From Optimization to Emergence: Collective Dynamics in Design and Inference

Hazhir Aliahmadi[1]

[1]Department of Physics, Engineering Physics, and Astronomy, Queen's University

June 24, 2026

From Observation to Laws

- Science starts with observation, but does not end with observation.

From Observation to Laws

- Science starts with observation, but does not end with observation.
- A law is not more data; it is the invariant structure that survives across many observations.

From Observation to Laws

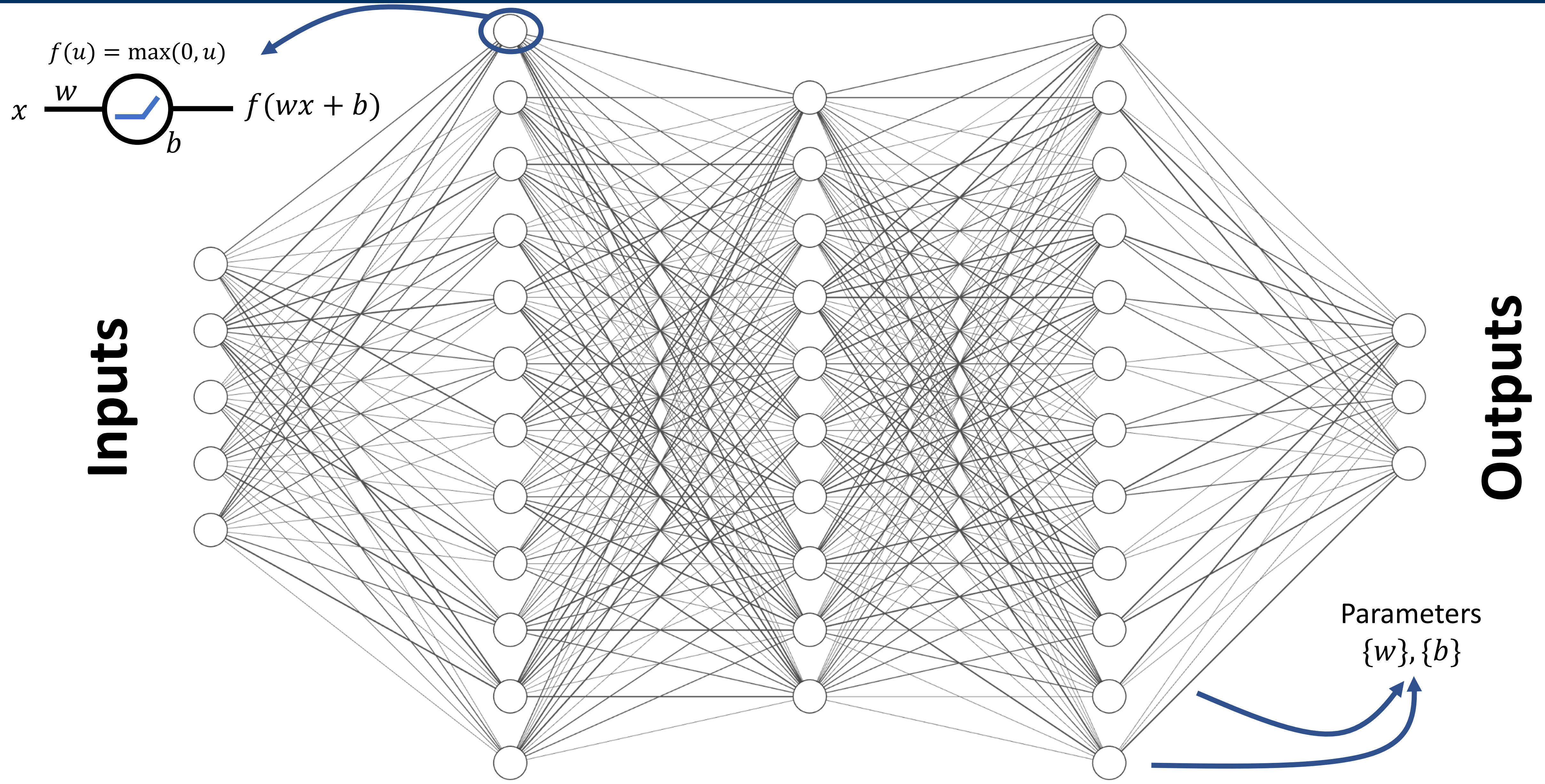
- Science starts with observation, but does not end with observation.
- A law is not more data; it is the invariant structure that survives across many observations.
- Symmetry decides what can survive; collective behaviour decides what can appear.

From Observation to Laws

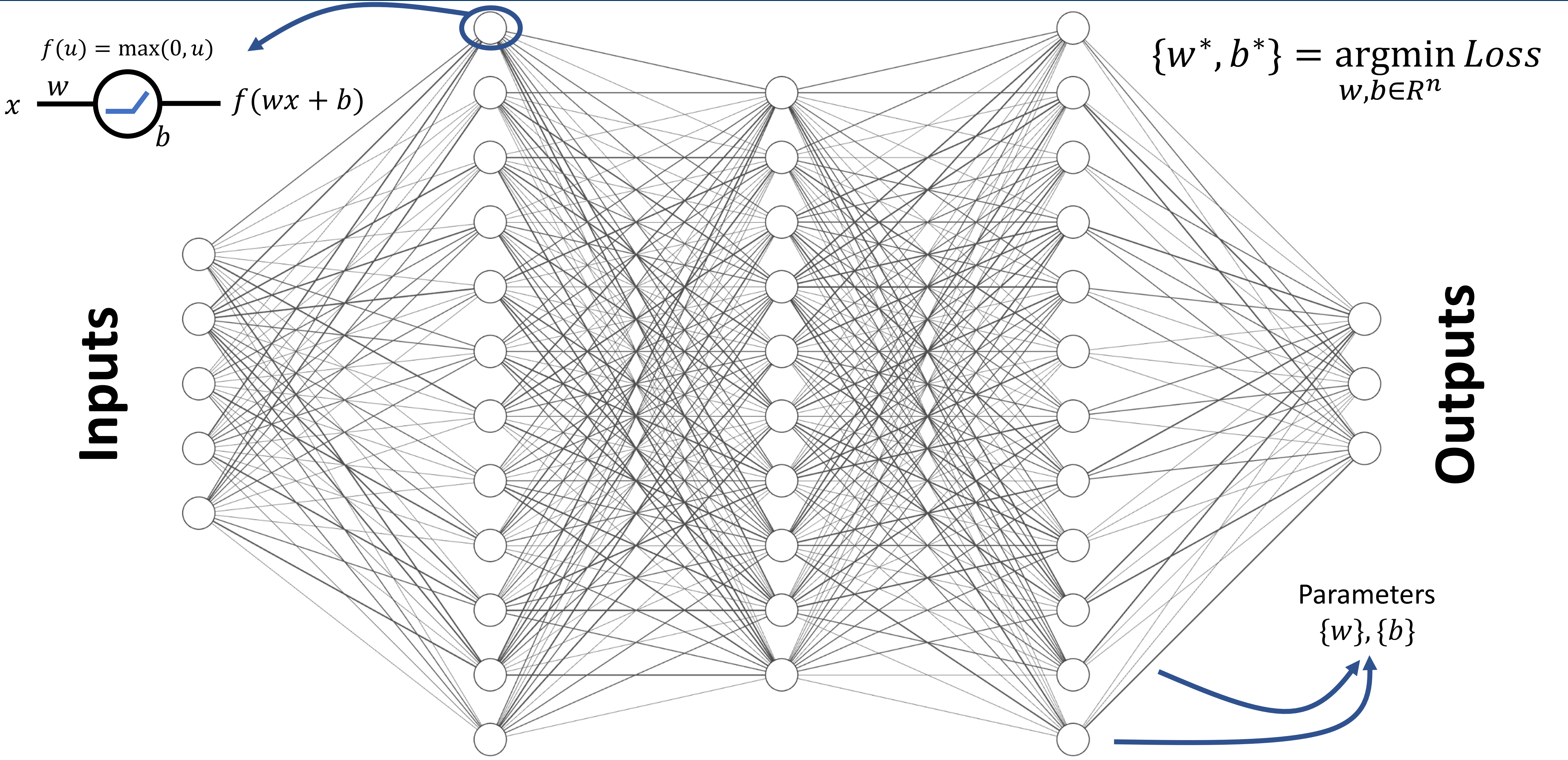
- Science starts with observation, but does not end with observation.
- A law is not more data; it is the invariant structure that survives across many observations.
- Symmetry decides what can survive; collective behaviour decides what can appear.

Laws emerge when the details collapse.

Neural networks: fit or law?



Neural networks: fit or law?



Neural networks: fit or law?

- A modern neural network is a huge nonlinear system with millions of degrees of freedom.

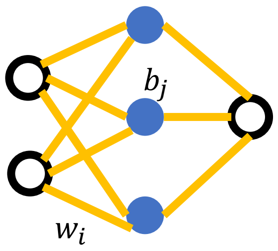
- Two concerns:

  Capacity without explanation
  More parameters make fitting easier, but they do not tell us what matters.

  Optimization without emergence
  If the result is only a fitted parameter vector, no collective law has appeared.

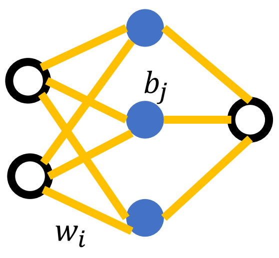
From One Minimum to an Ensemble

Network view

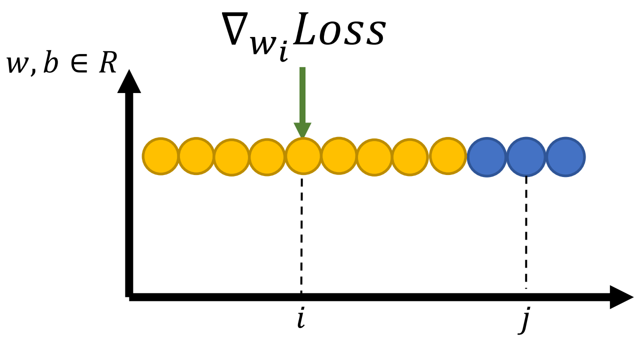


From One Minimum to an Ensemble

Network view

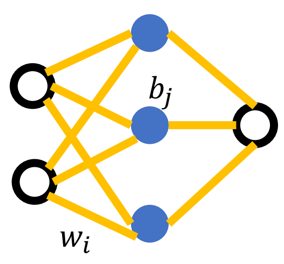


Parameter-as-particle view

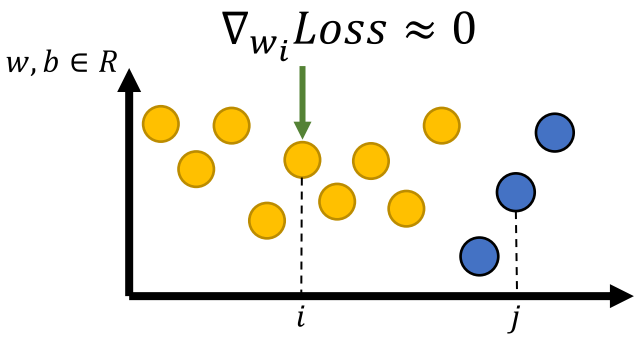


From One Minimum to an Ensemble

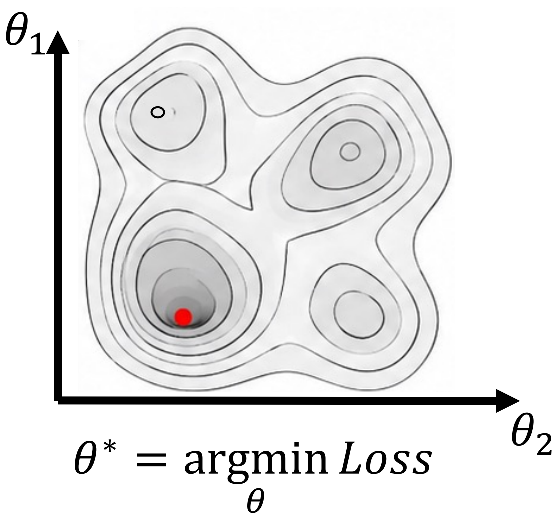
Network view



Parameter-as-particle view

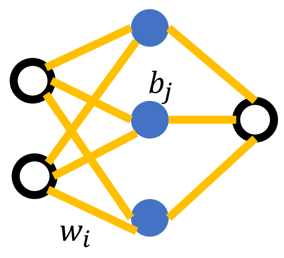


Landscape view

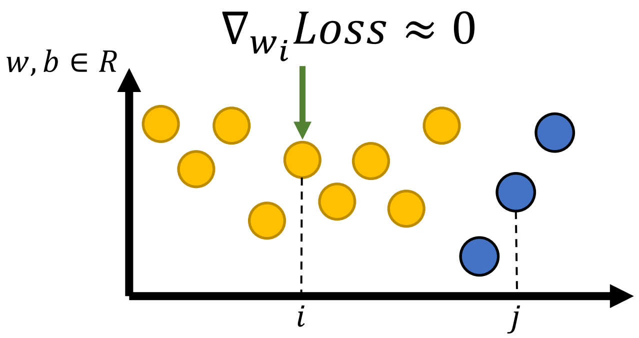


From One Minimum to an Ensemble

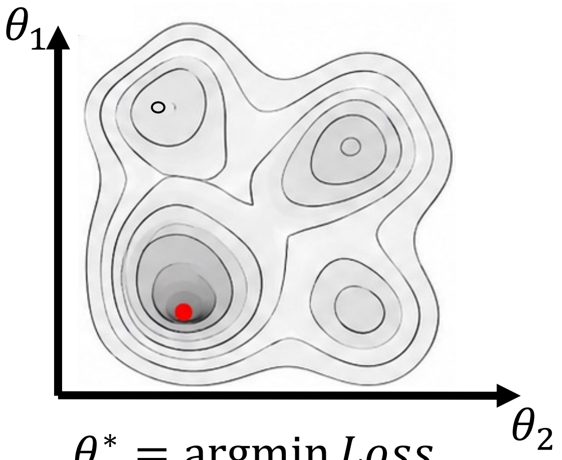
Network view



Parameter-as-particle view

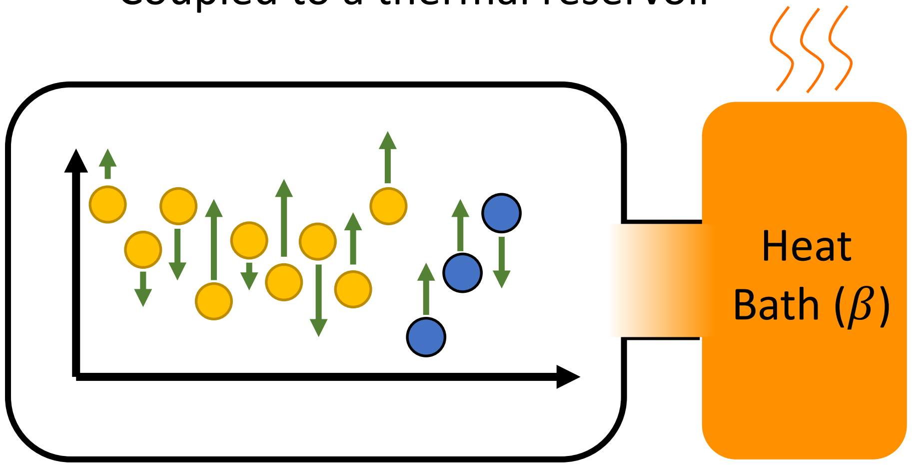


Landscape view



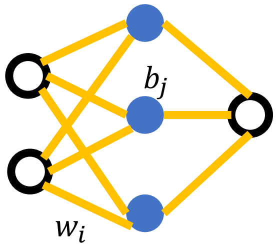
$$\theta^* = \underset{\theta}{\operatorname{argmin}}\, Loss$$

Coupled to a thermal reservoir

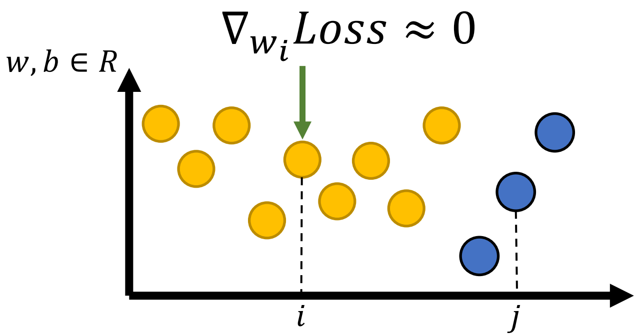


From One Minimum to an Ensemble

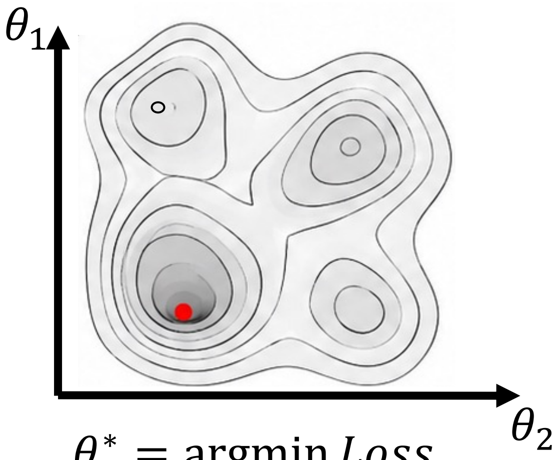
Network view



Parameter-as-particle view

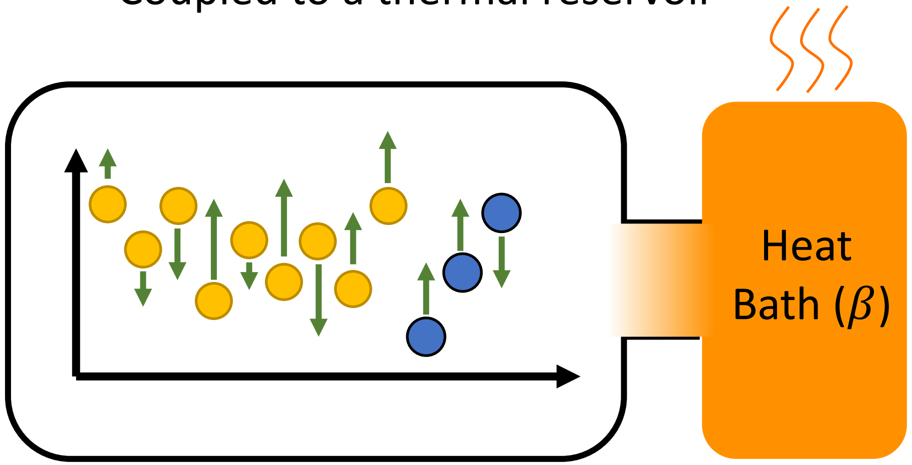


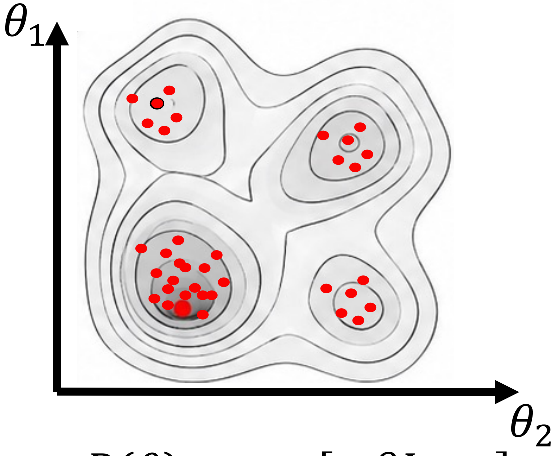
Landscape view



$$\theta^* = \underset{\theta}{\operatorname{argmin}}\, Loss$$

Coupled to a thermal reservoir

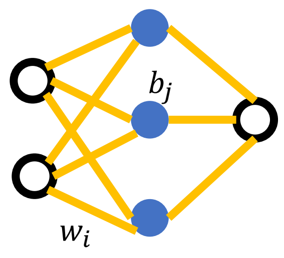




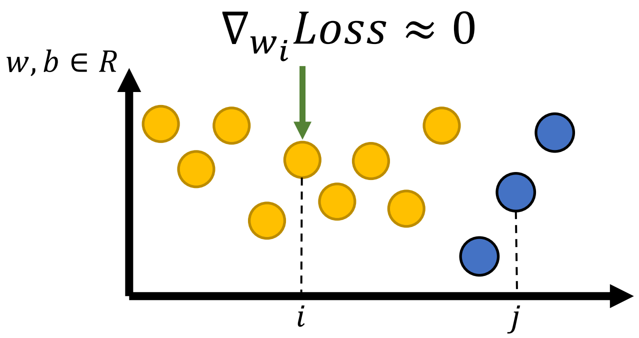
$$P(\theta) \propto \exp[-\beta Loss]$$

From One Minimum to an Ensemble

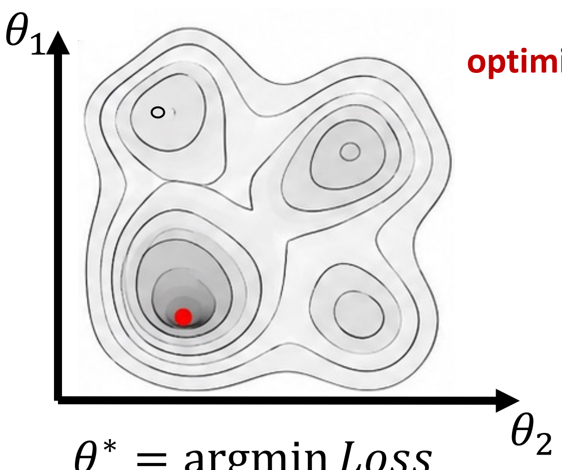
Network view



Parameter-as-particle view

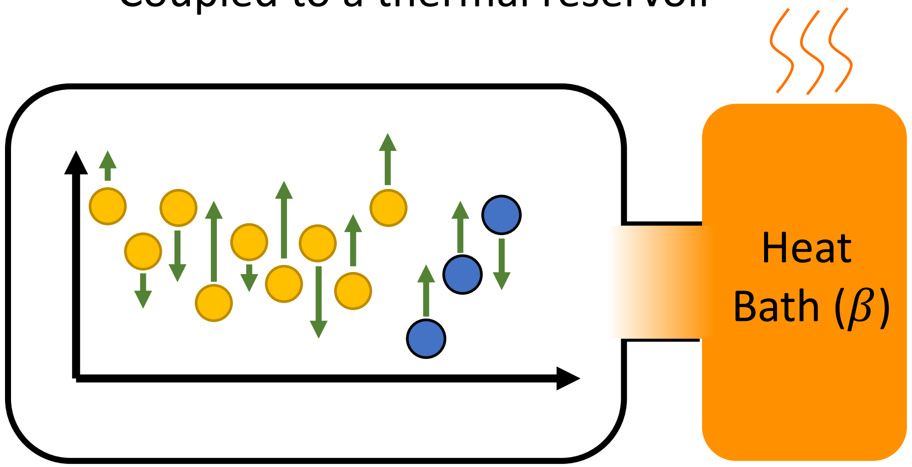


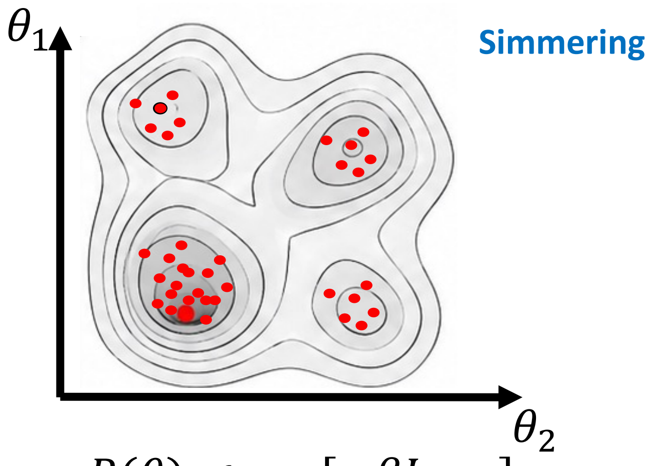
Landscape view



$$\theta^* = \underset{\theta}{\operatorname{argmin}}\, Loss$$

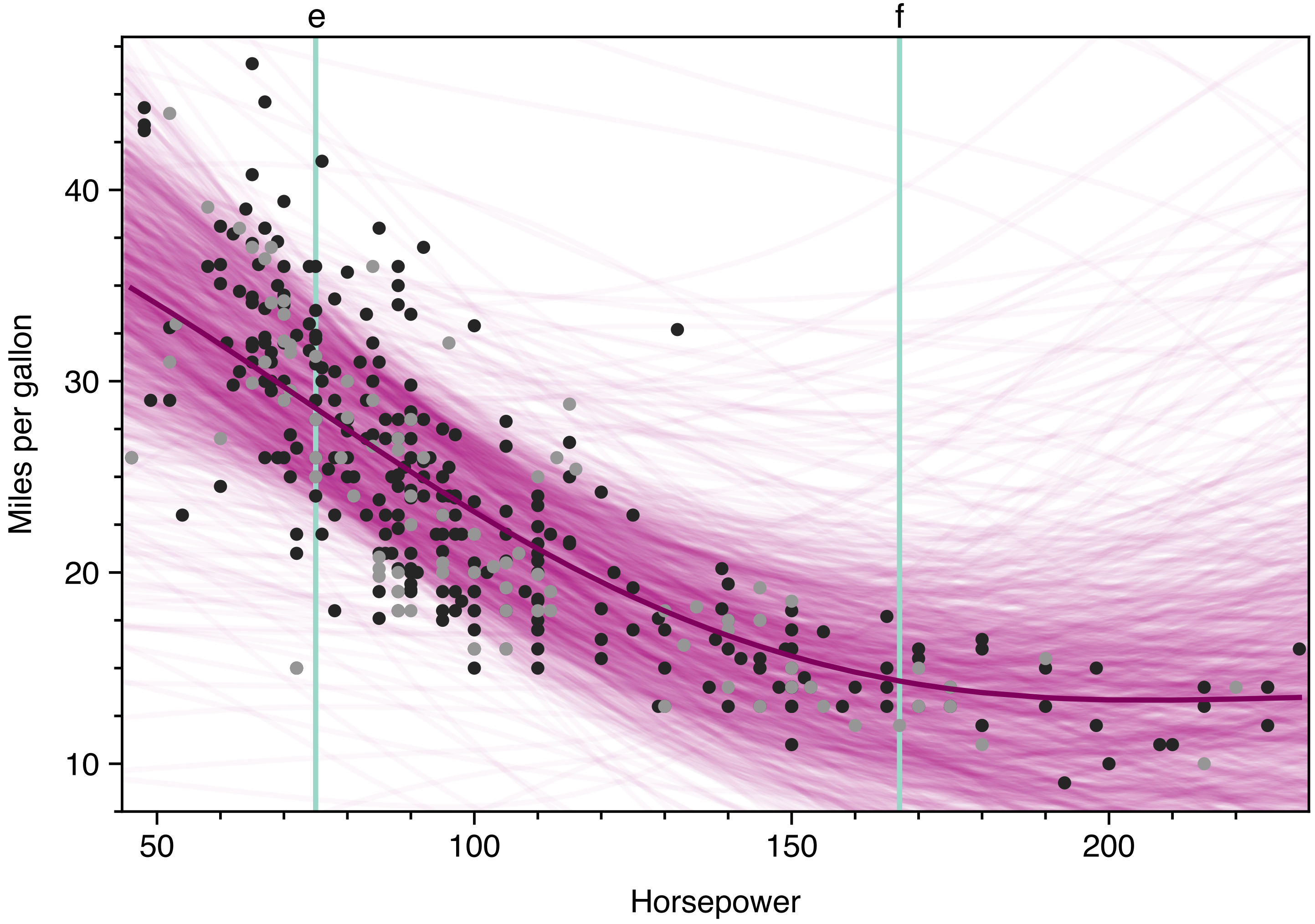
Coupled to a thermal reservoir





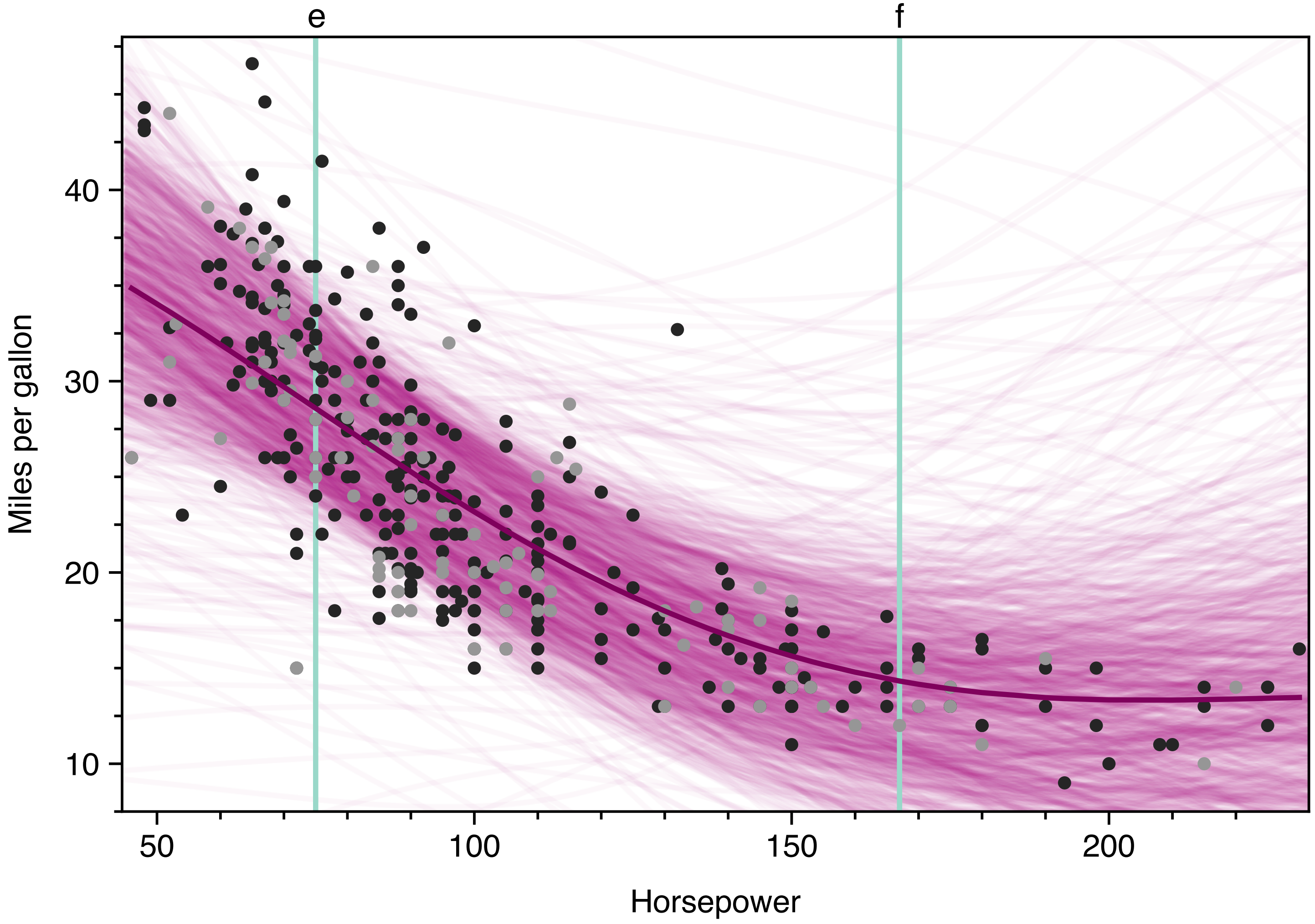
$$P(\theta) \propto \exp[-\beta Loss]$$

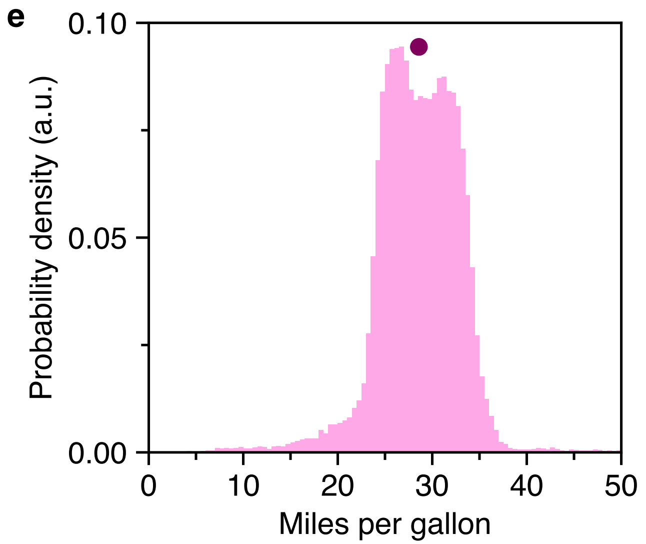
Regression

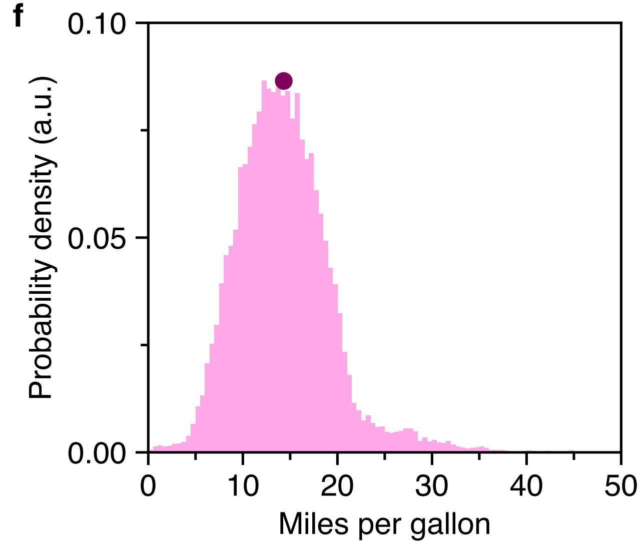


Babayan & **Aliahmadi**, with van Anders, *Nat. Commun.* **17**, 271 (2026)

Regression

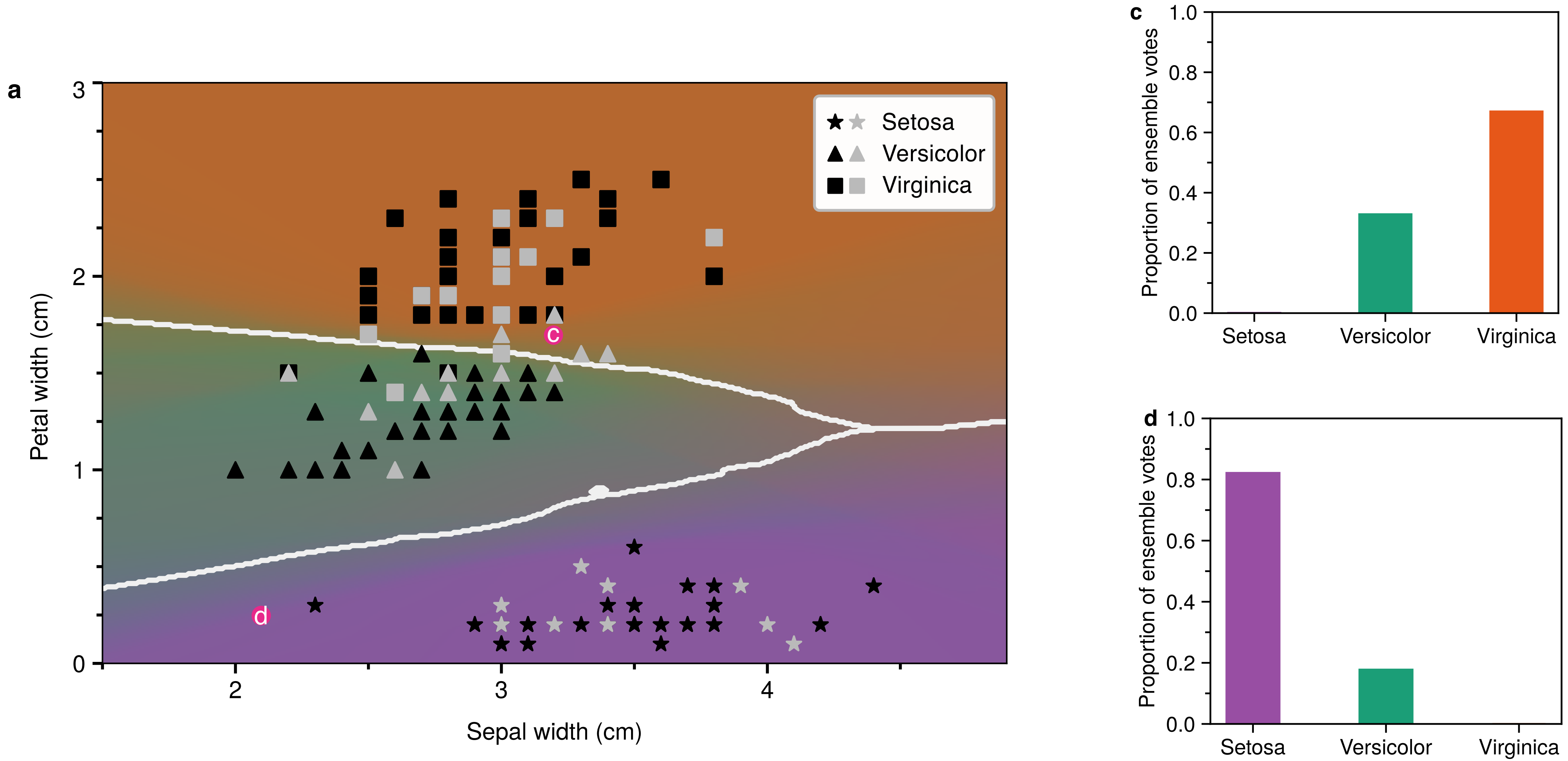






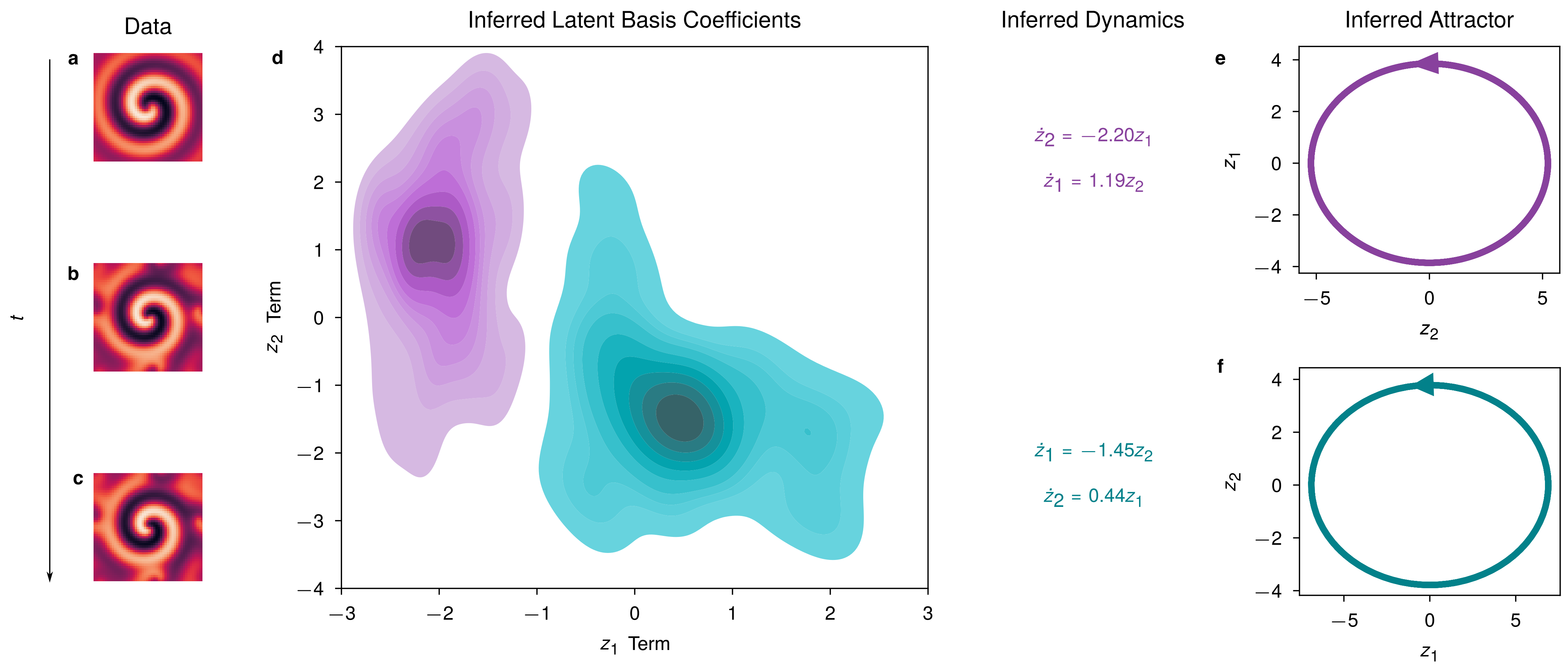
Babayan & **Aliahmadi**, with van Anders, *Nat. Commun.* **17**, 271 (2026)

Classification



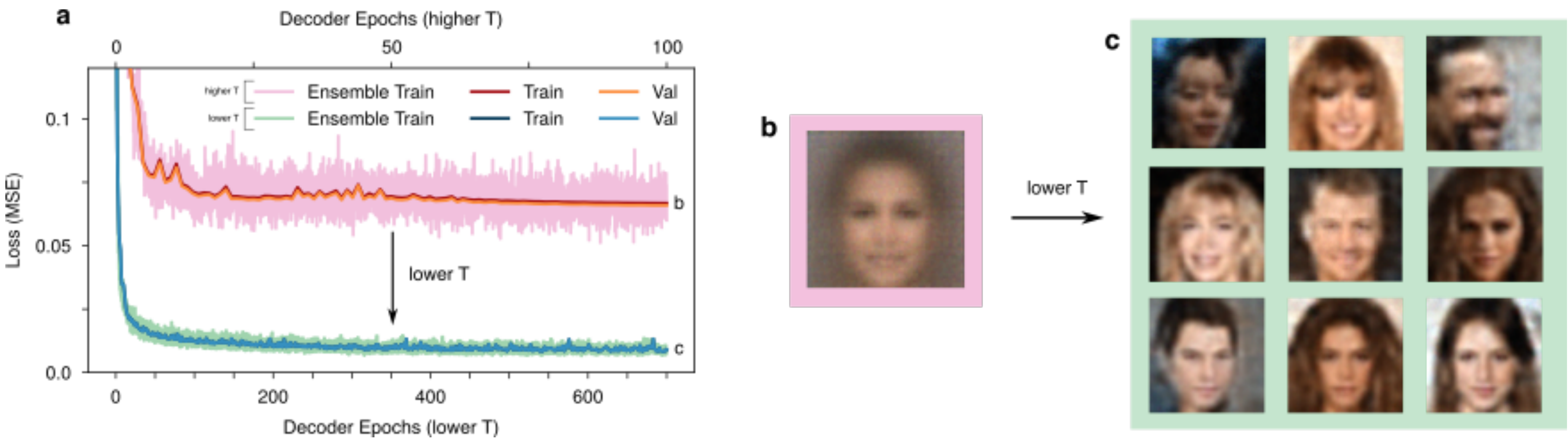
Babayan & **Aliahmadi**, with van Anders, *Nat. Commun.* **17**, 271 (2026)

Discovering governing equations



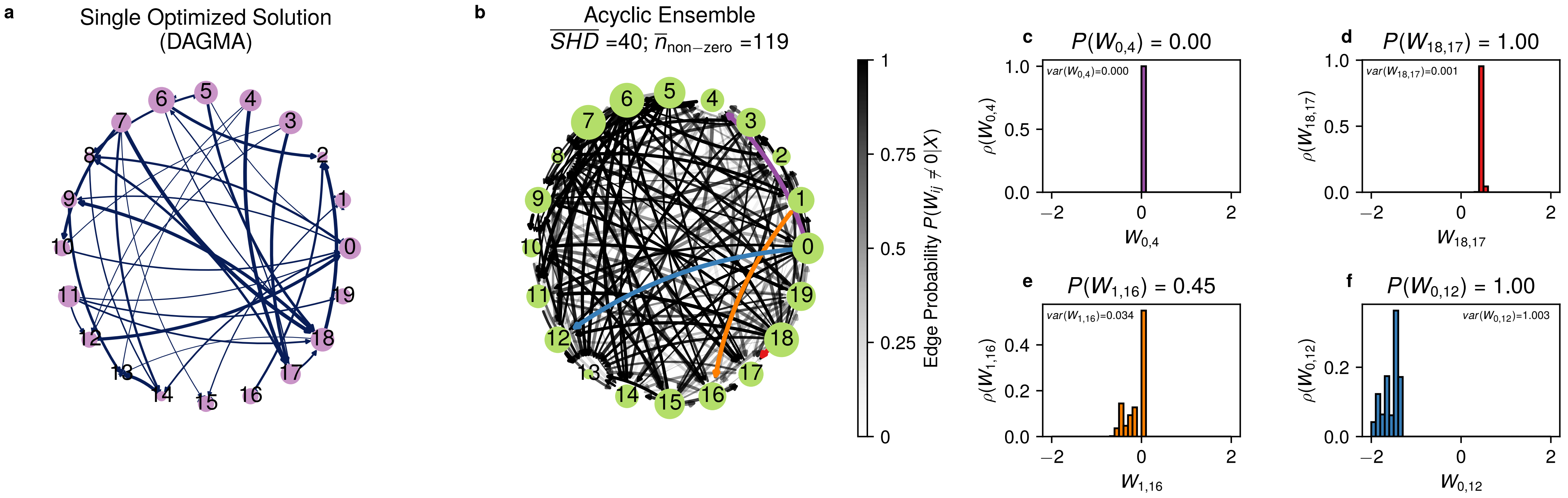
Aliahmadi & Babayan, with van Anders**, arXiv:2605.16164 (2026)**

Generative AI



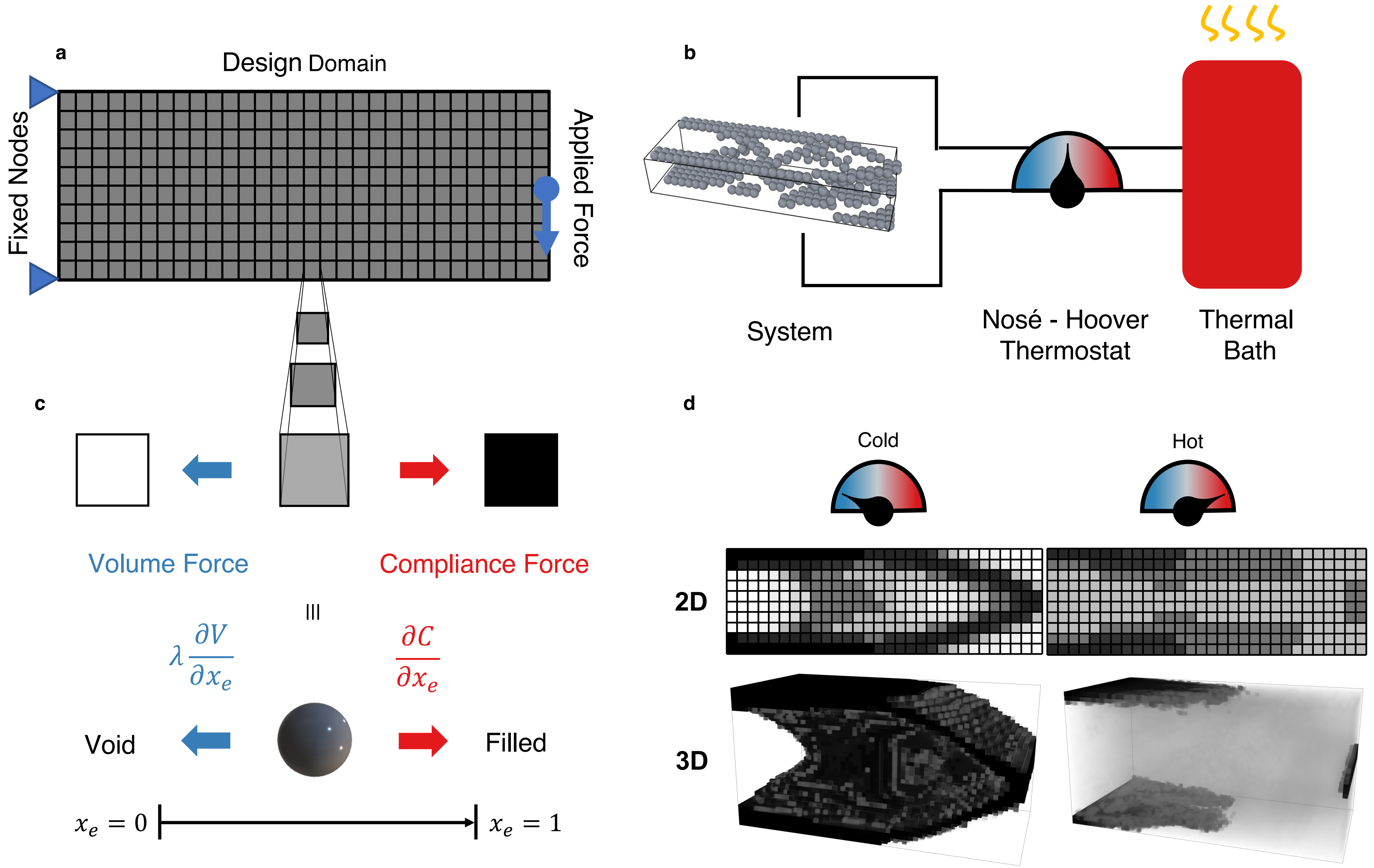
Aliahmadi *et al.*, arXiv:2605.16164 (2026)

Bayesian Network



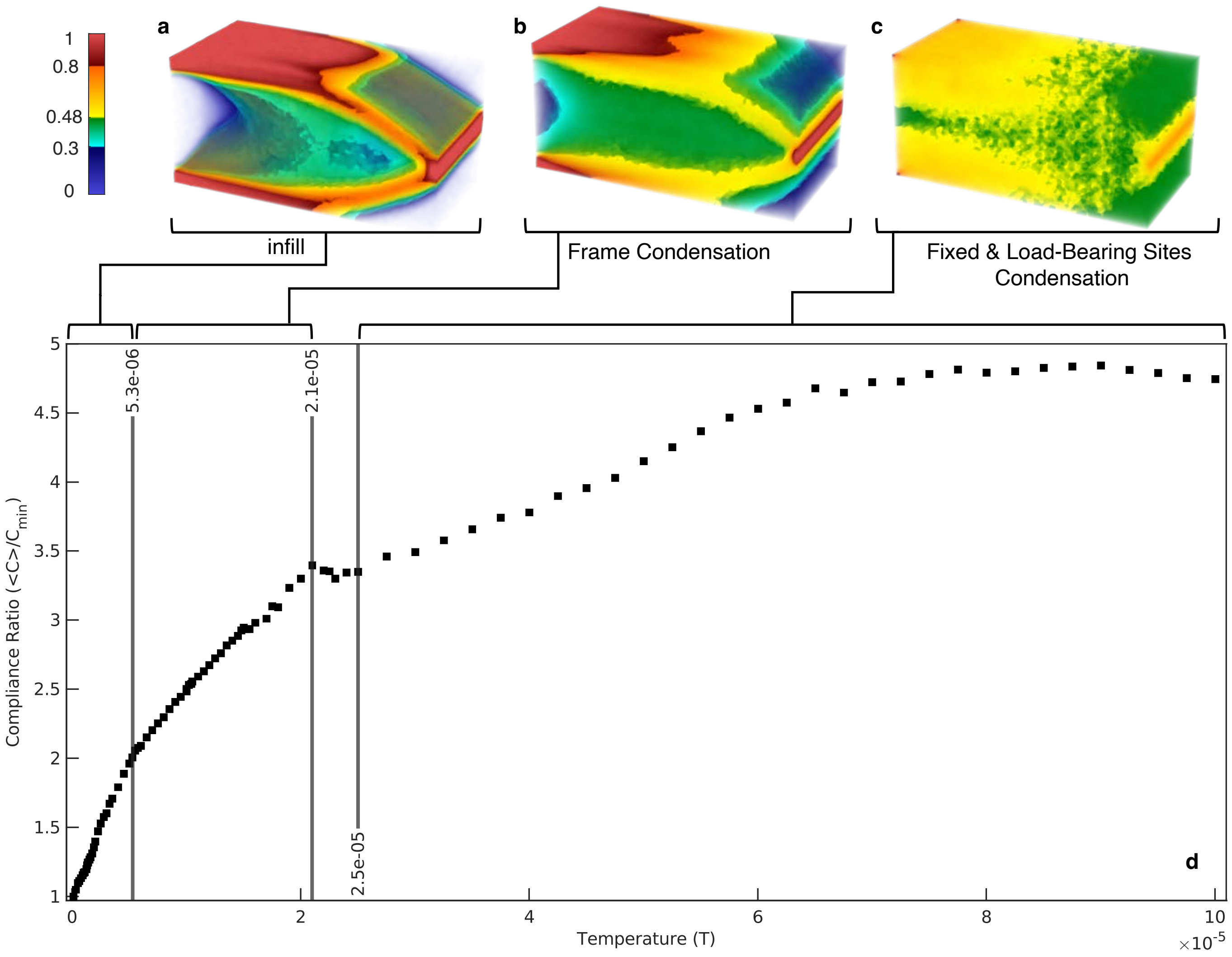
Aliahmadi *et al.*, arXiv:2606.06440 (2026)

Topology and Shape Optimization



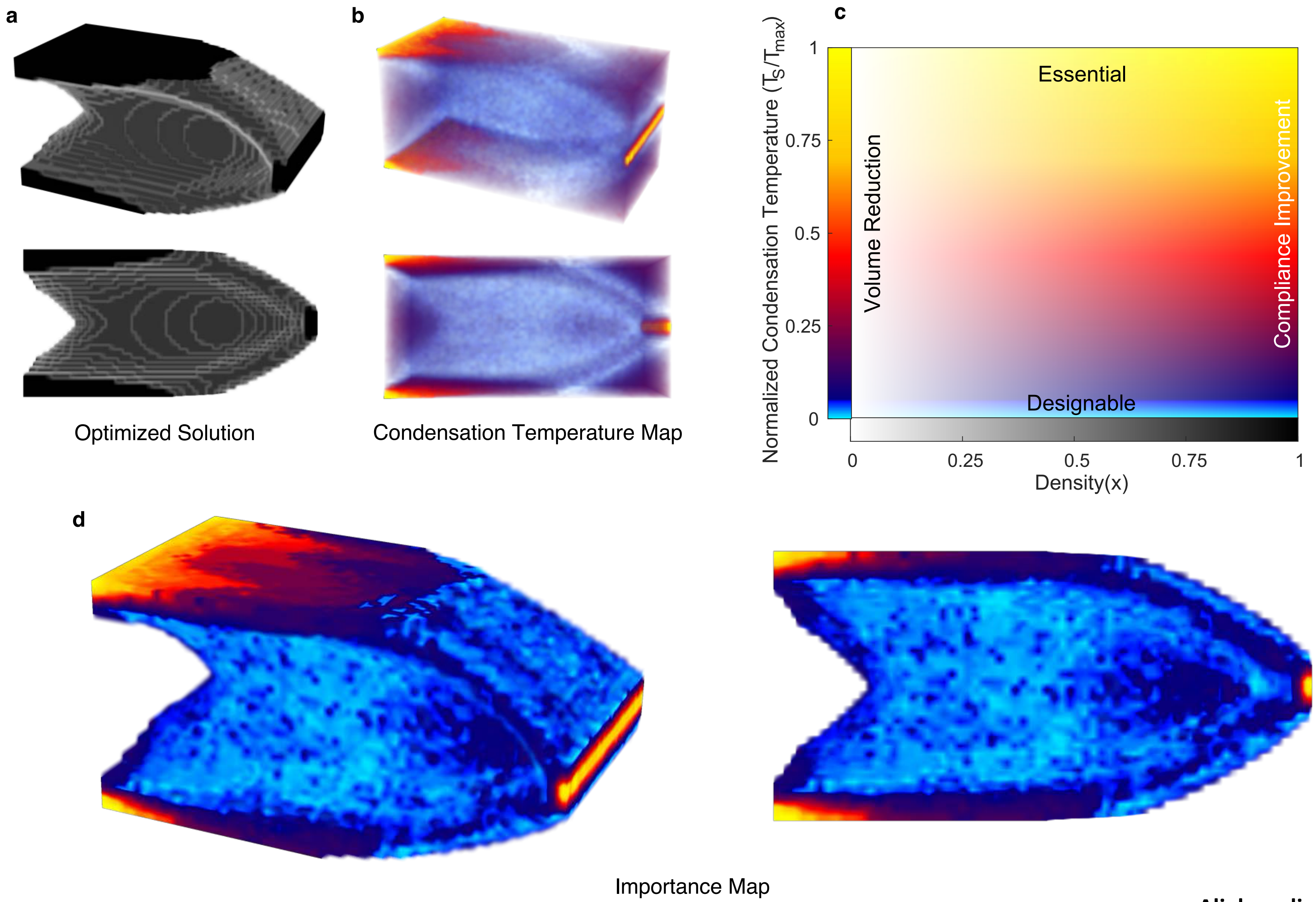
Aliahmadi *et al.*, arXiv:2503.16211 (2025)

Topology and Shape Optimization



Aliahmadi *et al.*, arXiv:2503.16211 (2025)

Topology and Shape Optimization



Aliahmadi *et al.*, arXiv:2503.16211 (2025)

Applications

- **Machine Learning & SciML**
- **Probabilistic Modelling & Causal Inference**
- **Dynamical Systems & Digital Twins**
- **Control & Decision-Making**
- **Computational Design**
- **Quantitative Finance**
- **…**

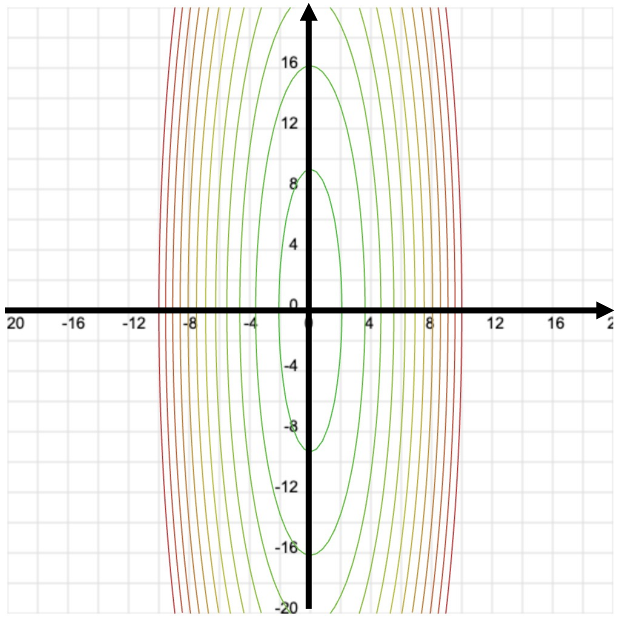
Free Energy Minimization

$$\min F = Loss - T \ln \Omega$$

Free Energy Minimization

$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ Sloppy Direction

Multiplicity



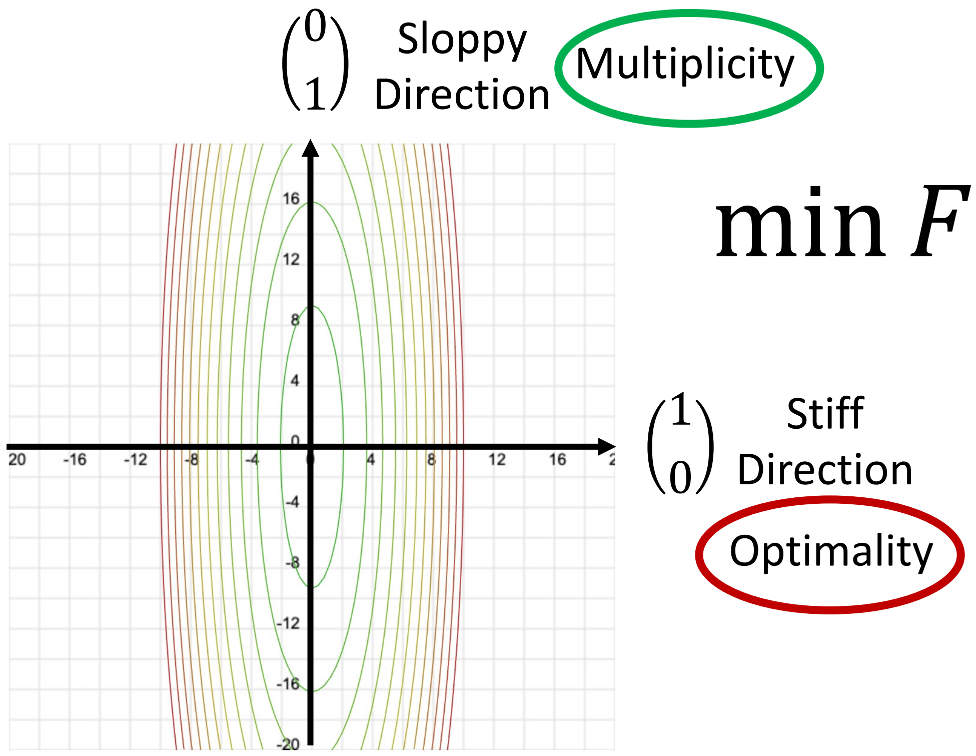
$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Stiff Direction

Optimality

Multiplicity

$$\min F = Loss - T \ln \Omega$$

Free Energy Minimization



Multiplicity

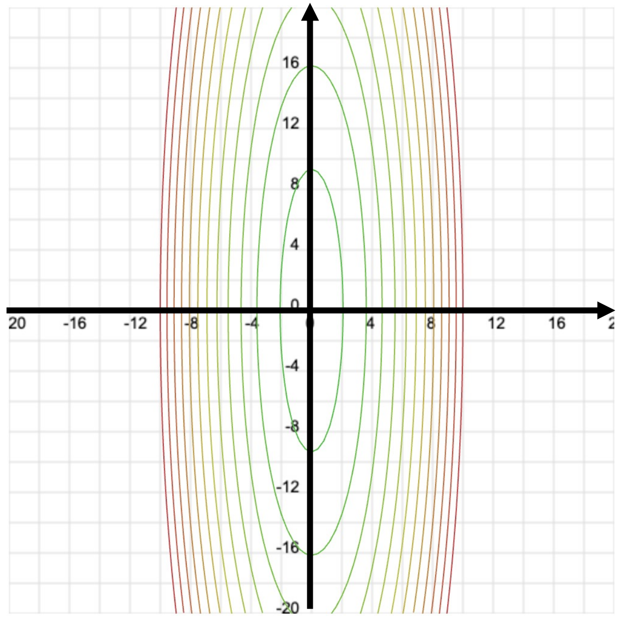
$$\min F = Loss - T \ln \Omega$$

Objective Function

Free Energy Minimization

$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ Sloppy Direction

Multiplicity



$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Stiff Direction

Optimality

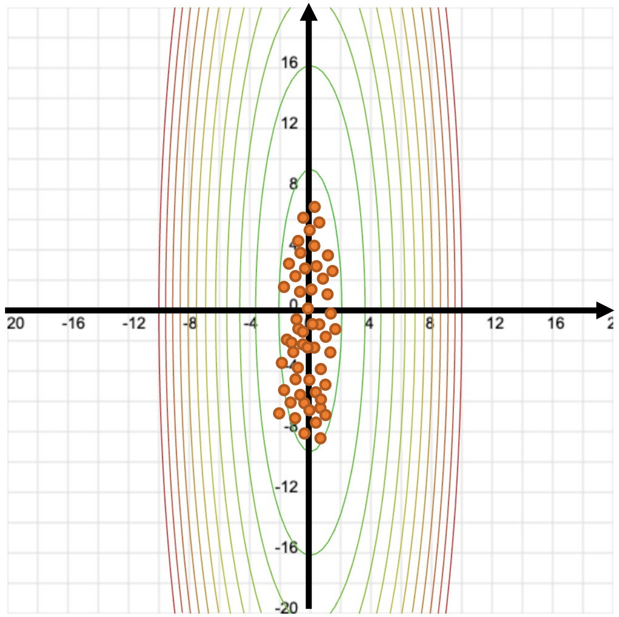
$$\min F = Loss - T \ln \Omega$$

Objective Function (Loss)

Regularizer ($T \ln \Omega$)

Free Energy Minimization

$\begin{pmatrix}0\\1\end{pmatrix}$ Sloppy Direction — Multiplicity



$\begin{pmatrix}1\\0\end{pmatrix}$ Stiff Direction — Optimality

$$\min F = \underbrace{Loss}_{\text{Objective Function}} - T \underbrace{\ln \Omega}_{\text{Regularizer}}$$

Objective Function

Regularizer

Free Energy Minimization $F = Loss - T \ln \Omega$

$$\min_{\theta} F(\theta)$$

Free Energy Minimization $F = Loss - T \ln \Omega$

$$\min_{\theta} F(\theta)$$

Collective Coordinate

$$\theta = \theta(w)$$

Free Energy Minimization $F = Loss - T \ln \Omega$

$$\min_{\theta} F(\theta)$$

Collective Coordinate which we DO NOT know

$$\theta = \theta(w)$$

Free Energy Minimization $F = Loss - T \ln \Omega$

$$Loss(w)$$

Model Coordinate

which we DO know

$$\min_{\theta} F(\theta)$$

Collective Coordinate

which we DO NOT know

$$\theta = \theta(w)$$

Free Energy Minimization $F = Loss - T \ln \Omega$

$$\operatorname{sim}_{w} Loss(w) \quad \equiv \quad \min_{\theta} F(\theta)$$

Model Coordinate which we DO know

Collective Coordinate which we DO NOT know

$\theta = \theta(w)$

Bayesian Inference on Collective Coordinate

$$F = Loss - T \ln \Omega \quad \xrightarrow{\beta = 1/T} \quad e^{-\beta F} = e^{-\beta Loss}\Omega$$

Bayesian Inference on Collective Coordinate

$$F = Loss - T \ln \Omega \quad \xrightarrow{\beta = 1/T} \quad \underbrace{e^{-\beta F}}_{p(\theta|Y)} = \underbrace{e^{-\beta Loss}}_{p(Y|\theta)} \underbrace{\Omega}_{p(\theta)}$$

$$p(\theta|Y) \propto p(Y|\theta)\, p(\theta)$$

A posteriori distribution | **Likelihood** | **A priori distribution**

We inferred
Not assumed

Bayesian Inference on Collective Coordinate

$$F = Loss - T \ln \Omega \quad \xrightarrow{\beta = 1/T} \quad \underbrace{e^{-\beta F}}_{p(\theta|Y)} = \underbrace{e^{-\beta Loss}}_{p(Y|\theta)} \underbrace{\Omega}_{p(\theta)}$$

$$p(\theta|Y) \propto p(Y|\theta)\, p(\theta)$$

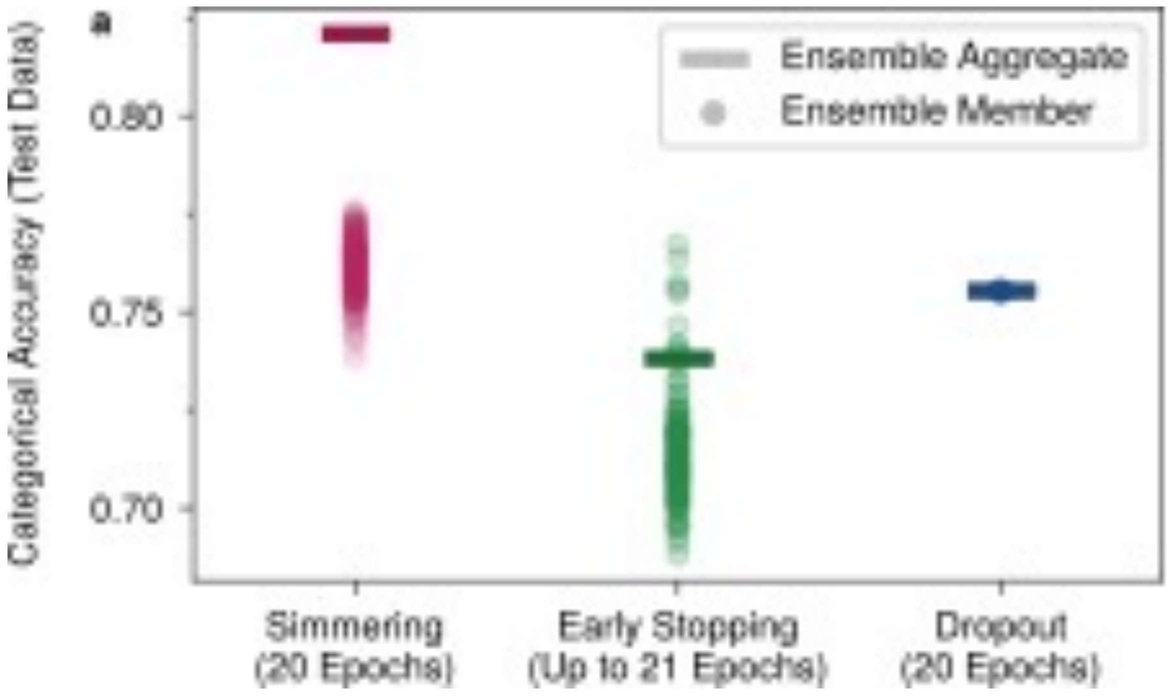
A posteriori distribution | **Likelihood** | **A priori distribution**

We inferred
Not assumed

Implicit Bayesian Inference

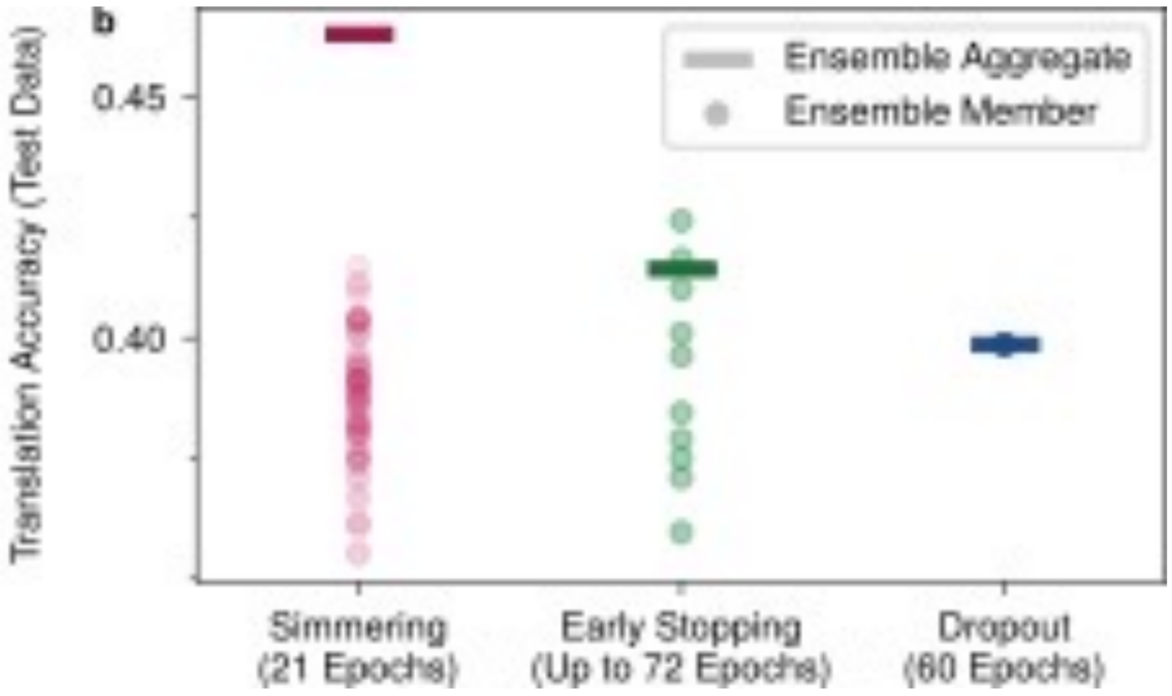
Implicit Bayesian Inference

CIFAR-10 classification problem



Representative Ensemble

Portuguese-English transcript translation (Transformer)

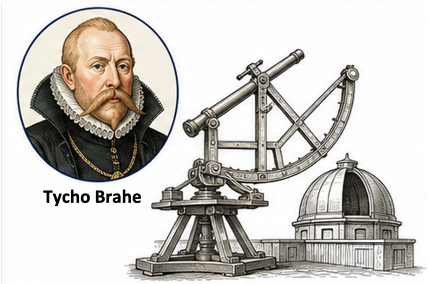


Representative Ensemble

Aliahmadi & Babayan, with van Anders**, arXiv:2605.16164 (2026)**

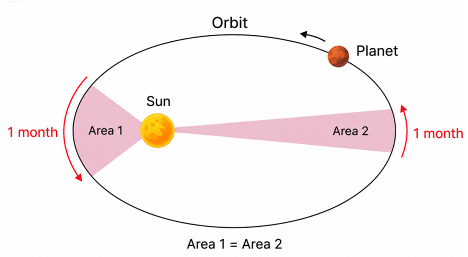
THE PAST

Measurements



Date	Mars longitude λ	Mars latitude β
1582-01-20	λ_1	β_1
1582-01-31	λ_2	β_2
1582-02-10	λ_3	β_3
...	...	...

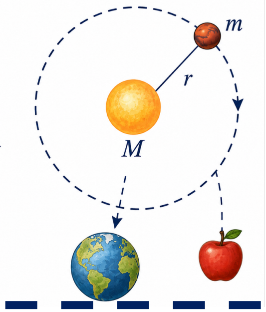
Patterns



Laws

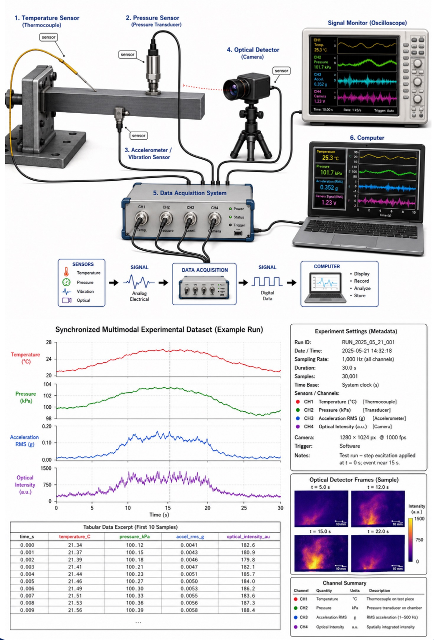
$$F = ma$$

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

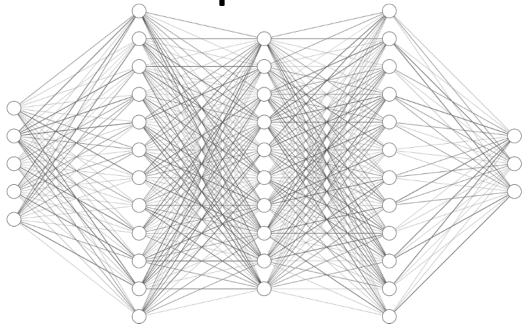


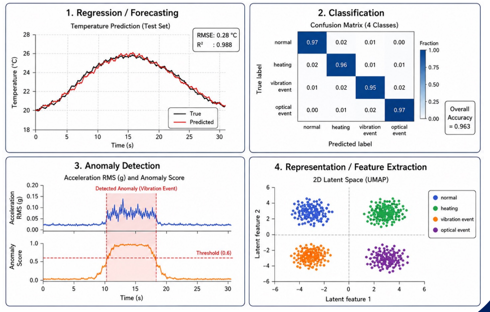
TODAY

Data



"Interpolation"





Understanding

Simmering Paper

Topology Optimization

Simmering Code

Generative AI

Bayesian Network