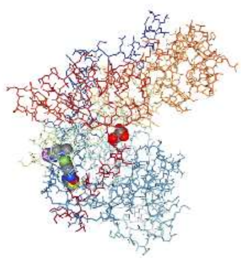
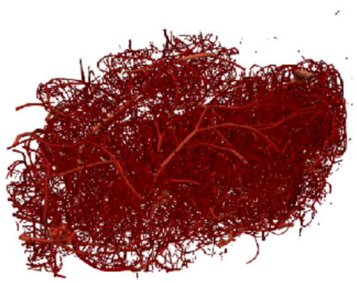
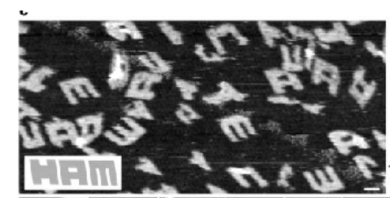
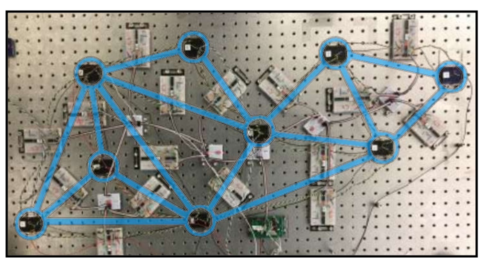
Many More is Different: A New Class of Matter

Andrea J. Liu
University of Pennsylvania









More is Different vs. Many More is Different

- "More is Different" from a few Anderson, Science (1972)
  - But usually many more is not much different from more in most condensed matter systems
- Systems with many more different from more
  - Brains: C. elegans (302 neurons) vs. honeybees ($\sim 10^6$ neurons) vs humans ($\sim 10^{11}$ neurons)
  - Digital neural networks: ChatGPT4 ($\sim 10^{14}$ parameters)

More is Different vs. Many More is Different

- "More is Different" from a few Anderson, Science (1972)
  - But usually many more is not much different from more in most condensed matter systems
- Systems with many more different from more
  - Digital neural networks: ChatGPT4 ($\sim 10^{12}$ parameters)
  - Brains: C. elegans (302 neurons) vs. honeybees ($\sim 10^{6}$ neurons) vs humans ($\sim 10^{11}$ neurons)

ADAPTIVE MATTER

More is Different vs. Many More is Different

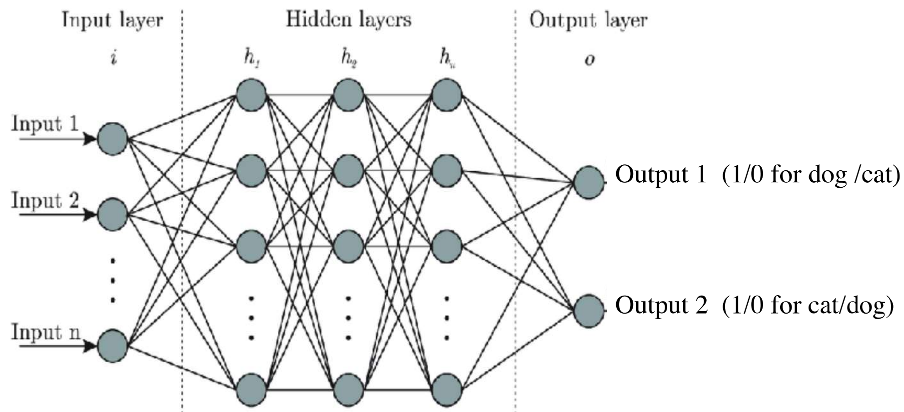
- "More is Different" from a few Anderson, Science (1972)
  - But usually many more is not much different from more in most condensed matter systems
- Systems with many more different from more
  - Digital neural networks: ChatGPT4 ($\sim 10^{12}$ parameters)
  - Brains: C. elegans (302 neurons) vs. honeybees ($\sim 10^{6}$ neurons) vs humans ($\sim 10^{11}$ neurons)

ADAPTIVE MATTER

new form of soft matter

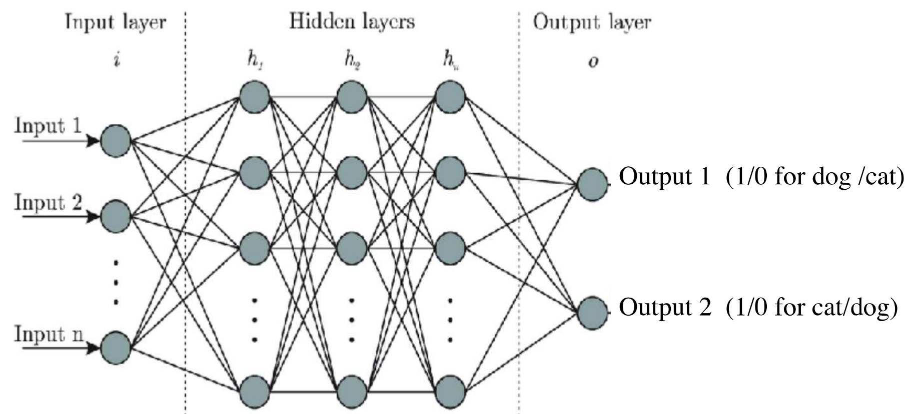
- Why is many more different in these systems?

Why is Many More Different in Digital Neural Networks?



- Network has "adaptive DOF" (parameters)
- How to adjust?
  - Evaluate cost function $\mathcal{C} = \sum_i (\mathrm{desired\ output_i} - \mathrm{free\ output_i})^2$
  - Minimize $\mathcal{C}$ by adjusting adaptive DOF

Why is Many More Different in Digital Neural Networks?

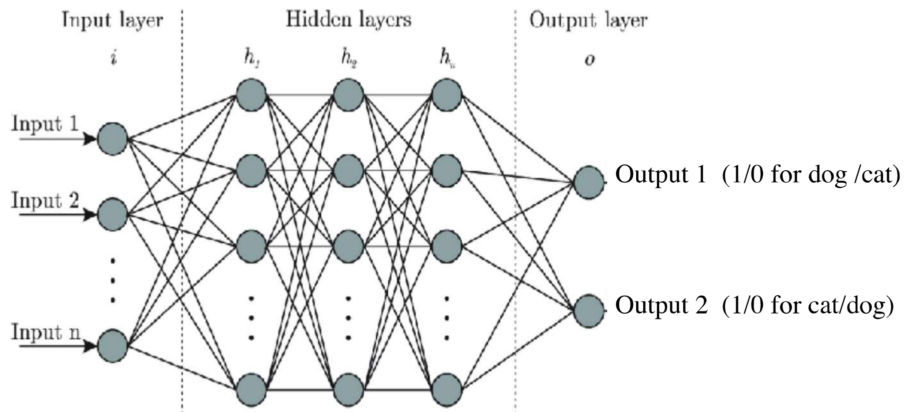


- Network has "adaptive DOF" (parameters)
- Minimize $\mathcal{C}$ by adjusting adaptive DOF

$$\mathcal{C} = \sum_i (\text{desired output}_\text{i} - \text{free output}_\text{i})^2$$

- Intrinsically disordered with correlations introduced by minimization of $\mathcal{C}$
- # constraints that can be satisfied increases with # adaptive DOF
- Many more is different bc # adaptive DOF increases with system size

Complex Functionality Arises From Correlated Disorder

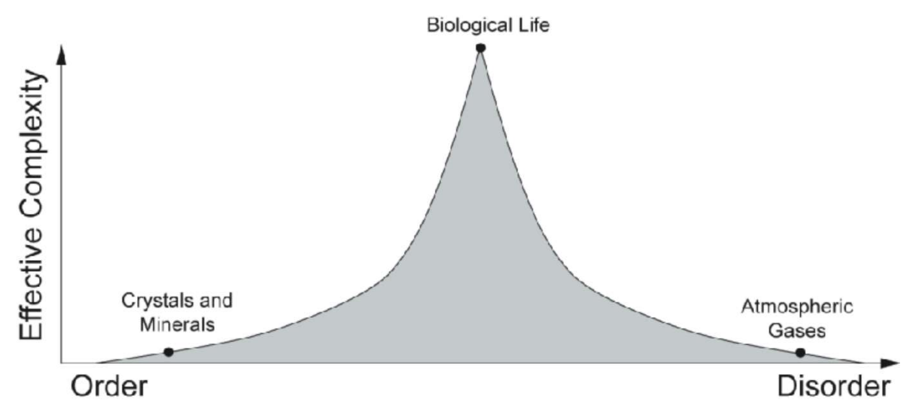


- Network parameters (adaptive degrees of freedom) are adjusted to satisfy constraints embodied in cost function

Gell-Mann/Lloyd 1996

$$\mathcal{C} = \sum_i (\mathrm{desired\ output_i} - \mathrm{free\ output_i})^2$$

- Minimize $\mathcal{C}$ by adjusting adaptive DOF
  - They are not fully ordered
  - They are not fully disordered



More is Different vs. Many More is Different

- "More is Different" from a few Anderson, Science (1972)
  - But usually many more is not much different from more in most condensed matter systems
  - No adaptive DOF!
- Systems with many more different from more
  - Brains: C. elegans (302 neurons) vs. honeybees ($\sim 10^6$ neurons) vs humans ($\sim 10^{11}$ neurons)
  - Digital neural networks: ChatGPT4 ($\sim 10^{14}$ parameters)
- Why is many more different in these systems?
  - # adaptive DOF increases with system size

- What physical systems besides brains might have this property?
- How can we design a simple physical system with this property?

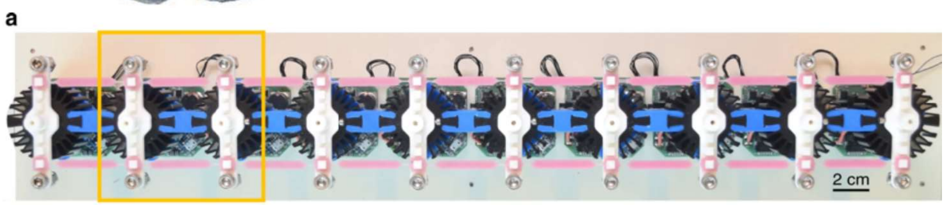
More is Different vs. Many More is Different

- “More is Different” from a few Anderson, Science (1972)
  - But usually many more is not much different from more in most condensed matter systems
  - No adaptive DOF!
- Systems with many more different from more
  - Brains: C. elegans (302 neurons) vs. honeybees (~10^6 neurons) vs humans (~10^{11} neurons)
  - Digital neural networks: ChatGPT4 (~10^{14} parameters)
- Why is many more different in these systems?
  - Because # adaptive DOF increases with system size

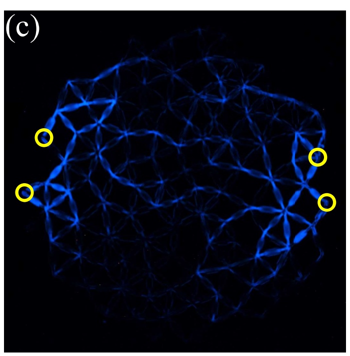
- What physical systems besides brains might have this property? Systems with many adaptive DOF (tunable interactions)

Systems with Adaptive DOF Tuned by External Brains/Processors

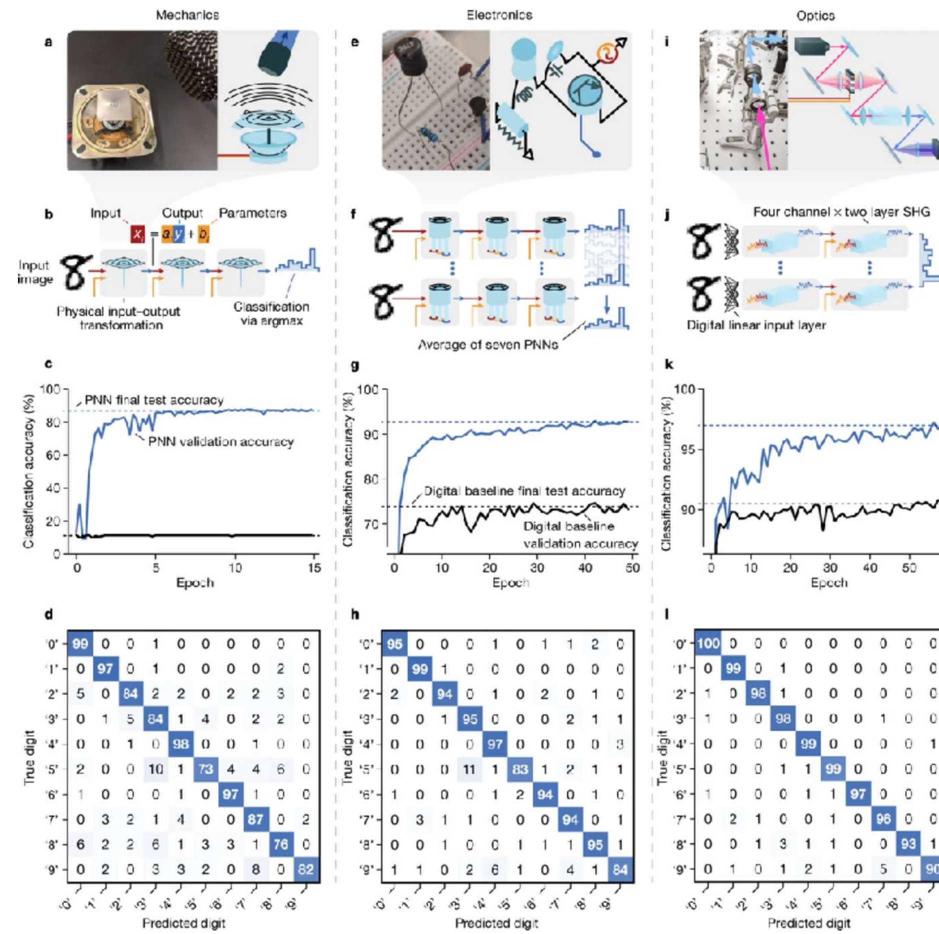
Rocks, Pashine,…Liu, Nagel
PNAS (2017)



Brandenburger, ..Coulais Nature Comm (2019)

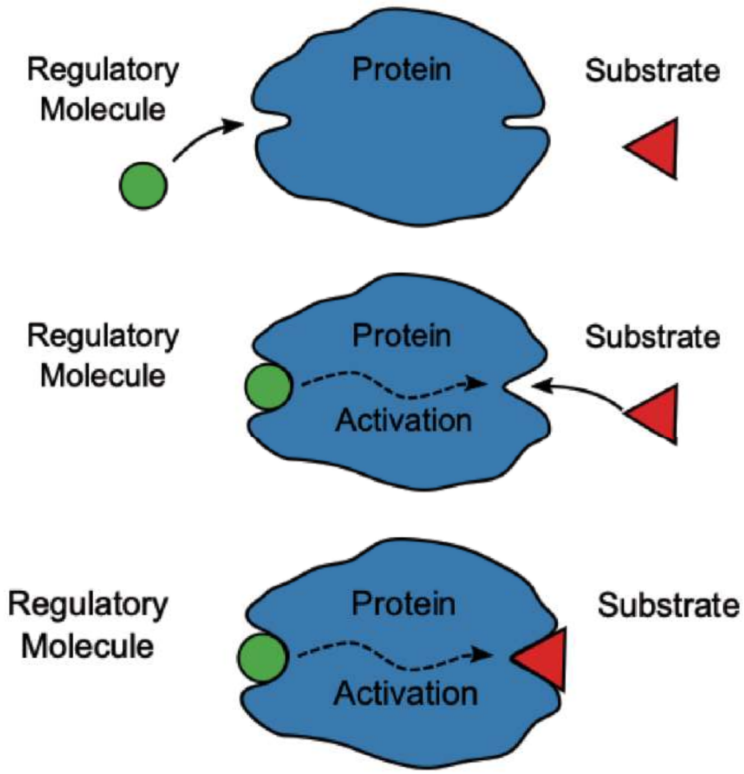


Pashine PRM (2021)



Wright, ..McMahon Nature (2022)

Living Systems With Adaptive DOF Tuned by Evolution: Allosteric Proteins

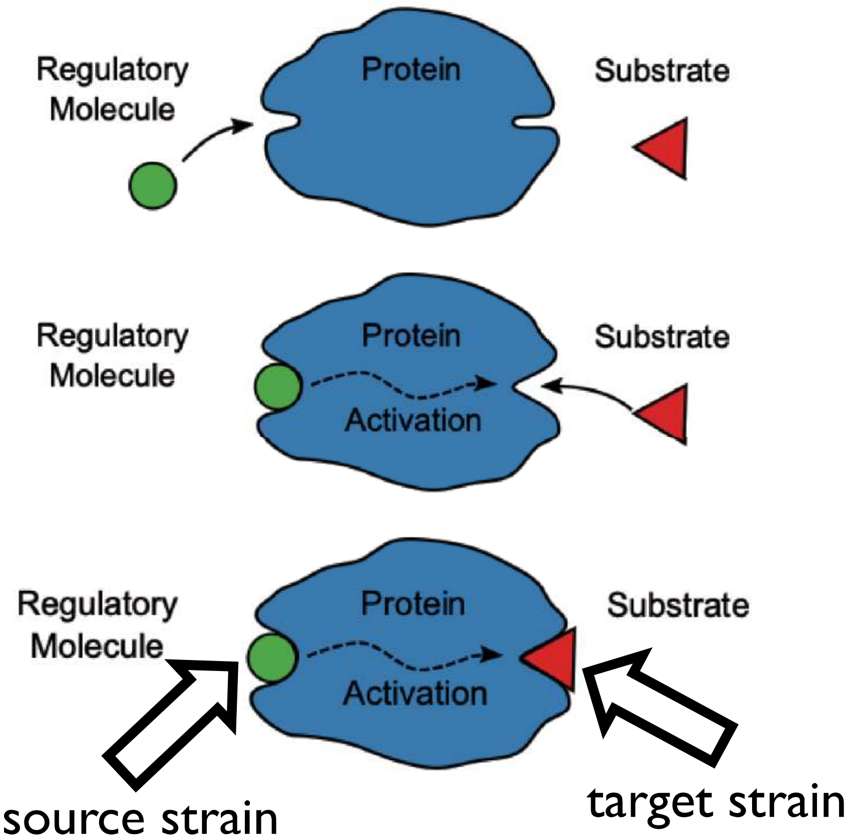


- Input: local strain at regulatory site
- Output: local strain
- Adaptive DOFs: choice of amino acid at each point in sequence

Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017
Rocks, Ronellenfitsch, Liu, Nagel, Katifori PNAS 2019
Hexner, Liu, Nagel PNAS (2020)
Stern, Hexner, Rocks, Liu PRX (2021)
Rocks, Katifori, Liu arXiv (2024)

Goodrich, Liu, Nagel PRL (2016)
Hexner, Liu, Nagel Soft Matter (2018)
Stern, Jayaram, Murugan Nat Comm (2018)
Pashine, Hexner, Liu, Nagel Sci Adv (2019)
Stern, Pinson, Murugan PRX (2020)
Stern, Arinze, Perez, Murugan PNAS (2020)
Arinze, Stern, Nagel, Murugan PRE (2023)

Living Systems With Adaptive DOF Tuned by Evolution: Allosteric Proteins



localized strain due to localized strain applied far away

- Input: local strain at regulatory site
- Output: local strain
- Adaptive DOFs: choice of amino acid at each point in sequence

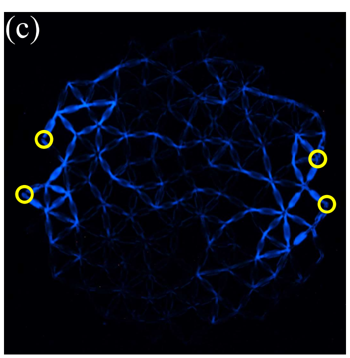
Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017
Rocks, Ronellenfitsch, Liu, Nagel, Katifori PNAS 2019
Hexner, Liu, Nagel PNAS (2020)
Stern, Hexner, Rocks, Liu PRX (2021)
Rocks, Katifori, Liu arXiv (2024)

Goodrich, Liu, Nagel PRL (2016)
Hexner, Liu, Nagel Soft Matter (2018)
Stern, Jayaram, Murugan Nat Comm (2018)
Pashine, Hexner, Liu, Nagel Sci Adv (2019)
Stern, Pinson, Murugan PRX (2020)
Stern, Arinze, Perez, Murugan PNAS (2020)
Arinze, Stern, Nagel, Murugan PRE (2023)

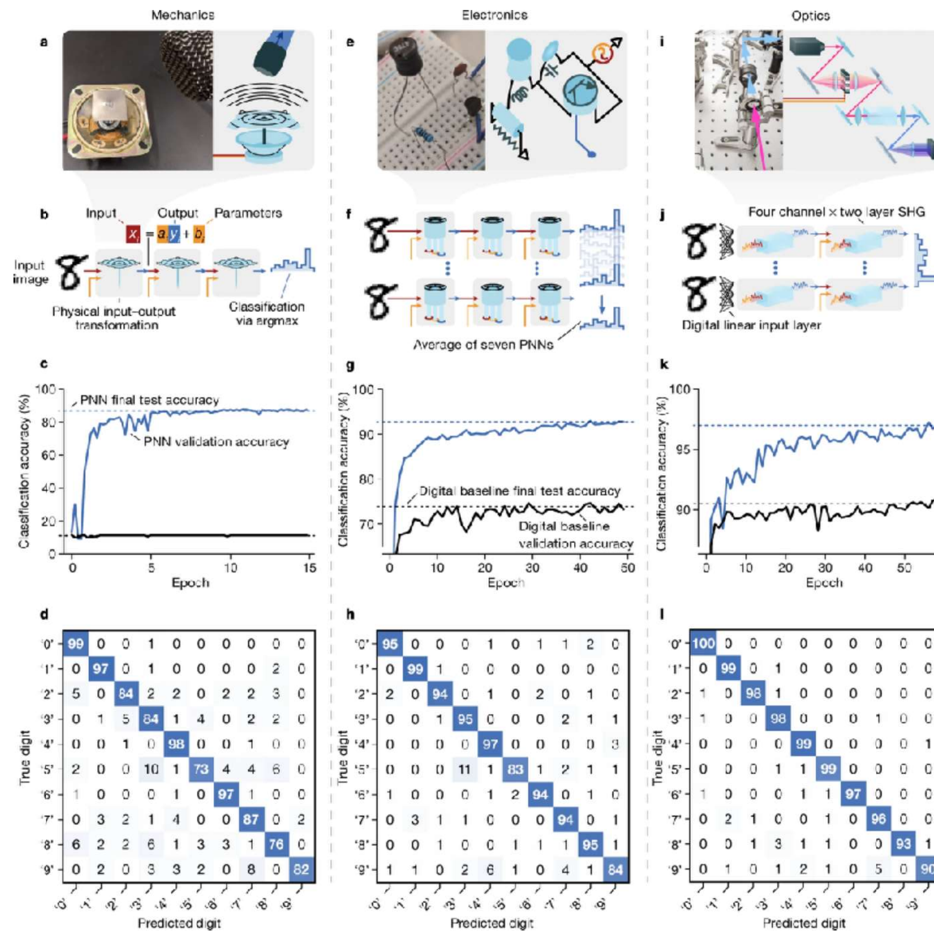
Systems with Adaptive DOF Tuned by External Brains/Processors

–

Rocks, Pashine,…Liu, Nagel
PNAS (2017)

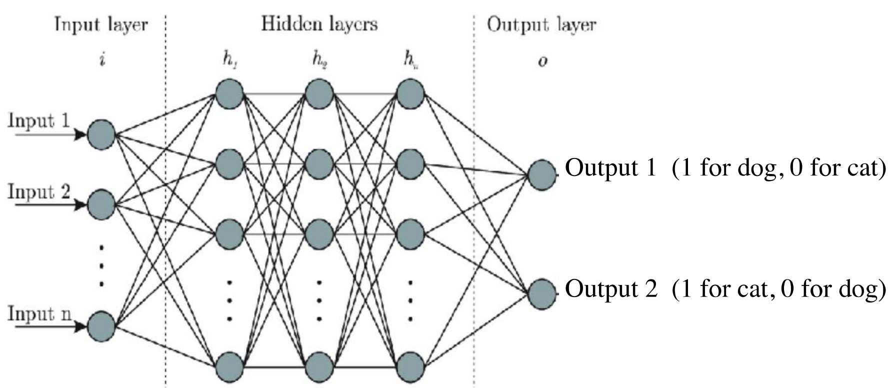


Pashine PRM (2021)

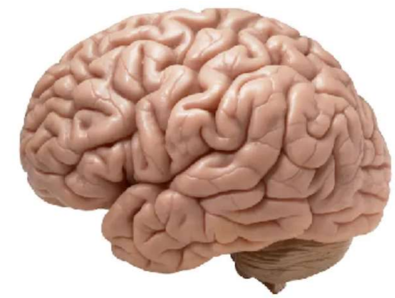


Wright, ..McMahon Nature (2022)

Problems with Usual External Brain/Processor: Not Scalable

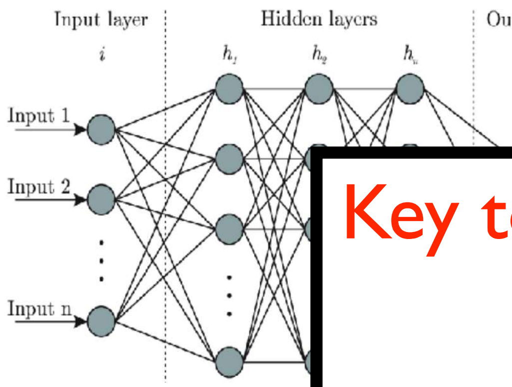


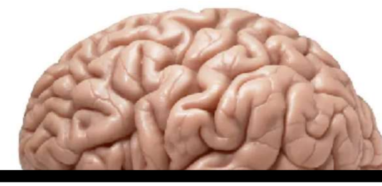
- Gradient descent on cost function requires knowing all network details
- Requires processor
- Costs much energy
  - Each ChatGPT text query for 100 words ~ 500 kJ + 17 oz of cooling water
  - Image generation ~10,000 kJ



- Neurons update without knowing what all other neurons are doing
- Doesn't require processor
- Relatively energy efficient: adult human brain uses ~ 2000 kJ/day

Problems with Usual External Brain/Processor: Not Scalable





Key to adjusting adaptive DOF in real time:

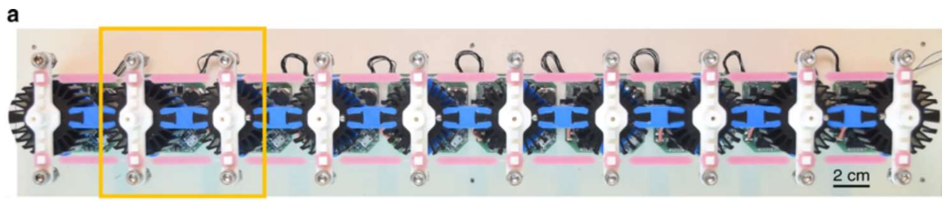
local rules

Stern, Murugan ARCMP (2023)

- Gradient descent on cost function requires knowing all network details
- Requires processor
- Costs much energy
  - Each ChatGPT text query for 100 words ~ 500 kJ + 17 oz of cooling water
  - Image generation ~10,000 kJ

- Neurons update without knowing what all other neurons are doing
- Doesn't require processor
- Relatively energy efficient: adult human brain uses ~ 2000 kJ/day

Systems with Adaptive DOF Implementing Local Rules With Processors

- Robotic swarms with tiny processor brains that control interactions based on local rules
- 

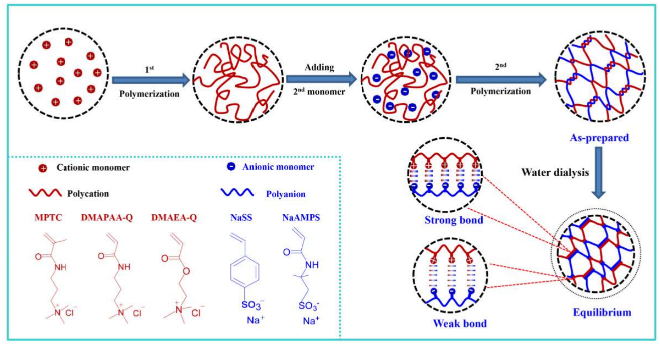
  Also I. Cohen, M. Z. Miskin, …

Brandenburger, ..Coulais Nature Comm (2019)

Systems That Adjust Themselves Using Local Rules

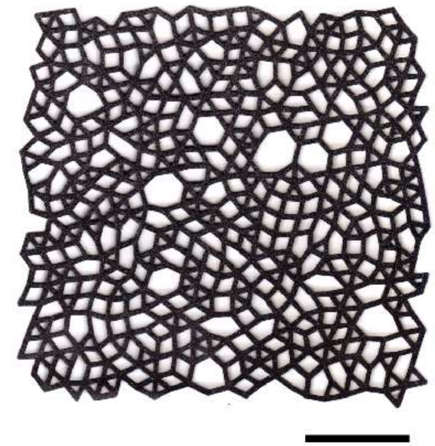
Vitrimers

Montornal, …Leibler Science (2011)

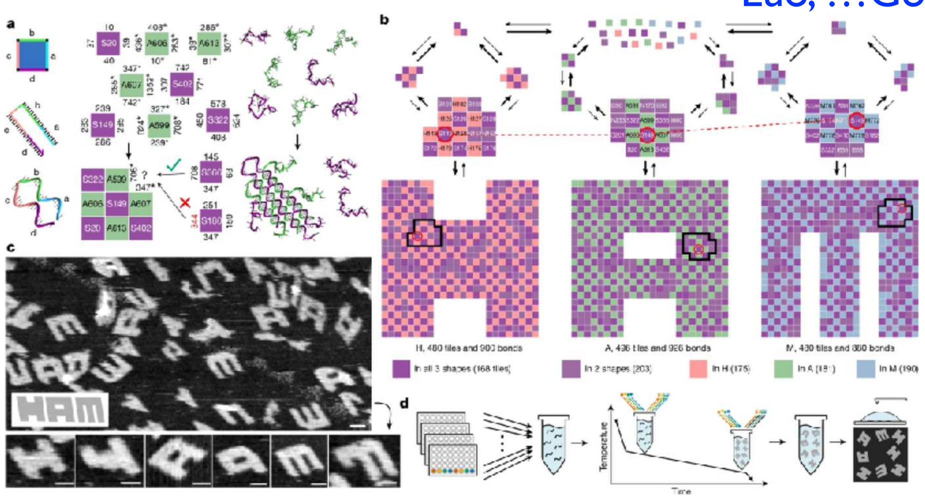


Double Self-Healing Networks

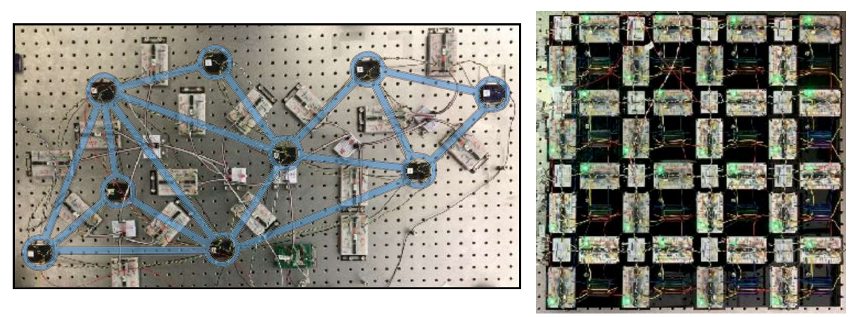
Luo, …Gong Adv Mat (2015)



Pashine, Hexner, Liu, Nagel
Sci Adv (2019)



Evans, …Winfree, Murugan Nature (2024)

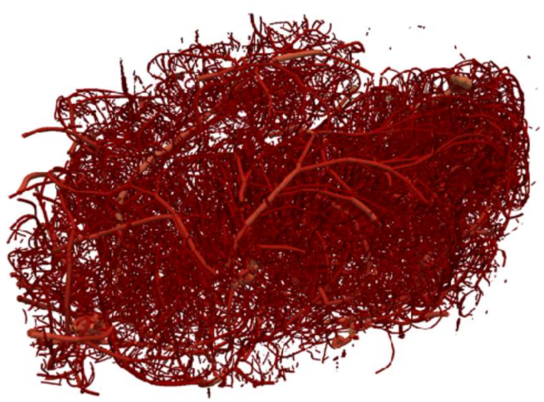


Contrastive Local Learning Networks

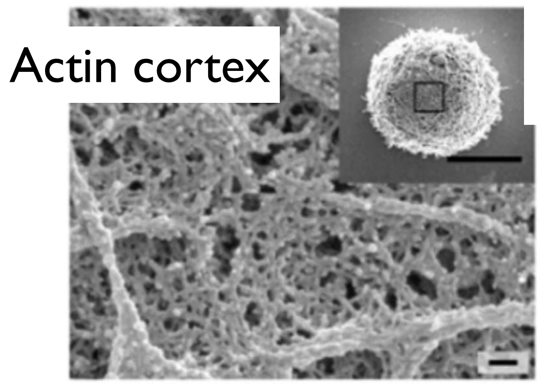
Dillavou, Stern, Liu, Miskin, Durian PRA (2022), PNAS (2024)

Systems That Adjust Themselves Using Local Rules

brain vasculature



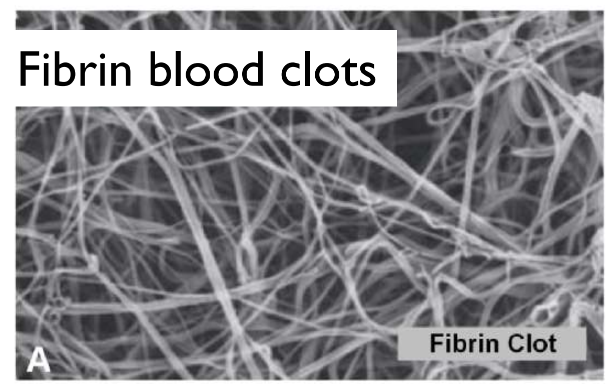
http://campar.in.tum.de/Main/ProjectVascularNetworks



Chugh et al Nat Cell Bio (2017)

Sharma et al Nat Phys (2016)

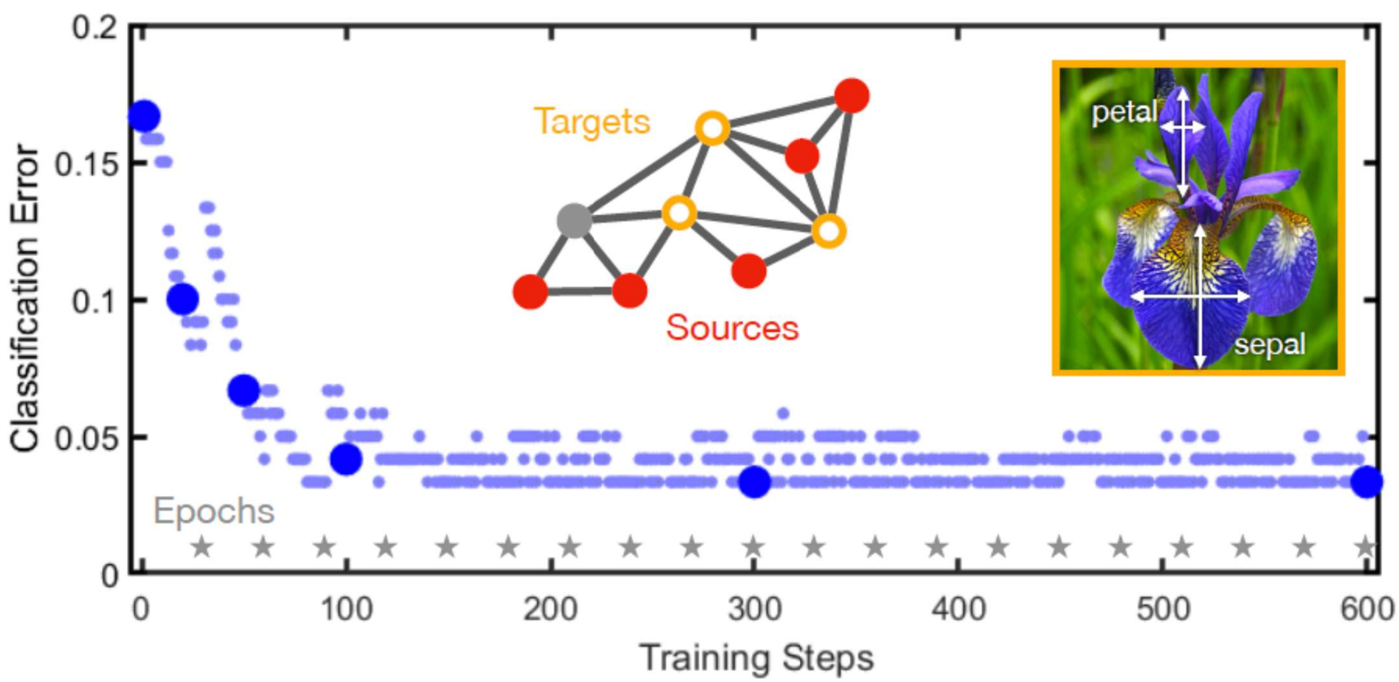
- Brain is 2% of body mass but uses 25% of O_2
- Input: pressure drop at arteries
- Output: pressure drop (enhanced blood current/ O_2) at specific local region
- Adaptive DOFs: conductances of network edges



Yasudasan, Averett Polym (2020)

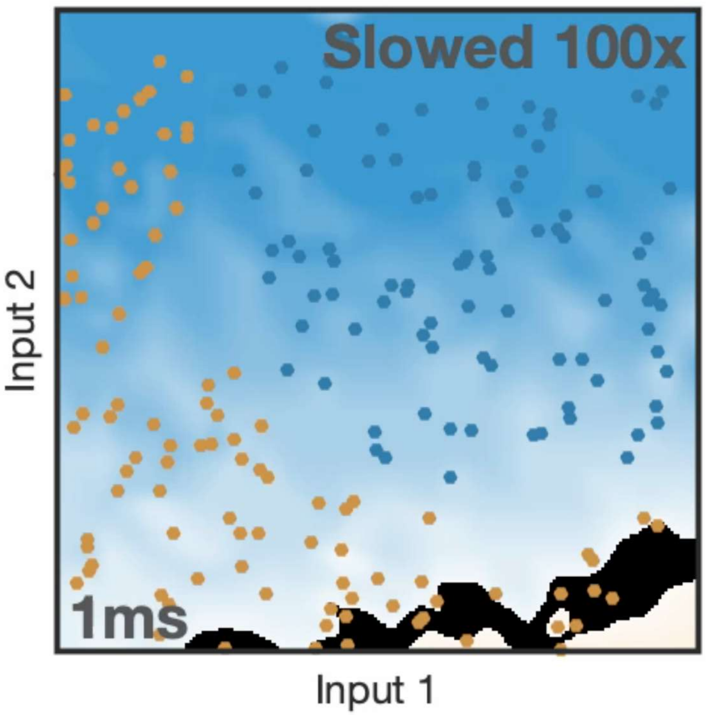
Classifying Irises

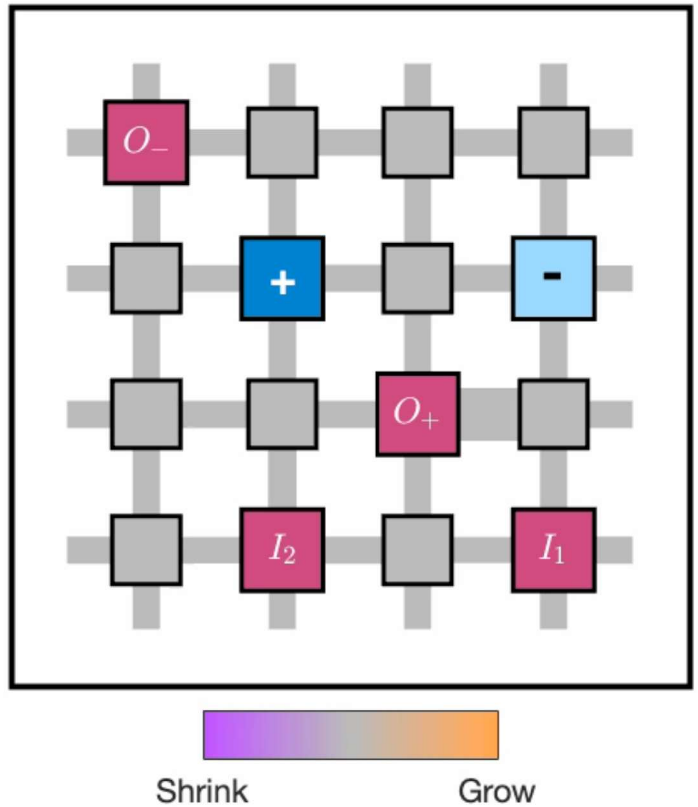
- 4 source nodes for measurements of 4 quantities
- 3 target nodes for 3 types of irises
- Classification accuracy = 97%! It really works!

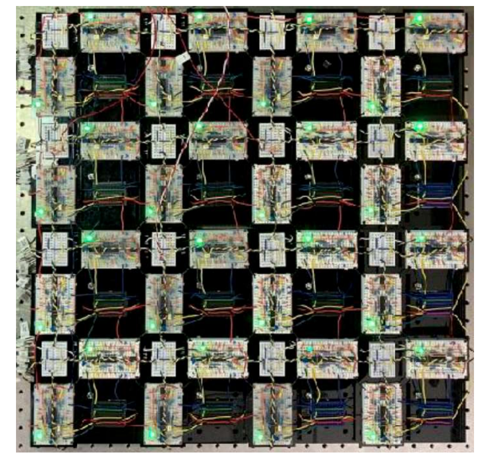


Dillavou, Stern, Liu, Durian PRApplied (2022)

Learning Nonlinear Classification

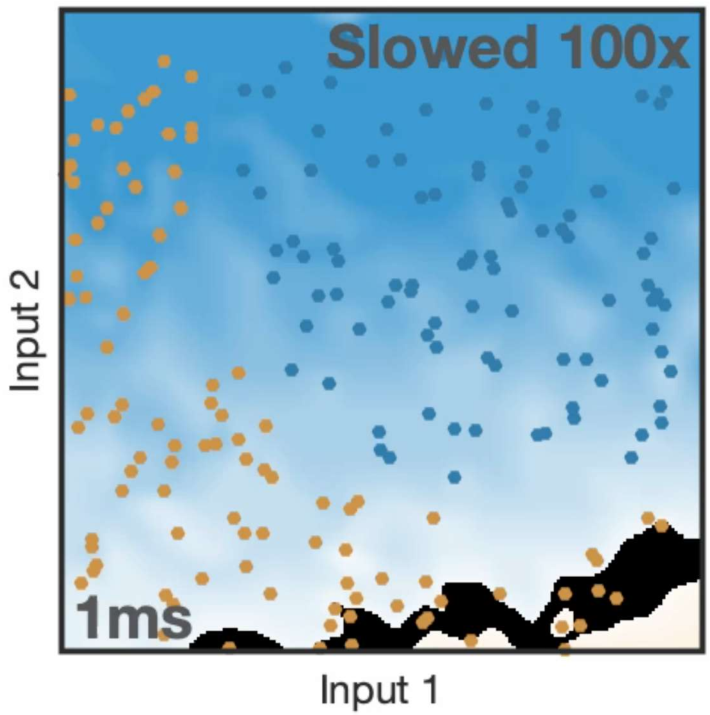


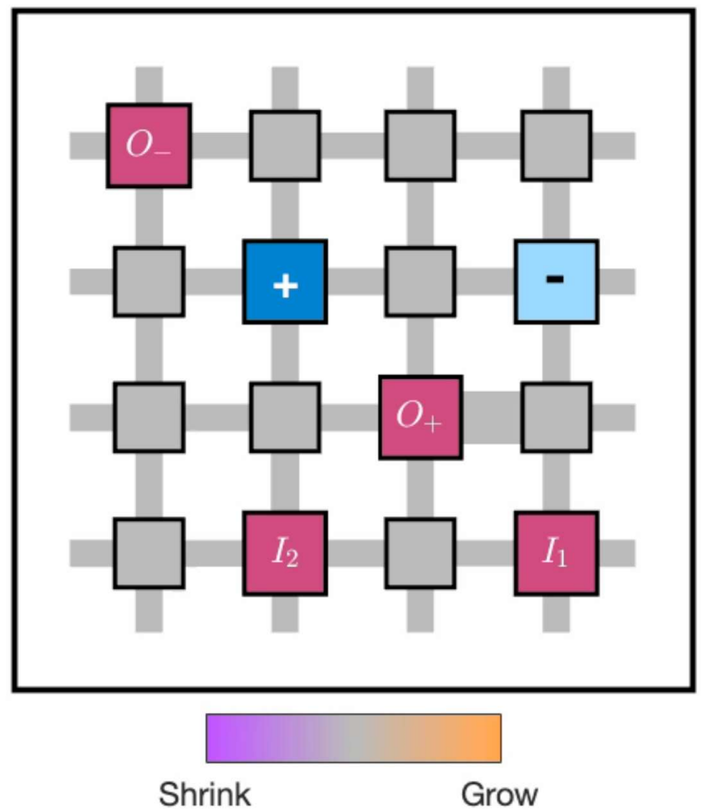


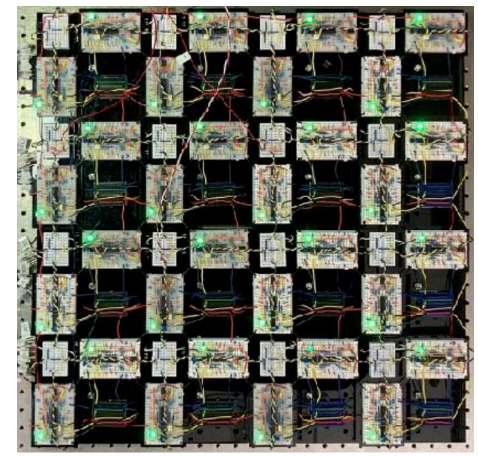


Dillavou, …Miskin, Durian PNAS (2024)

Learning Nonlinear Classification

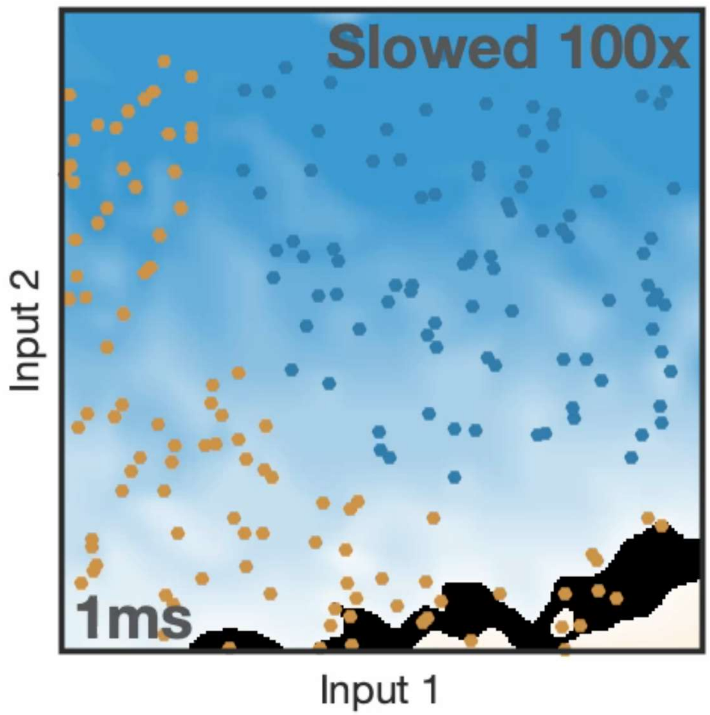


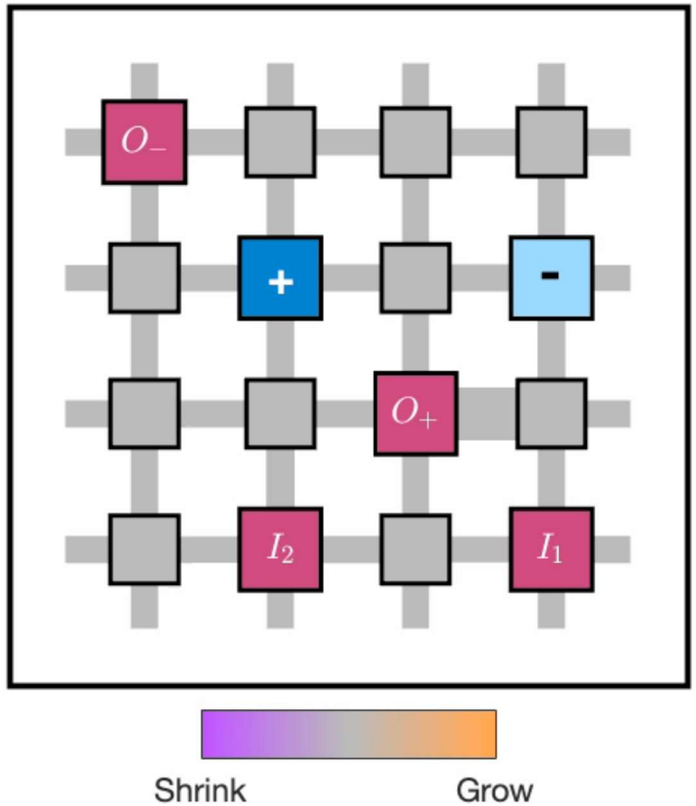




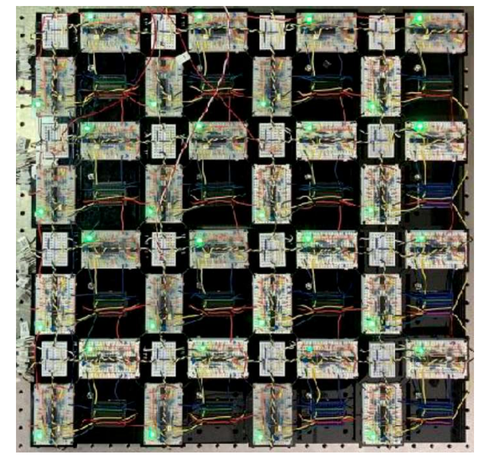
Dillavou, …Miskin, Durian PNAS (2024)

Learning Nonlinear Classification





Width of edges corresponds to conductance



Dillavou, …Miskin, Durian PNAS (2024)

Scaling Up

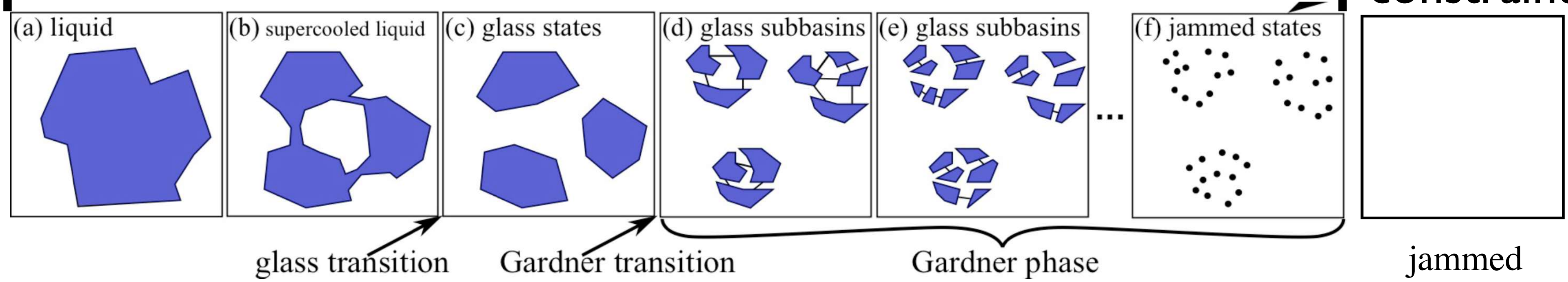
- Robustness to damage implies manufacturability
- Parameter changes are calculated in parallel by every edge
- Current ~10 pJ/parameter/inference
- Already comparable to ~1 pJ/parameter/inference on most energy-efficient Qualcomm AI chip
- Next step: printed circuit boards/chips for networks with ~100-200 edges (parameters)
- Then 3rd-gen design microfabrication of up to 10^7 edges on a chip
- If everything scales up as we hope
  - ~10^9 edges/chip
  - ~1 aJ/parameter/inference

General Strategy For Complex Functionality

- Hard spheres

Can satisfy constraints

Can't satisfy constraints

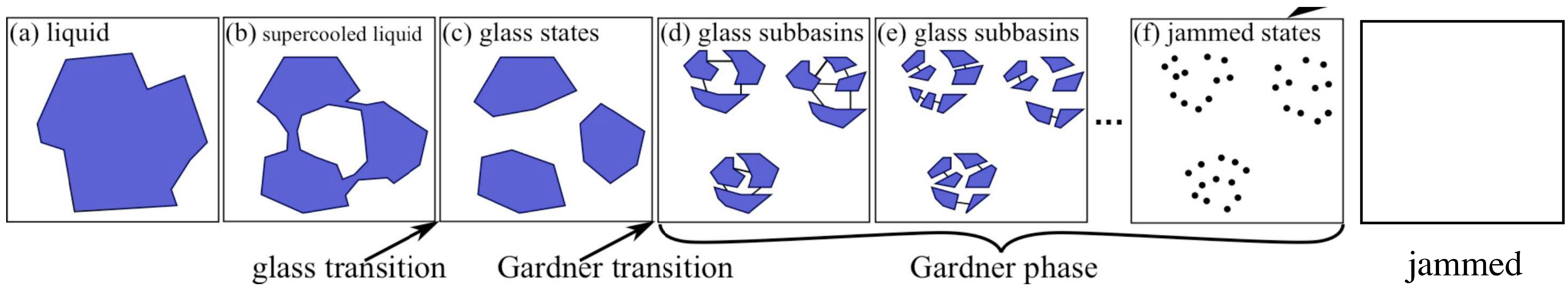


low density → high density

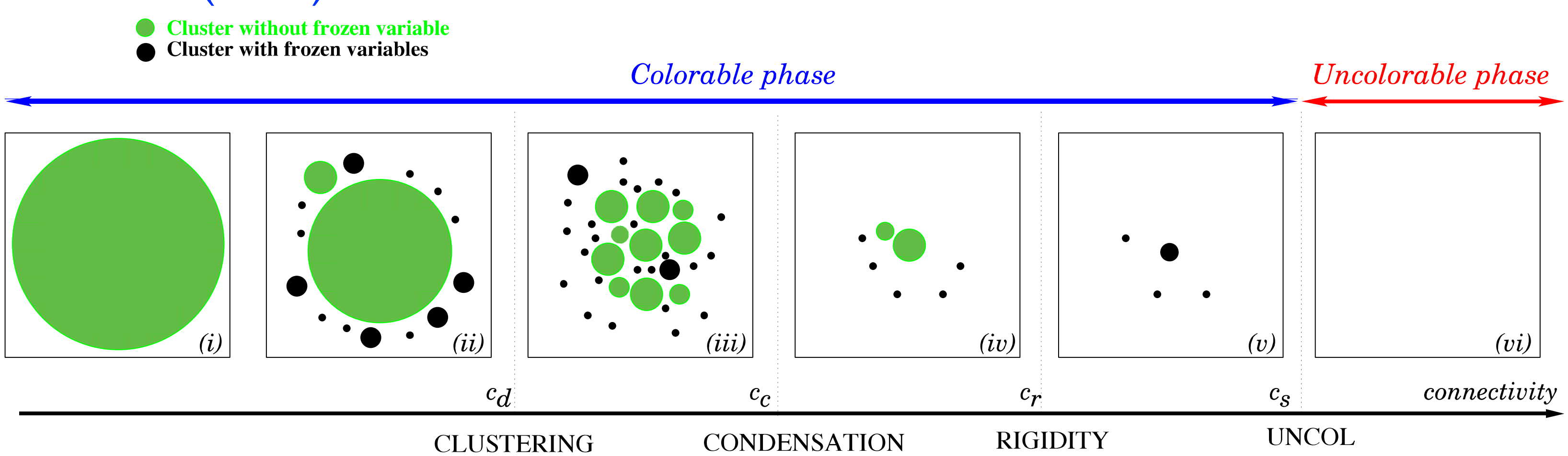
$$\frac{\#\text{constraints}}{\#\text{adaptive DOF}} \ll 1 \qquad \frac{\#\text{constraints}}{\#\text{adaptive DOF}} \approx 1 \qquad \frac{\#\text{constraints}}{\#\text{adaptive DOF}} \gg 1$$

Patrick Charbonneau, Jorge Kurchan, Giorgio Parisi, Pierfrancesco Urbani, Francesco Zamponi. Exact theory of dense amorphous hard spheres in high dimension. III. The full RSB solution. J. Stat. Mech: Theory and Expt (2014)

Compare to k-coloring Constraint-Satisfaction Problem



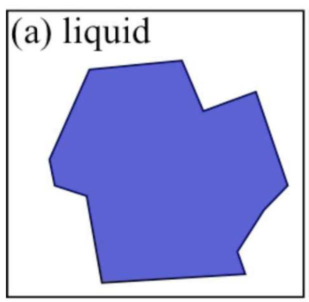
Charbonneau et al. (2014)



Krzakala, Zdeborova (2007)

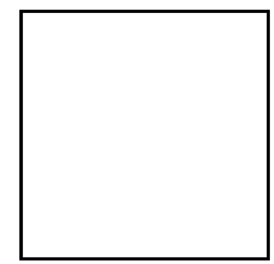
Constraint-Satisfaction Problems

Can satisfy constraints



non-universal stuff

Can't satisfy constraints



over-parameterized → **under-parameterized**

$$\frac{\#\text{constraints}}{\#\text{adaptive DOF}} \ll 1 \qquad \frac{\#\text{constraints}}{\#\text{adaptive DOF}} \approx 1 \qquad \frac{\#\text{constraints}}{\#\text{adaptive DOF}} \gg 1$$

Stay in this regime!

Overparameterization is Useful!

- AI/CLLNs work in over-parameterized regime
- Speculation: biology stays in overparameterized regime
  - Easiest to introduce extensive number of adaptive DOF
  - Lots of good solutions that are easy to find (every embryo is different!)
  - Good solutions are more generalizable (robust)
    - Few important directions in cost landscape, many unimportant, flattish directions
    - Possibility of representational drift
- Good solutions form ensemble that we can study as statistical physicists to understand how collective function emerges from microscopic parameter adjustments

Summary

- Adaptive matter is a new class of condensed matter
  - Many more can be different because more adaptive DOF can satisfy more constraints
  - Almost all examples are from living matter
  - Except: contrastive local learning networks (Dillavou/Durian et al), double networks (J-P Gong)
- Important for understanding function in adaptive matter
- Advantages of unifying framework for adaptive matter
  - Lots of good solutions that are easy to find (every embryo is different!)
  - Good solutions are more generalizable (robust)
  - Can local rules used in biological systems be used to design soft adaptive matter?
  - ….
- It's time for CMMR to face up to adaptive matter! Need unifying framework!

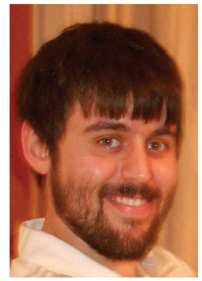
Thanks to

Menachem Stern

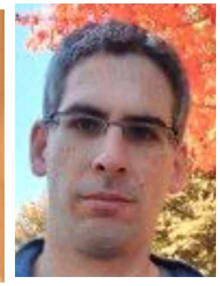
Sam Dillavou

Ben Pisanty

Jovana Andrejevic

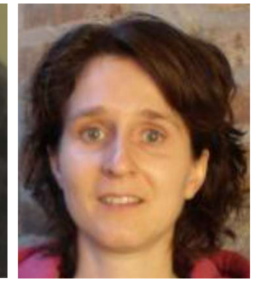


Jason Rocks

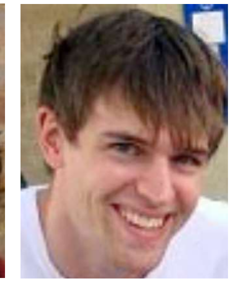


Daniel Hexner

Nidhi Pashine

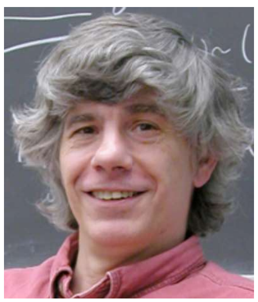


Irmgard Bischofberger



Carl Goodrich

Eleni Katifori



Sid Nagel

John Crocker

Doug Durian

Marc Miskin

DOE Biomolecular Materials

NSF DMR Materials Theory

NSF UPenn MRSEC

Simons Foundation