

机器学习密度泛函理论新进展

蒋鸿

北京大学化学与分子工程学院, 北京分子科学国家研究中心, 北京 100871

E-mail: jianghchem@pku.edu.cn

New advances in machine learning density-functional theory

Hong Jiang

Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

E-mail: jianghchem@pku.edu.cn

doi: [10.1360/CSB-2025-5605](https://doi.org/10.1360/CSB-2025-5605)

密度泛函理论(density functional theory, DFT)已成为复杂分子和材料体系计算模拟的核心工具,并在近年来蓬勃兴起的科学智能(AI for Science)前沿研究中发挥着关键作用^[1]. DFT以电子密度为基本变量,把复杂多电子薛定谔方程的求解以原则上精确的方式映射为一个有效单电子薛定谔方程的求解问题.所有的量子多体相互作用都包含在交换关联(XC)密度泛函之中,虽然其精确形式至今仍然未知,但经过数十年的不懈努力,科学家已发展了一系列能较好地兼顾精度和效率的密度泛函近似,形成了总体精度系统性提升的近似泛函“雅各天梯”^[2].另一方面,尽管在复杂分子和材料体系的理论模拟中已取得了巨大成功,现有近似交换相关泛函在实际应用中仍然面临精度不足的困难^[3].传统的近似泛函发展的基本思路包括满足精确条件和半经验拟合,或两者的结合^[4].近年来,机器学习(machine learning, ML)技术的兴起为超越传统近似泛函局限性提供了可能性.利用以神经网络为代表的机器学习模型强大的表示能力,机器学习DFT有望在精度上显著优于现有泛函近似同时仍能保持传统DFT方法在计算效率上的优势,因而近年来吸引了大量理论研究者的关注^[5-11].具有代表性的是由谷歌公司旗下的DeepMind团队于2021年发表的基于含约40万参数多层感知机神经网络的DM21泛函^[5],通过对由1116个实际分子反应和1074个涉及分数电荷/自旋虚拟体系的高精度反应数据集的训练,不仅在主族元素测试集上获得了很高的精度,同时也能对涉及化学键解离、电荷转移等传统泛函有严重困难的体系给出非常准确的预测.另一方面,机器学习DFT领域尽管已取得了令人瞩目的进展,但仍还没有发展出性能上普遍优于传统泛函并能

保持计算效率优势的高精度、实用性密度泛函近似.

Gavini团队在近期的*Science Advances*发表研究论文^[12],提出以简单原子和分子体系求解逆Kohn-Sham(inverse Kohn-Sham)方程所获得的精确电子密度、交换相关能和交换相关势作为训练数据,学习构建适用于一般多原子体系的神经网络交换相关泛函,为开发高精度机器学习DFT方法开辟了新途径,基本思路如图1所示.

Gavini团队首先通过高精度的组态相互作用(CI)计算获得了数值精确的基态电子密度,然后通过求解逆Kohn-Sham问题获得了对应的精确交换相关能和交换相关势.逆Kohn-Sham问题是指给定电子密度求解对应的交换相关势,在数学上可以表述为一个约束最小化问题,传统的求解方法存在很多数值计算上的困难. Gavini团队近年来发展了基于系统性收敛有限元基组的伴随态方法,可以确保了数值计算的稳定性和准确性.但是,无论是高精度CI计算,还是逆Kohn-Sham问题的求解,都要求非常高昂的计算成本,因此目前只能对电子数较少的原子和简单分子体系才能获得数值精确的交换关联势.

在获得精确的交换相关势和能量后, Gavini团队分别在局域密度近似(LDA)和广义梯度近似(GGA)框架内构建相应的神经网络DFT(NN-DFT)泛函近似——分别称为NNLDA和NNGGA.作者采用了在已有局域密度近似PW92和广义梯度近似PBE泛函基础上添加NN泛函修正项的设计逻辑.对于一般的泛函表达式 $E_{xc}^{NN}[\rho] = \int e_{xc}^{NN}[\rho](\mathbf{r})d\mathbf{r}$,相应的NNLDA和NNGGA泛函的被积函数表示为如下形式:

引用格式: 蒋鸿. 机器学习密度泛函理论新进展. 科学通报, 2025, 70: 5773-5775

Jiang H. New advances in machine learning density-functional theory (in Chinese). Chin Sci Bull, 2025, 70: 5773-5775, doi: [10.1360/CSB-2025-5605](https://doi.org/10.1360/CSB-2025-5605)

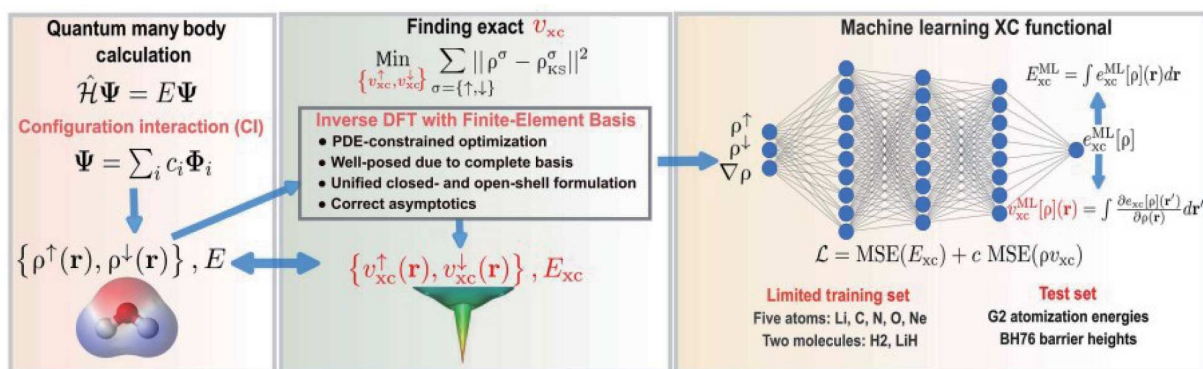


图 1 基于精确交换-相关势和能量训练神经网络交换-相关泛函方法的基本思路和工作流程示意^[12]. Copyright © 2025, The American Association for the Advancement of Science

Figure 1 Schematic illustration on the construction of the Neural network exchange-correlation functional based on exact Kohn-Sham XC potential and energy^[12]. Copyright © 2025, The American Association for the Advancement of Science

$$e_{xc}^{NNLDA}[\rho](\mathbf{r}) = e_{xc}^{PW92}(\rho(\mathbf{r}), \zeta(\mathbf{r})) + e_x^{UEG}[\rho](\mathbf{r})\phi(\mathbf{r})G^{NNLDA}(\rho(\mathbf{r}), \zeta(\mathbf{r})), \quad (1)$$

$$e_{xc}^{NNGGA}[\rho](\mathbf{r}) = e_{xc}^{PBE}(\rho(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r})) + e_x^{UEG}[\rho](\mathbf{r})\phi(\mathbf{r})G^{NNGGA}(\rho(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r})), \quad (2)$$

其中, e_{xc}^{PW92} 和 e_{xc}^{PBE} 分别是PW92-LDA和PBE-GGA近似泛函中的交换能密度, e_x^{UEG} 是均匀电子气(UEG)交换能密度, $\zeta(\mathbf{r}) = (\rho^\uparrow(\mathbf{r}) - \rho^\downarrow(\mathbf{r}))/\rho(\mathbf{r})$ 为相对自旋密度, $\phi(\mathbf{r}) = ((1 + \zeta(\mathbf{r}))^{4/3} (1 - \zeta(\mathbf{r}))^{4/3})/2$, $s(\mathbf{r}) = |\nabla\rho(\mathbf{r})|/[2(3\pi^2)^{1/3}\rho^{4/3}(\mathbf{r})]$ 为约化密度梯度. G^{NNLDA} 和 G^{NNGGA} 是分别为以 $\{\rho, \zeta\}$ 和 $\{\rho, \zeta, s\}$ 为输入描述符的前馈神经网络. 另外, 考虑到拟合原子和分子数据训练得到的NN交换相关泛函不满足UEG极限, 并可能在 $s \rightarrow 0$ 极限处表现出尖锐的梯度, 这对原子和分子的计算不会造成什么困难, 但可能在其他应用(比如固体)中导致难以收敛的情形, 因此研究团队也训练构建了如下明确满足UEG极限NNGGA-UEG泛函.

$$e_{xc}^{NNGGA-UEG}[\rho](\mathbf{r}) = e_{xc}^{PBE}(\rho(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r})) + e_x^{UEG}[\rho](\mathbf{r})\phi(\mathbf{r})\tanh(s(\mathbf{r})^{5/2}) G^{NNGGA}(\rho(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r})). \quad (3)$$

Gavini团队使用五个原子(Li、C、N、O、Ne)和两个简单分子(H_2 和LiH)以及均匀电子气(UEG)的精确交换相关势和能量数据训练了以上三个NN-DFT模型, 并在G2和GMTKN5等数据集上测试了它们在总能量、基态电子密度、热化学性质等方面的预测精度, 发现NN泛函总体性能上显著优于现有LDA和GGA泛函, 并在很多方面达到与meta-GGA(SCAN)和杂化泛函(B3LYP)相当的性能表现. 在总能量方面, 如图2所示, NNLDA和NNGGA不仅超越了广泛使用的LDA(PW92)和GGA(PBE)泛函, 甚至超过了更高层次的杂化泛函

B3LYP. 特别值得注意的是, NNGGA在每原子总能量上达到了小于2 kcal/mol的平均绝对误差(MAE), 非常接近化学精度的1 kcal/mol. 这意味着NN泛函在相对能量(如原子化能和势垒高度)方面的任何改进都将源于对单个系统能量的准确描述, 而不是依赖许多XC泛函中常见的误差系统抵消. 在热化学测试方面, 研究团队从GMTKN55数据集选择了色散作用贡献可忽略的19个子集来评估了NN泛函的性能. 这些子集包含了原子化能、电离势、电子亲和能、反应能和势垒高度等关键化学性质. 如图2所示, NNGGA和NNGGA-UEG在大多数子集上均优于PBE, 甚至在多数情况下也超过了SCAN和R2SCAN. 在WTMAD-1和WTMAD-2两个聚合指标上, NNGGA和NNGGA-UEG达到了与SCAN/R2SCAN几乎

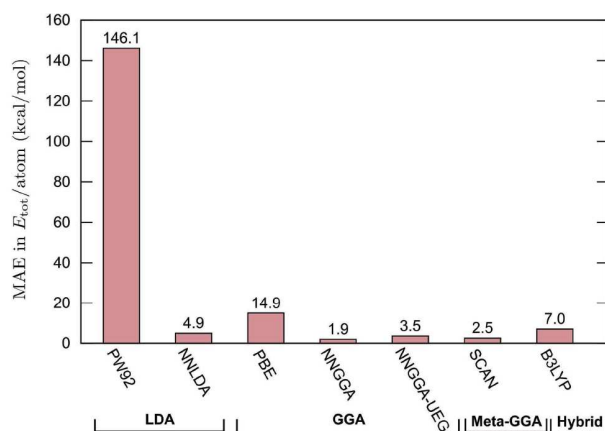


图 2 采用不同泛函近似计算G2数据集中所含由第一、二周期元素构成的97个分子总能量的平均绝对误差^[12]. Copyright © 2025, The American Association for the Advancement of Science

Figure 2 Mean absolute error (MAE) for the 97 molecules in the G2 dataset containing up to second-row elements calculated by different density functional approximations^[12]. Copyright © 2025, The American Association for the Advancement of Science

相同的精度. 考虑到NN泛函的训练过程中完全没有使用目标热化学量, 以上性能尤为引人注目.

Gavini团队的这项研究最值得关注的发现是, 精确交换相关势数据在训练构建强迁移性、数据高效机器学习泛函中的关键作用. 之前发展的机器学习DFT方法通常基于密度和/或总能量, 同时需要使用大量的高精度量子化学计算数据作为训练集. 而该工作仅使用了很少量简单体系的精确交换-相关势作为训练的目标, 代表了一种发展机器学习泛函的新思路. 作者认为该方法具有如下四方面的优势: 第一, 通过使用精确的电子密度和对应的精确交换-相关势, 该方法一定程度上嵌入了DFT的基本原理, 即Hohenberg-Kohn定理和Kohn-Sham假设, 从而获得了更强的泛化能力和数据效率; 第二, 使用交换相关势有助于隐式地辅助机器学习模型满足精确泛函的一些条件, 从而增强其泛化能力; 第三, 相较于电子密度, 交换相关势提供了一个更具指导性的学习目标, 因为密度的微小变化可能导致相应XC势的较大变化, 因此交换相关势具有更强的区分体系的能力, 从而包含更多的信息; 第四, 使用交换相关势作为学习目标规避了在训练过程中进行SCF计

算的必要性, 而后者在使用复杂的机器学习模型和/或较大的训练数据集的情形中会导致昂贵的训练成本.

尽管Gavini团队所发展的NN交换相关泛函在原子和分子系统上取得了显著成功, 但也仍然面临很多的挑战. 首先, 由于模型训练主要基于包含轻元素的简单原子和分子的数据, 这些泛函是否具备对含重元素分子体系和固体材料体系的外推能力尚未可知. 原则上, 通过在训练集加入重元素分子和固体体系的数据可以提高其模型泛化能力, 但重元素分子和固体体系的高精度电子密度的获得以及相应的逆Kohn-Sham方程的求解存在更大的挑战性. 其次, LDA和GGA水平的半局域泛函存在一些固有的局限性, 包括较为突出的自相互作用误差和静态相关误差, 这些困难在NNLDA和NNGGA中依然存在, 从而限制了它们对包含开壳层原子的强关联电子体系中的应用. 最后, 精确的交换相关势只有在严格的Kohn-Sham理论(即要求交换-相关势为空间坐标的局域函数)才有明确定义, 如何将类似的思路扩展到基于非局域交换-相关势的广义Kohn-Sham理论(如杂化泛函), 在概念和实践上仍需要做进一步的探索.

参考文献

- Huang B, von Rudorff G F, von Lilienfeld O A. The central role of density functional theory in the AI age. *Science*, 2023, 381: 170–175
- Zhang I Y, Xu X. On the top rung of Jacob's ladder of density functional theory: toward resolving the dilemma of SIE and NCE. *WIREs Comput Mol Sci*, 2021, 11: e1490
- Cohen A J, Mori-Sánchez P, Yang W. Challenges for density functional theory. *Chem Rev*, 2012, 112: 289–320
- Perdew J P. Artificial intelligence “sees” split electrons. *Science*, 2021, 374: 1322–1323
- Kirkpatrick J, McMorrow B, Turban D H P, et al. Pushing the frontiers of density functionals by solving the fractional electron problem. *Science*, 2021, 374: 1385–1389
- Zhuang Y, Gu Y, Zhang B, et al. Machine learning accurate exchange–correlation potentials for reducing delocalization error in density functional theory. *JACS Au*, 2025, 5: 4002–4010
- Liu X Y, Xiang Q X, Wang S R, et al. Machine learning parameters of optimally tuned range-separated hybrid functionals for transition metal complexes. *J Phys Chem Lett*, 2025, 16: 9672–9681
- Khan D, Price A J A, Huang B, et al. Adapting hybrid density functionals with machine learning. *Sci Adv*, 2025, 11: eadt7769
- Wang Y, Lin Z, Ouyang R, et al. Toward efficient and unified treatment of static and dynamic correlations in generalized Kohn–Sham density functional theory. *JACS Au*, 2024, 4: 3205–3216
- Pederson R, Kalita B, Burke K. Machine learning and density functional theory. *Nat Rev Phys*, 2022, 4: 357–358
- Ma H, Narayanaswamy A, Riley P, et al. Evolving symbolic density functionals. *Sci Adv*, 2022, 8: eabq0279
- Kanungo B, Hatch J, Zimmerman P M, et al. Learning local and semi-local density functionals from exact exchange-correlation potentials and energies. *Sci Adv*, 2025, 11: eady8962