通过扩散映射对拓扑神经网络量子态进行分类

Yanting Teng, [1] Subir Sachdev, [1] 和 Mathias S. Scheurer[2]

[1] *Department of Physics, Harvard University, Cambridge MA 02138, USA*

[2] *Institut für Theoretische Physik, Universität Innsbruck, A-6020 Innsbruck, Austria*

我们讨论并展示了一种无监督机器学习程序，用于检测量子多体系统中的拓扑序。我们使用受限玻尔兹曼机为低能谱定义一个变分试探态，并以概率随其变分能量呈指数衰减的方式对波函数进行采样；由此得到训练数据集，并将其作为扩散映射方案的输入。扩散映射为波函数提供了低维嵌入，揭示了超选择扇区是否存在，进而揭示拓扑序是否存在。我们表明，对于扩散映射，量子态的相似性度量可以用网络参数来定义，从而能够在多项式时间内高效计算。不过，必须谨慎处理可能存在的“规范冗余”。作为一个具体示例，我们将该方法应用于环面码。

arXiv:2301.02683v1 [quant-ph] 6 Jan 2023

I. 引言

近几年来，机器学习（ML）技术作为多体物理学中的新型工具，受到了广泛研究 [1–7]。机器学习已有各种重要应用，例如基于机器学习的多体波函数变分Ansatz、将机器学习应用于实验数据以提取关于底层物理学的信息、用于提高蒙特卡罗采样效率的机器学习方法，以及利用机器学习检测相变等。关于最后一类应用，近期的一个重点是拓扑相变¨– [8。这源于捕捉拓扑相变所面临的挑战：从定义上说，拓扑特征与数据集的全局连通性有关，而非样本的局部相似性。因此，除非数据集足够简单，使得拓扑上相连的样本对同时也具有局部相似性，或者所用的输入数据与底层拓扑不变量密切相关，否则许多标准机器学习技术都难以可靠地捕捉拓扑结构31]。

在这方面，文献 12提出的机器学习方法基于扩散映射（DM） [32–35], ，是一种特别有前景的途径，可用于学习拓扑相变；它能将高维数据嵌入低维子空间，使数据集中平滑连接的样本对映射到彼此接近的位置，而不相连的样本对则映射到相距较远的位置。因此，该方法抓住了拓扑学的核心概念。再加上它属于无监督方法，无需具备关于底层拓扑不变量的先验知识，因此特别适合用于拓扑相分类。因此，近期已有许多研究将这一方法应用于各种问题，例如不同的对称性保护拓扑系统（包括非厄米拓扑系统） [36–41], 实验数据 [39, 42], 多体局域化态 [43],和动力学 [44]；此外，也有研究探讨了基于扩散映射相结合的扩展方法结合路径搜索 [36] 以及量子计算方案 [45] 以实现加速的方法也已得到研究。

如上所述，物理学中机器学习的另一个备受关注的应用是神经网络量子态：正如参考文献46所提出的，神经网络可用于高效地参数化，并在许多情况下优化量子多体系统波函数的变分描述 [47–56]。具体而言，受限玻尔兹曼机（RBM） [4] 是这一领域中一种非常流行的神经网络结构。例如，环面码模型的基态 [57] 可以通过一个局域受限玻尔兹曼机试探态 [58], 精确表示，也就是说，只有相邻的自旋连接到同一个隐藏神经元。当加入参考文献58 中受限玻尔兹曼机试探态的额外非局域扩展后，研究表明，这也能为存在磁场时的环面码提供非常精确的变分描述 [59]。

在本研究中，我们将参考文献12中的DM方法与神经网络量子态相结合，旨在以无监督的方式捕捉相互作用量子多体系统中的拓扑序。我们采用参数为 Λ 的局域网络试探态，作为哈密顿量为 $\mathcal{H}$ 的系统低能子空间波函数 $|\Psi(\Lambda)\rangle$ 的变分描述。虽然我们也简要提及了生成态系综的其他可能方式，但主要聚焦于能量原理：我们对波函数进行采样，使得 $|\Psi(\lambda)\rangle$ 的概率正比于 $\exp(-\langle\mathcal{H}\rangle_\Lambda/T)$，其中 $\langle\mathcal{H}\rangle_\Lambda = \langle\Psi(\Lambda)|\,\mathcal{H}\,|\Psi(\Lambda)\rangle$。如图1(a)所示， $\mathcal{H}$ 低能谱中存在超选择扇区，这意味着当 T 足够小时，态系综会分解为彼此断开的状态子集（至少在系统大小固定且有限时）；无需预先提供标签，即可通过DM降维（随后进行 k 均值聚类）提取这些子集，从而识别拓扑序。当 T 足够大时，会纳入越来越多的高能态，所有扇区都相互连通，见图1(b)；通过对这些状态进行基于DM的嵌入，也可轻易揭示这一点。

重要的是，扩散映射是一种核方法，因为输入数据 x_l（在我们的情形下，状态 $|\Psi(\Lambda_l)\rangle$）

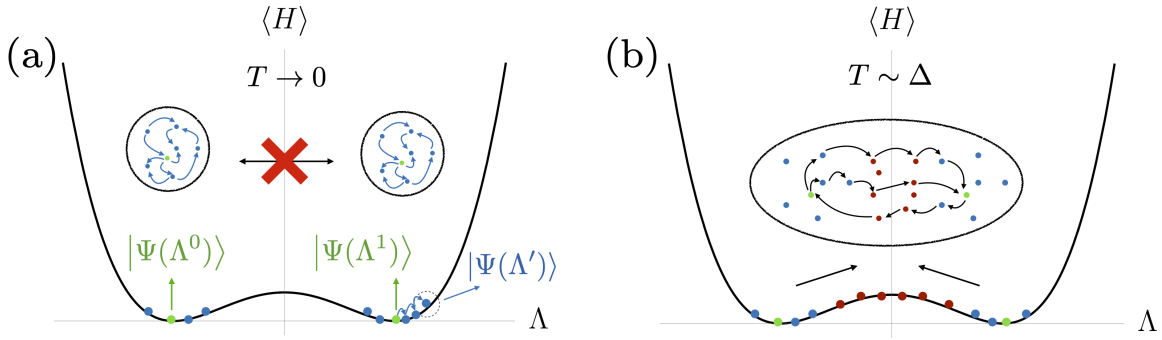

图 1. (a) “低能”系综示意图。从两个不同拓扑扇区中选取两个（或更多）初始态，$|\Psi(\Lambda^0)\rangle$ 和 $|\Psi(\Lambda^1)\rangle$，作为“种子”（绿色圆点）。圆点表示数据集（稍后输入扩散映射），该数据集由一组以网络参数 Λ标记的量子态构成。此数据集通过第 IIA 节和算法 1中概述的流程生成：下一个状态 Λ'（每个箭头处的蓝色圆点）由随机局域扰动提出，并根据能量期望值 $\langle H\rangle_{\Lambda'}$以一定概率接受。在小 T 的情况下，局域扰动无法将整个数据集连接起来，各拓扑扇区中的状态分别聚成簇（位于左、右两侧的谷地）。(b) “高能”系综示意图。状态采用与前述相同的算法生成，但使用较大的超参数 T（相较于能隙 Δ）。在此情况下，数据集包含一些低能态（蓝色圆点），也包含一些高能态（红色圆点）。由于高能态无法区分低能拓扑扇区，因此存在一些路径（椭圆形区域中圆点间的箭头所示），使得来自不同拓扑扇区的两个初始种子有效地“扩散”，并形成一个连通簇。

并非直接以高维向量的形式参与计算，而只通过相似性度量 $S(x_l, x_{l'})$ 参与计算，该度量用于比较两个样本 l 和 l' 的“相似”程度。在将扩散映射应用于拓扑分类问题时，它定义了样本的光滑变形（“同伦”）。我们讨论两种可能的度量。第一种就是波函数的量子力学重叠，$S_{\mathrm{q}}(\Lambda_l, \Lambda_{l'}) = |\langle\Psi(\Lambda_l)|\Psi(\Lambda_{l'})\rangle|^2$。虽然概念上简单明了，但在经典计算机上计算它的代价很高，因为需要进行重要性抽样。我们的网络试探态具有局域性，因此我们还可以构造另一种相似性度量，它可以表示为一个简单函数，其输入是描述待比较的两个状态的网络参数 Λ_l 和 $\Lambda_{l'}$。然而，这可能会引发一些微妙问题，因为具有不同 Λ 的两个状态可能对应于相同的波函数（相差一个全局相位）。我们将讨论如何高效地消除一般状态中的这些“规范冗余”。

我们以环面码为例说明这些方面，并明确展示该方法的成功 [57]。环面码是拓扑序的原型模型，此前也曾采用其他侧重点不同的机器学习技术进行研究 [15–18, 58–60]。我们展示了，DM算法无需先验知识，便能学习到环绕环面的底层环算符；在低 T下，这会产生四个簇，分别对应四个基态。在较大的 T下，这些簇会如预期般开始合并。有趣的是，DM仍能揭示数据集中与环算符期望值相关的底层结构。最后，我们还展示了施加磁场会使DM中的簇消失，从而捕捉到从拓扑序到禁闭相的转变。

本文其余部分安排如下。在第II节中，我们概述机器学习方法，包括所采用的局域网络量子态描述、系综生成、对文献12中DM方案的简要回顾，以及基于神经网络参数的相似性度量。随后以环面码模型为例，在第III节中详细讨论并阐释上述各个一般方面。最后，第IV 节给出明确的数值结果，并在第V节作出总结。

II. 通用算法

这里，我们首先以一般性的术语介绍并讨论我们的算法 [见图 2(a)]，随后将在后续章节中以环面码为例进行说明。考虑一个由 N 个量子比特或自旋组成的系统，其对应的算符为 $\{\hat{\boldsymbol{s}}\} = \{\hat{\boldsymbol{s}}_i, i = 1,\cdots,N\}, \hat{\boldsymbol{s}}_i = (\hat{s}_i^x, \hat{s}_i^y, \hat{s}_i^z)$，相互作用由局域且具有能隙的哈密顿量 $\hat{\mathcal{H}} = \mathcal{H}(\{\hat{\boldsymbol{s}}\})$ 控制。我们使用神经网络量子态 [46], 表示该系统的状态 $|\Psi(\Lambda)\rangle$

$$|\Psi(\Lambda)\rangle = \sum_{\boldsymbol{\sigma}} \psi(\boldsymbol{\sigma};\Lambda)\,|\boldsymbol{\sigma}\rangle\,, \tag{1}$$

其中 $\boldsymbol{\sigma} = \{\sigma_1, \sigma_2, ..., \sigma_N | \sigma_i = \pm 1\}$ 枚举局部计算基（例如 s^z基）中的物理自旋变量构型，而 Λ是网络 ψ 用来输出构型 $|\boldsymbol{\sigma}\rangle$对应的波函数振幅 $\psi(\boldsymbol{\sigma};\Lambda) = \langle\boldsymbol{\sigma}|\Psi(\Lambda)\rangle$ 的一组参数。由于物理希尔伯特空间会随系统大小呈指数增长，因此在选择网络架构（或试探态）ψ时，需要权衡表达能力与效率，使权重 Λ 能够以合理的精度近似态 $|\Psi(\Lambda)\rangle$，并

同时成为一种高效的表示（参数数量最少，且参数 Λ 随 N呈多项式增长）。为达到基态，或更一般地说，达到与低温物理相关的哈密顿量 $\mathcal{H}$ 的相关低能扇区，我们使用学习率 λ通过梯度下降，最小化由方程（1）定义的变分子空间中的能量，

$$\Lambda \to \Lambda - \lambda\, \partial_\Lambda \langle \hat{\mathcal{H}} \rangle_\Lambda , \quad \langle \hat{\mathcal{H}} \rangle_\Lambda = \langle \Psi(\Lambda) | \hat{\mathcal{H}} | \Psi(\Lambda) \rangle . \tag{2}$$

这里，量子力学期望值 $\langle \hat{\mathcal{H}} \rangle_\Lambda$ 采用重要性抽样进行计算（见附录 B）。

尽管希尔伯特空间中有指数多个状态，但局域哈密顿量的低能扇区预计只占据一个较小的子空间，其中的状态遵循面积律纠缠 [61, 62] ，而典型状态则遵循体积律 [63, 64]。受这些考虑启发，我们考察一类能够自然描述遵循面积律纠缠的量子态的网络。从图示上看，在这类网络中，连接隐藏神经元（表示权重 Λ）与物理自旋的连接是准局域 [51, 53–55]。在这种情况下，有

$$\psi(\boldsymbol{\sigma}, \Lambda) = \phi_1(\boldsymbol{\sigma}_1, \Lambda_1) \times \phi_2(\boldsymbol{\sigma}_2, \Lambda_2) \times \cdots , \tag{3}$$

其中，$\boldsymbol{\sigma}_\mathrm{J} = \{\sigma_k\}_{k\in\mathrm{J}}$ 表示相邻自旋的（重叠）子集，满足 $\cup_\mathrm{J} \boldsymbol{\sigma}_\mathrm{J} = \boldsymbol{\sigma}$ ； Λ_J 表示网络参数（权重和偏置）的子集，这些参数与 J中的物理自旋相连。

算法 1系综生成

过程 $(\{\Lambda\}_{n=1}^{N})$
  初始化：优化参数 Λ
  对于 k 个独立时刻 **执行**:
    对于 n 个采样步数 **执行**:
      提出新的参数 $\Lambda_p = f(\Lambda_t)$
      以能量决定的概率接受
$\langle \hat{\mathcal{H}} \rangle_\Lambda$ 和参数 T:
      $\Lambda_{t+1} = \mathbb{P}_{\mathrm{accept}}(\Lambda' | \Lambda;\, T)$
  返回 每个 k 的最后 m 个状态： $\{\Lambda_i | i = n - m, ..., n\}_k$

A. 数据集：网络参数系综

我们用于无监督检测拓扑序的数据集由一组波函数 $\{|\Psi(\Lambda)\rangle\}_l$构成，这些波函数由网络参数集合 $\{\Lambda\}_l$参数化。虽然根据具体应用，还可以采用其他选择，但我们生成该系综时，使状态 $|\Psi(\Lambda)\rangle$ 的相对出现概率由 $\hat{} \ \rho_T(\Lambda) = \exp(-\langle \mathcal{H} \rangle_\Lambda / T)/Z$给出，并带有适当的归一化因子 Z。因此，“类温度”超参数 T 较小时，对应“低能”系综； T 较大时，则参数化“高能”系综。

实际生成该系综时，我们首先通过优化参数 Λ （利用式（2））， 获得能量期望值最低的波函数。由于式（1）并不包含所有可能的状态，因此通常只能得到 $\mathcal{H}$的精确低能本征态的近似。不过，只要它能够涵盖系统的所有超选择扇区，以及连接这些扇区的（部分）高能状态，式（1）就足以满足我们检测拓扑序是否存在的目的。我们使用不同的初始条件，多次执行优化 $\Lambda \to \Lambda_l^0$，以获得若干“种子” Λ_l^0；这样做是为了确保每个超选择扇区都有一个低能代表态。理想情况下，数据集应直接从目标概率分布 ρ_T中采样，例如在能够访问有限温度下的实验系统时。本文采用一种受马尔可夫链启发的流程，基于ρ_T 为每个种子生成系综。具体而言，从状态 Λ出发，我们对随机选取的、与格点 J处自旋相关联的局域参数块提出更新，

$$\Lambda \to \Lambda' = \{\Lambda_1, \Lambda_2, \cdots, u(\Lambda_\mathrm{J}), \cdots, \Lambda_N\}, \tag{4}$$

其中，更新 u 仅取决于 Λ_J。给定当前参数 Λ 时，提议参数 Λ' 被接受的概率为

$$\mathbb{P}_{\mathrm{accept}}(\Lambda' | \Lambda; T) = \min\Big(1,\, e^{-\frac{\langle \hat{\mathcal{H}} \rangle_{\Lambda'} - \langle \hat{\mathcal{H}} \rangle_\Lambda}{T}}\Big). \tag{5}$$

这意味着，如果提议状态 $\Psi(\Lambda')$ 的能量期望值低于 $\Psi(\Lambda)$，则接受该提议；否则，以由玻尔兹曼因子决定的概率接受该提议。整个系综生成过程总结于算法 1。

B. 扩散映射

正如参考文献 12所提出的，扩散映射非常适合作为一种无监督机器学习算法，用于识别一组状态中是否存在超选择扇区以及其数量，例如上文定义的 $\{|\Psi(\Lambda)\rangle\}_l$。简要回顾扩散映射算法的核心思想 [32–35] 并引入相关记号，假设给定一个数据集 $X = \{x_l | l = 1, 2, ..., M\}$，其中包含 M 个样本 x_l。下文将讨论 $x_l = \Lambda_l$ 和 $x_l = |\Psi(\Lambda_l)\rangle$两种情况；第一种情况下，样本是用于参数化波函数的网络参数；第二种情况下，样本则是波函数本身。

为了直观地理解扩散映射，让我们定义状态 $x_l \in X$ 之间的扩散过程。状态 x_l 转移到 $x_{l'}$ 的概率由马尔可夫转移矩阵元素 $p_{l,l'}$ 定义。为了构造 $p_{l,l'}$，我们在状态 x_l 和 $x_{l'}$ 之间引入对称正定核 $k_\epsilon(x_l, x_{l'})$。接下来，转移概率矩阵

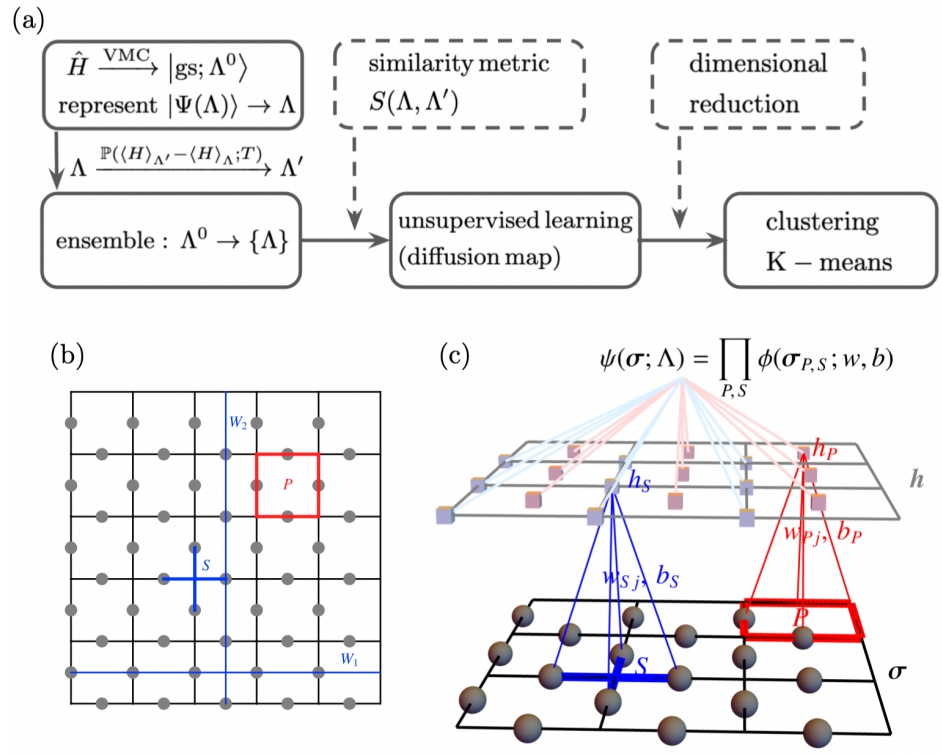

图2。（a）本工作所用机器学习算法的概览："种子" $\{\Lambda^0\}$ 通过变分蒙特卡罗计算得到（见附录B），接着使用马尔可夫链算法生成网络参数集合集（第 II A 节），然后在 DM 方法中使用相似性度量来定义核函数（第 II B 节和第 II C 节），最后对由主导 DM特征向量分量所张成子空间中的低维嵌入应用 k-均值算法。（b）环面码模型的正方晶格几何结构，其中量子比特$\hat{s}_i$ 定义在晶格的链路上（灰点）。方程（16）中的哈密顿量 [] 以算符$\hat{}$ $\mathcal{P}_P$ （由面元 P 上的自旋支撑，面元以红色方块表示）和星形算符$\hat{}$ $\mathcal{S}_S$ （由星形 S 上的自旋支撑，星形以蓝色链路表示）表示。沿 x（y）方向的两条蓝线表示沿直线路径 $\bar{x}$（$\bar{y}$）的威尔逊环算符$\hat{}$ W_1、$\bar{x}$（$\hat{}$$W_2$、$\bar{y}$）。（c）方程（17）中准局域试探态的示意图。该试探态是面元（或星形）上自旋的局域函数 ϕ 的乘积；当 $X = P$（S）为面元（或星形）时，该函数取决于参数 $\{w_{Xj}, b_X\}$。

$p_{l,l'}$ 定义为

$$p_{l,l'} = \frac{k_\epsilon(x_l, x_{l'})}{z_l}, \quad z_l = \sum_{l'} k_\epsilon(x_l, x_{l'}), \tag{6}$$

其中因子 z_l 确保概率守恒，$\sum_{l'} p_{l,\,l'} = 1\,\forall l$。随后，对转移概率矩阵进行谱分析，可以获得有关数据集 X的全局连通性的信息；在我们研究包含低能态的 X情境中，这使我们能够识别超选择扇区，进而确定拓扑序 [12]。为量化两个样本 x_l 和 $x_{l'}$ 之间的连接强度，引入了 $2t$ 步扩散距离 [32–35]，

$$D_{2t}(l,l') = \sum_{l''} \frac{1}{z_{l''}} [(p^t)_{l,l''} - (p^t)_{l',l''}]^2, \tag{7}$$

其中 p^t 表示转移概率矩阵 p的第 t 次矩阵幂。已有研究表明， D_{2t}可由转移矩阵 p的特征值 λ_n 和右特征向量 ψ_n 计算得到：其中$\sum_{l'} p_{l,\,l'} (\psi_n)_{l'} = \lambda_n (\psi_n)_l$，且按降序排列 $\lambda_n > \lambda_{n+1}$, it follows

$$D_{2t}(l,l') = \sum_{n=1}^{M-1} \lambda_n^{2t} [(\psi_n)_l - (\psi_n)_{l'}]^2 \tag{8}$$

经过直接代数运算后可得 [35]。从几何上看，这意味着若执行非线性坐标变换 $x_l \to \{(\psi_n)_l,\ n = 0, \ldots M-1\}$，扩散距离便可表示为欧几里得距离（以 λ_n 加权）。此外，由于全局连通性可从扩散距离的长时间极限 $t \to \infty$中体现出来，因此最大的特征值对描述连通性最为重要。更准确地说，我们选择一个形式如下的核函数 k_ϵ

$$k_\epsilon(x_l, x_{l'}) = \exp\left(-\frac{1 - S(x_l, x_{l'})}{\epsilon}\right), \tag{9}$$

其中 S 是一个局域相似性度量，满足 $S \in [0,1]$, $S(x_l, x_{l'}) = S(x_{l'}, x_l)$，且 $S(x,x) = 1$。这里，"局域"是指 $S(x_l, x_{l'}) = \sum_i \mathcal{S}_i(x_l, x_{l'})$，其中 $\mathcal{S}_i(x_l, x_{l'})$ 仅取决于 x_l 的构型以及

$x_{l'}$ 位于格点 i附近的构型。虽然我们将在下文第 II C 节讨论量子力学 N自旋/量子比特系统中 S 可能的显式形式，但对于由 N个自旋构成的经典系统，$x_l = \{\boldsymbol{S}_i^l, (\boldsymbol{S}_i^l)^2 = 1, i = 1, 2, \ldots, N\}$，一个自然的选择是 $S_{\rm cl}(x_l, x_{l'}) = \sum_i \boldsymbol{S}_i^l \cdot \boldsymbol{S}_i^{l'}/N$。在式 (9) 中，$\epsilon$ 起到“粗粒化”参数的作用，这是必要的，因为我们处理的只是有限数据集 X：对于给定的 X，通常预期当 $\epsilon \to 0$时，$k_\epsilon(x_l, x_{l'}) = p_{l,l'} = \delta_{l,l'}$，也就是说，如果 ϵ 足够小，所有样本彼此都不相似，且所有特征值 λ_n 都趋近于 1。反过来，当 $\epsilon \to \infty$ 时，粗粒化参数非常大，所有样本都彼此连通，$k_\epsilon(x_l, x_{l'}) \to 1$；由于$p_{l,l'} \to 1/M$，我们将得到 $\lambda_{n>0} \to 0$，而最大本征值 λ_0 始终为 1（这是概率守恒的结果）。对于介于这两个极限之间的 ϵ 取值，DM 谱包含关于 X的信息，包括其拓扑结构：如参考文献12所示，若 X 中存在 $k \in \mathbb{N}$ 个不同的拓扑等价类，则会体现为一段 ϵ取值范围，在该范围内 $\lambda_1, \ldots \lambda_{k-1}$均以指数精度（关于 ϵ）接近 1，且与 $\lambda_{n\geq k}$之间存在明显能隙。此外，在映射后的 $k-1$维空间中绘制不同样本 l 的图时，它们会依据所属的拓扑等价类聚类；这里采用标准欧几里得度量，例如可通过 k-均值算法实现。该空间为$\{(\psi_1)_l, (\psi_2)_l, \ldots, (\psi_{k-1})_l\}$。接下来，我们将使用这一方法来识别第 II A 节定义的波函数系综中的超选择扇区。不过，为此我们首先需要引入一个合适的相似性度量 S，下文将对此进行讨论。

C. 局部相似性度量

上述经典相似性度量 $S_{\rm cl} = \sum_i \boldsymbol{S}_i^l \cdot \boldsymbol{S}_i^{l'}/N$ 的自然推广方式，是取量子系统希尔伯特空间中的内积；前者可视为经典构型空间中的（欧几里得）内积。

$$S_{\rm q}(\Lambda_l, \Lambda_{l'}) = |\langle \Psi(\Lambda_l) | \Psi(\Lambda_{l'}) \rangle|^2. \tag{10}$$

虽然对于低秩量子态，利用量子模拟和计算装置可以高效估算这类或其他相关的保真度度量 [65–68], ，但估算 $S_{\rm q}$ 通常需要在经典计算机上执行计算开销很大的任务，因为这要求针对我们的变分过程对自旋构型进行采样。为了更高效地评估相似性度量，我们在此提出另一种途径，利用我们对 $\psi(\boldsymbol{\sigma}; \Lambda)$ 使用局域试探态这一事实，见式（3）。我们的目标是将相似性度量直接表示为

$$S_{\rm n}(\Lambda_l, \Lambda_{l'}) = \frac{1}{N_{\rm J}} \sum_{\rm J} f((\Lambda_l)_{\rm J}, (\Lambda_{l'})_{\rm J}), \tag{11}$$

其中 f 仅比较由 $_{\rm J}$表示的局域参数块，并且是一个可快速求值的函数，无需对自旋构型进行采样。此外，通过选择一个关于其参数对称的函数 f，即可确保 $S(x_l, x_{l'}) = S(x_{l'}, x_l)$；同时，只需设置 $N_{\rm J} = \sum_{\rm J}$并对 f 进行适当缩放，使得 $f \in [0, 1]$，即可轻松实现这一点。最微妙的条件是

$$S_{\rm n}(\Lambda_l, \Lambda_{l'}) = 1 \iff |\Psi(\Lambda_l)\rangle \propto |\Psi(\Lambda_l')\rangle, \tag{12}$$

因为这取决于用于 $\psi(\boldsymbol{\sigma}; \Lambda)$ 的具体网络架构，其中存在对权重 $\Lambda_l \to g[\Lambda_l]$ 的“规范变换” $g \in \mathcal{G}$，满足

$$|\Psi(\Lambda_l)\rangle = e^{i\vartheta_g} |\Psi(g[\Lambda_l])\rangle \tag{13}$$

对于某个全局相位 ϑ_g。我们希望确保

$$S_{\rm n}(\Lambda_l, \Lambda_{l'}) = S_{\rm n}(\Lambda_l, g[\Lambda_{l'}]) = S_{\rm n}(g[\Lambda_l], \Lambda_{l'}) \tag{14}$$

对于所有此类规范变换 $g \in \mathcal{G}$。保证式（14）成立的一种通用方法是将其替换为，

$$S_{\rm n}(\Lambda_l, \Lambda_{l'}) \longrightarrow \max_{g, g' \in \mathcal{G}} S_{\rm n}(g[\Lambda_l], g'[\Lambda_{l'}]). \tag{15}$$

然而，在实际操作中，由于相似性度量具有局域性，可能无需遍历 $\mathcal{G}$ 中的所有可能规范变换。接下来，我们将以环面码和一种特定的受限玻尔兹曼机变分试探态为例，说明这些规范变换，以及如何高效实现式（f 11）中的适当函数 和规范不变性（14）。

最后需要指出，尽管本文着重于应用扩散映射，但基于神经网络参数定义的相似性度量也可用于其他核方法，例如核主成分分析。根据底层数据集的结构，扩散映射相较于核主成分分析具有明显优势：前者确实捕捉了数据集的全局连通性，而非后者提取的方差最大的子空间。这就是为什么在一般数据集中，核主成分分析无法识别例如绕数，而扩散映射仍能良好发挥作用 [12]。具体而言，对于下文关于环面码的案例研究，我们发现，数据结构简单，因此在较小的 T 且没有磁场 $h = 0$时，核主成分分析也能识别拓扑扇区；然而，当开启 h 时，只有扩散映射能良好发挥作用，正如下文所讨论的。

III. 示例：环面码

现在，我们以环面码模型为例，展示基于扩散映射的机器学习算法 [57], ，该模型定义在 $L_x \times L_y$ 正方晶格上，每条键上都有自旋1/2算符或量子比特，见图2(b)，总计有 $N = 2L_xL_y$ 个自旋；在本文中，我们始终假设采用周期性边界条件。将晶格上一个基本正方形（顶点）边缘上的四个自旋统称为面元P (星形 S)，面元算符和星算符定义为

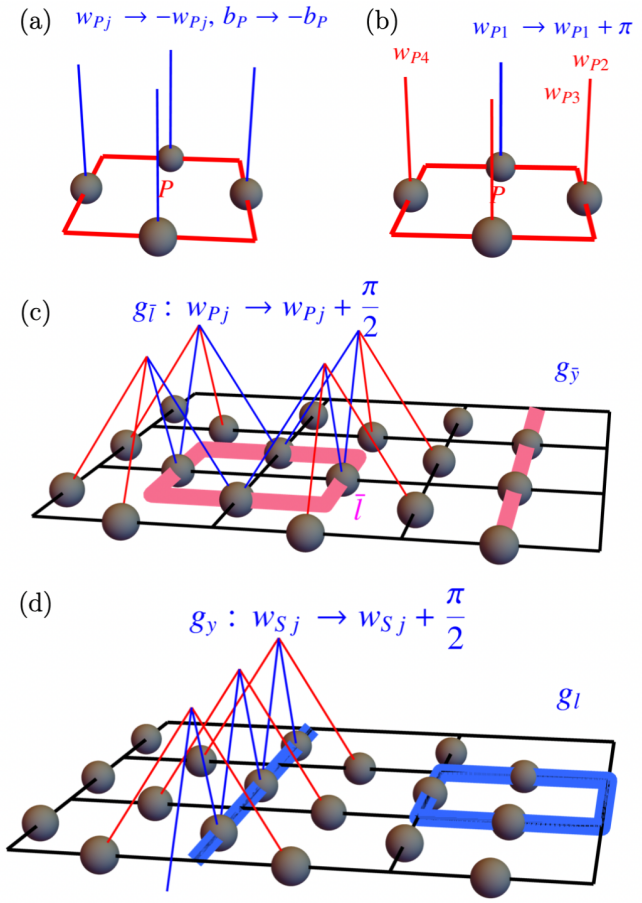

图 3. 式（17）中受限玻尔兹曼机试探态的规范自由度。以下变换只会导致全局相位变化：(a) 将一个面元（或星形，图中未示出）的所有参数乘以负号，见式（18a）；(b) 对单个参数进行 π 移位，见式（18b）和（18c）；(c) 对被弦 $\bar{l}$ 穿过的权重进行 $\pi/2$ 移位，其中弦 $\bar{l}$ 由式（ $g_{\bar{l}}$ 定义，见式（18e）。粉色直线表示对非可缩回路的变换，该回路记为 $g_{\bar{y}}$；(d) 与 (c) 相同，但针对直接晶格上的环路以及 g_l 和 $g_{\bar{y}}$，参见式（18d）。

分别为 $\hat{\mathcal{P}}_P = \prod_{i\in P}\hat{s}_i^z$ 和 $\hat{\mathcal{S}}_S = \prod_{i\in S}\hat{s}_i^x$。因此，环面码哈密顿量写作

$$\hat{H}_{\rm tc} = -J_P\sum_P \hat{\mathcal{P}}_P - J_S\sum_S \hat{\mathcal{S}}_S, \tag{16}$$

其中求和遍历$\hat{}$ $\hat{}$晶格上的所有面元和星。所有“稳定子” $\mathcal{P}_P$, $\mathcal{S}_S$ 彼此对易，且与哈密顿量对易。聚焦于 J_P, $J_S > 0$，基态是所有稳定子对应本征值为 $+1$ 的本征态。通过计数论证，并考虑约束 $\prod_S \mathcal{S}_S = \prod_P \mathcal{P}_P = \mathbb{1}$，可知在周期性边界条件下恰有四个简并基态。

为了以变分方式描述环面码模型（16）的基态和低能子空间，我们对其进行参数化——

化 $\psi(\boldsymbol{\sigma};\ \Lambda)$ 在式（1）中，并使用试探态

$$\psi_{\rm rbm}(\boldsymbol{\sigma};\Lambda) = \prod_P \cos(b_P + \sum_{j\in P} w_{Pj}\sigma_j) \times \prod_S \cos(b_S + \sum_{j\in S} w_{Sj}\sigma_j), \tag{17}$$

见参考文献58中提出的方法，其中每个面元 P （星形 S）都对应一个“偏置” b_P （b_S）和四个权重$w_{P,\ j}$ （$w_{S,\ j}$），此处所有这些量均取实数，即 $\Lambda = \{b_P, b_S, w_{P,j}, w_{S,j}\}$。该试探态可视为一个受限玻尔兹曼机 [46] （见附录A），如图2（c）所示，其几何性质与底层环面码模型相同。显然，式（17）定义了一个准局域试探态，因为它具有式（3）的形式，其中 $_j$枚举所有面元和星（因此 $N_{\rm j} = 2N$）。对于这个特定的试探态，如上文第IIC 节所述，规范变换 $g \in \mathcal{G}$由以下对参数 b_P、 b_S、 $w_{P,\ j}$ 和 $w_{S,\ j}$进行的操作生成：

1. 对于 X 任意一个面元或星形，将该面元或星形的所有偏置和权重乘以 -1 [见图3(a)],

$$g_{X,-}: b_X \to -b_X,\ w_{Xj} \to -w_{Xj}, \tag{18a}$$

使波函数保持不变 [$\vartheta_g = 0$，见式（13）]。

2. 将 π 与该面元或星形相关联的偏置或任意一个权重增加 X [见图3(b)],

$$g_{X,\pi,b}: b_X \to b_X + \pi, \tag{18b}$$
$$g_{X,\pi,j}: w_{Xj} \to w_{Xj} + \pi,\quad j \in X, \tag{18c}$$

会使 [$\vartheta_g = \pi$ 式（13）]整体多出一个负号。

3. 对于任意闭合回路 ℓ （或 $\bar{\ell}$），在直接（或对偶）晶格上，将$\frac{\pi}{2}$ 所有与弦穿过的自旋相连的星（面元）的权重增加[见图3(c-d)],

$$g_\ell: w_{Sj} \to w_{Sj} + \frac{\pi}{2},\quad Sj \in \ell, \tag{18d}$$
$$g_{\bar{\ell}}: w_{Pj} \to w_{Pj} + \frac{\pi}{2},\quad Pj \in \bar{\ell}, \tag{18e}$$

根据弦的长度，在公式（13）中分别得到 $\vartheta_g = 0$ 或 π 。注意，任意回路构型$\mathcal{L}$（其中可以包含任意数量的环路）都可以由集合 $\{g_S, g_P, g_{x,y}, g_{\bar{x},\bar{y}}\}$ 生成，其中g_S (g_P)在对偶（直接）晶格上围绕星形 S （面元 P）生成一个基本回路，见图 3（c、d）；而 $g_{x,\ y}$ ($g_{\bar{x},\ \bar{y}}$)则在直接（对偶）晶格上沿 x, y 方向生成一条非可缩回路。由于任意可缩回路的长度都是偶数，$\vartheta_g = 0$ 对于由g_S 和 g_P生成的任意弦变换都成立。在奇数晶格上，规范变换 $g_{x,\ y}$($g_{\bar{x},\ \bar{y}}$)涉及奇数个格点，因此会导致 $\vartheta_g = \pi$。

处理这种规范冗余的一种极低效方法，是采用 S_n 公式（11）中的一种选择，该选择不具有公式（18）中任一变换下的不变性；例如，直接取权重的欧几里得距离就属于这种情况，

$$S_{\rm eu}(\Lambda_l,\Lambda_{l'}) \propto ||\Lambda_l-\Lambda_{l'}||^2 = \sum_X\Big[(b_X^l-b_X^{l'})^2+\sum_{j\in X}(w_{Xj}^l-w_{Xj}^{l'})^2\Big],$$

其中对 X 的求和涉及所有面元和星形。若直接遍历所有可能的规范变换以求出式（15）中的最大值，原则上可以弥补规范不变性的缺失。然而，由于规范变换的数量随系统大小呈指数增长 N（这对上述定义的三类变换 1.-3. 均成立），这种方法在较大的 N下会变得非常耗费计算资源。幸运的是，试探态和相似性度量的局域性使我们能够构造出可更快速计算的相似性度量：例如，考虑

$$S_n(\Lambda_l,\Lambda_{l'}) = \frac{1}{2}+\frac{1}{10N}\sum_X \max_{\tau_X=\pm}\Big[\sum_{j\in X}\cos 2(\tau_X w_{Xj}^l - w_{Xj}^{l'})+\cos 2(\tau_X b_X^l - b_X^{l'})\Big], \tag{19}$$

显然，它满足 $S_n(\Lambda_l,\Lambda_{l'}) = S_n(\Lambda_{l'},\Lambda_l)$,$S_n(\Lambda_l,\Lambda_{l'}) \in [0,1]$，并且具有局域性 [它具有式（11）的形式，其中 j枚举所有 X]。关于规范不变性，首先注意到，选择 cos(·) 会立即带来对式（18a）的不变性。其次，对于每个 X，我们只需在两个取值（τ_X）之间求最大值，即可确保对式（18b）和（18c）的不变性；也就是说，该最大化过程只会使计算成本加倍。

不过，式（18d）和（18e）中的“弦”冗余尚未在式（19）中予以考虑。形式上，可以通过对所有可能的环配置取最大值来处理这一点，记作 $\mathcal{L}$，

$$S_{\rm str}(\Lambda_l,\Lambda_{l'}) = \frac{1}{2}+\frac{1}{10N}\max_{\mathcal{L}}\Big\{\sum_X \max_{\tau_X=\pm}\Big[\sum_{j\in X}\mu_{Xj}^{\mathcal{L}}\cos 2(\tau_X w_{Xj}^l - w_{Xj}^{l'})+\cos 2(\tau_X b_X^l - b_X^{l'})\Big]\Big\}, \tag{20}$$

其中 $\mu_{Xj}^{\mathcal{L}}=-1$ 若 Xj 位于包含在 $\mathcal{L}$ 中的回路上，$\mu_{Xj}^{\mathcal{L}}=1$ 则不然。尽管这类弦的数量呈指数增长，参考文献12提出了一种能够高效求得近似最大值的算法。在我们的例子中，该算法相当于随机选择一个面元 P或一个星形 S，或一个方向 $d=x$、y，然后在公式（19）中对 Λ_l 应用 g_S 或 g_P 或 $g_{d=x、y}$。如果相似度没有降低，就保留该变换；如果相似度降低，则舍弃该规范变换。重复此过程 N_g 次。在参考文献12中，N_g 在 10^3 和 10^4 之间就足以处理一个大型由 18×18 个方格晶格格点组成的系统（总计 $N=2\times 18^2$ 个量子比特）。此外，g_S和 g_P都是局域的，因此，评估规范变换引起的相似度变化只需耗费 $\mathcal{O}$（N^0）的工作量。

在下面的数值模拟中，不对环配置进行采样而直接使用式（19）$\mathcal{L}$ 就足够了。原因在于，对于我们所采用的、受马尔可夫链启发的 Λ_l 采样过程（见附录 C），对应这些回路变换的更新极少发生。此外，即使由于弦冗余，少数样本对被错误地分类为不同构型，扩散映射仍能正确捕捉全局连通性，从而判断拓扑扇区是否存在。

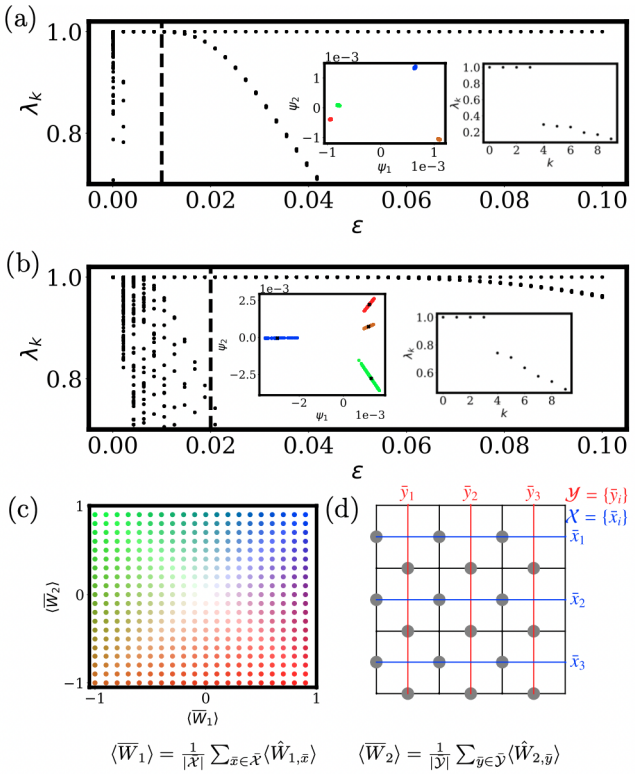

图 4. (a) 使用式（19）中的中性网络相似度量，在 $h=0$ 和 $T=0.1$ 处拓扑相的 DM 谱。左侧插图：对应的主导DM分量；颜色表示 (c-d) 中定义的环路可观测量期望值。右侧插图：按降序排列的 DM 谱，虚线标示 $\epsilon=0.01$ 处。(b) 与 (a) 相同，但使用式（10）中的精确重叠 $S_{\rm q}$ 作为度量。(c) (a) 和 (b) 左侧插图中的非局域环路值 $\langle W_1\rangle$，$\langle W_2\rangle$ 的色彩图。(d) 沿 x（y）方向的不同直线威尔逊环 $\hat{W}_{1,\,\bar{x}_i}$（$\hat{W}_{2,\,\bar{y}_i}$），分别用蓝色（红色）线表示。(c) 中色彩图的环路值是对所有直线环路期望值取空间平均值（如 $\langle W_1\rangle$，$\langle W_2\rangle$ 的方程所示）。

IV. 数值结果

接下来，我们将具体展示如何运用上述通用流程，探测并分析环面码中的拓扑序。我们从式（16）定义的纯环面码哈密顿量出发，使用式（17）中的变分受限玻尔兹曼机试探态。按照第 IIA 节所述的流程（另见算法 1），生成一组网络参数系综，系统大小为 $N=18$ 个自旋；系综生成所用的超参数以及包括式（ u 4）中形式在内的更多细节，见附录C。从现在起，所有能量均以 J_P 为单位，并令 $J_S=J_P=1$。

我们首先关注低能系综，并选择 $T=0.1$ 在方程（5）中的取值。对于方程（19）中的简单相似性度量，它可以在与系统大小 N呈线性关系的时间内精确计算。我们得到的扩散映射谱如图 4(a) 所示，它是方程（ ϵ 9）中 [12]的函数。我们观察到四个超选择扇区这一标志性特征：在 ϵ 的有限范围内，有四个特征值与 1 指数级接近。因此，样本（在我们的情形中即状态）与这四个扇区之间的对应关系，预期可在由前三个非平凡特征向量张成的投影子空间散点图中显现出来 $\psi_{1,\ 2,\ 3}$ [12]；注意，零阶特征向量（ψ_0）$_l=C$ 始终为常数，其特征值 $\lambda=1$ 由概率守恒决定。事实上，仅从前两个分量中便能看到这些簇，见图 4(a) 左侧插图。随后，将标准的 k均值算法应用于该投影子空间，以确定每个数据点所属的簇编号。为了验证机器学习算法是否根据四个物理扇区正确地对状态进行了聚类，我们计算每个状态对应的弦算符期望值，

$$\hat{W}_{1,\bar{x}}=\prod_{i\in\bar{x}}\hat{s}_i^x,\quad \hat{W}_{2,\bar{y}}=\prod_{i\in\bar{y}}\hat{s}_i^x, \tag{21}$$

其中 $\bar{x}(\bar{y})$ 是定义在对偶晶格上、沿 $x(y)$ 方向绕行的环路，如图 2(b) 中的蓝线所示。我们通过一组沿 $\mathcal{X}(\mathcal{Y})$ 方向绕行的直线环路的平均值，来量化一个态与物理扇区的关联；这些环路如图 $x(y)$ 中的蓝线（红线）所示。我们用图 4(c) 中定义的颜色编码，在图 4(a) 的插图中标示这一平均期望值$\langle W_1\rangle,\langle W_2\rangle$，确实可以看到聚类结果正确。

为证明这并非式（19）中相似性度量的特殊性质，我们使用式（10）中的完整量子力学重叠度量进行了相同分析，结果如图 4(b) 所示。定量细节有所变化，但正如预期，四个超选择扇区仍清晰可辨，且聚类结果正确。我们再次强调，式（19）中神经网络相似度量的计算精确计算 $\mathcal{O}(N)$] 比式（10）中的计算快得多 [精确计算 $\mathcal{O}(2^N)$，但我们可以在经典计算机上通过重要性抽样近似计算] 。不过请注意，一旦对所有样本计算出 S_n ，

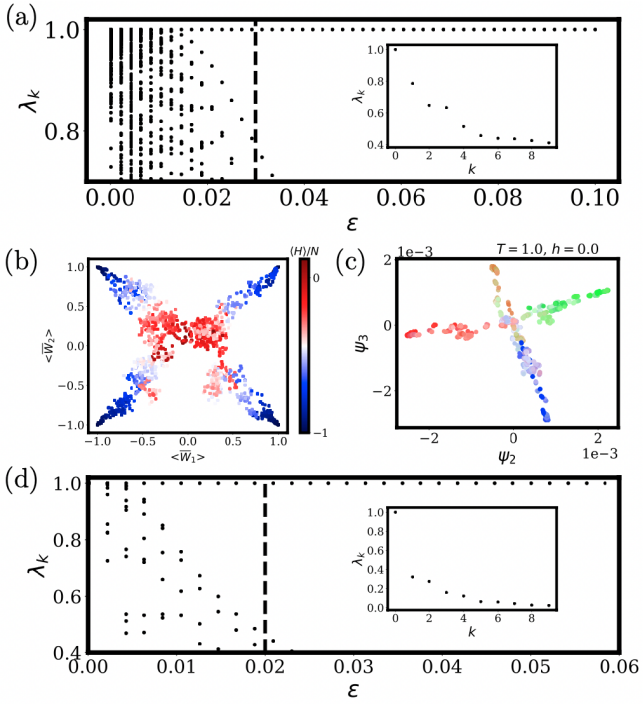

图 5. (a) 高能系综的 DM 谱，其中 $h=0$ 且 $T=1$。插图显示主图中虚线所示 $\epsilon=0.03$处的谱；(b) 空间平均直威尔逊环 $\langle W_{1(2)}\rangle$ [参见图 4(c-d)] ，对应于(a) 中状态的两个方向，其中颜色表示能量密度 $\langle H\rangle/N$；(c) 主导DM分量，点的颜色表示 $\langle W_{1(2)}\rangle$，采用图 4(d) 中的色图；(d) 使用量子度量 $h=1.0$ 且 $T=0.1$ 时平庸相的 DM 谱 S_{qo}。

实际上，这两种方法在基于DM的聚类中所需的计算时间相同。因此，假设存在一种量子模拟器，能够高效测量式（10）中的量子重叠，或任何其他可行的相似性度量，那么我们同样可以使用“测得的”相似性，通过DM方案高效地聚类超选择扇区。接下来，我们将证明，如果将能量足够高的状态纳入考虑，超选择扇区最终会彼此连通。为此，我们针对一个包含 $T=1$ 的系综重复相同的分析。如图5(a)所示，在所得的 DM 谱中，不存在任何 ϵ 值能使多个本征值（指数级地）接近 1，且与谱中其余部分之间存在明显能隙。这里我们再次使用了式（19）中的简化度量，但也已确认，使用重叠度量不会在定性上改变结果。为验证这是否是该数据集的正确结论，我们再次计算了系综中每个状态对应的式（21）中的环算符的期望值。如图5(b)所示，我们还用颜色表示每个状态的能量期望值。显然，四个低能量（蓝色）扇区（带有$|W_{1,2}|\simeq 1$）通过高能量（红色）状态（带有$|W_{1,2}|\ll 1$）彼此连通。这与DM结果一致，即

系综中的所有状态均彼此连通（拓扑序消失）。尽管如此，我们仍可以考察前三个非平凡DM分量 $\psi_{1,2,3}$。为简化展示，我们聚焦于图 5(c)中的二维投影。可以看到，DM揭示了数据中非常有趣的结构：四个叶瓣大致对应蓝、红、橙、绿四种颜色，它们分别与四个超选择扇区相关；更接近 $|W_{1,2}|=1$ （颜色较深）的状态则更靠近叶瓣尖端。最后请注意，颜色的排列方式使得红色与绿色、 [橙色与蓝色] 叶瓣位于两端，这符合预期，因为它们分别对应 $(W_1, W_2) \simeq (1,-1)$ 和 $(-1,1)$ [$(-1,-1)$ 和 (1,1)]。

摧毁拓扑序的另一条途径是施加磁场。为研究这一点，我们将环面码哈密顿量扩展为

$$\hat{H}'_{\rm tc} = \hat{H}_{\rm tc} - h\sum_i \hat{s}_i^z. \tag{22}$$

显然，在 $h \to \infty$ 的极限下，基态就是所有自旋都沿 $\hat{s}^z$ 方向极化的态，拓扑序也随之消失。从纯环面码模型（$h=0$）出发并施加 h，会降低“电荷”激发的能隙；这些激发由将环面码基态中的 $\mathcal{S}_S$ 从 $+1$ 翻转为 -1 定义。它们的凝聚会导致二级量子相变 [69–72]。

在讨论相变之前，我们先研究大h极限。首先注意到，式（17）中的试探态无需改变，因为它也能描述极化相。例如，将面元 P （以及星形 S）中最“北”（以及最“南”）的自旋分别记为$j_0(P)$ （以及 $j_0(S)$），则当 [时，自旋极化态得以实现；另见附录中的图 8（a）]

$$b_P = b_S = -\frac{\pi}{4}, \quad w_{Xj} = \begin{cases} \frac{\pi}{4}, & j = j_0(X), \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

事实上，在我们的式（17）受限玻尔兹曼机试探态中，自旋极化态有多种表示形式，其中包括并非仅由式（18）中的规范变换相互关联的表示。例如，自旋与面元和星形的对应关系 $j \to j_0(X)$ 可以改变，比如改用“最东侧”的自旋。如附录 A2中更详细的讨论所述，这种冗余源于式（ $\psi_{\rm rbm}(\boldsymbol{\sigma})$ 在式（17）中的乘积形式，以及 $\psi_{\rm rbm}(\boldsymbol{\sigma})$ 在只有一个 时恰好为零 j 满足 $\sigma_j=-1$；因此，这是自旋极化基态简单乘积性质的一种特殊表现。一般而言，除上述规范变换外，仍可能存在其他冗余，但我们不认为这种结构化的冗余集合会出现在一般状态中。解决这一问题有多种方法。最直接的方法是：对于一定比例的样本对 $S_{\rm n}$，将式（11）中的简单重叠度量替换为式（ $S_{\rm q}$ 10）中的直接重叠。若这一比例足够大，DM 算法便能识别出网络的簇

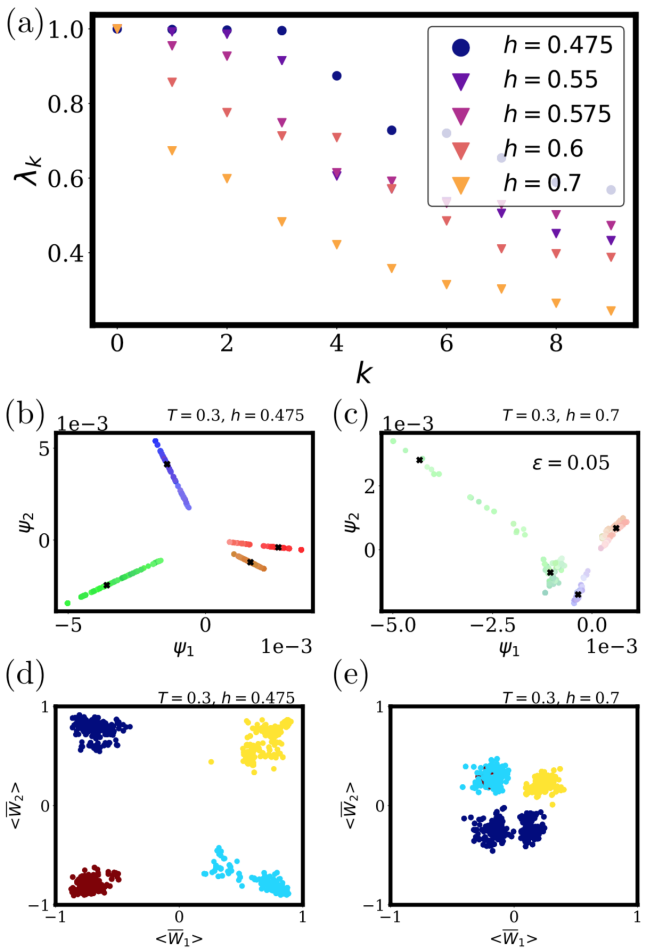

图 6. 有限场 h 下，具有 $T=0.3$ 的低能系综的 DM 谱。(a) 在 $\epsilon=0.05$ 处，不同场值 $h=0.475$、0.55、0.575、0.6、0.7 对应的前 10 个特征值。圆点标记（$h=0.475$）表明本征值谱具有四重简并，这是拓扑序的特征。相比之下，由三角形标记（$h \geq 0.55$）标出的谱不具有这种简并。通过观察简并本征值谱中出现能隙，确定转变场值为 $h_t \simeq 0.55$。这与我们在使用相同数据集 [时通过保真度观察到的结果一致，见附录B 1]。(b) 将 $h=0.475$ 的投影本征向量投影到前两个分量上。颜色表示 $\langle W_{1(2)}\rangle$，配色方案见图 4(c)。黑色叉号表示 k-均值中心。(c) 与 (b) 相同，但对应 $h=0.7$。(d) (b) 中状态沿两个方向的平均直线威尔逊环 $\langle W_{1(2)}\rangle$ 的期望值。颜色表示根据 (b) 所示特征向量在投影子空间中进行 $\overline{k}$-均值聚类得到的结果。(e) 与 (d) 相同，但对应 (c) 中所示的系综。

根据 $S_{\rm n}$ 实际上可能对应于相同波函数的不同参数。我们在附录 A3 中明确展示了这一点。不过，我们注意到，在这种情况下，核主成分分析将不再适用；它会错误地将连通样本分类为不同样本，因为它依据的是数据方差，而非连通性。为简化表述，正文中所有状态均使用 $S_{\rm q}$，并重点讨论扩散映射。

大磁场 $h = 1$和低温 $T = 0.1$下的 DM 谱如图5(d)所示。显然，不存在任何 ϵ 值能使多个本征值接近 1，同时与谱的其余部分保持能隙。这表明，正如预期的那样，磁场 h 导致了拓扑序的消失。

为使用我们的 DM 算法研究由 h 引起的相关相变，我们针对 h 的各种不同取值重复了相同的步骤。所选 h 对应的谱如图 6(a) 所示。我们看到，对于数据中的 $h = 0.55$，仍有四个扇区；而在 $h = 0.575$ 及更大的取值下，这些扇区则不存在。虽然相关的临界值 h 大于预期 [69–71]，但这并非 DM 算法的缺陷，而是由方程 (17) 中简单的局域变分试探态所导致的。通过计算保真度和环算符期望值，我们可以看到，对于我们的数据集而言，约为 $h = 0.55$ 的临界值才是预期答案（见附录 B 1）。更复杂的波函数试探态预计能给出更好的取值，但这并非本文的重点。更重要的是，我们从图 6(b) 可以看到，DM 对状态的聚类正确复现了依据平均环算符期望值进行的聚类 $\langle W_j \rangle$（同样以颜色标示）。或者，也可以从图 6(d) 中看到这一点，其中 $\langle W_j \rangle$ 标示了各个样本。使用四种不同颜色表示 DM 识别出的四个不同簇，可以看到所有状态都被正确聚类。正如根据特征值所预期的那样，在较大的 h 下已不再有明显的簇，见图 6(c)；尽管如此，在 ψ_1 中直接应用 k-均值聚类，仍能发现波函数中与 $\langle W_j \rangle$ 相关的一些残余结构，如图 6(e) 所示。

V. 总结与讨论

在这项工作中，我们介绍了一种用于研究具有拓扑序的量子相的无监督机器学习算法。我们使用神经网络参数高效地表示量子态系综，并按照其能量期望值对量子态进行采样。为了揭示量子态中超选择扇区的结构，我们采用了扩散映射这一降维技术，并给出了一个以网络参数定义的核函数。与基于波函数重叠（或其他量子力学相似性度量）的核函数不同，这种度量可以在经典计算机上高效计算（在多项式时间内）。

我们以准局域受限玻尔兹曼机（RBM）和处于外场中的环面码模型为例，展示了我们的一般算法；选择这种网络试探态的灵感来自先前的研究 [58, 59]，这些研究表明，低能谱可以用 RBM 高效表示。通过允许 RBM 网络具有空间非均匀性，我们识别出了该试探态的“规范对称性”，即网络中不会改变波函数的那组变化参数（平凡的全局相位因子除外）。我们精心设计了一种规范不变的相似性度量——这是一项关键性质；否则，以不同规范表示的相同波函数会被错误地判定为彼此不同。我们证明，所得的基于扩散映射的波函数无监督嵌入与环算符的期望值相一致；它正确捕捉了低能量和弱磁场下超选择扇区及拓扑序的存在，也捕捉了涉及较高能态和/或磁场增强时二者的缺失。我们还使用波函数的完整量子力学重叠作为相似性度量，验证了这些结果。

从更一般的层面来看，我们的分析凸显了扩散映射的以下两项关键性质：第一，当存在不同的拓扑扇区时，扩散映射的主导特征向量捕捉的是连通性，而非例如主成分分析所关注的方差。因此，即使由于使用近似相似性度量，部分波函数对被错误地判定为彼此不同，聚类仍能正确完成。这也是为什么将神经网络相似性度量（它在大场极限下还存在额外的态特定冗余）与一定比例的状态对之间的直接量子力学重叠相结合，就足以得到正确的分类。第二项关键性质是，扩散映射是一种核方法。这意味着实际的机器学习过程不需要将完整波函数作为输入；只需提供数据集中所有波函数对的核（或其某种度量）即可。我们通过在式（20）中适当定义网络相似性度量，利用这一点有效消除了状态的受限玻尔兹曼机参数化中的规范冗余。由于在经典计算机上计算完整的量子力学相似性度量（如波函数重叠）代价高昂，未来一个有趣的方向是利用新兴的量子计算资源，以量子力学方式计算相似性度量。随后可将其作为基于扩散映射的聚类的输入。

最后需要指出的是，本研究所采用的态系综是根据各状态相对于某个哈密顿量的能量进行采样得到的，这只是众多可能性之一。原则上，所提出的技术——使用关于量子多体波函数变分描述中网络参数的规范不变核来进行扩散映射聚类——可以更广泛地应用于任何感兴趣的系综。例如，若要考虑任意局域微扰，可以利用有限深度局域幺正电路生成一个系综。或者，也可以基于（Lindblad）时间演化生成一个系综，以探究拓扑序对时变微扰或与热浴耦合的稳定性。我们将此类可能性的研究留待未来工作。

VI. 代码和数据可用性

本研究中的蒙特卡罗模拟是在 JAX 中实现的 [73]。Python代码和数据将在以下网址提供：https://github.com/teng10/mltoric code/。

致谢

Y.T. 感谢与 Dmitrii Kochkov、Juan Carrasquilla、Khadijeh Sona Najafi、Maine Christos 和 Rhine Samajdar 的有益讨论。Y.T. 和 S.S. 感谢美国能源部通过资助项目 DE-SC0019030 提供的经费支持。M.S.S. 感谢 Joaquin F. Rodriguez-Nieva 此前在扩散映射 [12]方面的合作。本文的计算工作在 FASRC Cannon集群上完成，该集群由哈佛大学 FAS科学部研究计算组提供支持。

附录A：变分试探态：受限玻尔兹曼机

式（17）中的变分试探态是一种进一步受限的受限玻尔兹曼机（受限玻尔兹曼机），最初由参考文献 58 提出。受限玻尔兹曼机是玻尔兹曼机的一种受限类型，其“能量”函数 $E_{\mathrm{RBM}}(\boldsymbol{\sigma}, \boldsymbol{h}; \Lambda)$ 依赖于网络参数 Λ，其中 $\boldsymbol{\sigma}$ 是物理自旋，$\boldsymbol{h} = \{h_1, h_2, \cdots, h_N \mid h_i = \pm 1\}$ 是隐藏自旋（或隐藏神经元），二者均为伊辛变量。参数 Λ定义了物理自旋与隐藏自旋之间的耦合强度。受限玻尔兹曼机的限制在于，耦合仅存在于物理自旋 σ_i 和隐藏自旋 h_j 之间，耦合强度为$-w_{ij}$，因此“能量”函数的形式为$E_{\mathrm{RBM}}(\boldsymbol{\sigma}, \boldsymbol{h}; \Lambda) = -\sum_i a_i\sigma_i - \sum_i b_i h_i - \sum_{ij} w_{ij}\sigma_i h_j$。它是一种生成式神经网络，旨在基于玻尔兹曼因子对概率分布 $\mathbb{P}$ 进行建模，

$$\mathbb{P}(\boldsymbol{\sigma}; \Lambda) = \frac{1}{Z}\sum_{\boldsymbol{h}} e^{-E_{\mathrm{RBM}}(\boldsymbol{\sigma}, \boldsymbol{h}; \Lambda)}, \tag{A1a}$$

$$\text{normalization} \quad Z = \sum_{\boldsymbol{\sigma}, \boldsymbol{h}} e^{-E_{\mathrm{RBM}}(\boldsymbol{\sigma}, \boldsymbol{h}; \Lambda)}. \tag{A1b}$$

对于量子波函数振幅 $\psi(\boldsymbol{\sigma}; \Lambda)$ 的建模任务，可以将参数 Λ 扩展为复数，从而将受限玻尔兹曼机用作变分试探态。

进一步将参数限制为环面码模型中连接面元与星形几何的层间连接 [参见图 2(c)]，并令所有参数 Λ 均为纯虚数，我们便可恢复式（17）中的试探态$\tilde{}$（相差一个归一化因子 Z），

$$\begin{aligned}\psi(\boldsymbol{\sigma}; \Lambda) &= \frac{1}{\tilde{Z}}\sum_{X=P,S}\sum_{h_X=\pm 1} e^{-i\sum_X (w_{Xj}\sigma_j + b_X)h_X}, \\ &= \frac{1}{\tilde{Z}}\prod_{X=P,S}\cos\Big(\sum_{j\in X} w_{Xj}\sigma_j + b_X\Big).\end{aligned} \tag{A2}$$

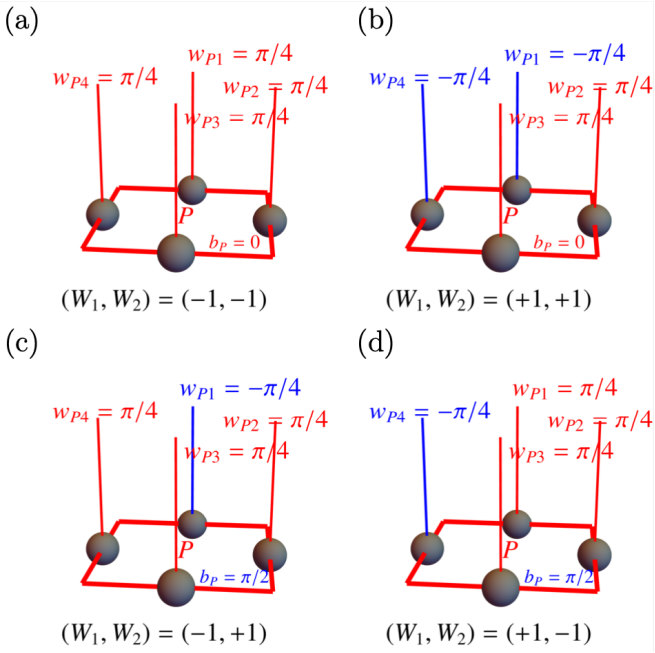

图 7. 环面码的四个基态在环算符本征基 [式（A4）] 中的受限玻尔兹曼机表示 W_1，W_2 式（A3a）。

cos(·) 因子来自对隐神经元求和；由于连接受到限制，试探态可分解为各个面元（星形）项的乘积。基于受限玻尔兹曼机试探态估算波函数的物理可观测量，需要进行蒙特卡罗抽样，这一过程将在附录 B中讨论。

1. 不同拓扑扇区中的基态表示

将式（16）中的环面码模型置于环面几何中，定义环算符很有用，

$$\hat{W}_1 = \prod_{i\in \bar{l}_x} \hat{s}_i^x, \quad \hat{W}_2 = \prod_{i\in \bar{l}_y} \hat{s}_i^x, \tag{A3a}$$

$$\hat{V}_1 = \prod_{i\in l_x} \hat{s}_i^z, \quad \hat{V}_2 = \prod_{i\in l_y} \hat{s}_i^z, \tag{A3b}$$

其中 $l_{x,\,y}$ 是沿 x、y 方向的非可缩回路，而 $l_{x,\,y}$ 则是在对偶晶格上的类似回路。注意，沿两个方向的环算符彼此不对易，例如 $\left[\hat{W}_1, \hat{V}_2\right] \neq 0$ 和 $\left[\hat{W}_2, \hat{V}_1\right] \neq 0$。然而，由于哈密顿量与这些环算符对易，即$\left[\hat{W}_{1,\,2}, \hat{H}_{\mathrm{tc}}\right] = \left[\hat{V}_{1,\,2}, \hat{H}_{\mathrm{tc}}\right] = 0$，因此基态子空间是四重简并的，并由环算符的特征向量张成。

假设我们定义本征基态 $|\psi_i\rangle$ ($\hat{i} = 0,\ 1,\ 2,\ 3$)

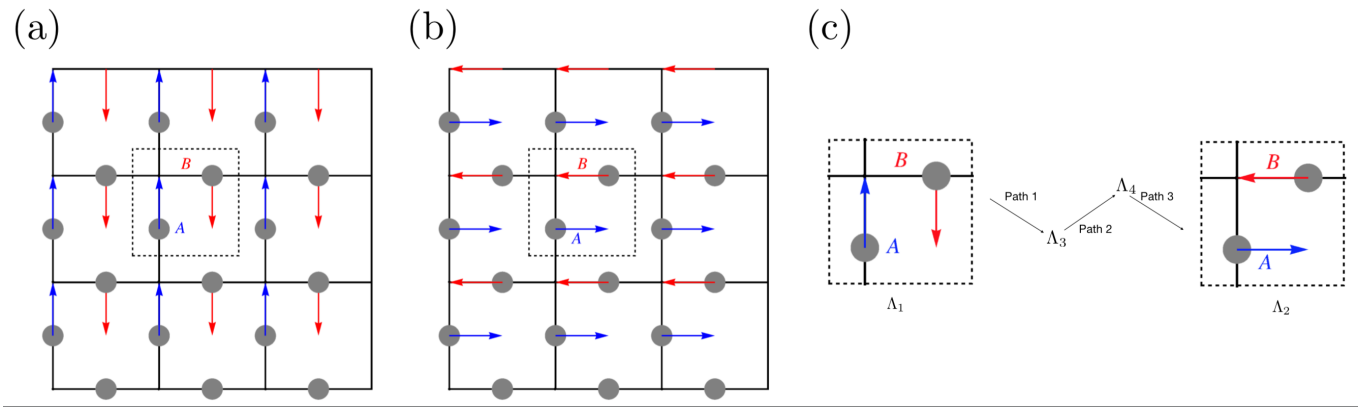

图 8. (a-b) 极化态的两种受限玻尔兹曼机表示，见式 (A8)。(c) 连接 (a-b) 中两个自旋表示的一条路径，具体见表 I。

跨越 $\mathcal{L}$,

$$\hat{W}_1|\psi_0\rangle = |\psi_0\rangle, \quad \hat{W}_2|\psi_0\rangle = |\psi_0\rangle, \tag{A4a}$$
$$\hat{W}_1|\psi_1\rangle = |\psi_1\rangle, \quad \hat{W}_2|\psi_1\rangle = -|\psi_1\rangle, \tag{A4b}$$
$$\hat{W}_1|\psi_2\rangle = |\psi_2\rangle, \quad \hat{W}_2|\psi_2\rangle = -|\psi_2\rangle, \tag{A4c}$$
$$\hat{W}_1|\psi_3\rangle = -|\psi_3\rangle, \quad \hat{W}_2|\psi_3\rangle = -|\psi_3\rangle. \tag{A4d}$$

公式（A2）中的受限玻尔兹曼机试探态可以表示 $W_{1,2}$ 的本征态，其特征值为 $(W_1, W_2) = (\pm 1, \pm 1)$。参考文献 [58] 给出了使用参数表示 $|\psi_3\rangle$ 的一种表示，

$$w_{Pj} = \frac{\pi}{4}, \quad b_P = 0, \quad w_{Sj} = \frac{\pi}{2}, \quad b_S = 0. \tag{A5a}$$

对于沿 x 和 y 方向的格点数为奇数的系统，可类似地通过将与星相关的权重固定为 $w_{Sj}=0$, $b_S=0$来构造其余三个简并态。然后，如图 7所示，通过改变 w_{Pj} 和 b_P 即可选出这四种状态。

2. 极化相中的网络参数冗余

在第 III节中，我们确定了一组规范变换，见式 (18)，它们使由式 (17) 中的受限玻尔兹曼机试探态参数化的一般波函数在相差一个全局相位 [式 (13)]的情况下保持不变。计算相似性度量 S_n时应考虑此类规范变换。此外，我们通过数值验证发现，对于在精确环面码波函数附近生成的状态，在通过式 (19) 中的 S_n 明确消除这类冗余后， S_n 是量子度量 S_q 的良好替代指标。然而，正如在讨论大 h 极限时所提及的，还存在一些态特定冗余，它们通常无法通过式 (18) 中的规范变换相互关联。

让我们在此以极化态 $|\Psi\rangle = |1, \cdots, 1\rangle_z$ 为例说明这类冗余性；该状态在 z基底中所有自旋均指向上方。注意，波函数试探态中 cos(·) 因子的数量与自旋数量相同。因此，我们可以通过将每个自旋分别分配给一个因子，并选择权重以确保所有自旋均指向上方，从而定义一种“覆盖”。任何这样的“覆盖”都是极化态的一种有效表示。例如，一种表示为:

$$b_P = b_S = -\frac{\pi}{4}, \quad w_{Sj} = \begin{cases} \frac{\pi}{4}, & j = j_s(S), \\ 0, & \text{otherwise}, \end{cases} \quad \text{and} \quad w_{Pj} = \begin{cases} \frac{\pi}{4}, & j = j_n(P), \\ 0, & \text{otherwise}. \end{cases} \tag{A6}$$

其中， $j_s(S)$ 表示星形 S 中最靠南的自旋， $j_n(P)$ 表示面元中最靠北的自旋P [，见图 8（a）]。自旋的任何此类覆盖都对应于一个极化态。例如，执行一次“旋转”会得到图 8（b）中的另一种覆盖。实际上，由于局域-z 基底中的大多数振幅为 0，波函数振幅所受的约束非常少，因此存在一组连续的权重来表示该极化态，所以对于完全极化态而言存在无限多种冗余表示。

为了说明这一点，我们来考虑一个最简单的例子，即只有两个自旋 [图中方框标出的区域。8(c)] 采用相同的受限玻尔兹曼机试探态，该试探态可轻松推广到更多自旋的情况。对于两个自旋，这种试探态表示为:

$$\psi_\Lambda(\sigma_A, \sigma_B) = \cos(b_S + w_{SA}\sigma_A + w_{SB}\sigma_B)\cos(b_P + w_{PA}\sigma_A + w_{PB}\sigma_B), \tag{A7}$$

其中，权重 $\Lambda = \{\Lambda_S = \{b_S, w_{SA}, w_{SB}\}, \Lambda_P = \{b_P, w_{PA}, w_{PB}\}\}$ 和 $\Lambda_{Xj} \in [0, \pi)$ 对于 $X = S$或 P完全确定双量子比特物理态。例如，以下两组权重 [Λ_1 和 Λ_2 在图中

路径 1 $\Lambda_1 \to \Lambda_3$	$w_{SB} = b_S + w_{SA} - \frac{\pi}{2}$ w_{SA}: $[0, \frac{\pi}{4})$ w_{SB}: $[\frac{\pi}{4}, -\frac{\pi}{4})$ b_S: $[-\frac{\pi}{4}, 0)$	Λ_P 固定 $w_{PA} = \frac{\pi}{4}, w_{PB} = 0, b_P = -\frac{\pi}{4}$	乘积 $\psi = \psi_S \times \psi_P$
$\cos(b_X + w_{XA} + w_{XB})$	$\neq 0$ if $b_S + w_{SA} \neq \frac{n}{2}\pi$, $n \in \mathbb{Z} \to 0 \to 1$	1	$\to 0 \to 1$
$\cos(b_X + w_{XA} - w_{XB})$	0		0 ✓
$\cos(b_X - w_{XA} + w_{XB})$	$\cos(2b_S - \frac{\pi}{2}) \to 0$	0	0 ✓
$\cos(b_X - w_{XA} - w_{XB})$		0	0 ✓
路径 2 $\Lambda_3 \to \Lambda_4$	Λ_S 固定 $w_{SA} = \frac{\pi}{4}, w_{SB} = -\frac{\pi}{4}, b_S = 0$	$w_{PB} = b_P - w_{PA} + \frac{\pi}{2}$ w_{PA}: $[\frac{\pi}{4}, 0]$, w_{PB}: $[0, \frac{\pi}{4}]$, $b_P = -\frac{\pi}{4}$	
$\cos(b_X + w_{XA} + w_{XB})$	1	1	1
$\cos(b_X + w_{XA} - w_{XB})$	0	$\cos(2w_{PA} - \frac{\pi}{2}) \to 0$	0 ✓
$\cos(b_X - w_{XA} + w_{XB})$	0		0 ✓
$\cos(b_X - w_{XA} - w_{XB})$		0	0 ✓
路径 3 $\Lambda_4 \to \Lambda_2$	$w_{SB} = -b_S + w_{SA} + \frac{\pi}{2}$ $w_{SA} = \frac{\pi}{4}, w_{SB}$: $(-\frac{\pi}{4}, 0], b_S$: $(0, -\frac{\pi}{4}]$	Λ_P 固定 $w_{PA} = 0, w_{PB} = \frac{\pi}{4}, b_P = -\frac{\pi}{4}$	
$\cos(b_X + w_{XA} + w_{XB})$	1	1	1
$\cos(b_X + w_{XA} - w_{XB})$		0	0 ✓
$\cos(b_X - w_{XA} + w_{XB})$	0		0 ✓
$\cos(b_X - w_{XA} - w_{XB})$		0	0 ✓

表 I. 从 Λ_1 到 Λ_2 的路径由三个步骤组成。路径 1（$\Lambda_1 \to \Lambda_3$）除了点 $w_{SB} = -\frac{\pi}{4}$ $b_S = 0$ $w_{SA} = 0$ 处外都是光滑的，在该点波函数消失。第一行中，红色箭头先递减至，然后再递增至 1，进一步说明了这一点。路径 2 和路径 3 均为光滑路径。最后一列显示，波函数 ψ 沿路径始终保持在极化态。

图 8(c)] 均可参数化极化态：

$$\Lambda_1 = \{b_S = -\frac{\pi}{4}, w_{SA} = 0, w_{SB} = \frac{\pi}{4}, b_P = -\frac{\pi}{4}, w_{PA} = \frac{\pi}{4}, w_{PB} = 0\}, \tag{A8a}$$

$$\Lambda_2 = \{b_S = \frac{\pi}{4}, w_{SA} = \frac{\pi}{4}, w_{SB} = 0, b_P = -\frac{\pi}{4}, w_{PA} = 0, w_{PB} = \frac{\pi}{4}\}, \tag{A8b}$$

$$\psi_{\Lambda_{1,2}} = \begin{cases} \sigma_A = \sigma_B = 1, \\ 0, \quad \text{否则。} \end{cases} \tag{A8c}$$

现在，为了说明连续冗余，我们在参数空间中构造一条路径，从 Λ_1 到 Λ_2。该路径由三个步骤组成 [图 8(c)],

$$\Lambda_1 \xrightarrow{\text{path 1}} \Lambda_3 \xrightarrow{\text{path 2}} \Lambda_4 \xrightarrow{\text{path 3}} \Lambda_2, \tag{A9}$$

其中，中间参数为：

$$\Lambda_3 = \{b_S = 0, w_{SA} = \frac{\pi}{4}, w_{SB} = -\frac{\pi}{4}, b_P = -\frac{\pi}{4}, w_{PA} = \frac{\pi}{4}, w_{PB} = 0\}, \tag{A10}$$

$$\Lambda_4 = \{b_S = 0, w_{SA} = \frac{\pi}{4}, w_{SB} = -\frac{\pi}{4}, b_P = -\frac{\pi}{4}, w_{PA} = 0, w_{PB} = \frac{\pi}{4}\}. \tag{A11}$$

沿着每个路径分量（表 I 中称为路径 1 至 3），改变 S（或 P）的参数，同时保持其他参数不变，并始终处于精确极化态。除路径 1 上一个奇异点外，该路径是连续的；在该奇异点处，波函数在 $\Lambda_{\text{singular}} = \{b_S = 0, w_{SA} = \frac{\pi}{4}, w_{SB} = -\frac{\pi}{4}, b_P = -\frac{\pi}{4}, w_{PA} = \frac{\pi}{4}, w_{PB} = 0\}$ 处消失。

3. 解析特殊冗余

在附录 A2中，我们明确展示了给定一个极化态时可能存在大量冗余。因此，为了使正文简洁，我们采用式（ S_q ）中的直接重叠 10作为相关度量

在有限场值下。正如正文所述，缓解相似性度量中冗余的一种直接方法是 S_n 在网络参数的式（19）中对其进行补充，加入直接重叠。通过结合这两种度量，我们能够降低计算直接重叠的计算成本

由于相似性易于计算，因此可将计算成本降低一定比例。更具体地说，我们定义一种混合度量 S_m ：将相似性度量~对中的随机一部分（比例为 f）$\{l, l'\}$ 替换为重标度重叠度量 S_q，使得：

$$S_m(l,l') = \begin{cases} \widetilde{S}_q(l,l') & \text{with probability } f, \\ S_n(l,l') & \text{with probability } 1-f. \end{cases} \tag{A12}$$

对重叠度量 S_q 进行如下重标度是必要的，因为我们希望将这两种度量置于同等地位，具体如下：

$$\widetilde{S}_q = \frac{S_q - n_q}{m_q - n_q} \cdot (m_n - n_n) + n_n, \tag{A13a}$$

$$m_q = \max(S_q), \quad n_q = \min(S_q), \tag{A13b}$$

$$m_n = \max(S_n), \quad n_n = \min(S_n). \tag{A13c}$$

例如，我们看到，重标度重叠的最小值与相似度的最小值相同~$\min(S_q) = \min(S_n)$。

在图 9中，我们展示了采用混合度量时，若以一定比例 $f = 0.4$替代，我们的扩散映射（DM）算法能够识别不同场值下超选择扇区的存在（较小场值对应的蓝色阴影区域$h = 0.475$ 和 $h = 0.55$）与缺失（$h = 0.7$），这与直接重叠算法的预测一致（见图 6）。我们注意到，在混合度量的情况下，扩散映射是一种自然的技术，因为该算法寻找的是连通性；而核主成分分析则无法识别这种转变（因为根据 S_n，有一定比例的波函数对被错误地视为彼此不相似，主成分分析的前几个分量仍显示出四个彼此分离的簇，直至磁场达到最大值，$h = 1$）。

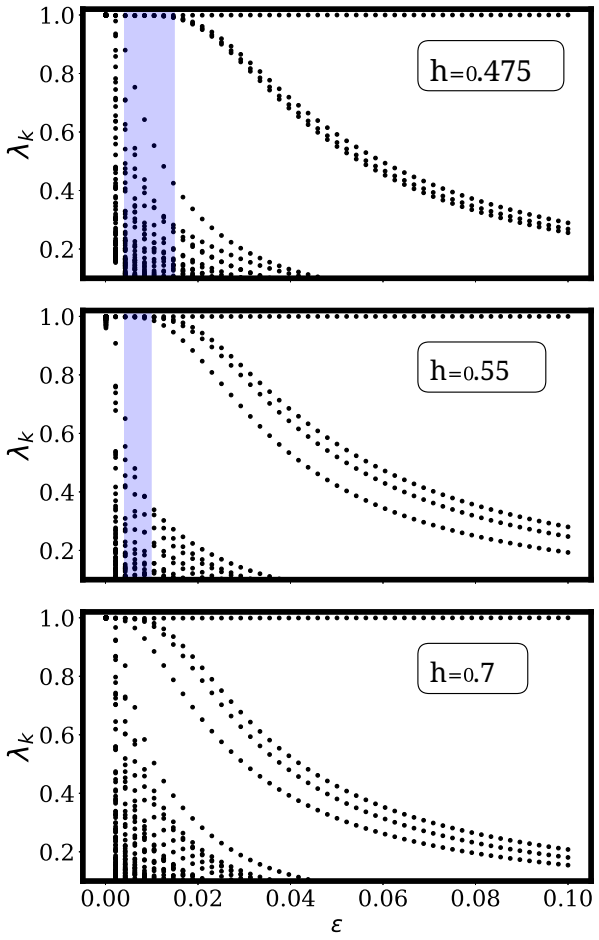

图 9 $h = 0.475, 0.55, 0.7$，$T = 0.3$，在使用混合相似性度量时 S_m 其占比为 $f = 0.4$ 见方程（A12）。蓝色阴影区域突出显示了存在一段 ϵ其简并本征值与衰减本征值之间存在谱隙，表明底层存在超选择扇区。随着场值接近转变场 h_c，这类区域的范围逐渐缩小，并在高场 $h = 0.7$下消失，表明不存在扇区。不同场值下的 DM谱，

附录B：使用变分蒙特卡罗进行优化

为求得基态 $|\Psi(\Lambda^0)\rangle \propto \sum_{\boldsymbol{\sigma}} \psi(\boldsymbol{\sigma};\, \Lambda^0)\,|\boldsymbol{\sigma}\rangle$，我们希望最小化能量期望值 $\langle E\rangle = \hat{}\langle\Psi|\, H\,|\Psi\rangle \,/\, \langle\Psi|\Psi\rangle$（本节省略变分参数$\Lambda^0$），该值根据变分原理以基态能量为下界。精确计算 $\langle E\rangle_{\text{exact}}$ 的代价很高，因为随着系统规模增大，求和需要遍历数量呈指数增长的自旋构型 $\boldsymbol{\sigma}$。这里我们使用变分蒙特卡罗（VMC）重要性抽样算法来估计这类期望值。其思路是计算不同构型之间的相对概率，并从真实的波函数概率密度 $|\psi(\boldsymbol{\sigma})|^2$中抽样，而无需计算所有 $\boldsymbol{\sigma}$对应的 $|\psi(\boldsymbol{\sigma})|^2$。执行该算法时，我们先初始化 M 个随机构型$\{\boldsymbol{\sigma}_i\}_{i=1}^{M}$，再基于前一个构型对每个构型进行随机游走，从而形成 M 条马尔可夫链。

具体而言，使用 Metropolis–Rosenbluth 算法 [74] 来提议下一个构型 $\boldsymbol{\sigma}_i'$ 根据函数$g(\boldsymbol{\sigma}'|\boldsymbol{\sigma})$，它与 c_i 局部相连。对于环面码模型，我们使用两种提议：自旋翻转和顶点翻转。这里，我们假设提出自旋翻转的概率为 p，提出顶点翻转的概率相应为 $1-p$，且在所有格点上发生的概率均相同：

$$g(\boldsymbol{\sigma}'|\boldsymbol{\sigma}) = \begin{cases} \frac{p}{n_s}, & \text{for spin flips} \\ \frac{1-p}{n_v}, & \text{for vertex flips} \end{cases} \tag{B1}$$

其中，n_s 和 n_v 分别是所有可能的自旋翻转和顶点翻转的数量。$\boldsymbol{\sigma}'$ 的接受概率由一个概率决定，

$$\mathbb{P}_{\text{accept}}(\boldsymbol{\sigma} \to \boldsymbol{\sigma}') = \min\left(|\frac{\psi(\boldsymbol{\sigma}')}{\psi(\boldsymbol{\sigma})}|^2,\, 1\right). \tag{B2}$$

随机游走将重复足够长的时间，使链 $\Sigma_{\rm MC} = \{\boldsymbol{\sigma}_f\}_{i=b}^{M}$ 末端的最终构型 $|\psi$近似于从概率分布 $\boldsymbol{\sigma}$ $(|^2)$ b 中抽取的样本。每条链中会舍弃一定数量的随机游走者，以减少链初始化带来的偏置。随后，可观测量 O 的期望值ˆ为，

$$\langle \hat{O} \rangle_{\rm MC} = \frac{\sum_{\boldsymbol{\sigma}} \psi(\boldsymbol{\sigma})^* \langle \boldsymbol{\sigma}|\hat{O}|\Psi\rangle}{\sum_{\boldsymbol{\sigma}} |\psi(\boldsymbol{\sigma})|^2}, \tag{B3a}$$

$$= \frac{\sum_{\boldsymbol{\sigma}} |\psi(\boldsymbol{\sigma})|^2 \frac{\langle \boldsymbol{\sigma}|\hat{O}|\Psi\rangle}{\psi(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}} |\psi(\boldsymbol{\sigma})|^2}, \tag{B3b}$$

$$= \frac{1}{M} \sum_{\boldsymbol{\sigma} \in \Sigma_{\rm MC}} \frac{\langle \boldsymbol{\sigma}|\hat{O}|\Psi\rangle}{\psi(\boldsymbol{\sigma})}. \tag{B3c}$$

ˆ 将算符 O 的局域值定义为，

$$O_{\rm loc} = \frac{\langle \boldsymbol{\sigma}|\hat{O}|\Psi\rangle}{\psi(\boldsymbol{\sigma})}, \tag{B4}$$

然后，蒙特卡罗估计即为马尔可夫链中各个ˆ局域值的平均值：$\langle O \rangle_{\rm MC} = \frac{1}{M} \sum_{\boldsymbol{\sigma} \in \Sigma_{\rm MC}} O_{\rm loc}$ 。

接下来，为了最小化 $\langle E \rangle$，我们可以用局域能量 $E_{\rm loc}$ 和波函数振幅导数 D_i来计算其相对于权重 Λ^0 的梯度：

$$\partial_{\Lambda_i} \langle E \rangle = \langle E_{\rm loc} D_i \rangle - \langle E_{\rm loc} \rangle \langle D_i \rangle \tag{B5a}$$

$$E_{\rm loc} = \frac{\langle \boldsymbol{\sigma}| H |\Psi\rangle}{\psi(\boldsymbol{\sigma})}, \quad D_i = \frac{\partial_{\Lambda_i} \psi(\boldsymbol{\sigma})}{\psi(\boldsymbol{\sigma})} \tag{B5b}$$

最后，我们使用学习率为 λ的梯度下降，

$$\Lambda_i \to \Lambda_i - \lambda \partial_{\Lambda_i} \langle E \rangle, \tag{B6}$$

以最小化能量期望值。梯度下降通过使用自适应 Adam 优化器来执行 [75]。我们重复这一训练步骤，直至达到经验收敛。

注意，受限玻尔兹曼机试探态可能会陷入局部极小值。为了找到环面码基态，我们将网络参数初始化为接近式（A5）中的解析解。

1. 保真度

为了在步长为 Δh 的有限场值 h 下求得近似基态，我们将权重初始化为前一场值 $h - \Delta h$ 对应的权重，然后使用当前优化得到的权重作为下一步 $h + \Delta h$ 的初始化。判断量子相变的一个有效方法是检查保真度 $\mathcal{F}(h)$，其定义如下 as,

$$\mathcal{F}(h) = |\langle \psi(h) | \psi(h + \Delta h) \rangle|^2. \tag{B7}$$

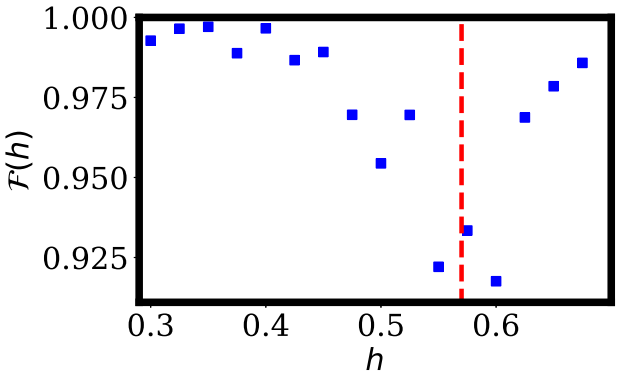

图 10. 保真度 $\mathcal{F}$ 随场 h的变化。红色虚线用于引导视线，保真度的低谷表明临界场值 $h_c \simeq 0.57$。

临界场 h_c 通过保真度的谷值来确定，这表明基态波函数发生了突变。对于受限玻尔兹曼机试探态，得到的场值为 $h_c \simeq 0.57$ (图 10 中虚线所示)。请注意，按照参考文献 59 中的方法，在试探态中加入环路期望值可以得到更精确的场值。

附录 C：系综生成

采用第1节所述的算法，我们可以通过设置超参数 $T = 0.1,\ 0.3,\ \mathbf{1},$ 生成与初始优化参数有所偏离的系综。系综的其他超参数选项包括独立链的数量 $k = 2$、每条链的长度 $n = 250$以及保留的样本数量 $m = n$。我们使用的参数提议函数以 p_m 的概率随机施加负号，或在单个自旋格点随机添加局部噪声 $\jmath$。更精确地说，

$$f(\Lambda, \xi) = \begin{cases} f_{-,\jmath}, & \text{with probability} : p_m, \\ f_{\text{local},\jmath}, & \text{with probability} : 1 - p_m, \end{cases} \tag{C1a}$$

$$f_{-,\jmath} = \begin{cases} -(\Lambda)_i, & i \in \jmath \\ (\Lambda)_i, & i \notin \jmath \end{cases} \tag{C1b}$$

$$f_{\text{local},\jmath} = \begin{cases} \text{uniform}(0, \xi) + (\Lambda)_{\rm i}, & \text{i} \in \jmath \\ (\Lambda)_i, & i \notin \jmath \end{cases} \tag{C1c}$$

在精确的环面码态中，$f_{-,\ \jmath}$ 对应于作用在σ_x 位置的算符 $\jmath$ ，用于产生一对 m 粒子。在平庸相中，取决于态的参数化方式，$f_{-,\ \jmath}$ 可能对应于位置 $\jmath$的单个自旋翻转。超参数选为 $p_m = 0.3$和 $\xi = 0.2$。在图11中，我们通过计算不同场值下的环路期望值$\langle W_{\jmath} \rangle$ 来展示这些系综。

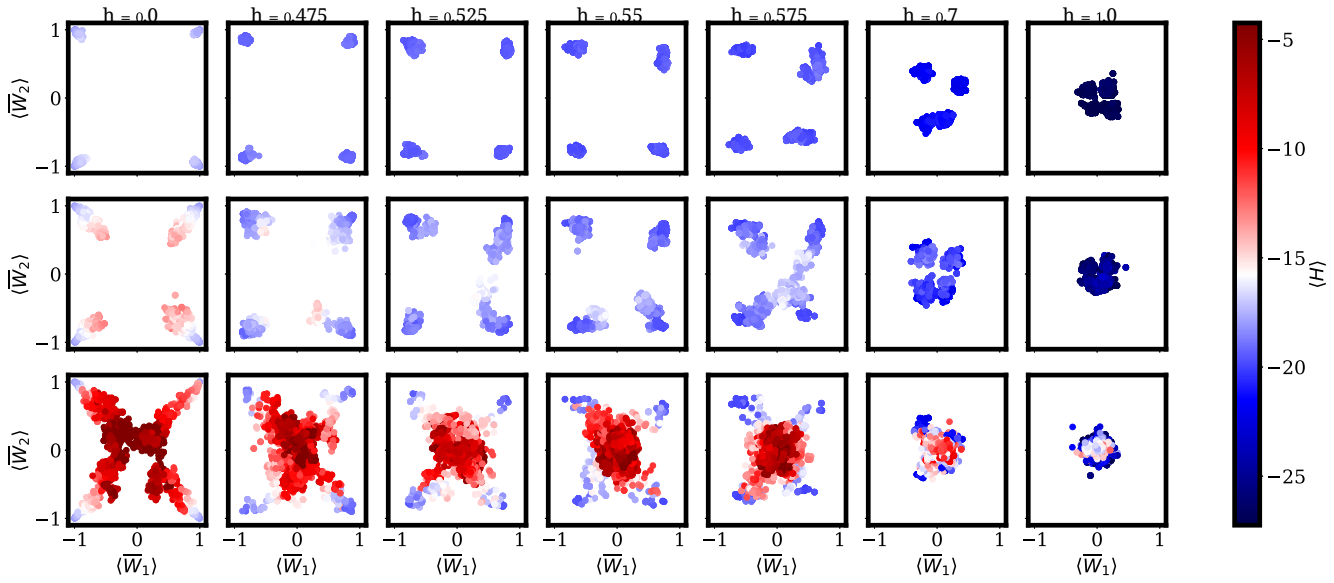

图 11. 展示了在 $N = 18$ 个自旋下，不同参数 T 和场 h 所对应的扩散过程。当 T 和 h 较小时，环路期望值 $\langle W_{1,2}\rangle$ 在二维平面上形成四个不同的簇。当所有场下 $T = 1.$ 较大，或在较高场 $h > 0.57$ 下 $T = 0.3$ 处于中间值时，这些簇会发生“扩散”，拓扑序随之消失。此类“扩散”过程可通过能量期望值 $\langle H\rangle$ 的颜色编码进行可视化。

[1] Pankaj Mehta、Marin Bukov、Ching-Hao Wang、Alexan- dre G. R. Day、Clint Richardson、Charles K. Fisher 和 David J. Sc hwab，《面向物理学家的机器学习入门：高偏置、低方差》，Physics Reports A High-Bias, Low-Variance Introduction to Machine Learn- ing for Physicists，810，1–124（2019）。[2]Giuseppe Carleo、Ignacio Cirac、Kyle Cranmer、Lau- rent Daudet、Maria Schuld、Naftali Tishby、Leslie Vogt- Maranto 和 Lenka Zdeborov´ a，《机器学习与物理科学》，Rev. Mod. Phys. 91，045002（2019）。[3]Sankar Das Sarma、Dong-Ling Deng 和 Lu-Ming Duan，《机器学习遇上量子物理学》，Physics Today **72**，48 54 2019 1903 03516 –（），[ ] arXiv：。physics.pop-ph。[4] Roger G. Melko、Giuseppe Carleo、Juan Carrasquilla 和 J. Ignacio Cirac，《量子物理学中的受限玻尔兹曼机》，Nature Physics 15，887–892（2019）。[5] Juan Carrasquilla，《量子物质的机器学习》，Advances in Physics: X 5，1797528（2020）。[6]Juan Carrasquilla 和 Giacomo Torlai，《如何利用神经网络研究量子多体物理学》，PRX Quantum 2，040201（2021）。[7] Anna Dawid、Julian Arnold、Borja Requena、Alexan- der Gresch、Marcin P-lodzie´n、Kaelan Donatella、Kim A. Nicoli、Paolo Stornati、Rouven Koch、Miriam B¨uttner、Robert Oku-la、Gorka Mu˜noz-Gil、Rodrigo A. Vargas- Hern´andez、Alba Cervera-Lierta、Juan Carrasquilla、Ve- dran Dunjko、Marylou Gabri´e、Patrick Huembeli、Evert van Nieuwenburg、Filippo Vicentini、Lei Wang、Sebas- tian J. Wetzel、Giuseppe Carleo、Eliška Greplov´a、Ro- man Krems、Florian Marquardt、Micha-l Tomza、Maciej Lewenstein 和 Alexandre Dauphin，《现代应用——机器学习在量子科学中的应用，”（2022），arXiv:2204.04198 [cond-mat, physics:quant-ph]. [8] Juan Carrasquilla 和 Roger G. Melko，“机器学习物质相”，Nature Physics 13，431–434（2017）。[9] Pengfei Zhang、Huitao Shen 和 Hui Zhai，“利用神经网络学习拓扑不变量”，Phys. Rev. Lett. 120, 066401（2018）。[10] Yi Zhang 和 Eun-Ah Kim，“用于机器学习的量子环路拓扑”，Phys. Rev. Lett. 118, 216401（2017）。[11] Matthew J. S. Beach、Anna Golubeva 和 Roger G. Melko，“在 Kosterlitz–Thouless 相变处机器学习涡旋”，Phys. Rev. B 97, 045207（2018）。[12] Joaquin F. Rodriguez-Nieva 和 Mathias S. Scheurer，“通过无监督机器学习识别拓扑序”，Nature Physics 15, 790–795（2019）。[13] Japneet Singh、Mathias S. Scheurer 和 Vipul Arora，“用于自旋系统采样和相变指示的条件生成模型”，SciPost Phys. 11, 043 2021( )。[14] Y.-H. Tseng 和 F.-J. Jiang，“Berezin- skii–kosterlitz–thouless 相变——基于基准测试的通用神经网络研究”，Results in Physics 33, 105134 2022( )。[15] Eliska Greplova、Agnes Valenti、Gregor Boschung、Frank Sch¨afer、Niels L¨orch 和 Sebastian D Huber，“使用预测模型无监督识别拓扑相变”，New J. Phys. 22, 045003（2020）。[16] Yi Zhang、Roger G. Melko 和 Eun-Ah Kim，“利用准粒子统计机器学习 2 量子自旋液体”，Phys. Rev. B 96, 245119（2017）。[17] Hsin-Yuan Huang、Richard Kueng、Giacomo Torlai、Vic-

tor V. Albert 和 John Preskill，“量子多体问题的可证明高效机器学习”，Sci- ence 377, eabk3333（2022）。[18] Nicolas Sadoune、Giuliano Giudici、Ke Liu 和 Lode Pollet，“多量子比特系统中物质相的无监督可解释学习”，（ 2022 2208 08850 ），[ ] arXiv: . cond-mat, physics:quant-ph .[19] Alex Cole、Gregory J. Loges 和 Gary Shiu，“利用持久同调进行可解释的相检测与分类”，arXiv 电子印本，arXiv:2012.00783（2020），arXiv:2012.00783 [cond-mat.stat-mech].[20] Dan Sehayek 和 Roger G. Melko，“ $\mathbb{Z}_2$ 规范理论的持久同调 ”，Phys. Rev. B 106, 085111（2022），arXiv:2201.09856 [cond-mat, physics:hep-th].[21] Niklas K¨aming、Anna Dawid、Korbinian Kottmann、Maciej Lewenstein、Klaus Sengstock、Alexandre Dauphin 和 Christof Weitenberg，“基于实验数据无监督机器学习拓扑相变”，《机器学习：科学与技术》2, 035037 2021( )。[22] Chi-Ting Ho 和 Daw-Wei Wang，“通过自监督机器学习方法稳健识别拓扑相变”，《新物理学杂志》23, 083021（2021）。[23]Min-Ruei Lin、Wan-Ju Li 和 Shin-Ming Huang，“基于四元数的机器学习在拓扑量子系统中的应用”，arXiv 电子印本（2022），arXiv:2209.14551 [quant-ph].[24] Gilad Margalit、Omri Lesser、T. Pereg-Barnea 和 Yu- val Oreg，“受重整化群启发的用于计算拓扑不变量的神经网络”，Phys. Rev. B 105, 205139（2022）。[25] Sungjoon Park、Yoonseok Hwang 和 Bohm-Jung Yang，“ 利用拓扑数据分析无监督学习拓扑相图 ”，Phys. Rev. B 105, 195115（2022）。[26]Ming-Chiang Chung、Tsung-Pao Cheng、Guang-Yu Huang 和 Yuan-Hong Tsai，“从纠缠角度深度学习二维手性 p-波超导体的拓扑相变”，Phys. Rev. B 104, 024506（2021）。[27] Yuan-Hong Tsai、Kuo-Feng Chiu、Yong-Cheng Lai、Kuan-Jung Su、Tzu-Pei Yang、Tsung-Pao Cheng、Guang-Yu Huang 和 Ming-Chiang Chung，“从纠缠方面深度学习拓扑相变：一种无监督方法”，Phys. Rev. B 104, 165108（2021）。[28] Alejandro Jos´e Ur´ıa- ´Alvarez、Daniel Molpeceres-Mingo 和 Juan Jos´e Palacios，“利用纠缠谱深度学习无序拓扑绝缘体”，arXiv 电子印本，arXiv:2201.13306（2022），arXiv:2201.13306 [cond-mat.dis-nn].[29] Paolo Molignini、Antonio Zegarra、Evert van Nieuwen- burg、R. Chitra 和 Wei Chen，“一种基于局部曲率、用于相互作用拓扑绝缘体的监督学习算法”，SciPost Phys. 11, 073（2021）。[30]Simone Tibaldi、Giuseppe Magnifico、Davide Vodola 和 Elisa Ercolessi，“从单粒子关联函数无监督和监督学习相互作用拓扑相”，arXiv 电子印本（2022），arXiv:2202.09281 [cond-mat.supr-con].[31] Andrea Tirelli 和 Natanael C. Costa，“通过拓扑数据分析学习量子相变”，Phys. Rev. B 104, 235146（2021）。

[32]R. R. Coifman、S. Lafon、A. B. Lee、M. Maggioni、B. Nadler、F. Warner 和 S. W. Zucker，《几何扩散作为调和分析与数据结构定义的工具：扩散映射》，美国国家科学院院刊 102，7426–7431（2005）。[33] Boaz Nadler、Stephane Lafon、Ioannis Kevrekidis 和 Ronald Coifman，《扩散映射、谱聚类与福克-普朗克算子的特征函数》，载于 *Advances in Neural Information Processing Systems* 18，第 卷（麻省理工学院出版社，2005）。[34] Boaz Nadler、St´ephane Lafon、Ronald R. Coifman 和 Ioannis G. Kevrekidis，《扩散映射、谱聚类与动力系统的反应坐标》，应用与计算调和分析 21，113–127（2006）。[35] Ronald R. Coifman 和 St´ephane Lafon，《扩散映射》，应用与计算调和分析 21，5–30（2006）。[36] Mathias S. Scheurer 和 Robert-Jan Slager，《无监督机器学习与能带拓扑》，Phys. Rev. Lett. 124，226401（2020）。[37] Yang Long、Jie Ren 和 Hong Chen，《拓扑声子学的无监督流形聚类》，Phys. Rev. Lett. 124，185501（2020）。[38] Li-Wei Yu 和 Dong-Ling Deng，《非厄米拓扑相的无监督学习》，Phys. Rev. Lett. 126，240402（2021）。[39] Yefei Yu、Li-Wei Yu、Wengang Zhang、Huili Zhang、Xiao- long Ouyang、Yanqing Liu、Dong-Ling Deng 和 L. M. Duan，《利用固态自旋对非厄米结态进行无监督学习的实验研究》，npj《量子信息》 **8**，116 2022 2112 13785 [ ] ( )，arXiv: . quant-ph .[40] Yanming Che、Clemens Gneiting、Tao Liu 和 Franco Nori，《通过无监督机器学习揭示拓扑量子相变》，Phys. Rev. B 102，134213（2020）。[41]En-Jui Kuo 和 Hossein Dehghani，《对对称性保护拓扑相变的无监督学习》，arXiv:2111.08747 [cond-mat, physics:quant-ph] （2021），arXiv:2111.08747 [cond-mat，physics:quant-ph]。[42]Eran Lustig、Or Yair、Ronen Talmon 和 Mordechai Segev，《利用流形学习识别实验中的拓扑相变》，Phys. Rev. Lett. 125，127401（2020）。[43]Alexander Lidiak 和 Zhexuan Gong，《利用扩散映射对量子相变进行无监督机器学习》，Phys. Rev. Lett. 125，225701（2020）。[44]Gaurav Gyawali、Mabrur Ahmed、Eric Aspling、Luke Ellert-Beck 和 Michael J. Lawler，《通过时间平均经典影子揭示强关联系统的微正则相图》，（ 2022），arXiv:2211.01259 [cond-mat，physics:quant-ph]。[45]Apimuk Sornsaeng、Ninnat Dangniam、Pantita Palit- tapongarnpim 和 Thiparat Chotibut，《用于非线性降维的量子扩散映射》，《物理评论A辑》104，052410（2021）。[46]Giuseppe Carleo 和 Matthias Troyer，《利用人工神经网络求解量子多体问题》，《科学》355，602–606（2017）。[47] Xun Gao 和 Lu-Ming Duan，《利用深度神经网络高效表示量子多体态》，《自然-通讯》8

, 662 (2017).

[48]Giuseppe Carleo、Yusuke Nomura 和 Masatoshi I mada，《利用深度神经网络构建量子多体系统的精确表征》，Nat Commun 9, 5322 (2018). [49] Sirui Lu、Xun Gao 和 L.-M. Duan，《利用受限玻尔兹曼机高效表征拓扑有序态》，Phys. Rev. B 99, 155136 (2019).[50]Or Sharir、Amnon Shashua 和 Giuseppe Carleo，《神经张量收缩与深度神经量子态的表达能力》，(2021), 10.48550/arXiv.2103.10293.[51] Jing Chen、Song Cheng、Haidong Xie、Lei Wang 和 Tao Xiang，《受限玻尔兹曼机与张量网络态的等价性》，Phys. Rev. B 97, 085104 2018( ).[52]Yusuke Nomura，《研究神经网络量子态中的网络参数》，(2022), arXiv:2202.01704 [凝聚态物理，物理学：物理学，物理学：量子物理].[53] Dong-Ling Deng、Xiaopeng Li 和 S Das Sarma，《神经网络态中的量子纠缠》，，17 (2017).[54] Zhih-Ahn Jia、Lu Wei、Yu-Chun Wu、Guang-Can Guo 和 Guo-Ping Guo，《浅层和深层量子神经网络态的纠缠面积律》，New J. Phys. 22, 053022 (2020). [55]Song Cheng、Jing Chen 和 Lei Wang，*Information Perspective to Probabilistic Modeling: Boltzmann Machines versus Born Machines*，技术报告 (2017) arXiv:1712.04144 [凝聚态物理，物理学：物理学，物理学：量子物理，统计].[56] Giacomo Torlai、Guglielmo Mazzola、Juan Carrasquilla、Matthias Troyer、Roger Melko 和 Giuseppe Carleo，《神经网络量子态层析》，Nature Physics 14, 447–450 (2018). [57] A. Yu Kitaev，《利用任意子实现容错量子计算》，Annals of Physics 303, 2–30 (2003), arXiv:quant-ph/9707021. [58] Dong-Ling Deng、Xiaopeng Li 和 S. Das Sarma，《机器学习拓扑态》，Phys. Rev. B 96, 195145 2017( ).[59] Agnes Valenti、Eliska Greplova、Netanel H. Lindner 和 Sebastian D. Huber，《作为可解释变分量子态的关联增强型神经网络》，(2021), arXiv:2103.05017 [凝聚态物理，物理学：量子物理].[60]Ze-Pei Cian、Mohammad Hafezi 和 Maissam Barkeshli，《从单个体态基态中提取威尔逊环算符和分数量子统计》，arXiv e-prints (2022), arXiv:2209.14302 [凝聚态物理：强关联电子]. [61] M. B. Hastings，《一维量子系统的面积律》，J. Stat. Mech. 2007, P08024–P08024 (2007). [62] F. Verstraete、M. M. Wolf、D. Perez-Garcia 和 J. I. Cirac，《投影纠缠对态的临界性、面积律与计算能力》，Phys. Rev. 《物理评论快报》96, 220601（2006）。[63] Michael M. Wolf、Frank Verstraete、Matthew B. Hast- ings 和 J. Ignacio Cirac，“量子系统中的面积律：互信息与关联”，《物理评论快报》100（2008），10.1103/Phys-RevLett.100.070502。[64] J. Eisert、M. Cramer 和 M. B. Plenio，“*Colloquium*：纠缠熵的面积律”，Rev. Mod. Phys. 82, 277–306（2010）。[65] Jing-Ling Chen、Libin Fu、Abraham A. Ungar 和 Xian- Geng Zhao，“N 态量子系统中两个状态之间的替代保真度度量”，《物理评论A辑》65, 054304 2002（）。[66] Paulo E. M. F. Mendon.ca、Reginaldo d. J. Napoli- tano、Marcelo A. Marchiolli、Christopher J. Foster 和 Yeong-Cherng Liang，“量子态之间的替代保真度度量”，《物理评论A辑》78, 052330（2008）。[67] Zbigniew Pucha-la 和 Jaros-law Adam Miszczak，“基于超保真度的迹距离界”，《物理评论A辑》79, 024302（2009）。[68] J. A. Miszczak、Z. Pucha-la、P. Horodecki、A. Uhlmann 和 K. Zyczkowski，“作为量子保真度界的次保真度与超保真度”，（2008），arXiv:0805.2037 [quant- ph]。[69] I. S. Tupitsyn、A. Kitaev、N. V. Prokof'ev 和 P. C. E. Stamp，“环面码模型与三维晶格规范希格斯模型相图中的拓扑多临界点”，Phys. Rev. B 82, 085114（2010）。[70]Simon Trebst、Philipp Werner、Matthias Troyer、Kirill Shtengel 和 Chetan Nayak，“拓扑相的崩溃：具有张力的环气体模型中的量子相变”，Phys. Rev. Lett. 98, 070602（2007）。[71]Fengcheng Wu、Youjin Deng 和 Nikolay Prokof'ev，“平行磁场中的环面码模型相图”，Phys. Rev. B 85, 195104（2012）。[72] Michael Schuler、Seth Whitsitt、Louis-Paul Henry、Subir Sachdev 和 Andreas M. L¨auchli，“有限尺寸环面谱中的量子临界点普适特征：窥探高维共形场论算符内容的窗口”，Phys. Rev. Lett. 117, 210401（2016），arXiv:1603.03042 [cond-mat.str-el]。[73]James Bradbury、Roy Frostig、Peter Hawkins、Matthew James Johnson、Chris Leary、Dougal Maclau- rin、George Necula、Adam Paszke、J ake VanderPlas、Skye Wanderman-Milne 和 Qiao Zhang，“JAX：Python+NumPy 程序的可组合变换”，（2018）。[74] Nicholas Metropolis、Arianna W. Rosenbluth、Mar- s hall N. Rosenbluth、Augusta H. Teller 和 Edward Teller，“快速计算机器进行的状态方程计算”，《化学物理学杂志》21, 1087–1092（1953）。[75]Diederik P Kingma 和 Jimmy Ba，“Adam：一种随机优化方法”，arXiv 预印本 arXiv:1412.6980（2014）。