

# 神经网络量子态及其应用

蒋文杰<sup>1</sup> 邓东灵<sup>1,2,†</sup>

(1 清华大学交叉信息研究院 北京 100084)

(2 上海期智研究院 上海 200232)

2021-01-26 收到

† email: dldeng@mail.tsinghua.edu.cn

DOI: 10.7693/wl20210202

## Neural network quantum states and their applications

JIANG Wen-Jie<sup>1</sup> DENG Dong-Ling<sup>1,2,†</sup>

(1 Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China)

(2 Shanghai Qi Zhi Institute, Shanghai 200232, China)

**摘要** 神经网络量子态是由人工神经网络所表示的量子态。得益于机器学习,尤其是深度学习近年来取得的突破性进展,神经网络量子态的研究得到了广泛的关注,成为当前的热点前沿方向。文章将介绍不同的神经网络量子态,其物理性质与典型应用场景,最新进展,以及面临的挑战。

**关键词** 人工神经网络, 量子多体问题, 量子纠缠, 贝尔不等式

**Abstract** Neural-network quantum states are states represented by artificial neural networks. Thanks to the dramatic progress achieved recently in the field of machine learning, especially deep learning, the study of neural-network quantum states has attracted tremendous attention across communities, and has become one of the most active directions of research. In this paper, we review different kinds of neural-network quantum states, their physical properties, and typical applications. In addition, we also discuss some most recent advances and future challenges along this direction.

**Keywords** artificial neural networks, quantum many-body problems, quantum entanglement, Bell inequality

## 1 引言

人工智能主要有三条发展路线:符号主义、连接主义与行为主义<sup>[1]</sup>。人工神经网络是连接主义的基石,也是最近几年深度学习取得突破进展的关键要素之一。它是受到生物大脑中信息处理模式的启发而提出的,最早可追溯到1943年由心理学家W. S. McCulloch与数理逻辑学家W. Pitts提出的神经元模型<sup>[2]</sup>。当前,基于神经网络的人工智能技术正在给人类文明的方方面面带来革命性的改变<sup>[3]</sup>:从语音、图像识别到引力波、黑洞探

测,再到数据挖掘、自动驾驶、医学诊断、证券市场分析,等等。2018年,计算机科学的最高奖——图灵奖也被授予三位人工智能科学家Yoshua Bengio, Geoffrey Hinton与Yann LeCun,以表彰他们在相关领域所做的突出贡献<sup>[4]</sup>。

另一方面,量子力学是现代物理最重要的基础理论之一<sup>[5]</sup>。其重要性广泛体现在我们的日常生活和科学探索中:从以电子计算机为代表的半导体工业到新奇的超导现象,从随处可见的化学电池到宇宙中神秘的黑洞,世间万物的变化规律都与量子力学密切相关。

然而,对量子系统尤其是量子多体系统的研究是非常困难的。实际研究中能够严格解析解决的问题很少,对于绝大部分问题的求解,我们只能依赖于数值方法。对于最一般的情形,数值方法需要消耗指数量级的计算资源,这对于规模较小的物理系统是可行的,但如果系统规模变大,这一指数级的要求在经典计算体系下就难以满足了<sup>[6]</sup>。1998年诺贝尔化学奖得主Walter Kohn将这一问题描述为“指数墙(exponential wall)”困难<sup>[7]</sup>。为此,物理学家做了大量的努力,发展了一系列计算方法,著名的蒙特卡罗算法以及重正化群算法就是典型代表。但是这些方法并不是通用的,分别有着各自的适用条件。比如蒙特卡罗算法在应用于一些有阻挫系统时会出现符号问题,从而使得算法需要指数级的时间;而密度矩阵重正化群算法一般仅适用于一维低纠缠系统。

在人工智能领域中,一个类似的问题是维度灾难(curse of dimensionality)。维度灾难最早是由动态规划先驱、著名应用数学家Richard E. Bellman提出,描述了高维与低维数据集截然不同的性质对计算问题带来的影响<sup>[8]</sup>:随着数据维度增加,有限规模的数据在空间中的分布会逐渐稀疏,从而失去统计意义。这就要求在一般情况下,我们需要非常大的数据规模来获得数据集的统计特征,但是这会对计算资源带来严重负担。经过多年的发展,人工智能领域提出了许多用于处理高维问题的方法和工具。人工神经网络就是一个应用非常广泛的例子,可以在一定程度上缓解维度灾难带来的困难。简单来说,人工神经网络可以看成是一个普适的函数拟合器。通过调节网络参数,它可以用来拟合任何光滑函数<sup>[9]</sup>。

由于指数墙困难和维度灾难的相似性,一个自然的想法是可以神经网络处理复杂的量子问题。如可以用神经网络识别不同的量子物态以及研究它们之间的相变(参见《物理》2017年第9期蔡子的专题文章)。另一方面,我们也可以用神经网络来表示量子态,其主要思想是把神经网络当成变分波函数,通过调节网络参数来逼近目标波函数(如多体系统的基态),进而求解所关心的物理问题。传统的量子多体变分波函数方法需要物

理学家针对所求解的问题设计特定的变分函数,而神经网络量子态方法可以使用相对普适的结构,对于先验知识的依赖程度较低。此外,人工智能领域里发展的一些优化神经网络的方法也可以用于神经网络量子态,提高算法效率。

近年来,通过神经网络量子态的方法求解量子多体问题受到了广泛关注<sup>[10-12]</sup>。当前,这是一个非常活跃的前沿研究方向。本文将介绍不同神经网络量子态的物理性质与典型应用场景,以及此方向的最新进展。所涉及的神经网络包括受限玻尔兹曼机,深度玻尔兹曼机,前馈神经网络,与循环神经网络等。典型应用包括:求解量子多体系统的基态及动力学演化,探测量子非定域性,量子层析,以及计算交错时序关联函数等。希望通过本文的讨论,读者能感受到神经网络量子态的魅力。众所周知,基于神经网络的智能程序AlphaGo<sup>[13]</sup>与AlphaFold<sup>[14]</sup>分别在围棋与预测蛋白质结构方面取得了里程碑式的突破。我们期望神经网络量子态能把这些突破延续到解决复杂的量子多体问题中来。

## 2 量子态的神经网络表示

在量子力学中,一个封闭的、不与外界产生关联的物理系统的全部可能状态组成一个希尔伯特空间,每个特定的物理状态由该空间中的一个矢量描述。希尔伯特空间在数学上是线性空间,因此在确定其基矢之后,每一个物理状态对应的矢量可以表示为选定基矢量的线性叠加。在实际物理问题中,我们经常需要处理包含多个子系统的情况,系统的希尔伯特空间维数为各子系统对应空间维数的乘积<sup>[15]</sup>。比如,假设我们需要描述包含 $N$ 个自旋粒子的量子系统,每一个粒子自旋可以取上下两种可能,其对应希尔伯特空间维数为2,那么整个系统的自旋状态就有 $2^N$ 种可能,从而总希尔伯特空间维数为 $2^N$ 。因此,表示最一般情况下的波函数需要指数量级的计算资源。这给数值求解量子多体问题带来了极大挑战。

幸运的是,人们关心的物理状态一般还受到某些限制,比如对称性的限制或者是某些物理观

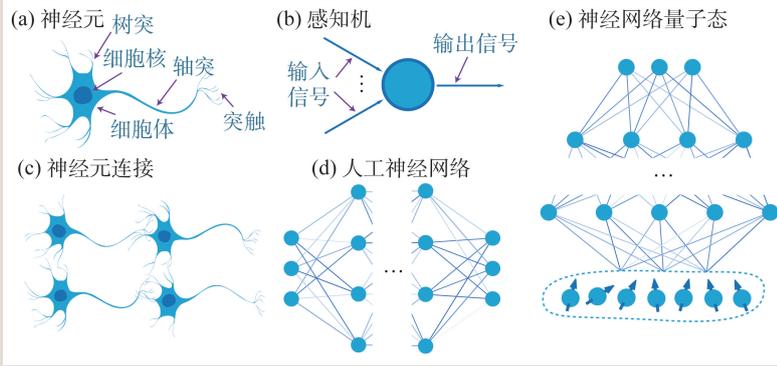


图1 神经网络量子态示意图 (a) 生物大脑中的神经元; (b) 感知机; (c) 生物神经网络; (d) 人工神经网络; (e) 量子态的神经网络表示

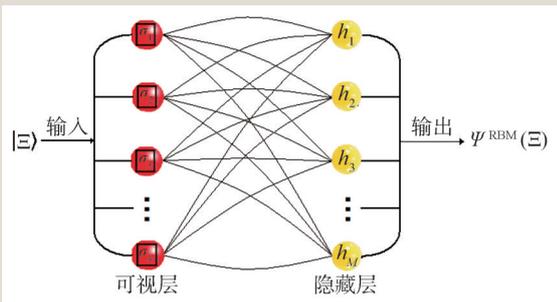


图2 量子态的受限玻尔兹曼机表示

测量的限制, 每一个子系统并不是完全独立的, 子系统状态会互相影响, 从而整体系统可能的状态只占据了希尔伯特空间中的很小一部分。原则上可以针对不同的物理系统, 利用具有特定结构的表示方法, 在使用相对较少的计算资源情况下表示这些物理状态<sup>[5]</sup>。著名的张量网络就是一个典型的例子<sup>[16]</sup>。物理中一般使用纠缠熵(entanglement entropy)来刻画量子系统之间的关联强度。张量网络可以有效表示纠缠熵满足面积定律(即纠缠熵与子系统的表面积成正比)的物理状态<sup>[17]</sup>。在这里, “有效”指的是只需要多项式量级的计算资源。另一个例子就是本文将要重点介绍的神经网络量子态。

神经网络由大量的节点(神经元)及它们之间的相互连接构成, 如图1所示。每个节点包含一种特定的输出函数, 称为激活函数。每两个节点间的连接代表对于通过该连接信号的加权值, 称为权重。神经网络就是通过这种方式来简单模拟人类的大脑。网络的输出则取决于网络的结构、

连接方式、权重和激活函数。神经网络中的神经元通常被排列成层状结构, 第一层被称为输入层, 数据由这一层输入。最后一层被称为输出层, 中间层被称为隐藏层。如果一个神经网络有多于两层的结构, 我们通常称其为深度神经网络, 基于此构建的机器学习模型称为深度学习。依据具体网络结构和信息传播方向的不同, 神经网络又可以分为很多种。常见的神经网络有前馈神经网络、卷积神经网络、玻尔兹曼机、循环神经网络等。本质上, 量子波函数是一个函数, 而神经网络是一个普适的函数拟合器。因此, 我们可以用神经网络表示量子态。

## 2.1 受限玻尔兹曼机

受限玻尔兹曼机(restricted Boltzmann machine, RBM)是一类应用非常广泛的神经网络, 其在数据降维、特征学习、图片生成、自然语言处理等场景中都有重要应用<sup>[18]</sup>。它是一个两层的神经网络, 其中一层称为可视层, 另一层称为隐藏层。可视层中的神经元可以连接到隐藏层, 而同层神经元之间不能相连。

考虑一个由  $N$  个量子比特组成的系统, 其量子态的一般形式为  $|\Phi\rangle = \sum_{\sigma} \psi(\sigma) |\sigma\rangle$ , 其中  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$  表示一个可能的构型。  $\psi(\sigma)$  可以看成是一个函数, 输入为  $\sigma$ , 输出为一个复数  $\psi(\sigma)$ , 表示分量  $|\sigma\rangle$  对应的振幅和相位信息。如图2所示, 可以用一个可视层有  $N$  个神经元(对应  $N$  个量子比特), 隐藏层有  $M$  个神经元的RBM来表示  $\psi(\sigma)$ <sup>[19]</sup>

$$\psi(\sigma) = \sum_{\mathbf{h}} e^{\sum_i a_i \sigma_i + \sum_j b_j h_j + \sum_{ij} w_{ij} \sigma_i h_j},$$

其中  $\mathbf{h} = (h_1, h_2, \dots, h_N)$  表示隐藏神经元的可能构型, 每个神经元有两个可能的取值  $\sigma_i = \pm 1$  和  $h_j = \pm 1$ ,  $a_i$ ,  $b_j$  与  $w_{ij}$  分别表示网络的偏置和连接

参数。为简单起见，我们称由受限玻尔兹曼机所表示的量子态为RBM态。

数学上可以证明，当 $M$ 取值足够大时，受限玻尔兹曼机可以以任意精度逼近任何光滑函数。因此，原则上量子态的受限玻尔兹曼机表示是完备的，任何量子态都可以用受限玻尔兹曼机表示。实际应用中， $M$ 一般随 $N$ 多项式增大，所以RBM表示量子态所需要的参数个数也是随 $N$ 多项式增加，而不是指数增加。如此，RBM态在解决某些量子多体问题的时候就可能可以绕过“指数墙”困难。

与张量网络表示不同，受限玻尔兹曼机可以有效表示具有大纠缠熵的量子态<sup>[20]</sup>。这得益于可视神经元与隐藏神经元的长程连接。事实上，我们可以解析构造一个满足纠缠熵体积定律(即纠缠熵与子系统体积成正比)的RBM态，其所包含的参数个数随 $N$ 仅为线性增加。而如果用常规的张量网络表示同样的量子态，所需的参数个数随 $N$ 是指数增加的。这体现了神经网络在表示大纠缠熵的量子态方面的独特优势。

如果我们限制只有近邻的可视神经元能连接到同一个隐藏神经元，这样可以进一步降低参数规模及优化难度，所得到的量子态称为短程RBM态。由于这个限制，任意可视神经元只关联到其附近的神经元。因此，所有短程RBM态都满足纠缠熵面积定律。

短程RBM可以严格表示一些有趣的奇异量子态，如拓扑态、超图态等。拓扑态的一个重要例子是环曲面码态(toric-code state)<sup>[21]</sup>，它是Alexei Kitaev提出的环曲面码哈密顿量(图3(a))的基态，在拓扑量子计算<sup>[22]</sup>与量子纠错中极为重要。图3(b)简要描述了环曲面码态的短程RBM表示。在环曲面码态的基础上，作用弦算符(即由不同格点泡利矩阵张量积所得算符)可以得到系统的激发态。如图3(c)所示，此激

发态含有4个准粒子，分别位于弦算符 $P_1^x$ 与 $P_1^z$ 两端。有意思的是，这些激发态都可以用短程RBM严格表示，且所需要的参数个数只随系统规模线性增加<sup>[23]</sup>。基于此，我们有一个直接推论，由短程RBM表示的环曲面码态及其激发态都满足纠缠熵面积定律。而传统方法证明此结论需要涉及比较复杂的数学工具(如商群)。

以上，我们讨论了量子系统纯态的RBM表示。实际中的量子系统不可避免地受到环境的干扰，其状态是一个混合态，需要用密度矩阵算符来描述。受限玻尔兹曼机也可以用来表示混合态<sup>[24]</sup>。需要指出的是，为满足密度矩阵半正定性的要求，表示混合态时受限玻尔兹曼机的参数需要满足特定的条件。此外，通过附加行列式或使用Grassmann代数的方法，受限玻尔兹曼机也可以用来表示费米子系统的量子态<sup>[25, 26]</sup>。

## 2.2 深度玻尔兹曼机

受限玻尔兹曼机可以有效表示一些有趣的量子态，但其表示能力有限。例如，它不能有效表示一些可以展示量子优势(quantum supremacy)的态，如二维团簇态通过特殊幺正变化所得的态<sup>[27]</sup>。这个结论可以从直观上理解，由于受限玻尔兹曼机简单的结构，其所表示的量子态可以通过有效算法求得。假设其能有效表示可以展示量子优势的态，这就意味着经典计算机可以有效模拟这个量子态，这与量子态可以展示量子优势是矛盾的。

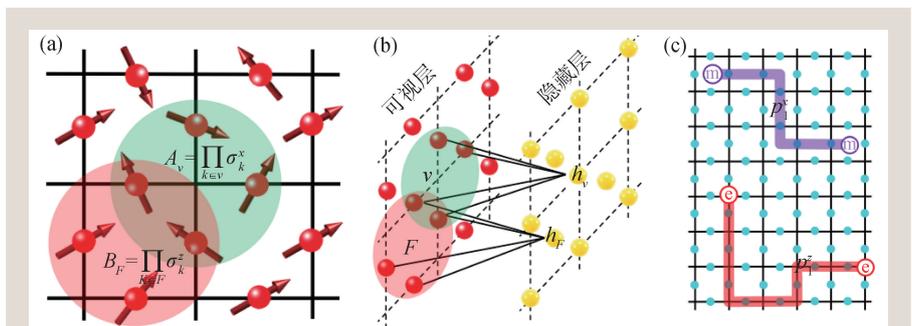


图3 受限玻尔兹曼机表示拓扑态 (a)环曲面码哈密顿量；(b)基态的RBM表示；(c)含4个准粒子的激发态

为加强受限玻尔兹曼机的表达能力，可以在原有网络上再加一层隐藏层，所得网络称为深度玻尔兹曼机(deep Boltzmann machine, DBM)。在计算复杂度理论中，一个被普遍接受但至今无法证明的猜想是复杂度的多项式层级不会塌缩，著名的  $P \neq NP$  猜想是这个猜想的一个特例。在假设以上猜想成立的情况下，可以证明：DBM 相比 RBM 在表达能力上可以有指数级的优势。存在一些量子态，如果用 RBM 表示需要指数级的参数，而 DBM 只需要多项式规模的参数<sup>[28]</sup>。

## 2.3 前馈神经网络

前馈神经网络是研究最早和最简单的神经网络之一，也是目前应用最广泛，发展最迅速的人工神经网络之一<sup>[18]</sup>。其神经元分层排列，每个神经元只与前一层神经元相连。信息从输入层逐层传递到输出层，单向传播无反馈。与受限玻尔兹曼机一样，前馈神经网络也可以用来表示量子态<sup>[29]</sup>。其输入层所含神经元数目对应所考虑量子系统的粒子数，输出层为单个神经元，输出一个复数，表示量子态对应分量的振幅和相位信息。

对于非常复杂的量子态，我们可以将波函数分作两个部分：波函数绝对值与对应的符号，并分别使用两个前馈神经网络表示。在实际应用中可以观察到，对于简单的量子态，前馈神经网络

可以准确学习到其对应的符号规则；对于一些复杂的量子态前馈神经网络也能学习到比较高的精度，这证实了利用前馈神经网络处理量子态的有效性<sup>[29]</sup>。

## 2.4 其他神经网络

人工智能领域针对不同的问题设计了多种多样的神经网络，原则上所有类型的神经网络都可以用于表示量子态。不同的网络有不同的结构，能有效表示的量子态以及网络训练的时间复杂度也不尽相同。在实际应用中，我们可以根据具体问题选择不同的神经网络<sup>[18]</sup>。比如循环神经网络(recurrent neural network, RNN)非常适合处理序列数据，在机器翻译、语音识别以及文本生成等领域有着广泛的应用。可以将多体系统中量子比特构型视为序列数据，从而利用循环神经网络来表示量子多体态<sup>[30]</sup>。卷积神经网络(convolutional neural network, CNN)则是另一类被广泛使用的深度神经网络，适合图像处理、行为认知、迁移学习等场景。文献[31]表明，卷积神经网络也可以用来表示量子态，如前面提到的环曲面码态。

## 3 神经网络量子态的应用

如前所述，人工神经网络可以非常有效地表示多体量子态，其在量子物理，尤其是解量子多体问题中有很广泛的应用。图4归纳了当前神经网络量子态的主要应用。接下来，我们简要介绍部分近期的相关工作，主要侧重于RBM量子态的应用。

### 3.1 求解量子系统基态和动力学演化

一个孤立封闭的量子系统可以由哈密顿量描述，其演化过程满足薛定谔方程。求解给定哈密顿量的基态和动力学演化是量子物理中常见的的基本问题。对于少数特殊的模型，如一维伊辛模型(Ising model)，其基态和动力学可以通过解析的方法严格求解。然而，实际研究中能



图4 神经网络量子态的应用

够解析求解基态和动力学的情况很少，我们需要依赖数值方法。

利用神经网络求解基态和动力学的核心想法是把神经网络量子态看成变分函数，通过梯度下降算法优化网络参数求解相应问题。以受限玻尔兹曼机为例，G. Carleo 和 M. Troyer 首先求解了几个典型的量子磁性模型(如伊辛模型、海森伯模型)的基态及动力学，并与传统的密度矩阵重正化群等方法进行了比较<sup>[19]</sup>。结果表明，神经网络的方法使用较少的参数就得到了相近精度的基态能量和动力学演化，这在一定程度上展示了神经网络方法的优越性。

值得指出的是，对于最一般情形求解基态和动力学演化可以证明是 NP 问题。因此，神经网络的方法也不能有效求解所有量子系统的基态和演化。当前的研究表明，其在解决涉及大量子纠缠与高维系统的问题中相比传统方法可能有优势，但是这一优势还没有得到确切的证明。如何判断给定哈密顿量的基态和动力学是否可以通过神经网络的方法有效求解是此领域里一个亟待解决的重要问题。这一问题的解决可能需要发展新的物理概念和数学工具。

### 3.2 交错时序关联函数

交错时序关联函数(out-of-time ordered correlator, OTOC)最早由 A. Larkin 与 Y. Ovchinnikov 在 1969 年研究超导理论时提出<sup>[32]</sup>。经过几十年的发展，OTOC 在表征量子混沌，量子信息置乱(information scrambling)，动力学相变等研究中都有重要应用。此外，其还可以给通过 Ads/CFT 对偶研究量子引力与黑洞带来新的启示。最近，实验测量 OTOC 也在离子阱、固态自旋、玻色-爱因斯坦凝聚等系统中实现。

考虑量子多体系统中在空间上分开的两个局域算符  $W$  与  $V$ ，其对应的 OTOC 定义为

$$F(t) = \langle W(t)^\dagger V^\dagger W(t) V \rangle,$$

其中  $W(t) = e^{iHt} W e^{-iHt}$  为  $W$  在海森伯绘景中的时间演化算符。不难看出，OTOC 的物理意义是描述一

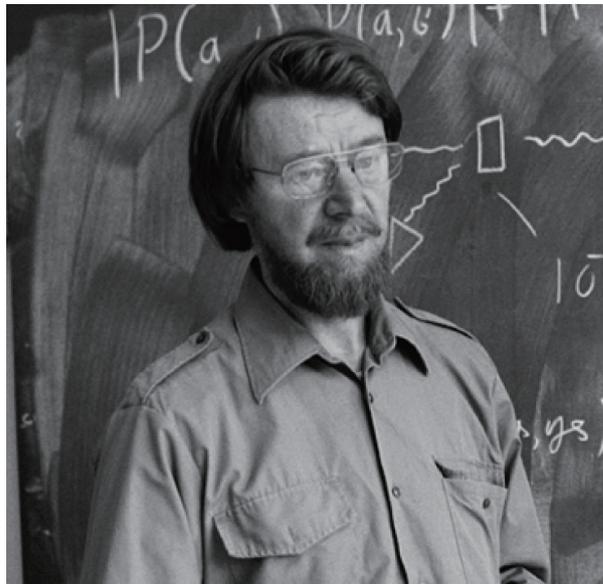


图5 约翰·斯图尔特·贝尔(John Stewart Bell, 1928.6.28—1990.10.1)。图片来源于网络

个局域扰动传播一段时间后，在另一个地方被探测到的物理现象。数值上，计算多体系统的 OTOC 是非常困难的，其复杂程度要大于基态或动力学演化的求解。文献[33]提出了 OTOC 的神经网络求解办法，其核心思想是把 OTOC 看成是两个随时间演化量子态的交叠，从而通过计算态的演化与交叠来求解。

### 3.3 量子非定域性

非定域性是量子系统一个非常奇特的性质，是量子物理与经典物理最核心的区别之一<sup>[34]</sup>。它描述比量子纠缠更强的关联——任何表现非定域性的量子态一定是纠缠的，反之则不一定成立。在实际应用中，量子非定域性是构建设备无关量子技术，如无条件安全量子密钥分配、自认证随机数产生器等，不可或缺的资源。对量子非定域性的思考和研究最早可以追溯到 20 世纪初发生在爱因斯坦和玻尔之间关于“上帝是否会掷骰子”著名争论<sup>[35]</sup>。1964 年，约翰·斯图尔特·贝尔(John Stewart Bell, 图 5)提出了著名的贝尔不等式<sup>[36]</sup>。从此，量子非定域性可以通过实验测试贝尔不等式的破坏来定量刻画。

然而，由于存在指数墙困难，在量子多体系

统中研究非定域性变得极为不易。在文献[37]中,本文作者之一把机器学习的方法引入到了量子多体非定域性的研究中。其核心思想是把量子多体系统中探测非定域性的问题转化为求解哈密顿量基态能量问题,从而可以利用上述神经网络量子态的方法处理。具体来说,对于给定的量子多体系统,其所有可能的经典关联组成一个高维空间的多面体,多面体的每一个面对应一个贝尔不等式,如图6所示。初始时,我们随机产生一个RBM量子态,其对于给定的观测量一般只表现出经典关联。通过不断地优化RBM的参数,其所表示的量子态逐渐超越多面体的一个面(即破坏对应的贝尔不等式),展示出量子非定域性。值得指出的是,神经网络量子态在探测多体非定域性问题中有独特优势,可以解决一些用传统方法无法解决或极为困难的问题,比如计算随机全关联系统贝尔不等式最大破坏值。

### 3.4 求解开放系统稳态和动力学

孤立量子系统演化过程遵循薛定谔方程,而实际系统往往不可避免地与环境发生作用,因此很多情况下并不能当作孤立系统处理。对于与环境弱耦合的开放系统,其状态的演化可以近似认为仅与当前状态相关,而与之之前的演化过程无关,从而通过玻恩—马尔可夫近似,可以导出开放系统所满足的演化方程,即Lindblad主方程<sup>[38]</sup>:

$$\dot{\rho} = \mathcal{L}\rho = -i[H, \rho] + \sum_j \frac{\gamma_j}{2} (2c_j \rho c_j^\dagger - c_j^\dagger c_j \rho - \rho c_j^\dagger c_j),$$

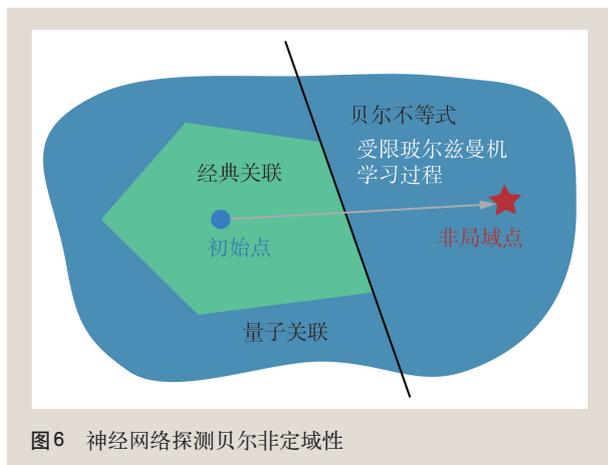


图6 神经网络探测贝尔非定域性

其中 $\mathcal{L}$ 为刘维尔超算符, $H$ 表示系统哈密顿量, $\rho$ 为密度矩阵, $c_j$ 与 $\gamma_j$ 分别表示耗散算符与耗散强度。

与孤立系统类似,神经网络的方法也可以用来求解开放量子系统的稳态与动力学演化,此时需要使用密度矩阵的神经网络表示<sup>[24]</sup>。与孤立系统不同的是,开放量子系统的能量不再是守恒量,因此不能通过对能量的变分来求解。但是,我们可以考虑优化变分近似演化与精确演化之间的距离或者通过Choi—Jamiołkowski同构把主方程转化为有效哈密顿量方程来求解。与之相关地,我们在文献[39]中进一步把神经网络的方法推广到了刘维尔能隙的求解中。

### 3.5 量子态层析

量子态层析(quantum tomography)是通过很多份相同量子态的测量来估计一个未知的量子态<sup>[15]</sup>。它是校准量子系统,检验量子操作的重要技术。同样的,由于希尔伯特空间维度随系统规模指数增大,多体系统的量子态层析也变得极为困难。以谷歌公司2019年实现量子优势的实验为例<sup>[40]</sup>,其量子线路涉及53个量子比特,对如此大规模的量子态层析最直接的方法需要确定 $2^{53} \approx 10^{16}$ 个参数,即使是存储这些参数也至少需要 $10^5$  TB的存储空间,远大于当前世界上最先进的超级计算机的内存空间。

神经网络可以有效表示部分量子态,其所需参数个数只随系统规模多项式增加。因此,通过神经网络的方法进行量子态层析只需要确定多项式规模的参数,可极大地减少所需资源。事实上,神经网络量子态层析已经被多篇论文提出<sup>[30, 41]</sup>,并得到了较多关注。最近,部分相关的理论方案也在实验上得到了验证<sup>[42]</sup>。

## 4 展望

神经网络量子态是最近几年蓬勃发展的交叉前沿方向。当前,这个方向的研究已经取得了一些令人兴奋的成果。然而,总体来说其发展还处

于初级阶段，很多重要基本问题亟待解决。首先，神经网络表达量子态为何有效及其局限性并未被完全理解。给定一个量子态，我们无法有效判断它是否可以用某个神经网络有效表达。这与矩阵直积态或更广的张量网络态在早期的发展相似。由于量子信息领域的飞速发展，现在我们知道量子纠缠是张量网络有效表示量子态的关键，也是判断具体问题能否用相关算法有效解决的前提。然而，量子纠缠并不是神经网络表达量子态的核心要素，理解神经网络有效性及局限性可能需要发展新的物理概念与数学工具。其次，神经网络量子态的“杀手级应用”还未发现。当前已有的研究大多数还处在原理演示阶段。人们发现了很多问题可以用神经网络解决，但这些

问题大多数也可以用传统方法处理，神经网络并不是唯一解决途径。再次，如何利用神经网络方法有效解决强阻挫量子磁性系统或强相互作用费米子系统相关问题仍是未解之谜。这些问题是量子多体物理中极为重要，却最具挑战性的问题。

总之，神经网络量子态在机器学习与量子物理之间架设了新的桥梁。这个方向的研究有利于不同学科之间的交叉融通，对机器学习与量子物理的发展都大有裨益。当前，这个方向正在蓬勃发展，挑战与机遇并存。鉴于AlphaGo与AlphaFold的成功，我们有理由期待未来神经网络量子态在解决复杂的量子多体问题中取得重要突破，大放异彩。

## 参考文献

- [1] Stuart R, Peter N. *Artificial Intelligence: A Modern Approach*. Prentice Hall Press, 2009
- [2] McCulloch W S, Pitts W. *Bull. Math. Biol.*, 1943, 5: 115
- [3] Jordan M T, Mitchell T M. *Science*, 2015, 349: 255; LeCun Y, Bengio Y, Hinton G. *Nature*, 2015. 521: 436.
- [4] <https://awards.acm.org/about/2018-turing>
- [5] Griffiths D, Schroeter D. *Introduction to Quantum Mechanics* (3rd ed.). Cambridge University Press, 2018
- [6] Coleman P. *Introduction to Many-Body Physics*. Cambridge University Press, 2015
- [7] Kohn W. *Rev. Mod. Phys.*, 1999, 71: 1253
- [8] Marimont R B, Shapiro M B. *IMA J. Appl. Math.*, 1979, 24: 59
- [9] Graupe D. *Principles of artificial neural networks*. World Scientific, 2013
- [10] Das Sarma S, Deng D L, Duan L M. *Phys. Today*, 2019, 72: 48
- [11] Melko R G, Carleo G, Carrasquilla J *et al.* *Nat. Phys.*, 2019, 15: 887
- [12] Carleo G *et al.* *Rev. Mod. Phys.*, 2019, 91: 045002
- [13] Silver D, Huang A, Maddison C *et al.* *Nature*, 2016, 529: 484
- [14] Senior A W, Evans R, Jumper J *et al.* *Nature*, 2020, 577: 706
- [15] Nielsen M A, Chuang I L. *Quantum and Quantum Information* (10th. ed.). Cambridge University Press, 2010
- [16] Orus R. *Ann. Phys. (N. Y.)*, 2014, 349: 117
- [17] Eisert J, Cramer M, Plenio M B. *Rev. Mod. Phys.*, 2010, 82: 277
- [18] Goodfellow I, Bengio Y, Courville A. *Deep Learning*. The MIT Press, 2016
- [19] Carleo G, Troyer M. *Science*. 2017, 355: 602
- [20] Deng D L, Li X, Das Sarma S. *Phys. Rev. X*, 2017, 7: 021021
- [21] Kitaev A Y. *Ann. Phys.*, 2003, 303: 2
- [22] Nayak C *et al.* *Rev. Mod. Phys.*, 2008, 80: 1083
- [23] Deng D L, Li X, Das Sarma S. *Phys. Rev. B*, 2017, 96: 195145
- [24] Alexandra N, Vincenzo S. *Phys. Rev. Lett.*, 2019, 122: 250501; Hartmann M J, Carleo G. *Phys. Rev. Lett.*, 2019, 122: 250502; Filippo V, Alberto B, Nicolas R *et al.* *Phys. Rev. Lett.*, 2019, 122: 250503
- [25] Huang Y, Moore J E. 2017, arXiv: 1701.06246
- [26] Glasser I *et al.* *Phys. Rev. X*, 2018, 8: 011006
- [27] Gao X, Wang S T, Duan L M. *Phys. Rev. Lett.*, 2017, 118: 040502
- [28] Gao X, Duan L M. *Nat Commun.*, 2017, 8: 662
- [29] Cai Z, Liu J. *Phys. Rev. B*, 2018, 97: 035116
- [30] Carrasquilla J, Torlai G, Melko R G *et al.* *Nat. Mach. Intell.*, 2019, 1: 155
- [31] Carrasquilla J, Melko R. *Nature Phys.*, 2017, 13: 431
- [32] Larkin A, Ovchinnikov Y N. *Sov. Phys. JETP*, 1969, 28: 1200
- [33] Wu Y, Duan L M, Deng D L. *Phys. Rev. B*, 2020, 101: 214308
- [34] Brunner N *et al.* *Rev. Mod. Phys.*, 2014, 86: 419
- [35] Einstein A, Podolsky B, Rosen N. *Phys. Rev.*, 1935, 47: 777
- [36] Bell J S. *Physics Physique Fizika*, 1964, 1: 195
- [37] Deng D L. *Phys. Rev. Lett.*, 2018, 120: 240402
- [38] Breuer H, Petruccione F. *The Theory of Open Quantum Systems*. Oxford University Press, 2007
- [39] Yuan D, Wang H, Wang Z *et al.* 2020, arXiv: 2009.00019
- [40] Arute F, Arya K, Babbush R *et al.* *Nature*. 2019, 574: 505
- [41] Torlai G *et al.* *Nature Phys.*, 2018, 14: 447
- [42] Palmieri A M, Kovlakov E, Bianchi F *et al.* *npj Quantum Inf.*, 2020, 6: 20