以该模型仍是离散数据的概率学习模型.

随着量子硬件操控水平的提高, 在量子芯片上实现量子生成对抗网络已经有相关的工作. Hu<sup>[30]</sup>等利用超导量子体系首次实现了具有量子数据、量子生成器、量子鉴别器的量子生成对抗网络的实验验证, 生成保真度高达 98.8%. 在 Hu 的实验中, 量子数据为单比特量子态, 生成器参数为量子门旋转

角度. 在优化  $D$  或  $G$  时采用梯度下降算法, 量子系统的操控和测量通过 FPGA (field programmable gate arrays) 实现, 线路梯度通过经典计算机估算得到. 该实验对于 NISQ 体系下的量子机器学习优势的探索具有重要意义. 此外, Huang 等<sup>[21]</sup>通过将测量量子生成器得到的数据作为经典图像灰度值, 完成了手写数字集 0 和 1 的生成对抗任务, 为量子比特资源受限条件下经典数据的生成提供了一个思路.

## 2.2 量子玻恩机

量子玻恩机<sup>[7,8]</sup> (quantum circuit Born machine, QCBM) 是基于波函数的概率解释的一种模型. 量子玻恩机将概率分布编码至初态上, 把量子态在计算基上的投影作为数据的概率. 量子玻恩机是一种隐式的生成模型, 因为我们实际上并不需要知道这个纯态的密度矩阵, 而只需要知道叠加态振幅的模值. 这样的设定对量子线路的结构限制也更少, 所以一般被认为具有更强的表示能力.

对目标概率分布  $\pi(x)$  进行采样得到数据集  $D = \{x\}$ , QCBM 通过参数量子线路把初态  $|0\rangle$  演化为  $|\psi_\theta\rangle$ , 并在计算基上的测量得到生成数据  $x \sim p_\theta(x) = |\langle x|\psi_\theta\rangle|^2$  的概率分布. 训练目标是让模型的概率分布  $p_\theta$  接近真实数据分布  $\pi(x)$ . QCBM 采用所谓的 MMD (Maximum Mean Discrepancy) 损失函数:

$$\mathcal{L} = \left\| \sum_x p_\theta(x) \phi(x) - \sum_x \pi(x) \phi(x) \right\|^2, \quad (3)$$

其中  $\phi$  将  $x$  映射到高维 Hilbert 空间. 进一步地, 定义核函数  $K(x, y) = \phi(x)^T \phi(y)$ , (3) 式即可避免在高维空间中进行运算, 等价于

$$\begin{aligned} \mathcal{L} = & \mathbb{E}_{x \sim p_\theta, y \sim p_\theta} [K(x, y)] - 2 \mathbb{E}_{x \sim p_\theta, y \sim \pi} [K(x, y)] \\ & + \mathbb{E}_{x \sim \pi, y \sim \pi} [K(x, y)]. \end{aligned} \quad (4)$$

核函数可以取为高斯形式:  $K(x, y) = \exp\left(-\frac{1}{2\sigma}|x - y|^2\right)$ . 直接计算上述损失函数的梯度比较困难, 但可以证明, 下式无偏地等价于梯度:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \theta_l^\alpha} = & \mathbb{E}_{x \sim p_{\theta+}, y \sim p_\theta} [K(x, y)] - \mathbb{E}_{x \sim p_{\theta-}, y \sim p_\theta} [K(x, y)] \\ & - \mathbb{E}_{x \sim p_{\theta+}, y \sim \pi} [K(x, y)] + \mathbb{E}_{x \sim p_{\theta-}, y \sim \pi} [K(x, y)], \end{aligned} \quad (5)$$



其中  $p_{\theta^+}(x)$  和  $p_{\theta^-}(x)$  是在参数  $\theta^\pm = \theta \pm \frac{\pi}{2} \mathbf{e}_l^\alpha$  下的模型输出概率,  $\mathbf{e}_l^\alpha$  是参数空间的第  $(l, \alpha)$  个单位矢量 ( $\alpha$  和  $l$  分别表示迭代的步数与当前迭代的第  $l$  个参数的下标).

对所有量子比特进行测量, 会得到计算基上的概率分布. 源数据被映射为经典的比特流数据, 与量子态的计算基一一对应. 这种映射是一种同构, 源数据空间和映射空间一样大, 因此数据集的数据必须是有限的. 对于  $N$  个数据的集合, 需要  $n \sim O(\log(N))$  个量子比特. 而生成线路中所用的量子门个数一般为  $O(\text{poly}(n))$ .

量子玻恩机是一种隐式生成模型, 模型本身并不涉及经典数据的特征维度, 甚至不涉及数据本身的形态. 在量子体系中不需要考虑数据的原本面貌是什么. 无论是图片、数字还是文字, 都会被编码为经典比特流数据. 一旦数据同构的方式确定了, 基于量子线路的优化体系就与原本的数据形态无关了.

利用量子玻恩机进行概率分布的学习和生成在条带图 (bars and stripes, BAS)<sup>[8,16,26,28]</sup> 任务中取得很好的结果, 对于理想的无误差测量, 量子玻恩机生成真实数据分布的比率为 95.7%<sup>[8]</sup>. 此外通过对量子玻恩机的非计算基测量, 可以得到更多的信息, 这种方法可用来生成先验概率分布, 如经典生成对抗网络的先验分布<sup>[31]</sup>. Cheng 等<sup>[32]</sup> 还从信息熵的角度比较了量子玻恩机与受限玻尔兹曼机<sup>[33,34]</sup>, 为理解基于概率的生成模型和基于能量的生成模型提供了一个新的视角. Coyle 等<sup>[35]</sup> 提出一种量子玻恩机的子类——量子伊辛玻恩机——并引入量子核函数<sup>[36]</sup> 来讨论量子学习优势.

### 2.3 量子玻尔兹曼机

与前述模型不同, 量子玻尔兹曼机<sup>[9–12]</sup> (quantum Boltzmann machine, QBM) 需要一种专用的量子计算机——量子退火机.

经典的玻尔兹曼机被广泛用作机器学习的基础工具, 例如深度信念网络<sup>[37,38]</sup> 和深度玻尔兹曼机<sup>[37,39]</sup>. 玻尔兹曼机是二进制数节点构成的概率生成网络, 由可见节点和隐藏节点组成, 记为  $z_a, a = 1, 2, \dots, N$ , 其中  $N$  是总节点数. 这里为了和后面的量子生成模型保持一致, 取  $z_a \in \{+1, -1\}$  而不是  $z_a \in \{+1, 0\}$ . 为区别可见节点和隐藏节点, 我们分

别记为  $z_v, z_i$ . 同时用向量符号  $\mathbf{v}, \mathbf{h}, \mathbf{z} = (\mathbf{v}, \mathbf{h})$  表示可见态、隐藏态和组合态. 在物理上其能量可以用伊辛模型表示:

$$E_{\mathbf{z}} = - \sum_a b_a z_a - \sum_{a,b} w_{ab} z_a z_b, \quad (6)$$

其中  $b_a, w_{ab}$  是在训练的过程中进行需要优化的参数. 定义负对数相似度损失函数:

$$\mathcal{L} = - \sum_{\mathbf{v}} P_{\mathbf{v}}^{\text{data}} \log P_{\mathbf{v}} = - \sum_{\mathbf{v}} P_{\mathbf{v}}^{\text{data}} \log \frac{\sum_{\mathbf{h}} \mathbf{e}^{-E_{\mathbf{z}}}}{\sum_{\mathbf{z}'} \mathbf{e}^{-E_{\mathbf{z}'}}}, \quad (7)$$

其中观测到可见态  $\mathbf{h}$  的概率为边际分布  $P_{\mathbf{v}} = Z^{-1} \sum_{\mathbf{h}} \mathbf{e}^{-E_{\mathbf{z}}}$ ,  $Z = \sum_{\mathbf{z}} \mathbf{e}^{-E_{\mathbf{z}}}$ . 那么玻尔兹曼机的任务就是优化哈密顿量的参量  $\theta \in \{b_a, w_{ab}\}$ , 使边际分布尽可能接近真实的数据分布  $P_{\mathbf{v}}^{\text{data}}$ . 这样就可以用梯度的方法来优化玻尔兹曼机:

$$\frac{\partial \mathcal{L}}{\partial \theta} = \sum_{\mathbf{v}} P_{\mathbf{v}}^{\text{data}} \frac{\sum_{\mathbf{h}} \partial_{\theta} E_{\mathbf{z}} \mathbf{e}^{-E_{\mathbf{z}}}}{\sum_{\mathbf{h}} \mathbf{e}^{-E_{\mathbf{z}}}} - \frac{\sum_{\mathbf{z}} \partial_{\theta} E_{\mathbf{z}} \mathbf{e}^{-E_{\mathbf{z}}}}{\sum_{\mathbf{z}} \mathbf{e}^{-E_{\mathbf{z}}}}, \quad (8)$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = \overline{\langle \partial_{\theta} E_{\mathbf{z}} \rangle_{\mathbf{v}}} - \langle \partial_{\theta} E_{\mathbf{z}} \rangle, \quad (9)$$

其中  $\langle \dots \rangle$  和  $\langle \dots \rangle_{\mathbf{v}}$  分别为对自由变量的玻尔兹曼均值以及对可见变量的玻尔兹曼均值.  $\overline{\langle \dots \rangle_{\mathbf{v}}} \equiv \sum_{\mathbf{v}} P_{\mathbf{v}}^{\text{data}} \langle \dots \rangle_{\mathbf{v}}$  是双重均值. 这样我们就得到更新法则:

$$\delta b_a = \eta (\overline{\langle z_a \rangle_{\mathbf{v}}} - \langle z_a \rangle), \quad (10)$$

$$\delta w_{ab} = \eta (\overline{\langle z_a z_b \rangle_{\mathbf{v}}} - \langle z_a z_b \rangle). \quad (11)$$

关于如何估算等式右边的两项均值, 在机器学习领域已经有较多的方案, 其中一个简化的玻尔兹曼机模型——称之为受限玻尔兹曼机 (restricted Boltzmann machine, RBM)——已经成功的应用于实例<sup>[40]</sup>.

量子玻尔兹曼机的出发点是使用作为物理实体的量子比特来替代经典的比特或节点. 然而不同的是, 量子力学体系的哈密顿量是  $2^N \times 2^N$  维的矩阵. 而且量子力学算符大多是不对易的, 这对于求解梯度十分不便. 横场伊辛模型的哈密顿量形式为:

$$H = - \sum_a \Gamma_a \sigma_a^x - \sum_a b_a \sigma_a^z - \sum_{a,b} w_{ab} \sigma_a^z \sigma_b^z, \quad (12)$$

其中  $\mathbf{I}, \sigma^z, \sigma^x$  为泡利矩阵, 下标  $a$  表示作用于第  $a$  个量子比特. 与经典玻尔兹曼机类似, 定义配

分函数和密度矩阵分别为:  $Z = \text{Tr}[e^{-H}]$ ,  $\rho = Z^{-1}e^{-H}$ . 对于给定的可见态 $|\mathbf{v}\rangle$ , 可定义边际分布  $P_{\mathbf{v}} = \text{Tr}[A_{\mathbf{v}}\rho]$ ,  $A_{\mathbf{v}} = |\mathbf{v}\rangle\langle\mathbf{v}| \otimes \mathcal{I}_h$ . 其中  $\mathcal{I}_h$  为作用在隐藏变量的单位阵.

那么类似地, 可以得到量子损失函数和对应梯度为:

$$\mathcal{L} = - \sum_{\mathbf{v}} P_{\mathbf{v}}^{\text{data}} \log \frac{\text{Tr}[A_{\mathbf{v}}e^{-H}]}{\text{Tr}[e^{-H}]}, \quad (13)$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = \sum_{\mathbf{v}} P_{\mathbf{v}}^{\text{data}} \left( \frac{\text{Tr}[A_{\mathbf{v}}\partial_{\theta}e^{-H}]}{\text{Tr}[A_{\mathbf{v}}e^{-H}]} - \frac{\text{Tr}[\partial_{\theta}e^{-H}]}{\text{Tr}[e^{-H}]} \right). \quad (14)$$

然而  $H$  和  $\partial_{\theta}H$  并不对易, 也就是  $\partial_{\theta}e^{-H} \neq -e^{-H}\partial_{\theta}H$ . 这样在求解期望值的时候就会十分困难. 尽管 Amin<sup>[9]</sup> 通过分析的方法估算除了损失函数的界, 在一些特定的条件下可以估算损失函数的梯度, 但与前述量子生成模型相比, 参数更新的成本已然增加了.

尽管如此, 量子玻尔兹曼机仍然可以用来完成监督学习任务, 可以生成概率分布, 可以进行鉴别学习. Amin 等<sup>[9]</sup> 用量子玻尔兹曼机实现了伯努利分布的概率生成. Kieferová<sup>[12]</sup> 用量子玻尔兹曼机设计了一种生成模型用以实现部分的量子态层析, 并且除非  $BQP = BPP$ , 经典计算机无法有效地模拟其训练过程. Dorband<sup>[41]</sup> 在 D-wave 量子退火机上演示了量子玻尔兹曼机的构建过程. 但是, 相较于其他量子生成模型, 量子玻尔兹曼机带来的近似误差是难以避免的, 同时也难以说明相较于经典版本的优势所在.

## 2.4 量子自编码器

在经典的数据处理中, 通过机器学习中的自编码器可以将数据从一个高维空间缩减到低维空间. 经典的自编码模型可以概括为: 通过参数化的神经网络来训练一个数据集, 其数据维度为  $(n+k)$ , 得到擦除  $k$  个特征或维度的新数据, 如果这些数据可以通过解码器 (神经网络) 重新生成接近于原数据的一组新数据, 那么留下的  $n$  维数据被称为压缩数据. 这个过程被称为数据压缩. 数据压缩对于处理稀疏特征数据集或高维数据集是十分直观和有效的手段. 受此启发, 量子自编码器<sup>[13–15]</sup> (quantum autoencoder) 被设计为量子态压缩的工具.

Romero 等<sup>[13]</sup> 提出的量子自编码器与经典类似: 记  $\{p_i, |\psi_i\rangle_{AB}\}$  是  $(n+k)$  个量子比特上的纯态

集合, 子系统 A 和子系统 B 分别占有  $n$  和  $k$  个比特. 记  $U^{\theta}$  为作用在  $(n+k)$  个比特上的幺正算符, 其中  $\theta = \{\theta_1, \theta_2, \dots\}$  是参数集合, 以及  $|a\rangle_{B'}$  为  $k$  比特寄存器的参考纯态. 通过经典的优化方式, 我们希望平均保真度函数取得最大值:

$$C_1(\theta) = \sum_i p_i \cdot F(|\psi_i\rangle, \rho_{i,\theta}^{\text{out}}), \quad (15)$$

其中保真度定义为  $F(|\psi_i\rangle, \rho_{i,\theta}^{\text{out}}) = \langle\psi_i|\rho_{i,\theta}^{\text{out}}|\psi_i\rangle$ ,  $\rho_{i,\theta}^{\text{out}}$  是线路经过解码后的密度矩阵:

$$\rho_{i,\theta}^{\text{out}} = (U^{\theta})_{AB'}^{\dagger} \text{Tr}_B[U_{AB}^{\theta}[\psi_{iAB} \otimes a_{B'}](U^{\theta})_{AB}^{\dagger}]U_{AB'}^{\theta} \quad (16)$$

其中  $\psi_{iAB} = |\psi_i\rangle\langle\psi_i|_{AB}$ , 以及  $a_{B'} = |a\rangle\langle a|_{B'}$ .

量子自编码器的核心在于, 如果想要有效地压缩量子态, 必须保证态  $|\psi_i'\rangle = U^{\theta}|\psi_i\rangle_{AB}$  是 A, B 两个子系统的直积态. 即当下式成立时, 才有  $C_1 = 1$ :

$$U^{\theta}|\psi_i\rangle_{AB} = |\psi_i^c\rangle_A \otimes |a\rangle_B, \quad (17)$$

其中  $|\psi_i^c\rangle_A$  是  $|\psi_i\rangle$  的压缩后的态.

对输出态进行量子态层析以确定其密度矩阵是非常困难的. 可以证明测量 B 和 B' 寄存器的态保真度等价于 AB 寄存器初态和末态的保真度 (我们称 B 寄存器输出态为冗态). 这样, 可以定义另一种平均保真度  $C_2(\theta)$ ,

$$C_2(\theta) = \sum_i p_i \cdot F(\text{Tr}_A[U^{\theta}|\psi_i\rangle_{AB} \langle\psi_i|_{AB} (U^{\theta})^{\dagger}], |a\rangle_B). \quad (18)$$

实际上求两个态的保真度, 有一种 SWAP 测试线路可以完成<sup>[42]</sup>.

通过参数量子线路来构建作用于 AB 系统的幺正变换, 并通过 SWAP 测试线路得到保真度信息. 取损失函数为  $\mathcal{L} = -\log(1 - C_2)$ , 采用经典方式进行优化, 可以得到保真度最高的参数线路, 这样就实现了数据压缩.

如图 2 所示, 量子态压缩模型可以由以下步骤来构建: 第一, 制备初态  $|\psi_i\rangle$  和参考态  $|a\rangle$ , 并假设制备是有效的; 第二, 构建参数线路, 初态在幺正算符  $U$  下演化; 第三, 通过 SWAP 测试得到冗态和参考态之间的保真度.

量子自编码器在氢分子态的应用是一个成功的例证<sup>[13]</sup>. 其方法是用 4 qubits 来模拟氢分子, 通过变分量子求解器 (VQE) 求出对于固定原子间距的量子态, 以此作为量子自编码器的输入态. 氢分子的态并非占满了 4 bit 的希尔伯特空间, 这就为量子态压缩提供了可能. 直观上看, 量子态压缩的

过程就像在整个希尔伯特空间中将若干子空间碎片像拼图一样拼接成紧密的空间. 量子自编码器在量子加法器方面也有一定的应用<sup>[14]</sup>, 这种加法器用以实现将两个相同比特数的量子子系统编码到一个新的子空间. Ding 等<sup>[43]</sup>借助超导量子体系在实验上验证了量子加法器的可行性, 实验保真度与理论值一致.

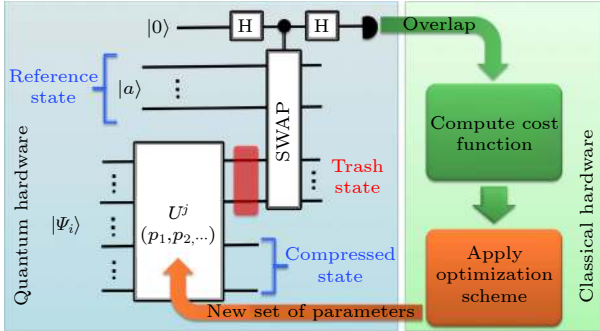


图2 量子自编码器的训练示意图, 其中  $\mathbf{p}$  表示变量, 摘自<sup>[13]</sup>

Fig. 2. Schematic representation of the hybrid scheme for training a quantum autoencoder where  $\mathbf{p}$  represents variables, image from<sup>[13]</sup>.

Pepper 等<sup>[44]</sup>采用了 Romero 等<sup>[13]</sup>提出的量子自编码器框架, 不同的是, 其数据压缩任务是将多维量子比特 (qutrit,  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle$  (已归一化)) 压缩为二维量子比特 (qubit) 态. 通过光量子实验体系, Pepper 等在实验上实现了 qutrits 的压缩, 并且证明了这样的实验优化系统对扰动是鲁棒的. Bondarenko 等<sup>[45]</sup>注意到经典数据通过经典自编码器可以有效地去除噪声, 因此量子自编码器同样有类似的作用. 通过构建以自旋翻转噪声和任意么正门噪声为主的 GHZ 含噪声量子态, Bondarenko 等验证了量子自编码器在量子数据去噪方面的作用.

如何判断量子态能否通过量子自编码器有效压缩一直受到关注, Huang 等<sup>[46]</sup>在理论上论证了实现有效量子自编码态压缩的条件, 并证明了若输入数据的最大线性无关向量的个数不超过隐变量空间 (压缩后的低维空间) 的维度, 那么量子自编码器就可以无损地将高维数据压缩至低维. Huang 等利用量子光学体系在实验上验证了其理论的结果. 此外 Cao 等<sup>[47]</sup>提出可在量子退火机上运行、通过噪声辅助的绝热模型, 同样可以完成量子自编码任务. 同时, 对量子态压缩的探索也使得量子梯度消失的现象——也即贫瘠高原的现象——得到进一步的研究<sup>[48]</sup>.

## 2.5 其他生成模型

从前述量子生成模型可以看出, 经典机器学习给我们带来了很多启发. 类似地, 经典领域中因子图 (factor graph) 可以作为概率生成模型, Gao 等<sup>[49]</sup>提出了一种量子生成模型 (quantum generative model, QGM), 并证明此模型可以高效地表示任意因子图. 并且通过理论证明至少存在一些特定的概率分布, 可以被 QGM 有效表示, 但不能被任意具有多项式量级变量的因子图有效表示. 如果多项式层次作为 P vs. NP 问题的泛化是不塌缩的话, QGM 相对于经典的因子图具有指数优势.

同时, 除了 QGM 的表示能力和泛化能力, Gao 等<sup>[49]</sup>还提出了一种通用的量子学习算法, 该算法利用张量网络、对多体纠缠量子态的父哈密顿量 (parent Hamiltonian) 构建、以及递归的量子相位估计进行学习. 显然希望这种算法对任意问题都能有多项式时间内解决是不现实的, 但作者证明了至少存在一些问题, QGM 相对于任意经典算法均有指数加速的效果, 前提是假设量子计算机不能被经典计算机有效地模拟.

直观地理解 QGM 中的指数加速效果: 机器学习中的生成模型旨在通过寻找潜在的概率分布, 对自然界任意的数据生成过程进行建模. 由于自然界受量子力学定律支配, 用经典生成模型中的概率分布对真实世界中的数据进行建模具有很大的局限性. Gao 等<sup>[49]</sup>提出的 QGM 模型, 用一个多体纠缠量子态的概率幅对数据的相关性进行参数化. 因为量子概率幅的相干性产生的现象比经典的复杂得多, QGM 才有可能在特定情况下实现性能的显著提升. QGM 使用图态  $|G\rangle$  表示, 由  $m$  个量子比特与图  $G$  构成. 如图 3 所示.

Gao<sup>[50]</sup>有另一篇工作同样旨在发掘量子相关性在生成模型中的作用. 其中比较了经典贝叶斯网络顺序模型与对应的一维张量网络, 展示了矩阵直积态 (matrix product states, MPSs) 比经典机器学习模型有更强的表示能力. 因为一维模型可以在经典计算机上被有效模拟, Gao 等采用数值模拟的方式测试了真实数据集的生成, 发现 MPS 生成能力比经典模型有提升. 此结果为理解基于 MPS 的机器学习算法提供了新的视角, 因为有一类张量网络子类不能有效地被经典计算机模拟, 但是可以在量子计算机上构建, 这就使得在机器学习中的生成模型框架下量子优势是可能存在的.



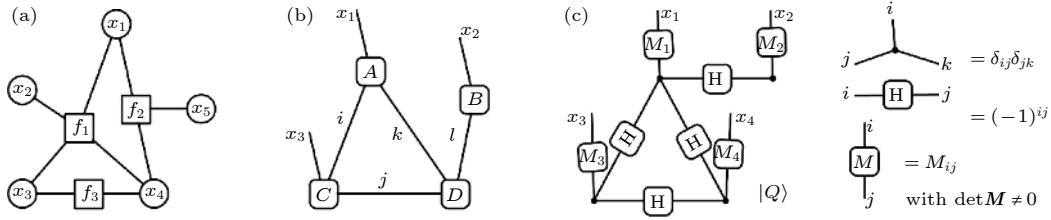


图 3 经典和量子生成模型 (a) 因子图表示; (b) 张量网络态表示; (c) 量子生成模型定义<sup>[49]</sup>

Fig. 3. Classical and quantum generative models: (a) Illustration of a factor graph; (b) illustration of a tensor network state; (c) QGM definition<sup>[49]</sup>.

除以上的量子生成模型外, 还有基于贝叶斯理论提出量子贝叶斯线路<sup>[16]</sup>. 这种量子线路引入辅助比特表示先验概率分布, 通过受控操作传递给数据比特寄存器. 这样可以极大地规避参数量子线路中出现的生成概率不均匀的问题.

### 3 量子生成模型展望

随着量子计算机硬件的不断发展, 量子生成模型也不只在量子虚拟机上得到实现, 基于真实量子芯片的生成模型受到了广泛的关注, 但是量子生成模型存在的问题仍待解决.

首先, 量子生成模型是否可以生成任意概率分布的数据? 对于连续变量的数据, 量子生成模型还无法像经典神经网络一样有效地生成. 同时, 数据是量子还是经典的决定了生成模型的选取. 对于经典数据, 利用 MPQCs 或 MPSs 生成量子态, 对量子态的测量得到了量子信息的经典统计结果. 如何更有效地生成量子态以及如何从量子态读取更多信息一直受到关注, 不同的量子生成模型给出了不同的策略, 然而对于任意连续变量数据的生成还需要进一步的探索.

其次, 量子生成模型作为量子机器学习的一部分, 如何体现量子计算的优势, 仍然是一个开放性问题. 对量子线路的表示能力的研究为我们提供了一个思路, 但是对于不同的生成模型很难给出统一的表示能力定义. 而 Gao 等<sup>[49,50]</sup>从量子相关性出发, 直观上将量子内在相关性的参数化来实现经典生成模型无法模拟的任务, 说明了量子优势在机器学习领域中的存在性.

对于量子生成模型, 还需要进一步探索. 在量子领域, 量子生成模型提供了 NISQ 体系下可执行的算法, 在机器学习领域, 它引入了量子计算和量子信息的理念, 使机器学习算法突破经典计算的

瓶颈成为可能.

感谢中国科学技术大学薛程博士和赵健博士的讨论.

### 附录 A 参数量子线路

量子计算机遵循量子力学定律进行逻辑运算和信息存取. 量子力学体系区别于经典体系之处在于, 由于叠加态和纠缠态的存在, 相同位数的量子数据比经典数据有着指数强的表示能力. 将量子信息转化为经典信息需要对量子态进行测量.

量子计算的基本计算单位是量子比特 (qubit), 单比特量子叠加态记为  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  (已归一化). 根据玻恩的几率解释, 对单比特量子态进行测量, 得到 0 态的概率为  $|\alpha|^2$ , 得到 1 态的概率为  $|\beta|^2$ . 对于多量子比特系统, 根据量子力学体系的完备性, 量子态可在计算基上展开, 由  $|\Phi\rangle = \sum_i a_i |i\rangle$  测得任意态  $|i\rangle$  的概率为  $|a_i|^2$ . 对于泡利算符  $\sigma_i$ ,  $i \in \{0, x, y, z\}$ , 其期望值可以通过算符的谱来计算, 即  $\langle \Phi | \sigma_i | \Phi \rangle = \langle \Phi | \sum_j \lambda_j |\psi_j\rangle \langle \psi_j| | \Phi \rangle$ , 其中  $\lambda_j$  为本征值,  $|\psi_j\rangle$  为本征态. 对于一个更复杂的哈密顿量  $H$ , 我们总能将其分解为若干泡利算符的直积, 再分别求出各个部分的期望值, 即  $H = \sum_{i,j,k,\dots,m} \alpha_{i,j,k,\dots,m} \cdot \sigma_i \otimes \sigma_j \otimes \sigma_k \cdots \otimes \sigma_m$ ,  $i,j,k,\dots,m \in \{0, x, y, z\}$ ,  $\alpha_{i,j,k,\dots,m}$  为泡利基的实系数.

通过测量量子态在计算基上的投影或哈密顿量的期望值得到的经典数据, 其处理策略也根据量子生成模型的不同而有各式的差异. 如对于量子玻恩机, 量子测量得到的数据被映射为数据集的概率分布, 显然这样对于小型数据集或稀疏数据集有表示优势. 而对于量子生成对抗网络, 测量得到的数据可直接作为生成数据以混淆源数据.

量子生成模型的核心在于利用量子计算机生成量子态, 对量子态进行测量得到经典的有效信息. 如何有效地生成量子态, 是我们面临的首要问题. 参数量子线路 (PQC) 为这一问题提供了一个思路.

参数量子线路一般由若干单量子门和多量子纠缠门组合得到, 单量子门可以是任意单比特么正门  $U_S(\theta)$ , 常用的有泡利旋转门  $R_X(\theta) = e^{-i\frac{\theta}{2}\sigma_x}$ ,  $R_Y(\theta) = e^{-i\frac{\theta}{2}\sigma_y}$ ,  $R_Z(\theta) =$



$e^{-i\frac{\theta}{2}\sigma_z}$ , 多量子门一般采用  $U_{CZ} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \sigma_z$ ,  $U_{\text{CNOT}} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \sigma_x$  等两比特门. 单量子门和多量子门交替的结构极大地提高了态的纠缠能力, 扩展了态的表示空间.

记量子线路的参数为  $\theta$ , 则参数量子线路可以用  $U(\theta)$  表示. 若要保证参数量子线路是有效的, 其参数总量和线路深度都有相应的限制. 一般设定参数的总量是多项式级的, 即  $n \sim O(\text{poly}(N))$ , 其中  $n$  为参数个数,  $N$  为量子比特个数. 同时线路的深度也设为多项式级的, 即  $L \sim O(\text{poly}(N))$ .

如何使参数量子线路有更强的表示能力一直受到关注. 最直观的方式是增加参数量子线路的深度, 构成 MPQCs (multiple parameterized quantum circuits). 然而 MPQCs 的线路深度至多为多项式级的才能满足有效性要求, 也即图 A1 中的 PQC 模块数是常数级的, 不随量子比特数目增长. 而且由于量子噪声的存在, 增加 PQC 的模块数会极大地影响量子态的保真度, 所以在真实的任务中往往会限制 PQC 的模块数. 除此之外还可以通过非计算基补充测量的方式, 来获取更多的数据信息, 这种方法在 QCAAN 线路<sup>[31]</sup>中被提到.

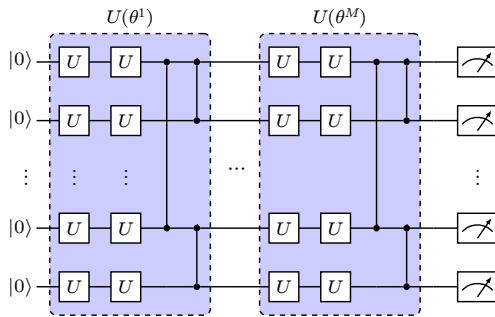


图 A1 参数量子线路示意图  
Fig. A1. Illustration of MPQCs.

对经典神经网络的表示能力的研究是一项非常重要的课题, 参数量子线路的表示能力直接关系到其能否体现量子优越性. Du 等<sup>[22]</sup>证明了在多项式层次结构保持的前提下, 具有  $O(\text{poly}(N))$  个门的参数量子线路, 其表示能力比具有  $O(\text{poly}(N))$  个参数的经典玻尔兹曼机要强. 尽管如此, 参数量子线路和经典神经网络的直接比较还有一些困难, 如何定义对量子体系和经典体系均适用的表示能力指标, 以及参数量子线路能否体现量子的优越性, 仍是开放性的问题.

## 参考文献

- [1] Zhu J Y, Krähenbühl P, Shechtman E, Efros A A 2016 *European Conference on Computer Vision*, Berlin, September 16, 2016 p597
- [2] Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, Kalchbrenner N, Senior A, Kavukcuoglu K 2016 arXiv: 1609.03499 [cs.LG]
- [3] Gómez-Bombarelli R, Wei J N, Duvenaud D, Hernández-Lobato J M, Sánchez-Lengeling B, Sheberla D, Aguilera-Iparraguirre J, Hirzel T D, Adams R P, Aspuru-Guzik A 2018 *ACS Cent. Sci.* **4** 268
- [4] Isola P, Zhu J Y, Zhou T, Efros A A 2016 arXiv: 1611.07004[cs.CV]
- [5] Dallaire-Demers P L, Killoran N 2018 *Phys. Rev. A* **98** 012324
- [6] Lloyd S, Weedbrook C 2018 *Phys. Rev. Lett.* **121** 040502
- [7] Benedetti M, Garcia-Pintos D, Perdomo O, Leyton-Ortega V, Nam Y, Perdomo-Ortiz A 2019 *npj Quantum Inf.* **5** 1
- [8] Liu J G, Wang L 2018 *Phys. Rev. A* **98** 062324
- [9] Amin M H, Andriyash E, Rolfe J, Kulchytskyy B, Melko R 2018 *Phys. Rev. X* **8** 021050
- [10] Khoshaman A, Vinci W, Denis B, Andriyash E, Amin M H 2019 *Quantum Sci. Technol.* **4** 014001
- [11] Benedetti M, Realpe-Gómez J, Biswas R, Perdomo-Ortiz A 2017 *Phys. Rev. X* **7** 041052
- [12] Kieferová M, Wiebe N 2017 *Phys. Rev. A* **96** 062327
- [13] Romero J, Olson J P, Aspuru-Guzik A 2017 *Quantum Sci. Technol.* **2** 045001
- [14] Lamata L, Alvarez-Rodriguez U, Martín-Guerrero J, Sanz M, Solano E 2018 *Quantum Sci. Technol.* **4** 014007
- [15] Li R, Alvarez-Rodriguez U, Lamata L, Solano E 2017 *Quantum Meas. Quantum Metrol.* **4** 1
- [16] Du Y, Liu T, Tao D 2018 arXiv: 1805.11089 [quant-ph]
- [17] Peruzzo A, McClean J, Shadbolt P, Yung M H, Zhou Z Q, Love P J, Aspuru-Guzik A, O'Brien J L 2014 *Nat. Commun.* **5** 4213
- [18] Nielsen M A, Chuang I L 2002 *Quantum computation and quantum information* (Cambridge: Cambridge University Press) pp221–225
- [19] Preskill J 2018 *Quantum* **2** 79
- [20] Harrow A W, Hassidim A, Lloyd S 2009 *Phys. Rev. Lett.* **103** 150502
- [21] Huang H L, Du Y, Gong M, Zhao Y, Wu Y, Wang C, Li S, Liang F, Lin J, Xu Y, Yang R, Liu T, Hsieh M H, Deng H, Rong H, Peng C Z, Lu C Y, Chen Y A, Tao D, Zhu X, Pan J W 2020 arXiv: 2010.06201 [quant-ph]
- [22] Du Y, Hsieh M-H, Liu T, Tao D 2020 *Phys. Rev. Research* **2** 033125
- [23] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, WardeFarley D, Ozair S, Courville A, Bengio Y 2014 *Proceedings of the 27th International Conference on Neural Information Processing Systems* **2** pp2672–2680
- [24] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A C 2017 arXiv: 1704.00028 [cs.LG]
- [25] Zoufal C, Lucchi A, Woerner S 2019 *npj Quantum Inf.* **5** 103
- [26] Zeng J, Wu Y, Liu J G, Wang L, Hu J 2019 *Phys. Rev. A* **99** 052306
- [27] Schuld M, Bergholm V, Gogolin C, Izaac J, Killoran N 2019 *Phys. Rev. A* **99** 032331
- [28] MacKay D J C 2002 *Information Theory, Inference & Learning Algorithms* (Cambridge: Cambridge University)
- [29] Situ H, He Z, Wang Y, Li L, Zheng S 2020 *Information Sciences* **538** 193
- [30] Hu L, Wu S H, Cai W, Ma Y, Mu X, Xu Y, Wang H, Song Y, Deng D L, Zou C L, Sun L 2019 *Sci. Adv.* **5** eaav2761
- [31] Rudolph M S, Toussaint N B, Katarbarwa A, Johri S, Peropadre B, Perdomo-Ortiz A 2020 arXiv: 2012.03924 v2 [quant-ph]
- [32] Cheng S, Chen J, Wang L 2018 *Entropy* **20** 583
- [33] Hinton G E, Sejnowski T J 1986 *In Parallel Distributed Processing: Explorations in the Microstructure of Cognition* **1** 282

- [34] Hinton G E 2012 *In Neural Networks: Tricks of the Trade* Berlin Heidelberg, Germany, 2012 p599
- [35] Coyle B, Mills D, Danos V, Kashefi E 2020 *npj Quantum Inf.* **6** 60
- [36] Hofmann T, Schölkopf B, Smola A J 2008 *Ann. Statist.* **36** 1171
- [37] Hinton G E, Sejnowski T J 1983 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* Washington D. C., USA, 1983 p448
- [38] Hinton G E, Osindero S, Teh Y-W 2006 *Neural Comput.* **18** 1527
- [39] Salakhutdinov R, Hinton G 2009 *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics* Florida, USA 2009 p448
- [40] Smolensky P 1986 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol.1) (Cambridge: MIT press) pp194–281
- [41] Dorband J E 2015 *12th International Conference on Information Technology-New Generations* Las Vegas, USA, April 13–15, 2015 p703
- [42] Buhrman H, Cleve R, Watrous J, Wolf R D 2001 *Phys. Rev. Lett.* **87** 167902
- [43] Ding Y, Lamata L, Sanz M, Chen X, Solano E 2019 *Adv. Quantum Technol.* **2** 1800065
- [44] Pepper A, Tischler N, Pryde G J 2019 *Phys. Rev. Lett.* **122** 060501
- [45] Bondarenko D, Feldmann P 2020 *Phys. Rev. Lett.* **124** 130502
- [46] Huang C J, Ma H, Yin Q, Tang J F, Dong D, Chen C, Xiang G Y, Li C F, Guo G C 2020 *Phys. Rev. A* **102** 032412
- [47] Cao C, Wang X 2021 *Phys. Rev. Applied.* **15** 054012
- [48] Cerezo M, Sone A, Volkoff T, Patrick L C, Coles J 2021 *Nat. Commun.* **12** 1791
- [49] Gao X, Zhang Z Y, Duan L-M 2018 *Sci. Adv.* **4** eaat9004
- [50] Gao X, Anschuetz E R, Wang S T, Cirac J I, Lukin M D 2021 arXiv: 2101.08354 v1 [quant-ph]

## SPECIAL TOPIC—Machine learning and physics

# Quantum generative models for data generation\*

Sun Tai-Ping<sup>1)</sup> Wu Yu-Chun<sup>1)2)†</sup> Guo Guo-Ping<sup>1)2)3)</sup>

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 17 May 2021; revised manuscript received 30 June 2021 )

### Abstract

In recent years, many generation-based machine learning algorithms such as generative adversarial networks, Boltzmann machine, auto-encoder, etc. are widely used in data generation and probability distribution simulation. On the other hand, the combined algorithms of quantum computation and classical machine learning algorithms are proposed in various styles. Especially, there exist many relevant researches about quantum generative models, which are regarded as the branch of quantum machine learning. Quantum generative models are hybrid quantum-classical algorithms, in which parameterized quantum circuits are introduced to obtain the cost function of the task as well as its gradient, and then classical optimization algorithms are used to find the optima. Compared with its classical counterpart, quantum generative models map the data stream to high-dimensional Hilbert space with parameterized quantum circuits. In the mapping space, data features are easier to learn, which can surpass classical generative models in some tasks. Besides, quantum generative models are potential to realize the quantum advantage in noisy intermediate-scale quantum devices.

**Keywords:** quantum machine learning, quantum generative models, quantum advantage

**PACS:** 03.67.–a, 03.67.Ac

**DOI:** 10.7498/aps.70.20210930

\* 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 Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB24030600), and the Anhui Initiative in Quantum Information Technologies, China (Grant No. AHY080000).

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