

Statistical Physics and AI

Outline

1. Statistical Physics, Phase Diagram and Machine Learning
2. Boltzmann Machine (BM)
3. Applications (Using Physics studies Physics)



CUHK-SZ



<https://mypage.cuhk.edu.cn/academics/yiyin/>



The Nobel Prize in Physics 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Physics 2024 to

John J. Hopfield

Princeton University, NJ, USA

Geoffrey Hinton

University of Toronto, Canada

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”

They trained artificial neural networks using physics

In 1982, John Hopfield invented a network that uses a method for **saving and recreating** patterns. He found inspiration **in physics' models...** The invention became important in, for example, image analysis.

In 1983–1985, Geoffrey Hinton used tools from **statistical physics to create the Boltzmann machine**, which can learn to **recognise characteristic elements** in a set of data. The invention became significant, for example, for classifying and creating images.

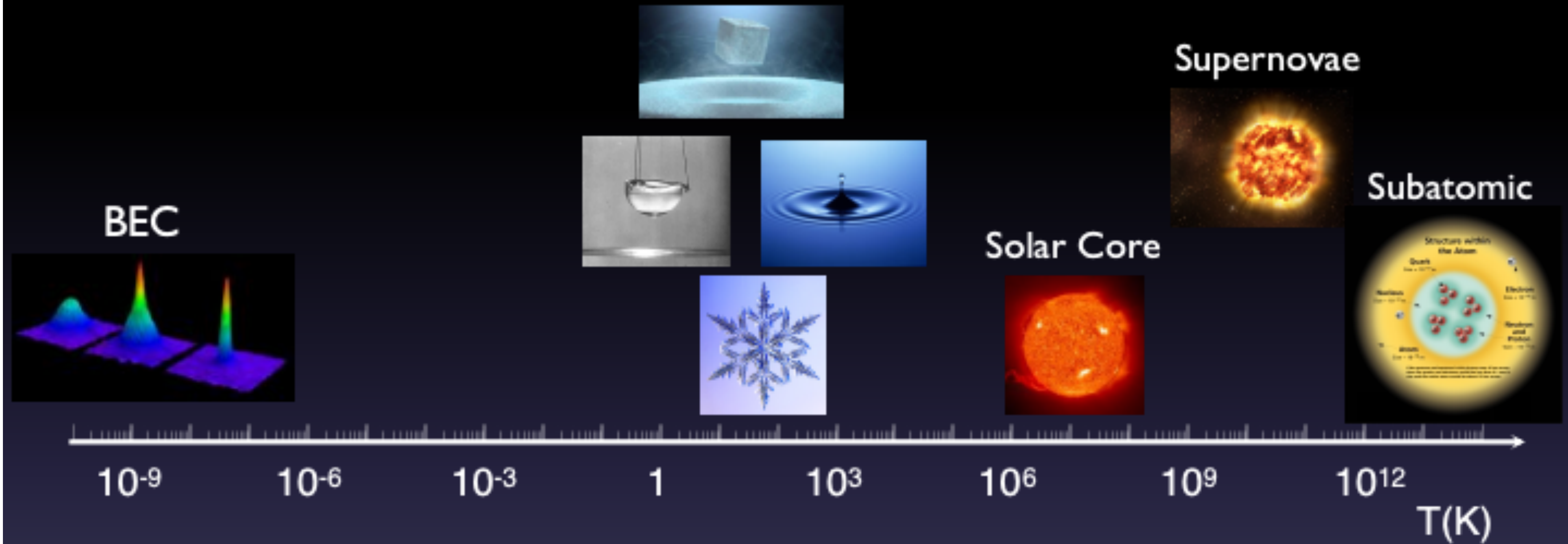
nobelprize.org

Outline

1. Statistical Physics, Phase Diagram and Machine Learning
2. Boltzmann Machine (BM)
3. Applications (Using Physics studies Physics)

I. Statistical Physics, Phase Diagram and Machine Learning

Phases of Matter

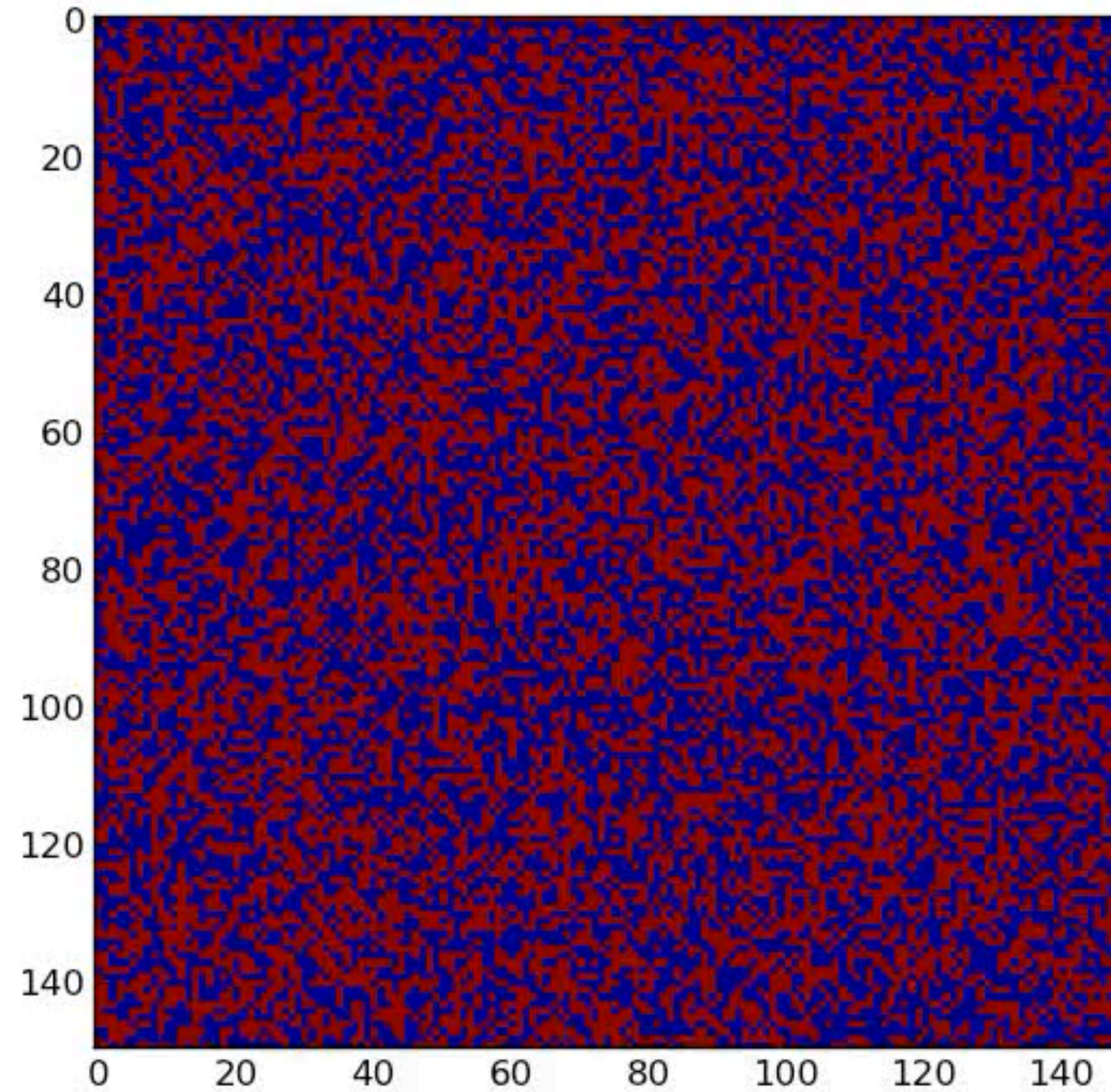


Simple Equation Rich Phenomena

“Often, people in some unjustified fear of physics say you can’t write an equation for life. Well, perhaps we can.” - Feynman

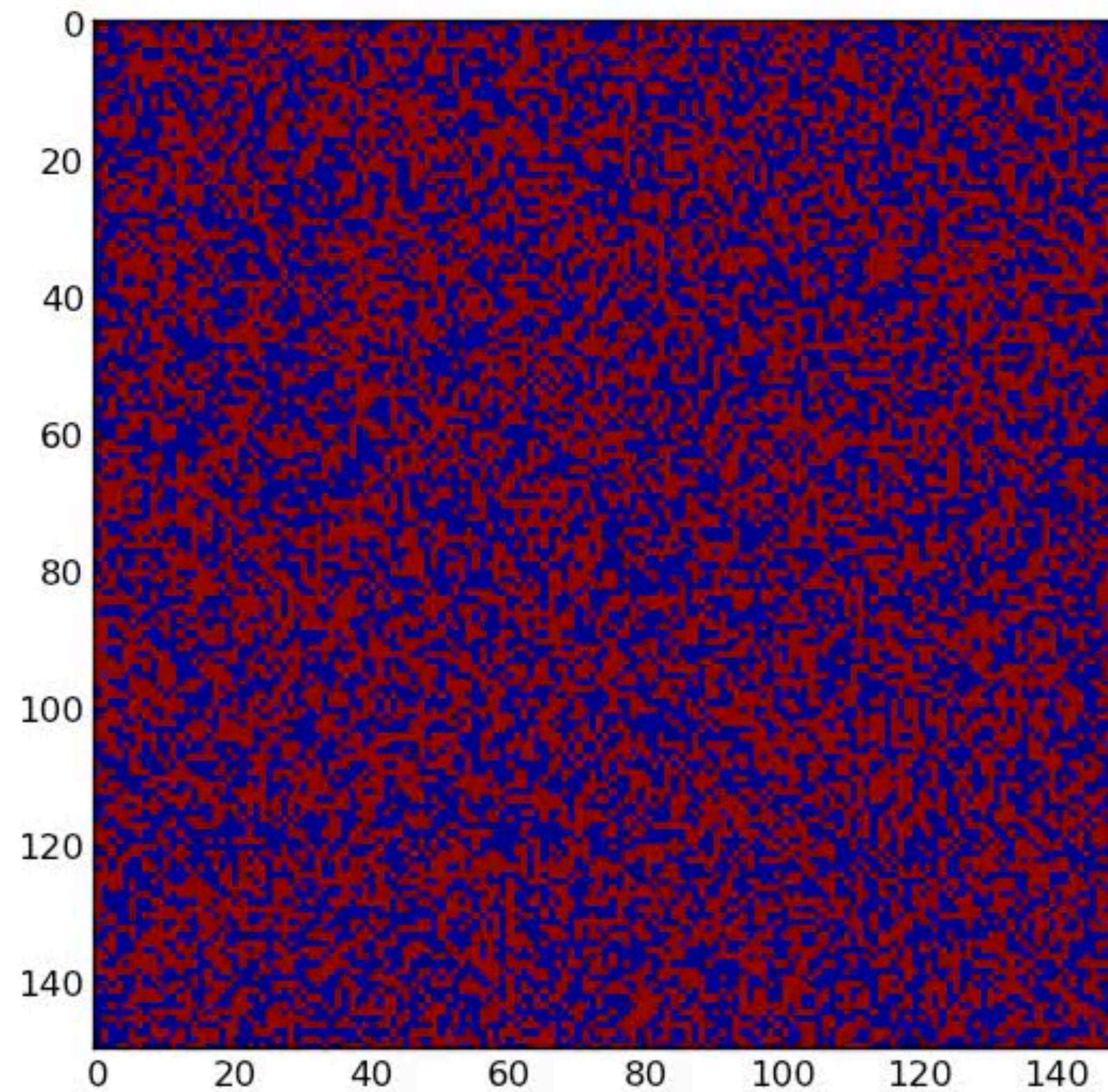
Low Temperature Phase

by B.D.Hammel



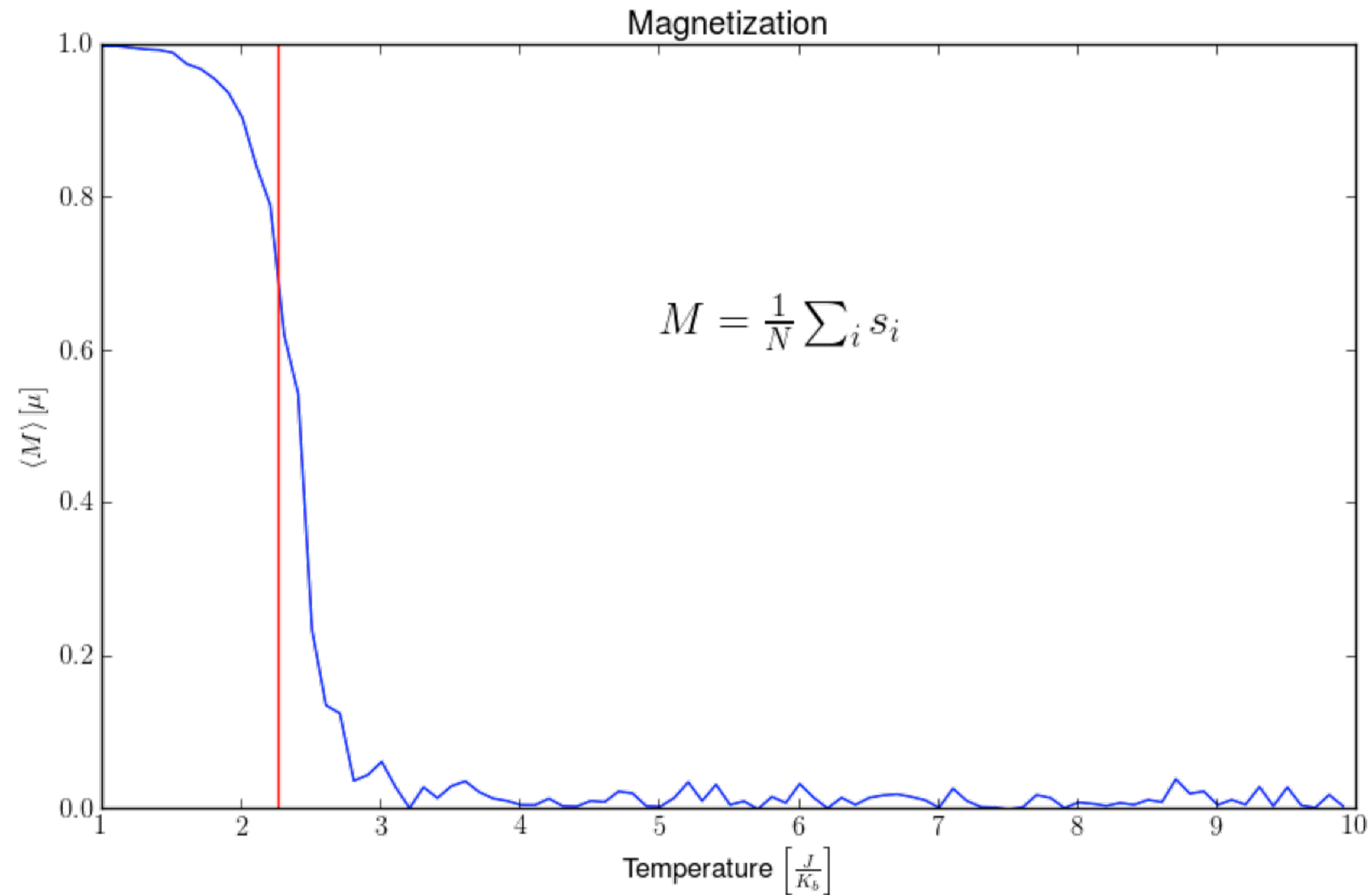
High-Temperature Phase

by B.D.Hammel

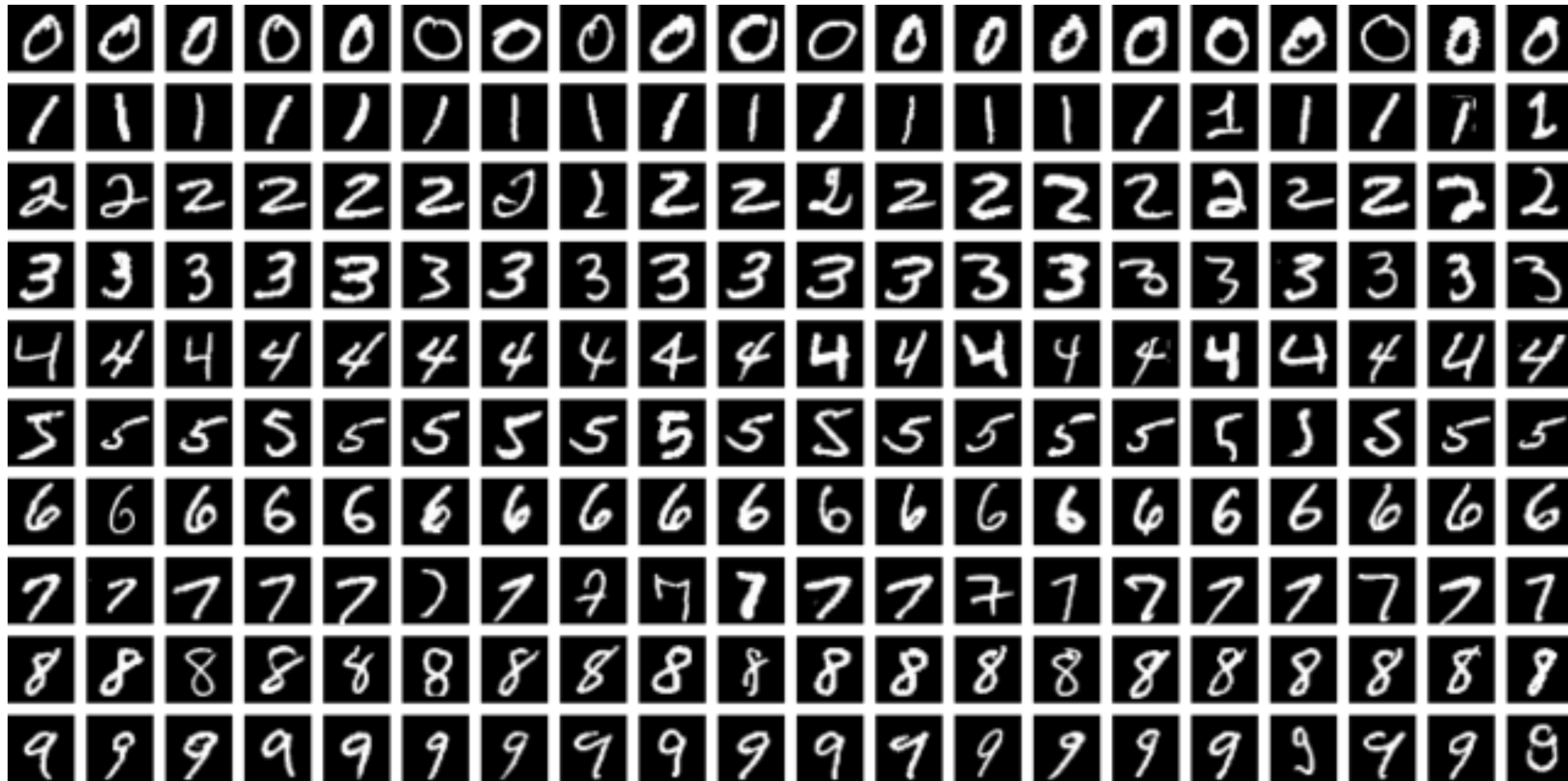


Phase Transition in Ising Model

by B.D.Hammel

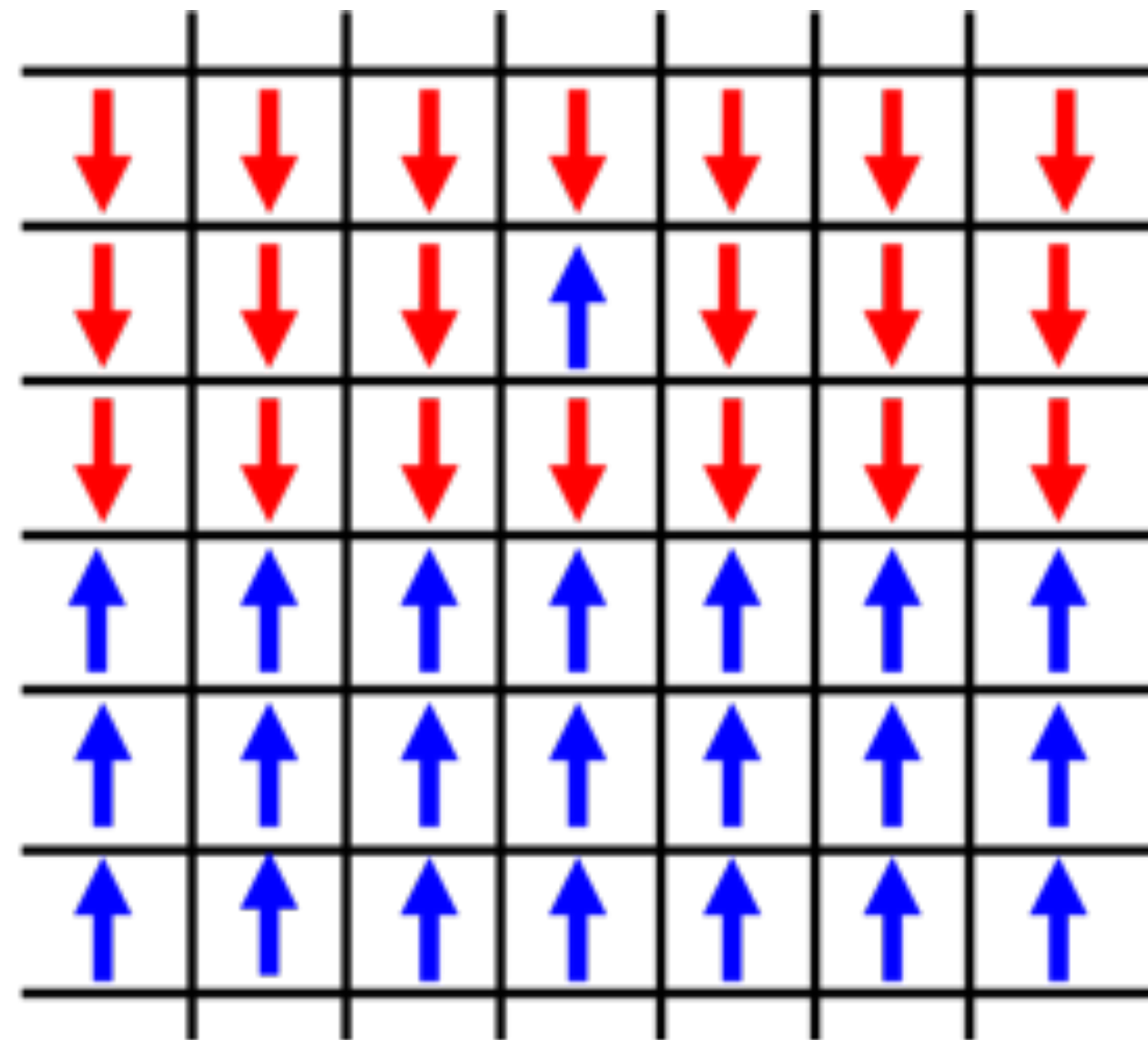


Learning Black-White Photos

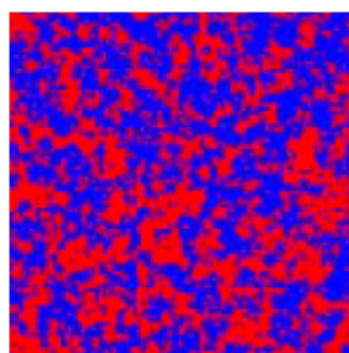


Sample images from MNIST test dataset

The Correspondence Between Spin Lattices and Photos



www.bdhammel.com



	C0	C1	C2	C3	C4	C5	C6	C7	C8
R0									
R1									
R2									
R3									
R4									
R5									
R6									
R7									
R8									

silabs.com

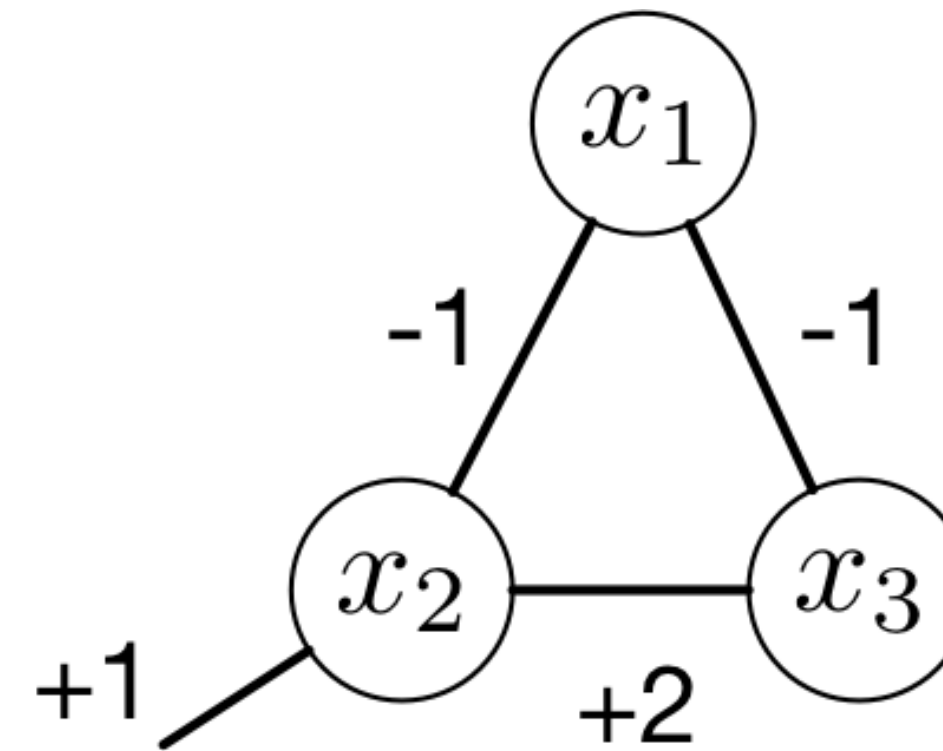


2. Boltzmann Machine (BM)

- Definition, sampling and training
- Zero-temperature limit: connection to neural network
- Restricted BM: a feature recorder

Example:

by Roger Grosse



x_1	x_2	x_3	$w_{12}x_1x_2$	$w_{13}x_1x_3$	$w_{23}x_2x_3$	b_2x_2	$H(\mathbf{x})$	$\exp(H(\mathbf{x}))$	$p(\mathbf{x})$
-1	-1	-1	-1	-1	2	-1	-1	0.368	0.0021
-1	-1	1	-1	1	-2	-1	-3	0.050	0.0003
-1	1	-1	1	-1	-2	1	-3	0.368	0.0021
-1	1	1	1	1	2	1	5	148.413	0.8608
1	-1	-1	1	1	2	-1	3	20.086	0.1165
1	-1	1	1	-1	-2	-1	-3	0.050	0.0003
1	1	-1	-1	1	-2	1	-1	0.368	0.0021
1	1	1	-1	-1	2	1	1	2.718	0.0158

$$\mathcal{Z} = 172.420$$

Application of Conditional Probability



Original

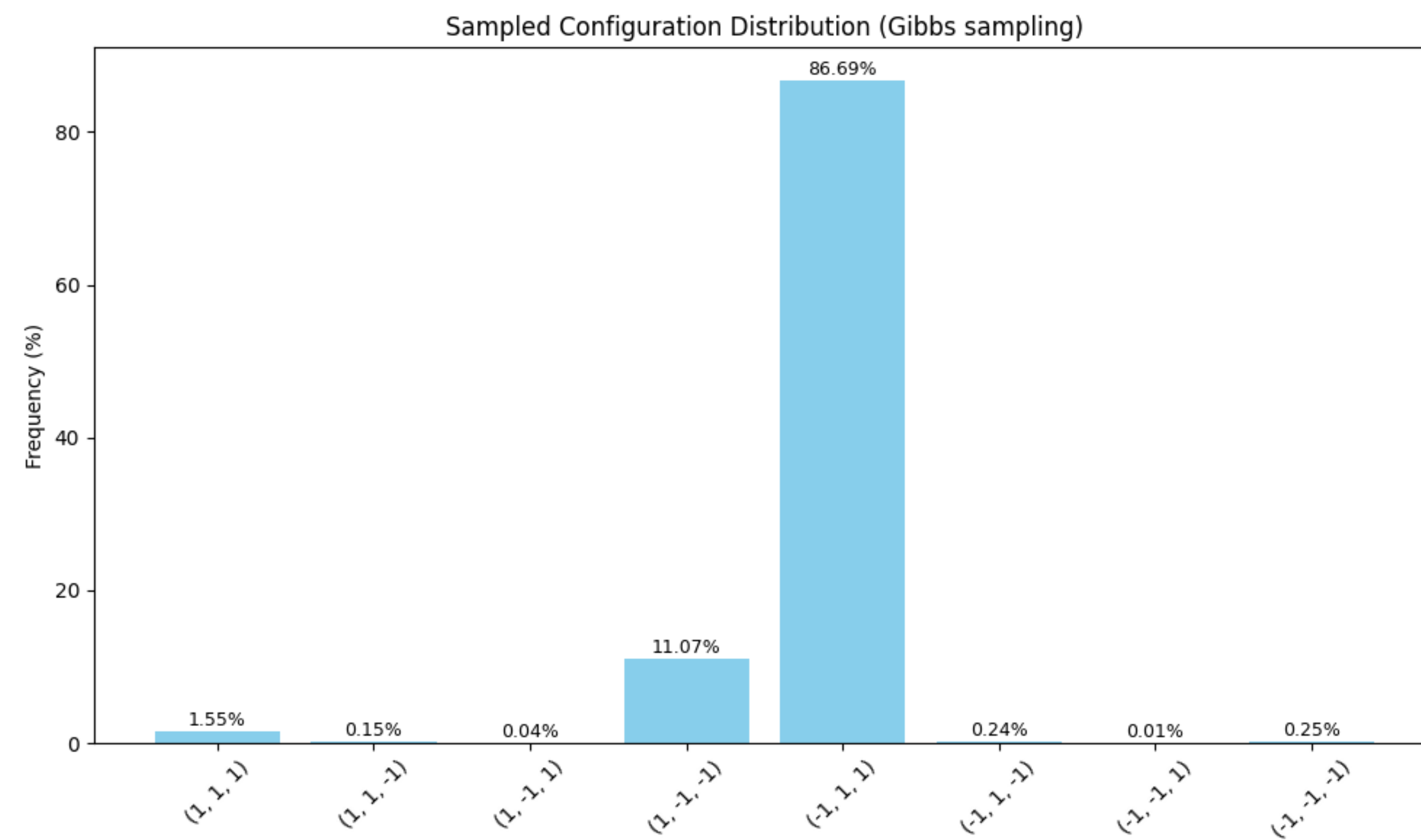
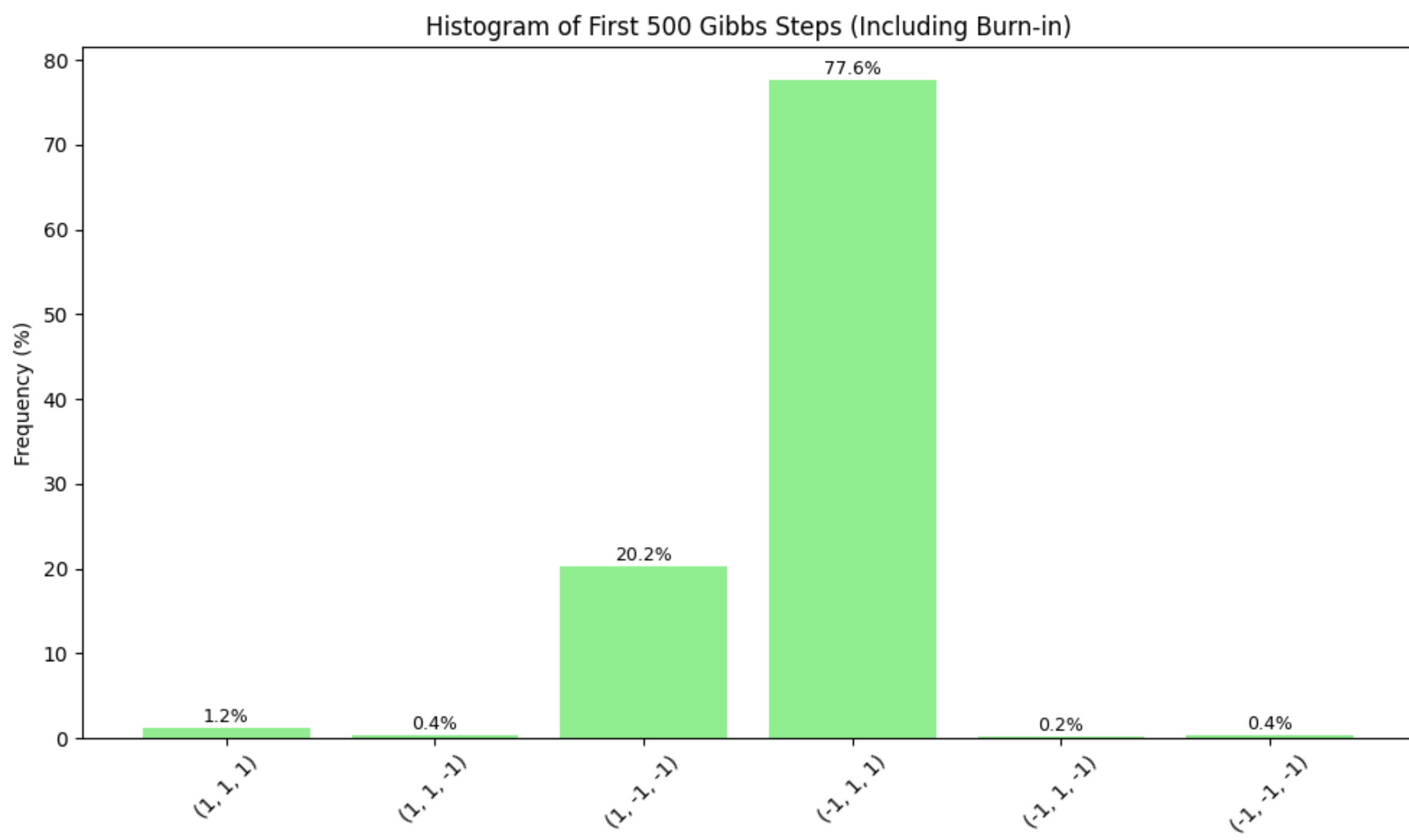


Degraded



Reconstructed

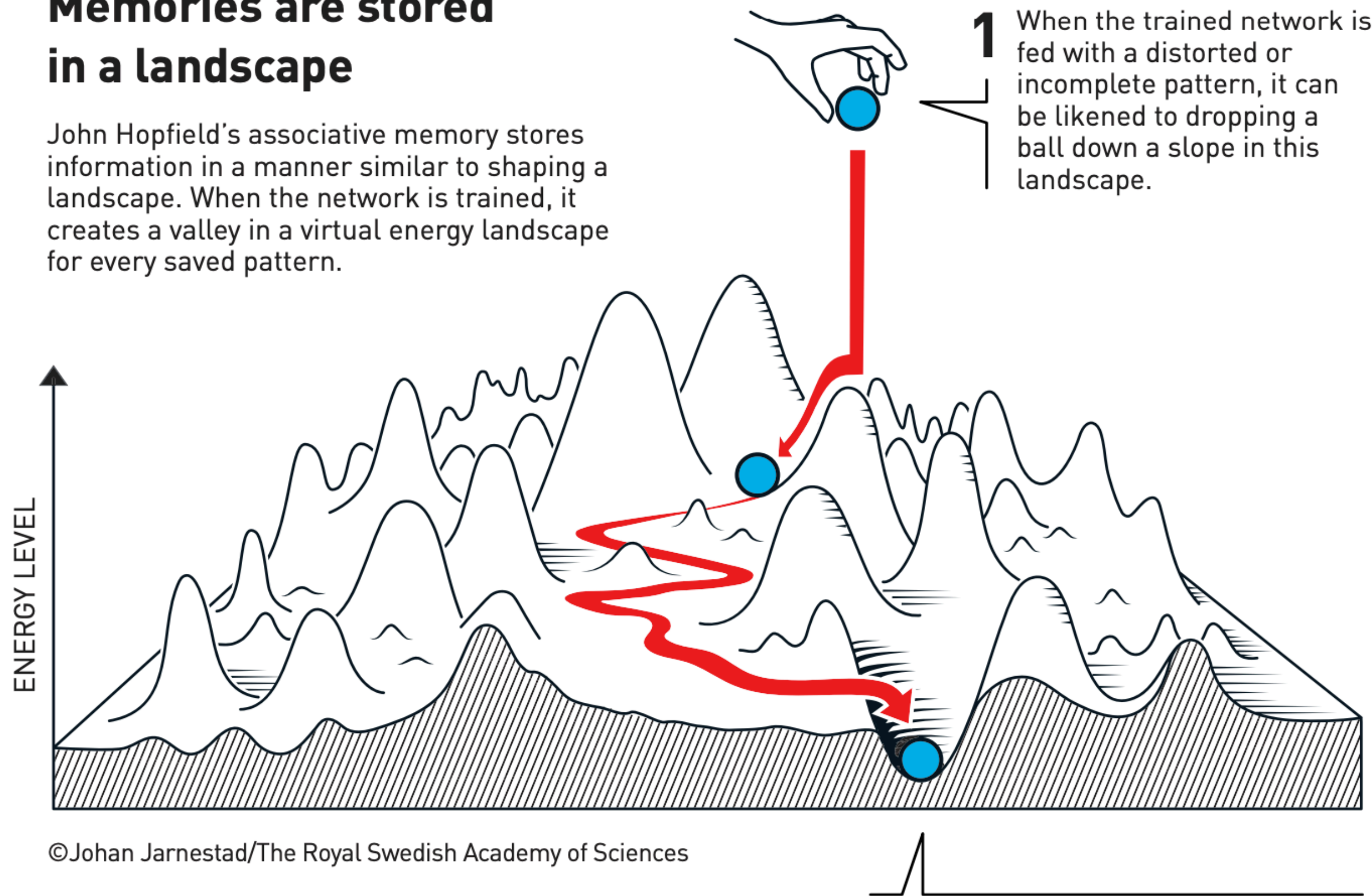
<https://staff.itee.uq.edu.au/janetw/cmc/chapters/Hopfield/>



Gibbs Sampling results

Memories are stored in a landscape

John Hopfield's associative memory stores information in a manner similar to shaping a landscape. When the network is trained, it creates a valley in a virtual energy landscape for every saved pattern.

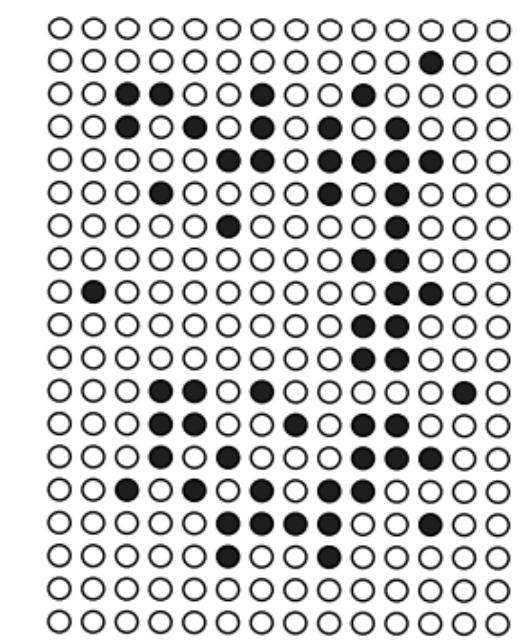


©Johan Jarnestad/The Royal Swedish Academy of Sciences

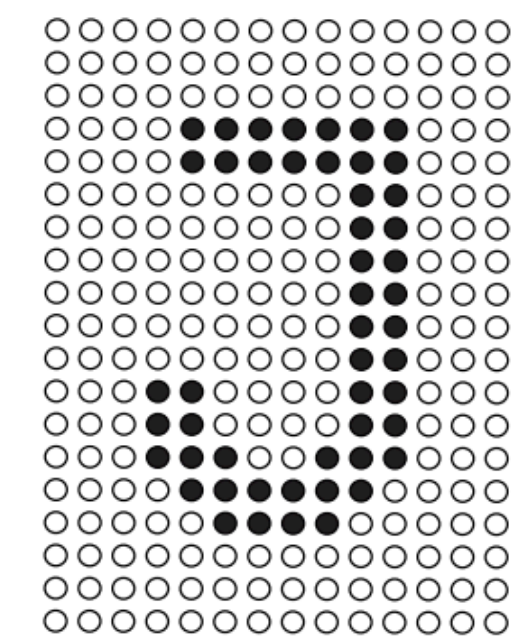
1 When the trained network is fed with a distorted or incomplete pattern, it can be likened to dropping a ball down a slope in this landscape.

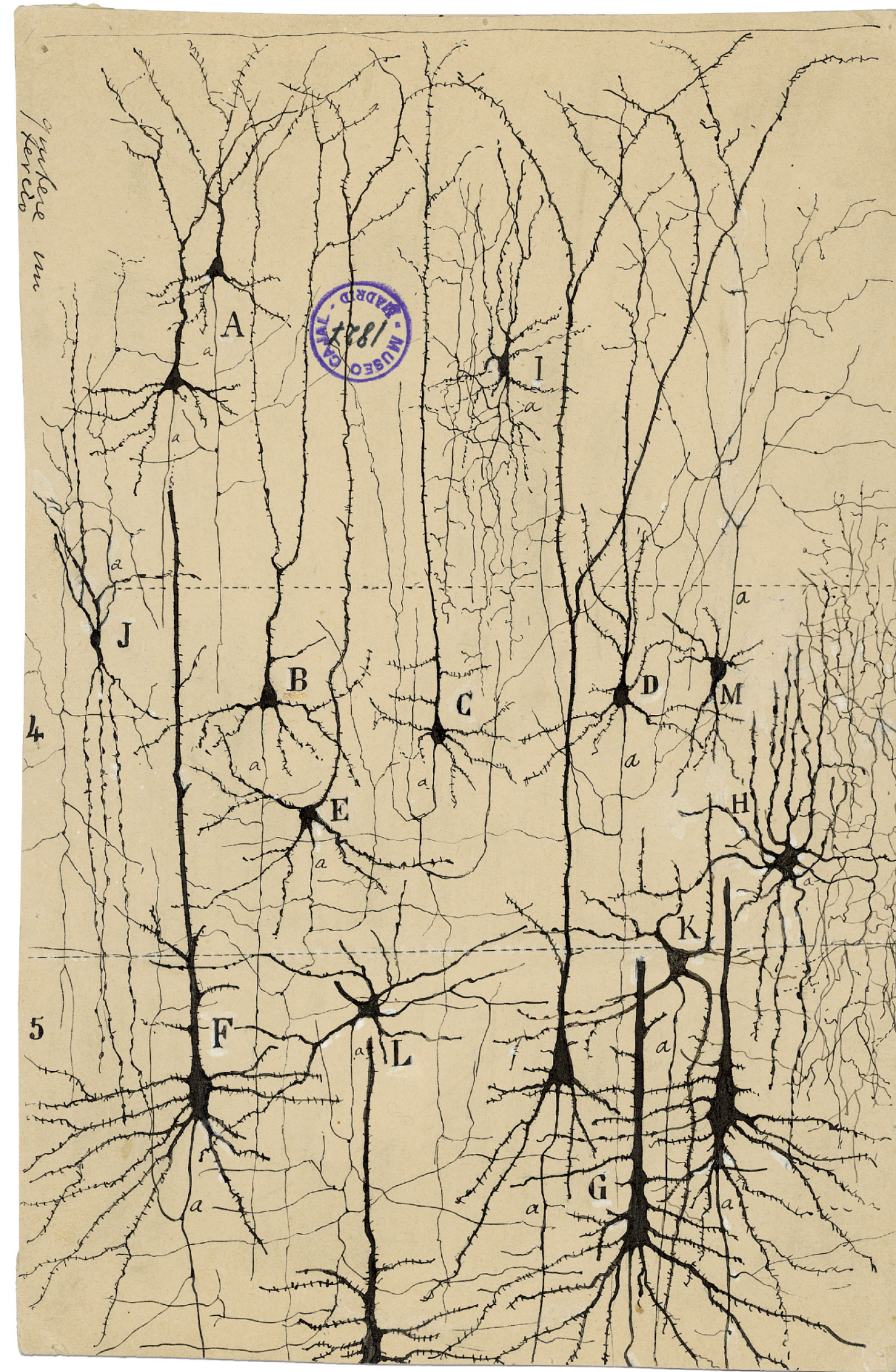
2 The ball rolls until it reaches a place where it is surrounded by uphills. In the same way, the network makes its way towards lower energy and finds the closest saved pattern.

INPUT PATTERN



SAVED PATTERN

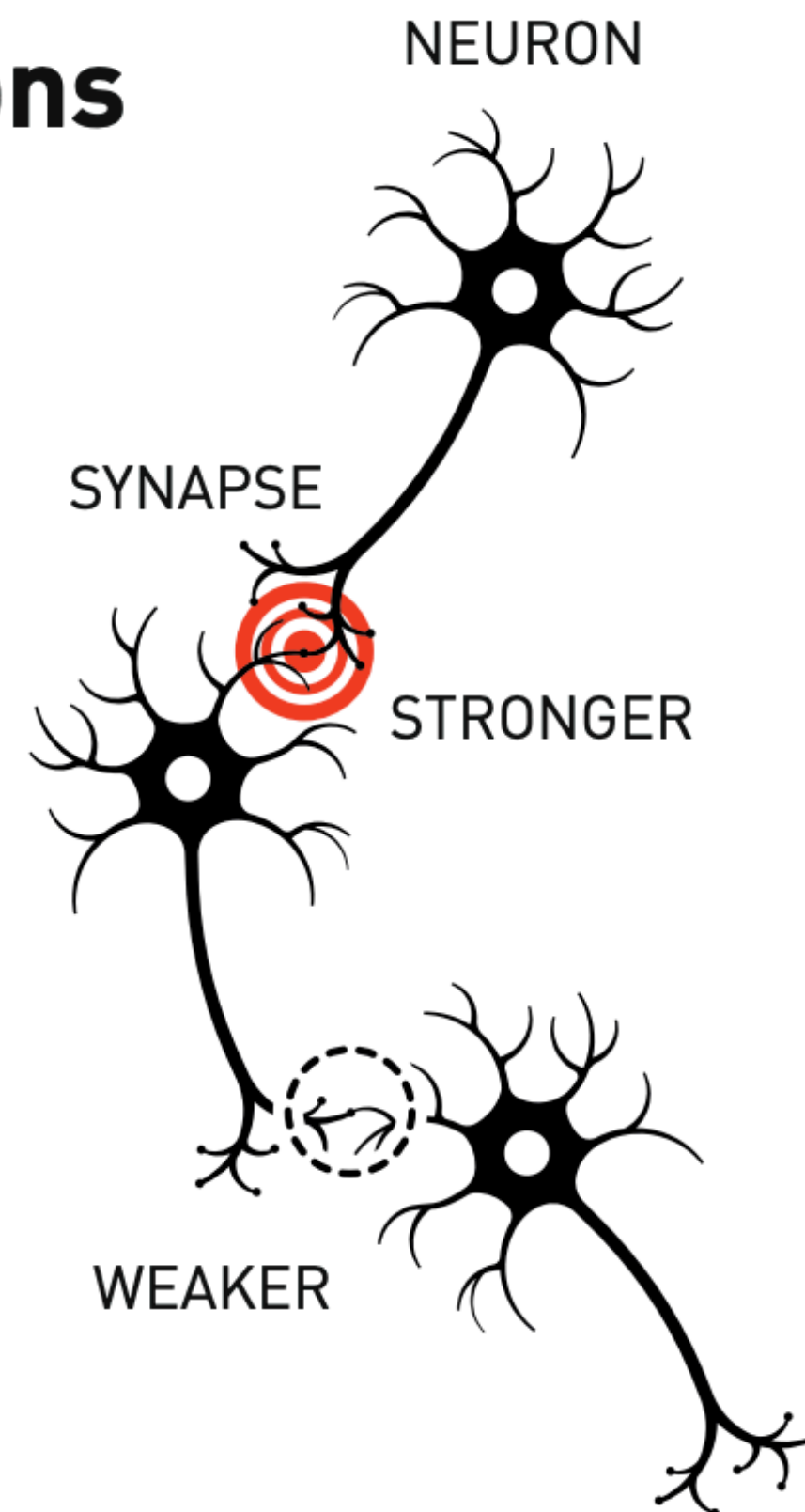




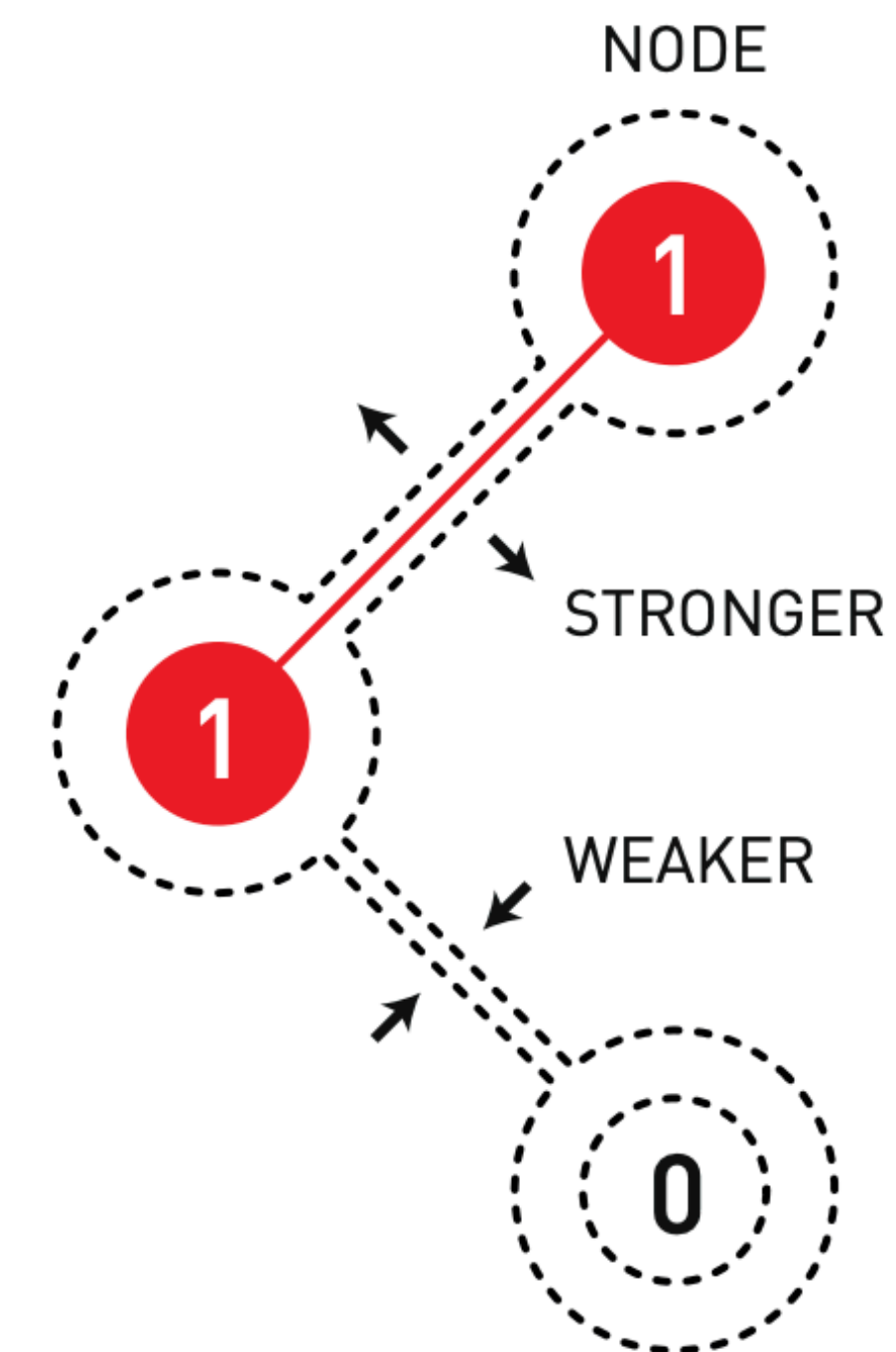
Drawing by Ramon y Cajal

Natural and artificial neurons

The brain's neural network is built from living cells, neurons, with advanced internal machinery. They can send signals to each other through the synapses. When we learn things, the connections between some neurons get stronger, while others get weaker.



Artificial neural networks are built from nodes that are coded with a value. The nodes are connected to each other and, when the network is trained, the connections between nodes that are active at the same time get stronger, otherwise they get weaker.





The Temperature Parameter

The default value of `temperature` is 1.0.

- We recommend users to set the `temperature` according to their use case listed in below.

USE CASE	TEMPERATURE
Coding / Math	0.0
Data Cleaning / Data Analysis	1.0
General Conversation	1.3
Translation	1.3
Creative Writing / Poetry	1.5

What is your favorite drink?

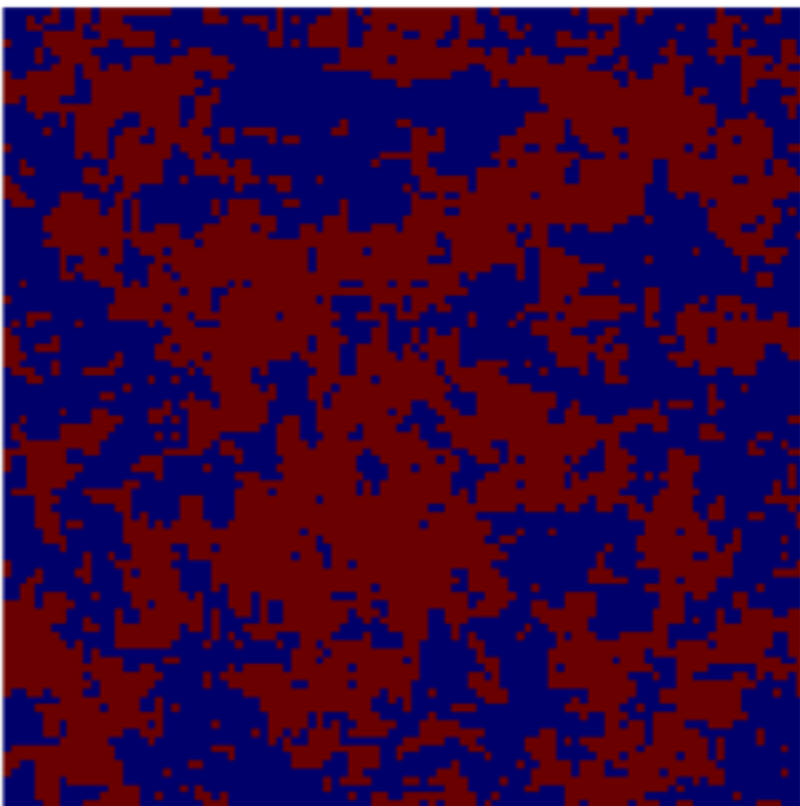
温度值	回答风格	示例回答
低温 (0.1-0.3)	严谨、客观	"作为AI，我没有个人偏好，但可以推荐水、咖啡或茶。"
中温 (0.5-0.7)	自然、略带互动	"虽然我没有味觉，但很多人喜欢冰爽的柠檬茶。你呢？"
高温 (0.8-1.0)	创意、幽默	"我幻想喝银河系星光奶茶！现实点的话...可能是魔法咖啡？"
超高温 (1.5)	疯狂、无逻辑	"最爱火山岩浆特饮配彗星尾粉末！不过建议你先试试氧气泡泡水~"

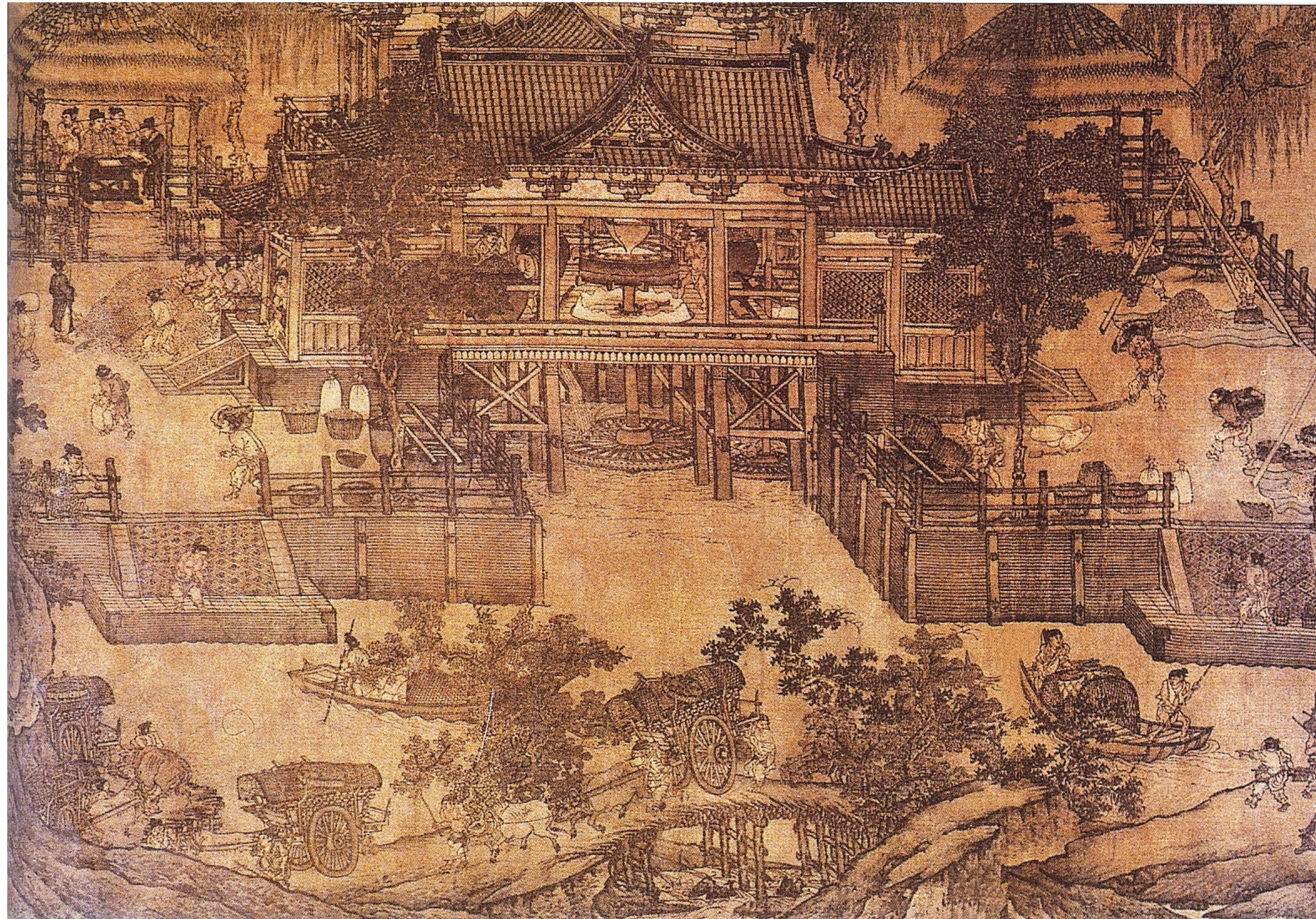
deepseek

1 [J/K_b]



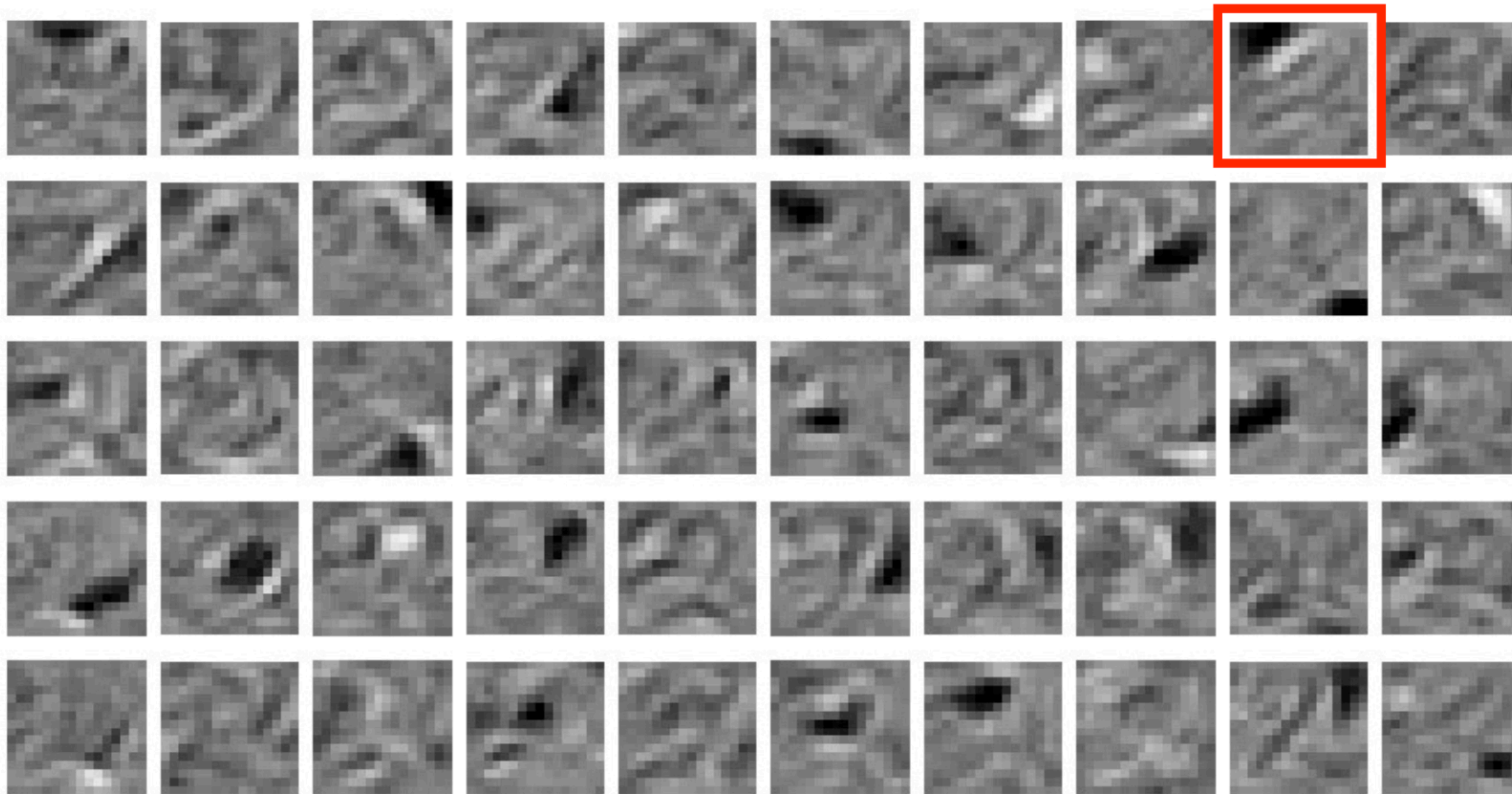
2.5 [J/K_b]





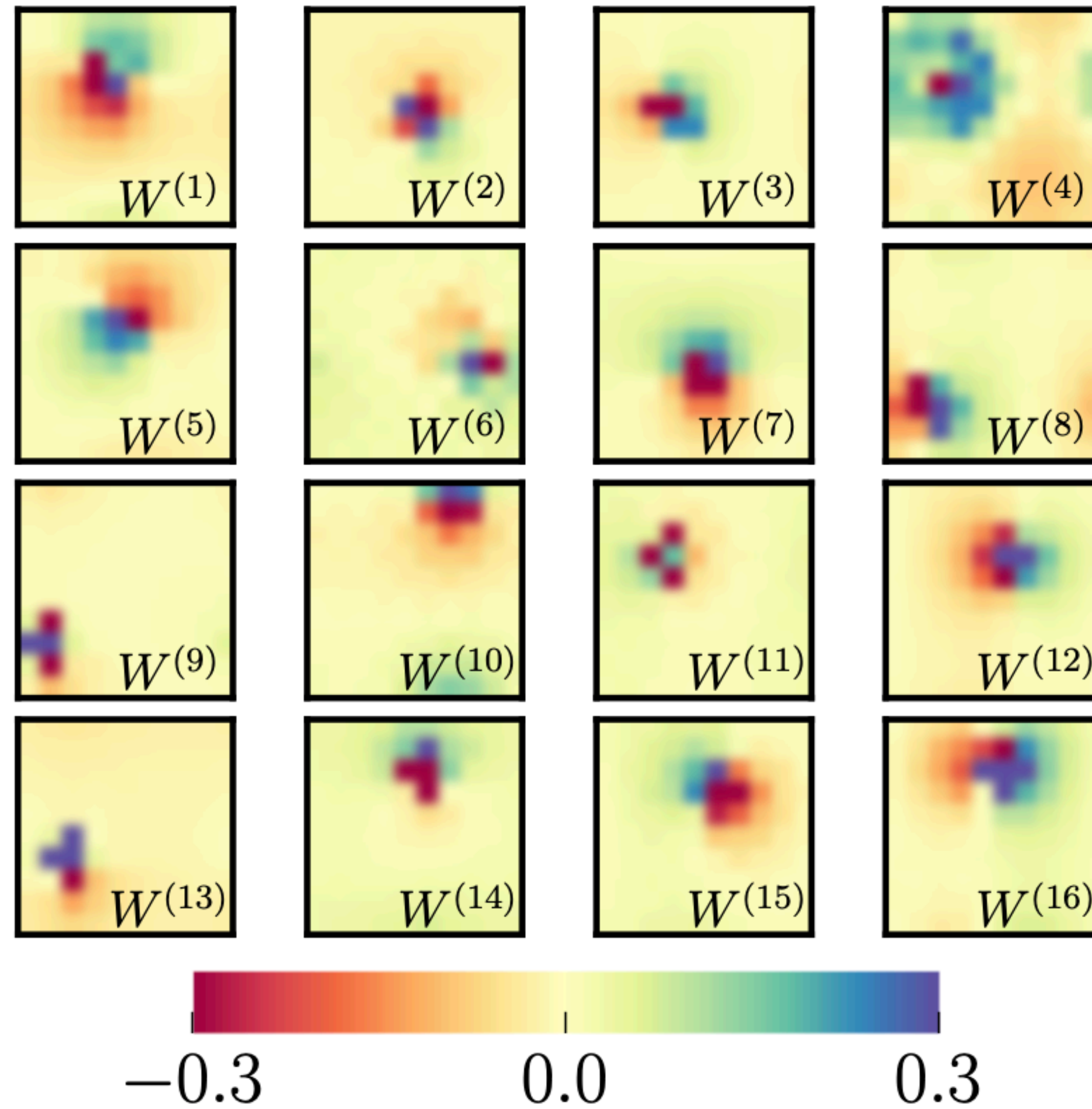
Projecting 3d objects generates rich 2d patterns

Hidden unit as a feature detector



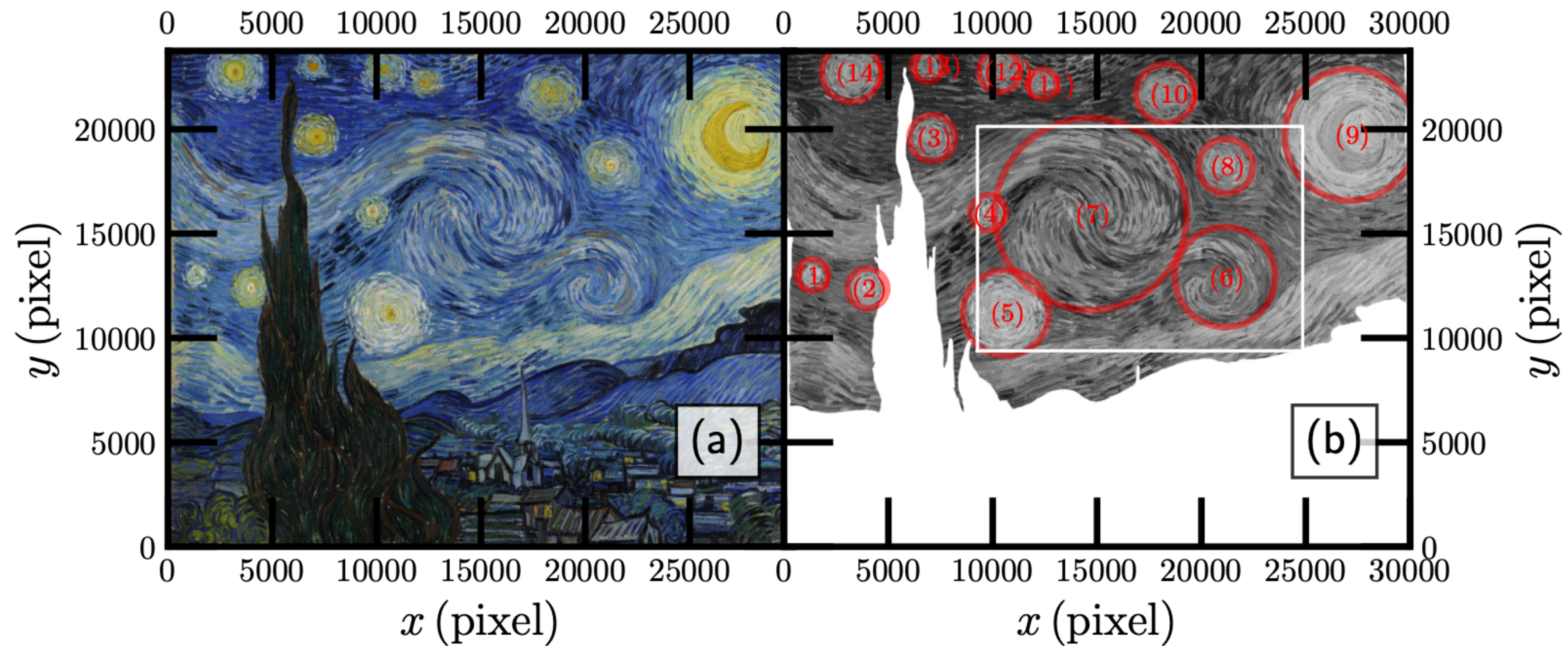
Heisenberg 2D

Carleo-Troyer, Science 2017



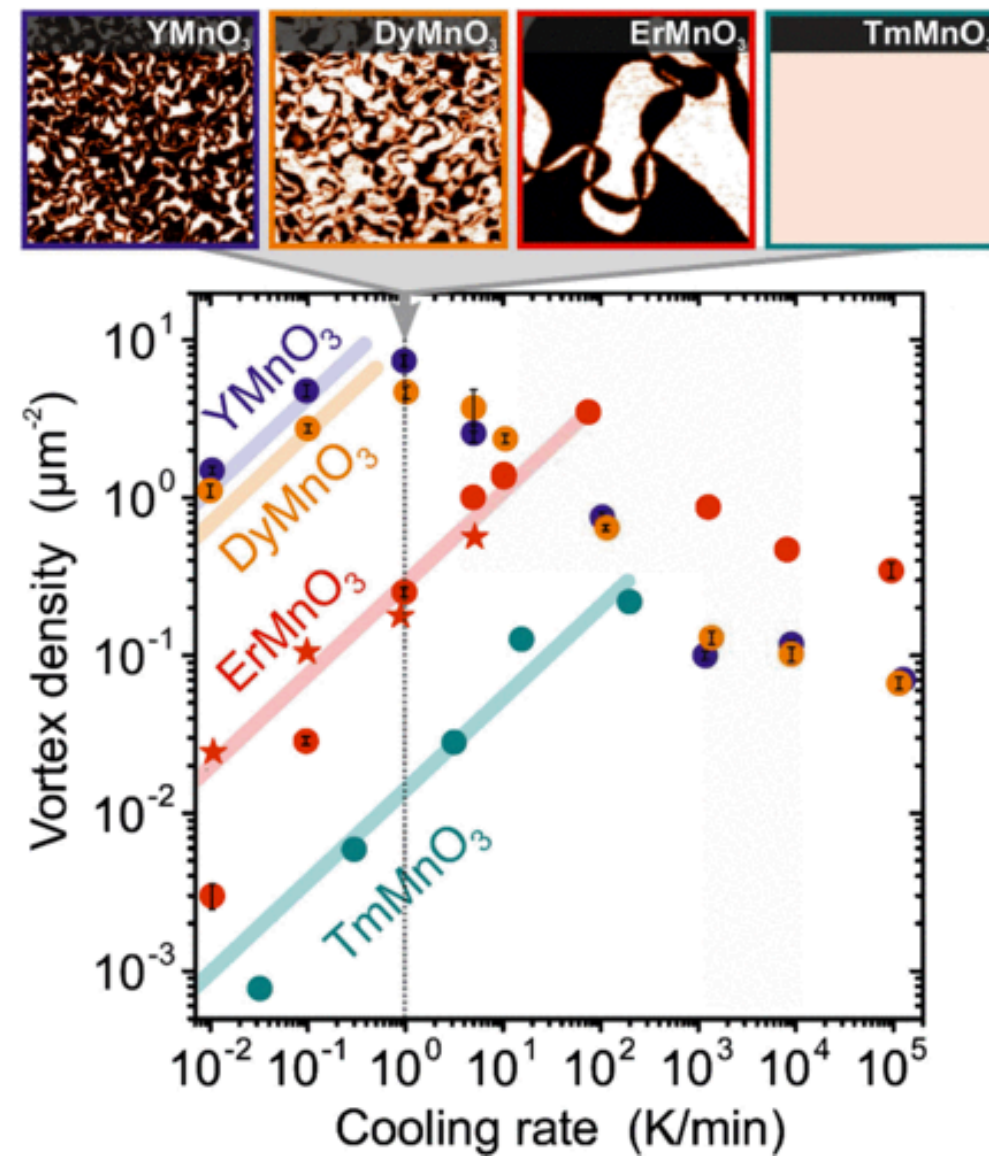
Non-equilibrium Phases

The Starry Night and Turbulence

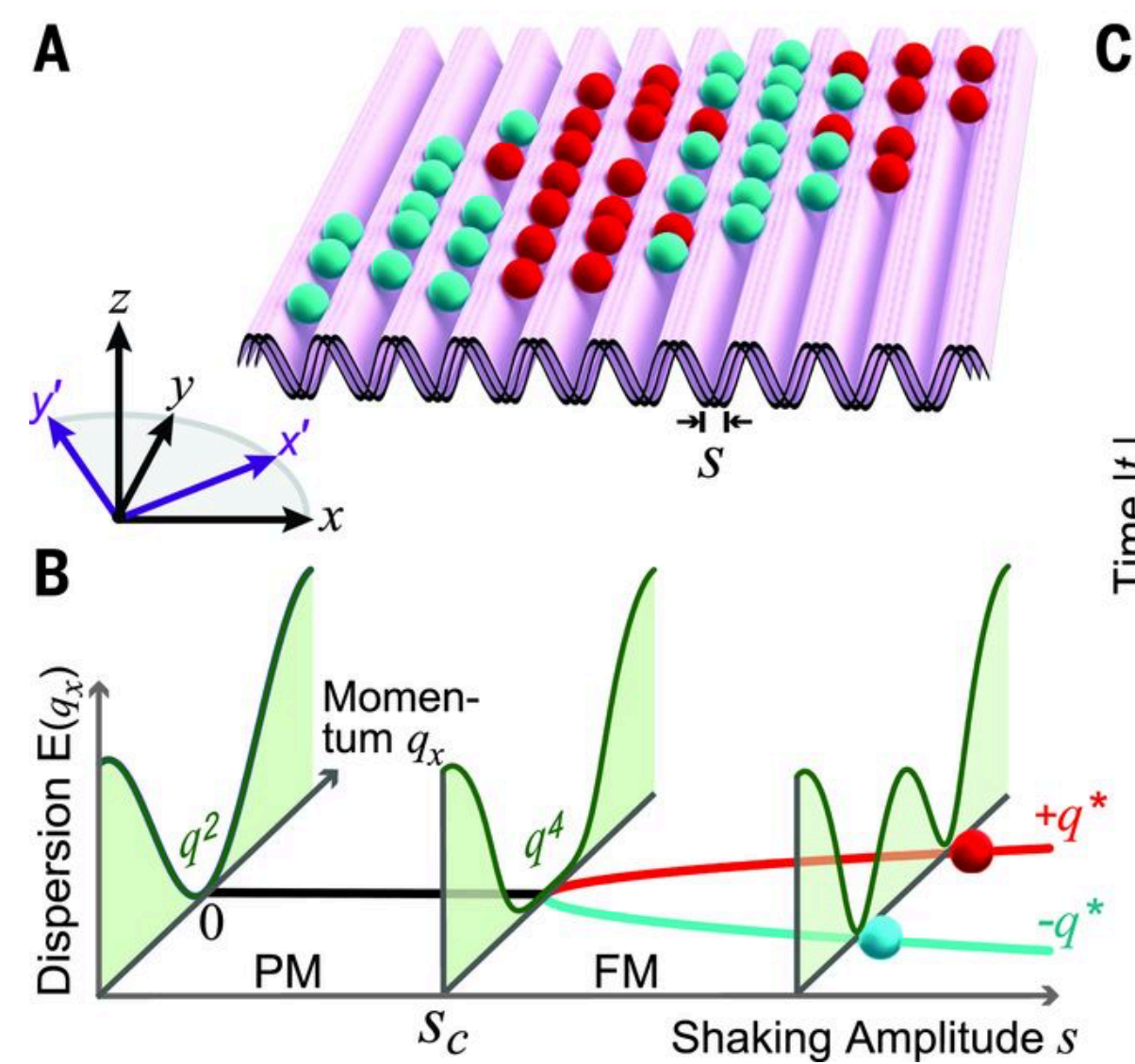


23/0.034/5

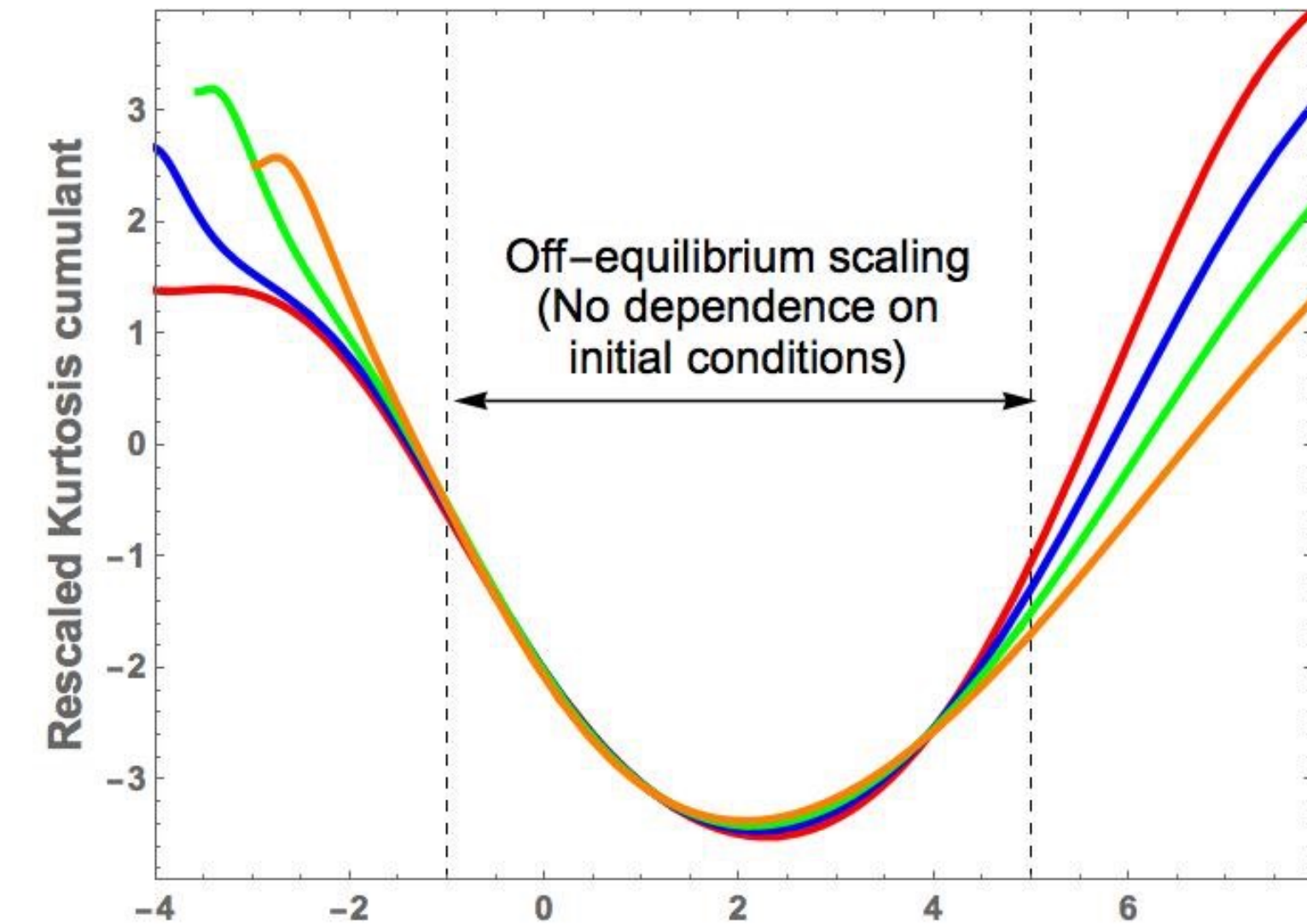
Kibble-Zurek scaling



(Bosons in a shaken optical lattice, W. Clark et al, Science'16)



Hexagonal Manganites, M. Griffin et al, Phys.Rev.X '12



(Conjecture for the QCD point, Mukherjee, Venugopalan and YY, PRL, 16.)

Kibble-Zurek scaling: critical fluctuations and correlations scale with off-equilibrium correlation length for dynamical critical systems

a beautiful demonstration of the unity of physics.



An image generated with Stable Diffusion 3.5 based on the text prompt a photograph of an astronaut riding a horse

Diffusive model for AI Image Generation

**Deep Unsupervised Learning using
Nonequilibrium Thermodynamics**

Jascha Sohl-Dickstein

Stanford University

JASCHA@STANFORD.EDU

Eric A. Weiss

University of California, Berkeley

EAWISS@BERKELEY.EDU

Niru Maheswaranathan

Stanford University

NIRUM@STANFORD.EDU

Surya Ganguli

Stanford University

SGANGULI@STANFORD.EDU

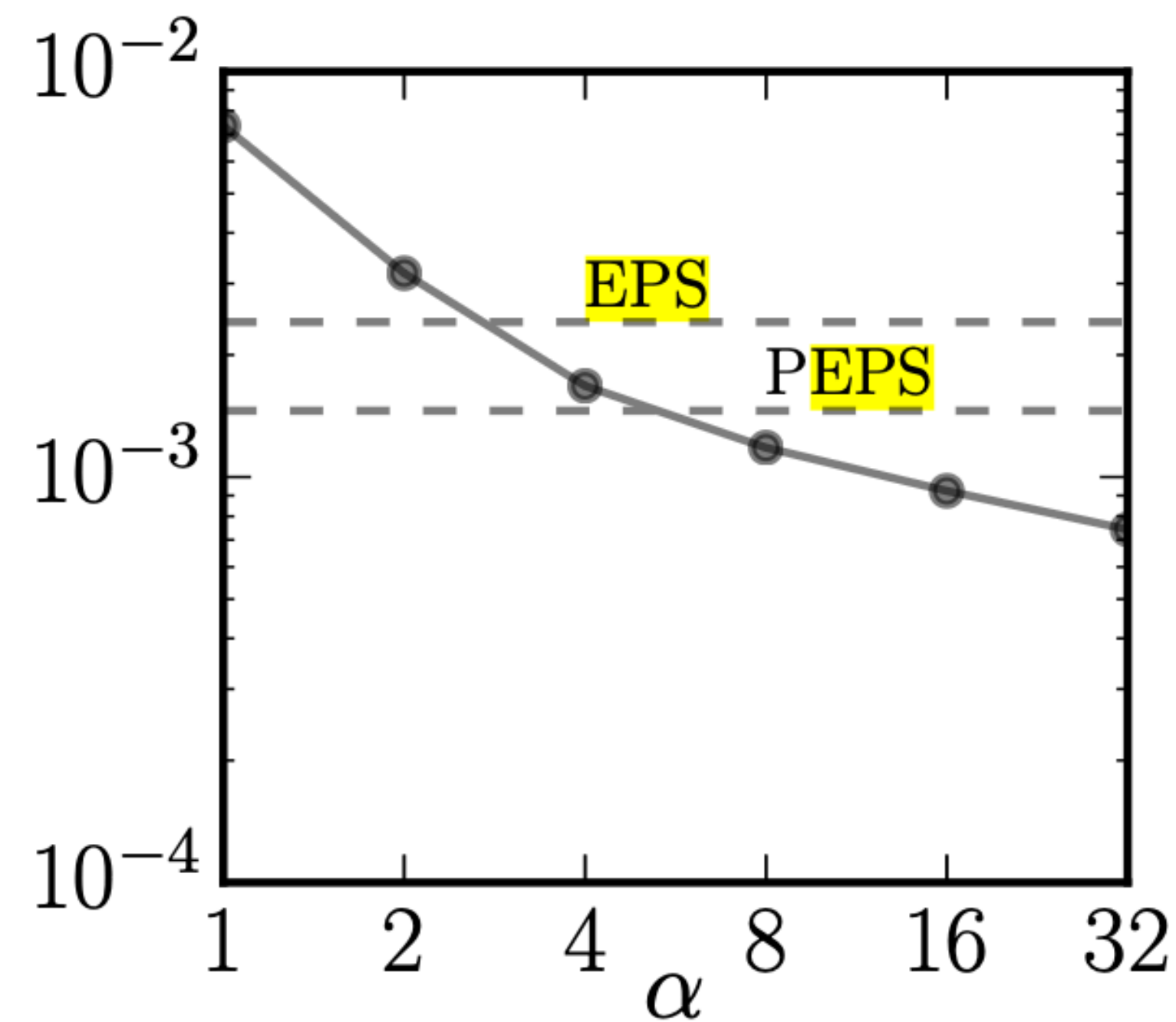
- Diffusion models are widely used for image denoising, image generation, and video generation based on non-equilibrium statistical physics

3.Applications (Using Physics studies Physics)

- Quantum many-body problems
- E.o.S. of quark matter

2D Heisenberg Model

Carleo-Troyer, Science 2017



- Arbitrary precision on the ground-state energy can be obtained upon increasing the hidden units density α

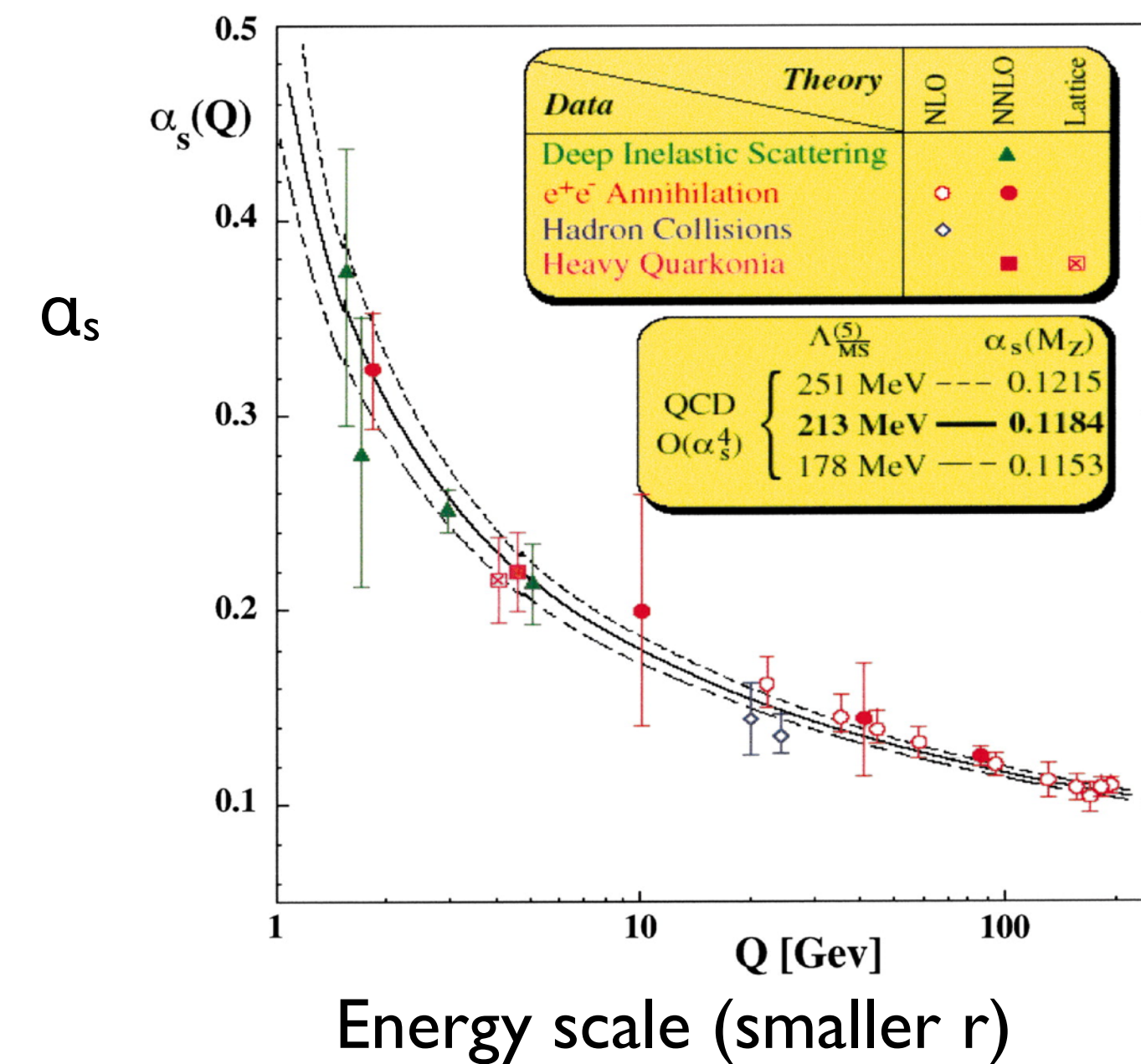
QCD E.o.S. and Restricted Boltzmann Machine



Lei Gao-Shenyu Yin-YY-Kai Zhou, in progress

QCD

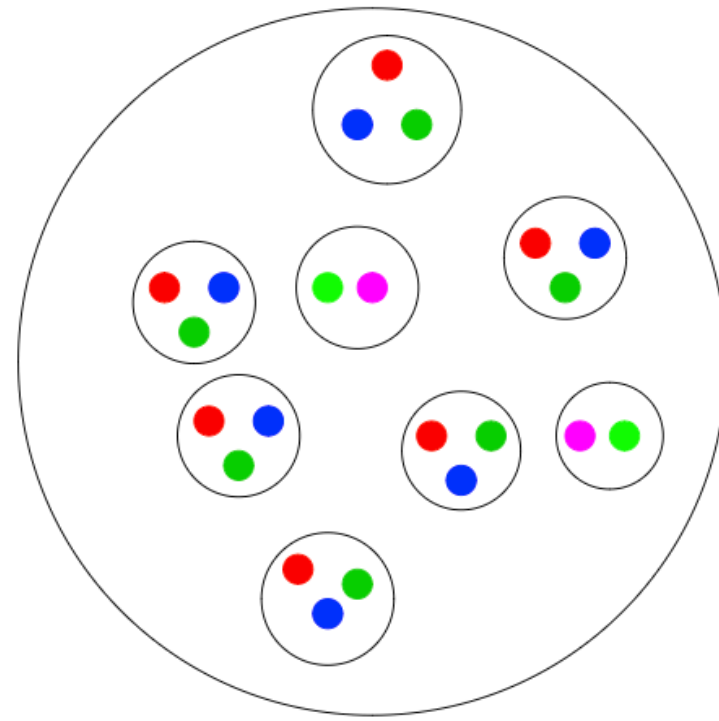
$$L_{QCD} = \sum_q \bar{\psi}_q \left(i \gamma_\mu D^\mu - m_q \right) \psi_q - \frac{1}{2} \text{Tr} \left[\bar{G}_{\mu\nu} \bar{G}^{\mu\nu} \right]$$



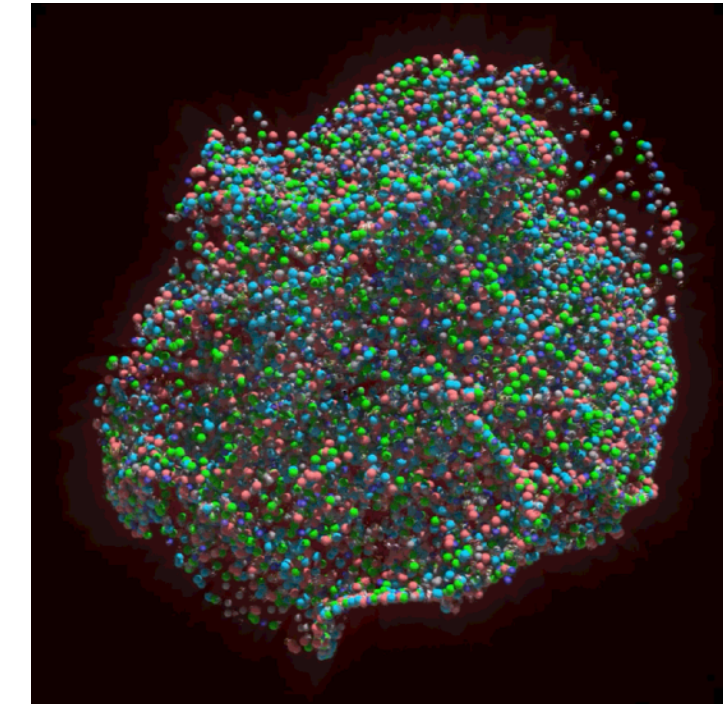
- The theory of quark, gluon and their interaction: Quantum ChromoDynamics (QCD)

Quark Matter

- In ordinary conditions, quarks and gluons are **bounded (confined)** within hadrons



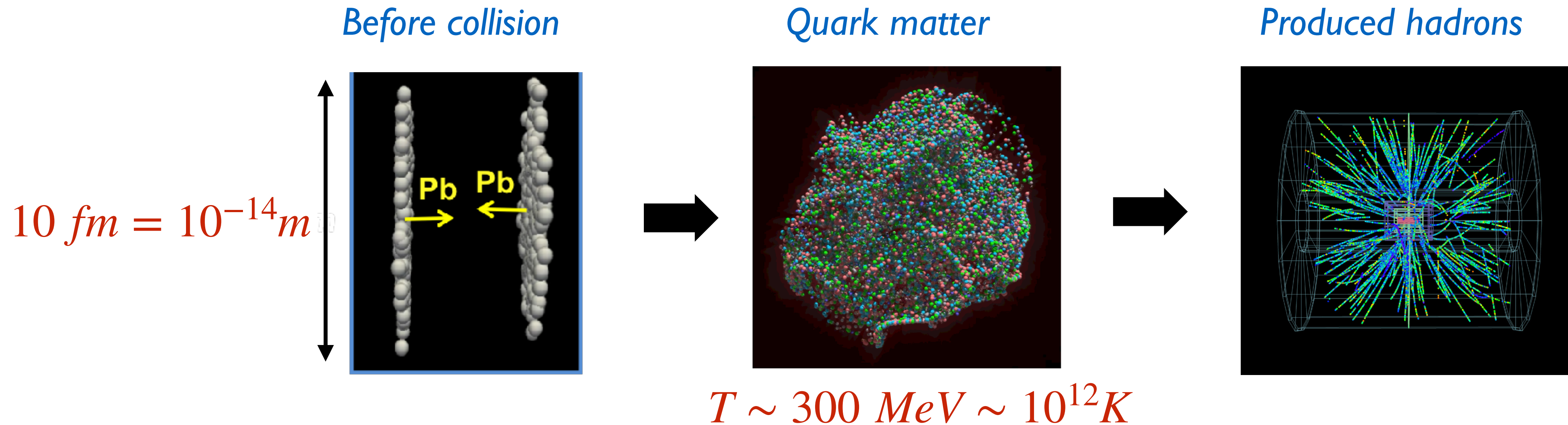
Increasing density
or temperature



- Quark matter: state of matter that exists at extreme temperature/density where quarks and gluons are **unbounded (deconfined)**
- Filled the early universe, may appear in the core of a neutron star

Heavy Ion Collisions

Image from CERN



- At the Relativistic Heavy-Ion Collider (RHIC) in the US since 2000 and at the Large Hadron Collider (LHC) in Europe since 2010
- Quark matter produced in those collisions, often called quark-gluon plasma, is **exceptionally tiny and hot**

Quark Liquid

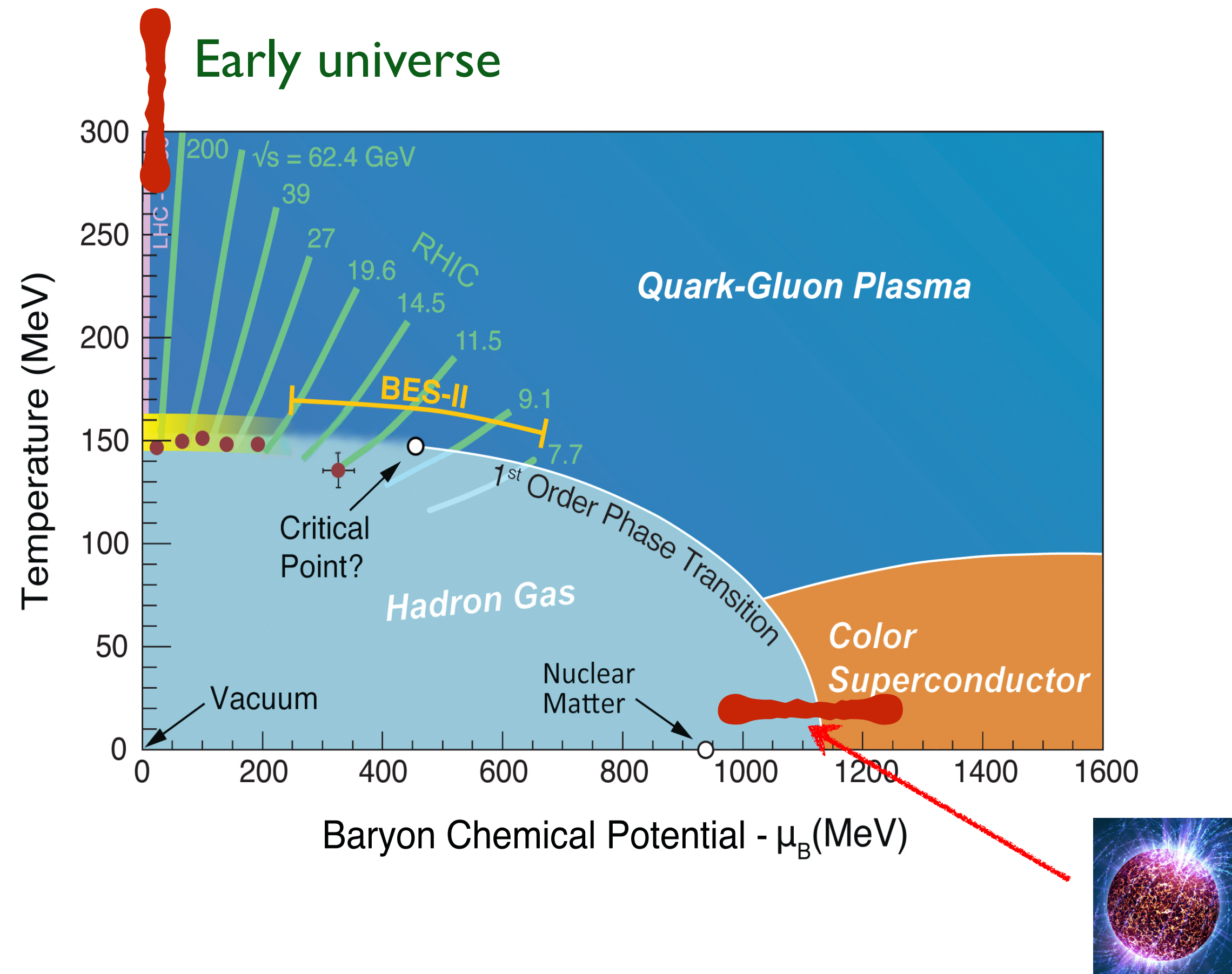
- In $T \rightarrow \infty$ limit, quark matter must be weakly coupled quark/gluon quasiparticles (**asymptotic freedom**)
- Those produced in heavy-ion collisions are found to behave as **a liquid** with **extraordinarily small shear viscosity** (measure of the “friction” in fluid flow): $\eta/s \sim 1/(4\pi)$

This value is close to that of some strongly-coupled gauge theories and too small to be explained in terms of quasi-particles

- Other systems **without apparent quasi-particles**: strongly correlated systems like strange metals or those near quantum phase transitions

Pushing the boundary of strongly correlated matter

QCD Phase Diagram

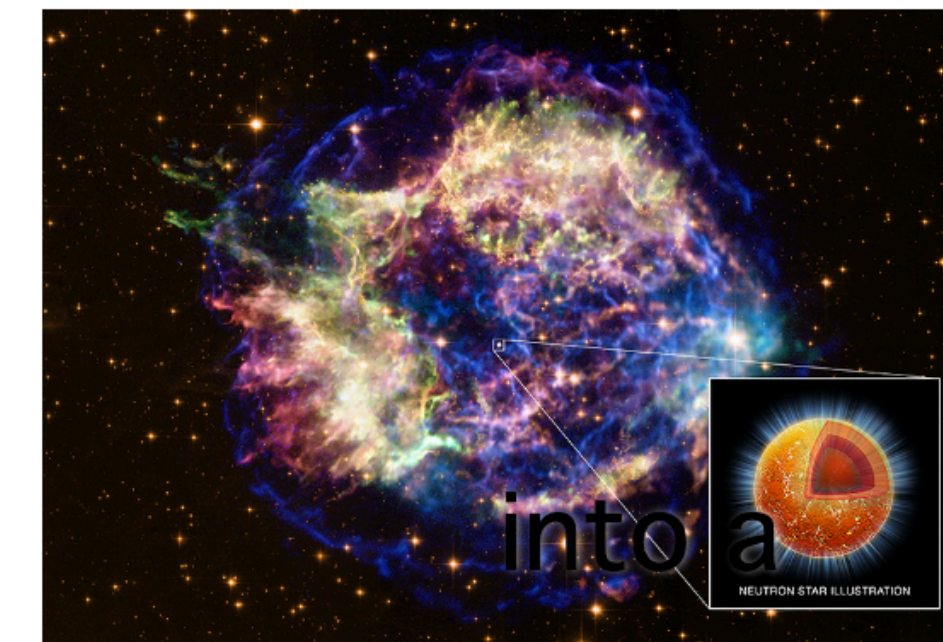


Rich phase structure governed by QCD theory

Neutron stars: death and birth

A neutron star is the collapsed core of a massive star ($\sim 8-20 M_{\odot}$) left behind after a supernova explosion; more massive stars collapse into black holes.

- the remnant neutron star compresses $\sim 1.4 M_{\odot}$ into a sphere only about 9-16 km across
- this material is crushed together so tightly that gravity overcomes the repulsive force between negatively charged electrons and positively charged protons
- the resulting structure of the star is complex, with a solid crystalline crust about 1 kilometer thick, enclosing a core of superfluid neutrons and superconducting protons
- above the crust exists both an ocean and atmosphere of much less dense material



Cassiopeia A

Tolman-Oppenheimer-Volkoff Equation

by Sophia Han

- for the interior of a **spherical, static, relativistic** star

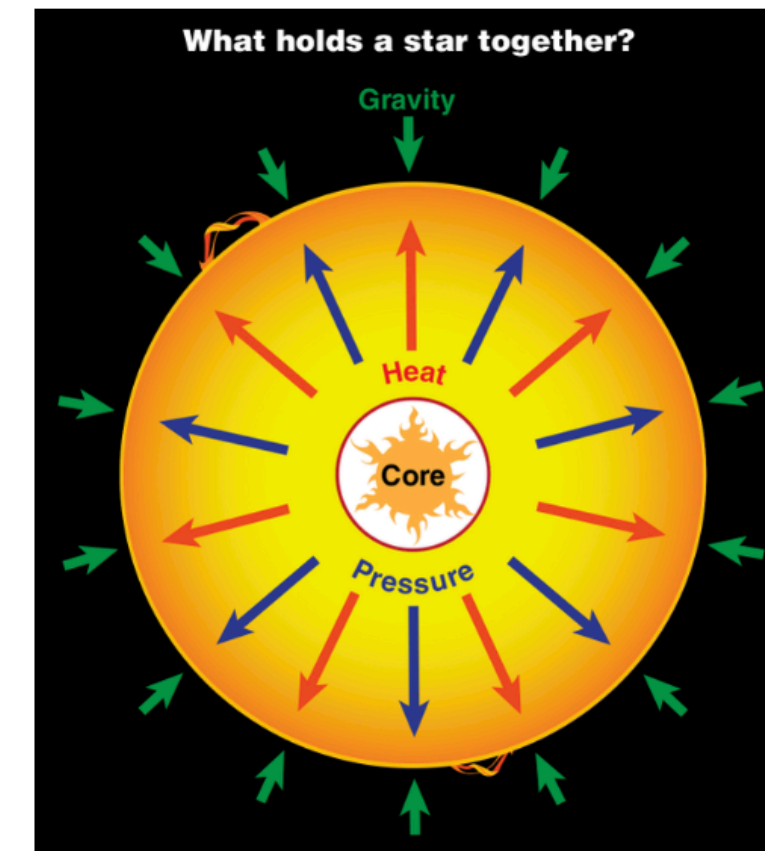
$$\frac{dp}{dr} = -\varepsilon(r) \frac{Gm(r)}{r^2} \left[1 + \frac{p(r)}{\varepsilon(r)} \right] \left[1 + \frac{4\pi r^3 p(r)}{m(r)} \right] \left[1 - \frac{2Gm(r)}{r} \right]^{-1}$$

$\nearrow \rho_{\text{central}}$
 $\nwarrow p(\varepsilon)$

where the enclosed mass is defined as

$$m(r) \equiv 4\pi \int_0^r \varepsilon(r) r^2 dr$$

For a given **equation of state (EoS)**, TOV Eqs. can be solved simultaneously for the radial distribution of **pressure** and that of **energy density**.



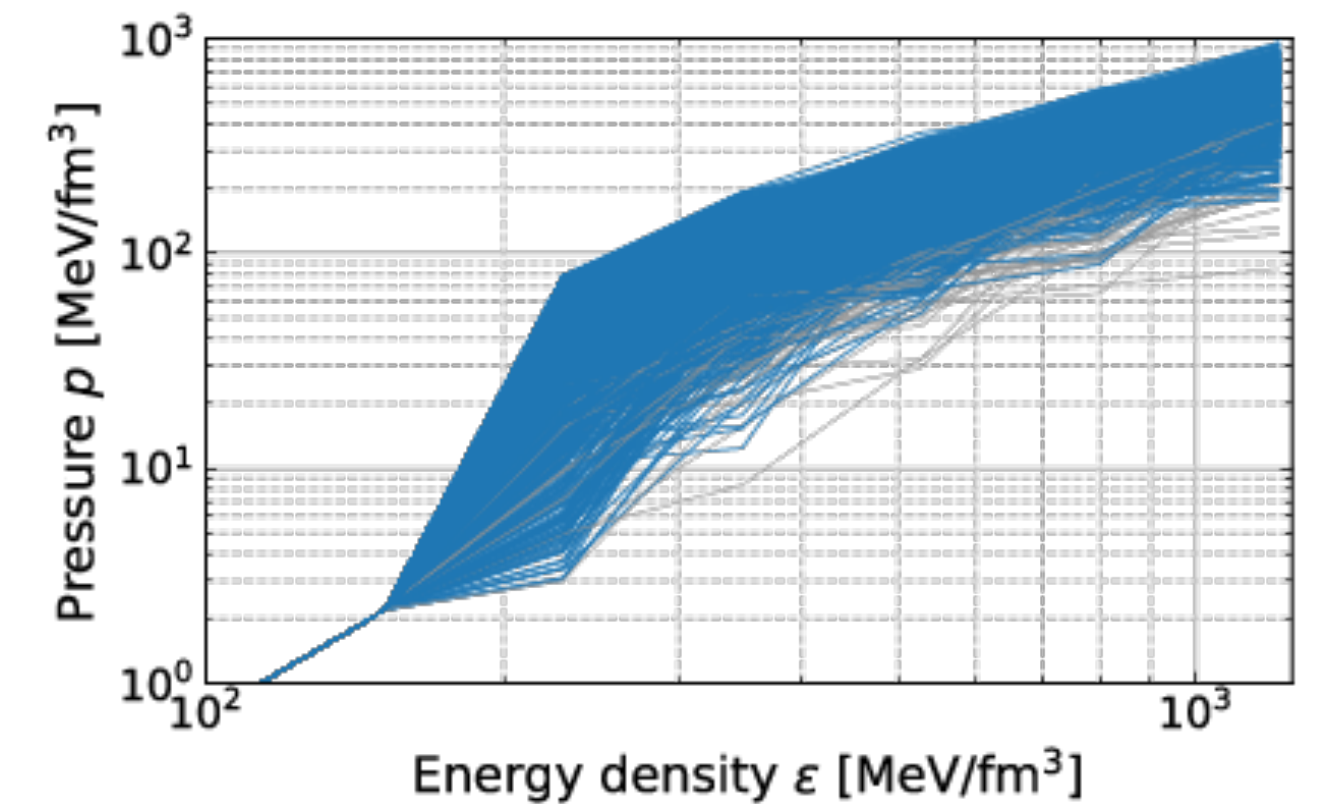
$p(r = 0) = p_{\text{central}}$
 star surface at zero pressure: $p(r = R) = 0$
 gravitational mass of the star: $m(r = R) \equiv M$

$\xrightarrow[M(p_{\text{central}})]{R(p_{\text{central}})} \mathbf{M(R)}$

Neutron Star EoS

- **Piecewise polytropic parametrization:** divide the density range into segments and use polytropic interpolation in each segment.

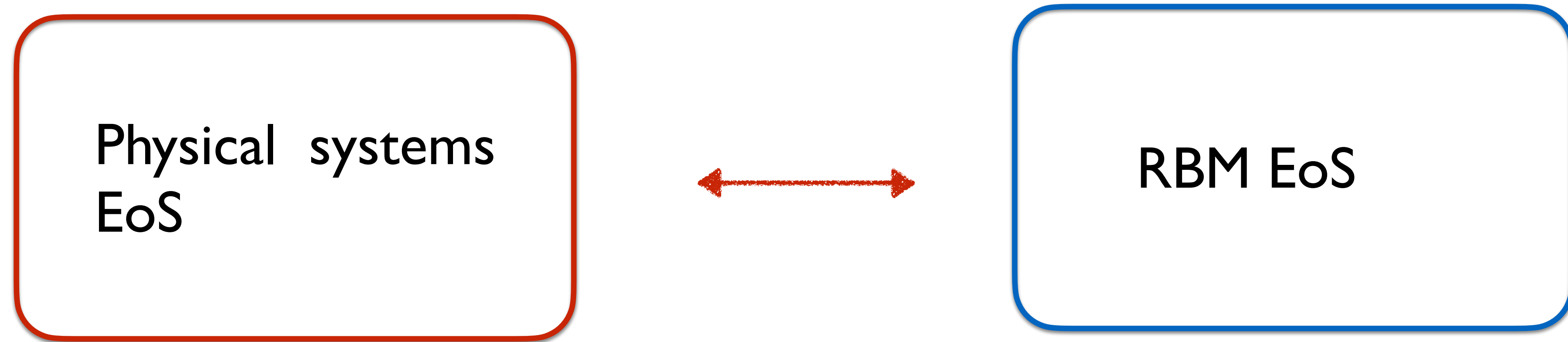
$$p = p_{i-1} + \sum c_{i,n} (\epsilon - \epsilon_i)^n$$



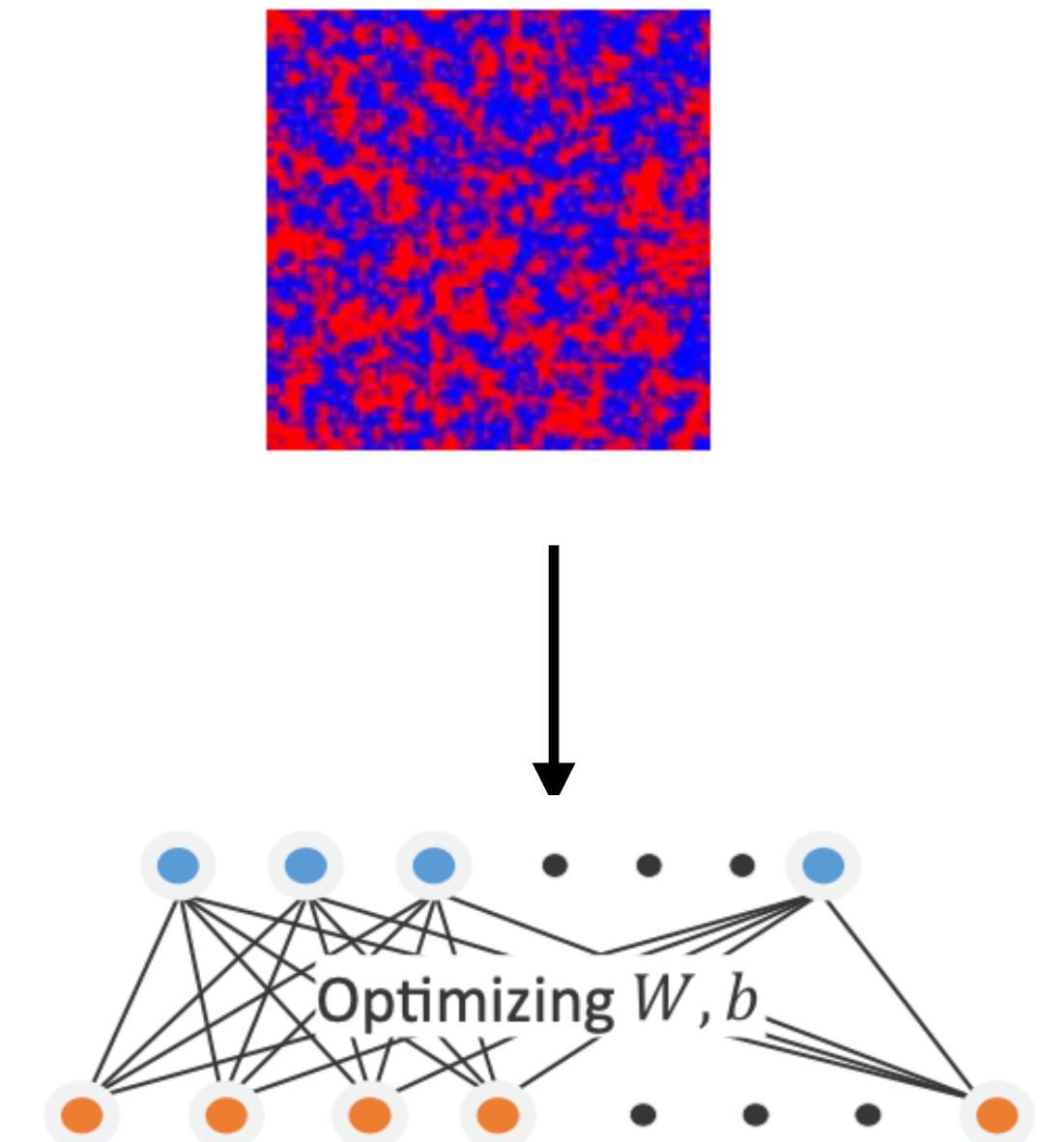
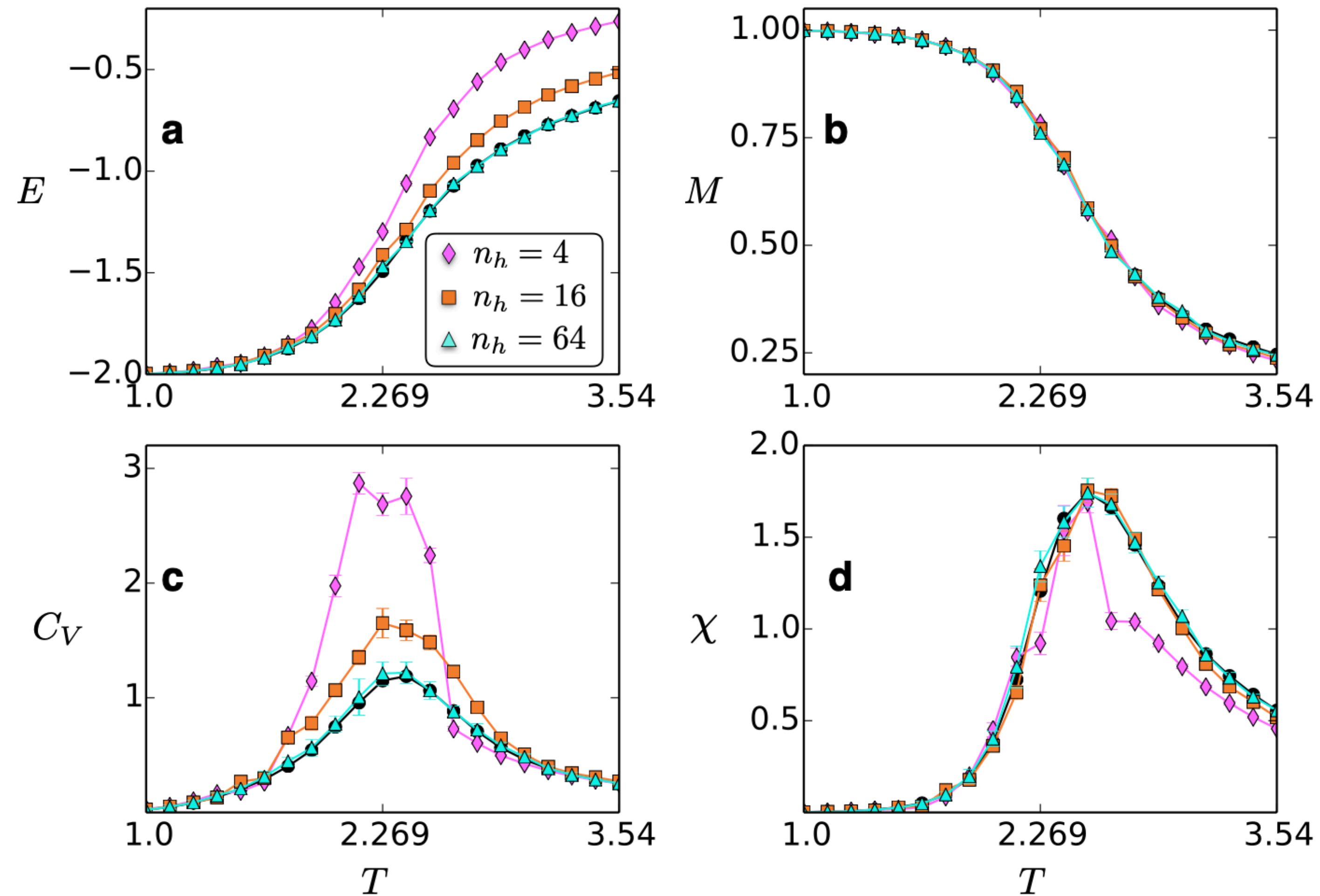
Fujimoto et al, JHEP21

- Number of parameters is not tied to the EoS structure
- The parameters lack clear physical interpretation
- **Alternatives?**

“Using Physics to Parametrize Physics”

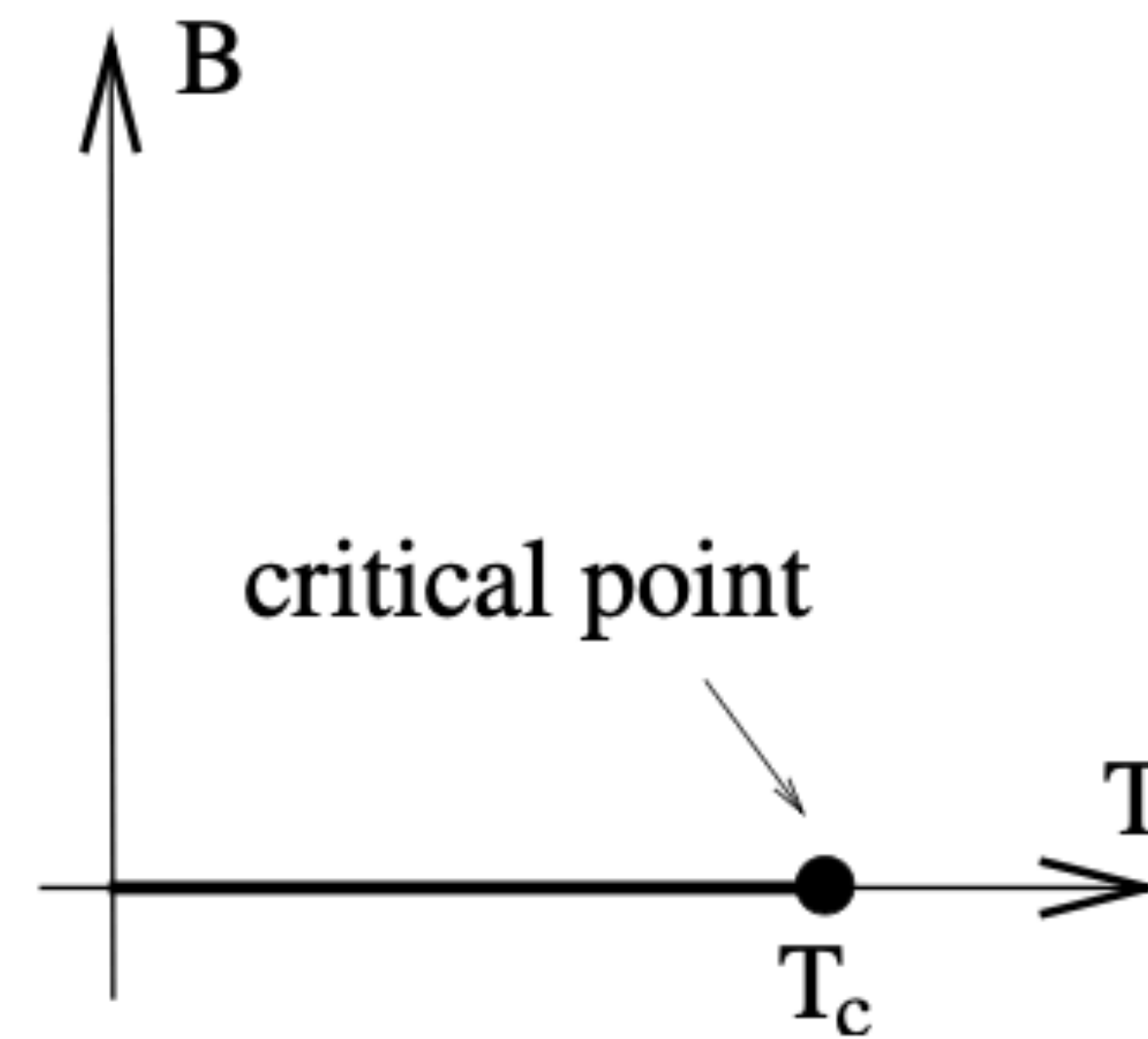
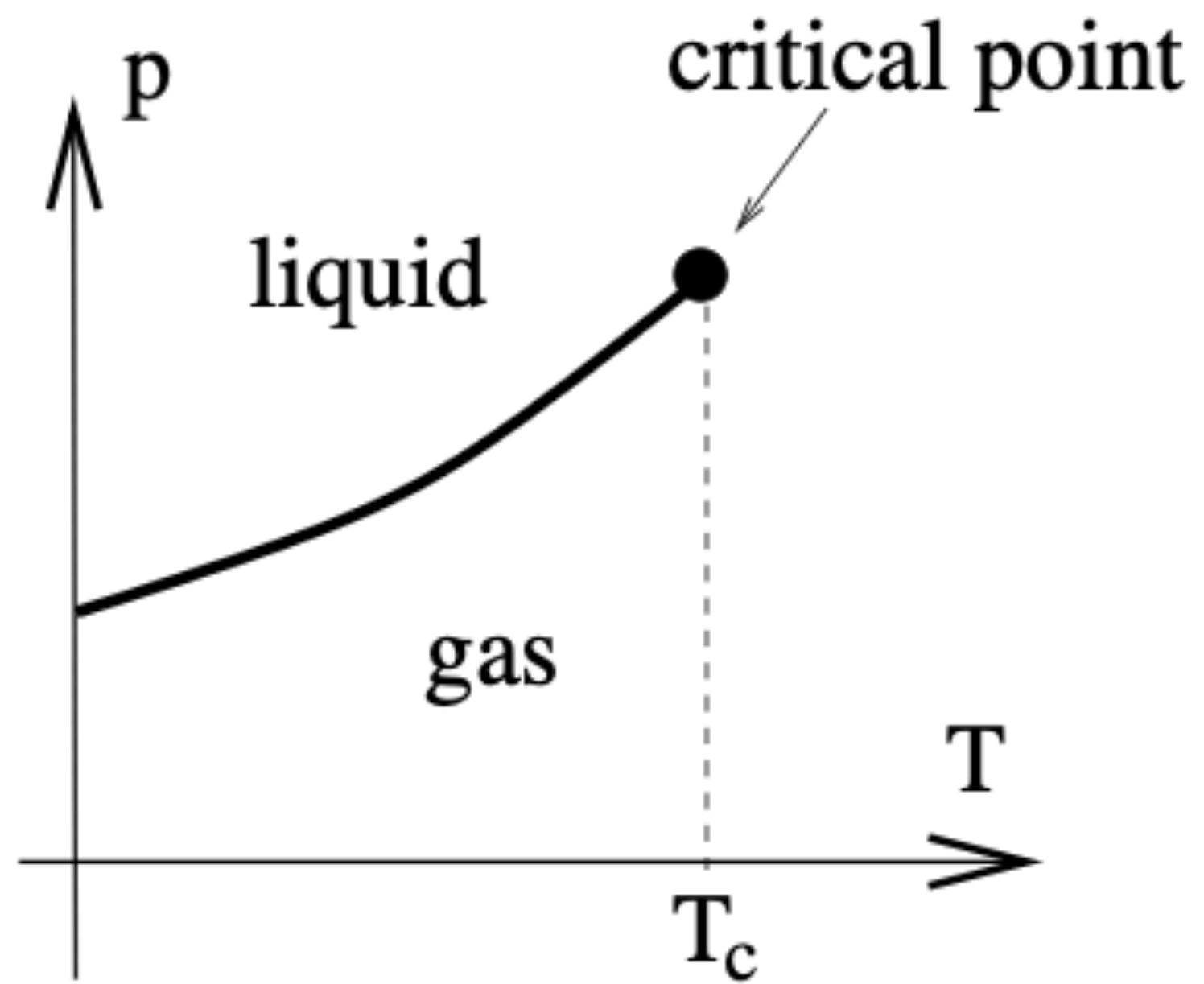


- Assumption: the E.o.S. of a class of physical systems can be mapped to **Restricted Boltzmann Machine (RBM)** EoS (generalizing mapping near the critical point)
- Parametrization EoS $\Leftrightarrow$ Parametrization RBM Hamiltonian
(Thermal stability is built-in)



- RBM faithfully reproduces the thermodynamics Ising model

Universality near the critical point

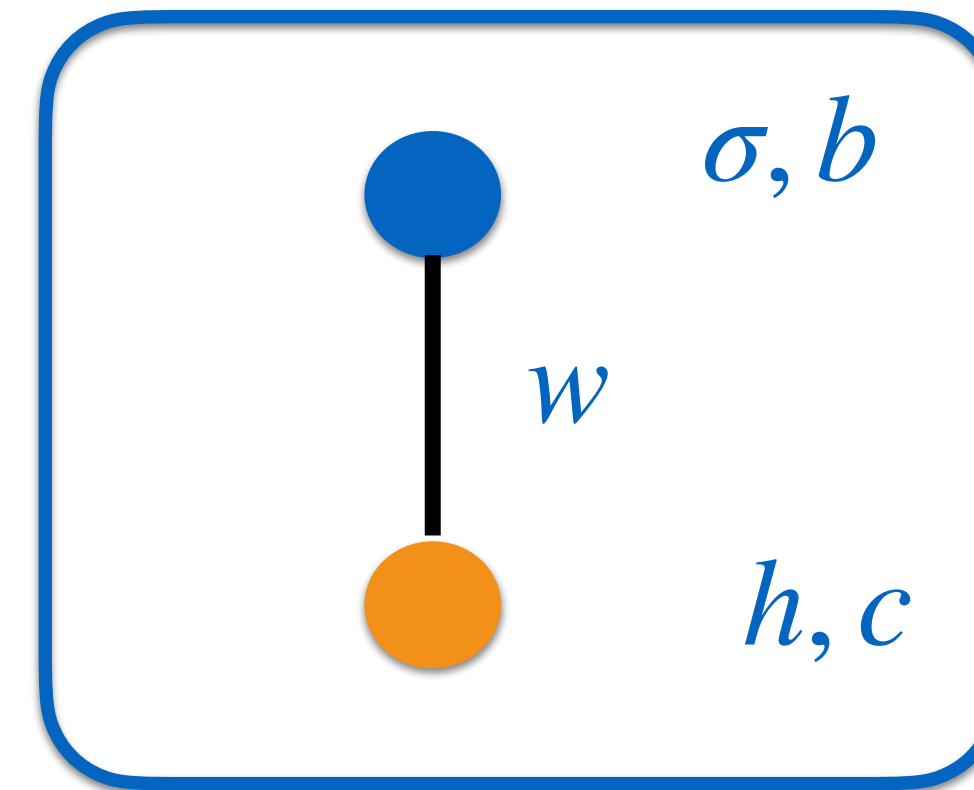


Learning Thermodynamics with RBM

- RBM faithfully reproduces the thermodynamics Ising model
Torlai-Melko, PRB2016
- Limitation: current RBMs match spin configurations, restricting generalization to other systems
- **Our proposal:** mapping at the partition function level

$$\log Z = \alpha \log Z_{RBM} + \beta$$

1d Ising model



- 1d Ising
- 1 visible + 1 hidden spin RBM
- **Analytically** mapping

$$w = \beta J, \quad b + c = \beta B$$

(2d Ising model in progress)

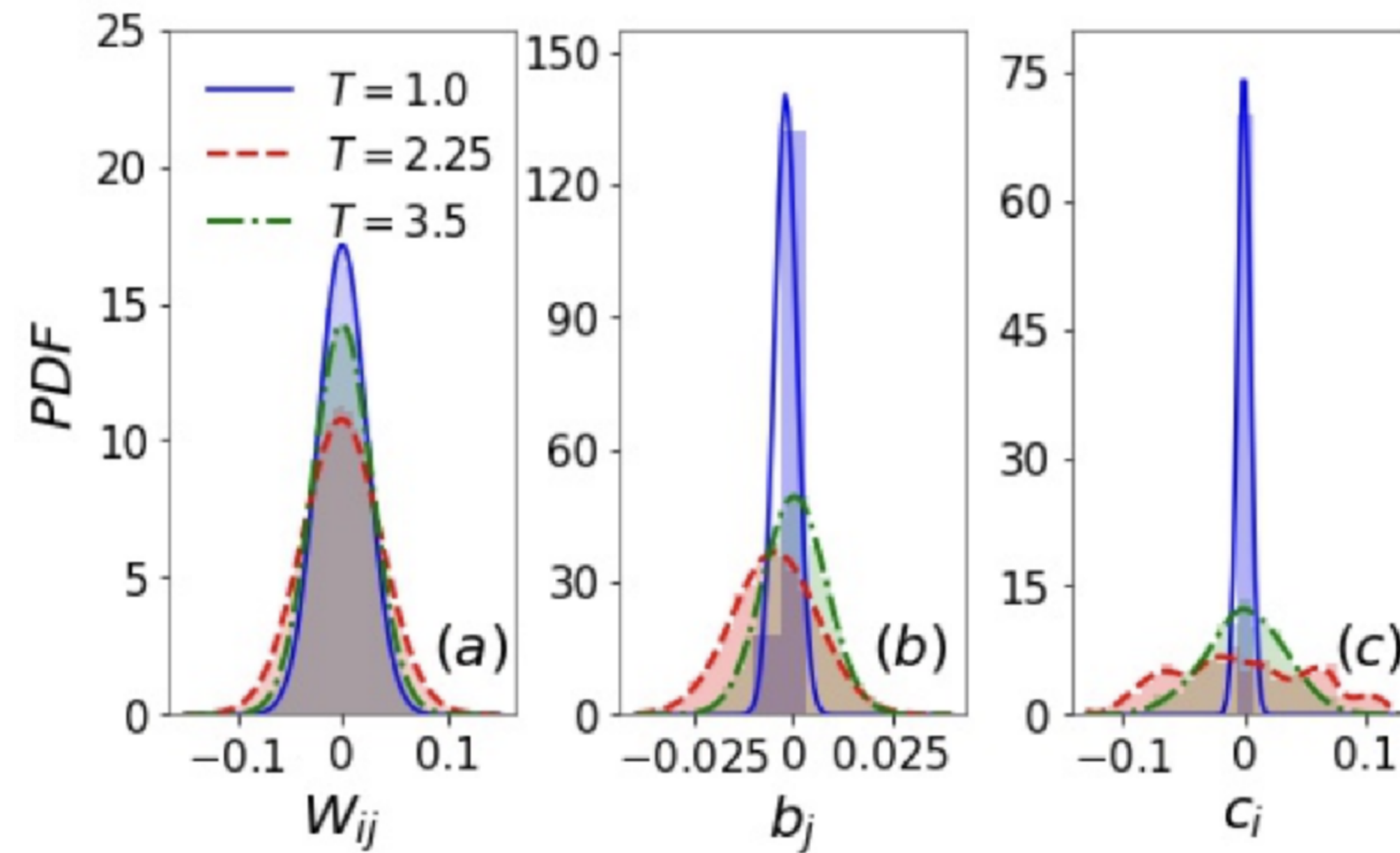
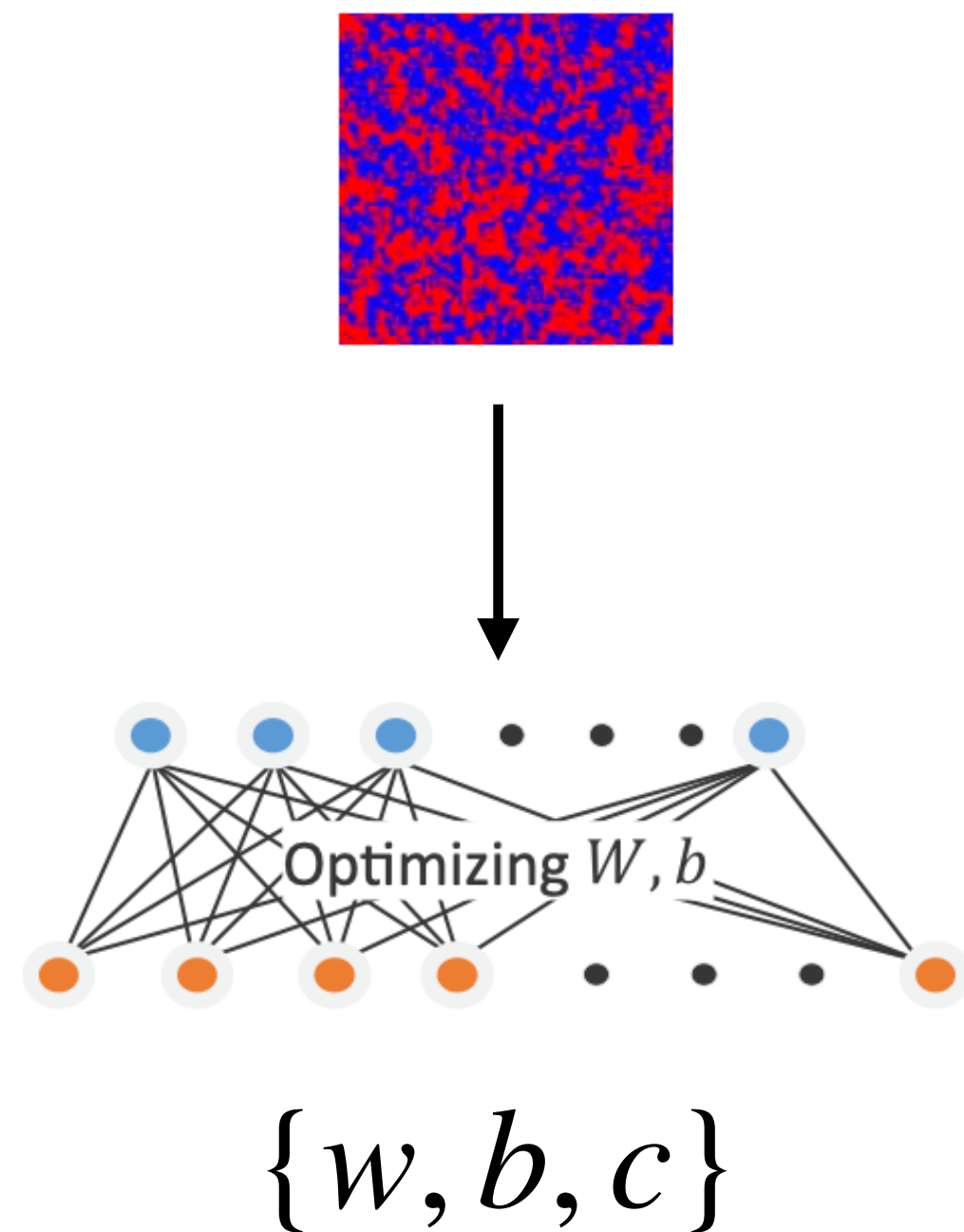
$$\sim \log \left[e^{\beta J} \cosh(\beta B) + e^{-\beta J} \sqrt{1 + e^{4\beta J} \sinh^2(\beta B)} \right]$$

$$\sim \log \left[e^w \cosh(b + c) + e^{-w} \cosh(b - c) \right]$$

$$\cosh(b - c) = \sqrt{1 + e^{4\beta J} \sinh^2(\beta B)}$$

$(B = 0 \leftrightarrow b = c = 0)$

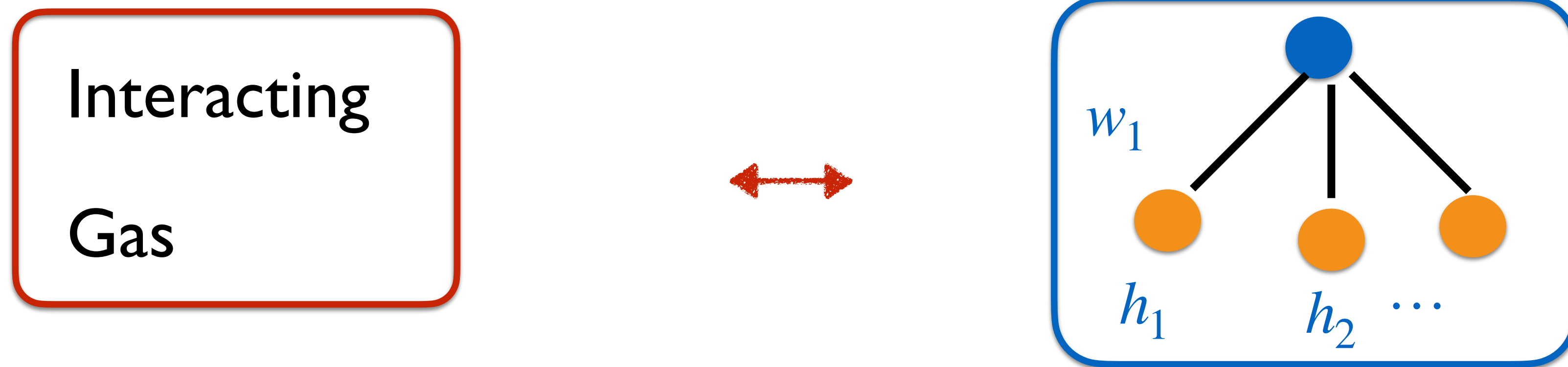
Matching Configuration vs Matching Partition Function



The histogram of RBM parameters with 400 hidden units that reproduces spin configurations of 2d Ising model at $B=0$ (Gu-Zhang entropy 2020)

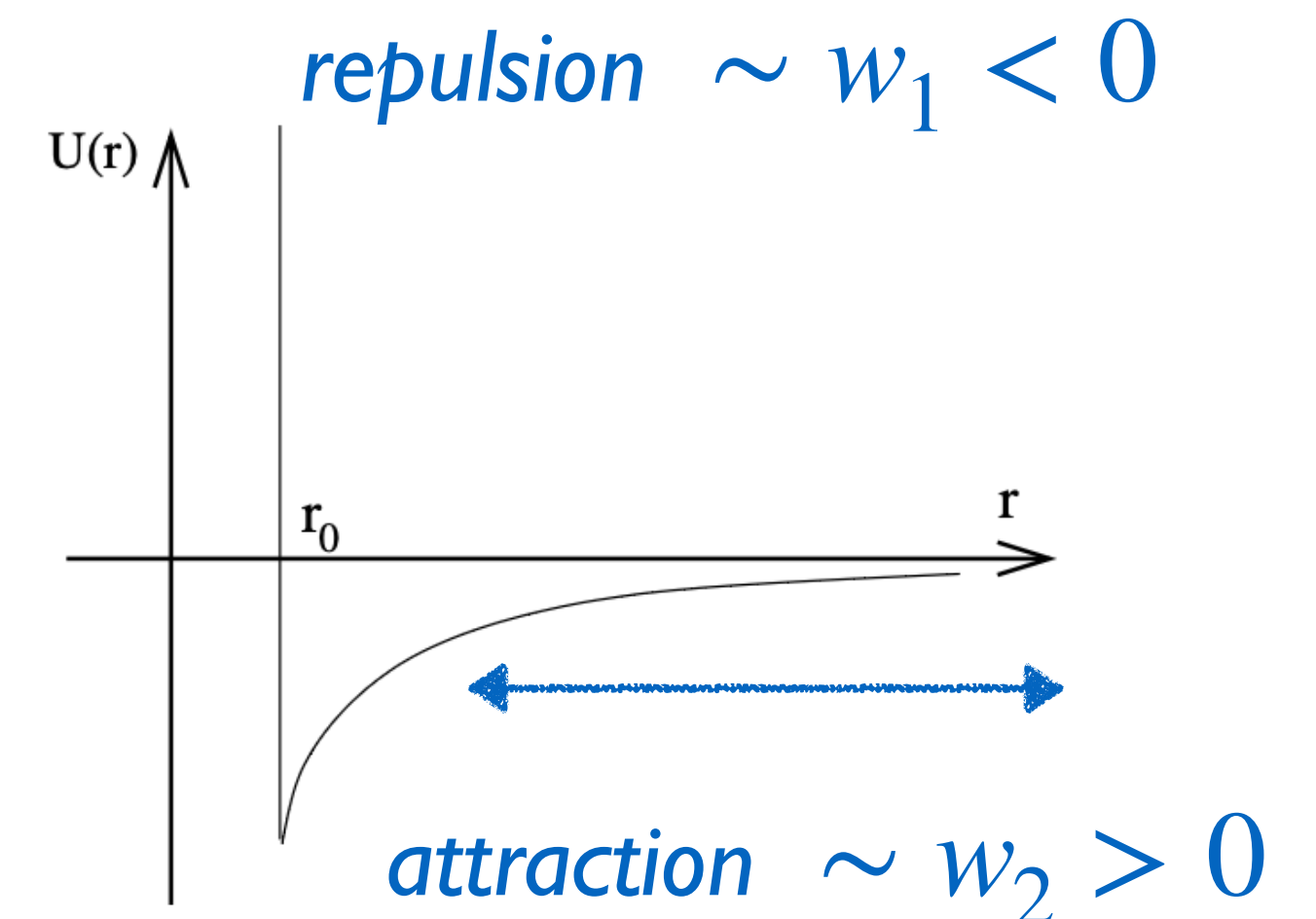
- Thermodynamic function matching only needs representative weights and "biases" → much **smaller** number of units.

Interacting Gas



- Ideal gas $\sim$ RMB with 1 visible spin (bias $\sim \mu$)
- Interaction corrections $\sim$ coupling hidden spins

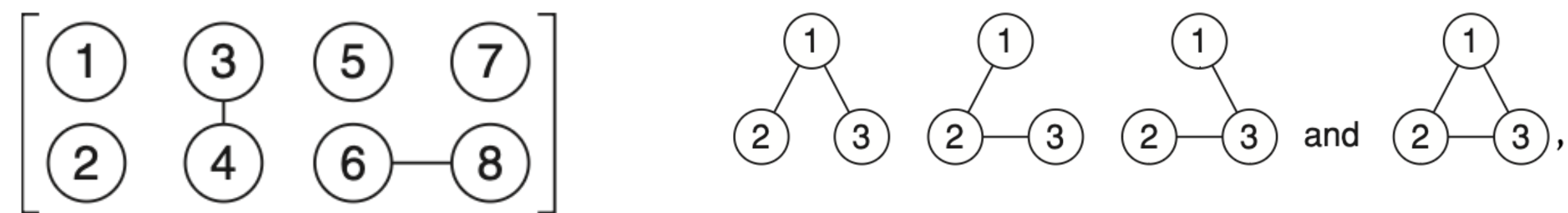
$$\int dr r^2 (e^{-\beta u(r)} - 1) \sim (e^{w_1} - 1) + (e^{w_2} - 1) + \dots$$



- Hard sphere gas: 2 hidden spins with w_1, w_2 describe repulsive and attractive regime

Cluster Expansion

$$\ln Z = \ln Z_0 + \sum_{\text{connected clusters}} (\text{cluster contribution})$$



- N-cluster effects and interaction potential features are encoded by hidden spins + additional visible spins

Summary and Outlook

Outlook

- Nature is a teaching machine. We just begin to learn AI from it.

Back-up

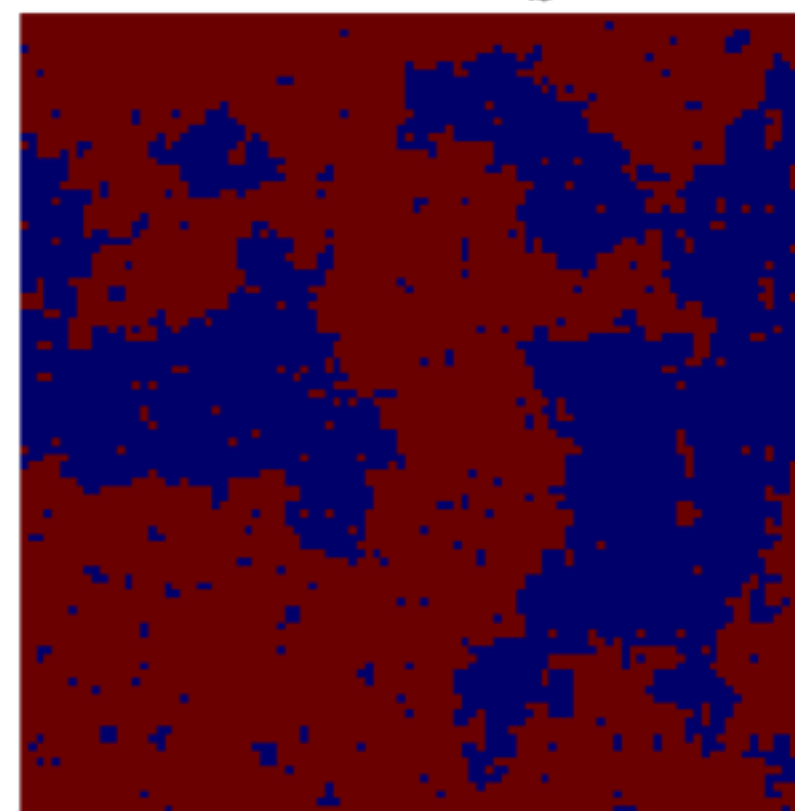
...“the world” is something like a great chess game being played by the gods, and we are observers of the game. We do not know what the rules of the game are; ... Of course, if we watch long enough, we may eventually catch on to a few of the rules.

Feynmann

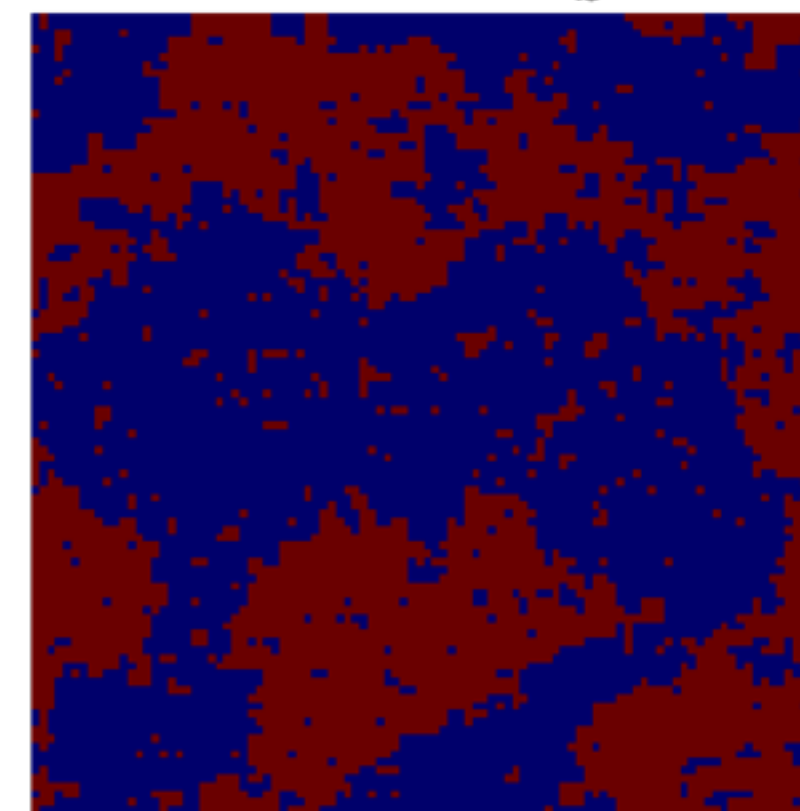
1 [J/K_b]



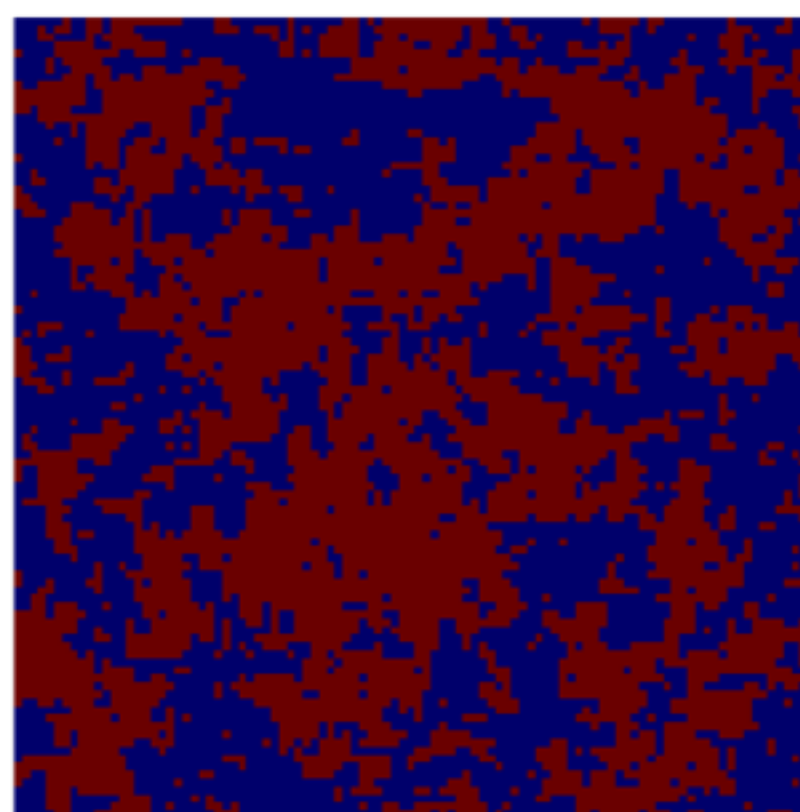
2 [J/K_b]



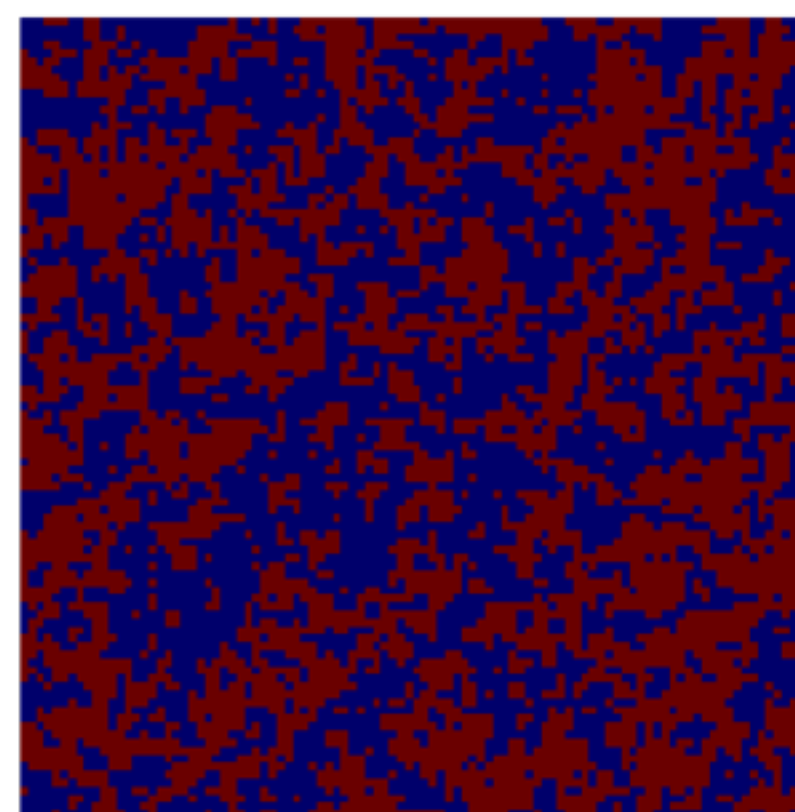
2.2 [J/K_b]



2.5 [J/K_b]



3 [J/K_b]



5 [J/K_b]

