frontiers | Frontiers in Physics

TYPE Perspective
PUBLISHED 08 October 2025
DOI 10.3389/fphy.2025.1641727

OPEN ACCESS

EDITED BY
Alex Hansen,
NTNU, Norway

REVIEWED BY
Raffaela Cabriolu,
Norwegian University of Science and Technology, Norway

*CORRESPONDENCE
Jean-Michel Tucny,
jeanmichel.tucny@uniroma3.it

RECEIVED 05 June 2025
ACCEPTED 16 September 2025
PUBLISHED 08 October 2025

CITATION
Tucny J-M, Durve M and Succi S (2025) Is the end of insight in sight?
Front. Phys. 13:1641727.
doi: 10.3389/fphy.2025.1641727

Is the end of insight in sight?

Jean-Michel Tucny[1]*, Mihir Durve[2] and Sauro Succi[2]

[1]Department of Civil, Computer Science, and Aeronautical Technologies Engineering, Università degli Studi Roma Tre, Rome, Italy, [2]Center for Life Nano- & Neuro-Science, Italian Institute of Technology (IIT), Rome, Italy

The rise of deep learning challenges the longstanding scientific ideal of insight—the human ability to understand phenomena by uncovering underlying mechanisms. From a physics perspective, we examine this tension through a case study: a physics-informed neural network (PINN) trained on rarefied gas dynamics governed by the Boltzmann equation. Despite strong physical constraints and a system with clear mechanistic structure, the trained network's weight distributions remain close to Gaussian, showing no coarse-grained signature of the underlying physics. This result contrasts with theoretical expectations that such networks might retain structural features akin to discrete dynamical systems. We argue that high predictive accuracy does not imply interpretable internal representations and that explainability in physics-informed AI may not always be achievable—or necessary. These findings highlight a growing divergence between models that predict well and those that offer insight.

KEYWORDS

explainable artificial intelligence (XAI), physics-informed neural networks (pinns), interpretability, Boltzmann equation, rarefied gas dynamics, machine learning, random matrix theory

1 Introduction

Recent advances in machine learning (ML), particularly through large language models (LLMs), have dramatically reshaped both science and society. These models now routinely tackle problems previously thought to be beyond reach, ranging from natural language understanding and protein folding to autonomous systems and symbolic reasoning [1–3]. Such progress introduces a fundamentally different approach to scientific discovery—one based not on physical insight into underlying mechanisms, but on data-driven optimization through a dense web of parameters. While physics-informed constraints can improve convergence [4], the learning process itself often remains opaque.

It no longer appears tenable to dismiss ML as a "glorified interpolator" or LLMs as "stochastic parrots" [5]. Instead, ML is beginning to challenge the very role of mechanistic understanding—or what has traditionally been called Insight—in scientific modeling. This tension raises the possibility of an "End of Insight" (EoI), a term coined by Strogatz [6], referring to the notion that certain scientific challenges may resist explanation in terms of simple governing principles, especially when they involve multiple interacting processes across vastly different scales in space and time.

This prospect is sad and perilous but plausible. Insight, as shaped by centuries of theory-driven physics, may not scale gracefully to problems such as epidemics, climate dynamics, or non-equilibrium statistical systems. ML, unconcerned with interpretability, may allow us to push the frontiers of knowledge in such domains—but without the perk of Insight and the intimate pleasure of "finding things out". This should not distract us from the fact that ML is still subject to a number of major limitations, especially when paired with the rising energy

cost of training ever-larger models, a trajectory that raises concerns about sustainability and rapidly diminishing returns [7].

In this paper, we contribute to this discussion through a focused case study: a physics-informed neural network (PINN) trained on a rarefied gas flow governed by the Boltzmann equation. The problem is highly structured, well understood, and modeled using both mechanistic equations and direct numerical simulation. Given these features—and the close connection between machine learning and discrete dynamical systems—we explore whether the network's parameters retain coarse signatures of the underlying physics. Rather than aiming to resolve interpretability, we use this controlled setting to test assumptions about what structured learning might look like when physical constraints are strongly present.

2 The basic ML procedure

The basic idea of ML lies in approximating a D-dimensional output y through recursive application of a nonlinear map [8]. For a neural network (NN) with input x, L hidden layers $z_1 \dots z_L$, each containing N neurons, and an output layer y, the update chain reads as in Equations 1–3:

$$z_0 = x \tag{1}$$

$$z_1 = f(W_1 x - b_1), \quad \dots, z_L = f(W_L z_{L-1} - b_L), \tag{2}$$

$$y = z_{L+1} = f(W_{L+1} z_l - b_{L+1}) \tag{3}$$

where W_l are $N \times N$ weight matrices, b_l are N-dimensional arrays of biases, and f is a nonlinear activation function. At each layer, the output is often normalized $\|z\| = 1$. The weights are updated via backpropagation, typically via a steepest descent as presented in Equation 4:

$$W' = W - \alpha \frac{\partial E}{\partial W} \tag{4}$$

where $E[W] = \|y_T - y\|^2$ is the loss function, W' are the updated weights and α is the learning rate.

2.1 Taming complexity

It is often claimed that, with enough data, ML can approximate virtually *any* target, whence the alleged demise of the scientific method [9, 10]. Put down in such bombastic terms, the idea is readily debunked by general considerations on the physics of complex systems, see for instance [11, 12]. Yet, ML does show remarkable proficiency in handling problems resistant to conventional modeling.

To understand why, we briefly examine the three main boosters of Complexity: Nonlinearity, Nonlocality and Hyper-Dimensionality.

2.1.1 Nonlinearity

Nonlinear systems exhibit two distinguishing and far-reaching features: i) they do not respond proportionally to input, and ii) they transfer energy (information) across scales. This makes them erratic and hard to predict, but also capable of emergent phenomena—complex behavior arising from simple rules, biology being a goldmine of such instances. While physics has developed mathematical tools to handle nonlinearity, these are often overwhelmed when couplings become too strong across vast scales, with weather forecasting being a prominent example. ML can definitely help such methods stretch their limits. However, at present, there is no clear evidence that it can systematically outperform them, especially when precision is in high demand, as is usually the case for scientific applications [13].

2.1.2 Nonlocality

In nonlocal systems, local behavior depends on distant states, often via long-range couplings. Although this interaction usually decays with the distance between the two regions, it cannot be ignored, no matter how far the interacting components are. A typical example from physics is classical gravitation, which is controlled by a potential decaying with the inverse power of the distance. The peculiarity of these systems is that they hardly reach a state of dynamic order known as "local equilibrium", usually controlled by a subset of "slow" variables living in a lower-dimensional manifold. Local equilibrium is the result of a neat scale separation between slow and fast variables, a feature which greatly simplifies the dynamics. Dynamics is notoriously much harder to capture than statistics and this is the reason why statistical physics is so effective in describing complex systems. With nonlocality in play, even statistical mechanics may remain hard to capture because of the aforementioned lack of scale separation between fast and slow modes. ML has shown promise in capturing such structures, particularly in identifying latent manifolds, though it remains an empirical rather than systematic approach [14].

2.1.3 Hyper-dimensionality

High-dimensional systems often suffer from the so-called curse of dimensionality (CoD), where the state space grows exponentially with the number of variables. Yet the real difficulty is subtler: due to nonlinearities, heterogeneities, and other structural constraints, important phenomena typically occur in sparse, low-volume regions of this vast space—what we might call the "golden nuggets". Locating these nuggets is exponentially hard, and this is where ML excels [15].

A deep neural network (DNN) with width $N = 10^3$ and depth $L = 10^2$ contains 10^8 weights and supports 10^{30} distinct computational paths between input and output. Through backpropagation, the training process effectively searches this massive space to adjust weights so as to populate those sparse regions where meaningful correlations live. As Geoffrey Hinton put it: "Boltzmann machines are nice, but the real thing is the backpropagation algorithm."

Modern ML applications such as DeepFold and LLMs now use up to 100 billion parameters—roughly the number of neurons in the human brain. But unlike our 20 W cerebral hardware, these models can require gigawatt-scale resources. It is estimated that next-generation chatbots will approach the gigawatt power demand, more than most existing power plants. This unveils the fundamental tension: ML systems trade Insight for brute-force optimization, and with it comes massive energy cost. The question is whether the End of Insight also implies the end of the energetic resources of planet Earth, in which case one has probably to think twice

before endorsing the "bigger is better" route undertaken by Big Tech companies [16].

The academic community is exploring ways to mitigate this, often with limited means. In the next section, we offer one such contribution: reframing ML as a class of discrete dynamical systems, namely, generalized diffusion-advection processes. This analogy allows weights to acquire physical meaning, potentially enabling more interpretable and energy-efficient learning strategies [17].

Let us describe the idea in more detail.

3 Machine learning and discrete dynamical systems

In a recent paper [17], the ML procedure was formally reinterpreted as a discrete dynamical system in relaxation form: more precisely, as a time-discretized neural integro-equation (NIDE) of the form shown in Equation 5:

$$\partial_t z = -\gamma(z - z^{eq}) \tag{5}$$

where $z = z(q,t)$ is the physical signal in spatial parameter q, and the local equilibrium defined in Equation 6:

$$z^{eq} = f(Z) \tag{6}$$

The mapping Z is a shorthand for the shifted linear convolution shown in Equation 7:

$$Z(q,t) = \int W(q,q')z(q')dq' - b(q) \tag{7}$$

where $b(q)$ is the bias function.

The procedure is quite transparent, both conceptually and mathematically: the solution $z(q,t)$ is attracted to a local equilibrium $z^{eq}(q,t)$, the target of the procedure, which is the result of a nonlinear deformation, via the activation functional, of the convoluted signal $Z(q,t)$. The former is linear and non-local; hence, it implies scale mixing while leaving amplitudes untouched. The nonlinear deformation responds to a criterion of amplitude selection but leaves scales untouched. For instance, $tanh(Z)$ leaves small amplitudes unaffected and saturates the large ones on both sides. Rectified Linear Unit (ReLU), on the other hand, leaves positive signals unchanged and sets negative ones to zero. Hence, the signal is first non-locally linearly convoluted and then locally and nonlinearly deformed in amplitude. This sequence is key for ML schemes as universal interpolators, especially in high-dimensional spaces. Once again, the three boosters of complexity are fully accounted for.

A simple Euler time marching of the Equation 5, as combined with a suitable discretization of the "space" variable q into a set of N discrete nodes, delivers Equation 8:

$$z_i(t+1) = (1-\omega)z_i(t) + \omega z_i^{eq}(t) \tag{8}$$

where $\omega = \gamma\Delta t$. Direct comparison with (5) shows that, with $\omega = 1$, this is precisely the forward step of the ML procedure with $L = T/\Delta t$ layers and N neurons per layer, with the initial condition $z(0) = x$ and output $y = z(T)$, T being the time span of the evolution.

Clearly, the result is highly dependent on the structure of the convolution kernel $W(q,q')$, whose discrete version is nothing but the weight matrix W_{ij}. In [17] it was noted that each kernel gives rise to a corresponding PDE and perhaps even low-order PDEs, such as advection-diffusion-reaction, with inhomogeneous, possibly time-dependent or even nonlinear coefficients, can give rise to pretty complex spatio-temporal patterns. Clearly, most common PDEs would lead to highly structured kernels, hence it was (naively) argued that inspection of real-life ML applications might show signatures of underlying structure. For instance, a simple advection-diffusion-equation in one spatial dimension would give rise to a tridiagonal-dominant weight matrix. The detection of such structural regularity in the weight matrices would offer a very valuable inroad to their explainability in the first place, let alone the energy savings resulting from a reduced set of weights.

The argument can be easily extended to more general PDEs, including strong inhomogeneities and nonlinearities, which could easily be implemented by convoluting local nonlinear combinations of the signal, as shown in Equation 9:

$$Z(q,t) = \int W(q,q')g(z(q'))dq' - b(q) \tag{9}$$

where $g(z)$ is a local activation function, independent of f.

For instance, by truncating the integral to the second moment, we would obtain $Z(q) = W_0(q)g(z) + W_1(q)\partial_x g(z) + W_2(q)\partial_{xx}g(z)$. In the above, the moments are defined as $W_k(q) = \int W(q,q')(q'-q)^k dq'$, and one may inspect their decay with increasing order to retain only a finite number in the sequence without seriously affecting the accuracy of the solution. The link between deep learning and PDEs is an active topic of research in the field [18].

4 Inspecting the weights of a PINN application to rarefied gas dynamics

The preceding considerations suggest that analyzing the weights of a trained network might offer insight into its internal logic, particularly when the problem is governed by a well-understood physical model. Let us test the idea by means of a concrete application. Recently, we trained a physics-informed neural network (PINN) on a body-force-driven rarefied gas flow through a 2D periodic array of cylinders in the laminar, isothermal and weakly compressible limit [19]. This problem has a well-defined structure governed by the Boltzmann equation (BE).

A key parameter in rarefied gas dynamics is the Knudsen number $Kn = \lambda/D$, defined as the ratio of the molecular mean free path λ to a characteristic length scale of the problem - in this case, the cylinder diameter D. The Knudsen number thus serves as a measure of rarefaction, characterizing the importance of non-equilibrium effects. In the continuum regime ($Kn \ll 1$), the Navier-Stokes equations provide an accurate description of the flow. However, as Kn increases, non-local effects due to the finite mean free path of molecules predominate, and momentum transfer is no longer only influenced by local velocity gradients. This nonlocal coupling is particularly evident in flows around curved surfaces.

Motivated by these challenges, we designed a neural network that takes as input the spatial coordinates $(x_i, y_i), i,j = [1,256]$ and the Knudsen number Kn, and outputs the velocity components u_x, u_y, pressure p and deviatoric components of

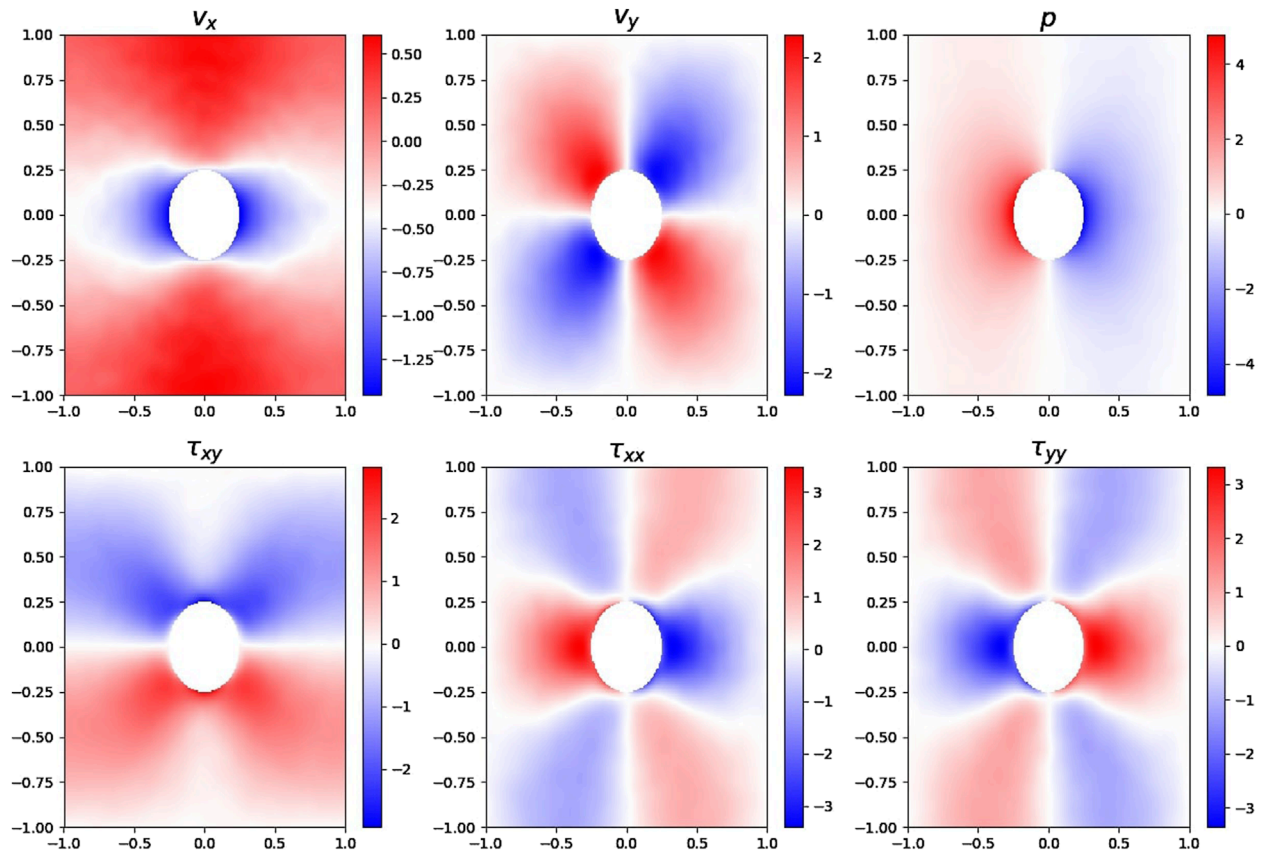

FIGURE 1
Normalized macroscopic fields predicted by the PINN for $Kn = 1$. Each output variable was standardized across the full training domain in x, y and Kn, resulting in zero mean and unit variance. This normalization, used to aid convergence, explains the presence of negative values in v_x despite the imposed positive body force in the x-direction. See Ref. [19] for details.

the stress tensors τ_{xy}, τ_{xx}, τ_{yy}, which are presented in Figure 1. This formulation captures the key physical quantities that characterize rarefied gas flows of industrial interest. While the input space is low-dimensional, the underlying physics is high-dimensional due to its dependence on the full velocity distribution function.

The network consists of a Fourier layer to impose periodic boundary conditions [20, 21], followed by nine hidden layers of 100 neurons with *tanh* as an activation function. To guide the network towards physical fidelity, the loss functions includes three terms: (1) the mean squared error with respect to DSMC data, normalized via Z-score scaling; (2) a no-penetration boundary condition; and (3) the continuum equation and the Cauchy momentum equations. The physics terms are collectively weighted at 10% relative to the data fidelity loss.

Additional pre-processing was required to ensure convergence. Pressure fields were debiased across the Knudsen range due to their artificial variation stemming from how Kn was numerically set. Likewise, velocity vectors lost orientation under normalization, which was necessary to stabilize training but severs their physical directional meaning. Finally, L1L2 regularization was applied to promote smoothness and broad participation across weights rather than sparse activation.

These design choices—though effective for learning—blur direct links between physical content and internal network representations. The basic question we pursue is whether, despite these compromises, the trained network retains any recognizable physical structure. Before discussing the results, let us first show that our problem does exhibit the three key properties we described as where neural networks should excel. To this purpose, let us recall basic facts about the Boltzmann equation (BE).

4.1 The Boltzmann equation

This equation describes the dynamics of the probability density function $f(x, v; t)$, encoding the probability of finding a particle (atom, molecule) around position x in space at time t with molecular velocity v. In one dimension, and neglecting external forces: the BE reads as in Equation 10

$$\partial_t f + v\partial_x f = Q(f, f) \tag{10}$$

The left hand side represents the free streaming of the molecules, while the right hand encodes molecular collisions via a quadratic integral in velocity space involving the product $f(v)f(v')$ of two colliding particles with velocities v and v' at (x, t). Equation 11 shows

it in full splendor:

$$Q(f,f) = \int P(v,w|v',w')[f(v)f(w) - f(v')f(w')]dwdv'dw' \quad (11)$$

where (v,w) and (v',w') are the pre- and post-collisional velocities, and $P(v,w|v',w')$ is the probability of such collision; by micro-reversibility this is the same as the probability of the inverse collision from (v',w') to (v,w). In a way, P can be interpreted as the weight kernel of the "kinetic Boltzmann machine", one which needs zero training since the physics supplies all the information it takes. In particular, the collision term is subject to mass-momentum-energy conservation laws, as shown in Equation 12:

$$\int Q(f,f)\{1,v,v^2\}dv = 0 \quad (12)$$

This structure embeds all three complexity boosters: nonlinearity via the quadratic collision term, nonlocality through the transport of information across space and velocity scales, and high dimensionality due to its formulation in six-dimensional phase space (plus time). While Q is local in physical space, it is nonlocal in velocity space, and its competition with the streaming term drives the system toward or away from local equilibrium f^{eq}, depending on the Knudsen number Kn. In the hydrodynamic limit ($Kn \to 0$), equilibrium dominates and the BE reduces to the Navier-Stokes equations. As Kn increases, non-equilibrium effects emerge, and molecular-scale information propagates over macroscopic distances.

Even more relevant to macroscopic observables, integration of the BE over velocities yields transport equations that are simultaneously nonlinear and nonlocal in physical space, such as the familiar convective term $u\partial_x u$, with $u(x,t) = \int vf(x,v,t)dv / \int fdv$. This emergent structure underpins the complexity of fluid turbulence and forms the basis for the powerful lattice formulations of the BE [22–25].

5 Learning the Boltzmann solutions via PINNs

The PINN described above is trained on numerical data from Direct Simulation Monte Carlo of the Boltzmann equation [26]. Given the problem's high physical structure and the inclusion of physics-informed loss terms, one might expect this to manifest in structured, interpretable weight patterns. However, as shown in Figure 2, the distribution of weights in the deepest layer closely resembles a zero-mean Gaussian. While small but statistically significant departures from normality are detected—excess kurtosis between −0.65 and −0.5, Kolmogorov–Smirnov (KS) distances between 0.015 and 0.03—these deviations do not amount to the emergence of any discernible physical structure. No clear trace of the governing equations appears to persist in the weight statistics.

One possible explanation for this mismatch lies in the conceptual assumptions behind the analogy between machine learning and discrete dynamical systems. Such analogies typically rely on the presence of an ordered metric structure among discrete coordinates q_i, with the weight matrices W_{ij} mediating local interactions, much like in finite-difference approximations to PDEs. In generic ML settings, where input nodes are abstract and unordered, this assumption rarely holds. But in our case, the input space (x_i, y_i, Kn) is spatially structured and sampled on a regular grid, making the presence of such order more plausible. If the network were reflecting the structure of a discrete Boltzmann operator, one might expect W_{ij} to develop a block-tridiagonal structure, at least in configuration space. The blocks themselves might remain irregular, owing to the stochastic nature of the DSMC in velocity space, but some coarse trace of this structure might reasonably be expected. Instead, our findings suggest that the

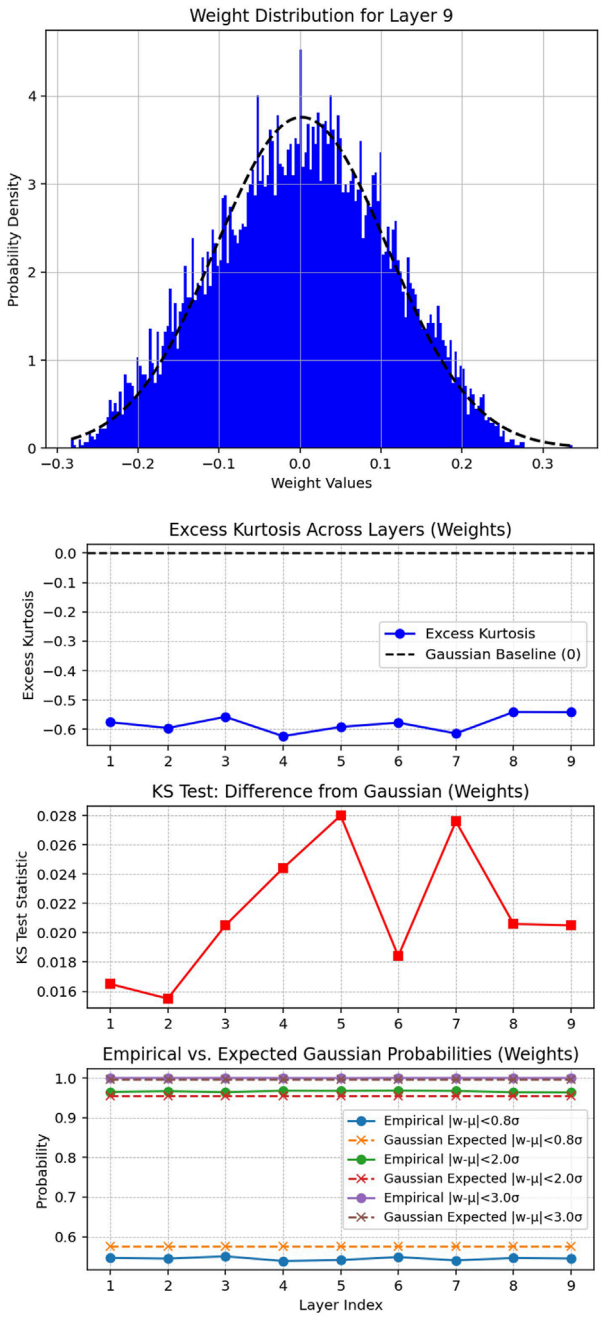

FIGURE 2
Distribution of weights in the deepest layer. The PDF is overlaid with a standard Gaussian. The excess kurtosis and KS statistical analysis indicates a weak, but statistically significant departure from normality.

training process scrambles the underlying order, leading to a diffuse, near-Gaussian parameter distribution that bears little resemblance to the physics it approximates. This supports the broader thesis that physics-informed ML and traditional simulation can arrive at the same predictive outputs via fundamentally distinct internal representations.

7 Tentative conclusions and outlook

The analysis of a PINN trained on a rarefied gas flow problem reveals a striking disconnect between the physical structure of the governing Boltzmann equation and the internal organization of the network. Despite being constrained by physics-informed losses, the network's weights resemble near-Gaussian distributions with no evident trace of the underlying integro-differential operator. This supports the view that machine learning and traditional simulation can offer functionally equivalent yet epistemologically distinct routes to the same solution.

That such a disconnect emerges even for a moderately complex and well-understood problem raises a deeper question: beyond a certain threshold of complexity, might Insight—as traditionally pursued in physics—become practically inaccessible? If so, the role of explainability must be rethought, not as a universal standard, but as a domain-dependent aspiration.

This need not be cause for alarm. A lack of interpretable structure at the parameter level does not imply that ML is unscientific—but it does suggest that physical knowledge and machine-learned representations follow fundamentally different logics. Bridging them may require new tools, not just to improve interpretability, but to reframe what interpretability itself should mean in AI-augmented science.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, upon reasonable request.

Author contributions

J-MT: Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. MD: Formal Analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Writing – review and editing. SS: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. SS is grateful to SISSA for financial support under the "Collaborations of Excellence" initiative and to the Simons Foundation for supporting several enriching visits. He also wishes to acknowledge many enlightening discussions with PV Coveney, A. Laio, D. Spergel and S. Strogatz. JMT is grateful to the FRQNT "Fonds de recherche du Québec - Nature et technologies (FRQNT)" for financial support (Research Scholarship No. 357484). SS and MD gratefully acknowledge funding by the European Union (EU) under the Horizon Europe research and innovation programme, EIC Pathfinder - grant No. 101187428 (iNSIGHT) and from the European Research Council ERC-PoC2 grant No. 101187935 (LBFAST).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. The large language model ChatGPT-4o was used to assist in improving the clarity and conciseness of the English in this manuscript. The authors take full responsibility for the content and interpretations presented.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

1. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Adv Neural Inf Process Syst* (2017) 30.

2. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with alphafold. *nature* (2021) 596:583–9. doi:10.1038/s41586-021-03819-2

3. NPO. All nobel prizes (2025).

4. Cai S, Mao Z, Wang Z, Yin M, Karniadakis GE. Physics-informed neural networks (pinns) for fluid mechanics: a review. *Acta Mechanica Sinica* (2021) 37:1727–38. doi:10.1007/s10409-021-01148-1

5. Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: can language models be too big? In: *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, FAccT 21*. New York, NY, USA: Association for Computing Machinery (2021). p. 610–23.

6. Strogatz S. The end of insight. In: *What is your dangerous idea? today's leading thinkers on the unthinkable*. New York, NW: Simon & Schuster (2007).

7. Succi S. Chatbots and zero sales resistence. *Front Phys* (2024) 12:1484701. doi:10.3389/fphy.2024.1484701

8. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* (2015) 521:436–44. doi:10.1038/nature14539

9. Anderson C. The end of theory: the data deluge makes the scientific method obsolete. *Wired Mag* (2008) 16:16.

10. Hasperué W. The master algorithm: how the quest for the ultimate learning machine will remake our world. *J Computer Sci & Technology* (2015) 15.

11. Coveney PV, Dougherty ER, Highfield RR. Big data need big theory too. *Philosophical Trans R Soc A: Math Phys Eng Sci* (2016) 374:20160153. doi:10.1098/rsta.2016.0153

12. Succi S, Coveney PV. Big data: the end of the scientific method? *Philosophical Trans R Soc A* (2019) 377:20180145. doi:10.1098/rsta.2018.0145

13. Coveney PV, Highfield R. Artificial intelligence must be made more scientific. *J Chem Inf Model* (2024) 64:5739–41. doi:10.1021/acs.jcim.4c01091

14. Strogatz SH. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Chapman and Hall/CRC (2024).

15. Poggio T, Mhaskar H, Rosasco L, Miranda B, Liao Q. Why and when can deep-but not shallow-networks avoid the curse of dimensionality: a review. *Int J Automation Comput* (2017) 14:503–19. doi:10.1007/s11633-017-1054-2

16. Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. *Neuron* (2017) 95:245–58. doi:10.1016/j.neuron.2017.06.011

17. Succi S. *A note on the physical interpretation of neural pde's*. arXiv preprint arXiv:2502.06739. (2025). doi:10.48550/arXiv.2502.06739

18. Han J, Jentzen A, E W. Solving high-dimensional partial differential equations using deep learning. *Proc Natl Acad Sci* (2018) 115:8505–10. doi:10.1073/pnas.1718942115

19. Tucny J-M, Lauricella M, Durve M, Guglielmo G, Montessori A, Succi S. Physics-informed neural networks for microflows: rarefied gas dynamics in cylinder arrays. *J Comput Sci* (2025) 87:102575. doi:10.1016/j.jocs.2025.102575

20. Dong S, Ni N. A method for representing periodic functions and enforcing exactly periodic boundary conditions with deep neural networks. *J Comput Phys* (2021) 435:110242. doi:10.1016/j.jcp.2021.110242

21. Lu L, Pestourie R, Yao W, Wang Z, Verdugo F, Johnson SG. Physics-informed neural networks with hard constraints for inverse design. *SIAM J Scientific Comput* (2021) 43:B1105–32. doi:10.1137/21m1397908

22. Succi S. *The lattice boltzmann equation: for complex states of flowing matter*. Oxford: Oxford University Press (2018).

23. Calì A, Succi S, Cancelliere A, Benzi R, Gramignani M. Diffusion and hydrodynamic dispersion with the lattice boltzmann method. *Phys Rev A* (1992) 45:5771–4. doi:10.1103/physreva.45.5771

24. Rybicki FJ, Melchionna S, Mitsouras D, Coskun AU, Whitmore AG, Steigner M, et al. Prediction of coronary artery plaque progression and potential rupture from 320-detector row prospectively ecg-gated single heart beat ct angiography: lattice boltzmann evaluation of endothelial shear stress. *The Int J Cardiovasc Imaging* (2009) 25:289–99. doi:10.1007/s10554-008-9418-x

25. D'Orazio A, Succi S. Boundary conditions for thermal lattice boltzmann simulations. In: *International conference on computational science*. Springer (2003). p. 977–86.

26. Bird GA. *Molecular gas dynamics, NASA STI/Recon technical report A 76* (1976).