

QiankunNet-Solid/DMET: 面向固体材料模拟的生成式神经网络量子态方法

马欢^{1†}, 付礼中^{1†}, 商红慧^{3*}, 杨金龙^{2,3*}

1. 中国科学技术大学未来技术学院, 合肥 230026

2. 中国科学技术大学, 合肥国家实验室, 合肥 230088

3. 中国科学技术大学, 精准智能化学全国重点实验室, 合肥 230026

† 同等贡献

* 联系人, E-mail: shh@ustc.edu.cn; jlyang@ustc.edu.cn

2025-03-22 收稿, 2025-05-30 修回, 2025-06-06 接受, 2025-06-09 网络版发表

国家自然科学基金(T2222026)、中国科学院战略性先导科技专项(XDB0450101)、量子科技创新计划(2021ZD0303306)和中国科学技术大学超级计算中心资助

摘要 神经网络量子态(neural network quantum states, NNQS)方法在量子多体系统模拟中展现出巨大潜力, 但其在固体材料中的应用仍面临周期性边界条件、长程相互作用与计算复杂度等关键挑战. 本文系统综述了基于Transformer架构的神经网络量子态方法QiankunNet的最新进展, 重点介绍了针对固体材料设计的QiankunNet-Solid和QiankunNet-DMET方法. QiankunNet-Solid通过二次量子化形式与自回归采样技术, 高效处理周期性体系的波函数表达与计算, 在一维氢链、石墨烯、硅及金刚石等体系中实现了与全组态相互作用(full configuration interaction, FCI)精度相当的计算结果. QiankunNet-DMET结合密度矩阵嵌入理论(density matrix embedding theory, DMET)与迁移学习策略, 通过局域化嵌入哈密顿量求解显著降低计算成本, 在强关联材料中预测了磁性序与电子密度分布, 与传统方法和实验数据符合. 本文进一步探讨了这些方法在高温超导、拓扑材料设计等领域的应用潜力, 并指出未来在提升系统规模、激发态计算与算法优化等方面的研究方向.

关键词 神经网络量子态, 量子蒙特卡罗, 密度矩阵嵌入理论, 周期性边界条件

1 神经网络量子态方法应用于固体材料的挑战

目前, 大多数神经网络量子态(neural network quantum states, NNQS)^[1~7]方法主要针对Heisenberg模型、Hubbard模型等物理模型或量子化学中的小分子体系, 用于计算能量的关联效应修正. 然而, 在材料领域, 同样有许多受到广泛关注的强关联系统, 例如高温超导材料、多铁性材料, 以及磁性材料、拓扑绝缘体等自

旋电子学材料. 关联材料体系的计算方法根据体系电子关联强度不同而有所区分. 对于弱至中等关联体系, GW方法^[8](G代表格林函数, W代表动态屏蔽库仑相互作用)是主流的准粒子能带计算方法. 其中, 非自洽G0W0^[9]计算因计算效率较高而被广泛采用, 但其精度较为依赖于作为起点的密度泛函理论(density functional theory, DFT)计算结果. 对于强关联体系(如强关联金属、具有金属-绝缘体转变特征的关联体系等), 动态平均场理论(dynamical mean-field theory, DMFT)^[10]

引用格式: 马欢, 付礼中, 商红慧, 等. QiankunNet-Solid/DMET: 面向固体材料模拟的生成式神经网络量子态方法. 科学通报, 2025, 70: 4015–4026

Ma H, Fu L, Shang H, et al. QiankunNet-Solid/DMET: a generative neural network quantum state method for solid material simulations (in Chinese). Chin Sci Bull, 2025, 70: 4015–4026, doi: [10.1360/CSB-2025-0315](https://doi.org/10.1360/CSB-2025-0315)

及其与DFT的结合方法(DFT+DMFT)^[11]已成为主流研究范式,能够较好地处理局域电子关联效应.量子蒙特卡罗(quantum Monte Carlo, QMC)^[12]方法虽然常被作为材料体系基态能量的参考基准,但其在能带结构计算中的应用仍面临较大技术挑战.值得注意的是,近期基于一阶与二阶耦合簇理论的运动方程方法(equation of motion coupled cluster with singles and doubles, EoM-CCSD)^[13]被证明在中等关联程度体系中可以给出与实验值或自洽的GW^{TC-TC}方法接近的带隙,为中等关联体系提供了新的研究思路^[14].而NNQS方法作为一种新兴的强关联体系计算方法,在其中实现周期性边界条件计算,从而将其应用于材料体系中,是解决耦合簇方法等微扰修正方法无法处理的强关联问题的新思路.固体体系中,周期性要求波函数在平移操作下保持协变性^[15],这对神经网络结构设计提出了额外约束;同时,固体体系中的长程相互作用和多k点计算需求进一步提升了对网络表达能力的要求.此外,在实际计算中为了消除有限体积效应带来的误差,常常需要通过不断增加k点采样密度来将计算结果外推至热力学极限,这一操作显著增加了对计算资源的需求.因此,神经网络量子态方法在模拟固体材料时,始终面临着在确保计算可行性的前提下,如何有效提升网络对复杂电子关联效应和周期性约束的适应性这一挑战.

目前,有数种NNQS方法已被报道可以适用于周期性体系.其中,一些方法采用了波函数的一次量子化表示,即将波函数直接表示为各个电子在实空间的位置坐标的函数^[16~19].这类方法可以通过将原子坐标嵌入到一系列周期性函数来实现周期性体系的计算.相关文献表明,此类方法已经可以得到与扩散蒙特卡罗(diffusion Monte Carlo, DMC)方法相当的基态能量.这类方法的复杂性在于,实空间中的费米子波函数具有交换反对称性,因此研究者必须精心设计网络结构,使得输出的波函数满足这一约束.常见的做法是将波函数取作多个含有可学习参数的行列式的和,一方面,这样的假设从物理出发,可能有助于使用尽量少的参数来描述复杂的多体波函数;另一方面,这种结构限制了网络的选择,给网络的设计带来了困难.针对周期性系统的二次量子化NNQS方法也已在文献^[20]中被报道,这一工作采用费米子到自旋的映射,将电子占据状态变换为自旋系统的组态,并使用受限玻尔兹曼机(restricted Boltzmann machine, RBM)作为波函数拟设.这一方法重新引入了量子化学中常用的基组,使得其可

以直接与传统量子化学方法进行比较.通过逐步选择一系列更大的基组,这一方法的计算精度可以被系统性地提高.此外,文献^[20]还提到,对于传统的网络结构,马尔可夫链蒙特卡罗(Markov chain Monte Carlo, MCMC)采样过程是NNQS方法的计算热点.在量子化学系统中,MCMC方法较低的采样率成为限制NNQS方法推广至大系统的主要瓶颈.由于二次量子化NNQS方法不需要显式考虑费米子交换反对称性,其网络结构设计的自由度更高;通过引入自回归网络、采用批量自回归采样(batched autoregressive sampling, BAS)的算法,可以大大减少采样过程的计算时间,从而将NNQS方法用于更大的体系.

QiankunNet^[6,7]是我们近期开发的基于Transformer的神经网络量子态方法.QiankunNet在分子体系中实现了较大规模系统的模拟,并且能够获得满足化学精度的基态能量等性质.为了应对神经网络量子态方法在固体体系模拟中的挑战,探索其在固体材料中的新应用,我们基于QiankunNet提出了两种针对周期性体系的扩展:QiankunNet-Solid^[21]和QiankunNet-DMET^[22].这两种方法在固体材料模拟中展现出良好效果,为神经网络量子态方法在材料研究领域的应用提供了新的技术路径.

2 QiankunNet-Solid方法:适用于周期性体系的神经网络量子态

在近期的工作中,我们将基于Transformer的神经网络量子态方法QiankunNet^[6,7]推广到了周期性材料体系,称作QiankunNet-Solid^[21].该方法引入了波函数的周期性并成功在神经网络量子态中考虑了长程关联.在周期性材料的基态能量、能带、态密度等性质的计算上取得了非常高精度的结果.

2.1 QiankunNet-Solid方法简介

QiankunNet-Solid方法的计算流程如图1所示.首先,利用原子基组合成的布洛赫原子轨道进行周期性Hartree-Fock (HF)计算,并以得到的单电子轨道为基组,使用二次量子化形式表达材料体系的哈密顿量.采用这种方式,体系的周期性特征就由这些周期性HF轨道表达出来,无须再单独考虑波函数的平移协变性.哈密顿量的信息以单双电子积分的形式存储.与分子体系相比,周期性体系的单双电子积分通常为复数而非实数,因此我们更改了哈密顿量数据结构以支持复数

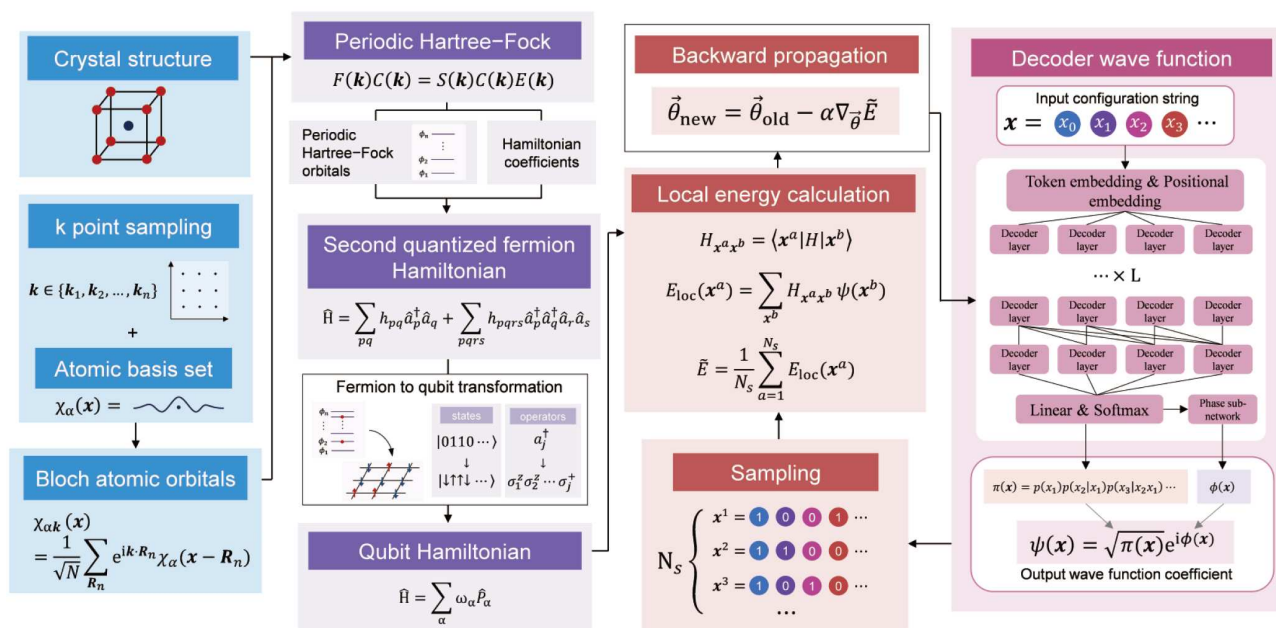


图1 QiankunNet-Solid的计算流程^[21]。QiankunNet波函数拟设以量子比特的组态作为输入, 经过类似GPT的Decoder网络输出波函数系数。每步迭代中, 首先使用批量自回归采样(BAS)生成样本, 然后利用二次量子化的哈密顿量计算总能量和梯度的期望值并优化网络参数, 直至能量收敛。Copyright © 2024 American Chemical Society

Figure 1 Workflow of QiankunNet-Solid^[21]. The QiankunNet wavefunction uses qubit configurations as input and outputs wavefunction coefficients through a GPT-like decoder network. In each iteration, batch autoregressive sampling (BAS) generates samples, the expectation values of total energy and gradients are evaluated using the second-quantized Hamiltonian, and network parameters are updated until energy convergence is achieved. Copyright © 2024 American Chemical Society

的哈密顿量系数。随后, 采用Jordan-Wigner变换^[23]将电子系统映射到自旋系统, 原始自旋轨道上的占据数表示被映射为同样数目的自旋比特的组态。QiankunNet-Solid采用一系列类似GPT的Decoder层作为波函数拟设, 与传统的神经网络结构(如卷积神经网络和循环神经网络)相比, 这种结构在捕捉输入序列的长程依赖性方面表现出色。网络以自旋系统的组态为输入, 经过嵌入层进行处理, 随后经过多个Decoder层得到波函数的振幅, 并额外使用一个多层感知机(multilayer perceptron, MLP)构成的子网络来表示波函数的相位。

在训练过程中, 网络首先根据波函数的概率幅采样一批自旋组态样本, 然后使用哈密顿量信息计算这一批次样本的能量期望值。这一能量期望值被用作损失函数, 通过反向传播可得到梯度的期望值, 并使用AdamW优化器对网络参数进行优化。在Hartree-Fock基组下、弱关联到中等关联程度的体系中, 基态波函数主要集中在Hartree-Fock态, 激发行列式的幅度较小, 使得传统MCMC算法的采样效率格外受限^[24]。然而, 由于GPT风格的Decoder网络自身具有自回归特性, 我

们可以采用自回归采样算法^[7,25]替代MCMC方法。在这种方法中, 每个样本的概率被分解为一系列条件概率的乘积, 而其中每个条件概率可以由网络单独得到。在生成样本的过程中, 根据这些条件概率对自旋组态的每一个或两个比特依次采样并计算下一步的条件概率, 直到生成整个样本。与MCMC方法相比, 这种算法等效于每次采样都以100%的接受概率生成样本^[16]。在QiankunNet中, 我们进一步采用了批量自回归采样方法, 显著提高了采样速度, 使得QiankunNet-Solid能够高效地在真实周期性体系上训练以得到体系的多电子波函数。

2.2 QiankunNet-Solid方法的数值计算和应用

为了展示这一方法的成效, 我们计算了数种一维(1D)、二维(2D)和三维(3D)真实周期性体系的基态能量^[21]。基于收敛的QiankunNet-Solid基态波函数, 还可以计算体系的多种信息; 例如, 采用运动方程方法, 可以计算弱关联到中等关联强度固体的准粒子能带谱, 以及态密度。这两类性质都是固体、材料系统中特有

且重要的性质,也是真实的计算材料学流程中必不可少的结果.图2和3展示了一维氢链、石墨烯以及硅晶体和金刚石晶体的基态能量计算结果,并和全组态相互作用(full configuration interaction, FCI)等多种方法的结果进行了对比.在每组数据中,QiankunNet-Solid结果与FCI结果的能量误差都在化学精度以内.尤其是对于一维氢链,当键长增大至一定程度时,受限Hartree-Fock (restricted Hartree-Fock, RHF)方法与受限耦合簇(restricted coupled cluster, RCC)方法失效,但QiankunNet-Solid结果仍然与参考值吻合,可见这一方法在弱关联至强关联范围内均可应用.图4展示了硅晶体的能带和态密度计算结果,并与基于FCI波函数的EoM方法计算结果进行了比较,可见QiankunNet-Solid结果与基于FCI波函数的运动方程方法计算结果吻合良好.这一结果显示QiankunNet-Solid具有满足固体和材料系统中基本的量子化学计算需求的能力.

3 QiankunNet-DMET方法在强关联材料中的应用

QiankunNet-Solid方法可以精确地获得能带、态密度等信息,但是由于计算复杂度的限制,QiankunNet-Solid方法难以拓展至大规模周期性系统.为此,我们将DMET方法^[26-32]与基于Transformer架构的QiankunNet求解器相结合,构建了QiankunNet-DMET框架^[22].DMET方法是量子嵌入方法(quantum embedding theory)^[26-31,33,34]的一种,通过量子嵌入策略将周期体系分割为“杂质-环境”子区域,利用DMET自洽迭代构建局部嵌入哈密顿量,并借助QiankunNet的高效采样与迁移学习技术,显著降低计算规模与训练成本.

3.1 QiankunNet-DMET方法框架

QiankunNet-DMET框架的核心是将全局系统划分为多个子区域(称为“杂质”),每个杂质嵌入更大的环境

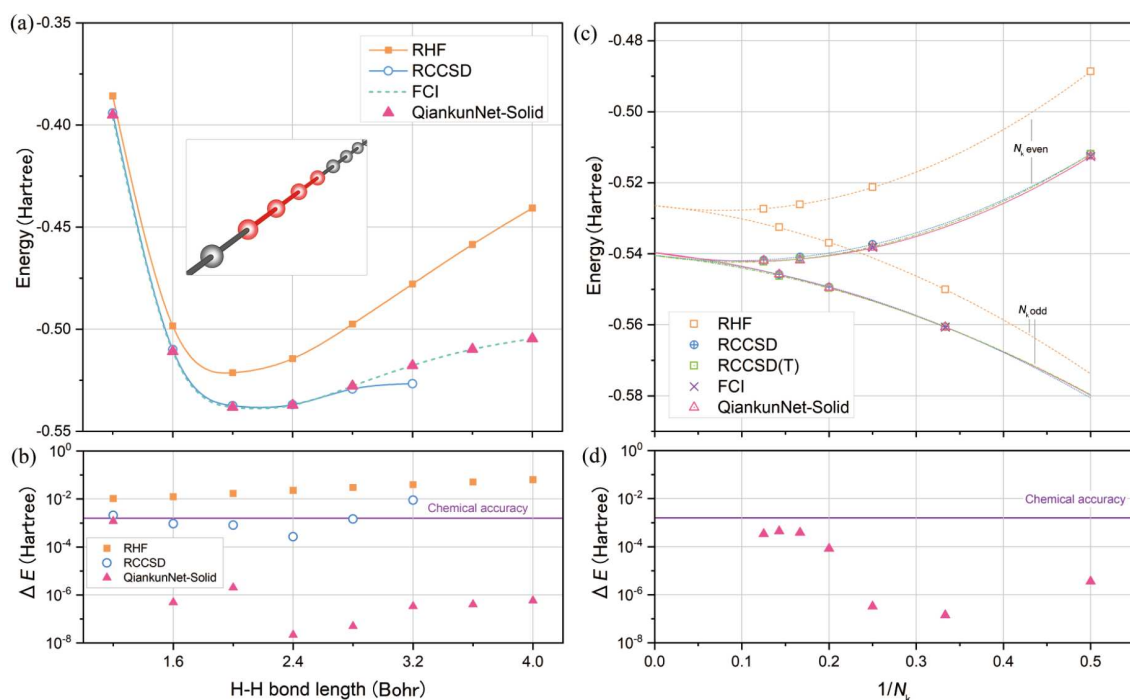


图2 QiankunNet-Solid对一维氢原子链的计算结果^[21]. (a) 分别采用RHF、RCCSD、FCI和QiankunNet-Solid方法计算的一维氢链在不同原子间距下的基态能量. 计算中使用了包含4个H原子的超胞、 $2 \times 1 \times 1$ 的k点. (b) 平均每个原子的能量与FCI结果的偏差. (c) 不同方法计算的一维氢链在原子间距为2 Bohr时, 基态能量收敛至热力学极限的趋势. (d) (c)中的平均每原子能量与RCCSD(T)的偏差. Copyright © 2024 American Chemical Society

Figure 2 The calculation results of QiankunNet-Solid for a one-dimensional hydrogen atomic chain^[21]. (a) Ground-state energies of a one-dimensional hydrogen chain at different interatomic distances, computed using RHF, RCCSD, FCI, and QiankunNet-Solid. A supercell containing 4 hydrogen atoms and a $2 \times 1 \times 1$ k-point mesh was used. (b) Deviation of the per-atom energy from the FCI results. (c) Convergence behavior of the ground-state energy toward the thermodynamic limit for a one-dimensional hydrogen chain at an interatomic distance of 2 Bohr. (d) Deviation of the per-atom energy from the RCCSD(T) results in (c). Copyright © 2024 American Chemical Society

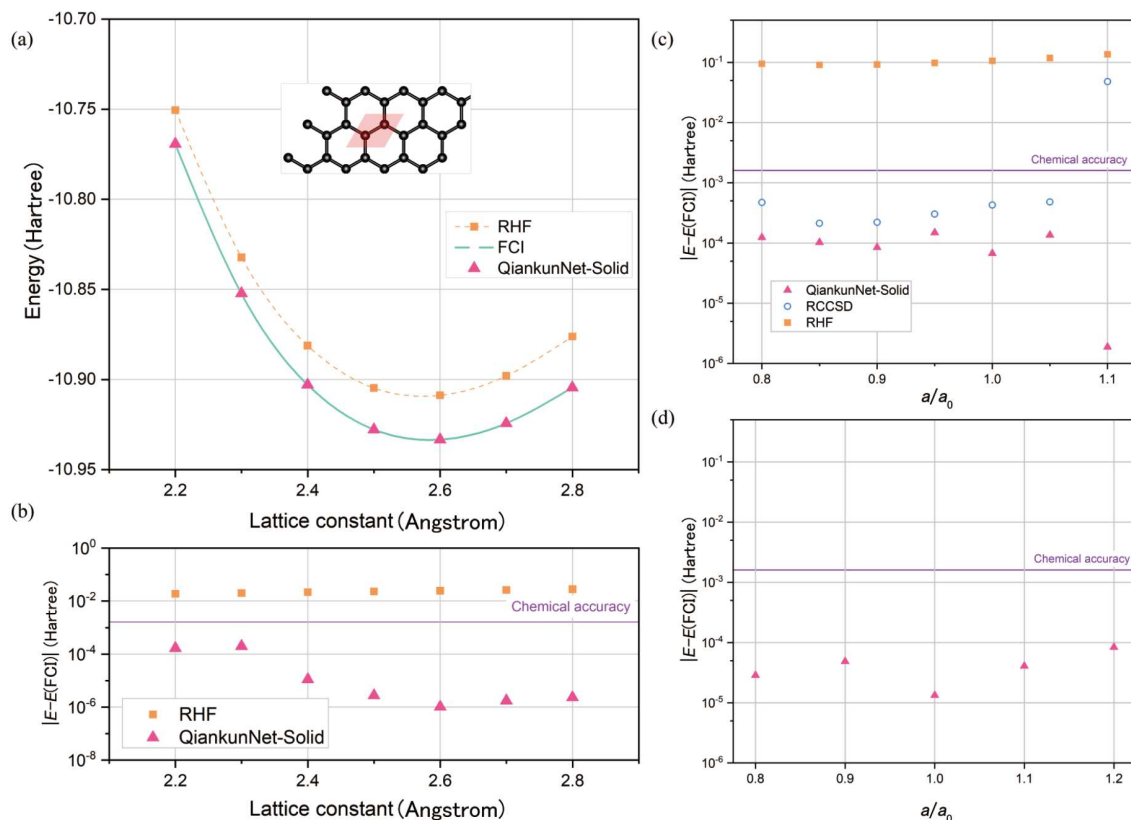


图 3 QiankunNet-Solid对石墨烯的计算^[21]. (a, b) 不同晶格常数下石墨烯的基态能量, 以及每晶胞的能量误差. 计算中采用 2×2 的k点采样, 并选取了包含16个自旋轨道的活性子空间. 本图中所有计算采用gth-szv基组. (c) 不同晶格常数下硅晶体的 Γ 点能量误差, 计算使用gth-szv基组, 每个晶胞包含16个轨道. (d) 不同晶格常数下金刚石晶体的 Γ 点能量误差, 计算使用gth-dzvp基组, 选取了包含16个轨道的活性空间. Copyright © 2024 American Chemical Society

Figure 3 The calculation results of QiankunNet-Solid for graphene^[21]. (a, b) Ground-state energies and per-cell energy errors of graphene at different lattice constants. Calculations were performed with a 2×2 k-point sampling and an active space containing 16 spin orbitals. All calculations in this figure employed the gth-szv basis set. (c) Energy errors at the Γ point for silicon crystals at different lattice constants, calculated using the gth-szv basis set with 16 orbitals per unit cell. (d) Energy errors at the Γ point for diamond crystals at different lattice constants, calculated using the gth-dzvp basis set and an active space containing 16 orbitals. Copyright © 2024 American Chemical Society

中, 并通过精确处理杂质与环境的相互作用实现电子结构的高效求解. QiankunNet被用于求解每个杂质子系统的多体波函数, 同时通过构建约化密度矩阵表征杂质及其邻近环境的关联效应.

QiankunNet-DMET计算的起点是固体材料的平均场计算, 并且通过局域化和分块, 将局域轨道划分为杂质和环境两部分. 杂质区域就是我们计算的核心区域, 而环境部分会被DMET方法约化. 一旦杂质区域被确定, 体系的密度矩阵可以被划分为: 杂质轨道间的相互作用、环境轨道间的相互作用, 以及杂质与环境间的相互作用这三部分. 在DMET方法中, 杂质轨道间的相互作用通常被保留, 而环境对杂质的影响通过密度矩阵的非对角部分(对应杂质轨道与环境轨道的相互作用)进行约化. 在具体操作中, DMET方法对密度矩阵的

非对角块 D_{IJ} 进行奇异值分解(singular value decomposition, SVD):

$$D_{IJ} = \sum_{\kappa} U_{I\kappa} \lambda_{\kappa} V_{\kappa J}^{\dagger}, I \in \text{env}, J \in \text{imp}, \quad (1)$$

基于SVD结果, 可以构建局域轨道到浴轨道的转换矩阵:

$$C = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & U_{I\kappa} \end{bmatrix}, \quad (2)$$

并且通过变换将哈密顿量从原来的局域基组变换到嵌入基组(由杂质轨道和浴轨道组成)上:

$$H_{\text{emb}} = C^{\dagger} H C. \quad (3)$$

通过这样的变换, 庞杂的环境轨道可压缩为仅含 N_{κ} 个虚轨道的“浴轨道”. 所谓浴轨道就是庞大的环

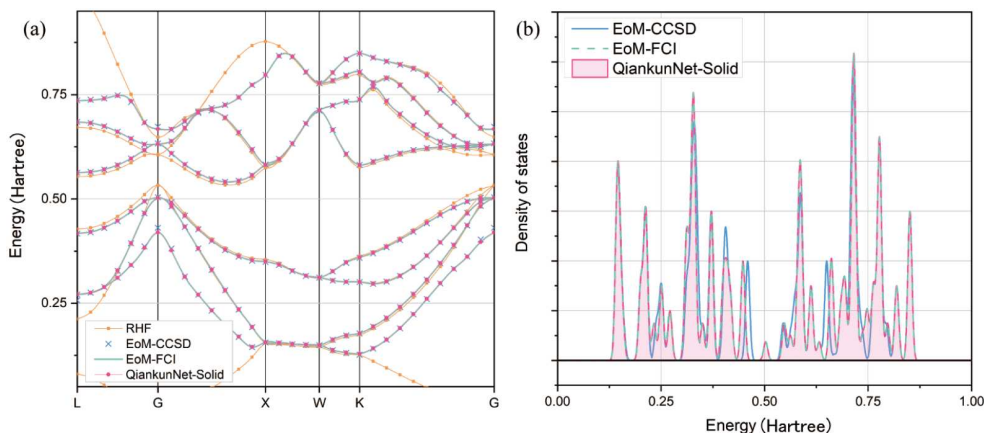


图4 (a) 采用QiankunNet-Solid以及数种其他方法计算得到硅的能带结构. 晶胞选取为包含两个硅原子的元胞, 每个晶格动量采用 $1 \times 1 \times 1$ 的k点采样, 针对每个晶格动量计算4个电子亲和能(EA)和4个电离势(IP), 得到费米能附近的8条准粒子能带. (b) 采用QiankunNet-Solid以及数种其他方法计算的硅晶体态密度. 晶格动量采取 $4 \times 4 \times 4$ 的均匀采样, 并在每个晶格动量点采用 $1 \times 1 \times 1$ 的k点采样, 对每个晶格动量计算费米能上下各4个激发能级. 两项计算均使用gth-szv基组^[21]. Copyright © 2024 American Chemical Society

Figure 4 (a) Band structure of silicon obtained by QiankunNet-Solid and several other methods. A primitive cell containing two silicon atoms was used, with $1 \times 1 \times 1$ k-point sampling for each crystal momentum. For each k-point, four electron affinities (EA) and four ionization potentials (IP) were computed, resulting in eight quasiparticle bands near the Fermi level. (b) Density of states of silicon calculated with QiankunNet-Solid and other methods. Crystal momenta were uniformly sampled with a $4 \times 4 \times 4$ grid, using $1 \times 1 \times 1$ k-point sampling at each crystal momentum. For each momentum, four excitation energies above and below the Fermi level were computed. Both calculations employed the gth-szv basis set^[21]. Copyright © 2024 American Chemical Society

境轨道的等效替代(具体细节详见文献[29,31]). 至此就是嵌入哈密顿量的构建过程, 对应图5中的 H_{emb} . 随后, 所有分块的嵌入哈密顿量通过QiankunNet求解, 得到基态能量和波函数(分别对应图5中的Energy和 $\Psi(x)$), 并由此计算单粒子与双粒子约化密度矩阵(图5中的1-RDM和2-RDM). 最终, 整合所有分块的信息, 构建全局体系的物性图像.

为确保自洽性, DMET算法需要进行两次迭代, 也就是图5中的 μ iteration和 u iteration. 第一次迭代的重点是电子数收敛, 即确保系统中的总电子数与原始电子数一致. 为此, 在嵌入哈密顿量中引入了一个额外的全局势 $\mu \sum \hat{a}_p^\dagger \hat{a}_p$ 用于调整体系中的总体电子数. 第二次迭代用于增强浴轨道的自洽构造, 通过在平均场计算中利用关联势 u 更新Fock矩阵来实现. 通过最小化平均场密度矩阵 $D_{pq}^{\text{mf}}(u)$ 与NNQS求解器得到的单粒子密度矩阵 D_{pq}^{NN} 之间的矩阵范数, 采用损失函数优化关联势 u :

$$\min \sum_{p,q} [D_{pq}^{\text{mf}}(u) - D_{pq}^{\text{NN}}]^2. \quad (4)$$

QiankunNet-DMET流程的关键挑战在于: DMET的两次迭代会生成大量嵌入哈密顿量, 导致计算时间显著增加. 然而, 迭代过程中哈密顿量的单电子与双电

子积分变化并不显著, 不会影响嵌入哈密顿量的整体结构, 表明所有哈密顿量具有相似的结构特征. 因此, 生成基态波函数的神经网络参数在不同迭代间差异较小. 基于此, 我们提出基于QiankunNet-DMET方法的迁移学习策略. 将前次迭代优化的神经网络参数作为后续迭代的初始值, 并保存每次迭代的最终参数用于重建网络. 由于初始参数(首次迭代除外)已接近最优解, 可降低神经网络的学习率并设置更严格的收敛阈值, 从而加速收敛并提升精度. 此外, 迁移学习策略可进一步拓展至具有相似轨道结构的碎片或不同体系. 例如, 一维氢链中不同键长的哈密顿量或金刚石中不同碳原子碎片, 其轨道排列模式相似时, 神经网络参数可跨体系迁移复用. 这一策略使初始训练步数减少50%以上, 显著降低了计算成本.

3.2 QiankunNet-DMET方法的数值验证与应用

为验证理论的有效性与精度, 我们将QiankunNet-DMET算法应用于从简单到复杂的一系列周期体系^[22], 相关结果如图6所示. 采用QiankunNet-DMET方法计算了H-H距离从0.6~2.0 Å的一维氢链的势能面, 结果与FCI方法一致, 特别是在H-H距离大于1.0 Å时达到化学精度, 而传统方法如CCSD方法在较大距离下通常低估

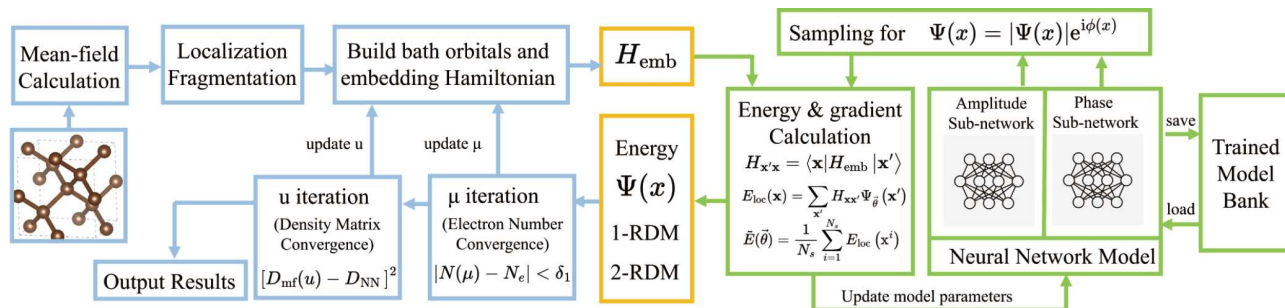


图5 QiankunNet-DMET算法流程图^[22]。DMET方法根据平均场计算的结果构建嵌入哈密顿量, 交给QiankunNet进行求解, 得到基态能量、波函数和约化密度矩阵。DMET方法利用QiankunNet的输出进行自洽迭代, 最终得到期望的物理性质。Copyright © 2024 Springer Nature

Figure 5 Flowchart of the QiankunNet-DMET algorithm^[22]. The DMET method constructs an embedded Hamiltonian based on mean-field calculation results. This Hamiltonian is then solved by QiankunNet to obtain the ground-state energy, wavefunction, and reduced density matrix. The DMET method subsequently uses QiankunNet's outputs to perform self-consistent iterations, ultimately converging to the desired physical properties. Copyright © 2024 Springer Nature

相关能。在金刚石晶体体系中, QiankunNet-DMET方法精确计算了系统能量, 并准确预测出平衡晶格常数, 精度优于传统CCSD求解器。在过渡金属氧化物体系中, QiankunNet-DMET方法对VO₂、MnO和NiO的磁性序进行了模拟, 成功预测出VO₂的铁磁序, 以及MnO、NiO的反铁磁(AF II)序, 计算结果如表1所示。可见, QiankunNet-DMET计算的磁矩和能量差均与实验和DFT数据高度一致, 验证了其在过渡金属氧化物中的适用性。在1T-TiSe₂体系中, QiankunNet-DMET方法通过比较正常有序结构与电荷密度波(charge density wave, CDW)畸变结构的能量和电子密度, 揭示了CDW畸变引起的Ti 3d与Se 4p轨道耦合增强以及周期性电子密度重分布。1T-TiSe₂的电荷密度分布计算结果如图6(e)所示, 计算结果与扫描隧道显微镜(scanning tunneling microscopy, STM)实验观测相符, 证明了该方法在处理低维集体现象中的高效性。

数值验证表明, QiankunNet-DMET方法不仅在处理含有3d过渡金属的强关联体系时表现出色, 而且在磁性序和电荷密度波等物性问题上与传统方法和实验数据符合良好。这一方法展现出的计算效率与精度优势为大规模强关联材料的从头算模拟提供了一条切实可行的新途径。

3.3 QiankunNet-DMET方法的计算效率提升

量子嵌入方法的优势在于它大大提升了计算效率、拓展了计算规模。通过将轨道划分为杂质和环境两部分, QiankunNet-DMET方法能高效地将周期性系统分割成多个片段, 大幅降低了计算成本, 同时也能有

效考虑各片段之间的关联。以氢链为例, 对不同k点数量下的情况进行测试: 氢链晶胞中包含2个氢原子, 且H-H距离固定为1.0 Å。我们统计了两种方法每一步的平均GPU时间, 计算时间测试结果见表2。结果表明, QiankunNet-DMET方法显著缩短了计算时间。造成这种计算时间差异的原因是多方面的, 其中主要原因是两种方法面对的希尔伯特空间的大小不同。对于QiankunNet-DMET方法, 由于量子嵌入方法的应用, 无论k点数目 N_k 取多少, QiankunNet-DMET都只需要处理一个4个轨道(包括两个H原子组成的分子轨道和2个浴轨道)构成的希尔伯特空间; 而对于QiankunNet-Solid方法, 采用 N_k 个k点的氢链对应了 $2 \times N_k$ 个轨道构成的希尔伯特空间, 包含的组态更加复杂, 并且随着k点的增加, 希尔伯特空间规模的增长是指数级的。从计算的角度来说, 更大的希尔伯特空间, 更复杂的组态, 要求使用更大规模的神经网络进行计算, 以增加QiankunNet的表达能力。同时由于哈密顿量轨道数的增多, 获取神经网络量子态的信息需要更大的计算量, 并且占用更多的显存。这些多种方面的原因影响了QiankunNet-Solid的计算效率。

QiankunNet-DMET方法的迁移学习策略也极大地提升了计算效率。图7(a)展示了一个DMET过程中, 除首个嵌入哈密顿量外, 初次需约8000步迭代才能收敛, 而后续加载先前模型的迭代步骤只需几百步迭代; 在 μ 迭代后期, QiankunNet求解器通常仅需几十步迭代即可收敛。

迁移学习策略同样可应用于不同片段或不同化学体系中相似嵌入哈密顿量的求解。例如, 在一维氢链势能面的计算中, 尽管H-H距离的变化会导致单电子和

表 1 QiankunNet-DMET计算得到的不同过渡金属氧化物的铁磁态(FM)与反铁磁态(AFM)的能量差、磁矩以及磁基态, 与文献参考值对比^[22]. Copyright © 2024 Springer Nature

Table 1 The energy differences between ferromagnetic (FM) and antiferromagnetic (AFM) states, magnetic moments, and magnetic ground states for various transition metal oxides (TMOs) calculated by the QiankunNet-DMET framework, alongside reference values from literature^[22]. Copyright © 2024 Springer Nature

金属氧化物	QiankunNet-DMET			参考值			文献
	ΔE	μ_B	磁基态	ΔE	μ_B	磁基态	
VO ₂	-105.54	1.043	FM	-89.4	1.05	FM	[35]
				-134.5	1.13	FM	[35]
				-	4.79	AFM	[36]
MnO	102.42	4.866	AFM	-	4.58	AFM	[37]
				94.33	-	AFM	[38]
				-	1.90	AFM	[36]
NiO	112.70	1.843	AFM	112.2	-	AFM	[39]
				104.0	1.725	AFM	[40]

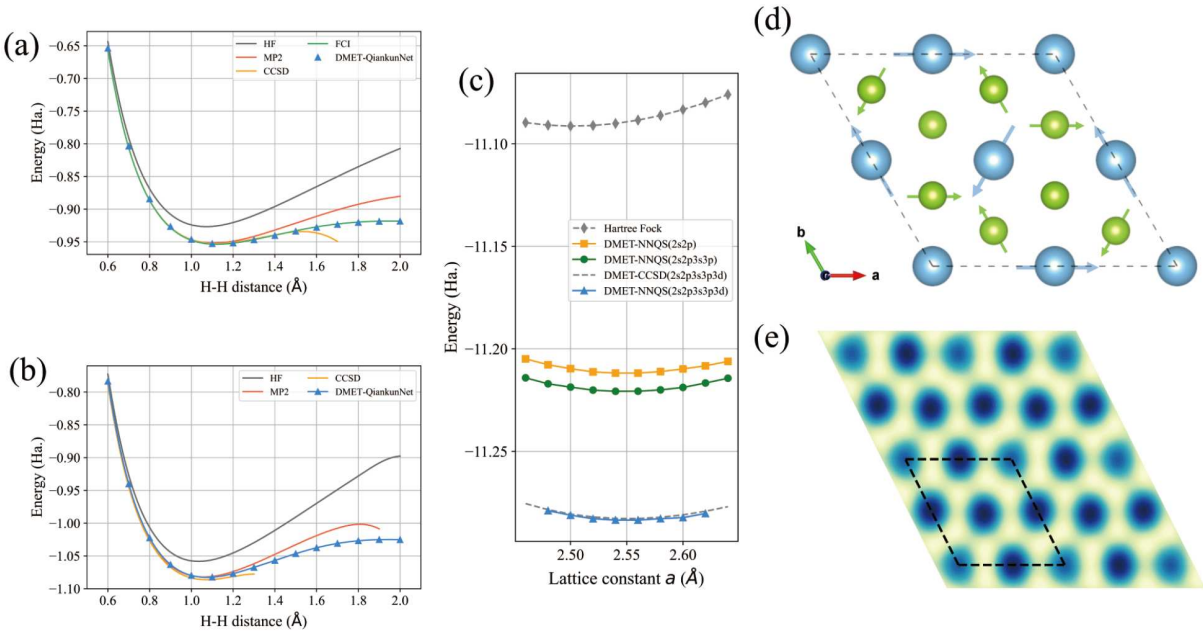


图 6 QiankunNet-DMET的数值计算结果^[22]. (a), (b) QiankunNet-DMET对一维氢链的势能面计算, 分别采用5/11个k点; (c) QiankunNet-DMET对金刚石晶体的计算结果; (d) 1T-TiSe₂的结构示意图和CDW不稳定性示意图; (e) QiankunNet-DMET计算的1T-TiSe₂的电荷密度波. Copyright © 2024 Springer Nature

Figure 6 Numerical results of QiankunNet-DMET^[22]. (a), (b) Potential energy surface calculations of a one-dimensional hydrogen chain using QiankunNet-DMET, with 5/11 k-points employed, respectively; (c) QiankunNet-DMET calculation results for diamond crystal; (d) schematic diagram of the structure and CDW instability of 1T-TiSe₂; (e) charge density wave of 1T-TiSe₂ calculated by QiankunNet-DMET. Copyright © 2024 Springer Nature

双电子积分的数值不同, 但嵌入哈密顿量总体的结构变化很小, 可以用同一个神经网络模型求解. 如图7(b)所示, 通过加载一个不同但相似的模型, 初始DMET迭代所需的步数从8000步降至不超过2000步. 第三种迁移学习场景涉及相似片段之间的迁移. 如图7(c)所示,

在金刚石的计算中, 将晶胞内两个碳原子划分为不同片段, 由于它们的哈密顿量相似, 便可采用迁移学习. 总之, 迁移学习策略显著提升了QiankunNet-DMET算法的效率, 使其效率提高数十倍. 对于一系列相似结构, 只需一次全面的QiankunNet训练即可显著降低求

表 2 QiankunNet-Solid方法和QiankunNet-DMET方法对不同 k 点数的一维氢链所需的单步迭代GPU时间(ms)^[22]. Copyright © 2024, Springer Nature

Table 2 GPU time per iteration (milliseconds) required by the QiankunNet-Solid and QiankunNet-DMET methods for a one-dimensional hydrogen chain with different numbers of k -points^[22]. Copyright © 2024, Springer Nature

N_k	QiankunNet-Solid	QiankunNet-DMET
3	40.73	34.99
5	318.86	33.49
7	373.04	32.60
9	599.24	33.81

解所有嵌入哈密顿量的计算成本, 从而使在更大系统上实施QiankunNet-DMET计算成为可能。

4 总结与展望

基于QiankunNet方法框架, 我们发展出两种针对固体材料特性的神经网络实现方案, 有效提升了神经网络量子态在周期性量子多体系统建模中的可操作性, 为拓展该方法在凝聚态体系中的应用提供了技术基准。QiankunNet-Solid基于二次量子化形式与周期性Hartree-Fock基组, 成功解决了波函数协变性、长程关联与热力学极限外推等难题, 在氢链、石墨烯、硅及金刚石等体系中得到了与FCI数据误差在化学精度以内的结果; QiankunNet-DMET通过引入量子嵌入方法, 将轨道划分为杂质和环境组分, 密度矩阵嵌入理论(DMET)方法为周期性体系的分区提供了一种高效方案, 从而

显著降低了计算成本。该方法不仅实现了体系的有效分块, 更重要的是为考虑区块间关联效应提供了可行途径。相较于传统的直接分块方法(如片段分子轨道法(FMO))在周期性体系处理上的局限性, DMET的分块方案具有更高的准确性和物理合理性。与已发展较为成熟的动力学平均场理论(DMFT)相比, DMET在物理图像表征层面具有相似性, 但计算复杂度显著降低。QiankunNet-DMET采用了迁移学习策略, 大幅降低求解嵌入哈密顿量计算成本, 在过渡金属氧化物(如 VO_2 、 MnO)和低维集体现象(如 TiSe_2 的电荷密度波)中精确预测磁性序与电子结构, 验证了其在强关联体系中的有效性。相较于传统神经网络量子态对周期体系的处理局限, 该方法为高温超导、拓扑绝缘体等复杂量子材料的研究构建了可扩展的计算框架。

QiankunNet-DMET和QiankunNet-Solid方法在不同情景下表现出各自的优势。QiankunNet-DMET方法的优势在于: 由于量子嵌入方法的引入, 可以拓展到较大的体系; 迁移学习的设计提升了计算效率; 增加 k 点并不显著影响计算复杂度。但QiankunNet-DMET方法也有相应的弊端, 该方法无可避免地带有由量子嵌入方法带来的误差, 当不同分块的轨道之间耦合较弱时, 这部分误差较小, 耦合较强时误差较大; 除此之外, QiankunNet-DMET更擅长捕捉局部性质, 长程相互作用被约化为杂质和浴轨道的相互作用, 难以精确量化; QiankunNet-DMET方法也更多用于基态性质的计算, 由于全局波函数的缺失, 在QiankunNet-DMET框架中

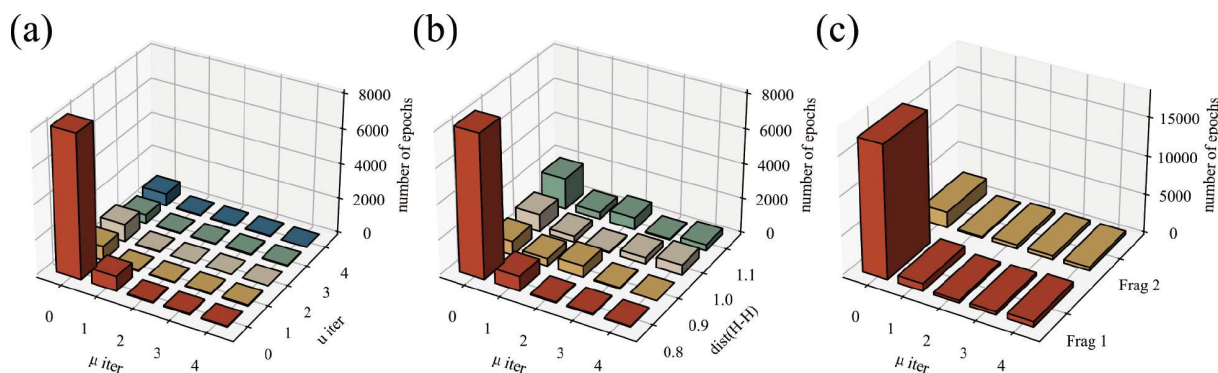


图 7 QiankunNet-DMET中涉及的各种不同迁移学习策略, 以及采取迁移学习策略后迭代步数的对比^[22]. 三种不同的迁移学习场景示意图: (a) 在DMET迭代过程中, 每个 μ 迭代步骤均加载前一步的模型, 同理, 每个 u 迭代步骤也加载前一步的模型; (b) 针对不同H-H距离下生成的氢链嵌入哈密顿量, 应用了迁移学习; (c) 在块状金刚石体系中, 不同相似片段之间采用了迁移学习。Copyright © 2024, Springer Nature

Figure 7 Various transfer learning strategies involved in QiankunNet-DMET and a comparison of iteration steps after applying transfer learning^[22]. Schematic illustrations of three different transfer learning scenarios: (a) during the DMET iteration process, the model from the previous step is loaded at each μ -iteration step, and similarly, the model from the previous step is also loaded at each u -iteration step; (b) transfer learning is applied to the embedded Hamiltonians of hydrogen chains generated at different H-H distances; (c) transfer learning is employed between different similar fragments in a bulk diamond system. Copyright © 2024, Springer Nature

计算非基态性质较为复杂. 并且, QiankunNet-DMET方法适用的体系规模也存在一定的限制, 如果在QiankunNet-DMET方法中分块过大, 则QiankunNet求解器需要很大的显存用于计算, 计算规模受显存的限制; 如果QiankunNet-DMET方法计算涉及的分块过多, 则DMET迭代不容易收敛. 相较于QiankunNet-DMET方法, QiankunNet-Solid方法直接求解体系的多体波函数, 避免了量子嵌入方法引入的误差, 其误差只来源于神经网络优化过程的收敛情况误差以及采样的随机性. 然而, QiankunNet-Solid方法需要完整考虑体系中各个自旋轨道之间的相互作用, 尤其在选取了多个k点时, 网络需要处理的总的自旋轨道数目随着k点数目呈线性增加. 一方面, QiankunNet中能量计算部分的复杂度随着轨道数目呈四次方增加; 另一方面, 在更多轨道数目的计算中, 每次迭代使用的采样量也需要一定程度的增加. 这不但在采样效率有限的情况下, 增加了生成样本、能量计算和梯度计算所需的时间, 而且增加的样本量也需要更大的内存和显存来处理. 这一问题的解决需要能量算法的进一步优化, 也需要采样生成和波函数计算方法的进一步并行化. 因此, QiankunNet-Solid方法目前尚只能用于中小规模、较少k点的计算, 在更大规模的问题上依然需要量子嵌入方法的帮助. 此外, QiankunNet-Solid方法可以依据运动方程理论, 对材料体系的能带谱和态密度等激发态性质进行计算. 不过, 运动方程理论采用了单粒子激发的近似, 因而只适用于弱关联至中等关联体系, 而在单粒子激发的近似失效的体系中则需要采用有约束的变分蒙特

卡罗等更严格的关联体系激发态计算方法. 未来, 我们将从多个方向进一步拓展QiankunNet系列方法的应用边界. 目前, 两种方法的计算规模受限于GPU显存, 我们计划研究分布式训练和轻量化网络架构, 以提升计算效率并扩展可模拟的体系规模. 通过对网络采用剪枝、量化等方式, 可能在几乎不损失精度的情况下减少计算开销; 同时也可能寻找更加简单、表达能力更强的网络结构, 来实现计算速度和精度的提升. 在分布式计算方面, 我们将基于GPU集群架构, 采用OpenMP和MPI结合, 对采样生成、能量计算、波函数计算等流程进行多卡并行优化, 进一步提升计算效率. 此外, 现有研究主要聚焦于材料的基态性质, 未来将增强对激发态和动力学性质的计算能力, 以支持光电磁响应等非平衡态研究. 在材料设计领域, QiankunNet-Solid能够高精度预测新型功能材料的电子结构, 而QiankunNet-DMET凭借量子嵌入方法, 能够处理更大尺度的体系. 两种方法相辅相成, 使得神经网络量子态在更多复杂材料中的模拟变成可能. QiankunNet-Solid在能源材料(如锂离子电池正极材料、固态电解质)和半导体材料性质计算(如带隙、有效质量及载流子迁移率计算)研究中展现出重要应用潜力. 对于复杂异质结构, 如量子阱、超晶格和二维材料异质结, QiankunNet-DMET的量子嵌入框架能够精确刻画界面电子重构和新奇物理现象. 展望未来, QiankunNet系列方法有望在强关联电子体系、量子材料设计及催化反应模拟等前沿领域发挥更大作用, 为揭示复杂量子现象和加速新材料发现提供有力的理论支撑.

参考文献

- Carleo G, Troyer M. Solving the quantum many-body problem with artificial neural networks. *Science*, 2017, 355: 602–606
- Deng D L, Li X, Sarma S D. Quantum entanglement in neural network states. *Phys Rev X*, 2017, 7: 021021
- Glasser I, Pancotti N, August M, et al. Neural-network quantum states, string-bond states, and chiral topological states. *Phys Rev X*, 2018, 8: 011006
- Sharir O, Shashua A, Carleo G. Neural tensor contractions and the expressive power of deep neural quantum states. *Phys Rev B*, 2022, 106: 205136
- Gao X, Duan L M. Efficient representation of quantum many-body states with deep neural networks. *Nat Commun*, 2017, 8: 662
- Shang H, Guo C, Wu Y, et al. Solving schrodinger equation with a language model. 2023, arXiv: 2307.09343
- Wu Y, Guo C, Fan Y, et al. NNQS-transformer: an efficient and scalable neural network quantum states approach for ab initio quantum chemistry. In: *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2023. 1–13
- Hedin L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys Rev*, 1965, 139: A796–A823
- Hybertsen M S, Louie S G. Electron correlation in semiconductors and insulators: band gaps and quasiparticle energies. *Phys Rev B*, 1986, 34: 5390–5413
- Georges A, Kotliar G, Krauth W, et al. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev Mod Phys*, 1996, 68: 13–125
- Anisimov V I, Poteryaev A I, Korotin M A, et al. First-principles calculations of the electronic structure and spectra of strongly correlated systems:

- dynamical mean-field theory. *J Phys-Condens Matter*, 1997, 9: 7359–7367
- 12 Hammond B L, Lester W A, Reynolds P J. Monte Carlo Methods in *Ab Initio* Quantum Chemistry. Singapore: World Scientific Publishing Co Pte Ltd., 1994
 - 13 Stanton J F, Bartlett R J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J Chem Phys*, 1993, 98: 7029–7039
 - 14 Moerman E, Miranda H, Gallo A, et al. Exploring the accuracy of the equation-of-motion coupled-cluster band gap of solids. *Phys Rev B*, 2025, 111: L121202
 - 15 Bloch F. Über die Quantenmechanik der Elektronen in Kristallgittern. *Z Physik*, 1929, 52: 555–600
 - 16 Cassella G, Sutterud H, Azadi S, et al. Discovering quantum phase transitions with fermionic neural networks. *Phys Rev Lett*, 2023, 130: 036401
 - 17 Pescia G, Han J, Lovato A, et al. Neural-network quantum states for periodic systems in continuous space. *Phys Rev Res*, 2022, 4: 023138
 - 18 Wilson M, Moroni S, Holzmann M, et al. Neural network ansatz for periodic wave functions and the homogeneous electron gas. *Phys Rev B*, 2023, 107: 235139
 - 19 Li X, Li Z, Chen J. *Ab initio* calculation of real solids via neural network ansatz. *Nat Commun*, 2022, 13: 7895
 - 20 Yoshioka N, Mizukami W, Nori F. Solving quasiparticle band spectra of real solids using neural-network quantum states. *Commun Phys*, 2021, 4: 106
 - 21 Fu L, Wu Y, Shang H, et al. Transformer-based neural-network quantum state method for electronic band structures of real solids. *J Chem Theor Comput*, 2024, 20: 6218–6226
 - 22 Ma H, Shang H, Yang J. Quantum embedding method with transformer neural network quantum states for strongly correlated materials. *npj Comput Mater*, 2024, 10: 220
 - 23 Jordan P, Wigner E. Über das Paulische Äquivalenzverbot. *Z Physik*, 1928, 47: 631–651
 - 24 Choo K, Mezzacapo A, Carleo G. Fermionic neural-network states for *ab-initio* electronic structure. *Nat Commun*, 2020, 11: 2368
 - 25 Barrett T D, Malyshev A, Lvovsky A I. Autoregressive neural-network wavefunctions for *ab initio* quantum chemistry. *Nat Mach Intell*, 2022, 4: 351–358
 - 26 Knizia G, Chan G K L. Density matrix embedding: a simple alternative to dynamical mean-field theory. *Phys Rev Lett*, 2012, 109: 186404
 - 27 Knizia G, Chan G K L. Density matrix embedding: a strong-coupling quantum embedding theory. *J Chem Theor Comput*, 2013, 9: 1428–1432
 - 28 Sun Q, Chan G K L. Quantum embedding theories. *Acc Chem Res*, 2016, 49: 2705–2712
 - 29 Wouters S, Jiménez-Hoyos C A, Sun Q, et al. A practical guide to density matrix embedding theory in quantum chemistry. *J Chem Theor Comput*, 2016, 12: 2706–2719
 - 30 Pham H Q, Hermes M R, Gagliardi L. Periodic electronic structure calculations with the density matrix embedding theory. *J Chem Theor Comput*, 2019, 16: 130–140
 - 31 Cui Z H, Zhu T, Chan G K L. Efficient implementation of *ab initio* quantum embedding in periodic systems: density matrix embedding theory. *J Chem Theor Comput*, 2019, 16: 119–129
 - 32 Cui Z H, Zhai H, Zhang X, et al. Systematic electronic structure in the cuprate parent state from quantum many-body simulations. *Science*, 2022, 377: 1192–1198
 - 33 Georges A. Strongly correlated electron materials: dynamical mean-field theory and electronic structure. In: AIP Conference Proceedings, 2004, 715: 3–74
 - 34 Yuan X, Zhang Y, Abtey T A, et al. VO₂: orbital competition, magnetism, and phase stability. *Phys Rev B*, 2012, 86: 235103
 - 35 Cheetham A K, Hope D A O. Magnetic ordering and exchange effects in the antiferromagnetic solid solutions Mn_xNi_{1-x}O. *Phys Rev B*, 1983, 27: 6964–6967
 - 36 Fender B E F, Jacobson A J, Wedgwood F A. Covalency parameters in MnO, α -MnS, and NiO. *J Chem Phys*, 1968, 48: 990–994
 - 37 Pask J E, Singh D J, Mazin I I, et al. Structural, electronic, and magnetic properties of MnO. *Phys Rev B*, 2001, 64: 024403
 - 38 Shanker R, Singh R A. Analysis of the exchange parameters and magnetic properties of NiO. *Phys Rev B*, 1973, 7: 5000–5005
 - 39 Twagirayezu F J. Density functional theory study of the effect of vanadium doping on electronic and optical properties of NiO. *Int J Comput Mat Sci*, 2019, 8: 1950007
 - 40 Wu Y, Xu X, Poletti D, et al. A real neural network state for quantum chemistry. *Mathematics*, 2023, 11: 1417

Summary for “QiankunNet-Solid/DMET: 面向固体材料模拟的生成式神经网络量子态方法”

QiankunNet-Solid/DMET: a generative neural network quantum state method for solid material simulations

Huan Ma^{1†}, Lizhong Fu^{1†}, Honghui Shang^{3*} & Jinlong Yang^{2,3*}

¹ School of Emerging Technology, University of Science and Technology of China, Hefei 230026, China

² Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China

³ Key Laboratory of Precision and Intelligent Chemistry, University of Science and Technology of China, Hefei 230026, China

† Equally contributed to this work

* Corresponding authors, E-mail: shh@ustc.edu.cn; jlyang@ustc.edu.cn

The rapid advancement of neural network quantum states (NNQS) has opened new avenues for simulating quantum many-body systems, yet their application to solid-state materials remains hindered by challenges such as periodic boundary conditions, long-range interactions, and computational scalability. This work introduces QiankunNet-Solid and QiankunNet-DMET, two innovative transformer-based NNQS frameworks tailored for simulating periodic and strongly correlated materials. Leveraging second quantization and autoregressive sampling, QiankunNet-Solid efficiently addresses wavefunction representation in periodic systems. By mapping electronic configurations to spin systems via Jordan-Wigner transformations and employing GPT-like decoder networks with multi-head attention mechanisms, it captures long-range dependencies in periodic lattices while maintaining translational symmetry. The method achieves chemical accuracy (errors <1 mHa) in predicting ground-state energies for 1D hydrogen chains, graphene, silicon, and diamond, matching full configuration interaction (FCI) benchmarks. Additionally, it enables accurate calculations of quasiparticle band structures and density of states through equation-of-motion theory, bridging the gap between traditional quantum chemistry methods and solid-state simulations.

To tackle strongly correlated systems, QiankunNet-DMET integrates density matrix embedding theory (DMET) with adaptive transfer learning strategies. This framework partitions periodic materials into localized impurity-environment subsystems, compressing environmental degrees of freedom via singular value decomposition (SVD) to construct low-dimensional embedded Hamiltonians. By reusing pre-trained neural network parameters across DMET iterations and structurally similar systems (e.g., hydrogen chains with varying bond lengths or carbon fragments in diamond), QiankunNet-DMET significantly reduces initial training steps and GPU time. In transition metal oxides such as VO₂, MnO, and NiO, the method predicts magnetic moments and energy differences between ferromagnetic (FM) and antiferromagnetic (AFM) states with sub-meV accuracy, aligning closely with experimental data. Furthermore, it resolves charge density wave (CDW) instabilities in 1T-TiSe₂, revealing orbital hybridization and electronic density redistributions consistent with scanning tunneling microscopy (STM) observations.

Numerical benchmarks underscore the framework's computational efficiency. QiankunNet-Solid demonstrates robustness across weak to strong correlation regimes, and QiankunNet-DMET exhibits near-constant computational costs regardless of k-point density, owing to its embedding nature. Comparative analyses highlight the synergy between the two methods: QiankunNet-Solid excels in high-precision small-system studies, while QiankunNet-DMET extends applicability to heterogeneous interfaces and bulk materials.

Looking ahead, the integration of QiankunNet-Solid and QiankunNet-DMET offers a scalable toolkit for exploring emergent quantum phenomena. Future directions include extending these methods to excited-state dynamics via constrained variational Monte Carlo and optimizing distributed training protocols using GPU clusters. Lightweight neural architectures, such as pruned transformer layers and quantized parameters, could further enhance efficiency without sacrificing accuracy. Potential applications span high-temperature superconductors, where precise modeling of cuprate electronic structures is critical, and topological materials requiring non-local correlation treatment. By combining quantum embedding strategies with machine learning, this work establishes a foundation for next-generation computational frameworks capable of unraveling complex quantum behaviors and accelerating the design of functional materials for energy storage, catalysis, and other important fields.

neural network quantum states, quantum Monte Carlo, density matrix embedding theory, periodic boundary conditions

doi: [10.1360/CSB-2025-0315](https://doi.org/10.1360/CSB-2025-0315)