Thanks to

Marco Cunha

Sadjad Arzash

Indrajit Tah

Daniel Hexner

Jason Rocks

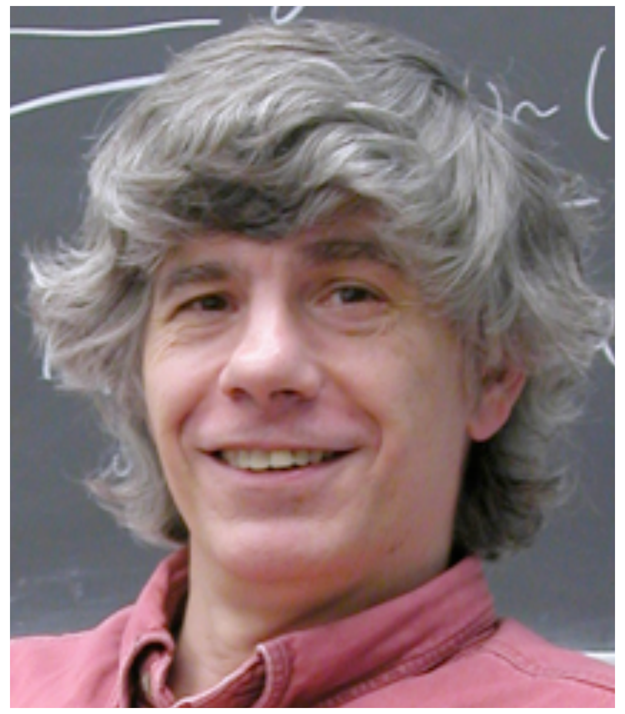
Carl Goodrich

Nidhi Pashine

Irmgard Bischofberger

Daniel Härtter

Janice Crawford



Sid Nagel

Eleni Katifori

John Crocker

Lisa Manning

Dan Kiehart

Christoph Schmidt

NSF DMR
Materials Theory

DOE
Biomolecular Materials

Simons Foundation

Center for Soft and Living Matter at Penn

- >60 faculty from >10 departments across STEM@ Penn
- Awards to faculty since January 2021 include
  - BMES Shu Chien Achievement Award (Dennis Discher)
  - NAE (Kate Stebe)
  - OSA Michael S. Feld Biophotonics Award (Arjun Yodh)
  - Sloan Fellowship, Packard Fellowship (Marc Miskin)
  - SIAM Early Career Prize (Paris Perdikaris)
  - APS Early Career Award for Soft Matter Research (Eleni Katifori)
  - MRS Fellow (Shu Yang, Eric Stach)
  - AAAS Fellow (Doug Durian, Karen Winey)
  - ACS Polymer Prize (Karen Winey)
  - Intel Outstanding Research Award (Chinedum Osuji)
  - ACS Outstanding Achievement Award in Nanoscience (Daeyeon Lee)
  - …

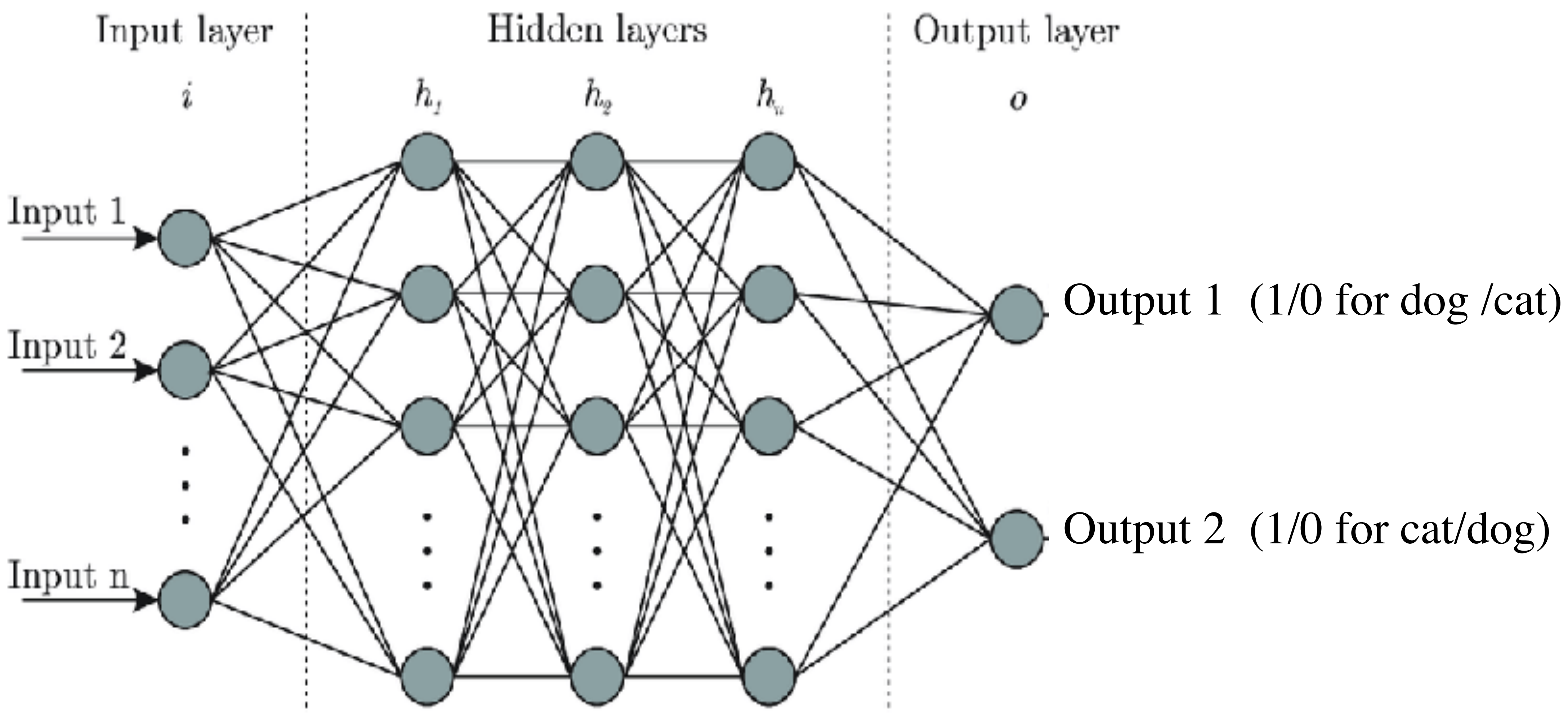
More is Different vs. Many More is Different

- "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 (~10^6 neurons) vs humans (~10^{11} neurons)
  - Digital neural networks: ChatGPT4 (~10^{14} parameters)

ADAPTIVE MATTER

- Why is many more different in these systems?
- What physical systems besides brains might have this property?

Why is Many More Different in Digital Neural Networks?



- Network has adjustable parameters
- How to adjust?
  - Evaluate cost function $\mathcal{C} = \sum_i (\text{desired output}_\text{i} - \text{free output}_\text{i})^2$
  - Minimize $\mathcal{C}$ by adjusting parameters
  - This is how system evolves

Constraint-Satisfaction Problems

- Each term in cost function is another constraint

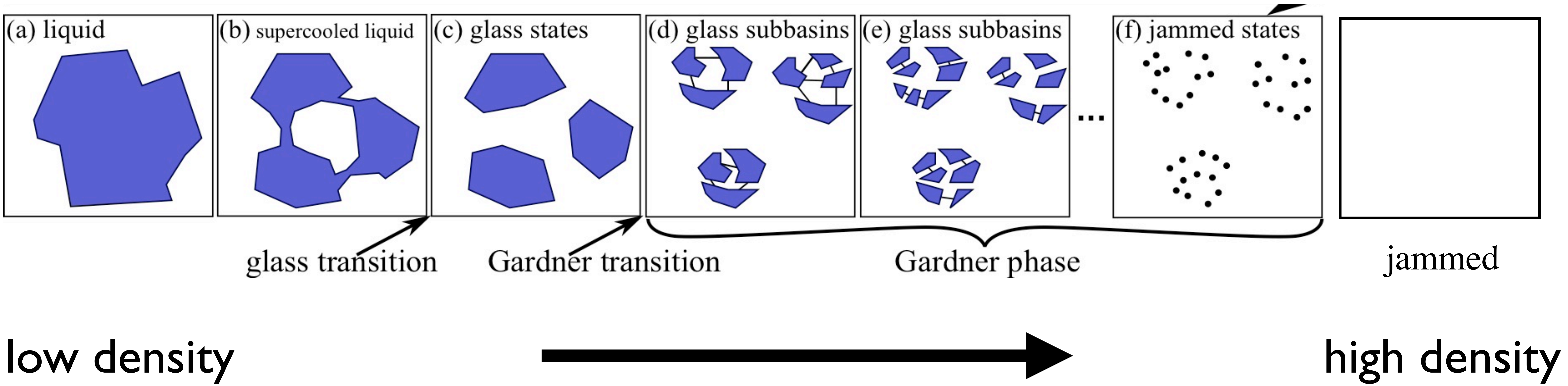
$$\mathcal{C} = \sum_i (\mathrm{desired\ output_i} - \mathrm{free\ output_i})^2$$

- Similar to problem of jamming. E = \sum (overlaps between particles)^2
- When all constraints are satisfied, no overlaps between particles (hard sphere configurations)
- Hard spheres solved in d=infinity so look at what happens there

GO TO SLIDES

Hard Sphere Constraint-Satisfaction Problem

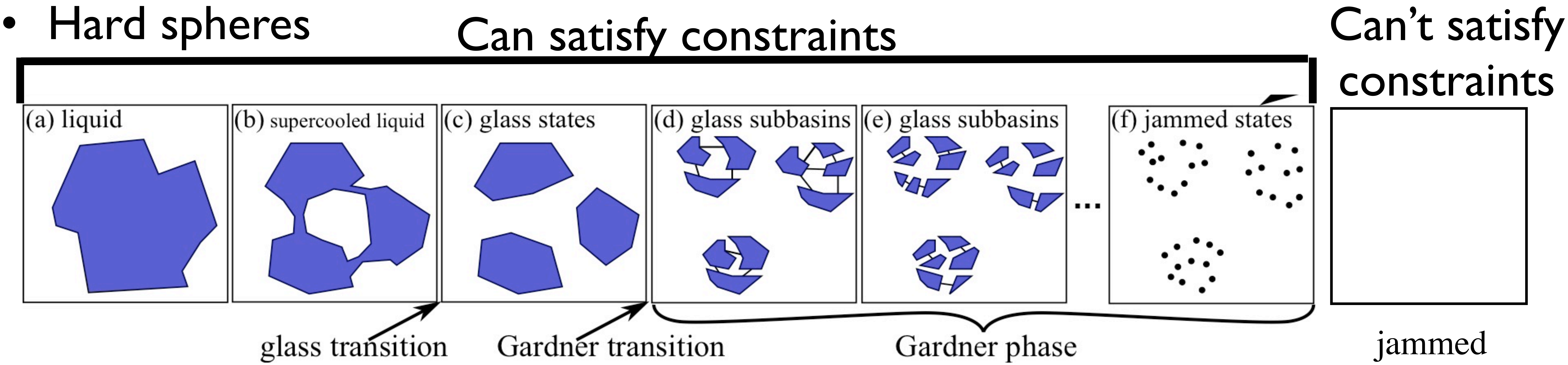
- Hard spheres



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. Journal of Statistical Mechanics: Theory and Experiment, 2014, 2014, pp.10009. 10.1088/1742- 5468/2014/10/P10009

Hard Sphere Constraint-Satisfaction Problem

- Hard spheres

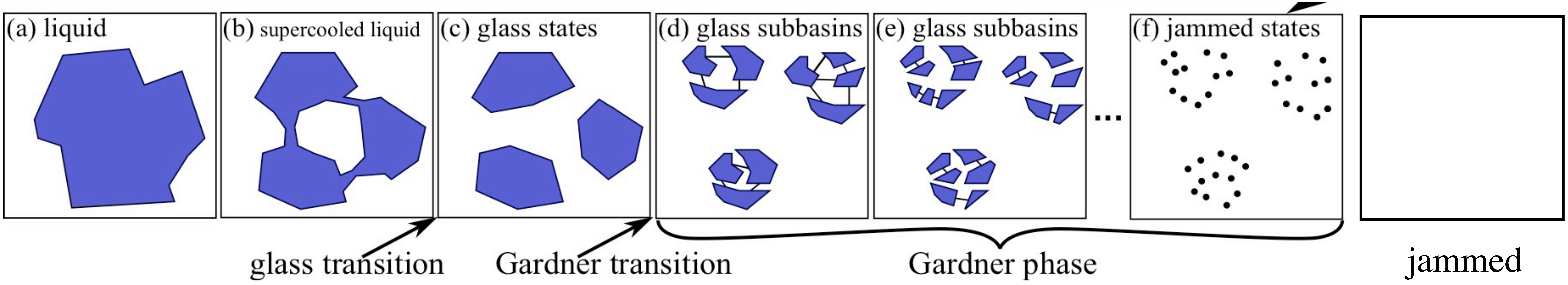


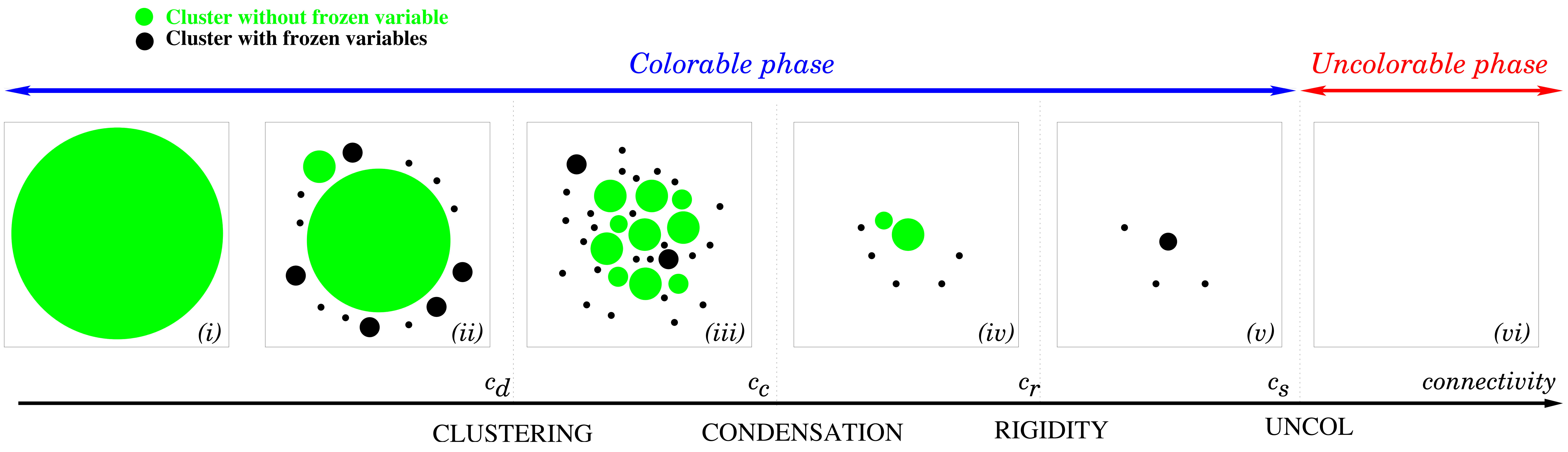
low density → high density

$$\frac{\text{number of constraints}}{\text{number of degrees of freedom}} \ll 1$$

$$\frac{\text{number of constraints}}{\text{number of degrees of freedom}} \gg 1$$

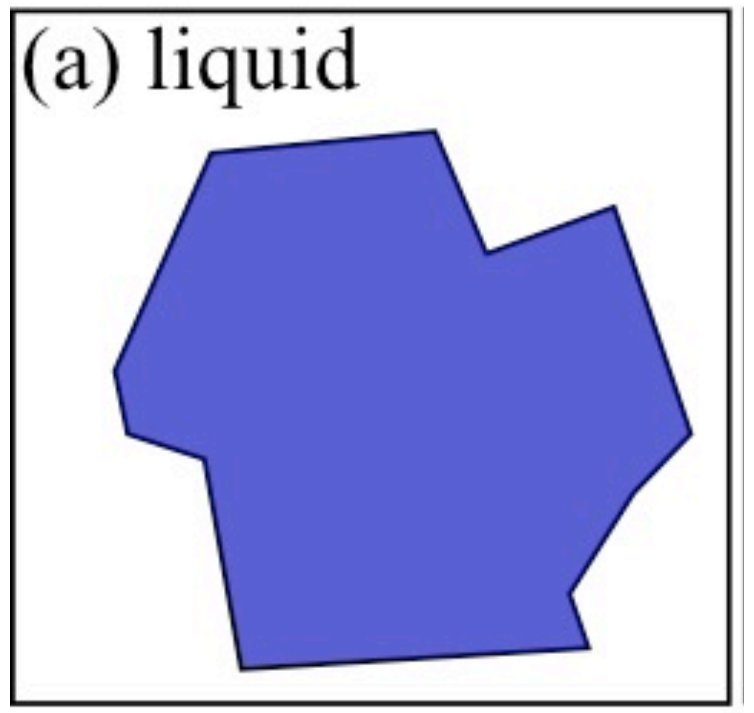
Constraint-Satisfaction Problems



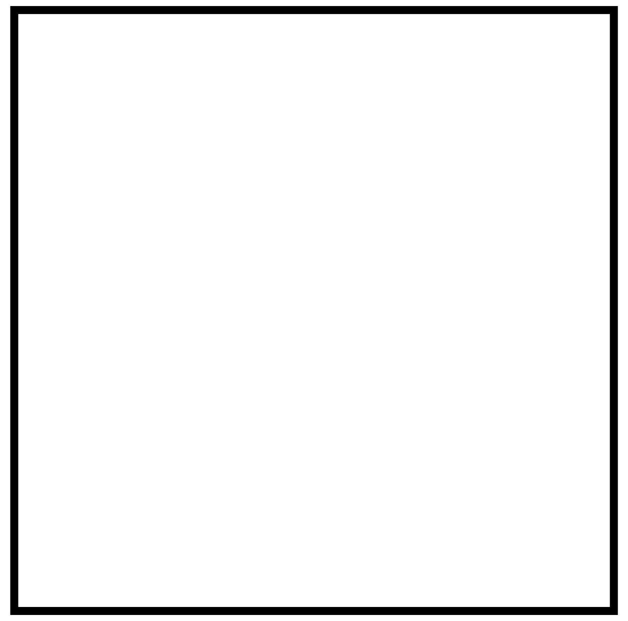


Constraint-Satisfaction Problems

Can satisfy constraints



Can't satisfy constraints



Non-universal stuff

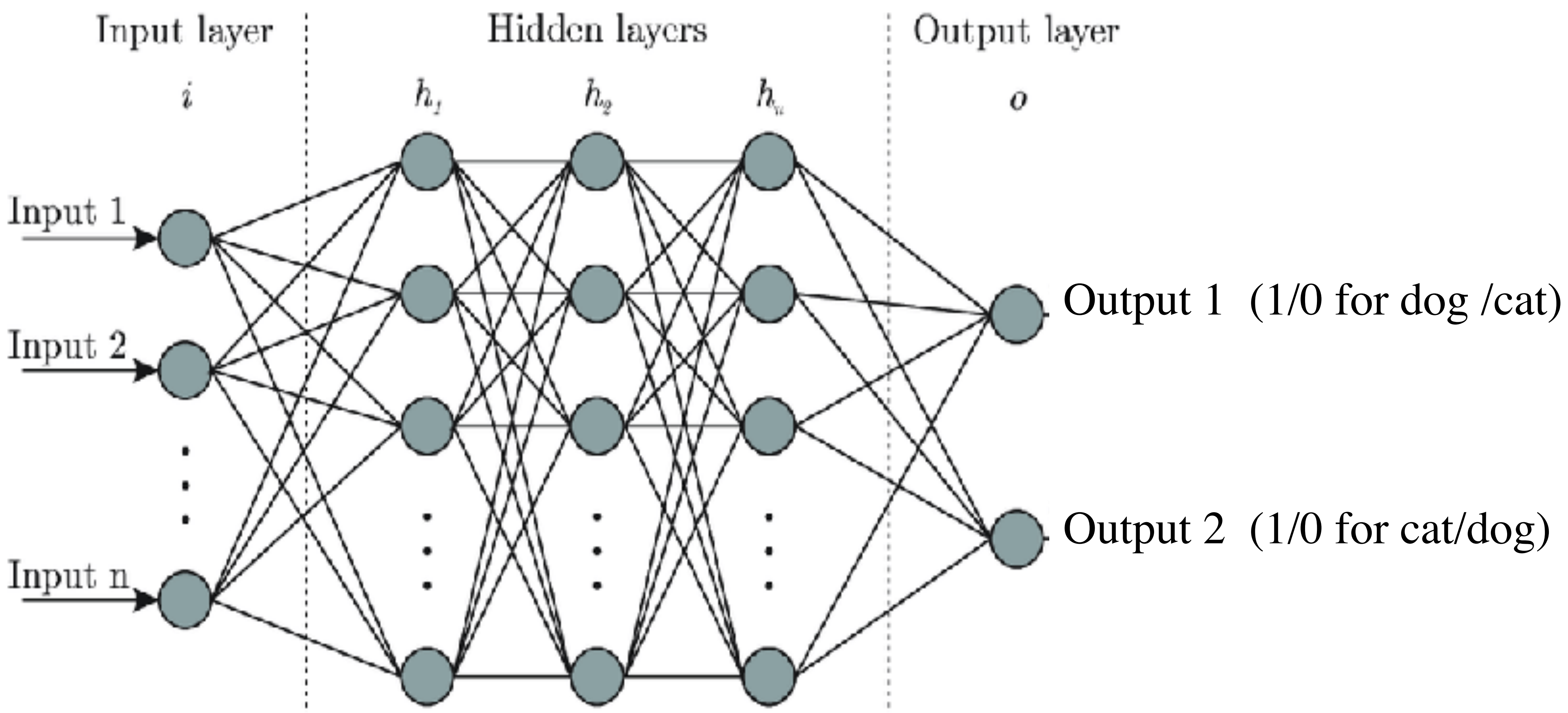
over-parameterized

$$\frac{\text{number of constraints}}{\text{number of degrees of freedom}} \ll 1$$

under-parameterized

$$\frac{\text{number of constraints}}{\text{number of degrees of freedom}} \gg 1$$

Why is Many More Different in Digital Neural Networks?



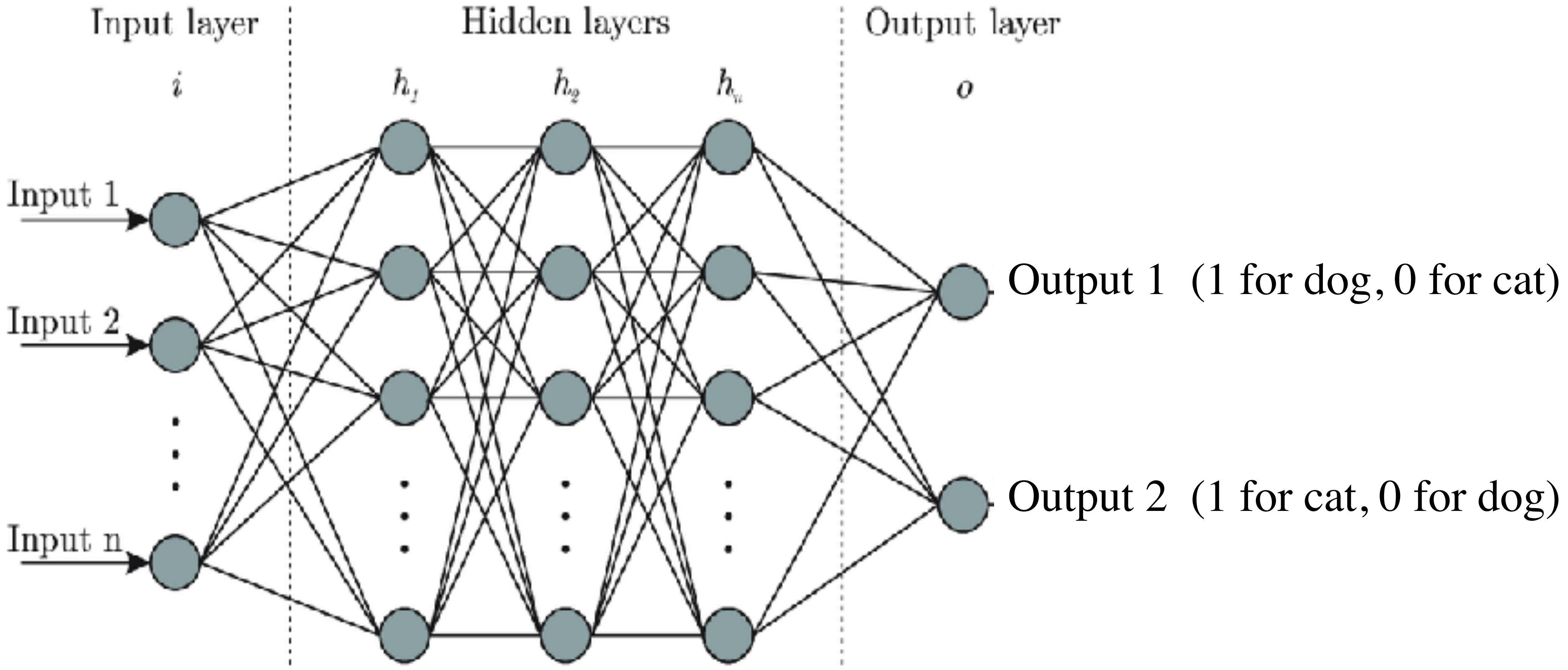
- Network has adjustable parameters
- How to adjust?
  - Evaluate cost function $\mathcal{C} = \sum_i (\text{desired output}_\text{i} - \text{free output}_\text{i})^2$
  - Minimize $\mathcal{C}$ by adjusting parameters
- # constraints that can be satisfied increases with # parameters
- Many more is different bc # parameters increases with system size

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 adjustable parameters!
- 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 adaptive matter?
  - # parameters increases with system size

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

Look Again to Digital Neural Networks

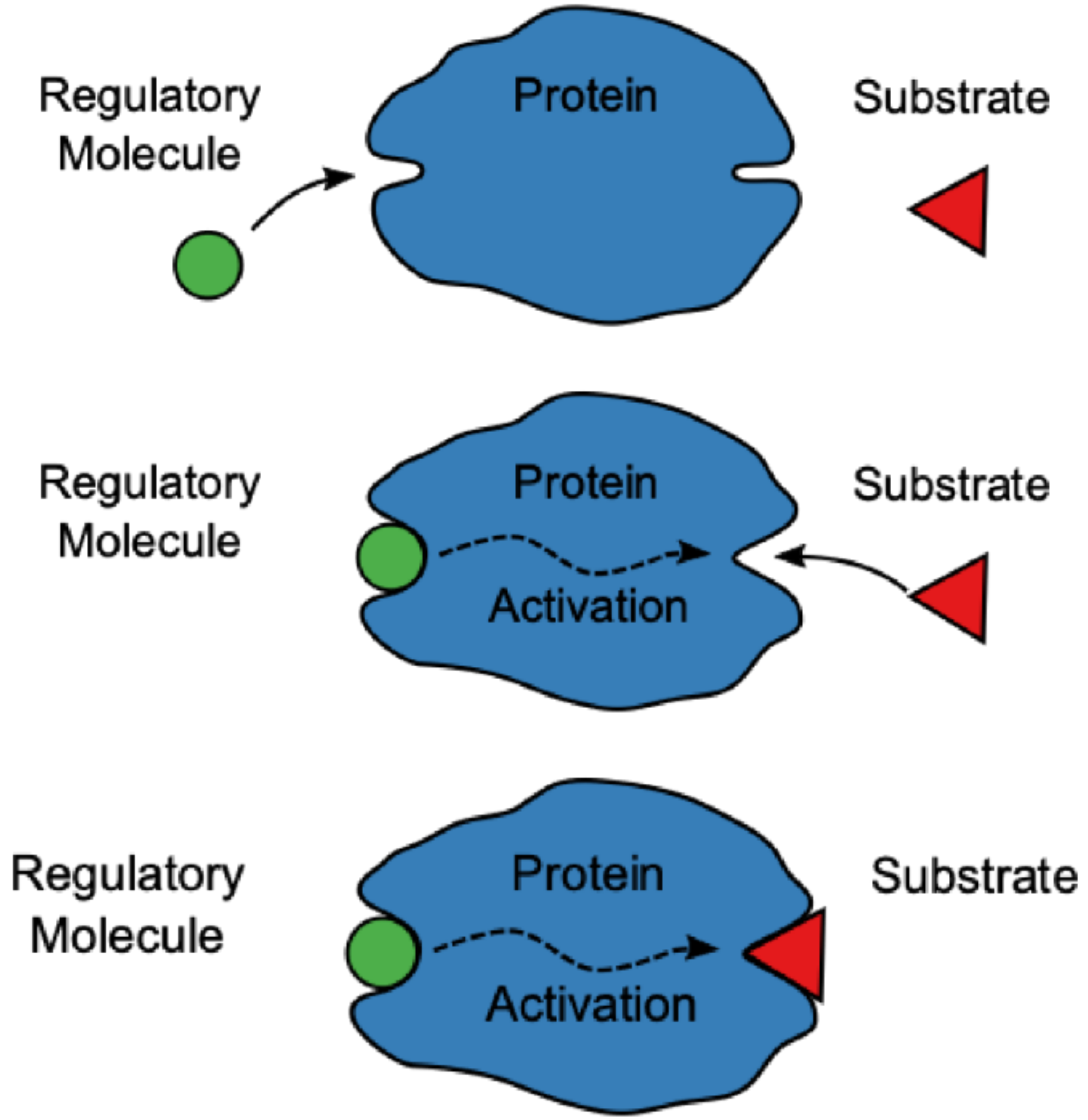


- They learn desired input/output relations—desired responses to stimuli
- Inverse input/output problems
- What real systems have desired responses to stimuli?
- Do they also have adjustable parameters?

Unifying Framework for Adaptive Matter: Definitions

- What is the function?
  - Not fitness but constraints that must be satisfied
  - Example of protein allostery (GO TO SLIDE)
- What are the parameters?
  - Level of description needs to be specified
    - e.g. protein allostery
      - Microscopic level (amino acids, atom positions)
      - Coarse-grained level (spring constants/equilibrium lengths, node positions) (GO TO SLIDE)
- What is process by which parameters are adjusted to achieve function?
  - protein allostery: evolution (GO TO SLIDE)
  - DNNs: global process where you need to know everything, e.g. gradient descent

Allosteric Proteins Have Desired Input/Output Relations

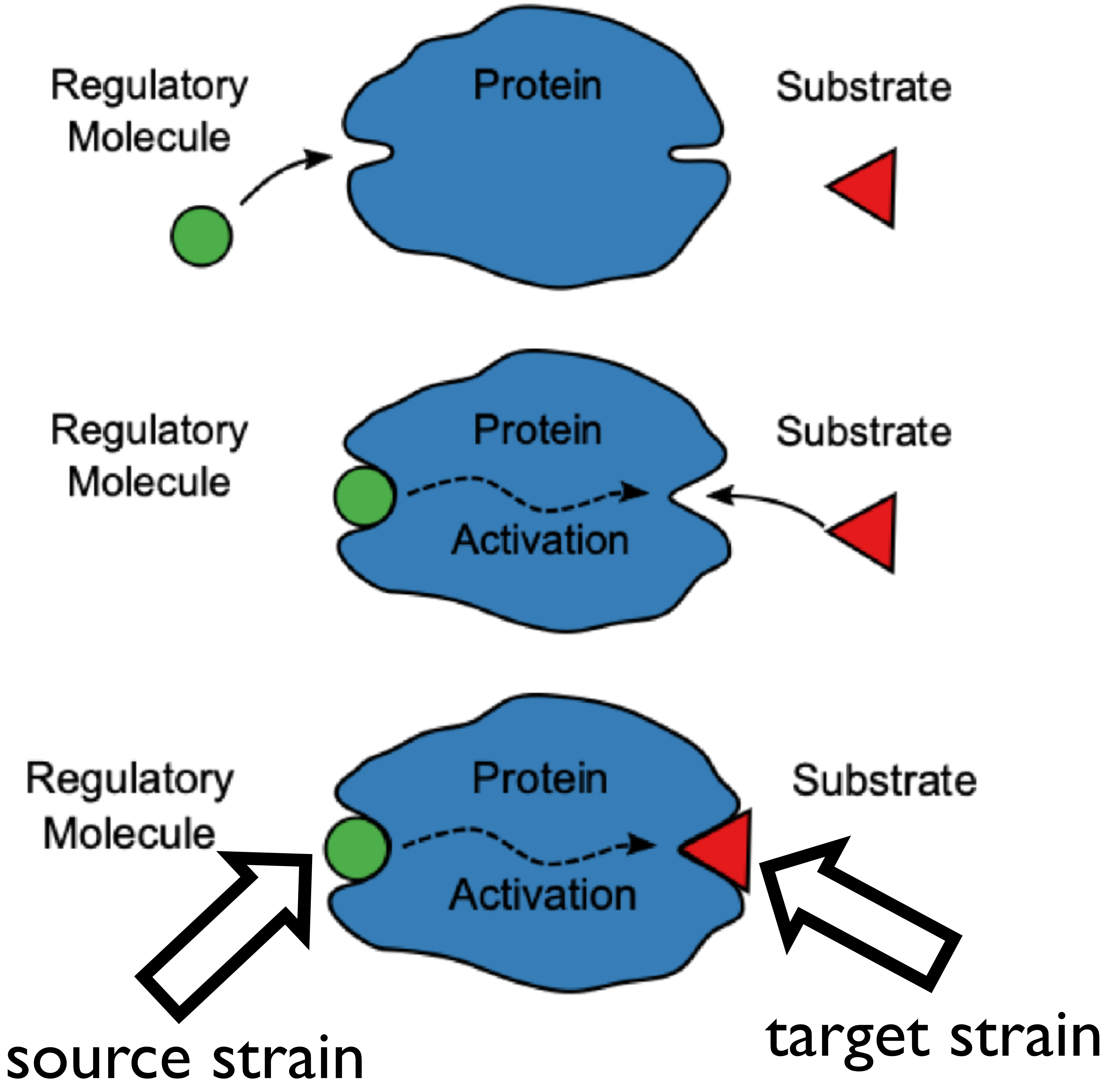


- Input: local strain at regulatory site
- Output: local strain

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)

Allosteric Proteins Have Desired Input/Output Relations



localized strain due to localized strain applied far away

- Input: local strain at regulatory site
- Output: local strain

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)

Unifying Framework for Adaptive Matter: Definitions

- What is the function?
  - Not fitness but constraints that must be satisfied
  - Example of protein allostery (GO TO SLIDE)
- What are the parameters?
  - Level of description needs to be specified
    - e.g. protein allostery
      - Microscopic level (amino acids, atom positions)
      - Coarse-grained level (spring constants/equilibrium lengths, node positions) (GO TO SLIDE)
- What is process by which parameters are adjusted to achieve function?
  - protein allostery: evolution (GO TO SLIDE)
  - DNNs: global process where you need to know everything, e.g. gradient descent
  - Brains: Hebbian rule is "local rule"—each neuron does not need to know what every other neuron is doing in order to change its synapses

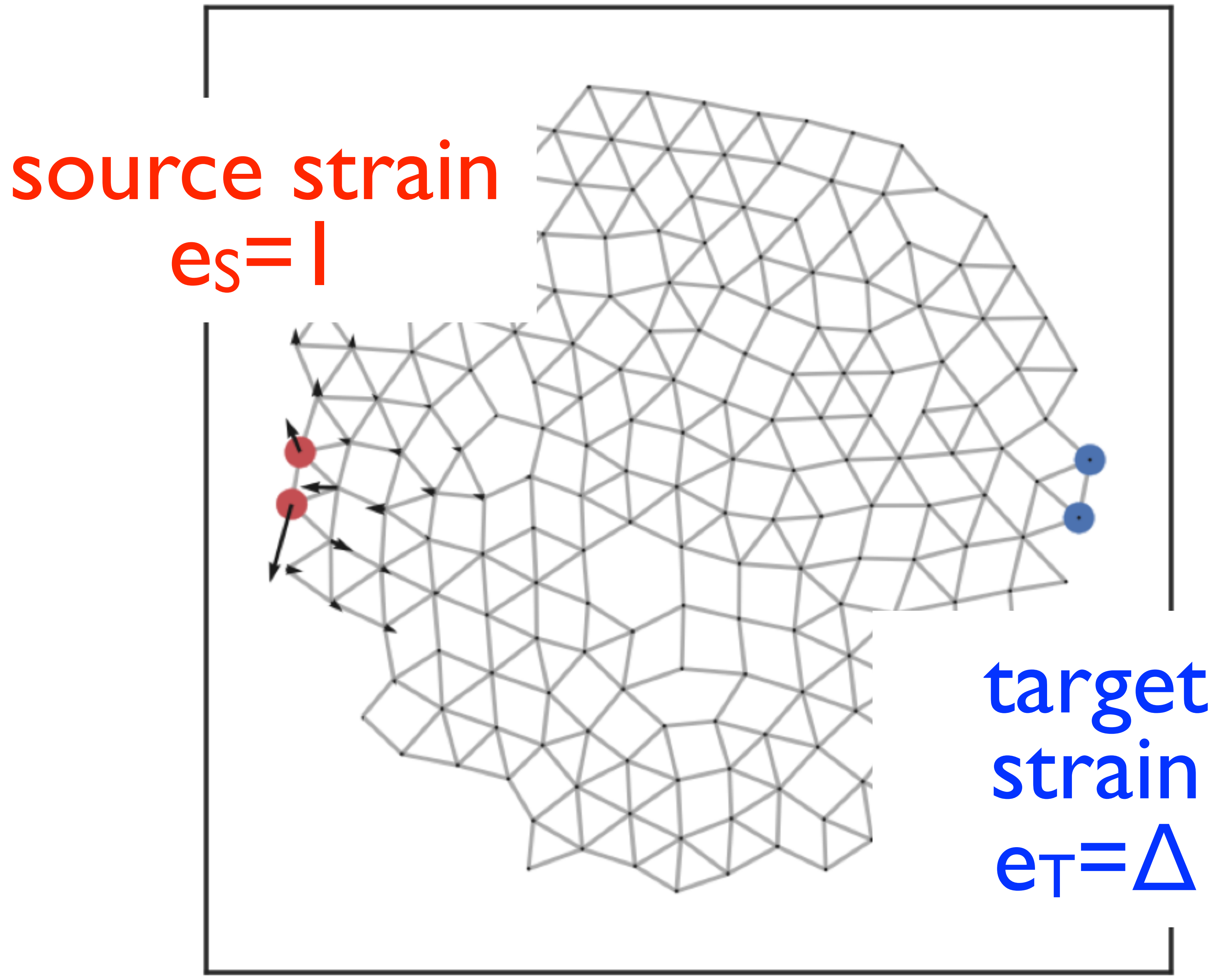
GO TO SLIDE

For Physical Systems with Lyapunov Function: Double Optimization

- Cost function: eg for allostery

$$\mathcal{C} = (\text{desired target strain } - \text{ free target strain})^2 = (\Delta - e_T)^2$$

- Parameters: edge stiffnesses/equil lengths/presence-absence…
- Physics requires minimization of elastic energy/free energy
- Physical DOF: node positions
- 1st optimzn: Minimize energy wrt physical DOF —required by physics!
- 2nd optimzn: Use gradient descent to minimize cost function wrt parameters



GO TO SLIDES

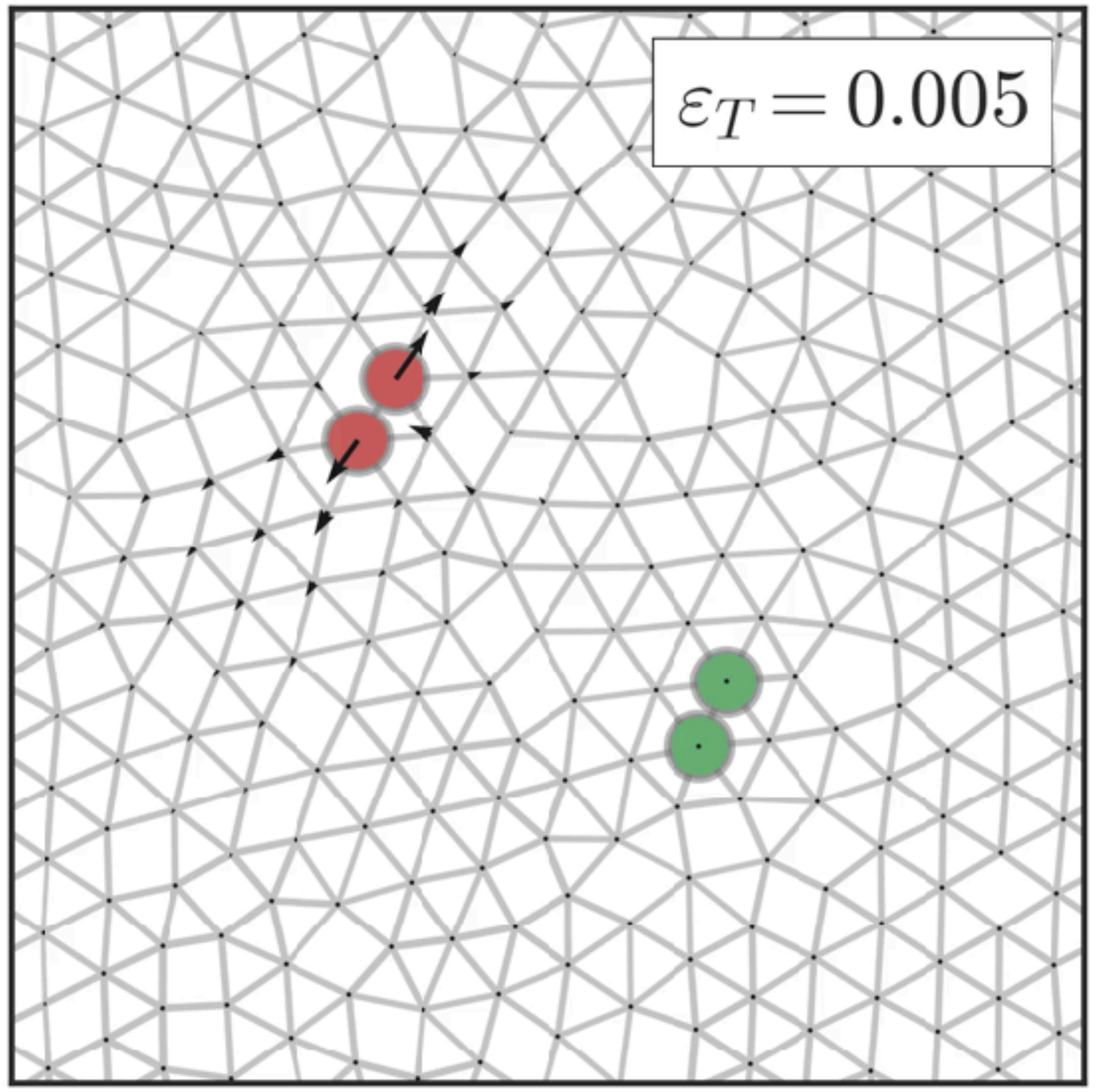
Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017

Double Optimization Works!

- Cost function: eg for allostery

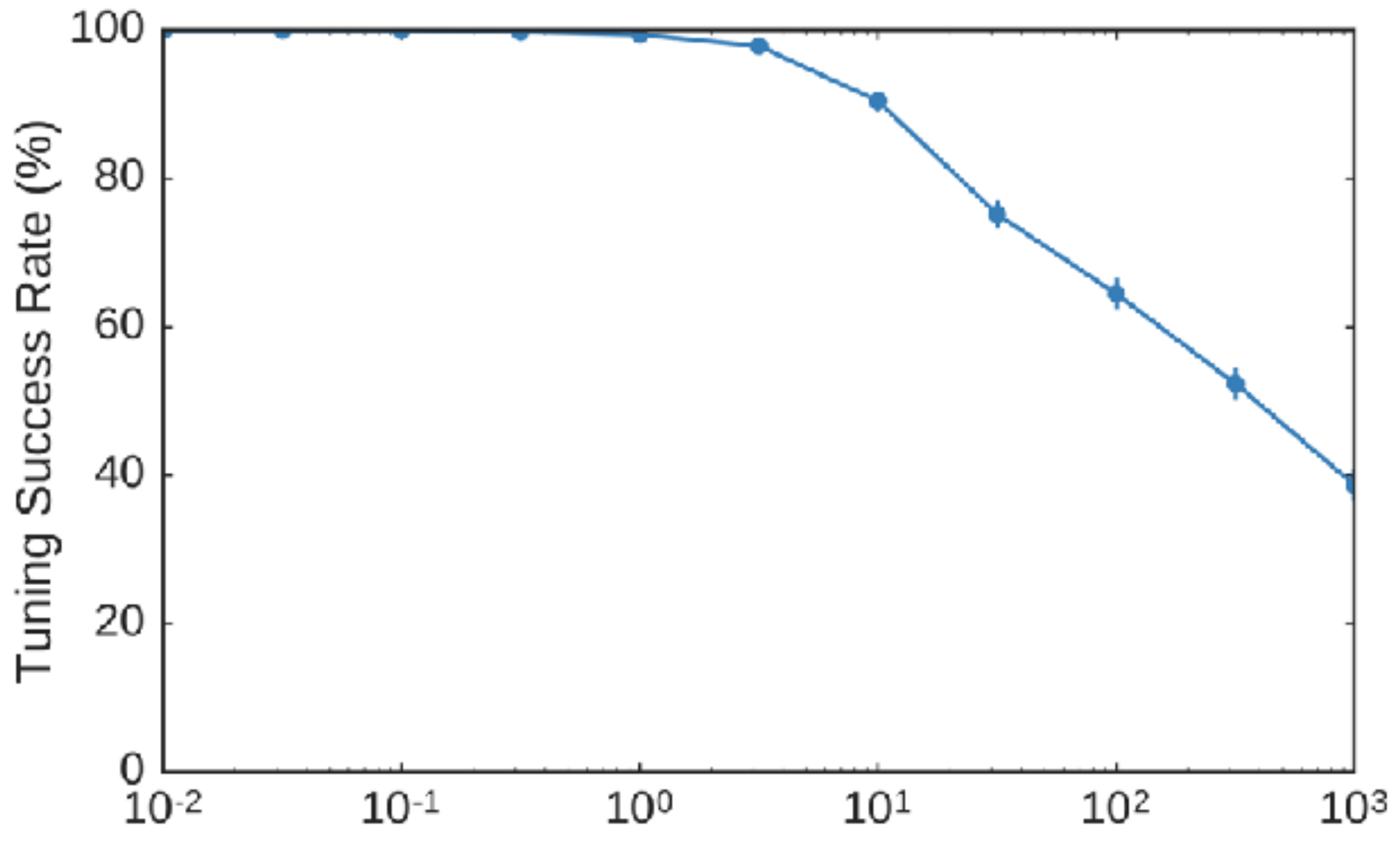
$$\mathcal{C} = (\text{desired target strain} \; - \; \text{free target strain})^2 = (\Delta - e_T)^2$$

- Adaptive DOF: edge stiffnesses/equil lengths/presence-absence…
- Physics requires minimization of elastic energy/free energy
- Physical DOF: node positions
- 1st optimzn: Minimize energy wrt physical DOF —required by physics!
- 2nd optimzn: Use gradient descent to minimize cost function wrt adaptive DOF

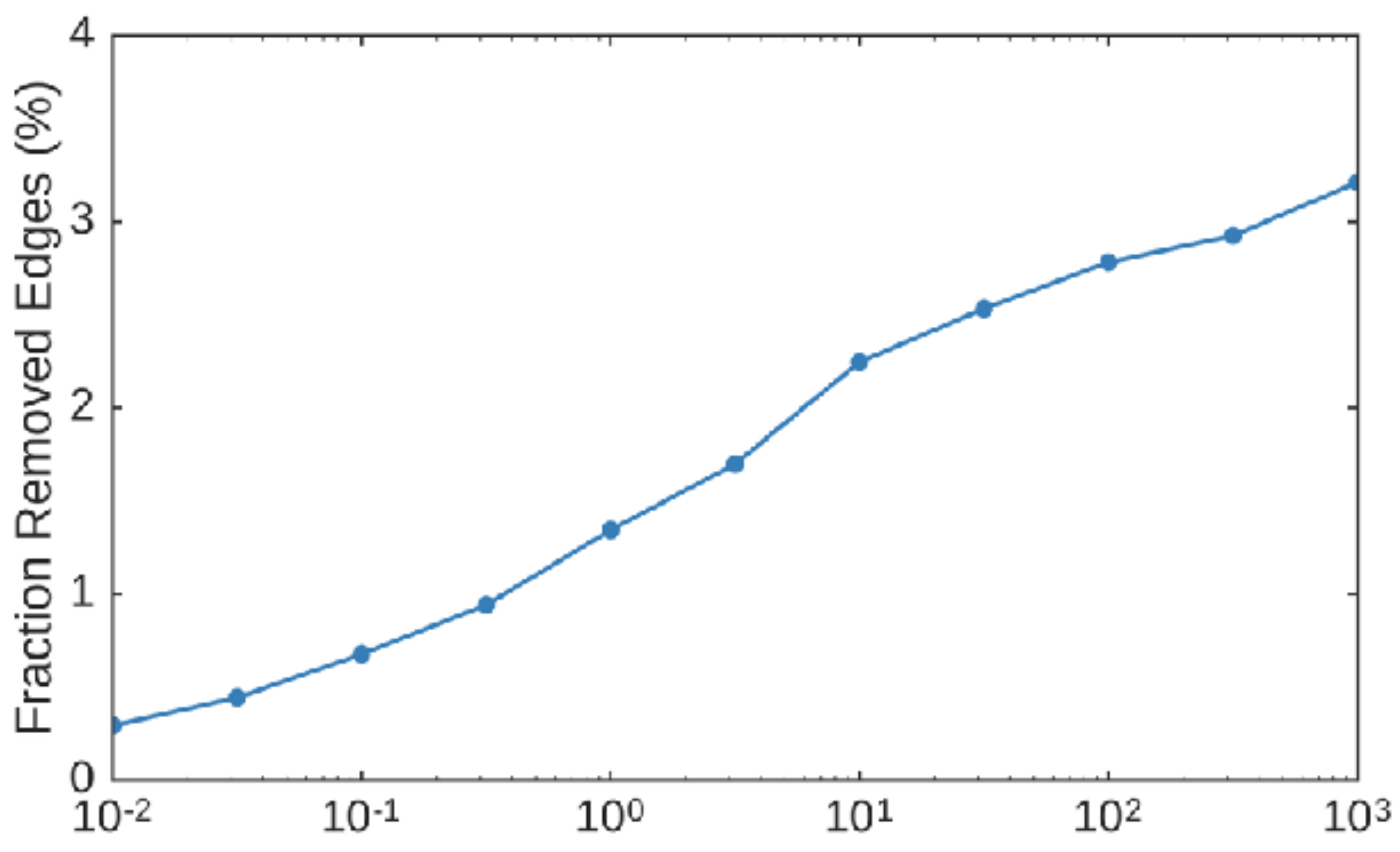


Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017

Double Optimization is Successful and Efficient



tuned target response Δ

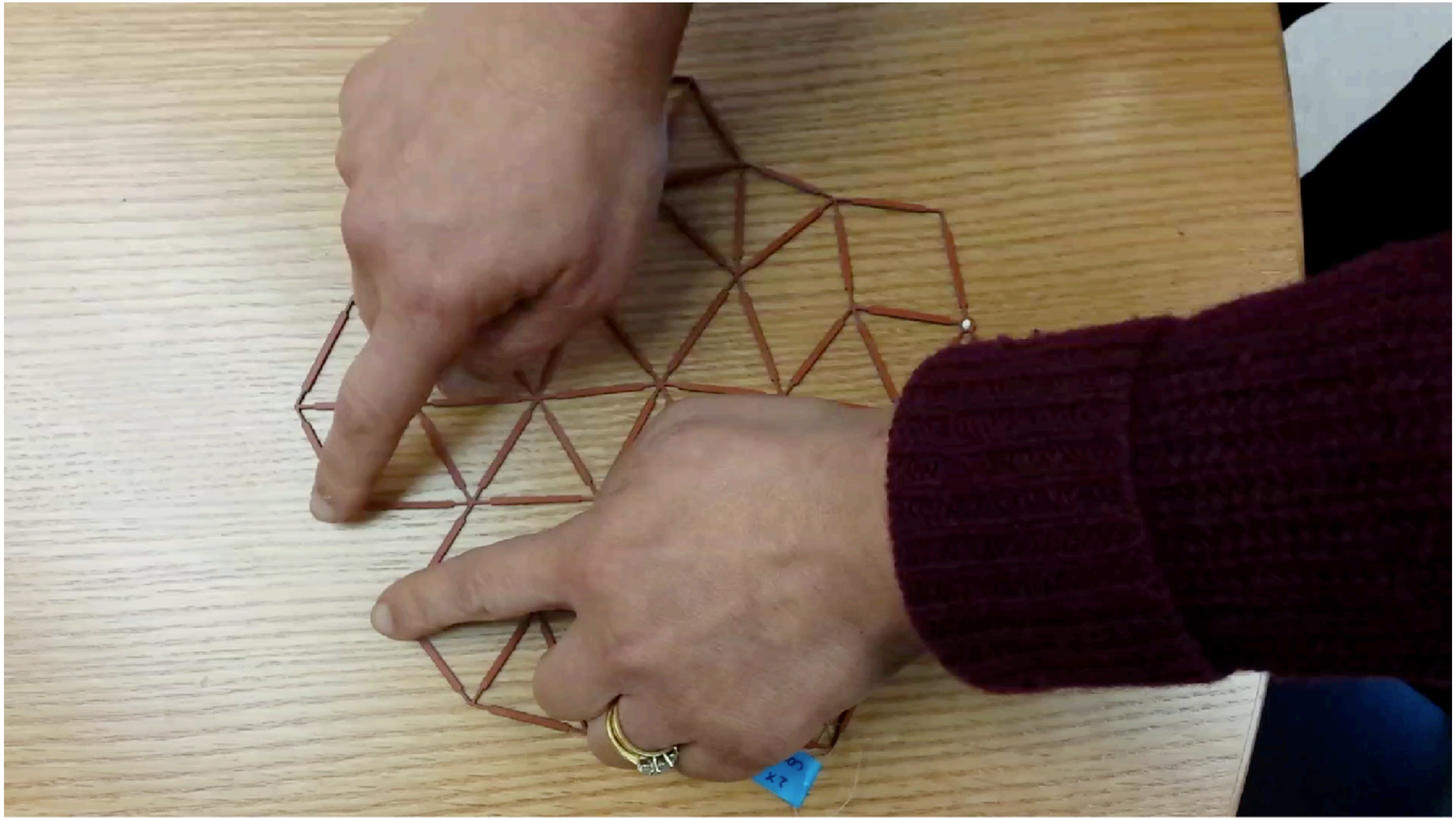


tuned target response Δ

- Can tune to Δ=1
  - with essentially 100% success
  - by pruning tiny fraction of bonds

Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017

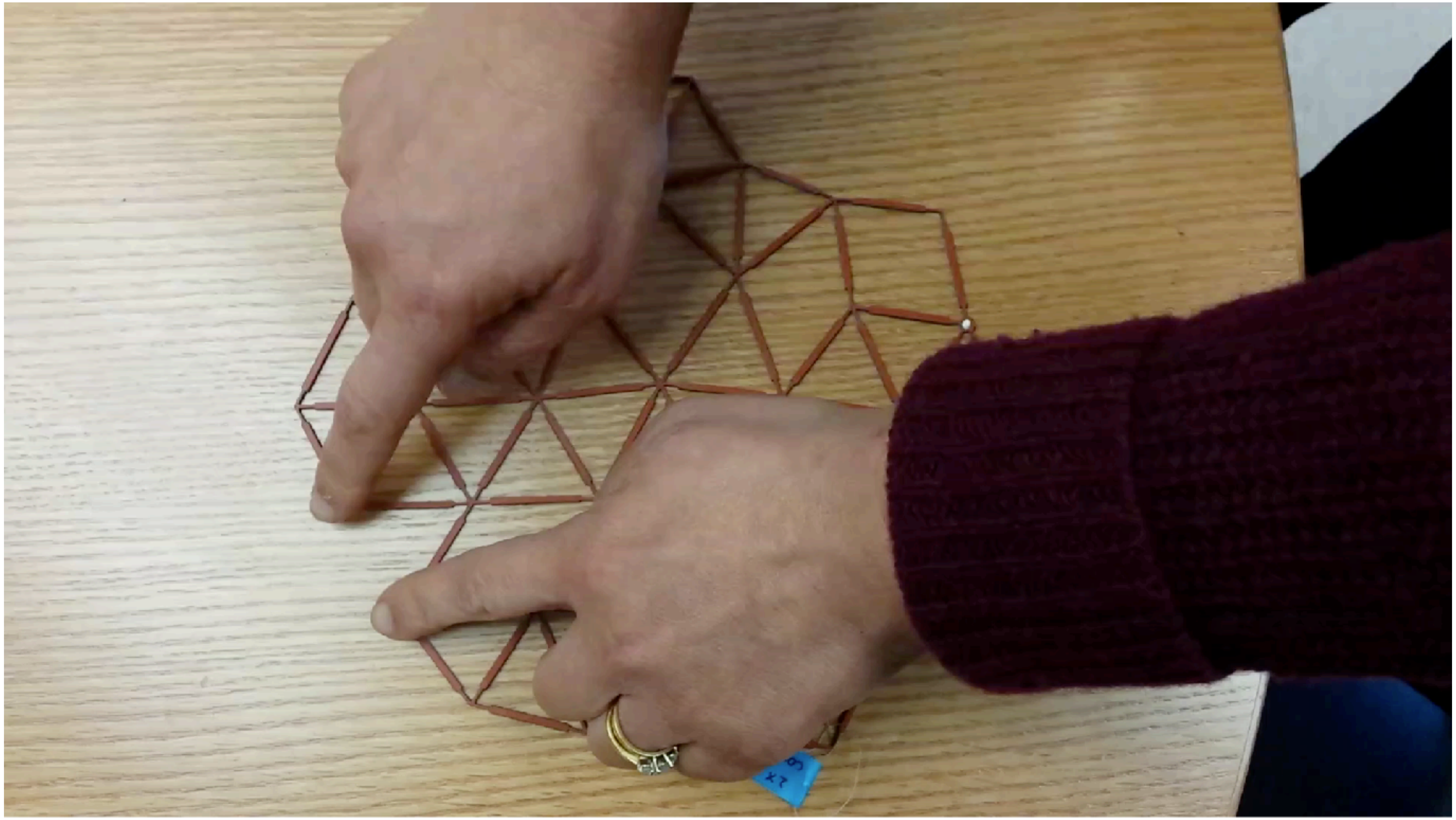
2D Lasercut Realizations In Real Life



Double optimization on computer yields
designs that are robust in the lab

Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017

2D Lasercut Realizations In Real Life



Double optimization on computer yields designs that are robust in the lab

Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017

3D-Printed Realizations In Real Life

Double optimization on computer yields designs that are robust in the lab

Rocks, Pashine, Goodrich, Bischofberger, Liu, Nagel PNAS 2017

Why is Adaptive Matter Framework Useful?

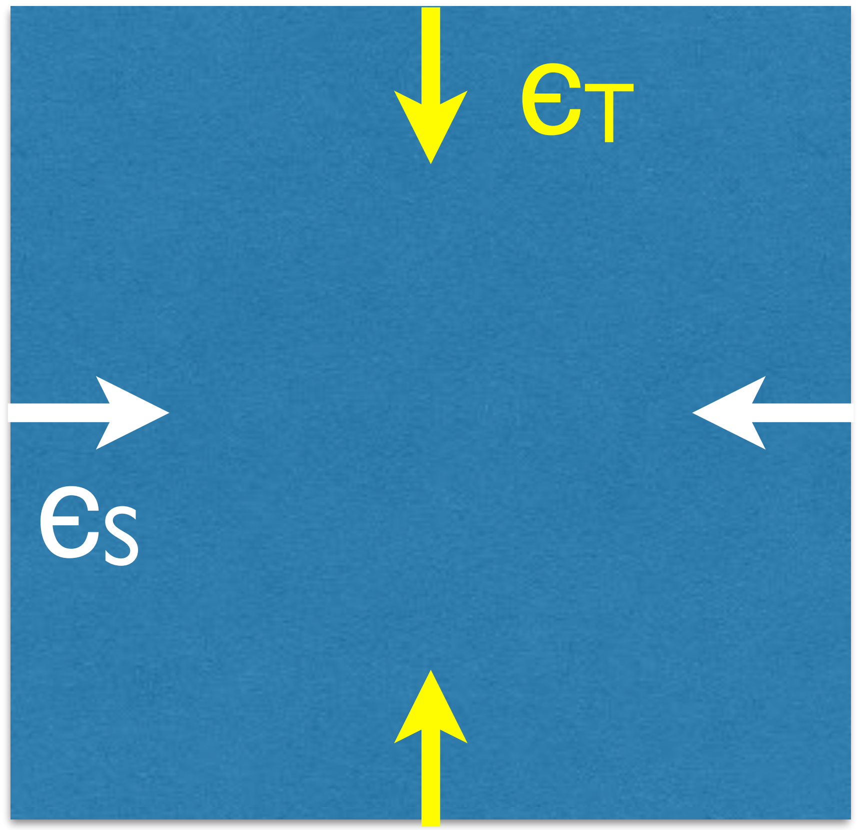
- Benefits of overparameterization
  - easy to reach a good solution
  - many good solutions
  - allows for stochasticity in the initial conditions to choose solution or subset of solutions
  - solutions typically generalizable, with many flattish directions in parameter space
    - robust to changes in most parameters
    - implies possibility of satisfying a different function (evolvability)
    - can remain good solution as inputs change
- Provides way of identifying (possibly quantifying) tradeoffs
  - how many parameters vs. metabolic cost of parameters
  - error vs. energy (metabolic) cost
  - error vs. robustness to noise or damage
  - error vs. adaptability to new tasks

Poisson's Ratio

Poisson's ratio

$$\nu = \frac{dB - 2G}{d(d-1)B + 2G}$$

Poisson's ratio
is an input/
output relation



$\epsilon_T = -\nu\epsilon_S$

Double optimzn
works here, too!

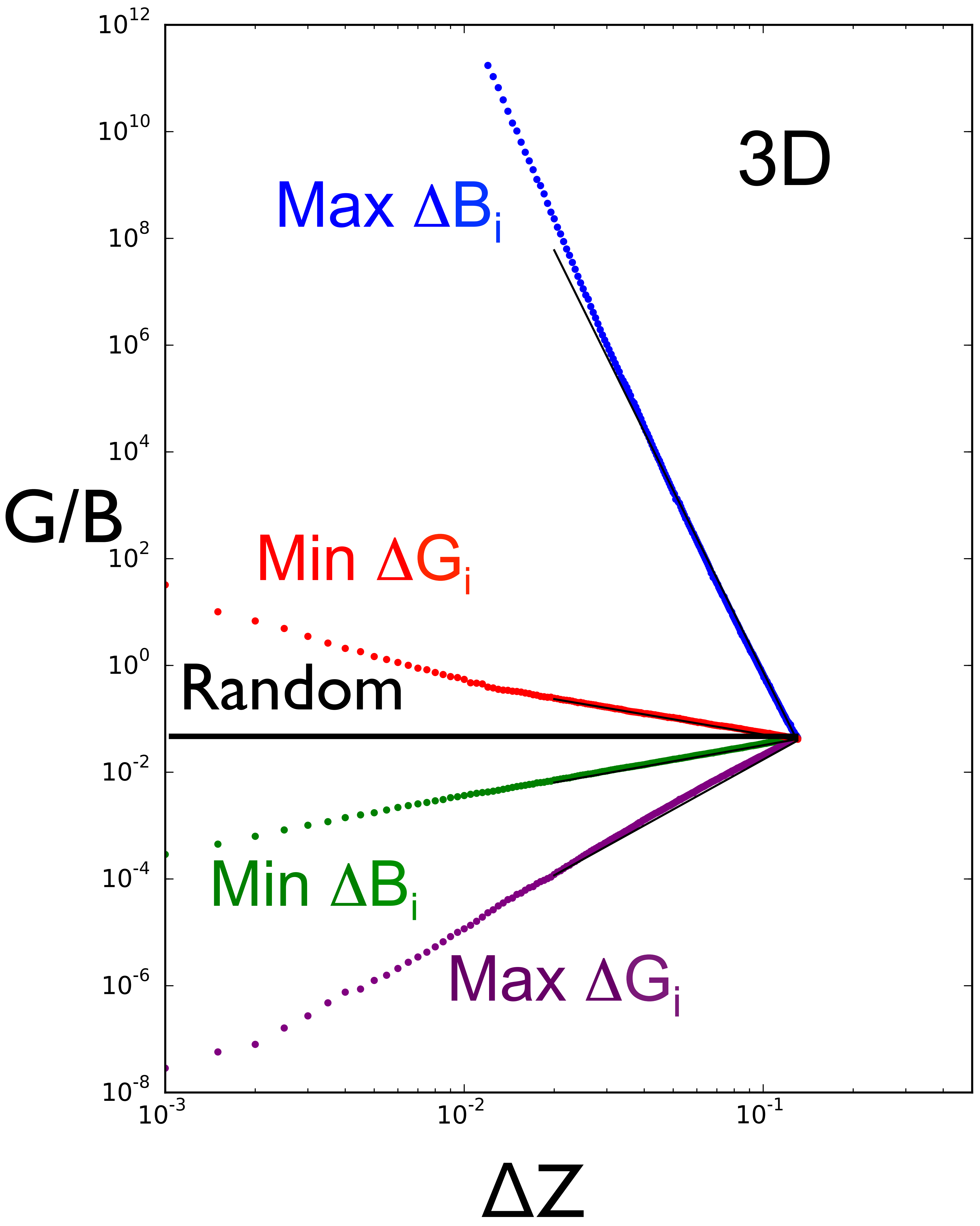
GO TO SLIDES

Goodrich/Liu/Nagel 2015
Hexner/Liu/Nagel 2018

Varying Correlations in Disorder Leads to Highly Malleable Behavior

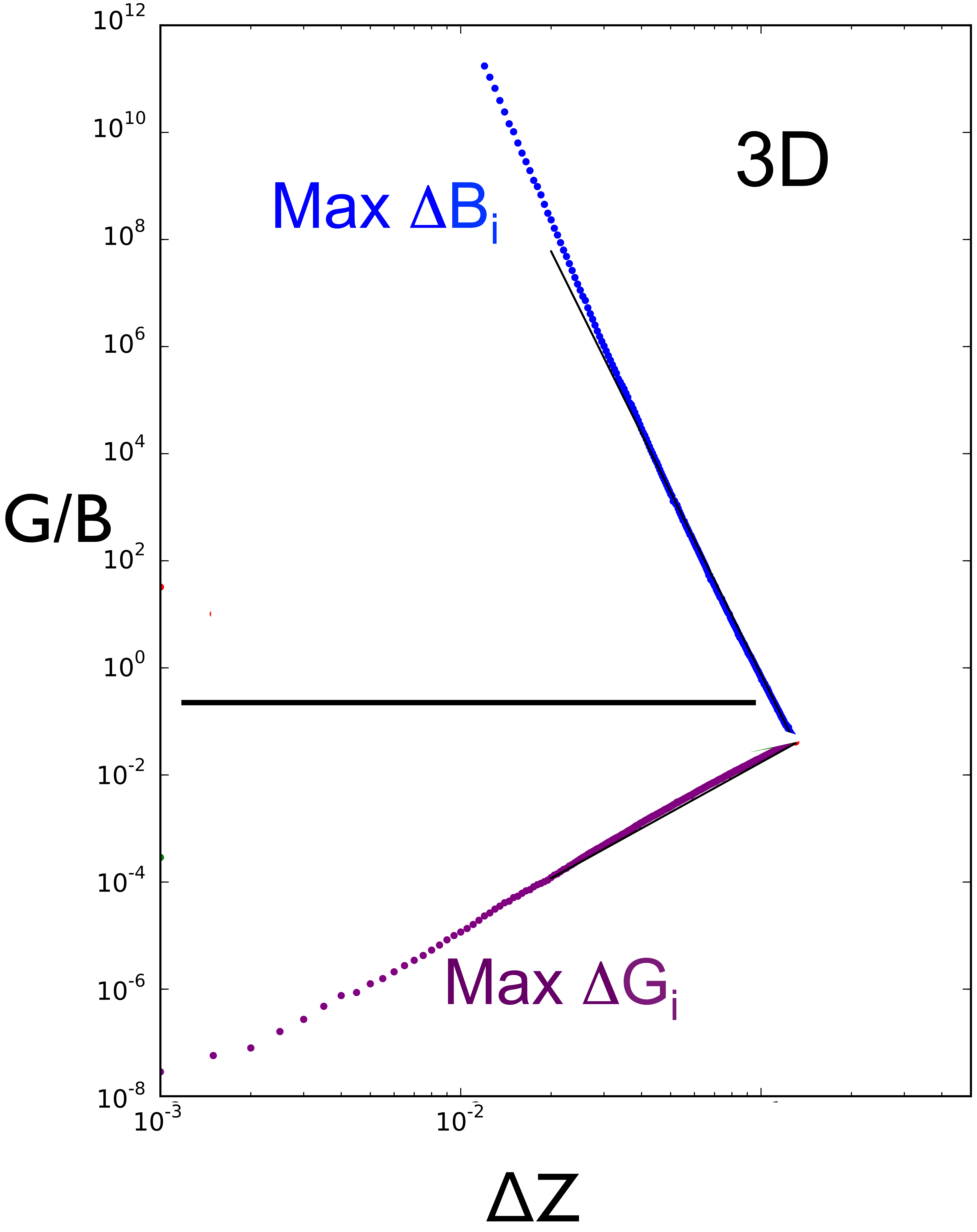
Poisson's ratio

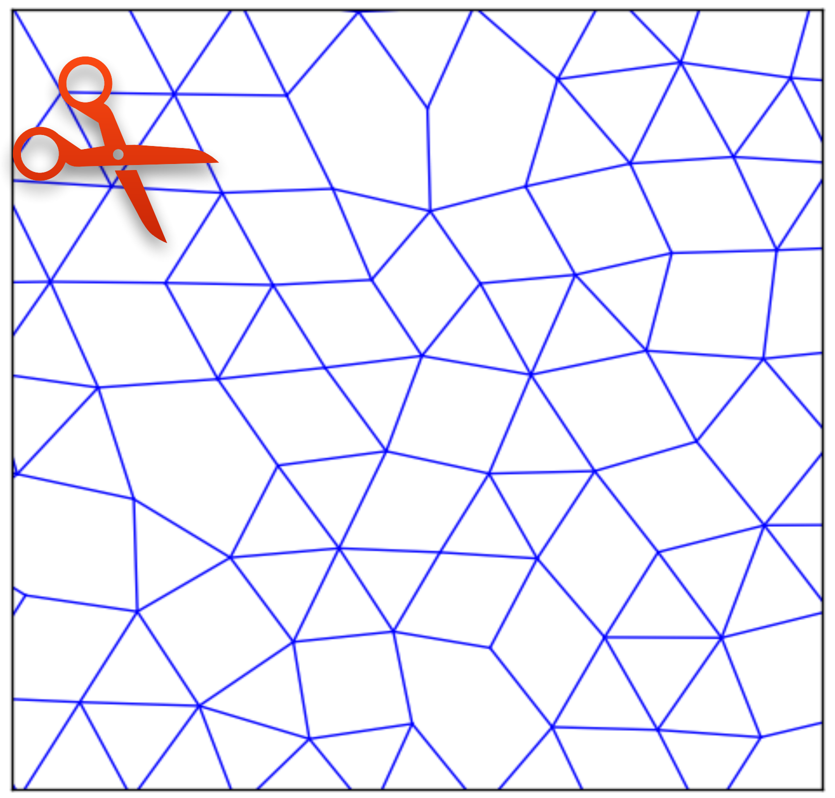
$$\nu = \frac{dB - 2G}{d(d-1)B + 2G}$$



Goodrich/Liu/Nagel PRL 2015
Hexner/Liu/Nagel Soft Matter 2018

Tuning Poisson Ratio with Double Optimization





- Prune max ΔB_i (gradient descent in B)

  $G/B \rightarrow \infty$, $\nu \rightarrow -1$, auxetic
- Prune max ΔG_i (gradient descent in G)

  $G/B \rightarrow 0$, $\nu \rightarrow d/2$

Goodrich PRL 2015
Hexner PRE 2018

Drawbacks of Global Gradient Descent for Double Optimization

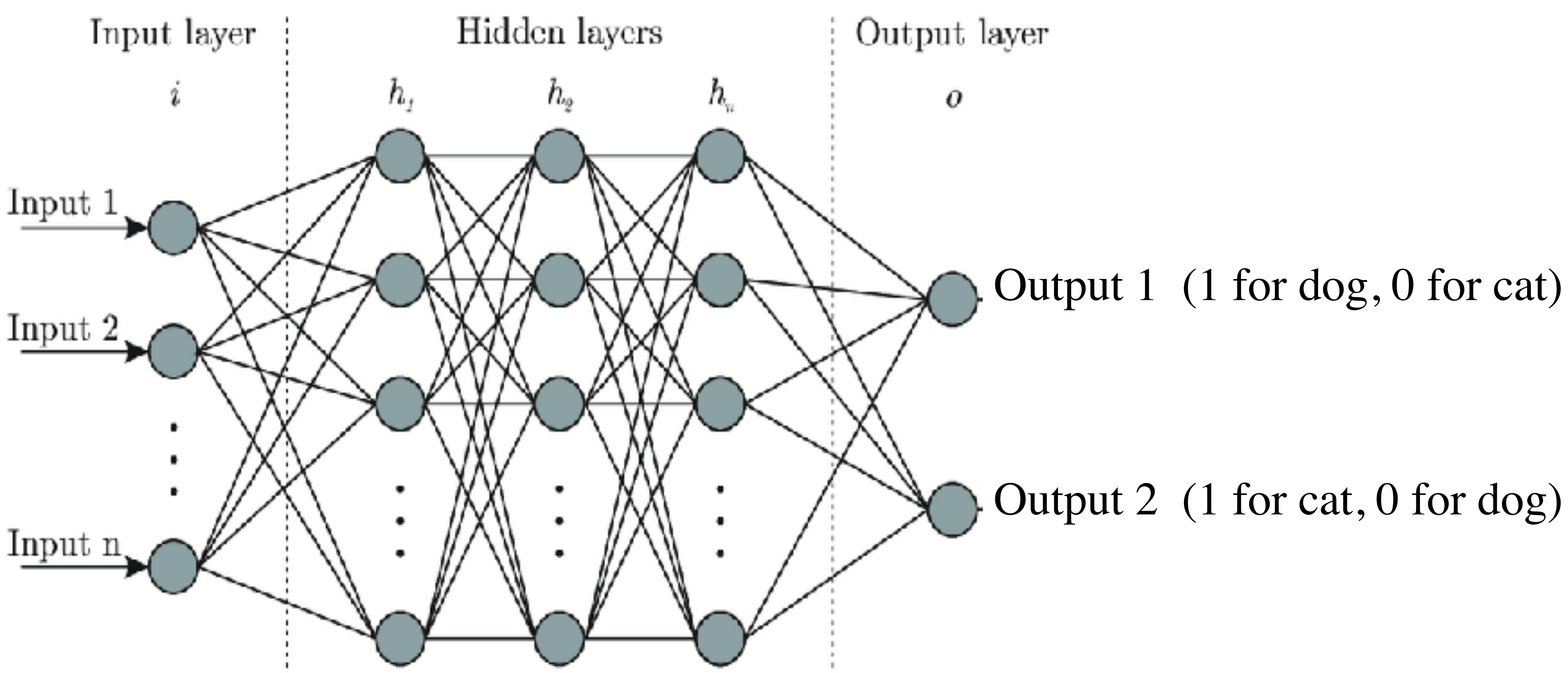
- Not scalable
  - Requires precise microscopic knowledge of system to evaluate cost function and gradient
  - Requires processor to calculate direction of gradient descent
  - requires ability to modify individual parameters
  - Design by minimizing cost function is inherently global and therefore requires processor

Drawbacks of Global Gradient Descent for Double Optimization

- Not scalable
  - Requires precise microscopic knowledge of system to evaluate cost function and gradient
  - Requires processor to calculate direction of gradient descent
  - requires ability to modify individual parameters
  - Design by minimizing cost function is inherently global and therefore requires processor

Can systems learn input/output relations
on their own, without a processor?

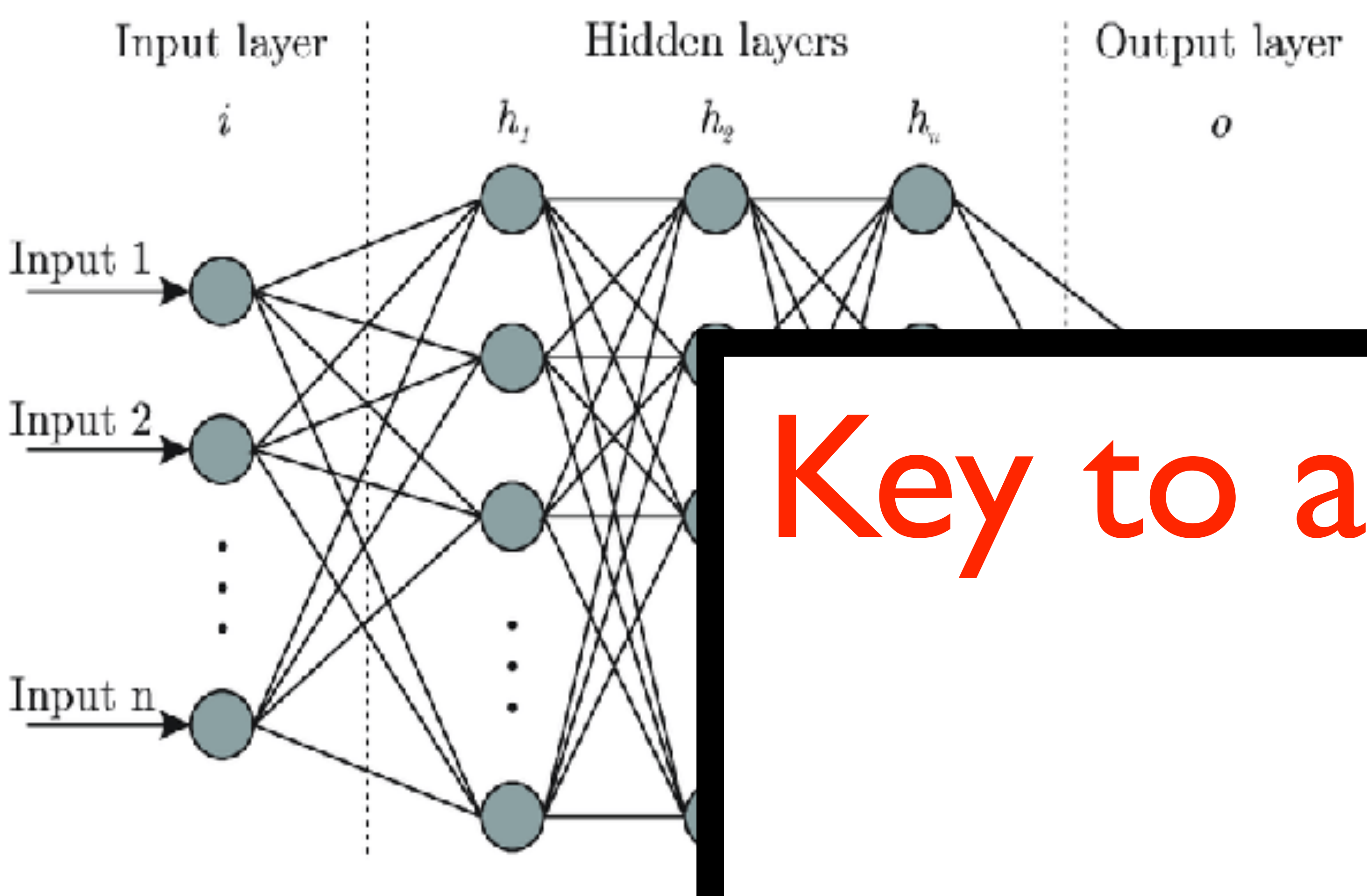
Digital vs. Real Neural Networks

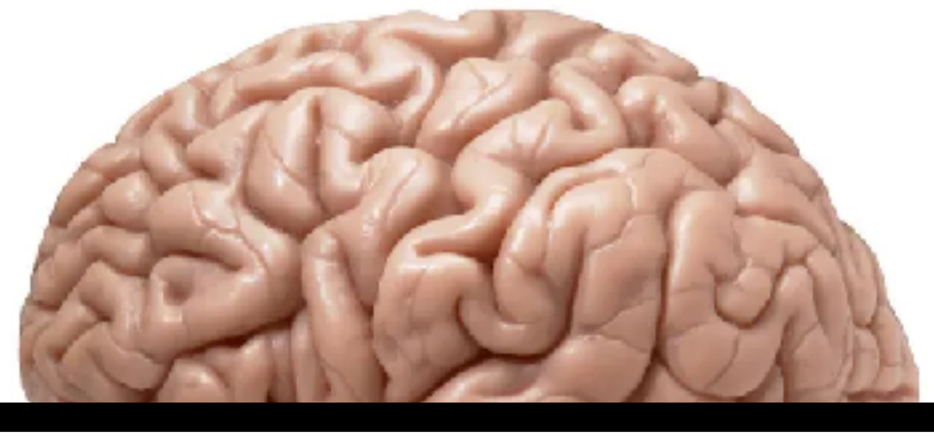


- Gradient descent on cost function requires knowing all network details
- Requires processor
- Costs much energy
  - Each ChatGPT text query ~ 200 kJ
  - Image generation ~10,000 kJ

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

Digital vs. Real Neural Networks





Key to adjusting adaptive DOF in real time:

local rules

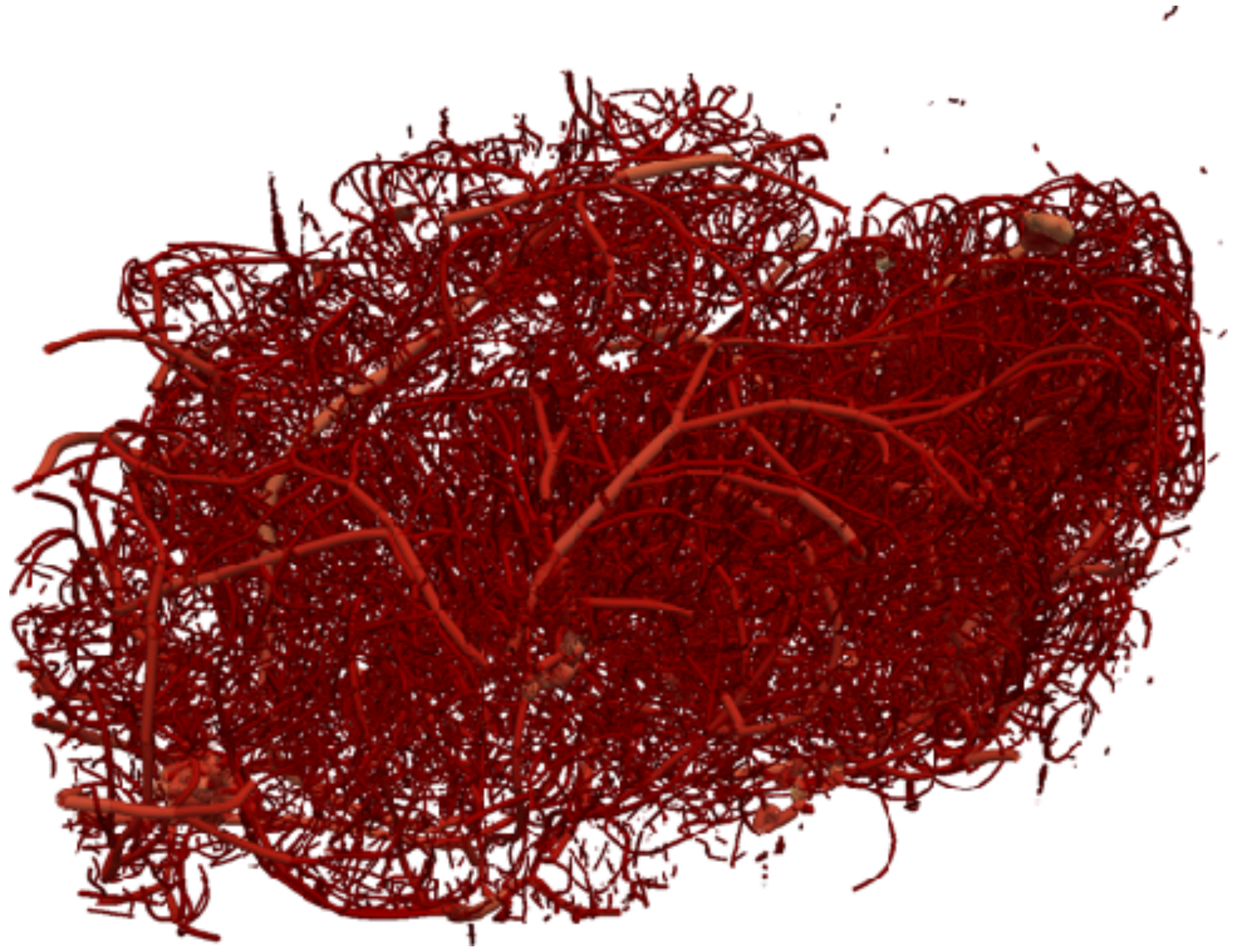
Stern, Murugan ARCMP (2023)

- Gradient descent on cost function requires knowing all network details
- Requires processor
- Costs much energy
  - Each ChatGPT text query ~ 200 kJ
  - 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

Brain Vasculature Has Desired Input/Output Relations

brain vasculature as flow network



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

Katifori

- 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

Rocks, Ronellenfitsch, Liu, Nagel, Katifori PNAS 2019
Rocks, Liu, Katifori PRR 2020
Rocks, Liu, Katifori PRL 2021
Stern, Hexner, Rocks, Liu PRX (2021)

Brain Vasculature

- Function: deliver desired pressure drop/current to designated location
- Parameters:
  - Level of description: low Re flow network of pipes (edges) connecting nodes
  - parameters: edge conductances. —extensive
- Process:
  - Involves astrocytes and reaction to low O2, but I don't know much about it
  - Has to use local rules
    - What are they?
    - How do they lead to the function?

Simple Models for Allosteric Proteins/Brain Vasculature

Mechanical: central force spring networks

$$E = \frac{1}{2}\sum_{\langle ij \rangle} k_{ij}\left(\hat{b}_{ij}\cdot(\vec{u}_i - \vec{u}_j)\right)^2 = \frac{1}{2}\vec{u}H\vec{u}$$

$$H\vec{u} = \vec{f}$$

nodes adjust positions so that network is at energy minimum

Flow networks with conductance on each edge:

$$P = \frac{1}{2}\sum_{\langle ij \rangle} k_{ij}(p_i - p_j)^2 = \frac{1}{2}\vec{p}L\vec{p}$$

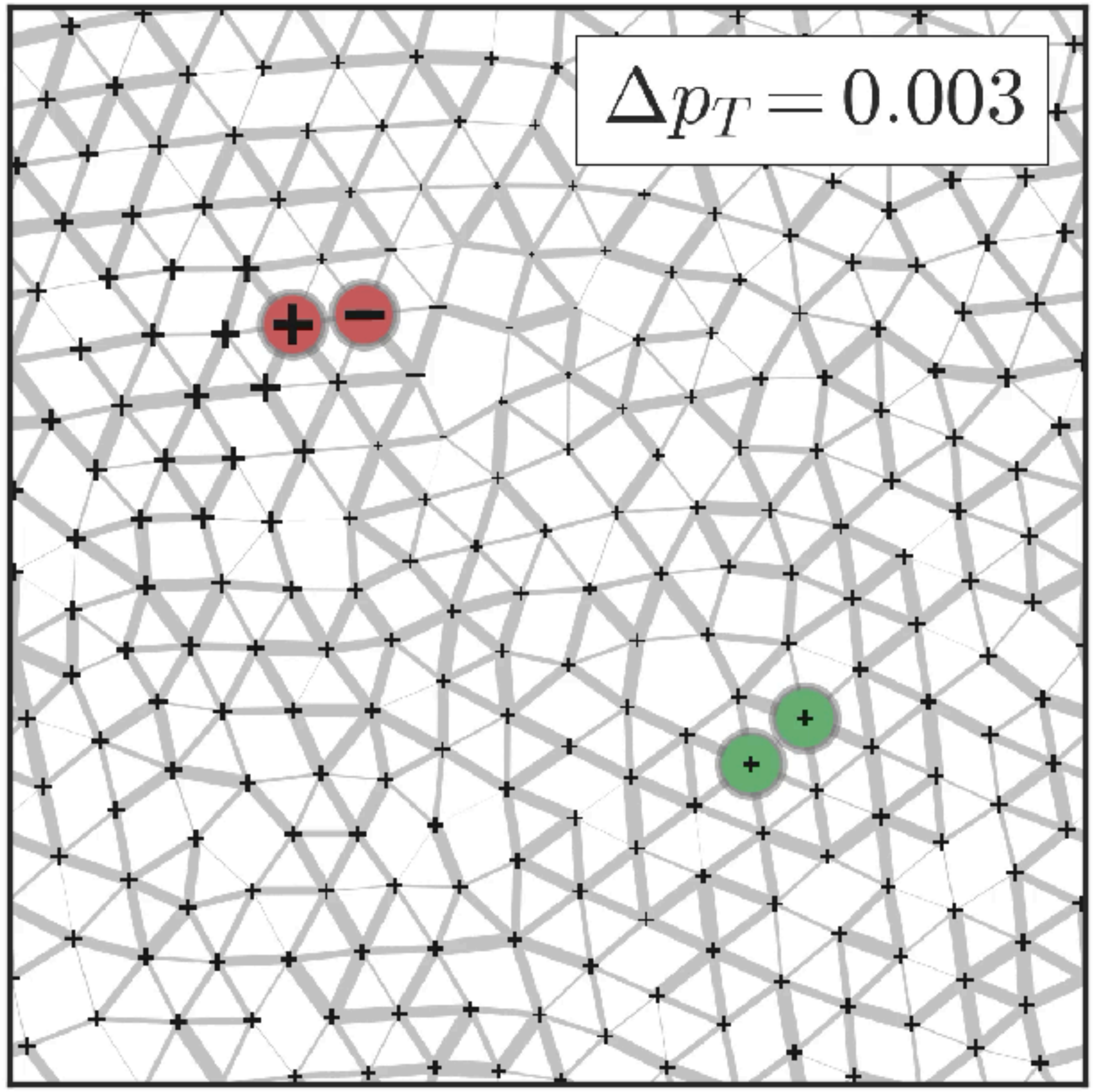
$$L\vec{p} = \vec{j}$$

flows on edges adjust so that network is at power loss minimum

- Flow networks are 1D mechanical networks without an embedding $\hat{b}_{ij}$ in space
- Low-Re-flow/linear-electrical-resistor networks are identical
- Pressure $\leftrightarrow$ Voltage

Double Optimization Works!

- Cost function: $\mathcal{C} = (\Delta - \Delta p_T)^2$
- Adaptive DOF: edge conductances
- Physics requires minimization of dissipated power
- Physical DOF: node pressures
- 1st optimzn: Minimize power wrt physical DOF —required by physics!
- 2nd optimzn: Use gradient descent to minimize cost function wrt adaptive DOF



Rocks, Ronellenfitsch, Liu, Nagel, Katifori PNAS 2019

More Examples: Physical Systems That Maintain High-Dim Homeostasis

- Homeostasis is not typical CS problem but are desired response to stimuli
  - many inputs (high-dimensional) that vary a lot
  - relatively few outputs (low-dimensional) that vary very little

- Example: rigidity homeostasis
  - Epithelial amnioserosa during dorsal closure during Drosophila development
  - Actin cortex

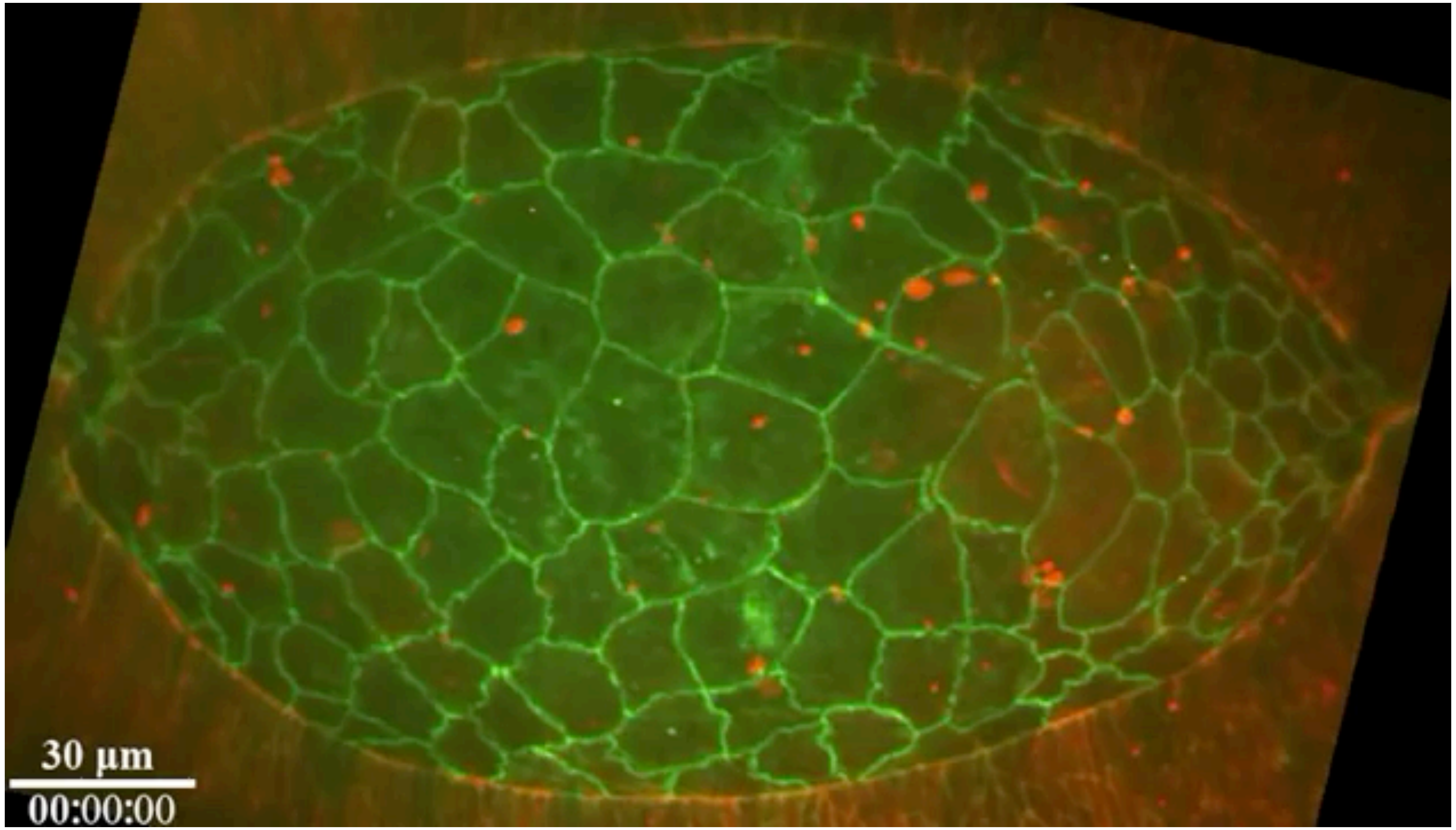
GO TO SLIDES

Maintaining Rigidity During Dorsal Closure

Rigidity under trying circumstances

Amnioserosa during dorsal closure of *Drosophila melangaster*

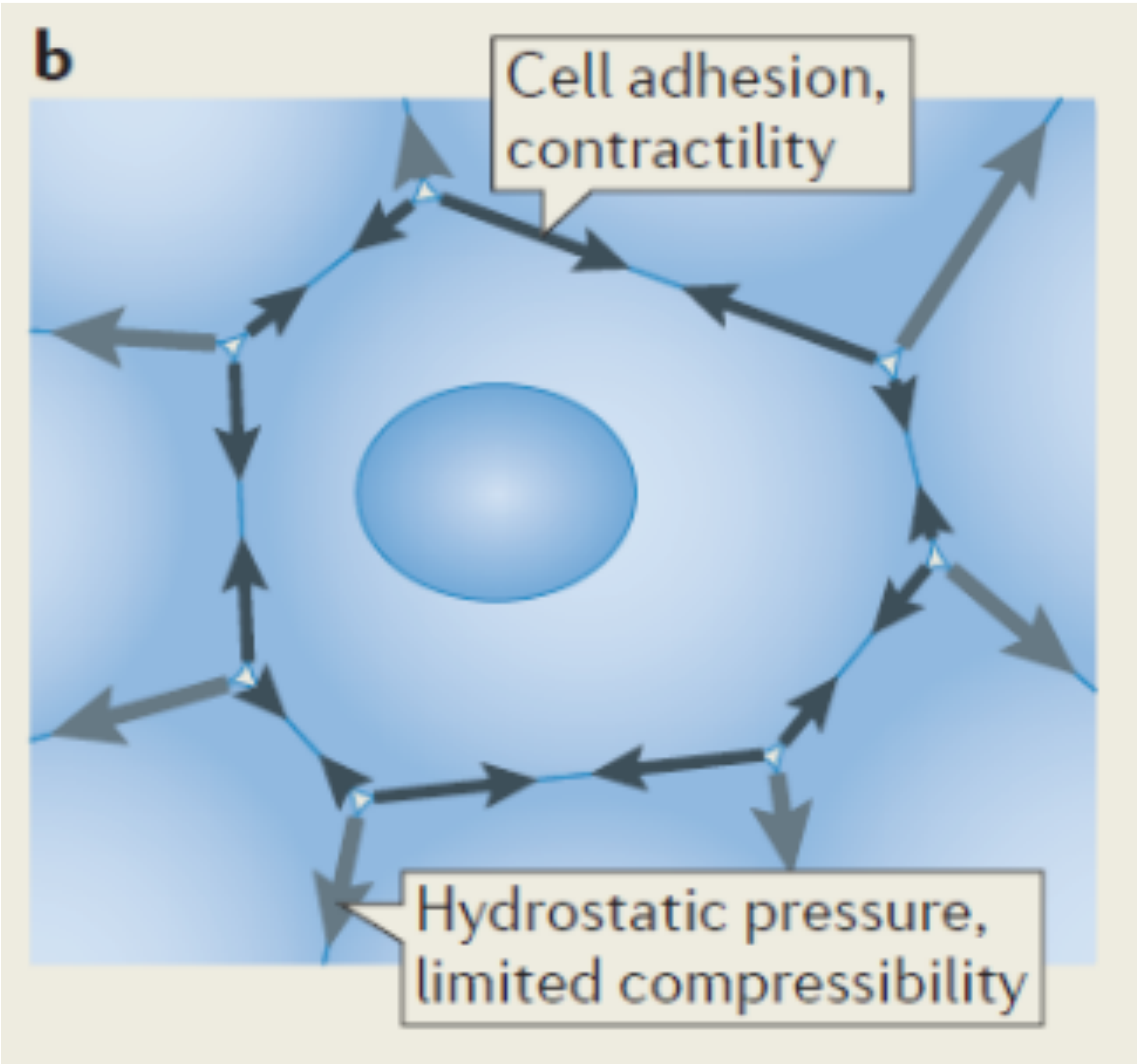
Tah/Haertter/Crawford/Kiehart/Schmidt/Liu



D. Kiehart

- Sheet of tissue one cell thick
- Covers dorsal opening between two cell sheets
- No cell rearrangements (T1 events) during entire process! AS is rigid
- Must adapt to extreme shrinking of tissue to remain rigid

Vertex Model of 2D (Epithelial) Tissue



$$E = k_p \sum_{cells} (p_i - p_0)^2 + k_a \sum_{cells} (a_i - a_0)^2$$

Cell's optimal perimeter

actomyosin contractility
adhesion/cortical tension

Cell's optimal area

cell incompressibility
preferred cell height

Key variables:

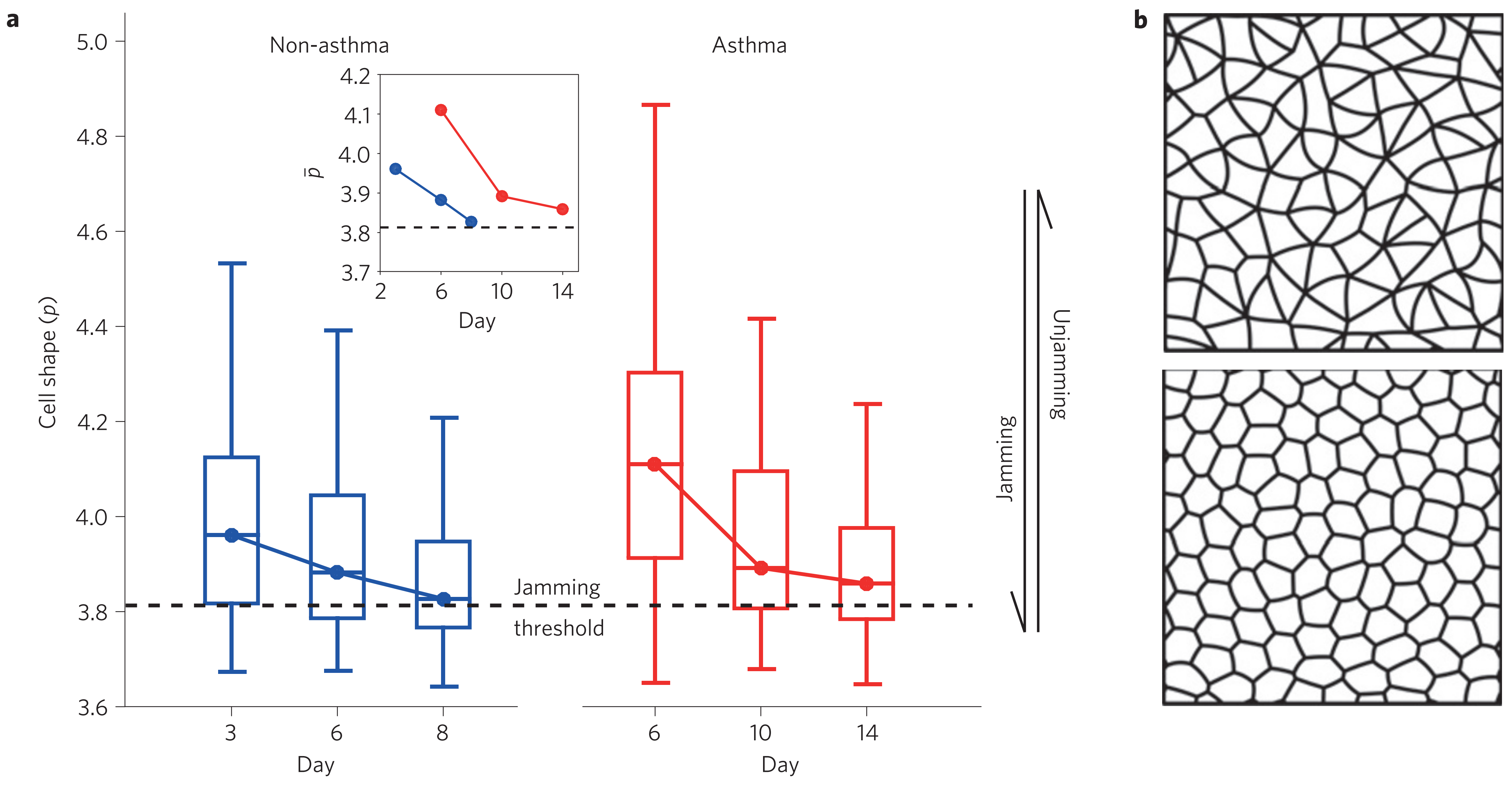
$q_0 = p_0/\sqrt{a_0}$ shape index

$r = k_a a_0/k_p$ stiffness ratio

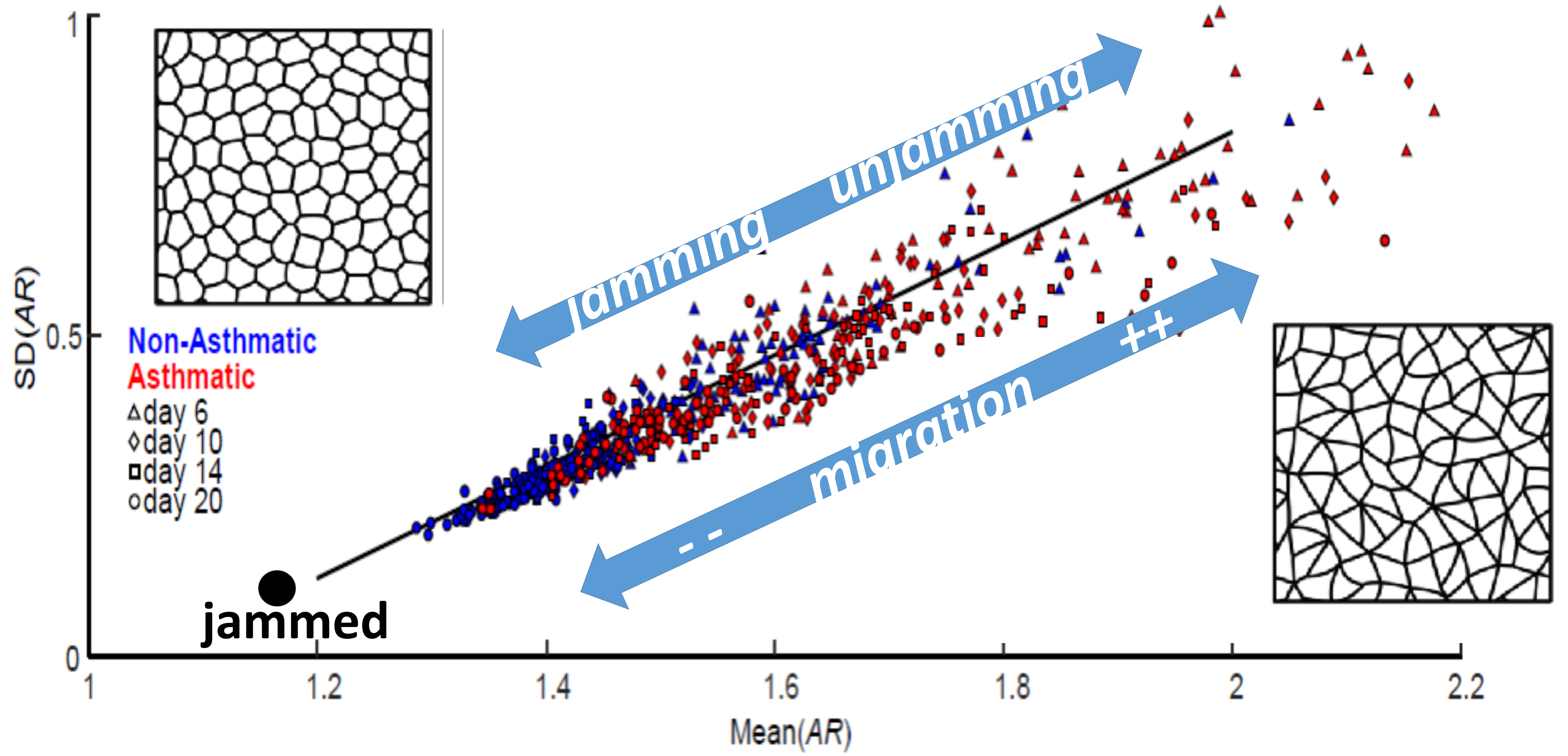
$$\tilde{E} = \sum_{cells} (q_i - q_0)^2 + r \sum_{cells} (\tilde{a}_i - 1)^2$$

T. Nagai, H. Honda, Philos. Mag. B 81, 699 (2001)
Hufnagel et al, PNAS vol. 104 (10) pp. 3835 (2007)
Farhadifar et al, Current Biology (2007)
Jülicher et al Phys. Rep. (2007)
Hilgenfeldt et al, PNAS 105 3 907–911 (2008)
Manning et al, PNAS (2010)
Staple et al EPJE 33 (2) 117 (2010)
Chiou et al PLOS Comp Bio 8 (5) e1002512 (2012)

Predicts Rigidity Transition Quantitatively!



Shape variation



Cell shape

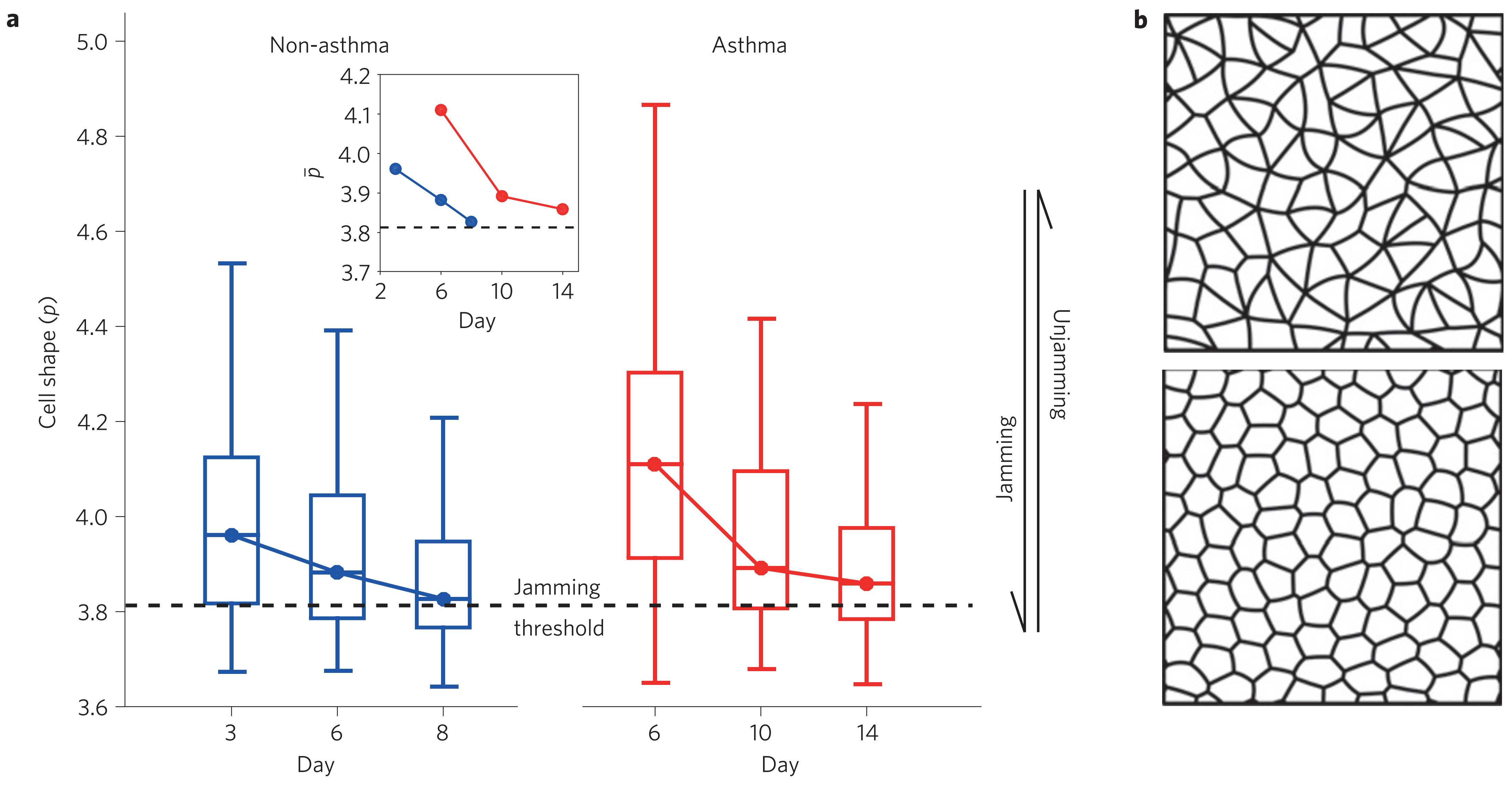
Park *et al Nat Phys* 2015

Critical $q_0^* \approx 3.81$

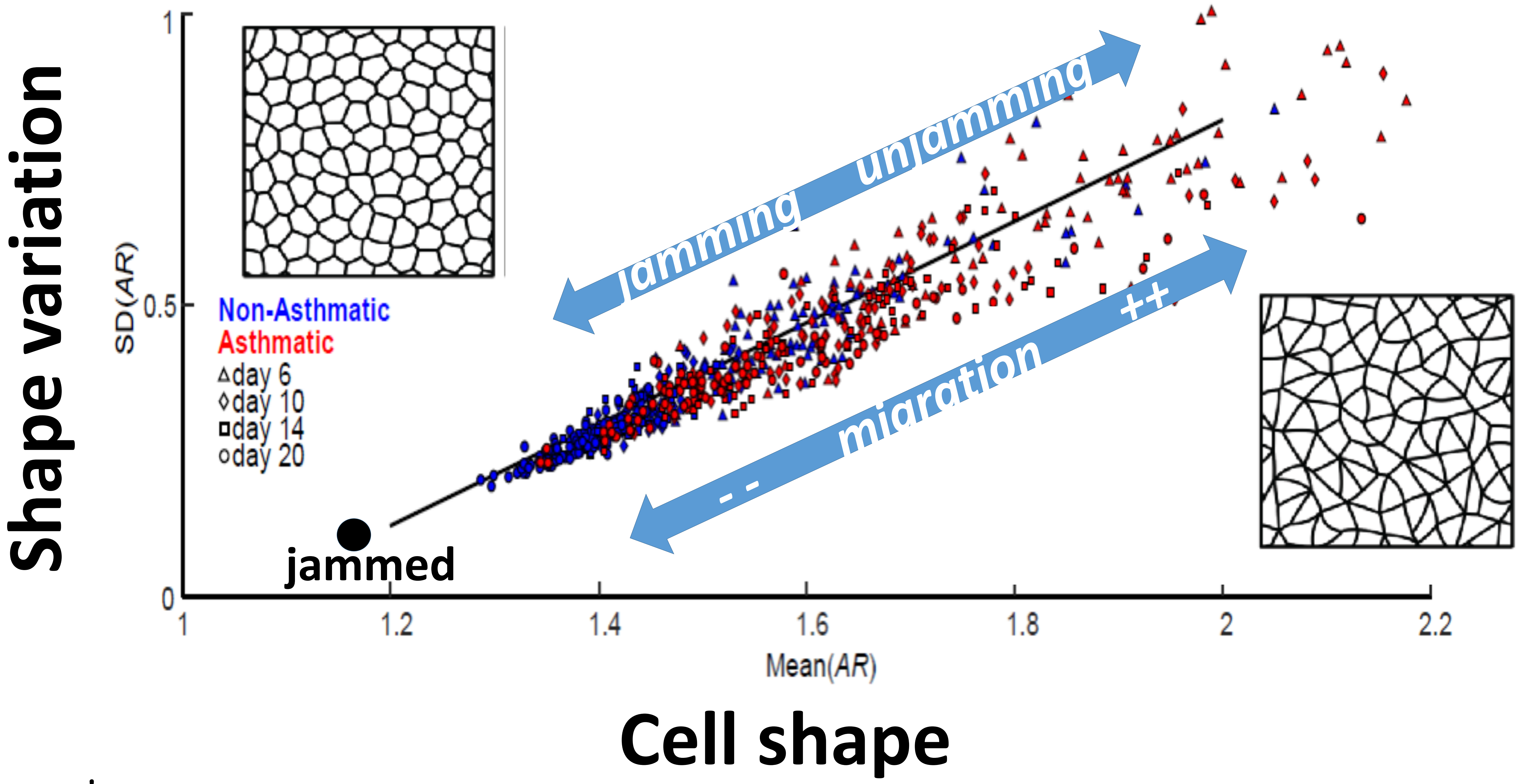
Vertex model accounts not only for cell *shape* but also shape *variation*

Park *et al Nat Phys* 2015
Atia *et al.*, *Nat Phys* 2018
Bi/Lopez/Schwarz/Manning, *Nat Phys* 2015

Predicts Rigidity Transition Quantitatively!



Park *et al Nat Phys* 2015



Vertex model accounts not only for cell *shape* but also shape *variation*

Critical $q_0^* \approx 3.81$

Park *et al Nat Phys* 2015
Atia *et al.*, *Nat Phys* 2018
Bi/Lopez/Schwarz/Manning, *Nat Phys* 2015

BUT

- Amnioserosa during dorsal closure has q ≅4.24 increasing to 4.36 and is always rigid!
- How?

Adaptive Degrees of Freedom (DOF)

- Recall vertex model

$$E = k_p \sum_{cells} (p_i - p_0)^2 + k_a \sum_{cells} (a_i - a_0)^2$$

- This energy is minimized by adjusting vertex positions as physical DOFs
- Let each cell have different preferred perimeter p_{0i} as parameters
- Double optimization: adaptive DOF couple physical DOF so that energy must be re-minimized as learning DOF adjust

GO TO SLIDES

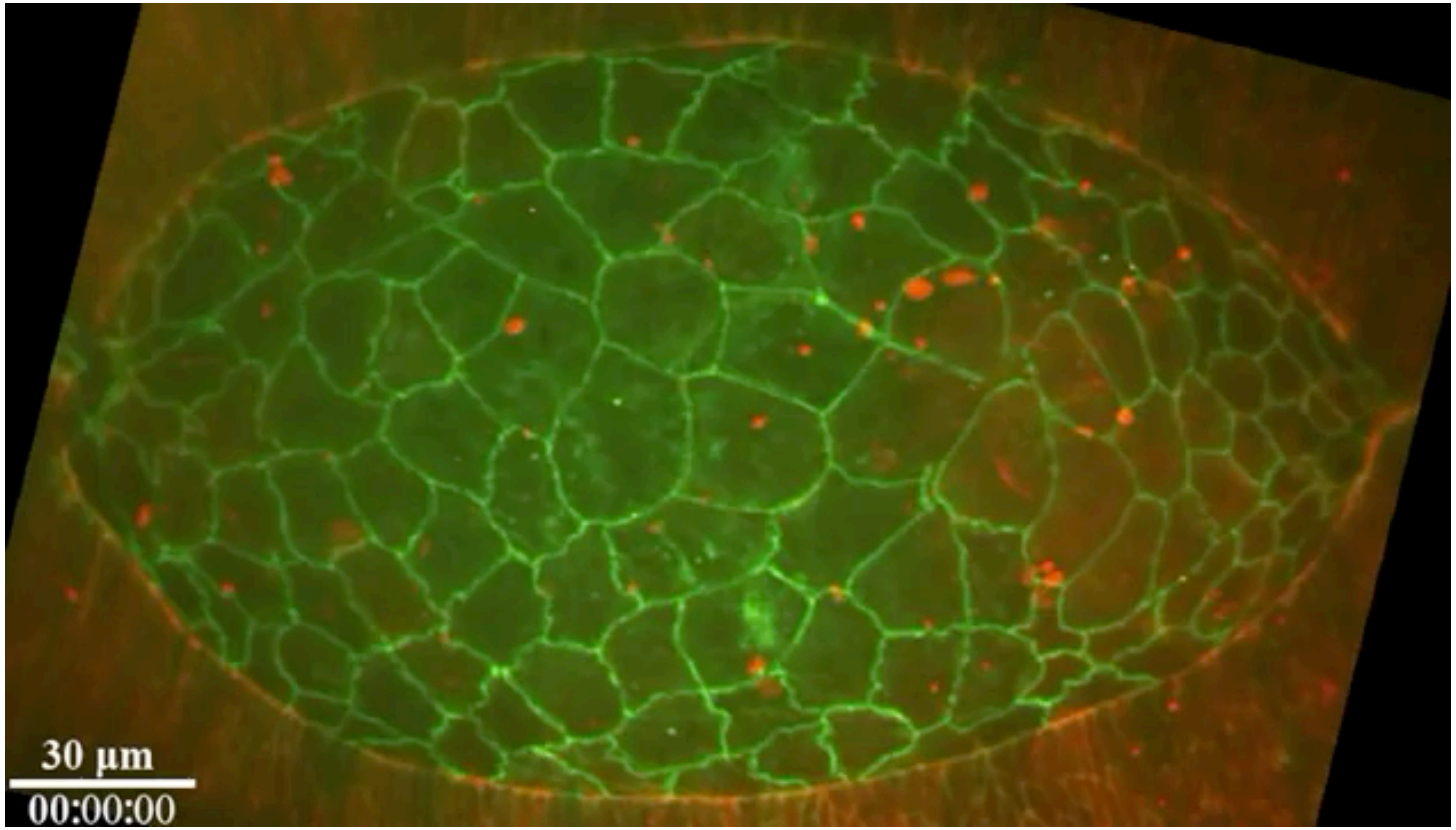
Tah, Haertter, Crawford, Kiehart, Schmidt, Liu arXiv (2023)

Maintaining Rigidity During Dorsal Closure

Rigidity under trying circumstances

Amnioserosa during dorsal closure of *Drosophila melangaster*

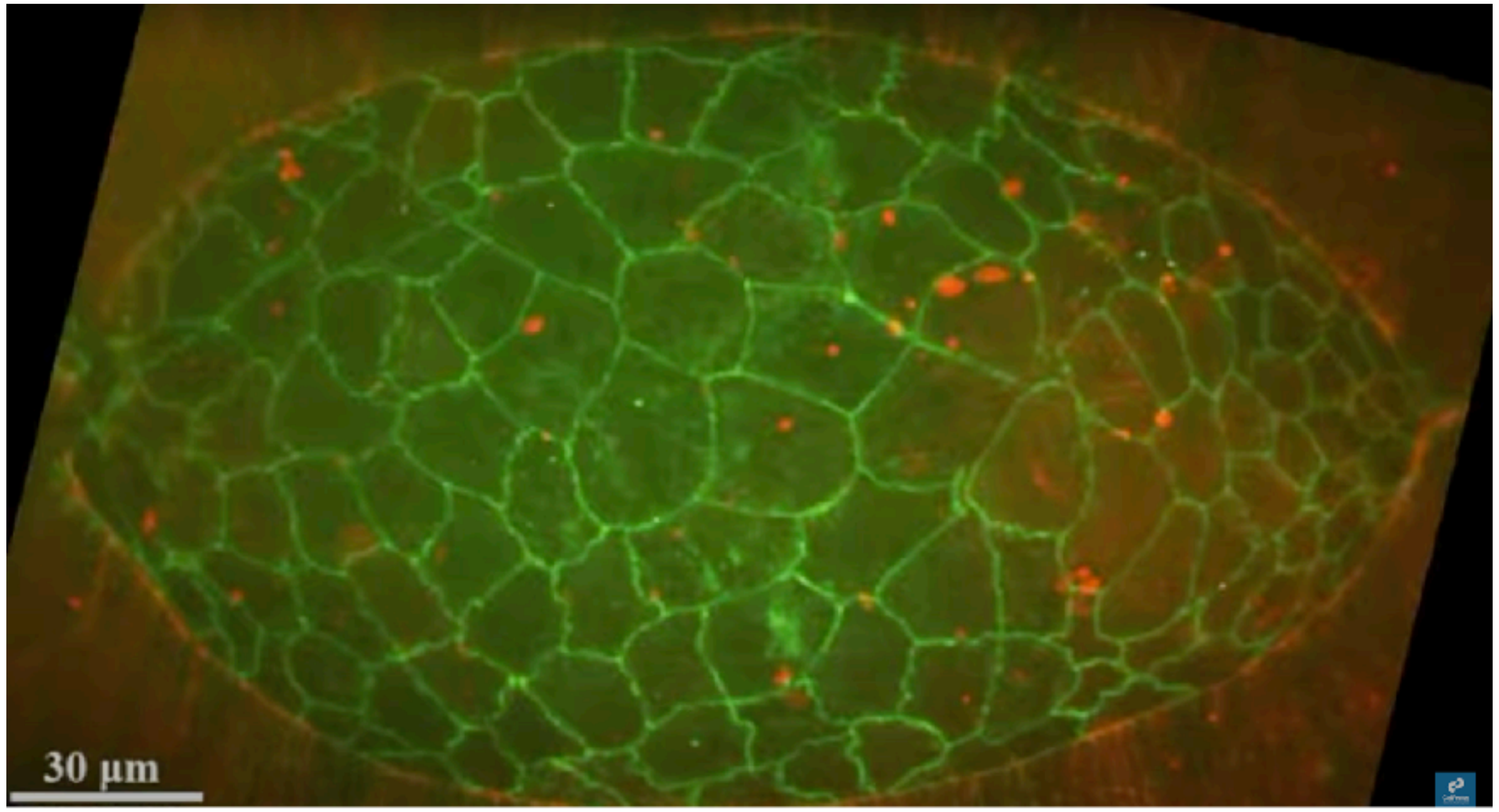
Tah/Haertter/Crawford/Kiehart/Schmidt/Liu

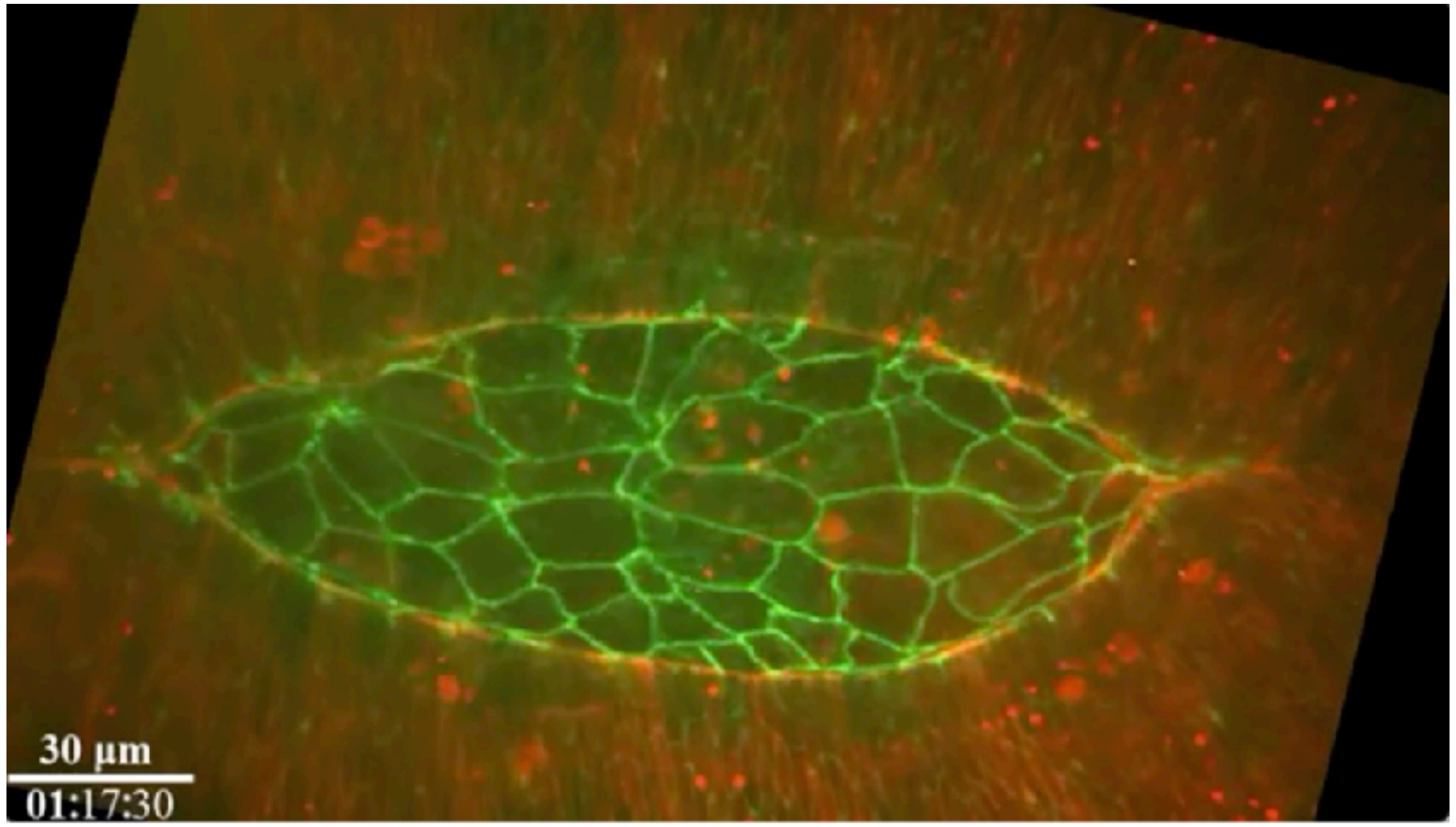


D. Kiehart

- Sheet of tissue one cell thick
- Covers dorsal opening between two cell sheets
- No cell rearrangements (T1 events) during entire process! AS is rigid
- Must adapt to extreme shrinking of tissue to remain rigid

What Local Rules?



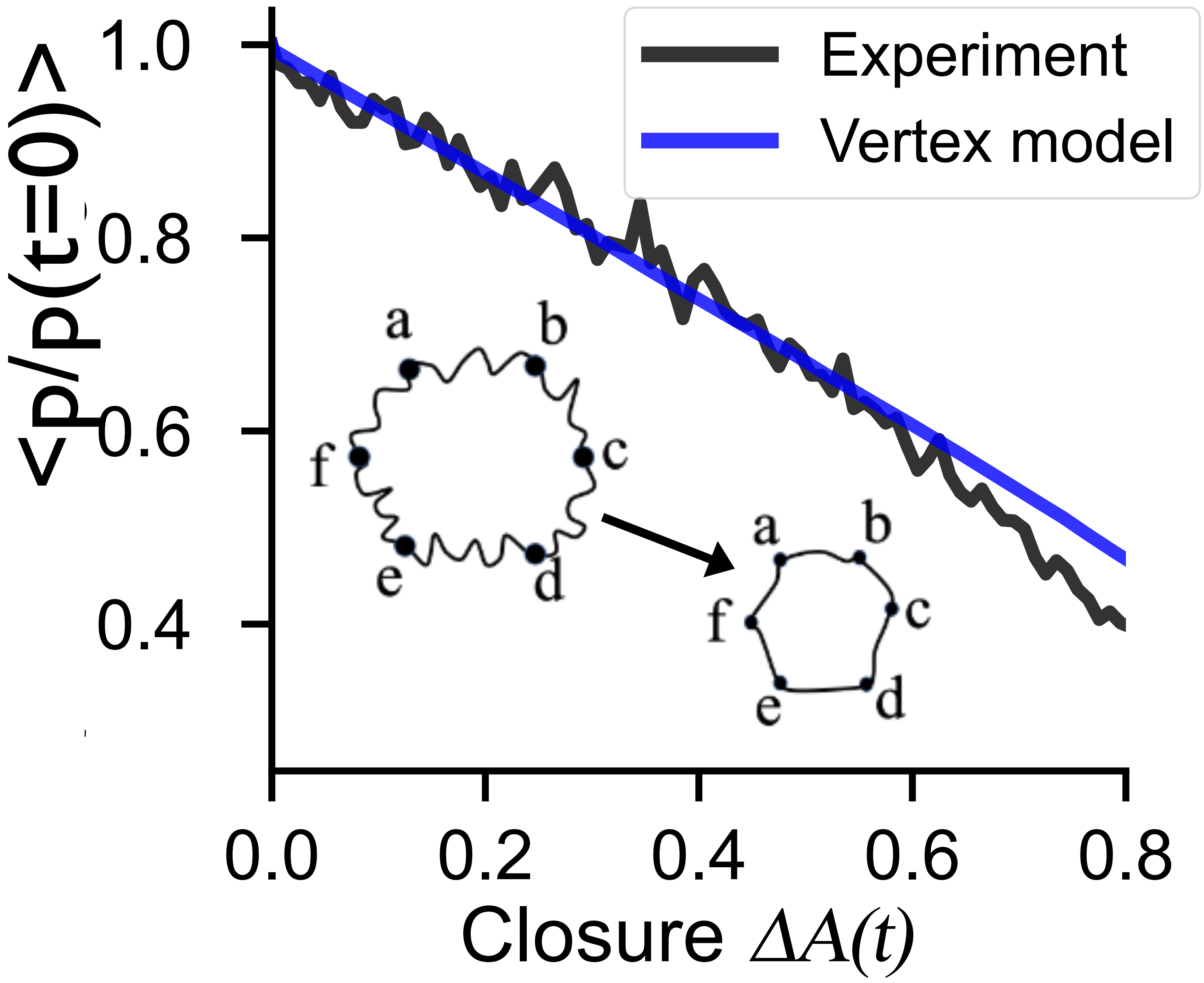


- Wide distbn of cell perimeters requires $\boldsymbol{\sigma}_{P0}$=0.45

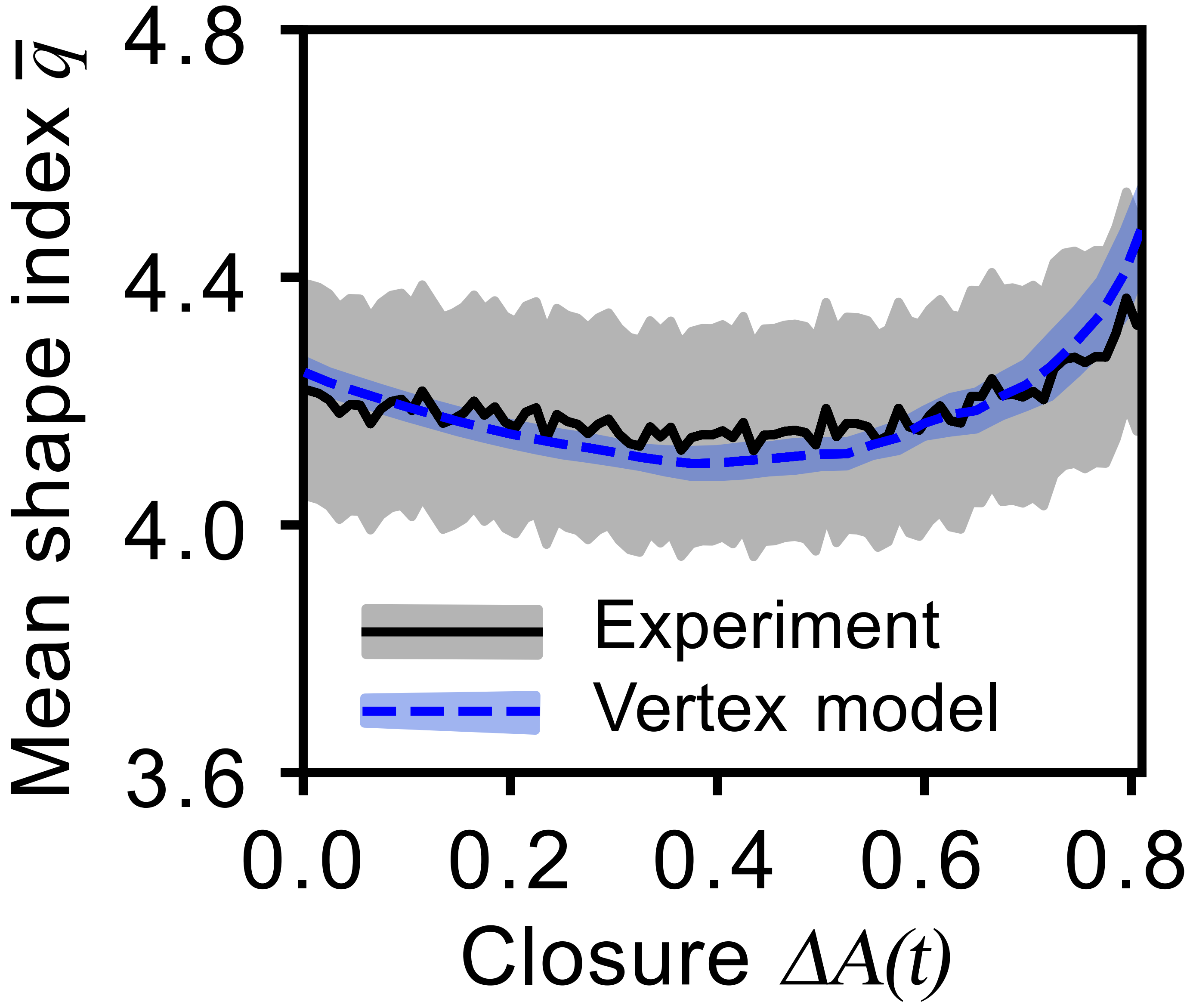
Li/Das/Bi PRL (2019)

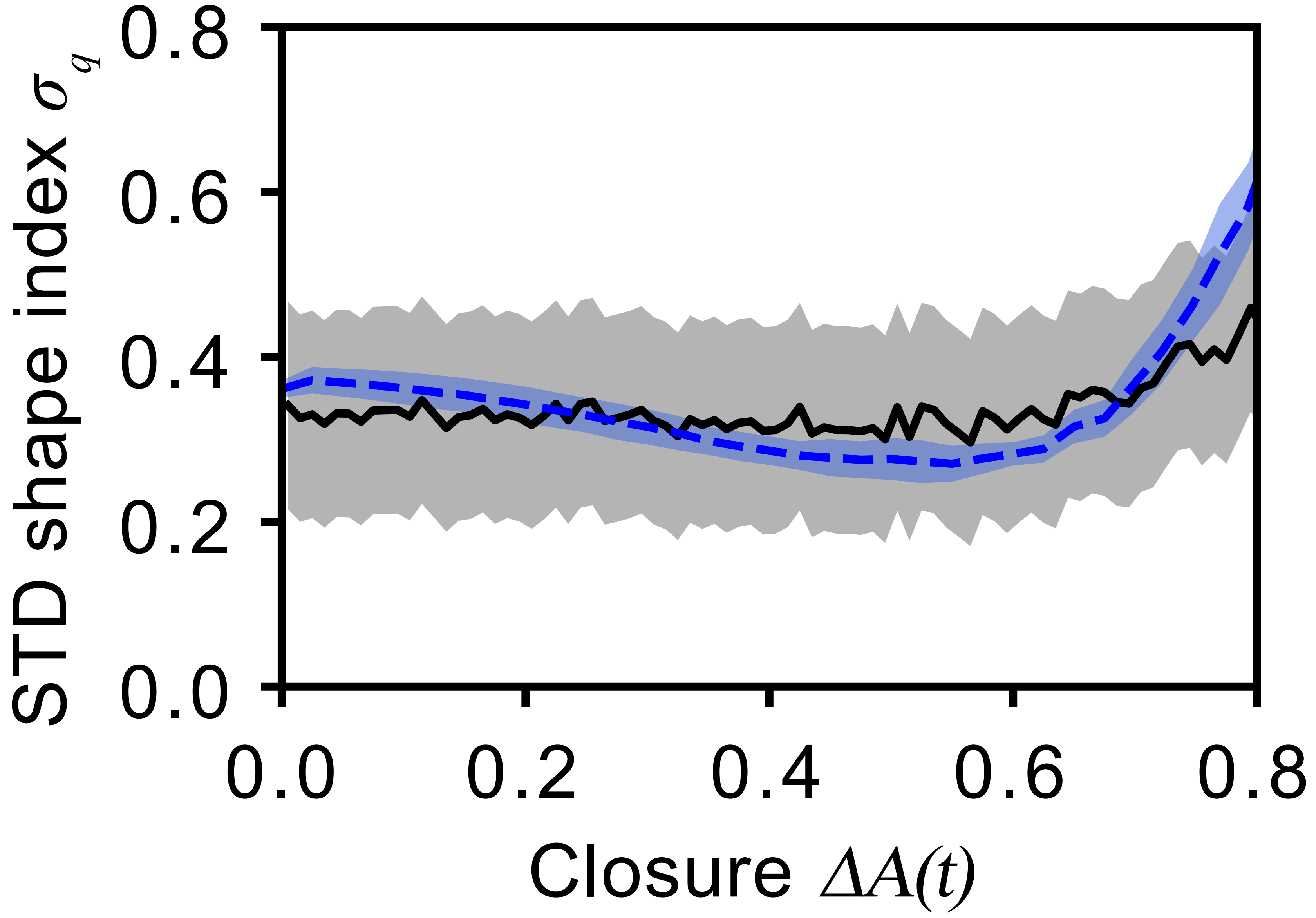
- Mean cell perimeter shrinks linearly during process
- Allow preferred cell perimeter to shrink at same rate

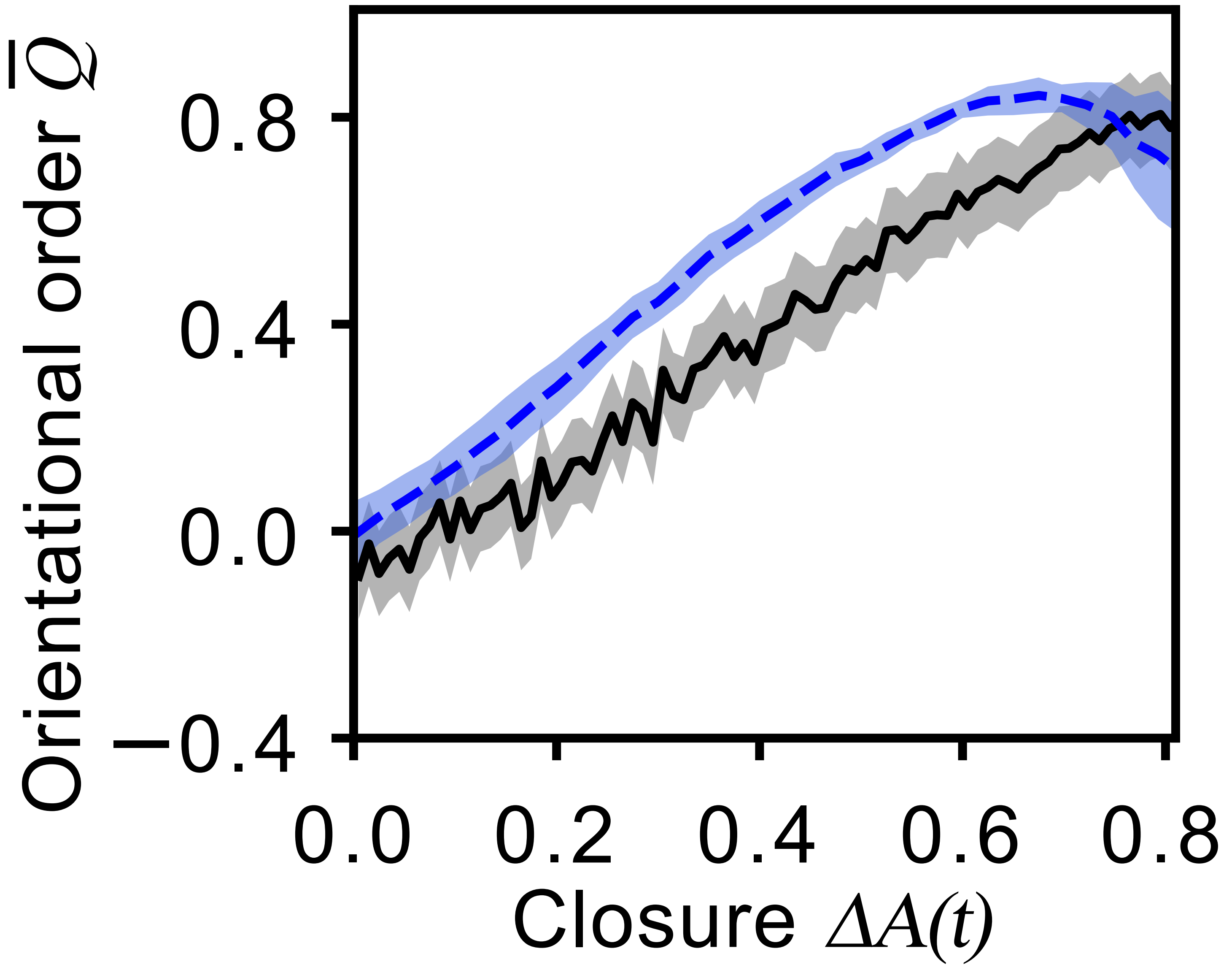
Tah, Haertter, Crawford, Kiehart, Schmidt, Liu arXiv (2023)



Adaptive Degrees of Freedom Are Changing During Dorsal Closure

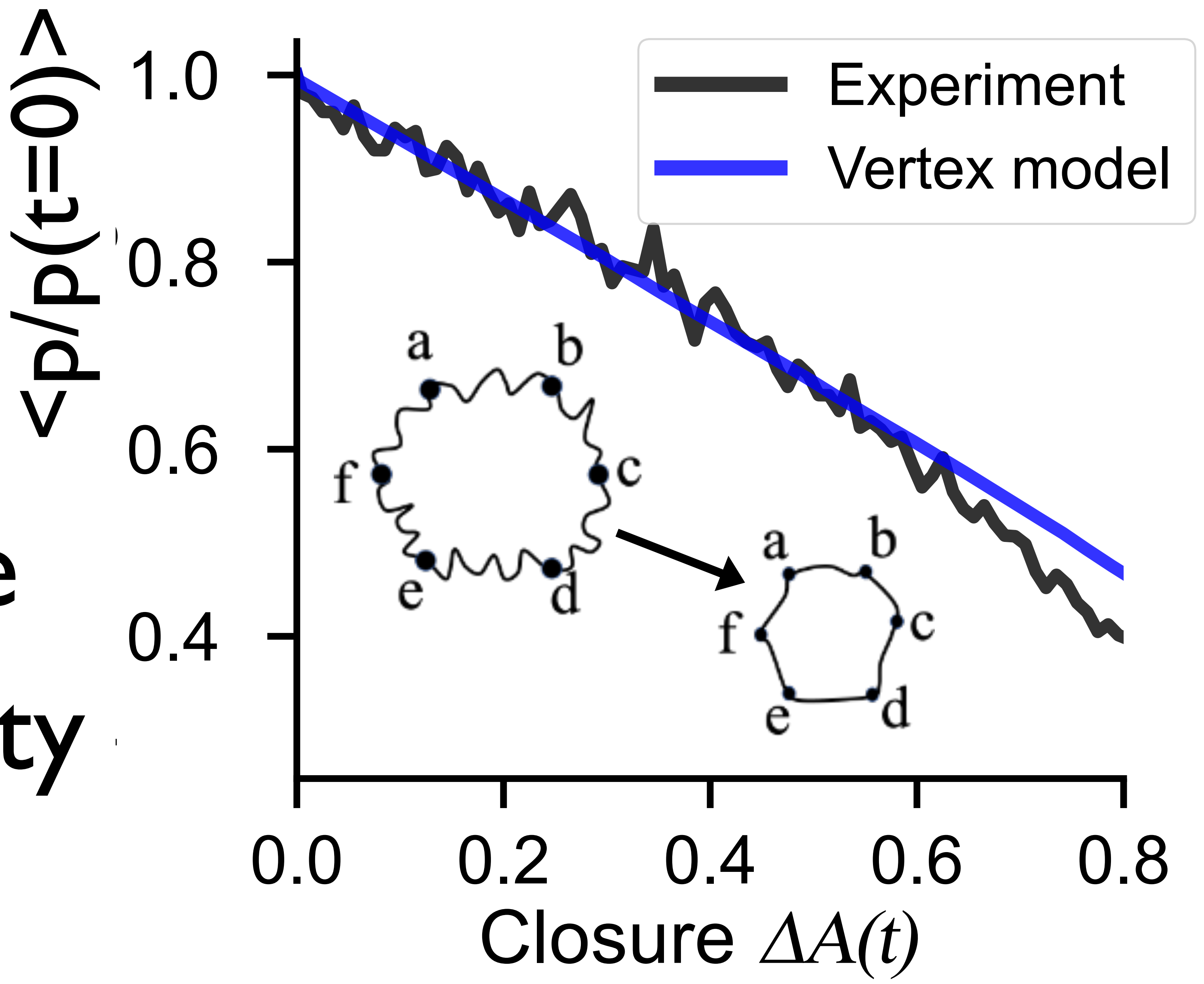






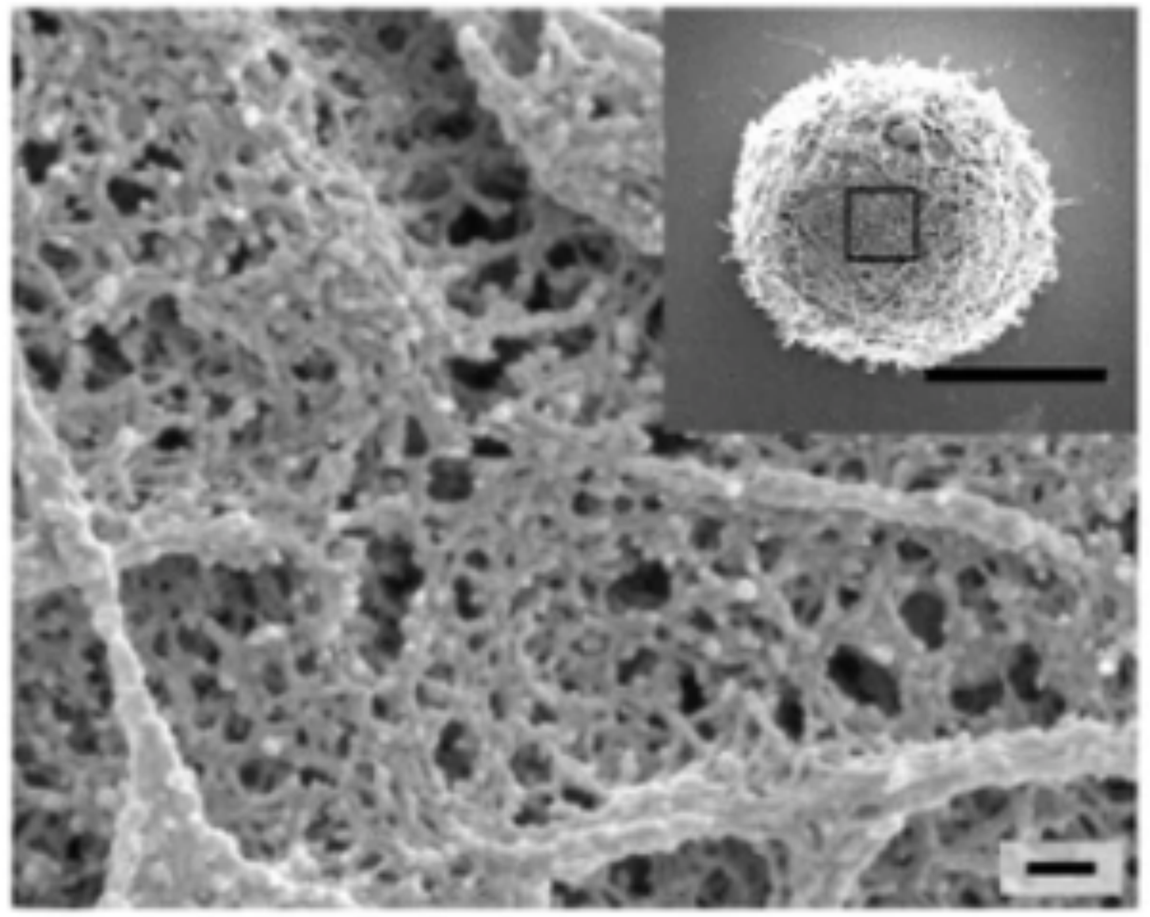
Tah, Haertter, Crawford, Kiehart, Schmidt, Liu arXiv (2023)

- Wide distbn of cell perimeters requires $\boldsymbol{\sigma_{p0}}$=0.45
- Mean cell perimeter shrinks linearly during process
- Allow preferred cell perimeter to shrink at same rate
- This is local rule that allows system to maintain rigidity

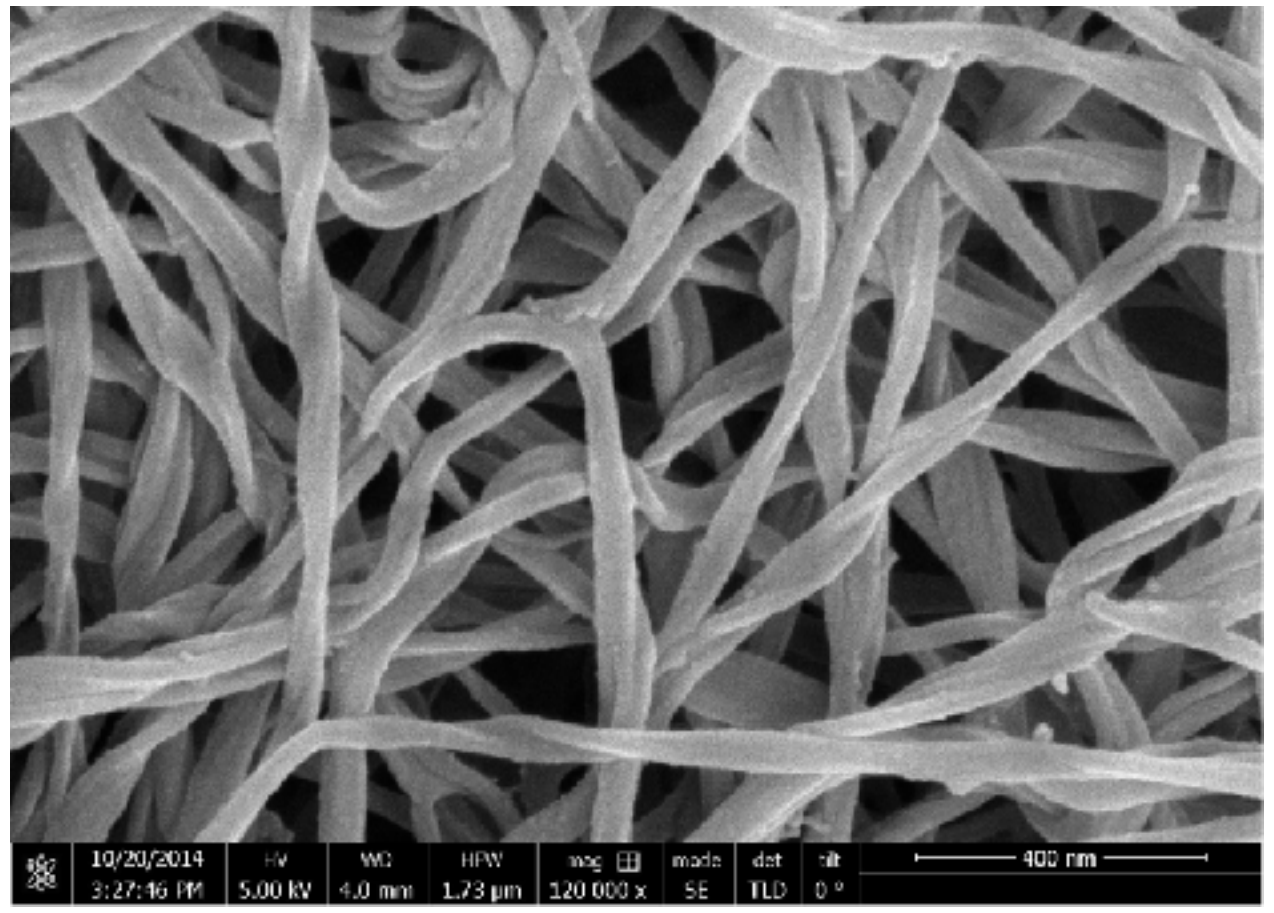


Quantitative agreement with experiments!

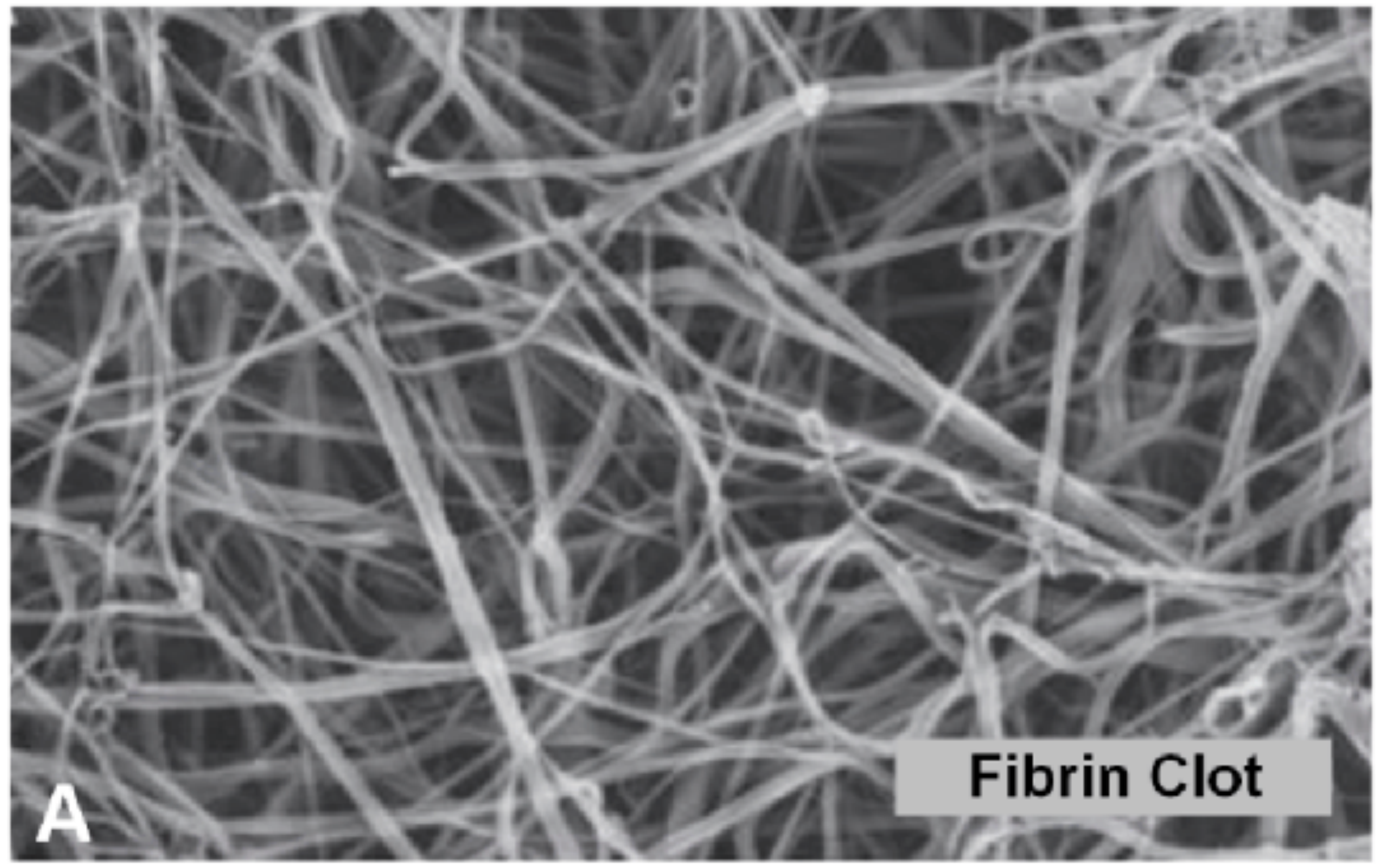
Biological Filament Networks That Confer Rigidity



Chugh et al Nat Cell Bio (2017)



Sharma et al Nat Phys (2016)

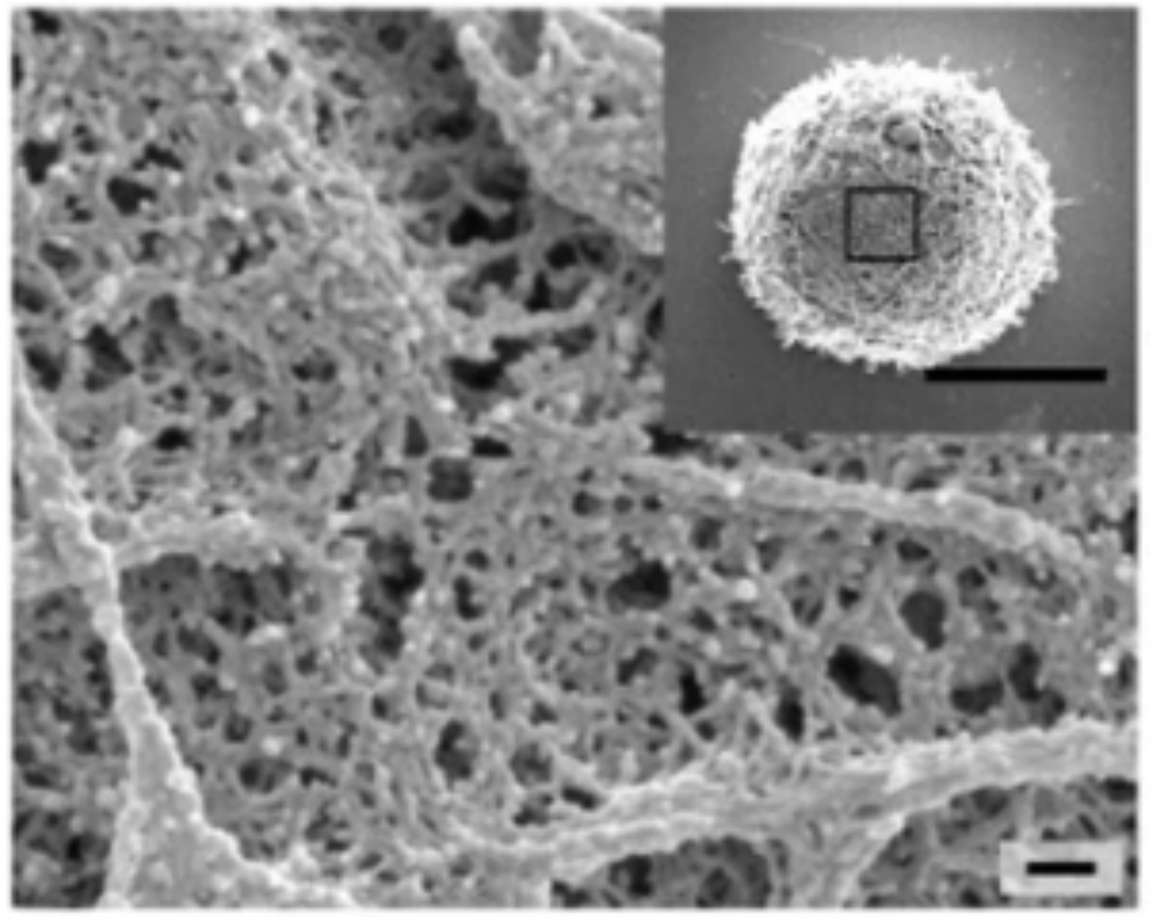


Yasudasan, Averett Polym (2020)

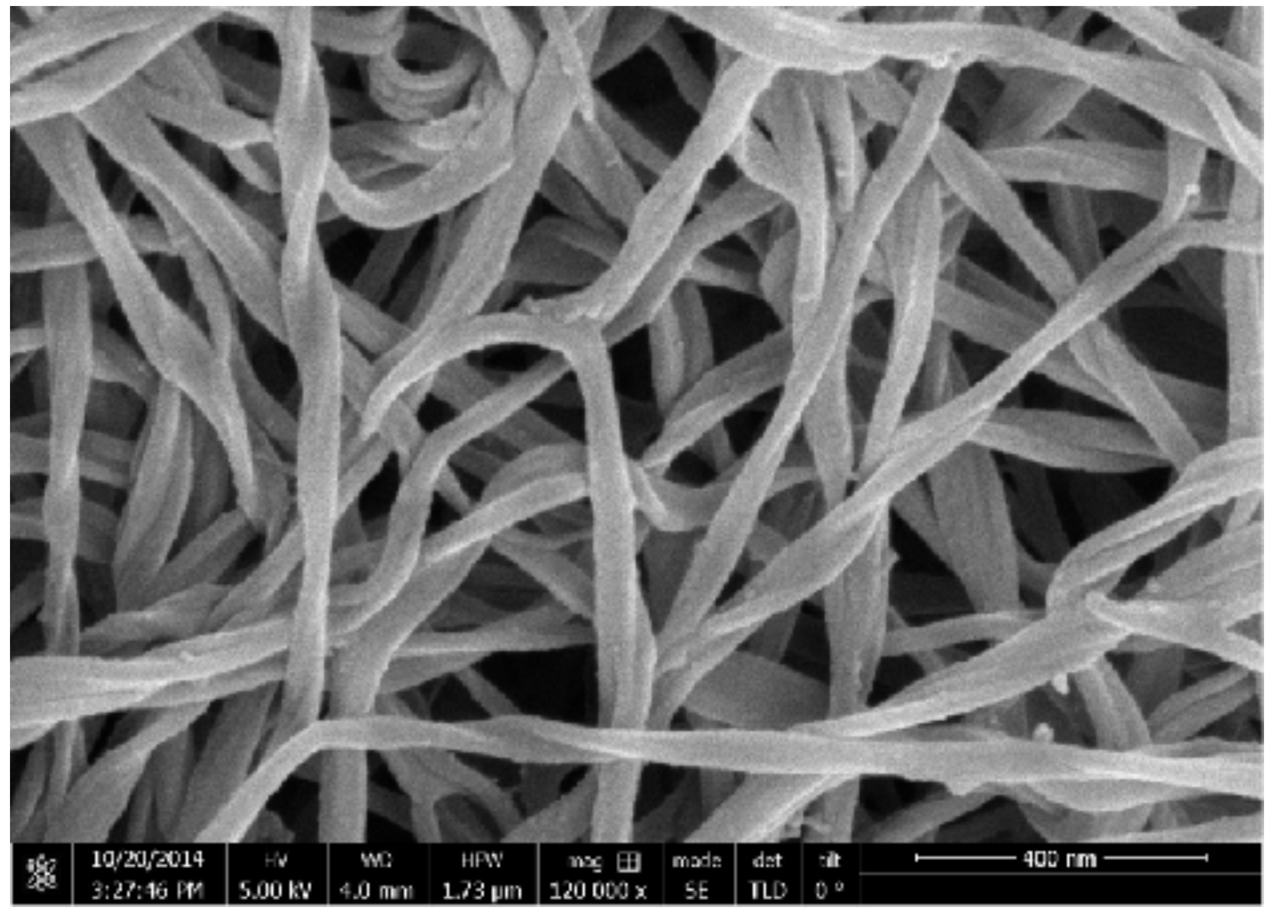
- They are rigid in a crazy way that is common across all 3 systems
- Under-coordinated networks $z \lesssim 4 < 6$
- Rigidified by prestress (tensegrity)—significant fraction of their stiffness
- Constant turnover with edges removed/added
    - Tension-inhibited pruning (cofilin, collagenase, filamin)
    - Filaments self-assemble/disassemble

Galvani Cunha, Liu, Crocker arXiv (2023)

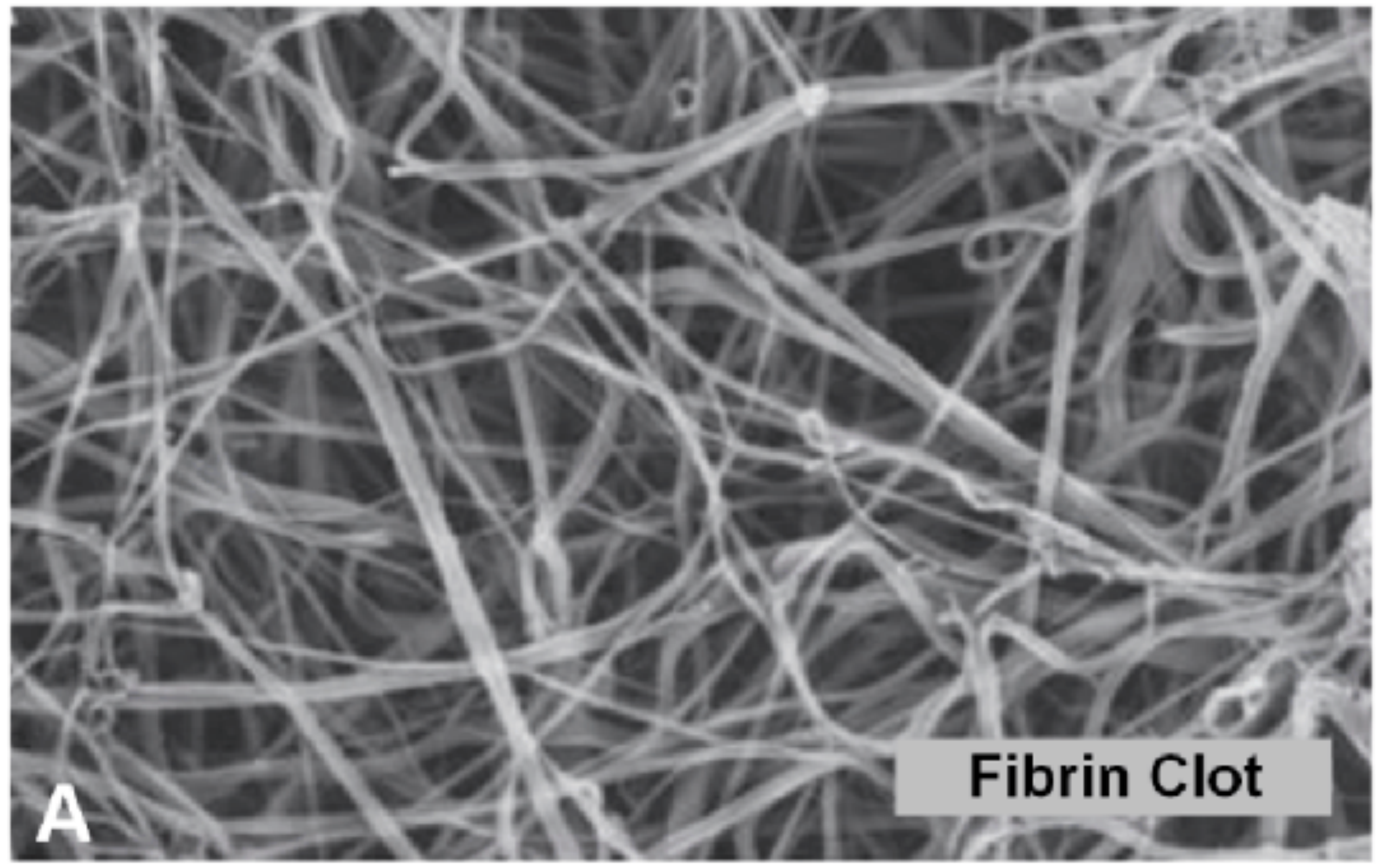
Biological Filament Networks That Confer Rigidity



Chugh et al Nat Cell Bio (2017)



Sharma et al Nat Phys (2016)



Yasudasan, Averett Polym (2020)

- They are rigid in a crazy way that is common across all 3 systems
- Under-coordinated networks $z \lesssim 4 < 6$
- Rigidified by prestress (tensegrity)—significant fraction of their stiffness
- Constant turnover with edges removed/added
  - Tension-inhibited pruning (cofilin, collagenase, filamin)
  - Filaments self-assemble/disassemble

use it or lose it

Galvani Cunha, Liu, Crocker arXiv (2023)

These Networks Have Desired Input/Output Relations

- They must maintain rigidity in face of constantly changing stresses
- Turnover creates parameters (presence/absence of edge)
- # parameters scales with system size—lots of parameters!
- Overparameterization implies
  - There many good solutions— each network is different but maintains rigidity
  - Systems tend to fall into minima with bigger basins of attraction—lots of flat directions in landscape that don't change rigidity much
- Representational drift—each network can maintain rigidity despite turnover
- solves ship of Theseus problem

GO TO SLIDE

These Networks Have Desired Input/Output Relations

- They must maintain rigidity in face of constantly changing stresses
- Turnover creates parameters (presence/absence of edge)
- # parameters scales with system size—lots of parameters!
- Overparameterization implies
  - There many good solutions— each network is different but maintains rigidity
  - Systems tend to fall into minima with bigger basins of attraction—lots of flat directions in landscape that don't change rigidity much
- Representational drift—each network can maintain rigidity despite turnover
- solves ship of Theseus problem

It's not crazy—it's BRILLIANT

Thinking of biological function, especially homeostasis, this way is potentially useful

GO TO SLIDE

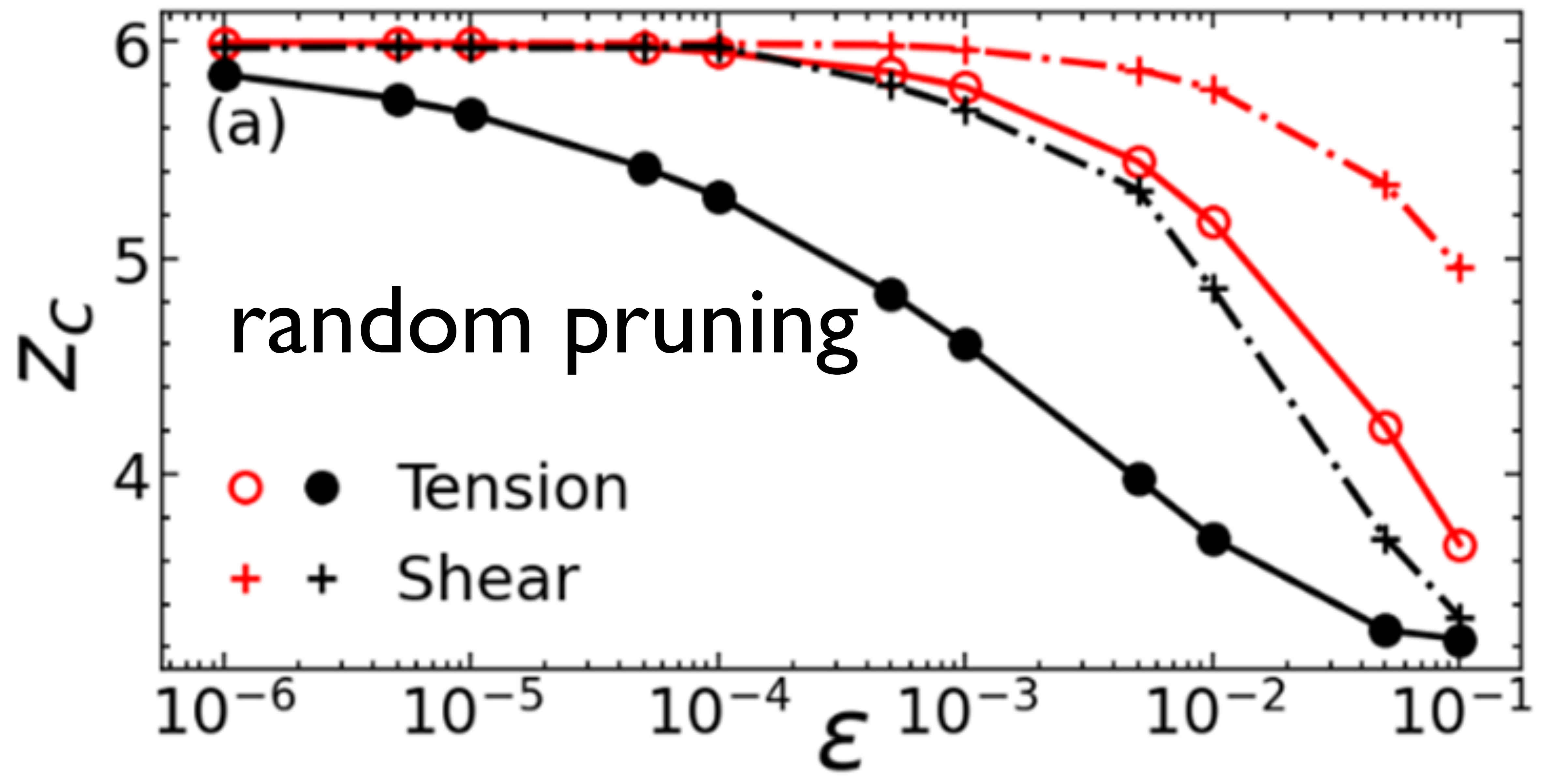
Low-Tension Pruning Leads to Lower Critical Coordinations

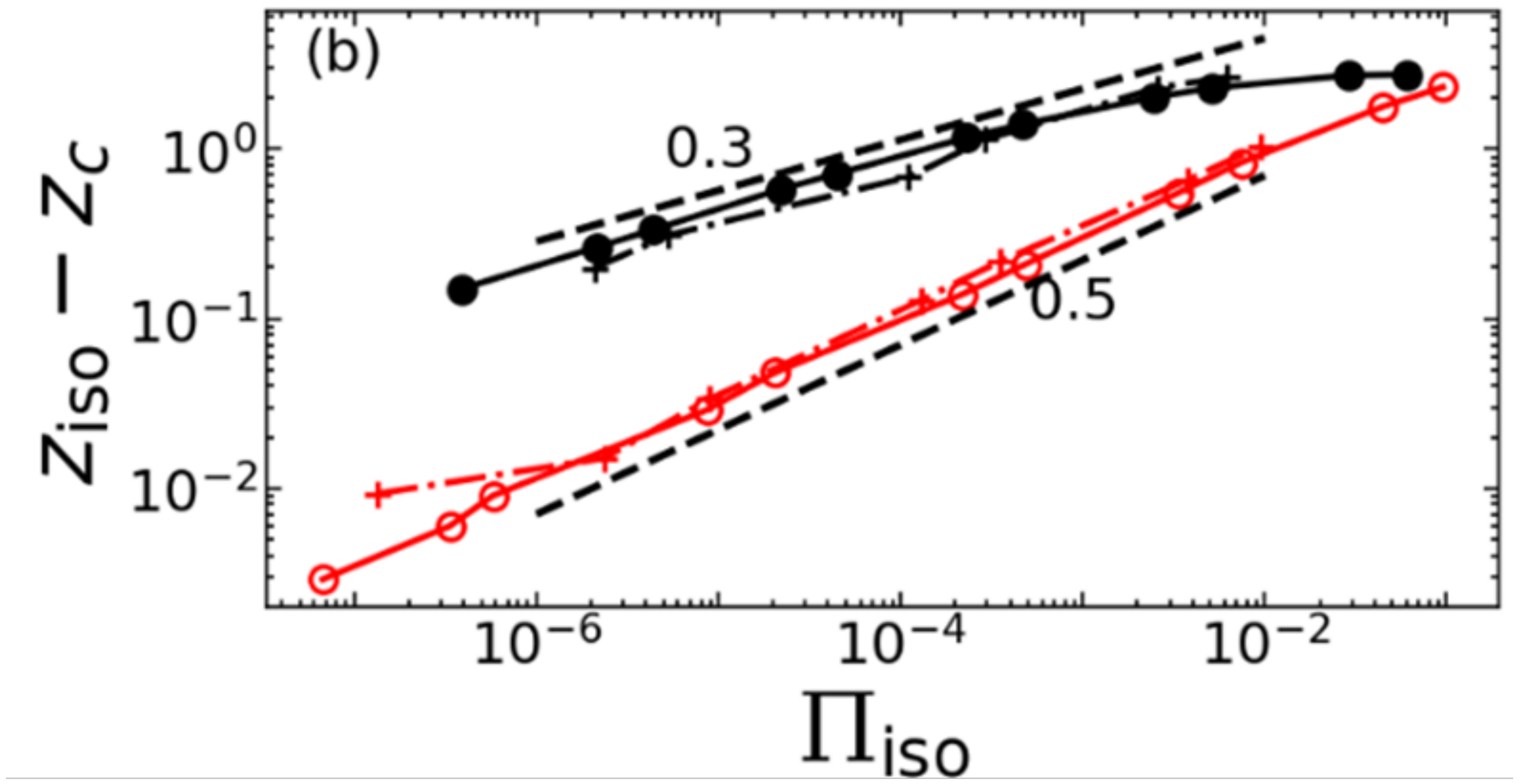
- Biologically-relevant $Z_c \sim 3$-4 much easier to reach with low-tension pruning than random pruning

Galvani Cunha/Liu/Crocker arXiv 2023

- Can only reach realistic ratios of prestress/stiffness with low-tension pruning

low-tension pruning





Low-Tension Pruning Allows Biologically-Realizable Networks

- Match Z, G for random vs. low-tension pruning
- Translate into real units: $\boldsymbol{\kappa}$~40pN/nm, ℓ~1μm

G~0.0044 $\boldsymbol{\kappa}/\ell$~16 kPa c.f. 20 kPa✓

$\boldsymbol{\Pi}$~0.0015 $\boldsymbol{\kappa}/\ell$~6 kPa. c.f. 5 kPa✓

- Compare top tension to breaking force:
  - experimental value for actin ~600pN
  - low-tension: 400pN✓
  - random: 900pN×