arXiv:2304.04725v2 [nucl-th] 8 Feb 2024

使用神经网络量子态对一维无自旋陷阱费米子系统进行机器学习

J. W. T. Keeble ,[1], M. Drissi ,[1, 2, †] . Rojo-Franc`as ,[3, 4] B. Juli´a-D´ıaz ,[3, 4] 和 A. Ri s [1, 5, 6]

[1] *Department of Physics, University of Surrey, Guildford GU2 7XH, United Kingdom* [2] *TRIUMF, 4004 Wesbrook Mall, Vancouver, BC V6T 2A3, Canada*

[3] *Departament de Física Quàntica i Astrofísica (FQA), Universitat de Barcelona (UB) c. Martí i Franquès 1 08028 Barcelona Spain* , , , , , ,

[4] *Institut de Ciències del Cosmos (ICCUB) Universitat de Barcelona (UB) c. Martí i Franquès 1*, 08028, *Barcelona, Spain* [5] *Departament de Física Quàntica i Astrofísica (FQA), Universitat de Barcelona (UB), c. Martí i Franquès 1*, 08028, *Barcelona, Spain*

[6] *Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB),c. Martí i Franquès 1*, 08028, *Barcelona, Spain*（日期：2024年2月9日）

我们计算了通过高斯势相互作用的全极化受限一维费米系统的基态性质。我们采用反对称人工神经网络（即神经量子态）作为波函数的 *Ansatz* ，并使用机器学习技术，通过变分方法最小化含有两个至六个粒子的系统能量。我们为这一玩具模型提供了与其他多体方法（包括精确对角化和哈特里-福克近似）的广泛基准比较。在广泛的相互作用强度范围内，神经量子态给出的能量最佳。我们发现，相互作用符号不同，基态也大不相同。在非微扰排斥区域，系统渐近地呈现晶体有序。相反，强吸引区域则显现出玻色化的迹象。神经量子态能够持续学习这些不同的相，并且参数数量几乎保持不变，计算时间随粒子数增加而仅有很小幅度的增长。

I. 引言

机器学习（ML）在 Science 中的兴起彻底改变了众多领域，从从头算 量子化学到宇宙学，机器学习都能直接从数据中“学习”，以理解物理现象 [1]。基于神经网络的学习算法以通用逼近定理（UAT）为基础，原则上能够以任意精度表示数据。然而，UAT 是存在性 定理，说明这类优势有可能实现，但并不一定指出如何实现这些优势 [2–5]。因此，研究人员可以广泛构建基于机器学习的算法来表征物理现象，其中最引人注目的是量子多体问题。

参考文献 [6]首次将机器学习应用于量子多体系统，重点关注离散系统。此后，人们已使用这些技术 [7–14] 处理了多种不同的物理系统。量子化学领域的应用，例如 FermiNet [15, 16] 或 PauliNet [17,18], , 为精确求解电子薛¨定谔方程 [19–23]铺平了道路。量子态与神经网络的结合催生了神经网络量子态这一全新领域（神经量子态） [9]。

基于机器学习的神经量子态方法通过将波函数表示为神经网络 [6]，以变分方式求解薛定谔方程。神经量子态的构造方式确保其显式满足多体波函数的对称性，例如费米子情形下的反对称性。迄今为止，有迹象表明，神经量子态能够以紧凑的方式压缩多体波函数中的相关信息 [6, 8, 15–17, 24–26]。神经量子态的实现形式多种多样，其中常采用第一量子化的实空间表述，利用蒙特卡罗（MC）技术对多粒子变量进行积分。这种方法本质上归结为变分蒙特卡罗（VMC） [27] ，并配以表达能力更强的*Ans¨atze*。此外，机器学习框架能够通过自动微分（AD）技术高效更新神经量子态的参数 [28]。直接获取多体波函数还有一项额外优势：原则上可以计算任意多体可观测量，并有可能模拟多体动力学 [6]。

本文的重点是，为在一个具有凝聚态物质研究意义的简单系统中求解多体薛¨定谔方程，提供神经量子态的最简实现。作为概念验证，我们将注意力转向一维系统；与三维系统相比，一维系统在计算上更为简单，并且研究对象是完全极化的（或无自旋的）费米子。我们假设费米子处于谐振阱中，并通过有限程相互作用刻画它们之间的成对相互作用。这项探索性工作是迈向构建能够描述更复杂核系统的神经量子态 [29–33]的垫脚石。

在某些情况下，玻色子和费米子可能具有相似的性质。这种现象被称为

* j.keeble@surrey.ac.uk
† mdrissi@triumf.ca

费米-玻色对偶性，在一维空间中尤为显著 [34, 35]。这种对偶性表现为强相互作用的费米子如同弱相互作用的玻色子，反之亦然 [36]。传统上，人们讨论这种对偶性时，关注的是具有接触相互作用的自旋1/2粒子 [37, 38]。对于完全极化费米子，泡利不相容原理要求波函数试探函数 关于粒子位置具有反对称性，因此在物理上禁止了S 波相互作用。相应的相互作用主要由 P 波、奇宇称项构成 [36, 39–42]。我们在此构建了一个玩具模型，通过采用偶宇称相互作用忽略了一些具有物理意义的信息，并对几种不同方法进行基准比较，以评估神经量子态（NQS）试探函数的质量。

有趣的是，如今世界各地的超冷原子实验室都已能够通过实验研究所考虑的量子多体问题 [43, 44]。从制备简并费米气体 [45]，开始，实验人员能够调节费米子之间的相互作用并研究状态方程 [46] ，甚至能够以受控的方式制备少体费米子系统 [47]。

在本文中，我们基准测试了参考文献 [15] 中的 FermiNet 神经量子态的极简实现，用于表示一维受陷阱束缚的费米子系统。在第 II 节中，我们定义了所考虑的系统和玩具模型哈密顿量。在第 III 节中，我们详细解释了神经量子态和用于基准测试的多体方法。在 IV 节中，我们对所得结果进行了详细分析。在第 V 节中，我们给出结论并展望未来研究。

II. 系统和哈密顿量

我们研究了一个由 A 个质量为m 的相同费米子组成的系统，它们被囚禁在频率为 ω的谐振阱中。下文中我们忽略自旋，假设该系统是完全极化的。我们关注有限范围的高斯粒子间相互作用，该相互作用在实空间中给出如下哈密顿量：

$$\hat{H} = -\frac{\hbar^2}{2m}\sum_{i=1}^{A}\nabla_i^2 + \frac{1}{2}m\omega^2\sum_{i=1}^{A}x_i^2 + \frac{V}{\sqrt{2\pi}\sigma}\sum_{i<j}\exp\left[-\frac{(x_i-x_j)^2}{2\sigma^2}\right]. \tag{1}$$

高斯相互作用由相互作用强度 V 和相互作用范围 σ 表征。我们选择这两个参数，使得在 $\sigma \to 0$ 的极限下，势能变为接触相互作用 $\to V\delta(x_i - x_j)$。

在本文余下部分，我们始终使用谐振子（组态相互作用）单位：长度以$a_{\mathrm{ho}} = \sqrt{\hbar/m\omega}$ 为单位定义，能量以 $\hbar\omega$为单位测量。

哈密顿量变为

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{A}\nabla_i^2 + \frac{1}{2}\sum_{i=1}^{A}x_i^2 + \frac{V_0}{\sqrt{2\pi}\sigma_0}\sum_{i<j}\exp\left[-\frac{(x_i-x_j)^2}{2\sigma_0^2}\right]. \tag{2}$$

重新定义相互作用范围，使得 $\sigma_0 = \sigma/a_{\mathrm{ho}}$。无量纲相互作用强度 V_0 与有量纲常数 V 的关系为 $V_0 = V/(a_{\mathrm{ho}}\hbar\omega)$。

对于无自旋费米子，有限范围相互作用是观察到相互作用效应的必要条件。事实上，在没有自旋的情况下，多体波函数关于空间变量 $\{x_i, i = 1, \ldots, A\}$ 是反对称的。对于粒子对 $i \neq j$，只要 $x_i = x_j$，波函数就会抵消。因此，接触相互作用不会对系统的能量产生贡献。直观地说，人们预期，与自旋态不同的情况下具有相同强度的相互作用相比，这类相互作用效应相对较小。

选择高斯形状因子来描述相互作用，主要是出于实际考虑 [37, 43]。首先，由于相关积分形式简单，高斯函数易于用于多体模拟，包括精确对角化 [48]以及随机方法 [49]。其次，如上所述，通过调节范围参数，可以利用归一化的高斯相互作用逼近接触相互作用极限 $\sigma \to 0$。第三，我们的最终目标是模拟核物理系统。有限程相互作用与核物理应用尤其相关，因为相互作用的范围与介子力载体的质量有关 [50]。核相互作用中也有采用高斯（或多个高斯函数之和）形状因子的例子，例如 Gogny 力 [50]。我们还注意到，高斯相互作用已被广泛用于分析多种多体系统，包括所谓的普适行为“高斯表征” [51]。

以往关于无自旋费米子的文献采用了奇宇称相互作用，在零程极限下其表现为 $V(x) \approx \overleftarrow{\partial_x}\delta(x)\overrightarrow{\partial_x}$。这是无自旋费米子相互作用的正确极限，无法通过高斯势再现。因此，我们的结果应被视为学术玩具模型，基于这一原因，我们着重关注方法学基准，而非实验应用。我们将奇宇称势在NQS框架下的研究留待今后开展。

A. 非相互作用情形

在进一步介绍如何考虑相互作用之前，我们先简要关注一下可解析求解的非相互作用情形。在没有自旋的情况下，每个费米子都可以占据一个能量为 $\epsilon_n = n + 1/2$ 的单粒子能级，该能级由一个量子数表征

n。单粒子波函数为谐振子本征态，

$$\varphi_n(x) = \mathcal{N}_n e^{-\frac{x^2}{2}} H_n(x) \,, \tag{3}$$

其中 $H_n(x)$ 是第 n 阶埃尔米特多项式，$\mathcal{N}_n = 1/\sqrt{2^n n! \sqrt{\pi}}$ 是归一化常数。多体波函数是一个纯斯莱特行列式。它可以分解为一个整体高斯包络乘以一个仅包含埃尔米特多项式的行列式：

$$\Psi(x_1, \ldots, x_A) = \frac{1}{\sqrt{A!}} \left[\Pi_{n=0}^{A-1} \mathcal{N}_n e^{-\frac{x_i^2}{2}} \right] \times \begin{vmatrix} H_0(x_1) & H_0(x_2) & \ldots & H_0(x_A) \\ H_1(x_1) & H_1(x_2) & \ldots & H_1(x_A) \\ \vdots & \vdots & \ddots & \vdots \\ H_{A-1}(x_1) & H_{A-1}(x_2) & \ldots & H_{A-1}(x_A) \end{vmatrix} . \tag{4}$$

进而，还可以用范德蒙德行列式进一步简化 [37, 52]。

多体系统的总能量只需将所有已占据单粒子态的能量相加即可得到，

$$E_A = \sum_{n=0}^{A-1} \epsilon_n = \frac{A^2}{2} \,, \tag{5}$$

并且它随 A^2 缩放。这些非相互作用基准在多种情况下都很有用，尤其是在构建神经量子态（NQS）试探函数时。具体而言，正如下文所述，我们先对神经网络进行预训练，使其拟合非相互作用解。这提供了一个受限、稳定且符合物理规律的初始结果，以此为起点，我们便可以开展计算量相对较大的变分模拟。

B. 密度矩阵

我们可以通过使用多体密度矩阵 [53, 54]，进一步刻画系统中的关联。A 体系统的单体密度矩阵（OBDM）定义为对多体波函数进行如下 $A-1$ 积分：

$$\rho(x_1', x_1) = A \int dx_2 \ldots dx_A \Psi^*(x_1', x_2, \ldots, x_A) \times \Psi(x_1, x_2, \ldots, x_A). \tag{6}$$

在 ρ 的空间表象中的对角化，

$$\int d\bar{x} \rho(x, \bar{x}) \phi_\alpha(\bar{x}) = n_\alpha \phi_\alpha(x), \tag{7}$$

由此得到所谓的自然轨道 $\phi_\alpha(x)$，以及占据数 n_α。α 原则上是一个从 $\alpha = 0$ 到无穷大的离散指标。单体密度矩阵的谱分解可展开如下：

$$\rho(x_1', x_1) = \sum_{\alpha=0}^{\infty} n_\alpha \phi_\alpha^*(x_1') \phi_\alpha(x_1). \tag{8}$$

我们采用如下归一化方式：$\sum_\alpha n_\alpha = A$。

在非相互作用或哈特里-福克（Hartree-Fock）基态中，可以发现

$$n_\alpha = \begin{cases} 1, & \alpha < A, \\ 0, & \alpha \geq A. \end{cases} \tag{9}$$

因此，公式（8）中的求和自然截断为 A 项。此外，对于非相互作用情形，自然轨道对应于公式（3）中的单粒子态。对非相互作用系统进行求和，可得单体密度矩阵具有以下形式

$$\rho(x_1', x_1) = \frac{e^{-\frac{x_1'^2}{2} - \frac{x_1^2}{2}}}{\sqrt{\pi}} \mathcal{R}^{(A)}(x_1', x_1), \tag{10}$$

其中 $\mathcal{R}^{(A)}(x_1', x_1)$ 是关于 $A-1$的最高阶为 x_1 和 x_1' [55–57]的多项式。

还可以证明，占据数由式(9描述的系统，其多体波函数是一个斯莱特行列式 [58]。换言之，与非关联值 $n_m = 1$ 和 $n_m = 0$ 的偏差为系统中的内禀关联提供了可靠的度量 [42]。

二体密度矩阵也是内禀关联的绝佳指标。它通常定义为 [53],

$$\Gamma(x_1', x_2'; x_1, x_2) = \binom{A}{2} \int dx_3 \ldots dx_A \Psi^*(x_1', x_2', \ldots, x_A) \times \Psi(x_1, x_2, \ldots, x_A) \,.$$

该矩阵的正定对角线元素提供了系统的对关联函数（PCF），

$$g(x_1, x_2) = \Gamma(x_1' = x_1, x_2' = x_2; x_1, x_2) \,, \tag{11}$$

它可以根据当另一个粒子位于 x_2 [54]时在位置 x_1 处找到一个粒子的概率进行直接的物理解释。在非相互作用情形下，也可以为这一对象找到闭式表达式。在这种情况下，关联函数具有如下类型的结构

$$g(x_1, x_2) = \frac{e^{-x_1^2 - x_2^2}}{\pi} (x_1 - x_2)^2 \mathcal{G}^{(A)}(x_1, x_2), \tag{12}$$

其中，$\mathcal{G}^{(A)}(x_1, x_2)$ 同样是关于 x_1 和 x_2 [57] 的最高阶为 $2(A-2)$ 的多项式。

III. 方法

现在我们介绍用于对少体系统求解结果进行基准比较的不同量子多体方法。首先介绍神经量子态试探函数，随后简要说明适用于 $A > 1$ 情形的精确对角化方法和Hartree-Fock近似。最后，我们介绍两种额外的数值方法，用于专门对 $A = 2$ 情形进行基准比较。

A. 神经网络量子态

1. *Architecture*

我们的神经量子态 试探函数是一个完全反对称的神经网络，其设计受到 FermiNet 实现的启发 [15]。我们的网络的输入是 A 系统中费米子的位置，$\{x_i, i = 1, \ldots, A\}$，输出则是多体波函数 $\Psi_\theta(x_1, \ldots\ldots, x_A)$，它取决于一系列网络权重和偏置，这些参数简洁地概括为一个多维变量 θ。我们在图1中给出了网络架构的示意图。该网络由 4个核心组件构成：等变层、广义 Slater 矩阵（GSM）、对数包络函数以及求和的带符号对数行列式函数。下面我们将依次介绍神经量子态的这 4 个核心组件。

为了使神经量子态满足反对称性，我们在整个网络中强制实现置换等变性。网络的输入是一个 A 维向量，$x \in \mathbb{R}^A$。等变层的输出 h_i 满足：输入变量 π 的任意置换都会相应地置换输出，即 $h_i(x_{\pi(1)}, \ldots, x_{\pi(A)}) = h_{\pi(i)}(x_1, \ldots, x_A)$。

为确保等变性，我们遵循文献 [59] 的方法。在输入层中添加一个置换不变特征进行预处理；在我们的情况下，该特征是多体位置的均值，$\mu^{(0)} = \frac{1}{A}\sum_{i=1}^{A} x_i$。因此，第一层的输入是 $x \in \mathbb{R}^A$ 与相应的平均位置 $\mu^{(0)} \in \mathbb{R}$ 的拼接。这定义了一个输入特征 $f^{(0)} \in \mathbb{R}^{A\times 2}$，使得

$$f_{ij}^{(0)} = \begin{cases} x_i & \text{if } j = 1 \\ \mu^{(0)} & \text{if } j = 2 \end{cases}. \tag{13}$$

网络的第一个等变层（$L = 1$）在图 1的灰色区域中显示。输入特征的每一行都通过一个共享层，以构建位置的中间 H维表示，$h^{(1)} \in \mathbb{R}^{A\times H}$ （图 1中的黄色区域）。该层由权重 $W^{(1)} \in \mathbb{R}^{H\times 2}$和偏置 $b^{(1)} \in \mathbb{R}^{1\times H}$ 组成的线性变换以及非线性激活函数构成。我们选择双曲正切激活函数，因为它是连续且可微的，而在计算多体动能时，这是必需的。具体而言，每一行 $h_i^{(1)} \in \mathbb{R}^{1\times H}$ 表示为

$$h_i^{(1)} = \tanh\left(f_i^{(0)} W^{(1)^\intercal} + b^{(1)}\right). \tag{14}$$

第二个等变层（$L = 2$）获取第一个层的输出 $h^{(1)}$及其按列平均值$\mu^{(1)} \in \mathbb{R}^H$，以定义新的输入特征 $f^{(1)} \in \mathbb{R}^{A\times 2H}$，使得

$$f_{ij}^{(1)} = \begin{cases} h_{ij}^{(1)} & \text{if } j \le H \\ \mu_i^{(1)} & \text{if } j > H \end{cases}. \tag{15}$$

与第一层类似，输入特征会通过一个权重为 $W^{(2)} \in \mathbb{R}^{H\times 2H}$ 的共享层，偏置 $b^{(2)} \in \mathbb{R}^{1\times H}$ 和相同的非线性激活函数。此外还添加了残差连接，使输出的每一行 $h_i^{(2)} \in \mathbb{R}^{1\times H}$ 都能明确地表示为

$$h_i^{(2)} = \tanh\left(f_i^{(1)} W^{(2)^\intercal} + b^{(2)}\right) + h_i^{(1)}. \tag{16}$$

在第二个等变层之后，$h^{(2)}$ 经过一个线性层（见图 1 中的米色区域），其使用共享权重 $W^{(M)} \in \mathbb{R}^{A\times H}$ 和偏置 $b^{(M)} \in \mathbb{R}^{1\times A}$ 输出一个矩阵 $M \in \mathbb{R}^{A\times A}$。每一行 $M_i \in \mathbb{R}^{1\times A}$ 读取。

$$M_i = h_i^{(2)} W^{(M)^\intercal} + b^{(M)}. \tag{17}$$

在这一阶段，尚未纳入渐近边界条件，例如波函数在无穷远处衰减。为此，我们采用一个包络函-数 [60], ，该函数在对数域中实现，以确保数值稳定性。对数包络矩阵的形式为

$$\ln(e_{ij}) = -\left(x_i W_j^{(e)}\right)^2, \tag{18}$$

其中 i 和 j 分别表示粒子和轨道的索引，而 $W^{(e)} \in \mathbb{R}^A$ 是用于确定每个轨道对数包络的学习权重。我们使用高斯包络，而不是参考文献 [15], 中的指数包络，它们更接近于方程（4）的非相互作用解。

然后，我们将 M 与相应的包络函数进行逐元素相乘，

$$\phi_{ij} = M_{ij} e_{ij}, \tag{19}$$

由此得到一个 GSM。该矩阵中的每个元素，$\phi_{ij} = \phi_j(x_i; \{x_{/i}\})$，都可理解为状态 j 上的广义单粒子轨道。该轨道不仅取决于粒子 i 的位置，还以置换不变的方式取决于其他粒子的位置，如记号 $\{x_{/i}\}$ [15] 所示。这一设计的重要优势在于，所有轨道都依赖于所有粒子的位置，从而能够高效地在波函数 [10, 15, 61],中编码回流关联，详见第 IIIA 4 节。

最后，我们对 GSM ϕ 取行列式，得到反对称波函数。这通常称为广义斯莱特行列式（GSD）。原则上，单个 GSD 足以表示任意反对称波函数 [62]。根据经验，我们还观察到，单个 GSD 几乎捕获了该系统中的所有关联 [30]。值得注意的是，GSD 通过 LU（下三角—上三角）分解在对数域中计算。采用这一方案是为了保证数值稳定性。

我们的计算框架具有通用性，可以调整等变层的数量 L 和隐藏节点的数量 H。我们的神经量子态使用 $L = 2$ 个等变层，每层包含 $H = 64$ 个隐藏节点。根据数值测试，这种选择在收敛方面表现最佳；相关测试见参考文献 [30]。我们的代码也可以处理多个 GSM，并且

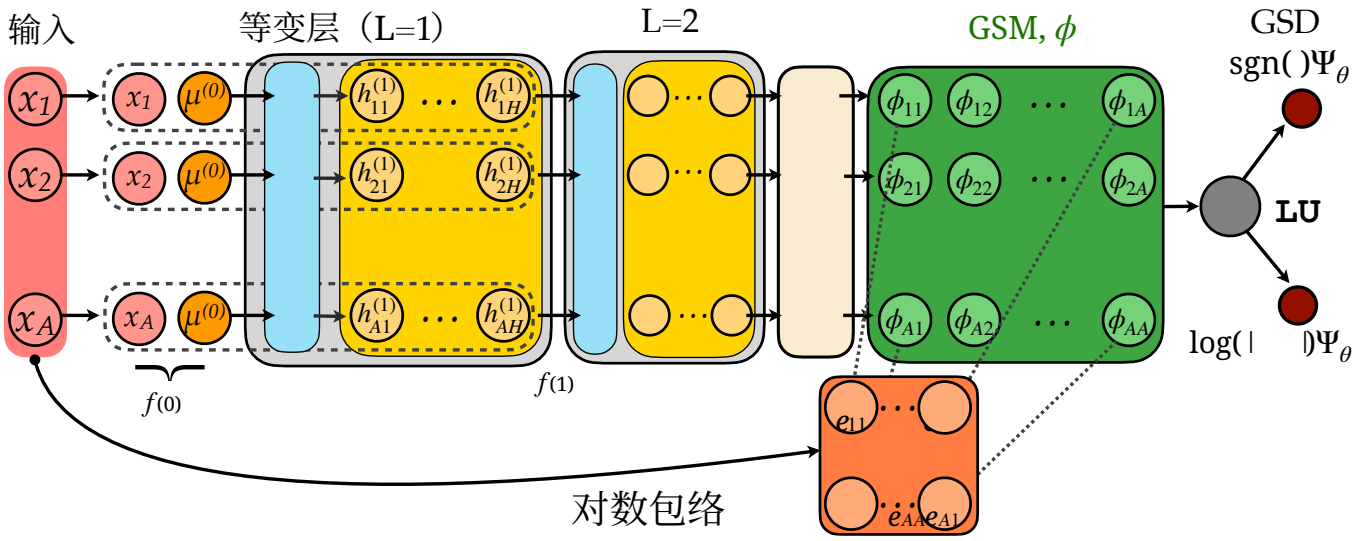

图 1. 本文用于 $L = 2$ 个含 H 个隐藏节点的等变层的神经量子态 *Ansatz*。网络的输入是$\{x_i, i = 1, \ldots, A\}$ 个粒子的位置。随后通过 2 个等变层（浅灰色区域）处理，以确保神经量子态在整个前向传播过程中保持等变性。灰色虚线表示共享层按行应用。其非线性激活函数为双曲正切函数（蓝色区域）。第二个中间等变层$h^{(2)}$ 的输出经过共享线性层，得到一个 $A \times A$ 矩阵 M，再与对数包络层（橙色区域）逐元素相乘，从而得到我们的 GSM ϕ （绿色区域）。最后，使用 LU 分解求出多体波函数 Ψ_θ的符号及其绝对值的对数。

广义斯莱特行列式。如果我们采用 D 个广义斯莱特行列式来描述系统，每个行列式都会通过包络权重对 D 的依赖关系，在公式 $W^{(e)}$ （18） [30] 中拥有各自的包络。在 $D > 1$ 情况下，广义斯莱特行列式通过带符号的 log-sum-exp 函数求和。关于 [30] 情况下数值实现的完整详细说明，请参阅参考文献 $D > 1$ 的附录 A。

2. *Variational Monte Carlo*

定义了神经量子态试探函数后，我们现在来讨论如何实现针对该问题的量子多体解。我们分两个阶段通过变分蒙特卡罗方法求解薛定谔方程：先针对初始目标波函数进行预训练，再对未知的基态波函数进行能量最小化。预训练步骤可视为一项监督学习任务，我们要求网络重现非相互作用系统的多体波函数。我们的目标是获得一个物理上合理的初始态，其空间范围与开启相互作用后的结果大致相似。为此，我们最小化损失函数，

$$\mathcal{L}^{\mathrm{Pre}}(\theta) = \int dx_1 \ldots dx_A \, |\Psi_\theta(x_1, \ldots, x_A)|^2 \times \sum_{ij} \left[\phi_i\left(x_j; \{x_{/j}\}\right) - \varphi_i(x_j)\right]^2,$$

其中 $\phi_i\left(x_j; \{x_{/j}\}\right)$ 是 GSM 的一个元素，而 φ_i 是第 i 个谐振子单粒子态，见公式（3）。损失函数可通过蒙特卡罗采样改写为

$$\mathcal{L}^{\mathrm{Pre}}(\theta) = \mathbb{E}_{X \sim |\Psi_\theta|^2} \left[\sum_{ij} \left[\phi_i\left(x_j; \{x_{/j}\}\right) - \varphi_i(x_j)\right]^2 \right], \tag{20}$$

其中，我们将 $X = (x_1, \ldots, x_A)$定义为一个服从多体波函数 $|\Psi_\theta|^2$ 的玻恩概率分布的 A 维随机变量（或游走者）。

首先，使用均值为零、方差为一的 $N_W = 4096$ 维高斯分布，将一组游走者独立分布在构型空间的原点处。随后，预训练阶段迭代 A维高斯分布。预训练阶段随后迭代 10^4个训练轮次。每个训练轮次包含三个不同阶段。首先，将N_W 个游走者通过 Metropolis-Hastings（MH）算法，按 NQS 的玻恩概率成比例地分布 [63, 64]。每个训练轮次运行 MH 算法 10 次迭代。其次，计算局部损失值并进行反向传播，以评估损失函数式（20）关于参数 θ的梯度。第三，使用参考文献 [65], 中的 Adam 优化器更新参数，超参数采用默认值，但学习率设为 10^{-4}。预训练完成后，NQS 即表示非相互作用情形的近似解。

我们变分蒙特卡罗方法的下一阶段是在强化学习框架下，最小化能量期望值以求得基态波函数。采用

标准的量子蒙特卡罗记法，能量期望值表示为局域能量的统计平均，

$$E(\theta)=\frac{\langle\Psi_\theta|\hat{H}|\Psi_\theta\rangle}{\langle\Psi_\theta|\Psi_\theta\rangle}=\mathbb{E}_{X\sim|\Psi_\theta|^2}\left[\Psi_\theta(X)^{-1}\hat{H}\Psi_\theta(X)\right]. \tag{21}$$

在统计不确定度范围内，该期望值符合变分原理，且 $E(\theta)$ 大于基态能量 $E_{\mathrm{g.s.}}$。

我们在对数域中计算动能，从而得到局域能量

$$\begin{aligned}E_L(X)&\equiv\Psi_\theta(X)^{-1}\hat{H}\Psi_\theta(X)\\&=-\frac{1}{2}\sum_{i=1}^{A}\left[\frac{\partial^2\ln|\Psi_\theta|}{\partial x_i^2}\bigg|_X+\left(\frac{\partial\ln|\Psi_\theta|}{\partial x_i}\right)^2\bigg|_X\right]\\&\quad+\sum_{i=1}^{A}\frac{x_i^2}{2}+\frac{V_0}{\sqrt{2\pi}\sigma_0}\sum_{i<j}e^{-\frac{(x_i-x_j)^2}{2\sigma_0^2}}.\end{aligned} \tag{22}$$

游走者通过 MH 算法进行传播。借助这一随机过程，游走者可以对波函数的玻恩概率所规定的任意概率分布进行采样。此外，我们遵循参考文献 [19] 中的方法，在 MH 采样器中动态调整提议分布的宽度。此举旨在确保 MH 算法每一步中，平均约有 50% 的游走者接受所提议的构型。这样一来，提议分布便能有效地随系统大小缩放，从而实现马尔可夫链更高效的热化。

为保证数值稳定性，我们遵循参考文献 [15]并采用 ℓ_1 范数裁剪，在此过程中计算一个“可接受”的局域能量值窗口。我们选择 ℓ_1 范数，因为它对离群值更加稳健。对于一批局域能量，我们计算局域能量中位数$\langle E_L\rangle$，以及与之关联的 ℓ_1 范数偏差 σ_{ℓ_1}。可接受窗口定义为范围 $\langle E_L\rangle\pm5\sigma_{\ell_1}$。窗口之外的任何值都会被替换为各侧所接受的最大值。

计算出能量期望值后，我们使用相对于 θ的梯度来更新试探函数 的参数，

$$\nabla_\theta E(\theta)=2\mathbb{E}_{X\sim|\Psi_\theta|^2}\left[\left(E_L(X)-E(\theta)\right)\nabla_\theta\ln|\Psi_\theta(X)|\right]. \tag{23}$$

能量的梯度通过对辅助损失函数应用反向模式自动微分来计算，

$$\mathcal{L}^{\mathrm{aux}}(\theta)=2\mathbb{E}_{X\sim\perp(|\Psi_\theta|^2)}\left[\perp\left(E_L(X)-\mathbb{E}_{X\sim|\Psi_\theta|^2}\left[E_L(X)\right]\right)\times\ln|\Psi_\theta(X)|\right], \tag{24}$$

其中 $\perp(\ldots)$ 表示分离函数 [66]。分离函数是一种常见的机器学习实现工具，能够方便地将梯度的解析计算与自动微分算法结合起来。在应用自动微分时，按定义，分离函数等价于恒等函数，但其导数为零，即 $\perp(x)=x$ 而 $\partial\perp(x)/\partial x=0$。虽然式（24）中的期望值不等于式（21）中的能量，但这两个表达式关于神经量子态参数的导数相等。与式（¨24）的直接实现相比，采用这一辅助损失函数的原因在于数值稳定性，因为它可以避免高阶导数的计算。可以证明，利用分离函数，计算梯度所需的最高阶导数是关于 x 动能的二阶导数。若直接对式（¨21）应用自动微分来计算能量的梯度，则会得到三阶混合导数。

与预训练阶段一样，我们使用 $N_W=4096$个游走者，并通过最小化能量进行 10^5 个训练轮次。每个训练轮次从 10 步 MH采样开始，通过公式（22）计算局域能量，并反向传播公式（24）中的局域梯度，以使神经量子态收敛至基态波函数。参数使用 Adam 优化器更新，除学习率设为 10^{-4} [65]外，其余超参数均采用默认值。图2 展示了一次最小化运行中能量的变化，参数为 $A=6$、$V_0=-20$ 和 $\sigma_0=0.5$。这是本文所展示的最具挑战性的情形之一：粒子数相对较多，且与非相互作用情形有较大偏差。图中同时展示了 E 的平均值（实线）以及公式（21）中积分的 99.7% 置信区间。我们发现，经过约 10000 个训练轮次后，能量与最终值的偏差在 0.15% 以内。能量稳步收敛至基态，并在中心值附近仅有较小振荡（见图2中的插图）。

我们假设，经过 10^5 训练轮次后，神经量子态达到基态，并将该阶段的参数 θ 冻结。为计算最终的能量估计值，我们遵循参考文献 [67] ，通过“阻塞”方法对多个批次进行计算。我们使用 10^4 个批次，每批次包含 4096 个游走者，总计 $\approx4.1\times10^7$ 个样本。我们移除了前述的 ℓ_1 范数裁剪，以免对最终的能量测量造成偏差。最终能量（及其标准差）通过对这些测量结果取平均值得到。总标准差也采用参考文献 [67]中的阻塞方法计算。单条链的标准差定义为：

$$\sigma_B^2=\sigma^2/(L_B/N_{corr}), \tag{25}$$

其中，σ_B 和 σ 分别为数据块 B 和所有样本的经验标准差。 L_B 是给定数据块 B 的长度，其相关长度为 N_{corr}。该标准差通常极小，不会影响我们的结论。本文中的所有观测量均使用这一扩展样本集计算，从而确保测量精度足够高。关于我们神经量子态的实现细节，更多内容请参见文献 [30] 以及 GitHub 仓库 [68]。

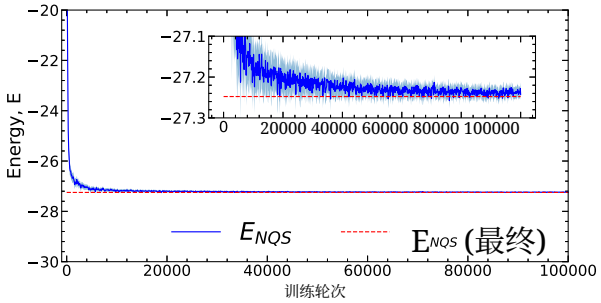

图 2. $A = 6$ 系统在 $V_0 = -20$ 和 $\sigma_0 = 0.5$条件下，能量期望值随训练轮次的收敛情况。神经量子态 $Ansatz$ 以蓝色实线表示。蒙特卡洛能量积分的 99.7%（3 个标准差）置信区间以蓝色阴影区域表示。通过阻塞方法计算得到的神经量子态最终能量以红色虚线表示。为清晰起见，收敛图中每隔 100th 个训练轮次显示一个数据点。

3. *Scaling with particle number*

现在，我们简要讨论 NQS 方法随粒子数 A的扩展性。为了评估扩展性，我们估算了 NQS 的四个主要层在前向传播中的内存和时间需求。前两个层（图 1中的灰色区域）包含偏置在内，分别有$H \times (2+1)$和 $H \times (2H+1)$个参数。需要强调的是，此处参数的数量与粒子数无关。这两个层是共享的，并作用于 A 个行向量，每个行向量对应一个不同的粒子，因此这些层的规模在 A中保持不变，而时间复杂度随 A线性增长。其余两个层，即 GSM 和对数包络层，包含偏置在内，分别有 $D \times (H+1)$ 个 $\times A$ 和 $D \times A$ 个参数。对于这两个层，规模随A线性增长，而时间复杂度随 A二次增长。最后，对 GSM 的行列式进行最终求值无需参数，但其时间复杂度按 $D \times A^3$增长。

实际上，我们将 $H = 64$、$L = 2$ 和 $D = 1$ 固定，而 A 从二变化到六。就内存而言，主要成本是存储中间层的权重矩阵（关于 H 呈二次关系），尽管其关于 A 保持不变。例如，$A = 2$ 神经量子态包含 8580 个参数，而 $A = 6$ 神经量子态包含 8844 个参数；由于参数数量与 A 呈线性关系，因此仅增加了 3%。关于计算成本，渐近主导项是行列式的求值（前向传播中关于 A 呈三次关系）。然而，在我们的模拟中，我们发现，对于所研究的系统尺寸，随着 A 增大，我们的神经量子态每训练轮次的墙钟时间增加相对较小。在 $A = 2$ 情况下，每训练轮次的墙钟时间约为 0.1 秒，而在 $A = 6$ 中则升至每训练轮次 0.3 秒。这种线性扩展表明，我们尚未达到渐近大 A 的行为。对我们的变分蒙特卡罗模拟进行进一步性能分析——表明，对于 $A \leq 6$，就墙钟时间而言，主要耗时实际上来自 MH 抽样 [69]。变分蒙特卡罗算法的这一部分比任何前向或反向网络传播都耗时得多，因此应在任何时间优化策略中予以解决。作为参考，我们所有实验均在单块 P4000 Quadro GPU 上运行；通过硬件升级，整体墙钟时间有望大幅缩短。

最后，在上述所有讨论中，我们都假设H 保持固定，而 A 不断增加。由此得到的温和标度律当然会受到网络表达能力下降的限制，因为 A 会增加。最低要求是 $A \leq H$，以确保 GSM 满秩，且神经量子态不为零。为检验网络的表达能力如何随 A变化，我们针对 $A = 12$ 系统进行了额外的数值测试，L 和 D取值保持不变，但改变了 H。虽然没有精确的基准方法可供比较，但我们发现，在这种情况下，对于$H > 16$，神经量子态试探函数 仍能达到低于相应 Hartree-Fock上界的能量。$H = 64$、128 和 256 对应的最终能量值基本已收敛。我们认为，这表明即使可能需要增加 H 以确保神经量子态的偏差仍可忽略，网络的表达能力依然能够得到控制。要更好地描述神经量子态试探函数的表达能力极限，还需针对更大的粒子数开展进一步探索。

4. *Backflow correlations*

我们的方法与以往的 ML 实现 [15, 31–33], ，尤其是 FermiNet，有多方面的不同。我们去除了卷积层，不考虑任何自旋依赖，使用高斯包络而非指数包络，并在广义斯莱特行列式之和方面实现了更好的数值稳定性。至关重要的是，我们也没有引入任何显式的 Jastrow 因子。正如下面将要看到的，这并不会损害相应波函数的质量。

我们神经量子态实现成功的最相关因素，是其融入回流关联的能力。在非相互作用情形和Hartree–Fock描述中，多体波函数是仅依赖于单粒子位置的轨道所构成的斯莱特行列式 [58, 60, 70]。在我们的神经量子态实现中，我们利用置换等变性与行列式层相结合的方式，为波函数表示引入了一种回流形式 [71–73]。这将单粒子轨道转换为准粒子轨道，使其以扩展的、置换不变的方式依赖于所有粒子的位置，并允许神经量子态波函数的节点面发生偏移 [74]。回流变换能够将关联融入多体波函数。具体而言，我们的神经量子态表示提供了一种灵活、数值成本低且

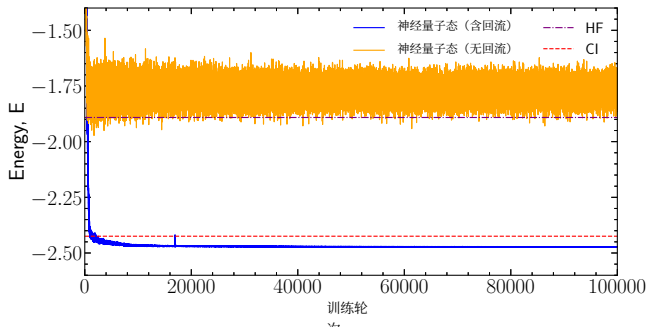

图 3. $A = 2$ 系统在 $V_0 = -20$ 和 $\sigma_0 = 0.5$条件下的能量期望值随训练轮次的收敛情况。标准神经量子态 $Ansatz$ （带回流）以蓝色实线表示，并优于精确对角化基准（红色虚线）。不带回流的神经量子态$Ansatz$ （橙色线）则收敛至 Hartree-Fock 解。

这些关联的相对通用表示 [10,15]，其可在变分蒙特卡罗过程中高效学习。

我们开展了一项富有启发性的数值实验，以阐明回流能力对于我们的神经量子态的重要性。神经量子态 试探函数 可以通过消除对 μ 在等变层中的依赖而轻松修改为不含任何回流的形式。该过程大致将模型的参数数量减少一半。此外，对于具有单个斯莱特行列式且不含回流变换的网络，可以预期变分过程会简单地得到 Hartree-Fock 行列式。这正是我们在图 3中观察到的结果，其中我们展示了 $A = 2$ 个粒子、相互作用强度为 $V_0 = -20$ 且范围为 $\sigma_0 = 0.5$时的能量最小化曲线。该图给出了标准神经量子态架构（含回流，蓝色实线）以及不含回流的修改版神经量子态（橙色实线）的结果。不含回流的神经量子态达到Hartree-Fock 能量，这是纯单粒子框架内相应的最优能量。同时，它还在能量最小值附近表现出明显的振荡。相比之下，在纳入回流关联后，神经量子态可以达到（事实上还优于）直接对角化基准，并且振荡极小。我们将此视为一个迹象，表明 试探函数中回流的存在是置换等变性的结果，使神经量子态具备准确表示超越平均场的费米子间关联的非凡能力。此外，加入回流后，能量的统计涨落也得到减弱，这表明回流对于实现准确的基态能量至关重要。

B. 直接对角化

我们采用直接对角化（或组态相互作用）方法来提供一个基准值。

对于 A体神经量子态。具体而言，我们使用谐振子作为单粒子基。我们将式（2）中的哈密顿量写成二次量子化形式，

$$\hat{H} = \sum_{\alpha=1}^{A} \epsilon_\alpha \hat{n}_\alpha + \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} V_{\alpha\beta,\delta\gamma} \hat{a}_\alpha^\dagger \hat{a}_\beta^\dagger \hat{a}_\gamma \hat{a}_\delta, \qquad (26)$$

其中 $V_{\alpha\beta,\ \delta\gamma}$ 是单粒子谐振子基组中的矩阵元，其解析表达式见参考文献 [75]。 $\hat{n}_\alpha = \hat{a}_\alpha^\dagger \hat{a}_\alpha$ 是数目算符，而 $\hat{a}_\alpha^\dagger$ ($\hat{a}_\alpha$) 是谐振子态的费米产生（湮灭）算符 α。

在我们的实现中，我们根据多体非相互作用能量截断基，这也定义了单粒子基的截断。对于所有粒子数，我们只考虑前 $N_{\max} = 20$ 个单粒子谐振子模式。这是当前资源条件所允许的最大态数，而我们的资源受限于相互作用矩阵元计算的精度。即使存在这些限制，排斥区域的结果仍已收敛。一般而言，吸引相互作用下的收敛速度较慢，例如参考文献 [76] 报告的零程情形。构建好多体基后，我们便构造哈密顿量矩阵，并使用标准 Lanczos 算法将其对角化。有关该方法的详情，请参阅参考文献 [48, 77]。

我们的计算受限于基截断$N_{\max}$，这意味着通过对角化获得的结果并不精确，只能提供能量上界。采用更大的基可以得到更好的结果，但多体基的维数随粒子数呈指数增长，因此对于大型系统而言并不实用。

由于采用谐振子基组，我们在排斥相互作用区域的能量计算中获得的误差小于吸引相互作用区域。正如我们将在下文看到的，这可以从系统的空间重排角度加以理解。然而，即使不改变单粒子基组，随着粒子数目的增加，结果的准确性也会降低。在本文中，我们只考虑了基态，但需要强调的是，该方法同样能够对激发态作出预测。

C. 哈特里-福克

我们通过考察 HF 基态提供了一个额外的基准，它是多体波函数的一种最简非关联试探函数 。与神经量子态的实现保持一致，我们在坐标空间中求解该问题。这种表示方式对于捕捉密度分布形状相对较大的变化尤其有用；采用固定基组的方法可能难以描述这些变化。

HF 轨道 $\phi_\alpha(x)$ 中的 $\alpha = 0, \ldots, A-1$ 均被完全占据，占据数为 $n_\alpha = 1$。所有这些。

其余态中，$\alpha \geq A$ 为空，$n_\alpha = 0$。HF 轨道用于构造单体密度矩阵，

$$\rho^{\mathrm{HF}}(x_1', x_1) = \sum_{\alpha=0}^{A-1} \phi_\alpha^*(x_1')\phi_\alpha(x_1), \tag{27}$$

以及相应的密度分布 $n^{\mathrm{HF}}(x) = \rho^{\mathrm{HF}}(x_1 = x,\ x_1' = x)$。轨道由 HF 方程获得，

$$\left[-\frac{1}{2}\nabla^2 + \frac{1}{2}x^2\right]\phi_\alpha(x) + \int d\bar{x}\Sigma(x,\bar{x})\phi_\alpha(\bar{x}) = \varepsilon_\alpha^{\mathrm{HF}}\phi_\alpha(x)\,. \tag{28}$$

对于有限程相互作用的情况，HF 方程是一组 A 积分-微分方程。HF 自能是直接项与（非局域）交换项之和，

$$\begin{aligned}\Sigma(x_1', x_1) =& \delta(x_1 - x_1')\int dx\, V(x_1 - x)n^{\mathrm{HF}}(x) \\ &+ V(x_1' - x_1)\rho^{\mathrm{HF}}(x_1', x_1)\,.\end{aligned} \tag{29}$$

相应的 HF 能量由单粒子能量之和计算得到；作为一致性检查，也可以通过对平均场进行直接积分，并加入相应的反对称性修正来获得 [78]。

我们通过在等距点离散网格上进行迭代，求解这组自洽方程。动能项采用傅里叶网格哈密顿量方法表示为矩阵 [79], ，只要网格范围远大于波函数的支集，该方法就能正常工作。平均场可以利用矩阵运算高效计算，且式（28）可化为一个本征值问题。对于小型系统，我们通常使用 $N = 200$ 个网格点，范围从 $x = -5$ 到 5。对于具有 $A > 4$的大型系统，网格范围扩展至 ± 6 ，并选择 $N = 240$个点，以保持相同的网格间距。关于 HF 方程求解的计算笔记本见参考文献 [80]。

D. $A = 2$的实空间解

在接下来的章节中，我们使用 $A = 2$ 系统来确定哪些相互作用强度和范围的取值值得关注。我们还采用了另外两种数值方法来专门求解这一问题。

对于 $A = 2$ 个粒子，式（1）中的哈密顿量尤其容易处理。按照文献 [41, 81]，我们引入质心（CoM）坐标，$R = \frac{1}{\sqrt{2}}(x_1 + x_2)$，以及相对坐标，$r = \frac{1}{\sqrt{2}}(x_1 - x_2)$。由此，哈密顿量 $\hat{H} = H_{\mathrm{CM}} + H_{\mathrm{rel}}$ 可分解为两个彼此对易的分量，

$$\hat{H}_{\mathrm{CM}} = -\frac{1}{2}\nabla_R^2 + \frac{1}{2}R^2, \tag{30}$$

$$\hat{H}_{\mathrm{r}} = -\frac{1}{2}\nabla_r^2 + \frac{1}{2}r^2 + \frac{V_0}{\sqrt{2\pi}\sigma_0}e^{-\frac{r^2}{\sigma_0^2}}\,. \tag{31}$$

质心分量就是一个谐振子。相对哈密顿量则包括一个谐振子项和高斯相互作用项。我们注意到，相对坐标中的高斯形状因子具有不同的宽度，且其有效宽度大于原始相互作用的宽度。

质心和相对哈密顿量可以分别对角化，分别得到本征态 $\varphi_\alpha(R)$和 $\psi_\beta(r)$，其本征能量分别为 ϵ_α 和 ε_β。对于相对坐标，我们需要求解以下本征值问题

$$\left[-\frac{1}{2}\nabla_r^2 + \frac{1}{2}r^2 + \frac{V_0}{\sqrt{2\pi}\sigma_0}\exp\left(-\frac{r^2}{\sigma_0^2}\right)\right]\psi_\beta(r) = \varepsilon_\beta\psi_\beta(r)\,. \tag{32}$$

由于泡利原理，唯一可接受的解是在相对坐标下呈反对称的解，$\psi_\beta(-r) = -\psi_\beta(r)$。这些解对应于 β的奇数值，因此基态必然对应于 $\beta = 1$，而非能量最低的空间对称 $\beta = 0$ 态。为了以数值方式求解这一问题，我们在均匀间隔的网格上离散化相对坐标。与HF情形一样，利用傅里叶网格哈密顿量方法 [79]将动能项离散化为矩阵。我们采用 $N = 200$ 个网格点，范围从 $r = -5$到5。该问题的计算笔记本也见参考文献 [80]。

如果已有 ψ_β 的数值解，则总二体波函数为其乘积，

$$\Psi_{\alpha,\beta}(x_1, x_2) = \varphi_\alpha\left(\frac{1}{\sqrt{2}}(x_1 + x_2)\right) \times \psi_\beta\left(\frac{1}{\sqrt{2}}(x_1 - x_2)\right). \tag{33}$$

基态的总能量由 $E = \epsilon_0 + \varepsilon_1 = \frac{1}{2} + \varepsilon_1$ 给出。我们只讨论该系统的基态，但请注意，此方法也可用于求得 $A = 2$ 的其余能谱。

E. $A = 2$ 的微扰理论

或者，也可以使用标准微扰理论工具求解能量 ε_β，在方程（32）中。我们将谐振子作为参考态，因此 $\psi_\alpha^{(0)} = \varphi_\alpha$ 以及 $\varepsilon_\alpha^{(0)} = \epsilon_\alpha$，其中 φ_α 和 ϵ_α 分别是谐振子本征态和本征值。随后将高斯相互作用项视为微扰。我们采用矩阵元 $V_{\alpha\beta} = \frac{V_0}{\sqrt{2\pi}\sigma_0}\left\langle\varphi_\alpha\middle|e^{-\frac{r^2}{\sigma_0^2}}\middle|\varphi_\beta\right\rangle$ [82]的已知解析表达式。零阶结果就是无相互作用的受限双粒子系统的本征值，即方程（5），其基态给出 $E^{(0)} = 2$。注意到反对称性约束迫使 α 和 β 取奇数值，我们得到总能量的一阶微扰理论（PT1）表达式为：

$$E^{(1)} = E^{(0)} + V_{11} = 2 + \frac{V_0}{\sqrt{2\pi}}\frac{\sigma_0^2}{(1+\sigma_0^2)^{3/2}}. \tag{34}$$

正如预期，该表达式关于微扰强度 V_0是线性的。能量对 V_0 的依赖性由 σ_0决定。对于 σ_0的小值，斜率呈二次增长。换句话说，偏离非相互作用情形的程度关于 σ_0是二次的。相反，对于σ_0的大值，斜率按 $1/\sigma_0$减小。这是意料之中的，因为按照构造，我们的相互作用项的参数化具有这样的 $1/\sigma_0$ 依赖性。

二阶（PT2）和三阶（PT3）微扰理论结果可以直接从相应的矩阵元中获得 $V_{\alpha\beta}$。我们在中间态求和中最多使用八个额外态，实际上这已足以使结果收敛。我们在参考文献 [80]中提供了一个计算笔记本。正如我们将在下文看到的，这些 PT 估算使我们能够找到非微扰效应尤为重要的参数空间区域。

IV. 结果

在本节中，我们展示了使用不同方法对 $A = 2$ 到 $A = 6$ 的无自旋费米子系统所获得的结果。我们首先在第一小节中讨论二体情形。这使我们能够探究结果对 V_0和 σ_0的依赖性。关于 $A > 2$ 的结果将在下一小节，即第IVB节中讨论。

A. 二体扇区

高斯玩具模型中的二体情形已经相对复杂。零程相互作用的解众所周知 [81], 而特定有限程相互作用的半解析解也已得到 [41]。

1. *Energy*

能量对 V_0 和 σ_0 的依赖关系与更重系统的情况相似。我们尤其关注寻找非微扰行为的区域，以检验 NQS 试探函数 在最具挑战性的情形下的表现。我们首先考察能量对相互作用范围 σ_0 的依赖关系。需要指出的是，我们在此主要进行理论探讨，所研究的相互作用范围和强度值未必能通过近期实验直接实现。

我们首先考虑一种吸引力很强的相互作用，其 $V_0 = -20$。基态能量的结果如图 4所示。在图中，实线表示 NQS试探函数，其中包含统计上的变分蒙特卡罗不确定性，通常小于 10^{-5} （单位相同）。我们还展示了精确对角化（虚线）、HF（紫色短点划线）以及实空间（实心圆）的解。为了分析

我们还用不同线型展示了微扰结果 PT1、

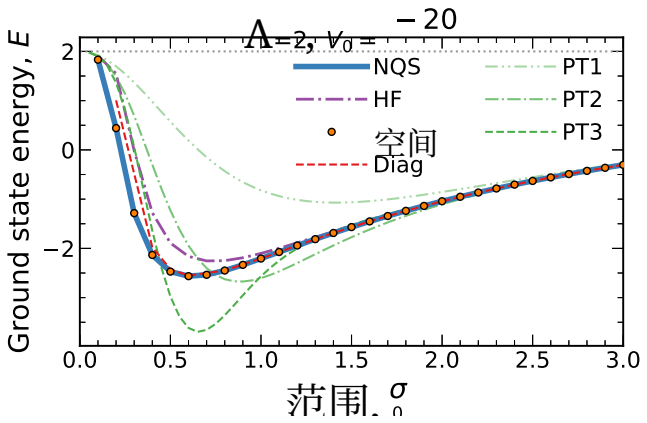

图 4. $A = 2$系统的基态能量随范围 σ_0 的变化，其中相互作用强度 $V_0 = -20$ 较大且具有吸引性。实线表示 $H = 64$、$L = 2$和 $D = 1$ 时的 NQS $Ansatz$ 能量。橙色圆点表示实空间解。短点划紫线表示哈特里-福克解。双点划线、点划线和虚线绿线分别表示一阶（PT1）、二阶（PT2）和三阶（PT3）微扰理论结果。水平点线表示非相互作用能量基准。

PT2 和 PT3 对能量的预测。

对于 $A = 2$ 个粒子，非相互作用情形对应于 $E = 2$, 在图中显示为一条水平点线。正如预期的那样，当范围趋于零时， $\sigma_0 \to 0$，所有方法都与该值一致。作为范围的函数，所有能量预测随后都会降低，达到取决于模型的最小值，最终增大并达到渐近 σ_0^{-1}行为。

不同方法之间的比较揭示了问题的复杂性。首先，我们注意到，在很宽的取值范围内，NQS 解与实空间解完全吻合。其次，能量关于 σ_0 的函数在$\sigma_0 \approx 0.5$ 附近达到最小值，并且在 V_0取该特定值时，能量为 $E \approx -2$ 。我们由此得出结论：对于 σ_0的很宽取值范围，网络都表现良好。

在很宽的取值范围内，NQS 的变分结果优于精确对角化和 Hartree-Fock 解。具体而言，Hartree-Fock 解的极小值稍浅，且出现在更大的 σ_0处。注意，Hartree-Fock 预测是由单粒子轨道构成的斯莱特行列式的最优解， $\phi_\alpha(x)$ [58, 60, 70]。Hartree-Fock 单粒子轨道依赖于单个位置。NQS 通过利用 GSM 广义轨道中的反流关联，优于这一解，详见IIIA4小节。在 σ_0取值非常小时，NQS 预测也优于精确对角化结果；此时，截断基底可能难以捕捉 $\sigma_0 \to 0$ 极限下出现的极窄结构。

微扰理论结果的收敛性

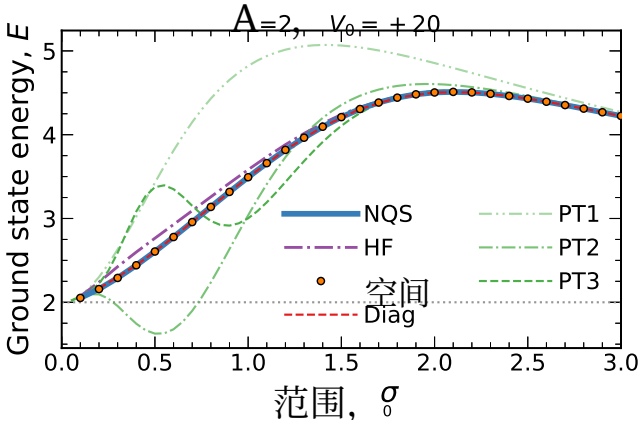

图 5。与图4 相同，但对于较大的排斥相互作用强度， $V_0 = +20$。

这表明了二体问题的“微扰性”。式（34）的 PT1（虚线双点线）结果在所有范围取值下都排斥性过强。只有远远超过其浅极小值（出现在 $\sigma_0 = \sqrt{2}$）时，该结果才接近 NQS 和实空间结果。对于较小的 σ_0，PT2 结果与 NQS 结果明显更接近，但在 σ_0 以下时，它们无法再现随 $\sigma_0 \approx 2$变化的正确关系。与 PT1 预测相反，当超过约 $\sigma_0 \approx 0.7$时，PT2 结果比 NQS 预测更具吸引性。PT3 也有类似情况：从 $\sigma_0 \approx 0.5$ 开始，其结果明显低于 NQS 解。我们发现，这些三阶结果仍与真实值相差甚远，尤其是对于 $\sigma \in [0.5, 0.9]$，其收敛模式相当不规则。这个参数空间中的困难区域或许可以成为检验 NQS 解的良好试验场。

通过考察一个非常强的排斥情形，我们确认这一范围区域确实具有特别显著的非微扰特征。我们在图 5中展示了 $A = 2$ 系统在 $V_0 = +20$取值下的基态能量预测结果。这里能量预测的形状明显不同。神经量子态能量在 $\sigma_0 = 0$处偏离非相互作用基线，并在 $E \approx 4$ 处$\sigma_0 = 2$增至约为最大值。与吸引情形一样，较大的范围会导致相对微扰性的结果，即 PT1、PT2 和 PT3 预测与精确基准方法一致。相反，当范围小于$\sigma_0 = 1$时，PT 预测呈现出复杂的模式。特别是，PT2 和 PT3 结果在 $\sigma_0 = 0.5$附近显示出异常振荡。总而言之，我们得出结论：对于数量级为 $V_0 \approx \pm 20$的相互作用强度，当相互作用范围的数量级为 $\sigma_0 = 0.5$时，会出现非微扰行为。虽然为简洁起见，我们在此不展示相关结果，但需要强调的是，对于具有 σ_0 的系统，NQS 和 HF 能量对范围 $A > 2$的依赖关系与图4和5中所示的情况非常相似。感兴趣的读者可以在参考文献 [30]中找到这些信息。

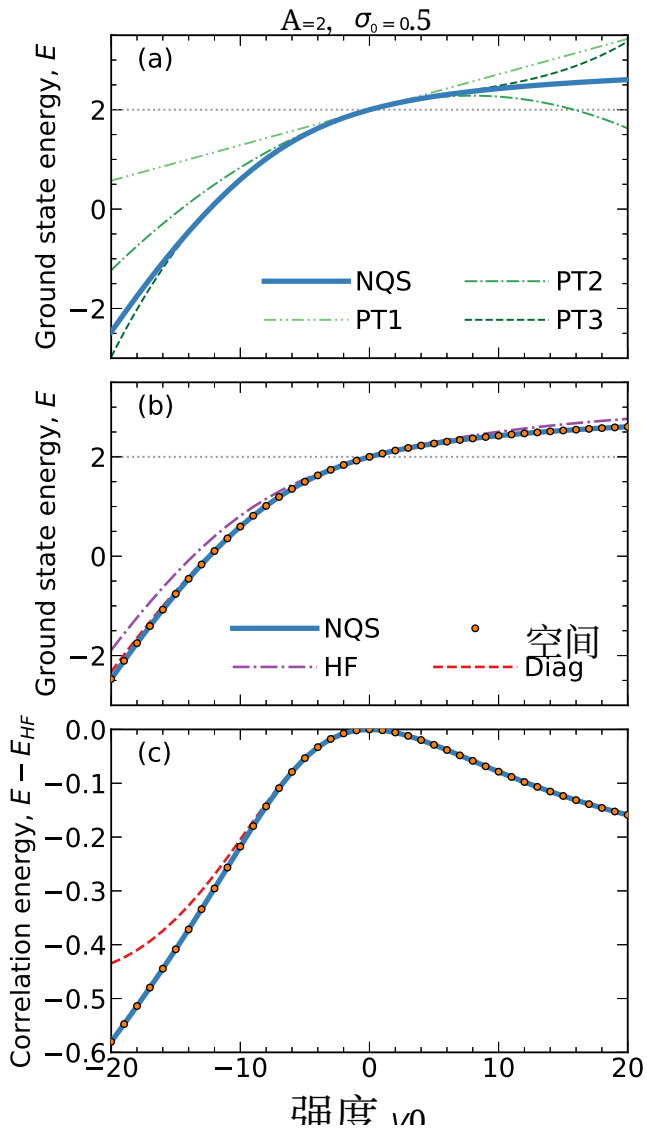

图 6. 在固定范围 $\sigma_0 = 0.5$ 下， $A = 2$ 系统的能量随相互作用强度 V_0 的变化。面板 (a)：将神经量子态基态能量（实线）与微扰理论结果进行比较。面板 (b)：将神经量子态基态能量（实线）与 $A = 2$系统的实空间解（实心圆）、直接对角化结果（红色虚线）以及 HF 近似（紫色短点划线）进行比较。面板 (c)：神经量子态、实空间解和直接对角化对关联能 $E_c = E - E_{\mathrm{HF}}$ 的预测。

对 σ_0 我们结果的依赖性进行分析后，我们得出结论：$\sigma_0 = 0.5$ 的取值处于非微扰区间之内。现在我们探究基态能量对相互作用强度的依赖性V_0 以分析神经量子态试探函数的性能。对于 $A = 2$ 的结果显示在图6的三个面板中。图6（a）展示了神经量子态基态能量（实线）与不同微扰理论结果之间的比较。这为了解系统的微扰性质提供了参考

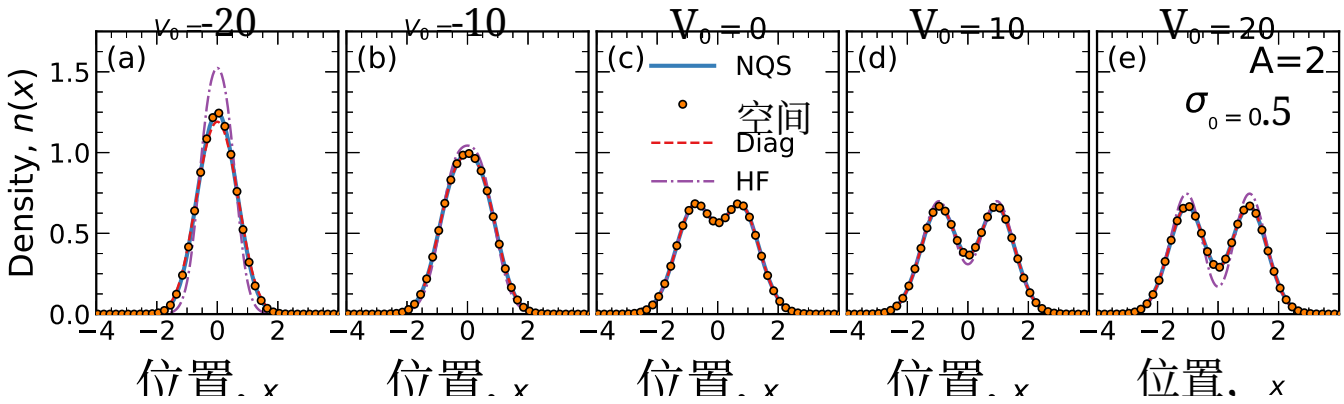

图 7。密度分布 $n\ (x)$ 的 $A = 2$ 系统随位置 x变化的情况。面板 (a) 至 (e) 展示了取值为 V_0 从 -20 到 20、步长为 10时的结果。参见图6 以了解图例的说明。

作为.....的函数 V_0，而不是 σ_0。根据式（34）得到的PT1 预测（双点划线）正如预期的那样，关于 V_0 呈线性关系，并且仅在 V_0 取小值时才能捕捉到关于 V_0的完整依赖关系，$|V_0| < 4$。PT2 结果（点划线）在更宽的强度范围内有效，但在 V_0取较大的负值（正值）时高估（低估）真实的基态能量。PT3 结果（短划线）在范围 $-15 < V_0 < 10$内表现良好，但超出该范围后则失效。再次，这一分析表明，V_0 最具吸引力和排斥力的取值处于非微扰区间。

我们在图中将 NQS 结果与更现实的基准进行比较。6 （b）。总体而言，我们发现实空间解与 NQS 预测之间具有极好的一致性。直接对角化结果也提供了良好的能量预测，尽管在耦合的大负值处观察到相对于NQS的轻微偏差。得益于回流关联，NQS 提供了比 HF 近似更好的变分结果。

现在我们来讨论能量对 V_0 的整体依赖关系。在非相互作用极限 $V_0 = 0$ 下，所有方法都与基准值 $E = 2$ 完全一致。在排斥侧，随着相互作用强度的绝对值增大，能量通常高于基准值，并且似乎趋于饱和，达到一个接近 $E \approx 2.5$ 的值。相比之下，在吸引侧，能量下降得相对较快。当达到 $V_0 \approx -10$ 左右时，系统能量变为吸引性。随后，随着 V_0 越来越具有吸引性，总能量的绝对值逐渐减小。从系统的能量特性可以立即看出，吸引区域与排斥区域截然不同。

这一差异在图 6(c) 中更加明显，该图展示了系统的关联能 E_c。我们遵循教科书中的惯例 [60]，并将 E_c 定义为给定方法得到的基态能量与 HF 预测之差，$E_c = E - E_{\mathrm{HF}}$。我们注意到，在二体层次上，这一量可能并不能提供太多有意义的信息，因为预计 HF 预测在 $A = 2$ 区间内表现不佳。

关联能从 $V_0 = 0$ 开始偏离零。在排斥侧，E_c 的绝对值缓慢增大。在 $V_0 = 20$ 处，我们得到 $E_c \approx -0.15$，约占总能量的 6%。相比之下，在吸引区间，NQS 预测为 $E_c \approx -0.6$，在同一取值下，这对总能量 $E \approx -2$而言是一个相当大的贡献。在 V_0 的两侧，关联能都没有饱和的迹象。我们还注意到，当 $V_0 < -10$ 时，直接对角化结果与实空间结果及 NQS 预测均不一致。正如接下来将看到的，这种差异是由模型空间截断造成的。

2. *Density profile*

为了进一步理解系统关联的起源，我们考察系统的密度分布，$n(x)$。图7展示了采用不同多体方法得到的密度分布，对应五种不同的 V_0取值，范围从 -20 到 20，步长为 10。图7(c) 展示了非相互作用结果，$V_0 = 0$。正如预期，所有方法的结果一致。密度分布在中心处有一个广为人知的凹陷，这是 $\alpha = 0$ 与 $\alpha = 1$ 谐振子轨道相互作用的结果。这种密度调制在费米子系统中很常见 [41]。图7(d) 和7(e) 展示了排斥区域中相同的密度分布，其相互作用取值分别为 $V_0 = 10$ 和 20。我们发现，密度中心处的凹陷减小，而密度分布的整体尺寸基本保持不变。实空间 $A = 2$ 问题的解、直接对角化方法和神经量子态的预测结果非常吻合。相比之下，Hartree-Fock 结果似乎高估了费米子结构，表现为凹陷更浅、密度极大值更高。换言之，Hartree-Fock 轨道似乎比其关联对应物更加局域化。在电子气的三维维格纳晶体中，也观察到了 Hartree-Fock 预测与完全关联预测之间类似的局域化模式差异 [83]。

从物理角度看，排斥区域呈现的图景类似于局域化或（维格纳）晶化 [23, 84–89]。随着粒子间相互作用强度增大，系统通过使粒子彼此间隔超过相互作用范围来最小化相互作用能。然而，如果粒子远离谐振子势阱中心，这样做可能会以势能增加为代价。因此，系统倾向于使轨道更强烈地局域化为周期结构，从而降低中间凹槽处的粒子间密度。

相比之下，吸引区间中的密度分布，如图 7（a）和 7（b）所示，其特征是结构相对不明显。费米子凹陷在 $V_0 \approx -10$ 附近消失；对于更大的吸引强度值，密度分布呈现为一个简单的峰，随着 V_0 变得更加负，其宽度变窄而高度增加。事实上，一维系统中的一项长期预测表明，具有强吸引相互作用的无自旋费米子系统应表现得像非相互作用玻色子系统 [35, 36, 39–42、90–92]。对于 $A = 2$ 系统而言，密度分布明确支持了这一假设。值得强调的是，神经量子态无需任何进一步调整就捕捉到了玻色化转变。我们还要强调，我们的结果在定性上与文献 [41], 的半解析解相似，其中使用了软核方形势阱相互作用。

总体而言，系统的密度分布极其有用，因为它为玩具模型的两种截然不同的相提供了直接证据。吸引区间在系统密度中显示出玻色子性质的迹象，这一预测与此前采用奇对称相互作用势的研究结果相符 [35, 36,39, 40, 42, 90–92]。关于宇称偶势的预测非常少 [41, 55], 但我们在玻色子相中的不同基准方法之间取得了一致结果，这表明这一图景相当稳健。此前一项在 $A = 2$ 情形 [41] 下采用偶宇称软核势的理论研究，在排斥相中发现了类似特征。在那里，粒子倾向于局域在相互作用范围之外，远离排斥核，从而使能量最小化。事实上，对于排斥相互作用，同一项研究表明，局域相对于费米子和玻色子是相同的 [41]。总而言之，现有的二体层次证据表明，对于偶宇称势，在吸引区间和排斥区域中都存在费米子-玻色子对偶性的迹象。

3. *Occupation numbers*

密度分布 $n(x)$ 只是单体密度矩阵的一个分量。我们可以使用目前为止讨论的基准方法获取单体密度矩阵。对于神经量子态，我们使用参考文献 $\rho(x_1', x_1)$ 中的幽灵粒子方法随机计算 [93]。对于实空间解，我们在均匀网格上对方程（6）执行积分。该解-随后即可在空间均匀网格上将式（7）作为矩阵对角化问题轻松求解。采用直接对角化方法，我们在二次量子化形式下计算 ρ ，其定义为

$$\rho_{\alpha\beta} = \langle \Psi | \hat{a}_\alpha^\dagger \hat{a}_\beta | \Psi \rangle \,, \tag{35}$$

其中， $|\Psi\rangle$是预先计算得到的基态波函数， $\rho_{\alpha\beta}$ 是 OBDM 的矩阵元， ρ。最后，我们对 ρ 进行对角化，从而在多体基组中得到本征值 n_α 。

图 8（a）展示了 $A = 2$ 系统的占据数随相互作用强度的变化。我们将 NQS 结果（实心圆）与实空间解（空心圆）和直接对角化方法（实线）进行对比。我们发现，这三种方法之间具有极好的一致性。需要指出的是，我们没有将结果与 HF 基准进行比较；HF 能够较好地描述系统的能量。然而，对于占据数而言，HF 方法仅限于完全占据态或完全未占据态，无法捕捉占据数相对于非相互作用情形的偏离。

对于 $A = 2$，自然轨道占据存在二重简并，NQS 能够无缝处理这一情况。空穴占据值 $n_{\alpha<A}$ 会随着 V_0 绝对值的增大而偏离 1。我们发现，变化速率取决于系统处于吸引侧还是排斥侧；这一结果与图 6（c）中显示的关联能相一致。在排斥侧，对于 $V_0 = 20$，占据数约为 $\approx 96\%$，表明系统的关联相对较弱。然而，随着 V_0 增大，占据数偏离 1 的趋势表明，即使处于晶体相，粒子也不能完全由局域束缚的单个轨道来描述 [41]。相比之下，在吸引情形下，对于 $V_0 \approx -20$ ，占据数为 $n_\alpha \approx 0.9$。这表明在玻色化相中，关联发挥着更强的作用。

图 8(b) 展示了“未占据”轨道的类似图示， $n_{\alpha \geq A}$。该图以对数刻度绘制，并包含 NQS 对角化中的前 14 个轨道。由于 NQS 方法中 OBDM 估计具有随机性，因此存在显著的统计噪声底，其量级为 $\approx 10^{-2}$。尽管这限制了我们能够得出的一些结论，但我们强调，NQS 多体波函数在变分意义上与其他估计结果相当（如果不是更优的话）。因此， n_α高于统计阈值的数值应能更忠实地表示真实的占据数。

与图8(a) 类似，我们发现排斥侧和吸引侧在定性上有所不同。在排斥区域，我们只看到一个占据数大于 10^{-3}的二重简并粒子态，其占据数达到接近 $n_\alpha \approx 0.02 - 0.03$ 的数值，条件是 $V_0 = 20$。相比之下，在吸引区间，我们发现两个占据数高于 10^{-3}的二重简并态。这里占据数较高的态在 $n_\alpha \approx 0.08$ 时达到 $V_0 = -20$。第二个二重简并态也得到了

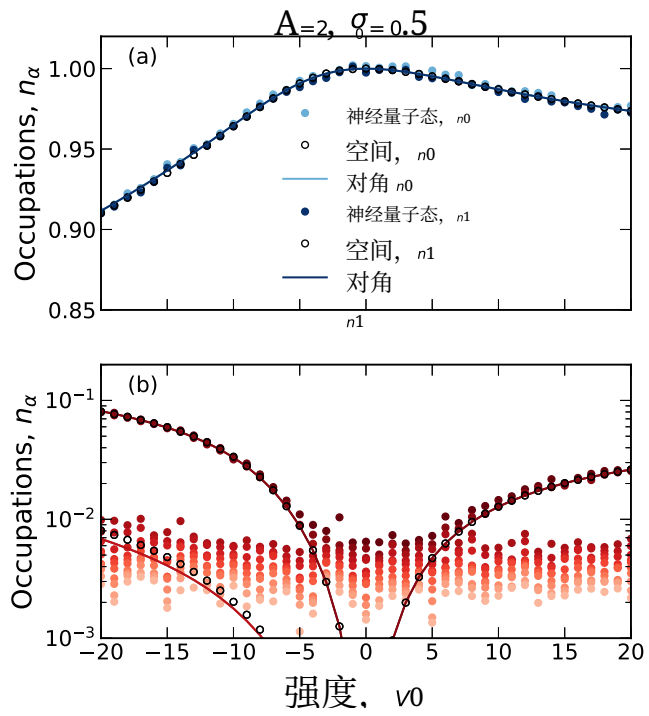

图 8。对于固定范围 $\sigma_0 = 0.5$，系统 $A = 2$ 的占据概率随相互作用强度的变化。面板(a) 显示费米面以下两个空穴态的占据概率。面板(b)以对数刻度显示粒子态的占据概率。我们将NQS结果（实心圆）与实空间结果（空心圆）进行比较。实线对应直接对角化结果。

精确对角化和实空间解的预测。其占据数约为占据数最高态的十分之一。神经量子态（NQS）无法分辨这一占据数，因为它落在统计噪声底之内。不过，总体而言，这一图景进一步印证了这样的观点：吸引侧和排斥侧的关联具有截然不同的性质。先验来看，玻色化相似乎具有更复杂的结构，其中混合了比相应晶体相更多的单粒子模式。

B. 少体扇区

现在，我们将讨论转向少体扇区中系统的结果，系统包含 $A = 3$ 至多$A = 6$ 粒子。我们继续将 NQS 结果与直接对角化方法以及哈特里–福克方法进行比较。$A > 2$没有实空间解。我们进一步指出，我们在二体系统中得出的若干物理结论在少体领域仍然成立。

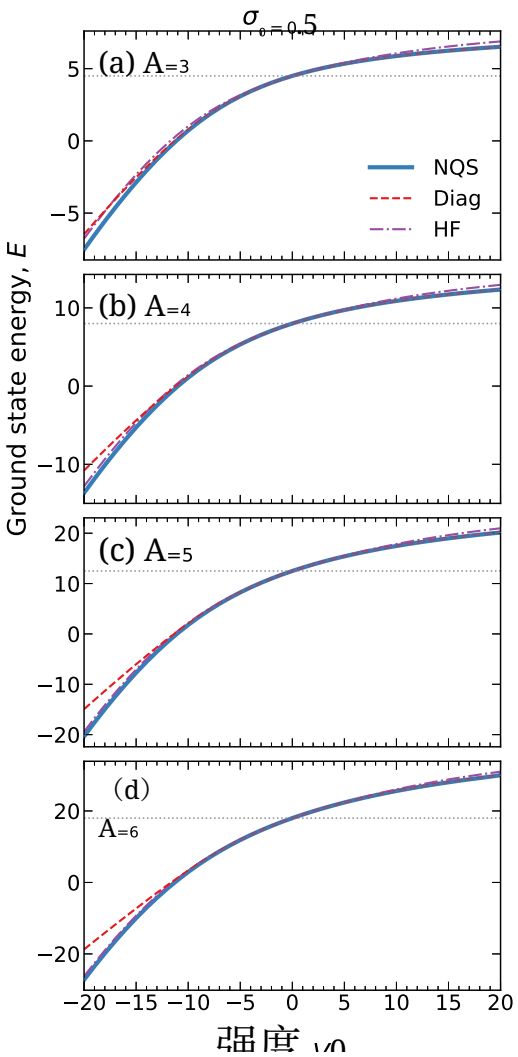

图 9. A粒子系统的基态能量随相互作用强度 V_0 的变化，其中范围固定为 $\sigma_0 = 0.5$。面板 (a) 至 (d) 分别对应由 $A = 3$ 至 $A = 6$ 个粒子组成的系统。关于图例的解释，请参见图 6 。

1. *Energy*

我们首先考察系统的能量。图9(a)-9(d) 展示了 $A = 3$ 至 6 个粒子的能量随 V_0的变化。我们重点关注范围取值为 $\sigma_0 = 0.5$的情形。所有这些系统的能量定性行为都与 $A = 2$ 情形非常相似。在排斥侧，系统能量缓慢上升，并趋于在约为非相互作用值 $1.25 - 1.6$倍处达到平台。在这一排斥范围内，粒子数 $E \approx A^{2.2}$的幂律标度能够很好地近似完整结果。当 V_0 变得极具排斥性时，能量趋于饱和，这与晶体的物理图像相符。稍后给出的密度分布进一步印证了这一图像。相较之下，在吸引区间，单粒子能量随 V_0下降得陡峭得多，其中

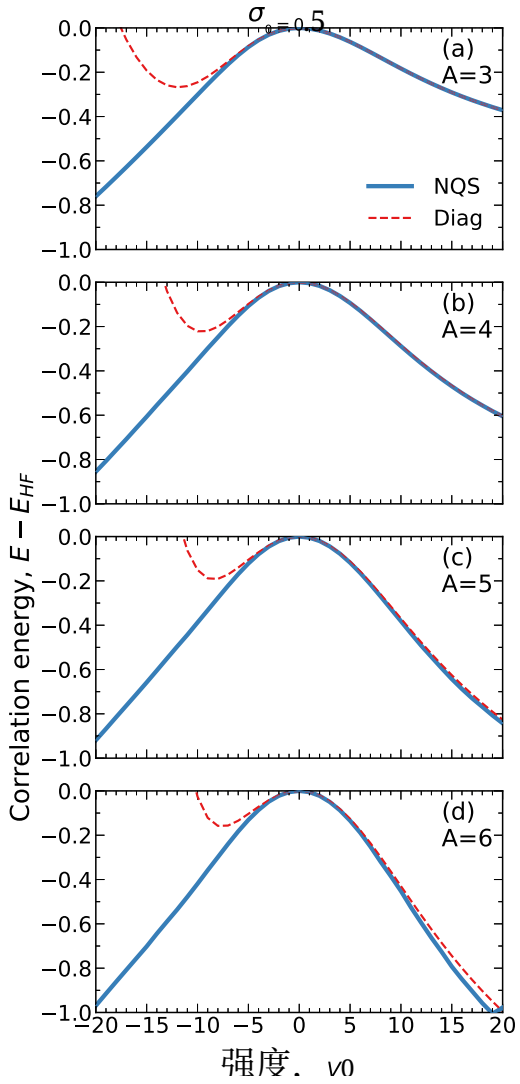

图10。 A粒子系统的关联能是相互作用强度 V_0 的函数，对于固定范围 $\sigma_0 = 0.5$。面板(a)至(d)分别对应 $A = 3$ 至 $A = 6$ 粒子的系统。关于图例的说明，请参见图 6。

没有显示出饱和迹象，因为 V_0 变得更加负。

关于基准方法，我们注意到两种不同的显著趋势。首先，HF 结果高估了基态能量，这符合预期，因为其中缺少回流。随着粒子数增加，HF 与 NQS 预测之间的差异变得不那么明显。换言之，随着粒子数增加，HF 近似在相对能量方面变得更加可靠。这并不奇怪，因为粒子数增加时，平均场图像应当更加适用。其次，对于 $V_0 < -10$，我们发现精确对角化方法得到的束缚总是弱于 NQS 预测。我们将此归因于有限模型空间；在强吸引区间，该空间并不足够。根据图11的结果，在玻色化区间，密度

该系统相比非相互作用谐振子情形要压缩得多。因此，可以预期需要更多的单粒子模式来描述系统的密度分布和能量性质。

为了进一步量化系统中关联的重要性，我们考察了系统中$A = 3$ 至 6 个粒子的关联能随 V_0 的变化，如图10所示。需要强调的是，所有面板采用相同的能量尺度。换言之，HF 与完整结果之间的差异似乎相对恒定，不受粒子数影响。在 $V_0 = 20$处的排斥区域，每向系统中增加一个粒子，能量就增加约 0。2 个单位。相比之下，在吸引区域，能量增长较为平缓。需要强调的是，在神经量子态中，非零关联能直接源于试探函数中的回流关联试探函数。值得注意的是，这种扩展使神经量子态能够捕捉关联能对 V_0的非对称依赖，这归因于系统所处的不同相。

在吸引侧，直接对角化方法对关联能的描述相对较差。对于低于 $V_0 < -10$ 或 -15的耦合常数，可得到 E_c 的非物理正值。在排斥侧，对于 $A \approx 6$，我们同样发现直接对角化方法对关联能的描述相对较差的迹象。话虽如此，很明显，直接对角化方法在排斥侧的表现要好得多，在能量的绝对值方面，与 NQS预测的偏差均在 0.05 以内。

2. *Density profiles*

检查系统的密度分布也非常有用， $n(x)$ ，对于包含 $A = 3$ 至 6 个粒子的系统。我们在图 11中总结了这些结果，其中各列对应不同的相互作用强度，各行对应不同的粒子数。我们比较了神经量子态（实线）、直接对角化方法（虚线）和 HF 近似（短点划线）的预测结果。

中间的面板对应非相互作用情形（$V_0 = 0$），显示出 A 个峰值叠加在类高斯的整体行为之上。这是预期的行为，由方程（3）和（27）所导致。在吸引区间，从 $V_0 \approx -10$ 开始，密度分布中的类峰值费米子结构消失。神经量子态结果显示出单个峰值，并且随着V_0 的吸引性增强，该峰变得更窄。这与 HF 预测一致；随着粒子数增加，HF 预测逐渐接近神经量子态结果。相比之下，直接对角化方法无法捕捉超过 $V_0 \leq -10$时的狭窄密度分布。在 $V_0 = -20$时，神经量子态和 HF 的密度分布几乎比非相互作用情形窄两倍。因此，很自然地会¨预期，直接对角化中的模型空间截断无法捕捉如此强烈的密度

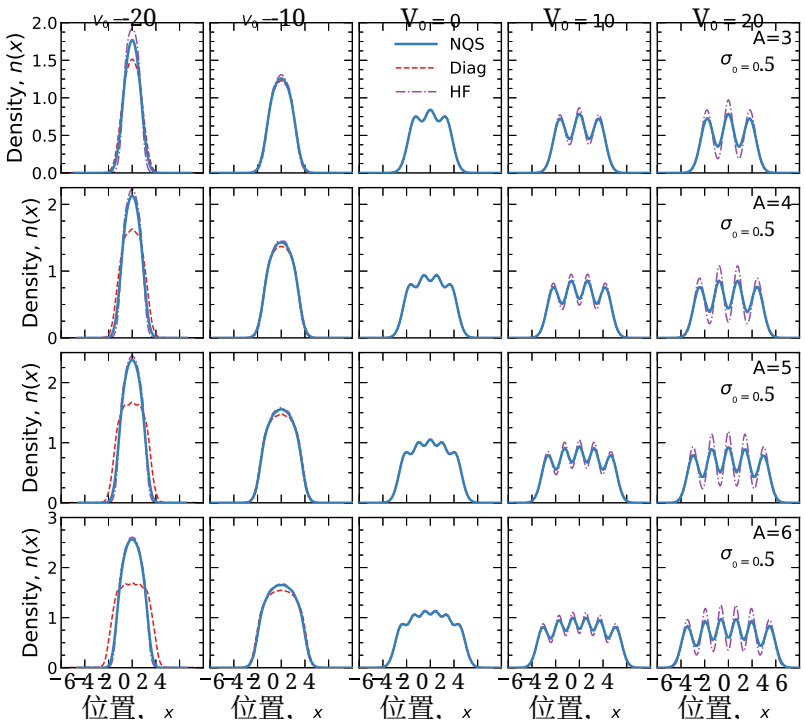

图 11. 密度分布 $n(x)$，作为位置的函数 x，对应不同的 V_0 取值和固定范围 $\sigma_0 = 0.5$。从左到右 V_0 从 -20 到20递增，以 10 个单位为间隔。从上到下，我们展示 $A = 3$、4、5 和 6 个粒子的结果。

重新排列。

在强排斥极限下，$V_0 \gg 0$相互作用有效地将费米子彼此分开，从而使系统完全晶化 [85, 88]。然而，单粒子轨道的分离不可能无限进行，因为随着 $|x|$ 的增加，谐振阱变得越来越有效。虽然 A密度峰值的位置几乎不变，但随着 V_0 的增加，它们之间的谷值变得越来越明显。在这一排斥区域中，直接对角化得到的密度分布结果与 NQS 预测一致。相比之下，HF 方法存在显著不足。HF 密度分布会高估（低估）峰值（谷值），且随着粒子数增加，偏差变得更加明显。我们强调，密度分布中的差异源于我们的 NQS 试探函数中存在回流关联。关于密度分布和系统平均尺寸的更多细节，请参见文献 [30]。

吸引区域和排斥区域的密度分布差异反映在系统的能量特性中。我们展示了总能量中不同部分随相互作用强度的变化V_0 见图12。从上到下的各面板对应不同的粒子数，从 $A = 2$ （顶部）到$A = 6$ （底部）。我们区分由动能 $\langle K\rangle = \langle\Psi|\sum_i -\frac{\partial_i^2}{2m}|\Psi\rangle/\langle\Psi|\Psi\rangle$、（外部）谐振子势能 $\langle U\rangle = \langle\Psi|\sum_i -\frac{x_i^2}{2}|\Psi\rangle/\langle\Psi|\Psi\rangle$以及相互作用能 $\langle V\rangle =$所产生的各项贡献。

$V_0/(\sqrt{2\pi}\sigma_0)\langle\Psi|\sum_{i<j} e^{-(x_i^2-x_j^2)/2\sigma_0^2}|\Psi\rangle/\langle\Psi|\Psi\rangle$。这些组件均使用神经量子态波函数概率以随机方式计算。

我们发现了一个相对独立于粒子数的共同图景。在非相互作用情形下，维里定理规定 $\langle E\rangle = 2\langle K\rangle = 2\langle U\rangle$（以及 $\langle V\rangle = 0$）。在排斥侧，随着 V_0 的增长，谐振子能量和相互作用能分量增长得非常缓慢。我们将这种缓慢增长解释为局域化的结果。当单粒子轨道局域化为 A 定义明确且等距的峰值时，费米子特征的锐化只会略微改变这两个分量。相比之下，动能随 V_0显著降低。事实上，对于粒子数 $A > 4$，我们观察到 $\langle K\rangle \leq \langle V\rangle$，这一条件通常用于表征（维格纳）晶体 [88]。

从不同能量分量的角度进行分析，表明吸引区间的能量表现相当不同。首先，我们发现，在 $V_0 \ll 0$ 的区间内，吸引总能量是两个数值很大、但符号相反的能量分量相互抵消的结果。一方面，随着 V_0 减小，动能 $\langle K\rangle$ 大幅增加，这是密度强烈重排的结果。另一方面，当所有粒子都被限制在陷阱中心附近、因而彼此也相距很近时，相互作用能 $\langle V\rangle$ 会变得极具吸引性。中心约束也使谐振子势能 $\langle U\rangle$几乎完全抵消。

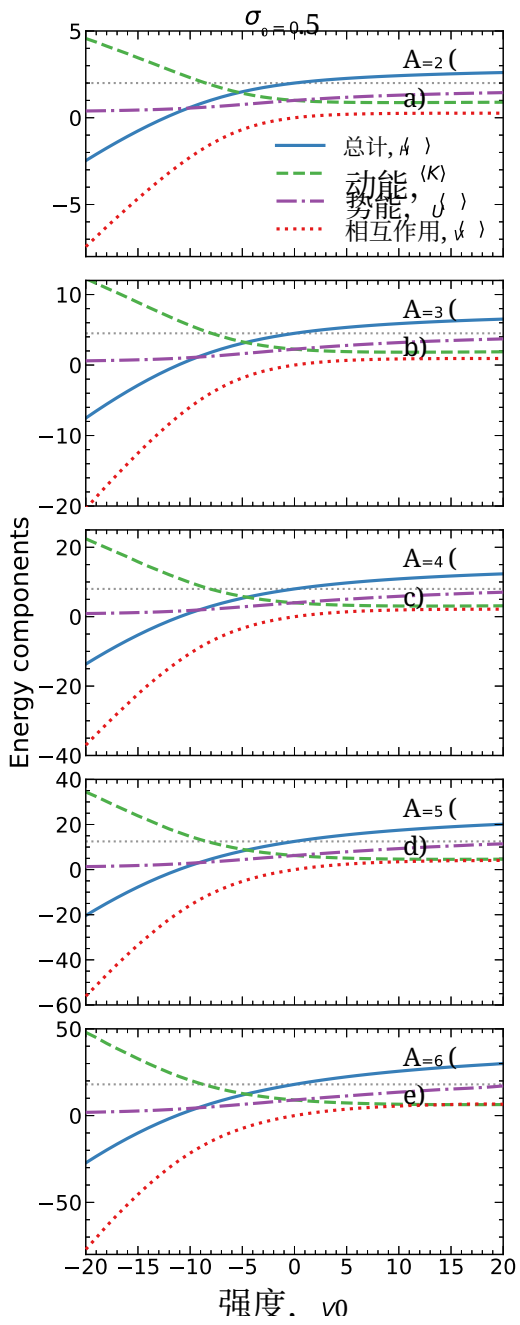

图12. NQS 基态能量的不同组成部分随相互作用强度 V_0 的变化，其中范围 $\sigma_0 = 0.5$ 固定。实线、虚线、点划线和点线分别表示总能量、动能、谐振子能量和相互作用能。面板 (a) 至 (e) 对应粒子数从 $A = 2$ 到 $A = 6$ 的系统。水平点线表示非相互作用基态能量。

此时，我们可以对能谱排斥侧和吸引侧的关联性质得出一些结论。显然，吸引侧的玻色子转变会导致密度分布发生剧烈重排。事实上，这种重排如此强烈，以至于基于非相互作用的直接对角化

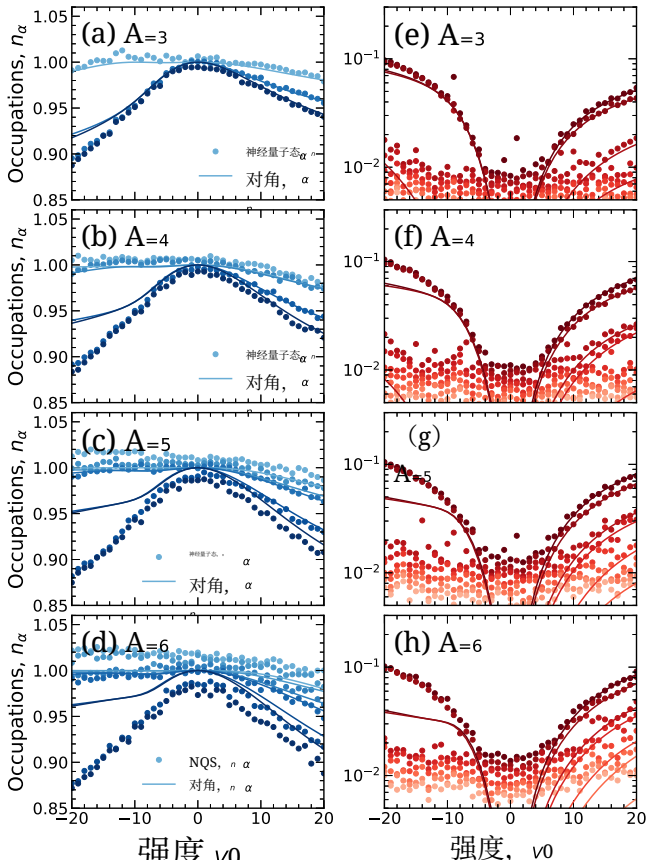

图13. 面板 (a)-(d)：固定范围 $\sigma_0 = 0.5$ 时，$A = 3$ （上排）至 $A = 6$ （下排）系统中空穴（$\alpha < A$）的占据概率随相互作用强度 V_0 的变化。我们将 NQS 结果（实心圆）与直接对角化预测（实线）进行比较。面板 (e)-(h)：粒子（$\alpha \geq A$）态（最高至 $\alpha = 14$）的相同结果，采用对数刻度。

作用于单粒子态的方法难以捕捉到这一点。准确描述这一转变需要改为在实空间（或者可能在动量空间）中求解系统。我们注意到，无自旋费米子的密度分布也呈现出定性相似的结果，$A = 3$ 以及 4 个无自旋费米子的相关结果由Schilling等人 [56]采用偶宇称谐和模型进行了报道。对于吸引相互作用，在该情形下也观察到了普适的玻色子峰。

相比之下，排斥区域中的密度重排主要导致单粒子态的局域化，从而夸大了相关的费米子特征。在采用谐振子模型 [56] 的少体系统中也发现了这一点。然而，在此极限下，Hartree-Fock 方法会高估费米子特征，这再次凸显了回流关联的重要性。

3. *Occupation numbers*

通过考察系统的占据数 n_α，我们可以进一步了解其关联结构，这些占据数对应于

不同的单粒子态 α。我们展示了 n_α 对于 $A=3$ （上排）至 $A=6$ （下排）的系统，见图13。左（右）侧面板对应于空穴（粒子）的占据概率， $n_{\alpha<A}$ ($n_{\alpha\geq A}$)。吸引区域和排斥区域的占据数呈现出明显的共同结构。在吸引情形下，神经量子态预测会出现一个双重简并态，随着 V_0 变得越来越负，该态的占据显著减少。这一结果与直接对角化对占据数的预测相符；后者也显示出双重简并的趋势，不过占据减少程度略小。我们再次将神经量子态与直接对角化预测之间的差异解释为截断模型空间所致。神经量子态和对角化的预测均表明，其余 $A-2$ 空穴态的占据率在 2% 的误差范围内为满占据，不过，神经量子态方法中的统计噪声底使得这一结论难以定量确定。

图 13(e)-13(h) 中所示的粒子态显示，在排斥区域中，对于 A 的所有取值，都会出现一个二重简并态。根据 NQS 的预测，这些态在 $\approx 10\%$ 处达到 $V_0=-20$ 的占据数。这些二重简并态让人联想到费米自由度的玻色化。尽管不存在活性自旋，这种现象仍会发生，因此只能在有限范围相互作用的情况下观测到。在一维（1D） p-波费米子 [42] 的变分计算中，也观测到了吸引区间内的二重简并结构。然而，这些模拟在描述排斥区域时更加困难，而 NQS 则可以无缝地处理该区域。

在排斥区域中， $V_0>0$，未观察到二重简并。单空穴态 [图13(a)-13(d)] 和单粒子态 [图13(e)-13(h)] 均各不相同，并且可以清楚地区分每个面板中排斥支上的 A 态，其占据数为 $n_\alpha > 5\times10^{-3}$。这一预测得到了直接对角化和神经量子态结果的共同支持。对于 $A=3$ 系统，在图13(a)中，我们发现出现了三个不同的单空穴态，其耗尽程度很小，占据数超过95%，直至 $V_0\approx20$。图13(e)中对应的单粒子态聚集成两个近乎简并的态，在 $V_0\approx20$处的占据数为 ≈5%，另有一个态的占据数为2%。随着 A增加，出现了类似的图景。对于 A的偶数值，我们发现 $A/2$ 态存在近似（但并不完全）简并。在单空穴情形中，占据数随着V_0 成对地稳步降低。类似地，粒子态的占据数也以成对的形式增加，且数值相近。对于奇数 A，图景基本相同，只是存在一个具有特定占据概率的奇数单粒子轨道。我们注意到，这一图景与文献 [56] 中针对具有谐振子型成对相互作用的一维自旋极化受限费米子所得到的结果存在相似之处。

对于单粒子态和单空穴态，统计噪声底约为 $\approx10^{-2}$。我们发现，该噪声底会随粒子数略有增加，增幅约为从 $A=3$ 变为 $A=6$时相差 2 倍。尽管可以在单粒子态中看到统计噪声，但其影响在单空穴态中更加明显，在这些态中，噪声会导致带有 $n_\alpha>1$ 的非物理涨落。我们在此分析中任意将单粒子态的数量限制为 14。对于 V_0的所有取值，噪声底都保持一致，这表明其来源是用于计算单体密度矩阵的随机方法。我们将单空穴态所显示的占据值范围限制为高于 5×10^{-3}，这突出显示了使用直接对角化方法获得的若干最低单空穴态（实线）。

4. *One-body density matrix*

到目前为止，我们讨论的都是系统的局域性质。现在，我们将注意力转向密度矩阵所反映的非局域结构。在本节及后续各节中，我们将讨论 NQS 结果，并请读者参阅附录，了解相应的 HF 和直接对角化结果。在图 14 中，我们展示了从 $A=2$ （顶行）到 $A=6$（底行）的基态波函数的单体密度矩阵 $\rho(x_1',x_1)$。不同列对应于从 $V_0=-20$ 到 20 的相互作用强度值（步长为 10）。中间各面板对应于非相互作用情形 V_0；结合式（27）（该式适用于非相互作用情形）与式（3）中的谐振子本征态，可以用相对简单的解析方式理解这一情形，最终得到式（10） [57]。在所有情形下，密度矩阵都完全为实数，并且在 $x_1'-x_1$ 平面上覆盖一个或多或少呈正方形的区域。这是式（27）求和中的单粒子项 $\varphi_n^*(x_1')\varphi_n(x_1)$ 共同作用的结果；这些项自然地为 $x_1'-x_1$ 平面的对角线 [94] 内外都提供了界限。这一点在式（10）中表现得十分清楚，其中 $\rho(x_1',x_1)$ 由 $x_1'-x_1$ 中的高斯包络乘以一个多项式调制得到。

此外，可以发现对角线附近存在一片相对狭窄的正值区域，在该区域内 $\rho(x'=x,x)\equiv n(x)>0$, 为正定。远离对角线时，在正负值交替出现的区域中会显现出明显的波纹。这些波纹反映了单粒子态的节点结构。事实上，在对角线一侧沿 $x_1'=-x_1$方向的符号变化次数，等于系统中的粒子数。方程（10）从沿 $A-1$方向的 2($x_1'=-x_1$) 阶多项式的根解释了这一现象。

在左侧面板所示的吸引情形中，系统的整体空间范围缩小了（见图11）。类似地，OBDM 在 $x_1'-x_1$ 平面上的支撑范围也有所缩小。此外，沿非对角方向与符号变化相关的峰值和谷值向对角线靠近。值得注意的是，随着相互作用强度变得更加负，这些干涉效应的幅度似乎也在增大。我们强调，

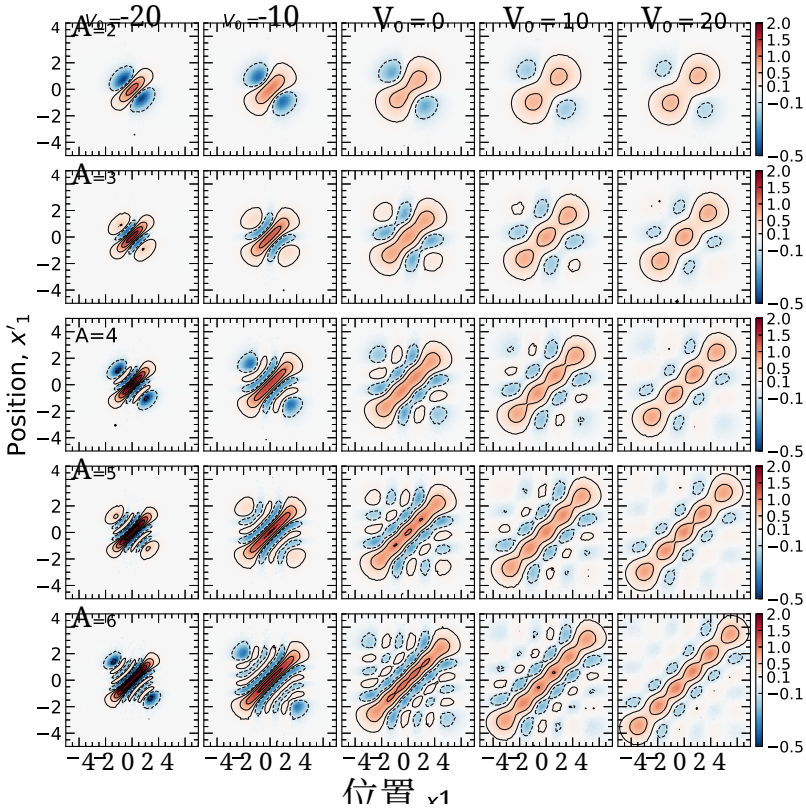

图 14. $\rho(x_1', x_1)$的单体密度矩阵密度等高线图，展示了从 $A = 2$ （顶行）到 $A = 6$ （底行）个粒子的系统。各列对应不同的相互作用强度，从 $V_0 = -20$ （左列）到 $V_0 = 20$ （右列）。等高线显示的密度矩阵值为 $-0.5, -0.1$、0.1、0.5、1.0、1.5和 2.0。所有面板中的颜色编码相同。

对这种强烈振荡的 OBDM 进行对角化，会得到具有占主导地位的双重简并结构的占据数，这更令人联想到相干图景。

在强排斥情形下，在右侧的面板中，可以清楚地区分出单粒子轨道沿对角线的局域化，其形式为定义良好且等距分布的峰值。在非相互作用区间和吸引区间中观察到的非对角振荡及符号变化，在此情形下都受到了显著抑制。例如，在 $V_0 = 20$处，只能观察到一系列定义良好的、沿平行于对角线方向排列的带负值的 $A - 1$ 准圆形区域。尽管沿非对角方向还存在额外的振荡，但其振幅要小得多，以至于在图的尺度下几乎无法辨别。

5. *Pair correlation function*

最后，我们在图 15 中展示了不同粒子数下，式（ g （x_1， x_2）的对关联函数（PCF）， g（x_1， x_2）11）。各面板的排列方式与图 14 相同。对关联函数提供了关于由系统的反对称性质以及相互作用 [54]所引起的关联的信息。函数 g（x_1， x_2）在吸引区域和排斥区域中有很大不同，但存在一些共

所有面板中的共同结构。具体而言，我们发现一个位于 $x_1 \approx x_2$ 附近的脊线，此处对关联函数消失。这是泡利不相容原理的结果；对于无自旋费米子，该原理施加了条件$g(x,\ x) = 0$。在非相互作用系统中，该条件由方程（$x_1 - x_2$）^{2} 中的（12）项强制满足。然而，这条“泡利脊”的宽度会随相互作用强度发生显著变化，这表明其中一部分也是相互作用效应的结果。在排斥侧，即中心右侧的面板中，可以清楚地发现，脊线随 V_0变宽。这一效应反映了图 7和 11中观察到的轨道空间局域化。我们预计，排斥侧的脊线大小对 V_0 以及相互作用的范围 σ_0都很敏感。当 V_0 更加排斥时，粒子会趋向于进一步局域化，从而在对关联函数中形成间距更大且定义更清晰的峰值。

相比之下，在图左侧面板所示的吸引区间内，我们发现，随着 V_0 变得更具吸引性，脊线会变窄；当 $A > 5$时，在带有 $V_0 = -20$的最左侧面板中，脊线已难以辨认。吸引侧的对分布函数整体较为平淡，除了脊线之外，在中心周围近似方形的区域内还呈现出极高的峰值。这表明，在距离中心位置数个单位范围内找到一个粒子的可能性很高

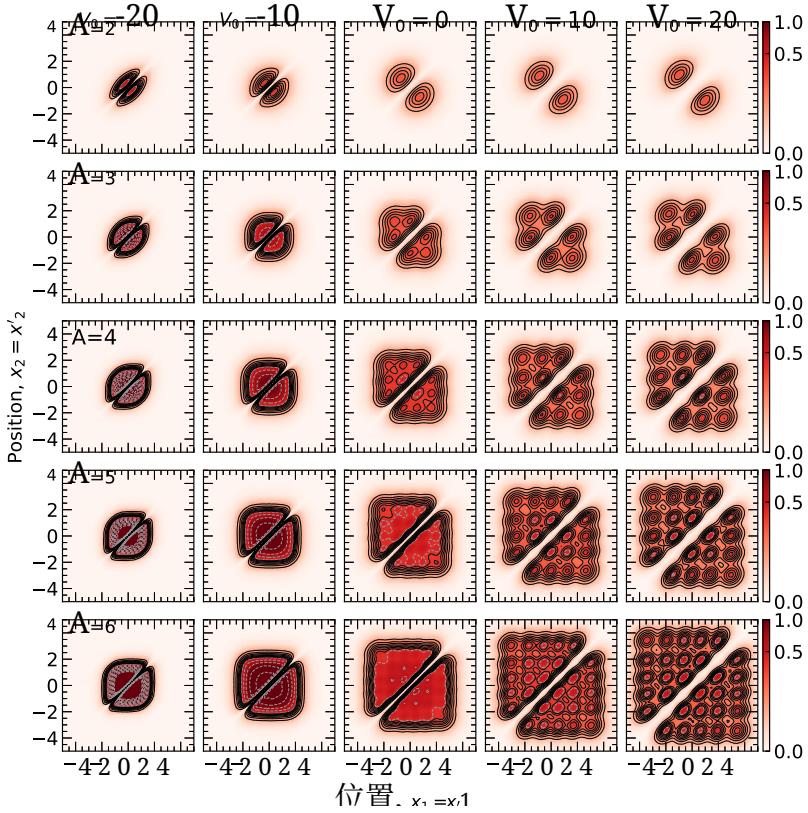

图15。对于从 $A = 2$ （顶行）到 $A = 6$ （底行）个粒子的系统， $g(x_1,\ x_2)$的对关联函数密度等高线图。各列对应不同的相互作用强度，从 $V_0 = -20$ （左列）到 $V_0 = 20$ （右列）。等高线以 0.05 为步长，从0到0.35（实线）绘制；并以 0.2为步长，从 0.4 到 2（虚线）绘制。所有面板采用相同的颜色编码。

，这一图景让人联想到玻色-爱因斯坦凝聚体。这一图景与具有软核势的两个无自旋费米子的对关联函数 [41] 以及参考文献中观察到的 P波费米子的关联结构相一致 [42]。

与这个相对平淡无奇的峰相比，$g(x_1, x_2)$ 在排斥侧具有丰富得多的结构，包含多个清晰可辨的峰值。这些峰值清楚地表明系统正在形成局域结构，并且在类似的一维少体维格纳晶体模拟中也观察到了这种现象 [85]。对于奇数粒子数，我们发现沿着 $x_1 = 0$方向，泡利脊的右侧（或左侧）总会有一个峰值。对于这些系统，密度在$x = 0$ (见图11)处有一个峰值，因此位于 $x_1 = 0$的一个粒子会有很大概率在固定间隔处找到 $A - 1$ 个粒子，间隔沿着 x_2方向。对于 $A = 3$，这发生在 $x_2 = \pm 2$处；而对于 $A = 5$ ，则大约发生在 $x_2 = \pm 1.5$ 和± 3处。

偶数个粒子的结构则相对不同。对于偶数个粒子，陷阱中心的密度分布会出现耗尽，见图 7 和11。因此，沿 $x_i = 0$ 的对关联函数不会显示任何峰值。相反，水平或垂直的 $A - 1$ 峰值出现在附近 $x_i \approx \pm 1$。而对于 $A = 4$ 粒子，在 $x_i = \pm 2.5$处发现了第二组三个峰值；对于 $A = 6$ 粒子，该组峰值局域于 $x_i = \pm 2$处。因此，定义明确的单粒子峰值会随着 A 的增加而更加压缩，这与图11的结果一致。

总而言之，所有系统的对关联函数都表明，在排斥区域存在一个定义良好的局域相，而在吸引区间则存在一个类玻色相。这种复杂行为是由一个包含回流关联的单一行列式所构成的神经量子态得到的。换言之，对于固定的粒子数，同一个神经量子态能够描述具有截然不同的空间范围、关联结构和底层物理图像的系统。我们将此视为神经量子态方法在描述凝聚态物质系统方面性能良好的一个积极信号。

V. 结论与展望

总而言之，我们应用 NQS 方法求解了无自旋自由度的一维系统中谐振子陷阱内相互作用费米子的薛定谔方程。我们采用一种受核

物理应用启发的玩具模型高斯形状因子来描述相互作用，并使其在零程极限下恢复非相互作用情形。我们特意选择了一组微扰理论可能难以轻易收敛的相互作用强度和范围。特别是，我们发现，将相互作用范围设为自然振子长度的一半是一个有趣的选择。

我们采用的神经量子态提供了一个完全反对称的试探函数，用于表示一维“无自旋”费米子。该神经量子态包含两个等变层，每层有 $H = 64$ 个隐藏节点。这些层将费米子的一维位置嵌入 H 维空间。随后，这些层被投影到一个单一的 GSD 中。这种方法从一开始就具备回流关联的优势，而且不需要显式使用 Jastrow 因子。从 $A = 2$ 个粒子增加到 $A = 6$ 个粒子时，神经量子态的参数数量几乎保持不变。通过将神经量子态与变分蒙特卡罗技术相结合，我们得到的波函数能够表示多体关联，并且无需对网络架构进行任何特定修改，就能准确描述截然不同的物理区间。

本文着重将这一新型神经量子态方法与其他知名的多体方法进行基准比较。对于 A 的所有取值，我们都将其与平均场 Hartree-Fock 方法和直接对角化方法进行比较。此外，在 $A = 2$ 情形下，我们可以获得实空间中的精确解，也可以对高斯相互作用进行微扰处理。对于 A 的所有取值，结果均表明，神经量子态能够高效地准确表示多体波函数。实空间连续体中的神经量子态方法具有变分性能这一优势，优于平均场方法，且不会像直接对角化方法那样受到基组截断效应的影响。我们通过若干性质来表征基准测试过程，包括基态能量、密度分布、单粒子占据数、单体密度矩阵以及对关联函数（PCF）。

NQS 能够准确描述不同物理区间中的基态波函数。特别是，当相互作用强度从强吸引变为强排斥时，我们发现了两种截然不同的图景。对于一维受限无自旋费米子系统的这两个相，我们给出了一种不依赖具体方法的统一解释。在吸引侧，系统的密度分布呈单峰，并以原点为中心，让人联想到玻色-爱因斯坦凝聚体。随着强度变得更加负，峰值会变得更窄、更高。这意味着系统发生了显著的密度重排，而基于谐振子基组的截断基组直接对角化方法难以捕捉这种变化。不过，当 A 足够大时，实空间实现的 Hartree-Fock 方法能够很好地描述这一相。在此区间，单体密度矩阵在非对角线上强烈振荡，对分布函数则呈现狭窄的方形轮廓。我们发现，一组简并的 2 个单粒子轨道被大幅耗尽，而其余 $A-2$ 个态则保持完全占据。

相比之下，当相互作用强度较大且为排斥时，我们发现了完全不同的图景。在此极限下，系统进入局域化晶体区间，费米子自发地呈周期性排列。这反映在密度分布中：其具有明确的周期结构，并呈现 A 个峰值；对关联函数（PCF）也显示出相应的晶体结构。从占据数来看，排斥区域的成对轨道呈现准简并。我们注意到，直接对角化方法能够很好地描述这一区间，但 HF 近似预测的密度分布具有夸大的局域化特征。

我们预见，这项工作在不久的将来会有多种应用，并将这些发展区分为应用层面和理论层面。在应用方面，我们在相对简单的费米子系统中发现了两种物理相的明确特征。对于无自旋系统，有限程相互作用效应不可或缺，并会导致吸引区域与排斥区域中截然不同的行为。在此，可以利用神经量子态研究尚未探索的物理特征，例如高阶密度矩阵。此外，还可以采用神经量子态波函数，运用标准蒙特卡罗技术计算其他相关算符的矩阵元 [61]。有了系统的完整波函数，我们还应能够刻画纠缠度量随相互作用参数变化的演化。采用奇宇称势进行神经量子态模拟也很有意思，因为这类势与以往无自旋费米子研究中采用的更现实的相互作用相似 [35, 37, 39]。我们强调，对于导数型奇宇称势，神经量子态方法的应用并不存在任何根本性限制。此外，还可以将这一方法扩展到具有硬边界条件的其他受限系统 [95], ，或扩展到采用周期性边界条件的无限系统 [20, 96, 97]。

在形式方面，许多方向都还有值得推进的工作。首先，可以尝试以更正式的方式量化波函数试探函数的质量。换言之，可以探究引入这一 NQS试探函数时丢失了哪些信息。一个可能的研究方向是量化 NQS 上的纠缠度量。在一维超冷原子的应用中，自由度与自旋有关，并且常在实验中受到操控。因此，需要一种明确且通用的自旋依赖型 NQS 表示法，可沿着近期在晶格系统 [98], 核子 [33], 电子气 [20, 96]或幺正费米子 [97], 中探索的方向发展，但也有望适用于 SU(N) 相互作用。将其与此处非微扰区间中丰富多样的实验结果进行比较，或可为检验 NQS试探函数 ¨试探函数的适用性提供肥沃的研究土壤。最后，我们的发现表明，最初在非相互作用情形下预训练的 NQS，最终至少能够学习量子系统的两种不同相。值得进一步分析这一结果是否具有普遍性，以及 NQS 是否确实能够跨越复杂的量子相变边界。

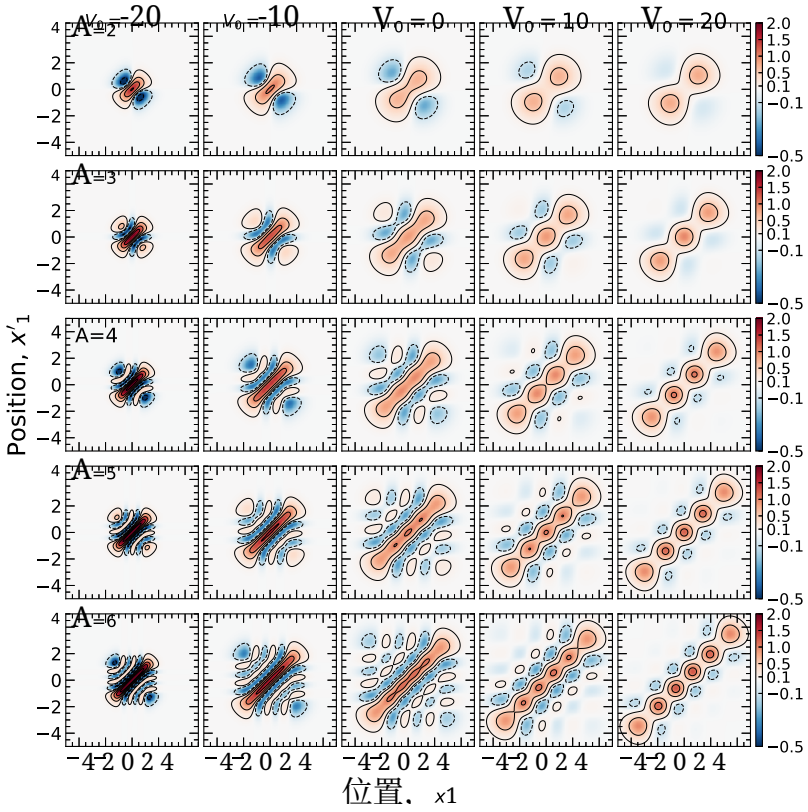

图 16。与图 14 中的 Hartree-Fock近似相同。

我们的最终目标是研究采用神经量子态的核物理系统。为此，我们需要扩展当前的技术，以处理三维系统。这还要求纳入自旋和同位旋自由度。迄今为止的所有研究都表明，通过采用神经量子态，少体扇区中的原子核可以在理论上得到研究 [29, 31–33]。我们在这项工作中发现的、随粒子数变化的良好标度关系，能否高效地迁移到核物理环境中，并对从头算核理论带来实质性改变，仍有待观察。

附录 A：非局域性质

在本附录中，我们考察了使用基准方法模型获得的非局域性质。具体而言，我们分析了在 HF 近似和直接对角化框架下获得的单体密度矩阵和对关联函数。一方面，这为正文中的 NQS结果提供了有用的交叉检验。另一方面，这使我们能够确定关联对系统非局域性质的影响。

1. Hartree-Fock

我们首先讨论图 16所示的 HF 近似下的单体密度矩阵。其格式和面板结构与图 14中的神经量子态数据相同：不同的行对应粒子数从 $A = 2$到 $A = 6$，不同的列则表示 $\sigma_0 = 0.5$的不同相互作用强度。定性而言，图 14和 16 非常相似。在吸引区间（中心左侧的面板）中，我们发现随着 V_0 变得更加负，系统会变得更加紧凑。与神经量子态的情况一样，HF 解在非对角线区域呈现出强烈的振荡行为。然而，与神经量子态解不同，HF 结果中的中心脊线更高，且密度矩阵的整体支撑区域呈现出更加清晰的方形。

中心列展示了神经量子态和 Hartree-Fock 解的相同结果，这与预期一致，因为这是非相互作用情形。不过，这对神经量子态解的数值精度而言是一项非平凡的检验；因此，它不仅重现了能量，也重现了非相互作用波函数的非局域性质。在排斥侧（中心右侧的面板所示），Hartree-Fock 结果也与图中展示的神经量子态数据非常相似。14。在 Hartree-Fock 解中，我们发现沿对角线集中的结构更多，解也更加局域，并且峰值清晰可辨

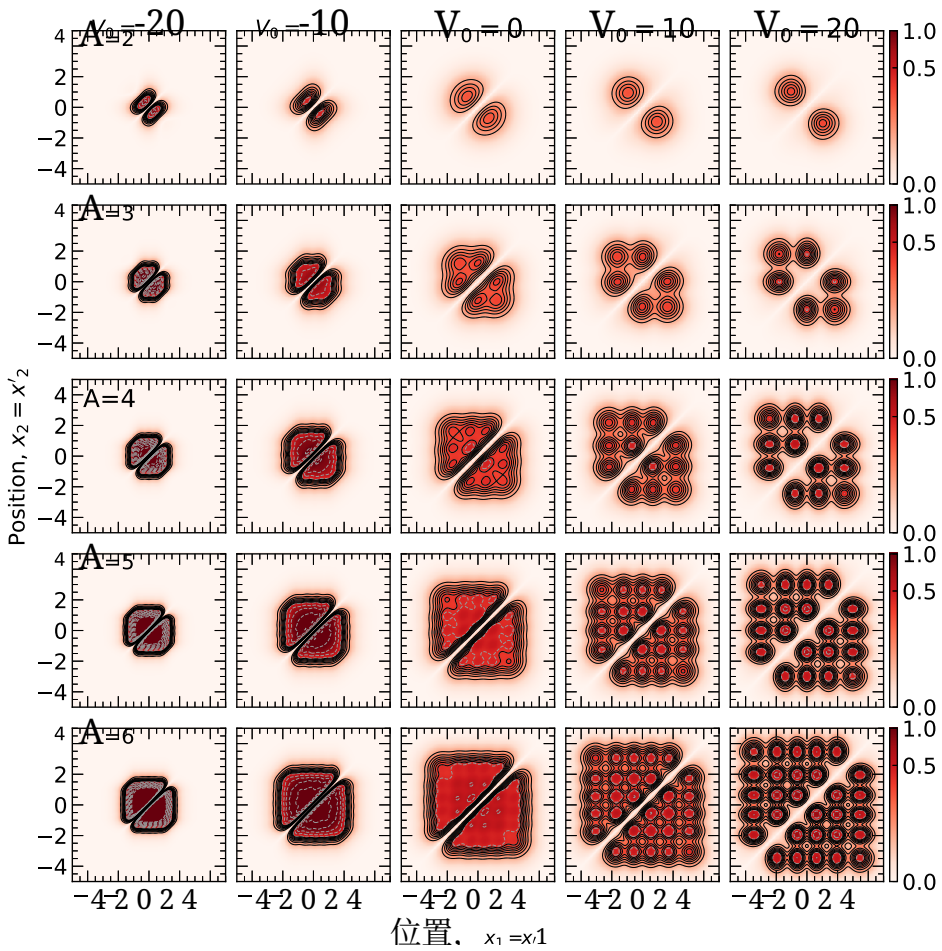

图 17。与图 15 中 Hartree-Fock近似的情况相同。

与神经量子态模拟相比。这与图中所示密度分布的发现相符。 11。在对角线之外，神经量子态结果显示出比 Hartree-Fock 预测略强的振荡行为。

现在讨论 HF 近似下的 PCF，如图 17所示。这里的格式与正文中的图 15相同。与 OBDM 的情况一样，两种情形的中间列相同。这再次表明，借助 NQS 试探函数，非相互作用情形处于可控范围内。这一基准测试表明，NQS试探函数 在二体层次不会丢失任何大量信息。

在吸引和排斥情形下，都可以观察到 HF 与神经量子态对分布函数之间的差异。在吸引情形下，HF 关联函数的支撑范围小于神经量子态的情形。这与相应的密度分布一致：HF 情形下的密度分布更窄。此外，在吸引情形下，HF 呈现出更尖锐、近似方形的行为，且泡利脊线也略窄。

在排斥情形下，Hartree-Fock 密度呈现出更加局域化的态。在对分布函数中，这体现为轮廓非常清晰的峰值。这些峰值 Hartree-Fock 分布中的峰值接近神经量子态（NQS）预测的峰值，但通常更圆润、更高。

2. 直接对角化

我们还计算了直接对角化模拟的单体密度矩阵和对关联函数。我们已经从能量学、局域密度和占据数的分析中得知，精确对角化解在吸引区间存在收敛问题。因此，我们预计会在这一区域看到差异；这些差异应被视为与有限模型空间截断相关的数值收敛性不足的迹象，而不是任何实际的物理局限性。

我们在图 18 中以与图 14和 16相同的格式展示了通过精确对角化获得的单体密度矩阵。正如预期，图中间和右侧面板几乎没有可观察到的差异。在此参数范围内，精确对角化结果已充分收敛，与神经量子态的吻合可作为可靠的定性基准。换言之，神经量子态波函数的一体非局域性质

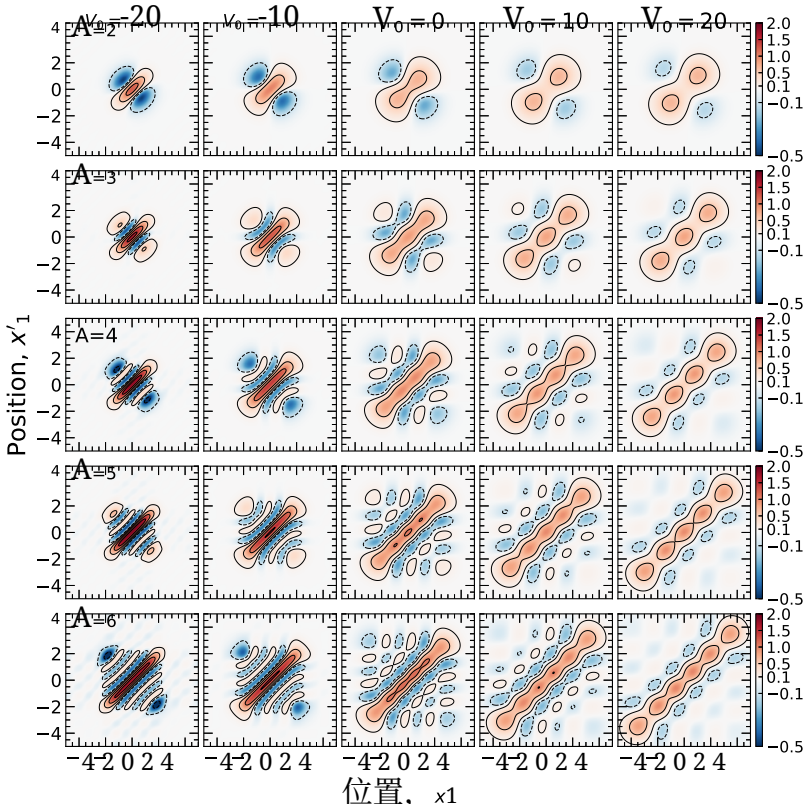

图 18. 与图 14相同，适用于直接对角化方法。

得到了很好的控制。

在吸引区间中，在中心左侧的图中，随着相互作用强度 V_0 和粒子数 A 增加，我们观察到差异也在增大。左下角的图中，$V_0 = -20$ 和 $A = 6$ 对这一情况的影响最大。我们发现密度矩阵在 $x_1 - x'_1$ 平面上具有扩展支撑。在对角线之外，神经量子态和精确对角化情况下的振荡次数相同，但后者呈现出更大、较不紧凑的振荡结构。我们推测，这是由于活性空间中缺少高能模，而这些模可能有助于描述更快的非对角振荡。

最后，图 19 展示了从精确对角化方法中提取的PCF。在非相互作用列和排斥列中，图15 中的 NQS 结果与直接对角化结果之间没有显著差异。这再次可以作为NQS波函数具有高质量的一个 संकेत。与 OBDM 的情况一样，NQS 与精确对角化方法之间的差异会随着A 增加，并且在 V_0取值更具吸引性时更加明显。与图 11 中的局域密度分布以及图 14中的OBDM的观察结果一致，直接对角化结果往往具有更紧凑的PCF，在对角线之外还呈现出一些非常微弱的振荡。我们再次将其归因于截断精确对角化方法中的多体基。

致谢

本工作得到英国科学技术设施委员会的资助，资助编号为 ST/L005743/1 和 ST/P005314/1；得到西班牙科学与创新部/西班牙国家研究署/10.13039/501100011033 资助的项目 PID2020-118758GB-I00 和 PID2020-114626GB-I00 的资助；得到西班牙科学与创新部/西班牙国家研究署/10.13039/501100011033 及欧洲社会基金“欧洲社会基金投资你的未来”资助的“Ram´on y Cajal”项目（编号 RYC2018-026072）的资助；得到西班牙科学与创新部/西班牙国家研究署/10.13039/501100011033 资助的“宇宙科学研究所 2020—2023 年玛丽亚·德·梅斯特雷卓越单位”项目（编号 CEX2019-000918-M）的资助，以及加泰罗尼亚自治区项目（编号 2021SGR01095）的资助。TRIUMF 通过与加拿大国家研究委员会签订的拨款协议获得联邦资金。本工作还获得西班牙政府经济事务与数字化转型部通过 QUANTUM ENIA 项目征集——Quantum Spain 项目提供的资金支持，并获得欧盟通过复苏、转型与韧性计划——下一代欧盟在数字西班牙2026议程框架下提供的资金支持。作者感谢 Artemisa 计算资源的支持，该资源由欧盟欧洲区域发展基金和瓦伦西亚自治区资助；同时感谢粒子物理研究所（IFIC，西班牙国家研究委员会—瓦伦西亚大学）提供的技术支持。

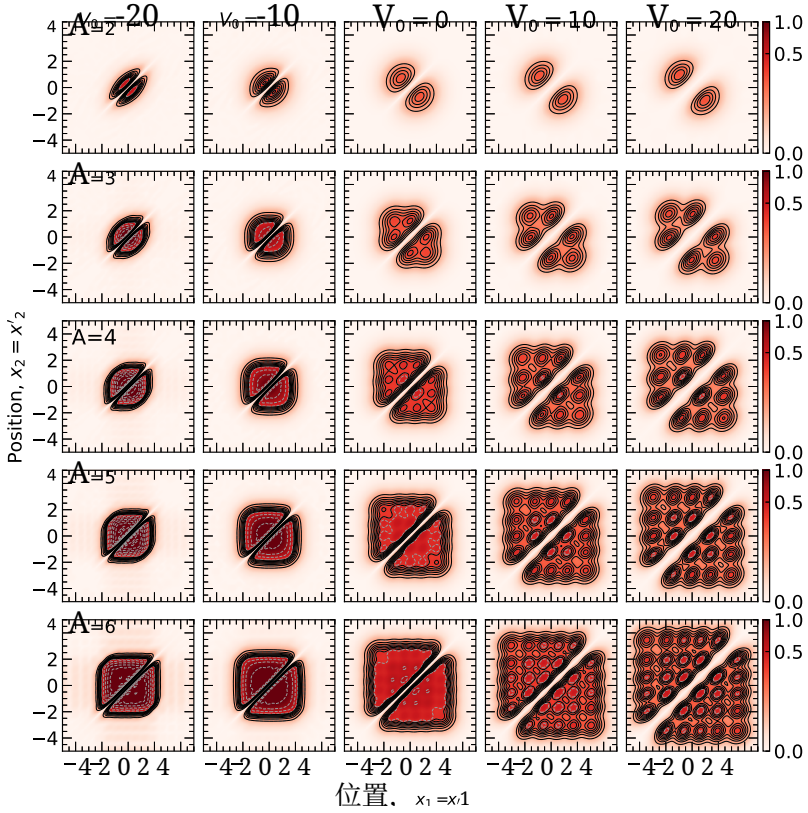

图 19. 与图 15相同，采用直接对角化方法。

[1] G. Carleo、I. Cirac、K. Cranmer、L. Daudet、M. Schuld、N. Tishby、L. Vogt-Maranto 和 L. Zdeborov´a，《现代物理学评论》91，045002（2019），arXiv:1903.10563。[2] G. Cybenko，《控制、信号与系统数学》2，303（1989）。[3] K. Hornik，《神经网络》4，251（1991）。[4] B. Irie 和 S. Miyake，载于 *ICNN*（1988），第 641–648 页。[5] A. N. Kolmogorov，载于 *Doklady Akademii Nauk*，第 114 卷（俄罗斯科学院，1957），第 953–956 页。[6] G. Carleo 和 M. Troyer，Science 355（2017），10.1126/sci- ence.aag2302，arXiv:1606.02318。[7] H. Saito，《日本物理学会志》86，093001（2017），arXiv:1707.09723。[8]H. Saito，《日本物理学会志》87，074002（2018），arxiv:1804.06521。[9] K. Choo、G. Carleo、N. Regnault 和 T. Neupert，Phys. Rev. Lett. 121，167204（2018），arXiv:1807.03325。[10] D. Luo 和 B. K. Clark，Phys. Rev. Lett. 122，226401（2019），arXiv:1807.10770。[11] M. Rigo、B. Hall、M. Hjorth-Jensen、A. Lovato 和 F. Pederiva，Phys. Rev. E 107，025310（2023），arxiv:2211.04614。[12] B. Fore、J. M. Kim、G. Carleo、M. Hjorth-Jensen、A. Lo- vato 和 M. Piarulli，Phys. Rev. Res. 5，033062（2023），arXiv:2212.04436。[13]P. Mujal，`Alex Mart´ınez Miguel、A. Polls、B. Juli´a- D´ıaz 和 S. Pilati，SciPost Phys. 10，073（2021），arXiv:2010年.03875.

[14] N. Cantero、和 S.Pilati，《物理评论》 E 102，0333012020年 2005 14290（），arXiv：..[15] D. Pfau、J. S. Spencer、A. G. D. G. Matthews 和 W. M.C. Foulkes，《物理评论研究》2，033429（2020年），arXiv:1909.02487。[16] J.S.Spencer、D. Pfau、A. Botev 和 W.M. C. Foulkes，（2020年），arXiv:2011.07125。[17] J.Hermann、Z. Sch

12, atzle和 F. 891 2020 1909 08423 Noe，《自然化学》（），arXiv：..[18] Z.Sch¨atzle、J. Hermann 和 F.No´e，《化学物理学杂志》154，124108（2021 年），arXiv:2010.05316。[19] M. Wilson、N. Gao、F. Wudarski、E. Rieffel 和N. M.Tubman，（2021年），arXiv:2103.12570。[20] M. Wilson、S. Moroni、M. Holzmann、N. Gao、F. Wudarski、T.Vegge 和 A.Bhowmik，Phys. Rev. B 107，235139（2023年），arXiv:2202.04622。[21] I. von Glehn、J.S.Spencer 和 D.Pfau，（2022年），arXiv:2211.13672。[22] M. Scherbela、R.Reisenhofer、L. Gerard、P.Marque- tand 和 P. Grohs，《自然计算科学》2，331（2022年），arXiv:2105.08351。[23] D. Luo、A. P. Reddy、T. Devakul 和L. Fu，（2023年），arXiv:2303.08162。[24] J. Hermann、J. Spencer、K. Choo、A. Mezzacapo、W. M. C. Foulkes、D. Pfau、G. Carleo 和F.No´e，《自然评论·化学》（2023年），arXiv:2208.12590。

[25] K. Choo、A.Mezzacapo和 G.Carleo，《自然·通讯》11，2368（2020年），arXiv:1909.12852。[26] F.Vicentini、D. Hofmann、A. Szab´o、D. Wu、C. Roth、C. Giuliani、G.Pescia、J. Nys、V. Vargas-Calder´on、N. As- trakhantsev 和G. Carleo，《SciPost物理代码库》，7（2022年）。[27] D. Ceperley、G. V.Chester 和M. H.Kalos，Phys. Rev.B16，3081（1977年）。[28] A. Paszke、S. Gross、S. Chintala、G.Chanan、E. Yang、Z. DeVito、Z. Lin、A.Desmaison、L. Antiga和 A. Lerer，载于 *NIPS Autodiff Workshop*（2017年）。[29] J.Keeble 和A. Rios，《物理快报B》809（2020年）arXiv:1911.13092。[30] J. W. T. Keeble，*Neural Network Solutions to the Fermionic Schrödinger Equation*，博士论文，萨里大学（2022年）。[31] C. Adams、G. Carleo、A. Lovato和 N. Rocco，Phys. Rev. Lett. 127，022502（2021年），arxiv:2007.14282。[32] A. Gnech、C. Adams、N. Brawand、G.Carleo、A. Lo-vato和 N. Rocco，《少体系统》63，7（2022年），arXiv:2108.06836。[33] A. Lovato、C. Adams、G. Carleo和 N. Rocco，《物理评论研究》4，043178（2022年），arXiv:2206.10021。[34] M. Girardeau，《数学物理杂志》1，516（1960年）。[35] M. Valiente，Phys. Rev. A 102，053304（2020年），arXiv:2009.00624。[36] M. D. Girardeau、H.Nguyen 和 M. Olshanii，《光学通讯》243，3（2004年），cond-mat/0403721。[37] T. Sowi´nski 和 M. ´Angel Garc´ıa-March，《物理学进展报告》82，104401（2019年），arxiv:1903.12189。[38] A.Minguzzi 和 P.Vignolo，《AVS量子科学》4，0271022022年 2201 02362（），arxiv:。。[39] M. Girardeau和 M. Olshanii，（2003年），cond- mat/0309396。[40] M.D. Girardeau 和 M.Olshanii，Phys.Rev.A 70，023608 2004年（）。[41] P. Ko´scik 和T. Sowi´nski，Sci. Rep. 8，48（2018年），arXiv:1707.04240。[42] P. Ko´scik 和 T.Sowi´nski，《新物理学杂志》22，093053（2020年），arXiv:2004.06610。[43] I. Bloch、J. Dalibard 和W. Zwerger，《现代物理学评论》80，885（2008年），arxiv:0704.3011。[44] S. Giorgini、L. P.Pitaevskii和 S.Stringari，《现代物理学评论》80，1215（2008年），arxiv:0706.3360。[45] B.DeMarco 和D. S. Jin，Science285，1703（1999年）。[46] N. Navon、S. Nascimb`ene、F. Chevy 和C. Salomon，Science 328，729（2010年），arxiv:1004.1465。[47] A.N. Wenz、G.Z¨urn、S.Murmann、I. Brouzos、T. Lompe和 S. Jochim，Science 342，457（2013年），arxiv:1307.3443。[48] A. Rojo-Franc`as、A.Polls 和 B. Juli´a-D´ıaz，数学 8，1196（2020年），arXiv:2005.10886。[49] Y.Suzuki 和 K. Varga，*Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems*（施普林格-Verlag柏林 海德堡，1998年）。[50] P.Ring 和P. Schuck，*The Nuclear Many-Body Problem*（施普林格，柏林，海德堡，1980年）。[51] A. Kievsky、M. Gattobigio、L. Girlanda 和M. Viviani，《核与粒子科学年度评论》71，465（2021年），arXiv:2102.13504。[52] L. Mitas，Phys. Rev. Lett. 96，240402（2006年），arXiv:cond- mat/0601485。[53] P.-O. L¨owdin，《物理评论》97，1474（1955年）。[54] M. J. Knight、H.M. Quiney 和 A.M. Martin，《新物理学杂志》24，053004（2022年），arXiv:2106.09187。[55] C.Schilling，《物理评论 A》88，042105（2013年），arXiv:1307.6858。[56] C. Schilling 和 R. Schilling，《物理评论A》93，021601(R)（2016年），arXiv:1508.04452。[57] J.Rius Casado，*Analytical and machine learning study of non-interacting one-dimensional spinless trapped fermionic systems*，巴塞罗那大学硕士论文（2023年）。[58] E. K.U.Gross、E.Runge和 O.Heinonen，*Many-Particle Theory* 1991（IOP 出版社，布里斯托尔，）。[59] A. Sannai、Y. Takai 和M. Cordonnier，（2019年），arXiv:1903.01939。[60] R. M. Martin、L.Reining 和 D. M.Ceperley，*Interacting Electrons: Theory and Computational Approaches*（剑桥大学出版社，2016年）。[61] F.Becca和 S.Sorella，*Quantum Monte Carlo Ap-proaches for Correlated Systems*（剑桥大学出版社，2017年）。{（剑桥大学出版社，2017年）。[62] M. Hutter，（2020年），arXiv:2007.15298。[63] N. Metropolis 和 S. Ulam，《美国统计协会杂志》44，335（1949年）。[64] W. K. Hastings，《Biometrika》57，97（1970年）。[65] D. P. Kingma 和 J. Ba，（2014年），arXiv:1412.6980。[66] S.-X. Zhang、Z.-Q. Wan 和 H. Yao，《物理评论研究》5，033041 2023年 1911 09117（），arXiv:。。。[67] A. Scemama、T. Leli`evre、G. Stoltz、E. Canc`es 和 M. Caffarel，《化学物理学杂志》125，114105（2006年），arXiv:cond-mat/0606439。[68] J. Keeble，“无自旋费米子”，https://github.com/ jwtkeeble/SpinlessFermions（2023年），GitHub 仓库。[69] A. Azzam，*Transfer Learning For Many-Body Systems*，巴塞罗那大学硕士论文（2023年），可访问 http://hdl.handle.net/2445/202717。[70] J.-P. Blaizot 和 G. Ripka，*Quantum theory of finite systems* 1986（麻省理工学院出版社，马萨诸塞州剑桥，）。[71] R. P. Feynman 和 M. Cohen，《物理评论》102，1189（1956年）。[72] M. Holzmann、B. Bernu 和 D. M. Ceperley，《物理评论 B》74，104510（2006年），arXiv:cond-mat/0605513。[73] M. Holzmann 和 S. Moroni，《物理评论 B》99，085121（2019年），arXiv:1910.07167。[74] D. M. Ceperley，《统计物理学杂志》63，1237（1991年）。[75] P. Mujal、E. Sarl´e、A. Polls 和 B. Juli´a-D´ıaz，《物理评论 A》96，043614（2017年），arXiv:1707.04166。[76] A. Rojo-Franc`as、F. Isaule 和 B. Juli´a-D´ıaz，《物理评论 A》105，063326（2022年）。[77] M. P-lodzie´n、D. Wiater、A. Chrostowski 和 T. Sowi´nski，（2018年），arXiv:1803.08387。[78] J. W. Negele 和 H. Orland，*Quantum Many-Particle Systems* 1987（Addison-Wesley，）。[79] C. C. Marston 和 G. G. Balint-Kurti，《化学物理学杂志》91，3571（1989年）。[80] A. Rios，“一维无自旋费米子”，https: //github.com/arnaurios/1D_fermions（2023年），GitHub 仓库。[81] T. Busch、B. G. Englert、K. Rzazewski 和 M. Wilkens，《基础物理学》28，549（1998年）。[82] B. L. Earl，《化学教育杂志》85，453（2008年）。[83] N. D. Drummond、Z. Radnai、J. R. Trail、M. D. Towler 和 R. J. Needs，《物理评论 B》69，085116（2004年），arXiv:0801.0377。[84] E. Wigner，《物理评论》**46**，1002 (1934)。

[85] K. Jauregui、W. H¨ausler和 B. Kramer，EPL 24，581（1993），arXiv:cond-mat/9305034。[86] B. Szafran、F. M. Peeters、S. Bednarek、T. Chwiej 和 J.Adamowski，《物理评论》B 70，035401（2004年），cond- mat/0405350。[87] I. Shapir、A.Hamo、S. Pecker、C. P. Moca、 ¨O. Leg- eza、G. Zarand和 S.Ilani，Science364，870（2019年），arXiv:1803.08523。[88] D. D. Vu 和S. Das Sarma，《物理评论》B 101，125113（2020年），arXiv:1901.00019。[89] N. Ziani、F. Cavaliere、K. Becerra 和 M. Sassetti，Crystals11（2020年），10.3390/cryst11010020。[90] B. E. Granger 和 D.Blume，《物理评论》Lett. 92，133202（2004年），cond-mat/0307358。[91]M. Valiente，《物理评论》A 103, L021302（2021年），arXiv:2009.00614.[92]I. Morera、B. Juli´a-D´ıaz 和 M. Valiente（2021年），arXiv:2103.16499.[93] W. L. McMillan，《物理评论》138, A442（1965年）。[94] A. Rios、B. Barker、M. Buchler 和 P. Danielewicz，Ann. Phys. 326, 1274（2011年），arXiv:1009.0215。[95] A. Radu 和 C. A. Duque，Sci. Rep. 12, 2535（2022年），arxiv:2109.03311.[96] G. Pescia、J. Nys、J. Kim、A. Lovato 和 G. Carleo（2023年），arXiv:2305.07240。[97] J. Kim、G. Pescia、B. Fore、J. Nys、G. Carleo、S. Gandolfi、M. Hjorth-Jensen 和 A. Lovato（2023年），arXiv:2305.07240.[98] J. Robledo Moreno、G. Carleo、A. Georges 和 J. Stokes，Proc. Natl. Acad. Sci. 119（2022年），arxiv:2111.10420。