



基于机器学习 J_1 - J_2 反铁磁海森伯自旋链相变点的识别方法

王伟 揭泉林

Identifying phase transition point of J_1 - J_2 antiferromagnetic Heisenberg spin chain by machine learning

Wang Wei Jie Quan-Lin

引用信息 Citation: *Acta Physica Sinica*, 70, 230701 (2021) DOI: 10.7498/aps.70.20210711

在线阅读 View online: <https://doi.org/10.7498/aps.70.20210711>

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

您可能感兴趣的其他文章

Articles you may be interested in

基于机器学习 J_1 - J_2 反铁磁海森伯自旋链相变点的识别方法*

王伟 揭泉林†

(武汉大学物理科学与技术学院, 武汉 430000)

(2021年4月14日收到; 2021年6月27日收到修改稿)

通过序参量来研究量子相变是比较传统的做法, 而从机器学习的角度研究相变是一块全新的领域. 本文提出了先采用无监督学习算法中的高斯混合模型对 J_1 - J_2 反铁磁海森伯自旋链系统的态矢量进行分类, 再使用监督学习算法中的卷积神经网络鉴别无监督学习算法给出的分类点是否是相变点的方法, 并使用交叉验证的方法对学习效果进行验证. 结果表明, 上述机器学习方法可以从基态精确找到 J_1 - J_2 反铁磁海森伯自旋链系统的一阶相变点、无法找到无穷阶相变点, 从第一激发态不仅能找到一阶相变点, 还能找到无穷阶相变点.

关键词: 海森伯 J_1 - J_2 模型, 机器学习, 神经网络, 相变

PACS: 07.05.Mh, 03.67.-a, 03.65.Vf

DOI: 10.7498/aps.70.20210711

1 引言

机器学习算法的改进和更快更强的硬件能够让研究者从庞大的数据中提取出有用的信息, 现在已广泛应用于许多研究领域, 这些发展对物理学也产生了广泛的影响, 特别是在分类量子相变上^[1-10]. 量子相变是系统在绝对零度下发生的一种相变, 它仅仅由系统的哈密顿量变化带来的量子扰动引起^[11,12]. 在量子力学中, 系统的物理状态可以表示为希尔伯特空间的一个矢量, 由于量子态通常已经归一化, 研究者经常使用保真度 (两个量子态的内积) 来研究量子相变, 它能反映两个态之间的相似程度, 保真率反映的是这种变化的速率, 这是一个纯粹几何学上的概念, 分析认为机器学习方法应该也能从态矢量中提取信息来反映系统相变, 其与保真度不同的是, 输出结果是该态矢量与模型训练数据中的所有同类态矢量的相似程度. 研究量子

相变, 除了能进一步丰富人们对低温磁性材料、铁电材料、超导体的认识外, 对其他学科和量子计算技术早期的实践都具有重要意义. 在一个很大的晶格中, 计算序参量来判断相变需要很大的计算量, 也需要有先验知识来构造一个合适的序参量. 机器学习为我们提供了一种新的识别相变点的方法.

Walker 等^[13] 使用无监督机器学习有效识别了材料结构变化的转变温度, Jadrich 等^[14] 研究表明不需要有任何先验知识来构造一个合适的序参量, 使用主成分分析法就能自动发现序参量来分类相, 从而简化了对相的常规分析. Canabarro 等^[15] 使用无监督机器学习识别了轴向次近邻伊辛模型的 3 个主要相. Carrasquilla 等^[16] 研究表明神经网络能够被用来编码物质相和识别强关联多体系统的相变. Ahmadreza 等^[17] 使用卷积神经网络高精度的定位了二维伊辛模型的相变点.

大量研究表明机器学习能够有效识别相变点, 然而当面对一个全新的系统时, 使用无监督学习得

* 国家自然科学基金 (批准号: 1217050658) 资助的课题.

† 通信作者. E-mail: qljie@whu.edu.cn

到的分类点是否是相变点仍需要计算序参量加以佐证, 当不知道相变点时, 也无法选择合适的分类点, 给训练数据集贴上标签, 训练神经网络模型. 这时就需要将非监督学习和监督学习结合起来, 可以先使用无监督学习算法找到分类点, 再使用稍微远离分类点的两侧数据集进行训练. 发现如果无监督学习得到的分类点为相变点, 通过增加分类点一侧的训练数据或者调整训练集的分布区域, 该点位置一般不会发生变化, 若其不是相变点, 其位置必发生变化, 由此, 就可以不需要任何先验知识, 不需要计算任何序参量, 直接通过机器学习的方法找到相变点. 本文采用上述机器学习方法可以从基态精确找到海森伯 J_1 - J_2 模型的一阶相变点、无法找到无穷阶相变点, 从第一激发态不仅能找到一阶相变点, 还能找到无穷阶相变点.

2 模型

本文研究的系统是一维反铁磁海森伯自旋链. 其哈密顿量是:

$$H = J_1 \sum_{\langle i,j \rangle} S_i \cdot S_j + J_2 \sum_{\langle\langle i,j \rangle\rangle} S_i \cdot S_j, \quad (1)$$

其中, J_1 表示最近邻相互耦合作用, J_2 表示次近邻相互耦合作用, $\langle i,j \rangle$ 表示取遍所有的最近邻, $\langle\langle i,j \rangle\rangle$ 表示取遍所有的次近邻, N 表示链长. 选择

周期性边界条件. 此模型最近邻和次近邻之间存在竞争, 在热力学极限下有两个相变点. Qian 等 [18–21] 研究表明在 $J_2/J_1 = 0.5$ 时, 此模型发生一阶相变, 该模型变为 Majumdar-Ghosh 模型, 此相变是由平移对称性破缺引起的, 在 $J_2/J_1 = 0.241$ 时, 此模型发生无穷阶相变, 当 $J_2/J_1 > 0.241$ 时, 体系是二聚化相, 当 $J_2/J_1 < 0.241$ 时, 体系是自旋液体相. 由图 1 可知, 此相变点对应于第一激发态的能级交叉点, 此时基态无能级交叉, 无法反映此相变. 需要特别说明的是, 本文计算的是有限长的系统, 因此由于系统的局限性, 从图 1 可以看出, 除了在此模型的相变点处有能及交叉外, 不同链长的系统在不同的地方也出现了能级交叉点, 这些假的相变点并没有反映真实的物理, 本文不做讨论. 下文讨论的机器学习方法的输出结果也反映了此情况, 这虽然在一定程度上对研究造成了干扰, 但并没有妨碍对系统真实相变点的识别.

3 无监督学习和深度学习学习方法

采用无监督学习识别分类点的方法为高斯混合模型 (GMM), GMM 是单一高斯概率密度函数的延伸 [22,23], 一维高斯分布模型的概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (2)$$

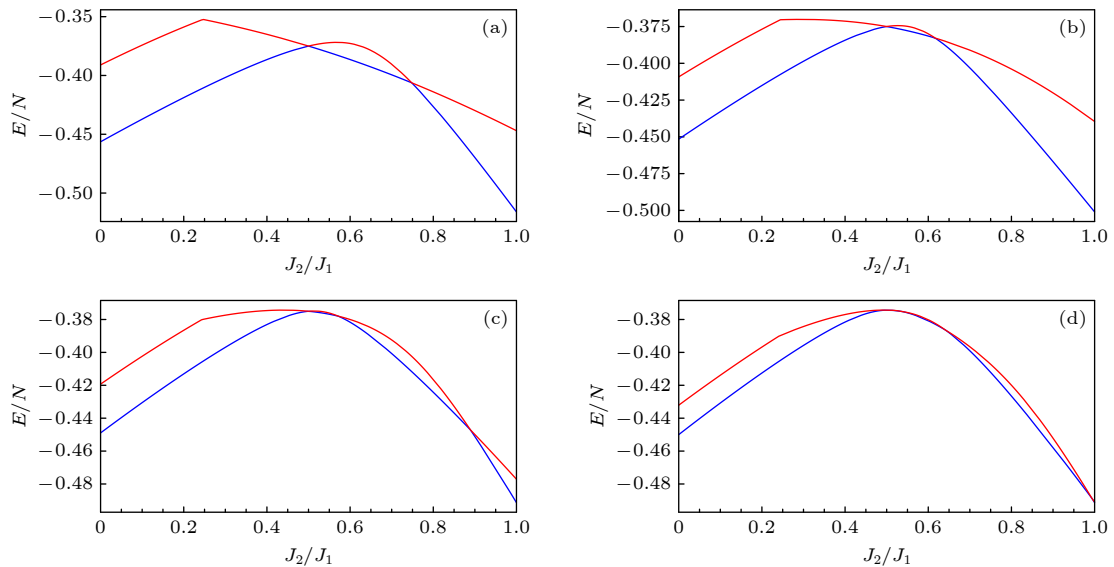


图 1 不同链长的 J_1 - J_2 海森伯自旋链系统基态与第一激发态能量随 J_2/J_1 的变化 (a) $N = 8$; (b) $N = 10$; (c) $N = 12$; (d) $N = 16$
 Fig. 1. Ground and first excited energy level diagram in J_1 - J_2 Heisenberg spin chain system with chain length: (a) $N = 8$; (b) $N = 10$; (c) $N = 12$; (d) $N = 16$.

其中 μ 为数据均值, σ 为数据标准差, 多维高斯分布模型的概率密度函数如下:

$$N(x; \mu, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}, \quad (3)$$

其中 μ 为数据均值, Σ 为数据协方差, D 为数据维度.

GMM是由 K 个单高斯模型组合而成的模型, 其概率密度函数为

$$p(x) = \sum_{i=1}^k \pi_k N(x; \mu_k, \Sigma_k), \quad (4)$$

其中 π_k 为混合系数, 满足 $\sum_{i=1}^k \pi_k = 1$.

GMM类似于聚类, 每一个高斯模型可以看作一种类别, 输入一个样本, 即可通过高斯概率密度函数计算其值, 然后通过比较其属于每个高斯模型的概率大小来判断高斯模型的类别. 使用GMM是因为高斯分布具备很好的物理性质以及良好的计算性能.

先通过GMM找到分类点, 然后通过卷积神经网络(CNN)进一步验证该点是否为相变点. CNN包括: 输入层、卷积层、池化层、全连接层和输出层. 数据集进入网络与卷积层多个卷积核卷积生成特征图, 池化层对特征图下采样, 生成新特征图, 全连接层将池化层生成的多个特征图融合、分类, 输出分类结果.

本文采用的CNN共包含5层^[24,25], 第1层为输入层, 输入数据为一维反铁磁海森伯模型的态矢量, 第2层为卷积层(卷积核大小为 4×4), 卷积后生成16个大小为 $2^{N/2} \times 2^{N/2}$ 的特征图(本文采用Relu激活函数). 第3层是池化层, 采用最大池化操作(采样窗口大小为 2×2 , 步长为2), 所以输出特征图尺寸减半, 即此刻特征图大小为 $2^{N/4} \times 2^{N/4}$, 但特征图数量仍为16个, 池化层的输出数据展为一维矢量后输入第4层全连接层(神经元个数设置为128). 为了避免系统参数更新过快导致过拟合, 本文在全连接层添加了dropout(dropout率设为0.5), 全连接层输出数据输入第5层输出层, 输出层神经元个数设置为 n (n 为类别数, 选取的训练数据为无监督学习得到的分类点两侧的数据, 采用独热编码(one-hot)将分类点一侧的训练数据标记为0, 另一侧标记为1, 因此本文 n 为2), 激活函数采用softmax函数^[1], 表达式为

$$f(x_j) = \frac{e^{x_j}}{\sum_{i=1}^n e^{x_i}}. \quad (5)$$

为了得到泛化能力较好的模型, 采用交叉验证的策略. 在交叉验证中, 将样本集随机划分为10等份, 使用其中的1份作为验证集而其余9份作为训练集, 通过正确率和交叉熵损失函数来衡量模型对于未知样本的预测能力.

采用远离相变点的训练集和验证集数据来训练和评估模型, 使用训练好的模型对测试数据进行预测.

4 结果与讨论

4.1 从基态寻找相变点

使用 $J_2/J_1 \in [0, 1)$ (链长 $N = 16$ 的取 $J_2/J_1 \in [0, 0.53)$, 因为取 $J_2/J_1 \in [0, 1)$ 会优先识别随系统尺寸变化而改变的相变点 $J_2/J_1 = 0.534$)的海森伯 J_1 - J_2 模型基态态矢量作为训练集, 生成 n 为2的GMM, 再用该模型对海森伯 J_1 - J_2 模型基态态矢量分类, 得到如图2(a)所示的分类结果, 可以看出 $J_2/J_1 = 0.5$ 是它的一个分类点, 其不随系统尺寸变化而改变.

接下来采用稍微远离 $J_2/J_1 = 0.5$ 的数据集作为训练集和验证集, 对该点是否是相变点进一步验证. 我们采用标记分别为0(采用one-hot编码)的 $J_2/J_1 \in [0.35, 0.45)$ 和为1的 $J_2/J_1 \in [0.55, 0.65)$ 的基态态矢量作为训练数据和验证数据(链长 $N = 16$ 的取 $J_2/J_1 \in [0.51, 0.61)$, 因为取 $J_2/J_1 \in [0.55, 0.65)$ 会识别随系统尺寸变化而改变的 $J_2/J_1 = 0.534$ 相变点), 训练CNN模型. 之后使用该模型对海森伯 J_1 - J_2 模型基态态矢量进行预测, 预测结果如图2(b)所示, 发现神经元输出在 $J_2/J_1 = 0.5$ 处发生跳变, 接下来采用标记分别为0的 $J_2/J_1 \in [0.3, 0.4)$ 和为1的 $J_2/J_1 \in [0.55, 0.65)$ 的基态态矢量和标记分别为0的 $J_2/J_1 \in [0.2, 0.3)$ 和为1的 $J_2/J_1 \in [0.55, 0.65)$ 的基态态矢量作为训练数据, 分别训练两个CNN模型, 再用模型对基态态矢量进行预测, 预测结果如图2(c)和图2(d)所示, 发现前后使用不同训练数据集训练的CNN对测试数据集的预测结果几乎没有发生变化, 因此, 基本就能确定 $J_2/J_1 = 0.5$ 为该系统相变点. 这样就在没有任何先验知识, 不知道系统哈密顿量, 能量的条件下从基态态

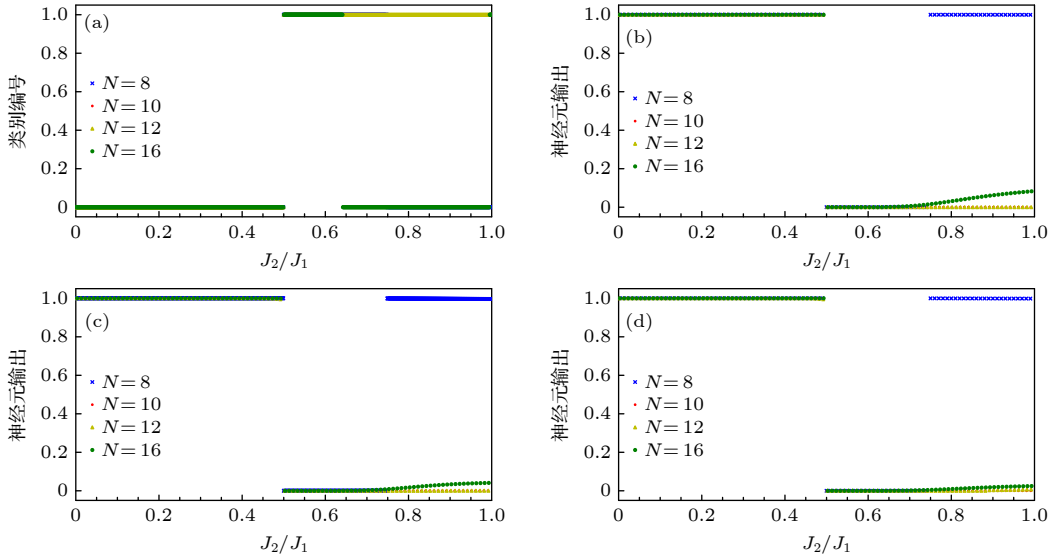


图 2 (a) 训练数据为 $J_2/J_1 \in [0, 1)$ 的海森伯 J_1 - J_2 模型基态矢量生成的 GMM 对基态矢量的分类结果; (b) 采用标记为 0 的 $J_2/J_1 \in [0.35, 0.45)$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65)$; (c) 标记为 0 的 $J_2/J_1 \in [0.3, 0.4)$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65)$; (d) 标记为 0 的 $J_2/J_1 \in [0.2, 0.3)$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65)$ 的基态态矢量作为训练数据, 训练所得的 CNN 模型对基态态矢量的预测结果

Fig. 2. (a) Ground state vector classification results of the GMM generated by the Heisenberg J_1 - J_2 model ground state vector with the training data of $J_2/J_1 \in [0, 1)$; (b) using the ground state vector of $J_2/J_1 \in [0.35, 0.45)$ marked as 0 and $J_2/J_1 \in [0.55, 0.65)$ marked as 1; (c) $J_2/J_1 \in [0.3, 0.4)$ marked as 0 and $J_2/J_1 \in [0.55, 0.65)$ marked as 1; (d) $J_2/J_1 \in [0.2, 0.3)$ marked as 0 and $J_2/J_1 \in [0.55, 0.65)$ marked as 1 as training data, the prediction results of the ground state vector by the trained convolutional neural network model.

矢量中直接找到了 J_1 - J_2 海森伯自旋链系统的一阶相变点。

同样使用 $J_2/J_1 \in [0, 0.5)$ 的海森伯 J_1 - J_2 模型基态态矢量作为训练集 (GMM 模型给出的不同链长的分类点不尽相同, 但其结果基本一样, 所以仅以链长 $N = 10$ 的态矢量的分类结果为例进行讨论), 生成 $n = 2$ 的 GMM, 再用该模型对海森伯 J_1 - J_2 模型基态态矢量分类, 得到如图 3(a) 所示的分类结果, 其在 $J_2/J_1 = 0.305$ 处出现分类点. 然后, 分别采用标记分别为 0 的 $J_2/J_1 \in [0.25, 0.3)$ 和为 1 的 $J_2/J_1 \in [0.35, 0.4)$ 的基态态矢量作为训练数据, 采用标记分别为 0 的 $J_2/J_1 \in [0.2, 0.25)$ 和为 1 的 $J_2/J_1 \in [0.35, 0.4)$ 的基态态矢量作为训练数据, 和采用标记分别为 0 的 $J_2/J_1 \in [0.2, 0.25)$ 和为 1 的 $J_2/J_1 \in [0.35, 0.4)$ 的基态态矢量作为训练数据 (标记为 1 的数据是标记为 0 数据的 5 倍), 分别训练 3 个 CNN 模型, 之后使用模型对系统基态的态矢量进行预测, 预测结果如图 3(b)—(d) 所示. 可以看出, CNN 给出的分类点随训练数据集数据的变化而发生改变, 不能给出确定的相变点, 因此不能通过该方法从基态态矢量中找到无穷阶相变点, 但能够有效排除无监督学习给出的假的相变点。

4.2 从第一激发态寻找相变点

同样, 使用 $J_2/J_1 \in [0, 1)$ 的海森伯 J_1 - J_2 模型第一激发态态矢量作为训练集, 生成 n 为 2 的 GMM, 再用该模型对海森伯 J_1 - J_2 模型第一激发态态矢量分类, 得到如图 4(a) 所示的分类结果, 可以观察到 $J_2/J_1 = 0.24$ 和 $J_2/J_1 = 0.5$ 可能是它的相变点. 接下来, 采用稍微远离 $J_2/J_1 = 0.5$ 数据集作为训练集和验证集, 对该点是否是相变点进一步验证. 采用标记分别为 0 的 $J_2/J_1 \in [0.35, 0.45)$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65)$ 的第一激发态态矢量作为训练数据, 采用标记分别为 0 的 $J_2/J_1 \in [0.3, 0.4)$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65)$ 的第一激发态态矢量作为训练数据, 和采用标记分别为 0 的 $J_2/J_1 \in [0.35, 0.45)$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65)$ 的第一激发态态矢量作为训练数据 (标记为 0 的数据是标记为 1 数据的 5 倍), 分别训练 CNN 模型, 训练得到的 3 个 CNN 模型预测结果如图 4(b)—(d) 所示, 发现训练好的 CNN 可以从测试数据中精确找到一阶相变点, 且其对测试数据集的预测结果几乎没有发生变化, 由此, 也可以通过第一激发态态矢量数据, 在没有任何先验知识, 不知道系统哈密顿量, 能量的条件下, 确定 $J_2/J_1 = 0.5$ 为 J_1 - J_2 海森伯自旋链系统相变点。

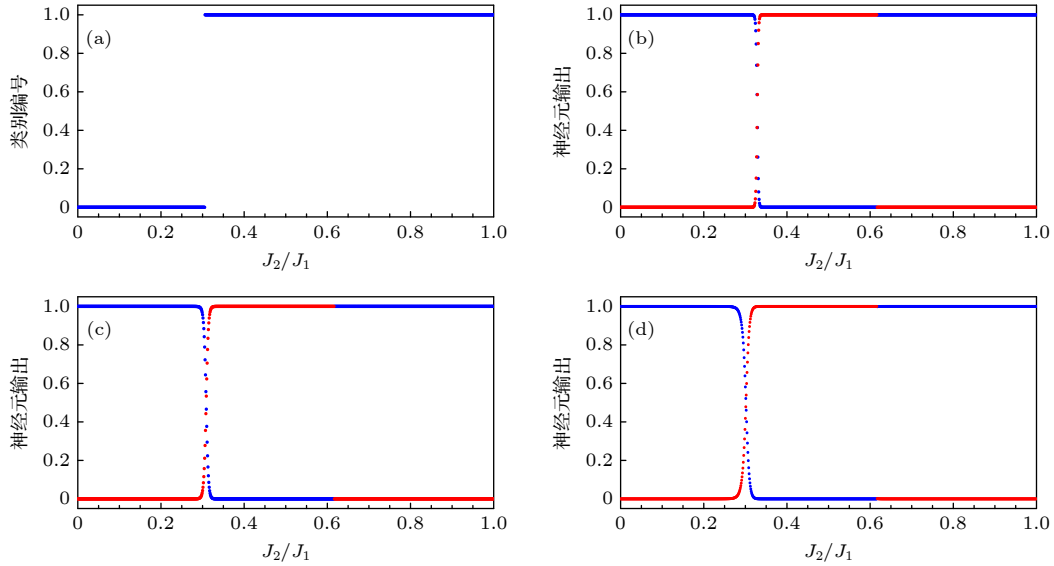


图 3 (a) 训练数据为 $J_2/J_1 \in [0, 0.5]$ 的海森伯 J_1 - J_2 模型基态态矢量生成的 GMM 对基态态矢量的分类结果; (b) 分别为采用标记为 0 的 $J_2/J_1 \in [0.25, 0.3]$ 和为 1 的 $J_2/J_1 \in [0.35, 0.4]$; (c) 标记为 0 的 $J_2/J_1 \in [0.2, 0.25]$ 和为 1 的 $J_2/J_1 \in [0.35, 0.4]$; (d) 标记为 0 的 $J_2/J_1 \in [0.2, 0.25]$ 和为 1 的 $J_2/J_1 \in [0.35, 0.4]$ (标记为 1 的数据是标记为 0 的 5 倍) 的基态态矢量作为训练数据, 训练所得的 CNN 模型对基态态矢量的预测结果

Fig. 3. (a) Ground state vector classification results of the GMM generated by the Heisenberg J_1 - J_2 model ground state vector with the training data of $J_2/J_1 \in [0, 0.5]$; (b) respectively using the ground state vector of $J_2/J_1 \in [0.25, 0.3]$ marked as 0 and $J_2/J_1 \in [0.35, 0.4]$ marked as 1; (c) $J_2/J_1 \in [0.2, 0.25]$ marked as 0 and $J_2/J_1 \in [0.35, 0.4]$ marked as 1; (d) $J_2/J_1 \in [0.2, 0.25]$ marked as 0 and $J_2/J_1 \in [0.35, 0.4]$ marked as 1 (the data marked as 1 is 5 times as much as the data marked as 0) as training data, the prediction results of the ground state vector by the trained convolutional neural network model.

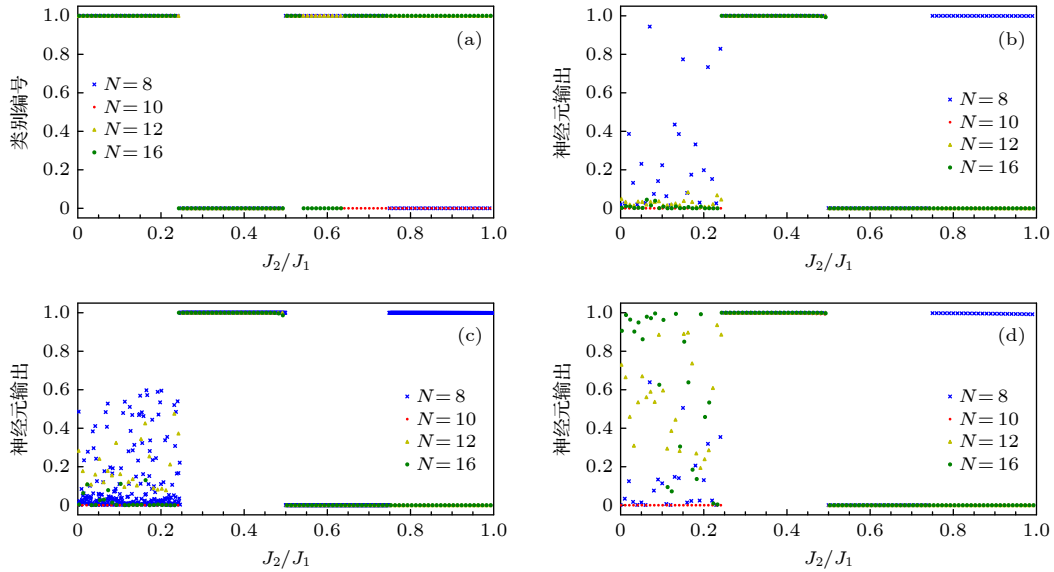


图 4 (a) 训练数据为 $J_2/J_1 \in [0, 1]$ 的海森伯 J_1 - J_2 模型第一激发态态矢量生成的 GMM 对第一激发态态矢量的分类结果; (b) 分别为采用标记为 0 的 $J_2/J_1 \in [0.35, 0.45]$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65]$; (c) 标记为 0 的 $J_2/J_1 \in [0.3, 0.4]$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65]$; (d) 标记为 0 的 $J_2/J_1 \in [0.35, 0.45]$ 和为 1 的 $J_2/J_1 \in [0.55, 0.65]$ (标记为 1 的数据是标记为 0 的 5 倍) 的第一激发态态矢量作为训练数据, 训练所得的 CNN 模型对第一激发态态矢量的预测结果

Fig. 4. (a) The first excited state vector classification results of the GMM generated by the Heisenberg J_1 - J_2 model first excited state vector with the training data of $J_2/J_1 \in [0, 1]$; (b) using the first excited state vector of $J_2/J_1 \in [0.35, 0.45]$ marked as 0 and $J_2/J_1 \in [0.55, 0.65]$ marked as 1; (c) $J_2/J_1 \in [0.3, 0.4]$ marked as 0 and $J_2/J_1 \in [0.55, 0.65]$ marked as 1; (d) $J_2/J_1 \in [0.35, 0.45]$ marked as 0 and $J_2/J_1 \in [0.55, 0.65]$ marked as 1 (the data marked as 1 is 5 times as much as the data marked as 0) as training data, the prediction results of the first excited state vector by the trained convolutional neural network model.

同样,采用稍微远离 $J_2/J_1 = 0.24$ 数据集作为训练集和验证集,对该点是否是相变点进一步验证.采用标记分别为 0 的 $J_2/J_1 \in [0.1, 0.2)$ 和为 1 的 $J_2/J_1 \in [0.3, 0.4)$ 的第一激发态态矢量作为训练数据,训练 CNN 模型,再用训练好的 CNN 模型对海森伯 J_1 - J_2 模型第一激发态态矢量的类别进行预测,预测结果如图 5(a) 所示,我们发现训练好的 CNN 可以从测试数据中精确找到无穷阶相变点,接着我们又采用标记分别为 0 的 $J_2/J_1 \in [0, 0.1)$ 和为 1 的 $J_2/J_1 \in [0.3, 0.4)$ 的第一激发态态矢量作为训练数据,训练 CNN 模型.再用训练好的 CNN 模型对海森伯 J_1 - J_2 模型第一激发态态矢量的类别进行预测,预测结果如图 5(b) 所示,发现训练好的 CNN 对测试数据集的预测结果几乎没有发生任何变化.由此基本就能确定 $J_2/J_1 = 0.24$ 为 J_1 - J_2 海森伯自旋链系统相变点.这样就在没有任何先验知识,不知道系统哈密顿量,能量的条件下从第一激发态态矢量中找到了该系统无穷阶相变点.

使用 $J_2/J_1 \in [0, 0.24)$ 的海森伯 J_1 - J_2 模型第一激发态态矢量作为训练集 (GMM 模型给出的不同链长的分类点不尽相同,但其结果基本一样,所以仅讨论链长 $N = 10$ 的态矢量的分类结果),生成 n 为 2 的 GMM,再用该模型对海森伯 J_1 - J_2 模型第一激发态态矢量分类,得到如图 6(a) 所示的分类结果,可以观察到 $J_2/J_1 = 0.125$ 可能是该模型的相变点.采用标记分别为 0 的 $J_2/J_1 \in [0.05, 0.1)$ 和 1 的 $J_2/J_1 \in [0.15, 0.2)$ 的第一激发态态矢量作为训练数据,采用标记分别为 0 的 $J_2/J_1 \in [0, 0.05)$ 和 1 的 $J_2/J_1 \in [0.15, 0.2)$ 的第一激发态态矢量作为训练数据,分别训练 CNN 模型,训练得到的两个 CNN 模型对测试集的预测结果如图 6(b), 图 6(c)

所示,我们观察得到 CNN 模型的预测结果在 $J_2/J_1 \in [0, 0.24)$ 区域内未出现明显的间断点,且预测结果随训练数据集的变化发生改变,因此 $J_2/J_1 = 0.125$ 不是该模型的相变点.说明监督学习能够有效排除无监督学习模型提供的假相变点,有效提高仅通过机器学习找系统相变点的准确率.

使用 $J_2/J_1 \in [0.25, 0.5)$ 的海森伯 J_1 - J_2 模型第一激发态态矢量作为训练集,生成 n 为 2 的 GMM,再用该模型对海森伯 J_1 - J_2 模型第一激发态态矢量分类,得到如图 6(d) 所示的分类结果,可以观察到 $J_2/J_1 = 0.385$ 可能是该模型的相变点.仍采用上述方法,使用不同训练数据分别训练 CNN 模型,用训练所得的 CNN 模型对测试集进行预测,同样可观察得到 CNN 模型的预测结果在 $J_2/J_1 \in [0.25, 0.5)$ 区域内未出现明显的间断点,且预测结果随训练数据集的变化发生改变,因此 $J_2/J_1 = 0.385$ 不是该模型的相变点.

4.3 可视化态矢量

采用线性函数归一化方法将态矢量里的值都转换到 $[0, 255]$ 的范围,公式为

$$X = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \cdot 255. \quad (6)$$

将得到的每一个数据转换为 $2^{N/2} \times 2^{N/2}$ (N 为链长,本文 $N = 10$) 的灰度图 (全黑为 0, 全白为 255).

$J_2/J_1 = 0.44, 0.49, 0.51, 0.58$ 的基态态矢量灰度图如图 7 所示,观察发现基态矢量灰度图在相变点前和相变点后都变化缓慢,但在相变点前后发生了突变.第一激发态态矢量灰度图如图 8 所示,观察得到第一激发态态矢量灰度图在无穷极相变点和一阶相变点前后也都发生了突变.由此能通过

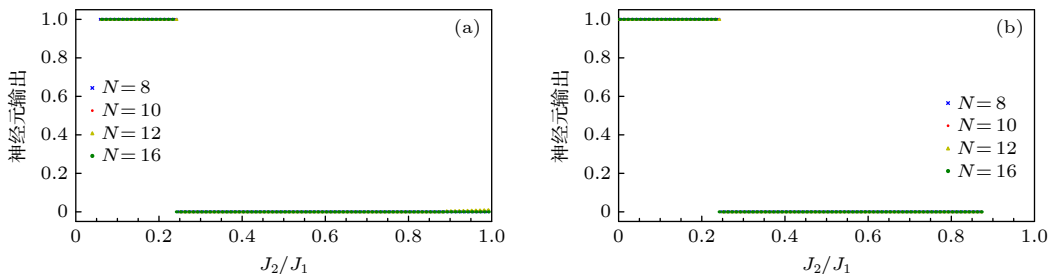


图 5 (a) 采用标记为 0 的 $J_2/J_1 \in [0.1, 0.2)$, 标记为 1 的 $J_2/J_1 \in [0.3, 0.4)$; (b) 标记为 0 的 $J_2/J_1 \in [0, 0.1)$, 标记为 1 的 $J_2/J_1 \in [0.3, 0.4)$ 的第一激发态态矢量作为训练数据,训练所得的 CNN 模型对第一激发态态矢量的预测结果

Fig. 5. (a) Using the first excited state vector of $J_2/J_1 \in [0.1, 0.2)$ marked as 0 and $J_2/J_1 \in [0.3, 0.4)$ marked as 1 as training data, the prediction results of the first excited state vector by the trained convolutional neural network model.

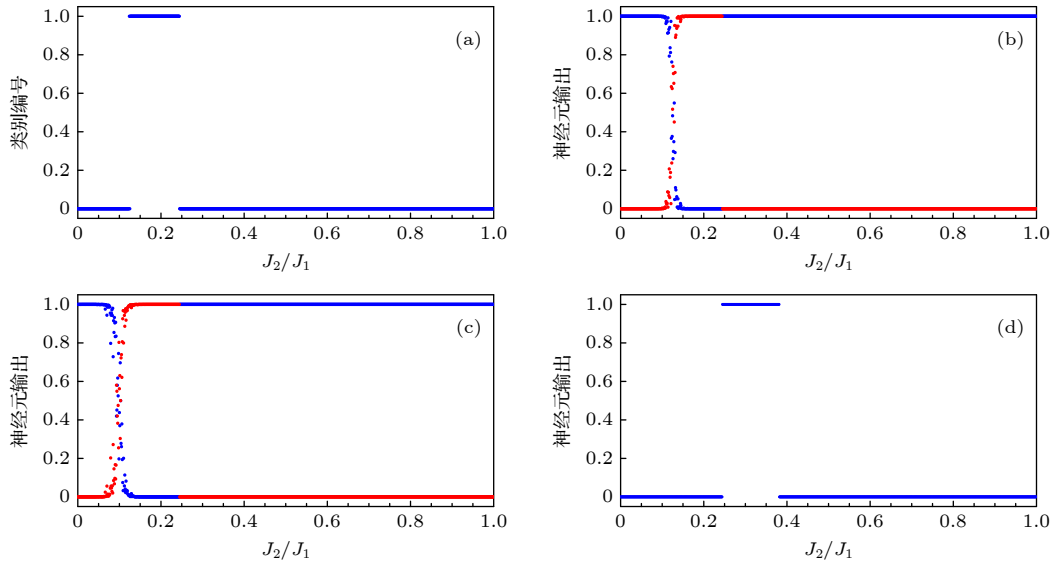


图 6 (a) 训练数据为 $J_2/J_1 \in [0, 0.24]$ 的海森伯 J_1 - J_2 模型第一激发态态矢量生成的 GMM 对第一激发态态矢量的分类结果; (b), (c) 分别为采用标记为 0 的 $J_2/J_1 \in [0.05, 0.1]$ 和为 1 的 $J_2/J_1 \in [0.15, 0.2]$ 的第一激发态态矢量作为训练数据, 训练所得的 CNN 模型对第一激发态态矢量的预测结果; (d) 训练数据为 $J_2/J_1 \in [0.25, 0.5]$ 的海森伯 J_1 - J_2 模型第一激发态态矢量生成的 GMM 对第一激发态态矢量的分类结果

Fig. 6. (a) The first excited state vector classification results of the GMM generated by the Heisenberg J_1 - J_2 model first excited state vector with the training data of $J_2/J_1 \in [0, 0.24]$; (b) (c) respectively using the first excited state vector of $J_2/J_1 \in [0.05, 0.1]$ marked as 0 and $J_2/J_1 \in [0.15, 0.2]$ marked as 1; $J_2/J_1 \in [0, 0.05]$ marked as 0 and $J_2/J_1 \in [0.15, 0.2]$ marked as 1 as training data, the prediction results of the first excited state vector by the trained convolutional neural network model; (d) he first excited state vector classification results of the GMM generated by the Heisenberg J_1 - J_2 model first excited state vector with the training data of $J_2/J_1 \in [0.25, 0.5]$.

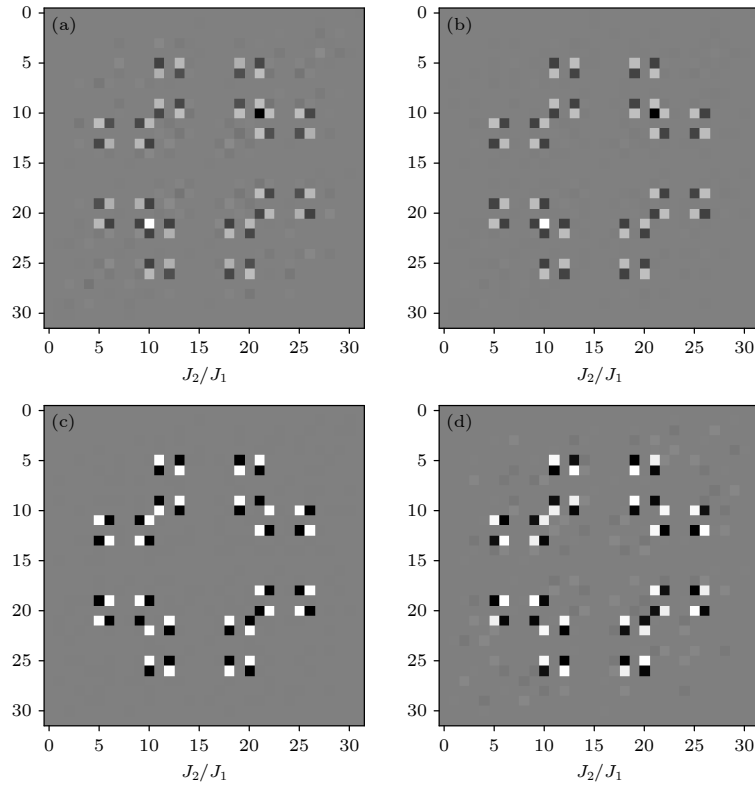


图 7 海森伯 J_1 - J_2 模型基态态矢量变换而来的灰度图 (a) $J_2/J_1 = 0.44$; (b) $J_2/J_1 = 0.49$; (c) $J_2/J_1 = 0.51$; (d) $J_2/J_1 = 0.58$

Fig. 7. Gray scale images transformed from the ground state vector of the Heisenberg J_1 - J_2 model: (a) $J_2/J_1 = 0.44$; (b) $J_2/J_1 = 0.49$; (c) $J_2/J_1 = 0.51$; (d) $J_2/J_1 = 0.58$.

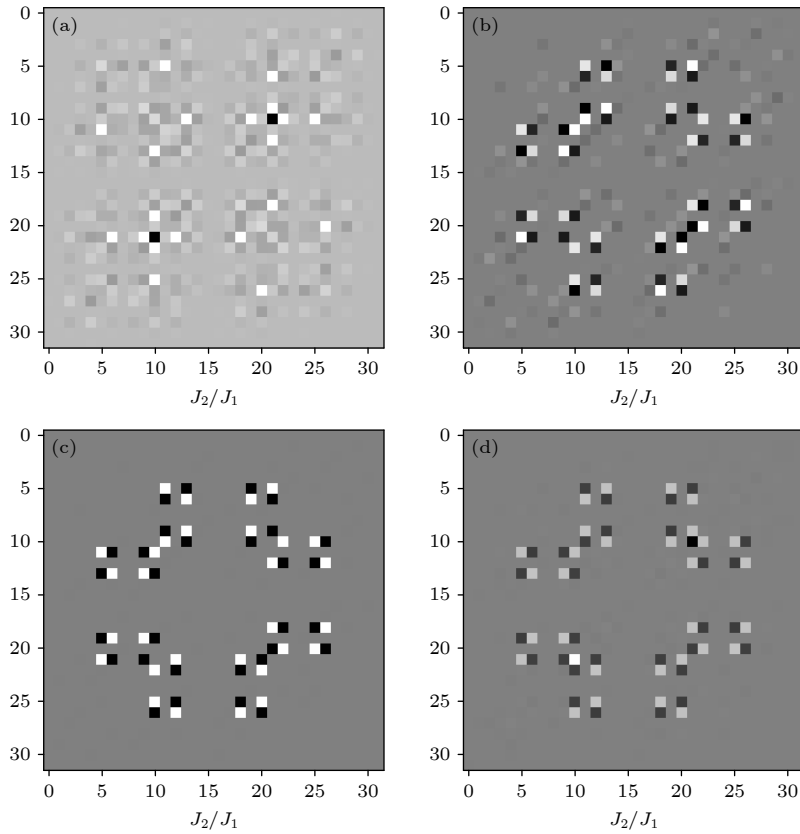


图 8 海森伯 J_1 - J_2 模型第一激发态态矢量变换而来的灰度图 (a) $J_2/J_1 = 0.24$; (b) $J_2/J_1 = 0.25$; (c) $J_2/J_1 = 0.49$; (d) $J_2/J_1 = 0.51$
 Fig. 8. Gray scale images transformed from the first excited state vector of the Heisenberg J_1 - J_2 model: (a) $J_2/J_1 = 0.24$; (b) $J_2/J_1 = 0.25$; (c) $J_2/J_1 = 0.49$; (d) $J_2/J_1 = 0.51$.

态矢量灰度图直接观测到一维次近邻海森伯 J_1 - J_2 模型的相变点. 机器学习方法也正是因为能提取出态矢量这些特征信息的变化, 才能够有效找到相变点.

5 结 论

本文将无监督学习与监督学习结合起来, 从一维次近邻海森伯 J_1 - J_2 模型的基态中找到了该模型的一阶相变点, 从第一激发态不仅找到了一阶相变点, 还找到了无穷阶相变点, 间接说明了第一激发态可能比基态蕴含更多信息, 将态矢量可视化说明了机器学习算法的可靠性.

研究表明, 机器学习能够在没有任何先验知识、不知道系统哈密顿量、能量的条件下, 直接从庞大的态矢量样本集合中找出一些可能存在的相变点, 这可能为研究未知系统时提供一些帮助; 监督学习能够对无监督学习给出的相变点加以验证, 去除假的相变点.

由于本文仅研究了一维海森伯 J_1 - J_2 模型, 此方法的鲁棒性和推广能力还未知, 需要进一步研究.

参考文献

- [1] Lei W 2016 *Phys. Rev. B.* **94** 195105
- [2] Wetzel S J 2017 *Phys. Rev. E* **96** 022140
- [3] Huang L, Wang L 2017 *Phys. Rev. B* **95** 035105
- [4] Phiala E S, Daniel T, William D 2018 *Phys. Rev. D* **97** 094506
- [5] Lakovlev I A, Sotnikov O M, Mazurenko V V 2018 *Phys. Rev. B* **98** 174411
- [6] Dong X Y, Pollmann F, Zhang X F 2019 *Phys. Rev. B* **99** 121104
- [7] Tan D R, Jiang F J 2020 *Phys. Rev. B* **102** 224434
- [8] Tan D R, Li C D, Zhu W P, Jiang F J 2020 *New J. Phys.* **22** 063016
- [9] Maskara N, Buchhold M, Endres M 2021 arXiv: 2103.15855 [quant-ph]
- [10] Tanja D 2021 arXiv: 2103.07236[quant-ph]
- [11] Sondhi S L, Girvin S M, Carini J P, Shahar D 1997 *Rev. Mod. Phys.* **69** 315
- [12] Bulla R, Vojta M 2003 *Rep. Prog. Phys.* **66** 2069
- [13] Walker N, Tam K M, Novak B, Jarrell M 2018 *Phys. Rev. E.* **98** 053305
- [14] Jadrich R B, Lindquist B A, Pineros W D, Truskeet T M 2018 *J. Chem. Phys.* **149** 194109
- [15] Canabarro A, Fanchini F F, Malvezzi A L, Pereira R, Chaves R 2019 *Phys. Rev. B* **100** 045129
- [16] Carrasquilla J, Melko R G 2017 *Nat. Phys.* **13** 431
- [17] Ahmadreza A, Michel P 2020 arXiv: 2007.09764 [cond-mat.stat-mech]

- [18] Chitra R, Pati S, Krishnamurthy H R, Sen D, Ramasesha S 1995 *Phys. Rev. B* **52** 6581
- [19] Castilla G, Chakravarty S, Emery V J 1995 *Phys. Rev. L* **75** 1823
- [20] Shu C, Li W, Shi J G, Wang Y P 2007 *Phys. Rev. E* **76** 061108
- [21] Qian X F, Shi T, Li Y, Song Z, Sun C P 2005 *Phys. Rev. A* **72** 012333
- [22] Zhou Z H 2016 *Machine Learning* (Beijing: Tsinghua University Press) p206 (in Chinese) [周志华 2016 机器学习 (北京: 清华大学出版社) 第206页]
- [23] Li H 2012 *The Elements of Statistical Learning* (Beijing: Tsinghua University Press) p162 (in Chinese) [李航 2012 统计机器学习 (北京: 清华大学出版社) 第162页]
- [24] Xu Q W, Wang P P, Zeng Z J, Huang Z B, Zhou X X, Liu J M, Li Y, Chen S Q, Fan D Y 2020 *Acta Phys. Sin.* **69** 014209 (in Chinese) [徐启伟, 王佩佩, 曾镇佳, 黄泽斌, 周新星, 刘俊敏, 李瑛, 陈书青, 范滇元 2020 物理学报 **69** 014209]
- [25] Goodfellow I, Bengio Y, Courville A (translated by Zhang SJ, Li Y J, Fu T F, Li K) 2017 *Deep Learning* (Beijing: The People's Posts and Telecommunications Press) pp143–317 (in Chinese) [伊恩·古德费洛, 约书亚·本吉奥, 亚伦·库维尔 著 (赵申剑, 黎晟君, 符天凡, 李凯 译) 2017 深度学习 (北京: 人民邮电出版社) 第143—317页]

Identifying phase transition point of J_1 - J_2 antiferromagnetic Heisenberg spin chain by machine learning*

Wang Wei Jie Quan-Lin[†]

(School of Physics and Technology, Wuhan University, Wuhan 430000, China)

(Received 14 April 2021; revised manuscript received 27 June 2021)

Abstract

Studying quantum phase transitions through order parameters is a traditional method, but studying phase transitions by machine learning is a brand new field. The ability of machine learning to classify, identify, or interpret massive data sets may provide physicists with similar analyses of the exponentially large data sets embodied in the Hilbert space of quantum many-body system. In this work, we propose a method of using unsupervised learning algorithm of the Gaussian mixture model to classify the state vectors of the J_1 - J_2 antiferromagnetic Heisenberg spin chain system, then the supervised learning algorithm of the convolutional neural network is used to identify the classification point given by the unsupervised learning algorithm, and the cross-validation method is adopted to verify the learning effect. Using this method, we study the J_1 - J_2 Heisenberg spin chain system with chain length $N = 8, 10, 12, 16$ and obtain the same conclusion. The first order phase transition point of J_1 - J_2 antiferromagnetic Heisenberg spin chain system can be accurately found from the ground state vector, but the infinite order phase transition point cannot be found from the ground state vector. The first order and the infinite order phase transition point can be found from the first excited state vector, which indirectly shows that the first excited state may contain more information than the ground state of J_1 - J_2 antiferromagnetic Heisenberg spin chain system. The visualization of the state vector shows the reliability of the machine learning algorithm, which can extract the feature information from the state vector. The result reveals that the machine learning techniques can directly find some possible phase transition points from a large set of state vector without prior knowledge of the energy or locality conditions of the Hamiltonian, which may assist us in studying unknown systems. Supervised learning can verify the phase transition points given by unsupervised learning, thereby indicating that we can discover some useful information about unknown systems only through machine learning techniques. Machine learning techniques can be a basic research tool in strong quantum-correlated systems, and it can be adapted to more complex systems, which can help us dig up hidden information.

Keywords: Heisenberg J_1 - J_2 model, machine learning, neural network, phase transition

PACS: 07.05.Mh, 03.67.-a, 03.65.Vf

DOI: 10.7498/aps.70.20210711

* Project supported by the National Natural Science Foundation of China (Grant No. 1217050658).

[†] Corresponding author. E-mail: qljie@whu.edu.cn