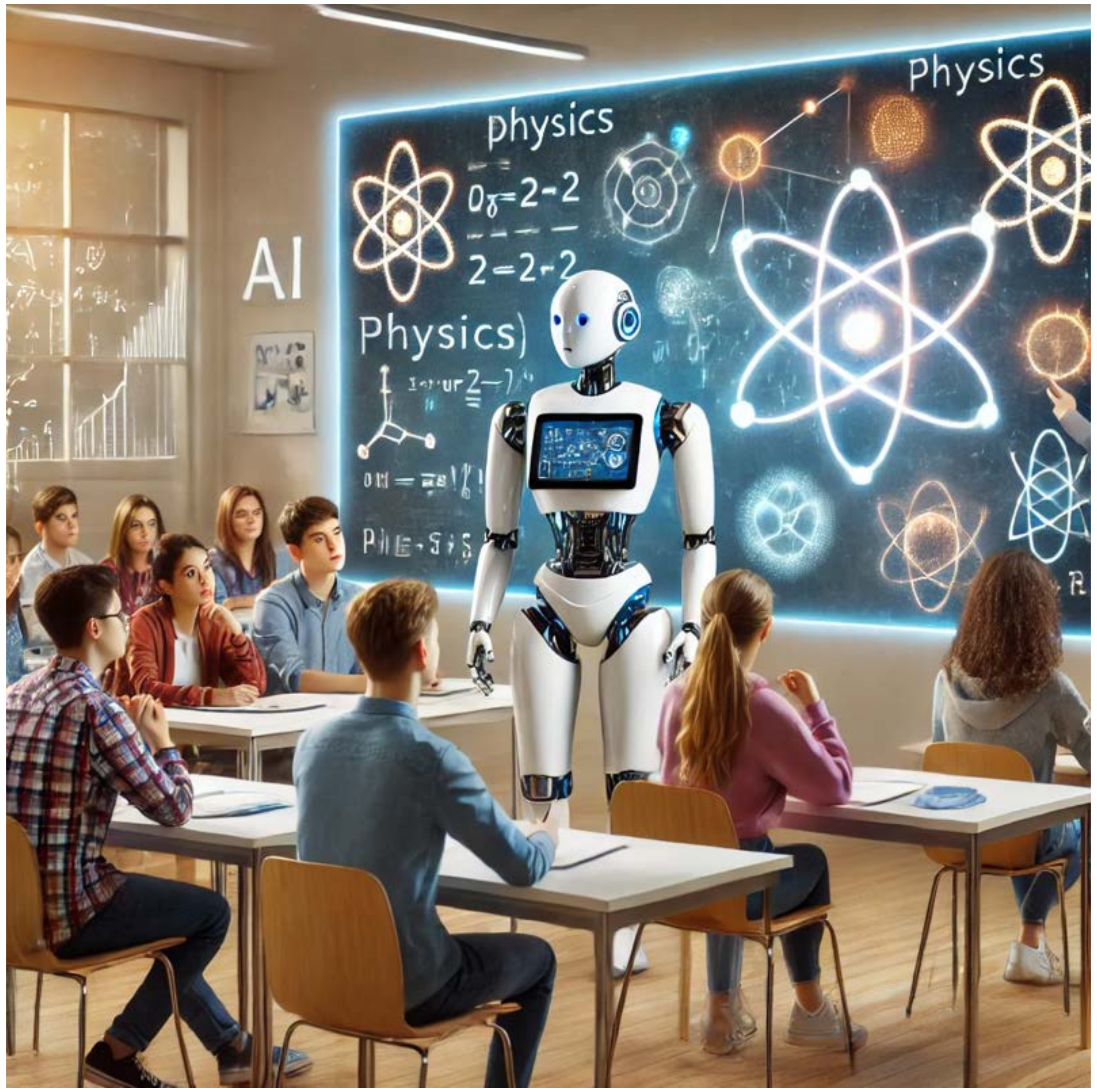
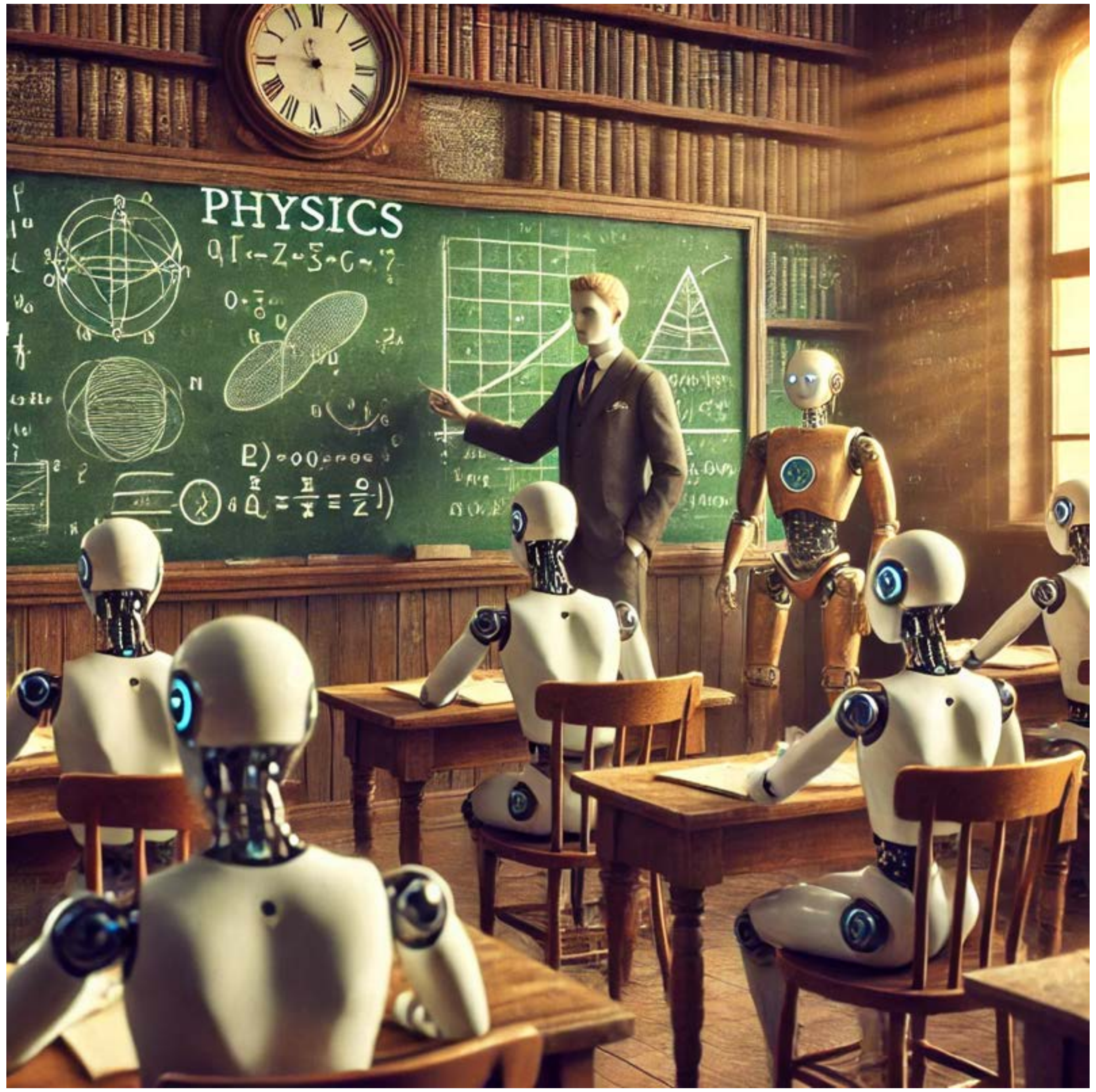
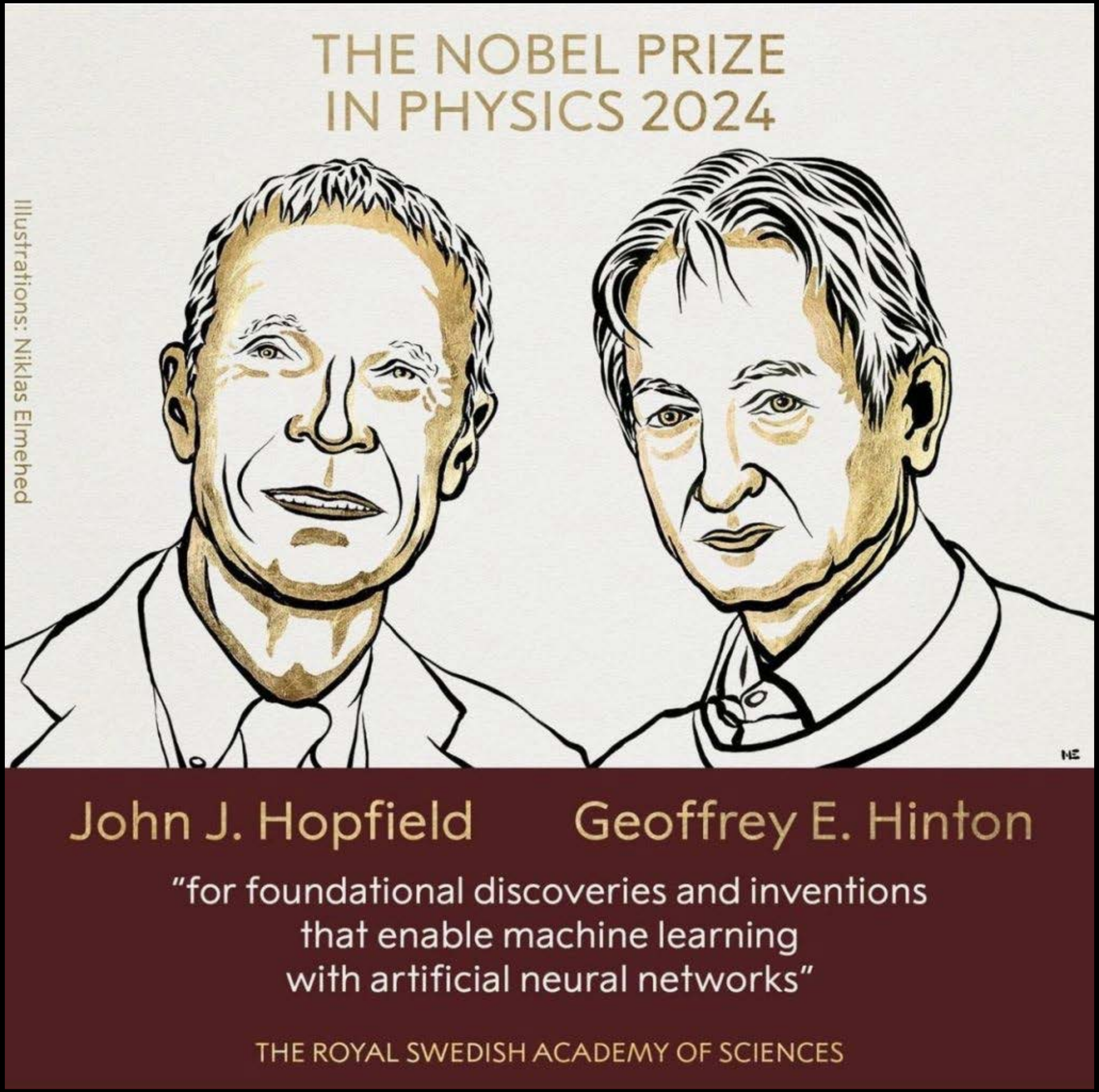
AI for Physics or Physics for AI?

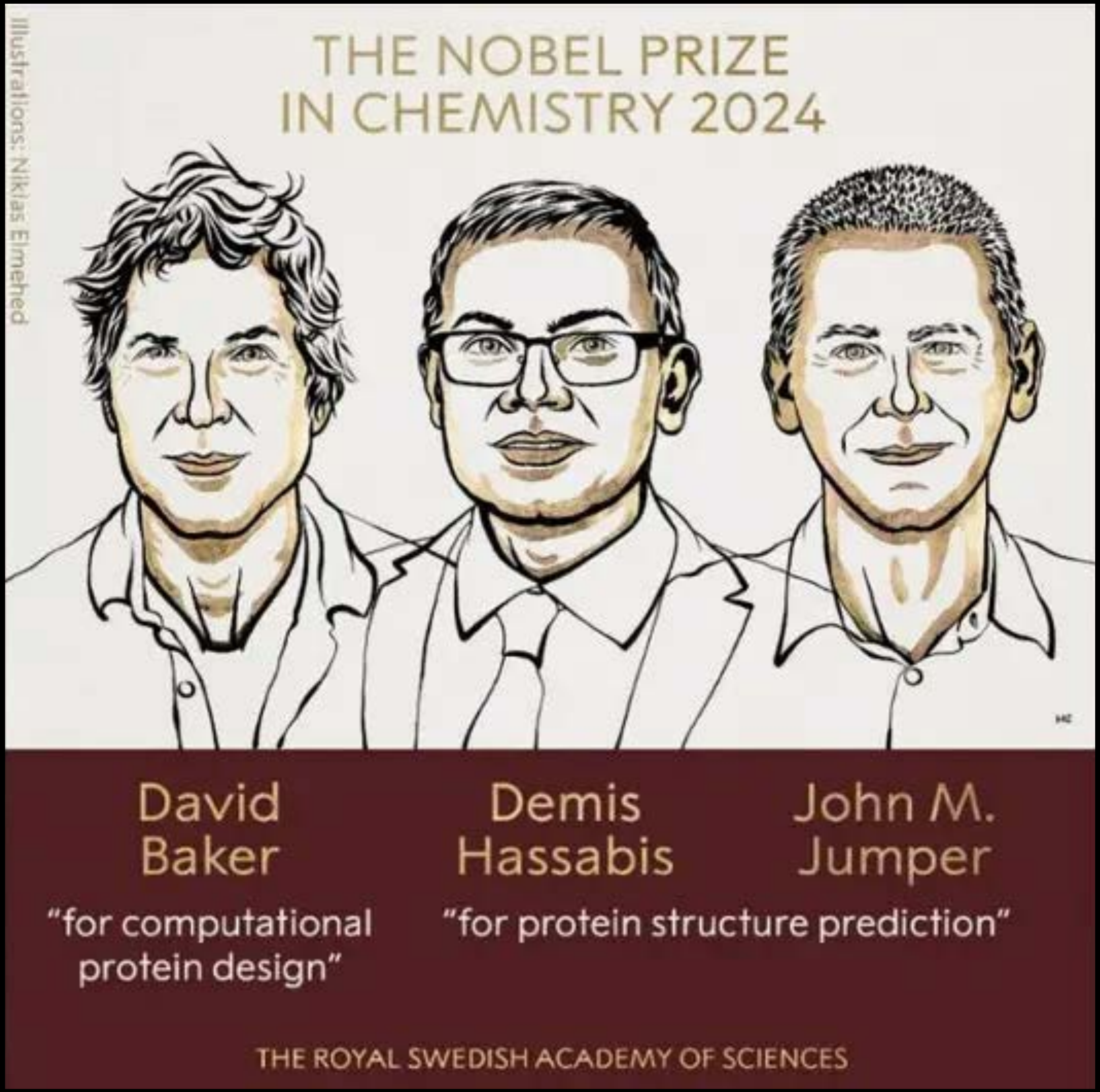




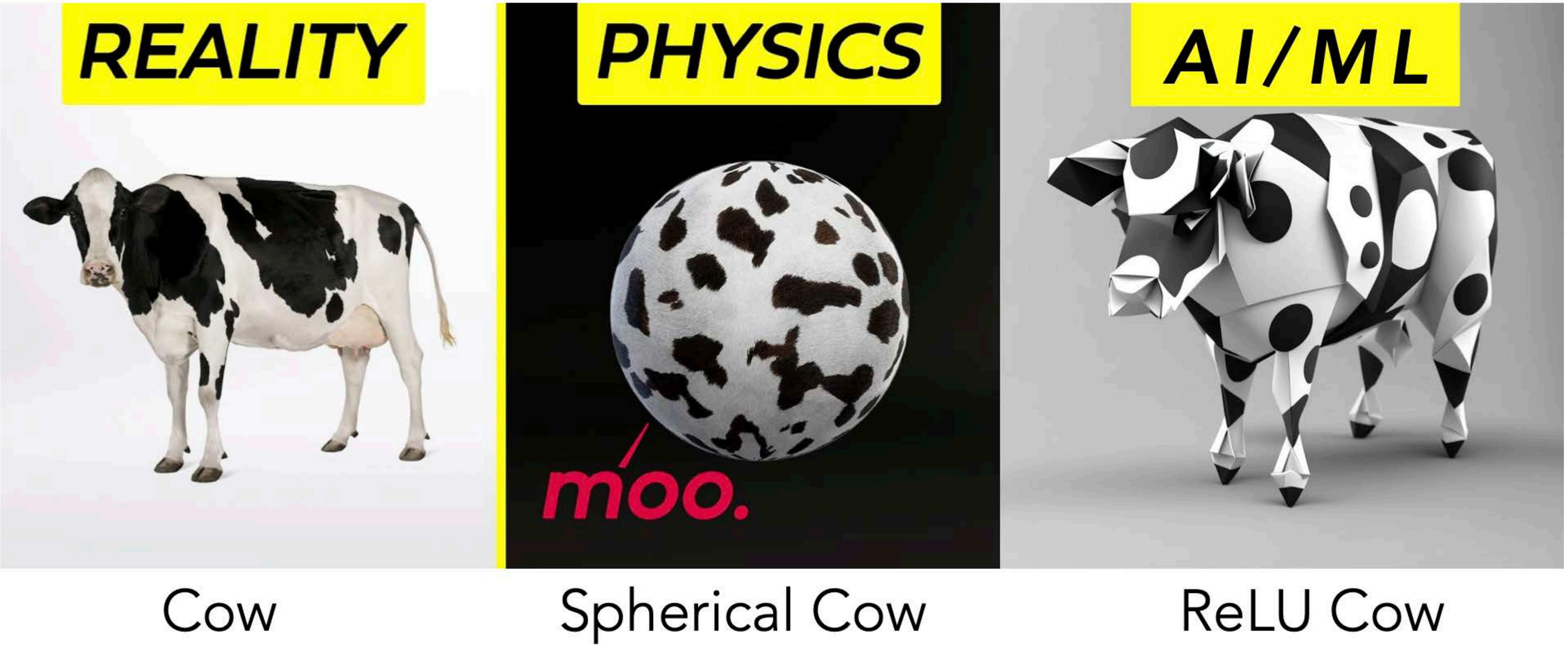
Christoph Weniger
Gravitational and Astropartical Physics Amsterdam (GRAPPA)
Institute of Physics, University of Amsterdam

SURF erc GRAPPA
GRavitation AstroParticle Physics Amsterdam


THE NOBEL PRIZE
IN PHYSICS 2024
Illustrations: Niklas Elmehed
John J. Hopfield
Geoffrey E. Hinton
“for foundational discoveries and inventions
that enable machine learning
with artificial neural networks”
THE ROYAL SWEDISH ACADEMY OF SCIENCES


THE NOBEL PRIZE
IN CHEMISTRY 2024
Illustrations: Niklas Elmehed
David
Baker
Demis
Hassabis
John M.
Jumper
“for computational
protein design”
“for protein structure prediction”
THE ROYAL SWEDISH ACADEMY OF SCIENCES

AI and Physics: Different approaches to predictions



Physics
- Differential equations describe constituent dynamics
- Often based on **idealisations**
- Typically interpretable

AI/ML
- Deep neural networks learn useful representations
- Often **task dependent**
- Can generalise from large amounts of data

Deep learning

Training
(evaluating examples & adjusting connections)

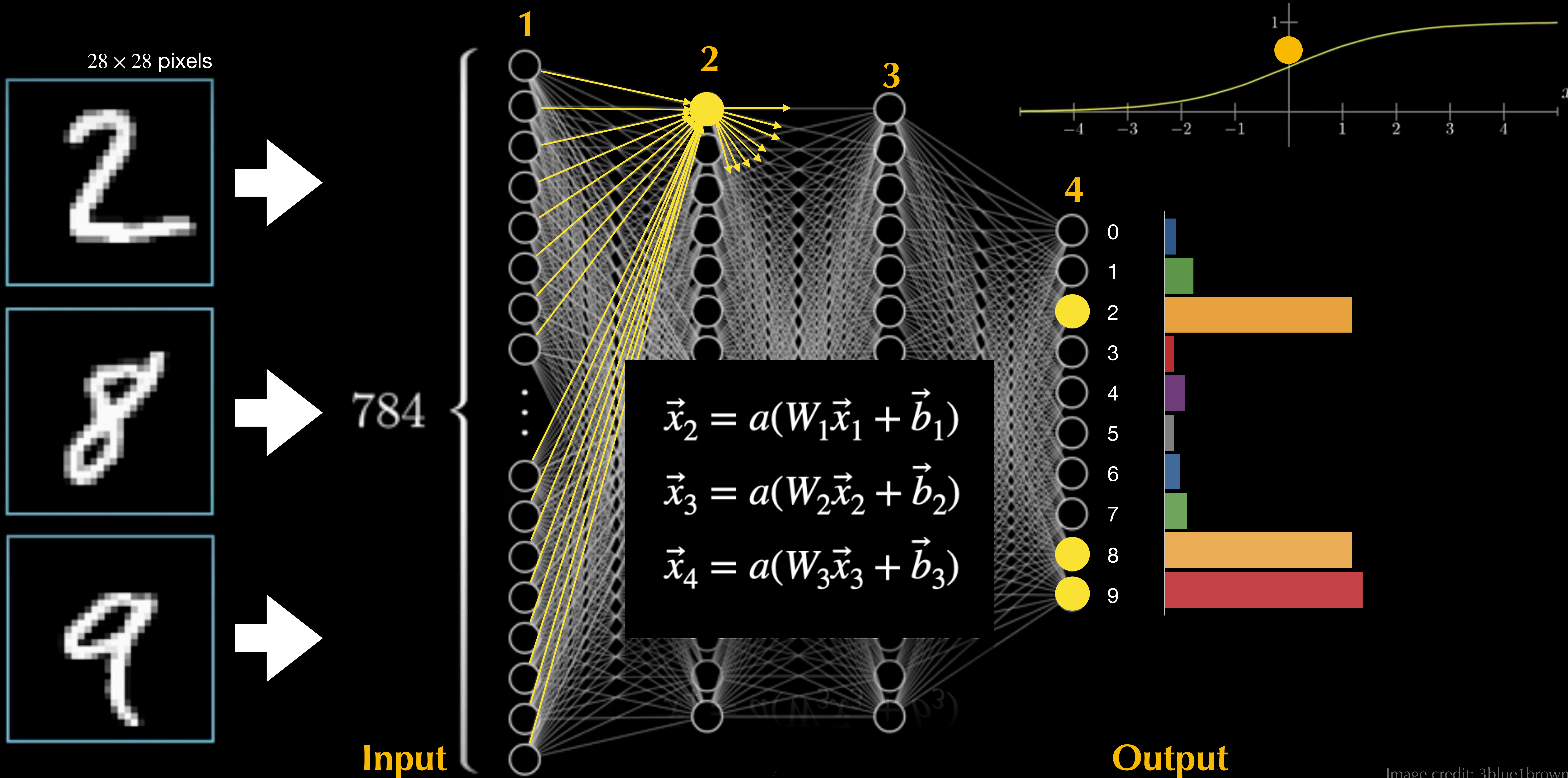
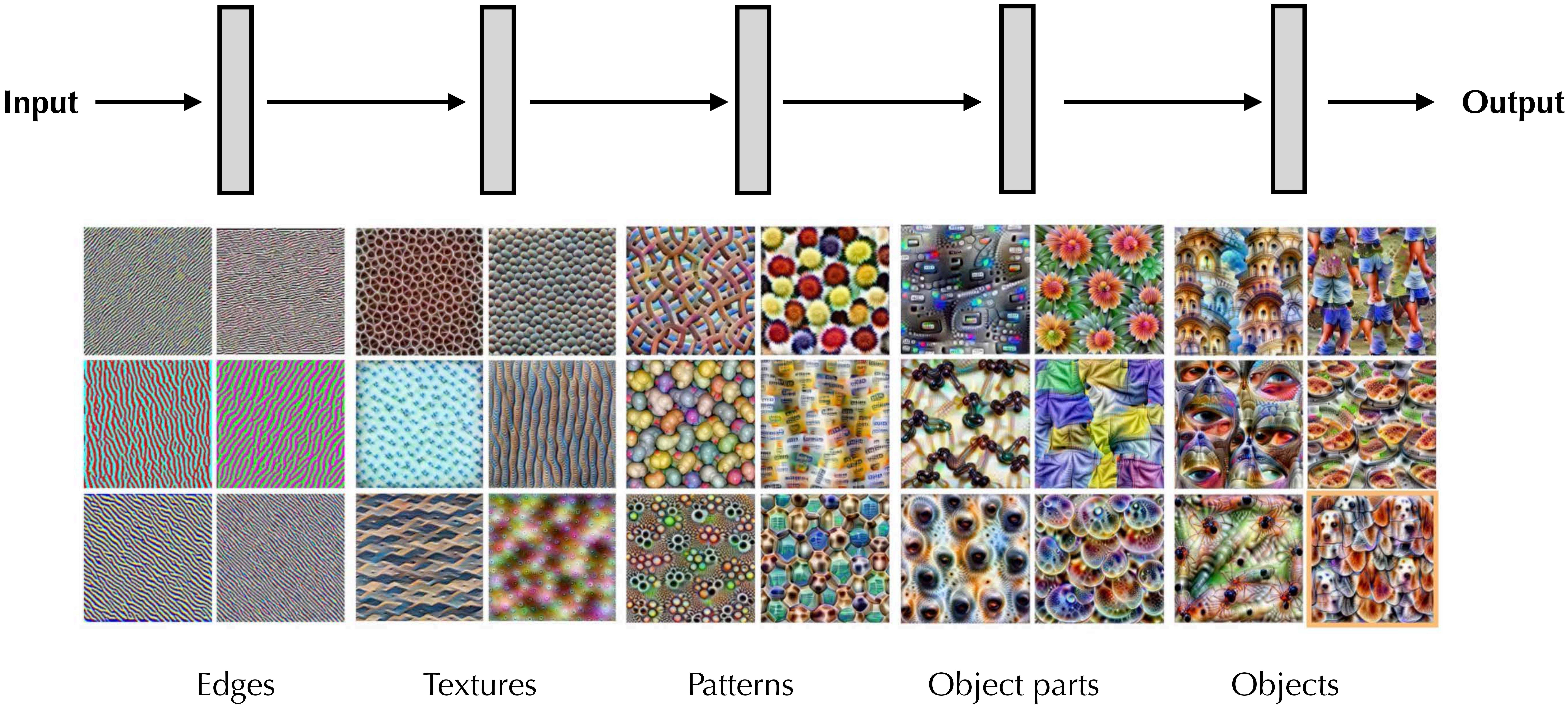


Image credit: 3blue1brown

Hierarchical Feature Learning in Neural Networks

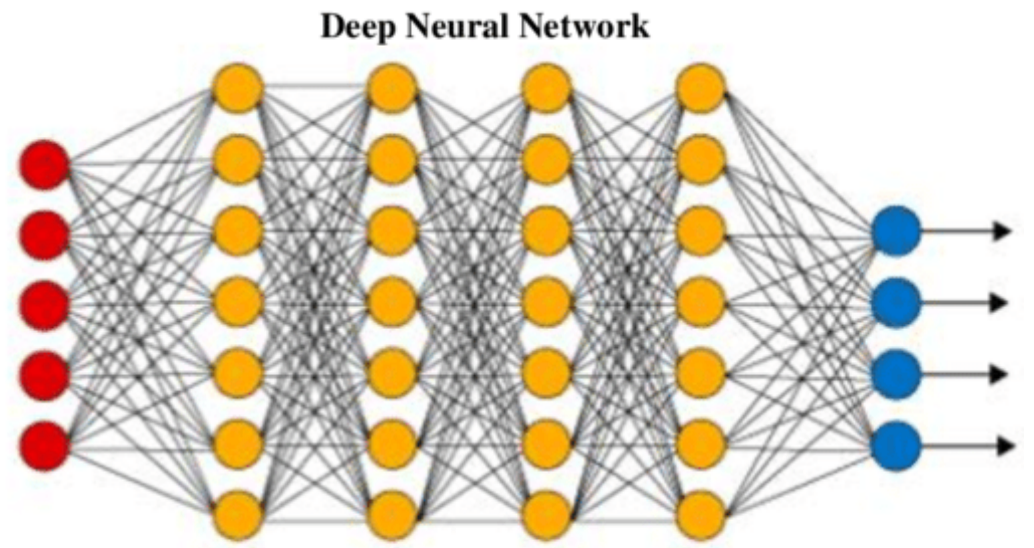
From Low-Level Edge Detection to High-Level Object Representations



https://www.isikdogan.com/blog/transfer-learning.html

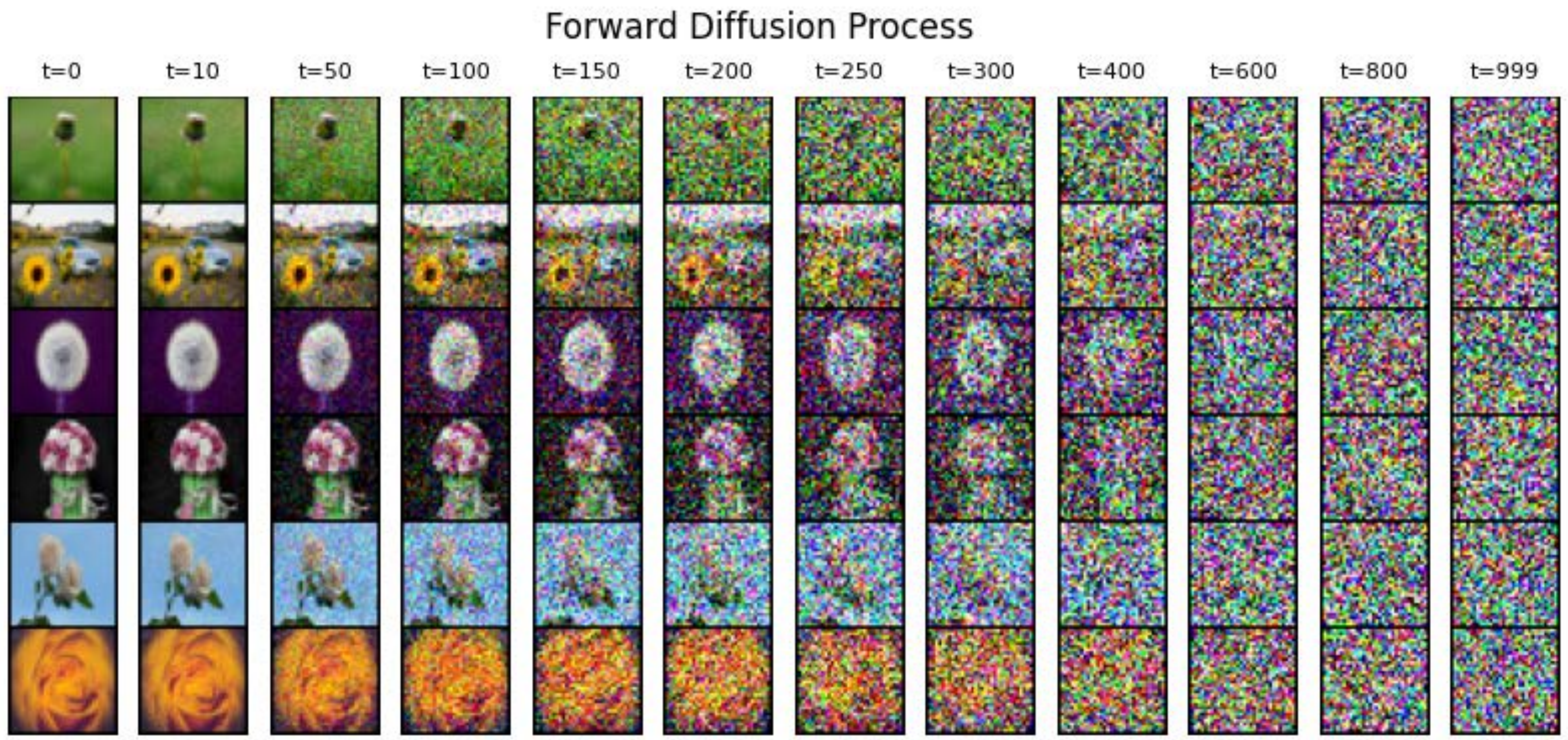
1) Emergent generative AI

Text and image generation via autoregressive models



$\mathbf{Q}$ = "Explain emergent AI in 10 words."

$\mathbf{w}_1$="AI"
$\mathbf{w}_2$="developing"
$\mathbf{w}_3$="new"
$\mathbf{w}_4$="behaviors"
$\mathbf{w}_5$="or"
$\mathbf{w}_6$="properties"
$\mathbf{w}_7$="not"
$\mathbf{w}_8$="explicitly"
$\mathbf{w}_9$="programmed"
$\mathbf{w}_{10}$="in."
$\mathbf{w}_{11}$=STOP



What are probable next words given the thousand previous words?

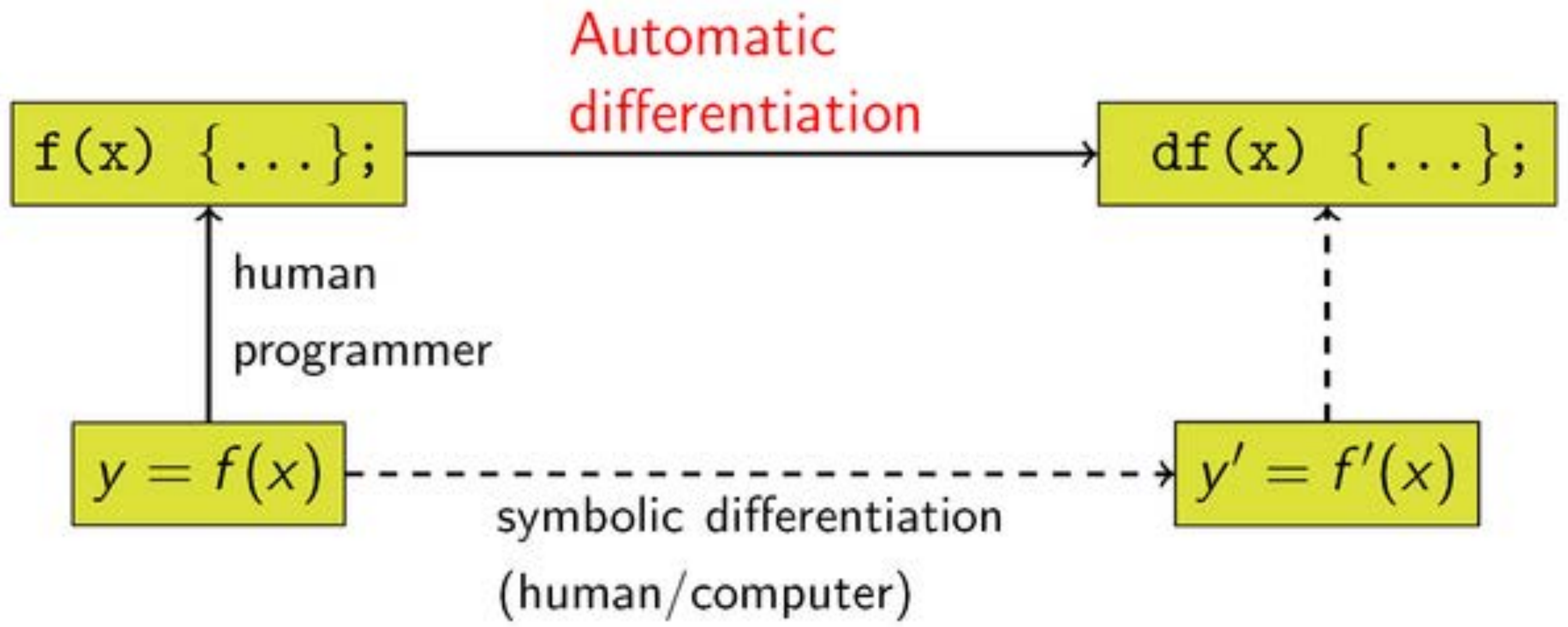
$$p_\phi(\mathbf{w}_k \mid \mathbf{w}_{k-1}, \mathbf{w}_{k-2}, \ldots, \mathbf{w}_{k-1000}; \mathbf{Q})$$

Given a noisy image, generate a slightly less noisy version and iterate.

$$p_\phi(\mathbf{m}_k \mid \mathbf{m}_{k-1}; \mathbf{Q})$$

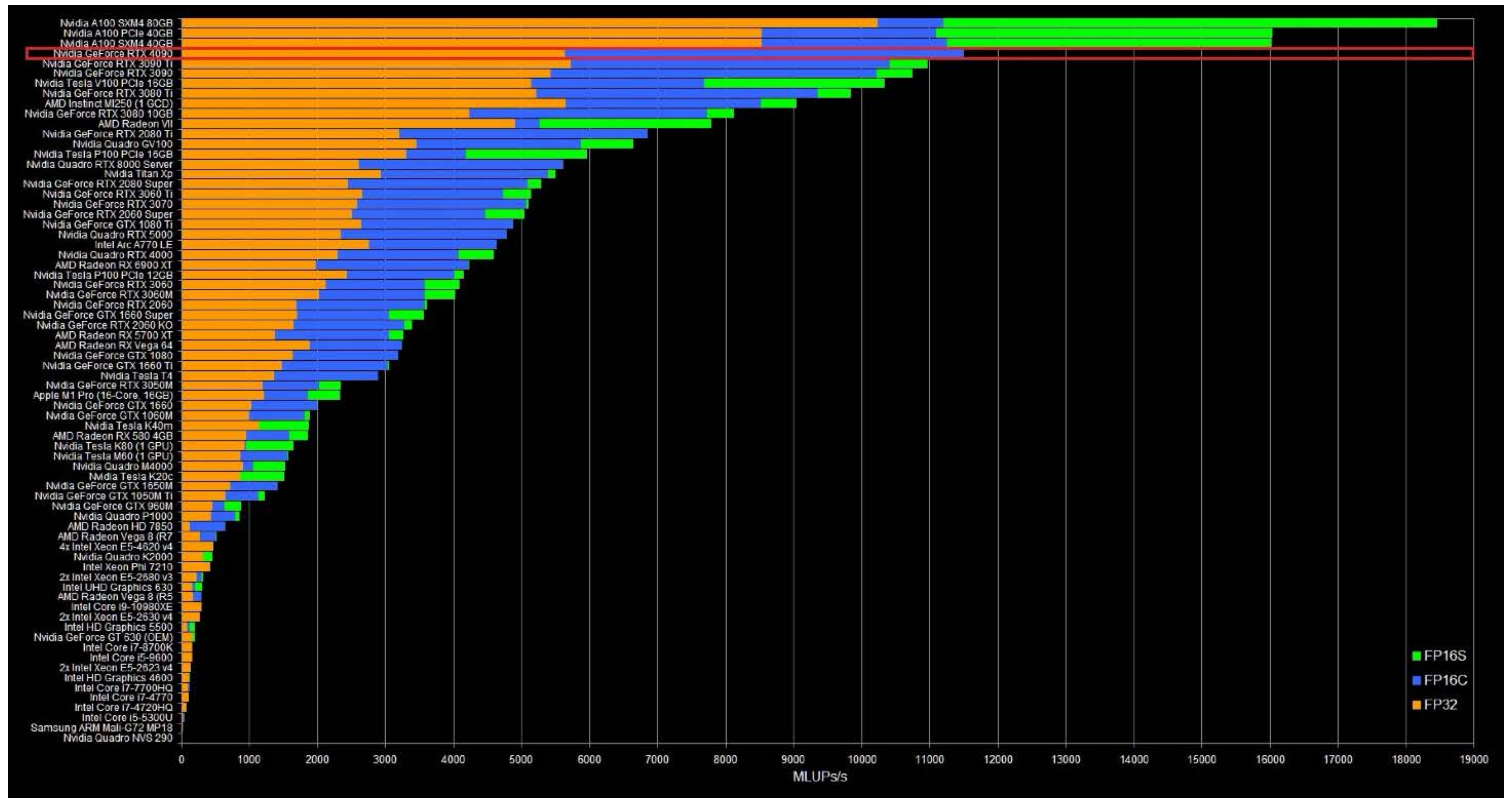
2) AI technology stack

(1) Automatic differentiation frameworks

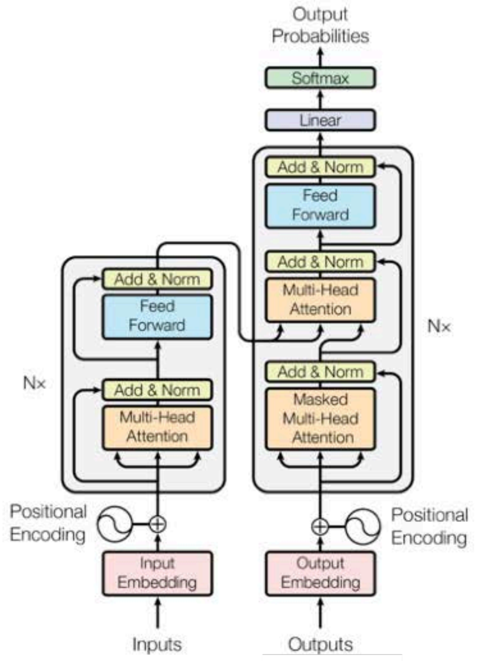


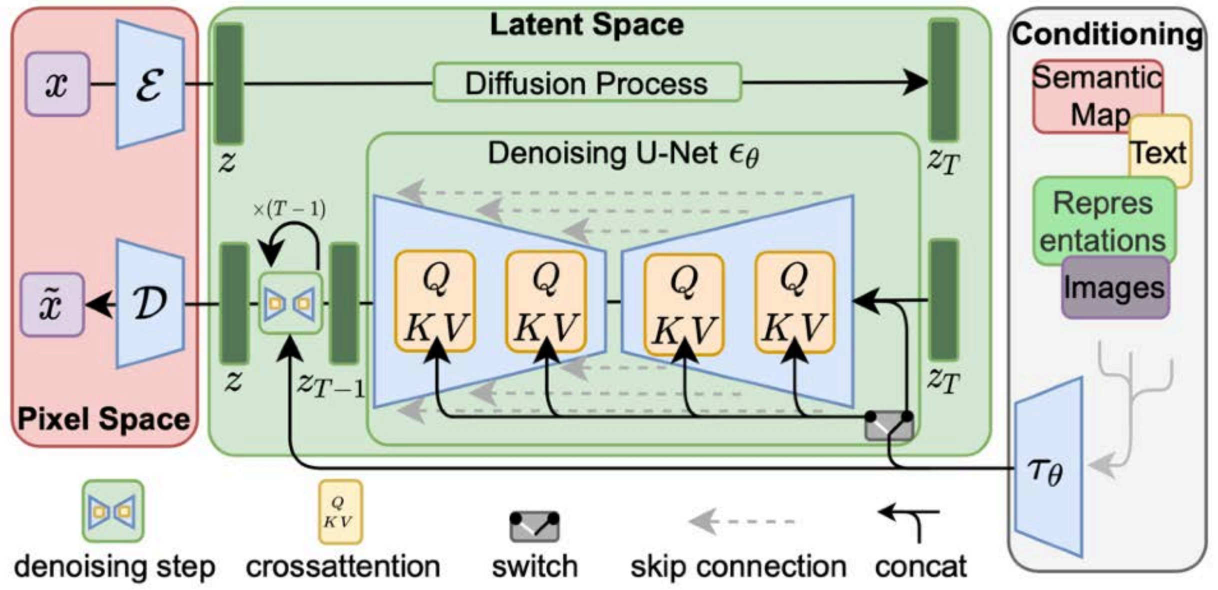
Zygote PyTorch

(2) Hardware (GPUs, TPUs, FPGAs, ...)



(3) Algorithms & Architectures





AI for Physics and Physics for AI

Optimisation, Variational Principles

Information Theory, Entropy

Energy-based models

Bayesian & Monte Carlo methods

Equivariance & Symmetries

GPU-acceleration of simulations

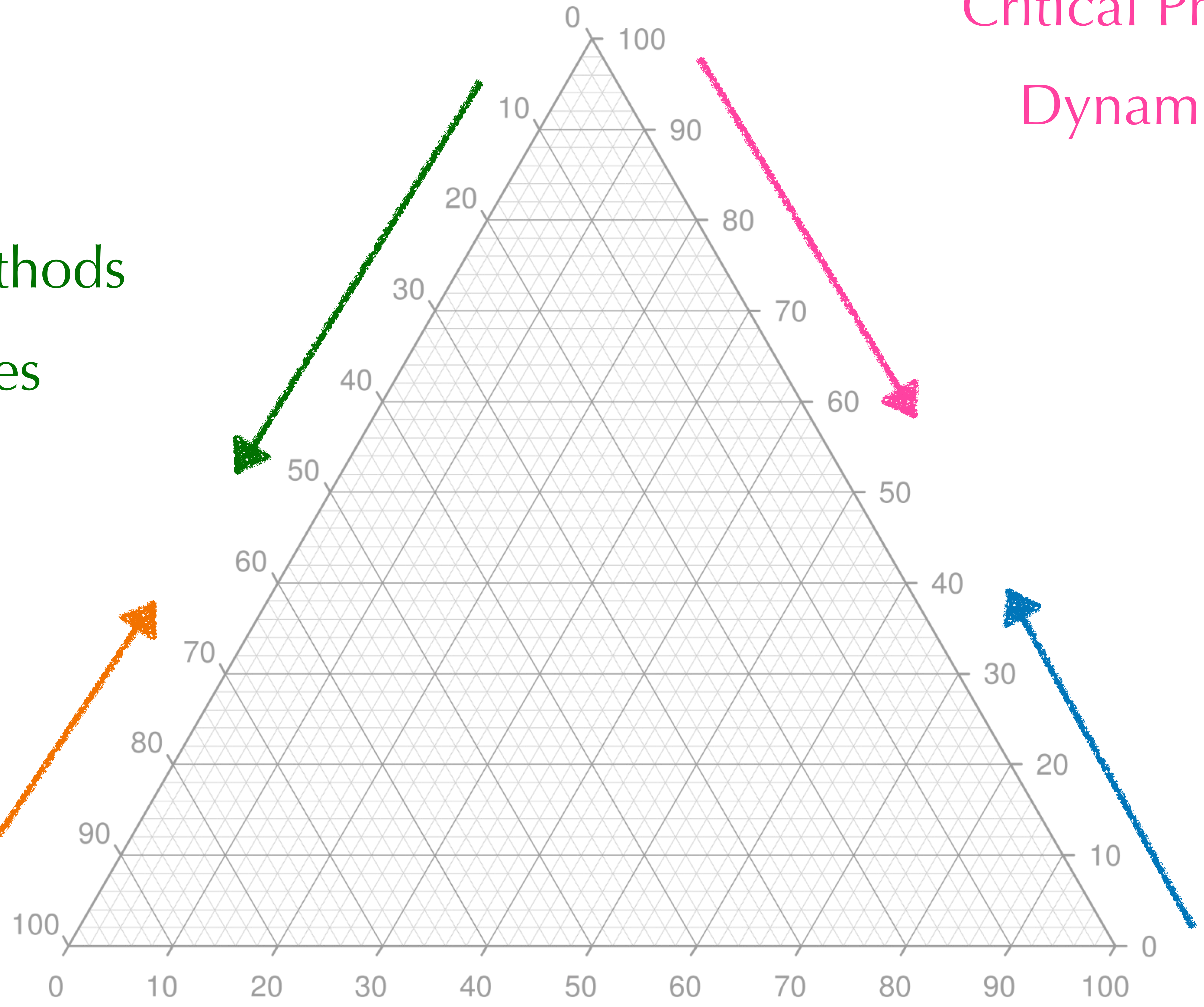
Automatic differentiation

Pattern recognition

Automatisation

Prediction

Physics



Critical Phenomena

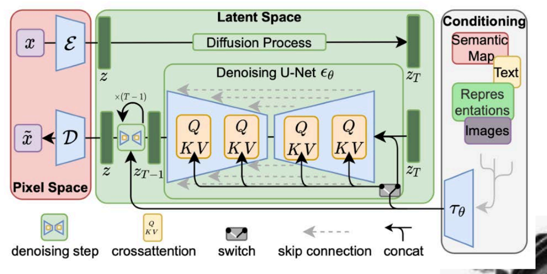
Dynamical systems

Education

Data interpretation

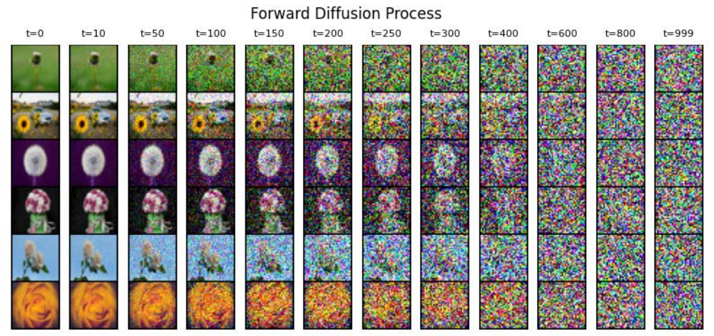
Idea generation

Coding



AI technology

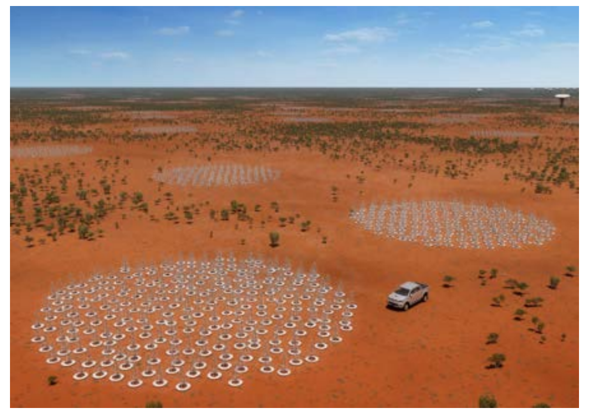
Emergent AI



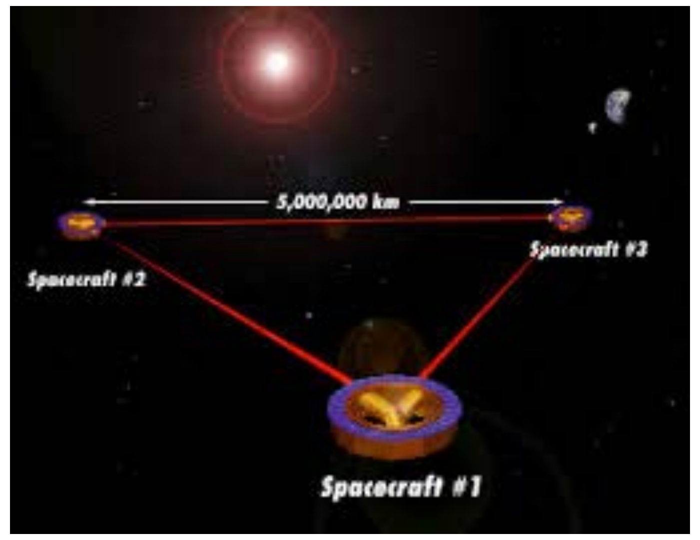
My Research

Probes of the dark Universe

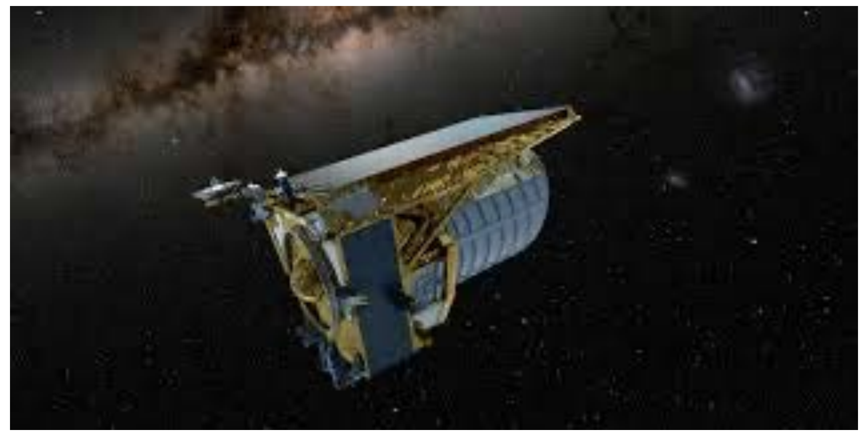
Inflation, dark energy, dark matter, strong-field gravity



21cm cosmology



Space-based GW



Galaxy surveys

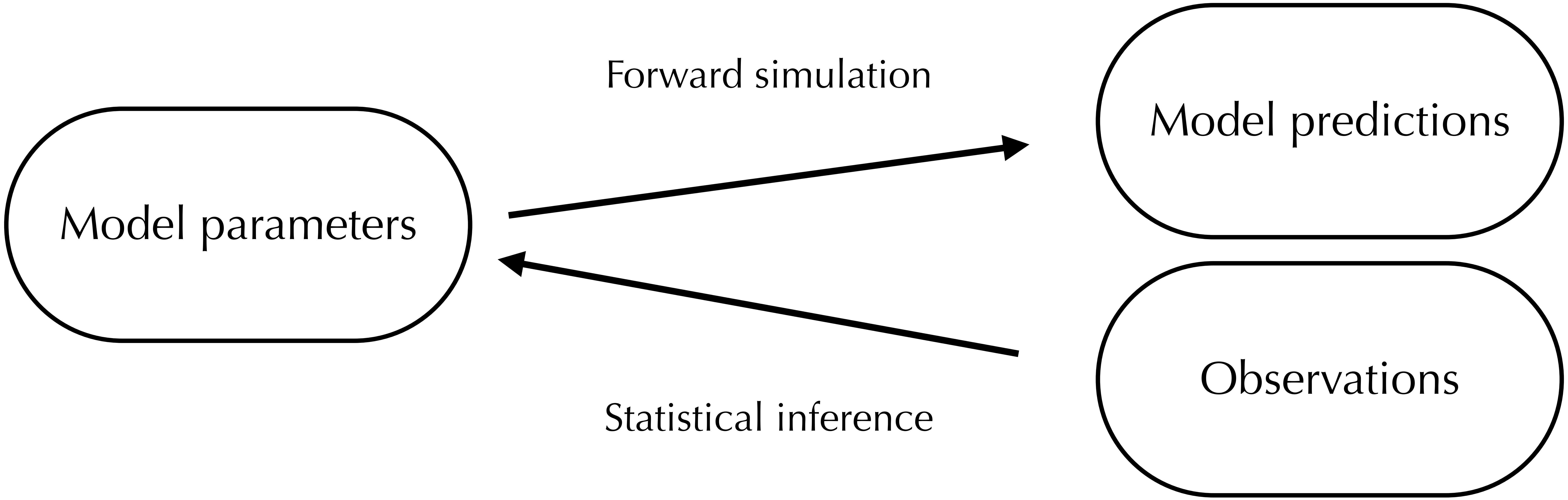
Ground-based GW

Image credit: ESA

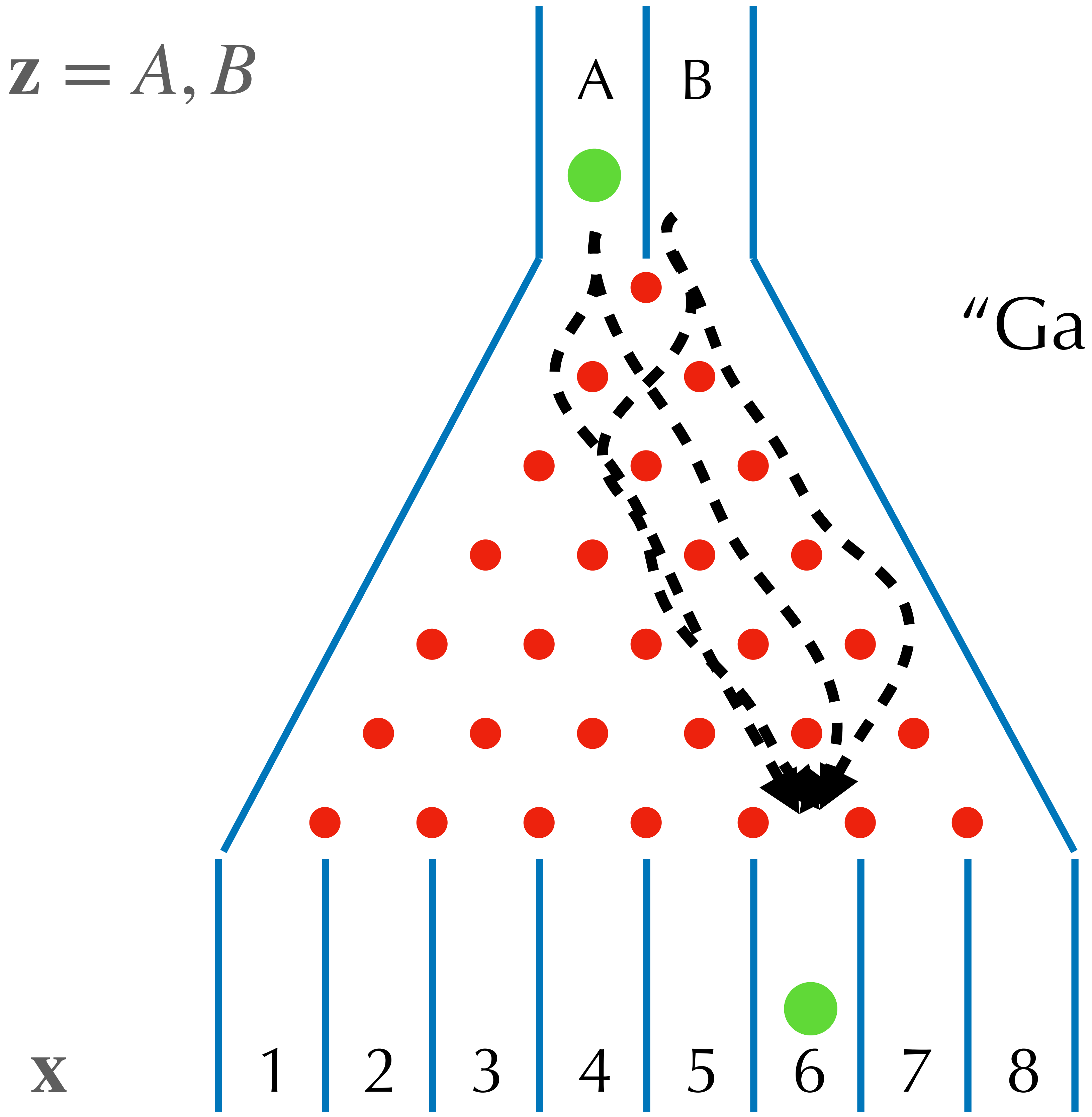
Example: Strong gravitational lensing

Credit: ALMA (ESO/NRAO/NAOJ), L. Calçada (ESO), Y. Hezaveh et al.

The inverse problem (aka statistical inference)



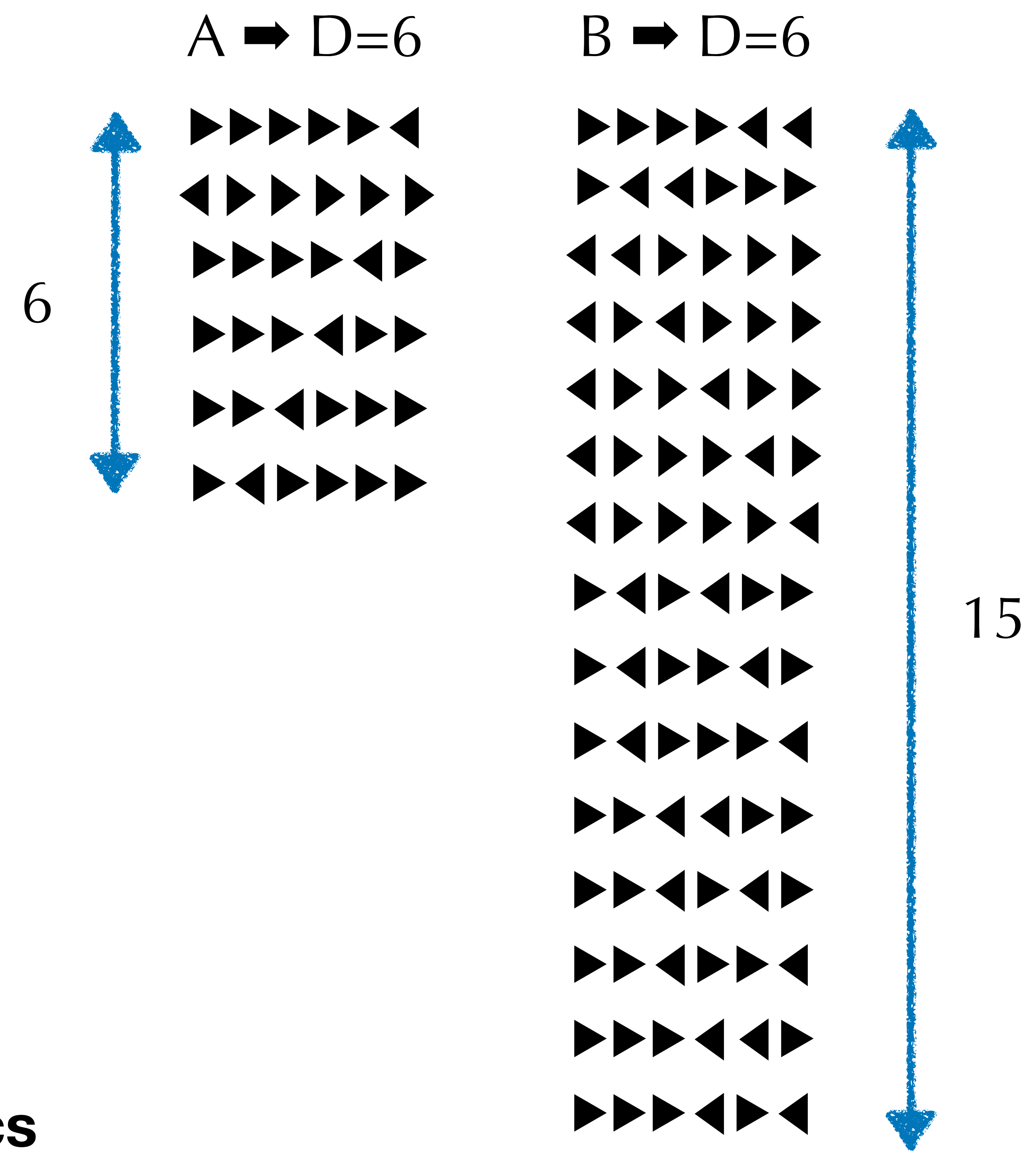
Traditional statistical inference



$D = 6$

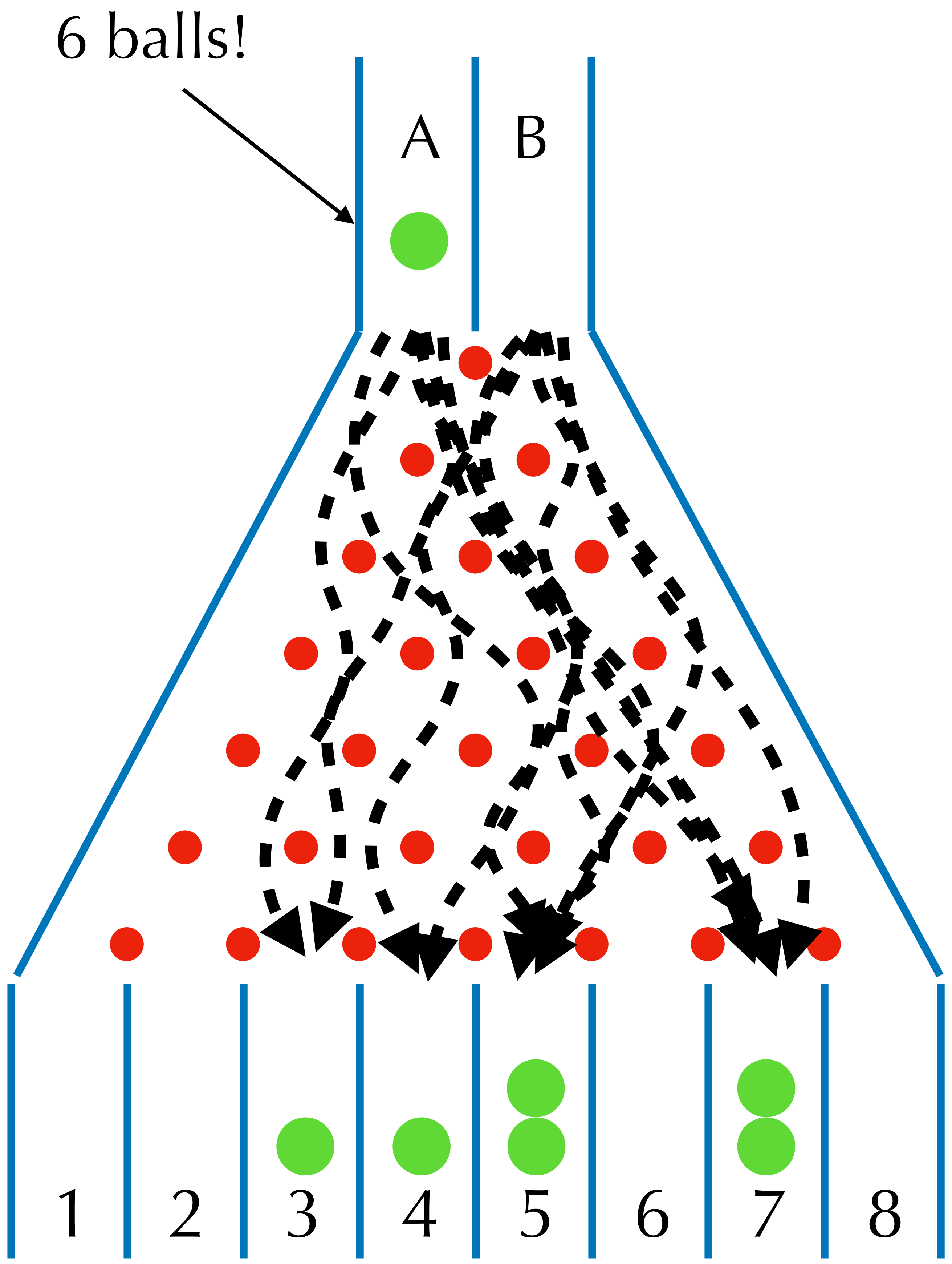
“Galton Board”

Bayesian statistics
(T. Bayes, 1763)



$$P(B \mid D = 6) = \frac{15}{15 + 6} \simeq 71\,\%$$

Traditional statistical inference



$D = (3,4,5,5,7,7)$

A ➡ $D = (3,4,5,5,7,7)$

(▶▶◀▶▶▶, ▶▶◀▶▶▶, ▶▶◀▶▶▶,
▶▶◀▶▶▶, ▶▶◀▶▶▶, ▶▶◀▶▶▶)
...
...
...
...

67500

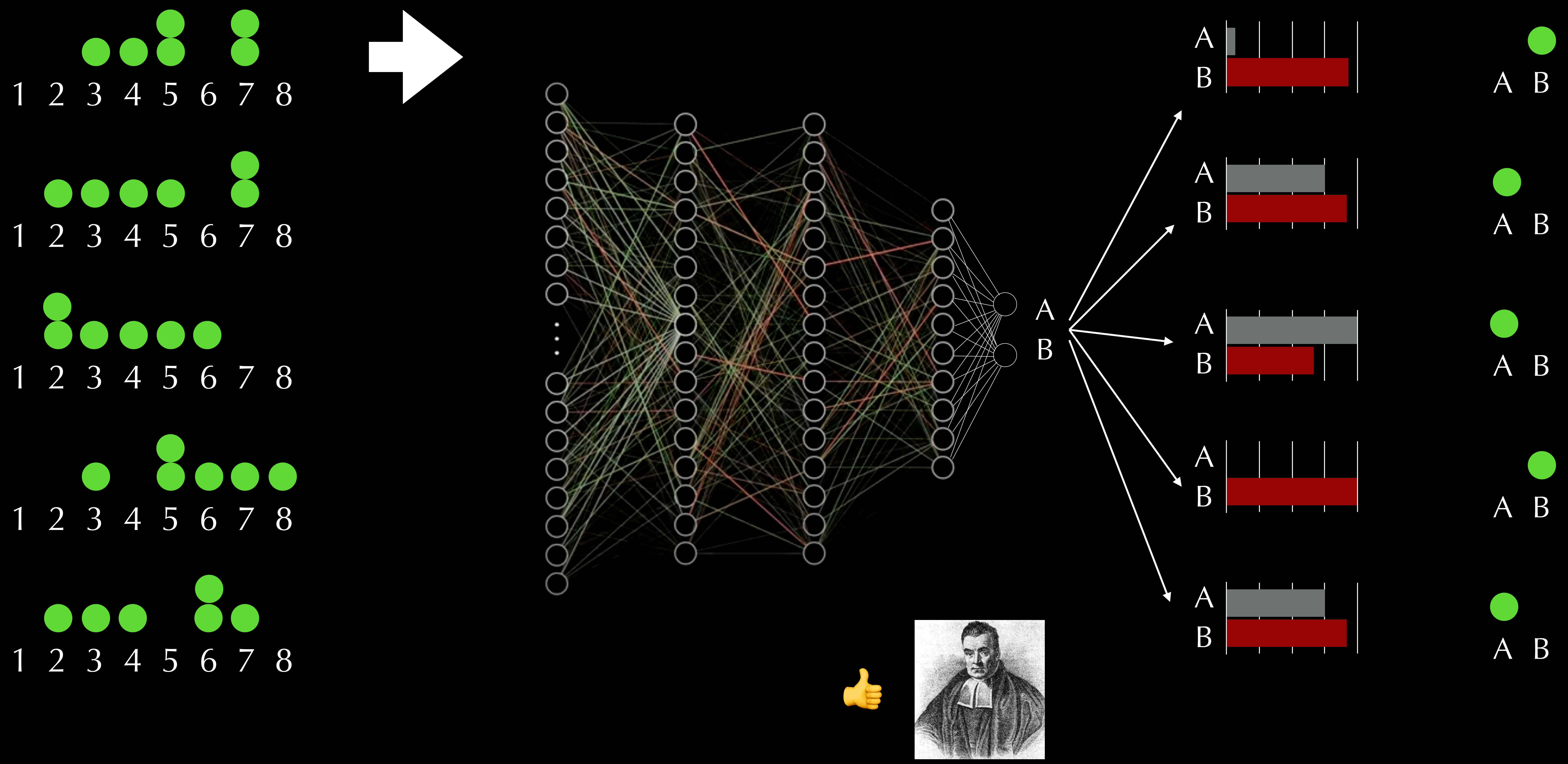
B ➡ $D = (3,4,5,5,7,7)$

(▶▶◀▶▶▶, ▶▶◀▶▶▶, ▶▶◀▶▶▶,
▶▶◀▶▶▶, ▶▶◀▶▶▶, ▶▶◀▶▶▶)
...
...
...
...

1.296.000

$$P(B \mid D = (3,4,5,5,7,7)) = \frac{1.296.000}{1.296.000 + 67500} \simeq 93\,\%$$

AI-assisted statistical inference

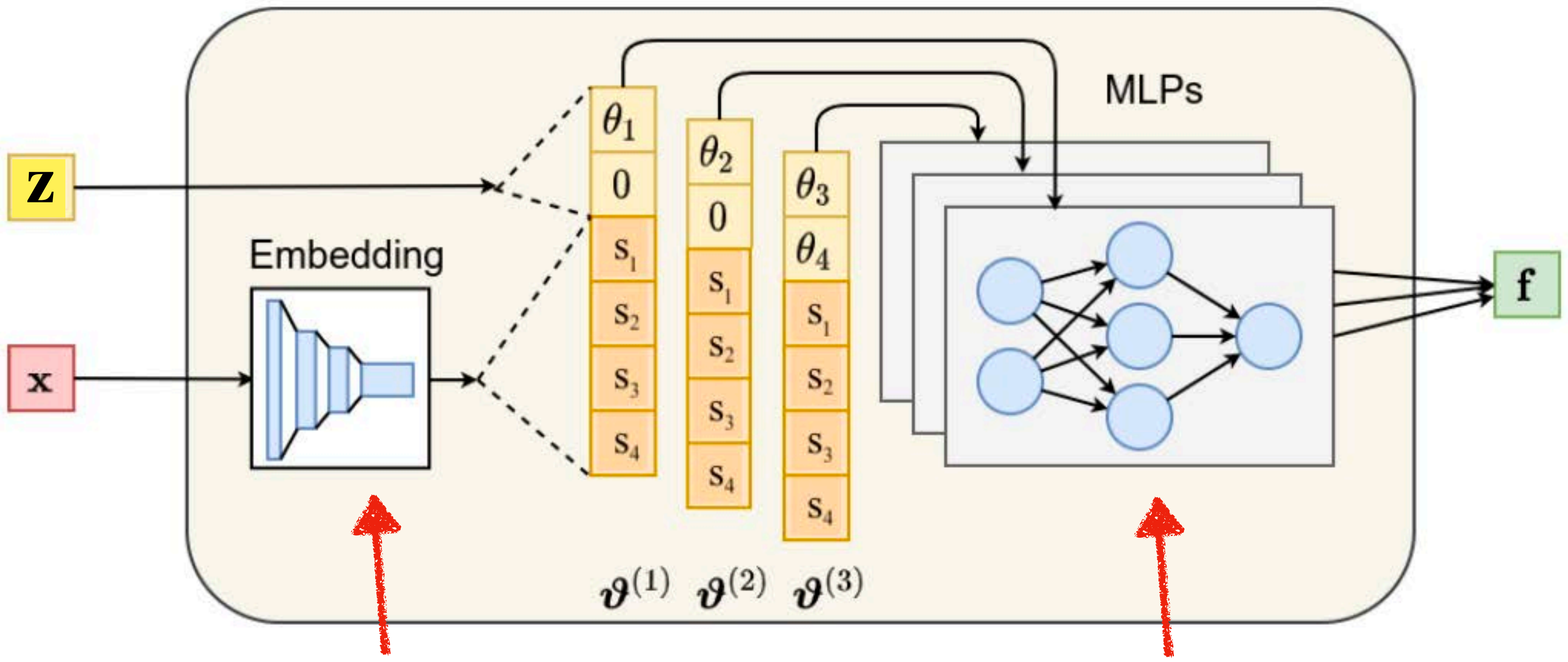


SBI with Neural Ratio Estimation*

Neural ratio estimation (NRE)

Train a neural network to discriminate

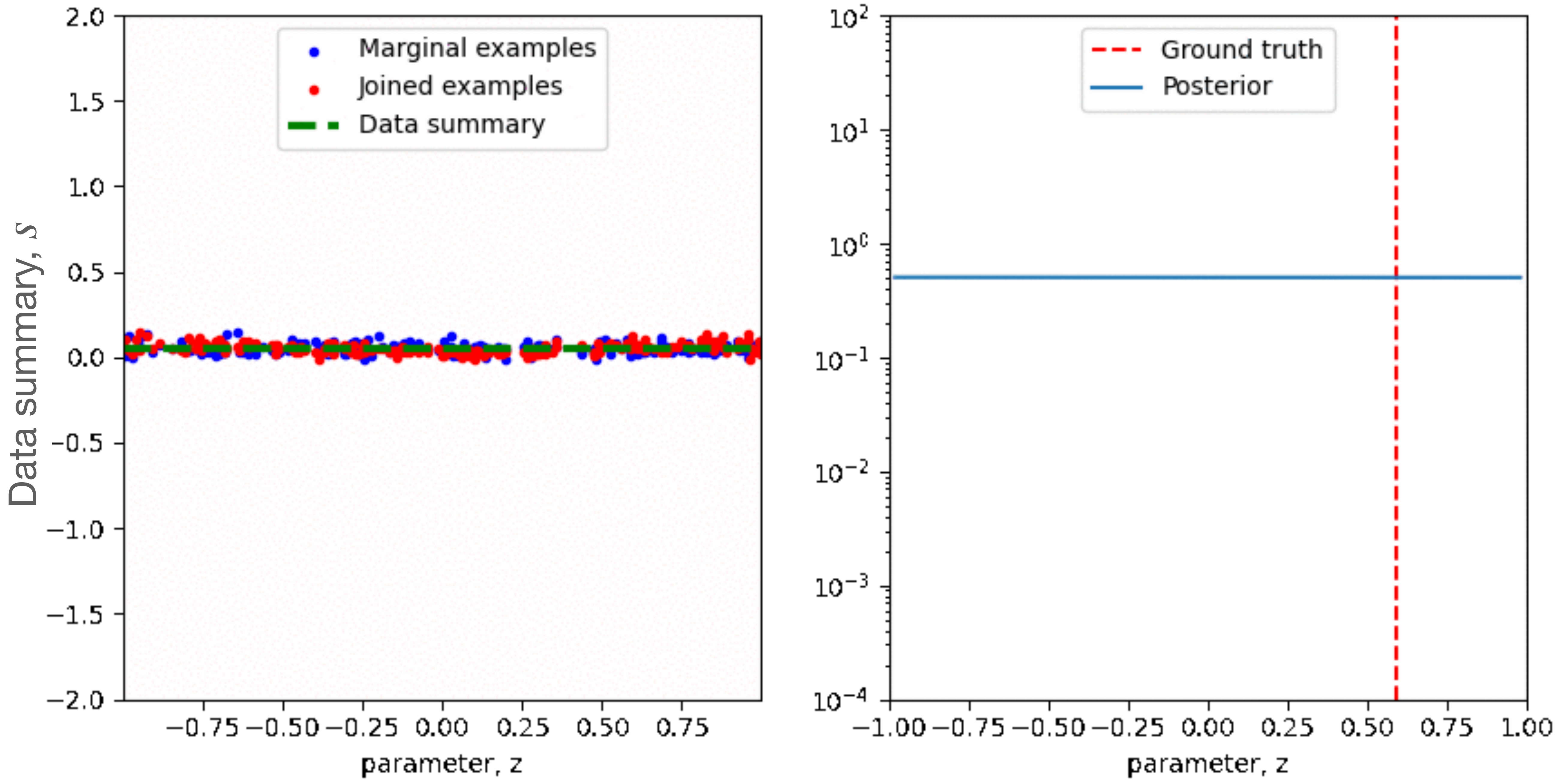
- Real sims: $z, \mathbf{x} \sim p(\mathbf{x}\,|\,z)p(z)$
- Scrambled sims: $z, \mathbf{x} \sim p(\mathbf{x})p(z)$



Embedding: Learns informative data summaries

MLP: Learns correlation between data summary and parameters

$$f_1 = \ln \frac{p(z_1\,|\,\mathbf{x})}{p(z_1)} \qquad f_2 = \ln \frac{p(z_2\,|\,\mathbf{x})}{p(z_2)} \qquad f_3 = \ln \frac{p(z_3, z_4\,|\,\mathbf{x})}{p(z_3, z_4)}$$



Note: After training, results are **amortised**, meaning that we can immediately access posteriors for any new observation $\mathbf{x}$.

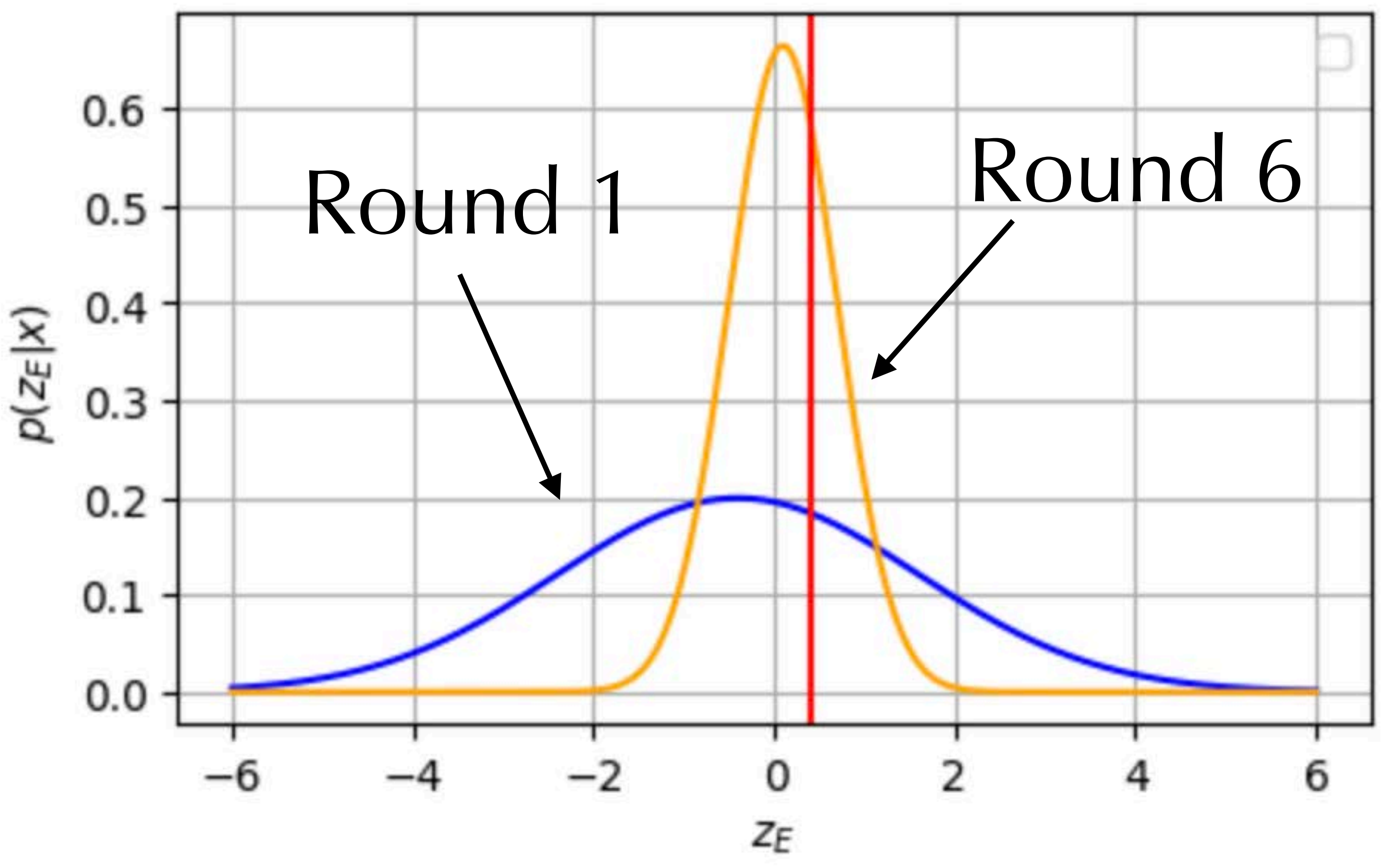
Miller+ 2011.13951, 2107.01214 - swyft & TMNRE

Sequential implicit inference

Papamakarios & Murray 1605.06376, Hermans+ 2020, Durkan+ 2020, Delaunoy+ 2022, Miller+ 2022

Sequential inference proceeds in multiple focus rounds

```
Initialise q_ϕ^(0)(z | x) = p(z)
For r = 1,…,R
  Simulate training data via x, z ~ p(x | z) q_ϕ^(r−1)(z | x_o)
  Train q_ϕ^(r)(z | x)  # Include some prior correction
Return q_ϕ^(R)(z | x) as posterior
```



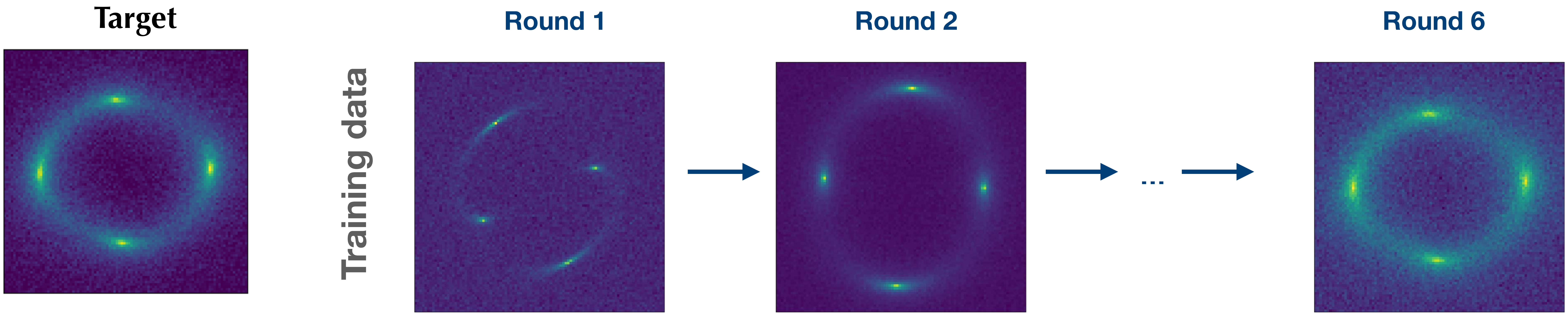
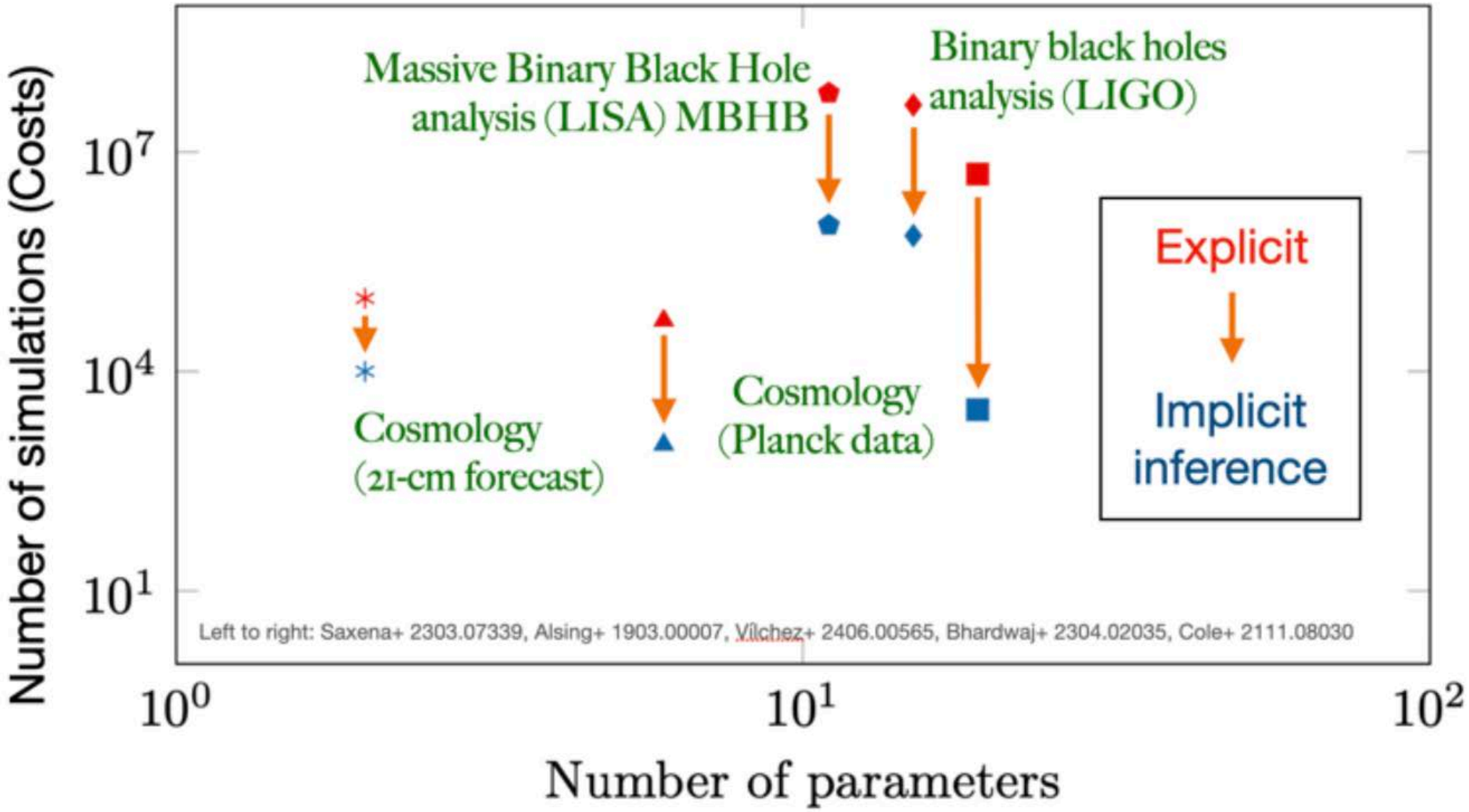


Image credit: Noemi Anau Montel

Potential benefits of implicit inference

Simulation-efficiency & scalability

Increases simulation-efficiency
(More information extracted per simulator run)



Decouples costs from model complexity
(More parameters do not always cause extra costs)

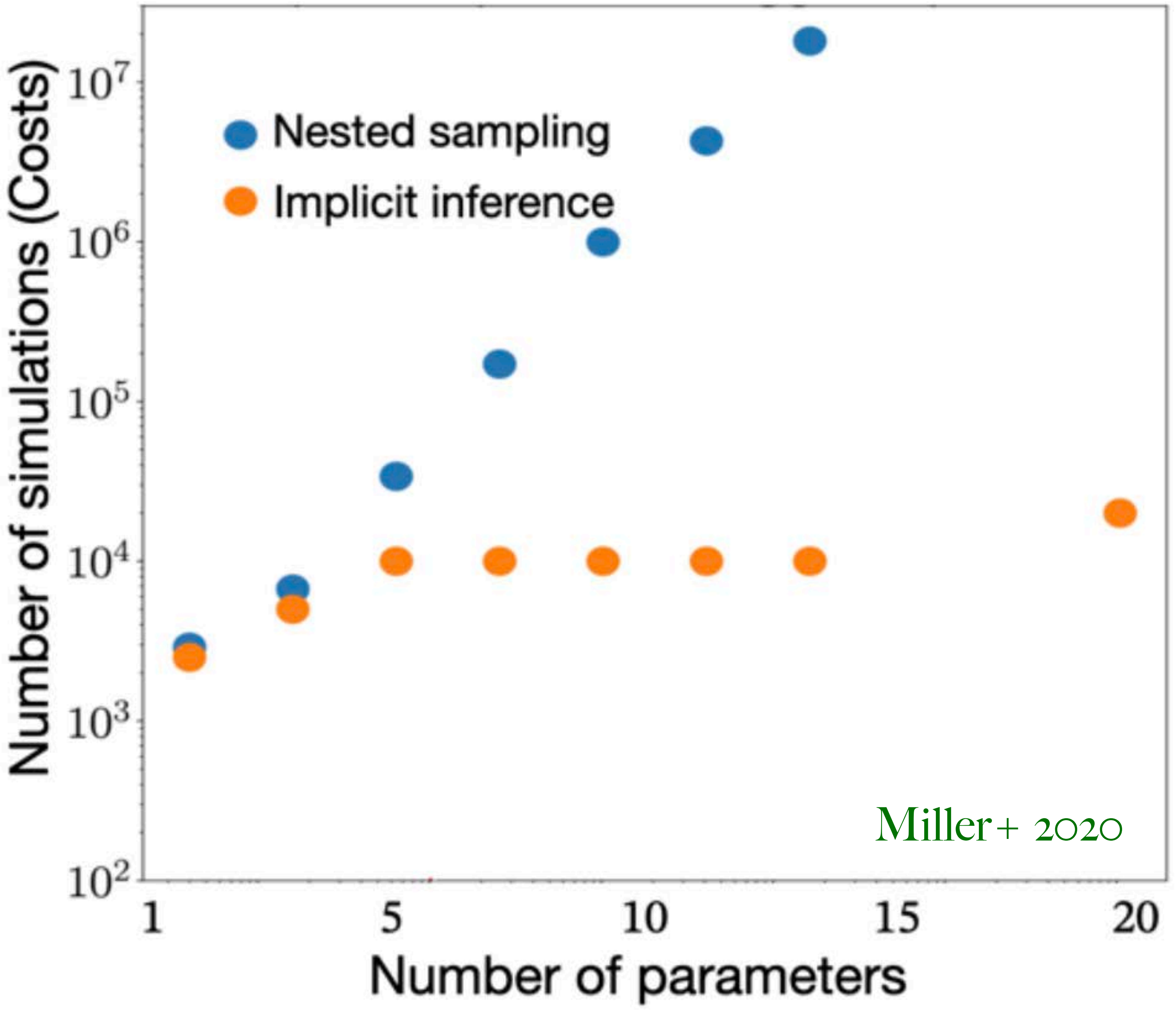


Figure credit: Noemi Anau Montel

Energy costs of analysing data

- **One open air concert (per 10.000 participants): ~1 MWh**
- One intercontinental flight (per person): ~3 MWh
- Amsterdam Karlsruhe train round-trip (per person):~100kWh

- Analysis of a non-standard cosmological model: ~100kWh
- Global analysis of multiple datasets (assuming 1 Mio CPU hours): ~7MWh

AI-based implicit inference
Non-standard cosmological model analysis
~1kWh

Source: Back-of-the-envelope estimates using ChatGPT 1o

Summary

- There are numerous connection points between AI technology, emergent AI and physics. This is a very dynamic field!
- The adoption of AI for physics research is still at an early stage. Examples include
  - Pattern recognition and detection.
  - Acceleration of computations and predictions.
  - Optimisation and discovery.
- AI is particularly effective for solving “inverse problems”, and is becoming quickly an indispensable tool for analysing current and upcoming astrophysical and cosmological data.

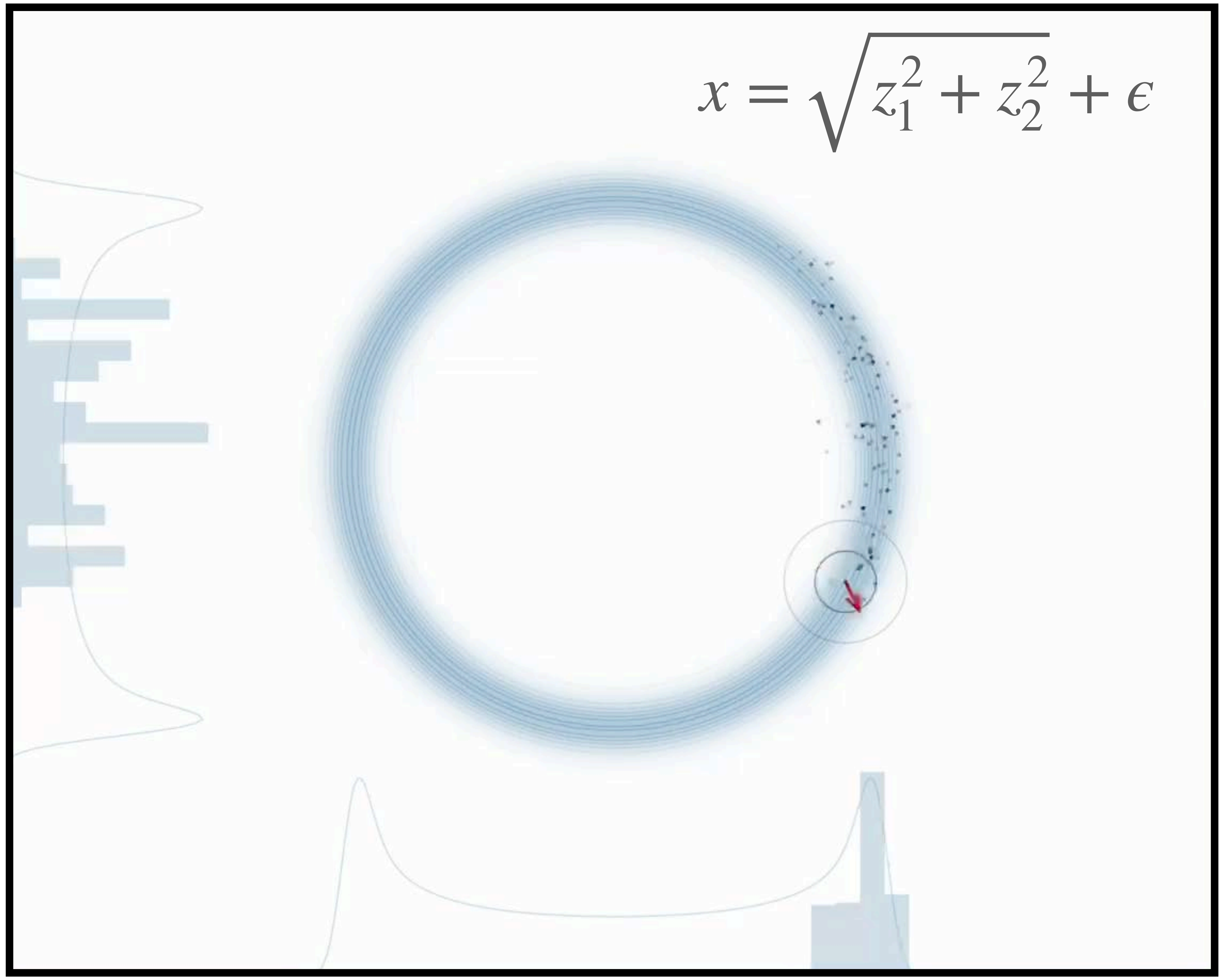
Thank you!

Backup

Numerical and Computational Approaches

Sampling based approaches - Markov Chain Monte Carlo (MCMC)

- **Challenges:** Non-trivial target distributions
- **Convergence**: Assessment through auto-correlation plots Gelman-Rubin statistics
- **Practical limitations**: MCMC might have problems with high-dimensional or multi-modal distributions. However, they are widely used in physics/astronomy



Speagle 1909.12313 (MCMC review)

https://chi-feng.github.io/mcmc-demo/app.html?algorithm=RandomWalkMH