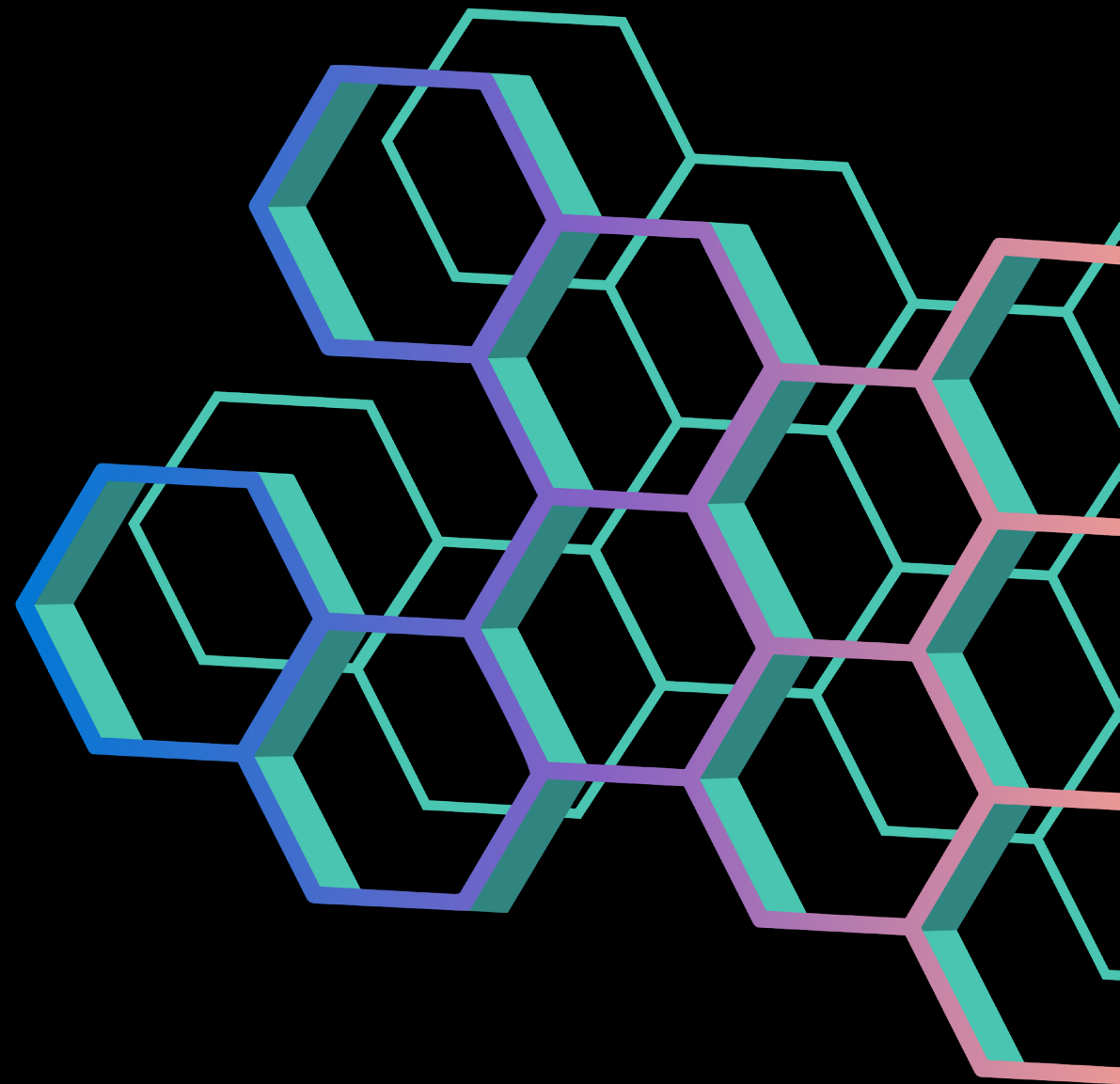


AI for Science

Prof Chris Bishop FREng FRSE FRS
Technical Fellow, *Microsoft Research AI for Science*



The greatest scientific discovery ever made

$$\left(-\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_{j \neq i}^N \sum_{i=1}^N \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^N \sum_{\alpha=1}^K \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \right) \Psi = E\Psi$$

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i$$

$$m_i \frac{d\mathbf{v}_i}{dt} = \mathbf{F}_i$$

$$F = G \frac{m_1 m_2}{r^2}$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right)$$

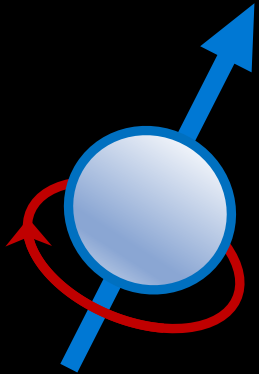
Magnetic dipole moment of the electron

Theory:

$$g/2 = 1.001\,159\,652\,181\,13 (\pm 86)$$

Experiment:

$$g/2 = 1.001\,159\,652\,180\,73 (\pm 28)$$



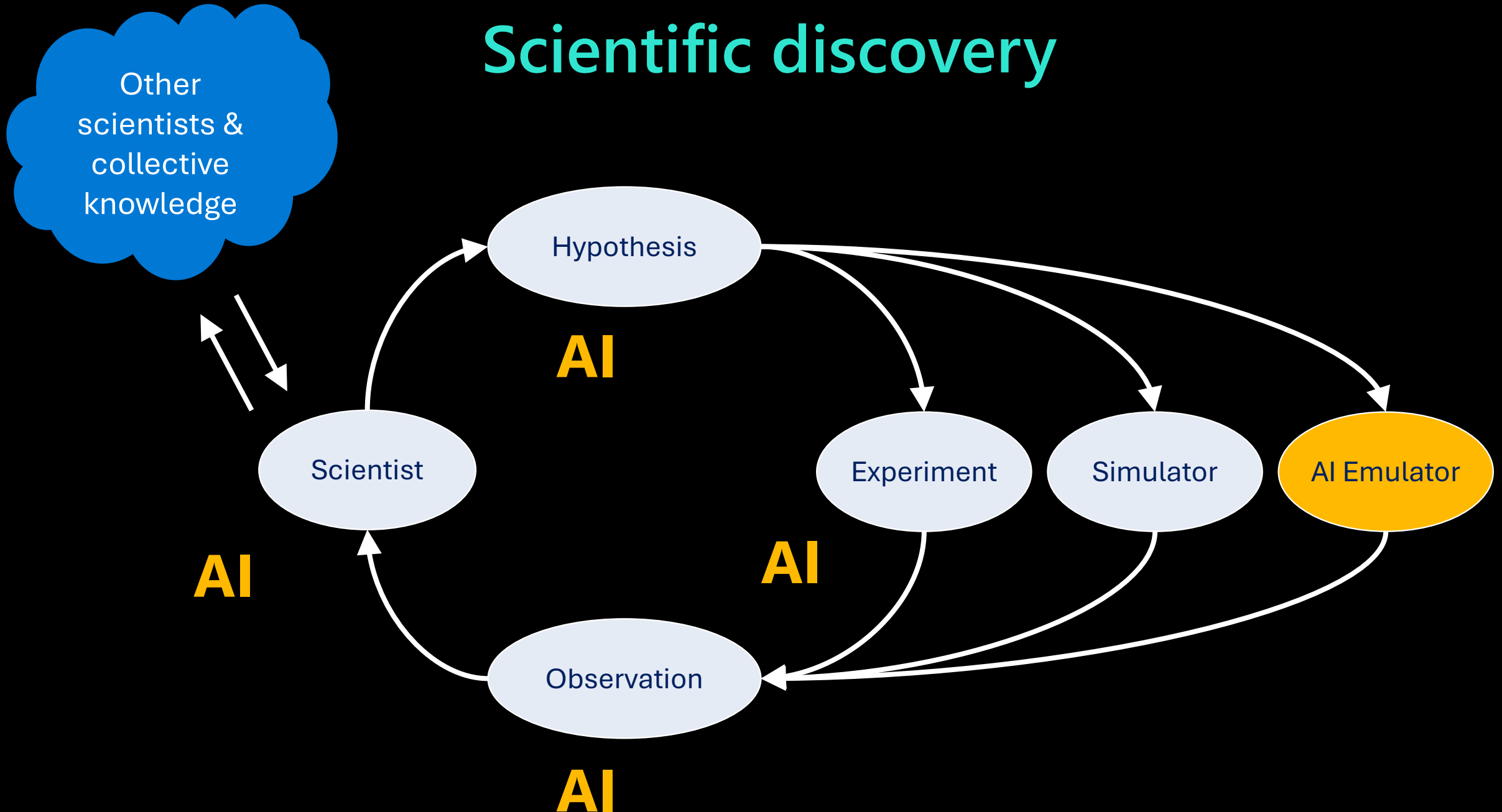


The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known.

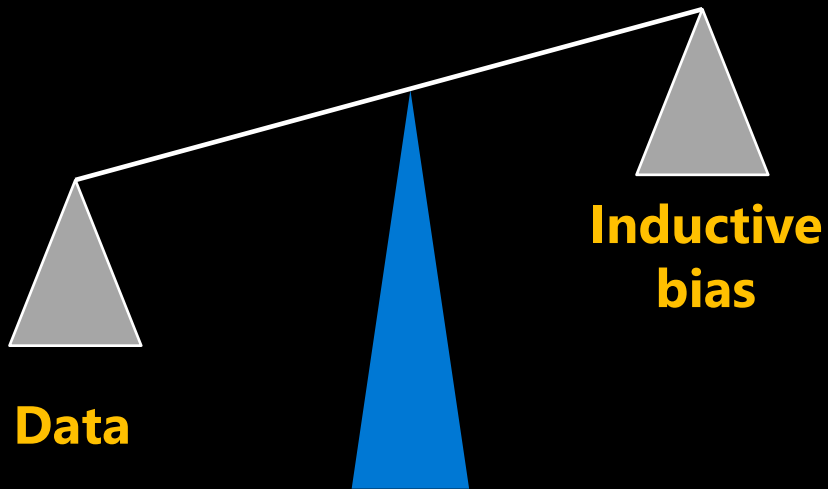
The difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

Paul Dirac, 1929

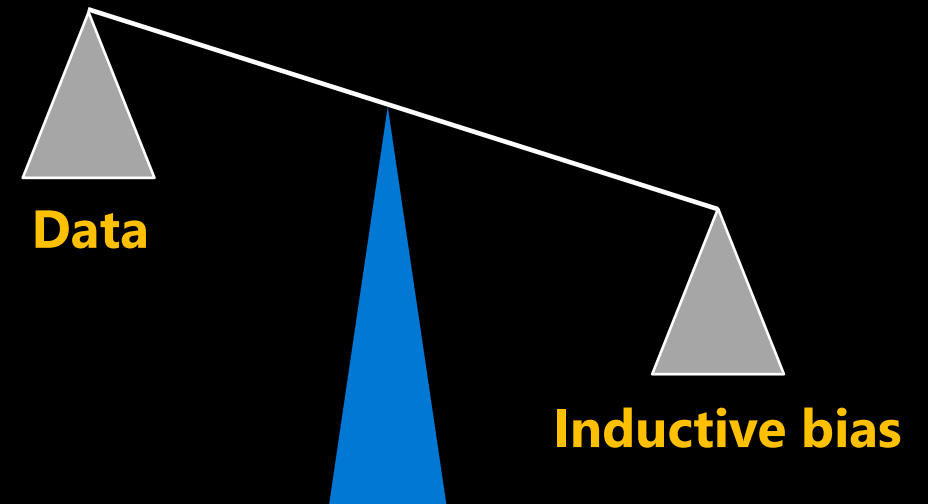
Scientific discovery



The 'no-free-lunch' theorem



Large language models

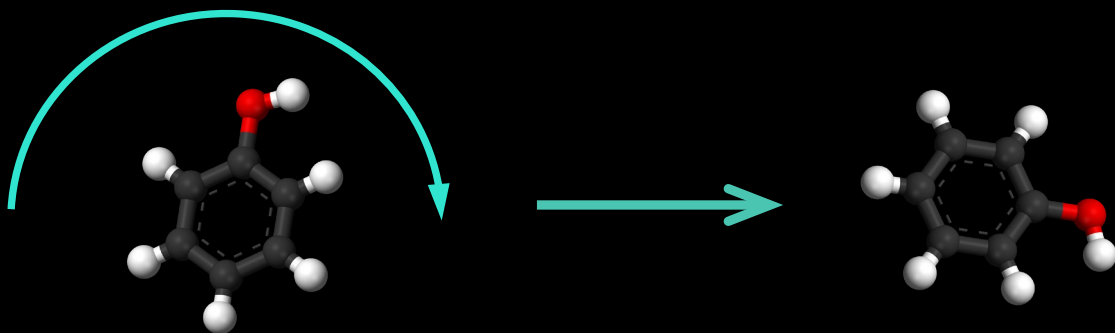


Scientific AI models

Inductive bias

$$\left(-\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_{j \neq i}^N \sum_{i=1}^N \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^N \sum_{\alpha=1}^K \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \right) \Psi = E\Psi$$

Invariance, equivariance



Geometrical deep learning

Conservation laws:

Energy
Mass
Momentum
Electric charge
...

The Bitter Lesson

Rich Sutton

March 13, 2019

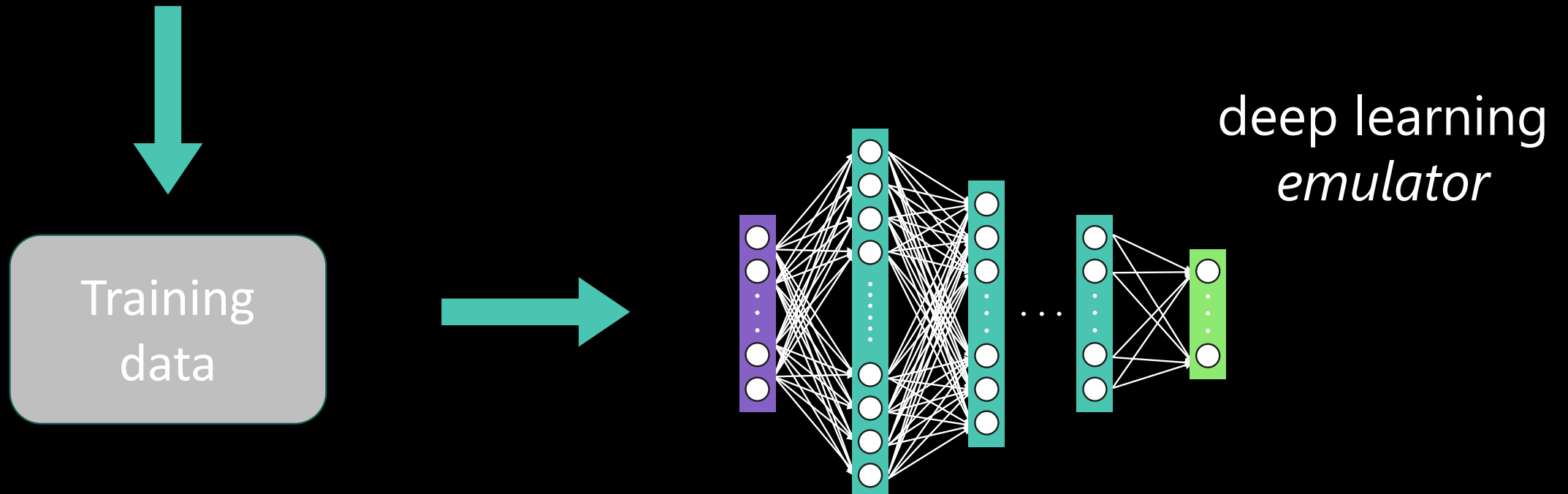
The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin. The ultimate reason for this is Moore's law, or rather its generalization of continued exponentially falling cost per unit of computation. Most AI research has been conducted as if the computation available to the agent were constant (in which case leveraging human knowledge would be one of the only ways to improve performance) but, over a slightly longer time than a typical research project, massively more computation inevitably becomes available. Seeking an improvement that makes a difference in the shorter term, researchers seek to leverage their human knowledge of the domain, but the only thing that matters in the long run is the leveraging of computation. These two need not run counter to each other, but in practice they tend to. Time spent on one is time not spent on the other. There are psychological commitments to investment in one approach or the other. And the human-knowledge approach tends to complicate methods in ways that make them less suited to taking advantage of general methods leveraging computation. There were many examples of AI researchers' belated learning of this bitter lesson, and it is instructive to review some of the most prominent.

In computer chess, the methods that defeated the world champion, Kasparov, in 1997, were based on massive, deep search. At the time, this was looked upon with dismay by the majority of computer-chess researchers who had pursued methods that leveraged human understanding of the special structure of chess. When a simpler, search-based approach with special hardware and software proved vastly more effective, these human-knowledge-based chess researchers were not good losers. They said that "brute force" search may have won this time, but it was not a general strategy, and anyway it was not how people played chess. These researchers wanted methods based on human input to win and were disappointed when they did not.

A similar pattern of research progress was seen in computer Go, only delayed by a further 20 years. Enormous initial efforts went into avoiding search by taking advantage of human knowledge, or of the special features of the game, but all those efforts proved irrelevant, or worse, once search was applied effectively at scale. Also important was the use of learning by self-play to learn a value function (as it was in many other games and even in chess

Deep learning emulators

$$\left(-\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{2} \sum_{j \neq i}^N \sum_{i=1}^N \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^N \sum_{\alpha=1}^K \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \right) \Psi = E\Psi$$



Benefits of emulators

Train on *synthetic* data generated by simulation

Data is richly and perfectly labelled

No GDPR or privacy issues

Quantity limited “only” by compute budget

Exploitation of under-utilized compute resources

Massive acceleration compared to simulation $\approx 10^3 - 10^{4+}$



A foundation model for the Earth system

<https://doi.org/10.1038/s41586-025-09005-y>

Received: 28 May 2024

Accepted: 9 April 2025

Published online: 21 May 2025

Open access

 Check for updates

Cristian Bodnar^{1,2,11}, Wessel P. Bruinsma^{1,11}, Ana Lucic^{1,3,11}, Megan Stanley^{1,11}, Anna Allen^{4,11}, Johannes Brandstetter^{1,5}, Patrick Garvan¹, Maik Riechert¹, Jonathan A. Weyn⁶, Haiyu Dong⁶, Jayesh K. Gupta^{2,7}, Kit Thambiratnam⁶, Alexander T. Archibald⁴, Chun-Chieh Wu⁸, Elizabeth Heider¹, Max Welling^{1,3}, Richard E. Turner^{1,4,9} & Paris Perdikaris^{1,10}✉

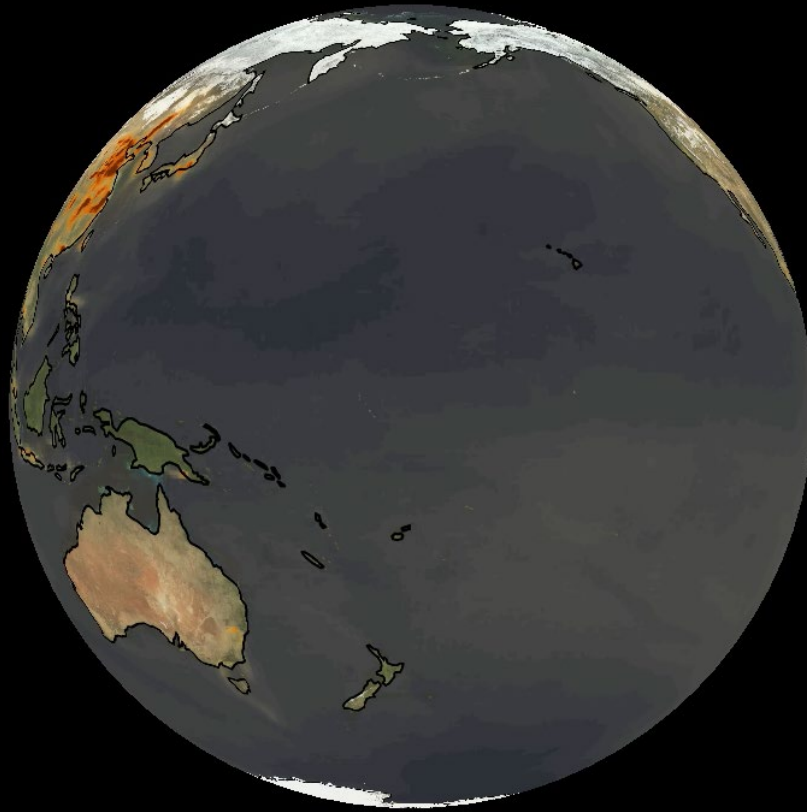
Reliable forecasting of the Earth system is essential for mitigating natural disasters and supporting human progress. Traditional numerical models, although powerful, are extremely computationally expensive¹. Recent advances in artificial intelligence (AI) have shown promise in improving both predictive performance and efficiency^{2,3}, yet their potential remains underexplored in many Earth system domains. Here we introduce Aurora, a large-scale foundation model trained on more than one million hours of diverse geophysical data. Aurora outperforms operational forecasts in predicting air quality, ocean waves, tropical cyclone tracks and high-resolution weather, all at orders of magnitude lower computational cost. With the ability to be fine-tuned for diverse applications at modest expense, Aurora represents a notable step towards democratizing accurate and efficient Earth system predictions. These results highlight the transformative potential of AI in environmental forecasting and pave the way for broader accessibility to high-quality climate and weather information.

Earth system forecasts are indispensable tools for human societies, as evidenced by recent natural events such as the floods in Valencia, the air quality crisis in New Delhi and hurricanes Helene and Milton in the eastern United States. Such systems not only provide crucial early warnings for extreme events but are also important for diverse applications ranging from agriculture to healthcare to global commerce. Modern Earth system predictions rely on complex models developed using centuries of accumulated physical knowledge, providing global forecasts of diverse variables for weather, air quality, ocean currents, sea ice and hurricanes.

medium-range weather forecasting at 0.25° resolution, leaving substantial gaps in other essential areas, including ocean dynamics, wave modelling and atmospheric chemistry. Furthermore, the potential for machine learning to outperform complex extreme weather prediction systems, which at present rely on human analysis of a wide range of models, remains underexplored.

In this paper, we introduce Aurora, a foundation model for the Earth system, capable of tackling a variety of forecasting tasks. Taking inspiration from recent successes of foundation models in other fields^{4,5}, we first pretrain Aurora on more than one million hours of diverse Earth

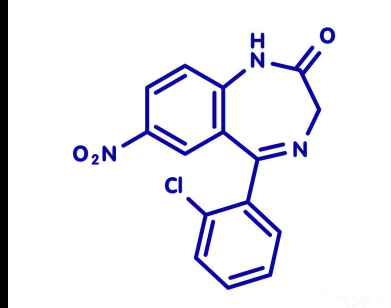
Aurora



Molecules and materials



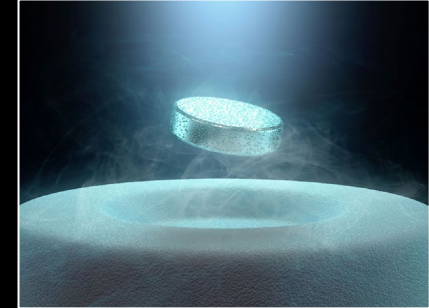
Semiconductors



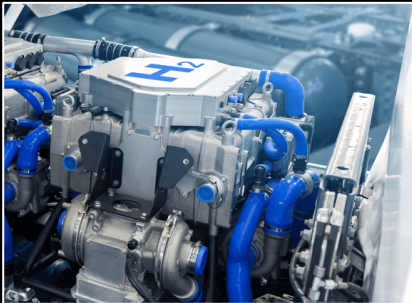
Drugs



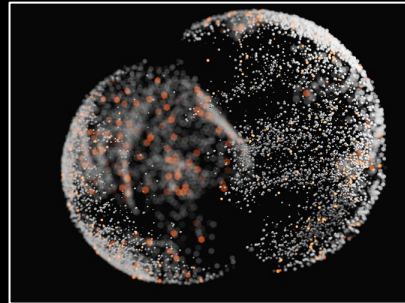
CO₂ capture



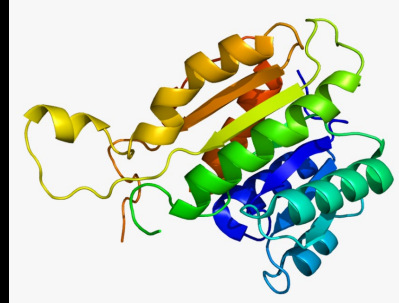
Superconductors



Energy storage



Catalysts

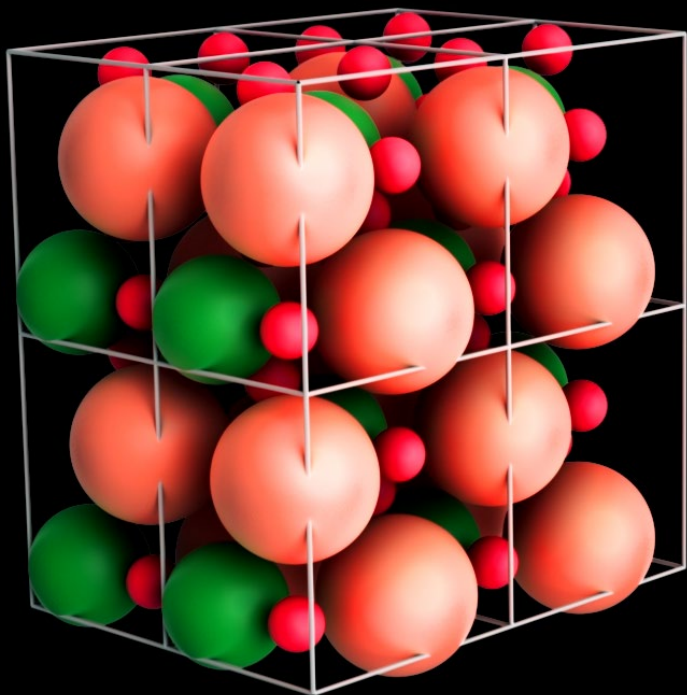


Proteins



Batteries

MatterGen



nature

Article

A generative model for inorganic materials design

<https://doi.org/10.1038/s41586-025-08628-5>

Received: 17 January 2024

Accepted: 10 January 2025

Published online: 16 January 2025

Open access

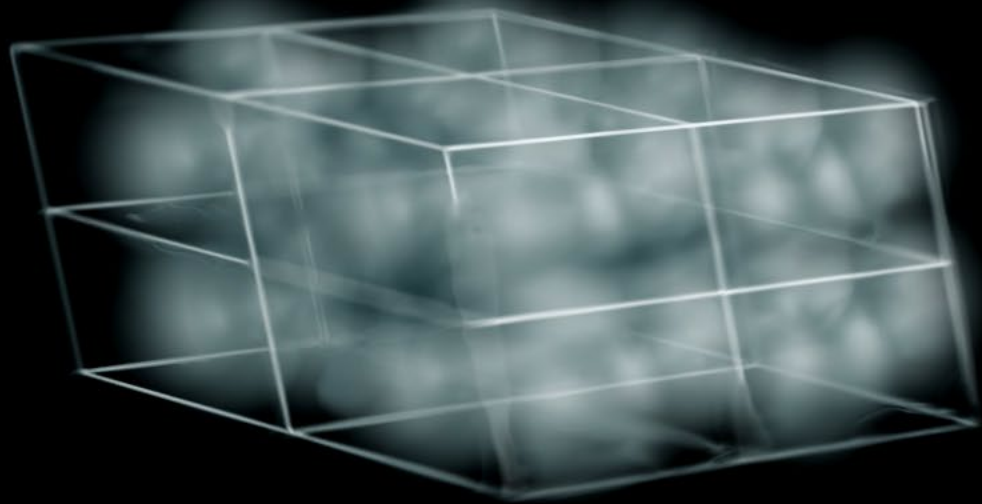
 Check for updates

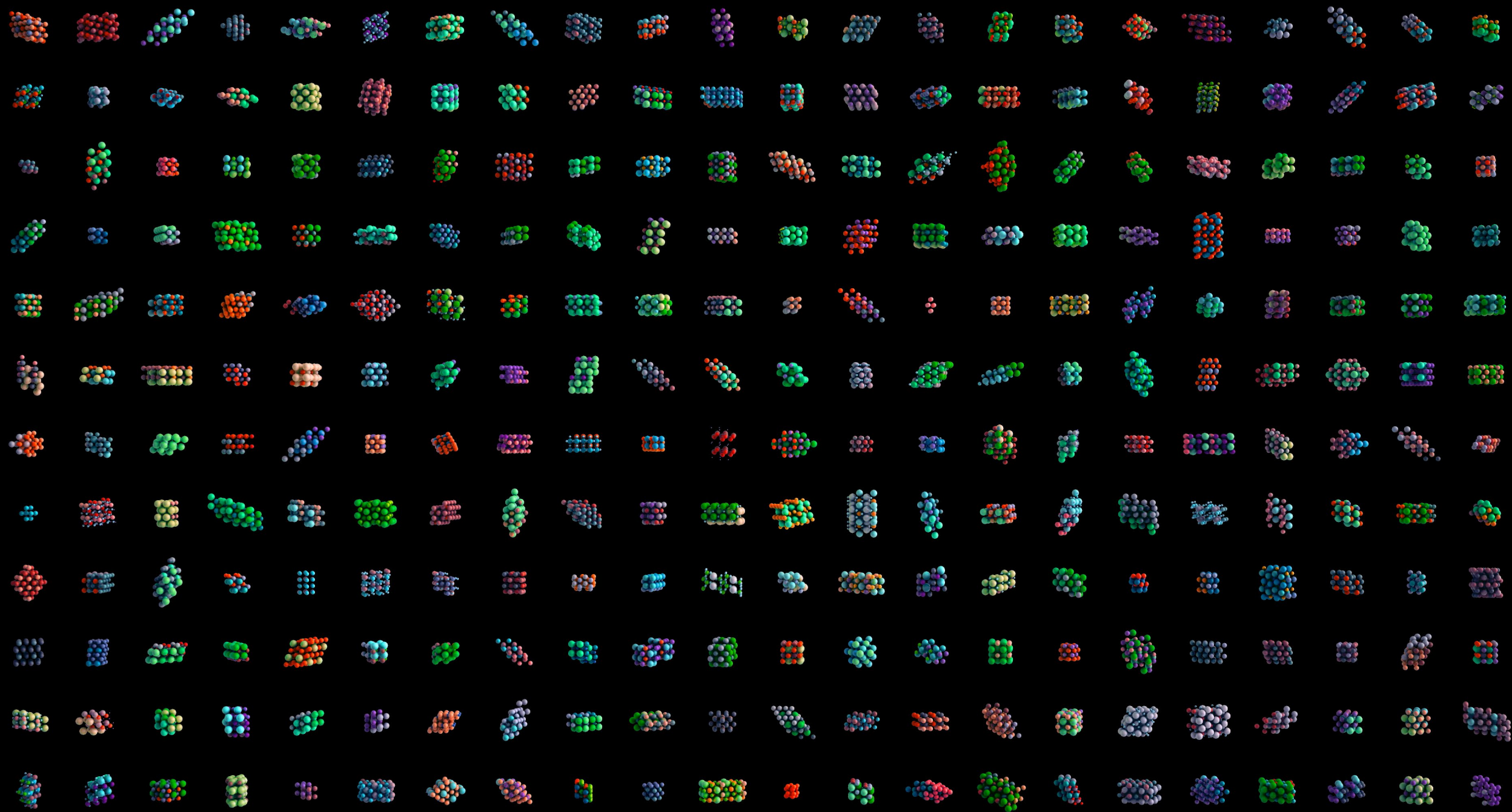
Claudio Zeni^{1,8}, Robert Pinsler^{1,8}, Daniel Zügner^{2,8}, Andrew Fowler^{1,8}, Matthew Horton^{3,8}, Xiang Fu¹, Zilong Wang⁴, Aliaksandra Shysheya¹, Jonathan Crabbé¹, Shoko Ueda¹, Roberto Sordillo¹, Lixin Sun¹, Jake Smith³, Bichlien Nguyen³, Hannes Schulz², Sarah Lewis¹, Chin-Wei Huang⁵, Ziheng Lu⁶, Yichi Zhou⁷, Han Yang⁶, Hongxia Hao⁶, Jielan Li⁶, Chunlei Yang⁴, Wenjie Li⁴, Ryota Tomioka^{1,8}✉ & Tian Xie^{1,8}✉

The design of functional materials with desired properties is essential in driving technological advances in areas such as energy storage, catalysis and carbon capture^{1–3}. Generative models accelerate materials design by directly generating new materials given desired property constraints, but current methods have a low success rate in proposing stable crystals or can satisfy only a limited set of property constraints^{4–11}. Here we present MatterGen, a model that generates stable, diverse inorganic materials across the periodic table and can further be fine-tuned to steer the generation towards a broad range of property constraints. Compared with previous generative models^{4,12}, structures produced by MatterGen are more than twice as likely to be new and stable, and more than ten times closer to the local energy minimum. After fine-tuning, MatterGen successfully generates stable, new materials with desired chemistry, symmetry and mechanical, electronic and magnetic properties. As a proof of concept, we synthesize one of the generated structures and measure its property value to be within 20% of our target. We believe that the quality of generated materials and the breadth of abilities of MatterGen represent an important advancement towards creating a foundational generative model for materials design.

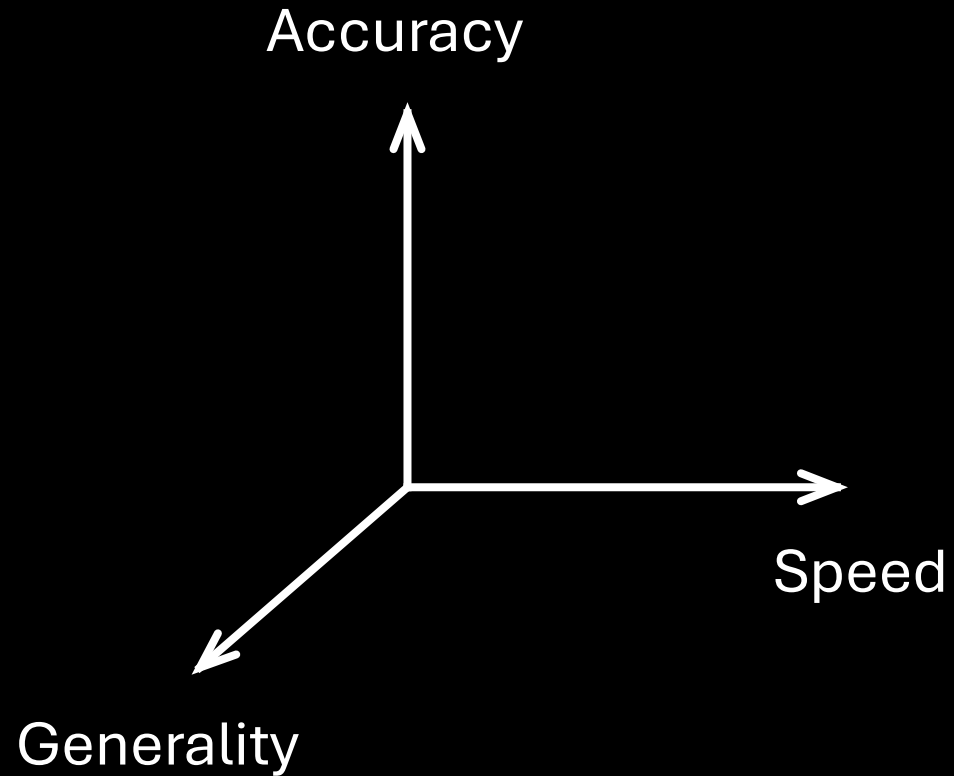
The rate at which we can discover better materials has a substantial impact on the pace of technological innovation in areas such as carbon capture, semiconductor design and energy storage^{1–3}. Traditionally, most materials have been discovered through experi-

fall short of producing stable materials according to density functional theory (DFT) calculations^{4,5,31}, are constrained by a narrow subset of elements^{7,9} and/or can only optimize a very limited set of properties, mainly formation energy^{4,5,8,11,31,32}.





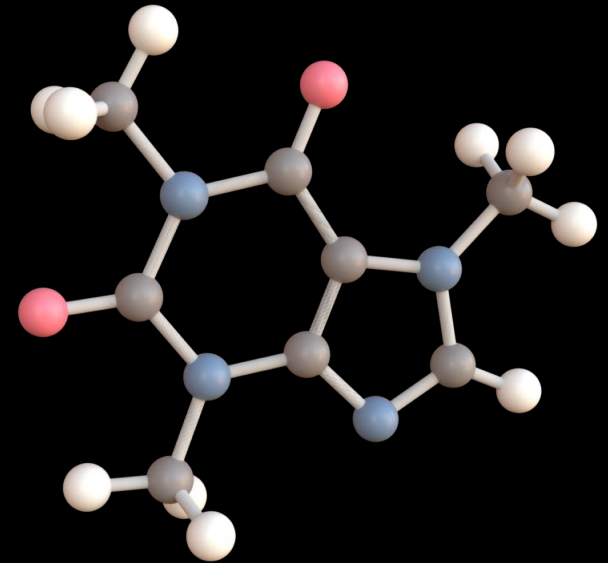
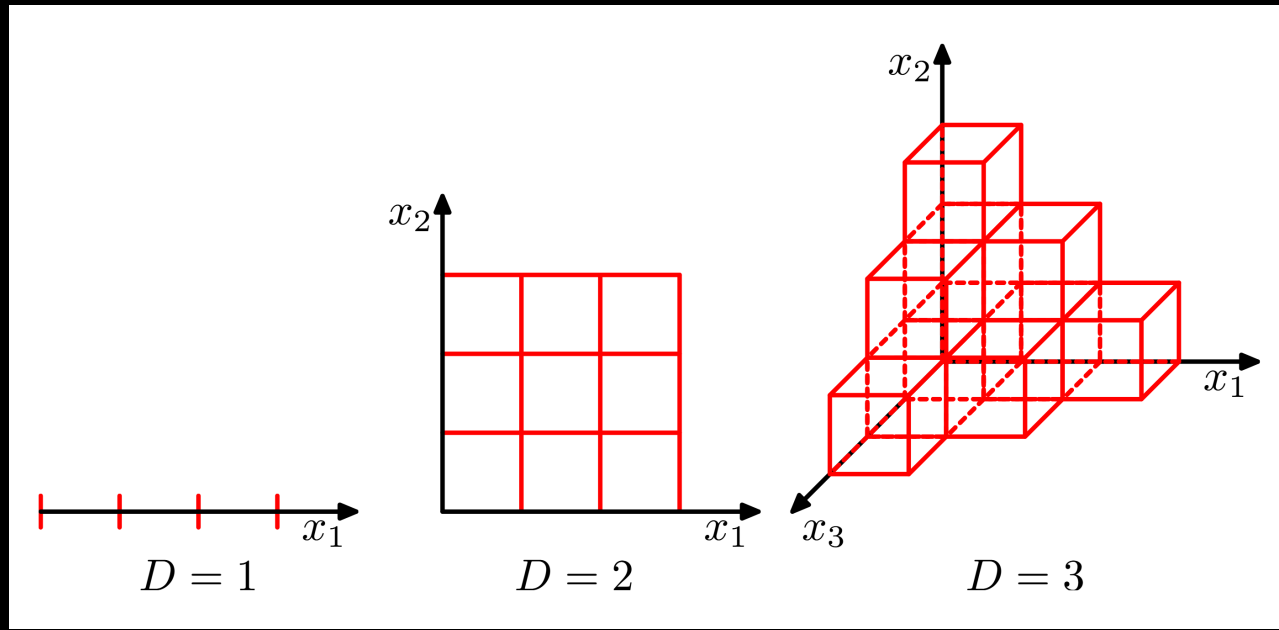
Fundamental trilemma



Chemical accuracy: 1kcal/mol

Wavefunction

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N)$$



DFT: density functional theory (1965)

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) \rightarrow \rho(\mathbf{x})$$

Computational cost $\exp(N)$

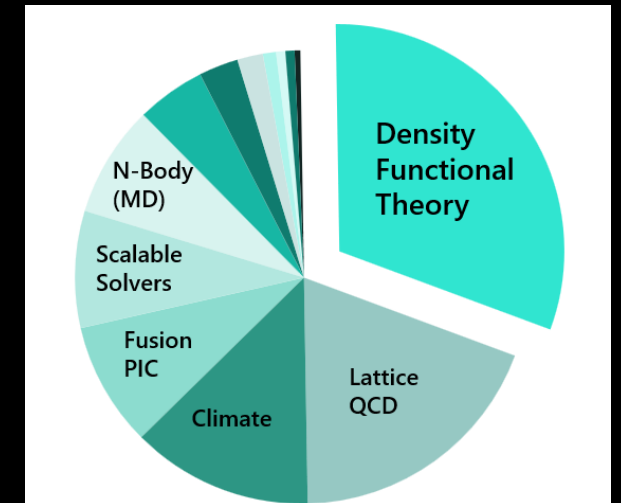
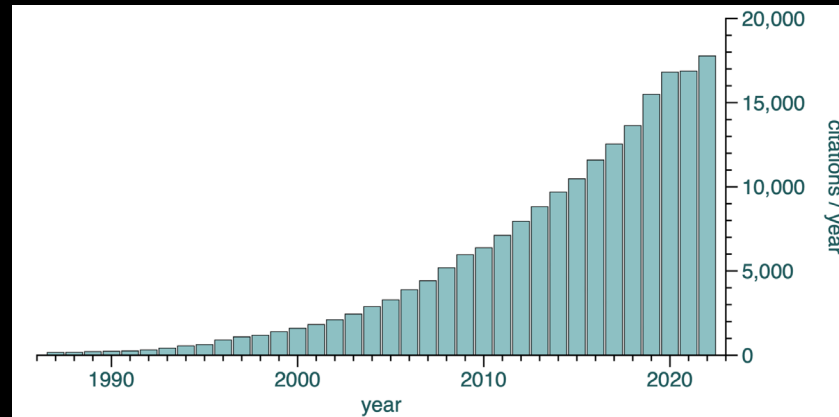


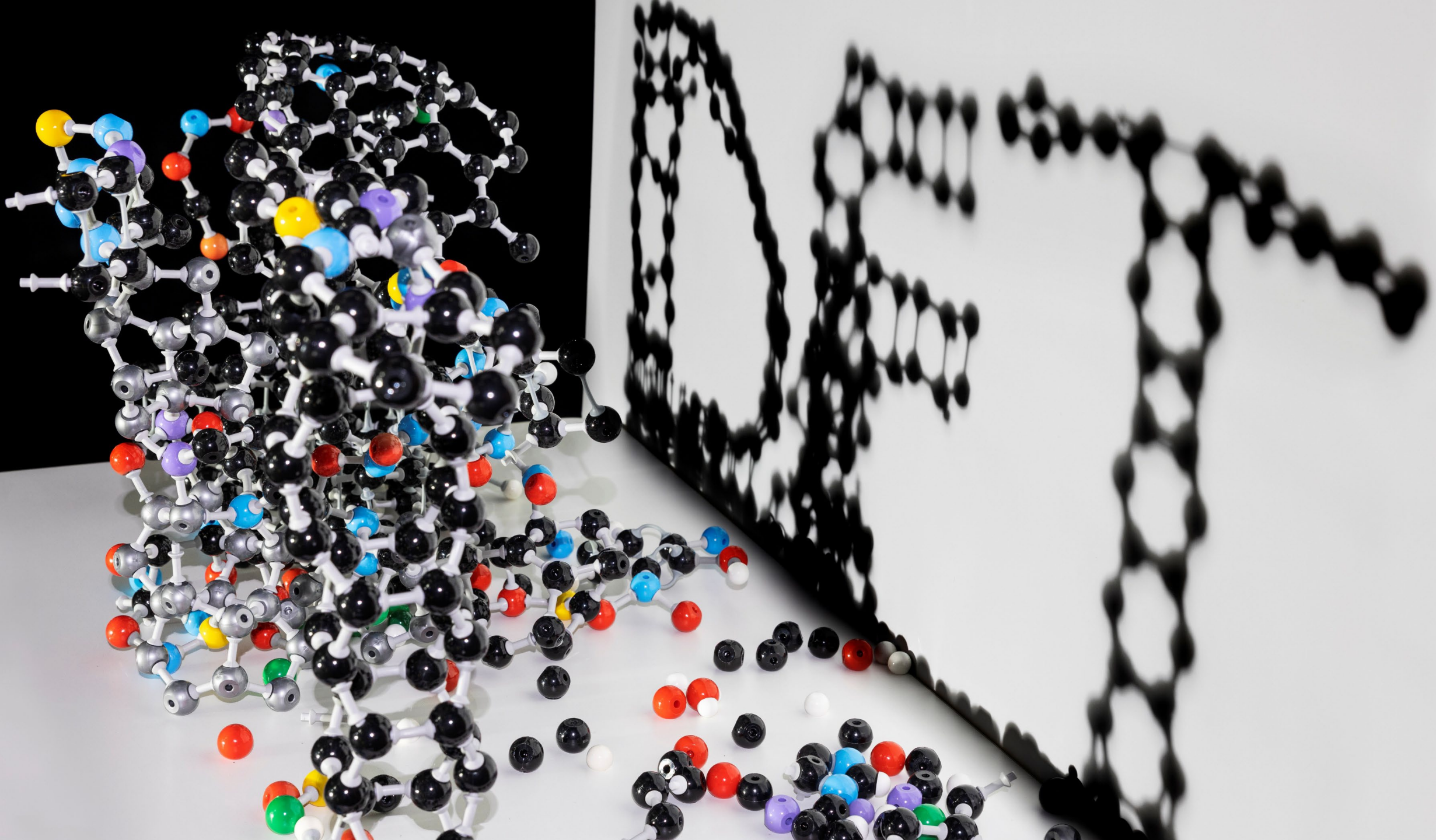
Computational cost N^3

Exact!



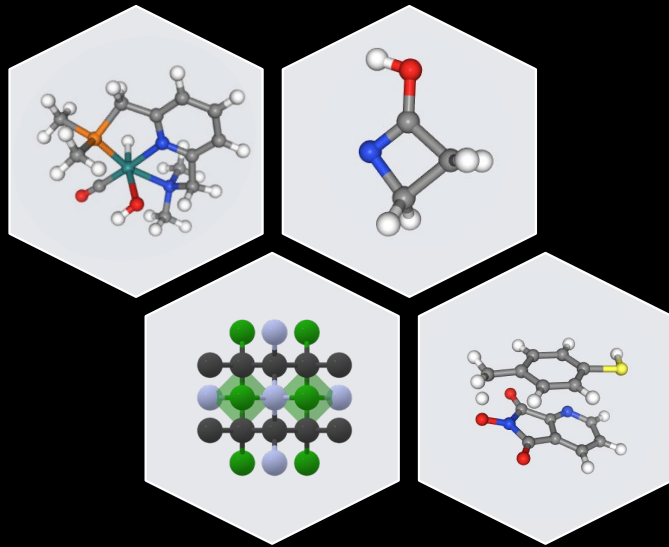
Photo from the Nobel
Foundation archive.
Walter Kohn



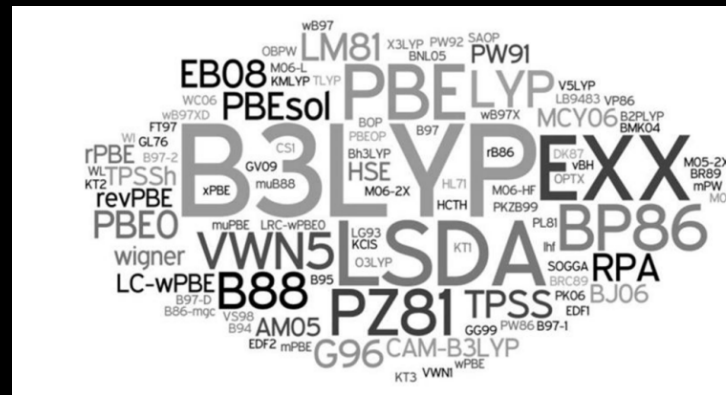


What is the catch?

Exchange-correlation functional $E_{xc}[\rho]$



It's universal!



Handcrafted functionals with limited accuracy – “Jacob’s ladder”

Science

Current Issue First release papers Archive About ▾

HOME > SCIENCE > VOL. 298, NO. 5594 > IN PURSUIT OF THE "DIVINE" FUNCTIONAL

| PERSPECTIVES

f t in

In Pursuit of the "Divine" Functional

ANN E. MATTSSON [Authors Info & Affiliations](#)

SCIENCE • 25 Oct 2002 • Vol 298, Issue 5594 • pp. 759-760 • DOI:10.1126/science.1077710

187

CH

SCIENCE'S COMPASS

PERSPECTIVES

PERSPECTIVES: DENSITY FUNCTIONAL THEORY



In Pursuit of the "Divine" Functional

Ann E. Mattsson

Paúl Dirac reputedly said that the Schrödinger equation (SE) marked the end of chemistry: All answers could be calculated from the SE. The SE can indeed be solved exactly for some systems, such as the hydrogen atom, and numerically to arbitrary precision for systems with a small number of electrons. However, the many-body character of the SE sets a fundamental limit on the size of solvable systems (*1*).

In a seminal paper, Hohenberg and Kohn (*2*) showed in 1964 that the SE, formulated as an equation of an *N*-electron wave function of $3N$ variables, could be reformulated as an equation of the electron

lated exactly, just as the properties of small systems can be calculated from the SE.

The first step toward the divine functional was the local density approximation (LDA), where only the electron density at a point in the system is used for determining this point's contribution to the total exchange-correlation energy of the system (*4, 6, 7*). LDA assumes a homogeneous electron gas but also works amazingly well for inhomogeneous systems. It is still used extensively.

Attempts to determine why the LDA

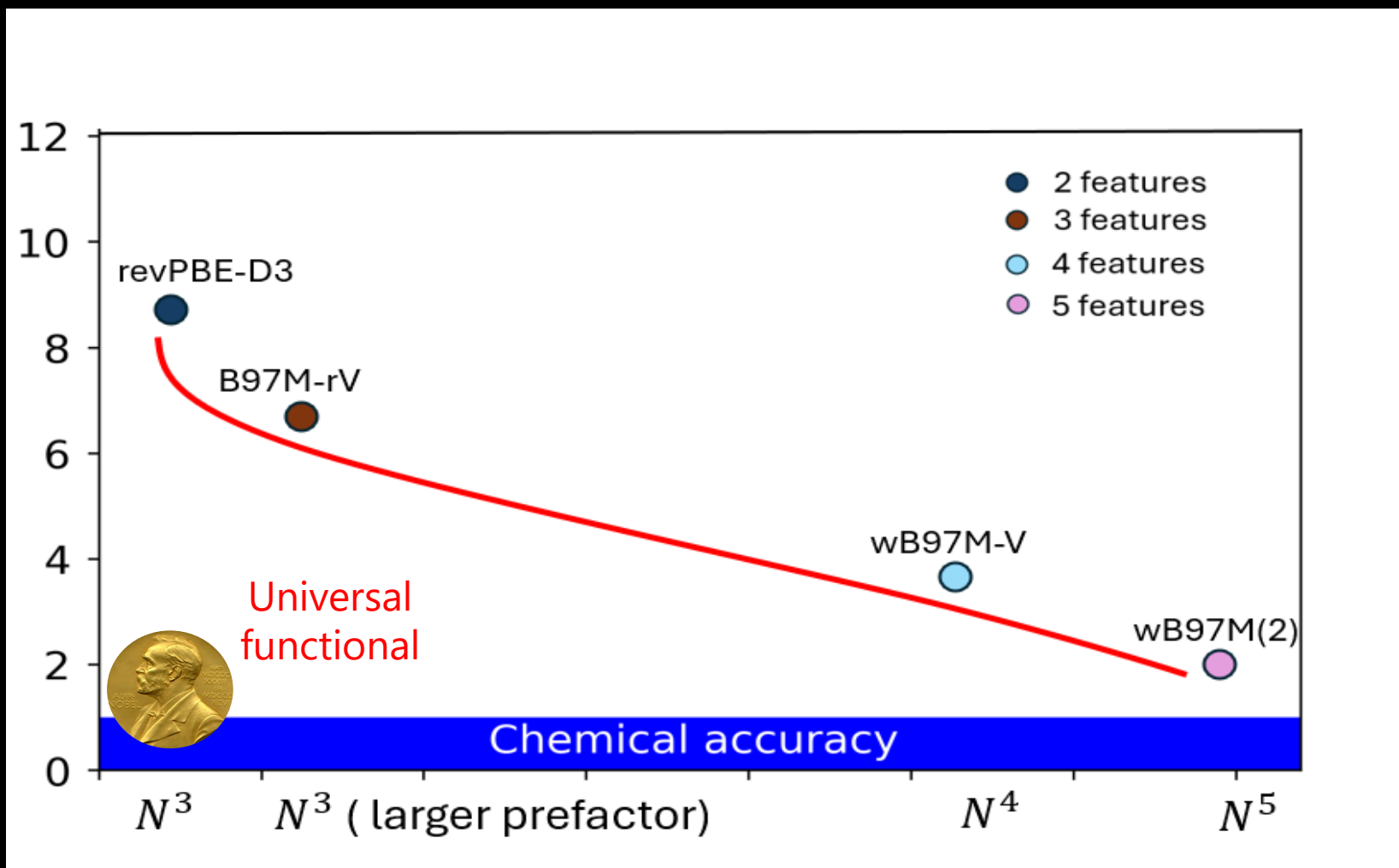
sively used in gas-phase catalysis, where the metallic character of the catalyst makes traditional SE methods inappropriate, and for determining phase diagrams (e.g., for steels).

In recognition of the importance of DFT, Walter Kohn was awarded half of the 1998 Nobel Prize in Chemistry (the other half was awarded to John Pople) (*10*). Today numerous approximate functionals are in use (*11*)—a sign of the method's utility but also an indication that none is suitable for all systems. As the complexity of systems investigated by DFT grows (see the figure), the task of choosing the right functional becomes increasingly difficult.

Although DFT is an invaluable tool in investigations of complicated systems, we cannot yet rely solely on this tool but need to combine it with, for example, some experimental data. Furthermore, recent studies involving surfaces have revealed unexpectedly poor and inconsistent energy results (*12, 13*). Errors are much harder to estimate for DFT than for SE contributions, making it difficult to know when DFT is likely not to



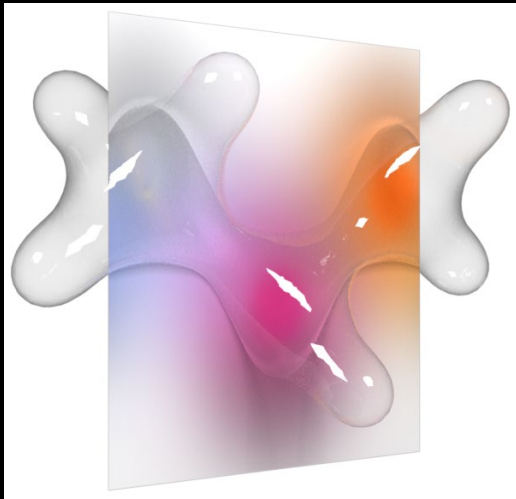
The DFT zoo



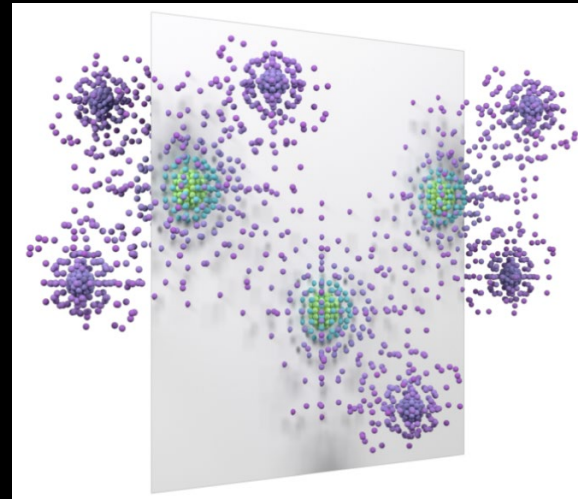


Skala functional

[arXiv:2506.14665](https://arxiv.org/abs/2506.14665)



Electron density



Non-locality through
message-passing



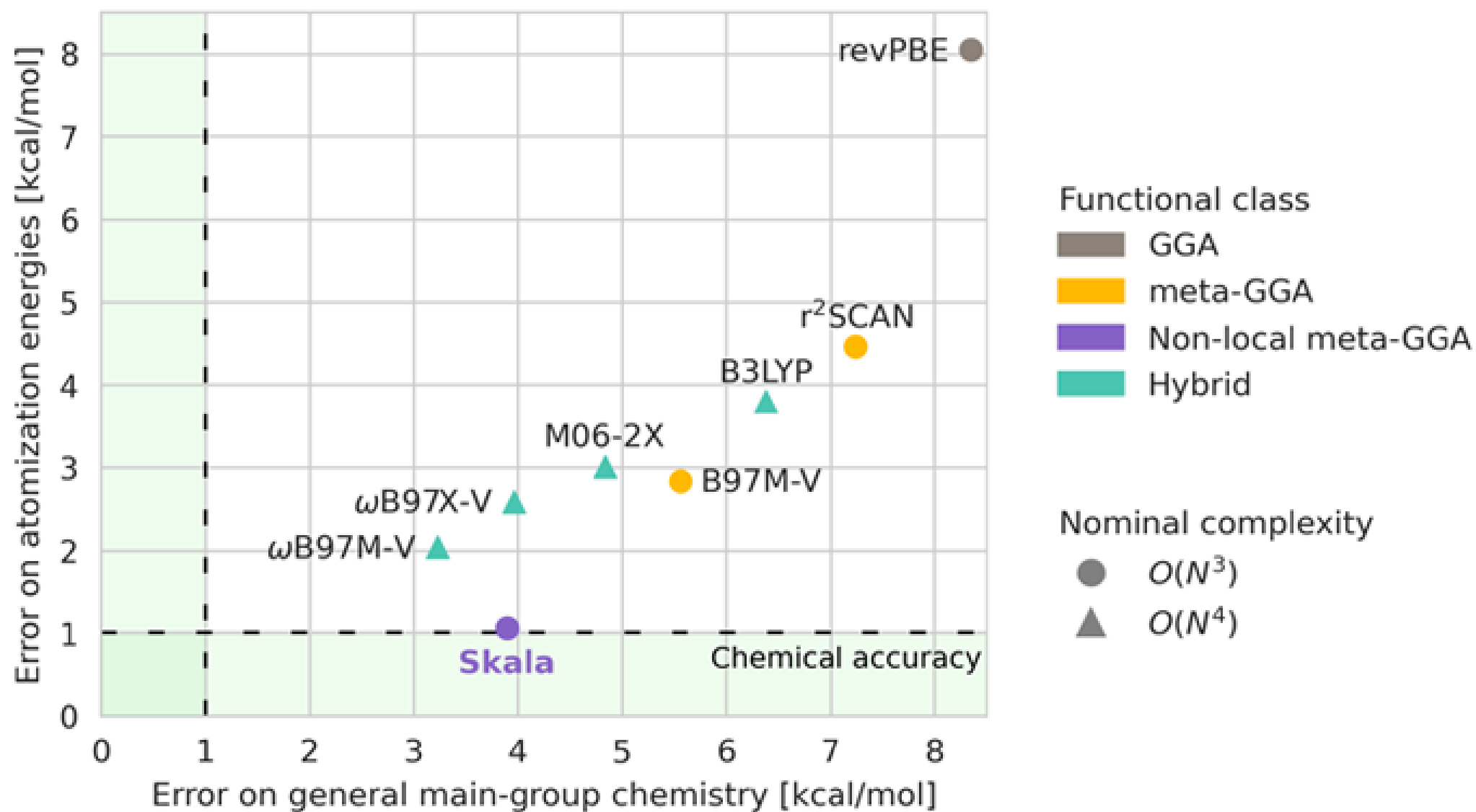
E_{xc}



Energy



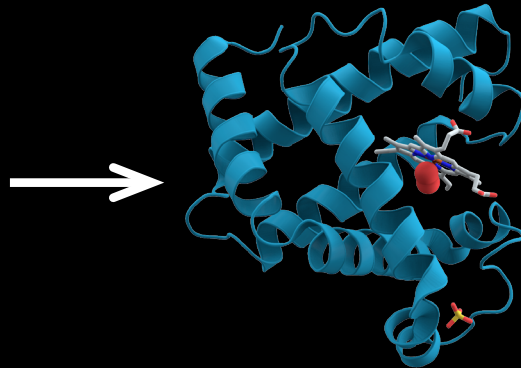
Hartree energy, KS kinetic energy, electron-nucleus Coulomb potential



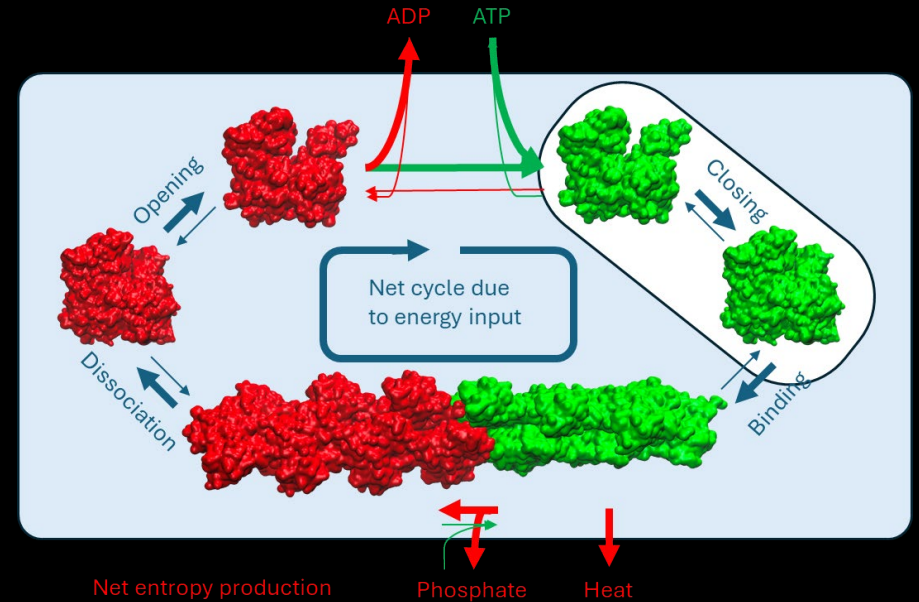
Molecular biology

```
12853800 TCCTTTCACACGCTTCCAAACCTACAGTTTCTCCACCTATATGTTTCAGgttcgtctttagttt
tcacttattgggtttcttctcaattgtgaacagAGTTTCAATTGGGAGTCATGGAAGAAAGAAGGAGGAT
12853600 ACATAGCCAAACGCTGGAAATCACTCATCTTTGGCTTCCTCCCTCTTCCTCAATCGTTGCTCTGAAAGgttc
taccactctgttactcagcaatcttcattctctcagGTTACTTACCGGGAAGCTATACGATCTAAACAG
12853400 AATCGTTAATCAAAGCGTTGAATCAAAAAGGAATAAAAGCTTTGGCTGATATAGTGATTAAACCACAGAAC
CTGTTATTTGCAAGGTGGGACTTCGATGATCGTCTTGATGGGATCCTTCCTTTGCTGCGCCAAATGAC
12853200 ACCGGAGGAGATTTTGTATGGAGCGCCGACATCGACACCTTAAACCTAGAGTTTCAGAAAGAGTTGTCCG
GATTCCTATGGTTGGAGATTTGATTATGTTGAGGTTATGCACTTTCATCACCAGAAATATACGTTCAAGgtt
12853000 aacagtatagtatataaagaacataggttgagataattattactattagtatataaagtatcagaggt
ataagaacataaagtcgaatgcaatcaataaagaatatataaagaagttcactactgattatgtgataaat
12852800 ACCGATTTTCCGCTGGCTGAGAAATGGGACGATATGAAGTACGGAGGAGACGGGAAACTAGACTATGAT
TGGATCGAGGAAGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
12852600 ACTCGGAGGAAACCGCTGGTATGATAGGAATCATGGCGGGAAGCGTGTCAATTCATAGATAGTAAAC
TGATAAAGTCTTGCCTTGGATACGTTTATATACCTTACTCATCCAGGAACCTCTGCAATTgaagtatcattt
12852400 aatcttgtgatagtattttttgttgcagTTTATAATCATTACATAGAAATGGGACTAAAGAGAGCA
ATGGGATTGGTAGCACAGCTCTGTAAACGATAAAGCGGAGAGCGGATCTCTACTTTGGCTATGATTGA
12852200 GCAAGATGGGGAACACTTGTTCCTTCAATTTTGTCTTAGCTTATTCAGGCCTTGACTTTGCTGTCTGG
agaaaagtaatcgaatgatcttcttcttttaataaaacattttggcagtatctaagatatgtataat
```

Sequence

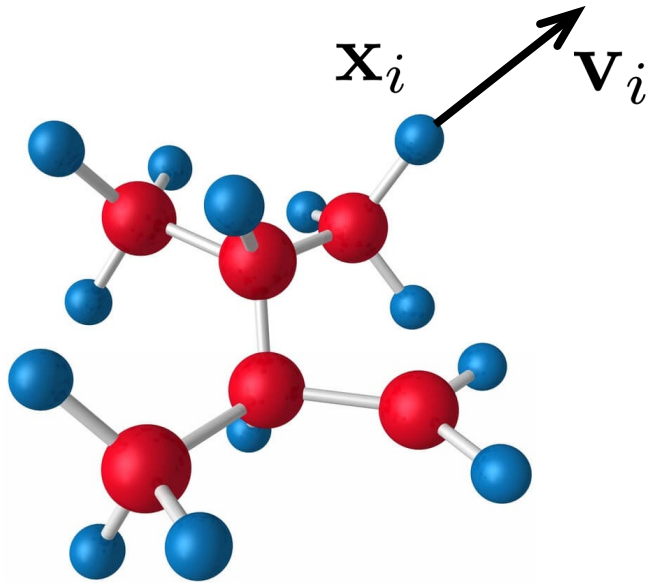


Structure



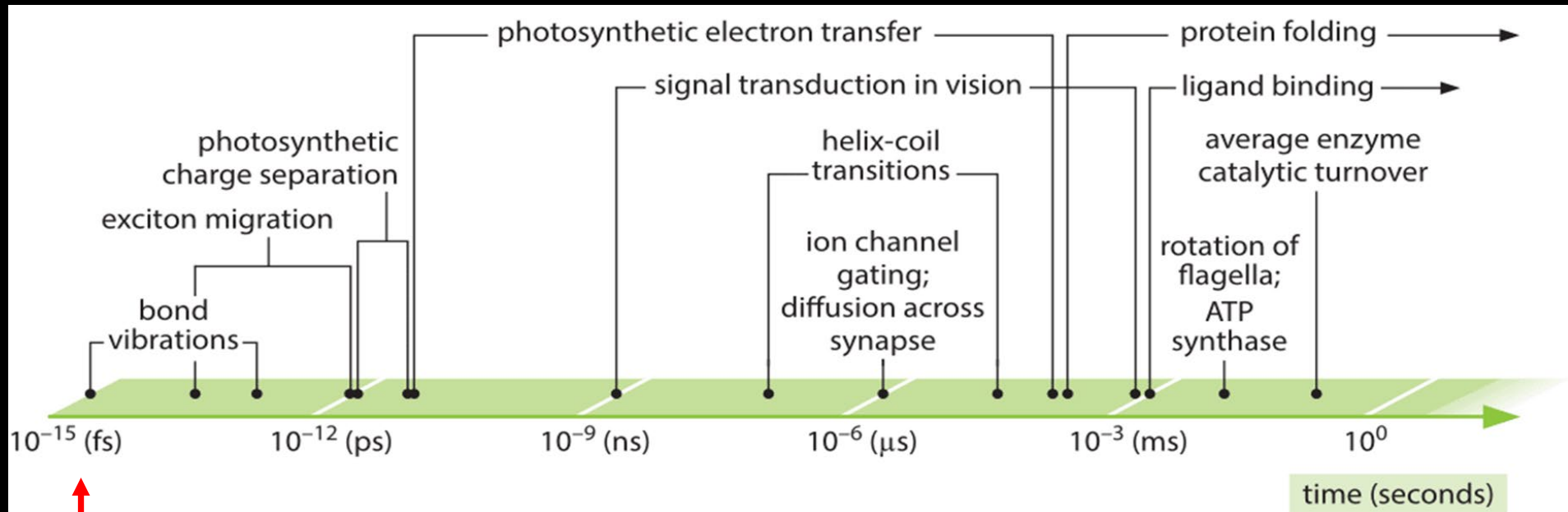
Function

Molecular dynamics (MD) simulation



$$\begin{aligned}\frac{d\mathbf{x}_i}{dt} &= \mathbf{v}_i \\ m_i \frac{d\mathbf{v}_i}{dt} &= -\nabla_i U(\mathbf{x}_1, \dots, \mathbf{x}_N)\end{aligned}$$

Curse of sequentiality

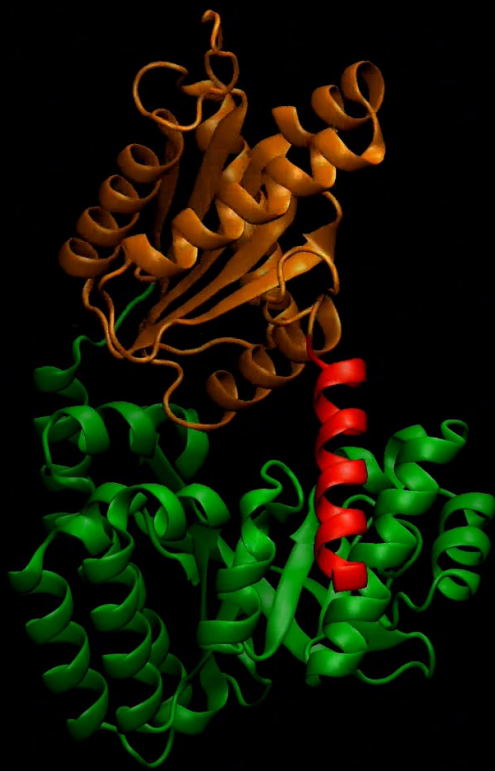


Integration step

desired simulation timespan

Simulations with current MD technology require 10^6 to 10^{15} sequential steps
But: chips no longer become faster for sequential computation

AI²BMD



nature

Article

Ab initio characterization of protein molecular dynamics with AI²BMD

<https://doi.org/10.1038/s41586-024-08127-z>

Received: 31 March 2023

Accepted: 26 September 2024

Published online: 6 November 2024

Open access

 Check for updates

Tong Wang^{1,2,3}✉, Xinheng He^{1,2}, Mingyu Li^{1,2}, Yatao Li^{1,2}, Ran Bi¹, Yusong Wang¹, Chaoran Cheng¹, Xiangzhen Shen¹, Jiawei Meng¹, He Zhang¹, Haiguang Liu¹, Zun Wang¹, Shaoning Li¹, Bin Shao^{1,3}✉ & Tie-Yan Liu¹

Biomolecular dynamics simulation is a fundamental technology for life sciences research, and its usefulness depends on its accuracy and efficiency^{1–3}. Classical molecular dynamics simulation is fast but lacks chemical accuracy^{4,5}. Quantum chemistry methods such as density functional theory can reach chemical accuracy but cannot scale to support large biomolecules⁶. Here we introduce an artificial intelligence-based ab initio biomolecular dynamics system (AI²BMD) that can efficiently simulate full-atom large biomolecules with ab initio accuracy. AI²BMD uses a protein fragmentation scheme and a machine learning force field⁷ to achieve generalizable ab initio accuracy for energy and force calculations for various proteins comprising more than 10,000 atoms. Compared to density functional theory, it reduces the computational time by several orders of magnitude. With several hundred nanoseconds of dynamics simulations, AI²BMD demonstrated its ability to efficiently explore the conformational space of peptides and proteins, deriving accurate^{3J} couplings that match nuclear magnetic resonance experiments, and showing protein folding and unfolding processes. Furthermore, AI²BMD enables precise free-energy calculations for protein folding, and the estimated thermodynamic properties are well aligned with experiments. AI²BMD could potentially complement wet-lab

AAAS

Science



PROTEIN SHAKE

Generative deep learning
powers emulation of protein
ensembles p. 700

14 AUGUST 2025

Can synthetic organs and
AI replace animal testing? p. 676

Plant volatiles enhance pathogen defense
through soil feedback pp. 680 & 698

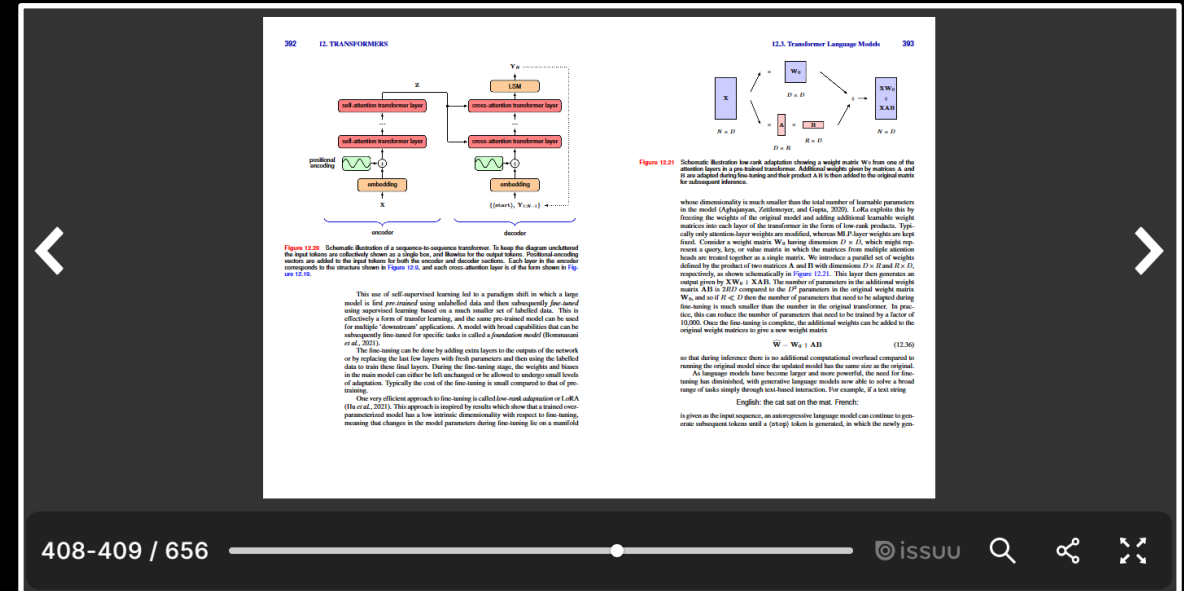
Diamond is not the only carbon allotrope
in rings p. 708

Christopher M. Bishop
with Hugh Bishop

Deep Learning

Foundations
and Concepts

 Springer



Hardback: 'stitched signatures'

Springer-Nature's best-selling book of 2024 and 2025!