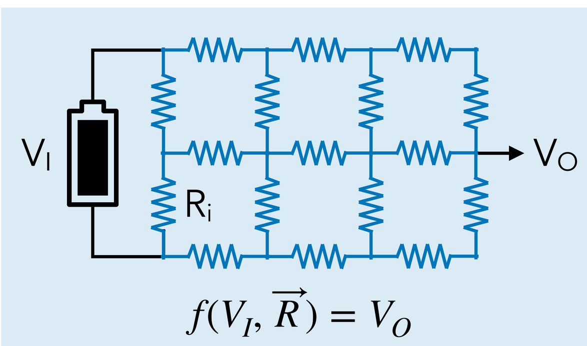
A Physics-Driven Self-Learning Transistor Network

Sam Dillavou[1], Benjamin Beyer[1], Menachem Stern[1], Marc Z Miskin[2], Andrea J Liu[1], Douglas J Durian[1]

[1]Dept of Physics & Astronomy [2]Dept of Electrical and Systems Engineering

University of Pennsylvania

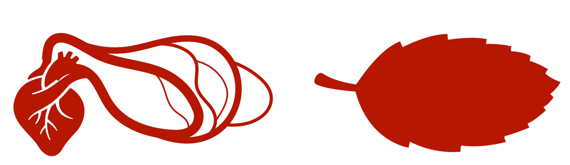
We build 'learning materials.'



Our system is an ensemble of variable resistors that self-adjust based on local rules to produce a desired output from a given input.

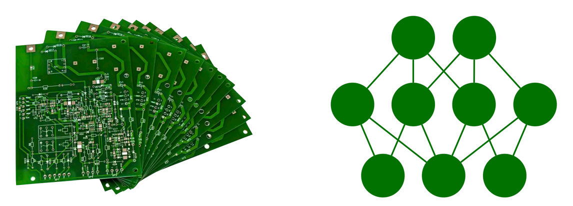
That is, it learns by example.

Active, out-of-equilibrium system with complex dynamics where learning is an emergent property.



Biological systems

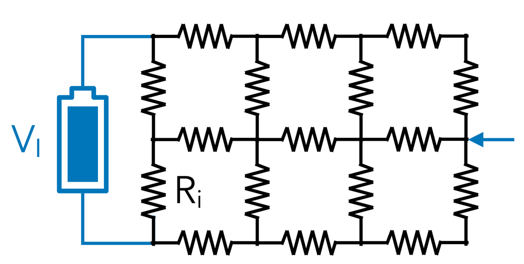
This is in contrast to the typical ways we construct desired functions:



by design: circuitboard — inflexible, prone to damage

by calculation: neural network — power hungry, inefficient

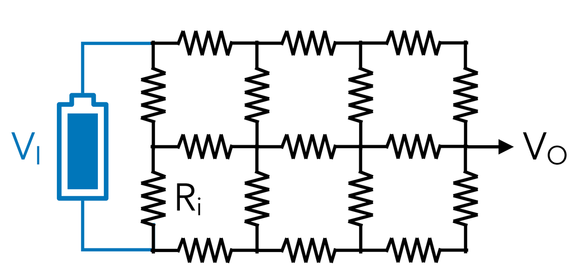
Our system uses physics + contrastive learning.



Clamp inputs & outputs

We build twin networks so we can access two electrical states.

In the free state, physics generates the output V_o. In the clamped state, we enforce the desired output V_D. [1]



Free: apply inputs.

$$\Delta R_i = -\frac{\partial}{\partial R_i}[P_C - P_F] = \left[(V_i^C)^2 - (V_i^F)^2\right]R_i^{-2}$$

The edges self-adjust to minimize the power difference between the two states.

Using this local rule (and no outside help), the ensemble creates the desired functionality. That is, learning is decentralized and physics-driven.

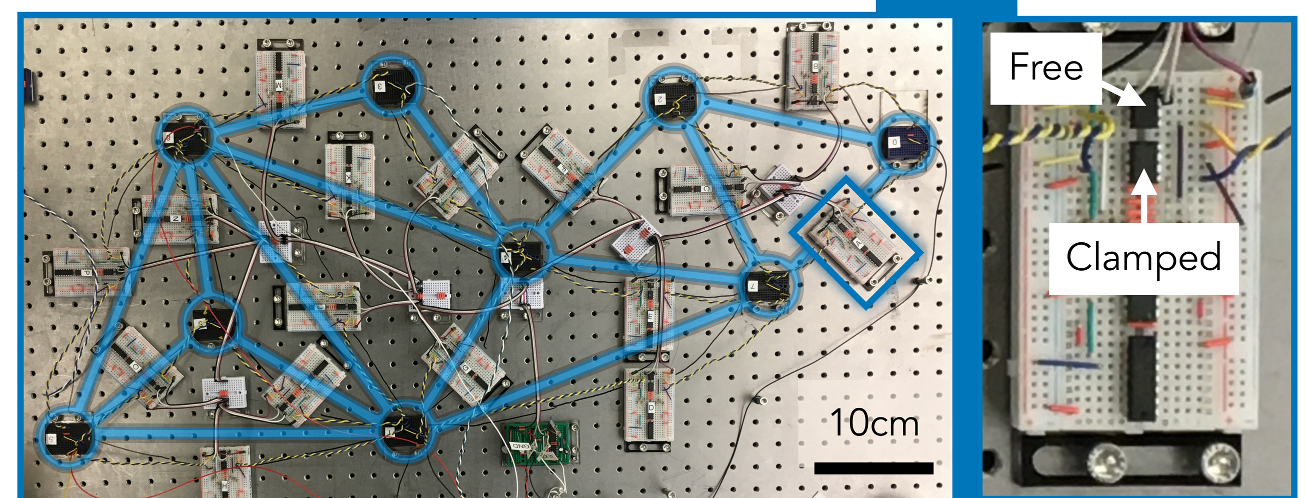
[1] M Stern, et. al. *PRX* 2021
[2] S Dillavou, et. al. *PRApplied* 2022
[3] JF Wycoff, S Dillavou et. al. *J Chem Phys* 2022
[4] M Stern, S Dillavou, et. al. *PRR* 2022

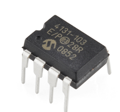
Our method is built on many works studying learning, e.g.
Contrastive learning: JR Movellan 1991
Equilibrium propagation: Scellier & Bengio 2017
Tuning in silico networks: Goodrich et. al. *PRL* 2015
Directed aging: Pashine et. al. *Sci Adv* 2019

Thank you to our funding sources

Nat'l Science Foundation — Dept of Energy — Simons Foundation

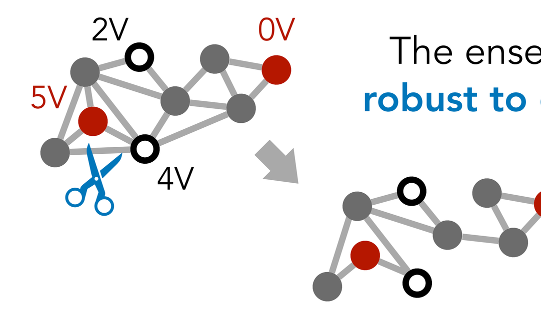
Generation 1 Network [2]





Digital variable resistors: discrete resistances only.

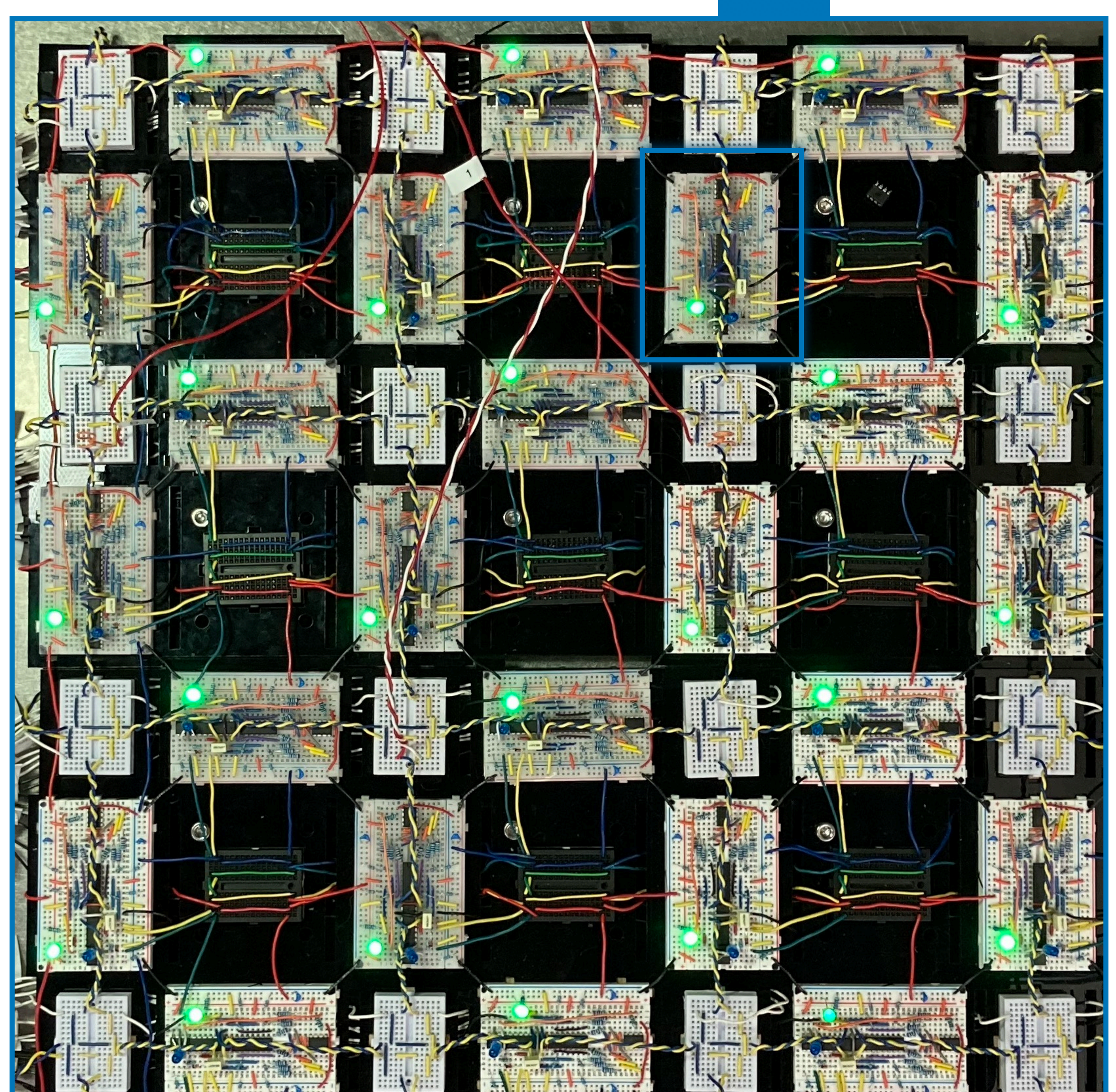
Network able to classify flowers from measurements (iris dataset >95%)



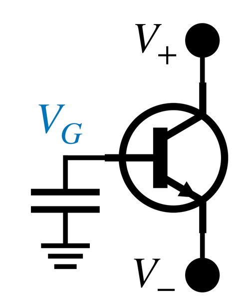
The ensemble is robust to damage.

Gen 1 system already useful for investigating physical learning: system learns better with desynchronous updates, and learns faster out of equilibrium. [3,4]

Generation 2 Network



This system learns by itself.



Edges are transistors, making them nonlinear and continuously variable.

$R \approx a e^{-b(V_G - V_0)}$ $\quad V_0 \approx \min(V_1, V_2)$

Nonlinearity is essential for complex functionality, and is behind the success of neural networks.

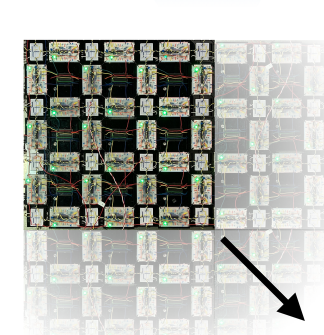
		Input 1: True	Input 1: False
Input 2	False	True	False
	True	False	True

XOR truth table

The network is able to learn a canonical example of nonlinearity, XOR. We can observe the entire learning process.

With this network we will investigate dynamics of physical learning: power vs accuracy, effects of bias, noise, damage, hallmarks of trained networks, etc, and build useful tools like smart sensors.

Our system is already 10x faster than a simulation of itself on modern CPU: 300μs vs 3ms/step.



Looking forward, we aim to scale up our networks to create **neuromorphic hardware** with the potential to be **faster and more energy efficient** than GPU-based machine learning. These scaled up systems will also allow us to study more complex physical learning problems in a transparent system.

Visualizing Learning