Machine Learning from the Perspective of Physics

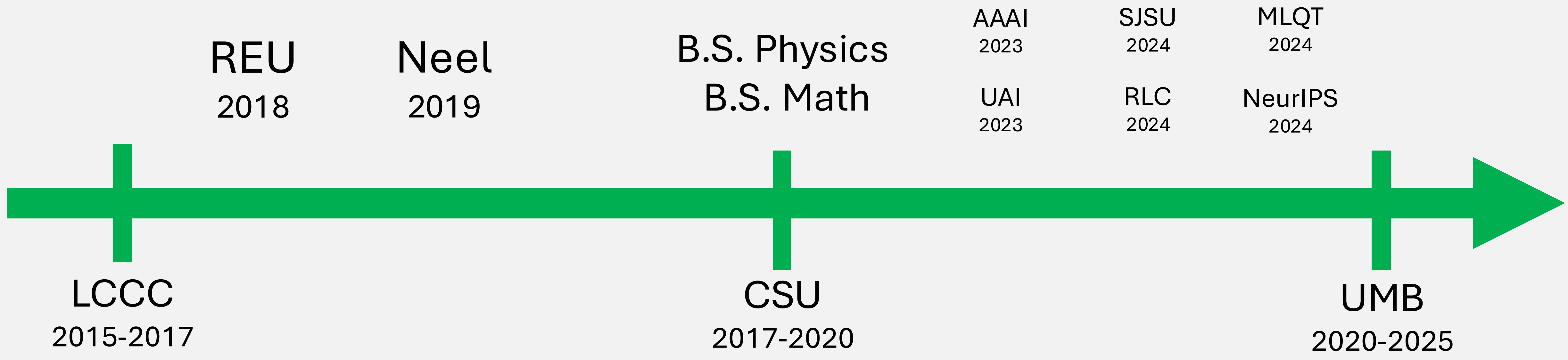
Jacob Adamczyk

Outline

- My Trajectory
- What is AI?
- AI in Physics
- Physics in AI
- Reinforcement Learning
- My Research
- The Future

My Trajectory

My Journey

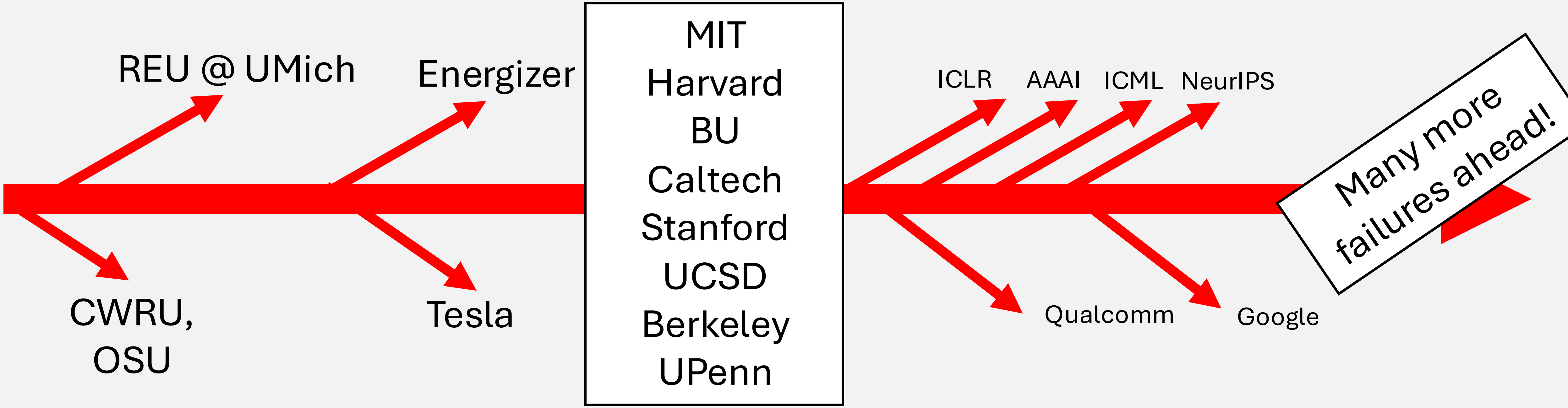


- Research with Dr. S (Microgels)
- Research with Dr. Kaufman (Stat Mech)
- Research with Dr. Heus (LES)
- Research with Dr. Stella-Gold (Lie Theory)

- Honors College
- SPS Involvement
- Weekly Physics Questions
- Travel to NOURS, OSAPS, APS

- Sigma Pi Sigma Induction
- Machine Learning Club (Nikša)
- Learning how to do research
- First paper with Dr. S

My Journey


REU @ UMich
Energizer
CWRU,
OSU
Tesla
MIT
Harvard
BU
Caltech
Stanford
UCSD
Berkeley
UPenn
ICLR
AAAI
ICML
NeurIPS
Qualcomm
Google
Many more
failures ahead!

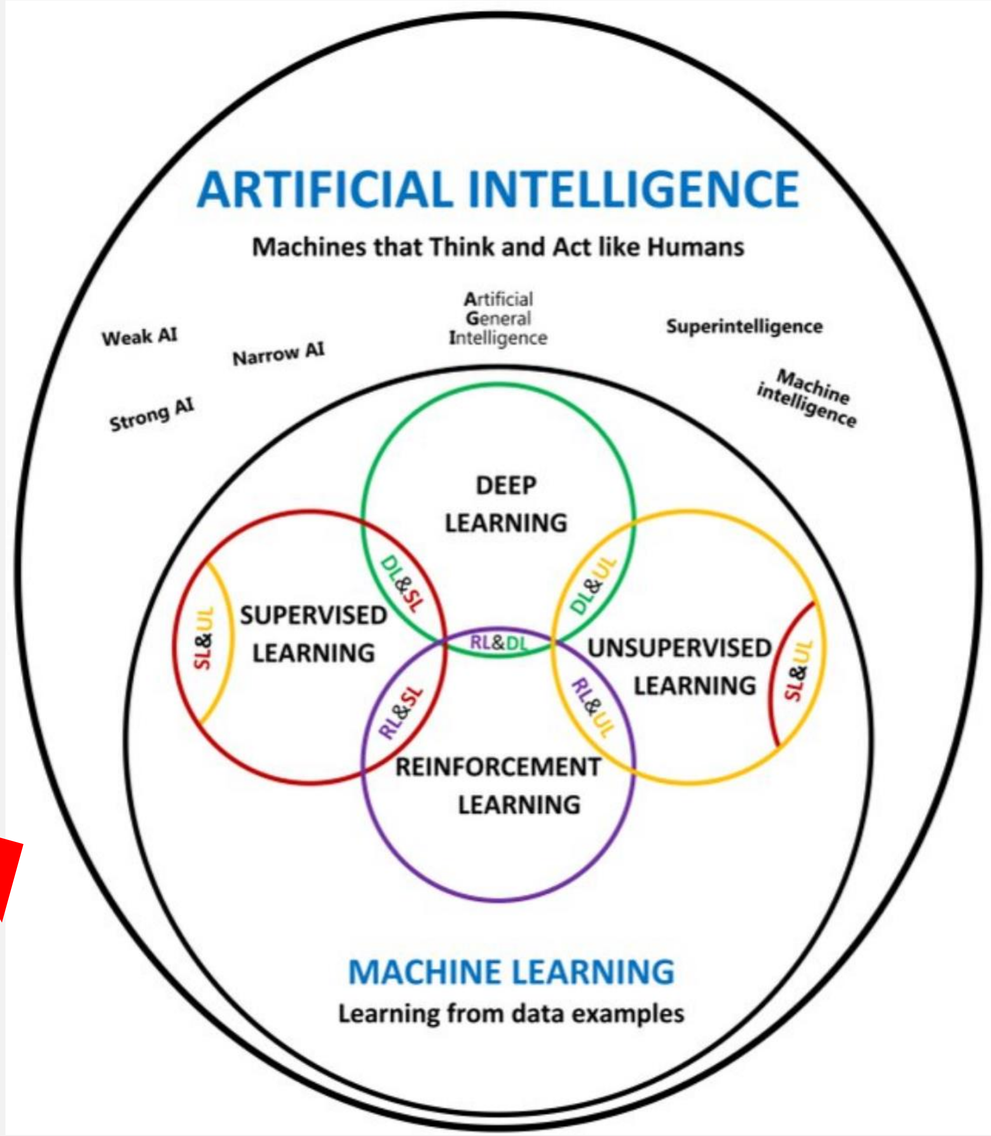
What is AI?

What is AI?

A general term for any "intelligent" system

- Yesterday, Rule-based GOFAI
- Today, learning by GD is the rage
- Tomorrow, "zero-shot in-context learning by 100T param. GPT"

Useless diagram

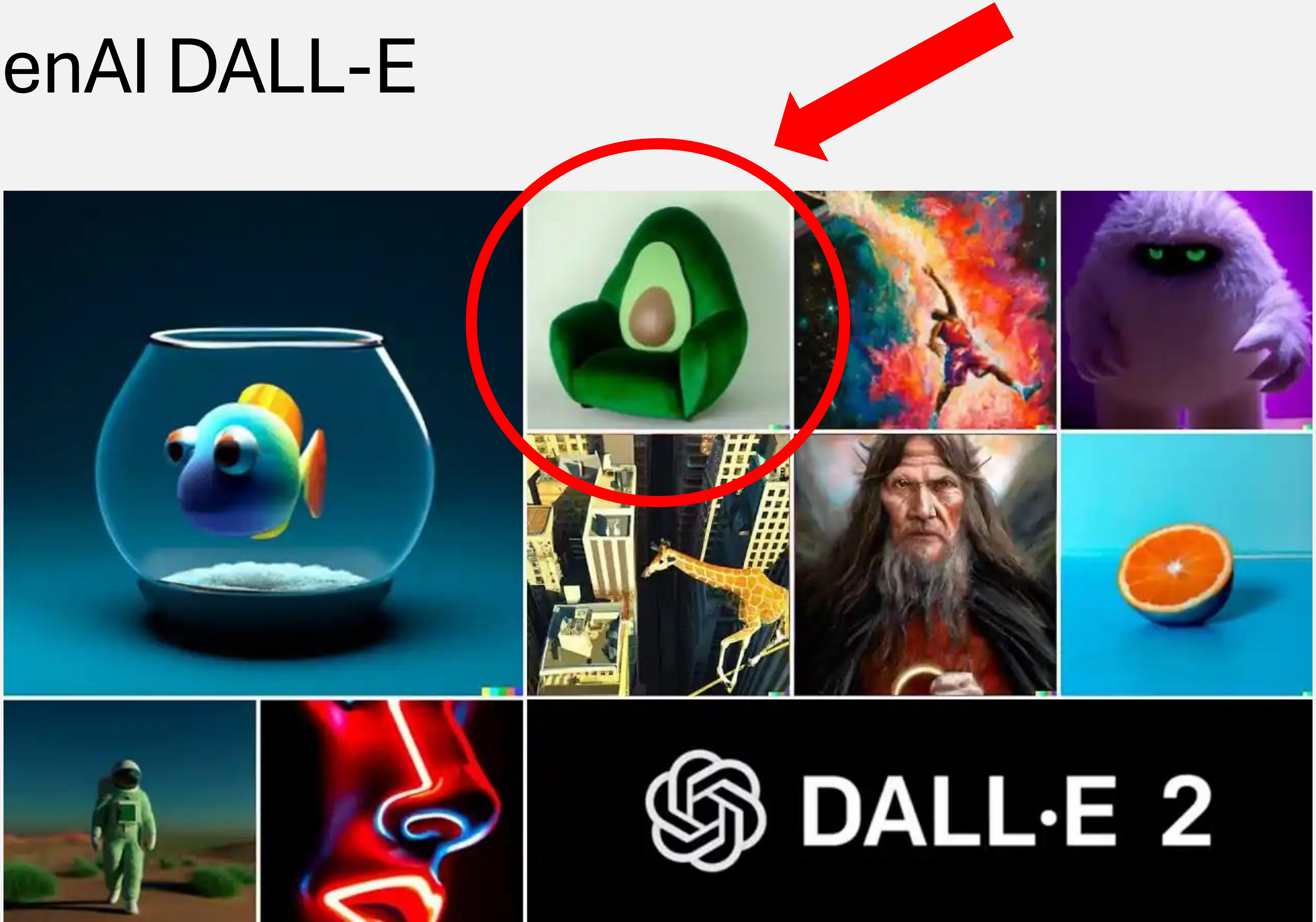


Instead of attempting to define,
let’s look at some examples

Cool Breakthroughs in AI

OpenAI DALL-E

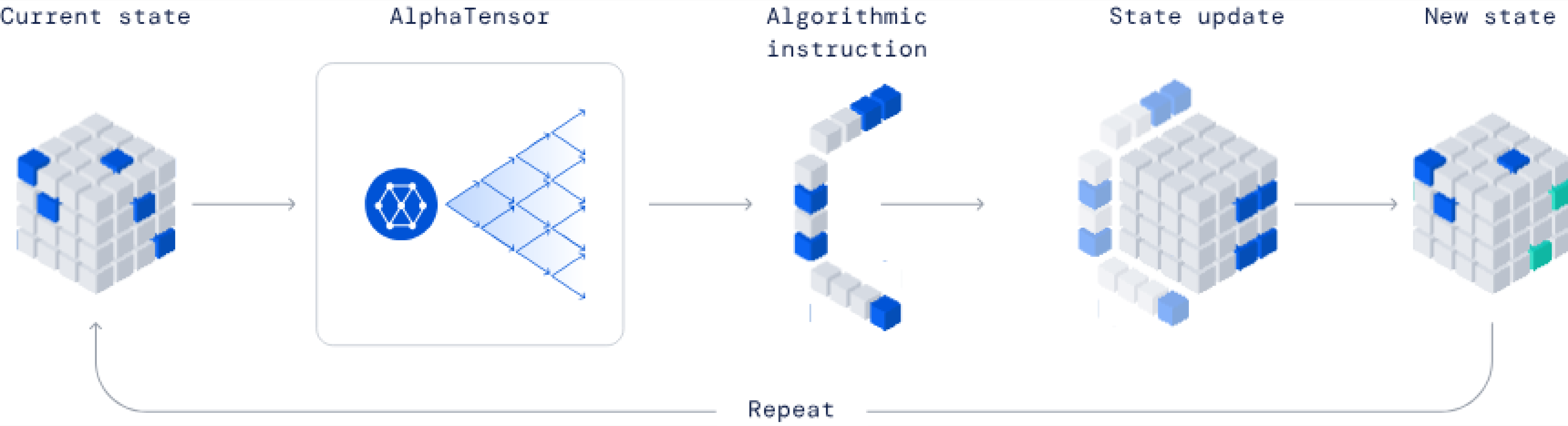
OpenAI DALL-E



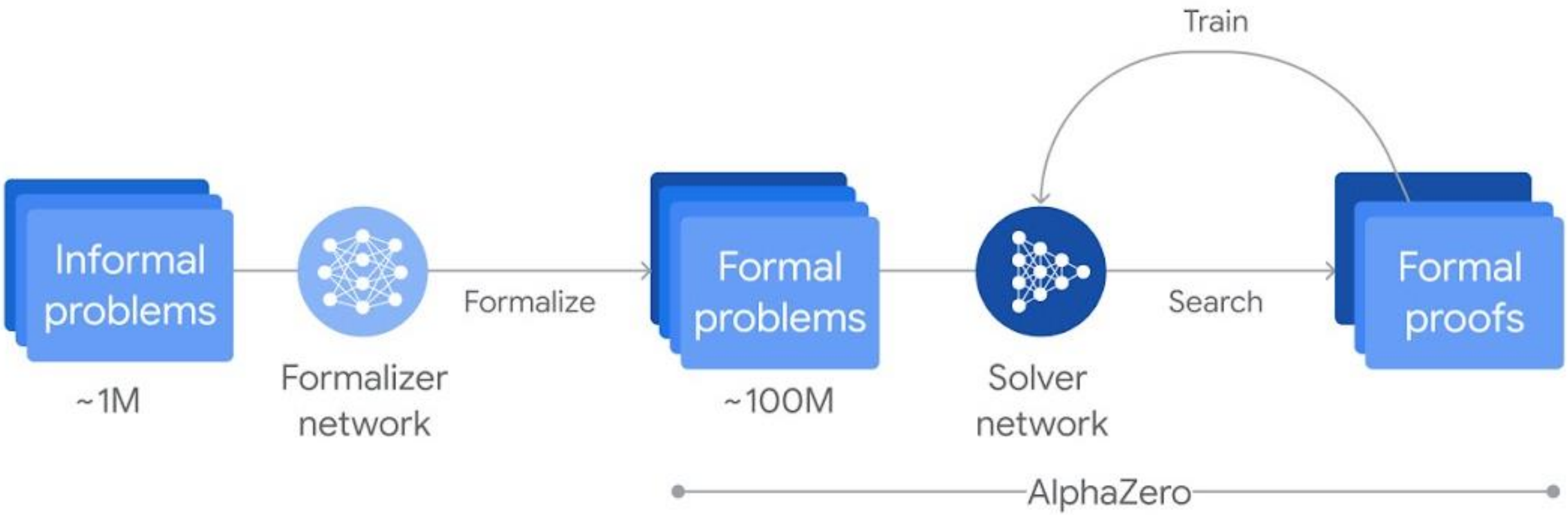
Music Generation (Google SeaNet)

(Try suno.com!)

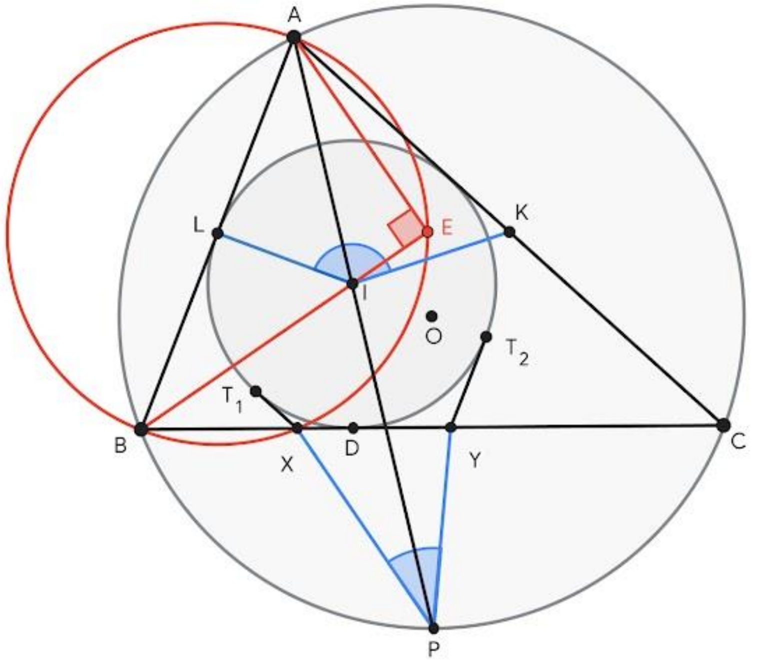
Algorithm Discovery



IMO

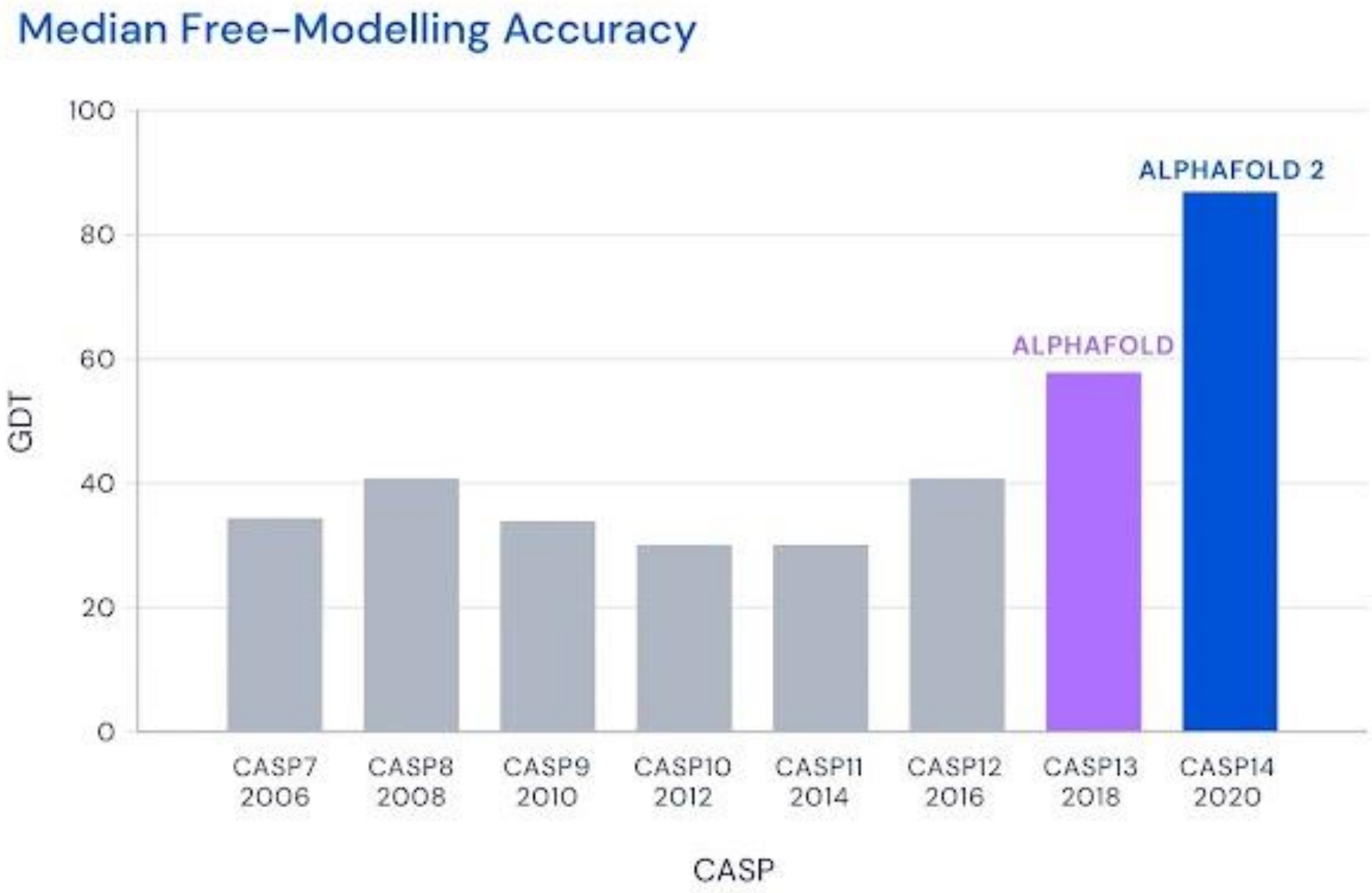


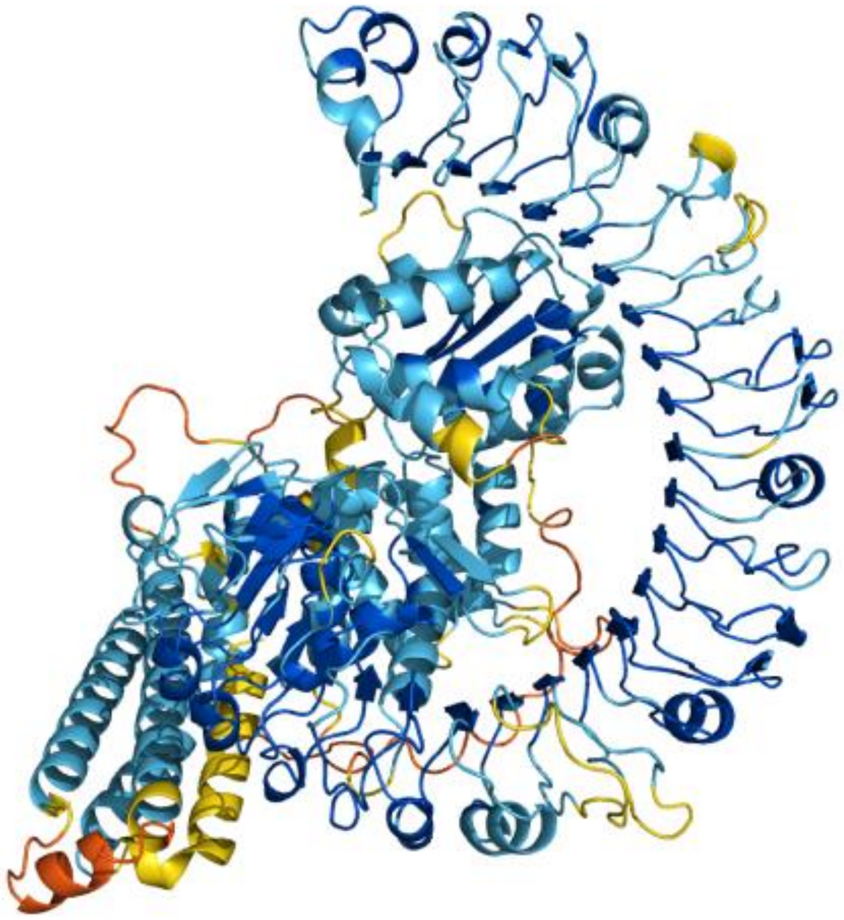
Score on IMO 2024 problems

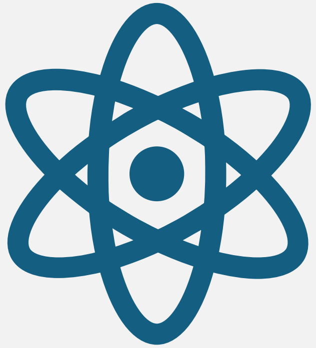


Video Generation (OpenAI Sora)

AlphaFold


Median Free-Modelling Accuracy
100
80
60
40
20
0
GDT
ALPHAFOLD 2
ALPHAFOLD
CASP7 2006
CASP8 2008
CASP9 2010
CASP10 2012
CASP11 2014
CASP12 2016
CASP13 2018
CASP14 2020
CASP



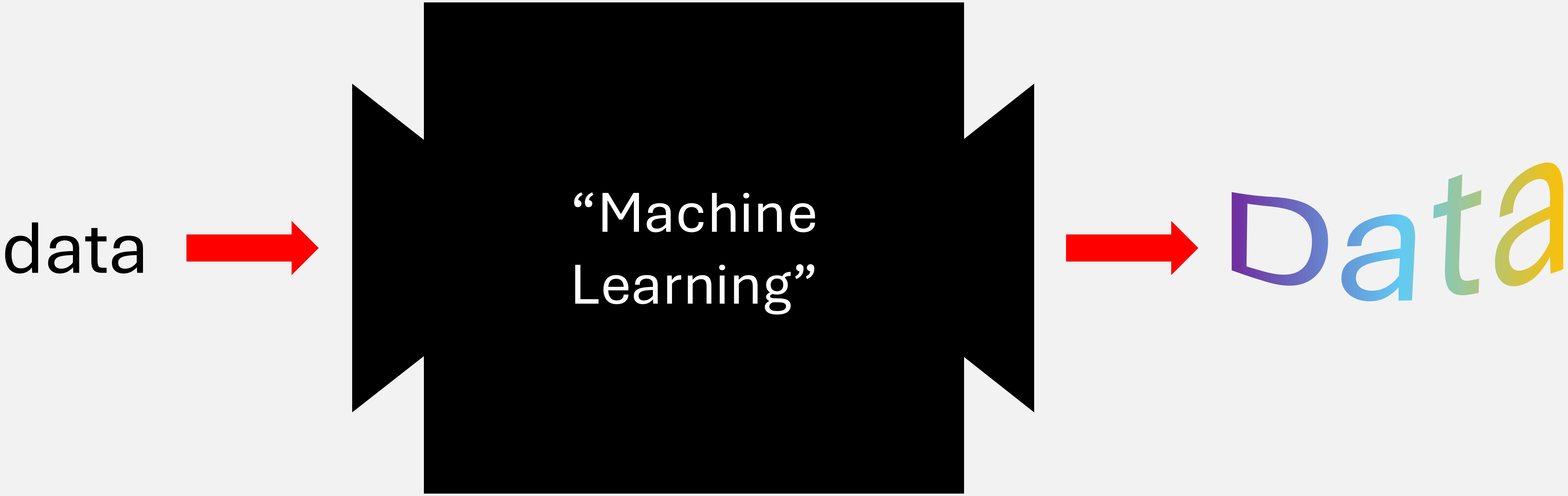


AI in Physics

Physics in AI

How Do Physicists Use AI?

“Experimental” Physicist

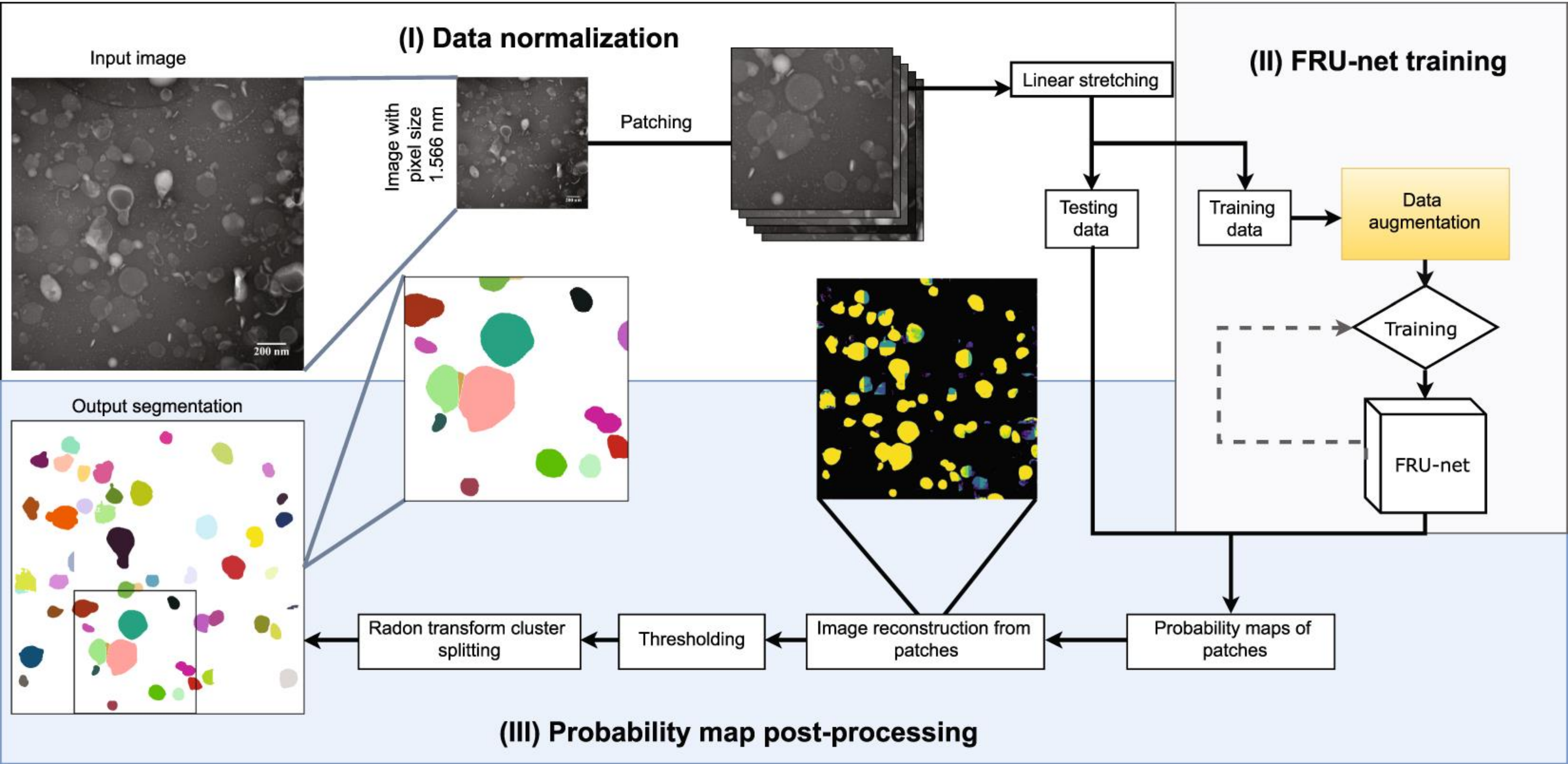


Examples

- Segmentation of images
- Data filtering
- Anomaly detection
- Data generation
- Predict material properties (T_c?)

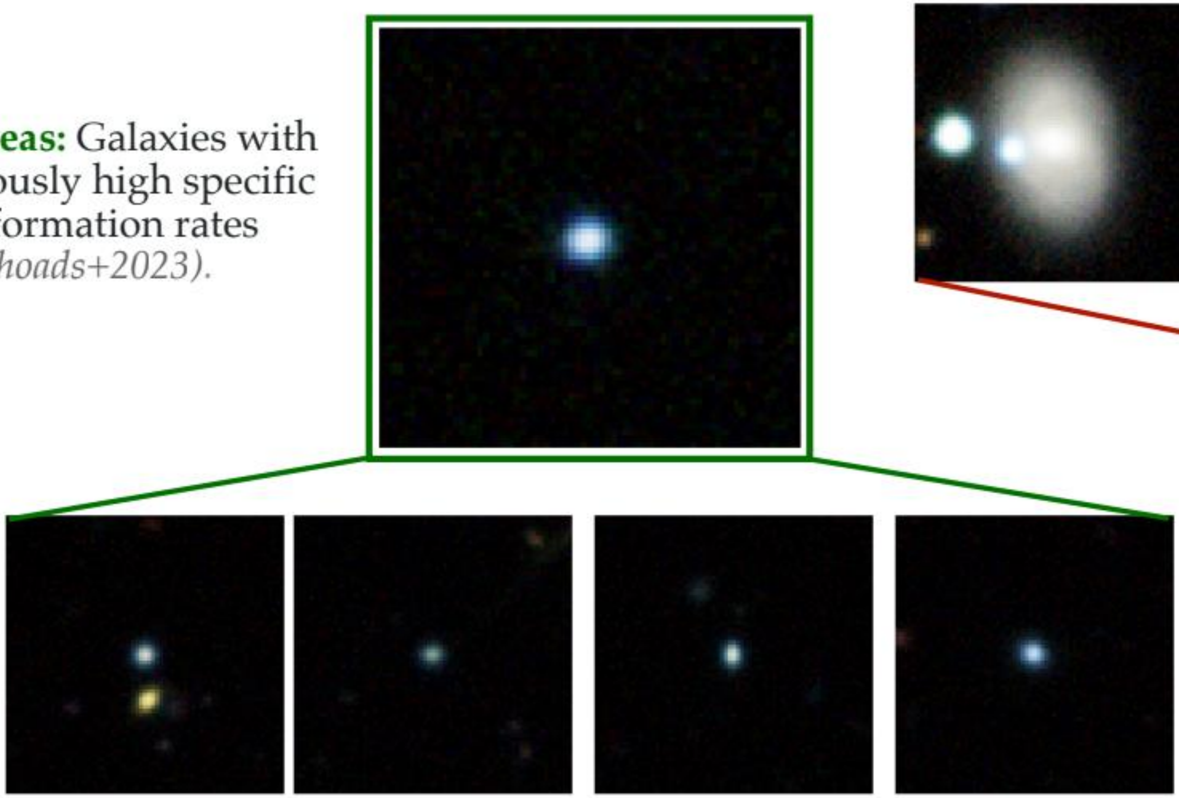
Visit the IAIFI website to see a lot of cool research!

Gómez-de-Mariscal et. al. 2019
…"Segment Anything"…



Recovering Galaxy Anomalies in the Latent Space

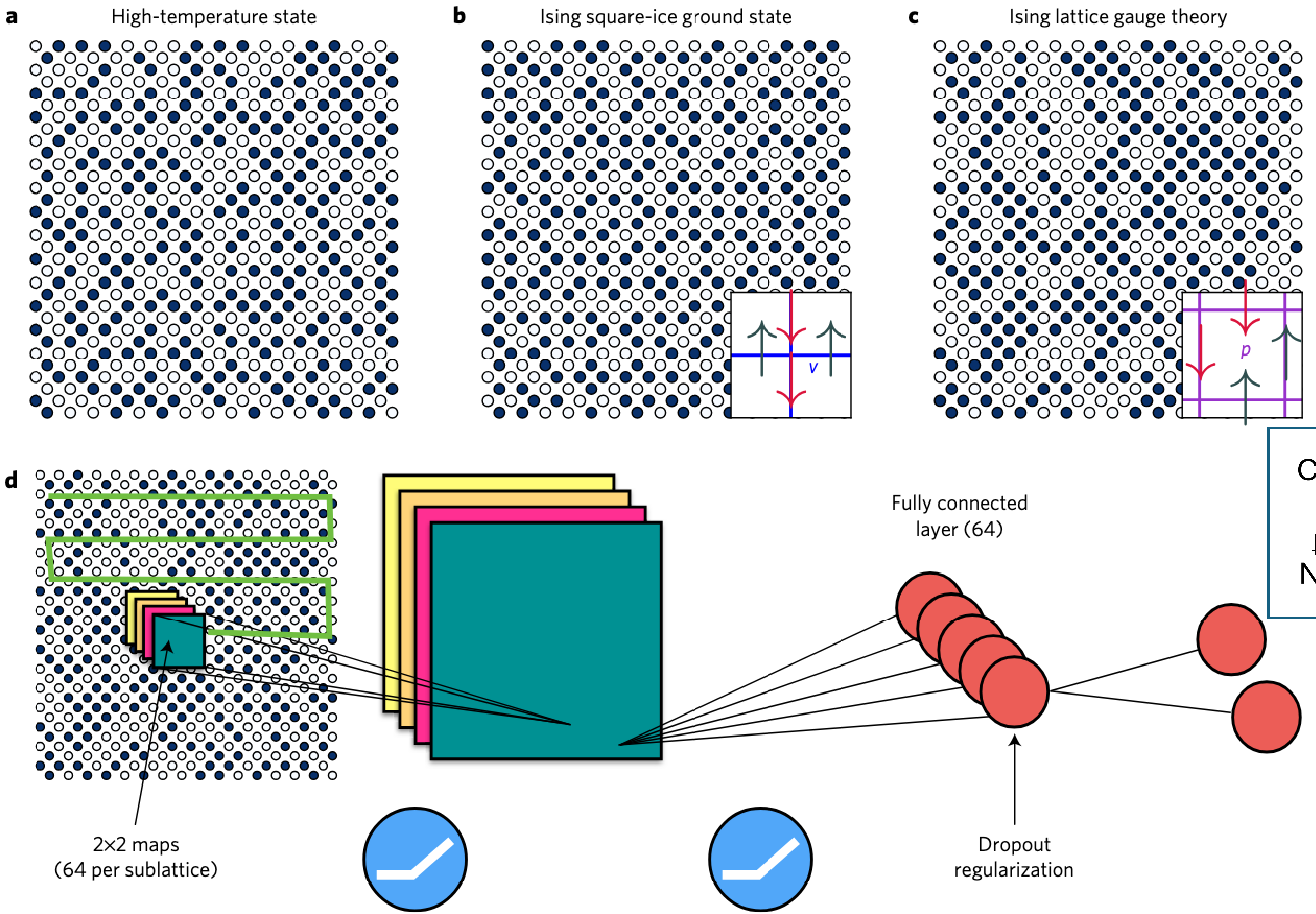
Green Peas: Galaxies with anomalously high specific star-formation rates *(Rhoads+2023)*.



Red Spirals: Galaxies with anomalously low ongoing star-formation *(Masters+2010)*.

(Gagliano & Villar+23 NeurIPS)

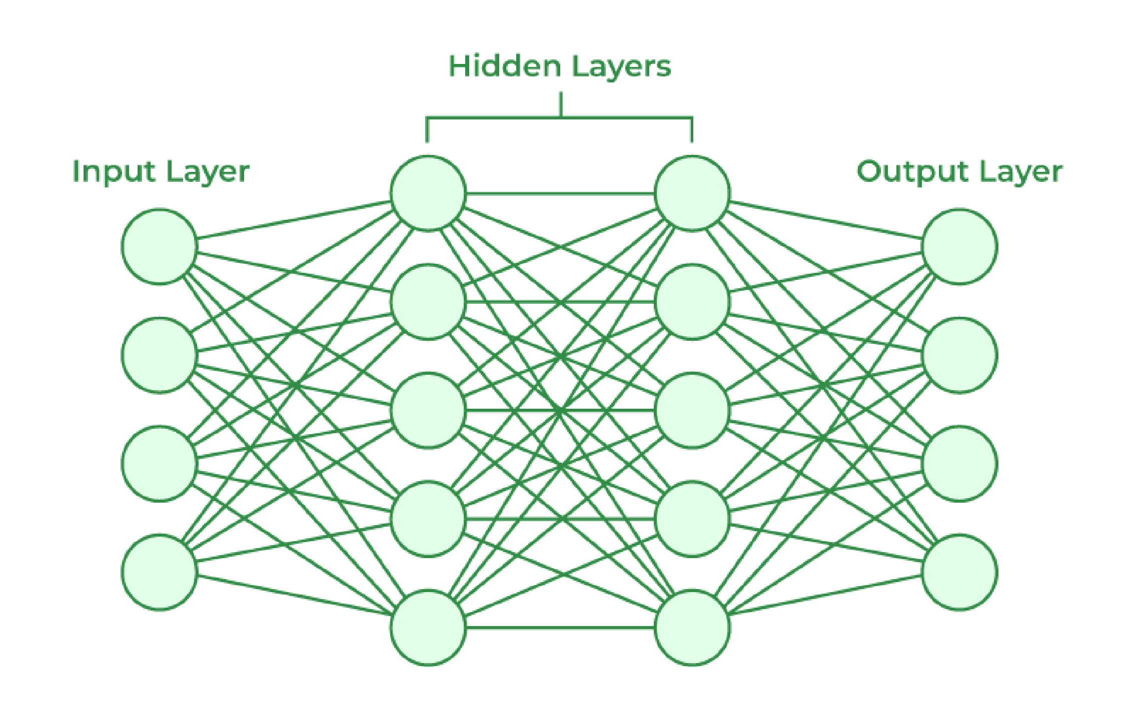
Soon-to-be applied to the SN problem - stay tuned!



Carrasquilla & Melko "Machine learning phases of matter", Nature physics 2017

“Computational” Physicist

A machine that can be “understood” and “engineered”



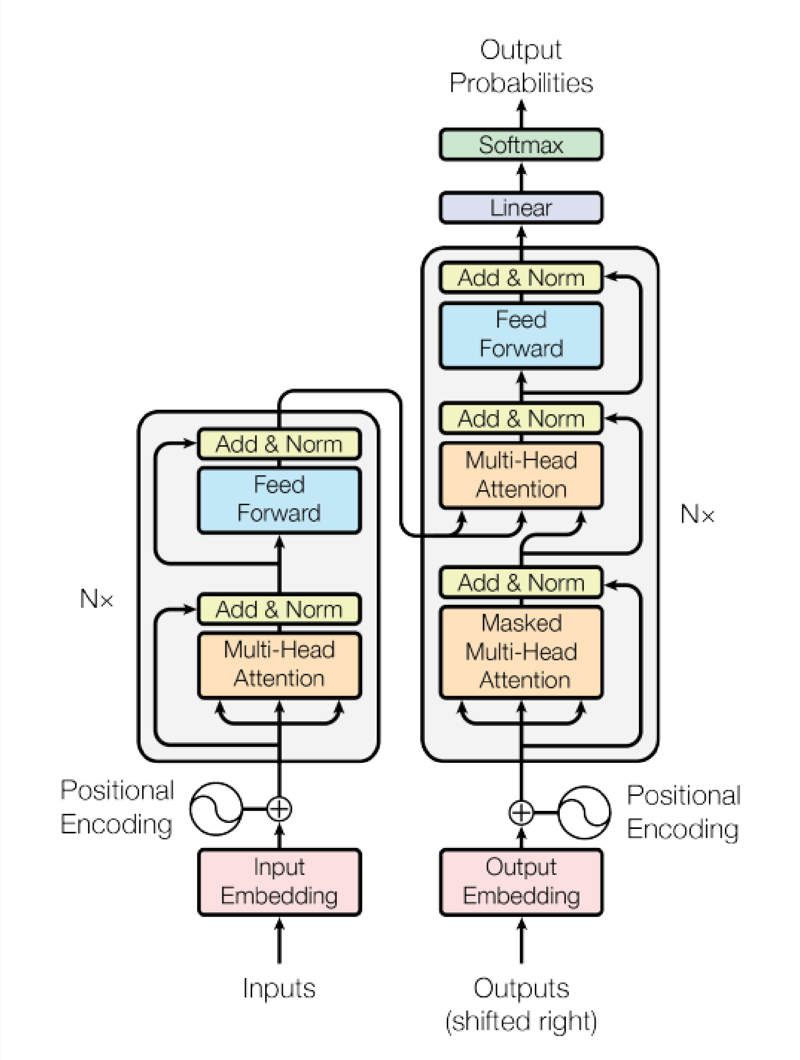
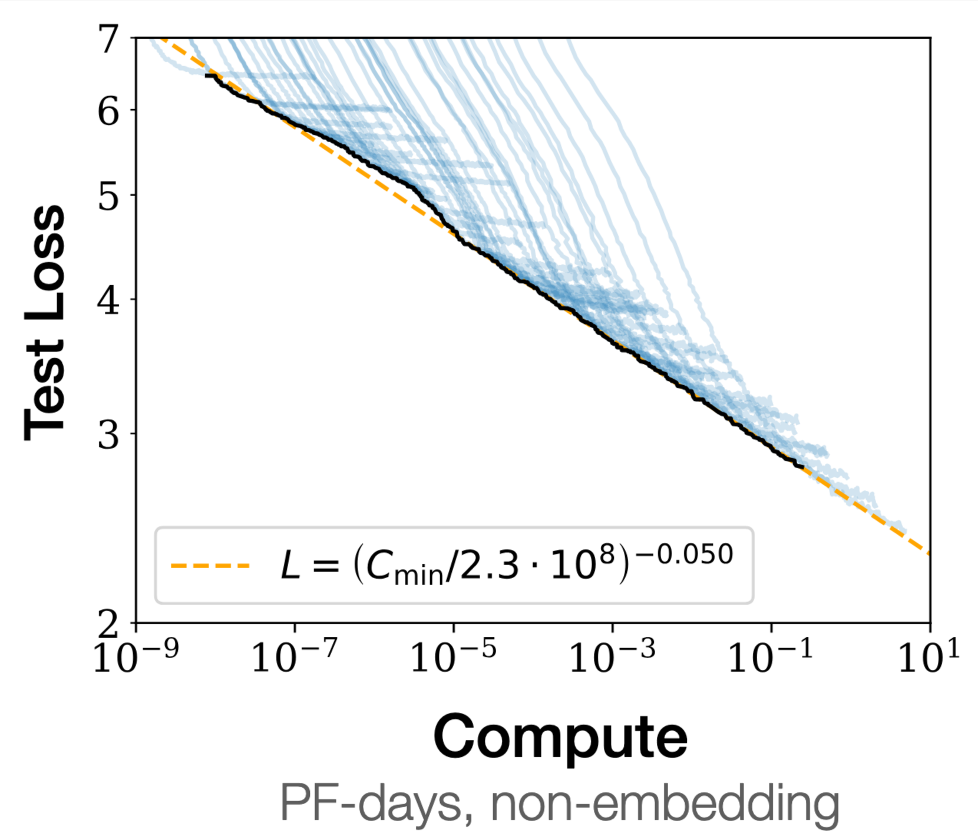
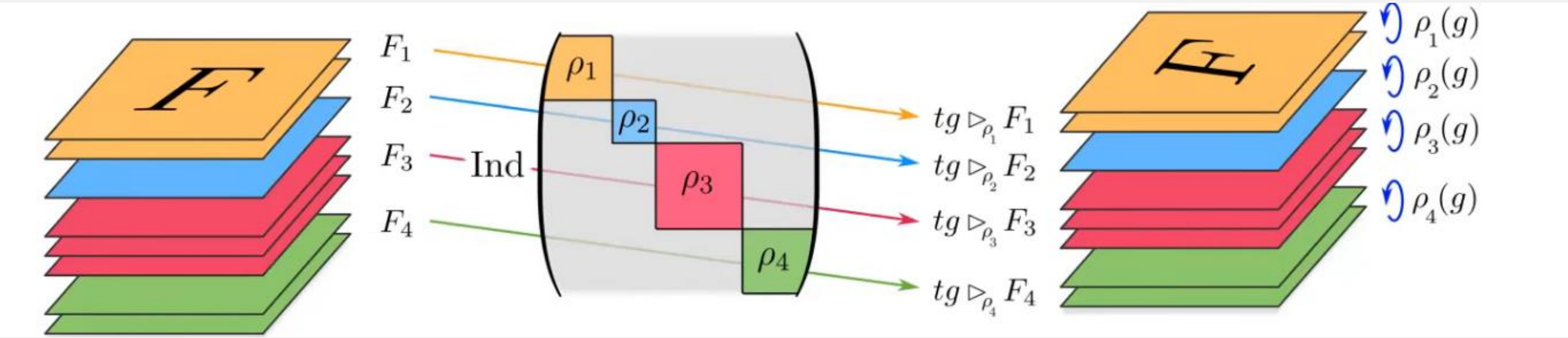


Figure 1: The Transformer - model architecture.

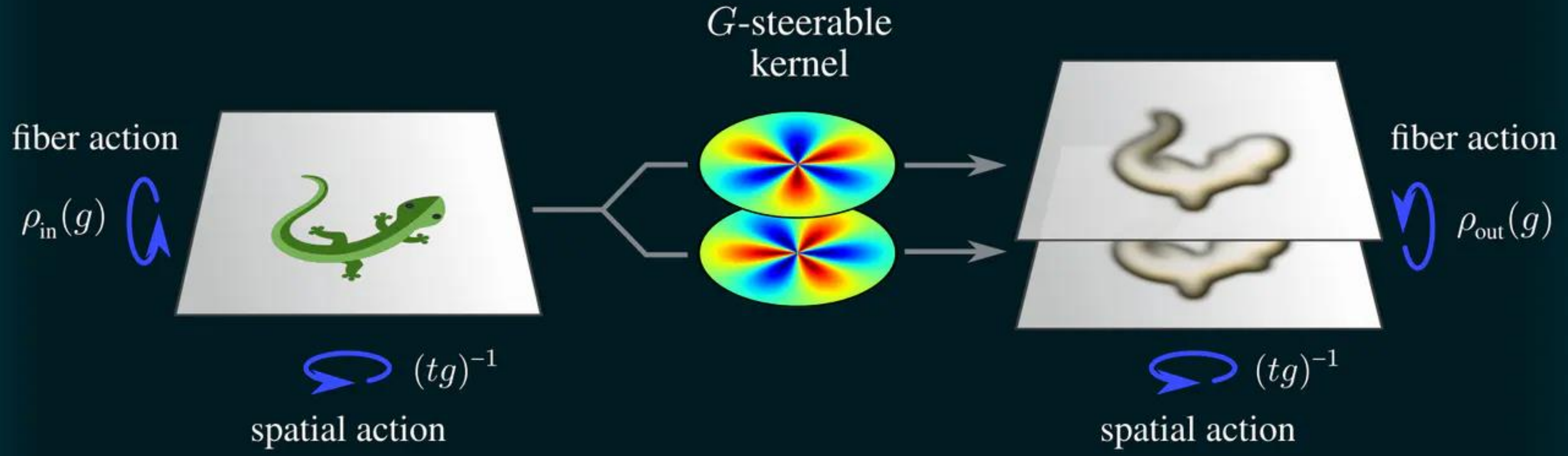
“Computational” Physicist

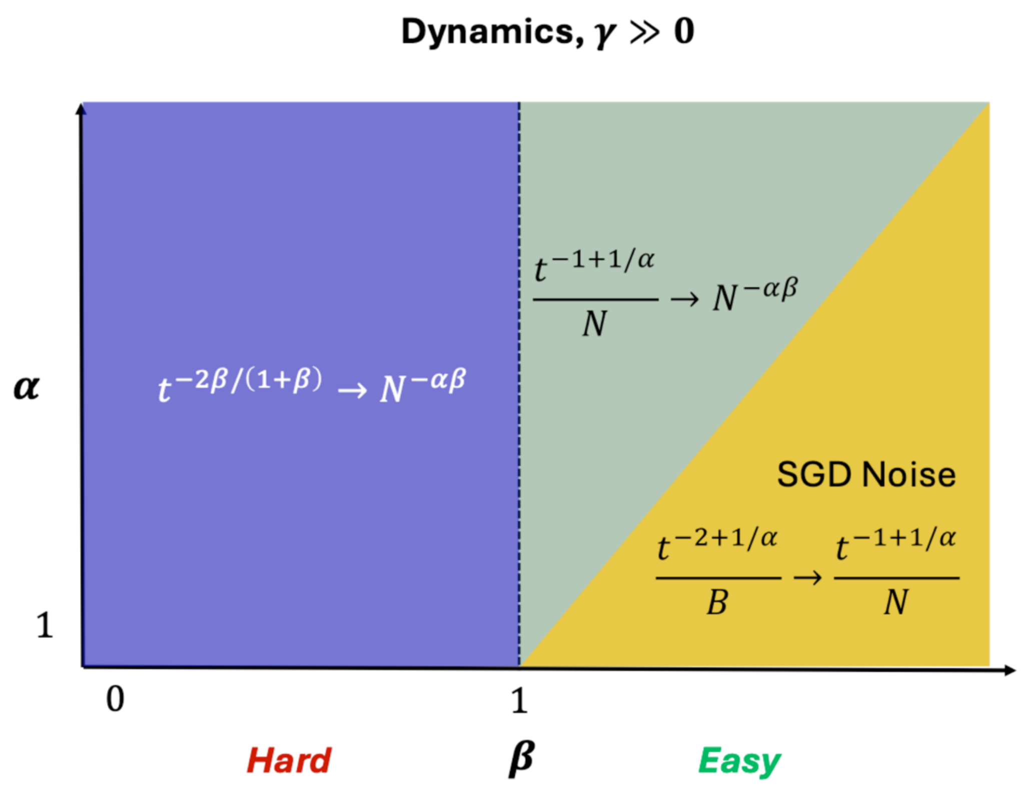
- Scaling Laws
- Invent new architectures
- “Geometric” Machine Learning
- New training objectives

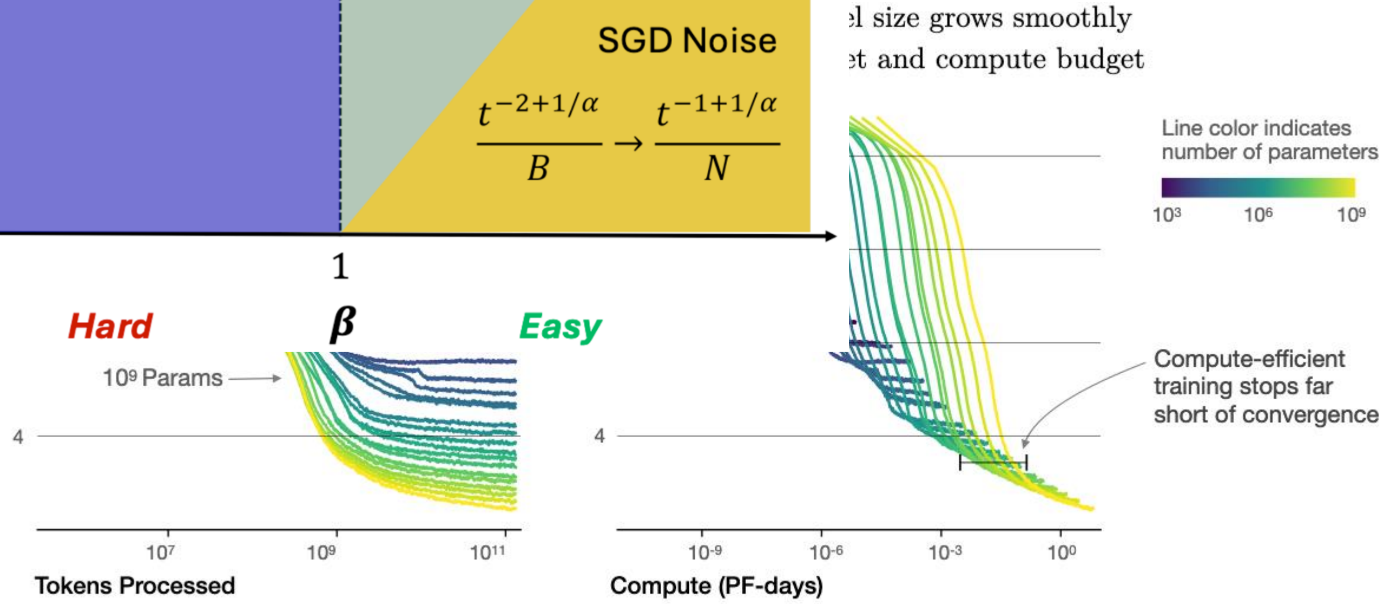


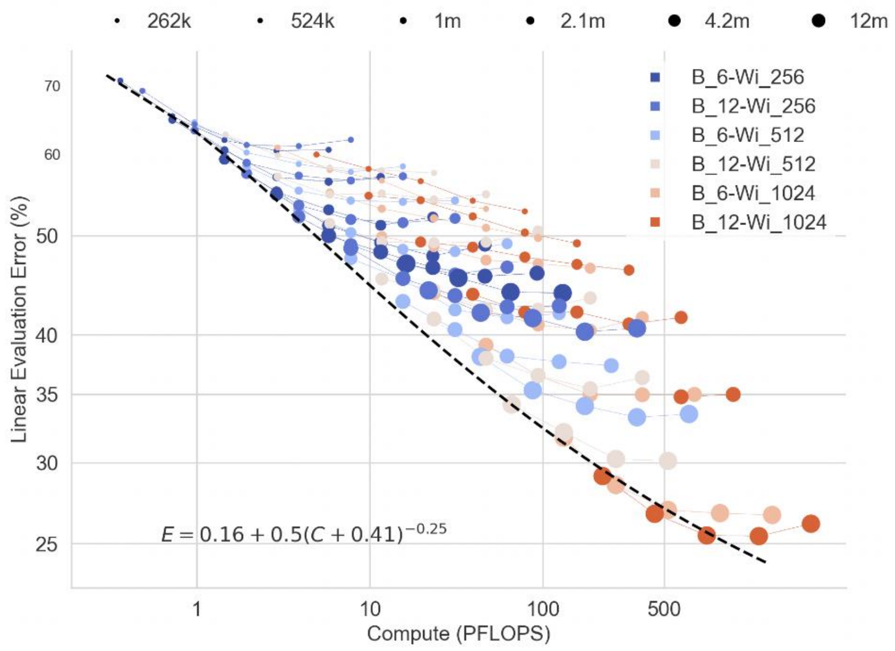


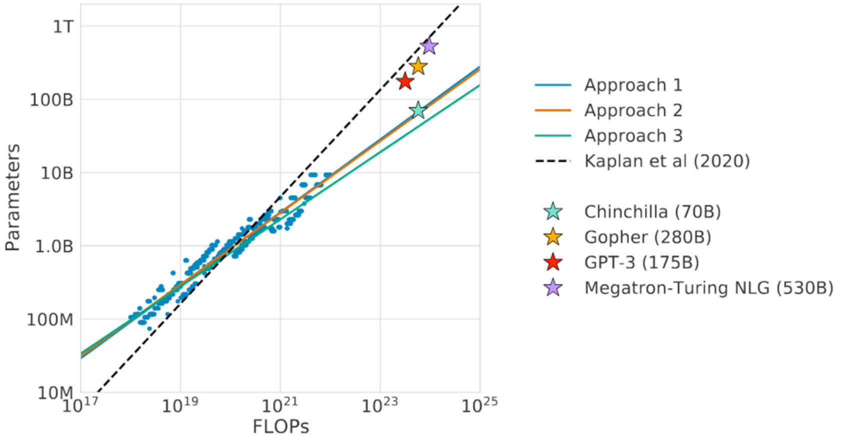
Maurice Weiler, Max Welling, Taco Cohen







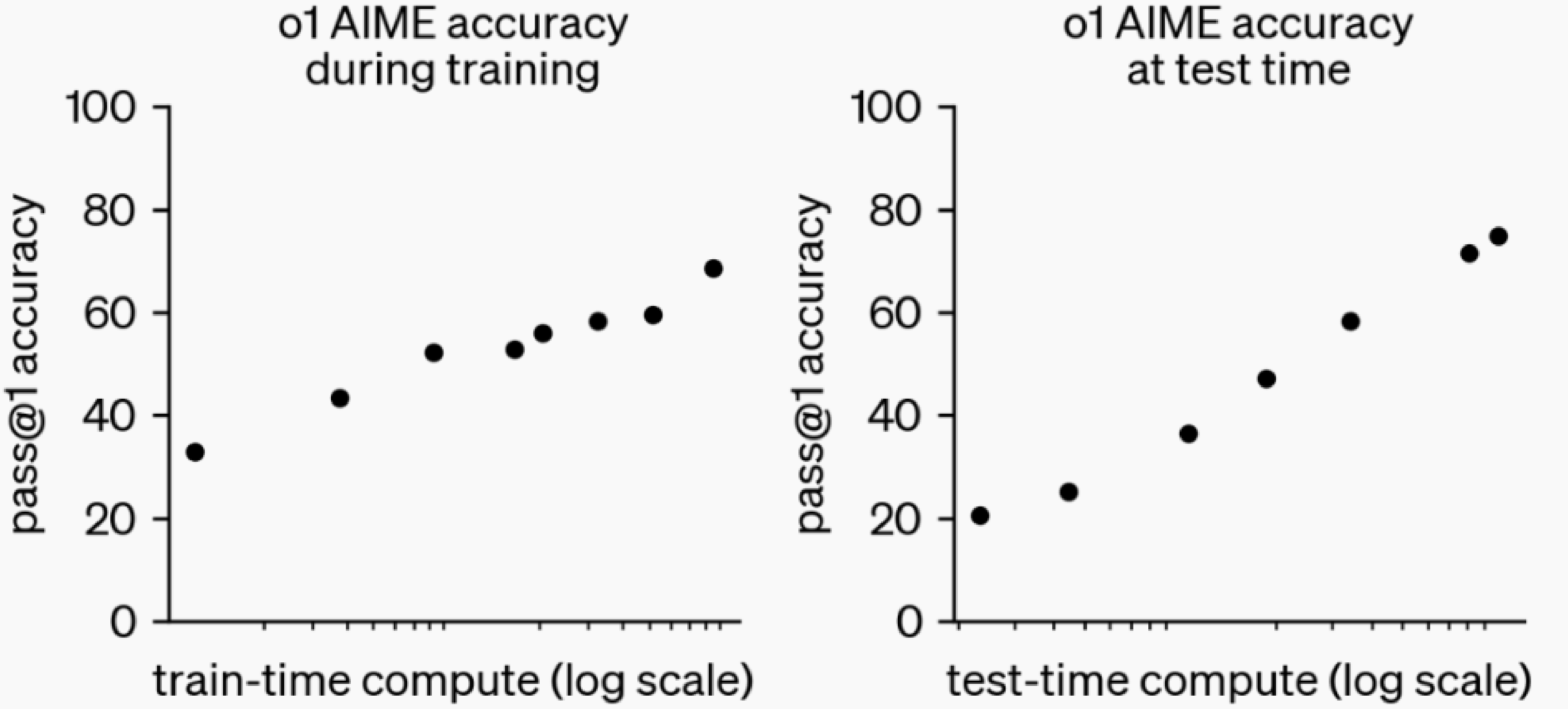




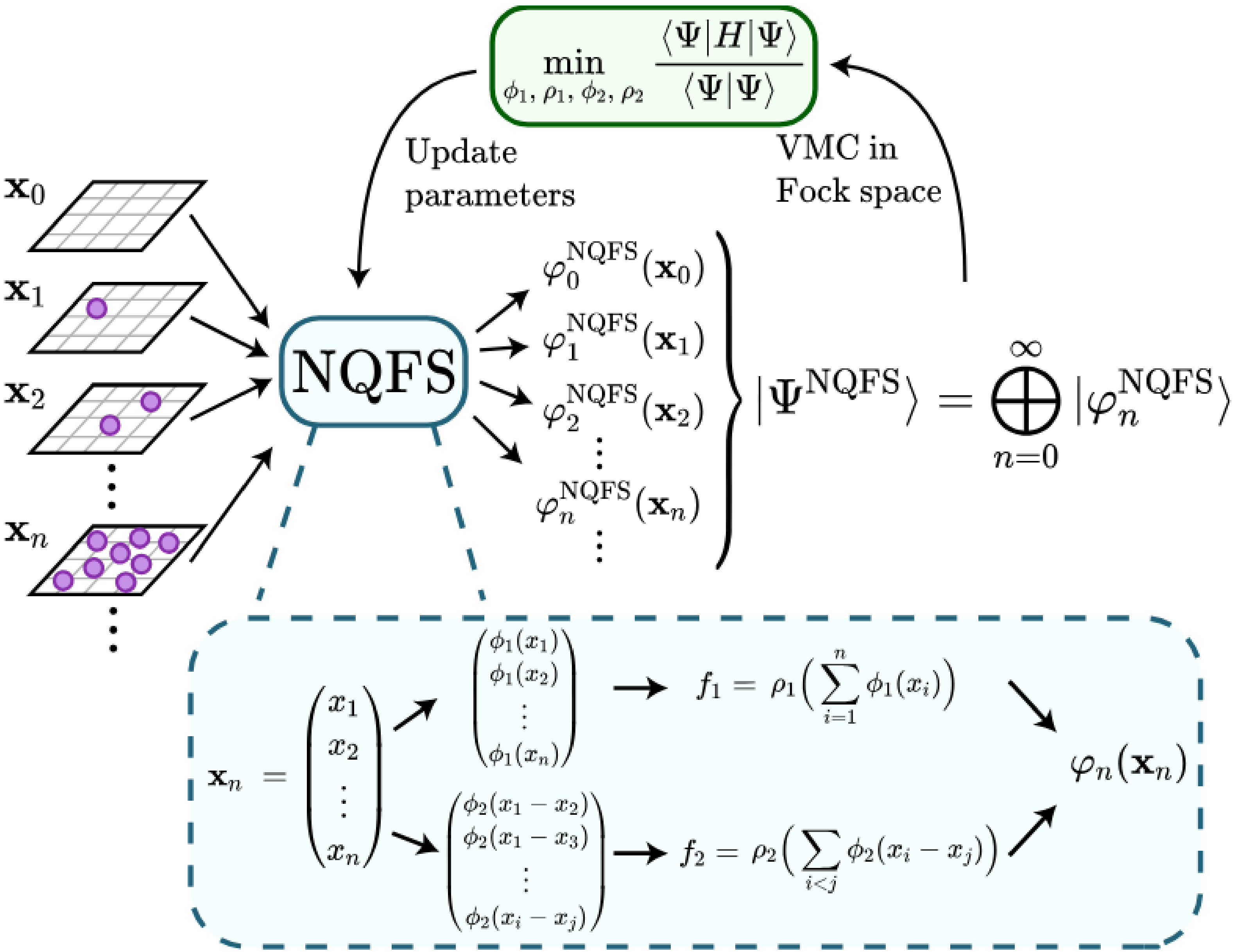
$$L(N, D) = \left[\left(\frac{N_c}{N} \right)^{\frac{\alpha_N}{\alpha_D}} + \frac{D_c}{D} \right]^{\alpha_D}$$

$$N \propto C^{\alpha_C^{\min}/\alpha_N}, \quad B \propto C^{\alpha_C^{\min}/\alpha_B}, \quad S \propto C^{\alpha_C^{\min}/\alpha_S}, \quad D = B \cdot S$$

$$\mathcal{L}(t, N) = c_t t^{-r_t} + c_N N^{-r_N} + \mathcal{L}_\infty,$$



o1 performance smoothly improves with both train-time and test-time compute



-FT

nman rules from Section (3.2) in a few single layer NN archi-
vidth and i.i.d. parameters, and evaluate the leading order in
N-FT action. The quartic coupling is

$$\cdots d^d y_4\, G_c^{(4)}(y_1,\cdots,y_4) G_c^{(2)}(y_1,x_1)^{-1}\cdots G_c^{(2)}(y_4,x_4)^{-1} + \text{perms}\Big], \tag{3.45}$$

$y_1)^{-1}$ involves differential operators, we use the methods from
4.

ture introduced earlier, $\phi(x) = W_i^1 \cos(W_{ij}^0 x_j + b_i^0)$, for $W^1 \sim$
$^2_{W_0}/d)$, and $b^0 \sim \text{Unif}[-\pi,\pi]$. We will consider the case where
ent and non-Gaussianities arise due to finite N corrections. To
uartic coupling for this NNFT, let us first compute the inverse
starting from the 2-pt function

$$G_{c,\text{Cos}}^{(2)}(x_1,x_2) = \frac{\sigma_{W_1}^2}{2} e^{-\frac{\sigma_{W_0}^2 (x_1-x_2)^2}{2d}}, \tag{3.46}$$

$y\, G_{c,\text{Cos}}^{(2)}(x,y)^{-1} G_{c,\text{Cos}}^{(2)}(y,z) = \delta^d(x-z)$. Translation invariance
lta function constraints $G_{c,\text{Cos}}^{(2)}(x,y)^{-1}$ as a translation invariant
; a Fourier transformation of the 2-pt function and its inverse
verse Fourier transformation, we obtain

$$^{2)}_{,\text{Cos}}(x,y)^{-1} = \frac{2\sigma_{W_0}^2}{\sigma_{W_1}^2 d} e^{-\frac{\sigma_{W_0}^2 \nabla_x^2}{2d}} \delta^d(x-y), \tag{3.47}$$

2.1 First fluctuation-dissipation relation

Applying the master equation (FDT) to the linear observable,

$$\langle \boldsymbol{\theta} \rangle = \left\langle \left[\!\left[\boldsymbol{\theta} - \eta \boldsymbol{\nabla} f^{\mathcal{B}} (\boldsymbol{\theta}) \right]\!\right]_{\text{m.b.}} \right\rangle = \langle \boldsymbol{\theta} \rangle - \eta \left\langle \boldsymbol{\nabla} f (\boldsymbol{\theta}) \right\rangle \,. \tag{7}$$

We thus have

$$\langle \boldsymbol{\nabla} f \rangle = 0 \,. \tag{8}$$

This is natural because there is no particular direction that the gradient picks on average as the model parameter stochastically bounces around the local minimum or, more generally, wanders around the loss-function landscape according to the stationary distribution.

Performing similar algebra for the quadratic observable $\langle \theta_i \theta_j \rangle$ yields

$$\langle \theta_i \left(\partial_j f \right) \rangle + \langle \left(\partial_i f \right) \theta_j \rangle = \eta \left\langle \widetilde{C}_{i,j} \right\rangle \,. \tag{9}$$

In particular, taking the trace of this matrix-form relation, we obtain

$$\langle \boldsymbol{\theta} \cdot \left(\boldsymbol{\nabla} f \right) \rangle = \frac{1}{2} \eta \left\langle \operatorname{Tr} \widetilde{\boldsymbol{C}} \right\rangle \,. \tag{FDR1}$$

Use of Physics in AI

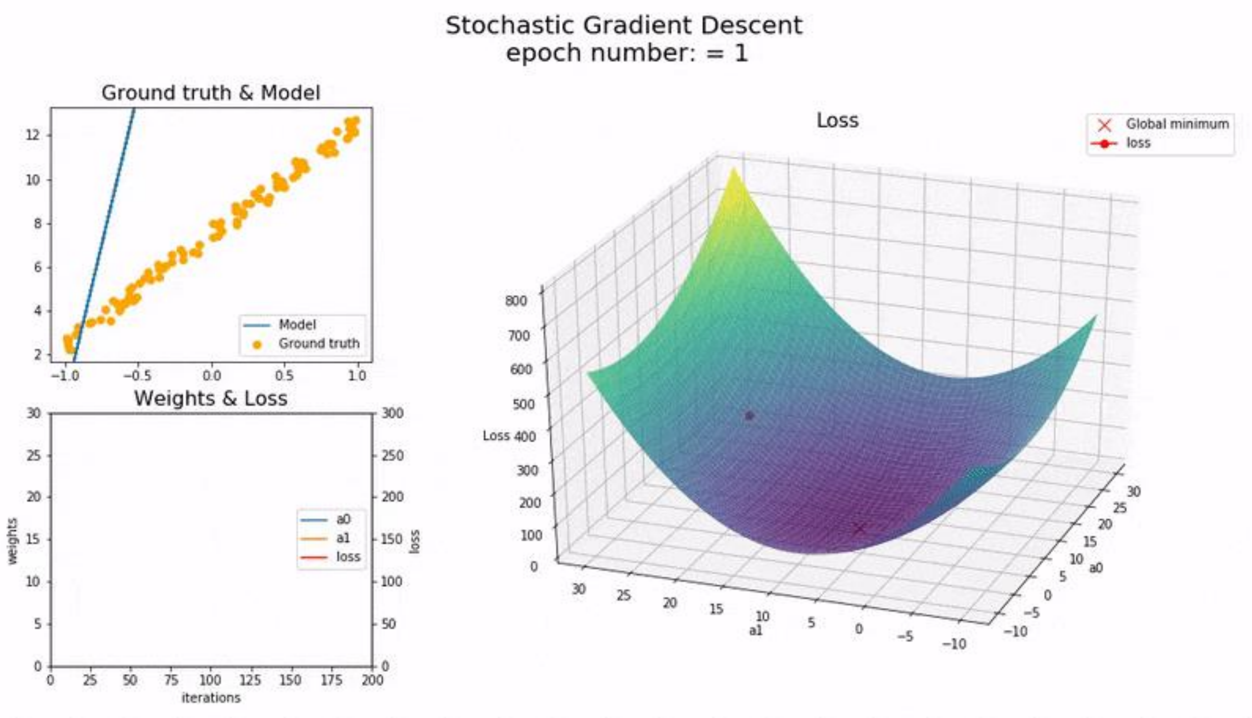
Historically

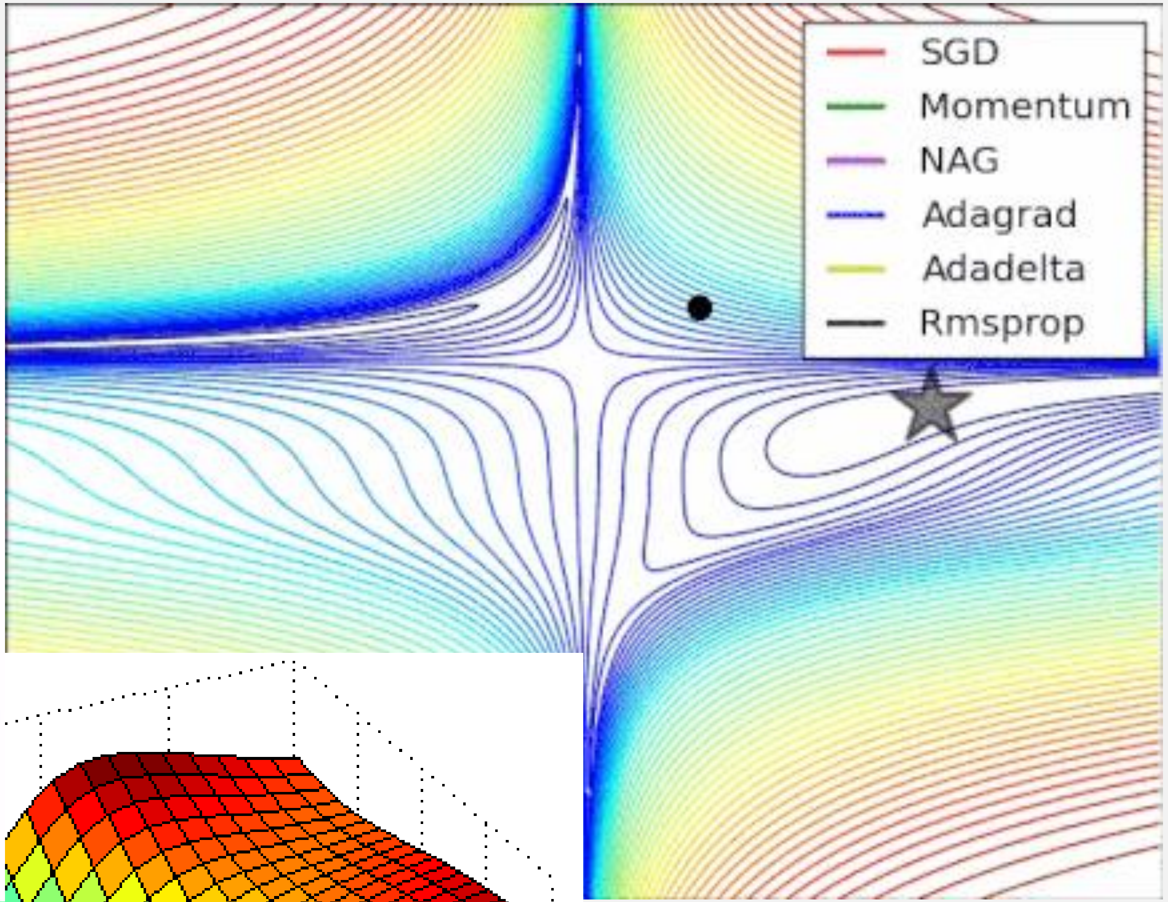
- Hopfield Networks
- Boltzmann Machines

More recently:

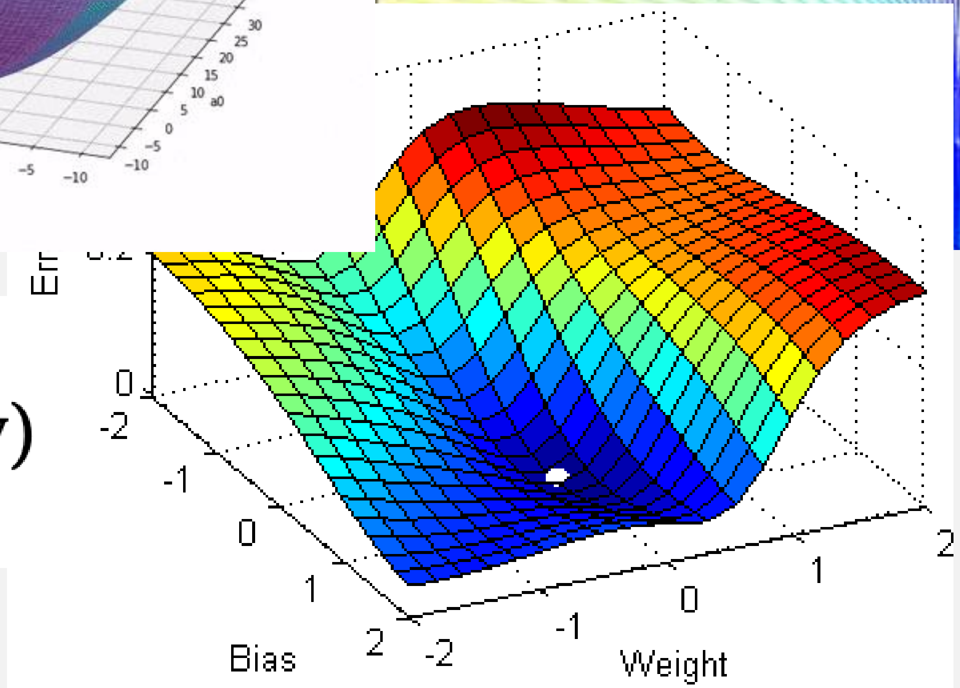
- Mean field approaches (and beyond)
- “Glassy” Phases

Physics of SGD





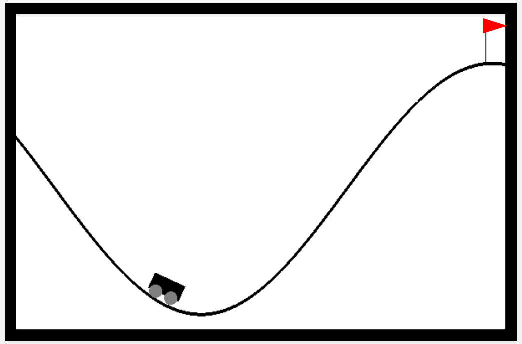
$$m\frac{\mathrm{d}^2\mathbf{w}}{\mathrm{d}t^2}+\mu\frac{\mathrm{d}\mathbf{w}}{\mathrm{d}t} = -\nabla_{\mathrm{w}} E(\mathbf{w})$$



Newton's 2nd Law!!

Reinforcement Learning

Reinforcement Learning



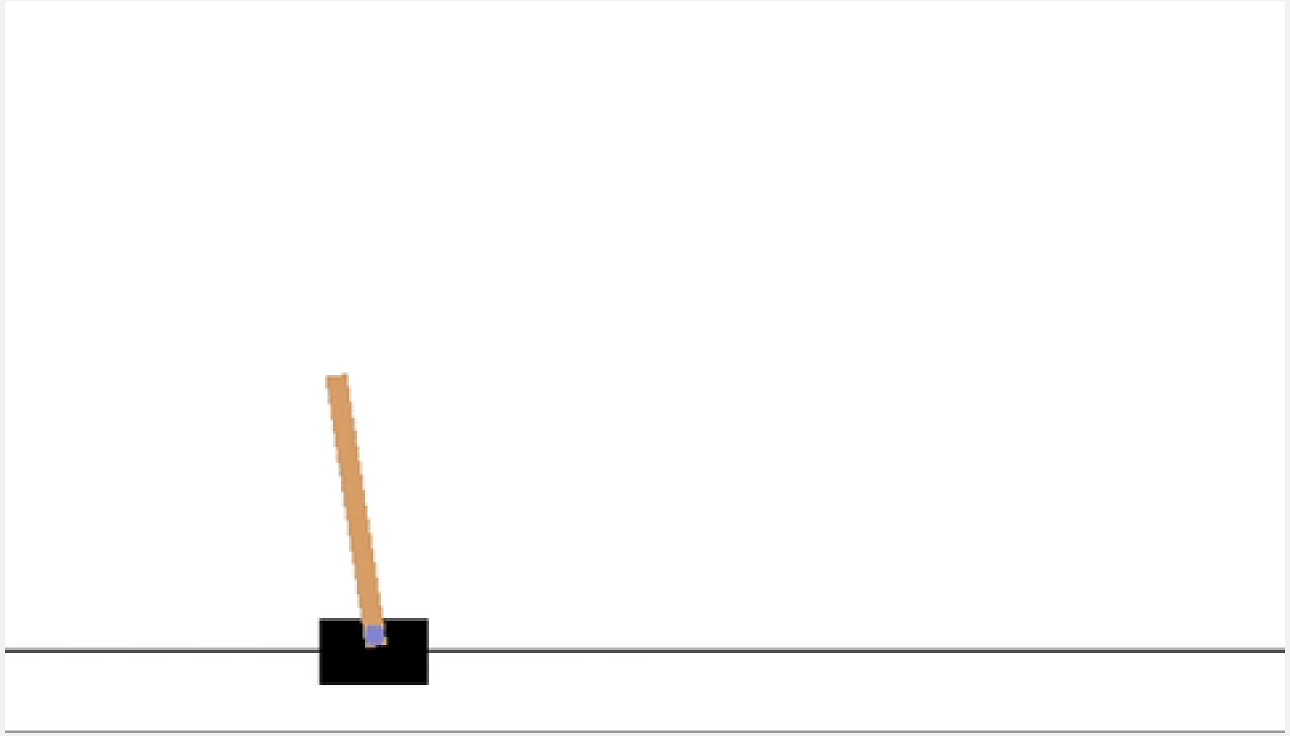
Reinforcement Learning (RL) is a paradigm created to solve sequential decision-making problems

Basic Ideas:

- An agent interacts with an **environment**
- Positive behaviors are reinforced relative to negative behaviors
  - Reinforcement is implemented via a **reward function**
- After many interaction-reinforcement cycles, the agent should learn to “successfully” interact with the environment

CartPole-v1

Easiest RL environment with continuous state space



CartPole-v1

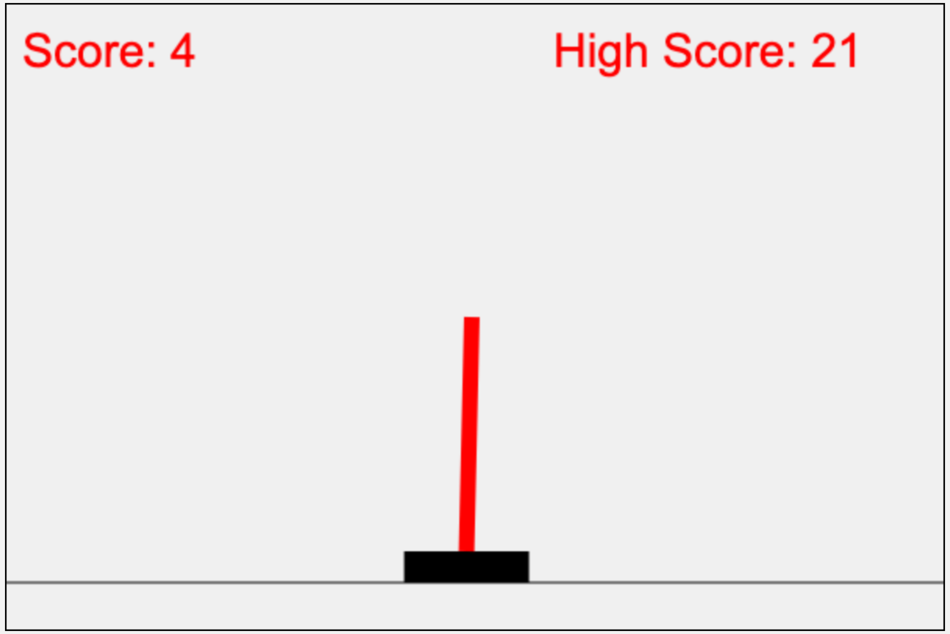
Easiest RL environment with continuous state space

Easy mode
(human-friendly)

Hard mode
(RL-“friendly”)

CartPole-v1

Easiest RL environment with continuous state space

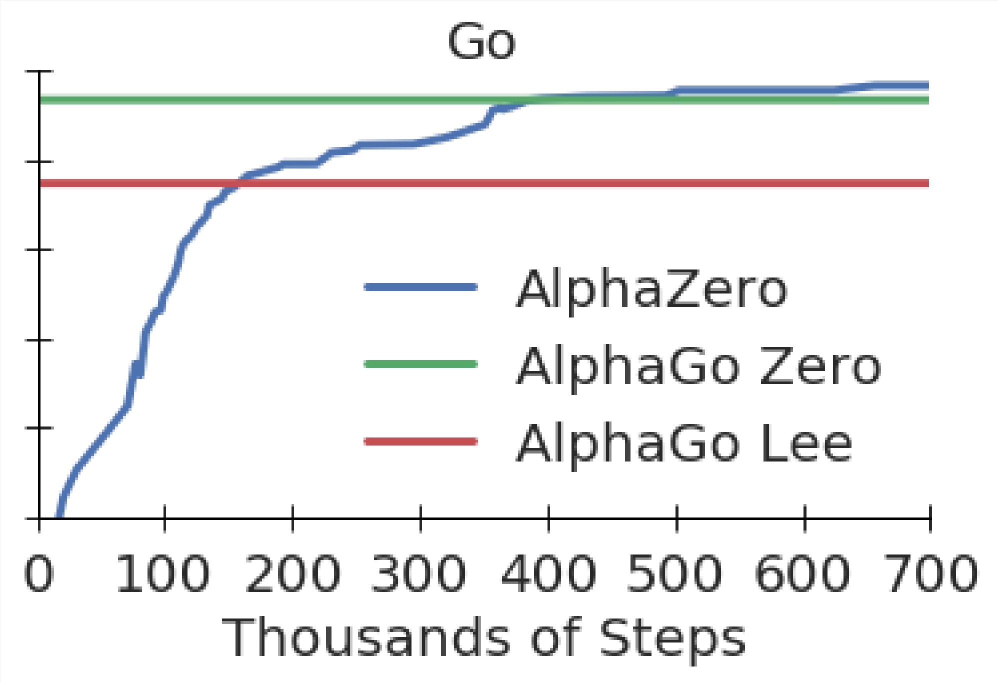


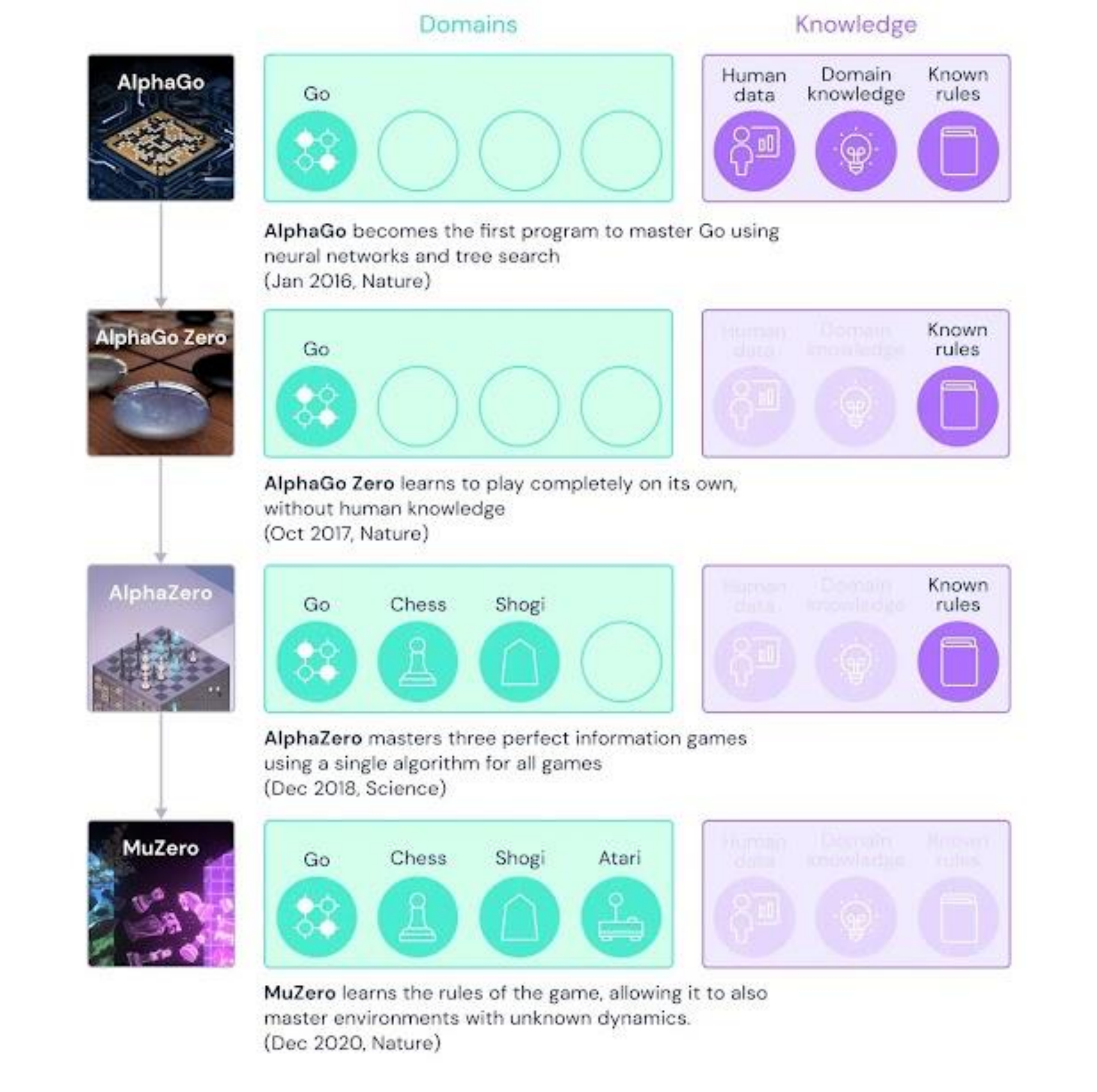
Easy mode
(human-friendly)

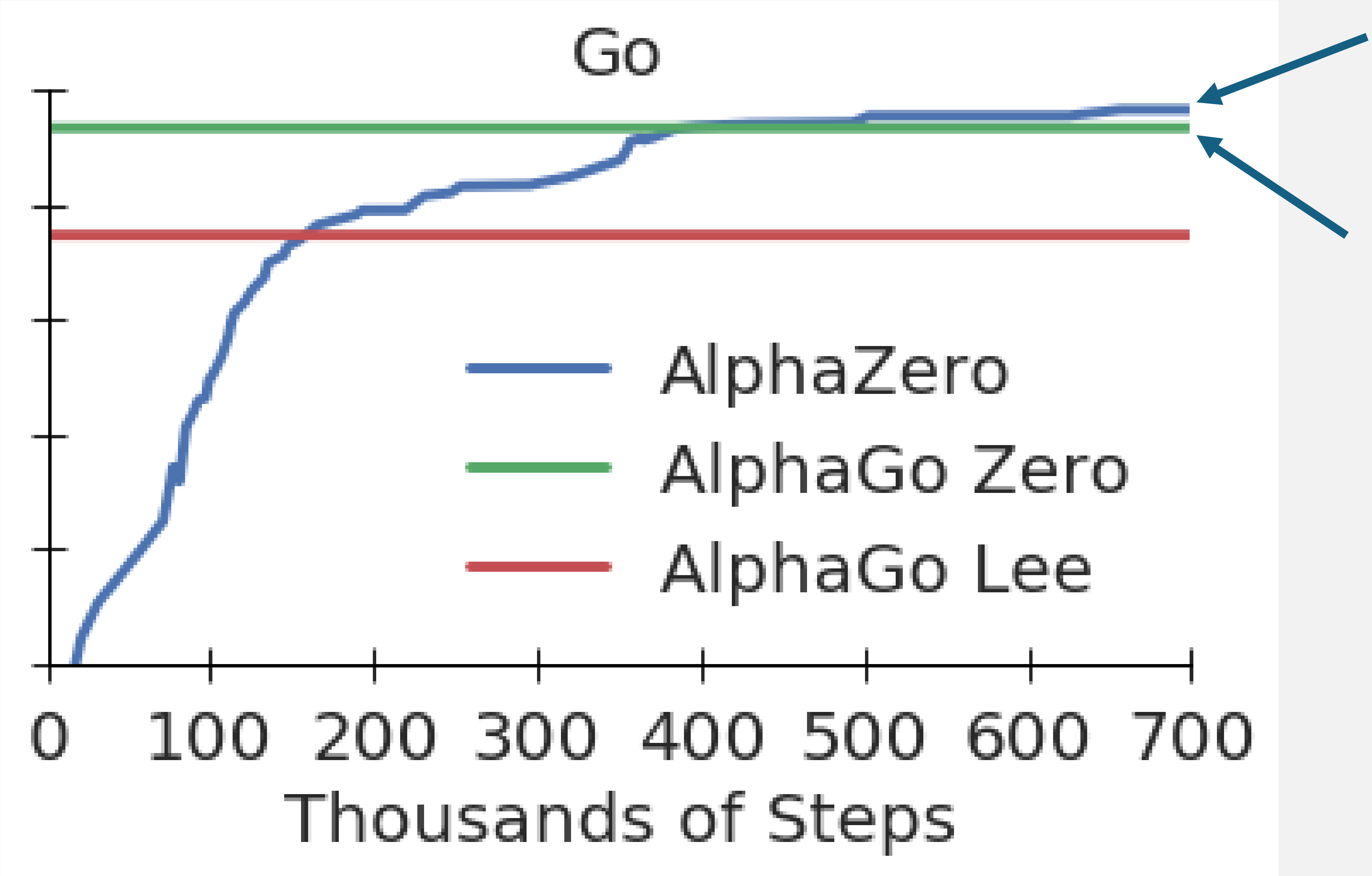
Score: 0 High Score: 10
Cart position: 259.91
Cart velocity: -0.78
Pole angle: -0.02
Pole angular velocity: 0.06

Hard mode
(RL-“friendly”)

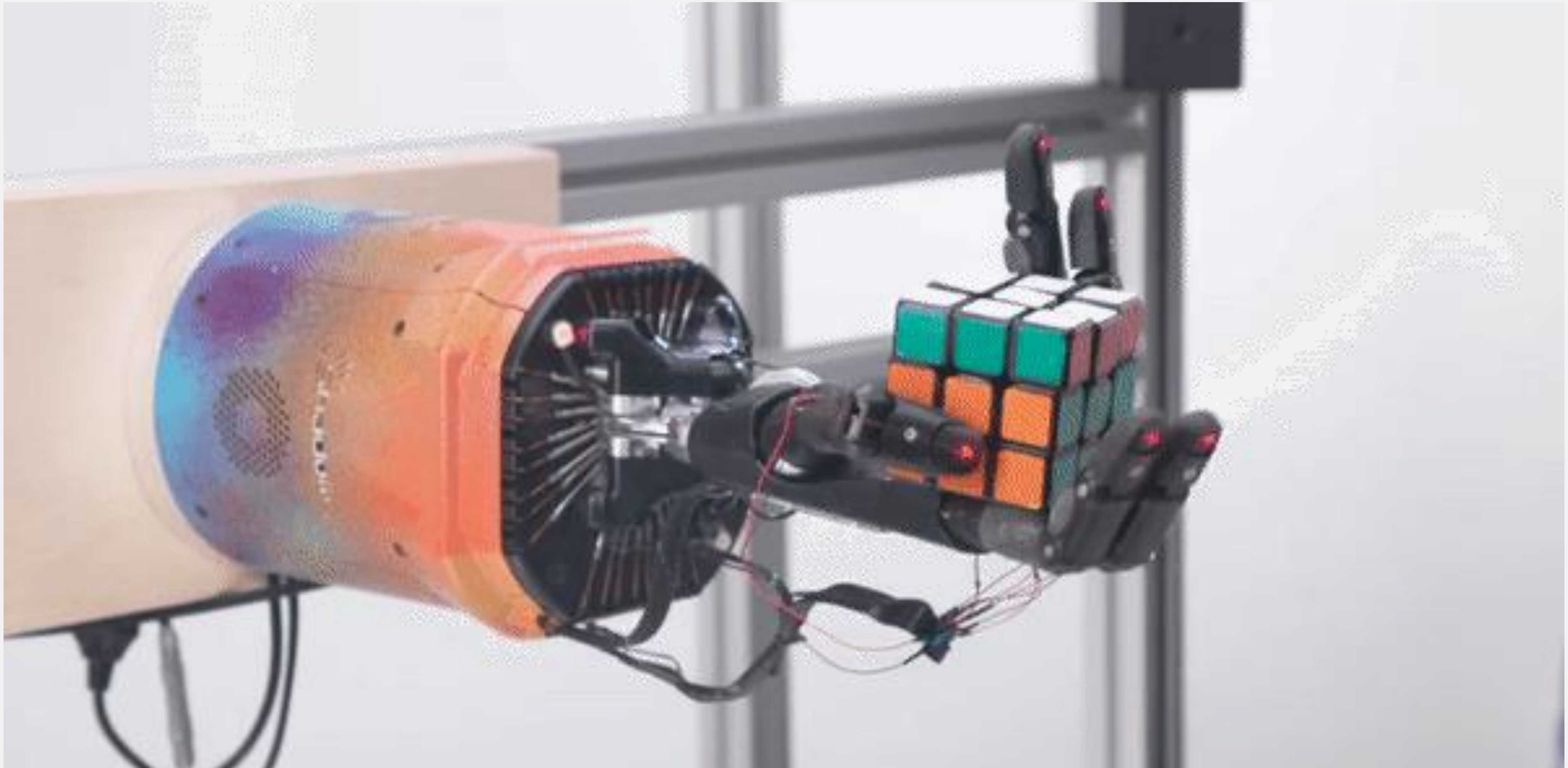
Breakthroughs in RL

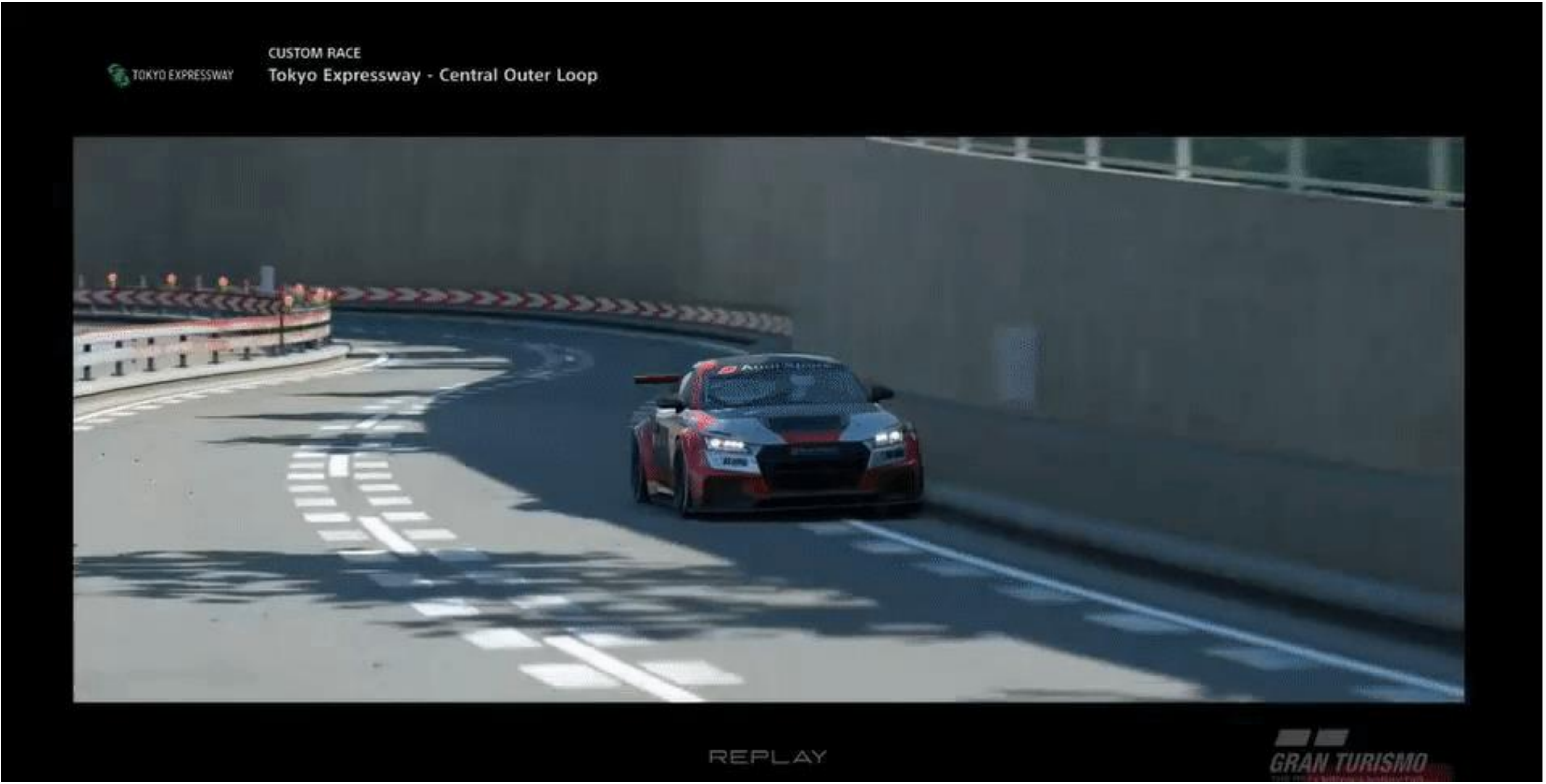




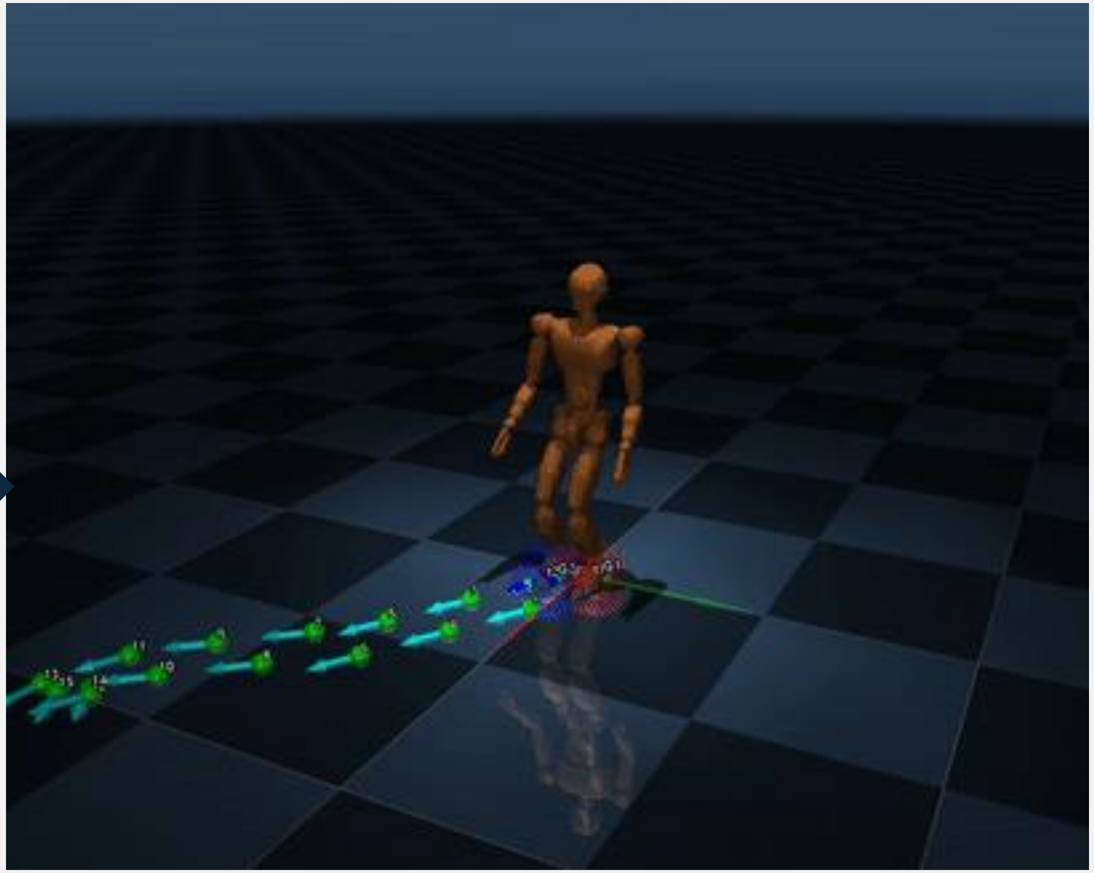

Go
AlphaZero
AlphaGo Zero
AlphaGo Lee
0
100
200
300
400
500
600
700
Thousands of Steps

Breakthroughs in RL




TOKYO EXPRESSWAY
CUSTOM RACE
Tokyo Expressway - Central Outer Loop
REPLAY
GRAN TURISMO

How?



Burgeoning Field with Bountiful Bridge

My Research

Q: What is the core object in stat. mech.?

A: Partition Function

So what?, $\mathcal{Z}(\beta)$

- Ubiquitous in any sampling problem
- Derivatives give CGF, Sensitivity, Phase Transitions
- Free energy (solution to optimization problem)
- Bogoliubov inequality
- Donsker-Varadhan
- Duality to entropy
- Linear algebra connections

Now what?

The partition function (normalization const.) counts things.

<u>Counting things is hard.</u>

Techniques have been developed in physics (and CS) to count things more easily:

<u>Technique 1:</u>

Re-weight via Boltzmann factor / “importance sampling”

and count everything!

(constrained → unconstrained!)

Stat mech of RL

- We have shown (via transfer matrix + prob. inf.) the optimal value function $Q(s,a)$ for undiscounted case ($\gamma = 1$) can be interpreted as a conditional free energy
- The SCGF θ is the "bulk" free energy

$$\beta Q(s,a) = -N\beta\theta + \log u(s,a) + O(\dots)$$

Where $u(s,a)$ is the Perron root's ($\rho = e^{-\beta\theta}$) corresponding left eigenvector.

$$\pi^*(a|s) \propto u(s,a)$$

Entropy regularized RL using large deviation theory. A. Arriojas, **JA**, S. Tiomkin, R. V. Kulkarni *(PRR, 2023)*

Stat mech of RL

- We have shown (via transfer matrix + prob. inf.) the optimal value function $Q(s,a)$ for undiscounted case ($\gamma = 1$) can be interpreted as a conditional free energy
- The SCGF θ is the "bulk" free energy

$$\beta Q(s,a) = -N\beta\theta + \log u(s,a) + O(\ldots)$$

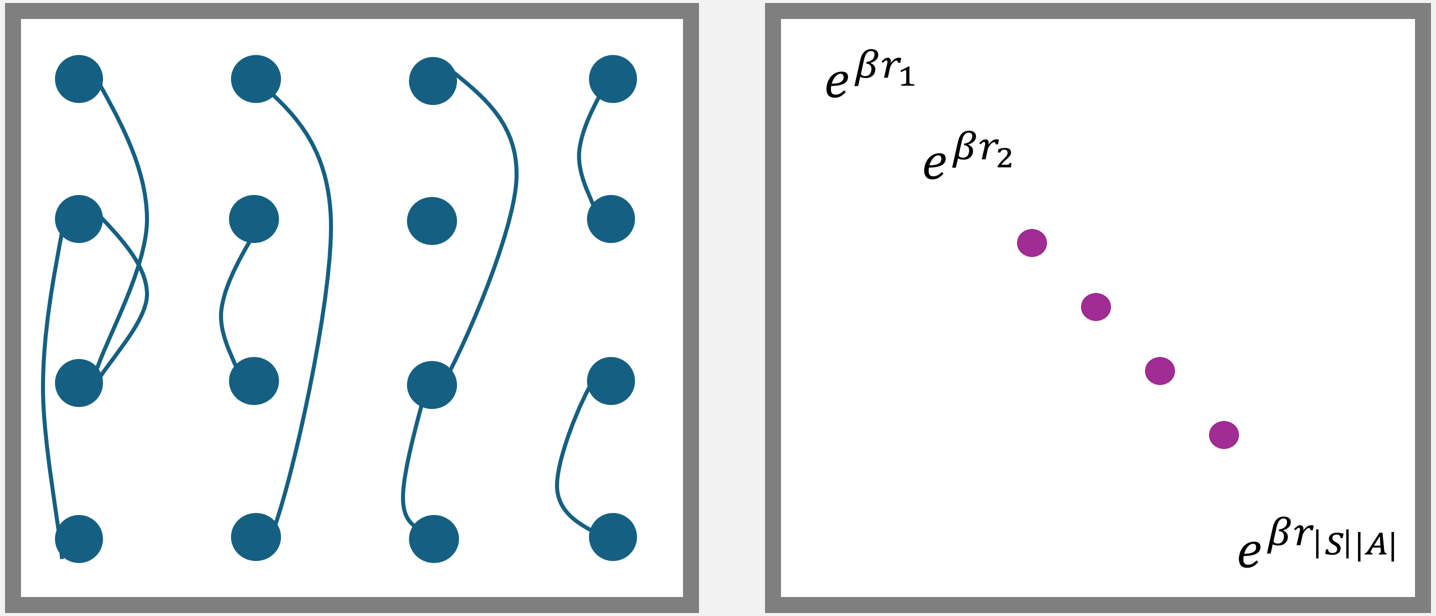
Where $u(s,a)$ is the Perron root's ($\rho = e^{-\beta\theta}$) corresponding left eigenvector.

$$\tilde{\mathrm{P}}(s',a'|s,a) = p(s'|s,a)\pi_0(a's')e^{\beta r(s,a)}$$

Entropy regularized RL using large deviation theory. A. Arriojas, **JA**, S. Tiomkin, R. V. Kulkarni *(PRR, 2023)*

Stat mech of RL

$\tilde{P} =$



This matrix can be used to generate the desired trajectories!

Entropy regularized RL using large deviation theory. A. Arriojas, **JA**, S. Tiomkin, R. V. Kulkarni *(PRR, 2023)*

RL framework using large deviations

- Analytical solution for RL problem using large deviation theory
- Average Reward → Perron-Frobenius eigenvalue of tilted matrix
- Optimal Policy → Perron-Frobenius eigenvector of tilted matrix

PHYSICAL REVIEW RESEARCH **5**, 023085 (2023)

Entropy regularized reinforcement learning using large deviation theory

Argenis Arriojas,[1,*] Jacob Adamczyk,[1] Stas Tiomkin,[2] and Rahul V. Kulkarni[1,†]
[1]*Department of Physics, University of Massachusetts Boston, Boston, Massachusetts 02125, USA*
[2]*Department of Computer Engineering, San Jose State University, San Jose, California 95192, USA*

Solution for Stochastic Dynamics

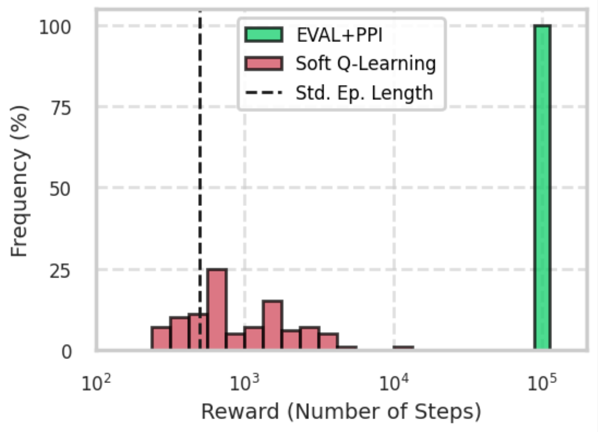
- Solution for stochastic dynamics is challenging because of constraint on system dynamics (fixed).
- Constrained problem can be solved by mapping to a distinct *unconstrained* problem with the same optimal policy

Bayesian Inference Approach for Entropy Regularized Reinforcement Learning with Stochastic Dynamics

Argenis Arriojas[1] **Jacob Adamczyk**[1] **Stas Tiomkin**[2] **Rahul V Kulkarni**[1]

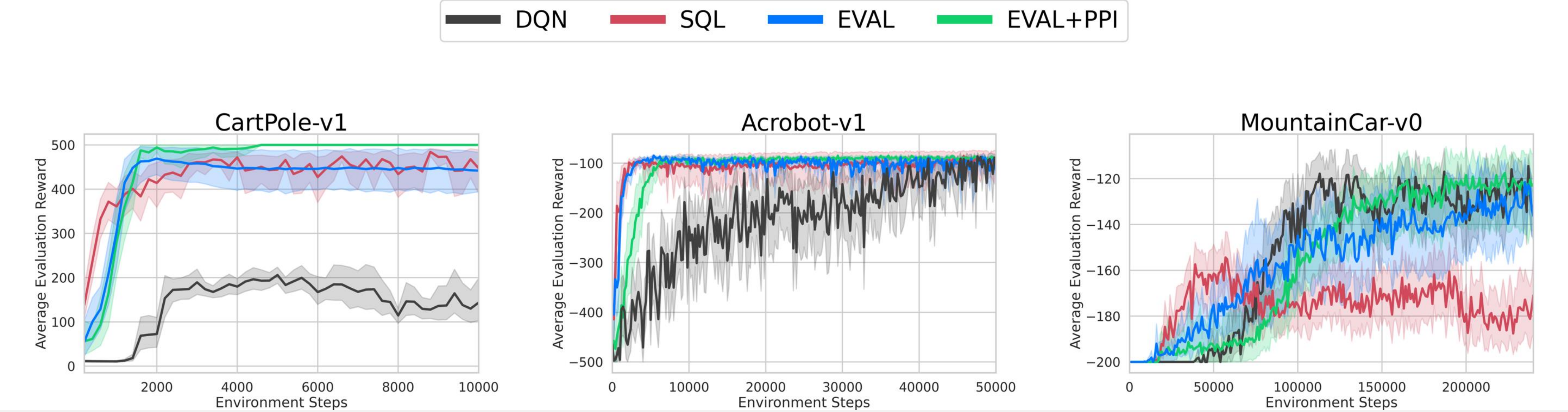
[1]Department of Physics, University of Massachusetts Boston, Boston, Massachusetts, USA
[2]Department of Computer Engineering, San Jose State University, San Jose, California, USA

Eigenvector Learning



- Novel algorithms with promising results

"EVAL: EigenVector-based Average-reward Learning" (under review)



Reward shaping and compositionality

- Motivated by Jarzynski relation → Set up focusing on Free Energy differences
- Reward shaping for entropy-regularized RL, applications for compositionality in RL

The Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23)

Utilizing Prior Solutions for Reward Shaping and Composition in Entropy-Regularized Reinforcement Learning

Jacob Adamczyk[1], Argenis Arriojas[1], Stas Tiomkin[2], Rahul V. Kulkarni[1]

[1]Department of Physics, University of Massachusetts Boston
[2]Department of Computer Engineering, San José State University
jacob.adamczyk001@umb.edu, arriojasmaldonado001@umb.edu, stas.tiomkin@sjsu.edu, rahul.kulkarni@umb.edu

Relating two free energies by a third

We show that for energies related by $\tilde{E} = E + \Delta E$,

$$\tilde{F} = F + F_\Delta$$

Where $F_\Delta = \sum_\sigma p(\sigma) e^{-\beta \Delta E(\sigma)}$

- The free energy for a system with energy ΔE and prior distribution $p(\sigma)$ (the configurational distribution for the system with energy $E(\sigma)$

Moreover, F_Δ and $\tilde{F}$ share the same eq. distribution:

$$p_\Delta(\sigma) = \tilde{p}(\sigma) \propto e^{-\beta \tilde{E}(\sigma)}$$

Relating two free energies by a third

We show that for energies related by $\tilde{E} = E + \Delta E$,

$$\tilde{F} = F + F_\Delta$$

Where $F_\Delta = \sum_\sigma p(\sigma) e^{-\beta \Delta E(\sigma)}$

- The free energy for a system with energy ΔE and prior distribution $p(\sigma)$ (the configurational distribution for the system with energy $E(\sigma)$

Moreover, $F_\Delta$and $\tilde{F}$ share the same eq. distribution:

$$p_\Delta(\sigma) = \tilde{p}(\sigma) \propto e^{-\beta \tilde{E}(\sigma)}$$

Simple Proof

$$\tilde{Z} = \sum_{\sigma} e^{-\beta \tilde{E}(\sigma)} = Z \sum_{\sigma} \left(\frac{1}{\sum_{\mu} e^{-\beta E(\mu)}} \right) e^{-\beta (E(\sigma) + \Delta E(\sigma))}$$

$$\tilde{Z} = Z \sum_{\sigma} \left(\frac{e^{-\beta E(\sigma)}}{\sum_{\mu} e^{-\beta E(\mu)}} \right) e^{-\beta \Delta E(\sigma)} = Z \sum_{\sigma} p(\sigma)\, e^{-\beta \Delta E(\sigma)}$$

$$\tilde{Z} = Z \cdot Z_{\Delta}$$

$$\tilde{F} = F + F_{\Delta}$$

Simple Proof

$$\tilde{Z} = \sum_{\sigma} e^{-\beta \tilde{E}(\sigma)} = Z \sum_{\sigma} \left(\frac{1}{\sum_{\mu} e^{-\beta E(\mu)}} \right) e^{-\beta (E(\sigma) + \Delta E(\sigma))}$$

$$\tilde{Z} = Z \sum_{\sigma} \left(\frac{e^{-\beta E(\sigma)}}{\sum_{\mu} e^{-\beta E(\mu)}} \right) e^{-\beta \Delta E(\sigma)} = Z \sum_{\sigma} p(\sigma)\, e^{-\beta \Delta E(\sigma)}$$

$$\tilde{Z} = Z \cdot Z_{\Delta}$$

$$\tilde{F} = F + F_{\Delta}$$

Gibbs-Bogoliubov Inequality

$\tilde{F} = F + F_\Delta$

- Considering the variational form for F_Δ we use the prior as the variational guess:

$$F_\Delta = \inf_q \left[\langle \Delta E \rangle_q + \beta^{-1} KL(q|p)\right]$$

$$F_\Delta \leq \langle \Delta E \rangle_p$$

- Combined with the previous result, we arrive at

$$\tilde{F} \leq F + \langle \Delta E \rangle_{p(\sigma)}$$

Gibbs-Bogoliubov Inequality

Q functions (conditional free energy)

- Same result holds, even while considering trajectories conditioned on initial (*state*, *action*) pairs and *discounting* over trajectories:

$$\tilde{Q}(s,a) = Q(s,a) + K(s,a)$$

Where K has an analogous definition to F_Δ:

- as reward, it takes $\tilde{r}(s,a) - r(s,a)$
- as a prior distribution, K is wrt the former's optimal policy:

$$\pi_0^{(K)} \doteq \pi^*$$

Q functions (conditional free energy)

- Same result holds, even while considering trajectories conditioned on initial (*state, action*) pairs and *discounting* over trajectories:

$$\tilde{Q}(s,a) = Q(s,a) + K(s,a)$$

Where K has an analogous definition to F_Δ:

- as reward, it takes $\tilde{r}(s,a) - r(s,a)$
- as a prior distribution, K is wrt the former's optimal policy: $\pi_0^{(K)} \doteq \pi^*$

$$\tilde{Q}(s,a) \geq Q(s,a) + \mathbb{E}_{\tau|(s,a)\sim\pi^*}(\tilde{r} - r)$$

Q functions (conditional free energy)

- Same result holds, even while considering trajectories conditioned on initial (*state, action*) pairs and *discounting* over trajectories:

$$\tilde{Q}(s,a) = Q(s,a) + K(s,a)$$

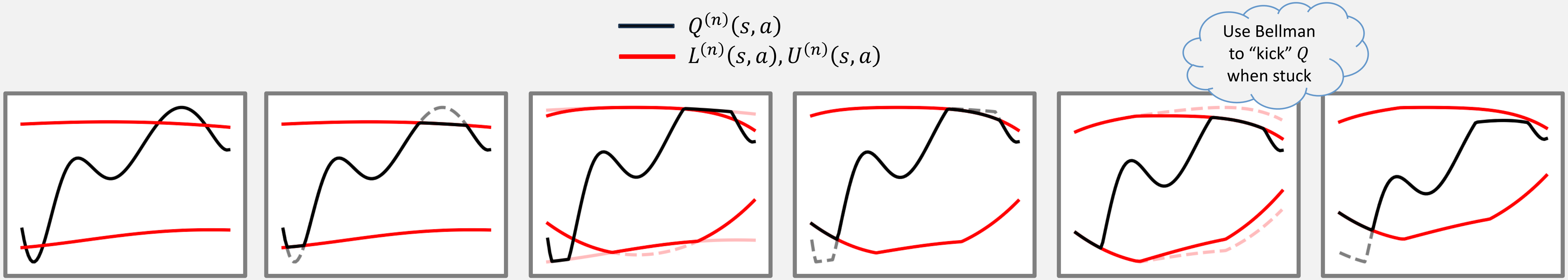
K and $\tilde{Q}$ have same optimal policy: $\pi_K^* = \tilde{\pi}^*$

Where K has an analogous definition to F_Δ:

- as reward, it takes $\tilde{r}(s,a) - r(s,a)$
- as a prior distribution, K is wrt the former's optimal policy: $\pi_0^{(K)} \doteq \pi^*$

$$\tilde{Q}(s,a) \geq Q(s,a) + \mathbb{E}_{\tau|(s,a)\sim\pi^*}(\tilde{r} - r)$$

Learning via clipping based on bounds



RLJ | RLC 2024

Boosting Soft Q-Learning by Bounding

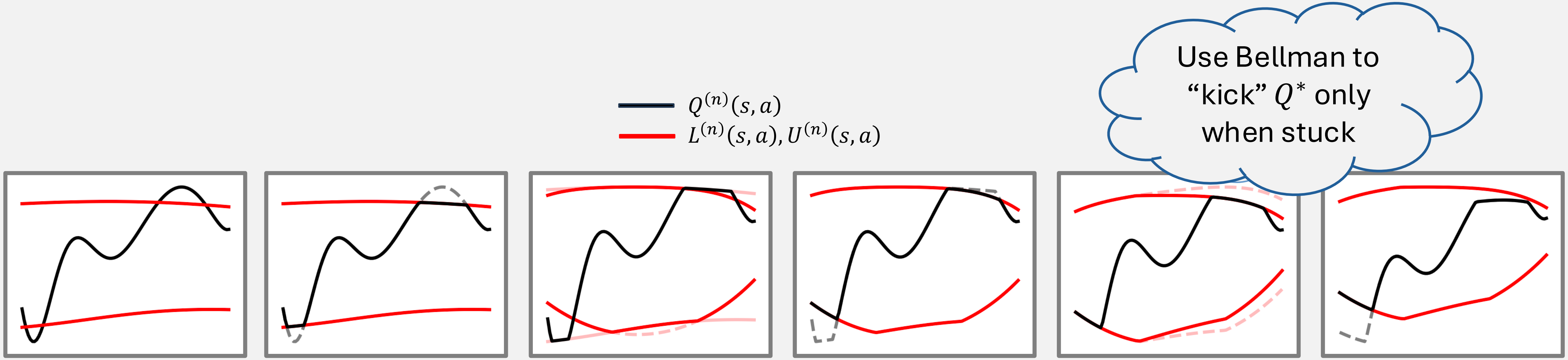
Jacob Adamczyk
jacob.adamczyk001@umb.edu
Department of Physics
University of Massachusetts Boston

Volodymyr Makarenko
volodymyr.makarenko@sjsu.edu
Department of Computer Engineering
San José State University

Stas Tiomkin
stas.tiomkin@sjsu.edu
Department of Computer Engineering
San José State University

Rahul V. Kulkarni
rahul.kulkarni@umb.edu
Department of Physics
University of Massachusetts Boston

Learning via clipping based on bounds



Clipping excludes invalid Q^*,
whereas Bellman pulls you toward Q^*

The Future

Future Plan for RL

1. Establish a general framework / dictionary that maps between deep RL and NESM research
2. Exploit positive feedback loop
3. Profit

RL for stat mech (opp. direction)

- Learn free energy
- Improvements over SA
- Learn the large deviation rate function

Recent Work

- All results have relied on left eigenvector
  - Right eigenvector contains info about a “backward”/dual problem
- Can be learned simultaneously
- Forward-backward leads to detailed balance results

Career Trajectory

Sony AI

Thank You