北京大学
PEKING UNIVERSITY

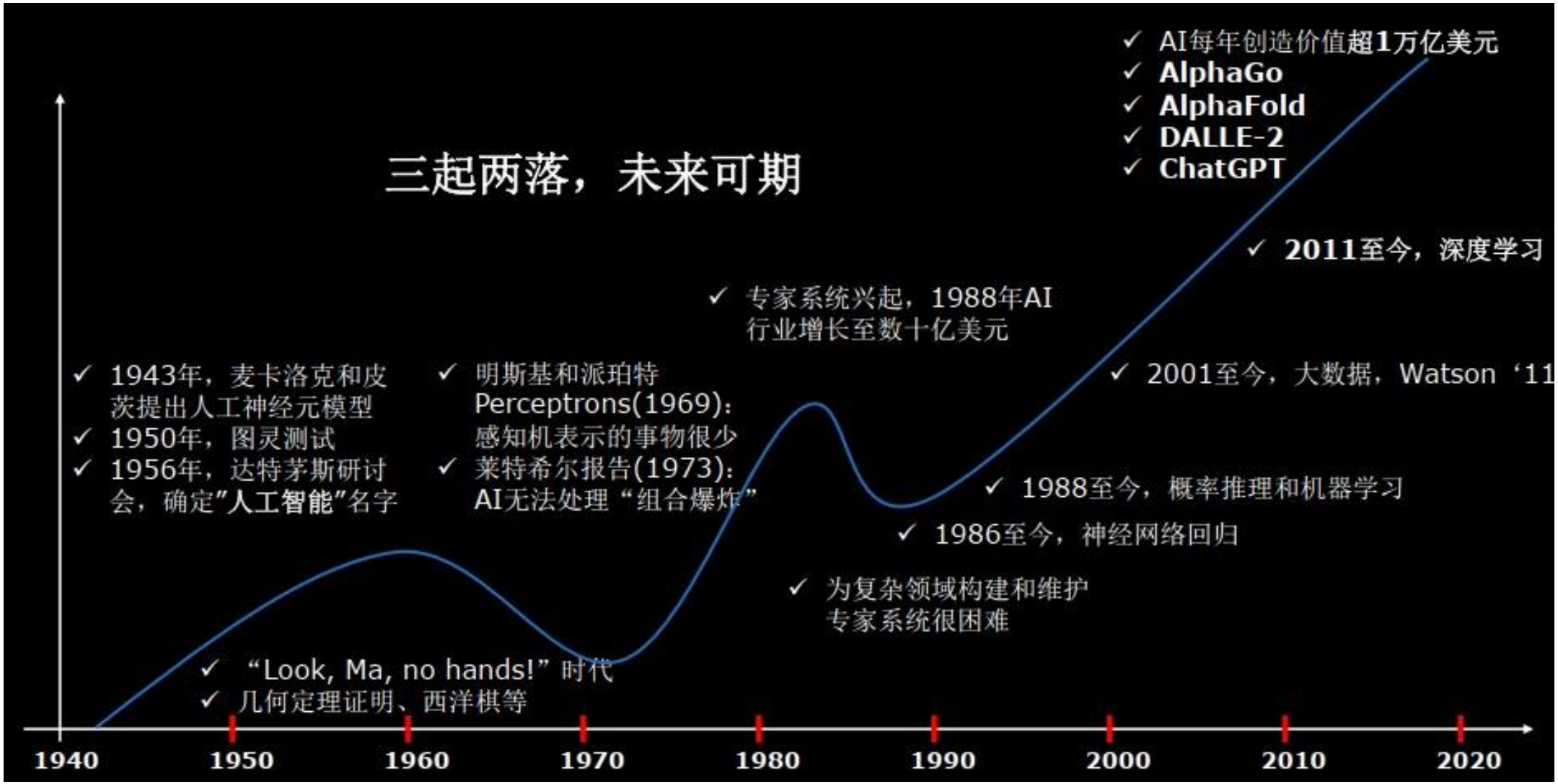
AI与凝聚态物理

陈 基

2026-11-02

物理与人工智能课程讲座

AI


三起两落，未来可期
✓ AI每年创造价值超1万亿美元
✓ AlphaGo
✓ AlphaFold
✓ DALLE-2
✓ ChatGPT
✓ 2011至今，深度学习
✓ 专家系统兴起，1988年AI行业增长至数十亿美元
✓ 1943年，麦卡洛克和皮茨提出人工神经元模型
✓ 1950年，图灵测试
✓ 1956年，达特茅斯研讨会，确定"人工智能"名字
✓ 明斯基和派珀特Perceptrons(1969)：感知机表示的事物很少
✓ 莱特希尔报告(1973)：AI无法处理"组合爆炸"
✓ 2001至今，大数据，Watson '11
✓ 1988至今，概率推理和机器学习
✓ 1986至今，神经网络回归
✓ 为复杂领域构建和维护专家系统很困难
✓ "Look, Ma, no hands!"时代
✓ 几何定理证明、西洋棋等
1940
1950
1960
1970
1980
1990
2000
2010
2020

来自马滟青

凝聚态物理

① **萌芽期（19 世纪末–1920s）：**从经典晶体性质与 "比热之谜" 出发，Drude 电子模型、Einstein 与 Debye 的比热量子理论、Born 的晶格动力学为固体理论奠基。

② **量子奠基（1925–1940s）：**量子力学催生能带理论（Bloch、Sommerfeld、Wilson），固体量子理论成型；1947 年晶体管发明开启半导体时代。

③ **黄金时代（1940s–1960s）：**费米液体理论、BCS 超导理论、Anderson 局域化、Ginzburg–Landau 理论，Born–Huang 晶格动力学，使 "固体物理" 体系成熟。

④ **学科确立与爆发（1967–1980s）：P.W. Anderson 于 1967 年正式以 "凝聚态物理" 命名这一学科**；随后重整化群、KT 相变、整 / 分数量子霍尔效应、高温超导等密集爆发。

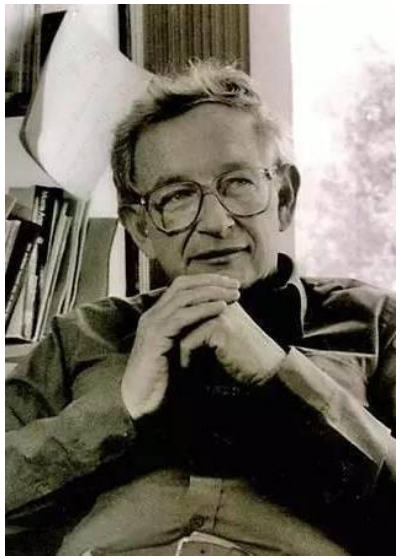
⑤ **现代前沿（1990s–今）：**学科正向量子材料结构、电子物态的调控与应用延伸，与其他学科交叉发展。

1967 年，时任剑桥大学 **Cavendish 实验室**的 Anderson 与 Heine 一道，将他们研究组的名称从 **"Solid State Theory（固体理论）" 改为 "Theory of Condensed Matter（凝聚态理论）"**，理由是原名称无法涵盖他们关心的液体、无序体系、多体现象等。
1978 年，美国物理学会（APS）的 "固体物理分部" 正式更名为 "凝聚态物理分部"，使这一术语成为学科的标准命名。

凝聚态物理

凝聚态物理是研究大量微观粒子（原子、电子、离子等）通过相互作用凝聚形成的宏观物质体系的物理性质与规律的学科。

1972 **More Is Different:** Broken symmetry and the nature of the hierarchical structure of science.



Philip Anderson

多者异也

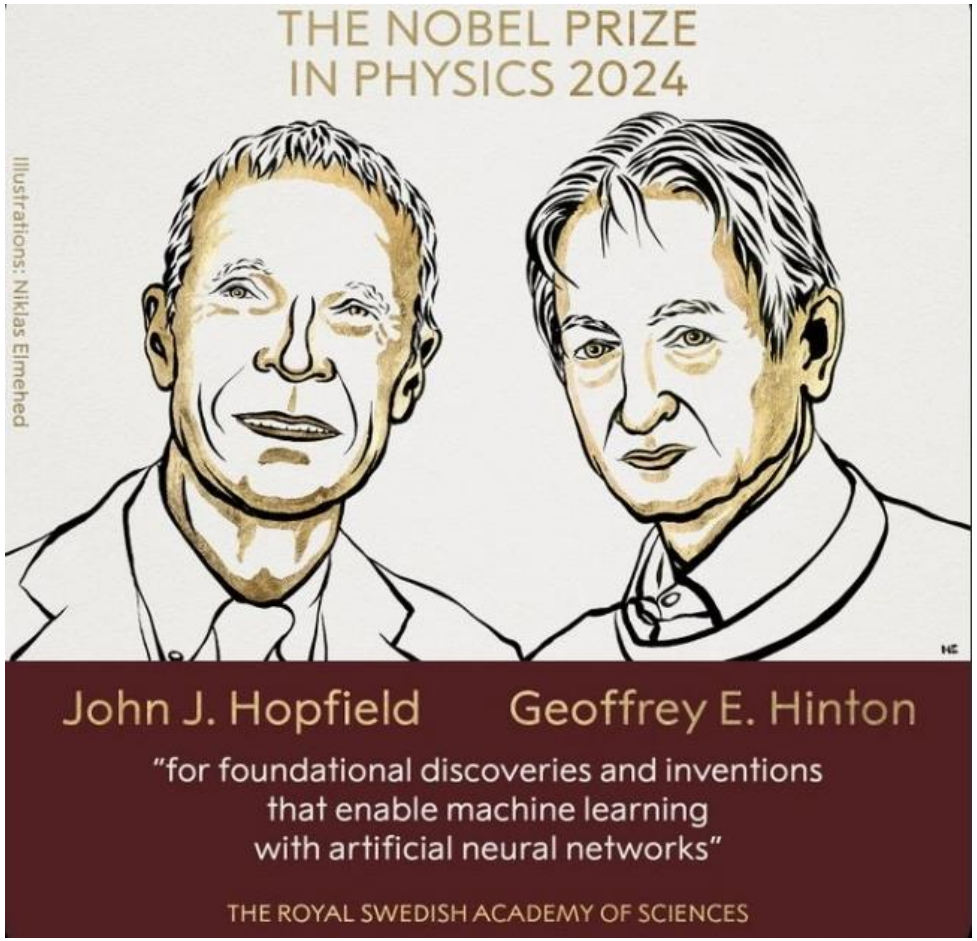
according to the idea: The elementary entities of science X obey the laws of science Y.

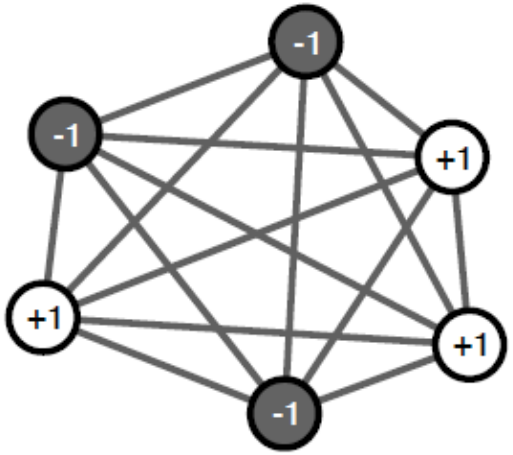
X	Y
solid state or many-body physics	elementary particle physics
chemistry	many-body physics
molecular biology	chemistry
cell biology	molecular biology
•	•
•	•
•	•
psychology	physiology
social sciences	psychology

凝聚态物理（诺奖代表）

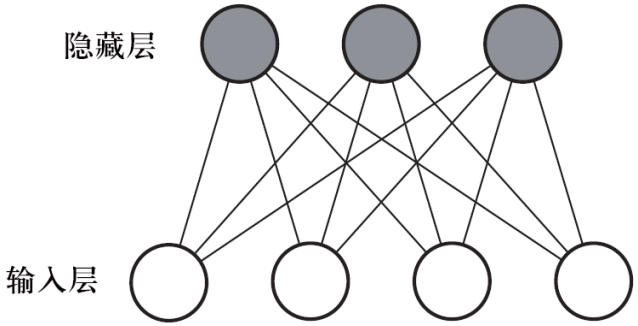
获奖年份	成果内容	获奖者	成果年份
1977	无序体系电子局域化、莫特金属 - 绝缘体转变、磁性理论	Anderson, Mott, Van Vleck	1977
1985	整数量子霍尔效应	von Klitzing	1980
1986	扫描隧道显微镜（STM）发明	Binnig, Rohrer	1981
1987	铜氧化物高温超导（La-Ba-Cu-O）	Bednor, Müller	1986
1996	^{3}He超流的实验发现	Lee, Osheroff, Richardson	1972
1998	分数量子霍尔效应	崔琦, Störmer, Laughlin	1982
2001	稀薄原子气体玻色–爱因斯坦凝聚（BEC）实验实现	Cornell, Wieman, Ketterle	1995
2003	超导体和超流理论	Abrikosov, Ginzburg, Leggett	1950s-1970s
2007	巨磁电阻效应（GMR）	Fert, Grünberg	1988
2016	KT 拓扑相变、Haldane 猜想	Thouless, Kosterlitz, Haldane	1972, 1983
2021	自旋玻璃、无序复杂系统的复本理论	Parisi	1970s
2025	超导电路中的宏观量子隧穿与能级量子化	Clarke, Devoret, Martinis	1980s

凝聚态物理 for AI





Hopfield网络 （1982）

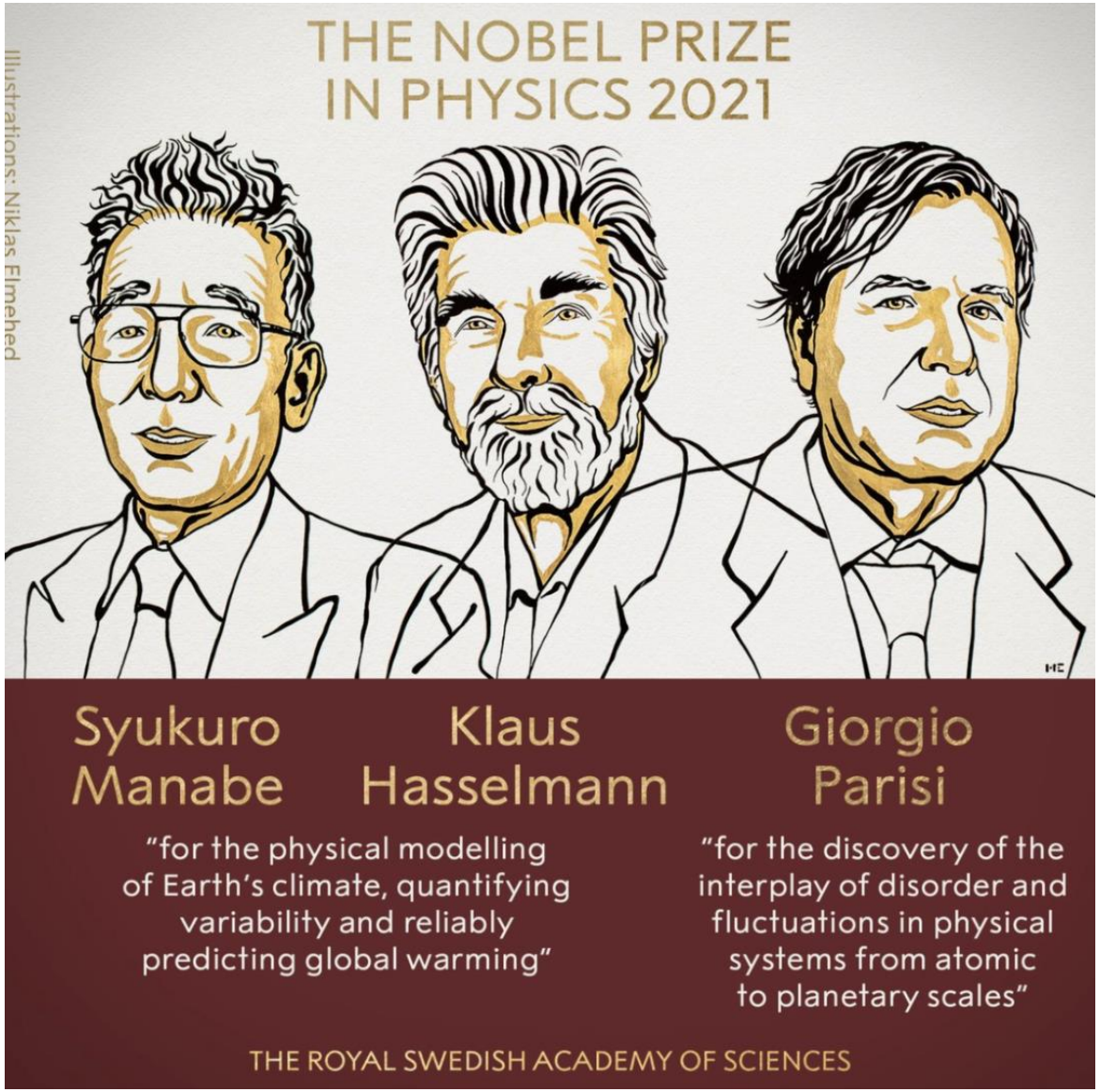


玻尔兹曼机（1985）

自旋玻璃模型

$$\mathrm{H} = -\sum_{<i,j>} J_{ij}\sigma_i\sigma_j$$

凝聚态物理 for AI



乔治·帕里西（Giorgio Parisi）获得2021年诺贝尔物理学奖最主要的工作是“发现了从原子到行星尺度的物理系统中无序和波动的相互作用”。

其中最主要的工作是1980年左右一系列对自旋玻璃的研究工作，他提出了复本对称破缺的概念并将其应用到Sherrington-Kirkpatrick自旋玻璃模型中，给出了平衡态的解。

他的工作对与理解神经网络的复杂行为和动力学特性，例如神经网络中的信息存储、检索机制和神经网络的优化都有重要的意义。

凝聚态物理 for AI

伊辛模型 $$H = -J\sum_{<i,j>}\sigma_i\sigma_j - h\sum_i \sigma_i$$

模型：它假设整个体系可以分为分立的格点，每个格点具有一定的自旋。整个体系的性质主要由相邻格点之间的自旋交换相互作用和每个格点上自旋受外磁场的作用决定。

背景：19 世纪末，物理学家已观察到铁磁体（如铁、镍）在高温下会失去磁性，低温下恢复磁性，但无法从微观层面解释。当时统计力学已能描述理想气体等简单系统，但面对磁体材料这样的“大量粒子通过相互作用形成集体行为”的复杂系统，缺乏简洁的数学模型。

历史：1920年，Ernst Ising在其博士论文（在导师Wilhelm Lenz的指导下）提出，并实现了一维模型的求解。1944年，昂萨格完成了二维模型的求解。1952年，杨振宁、李政道 基于伊辛模型的研究提出李-杨零点理论。

凝聚态物理 for AI

自旋玻璃模型 $$H = -\sum_{<i,j>} J_{ij}\sigma_i\sigma_j$$

自旋玻璃模型是凝聚态物理中描述**无序磁性系统**的经典模型。

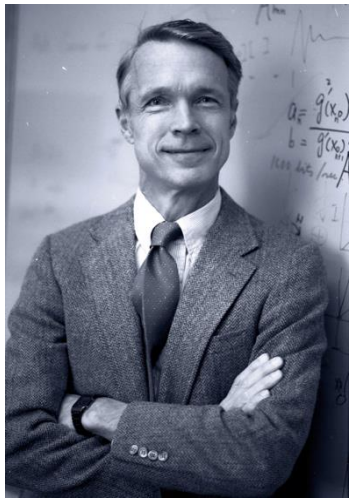
自旋玻璃：合金中由于杂质的存在导致相邻自旋间的交换相互作用是**随机且竞争的**，这种“相互作用竞争”使系统无法形成规则有序结构，低温下呈现独特的“玻璃态”（类似玻璃的非晶无序，无长程磁有序）。

自旋玻璃的研究始于 20 世纪 60 年代末 70 年代初。

1970 年，Harry Coles首先用“自旋玻璃”一词来描述稀释合金 AuCo 的特殊磁性。

同年，Philip W. Anderson也用自旋玻璃从理论上说明 CuMn 合金的特殊磁性。

John Hopfield



PHYSICAL REVIEW VOLUME 112, NUMBER 5 DECEMBER 1, 1958

Theory of the Contribution of Excitons to the Complex Dielectric Constant of Crystals*†

J. J. HOPFIELD‡
Physics Department, Cornell University, Ithaca, New York

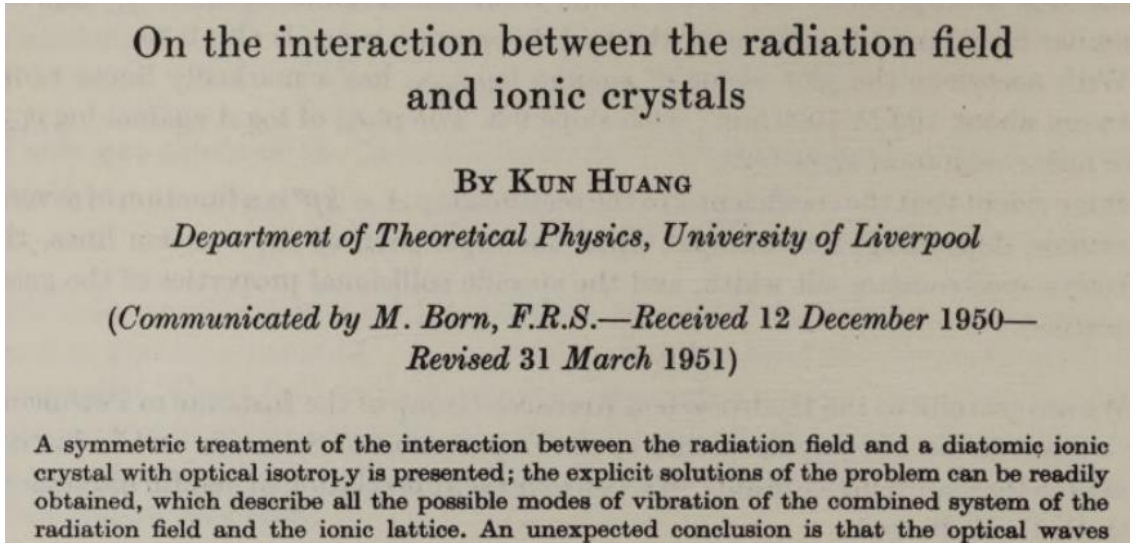
(Received July 16, 1958)

creation operators for the polarization field. The polarization field "particles" analogous to photons will be called "polaritons." (Excitons will be shown to be one kind of polariton in Sec. IV. Optical phonons are another example of polaritons.) The sum is taken over

- 1958，康奈尔大学，物理学博士
- 1958-1969，贝尔实验室，巴黎高师，伯克利，普林斯顿；凝聚态物理；与 D. G. Thomas 发展激子极化激元理论，半导体的光发射与吸收等光与物质相互作用问题；
- 1970-1980，普林斯顿物理系，生物物理；蛋白质动力学校对，血红蛋白协同效应，蛋白内电子隧穿和电子转移理论
- 1980-1997，加州理工，化学与生物学，计算神经系统；Hopfield 网络。
- 1997后，普林斯顿，分子生物学

黄昆





On the interaction between the radiation field and ionic crystals

By Kun Huang

Department of Theoretical Physics, University of Liverpool

(*Communicated by M. Born, F.R.S.—Received* 12 *December* 1950—*Revised* 31 *March* 1951)

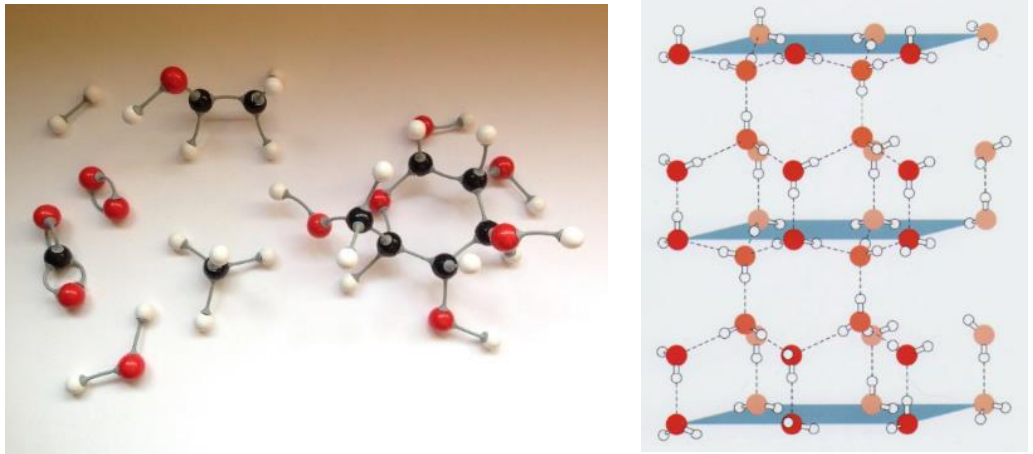
A symmetric treatment of the interaction between the radiation field and a diatomic ionic crystal with optical isotropy is presented; the explicit solutions of the problem can be readily obtained, which describe all the possible modes of vibration of the combined system of the radiation field and the ionic lattice. An unexpected conclusion is that the optical waves

1941年，燕京大学，物理学学士
1941-1945，西南联大
1945-1948，布里斯托尔大学，博士，导师Mott，访问Born
1948-1951，利物浦大学，博士后
1951-1966，北京大学物理系
1977-1983，中科院半导体所
黄散射，黄方程 + 极化激元概念，黄-里斯理论
《晶格动力学理论》
2001国家最高科学技术奖

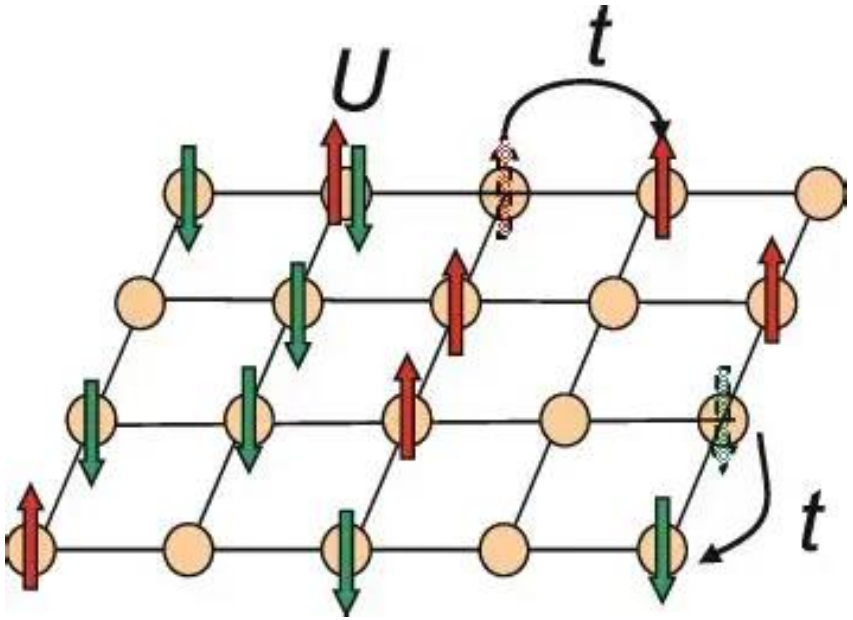
AI与凝聚态物理

本课程：AI for 凝聚态物理

凝聚态物理的前沿

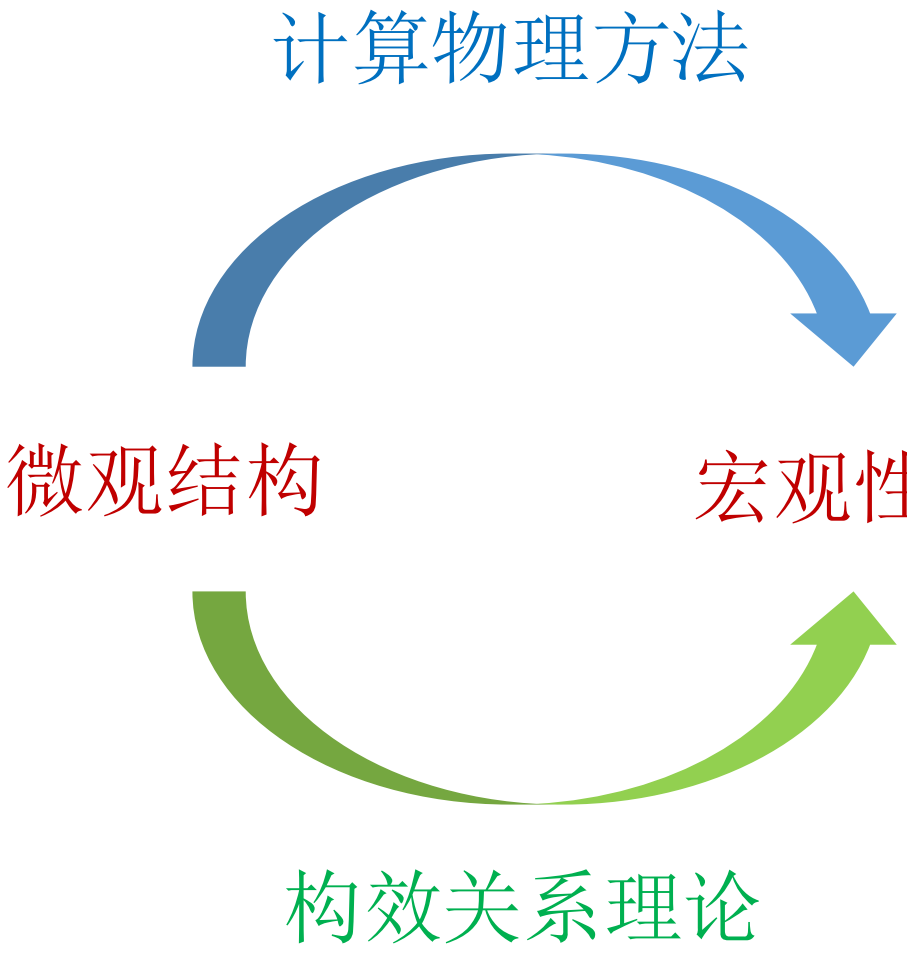


由具体原子组成的真实物质体系



量子多体模型

AI for 凝聚态物理



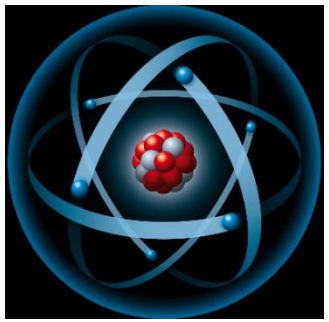
赋能

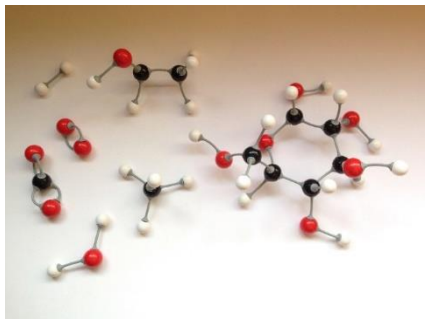
- 大数据
- 监督式学习
- 非监督式学
- 符号回归
- 深度学习
- 生成模型
- 强化学习
- 大语言模型

内容提纲

1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

第一性原理计算



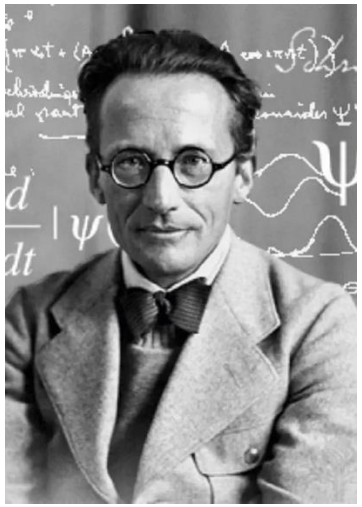


$$i\hbar\frac{\partial\psi}{\partial t}=\hat{H}\psi$$

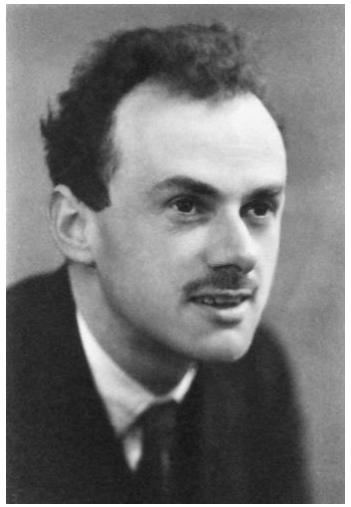
$$i\hbar\frac{\partial\psi}{\partial t}=\hat{H}\psi$$



Werner Heisenberg



Erwin Schrödinger



Paul Dirac

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

第一性原理计算

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N_e}\nabla_i^2 - \frac{1}{2}\sum_{I}^{N_n}\nabla_I^2 - \sum_{i}^{N_e}\sum_{I}^{N_n}\frac{Z_I}{r_{iI}} + \frac{1}{2}\sum_{i\neq j}^{N_e}\frac{1}{r_{ij}} + \frac{1}{2}\sum_{I\neq J}^{N_n}\frac{Z_I Z_J}{r_{IJ}}$$

Born Oppenheimer Approximation

$$\hat{H} = \hat{H}_e(R) + \hat{H}_R(r)$$

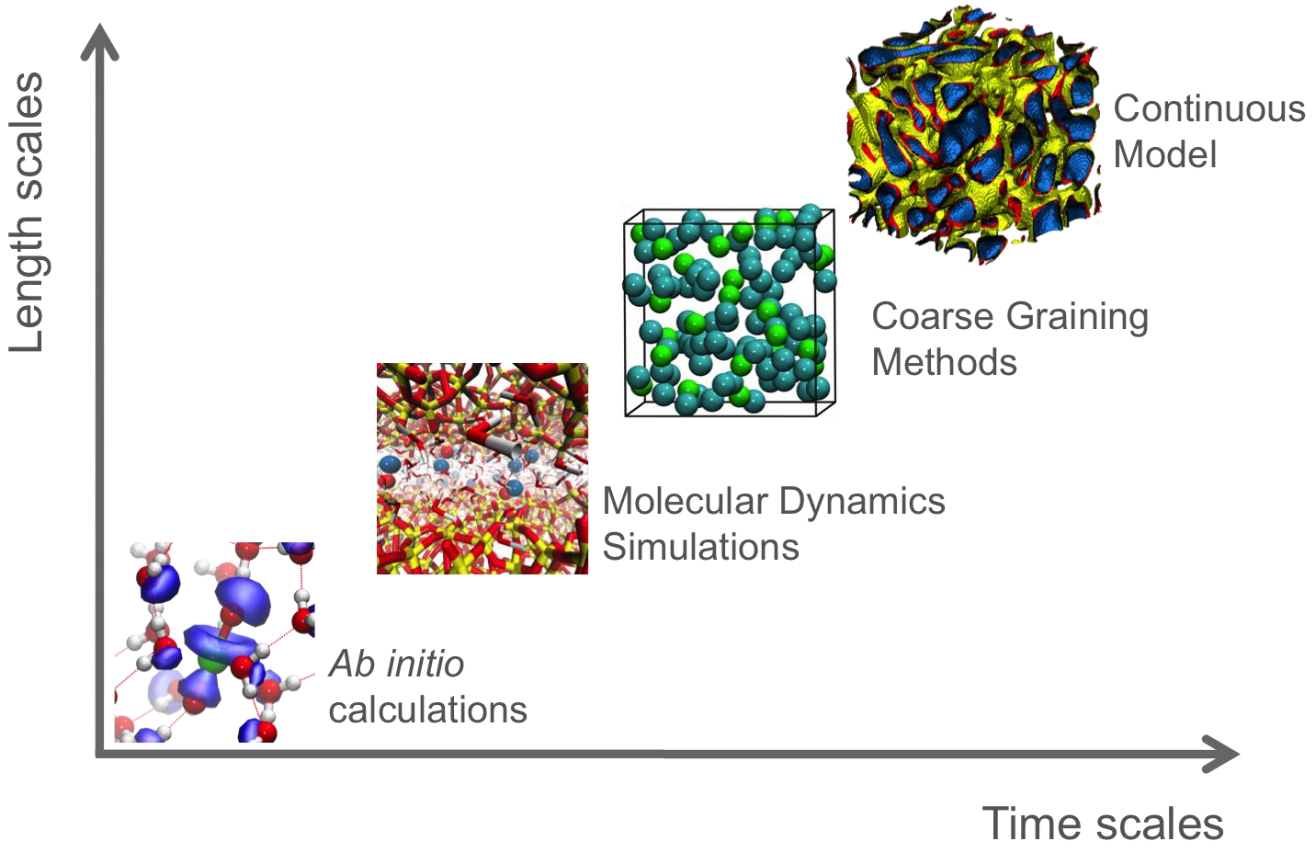
$$\hat{H}_R = -\frac{1}{2}\sum_{I}^{N_n}\nabla_I^2 + V_{PES}(R)$$

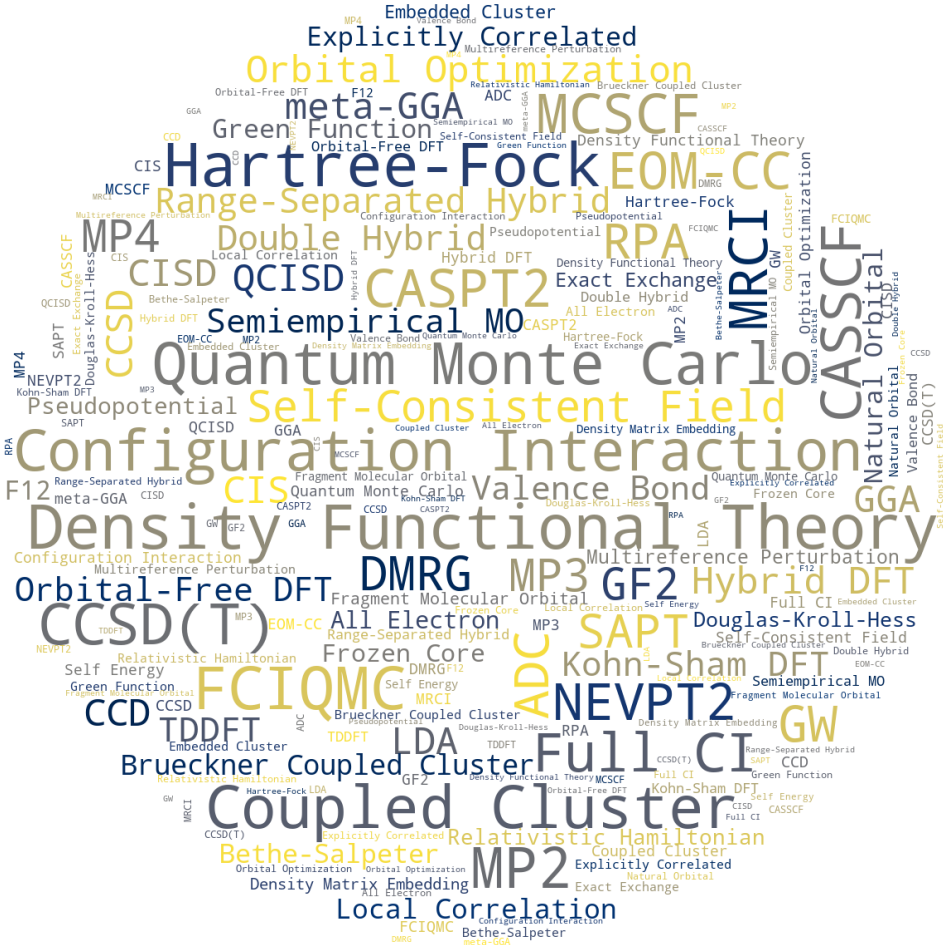
$$\hat{H}_e(R) = -\frac{1}{2}\sum_{i}^{N_e}\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}^{N_e}\frac{1}{r_{ij}} + V_{ext}(r,R)$$

第一性原理计算的核心问题：
电子结构（电子哈密顿量求解）

注：超越玻恩-奥本海默的“全量子化”计算也是凝聚态物理的前沿

第一性原理计算

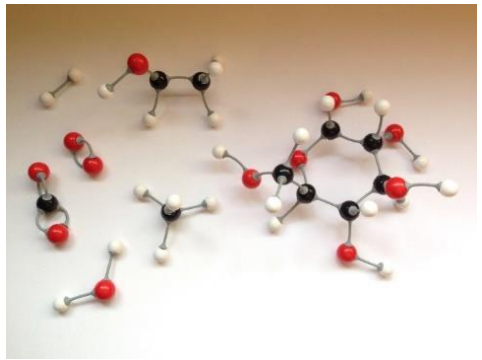

Length scales
Continuous Model
Coarse Graining Methods
Molecular Dynamics Simulations
Ab initio calculations
Time scales


Embedded Cluster
Explicitly Correlated
Orbital Optimization
meta-GGA
MCSCF
Green Function
Density Functional Theory
Hartree-Fock
EOM-CC
Range-Separated Hybrid
Hartree-Fock
MP4
Double Hybrid
RPA
CISD
QCISD
CASPT2
Exact Exchange
MRCI
CASSCF
Semiempirical MO
Quantum Monte Carlo
Natural Orbital
Self-Consistent Field
Pseudopotential
Configuration Interaction
CIS
Valence Bond
GGA
Density Functional Theory
Orbital-Free DFT
DMRG
MP3
GF2
Hybrid DFT
CCSD(T)
All Electron
Douglas-Kroll-Hess
Frozen Core
Kohn-Sham DFT
FCIQMC
ADC
SAPT
NEVPT2
CCD
TDDFT
LDA
GW
Brueckner Coupled Cluster
Full CI
Coupled Cluster
Relativistic Hamiltonian
Bethe-Salpeter
MP2
Density Matrix Embedding
Local Correlation

第一性原理计算

代表性方法	范式转变
波函数方法 1930s	从解析到数值
密度泛函理论 1960s	从波函数到密度
量子蒙特卡洛 1970s	从确定性到随机性

（1）波函数：从解析到数值



$$\hat{H}\psi(r;R) = E\psi(r;R)$$

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N_e}\nabla_i^2 - \sum_{i}^{N_e}\sum_{I}^{N_n}\frac{Z_I}{r_{iI}} + \frac{1}{2}\sum_{i\neq j}^{N_e}\frac{1}{r_{ij}}$$

Hartree-Fock

$$\Psi_{HF} = \frac{1}{\sqrt{N!}}\det\begin{vmatrix} \psi_1(r_1) & \psi_2(r_1) & \cdots & \psi_N(r_1) \\ \psi_1(r_2) & \psi_2(r_2) & \cdots & \psi_N(r_2) \\ \cdots & \cdots & \cdots & \cdots \\ \psi_1(r_N) & \psi_2(r_N) & \cdots & \psi_N(r_N) \end{vmatrix}$$

post Hartree Fock

$$\Psi_{exact} = \sum_{J=1}^{L} c_J \Psi_J$$

微扰法，耦合簇，组态相互作用

（2）密度泛函：从波函数到密度

The Nobel Prize in Chemistry 1998

for his development of the density-functional theory

密度泛函理论

求“电子密度分布”

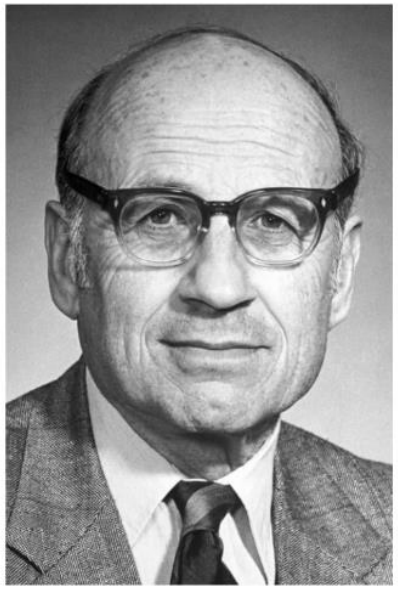


Photo from the Nobel Foundation archive.

Walter Kohn

Prize share: 1/2

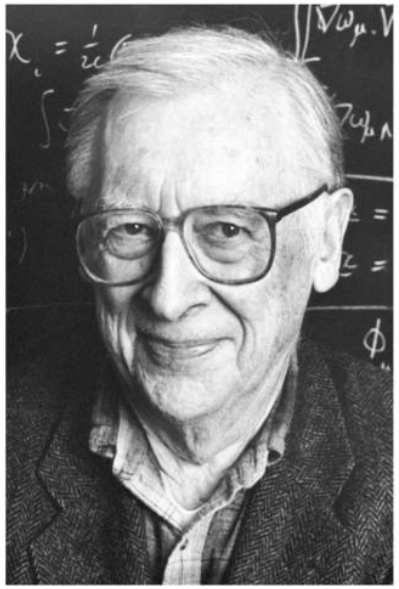


Photo from the Nobel Foundation archive.

John A. Pople

Prize share: 1/2

for his development of computational methods in quantum chemistry

量子化学方法

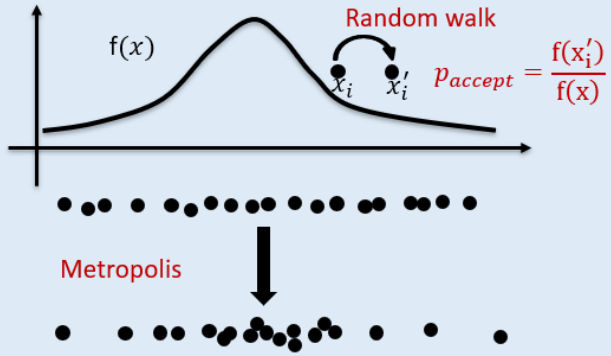
求“薛定谔方程的波函数”

量子蒙特卡洛：从确定性到随机性

Variational MC
VMC

Probability　　local energy E_L

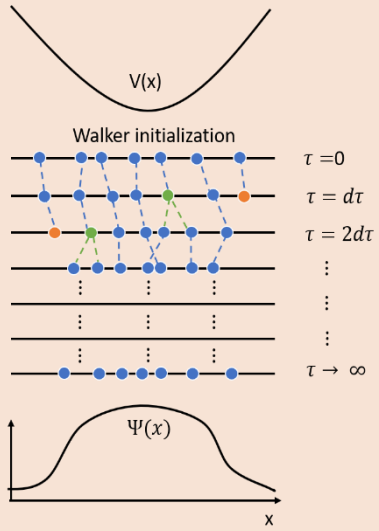
$$E_V = \frac{\int |\Psi_T(R)|^2 \frac{\hat{H}\Psi_T(R)}{\Psi_T(R)} dR}{\int |\Psi_T(R)|^2 dR} \geq E_0$$



Projection based MC
DMC etc.

Imaginary time evolution:
ground state projection

$$\Psi(R,\tau) = \Sigma_i \phi_i(R) e^{-(E_i - E_0)\tau} \xrightarrow{\tau \to +\infty} \phi_0(R)$$



内容提纲

1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

第一性原理计算

代表性方法	范式转变
波函数方法 1930s	从解析到数值
密度泛函理论 1960s	从波函数到密度
量子蒙特卡洛 1970s	从确定性到随机性
机器学习 2017以来	从计算到学习

第一性原理计算

Representation (ansatz) + solution (algorithm)

What mathematic form should be used to represent the wavefunction?

How to determine (the undetermined parameters) of the wavefunction?

第一性原理计算

$$\Psi_{HF/KS} = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_N(x_1) \\ \vdots & \ddots & \vdots \\ \phi_1(x_N) & \cdots & \phi_N(x_N) \end{bmatrix}$$

Self Consistent Field (solving linear equation)
Hartree-Fock
Density Functional Theory

$$\hat{F}\phi_i = \epsilon_i \phi_i$$

$$\Psi_{CI/CC} = \hat{\Omega}\Psi_{\mathrm{HF/KS}}$$

Solving linear/non-linear equation/Monte Carlo

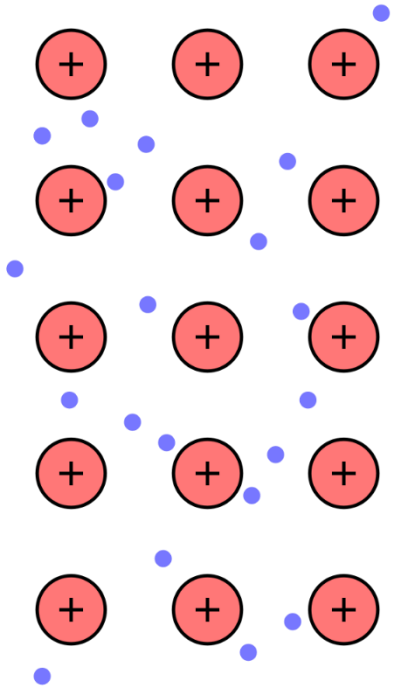
$$\hat{H}\boldsymbol{c} = ES\boldsymbol{c} \qquad f(t_1, t_2, \ldots) = 0$$

$$\Psi_{SJ} = e^{J(x_1,\ldots,x_N)} \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_N(x_1) \\ \vdots & \ddots & \vdots \\ \phi_1(x_N) & \cdots & \phi_N(x_N) \end{bmatrix}$$

Imaginary time evolution
with Monte Carlo

波函数拟设

How to represent a many-body wavefunction of electrons?



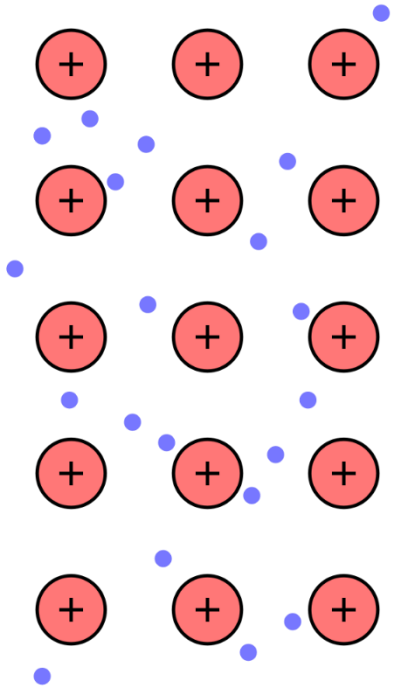
Electron information:
(spin) coordinates
$x_1, \ldots, x_N$

Ansatz

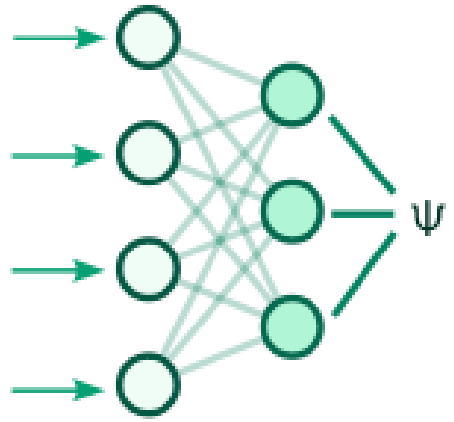
$\Psi(x_1, \ldots, x_N)$

神经网络波函数

Neural Network: a powerful representation architecture



Electron information:
(spin) coordinates
$x_1, \ldots, x_N$

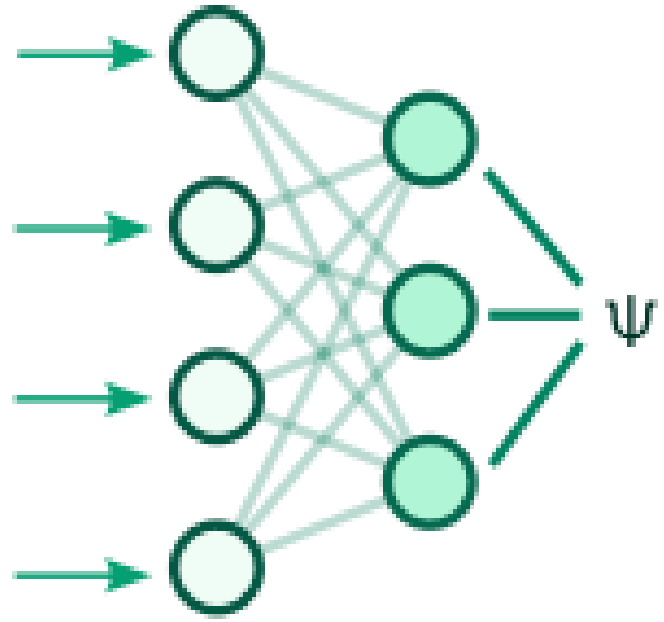


Neural Network
Wavefunction

$\Psi(x_1, \ldots, x_N)$

神经网络波函数

Neural Network: a powerful representation architecture



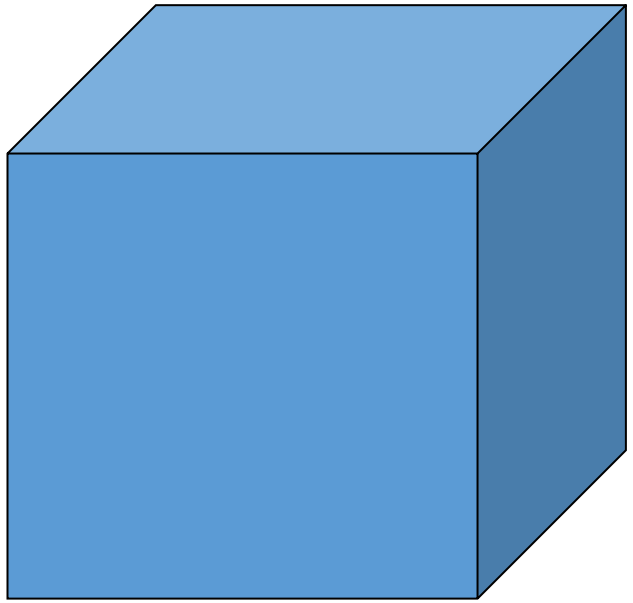
$$\Psi_{HF/KS} = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_N(x_1) \\ \vdots & \ddots & \vdots \\ \phi_1(x_N) & \cdots & \phi_N(x_N) \end{bmatrix}$$

$$\phi_j(\mathbf{x}_i) \rightarrow \phi_j(\mathbf{x}_i, \{\mathbf{x}_{k\neq i}\}; \theta_j)$$

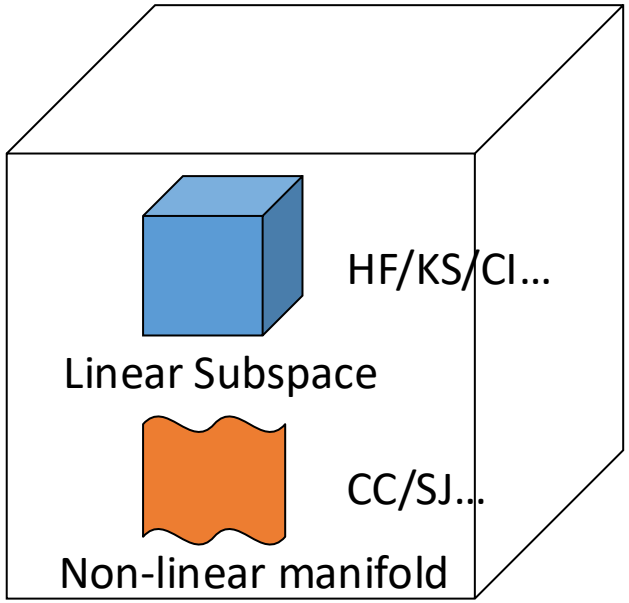
$$\Psi_{NN} = \begin{bmatrix} \phi_1(x_1, \{x_{k\neq 1}\}) & \cdots & \phi_N(x_1, \{x_{k\neq 1}\}) \\ \vdots & \ddots & \vdots \\ \phi_1(x_N, \{x_{k\neq N}\}) & \cdots & \phi_N(x_N, \{x_{k\neq N}\}) \end{bmatrix}$$

神经网络波函数

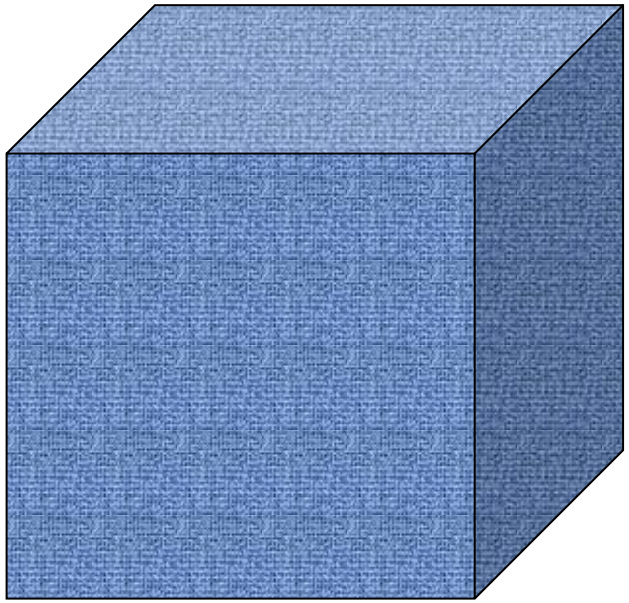
Full Hilbert Space



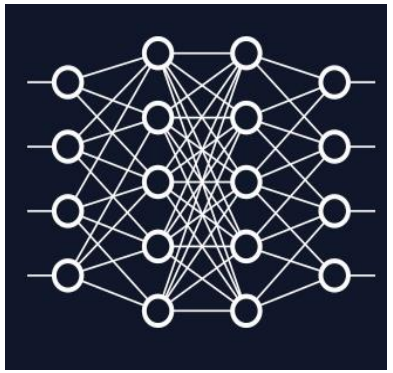
Traditional Wavefunction



Neural Network wavefunction

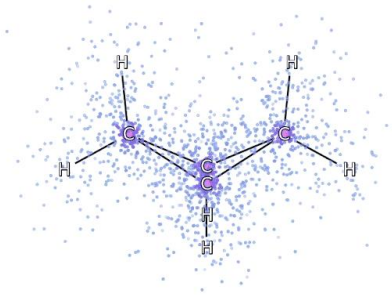


神经网络波函数的训练

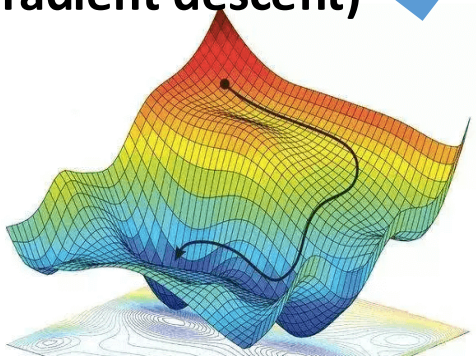


Neural network
wavefunction ansatz

Integration to obtain energy
(Monte Carlo)



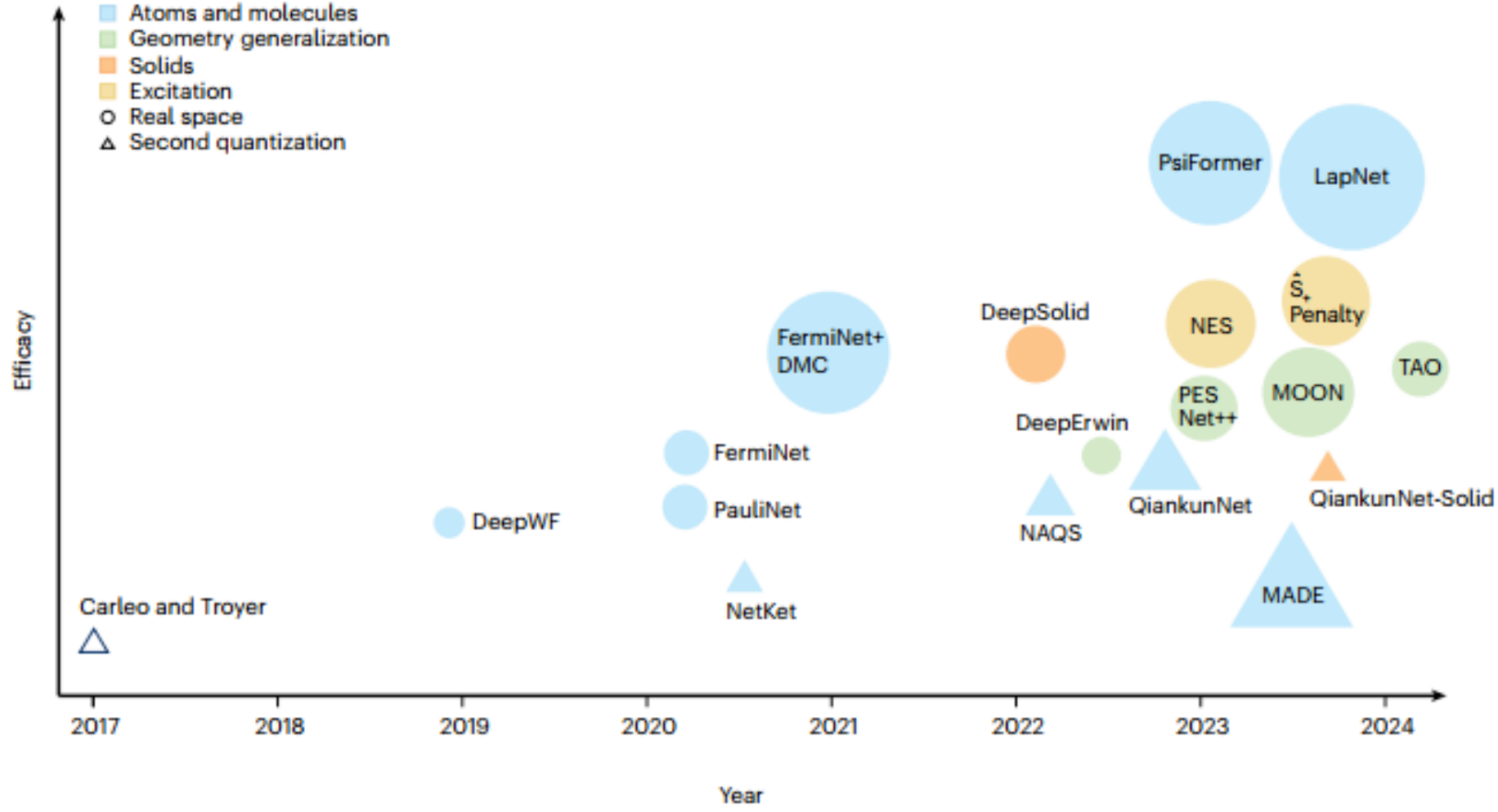
Computing the gradient of wavefunction
parameters to the energy



Wavefunction optimization
(Stochastic gradient descent)

Reinforcement Learning, no external data, energy is the natural loss!

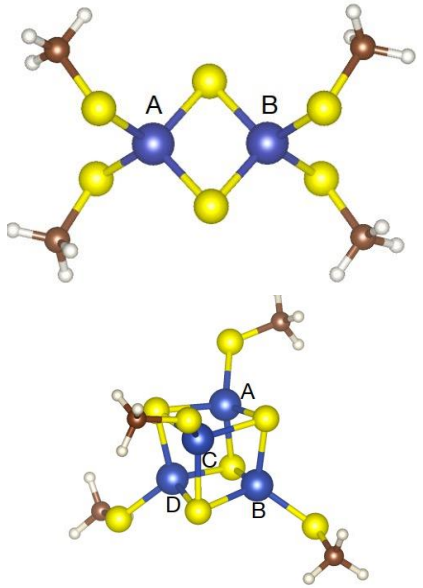
神经网络波函数方法的发展

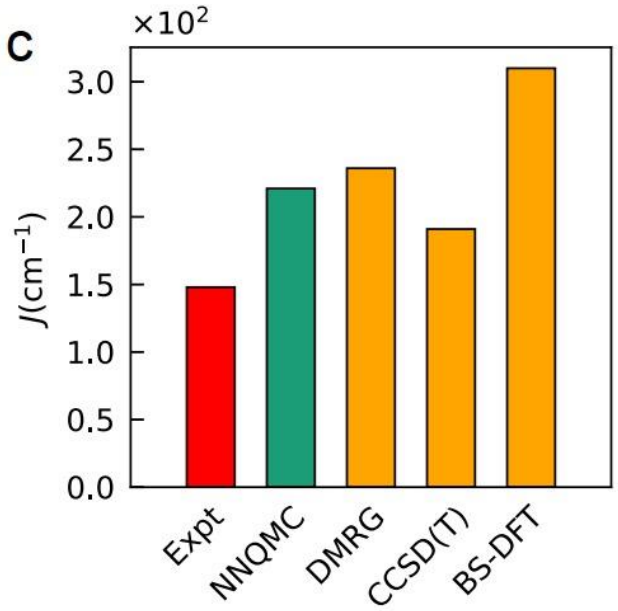


Tang, Chen, ..., Duan, JC, Xu, *Nat. Comput. Sci.* 5, 1133 (2025)

深度学习波函数方法的应用（分子）

Benchmark energy and magnetic coupling of iron-sulfur clusters

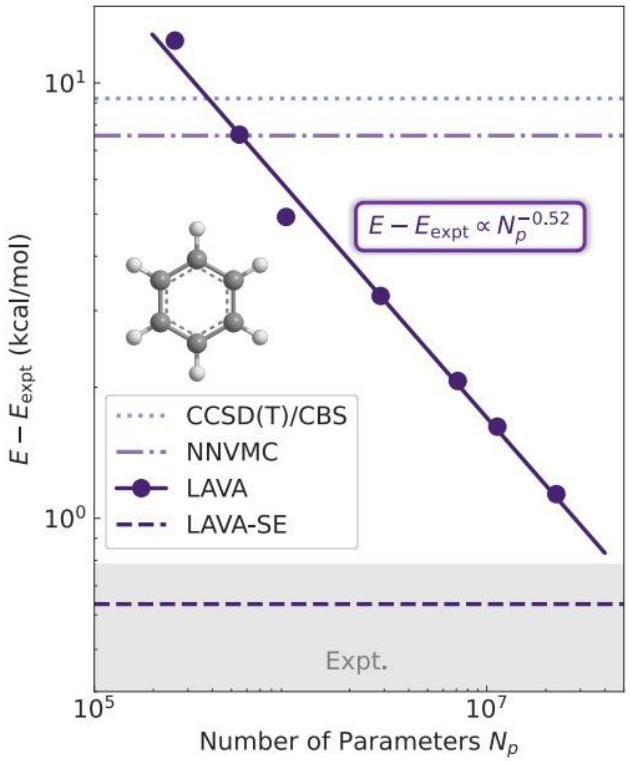




268 electrons (with local Pseudopotential)

Fu et al. arXiv:2505.19909

Approaching neural scaling laws with new optimizer

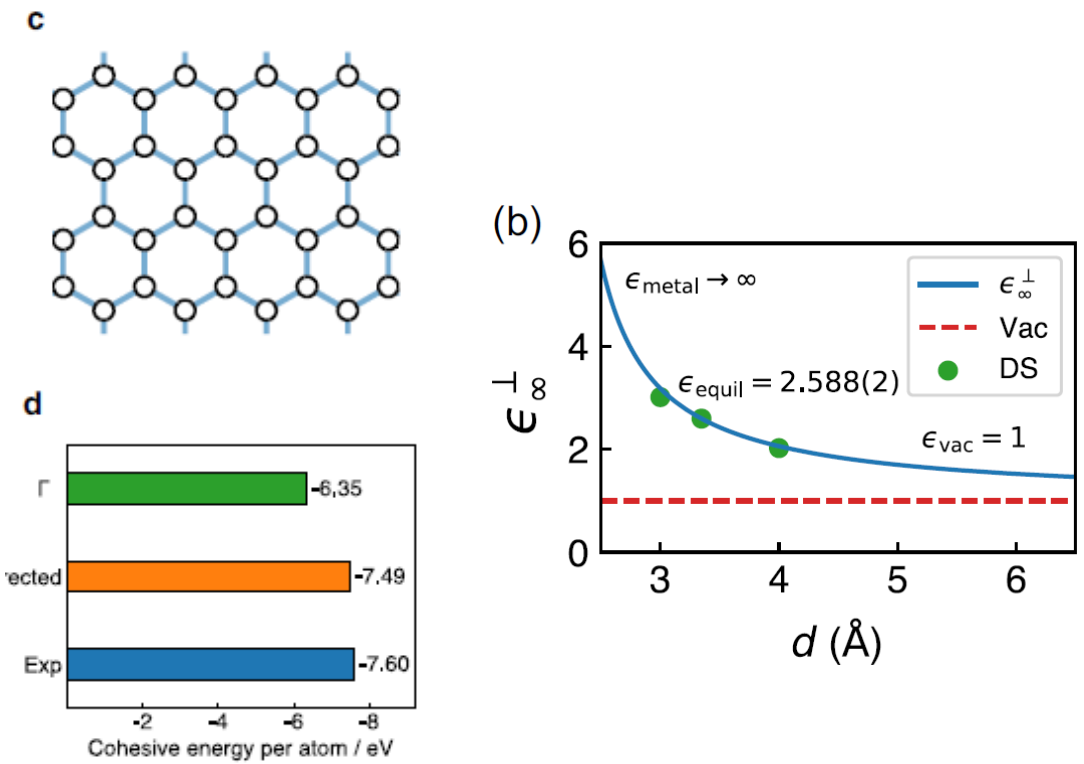


KJ/mol accuracy

Jiang and Wen et al. arXiv:2508.02570

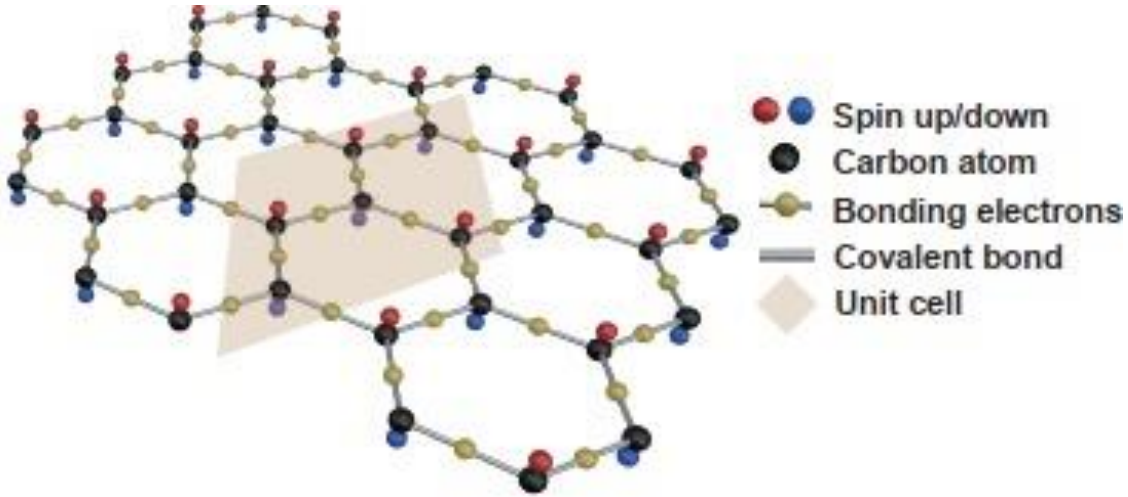
深度学习波函数方法的应用（固体材料）

Benchmark cohesive energy and electric polarization



Li, Li, JC, ***Nat. Commun.*** 13, 7895 (2022)
Li, Qian, JC, ***Phys. Rev. Lett.*** *132, 176401* (2024)

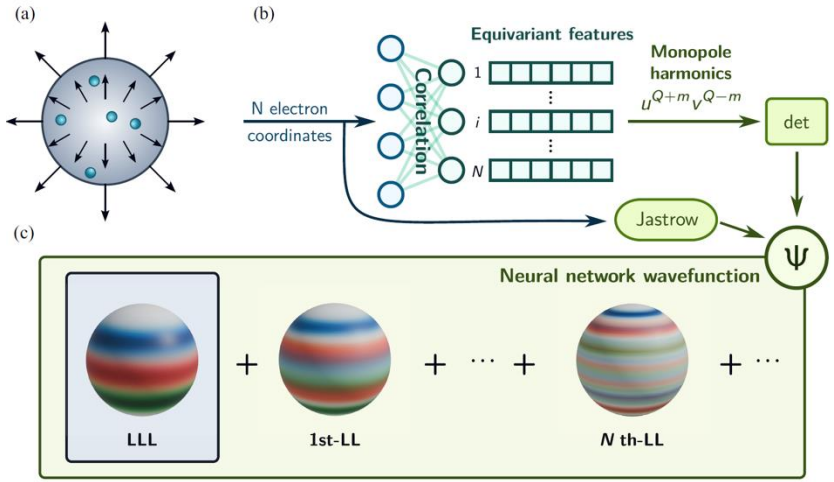
Valence bond structure of graphene: a particle view of many-body electronic structure with neural network wavefunction.



Wang, Fu, Li, Ren, JC, arXiv 2508.13751 (2025)

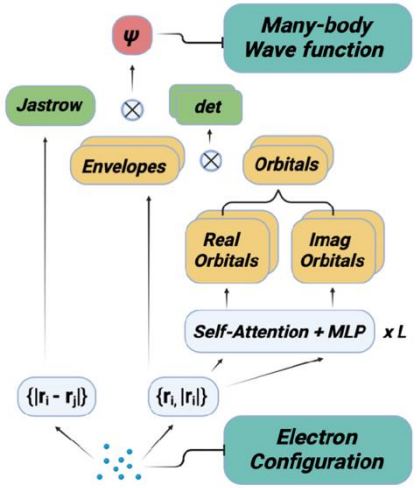
深度学习波函数方法的应用（量子霍尔效应）

球面几何



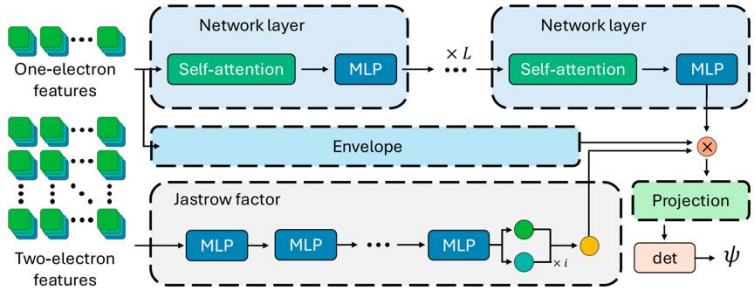
Qian, Zhao, Zhang, Xiang, Li, JC, PRL 2025, 134 176503

圆盘几何



Teng, Dai, Fu, PRB 2025, 111, 205117

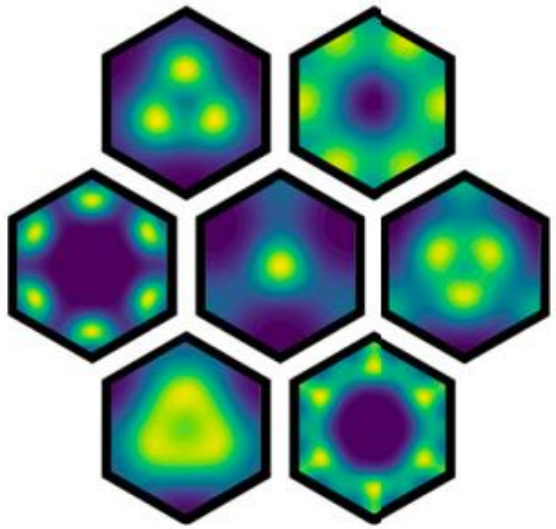
环面几何



Fadon, …, Foulkes, arXiv:2512.15872

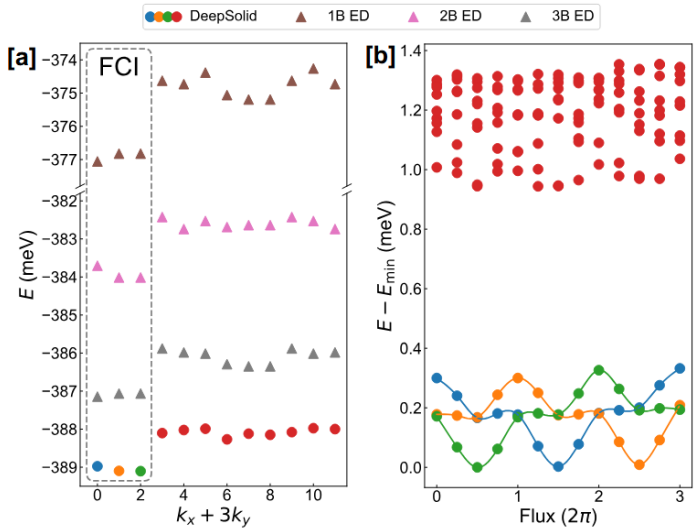
深度学习波函数方法的应用（莫尔体系拓扑关联态）

广义韦格纳晶体态



Li, , Qian, Ren, Xu, JC, Commun Phys 2025, 8, 364

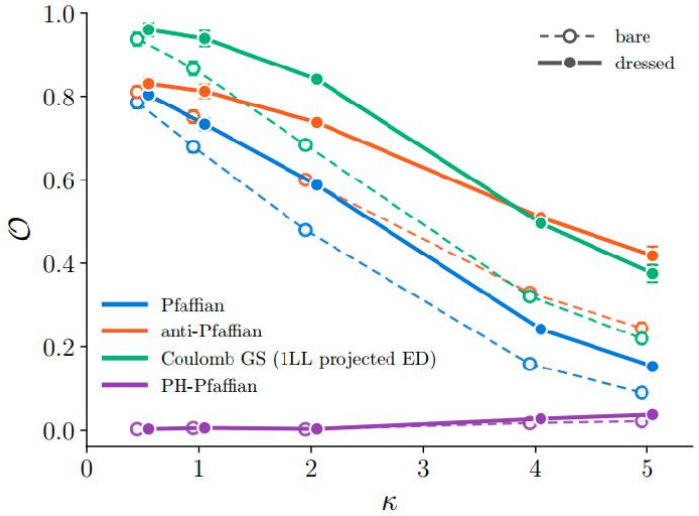
分数陈绝缘体态



Li,…, Wu, JC, Ren, arXiv:2503.11756

Luo, Zaklama, Fu, arXiv:2503.13585

非阿贝尔量子霍尔态



Graham, Abouelkoumsan, Fu, arXiv:2610.18967

内容提纲

1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

密度泛函模型和理论的发展

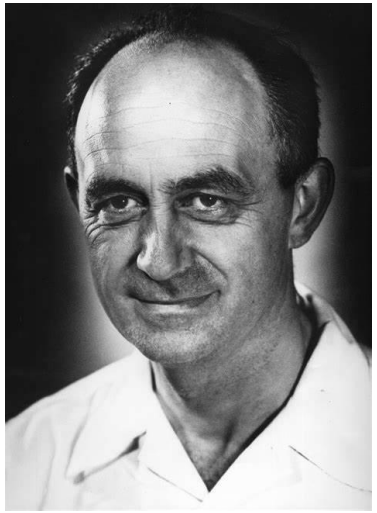
阶段	时间	主要发展	代表人物或工作
前驱探索	1927-1964	经验型能量密度泛函模型	Thomas, Fermi, von Weizsäcker, Wigner, Dirac, Slater
理论奠基	1964-1970s	建立严格的理论框架	Kohn, Hohenberg, Sham
通用泛函	1980-1990s	局域密度近似LDA和广义梯度近似GGA的通用参数化，	Perdew, Zunger, Becke, Lee, Yang, Parr
高阶泛函	1990s-2010	meta-GGA, hybrid functional, dispersion correction, explicit correlation	Perdew, Becke, Sun, Truhler, Head-Gordon, Grimme, Michaelides, Tkachenko, Schaffler
机器学习泛函	2010至今	复杂网络参数化泛函加数据驱动训练	DM21, Skala

经验型能量密度泛函模型

Thomas Fermi model, 1927



Llewellyn Thomas

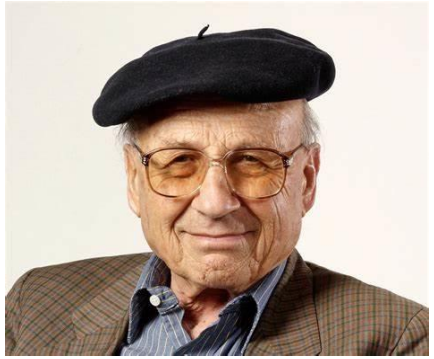


Enrico Fermi

体系能量是电子密度的泛函

$$E_{TF}[n(r)] = \frac{3}{10}(3\pi^2)^{2/3}\int n^{5/3}(r)\,dr - Z\int \frac{n(r)}{r}dr + \frac{1}{2}\iint \frac{n(r_1)n(r_2)}{r_{12}}dr_1 dr_2$$

密度泛函理论



Walter Kohn



Pierre Hohenberg



Lu Jeu Sham

Hohenberg Kohn Theorems, 1964

1. 薛定谔方程解的基态能量是电子密度的唯一泛函；

$$E = E[\rho(r)]$$

2. 最小化能量泛函的电子密度是薛定谔方程的解对应的真实的电子密度；

$$\rho_{true}(r) = \arg\min E[\rho(r)]$$

Kohn Sham equation, 1965

$$\left\{-\frac{\hbar^2}{2m}\nabla^2 + V_{ext} + V_H + V_{xc}\right\}\psi_i(r) = \varepsilon_i\psi_i(r)$$

$\psi_i(r)$: Kohn-Sham orbital
ε_i : orbital energy

V_{xc} : exchange correlation potential

$$E[\rho] = T_s[\rho] + E_{\text{ne}}[\rho] + E_{\text{Hartree}}[\rho] + E_{\text{xc}}[\rho].$$

Walter Kohn

1998 Nobel prize in Chemistry "for his development of the density functional theory"

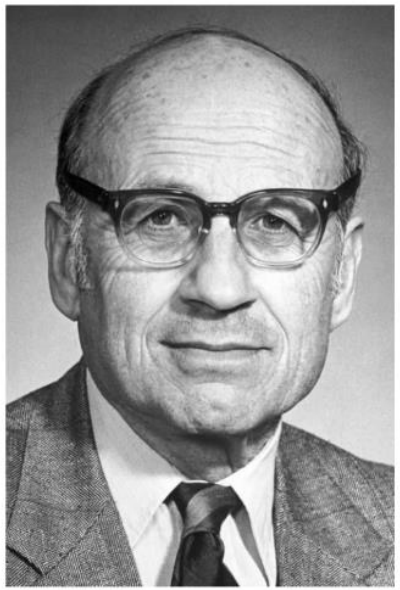


Photo from the Nobel Foundation archive.

Walter Kohn

Prize share: 1/2

- University of Toronto: BSc in Mathematics and Physics (1945), MA in Applied Mathematics (1946)
- Harvard University: PhD in Nuclear Physics (1948), supervisor: Julian Schwinger (Nobel Laureate in Physics, 1965)
- Postdoctoral research: Niels Bohr Institute, Copenhagen
- 1950: Professor of Physics, Carnegie Institute of Technology (now Carnegie Mellon University)
- 1960s: Professor, University of California, San Diego
- 1979: Professor of Physics, University of California, Santa Barbara (UCSB)

Other major scientific contributions :

- KKR method (Korringa–Kohn–Rostoker, 1954) (solid state physics)
- Kohn anomaly (1959) (solid state physics)
- Effective-mass theory (semiconductor physics)
- Hulthen–Kohn variational principle (quantum mechanics)

通用泛函的参数化

Perdew-Zunger (PZ) LDA, 1981

$$E_{xc}^{LDA} = E_{xc}^{LDA}[n(r)] = \int n(r)\varepsilon_{xc}(n)\,\mathrm{d}r$$

$$\varepsilon_{xc}(n) = \varepsilon_x(n) + \varepsilon_c(n)$$

The exchange part is given by Dirac.

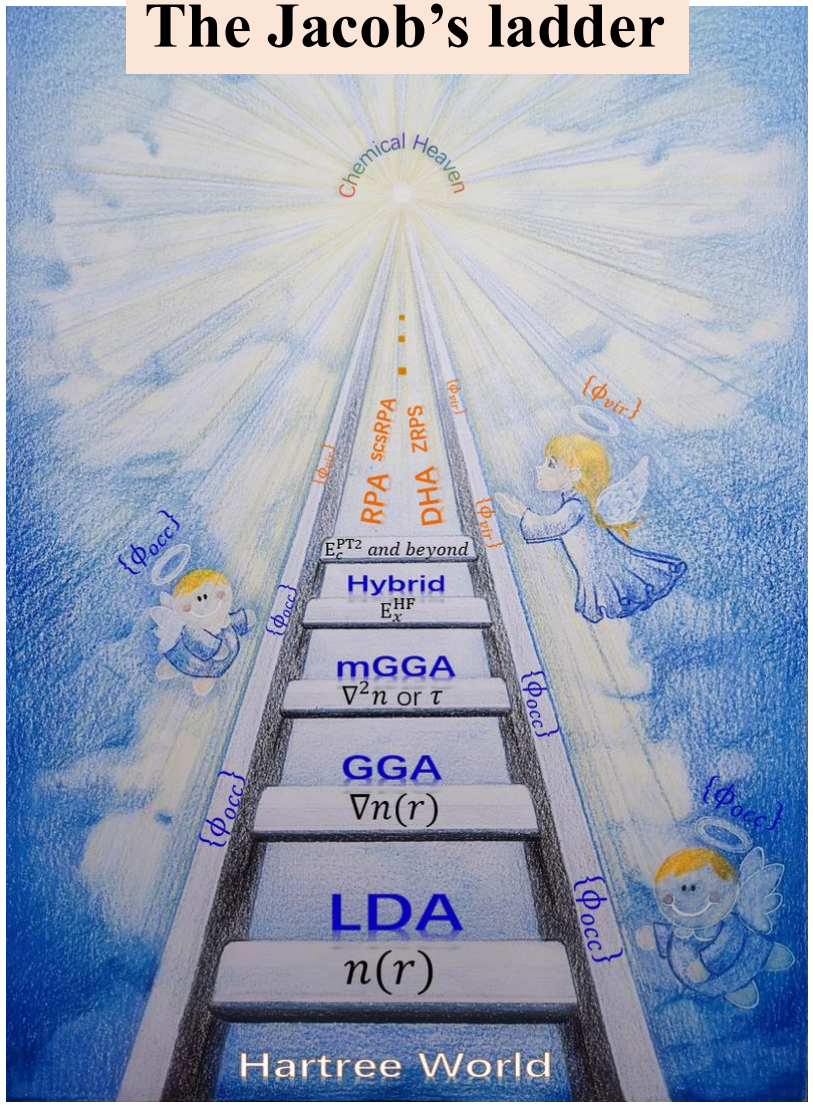
$$\varepsilon_x(n) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} n(r)^{1/3}$$

Parameters are fitted against Ceperley-Alder 1980 QMC data on electron gas

$$\varepsilon_c(r_s) = \begin{cases} -0.1423/(1 + 1.0529\sqrt{r_s} + 0.3334 r_s), & r_s \geq 1 \\ -0.0480 + 0.03114 \ln r_s - 0.0116 r_s + 0.0020 r_s \ln r_s, & r_s < 1 \end{cases}$$

深度学习密度泛函（交换关联泛函）

The Jacob's ladder



WIREs Computational Molecular Science 2020, 11 (1)

如何得到严格的交换关联泛函？

$$E_{xc}[\rho(r), \psi_i(r), \ldots] \qquad V_{xc} = \frac{\delta E_{XC}}{\delta n(r)}$$

理论设计

- 考虑理论的约束
- 特定情况下满足的解析形式
- 特殊体系的严格解

数据拟合

- 实际体系的更精确的计算结果

（监督式）机器学习：交换关联泛函的拟合

深度学习密度泛函（交换关联泛函）

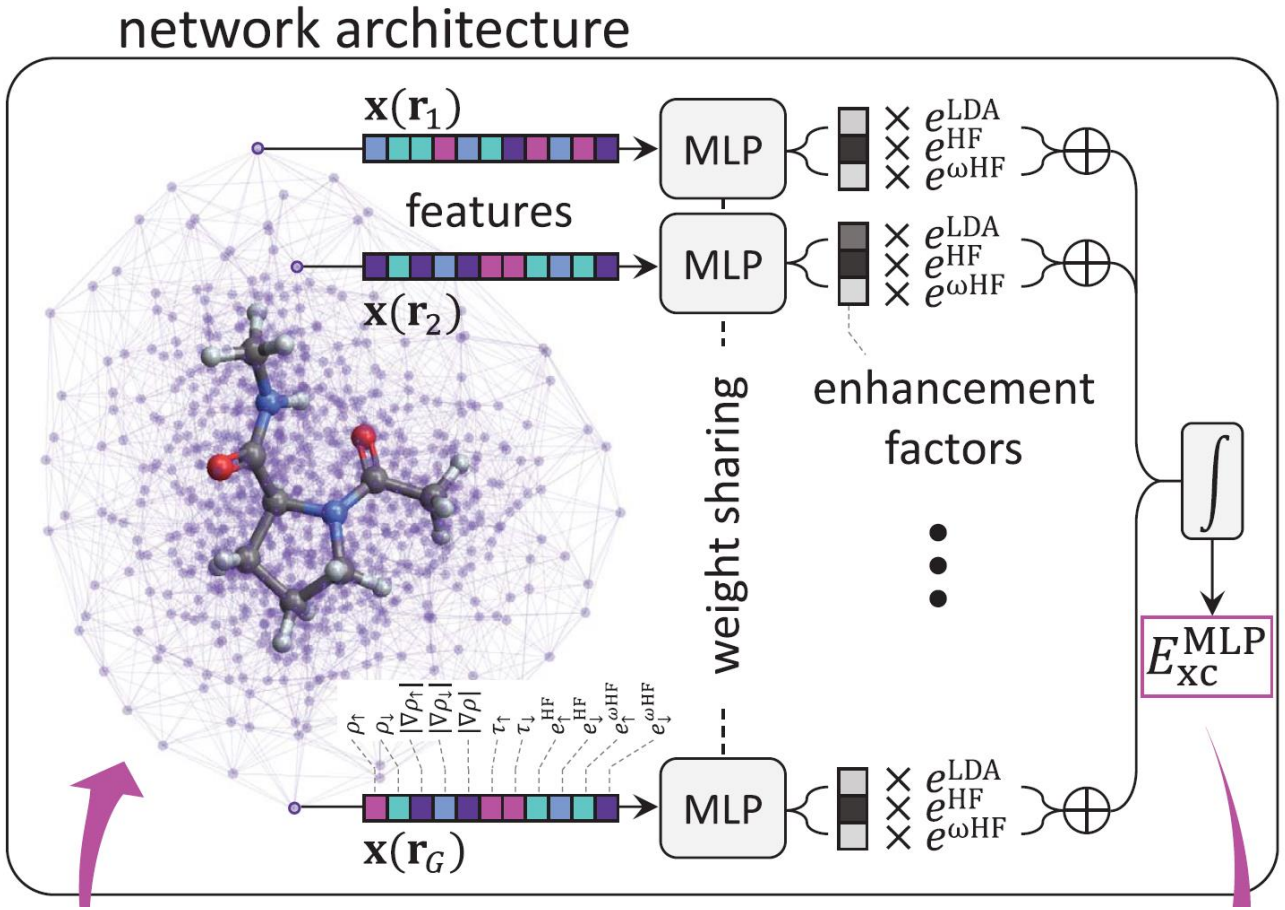
DeepMind: DM21

多层感知机

杂化泛函形式

$$E_{\mathrm{xc}}^{\mathrm{DM21}} = E_{\mathrm{xc}}^{\mathrm{MLP}} + E_{\mathrm{D3(BJ)}}.$$

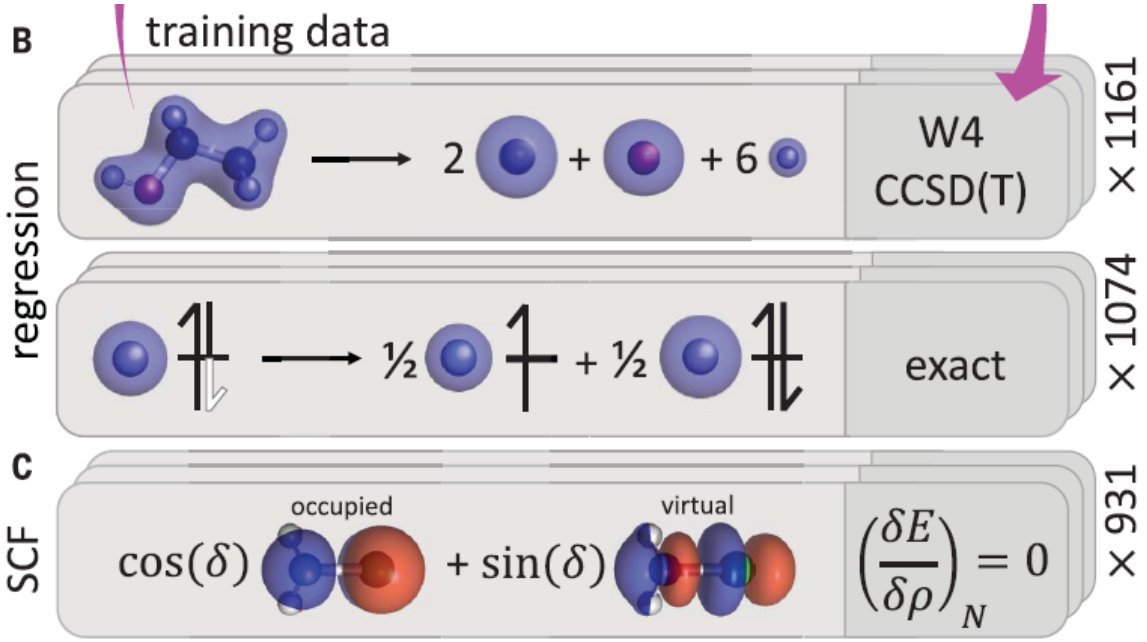
$$E_{\mathrm{xc}}^{\mathrm{MLP}}[\rho] = \int \mathbf{f}_\theta(\mathbf{x}(\mathbf{r})) \cdot \begin{bmatrix} e_{\mathrm{x}}^{\mathrm{LDA}}(\mathbf{r}) \\ e^{\mathrm{HF}}(\mathbf{r}) \\ e^{\omega\mathrm{HF}}(\mathbf{r}) \end{bmatrix} \mathrm{d}^3\mathbf{r}.$$



Kirkpatrick et al., Science 374, 1385–1389 (2021)

深度学习密度泛函（交换关联泛函）

DeepMind: DM21



为训练该泛函，优化目标由两项损失函数相加构成：一项是用于学习交换关联能本身的回归损失；另一项为梯度正则项，用以保证训练完成后，该泛函的泛函导数可用于自洽场（SCF）计算。

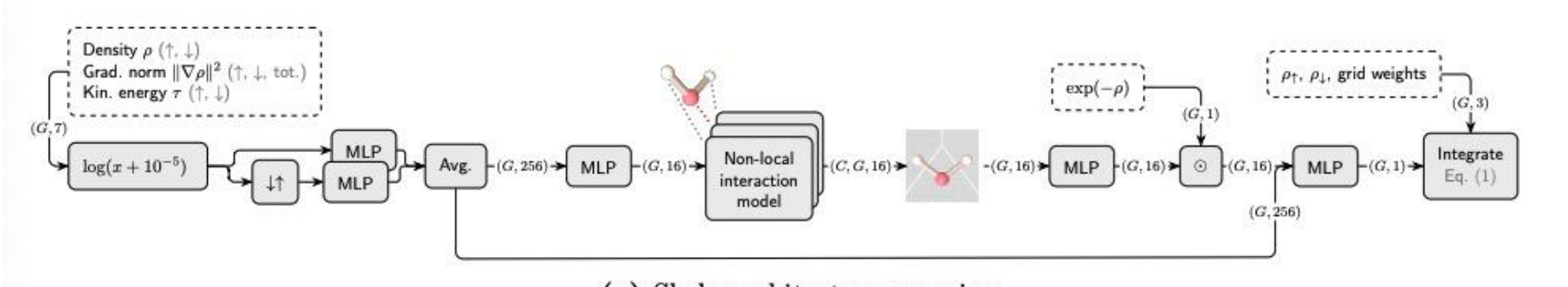
$$\mathcal{L} = \mathbb{E}_r\left[\left(\Delta E_{\mathrm{xc},r}^{\mathrm{DM21}} - \Delta E_{\mathrm{xc},r}^{*}\right)^2\right] + \lambda \mathbb{E}_s\left[\delta E_{\mathrm{SCF},s}^2\right].$$

Kirkpatrick et al., Science 374, 1385–1389 (2021)

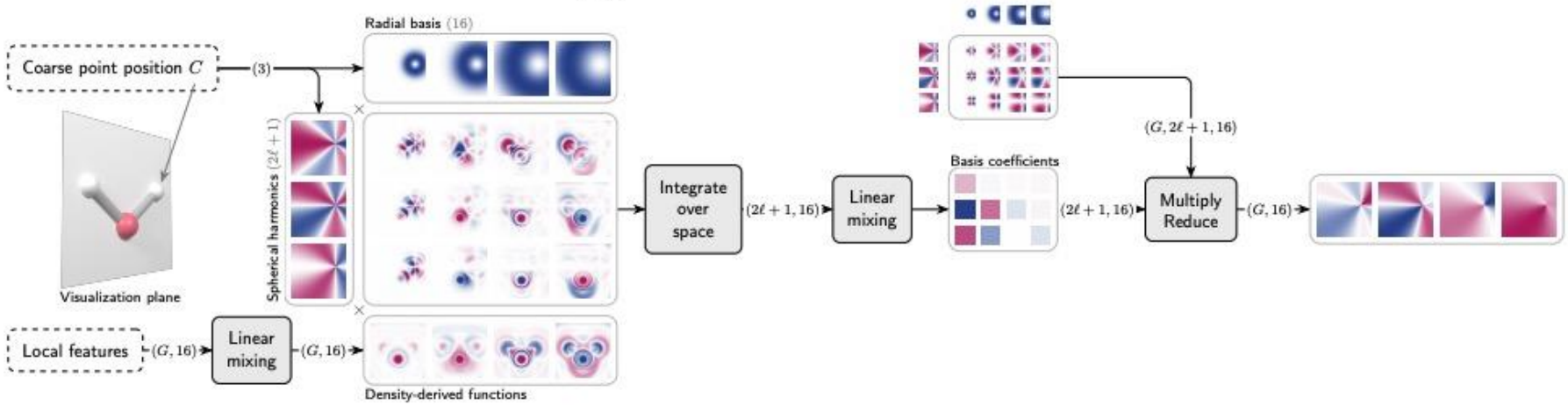
深度学习密度泛函（交换关联泛函）

Microsoft: Skala

兼具杂化泛函级别的精度与半局域泛函的计算开销



(a) Skala architecture overview

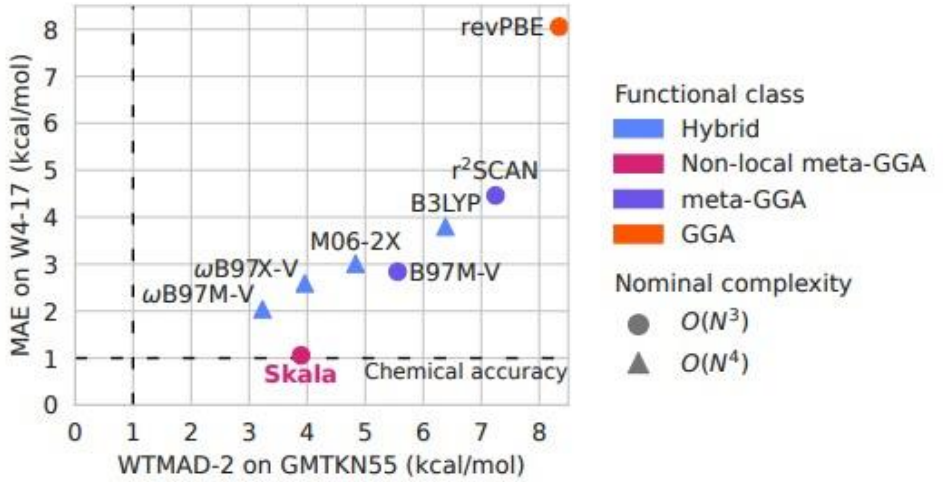


(b) Non-local interaction model

Luise et al., arXiv:2506.14665 (2025)

深度学习密度泛函（交换关联泛函）

Microsoft: Skala



(a)

	W4-17 (full) [MAE]	W4-17 (single ref.) [MAE]	GMTKN55 [WTMAD-2]
revPBE	8.05	7.22	8.34
r²SCAN	4.46	3.84	7.25
B97M-V	2.84	2.52	5.56
B3LYP	3.80	3.73	6.38
M06-2X	3.01	2.39	4.83
ωB97X-V	2.59	2.14	3.96
ωB97M-V	2.04	1.66	3.23
Skala	1.06	0.85	3.89

(b)

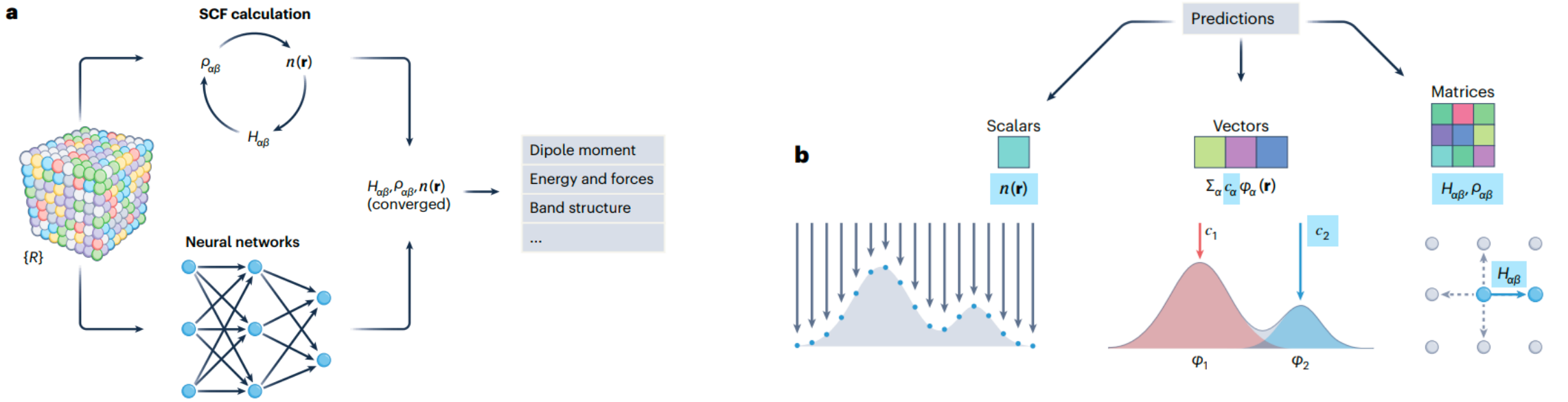
Figure 3: (a): The plot's horizontal axis shows weighted total mean absolute deviation (WTMAD-2) on the GMTKN55[14] test set for general main group thermochemistry, kinetics and non-covalent interactions. The vertical axis shows mean absolute error on the diverse atomization energies test set W4-17[66]. Skala performs similarly to the best-performing hybrid functionals, and reaches near chemical accuracy (1 kcal/mol) on W4-17. (b): Shows the precise errors (in kcal/mol) on W4-17 and GMTKN55, corresponding to the numbers in the plot. For W4-17, the table shows both the MAE on the full set (shown in the plot) as well as on the set of 183 single-reference structures with %TAE[(T)] < 10%.[72] All functionals, including Skala, were evaluated with a D3(BJ) correction, except for those with the VV10[53] correction, indicated with "-V".

Luise et al., arXiv:2506.14665 (2025)

深度学习密度泛函（交换关联泛函）

	DM21	Skala
泛函层级	杂化泛函	Meta-GGA
输入特征	包含密度和 KS 轨道衍生信息	密度、密度梯度、动能密度
网络结构	逐网格点独立 MLP	显式的跨网格消息传递模块
物理约束和训练集	分数电荷、分数自旋数据集上训练，强制满足能量分段线性条件	CCSD (T) 高精度分子能量数据集
训练理念	强调严格约束	优先保证SCF稳定和计算效率
适用范围	键解离、电荷转移体系	主族分子

深度学习密度泛函（哈密顿量）

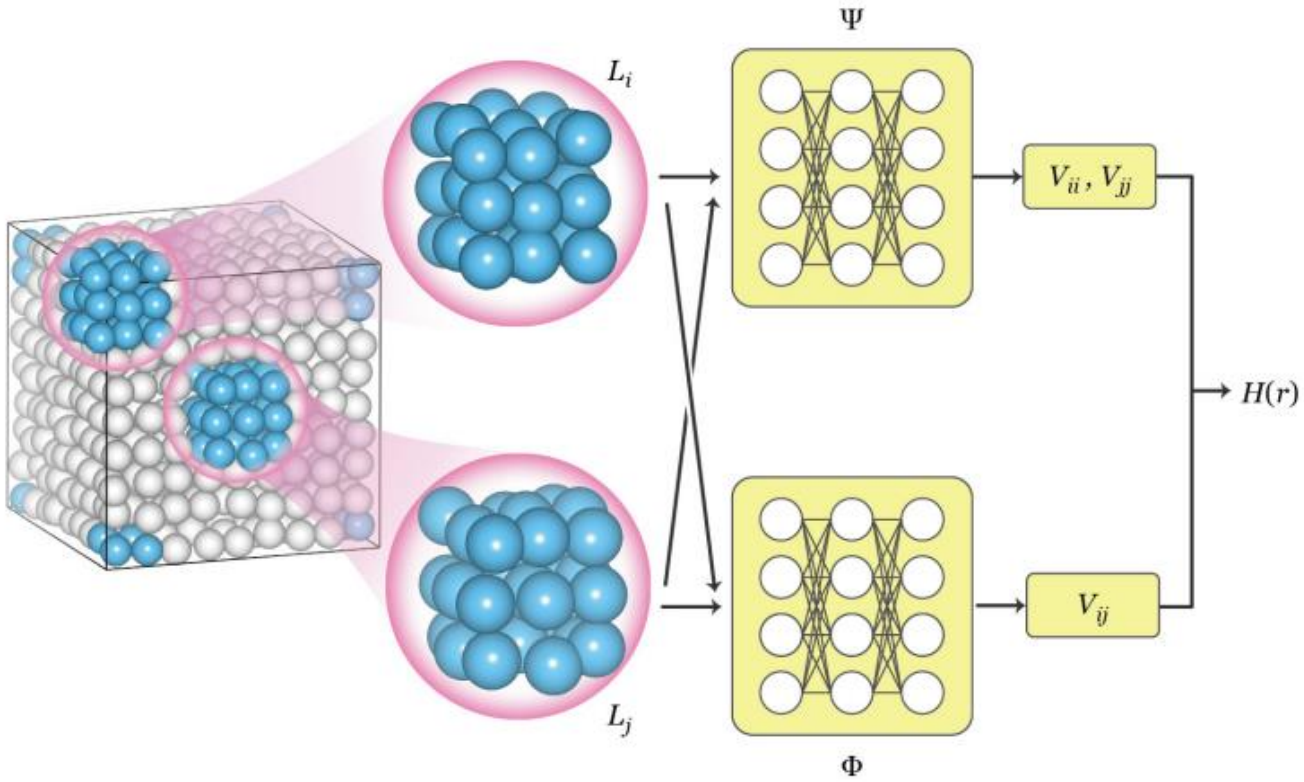


Tang, Chen, ..., Duan, JC, Xu, *Nat. Comput. Sci.* 5, 1133 (2025)

深度学习密度泛函（哈密顿量）

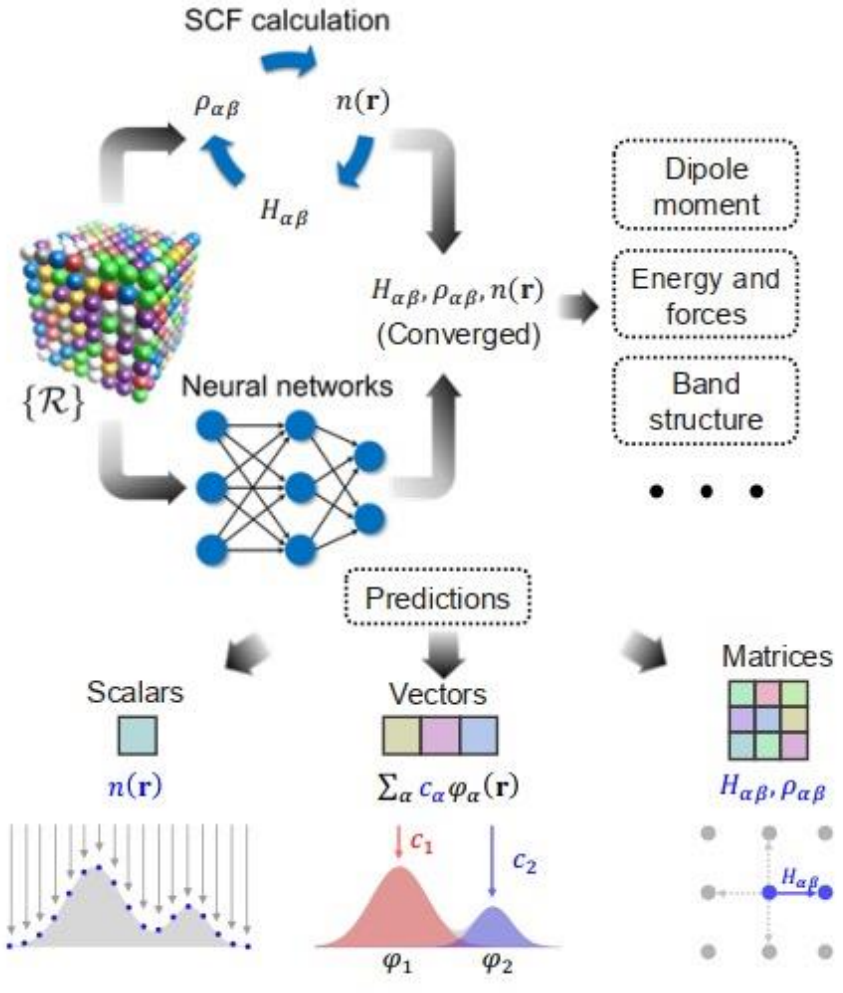
紧束缚（近似）哈密顿量

$$H(\mathbf{r}) = \sum_{i \neq j} V_{ij}(\mathbf{r}) c_i^{\dagger} c_j + \sum_{i} V_{ii}(\mathbf{r}) c_i^{\dagger} c_i + \text{H.c.},$$

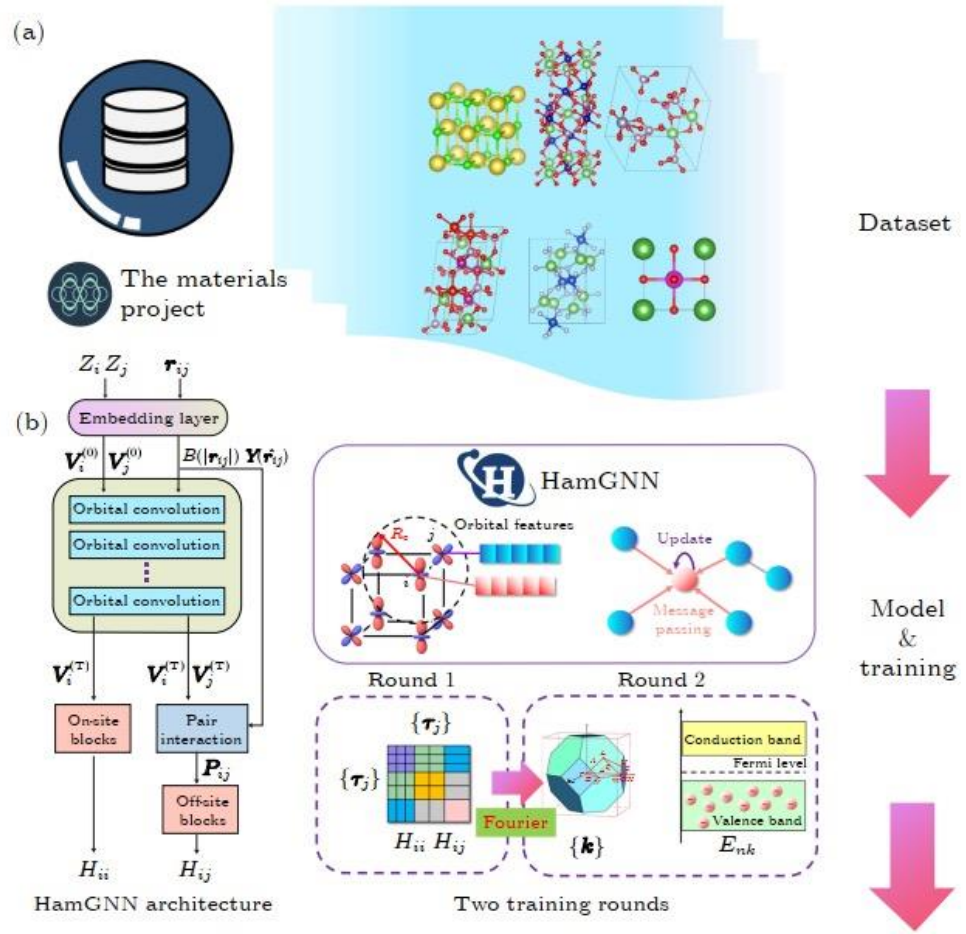


冯济、顾强强等 Science Bulletin 67, 29-37 (2022)

深度学习密度泛函（哈密顿量）

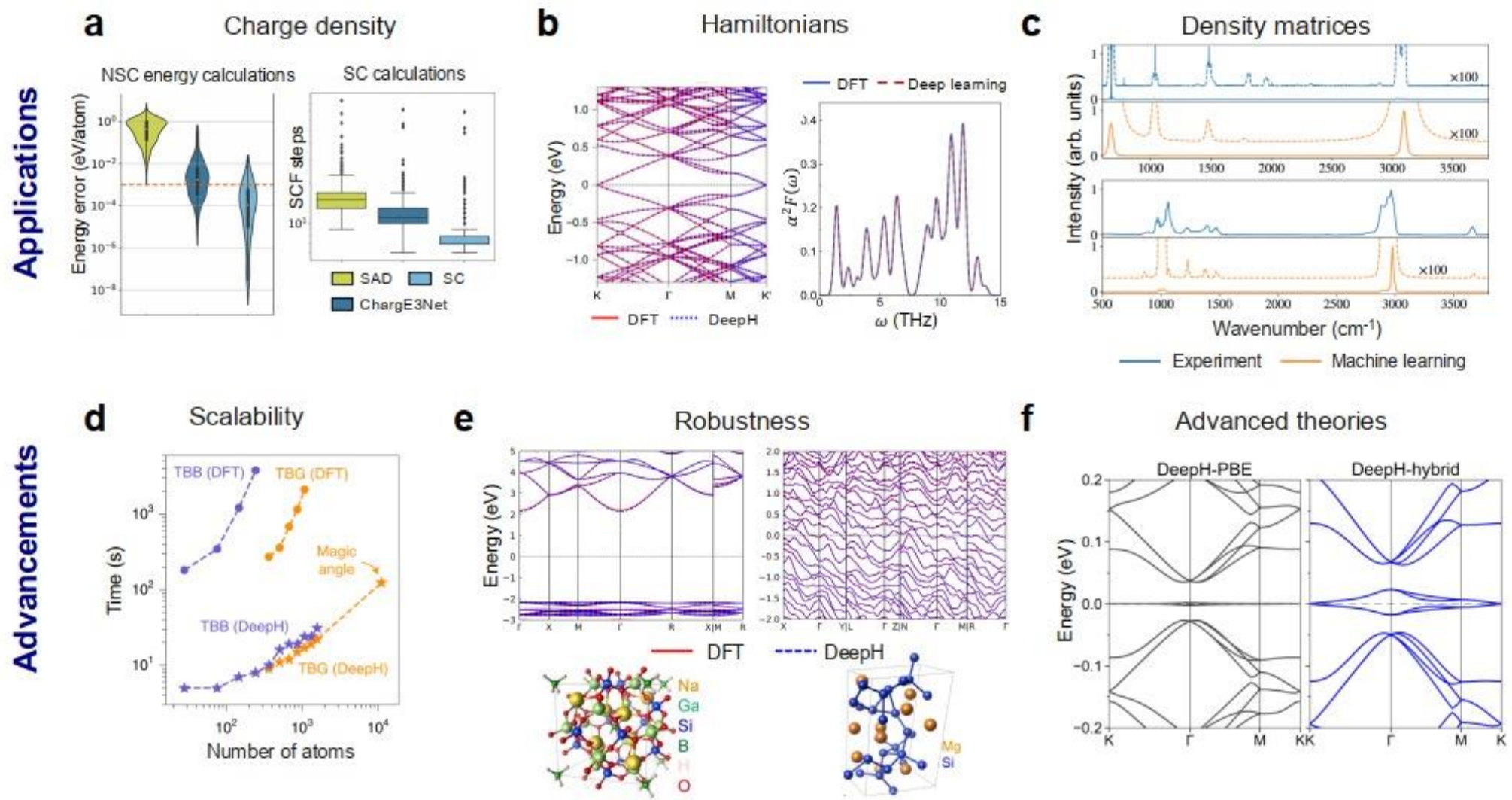


徐勇、段文晖等Nat. Comput. Sci. 2, 367–377 (2022)



向红军、龚新高等Chin. Phys. Lett. 41, 077103 (2024)

深度学习密度泛函（哈密顿量）



小结（1）：AI for 第一性原理计算

1：精度问题

- 深度学习波函数（非监督/强化学习）
- 深度学习密度泛函（交换关联泛函）（监督式学习）

2：效率问题

- 深度学习密度泛函（哈密顿量）（监督式学习）

内容提纲

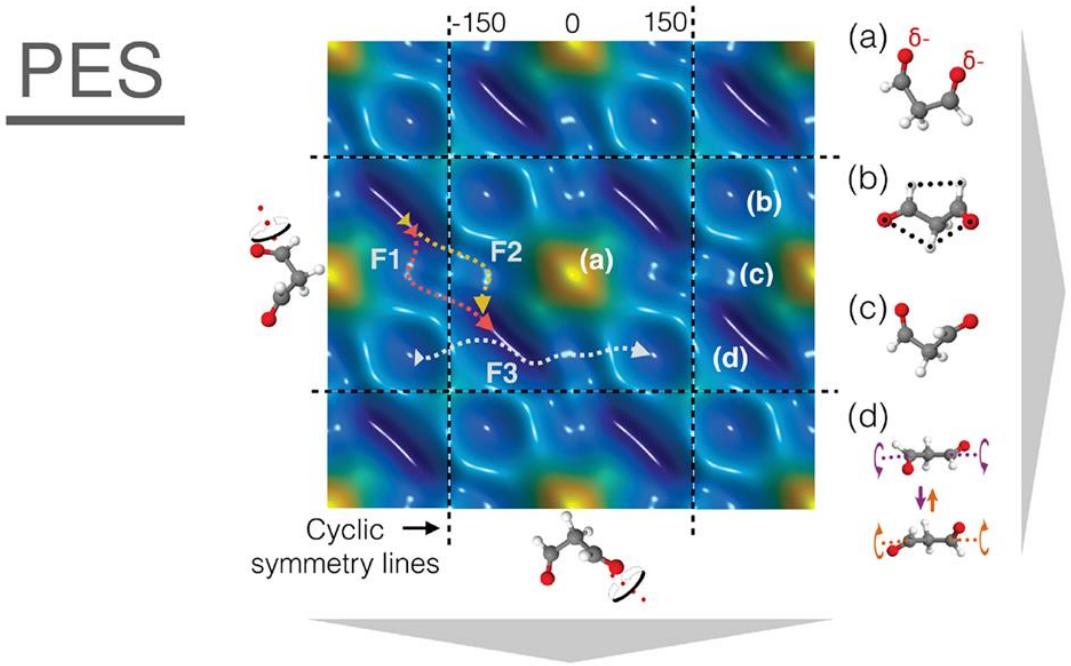
1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

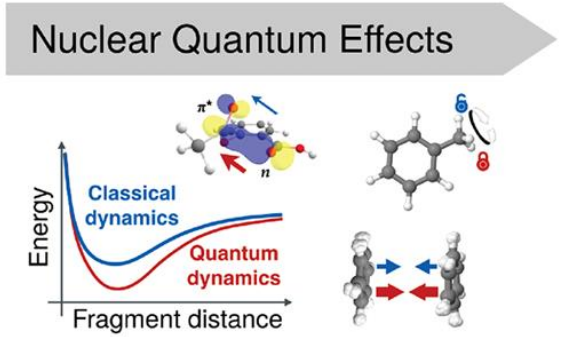
分子动力学（molecular dynamics）

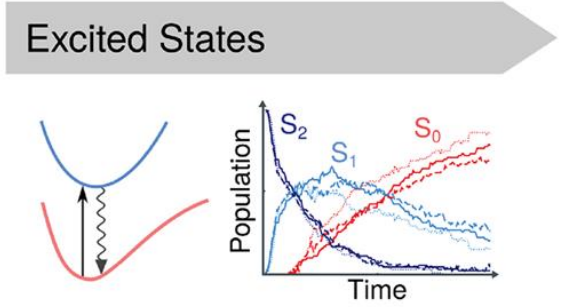
物理过程与势能面

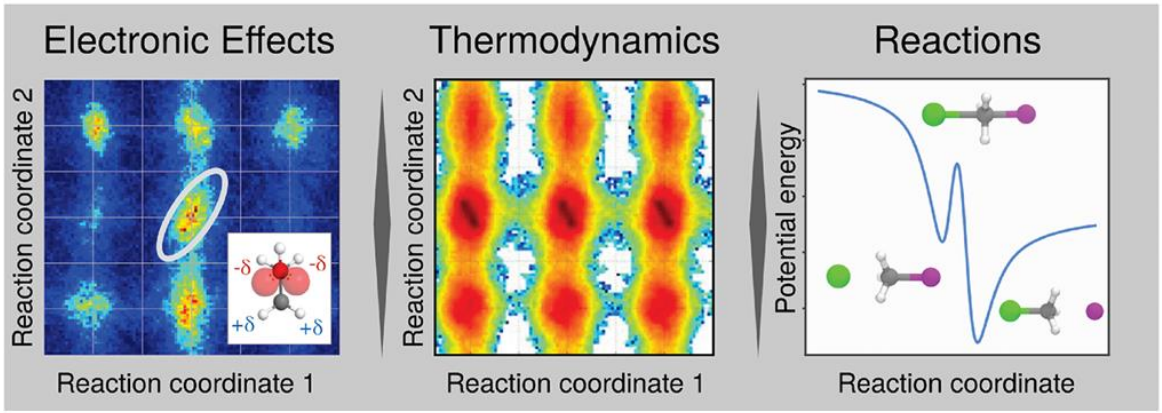
许多重要的物理现象都是关于
实际物质系统的演化过程

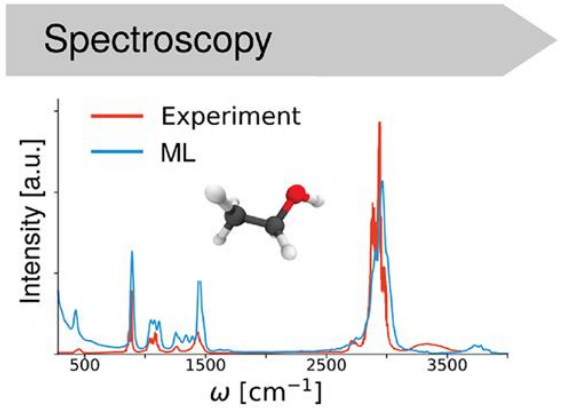
而决定演化过程的归根结底是势能面
Potential Energy Surface (PES)











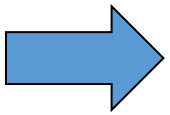
分子动力学（molecular dynamics）

哈密顿方程

$$H(p,q) = \mathrm{T(p)} + \mathrm{V(q)}$$

$$\dot{p} = -\partial_q H(p,q)$$

$$\dot{q} = \partial_p H(p,q)$$



Velocity Verlet算法

$$p_{n-\frac{1}{2}} = p_{n-1} - \frac{\tau}{2}\nabla \mathrm{V}(\mathrm{q_{n-1}})$$

$$q_n = q_{n-1} + \tau \nabla T(p_{n-\frac{1}{2}})$$

$$p_n = p_{n-\frac{1}{2}} - \frac{\tau}{2}\nabla \mathrm{V}(\mathrm{q_n})$$

V(q): 势能面是分子动力学的“动力来源”，来自粒子间相互作用

分子动力学中的力场模型发展

时间阶段	力场模型	代表性人物和工作
1950-1970	对势	硬球势，Lennard-Jones势，库伦势，Born-Meyer势
1970-1990	通用分子力场模型	Amber (Kollman), CHARMM (Karplus), OLPS (Jorgensen)
1990-2010	高精度力场	多体力场，反应力场，极化力场
2010至今	机器学习力场	Behler-Parrinello, GAP, SchNet, DeepPotential, MACE

内容提纲

1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

分子动力学：机器学习力场

第一性原理	一般有DFT、QMC和WF方法，精度最高 计算代价大，限制了计算体系的规模 （几十到几百原子）
传统力场	速度最快，计算规模可以达到数万原子以上 计算精度不高，且解析形式限制了其泛用性
机器学习力场	无特设的解析形式，具有第一性原理的计算精度 速度接近于传统力场，可计算的尺度大

机器学习力场的发展

开创阶段（2007 年之前）

•采用简单机器学习模型拟合小分子势能面，局限于小体系，无法推广到凝聚相材料。

•里程碑工作是 2007 年提出的 Behler–Parrinello 高维神经网络势。

第二阶段（2010–2018）

•2010 年：Csányi 团队提出基于高斯过程回归的 GAP 模型。

•2017 年：Schütt 等人提出基于连续滤波卷积神经网络的 SchNet。

•2018 年：张林峰等人提出基于深度势框架的 DeepMD。

第三阶段（2018 年至今）

•等变图神经网络（EGNN）成为主流。代表性模型包括 NequIP、MACE 、PaiNN、DPA4，这类模型无需人工预定义描述符，可以自动构建旋转等变表示。等变网络结构天然满足原子体系的物理对称性，拥有更高的精度与泛化能力。

•当前研究热点集中在提升外推能力、引入长程静电相互作用，以及开发适用于广阔化学空间的通用预训练机器学习原子势（MLIP）。

分子动力学：机器学习力场

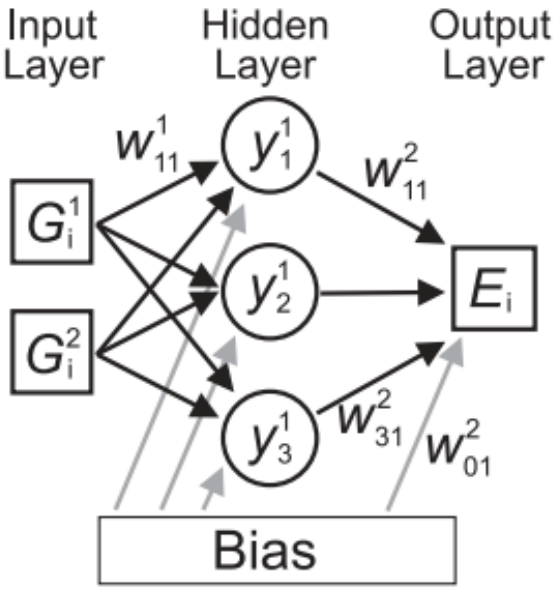
2007 年 Behler 与 Parrinello 的工作首先提出了高维神经网络势。

不同于具有固定解析能量函数的传统经典力场Behler–Parrinello 模型是一种数据驱动势，完全基于第一性原理密度泛函理论（DFT）数据集训练得到。

体系总能量被分解为各个原子能量之和，其中每个原子的能量仅由其局域化学环境决定

$$E_{\mathrm{tot}} = \sum_i E_i$$

$$E_i = f_a^2\left[w_{01}^2 + \sum_{j=1}^{3} w_{j1}^2 f_a^1\left(w_{0j}^1 + \sum_{\mu=1}^{2} w_{\mu j}^1 G_i^\mu\right)\right].$$



分子动力学：机器学习力场

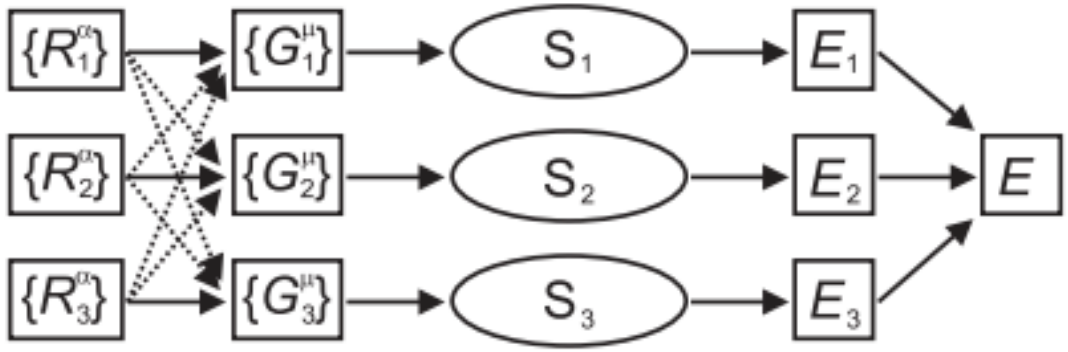


FIG. 2. Structure of the neural network as applied in this Letter to a system containing three atoms. The Cartesian coordinates of atom i are given by R_i^α. These are transformed to a set of μ symmetry function values G_i^μ describing the local geometric environment of atom i, which depends on the positions of all atoms in the system as indicated by the dotted arrows. The symmetry function values of atom i then enter the subnet S_i yielding the energy contribution E_i of atom i to the total energy of the system E. The structure of the subnets corresponds to the neural network shown in Fig. 1.

In order to define the energetically relevant local environment we employ a cutoff function f_c of the interatomic distance R_{ij}, which has the form

$$f_c(R_{ij}) = \begin{cases} 0.5 \times \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1\right] & \text{for } R_{ij} \leq R_c, \\ 0 & \text{for } R_{ij} > R_c. \end{cases} \quad (3)$$

Radial symmetry functions are constructed as a sum of Gaussians with the parameters η and R_s,

$$G_i^1 = \sum_{j \neq i}^{\text{all}} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}). \quad (4)$$

The summation over all neighbors j ensures the independence of the coordination number.

Angular terms are constructed for all triplets of atoms by summing the cosine values of the angles $\theta_{ijk} = \frac{\mathbf{R_{ij}} \cdot \mathbf{R_{ik}}}{R_{ij} R_{ik}}$ centered at atom i, with $\mathbf{R_{ij}} = \mathbf{R_i} - \mathbf{R_j}$,

$$G_i^2 = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{all}} (1 + \lambda \cos\theta_{ijk})^\zeta \times e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}), \quad (5)$$

分子动力学：机器学习力场



Michele Parrinello
SISSA (Trieste), IBM Zurich, Max Planck Institute for Solid State Research, ETH Zurich and Università della Svizzera Italiana

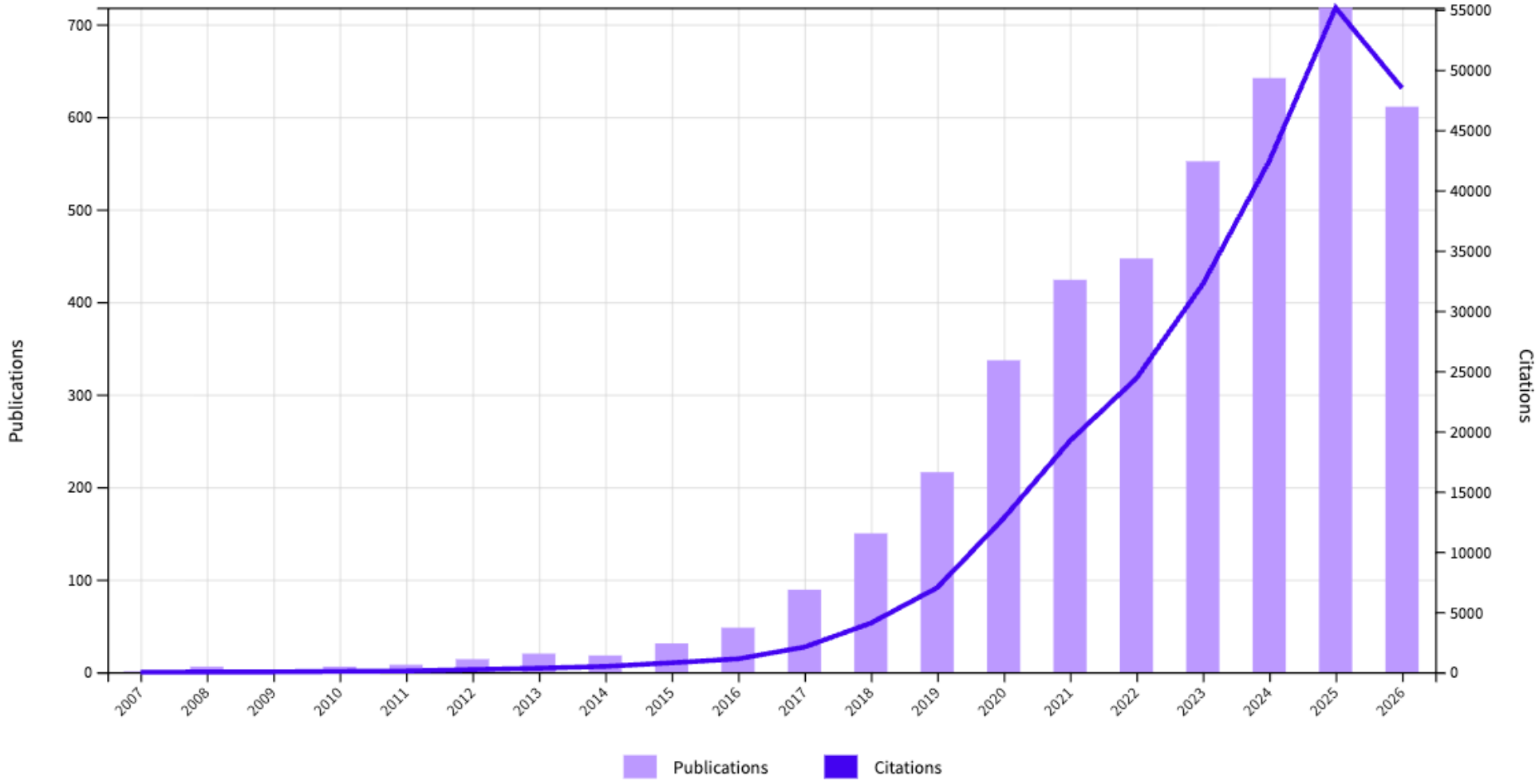
Major research contributions

- Car–Parrinello ab initio molecular dynamics
- Parrinello–Rahman method
- Metadynamics
- Behler–Parrinello neural network potential

Major awards

•APS Rahman Prize for Computational Physics (1995)
•Dirac Medal of the International Centre for Theoretical Physics (2009, shared with Roberto Car)
•Berni J. Alder CECAM Prize (2010)
•Enrico Fermi Prize of the Italian Physical Society (2012)
•Benjamin Franklin Medal in Chemistry (2020)
•Erwin Schrödinger Institute Medal (2025)

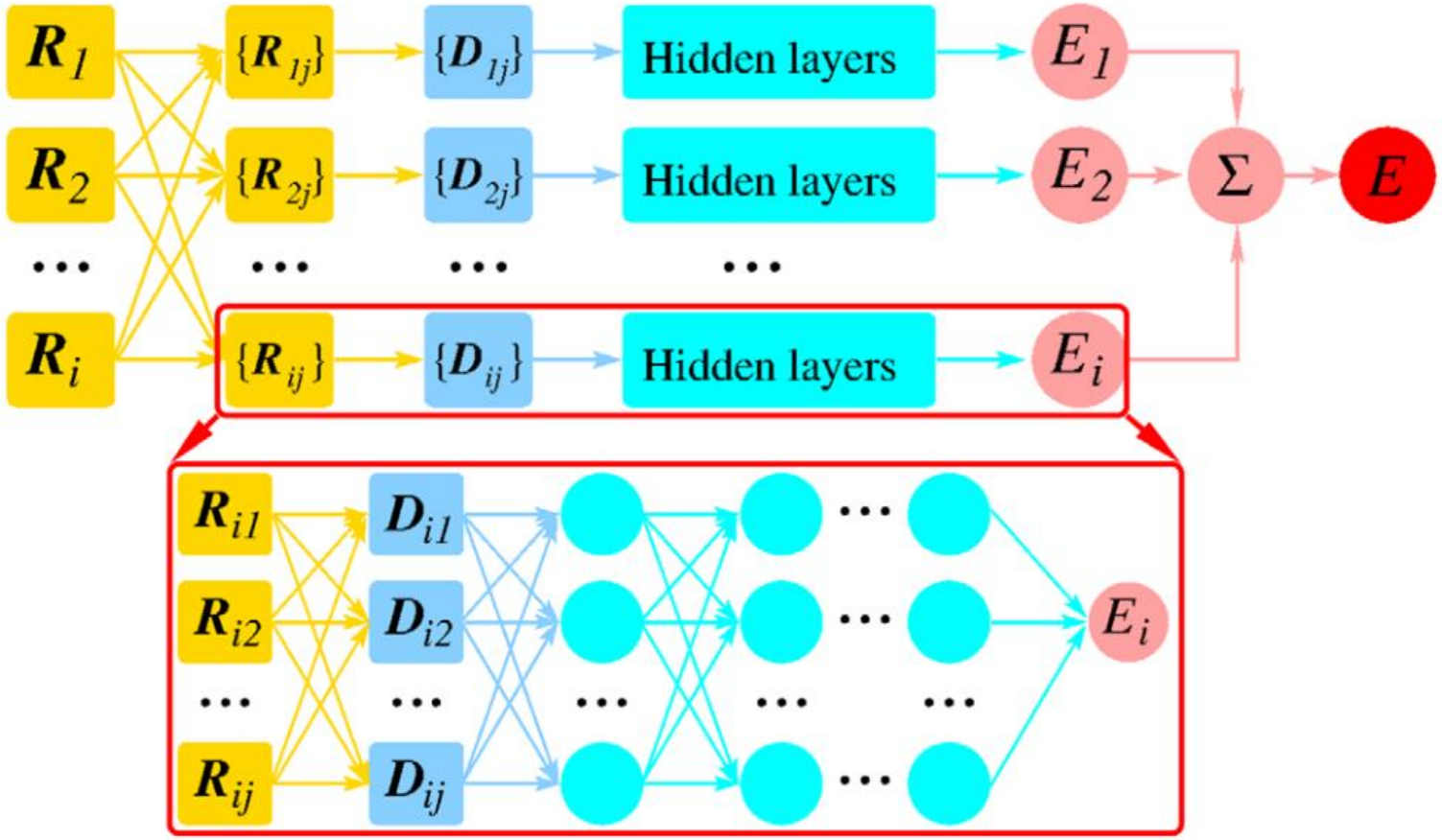
分子动力学：机器学习力场



分子动力学：机器学习力场

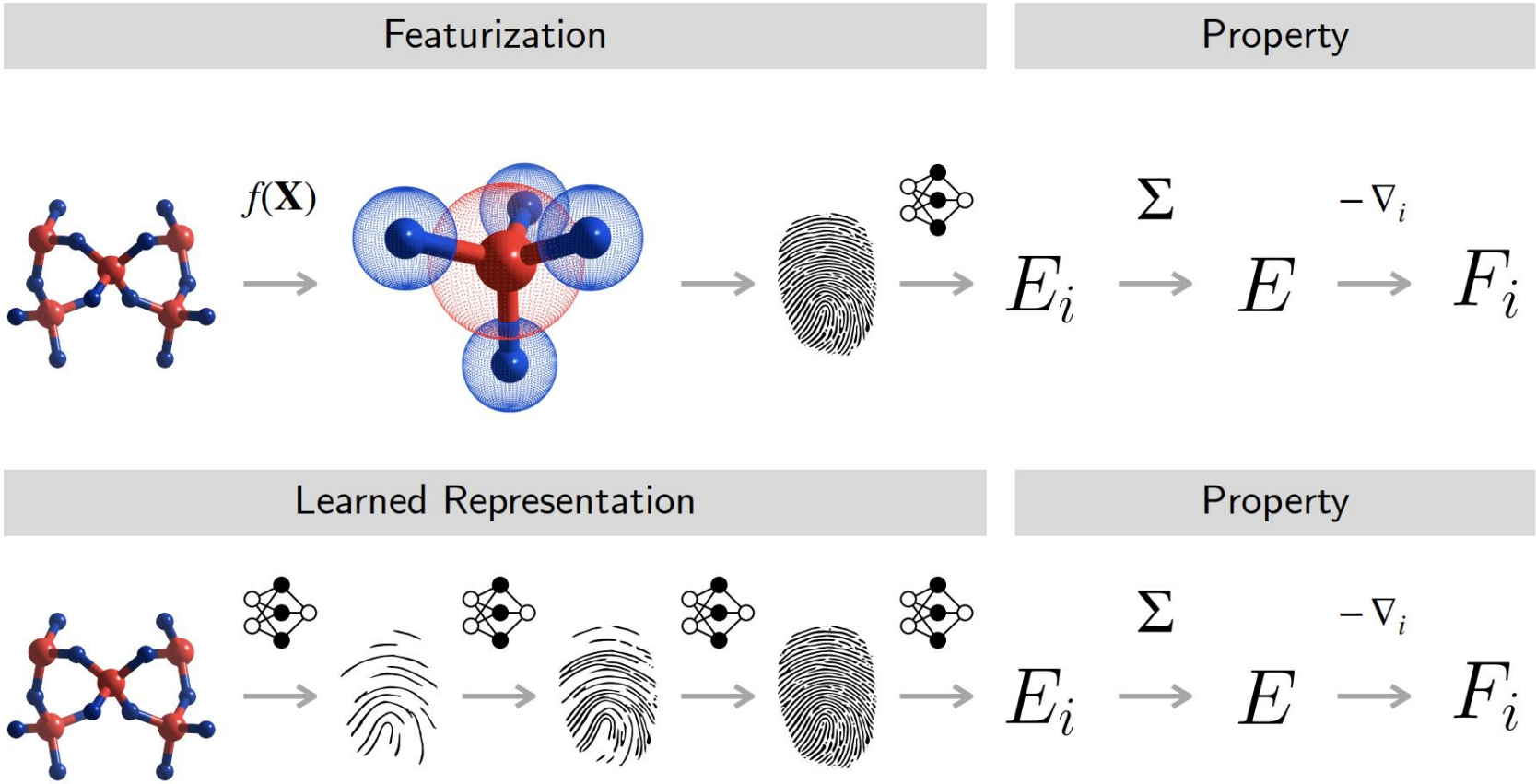
张林峰等：DeepPotential

特征设计实现E(3)不变：
在中心原子处，选两个不共线近邻，做正交化，手动建立一个依附局域原子环境的局部正交坐标系，把所有近邻相对坐标投影到这个局部坐标系，在局部坐标系下的坐标分量，天然不随全局坐标系旋转而改变。



分子动力学：机器学习力场

对分子信息的卷积操作：图神经网络GNN

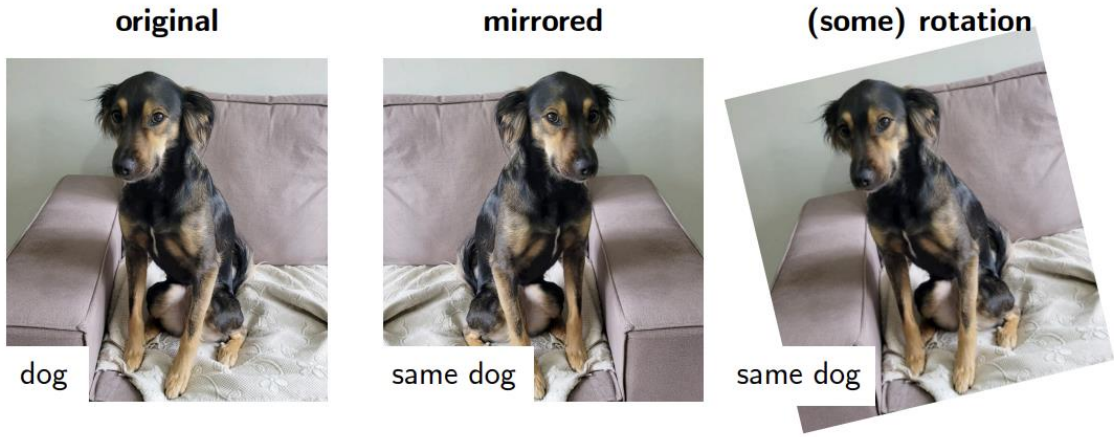


利用图卷积（**graph convolution**）进行特征提取后，
我们可以构建一个端到端的神经网络力场模型，无需特征工程。

分子动力学：机器学习力场

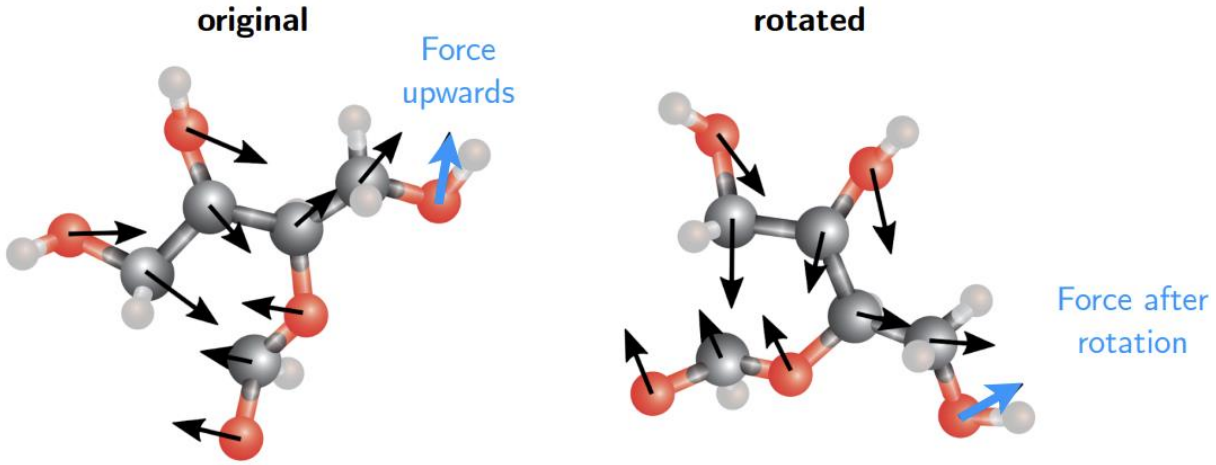
等变神经网络

Invariance



Equivariance

$$D_Y[g]f(x) = f(D_X[g]x) \quad \forall g \in G, \forall x \in X$$

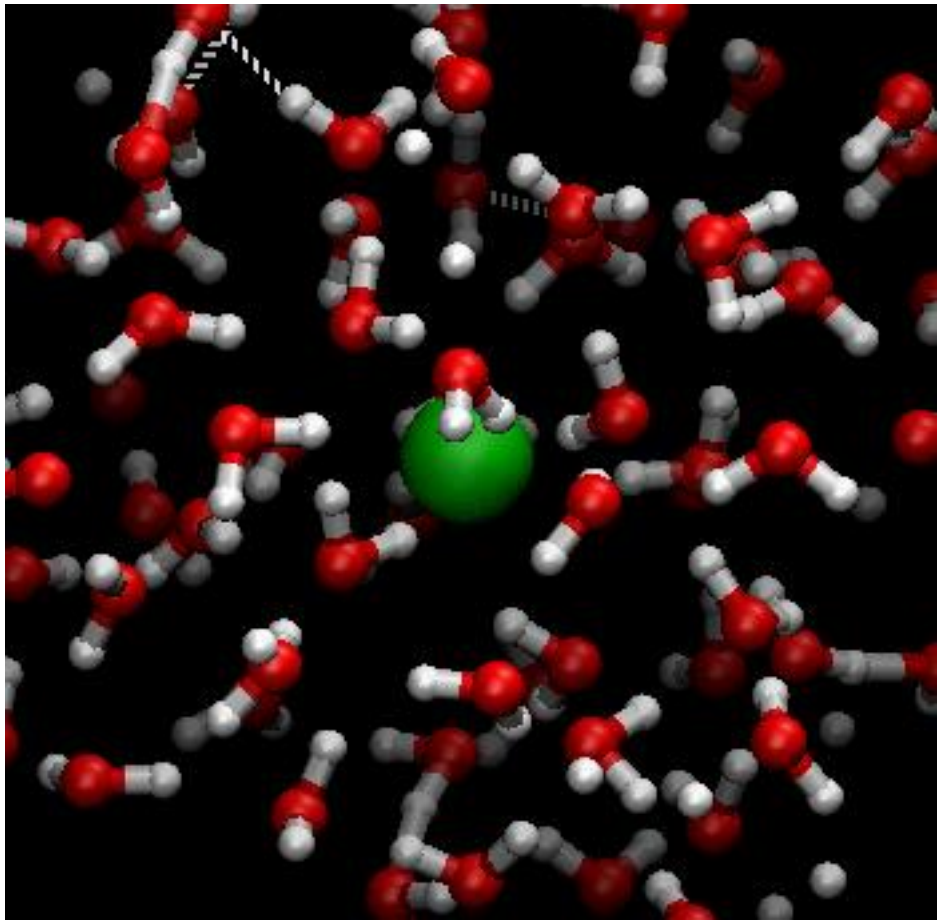


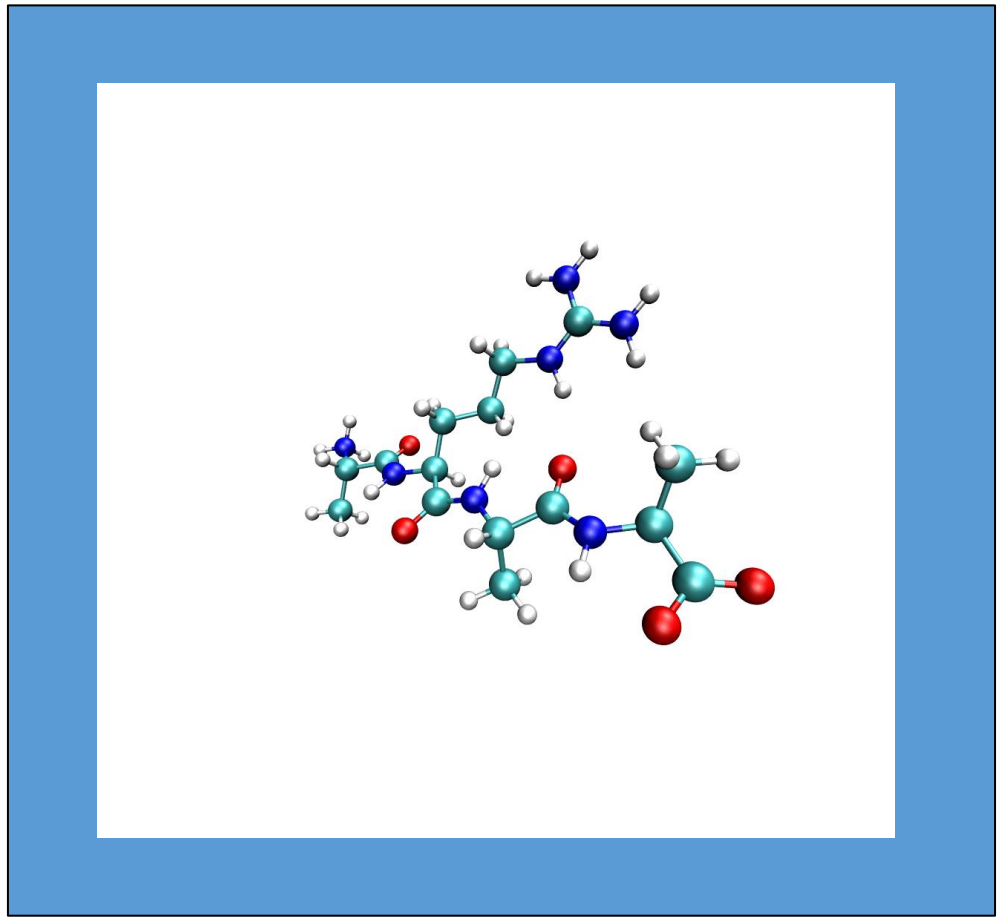
等变(equivariant)与不变(invariant)相对，指当对系统施加一个变换时，其中的某些特征会随着一个同样的变换而变换。

内容提纲

1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

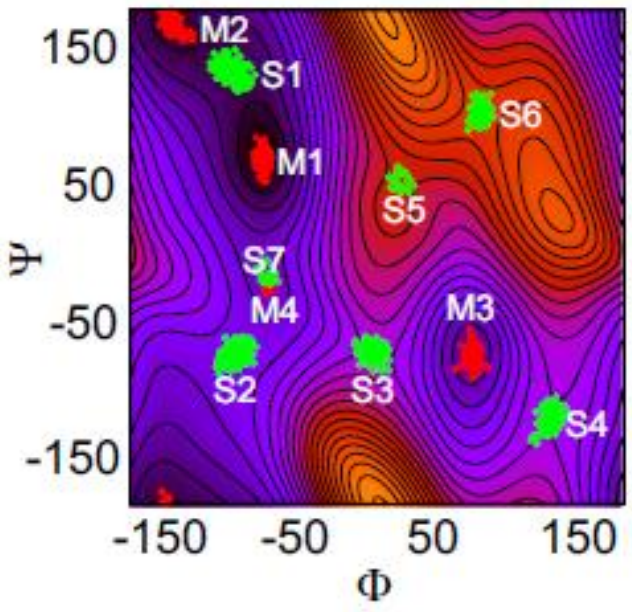
轨迹分析



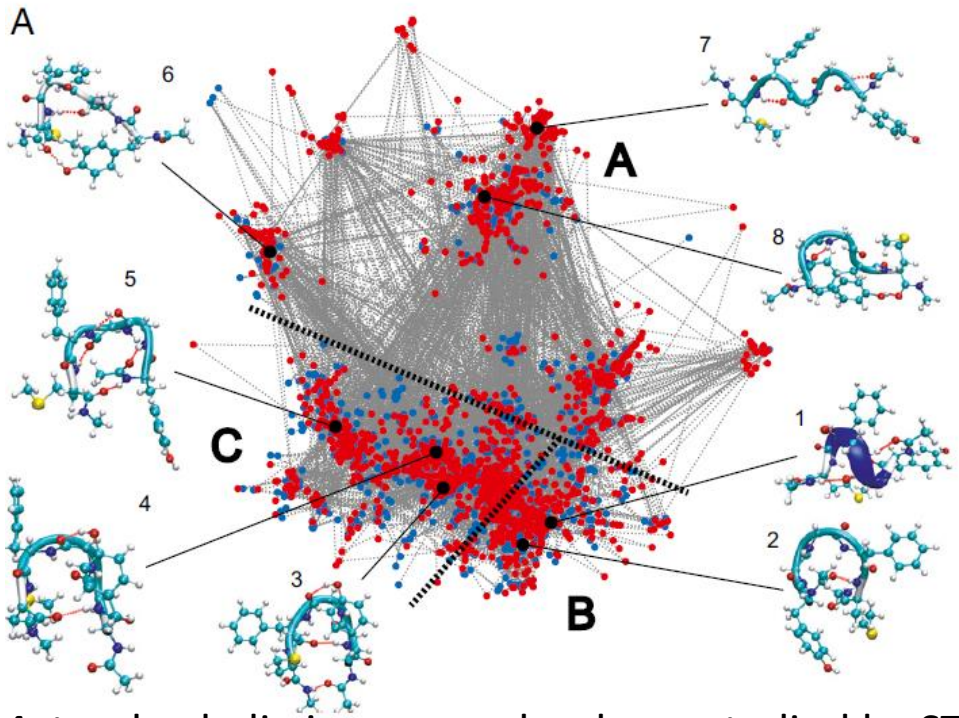


分析复杂结构

Example: Locating landmarks on high-dimensional free energy surfaces



FES as a function of the dihedral angles Φ and Ψ



Met-enkephalin in vacuum has been studied by START with the 10 Ramachandran dihedral angles as CVs.

Chen et al. PNAS 112, 3235 (2015)

加速采样

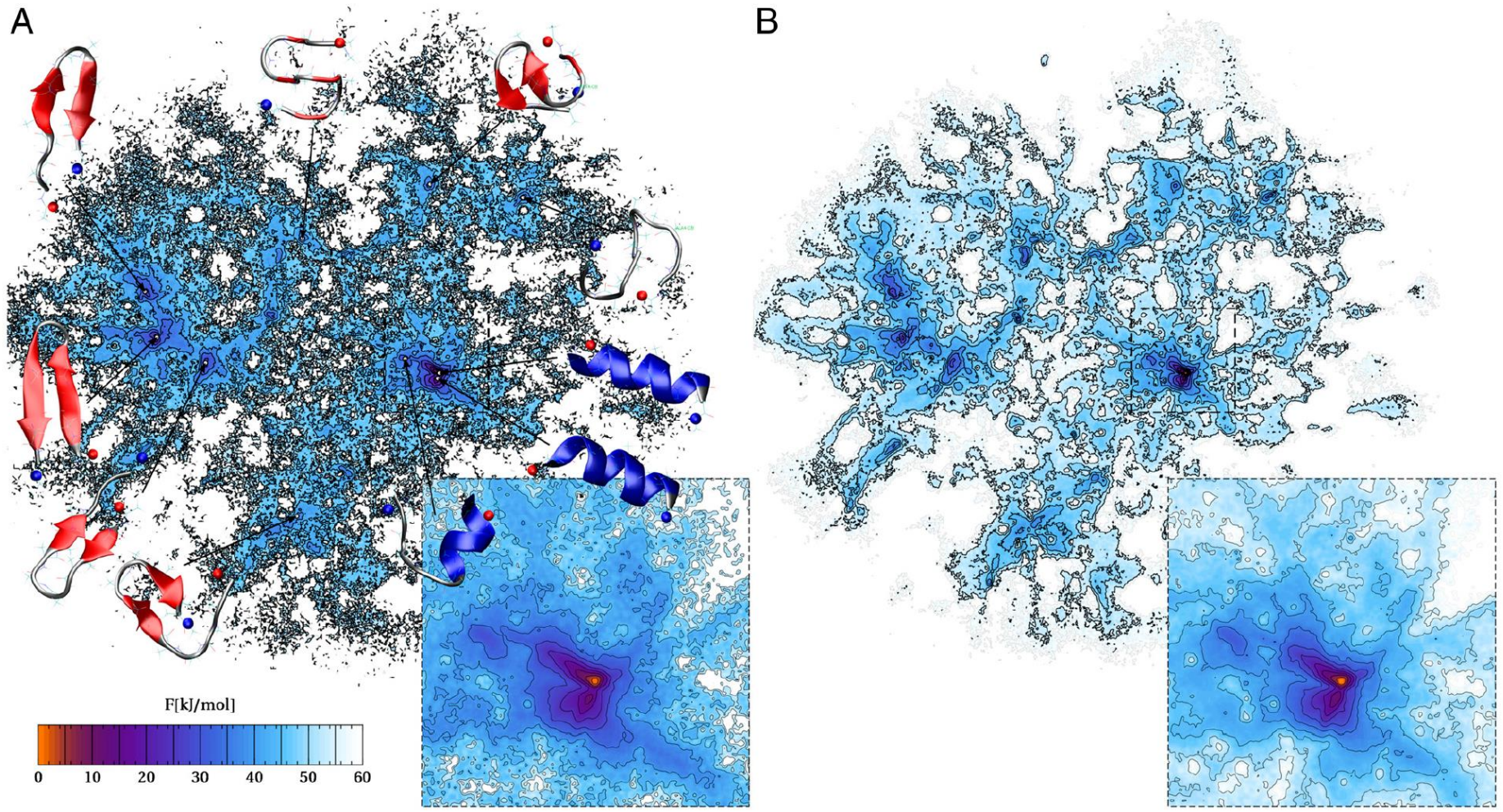


Fig. 4. The free-energy landscape for ala12 in implicit solvent calculated from parallel tempering (*A*) and field-overlap metadynamics (*B*) simulations. Here the free energy is shown as a function of the sketch-map coordinates and is seen to be very rough. In contrast to Fig. 3 each of the highlighted structures lies in a separate basin in the free-energy surface. In these structures the red and blue balls indicate the positions of the N and C termini, respectively.

降维与聚类

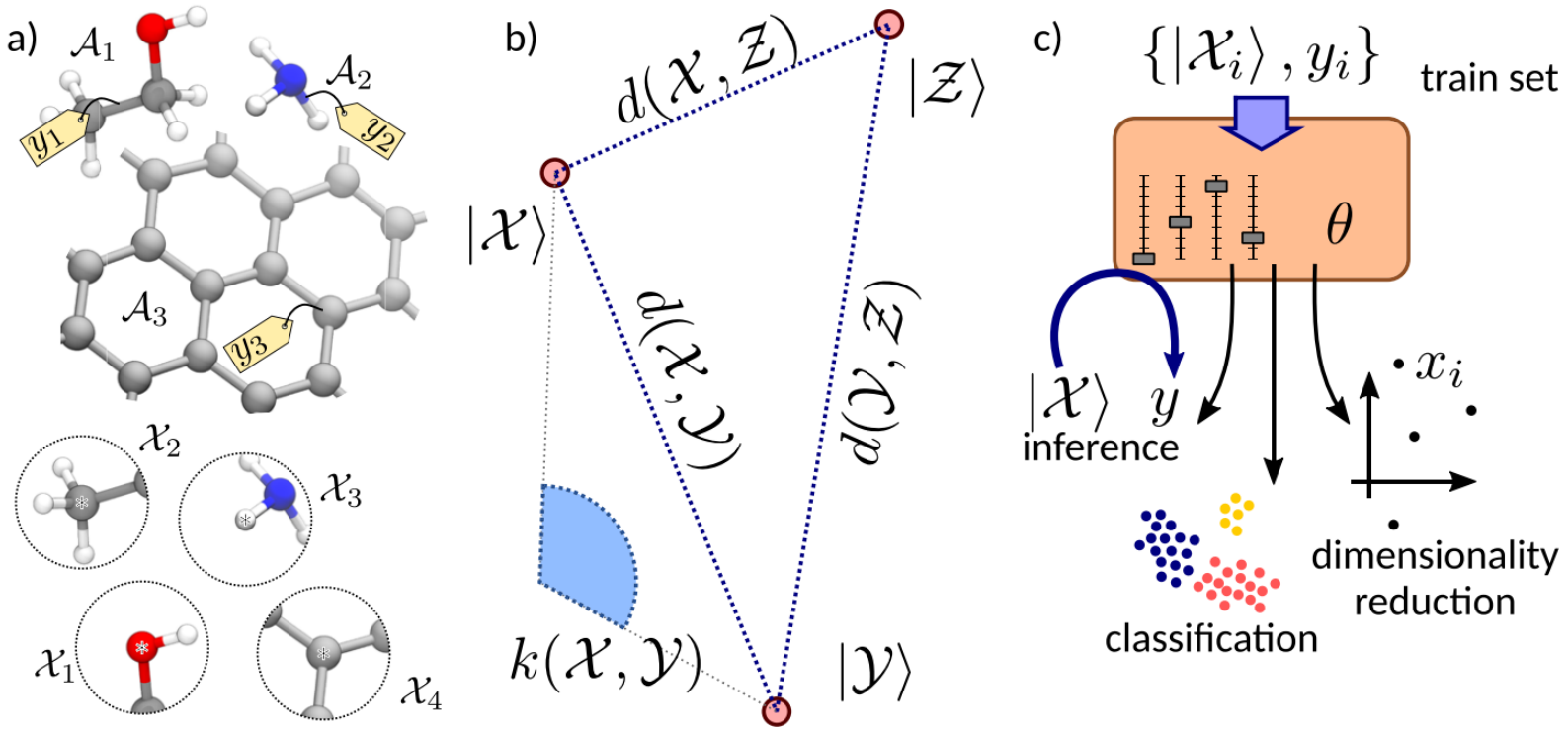
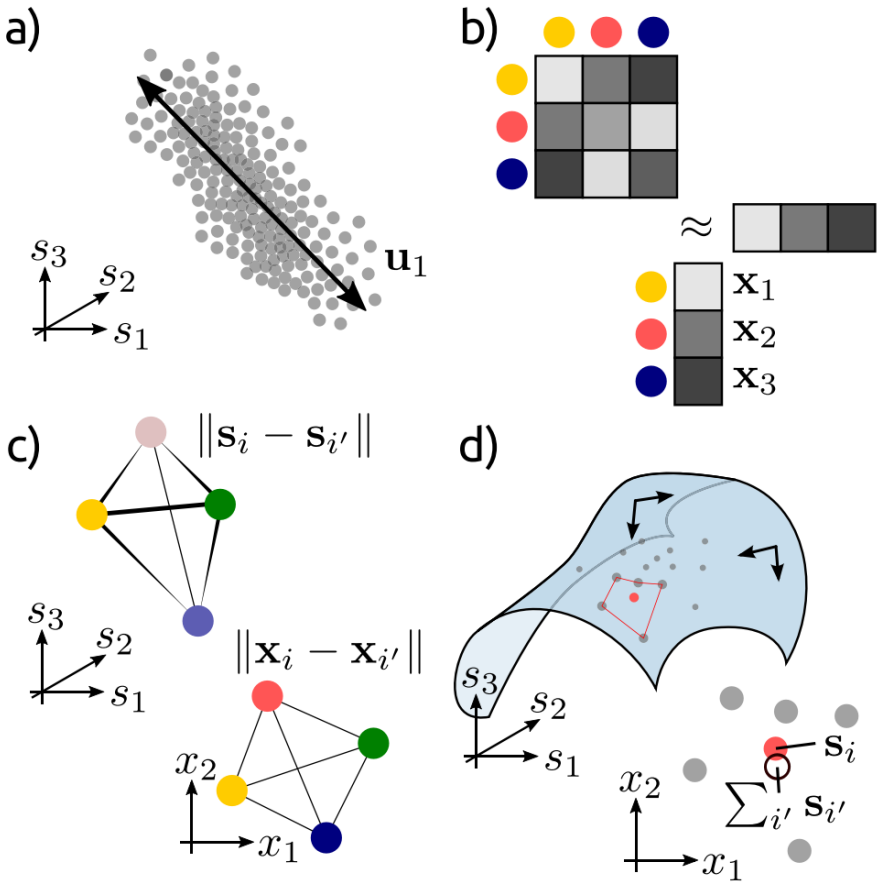


FIG. 1. A schematic representation of the main ingredients of atomistic machine learning. (a) Structures $\mathcal{A}$ or local environments $\mathcal{X}$ are the inputs of the model, possibly with labels y associated with them. (b) The inputs are associated with a mathematical representation, in terms of vectors of features $|\mathcal{X}\rangle$, a measure of similarity d, or a kernel k. (c) The machine-learning model, controlled by a series of hyperparameters θ, is trained based on a set of inputs. It can be then used for a number of machine-learning tasks.

J. Chem. Phys. 150, 150901 (2019)

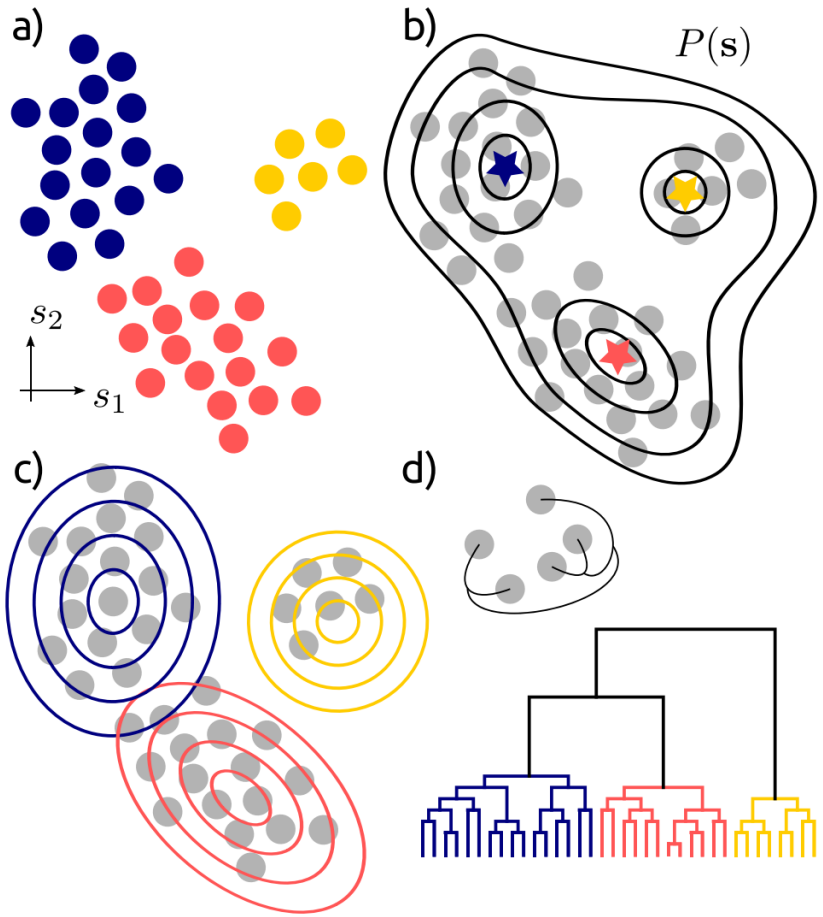
降维算法



Dimensionality reduction

FIG. 4. A summary of the main strategies underlying dimensionality reduction techniques. (a) Several methods (starting from PCA) attempt to identify the low-dimensional subspace that captures the largest fraction of the input data variance. (b) This is equivalent to finding the best $\mathcal{L}^2$-norm approximation of the Gram matrix and generalizes to kernel methods. (c) Multidimensional scaling and related approaches attempt to reproduce the similarity between high-dimensional data points in low dimension. (d) Embedding methods explicitly try to preserve local relations between points, under the assumption that they lie on a (locally) low-dimensional manifold.

聚类算法



Clustering

Clustering algorithms aim to recognize groups of input points that are related to each other and different from other groups of inputs. For instance, clusters could represent different classes of molecules, or configurations of a given system that are separated by a sparsely populated, or seldom accessed region

FIG. 3. A summary of clustering techniques. (a) A set of points in a finite-dimensional feature space are clustered together in a way that reflects some underlying common characteristic. (b) Density-based clustering identifies maxima in the probability distribution of inputs in feature space. (c) Distribution-based clustering determines a model of the data distribution as a combination of cluster probabilities. (d) (Hierarchical) linkage clustering determines work by accretion of clusters starting on inputs that are closest together in input space.

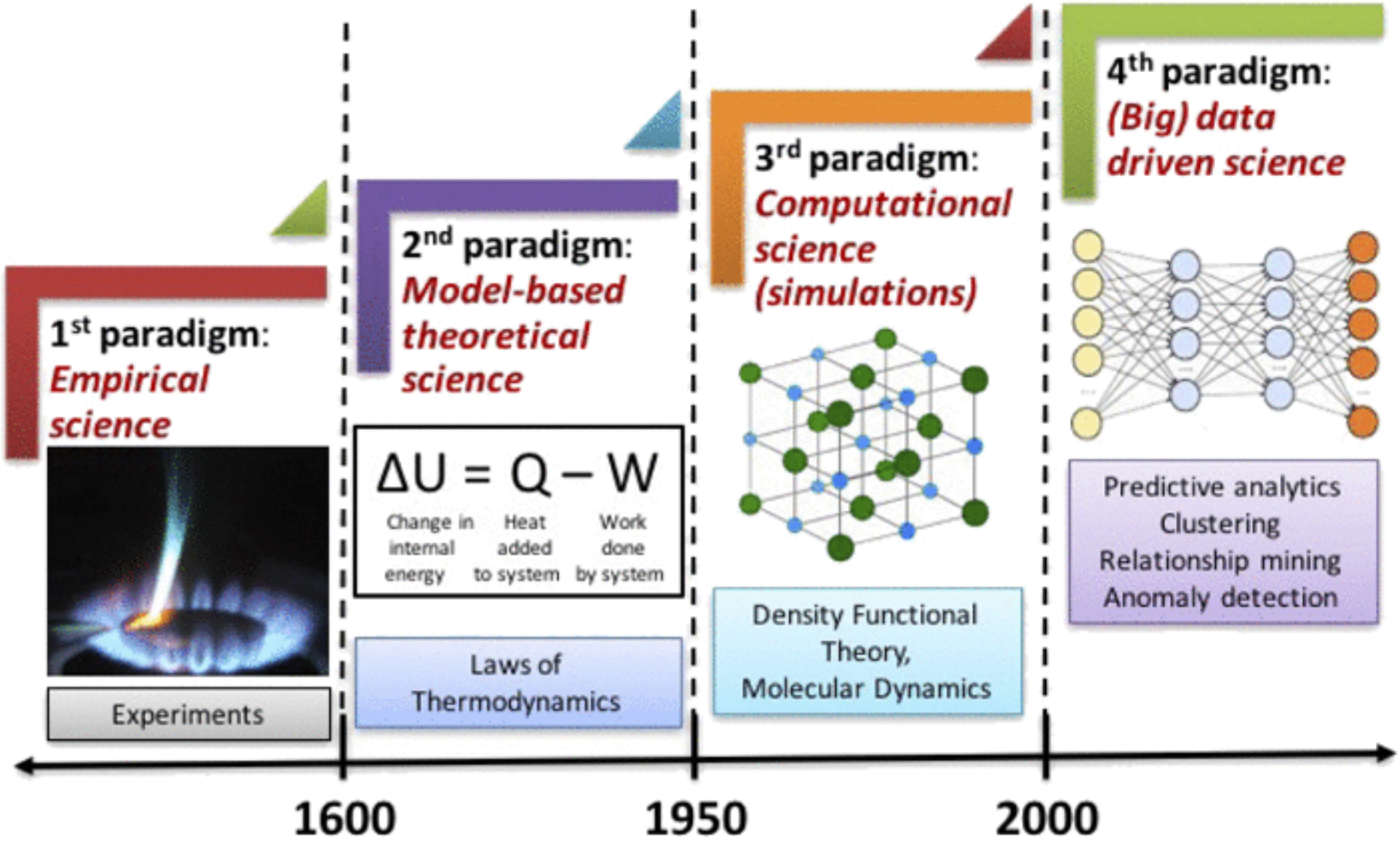
小结（2）：AI for 分子动力学模拟

1：精度问题	机器学习力场 材料的建模问题 监督式学习
2：效率问题	降维与聚类 序参量和增强采样 非监督式学习

内容提纲

1. AI for 第一性原理计算
   A. 深度学习波函数方法
   B. 深度学习密度泛函
2. AI for 分子动力学模拟
   A. 机器学习力场
   B. 降维与聚类
3. AI for 构效关系探索

物态调控与材料研发

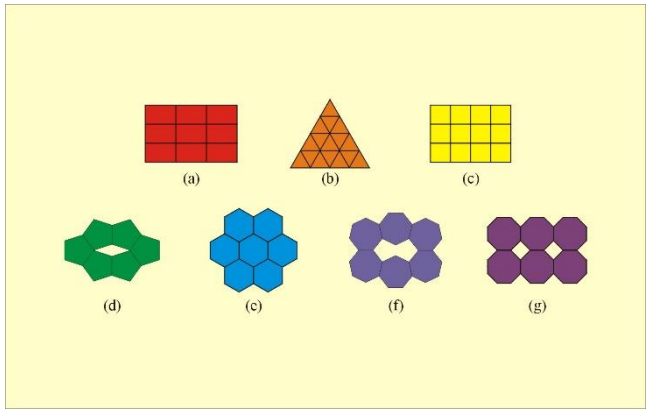

1st paradigm:
Empirical science
2nd paradigm:
Model-based theoretical science
3rd paradigm:
Computational science (simulations)
4th paradigm:
(Big) data driven science
ΔU = Q – W
Change in internal energy
Heat added to system
Work done by system
Predictive analytics
Clustering
Relationship mining
Anomaly detection
Experiments
Laws of Thermodynamics
Density Functional Theory, Molecular Dynamics
1600
1950
2000

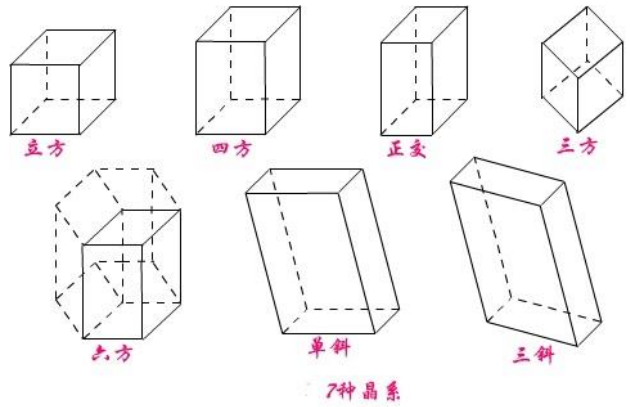
构效关系

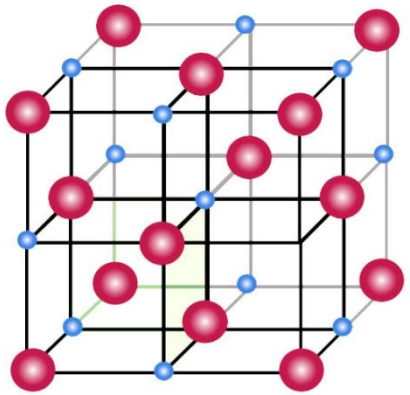
构效关系研究的旧范式

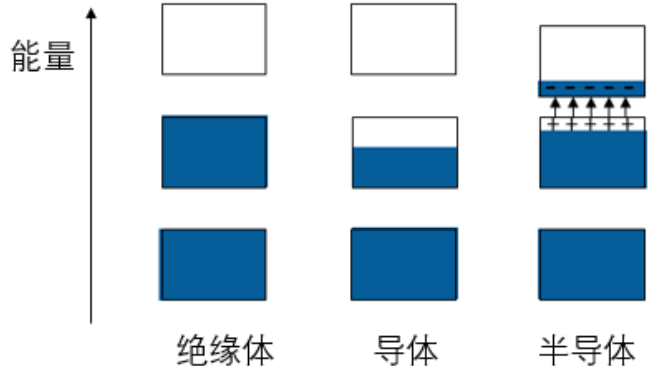
固体物理：周期性、晶格、原子结构、原子间相互作用、声子、能带、原子磁矩

关键词：简化与近似





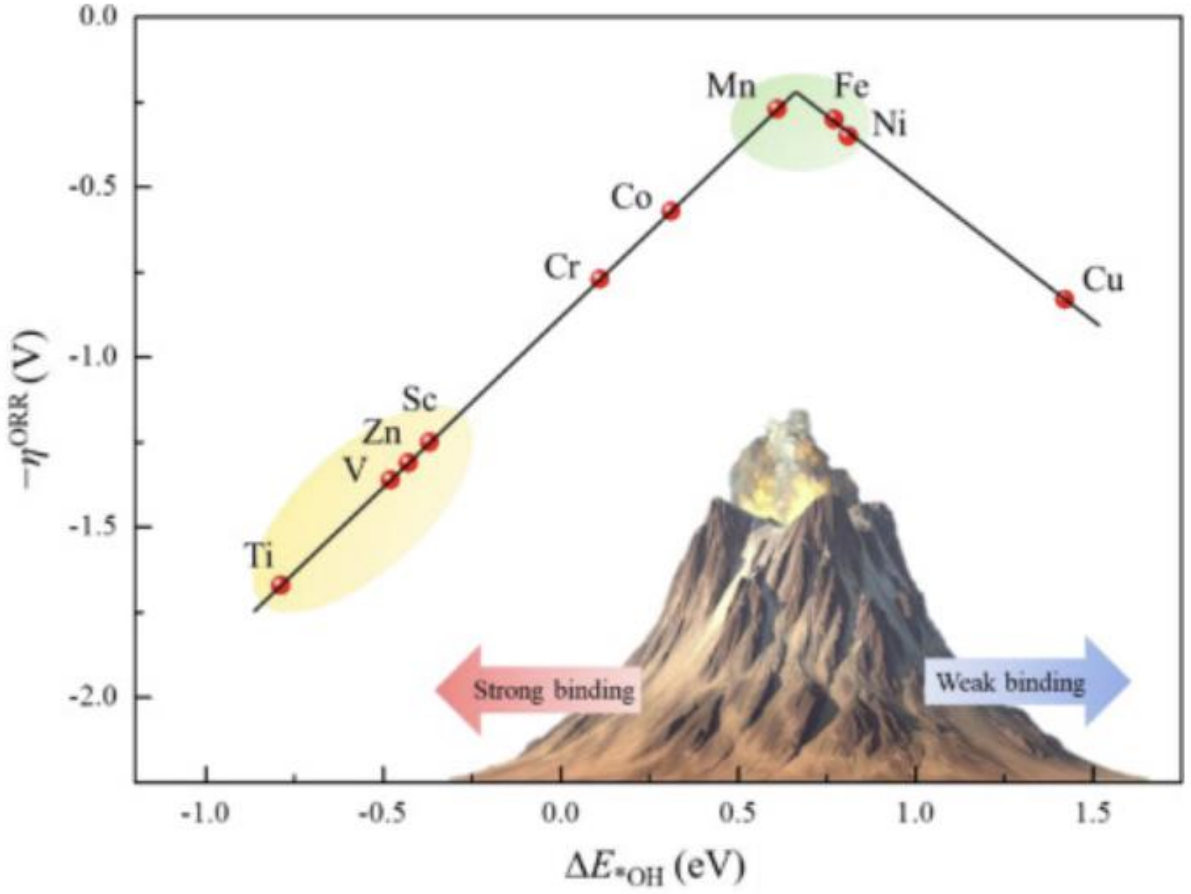




描述符：中间量

理解和建立“描述符”和“物性”之间的因果关系

构效关系



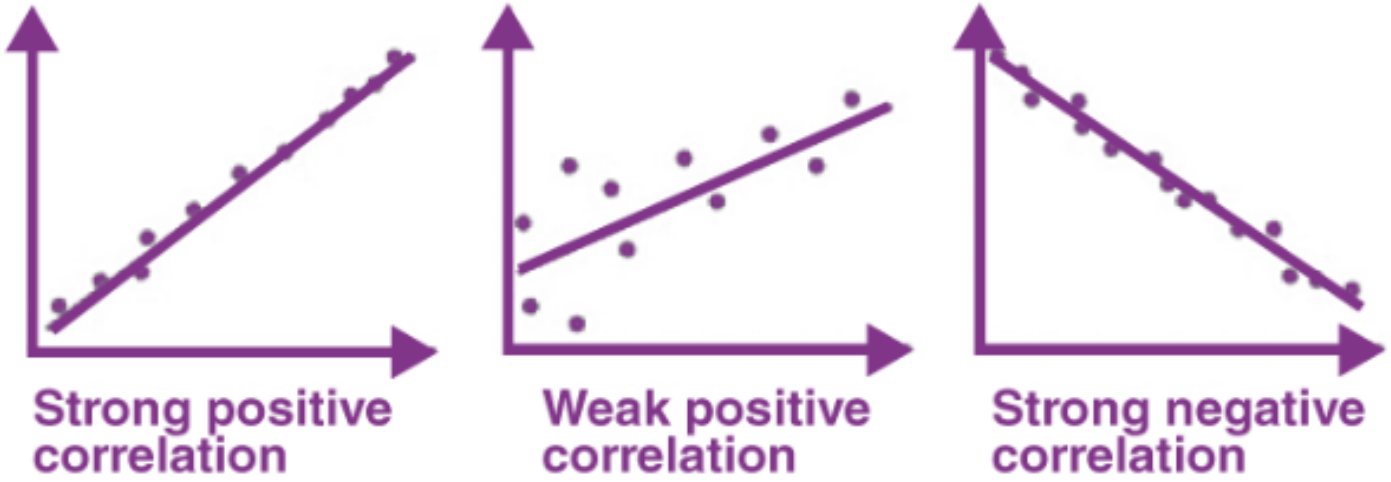
“化学活性”与“吸附能力”之间的“火山图”关系
因果关系：电子d带中心理论

构效关系

大数据和人工智能主导的新研究范式

寻找“规则”和“基因”

重“关联性”，轻“因果关系”



构效关系探索：数值回归

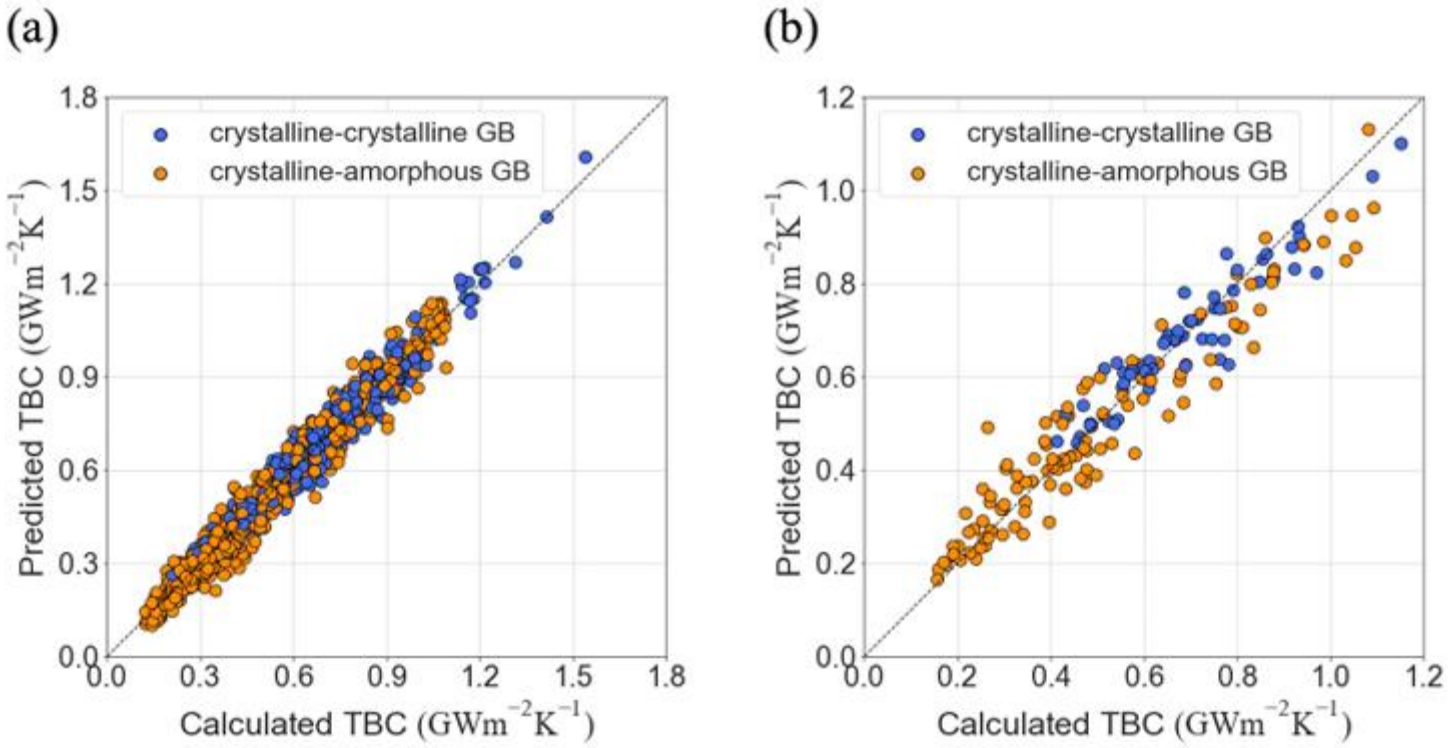


Table 1. Six-fold cross-validation results for thermal boundary conductance prediction of all training structures, ordered by root mean squared error (RMSE).

Model	RMSE (GWm^{-2}K^{-1})	MAPE
Ensemble	0.064	9.9%
ExtraTrees	0.078	11.3%
KNN	0.081	12.4%
GradientBoosting	0.087	15.6%
CatBoost	0.091	14.4%
AdaBoost	0.094	17.5%
XGBoost	0.098	19.8%
RandomForest	0.100	18.6%
HistGB	0.107	19.3%
LightGBM	0.108	23.0%
Linear regression	0.187	28.8%

Mean absolute percentage error (MAPE) was calculated for interpretation purpose. The ensemble model was constructed by averaging individual predictions from the five best models.

界面热导的机器学习预测模型

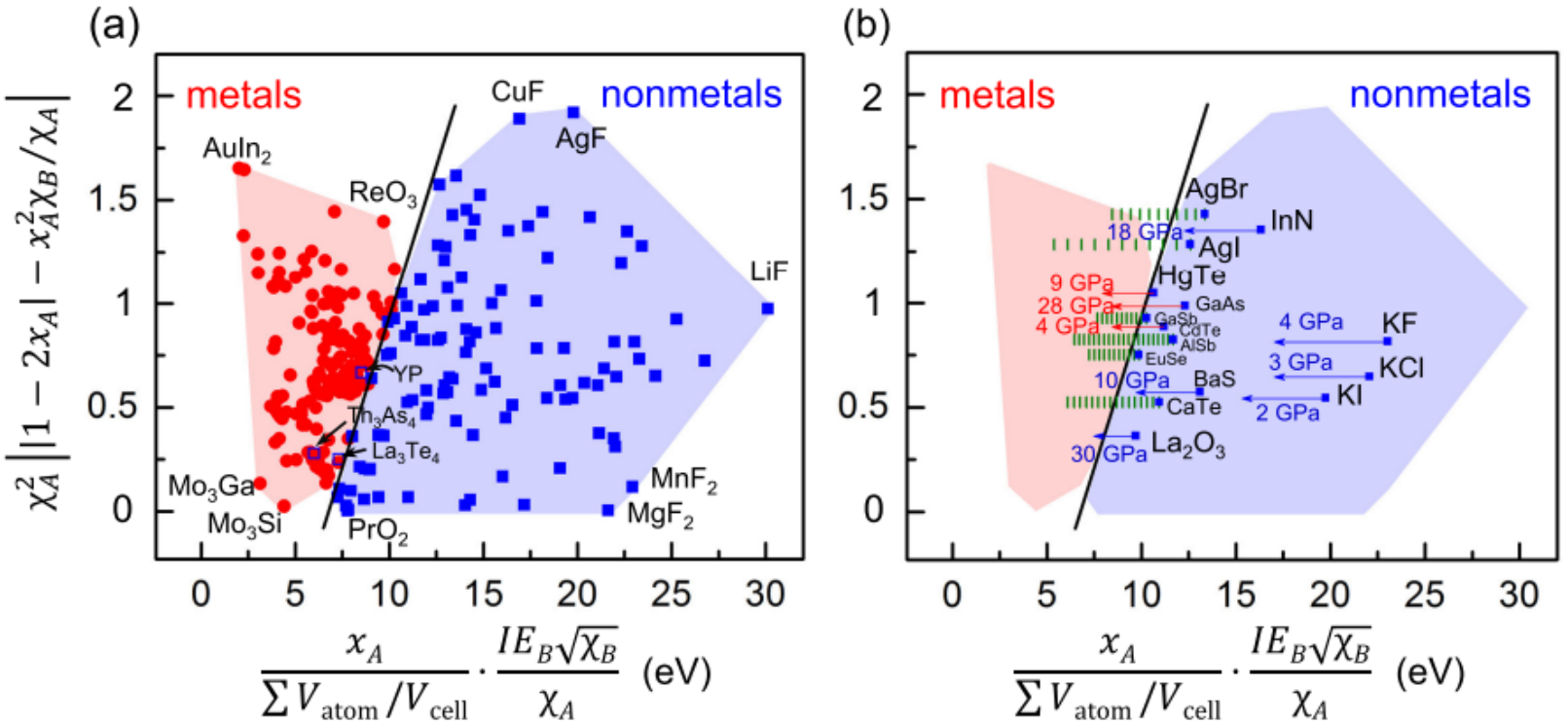
npj Computational Materials (2022) 8:219

构效关系探索：符号回归

$$\underset{c\in\mathbb{R}^M}{\operatorname{argmin}}\|\boldsymbol{P}-\boldsymbol{D}\boldsymbol{c}\|_2^2+\lambda\|\boldsymbol{c}\|_1,$$

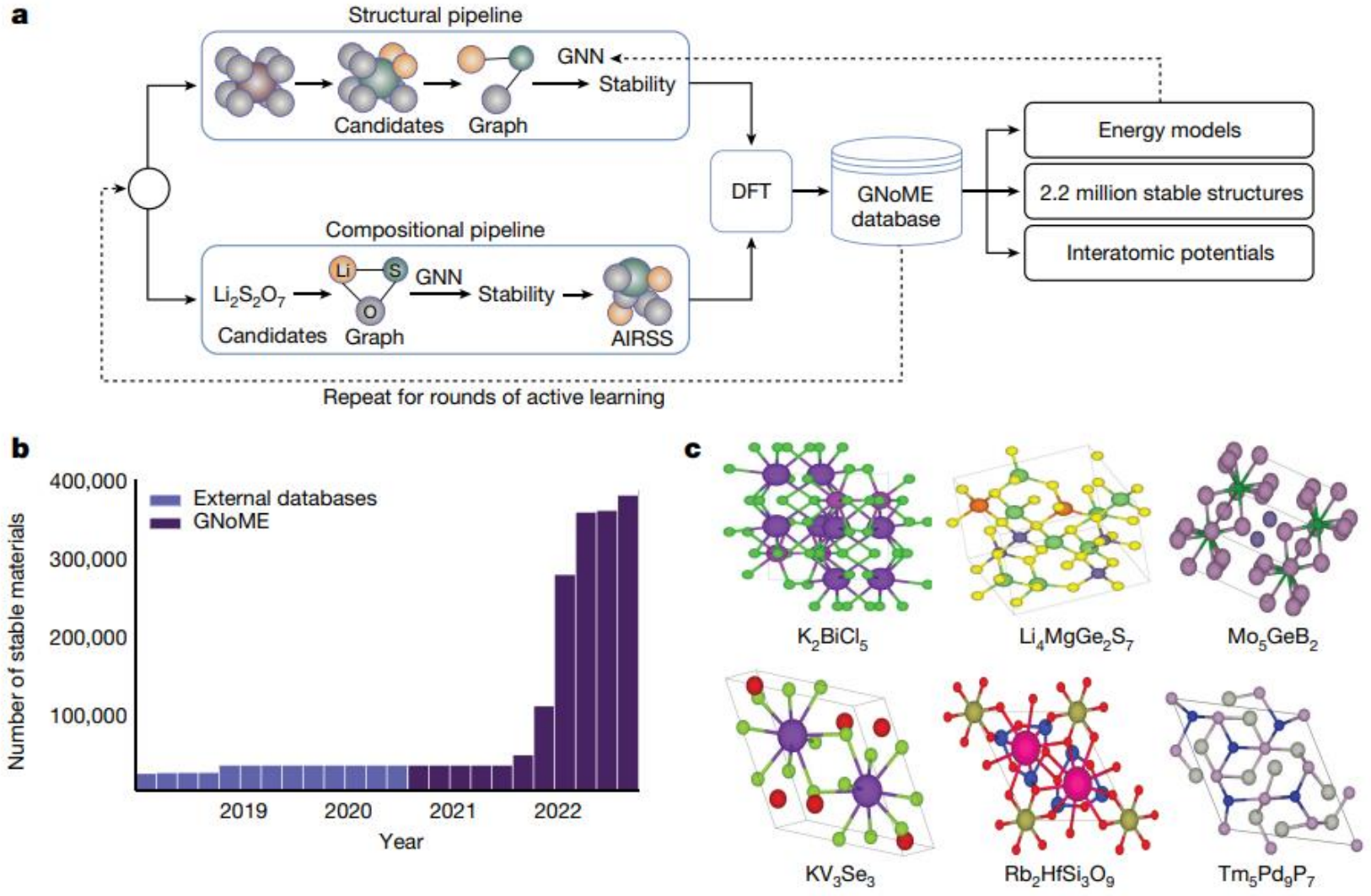
P: 物性特征
D: 描述符矩阵，是单个描述符d的组合
c: 模型参数

$$\hat{\boldsymbol{H}}^{(\mathrm{m})}\equiv\left\{I,+,-,\times,/,\exp,\log,|-|,\sqrt{},^{-1},^{2},^{3}\right\}[\phi_1,\phi_2],$$



Phys. Rev. Lett. 114, 105503 (2015); Phys. Rev. Mater. 2, 083802 (2018)

反向结构设计：生成模型



Nature 624, 80–85 (2023)

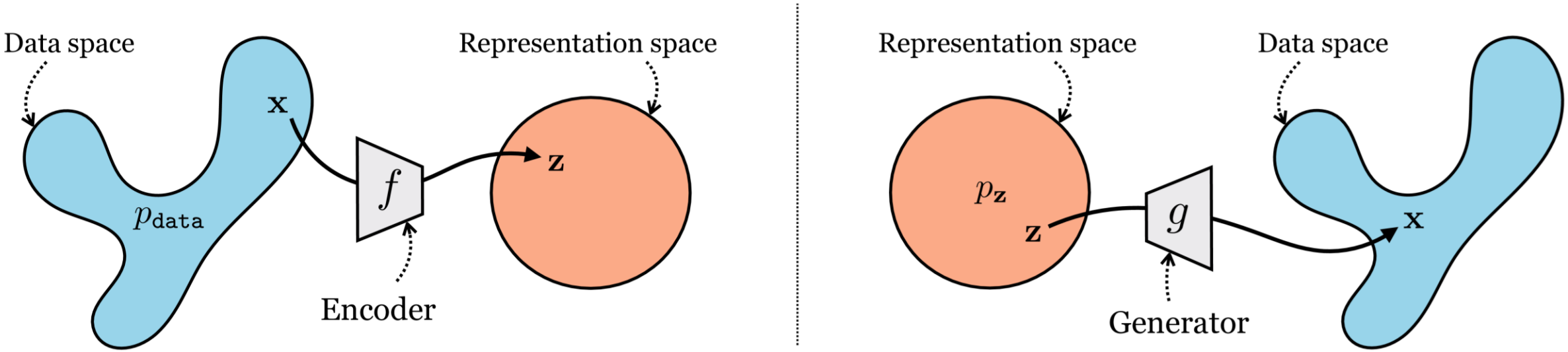
生成模型

学会模仿数据规律，然后自己创造出和它长得像的新东西的模型。



生成模型

原理：学习一个目标数据的概率分布，然后从这个分布里独立采样出新样本。



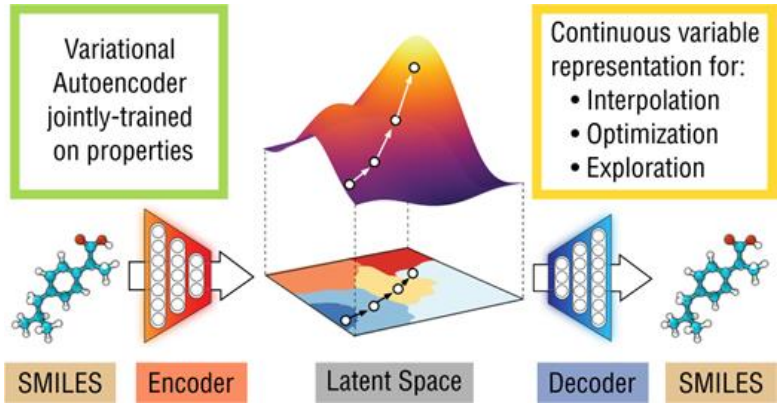
意义：替代传统（实验和计算）方法来更高效地产生接近真实的数据。

生成模型

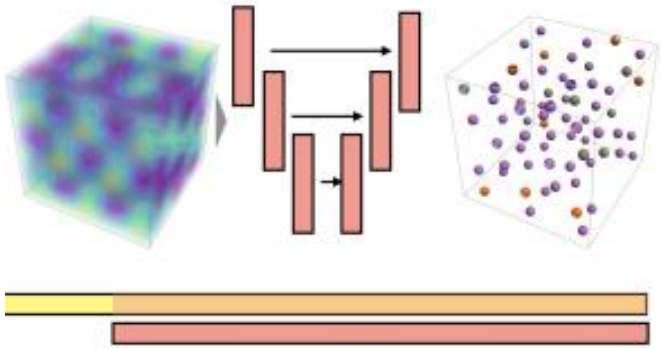
技术发展

1985 玻尔兹曼机 | 2013 变分自编码器 | 2014 对抗生成网络 | 2015 扩散模型 | 2017 Transformer | 2023 流匹配

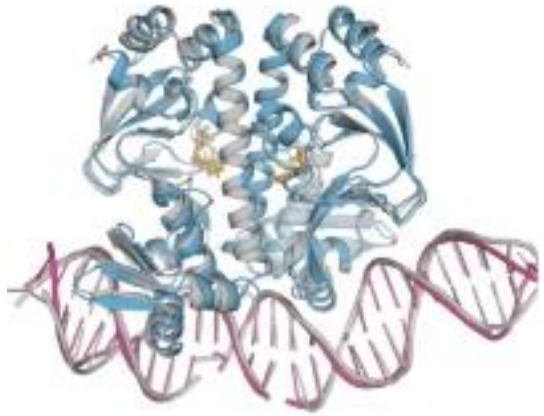
应用：分子、近体、蛋白



Gómez-Bombarelli et al.
ACS Cent. Sci. 268–276 (2018)



Hoffmann et al. arXiv:1909.00949

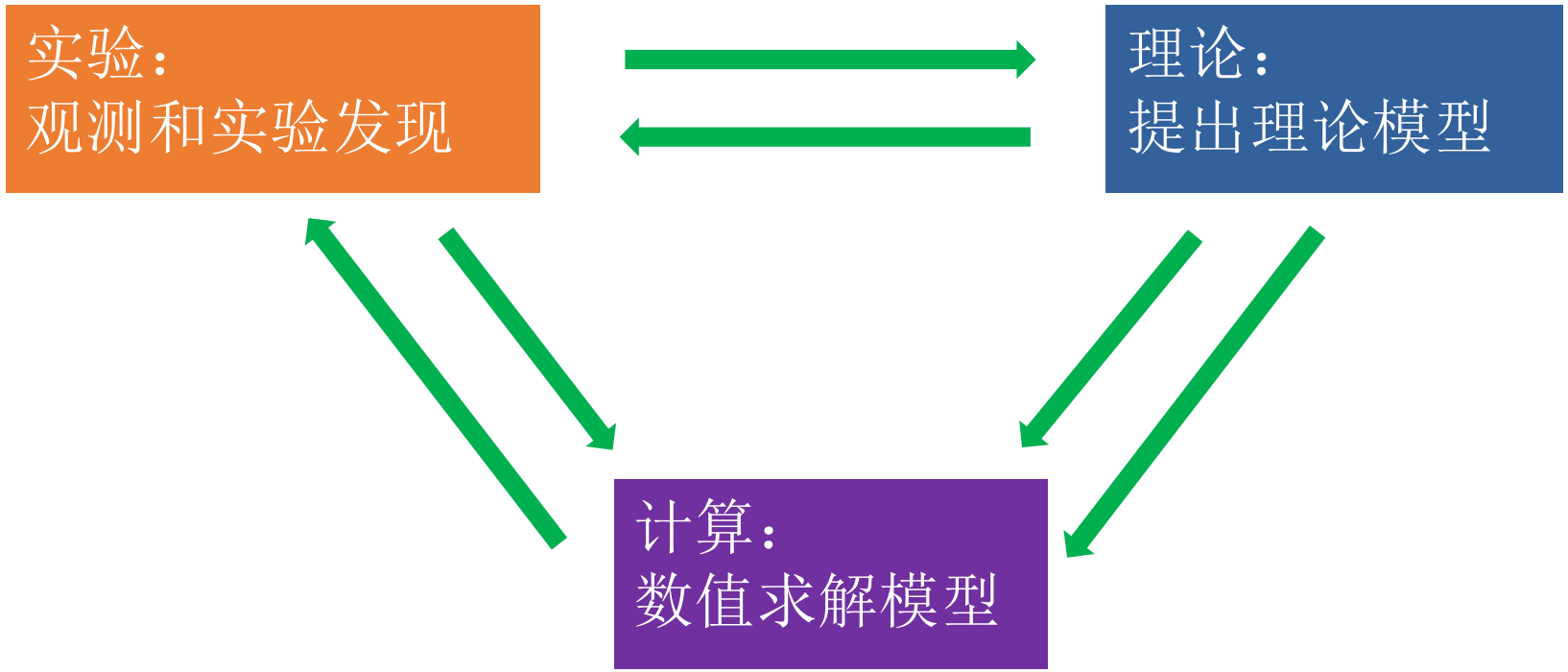


AlphaFold3
Nature 630, 493 (2024)

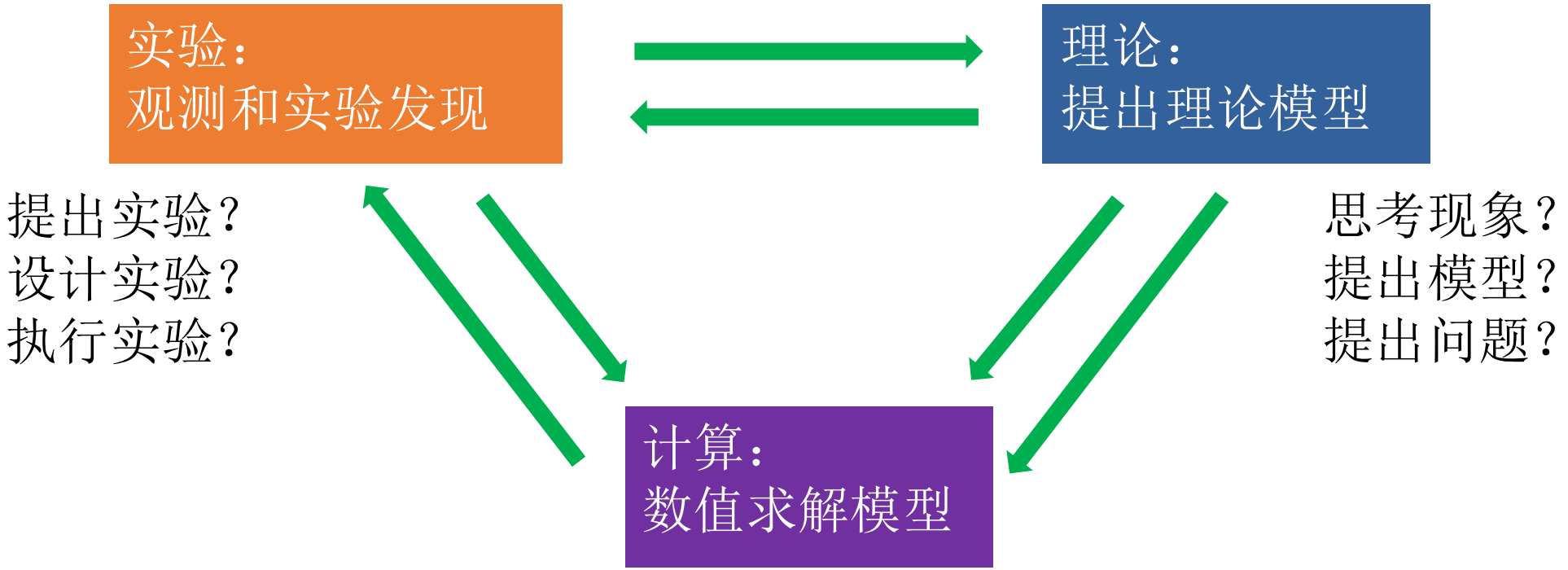
小结（3）：AI for 构效关系

1：精度问题	数值回归 符号回归 监督式学习
2：效率问题	生成模型 监督式学习+强化学习

凝聚态物理（AI史前时代）

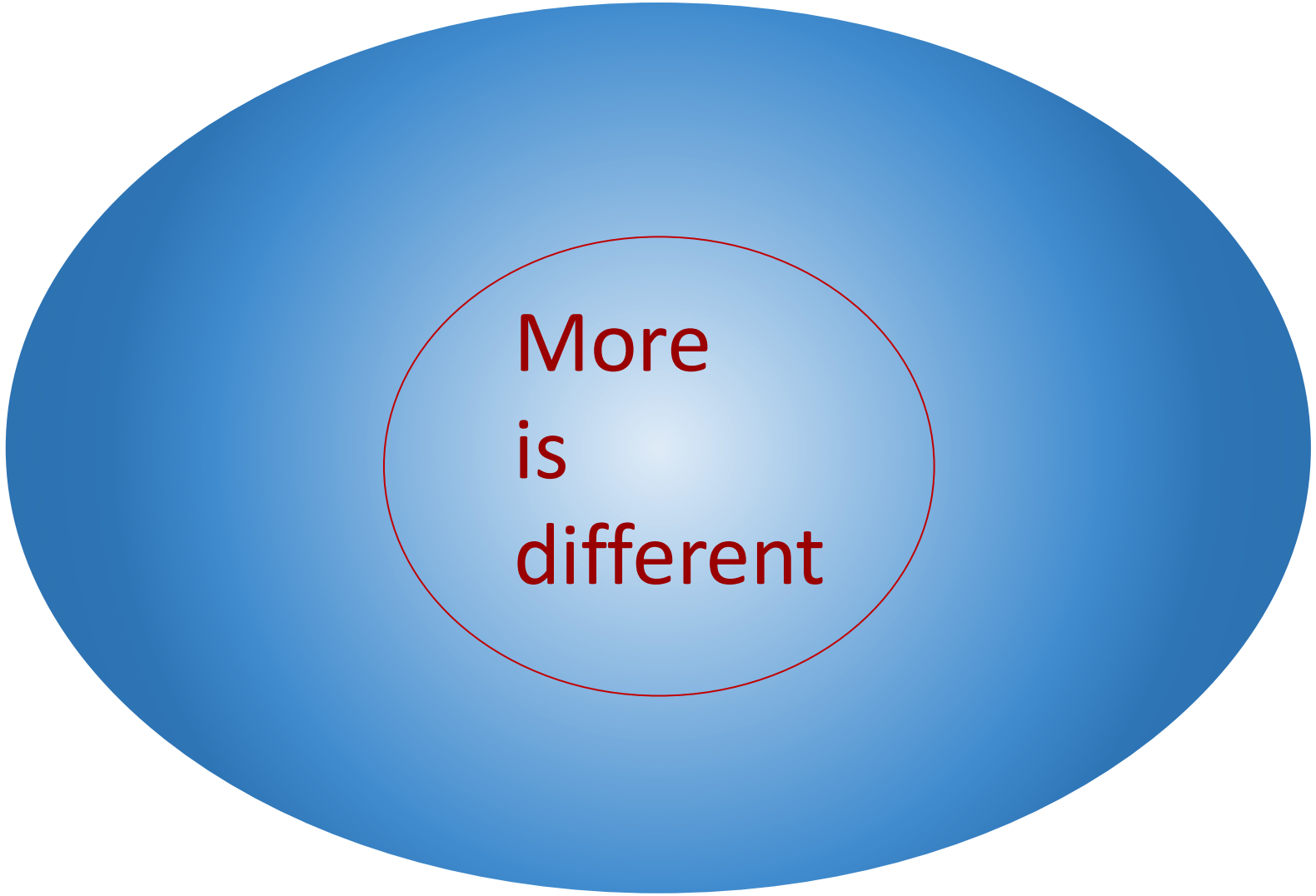


凝聚态物理（AI时代）



✓改进计算方法
✓提升求解能力
✓提升数据分析能力

AI与凝聚态物理



黄昆



“一是要学习知识，二是要创造知识，归根结底在于创造知识”

“学习知识不是越多越好，越深越好，而是应当与自己驾驭知识的能力相匹配。”

课后作业（5分）

作业：与AI聊一聊“AI与凝聚态物理”

具体要求及提交方式见课程网站说明