Machine learning based on physical dynamics

Florian Marquardt
Max Planck Institute for the Science of Light and
Friedrich-Alexander Universität Erlangen-Nürnberg

works with:
Victor Lopez-Pastor (Phys. Rev. X 2023)
Clara Wanjura (arXiv 2023)

DEEP LEARNING REVOLUTION

2012 IMAGENET

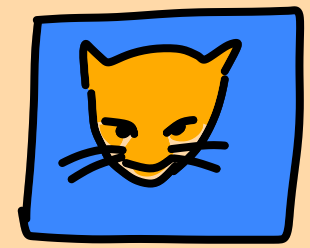
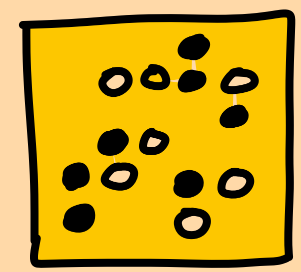


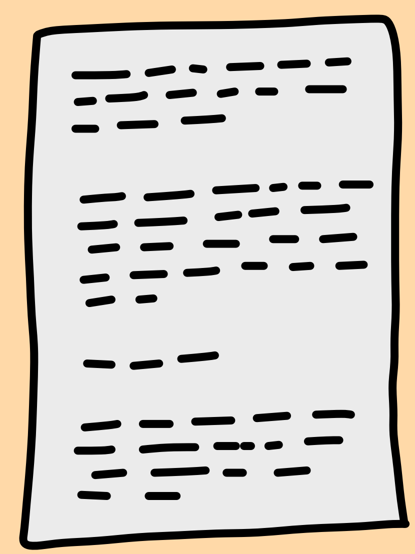
IMAGE RECOGNITION

2017 ALPHA ZERO

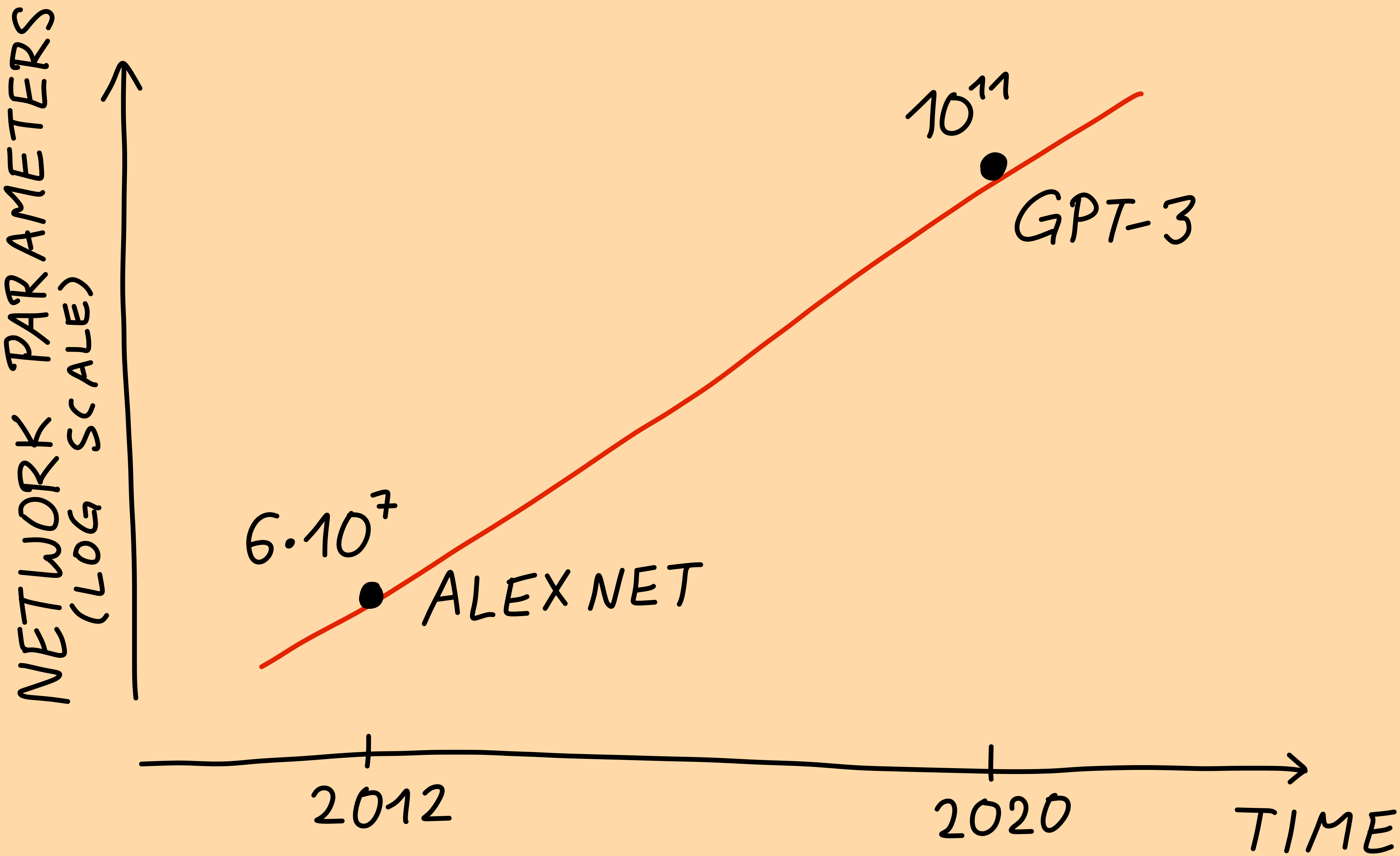


PLAYING GAMES

2020 GPT-3



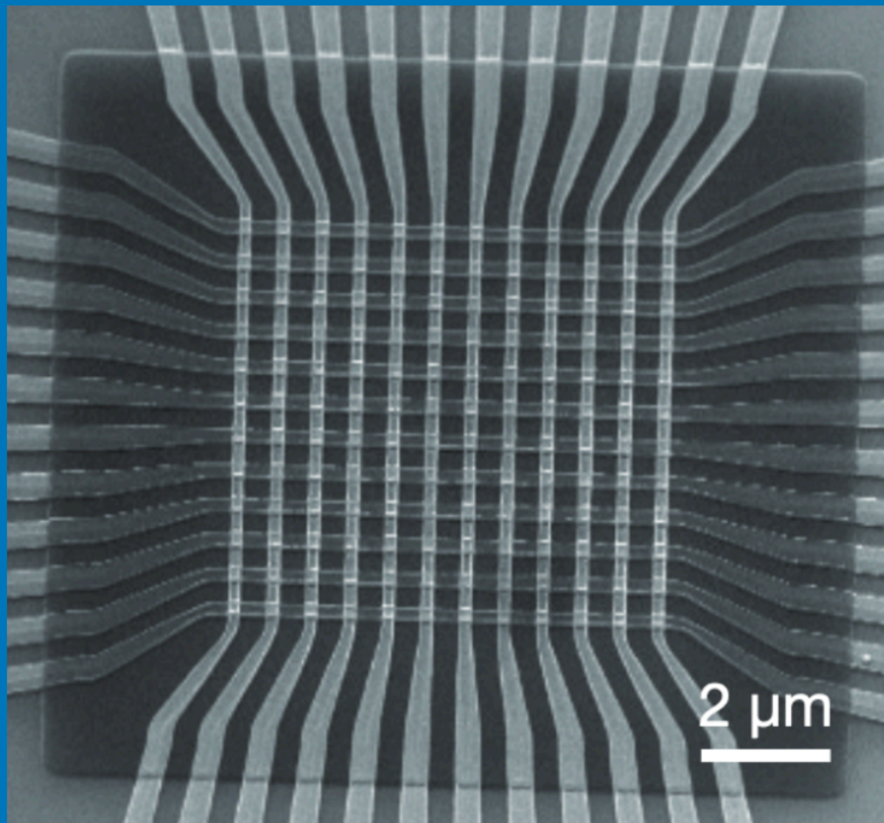
WRITING TEXT



→ CAN WE BUILD BETTER HARDWARE?
FAST, HIGHLY PARALLEL, ENERGY-EFFICIENT

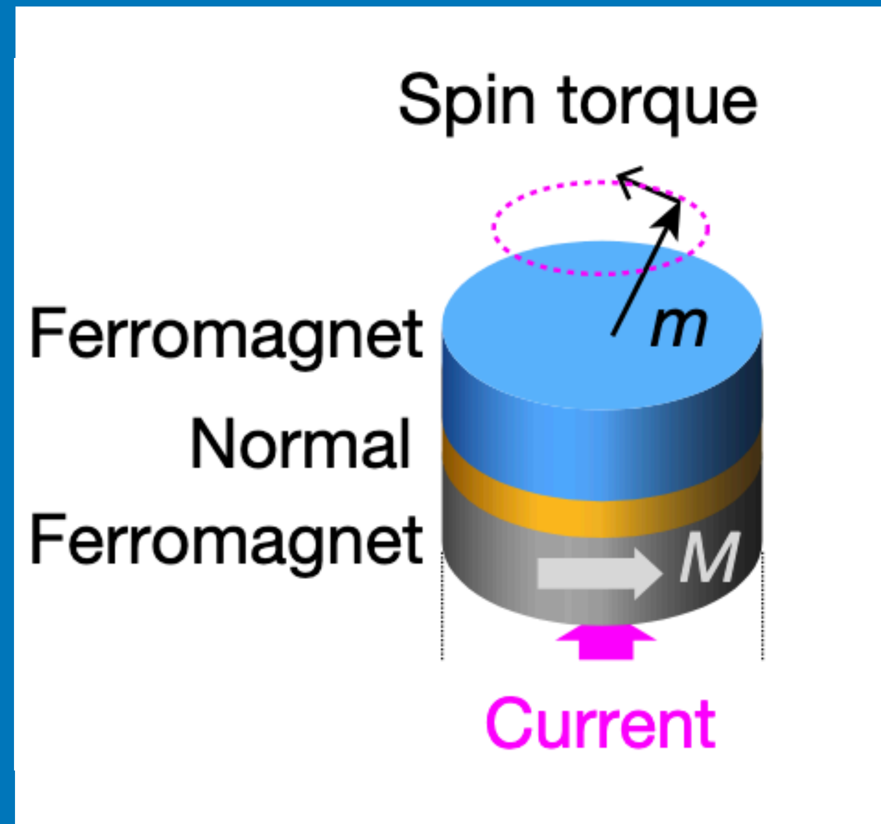
→ NEUROMORPHIC COMPUTING

NEUROMORPHIC COMPUTING



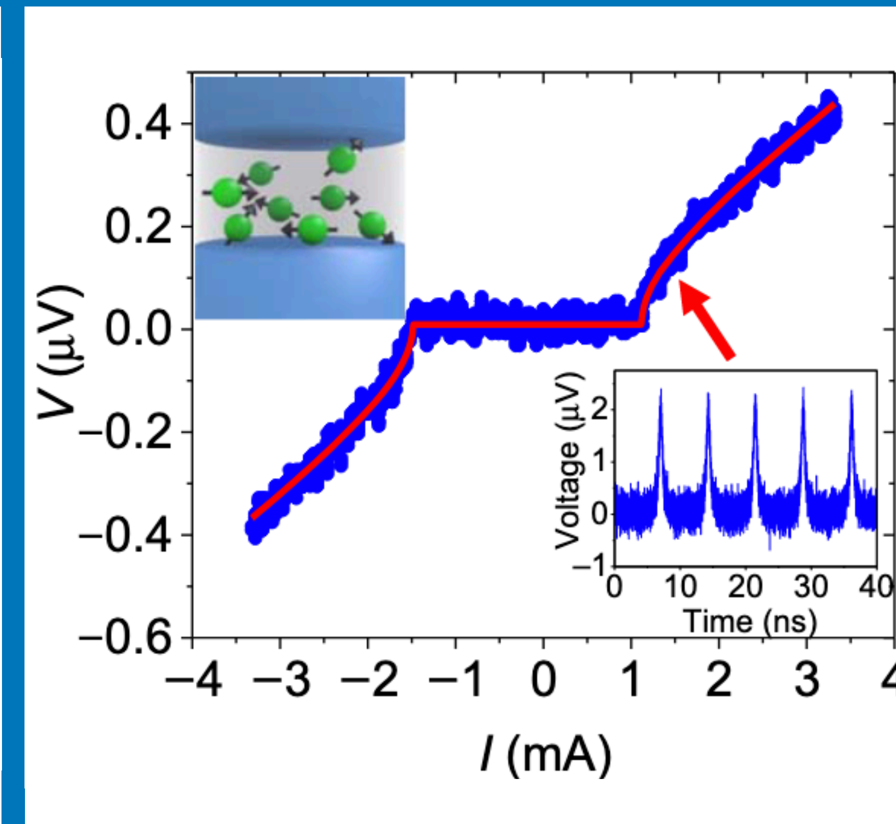
(Prezioso et al 2015)

MEMRISTORS



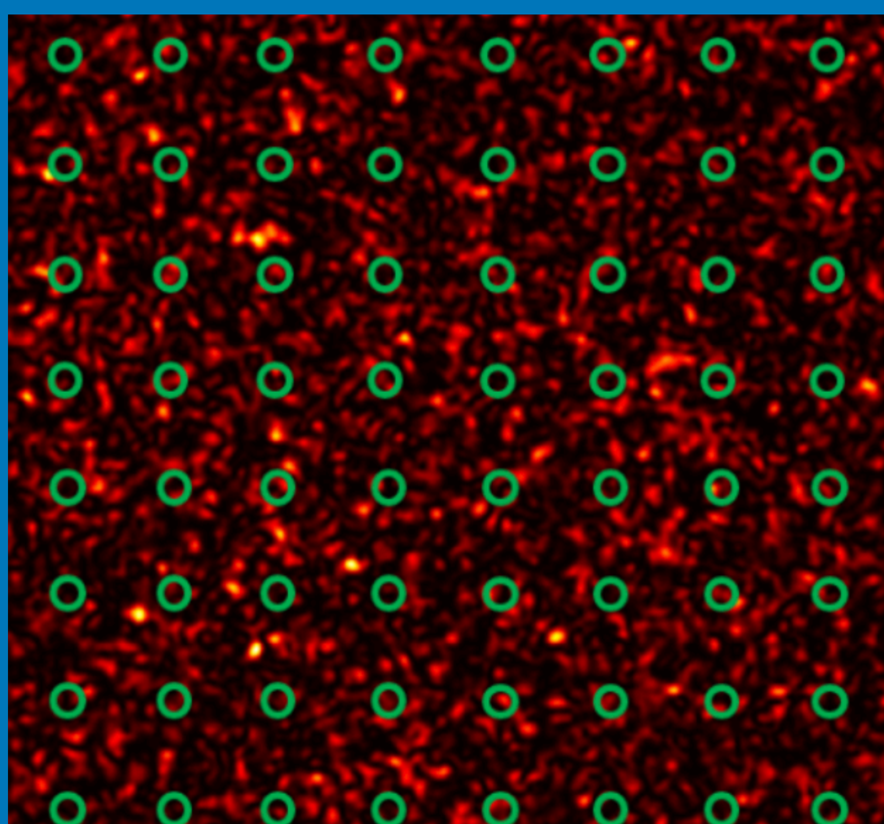
(Torrejon et al 2017)

SPINTRONIC OSCILLATORS



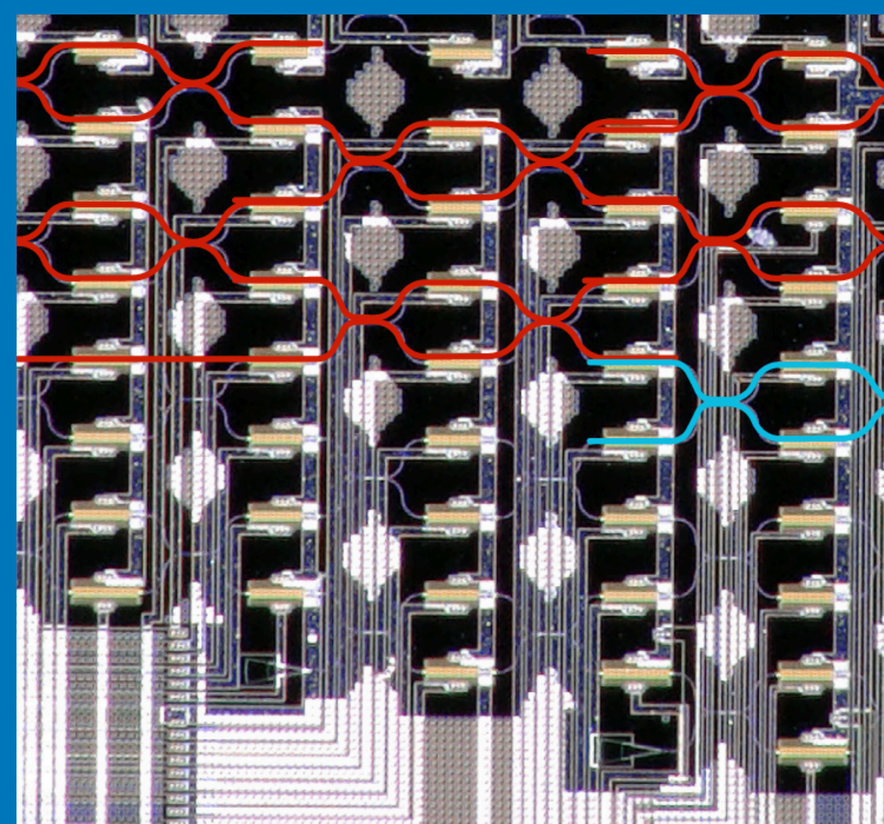
(Schneider et al 2018)

MAGN. JOSEPHSON JUNCTIONS



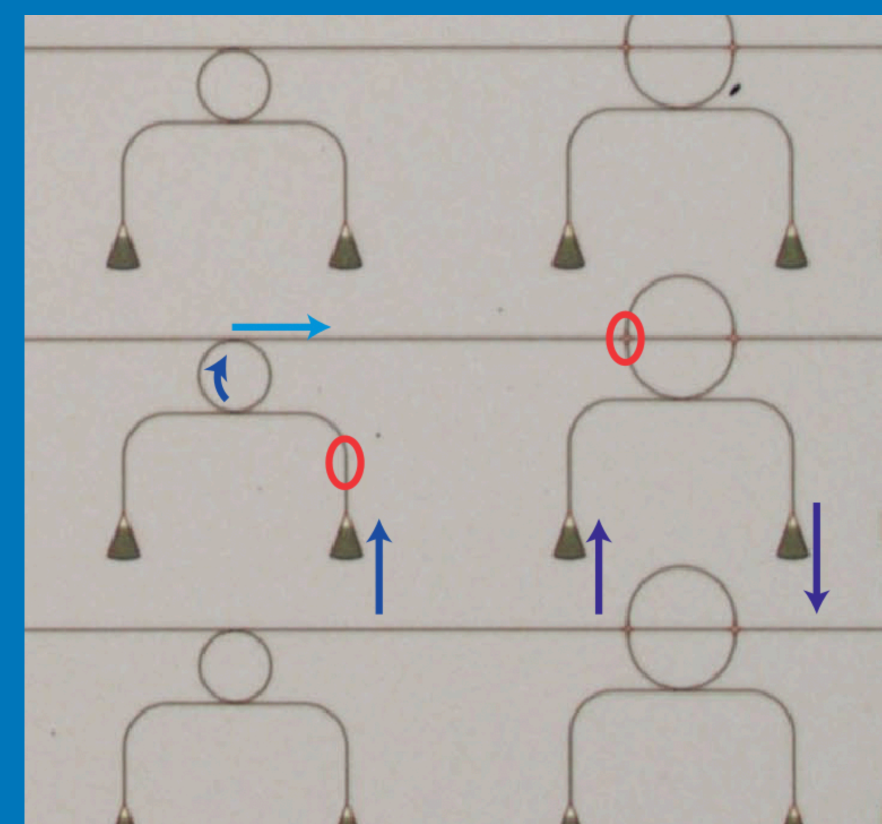
(Dong et al 2019)

OPTICS IN COMPLEX MEDIA



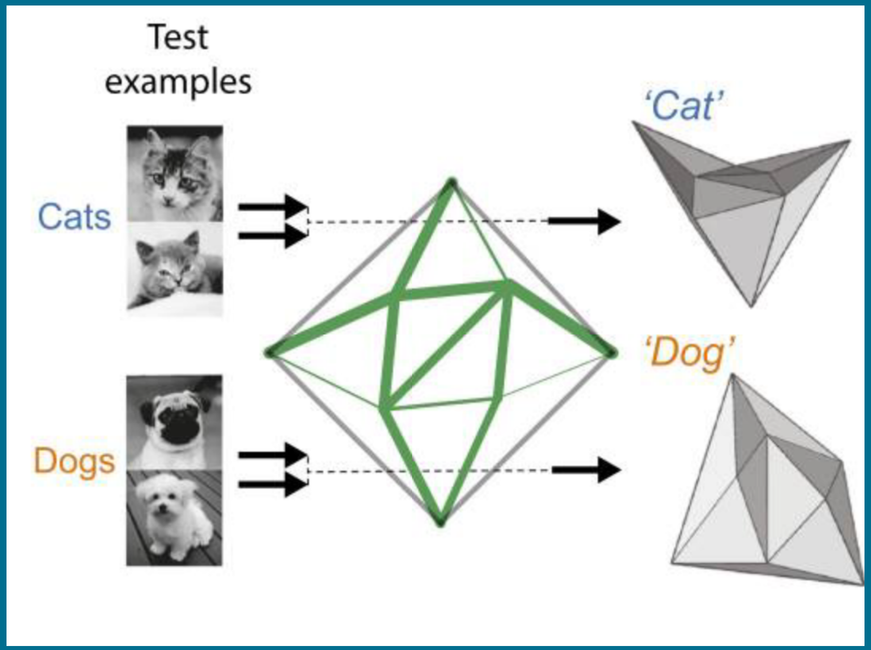
(Shen et al 2017)

TUNEABLE INTERFEROMETER



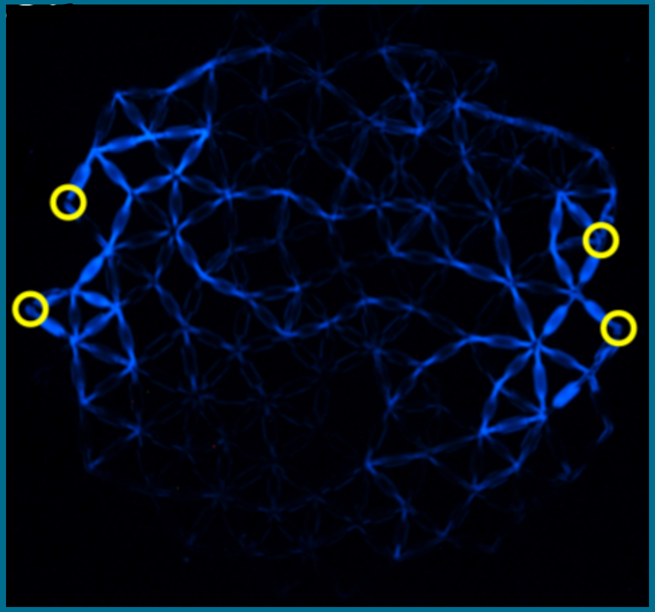
(Feldmann et al 2019)

PHASE-CHANGE MATERIALS



(Stern et al 2020)

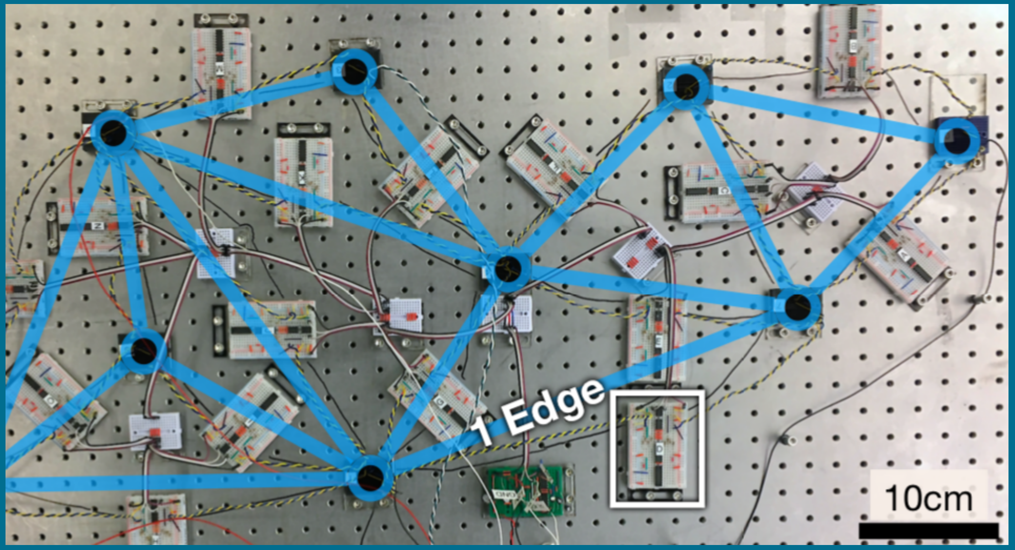
LEARNING
TO FOLD



(Pashine 2021)

MECHANICAL
NETWORKS

RESISTOR
NETWORKS

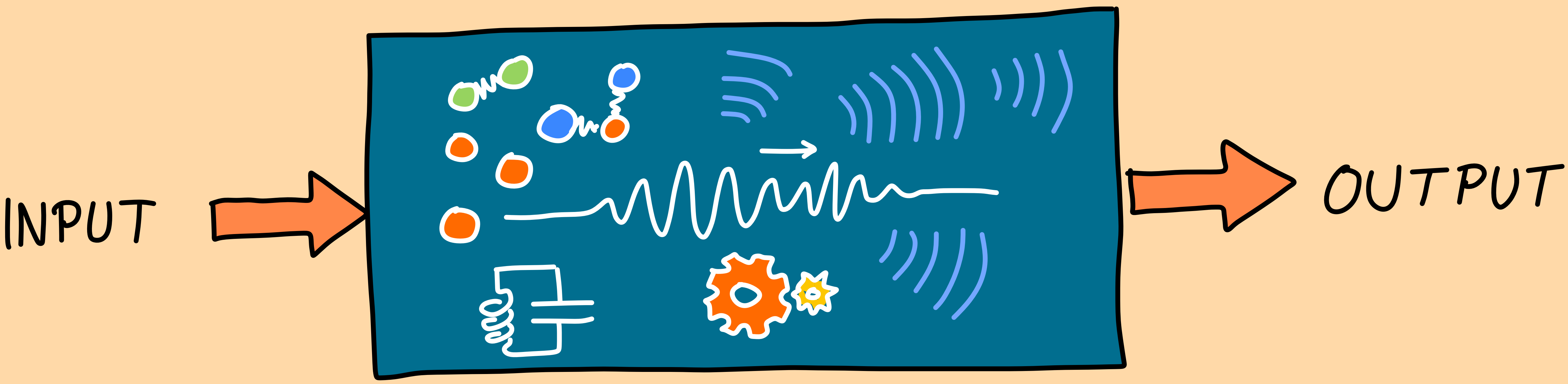


(Dillavou et al 2022)

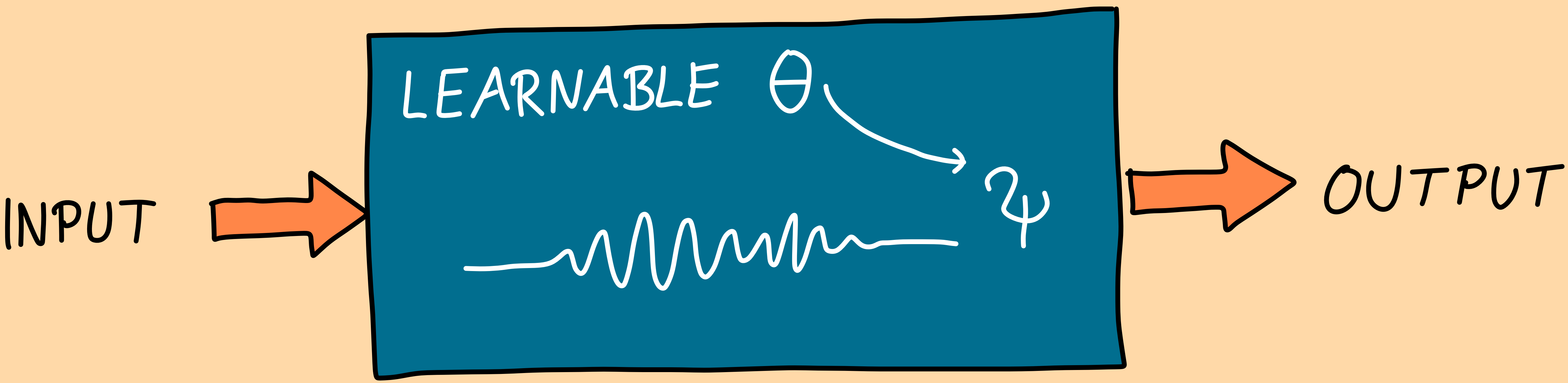
A new way to learn: Hamiltonian Echo Backpropagation

Victor Lopez-Pastor & F.M.
Phys. Rev. X 13, 031020 (2023)

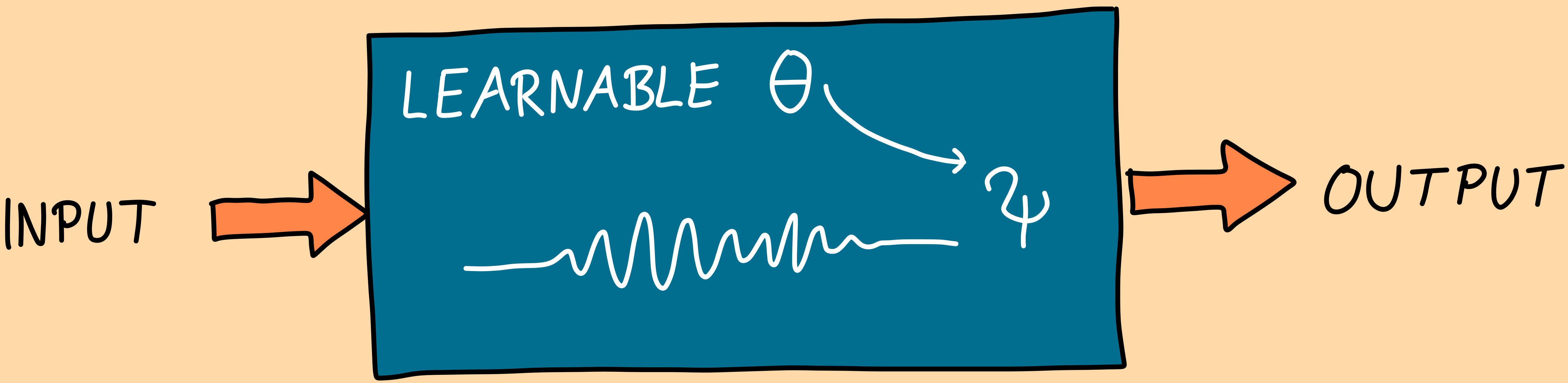
PHYSICAL LEARNING MACHINE



PHYSICAL LEARNING MACHINE



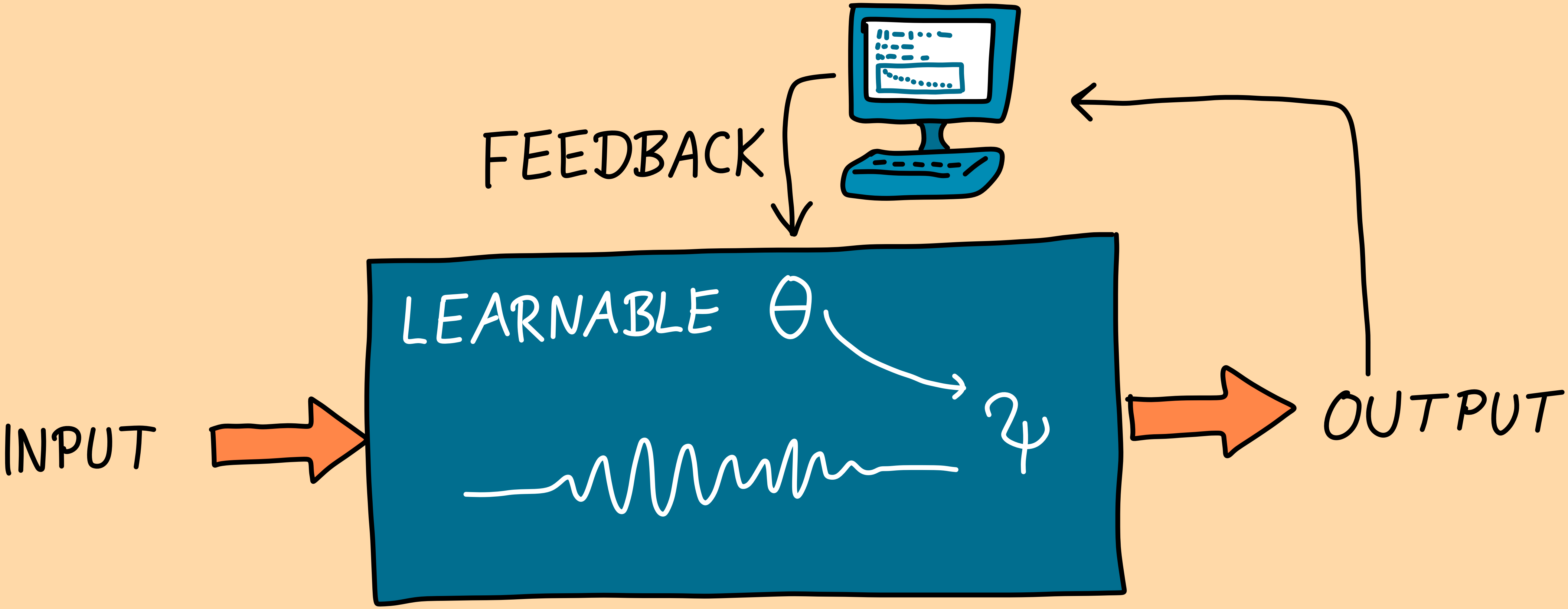
PHYSICAL LEARNING MACHINE



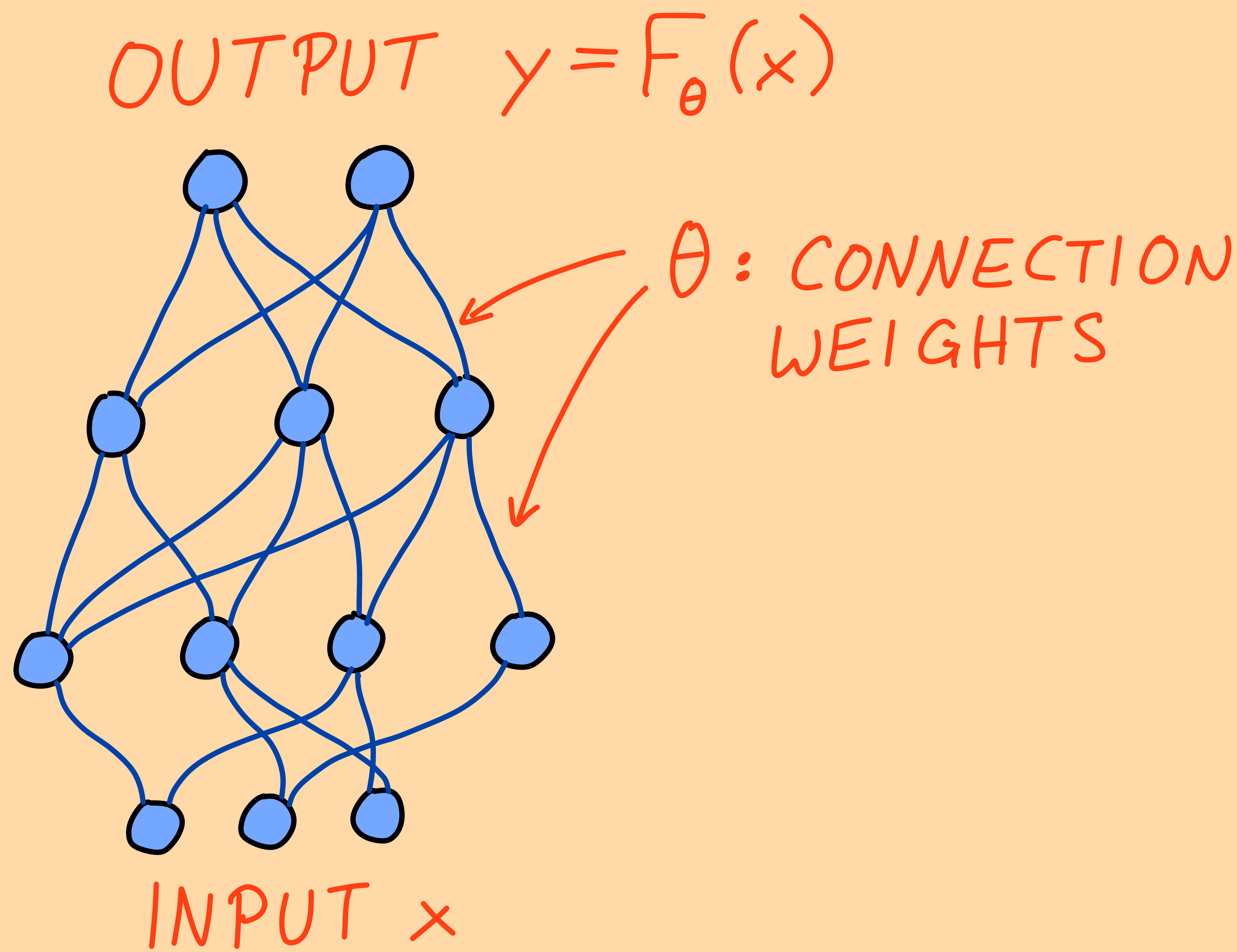
HOW TO LEARN?

(FIND RIGHT θ TO GET DESIRED INPUT → OUTPUT)

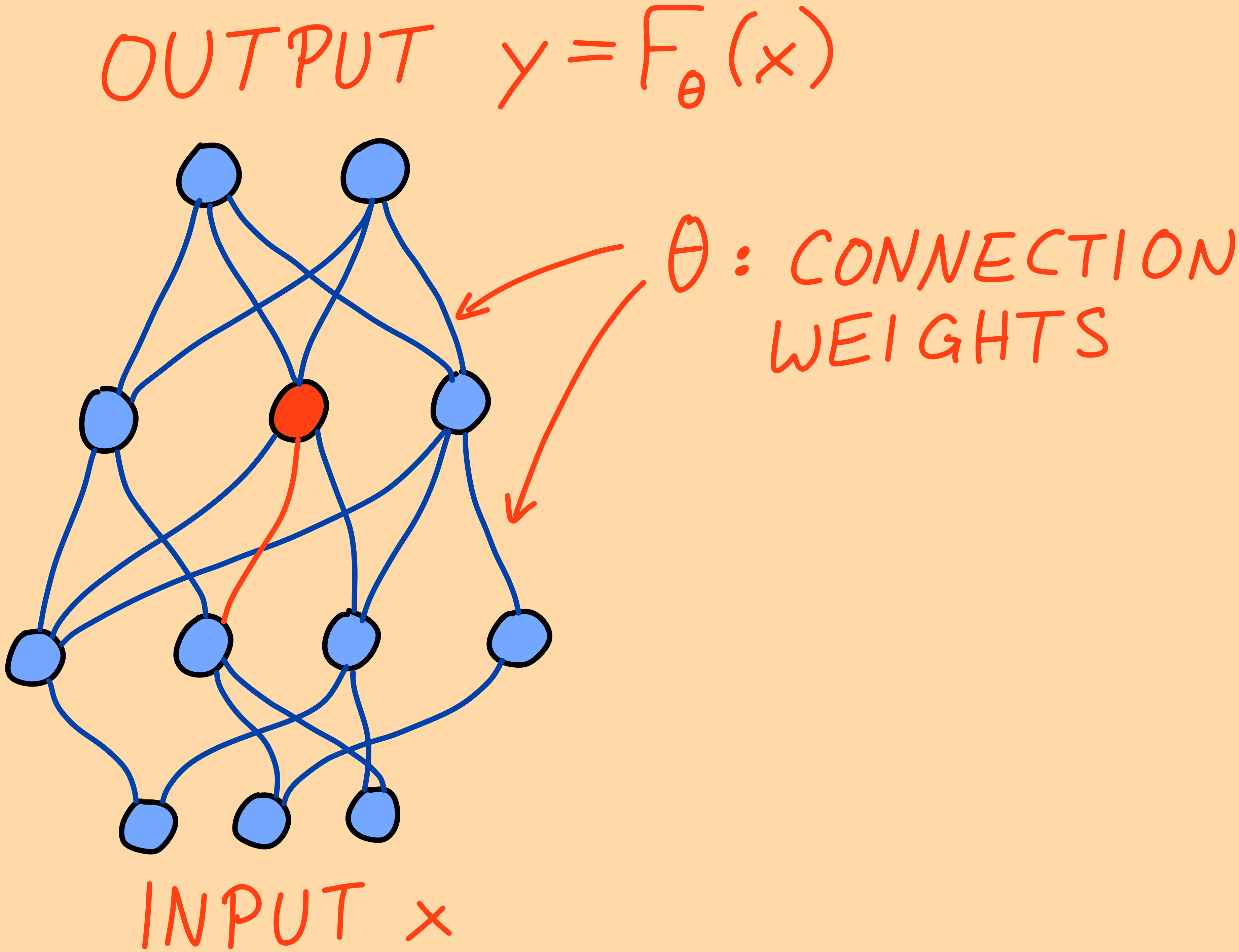
PHYSICAL LEARNING MACHINE



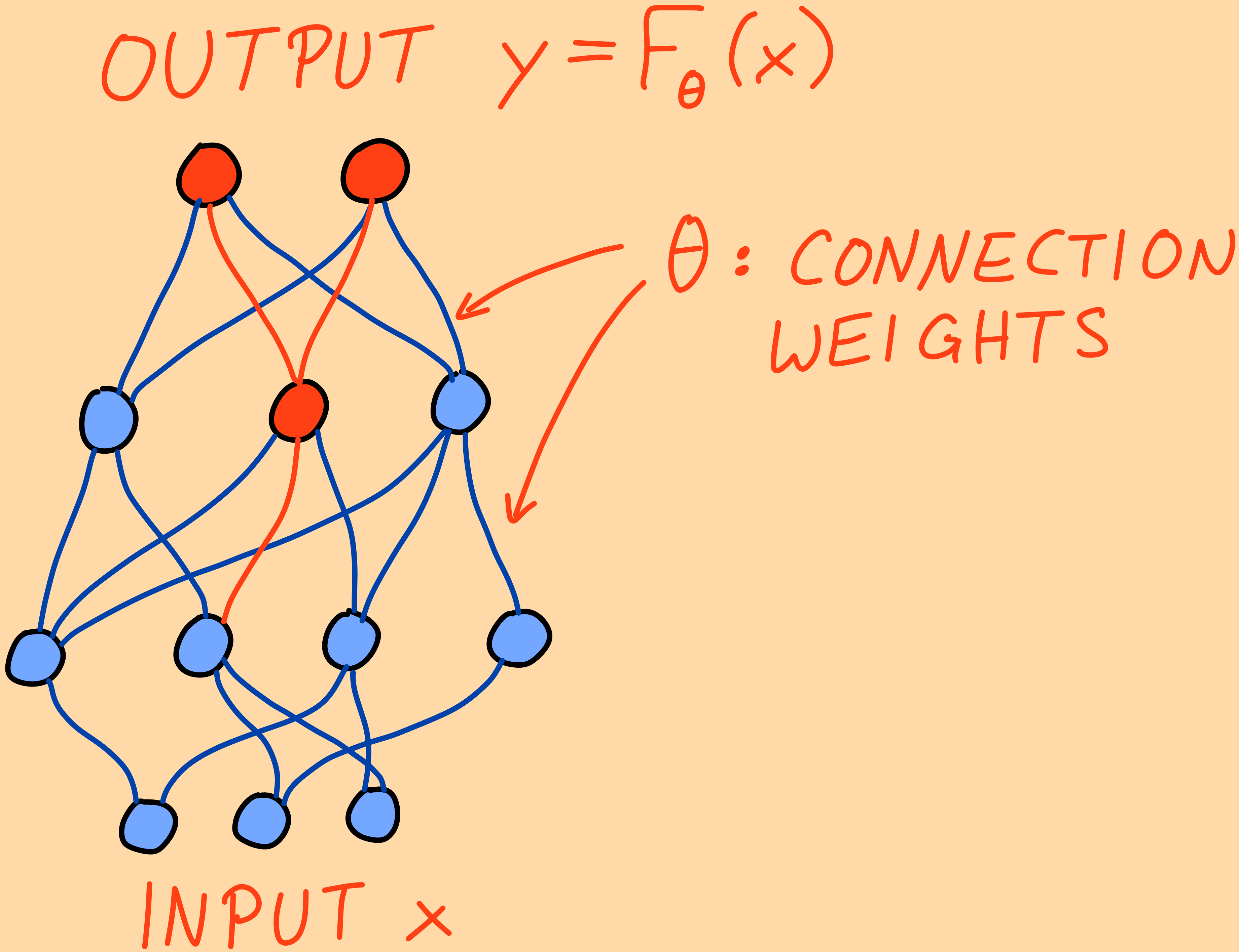
ARTIFICIAL NEURAL NETWORKS



ARTIFICIAL NEURAL NETWORKS

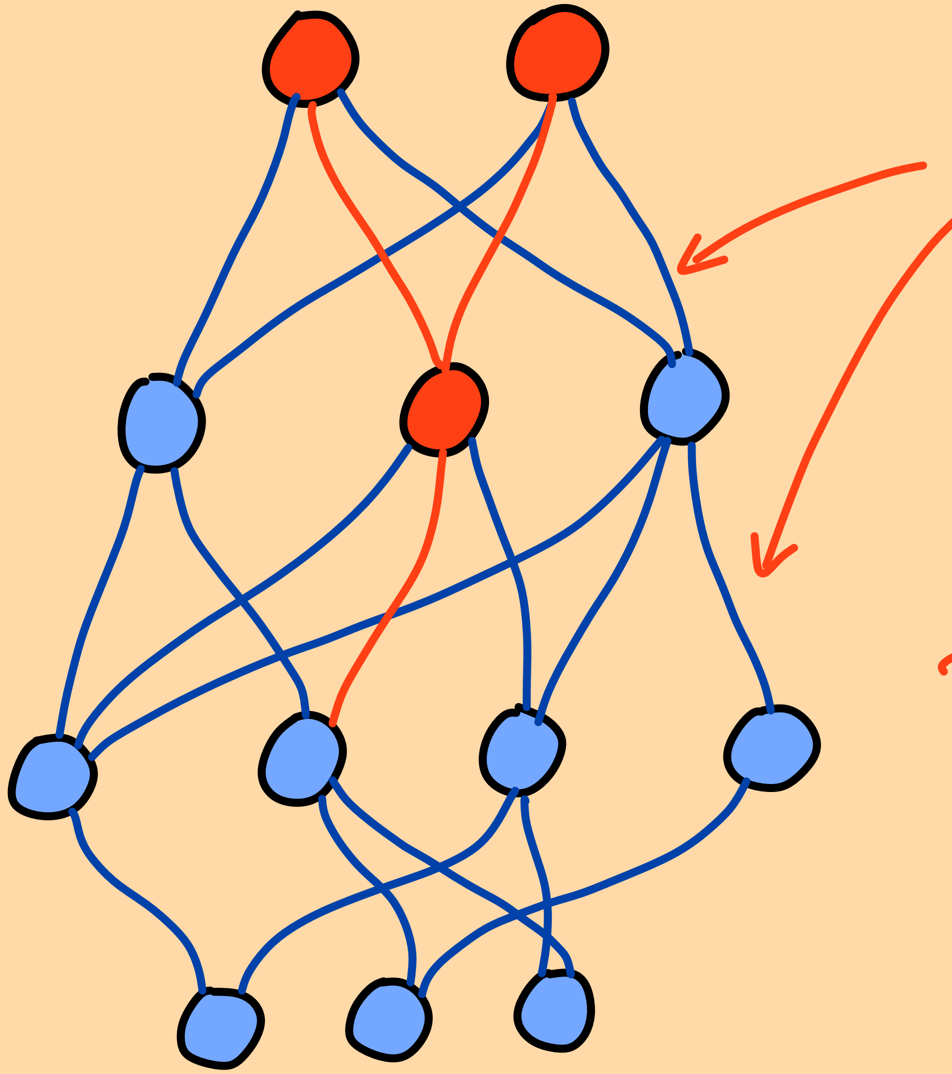


ARTIFICIAL NEURAL NETWORKS



ARTIFICIAL NEURAL NETWORKS

OUTPUT $y = F_\theta(x)$



INPUT x

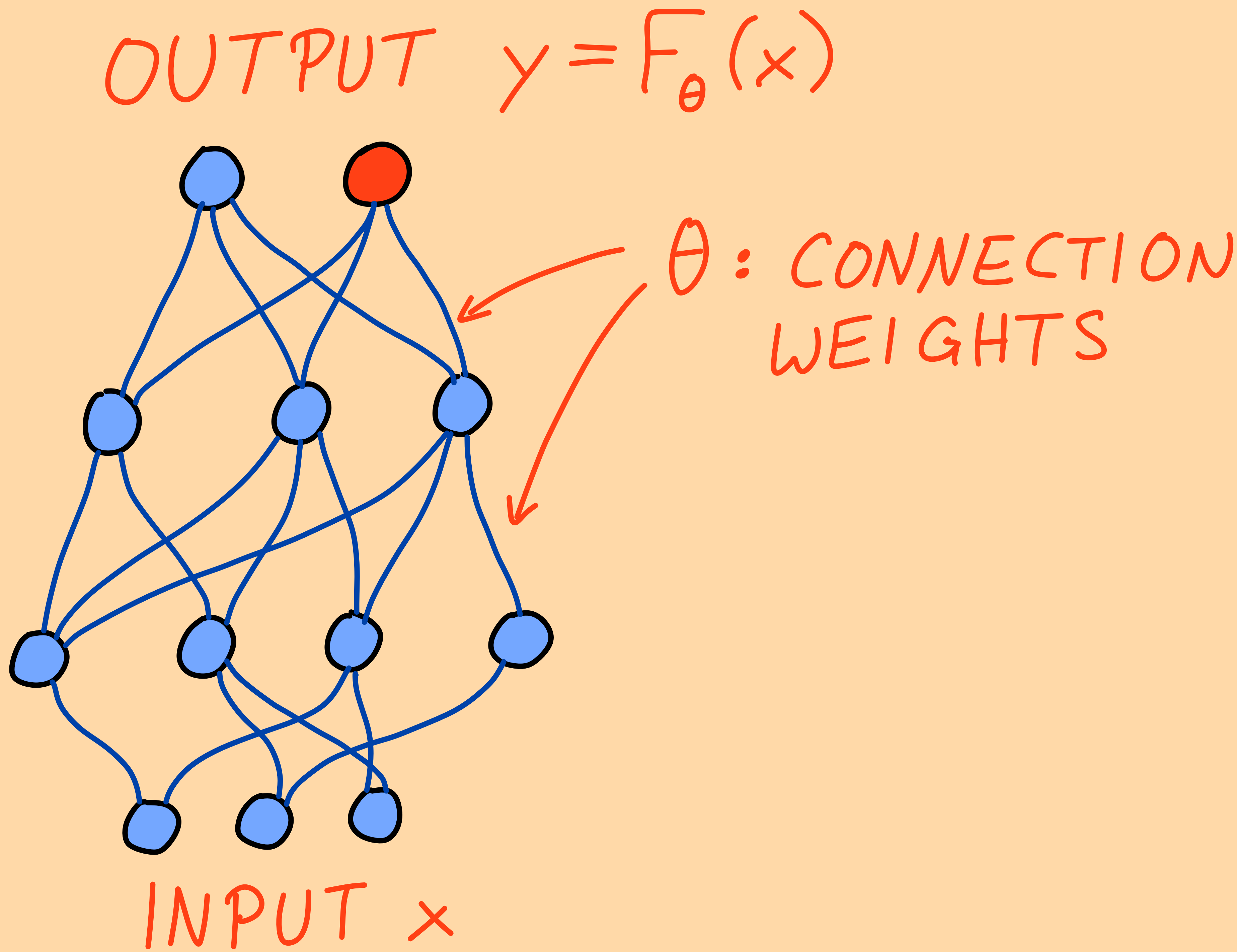
θ : CONNECTION WEIGHTS

PARAMETER SHIFT METHOD:

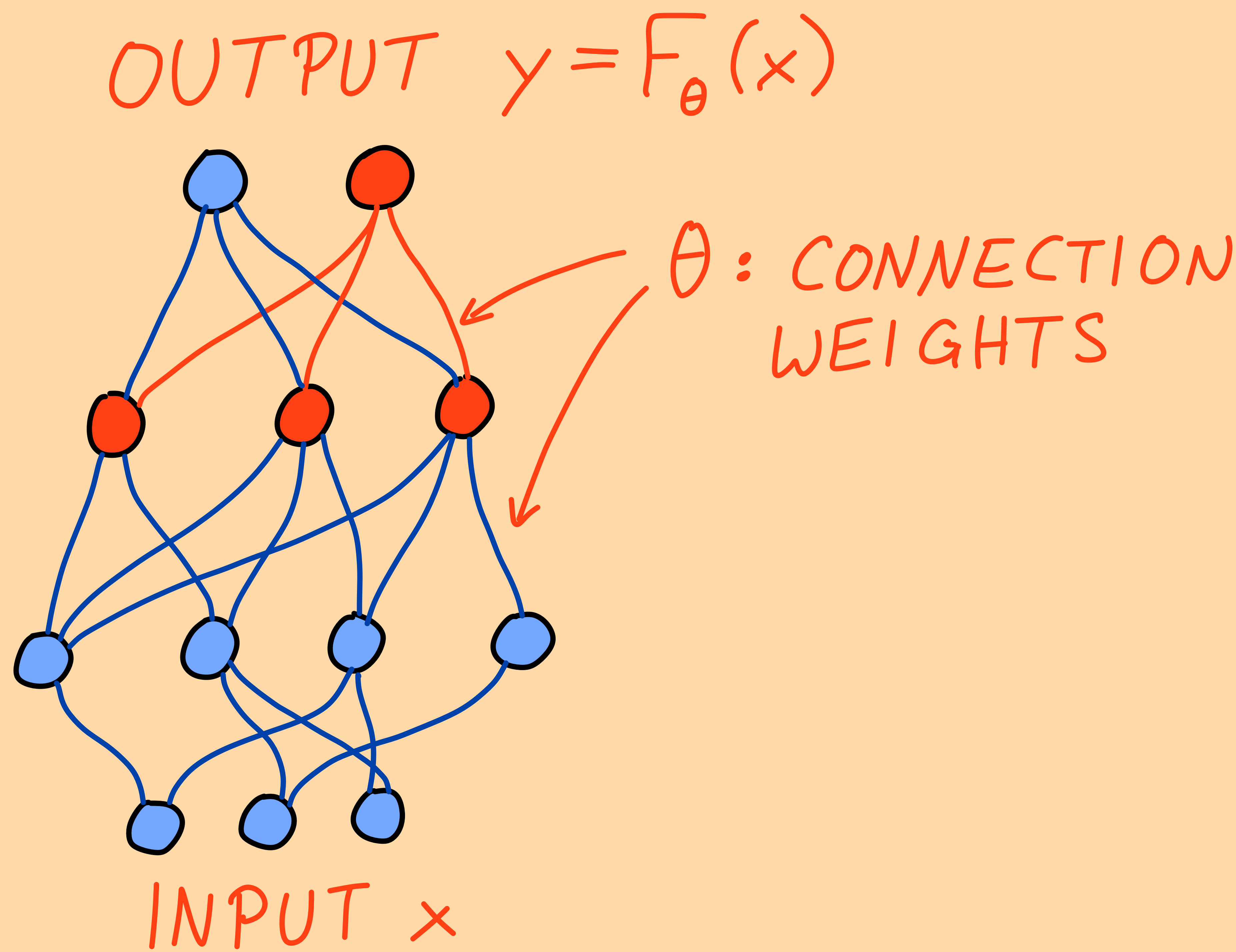
- ALWAYS POSSIBLE
- VERY <u>INEFFICIENT</u>

  EFFORT ~ #PARAMS

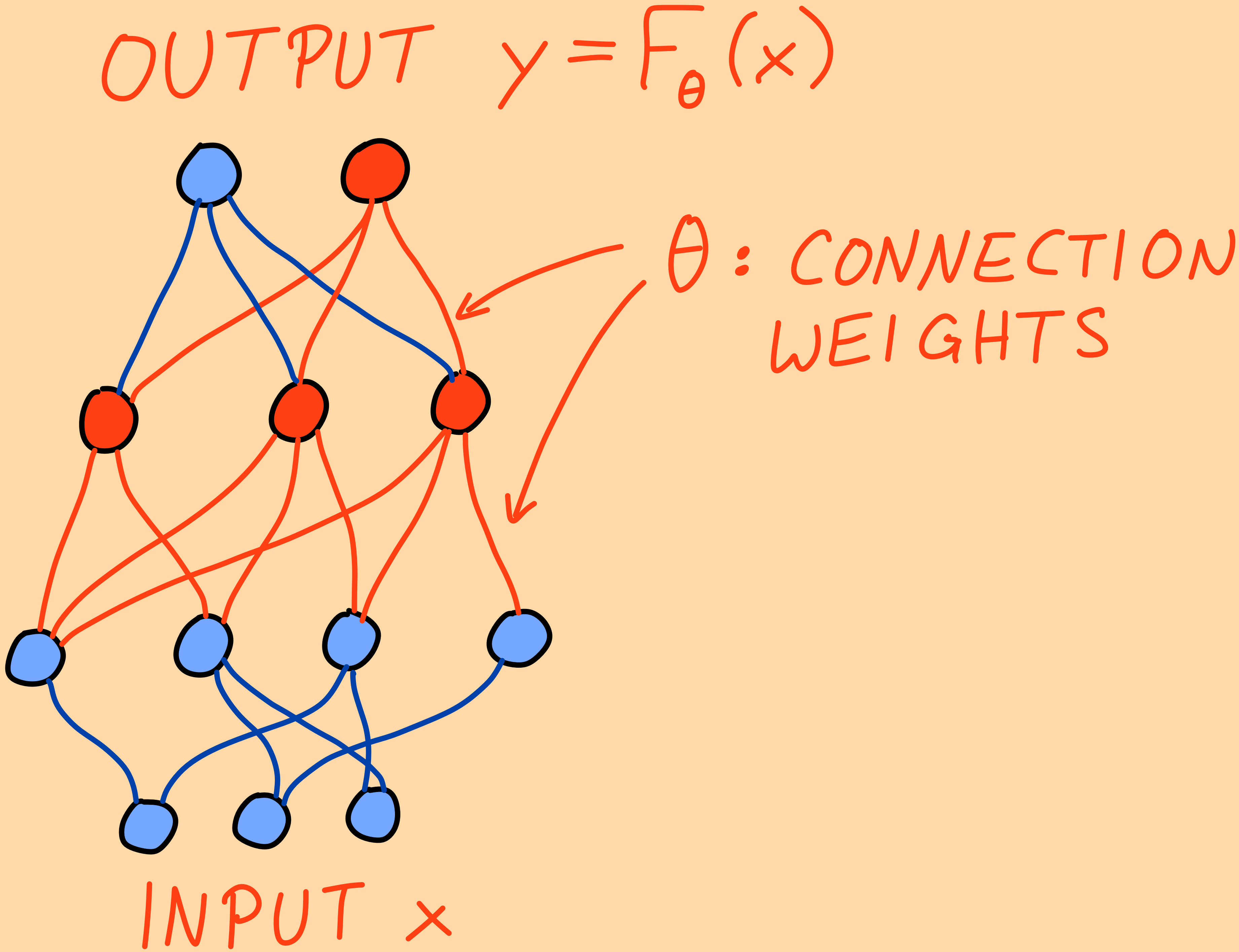
ARTIFICIAL NEURAL NETWORKS



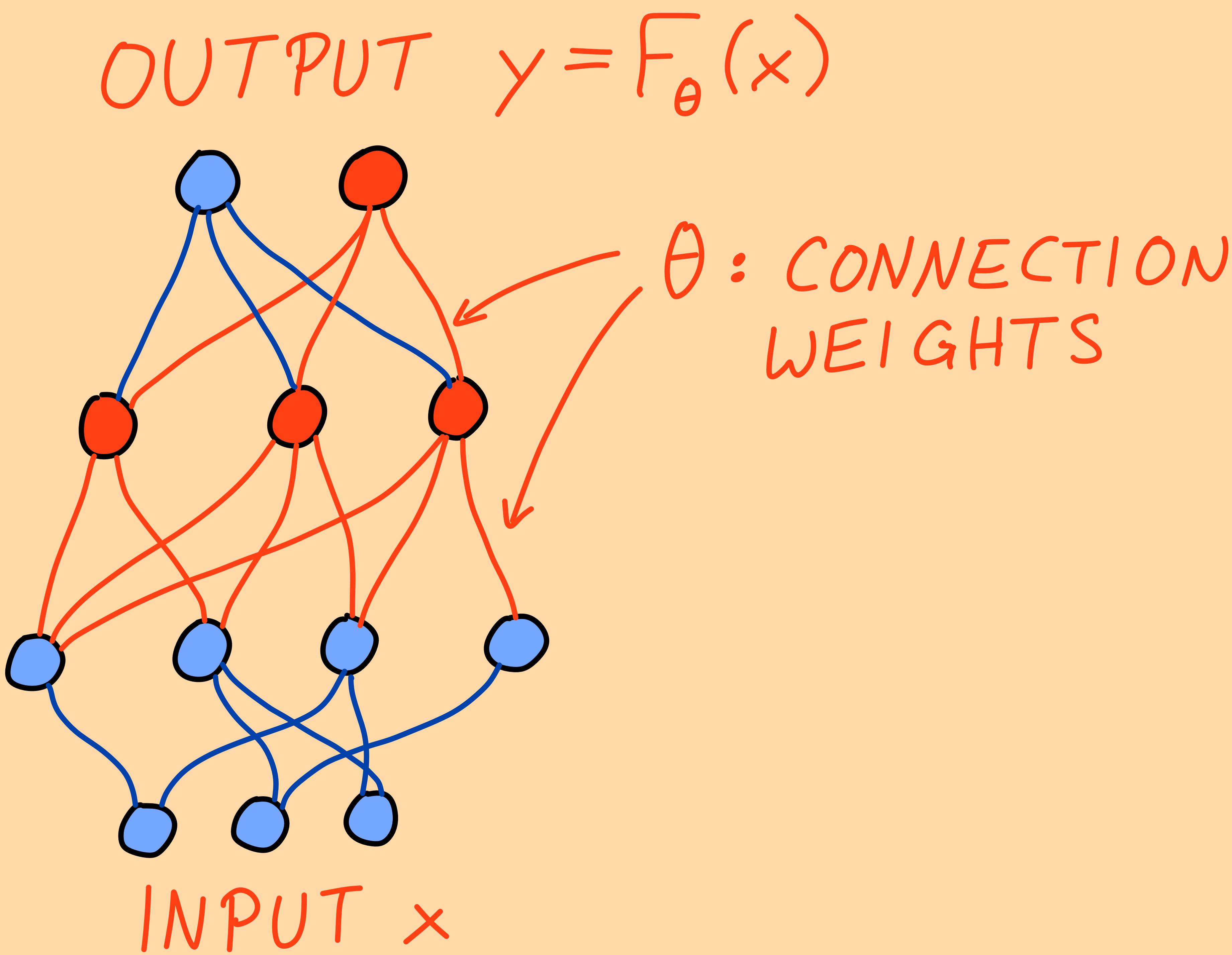
ARTIFICIAL NEURAL NETWORKS



ARTIFICIAL NEURAL NETWORKS

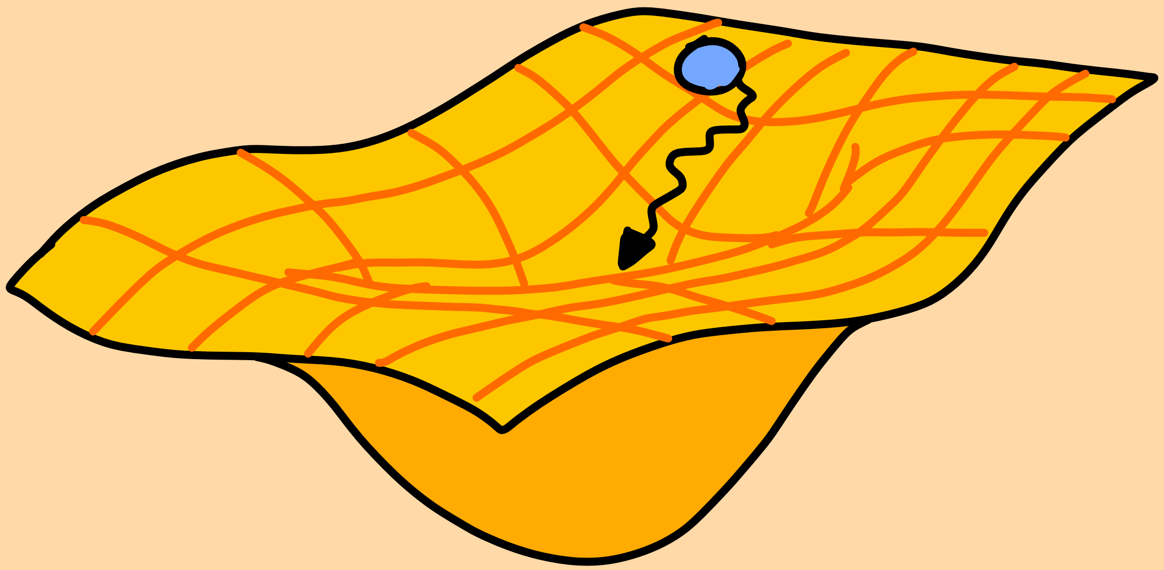


ARTIFICIAL NEURAL NETWORKS



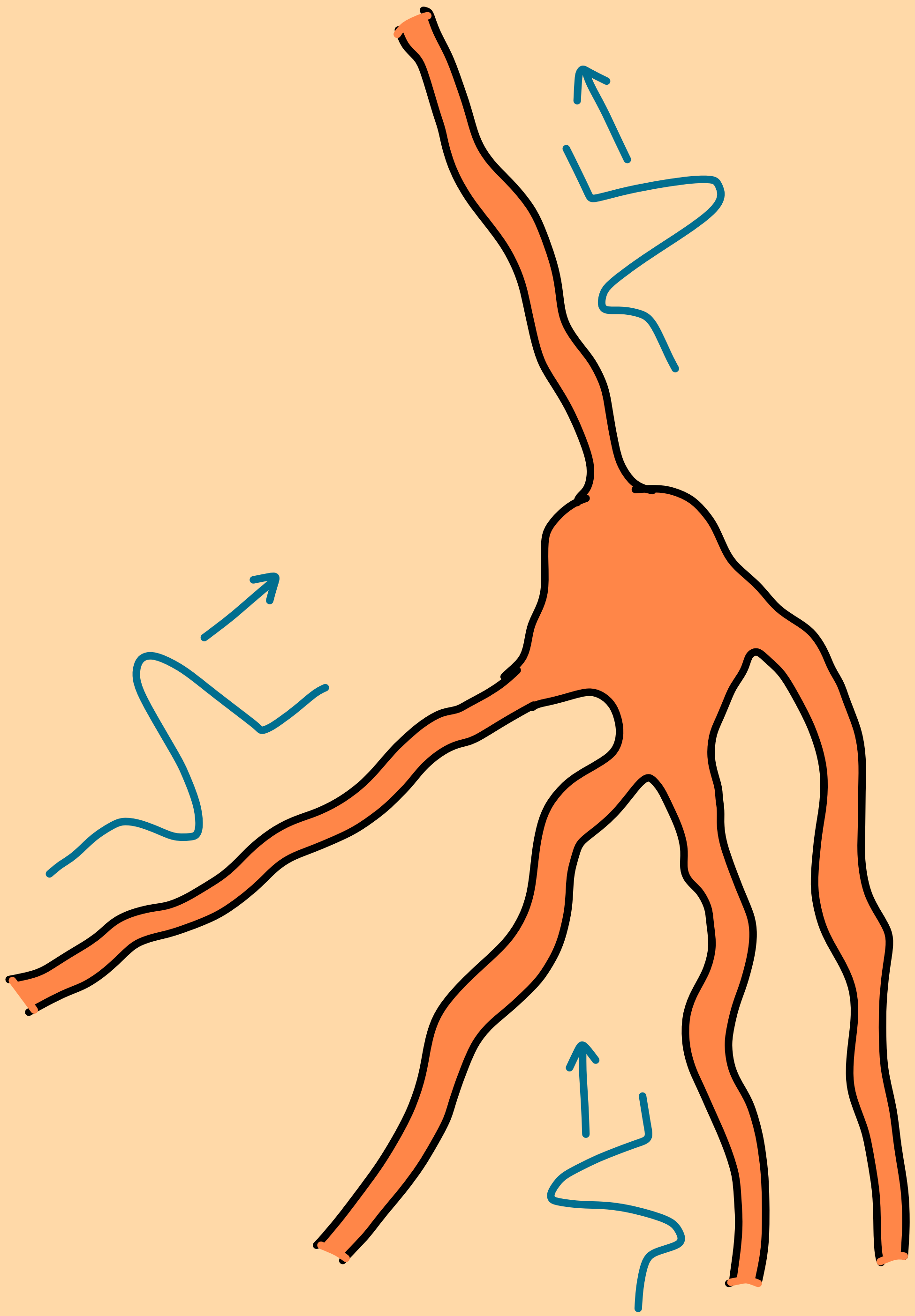
"BACKPROPAGATION" : <u>ALL</u> GRADIENTS FOR PRICE OF ONE EVALUATION

OPTIMIZATION-BASED RULES



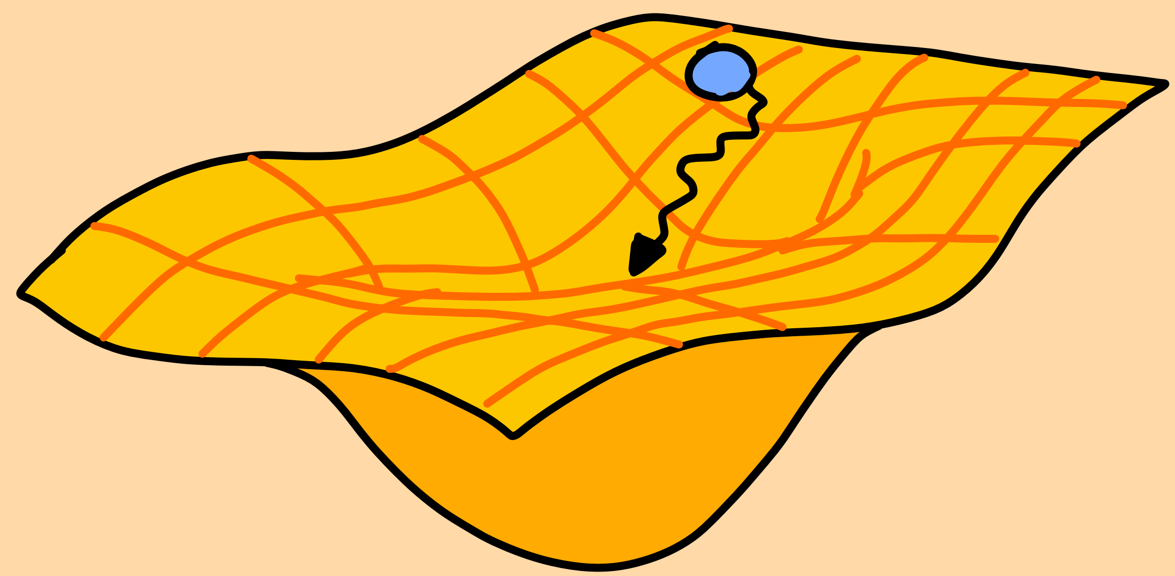
GRADIENT DESCENT OF COST FUNCTION

BIOLOGY-INSPIRED LEARNING RULES



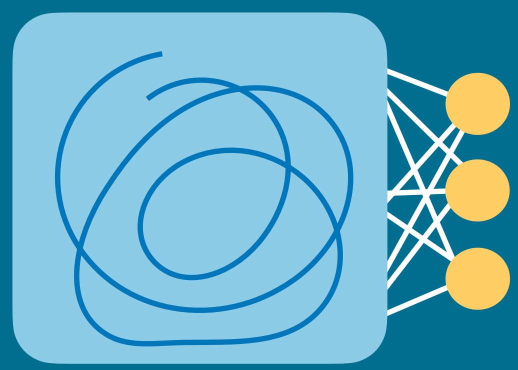
"NEURONS THAT FIRE TOGETHER WIRE TOGETHER"

OPTIMIZATION-BASED RULES

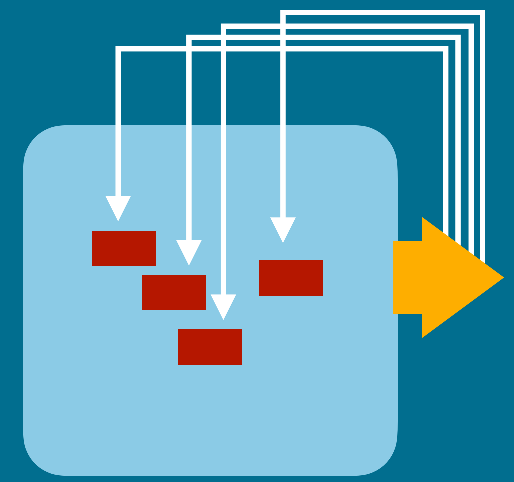


GRADIENT DESCENT OF COST FUNCTION

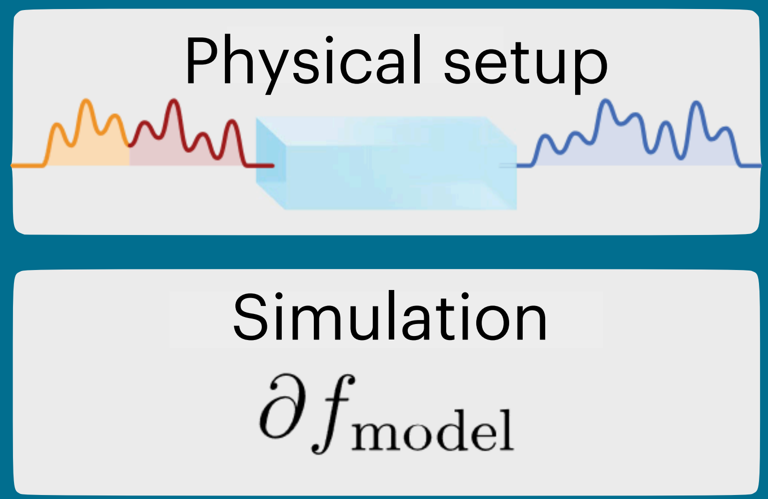
Training neuromorphic devices



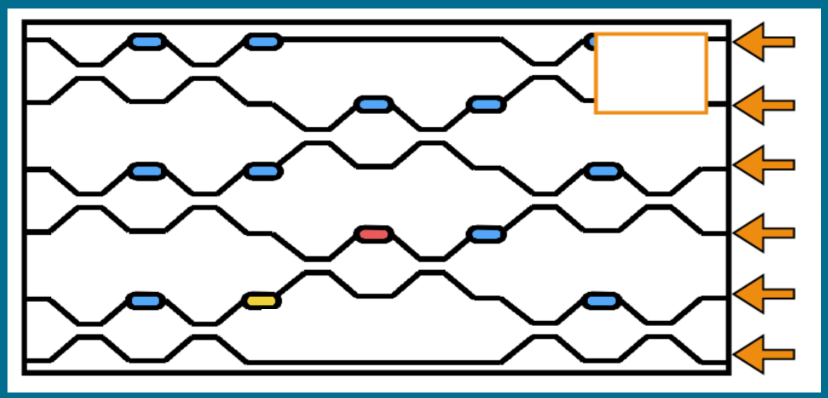
Reservoir computing



Parameter shift method

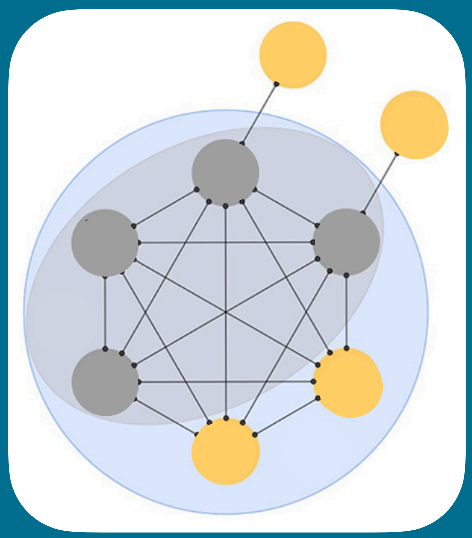


Hybrid method
Wright,Onodera,...,
McMahon 2022



Physical backprop. for special systems

Wagner & Psaltis 1987
Hughes,...,Fan 2018
Guo,...,Lvovsky 2021



Equilibrium Propagation

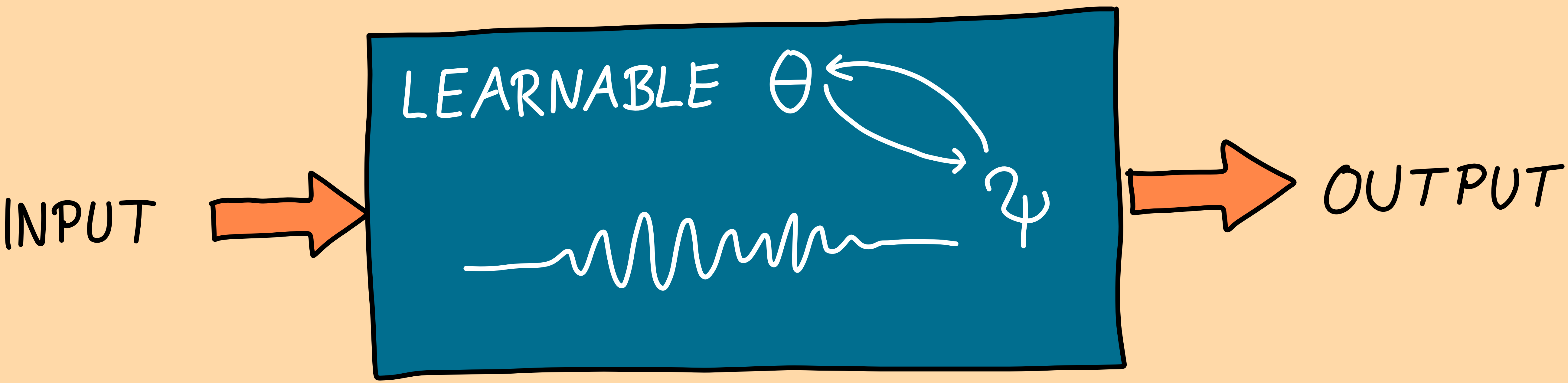
Scellier,..., Bengio 2017

Stern,...,Liu 2021

Physical backpropagation
and
Physical parameter update

PHYSICAL SELF-LEARNING MACHINE

AUTONOMOUS, NO FEEDBACK



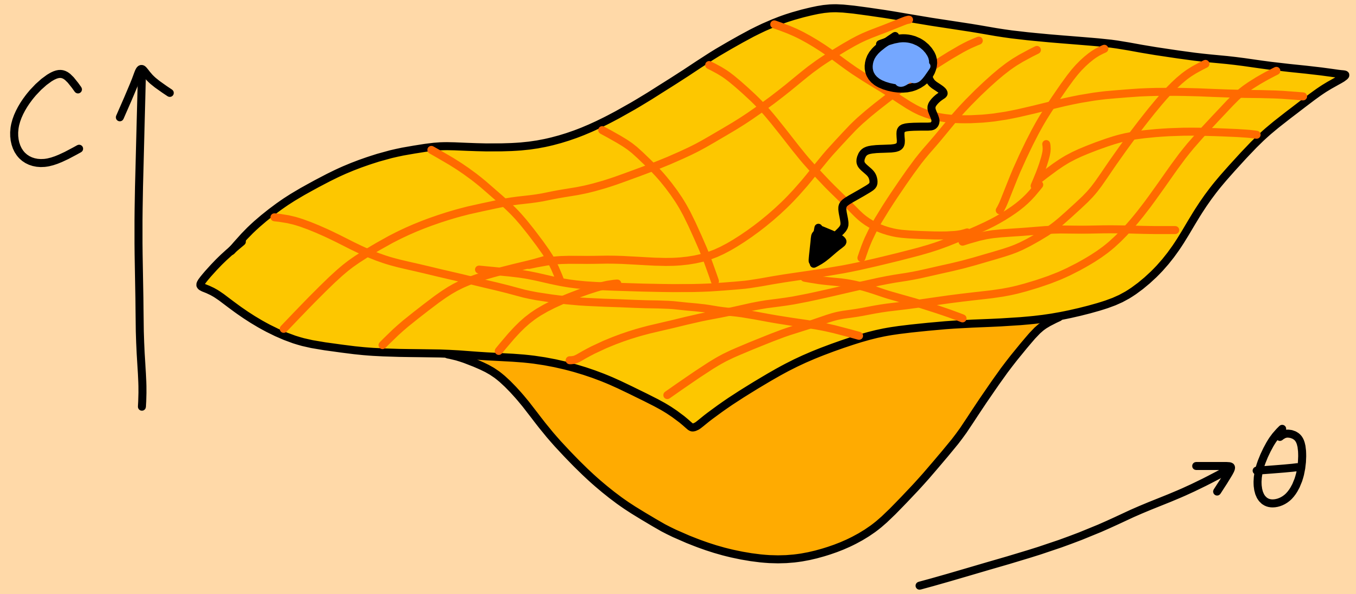
WE WANT: PHYSICAL BACKPROPAGATION AND PHYSICAL LEARNING UPDATE

GOAL: MINIMIZE COST FUNCTION

$$C = (\text{OUTPUT} - \text{TARGET})^2$$

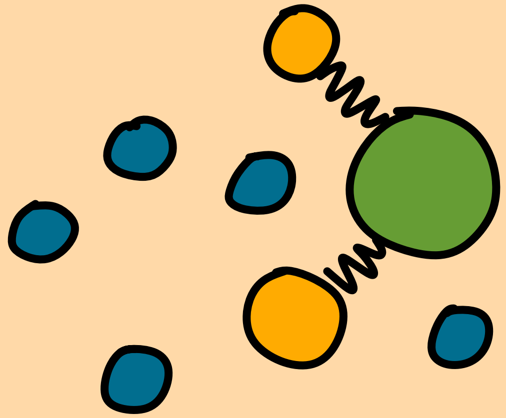
$$\theta_{NEW} = \theta_{OLD} - \eta \frac{\partial C}{\partial \theta}$$

GRADIENT DESCENT



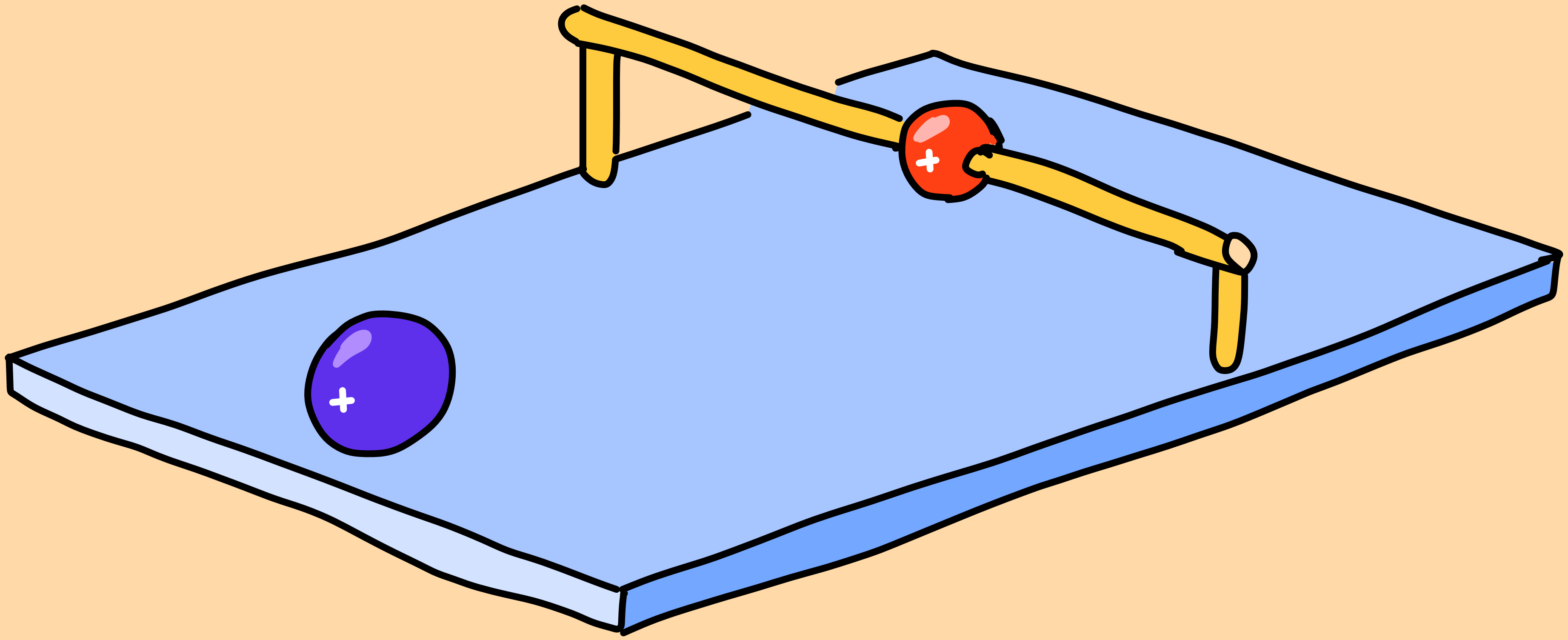
HERE:

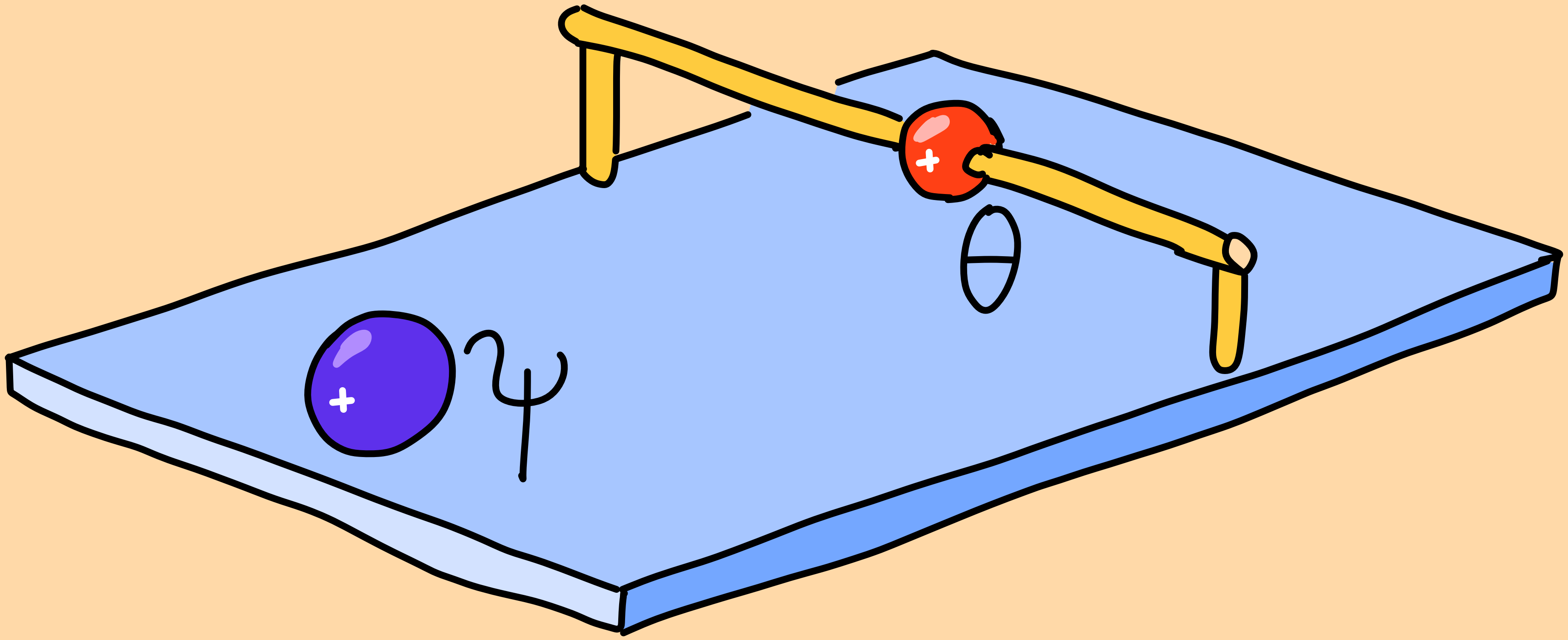
ARBITRARY TIME-REVERSAL-INVARIANT HAMILTONIAN

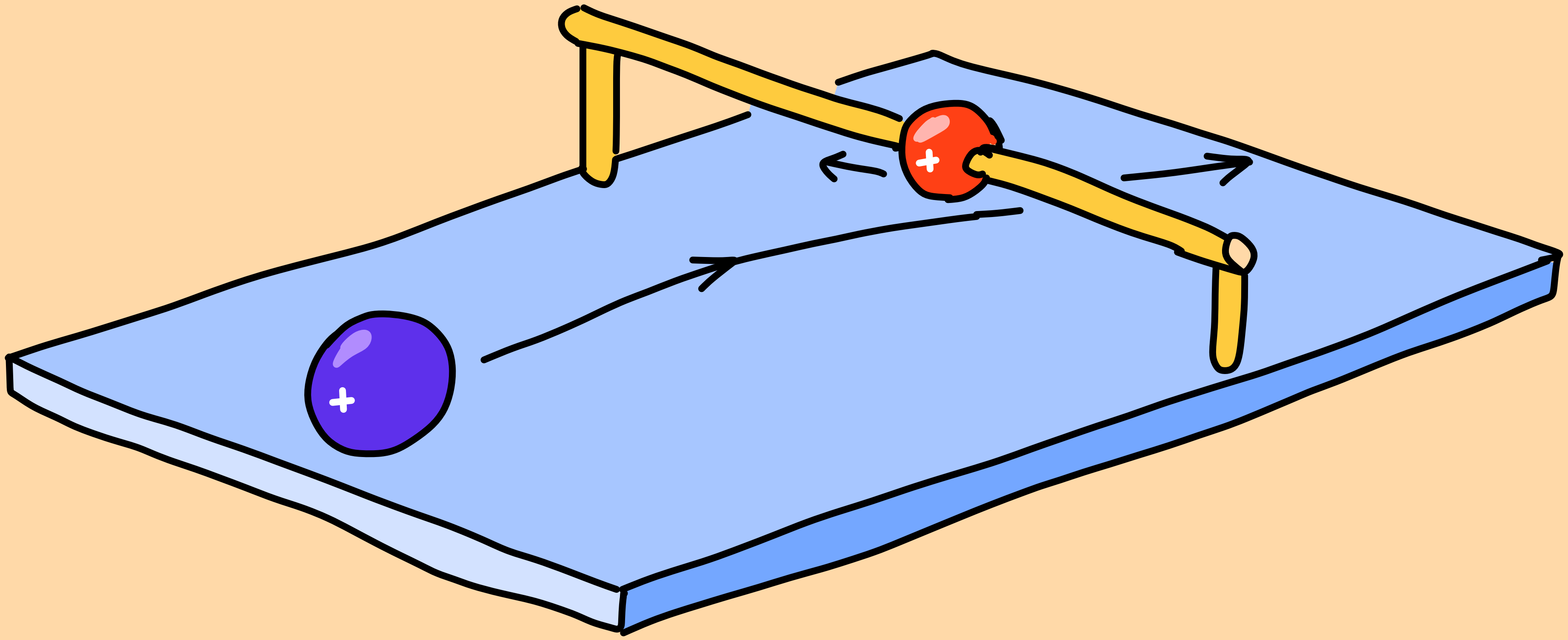


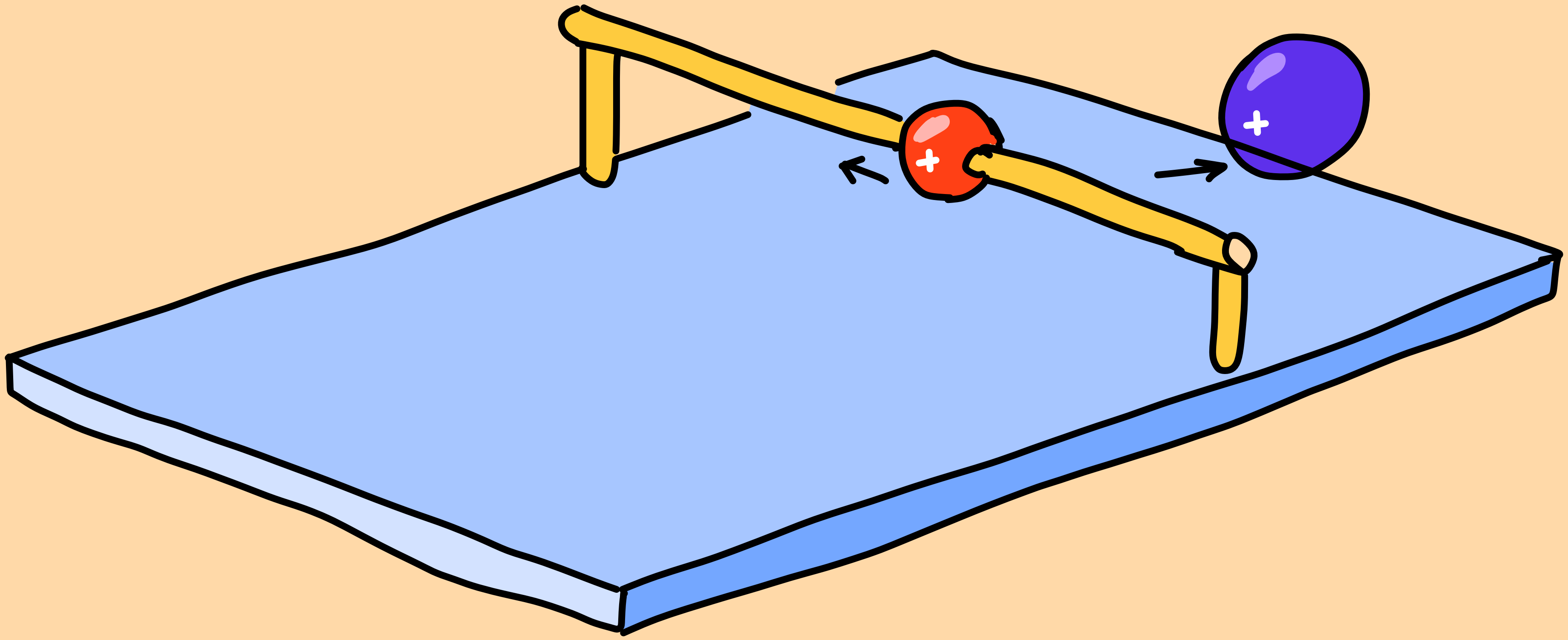
WITH TIME-REVERSAL OPERATION (!)

THE SELF-LEARNING PINBALL MACHINE

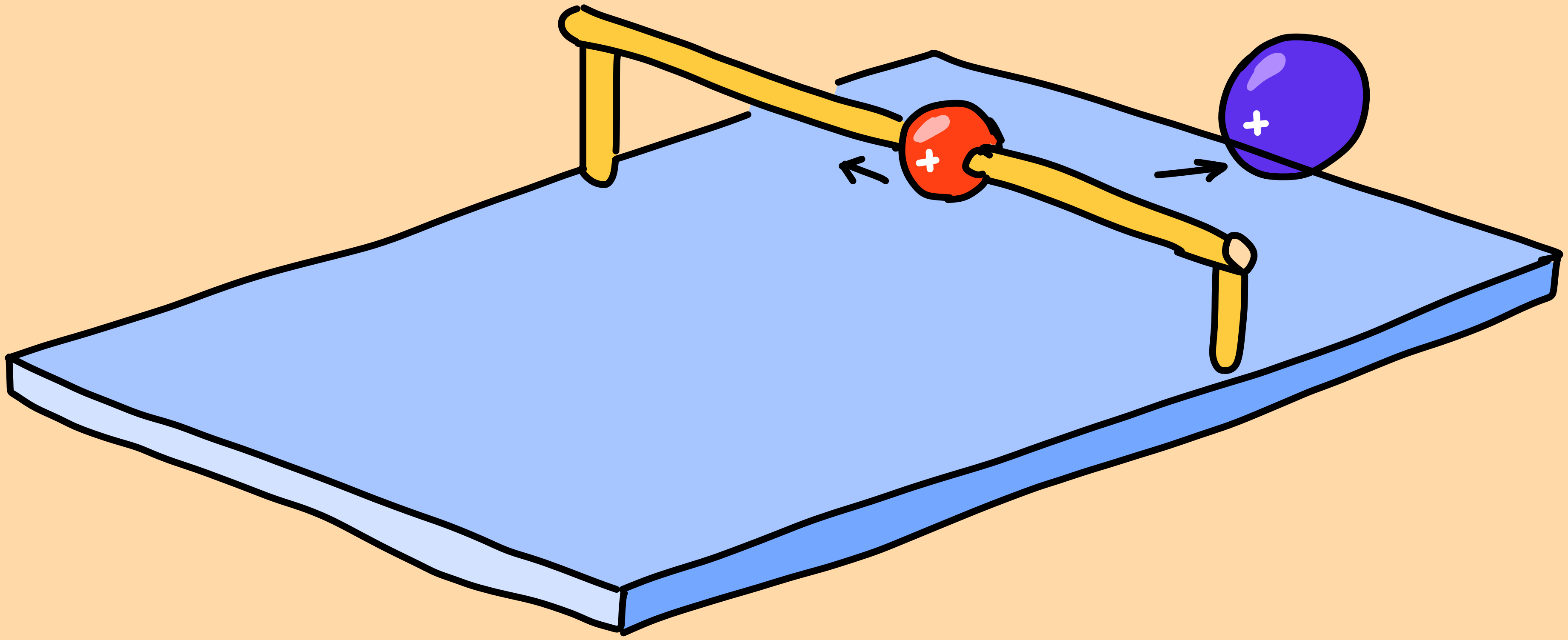


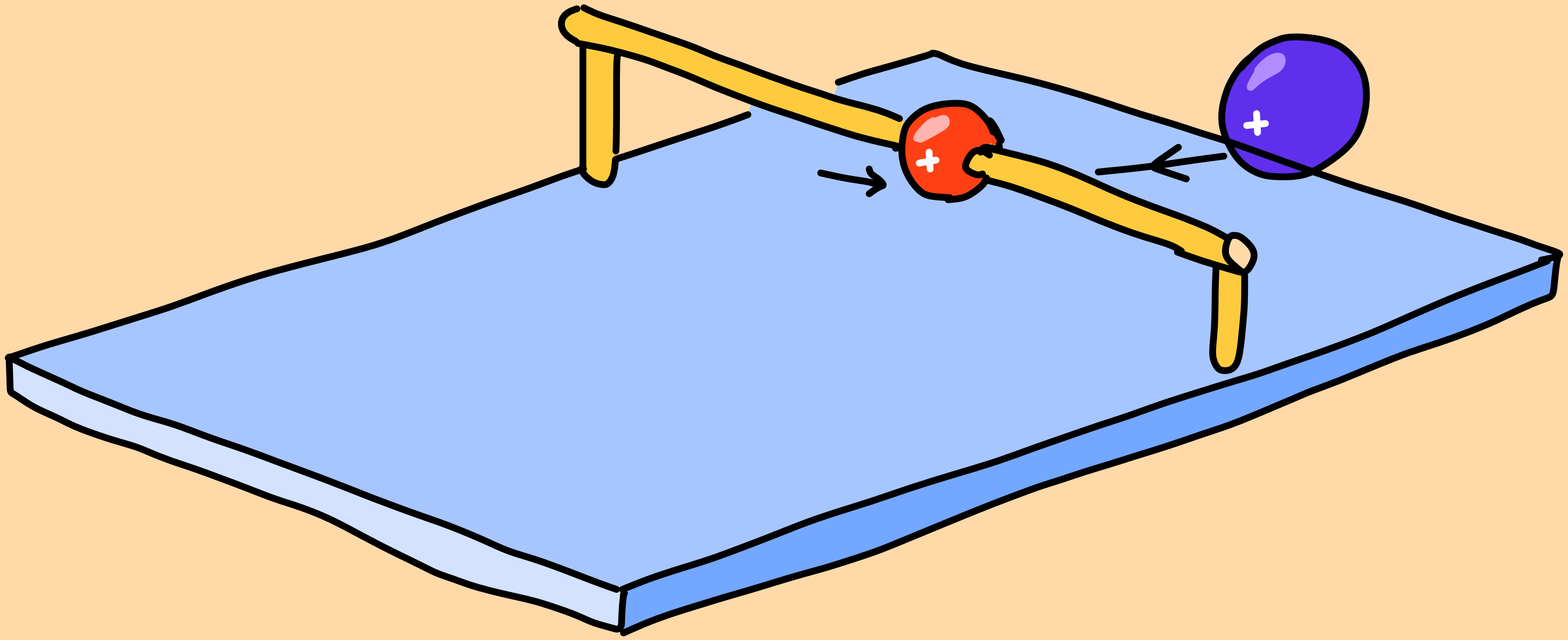

ψ
θ

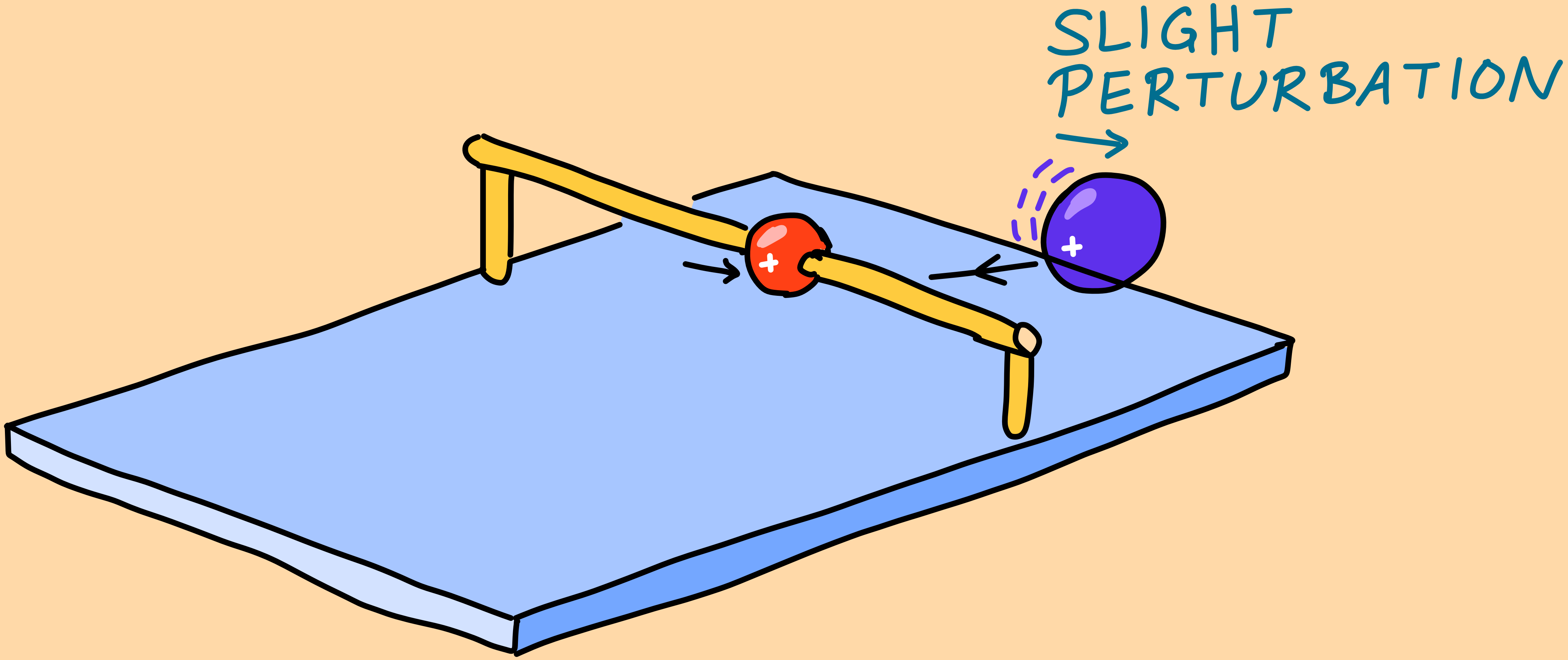


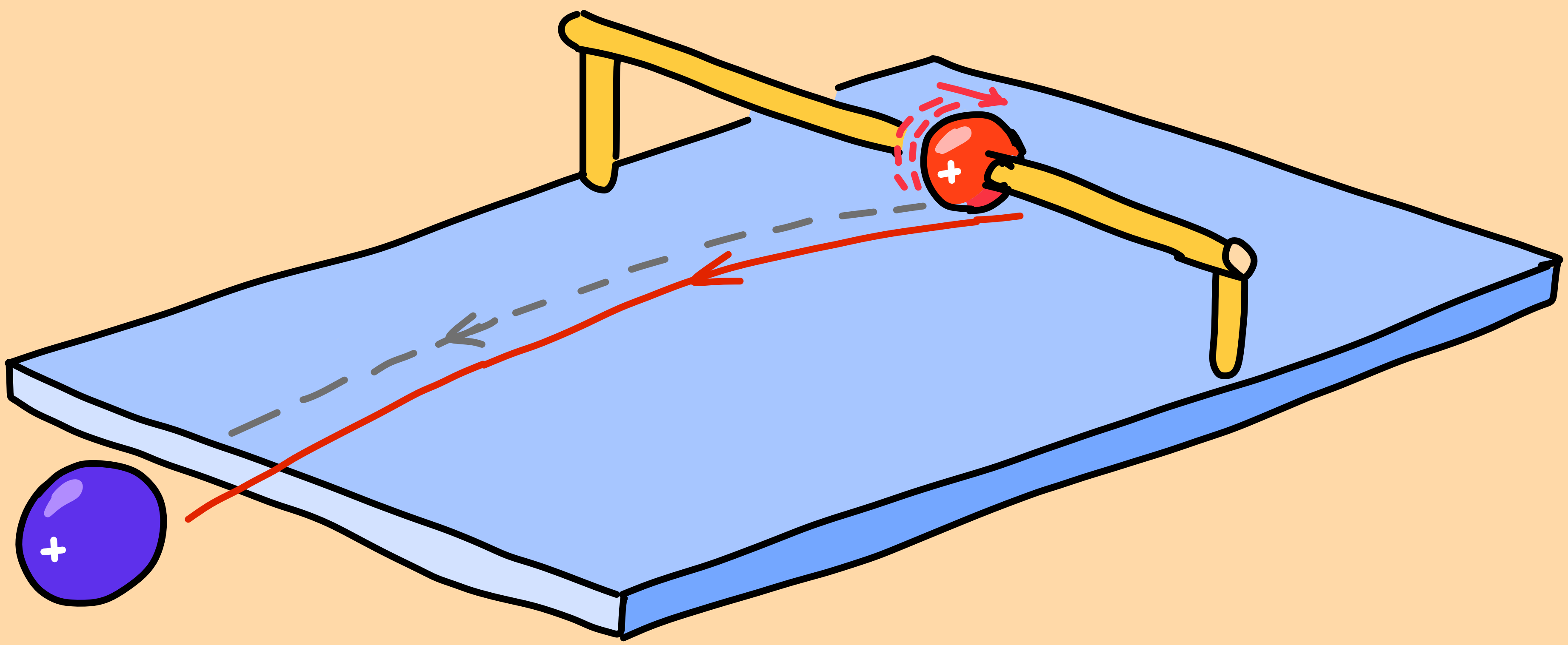


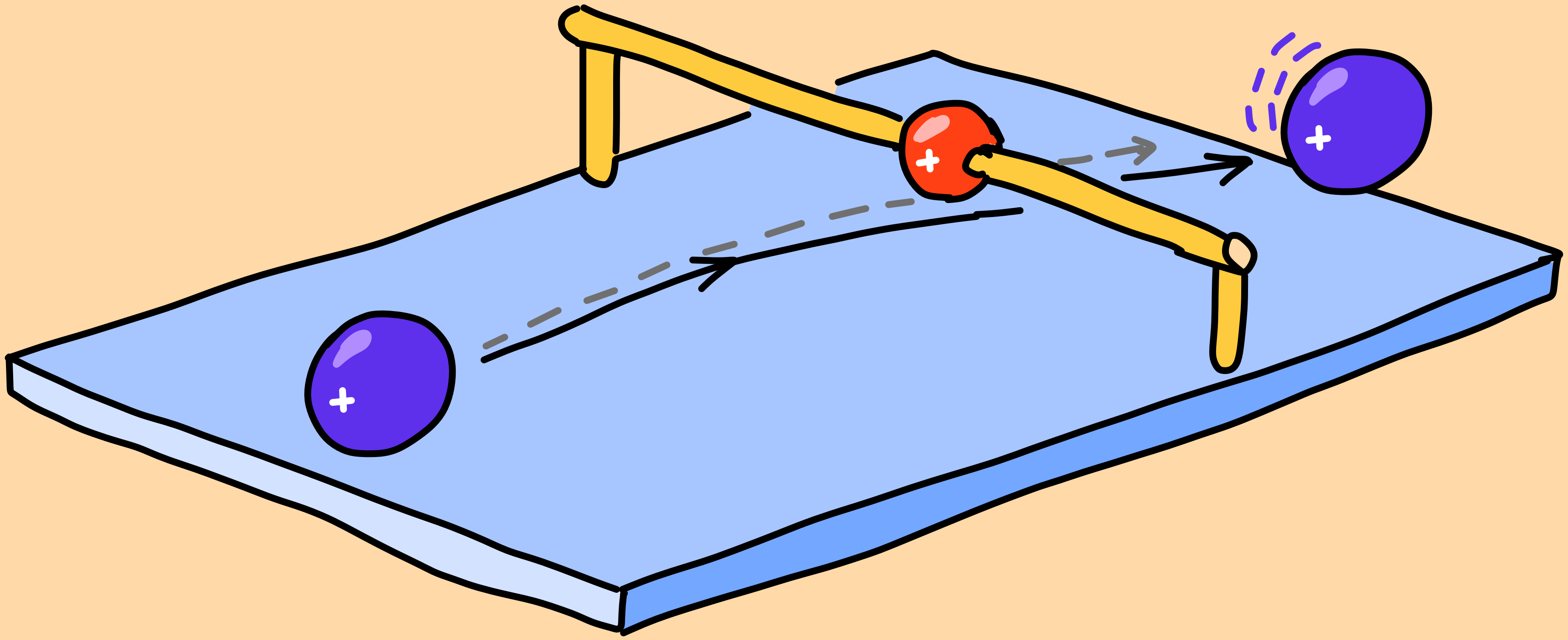
NOW: TIME-REVERSAL OPERATION



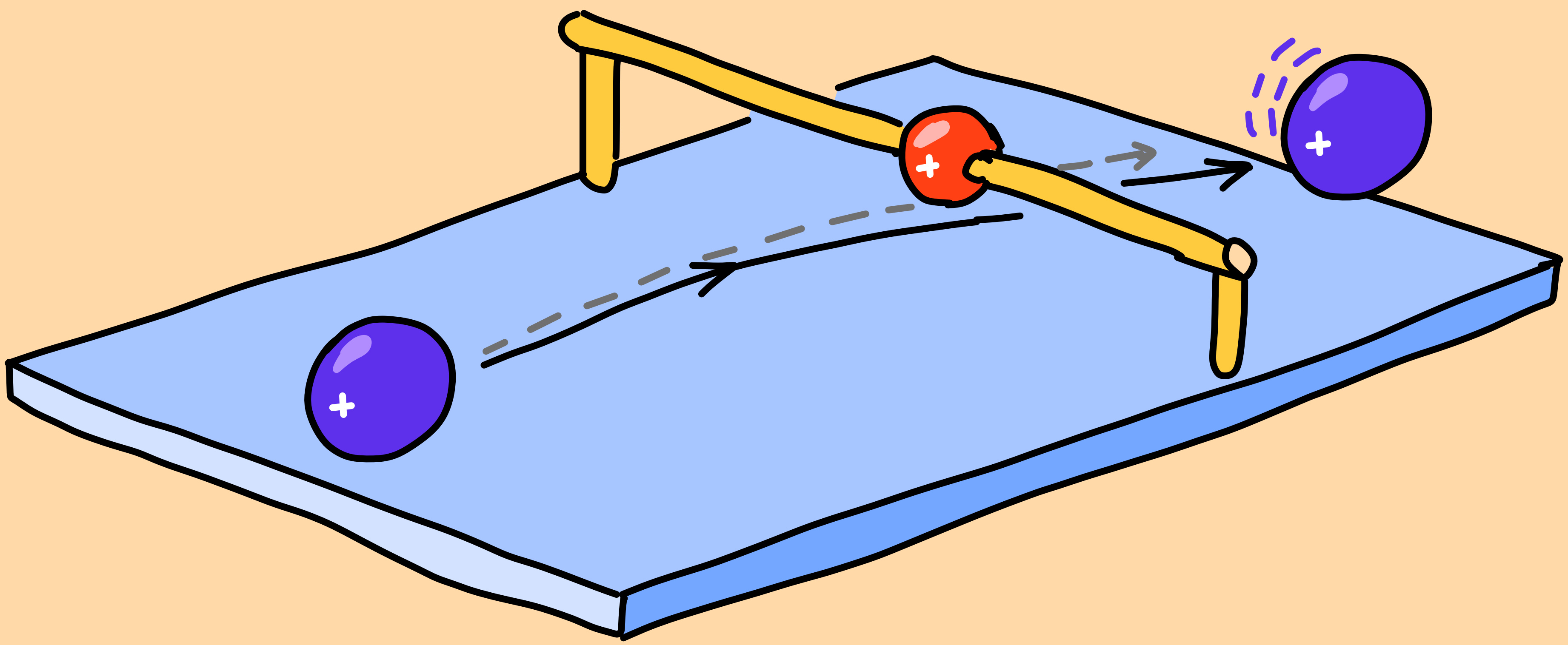


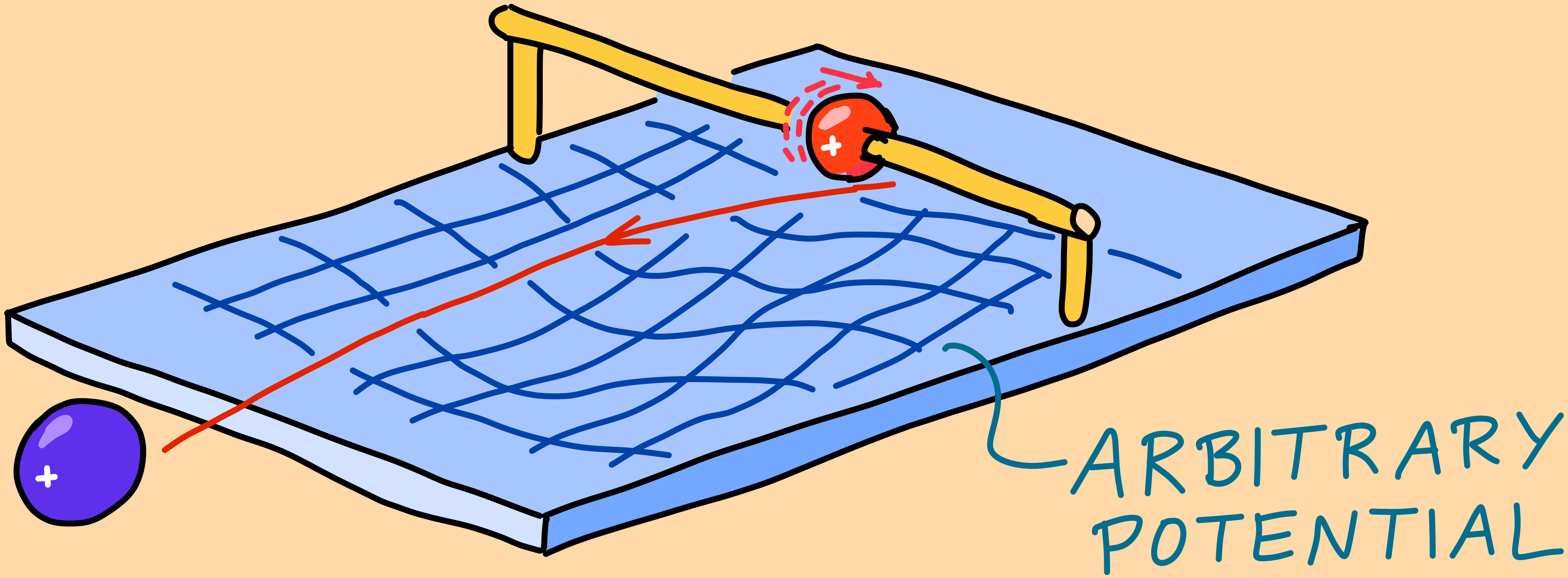

SLIGHT
PERTURBATION



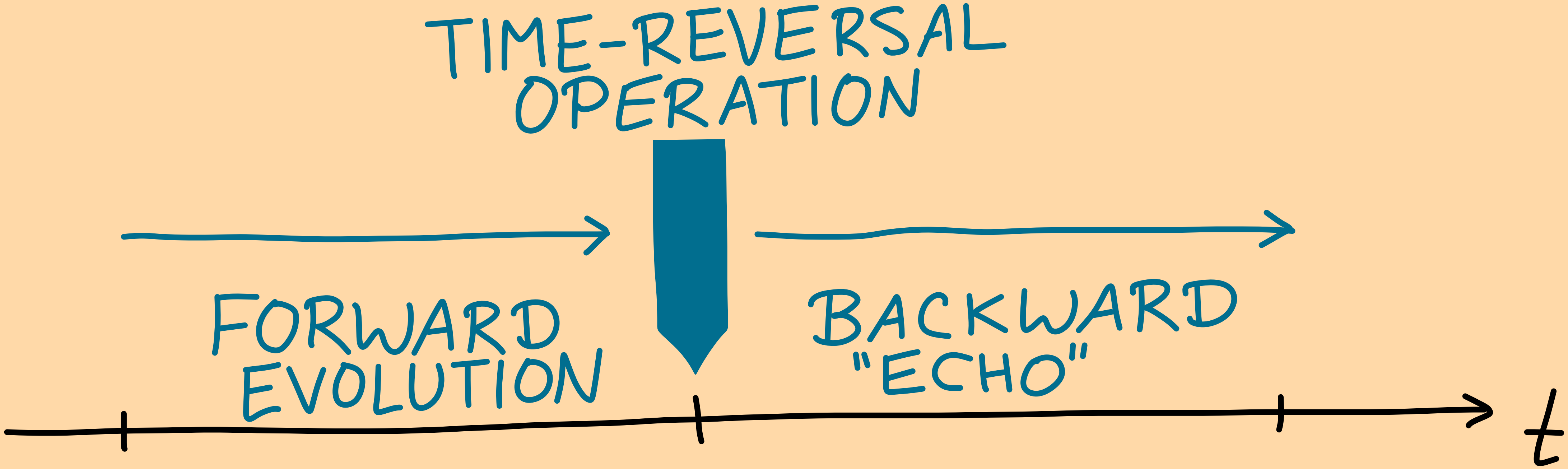


DYNAMICS CHANGED IN THE RIGHT WAY ⇒ "LEARNING"!

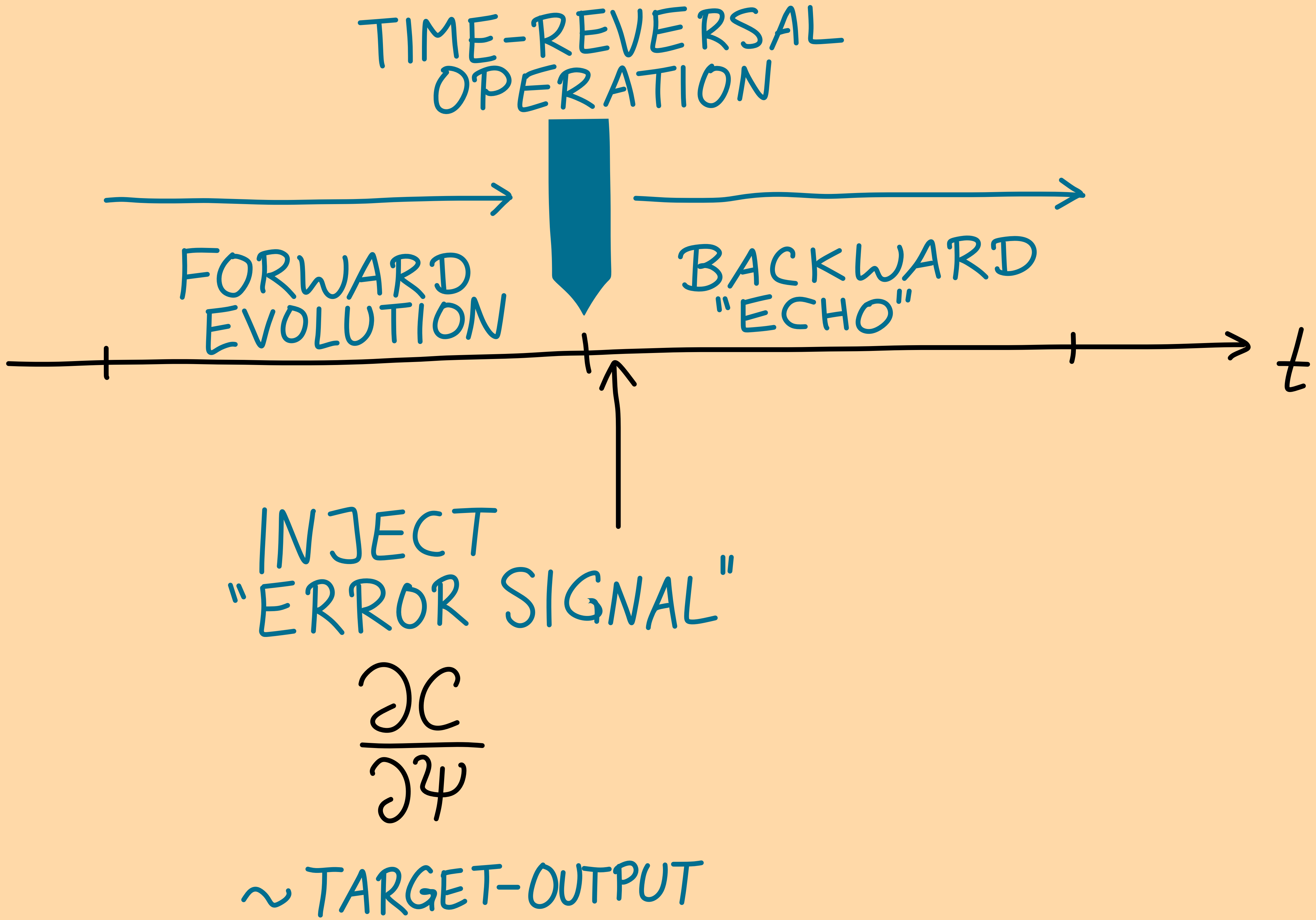



ARBITRARY
POTENTIAL

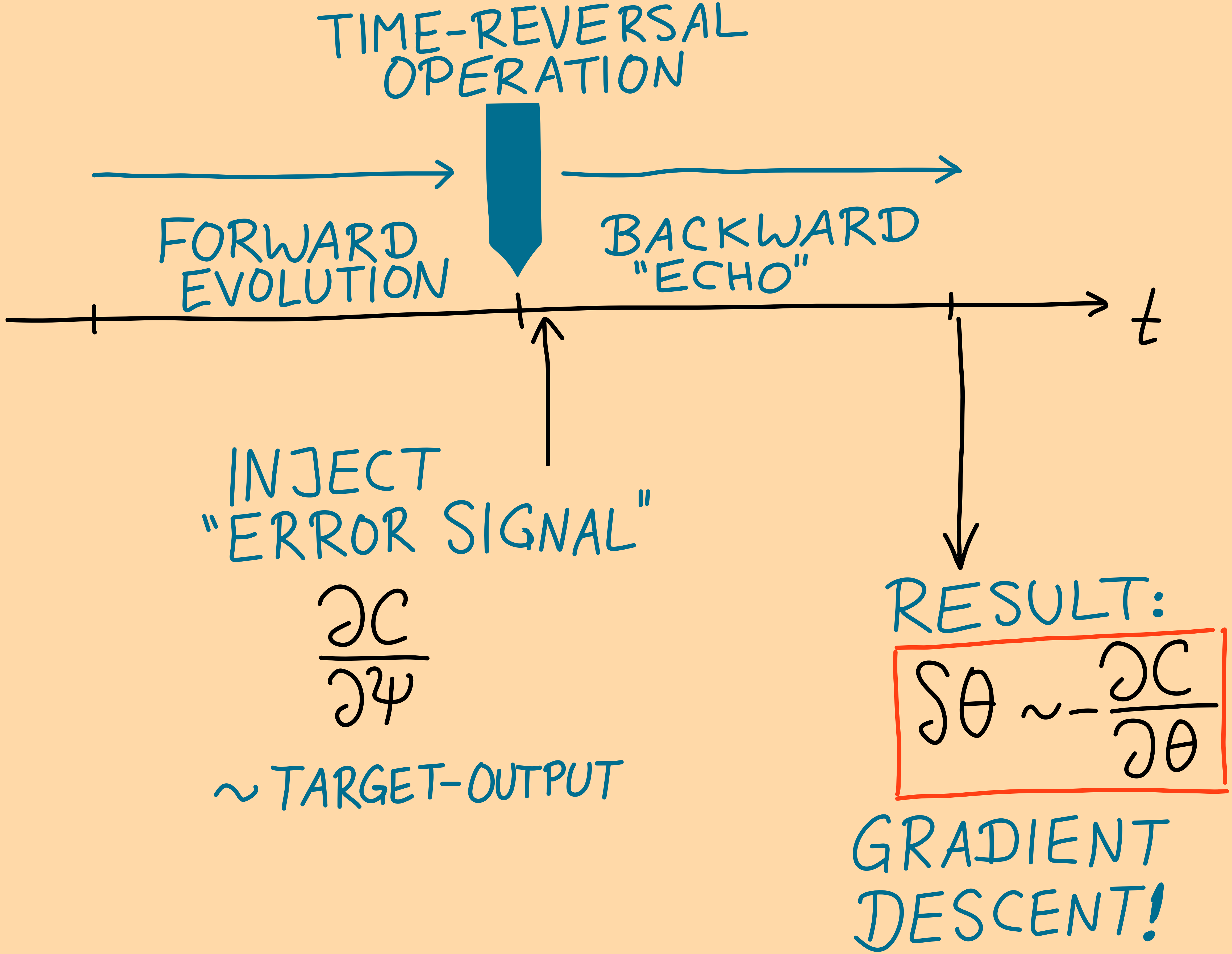
HAMILTONIAN ECHO BACKPROPAGATION



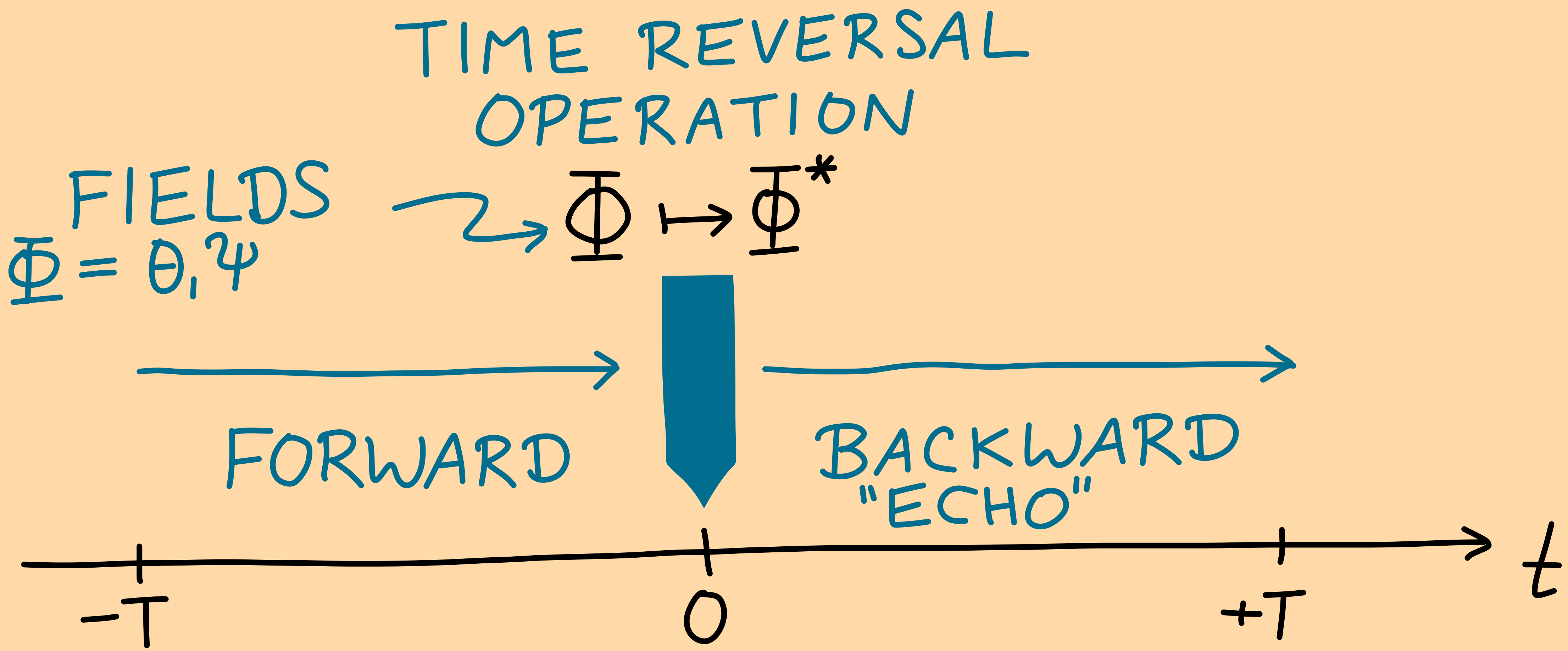
HAMILTONIAN ECHO BACKPROPAGATION



HAMILTONIAN ECHO BACKPROPAGATION



BRIEFLY: THE MATH



TIME-REVERSAL INVARIANCE

$$\Rightarrow \quad \Phi_{ECHO}(t) = \Phi^*(-t)$$

Victor Lopez-Pastor & F.M. , Phys. Rev. X 13, 031020 (2023)

$$\frac{\partial C}{\partial \theta} = ?$$

NEEDED:

PERTURBED FORWARD EVOLUTION

$$\mathcal{G}_{\Phi}(0,-T)^{+}$$

GREEN'S FUNCTION

ACCESSIBLE:

PERTURBED BACKWARD EVOLUTION

$$\mathcal{G}_{\Phi_{ECHO}}(T,0)$$

RELATED DUE TO TIME-REVERSAL-INVARIANCE

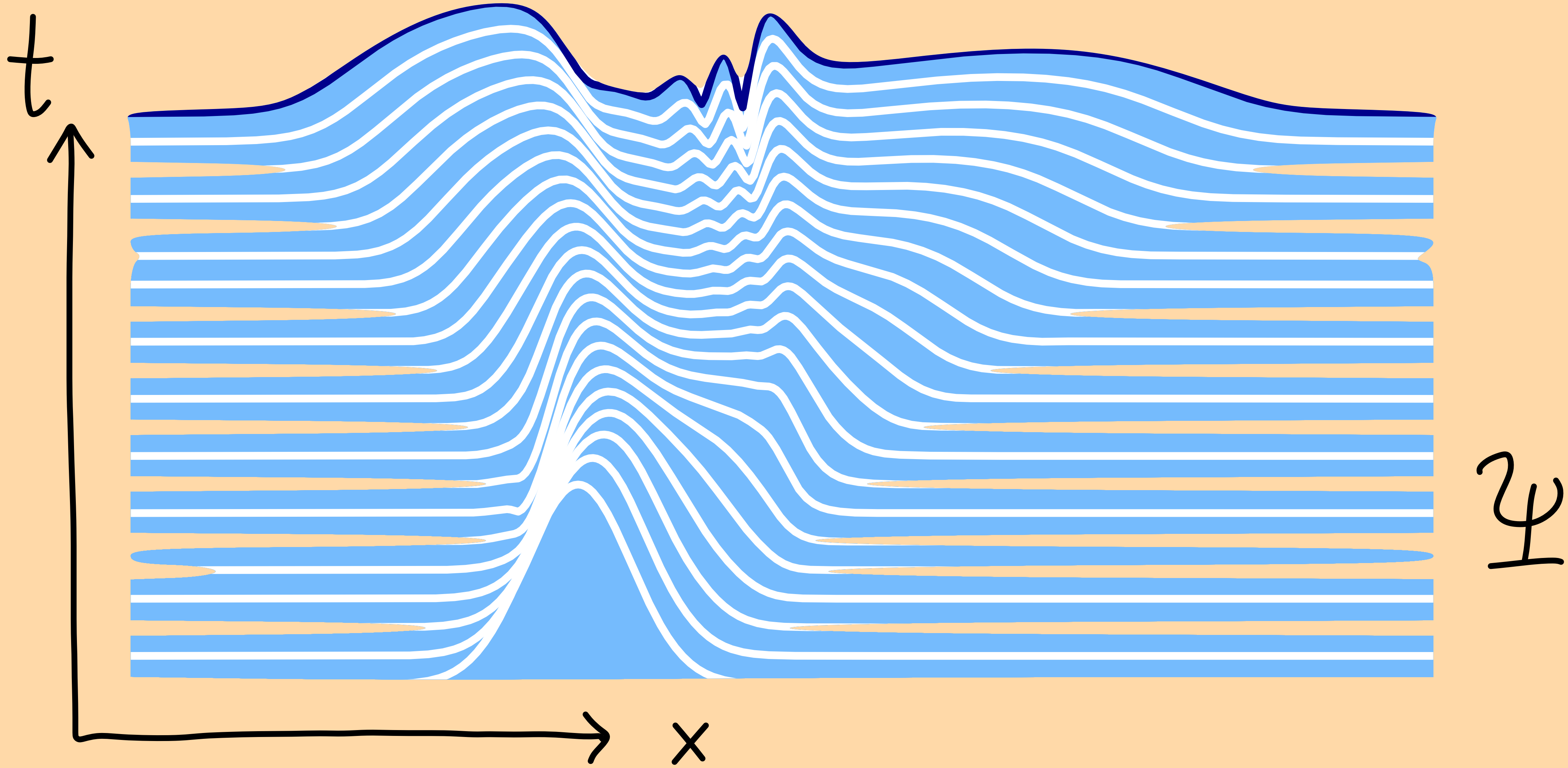
$\Rightarrow \ldots \Rightarrow$

RESULT:

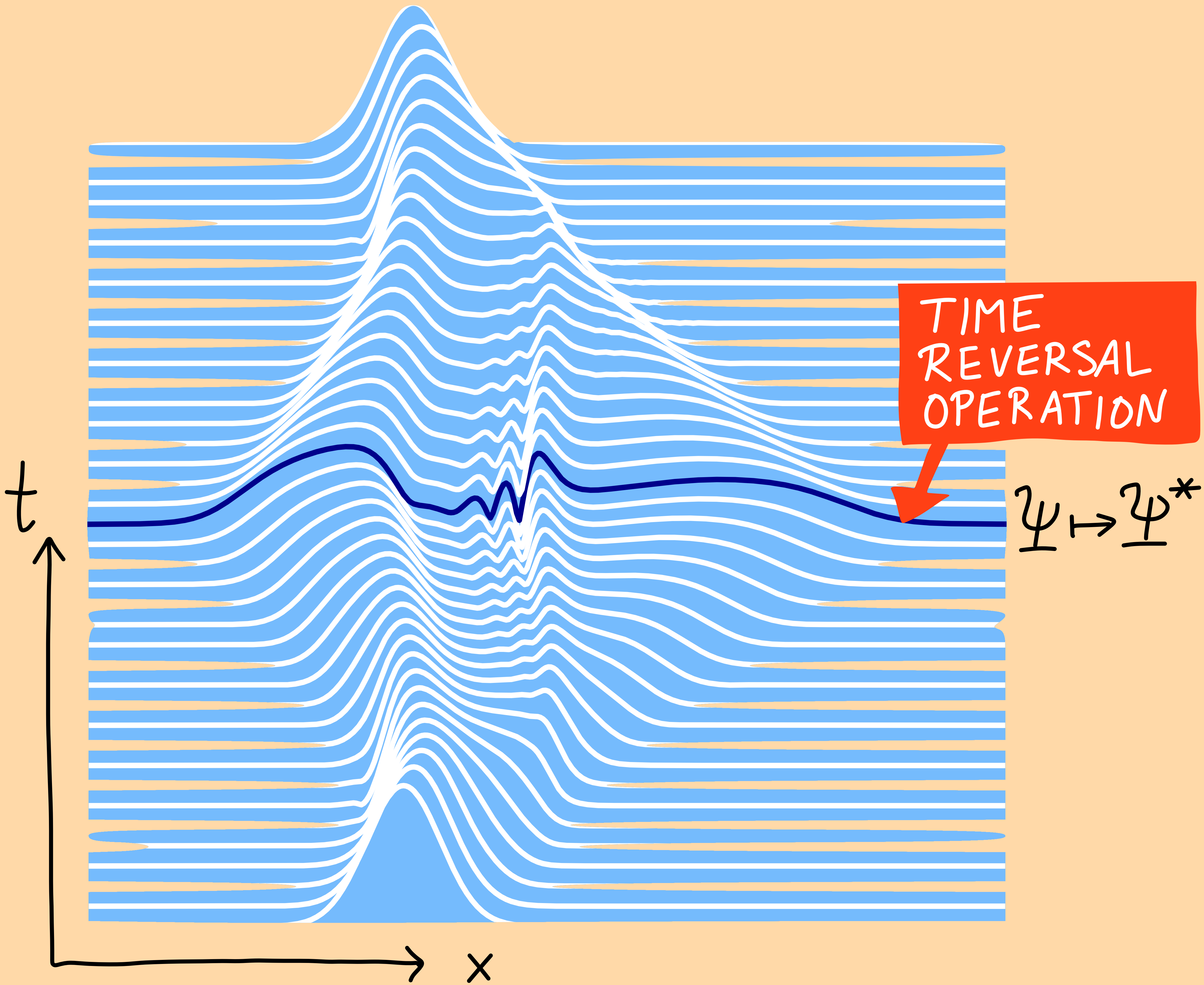
$$\delta\theta \sim -\frac{\partial C}{\partial \theta}$$

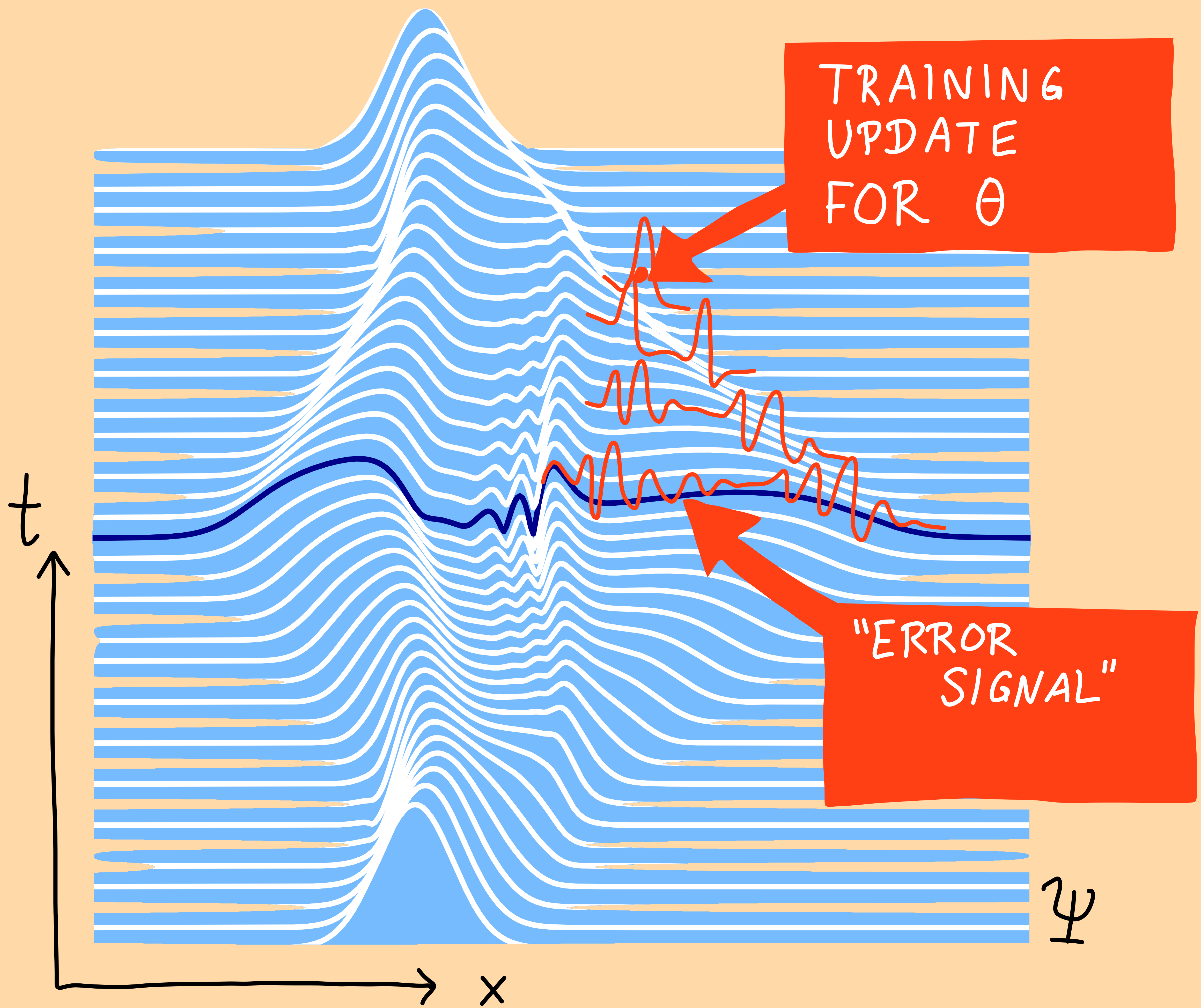
cf adjoint method

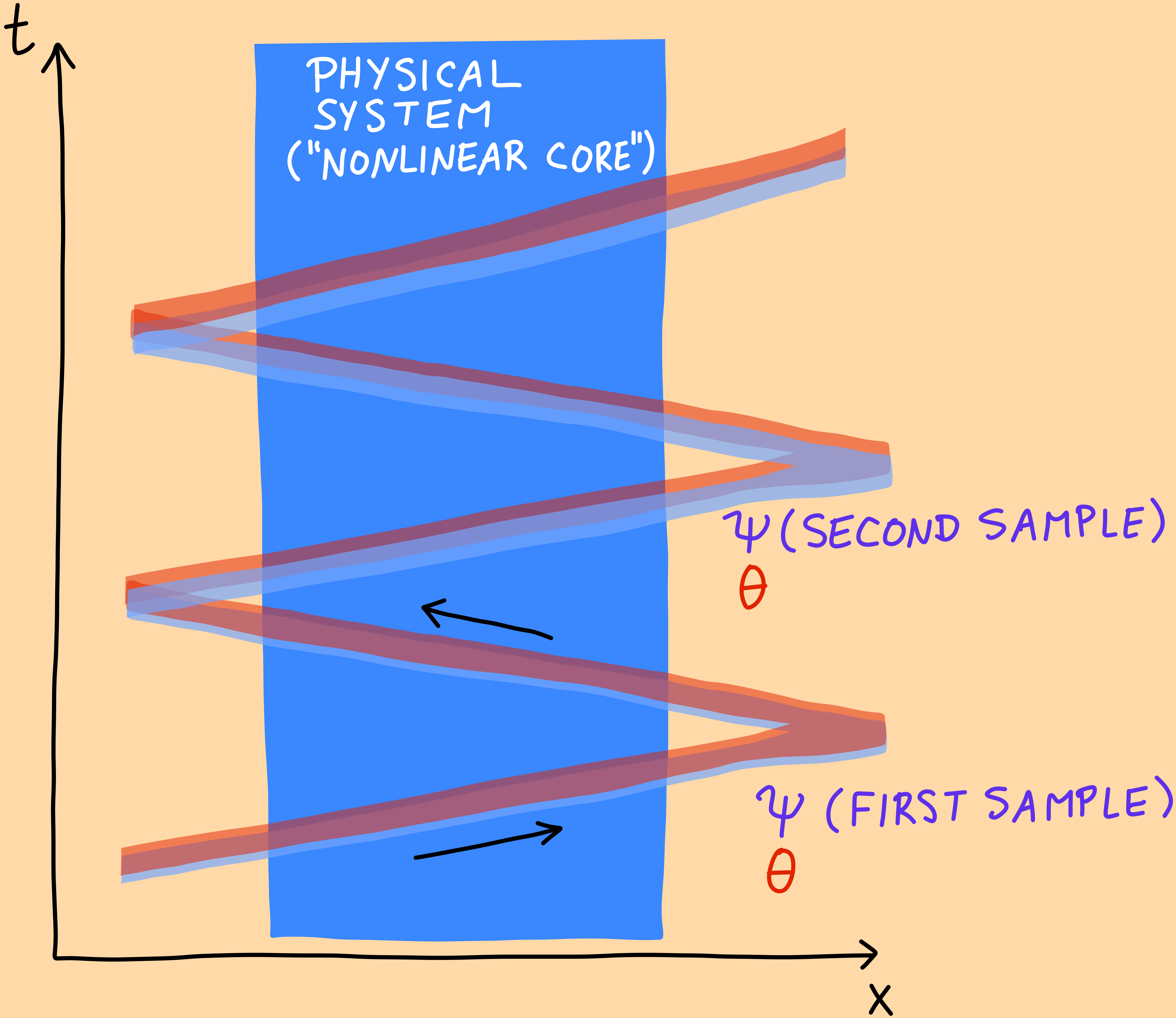
SELF-LEARNING NONLINEAR WAVE FIELDS



t
x
Ψ


TIME REVERSAL OPERATION
t
x
Ψ ↦ Ψ*


TRAINING UPDATE FOR θ
"ERROR SIGNAL"
t
x
Ψ


t
PHYSICAL SYSTEM ("NONLINEAR CORE")
ψ (SECOND SAMPLE)
θ
ψ (FIRST SAMPLE)
θ
x

COMPARE...

PSALTIS ET AL (1987 ff)

NONLIN. OPTICS SELF-LEARNING

BUT: NEED TO ENGINEER FORWARD vs BACKWARD TRANSMISSION

HUGHES, ..., FAN (2018)

GENERAL PHYSICAL BACKPROP

BUT: NO PHYSICAL UPDATE

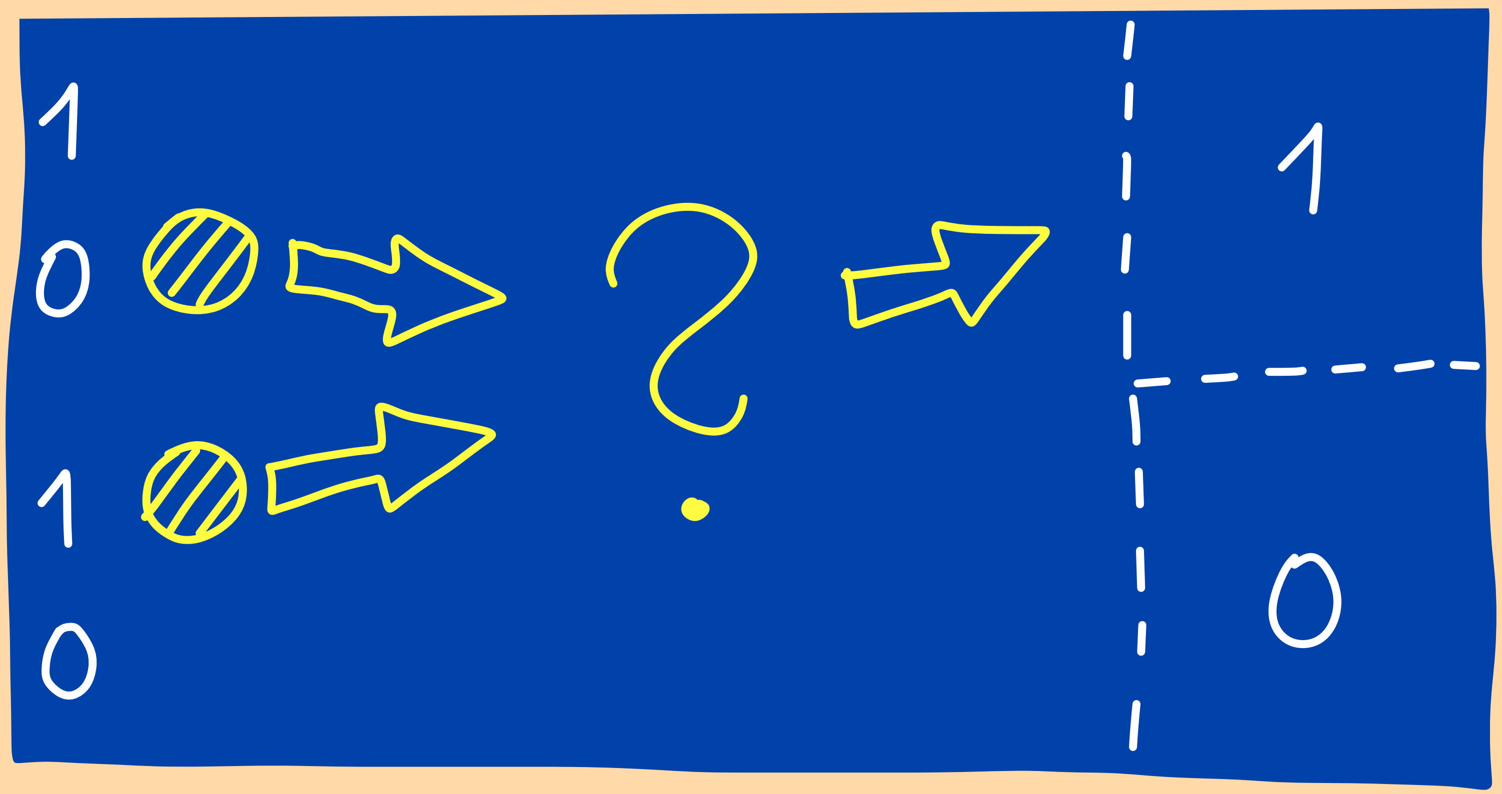
SUMMARY: HAMILTONIAN ECHO BACKPROPAGATION

- ✓ PHYSICAL BACKPROPAGATION
- ✓ PHYSICAL LEARNING UPDATE (VIA INTRINSIC DYNAMICS)
- ✓ HAMILTONIAN-INDEPENDENT (ANY TIME-REVERSAL-INVARIANT SYSTEM)
- ✓ NO ACCESS TO INTERNALS OF "NONLINEAR CORE" = PHYSICAL SYSTEM NEEDED
- ✓ AGNOSTIC OF PHYSICS PLATFORM

Victor Lopez-Pastor & F.M. , Phys. Rev. X 13, 031020 (2023)

FIRST NUMERICAL EXAMPLES

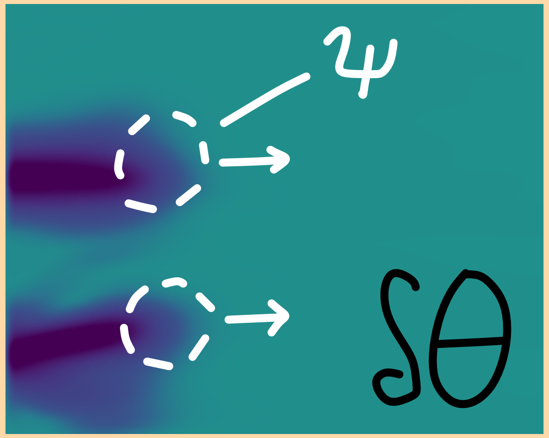
LEARNING XOR



COUPLED NONLINEAR WAVES

$$i\dot{\Psi} = \frac{\beta}{2}\Delta\Psi + (\chi\theta + \underbrace{g|\Psi|^2}_{\text{KERR}})\Psi$$

$$i\dot{\Theta} = i\Omega\pi_\theta + \chi|\Psi|^2$$


ψ
δθ

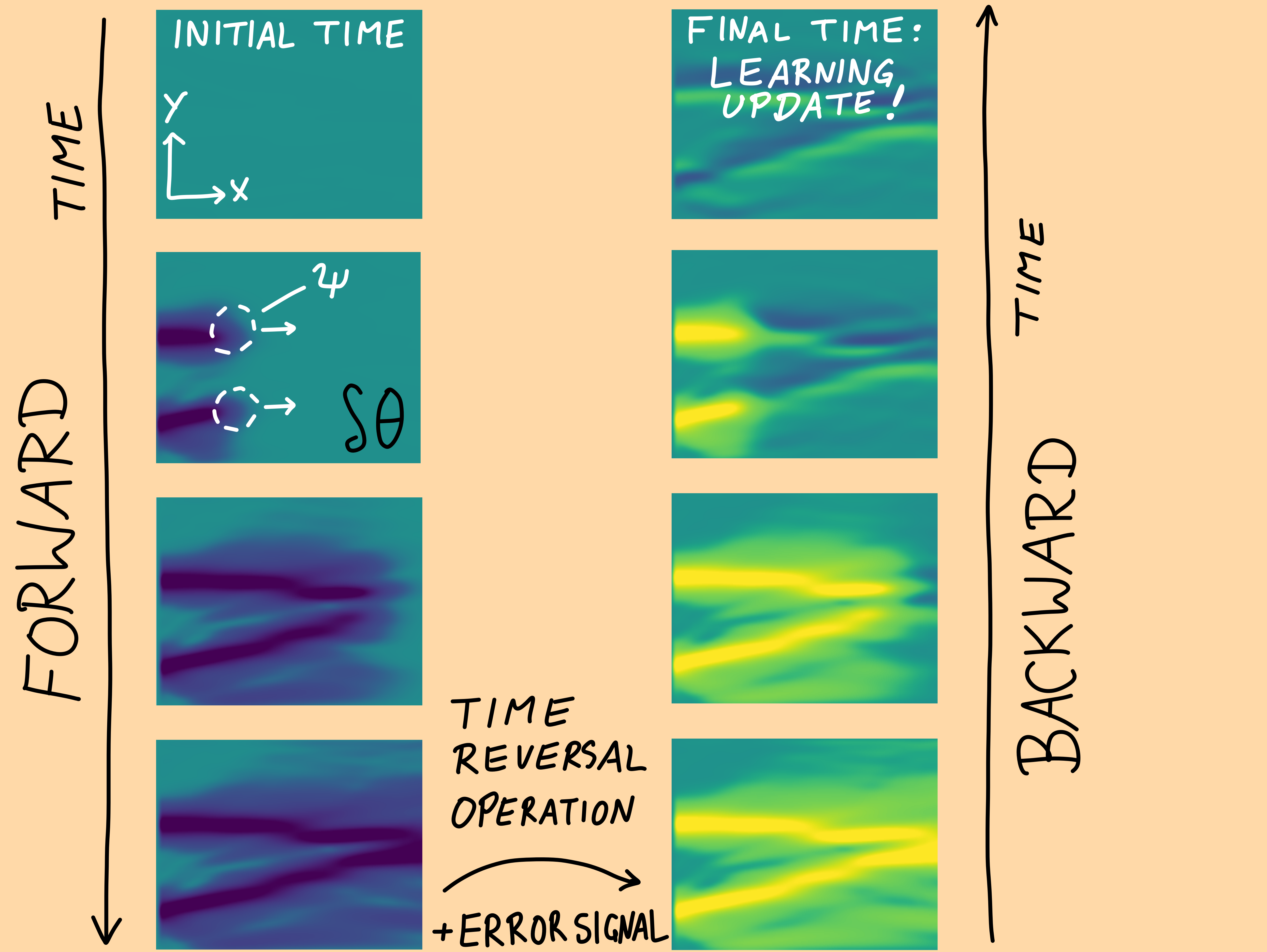
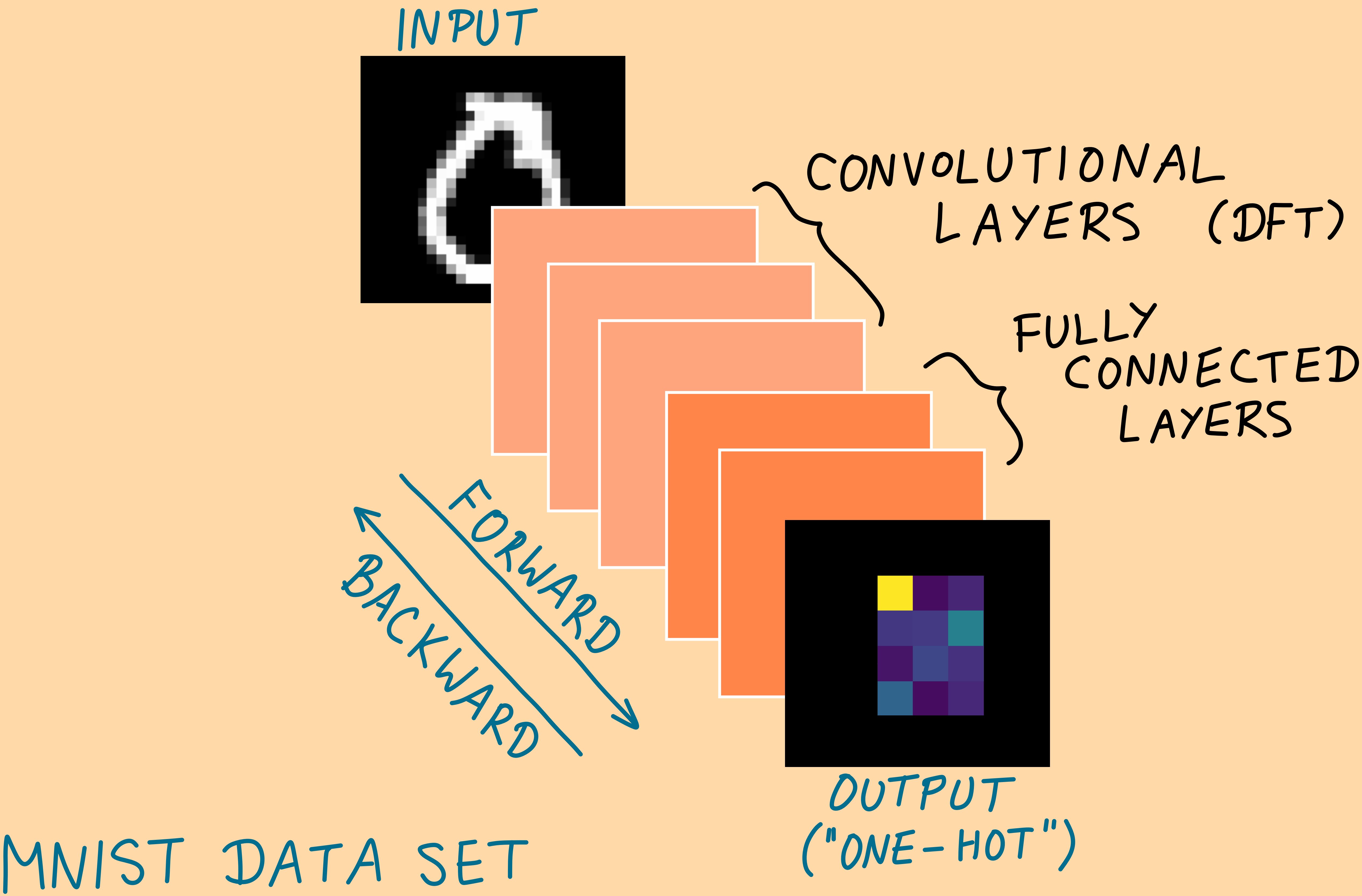
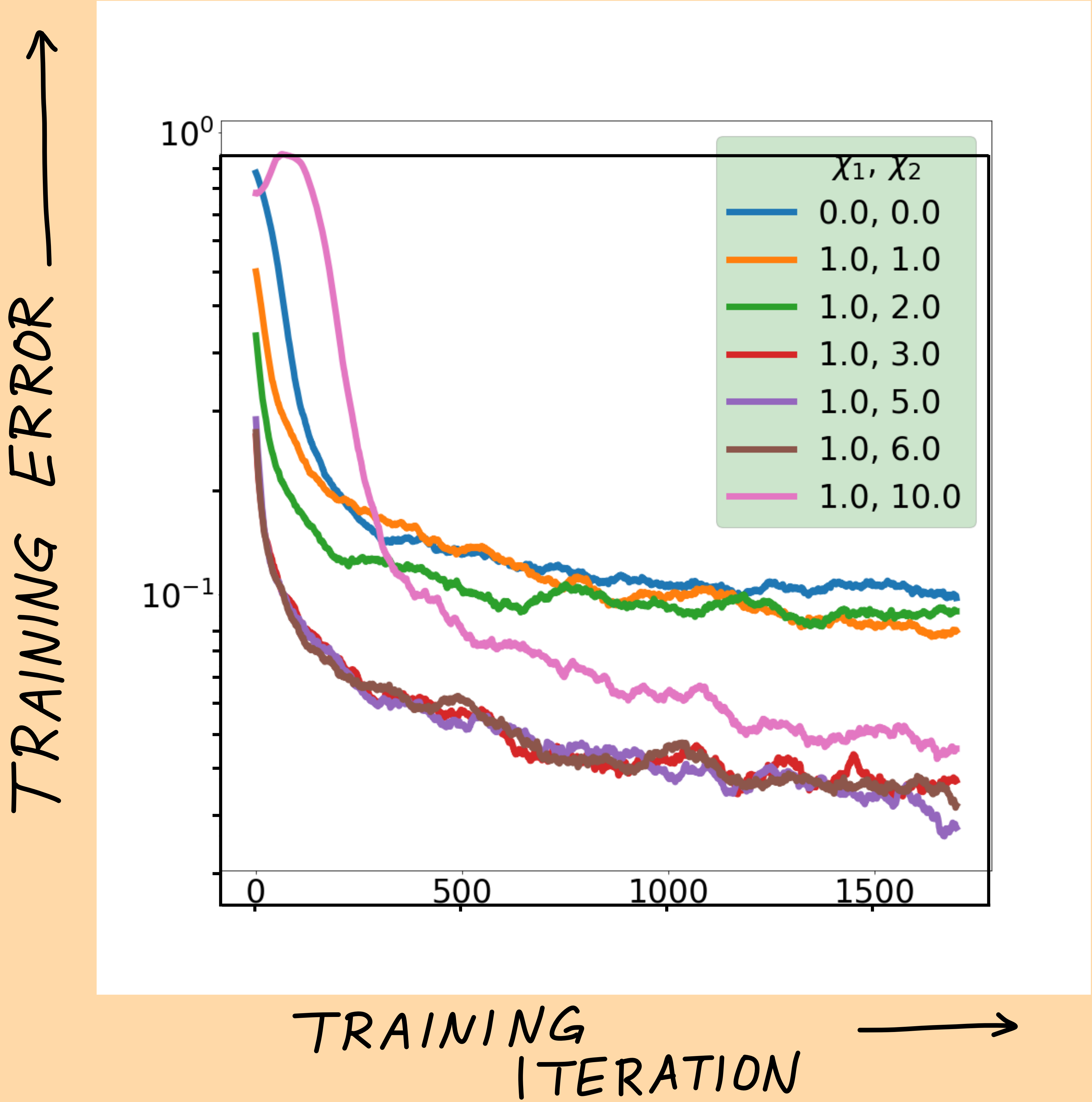

INITIAL TIME
Y
X
ψ
δθ
TIME
FORWARD
FINAL TIME: LEARNING UPDATE!
TIME
BACKWARD
TIME REVERSAL OPERATION
+ERROR SIGNAL

IMAGE CLASSIFICATION





Victor Lopez-Pastor & F.M. , Phys. Rev. X 13, 031020 (2023)

POSSIBLE EXPERIMENTAL PLATFORMS

REQUIREMENTS & CHALLENGES

→ NONLINEAR
→ LOW LOSS
(→ RE-AMPLIFICATION)

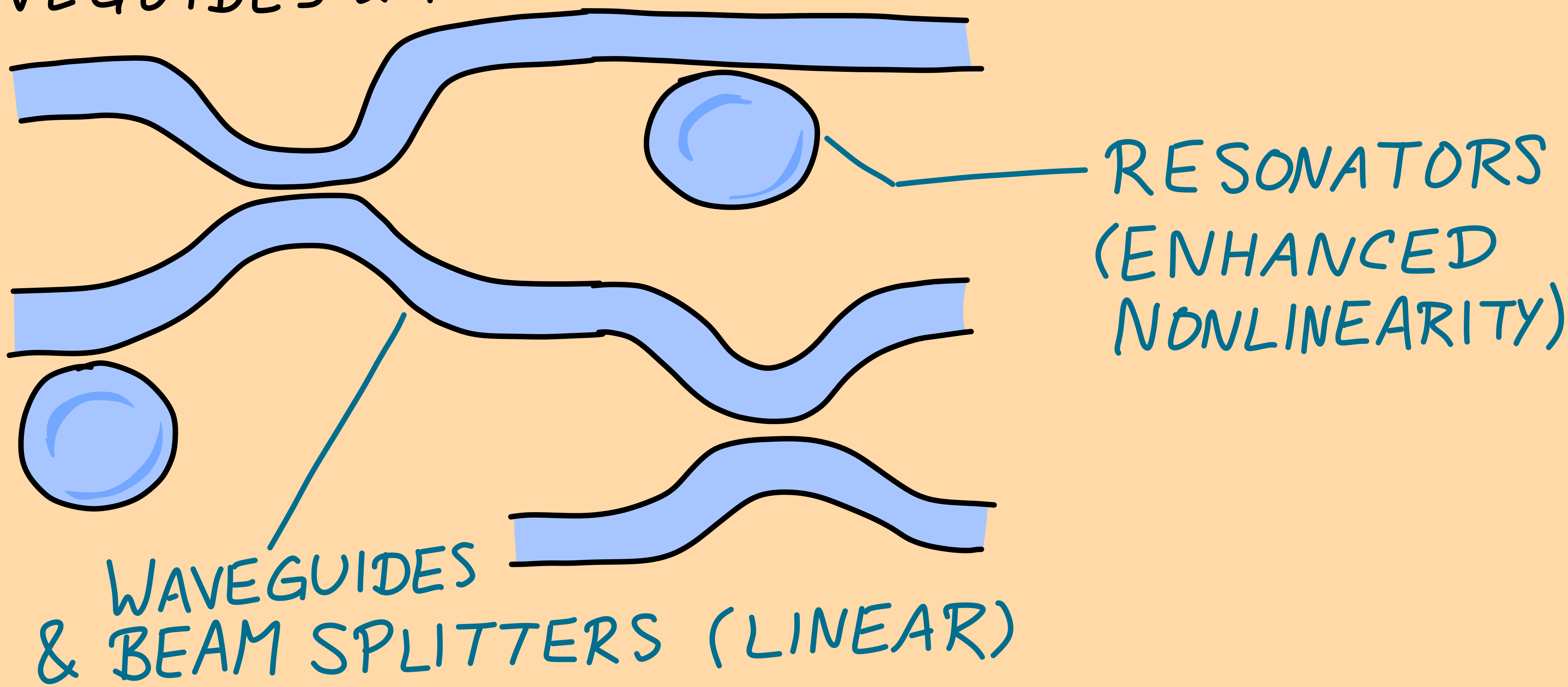
→ TIME-REVERSAL OPERATION [& "DECAY STEP"]
→ NOISE, NONIDEALITY IN THESE OPERATIONS
(→ CALIBRATION, ROBUSTNESS)

→ LONG-TERM STORAGE OF LEARNING FIELD
(→ READOUT)

NONLINEAR OPTICS

EXAMPLE: INTEGRATED PHOTONICS

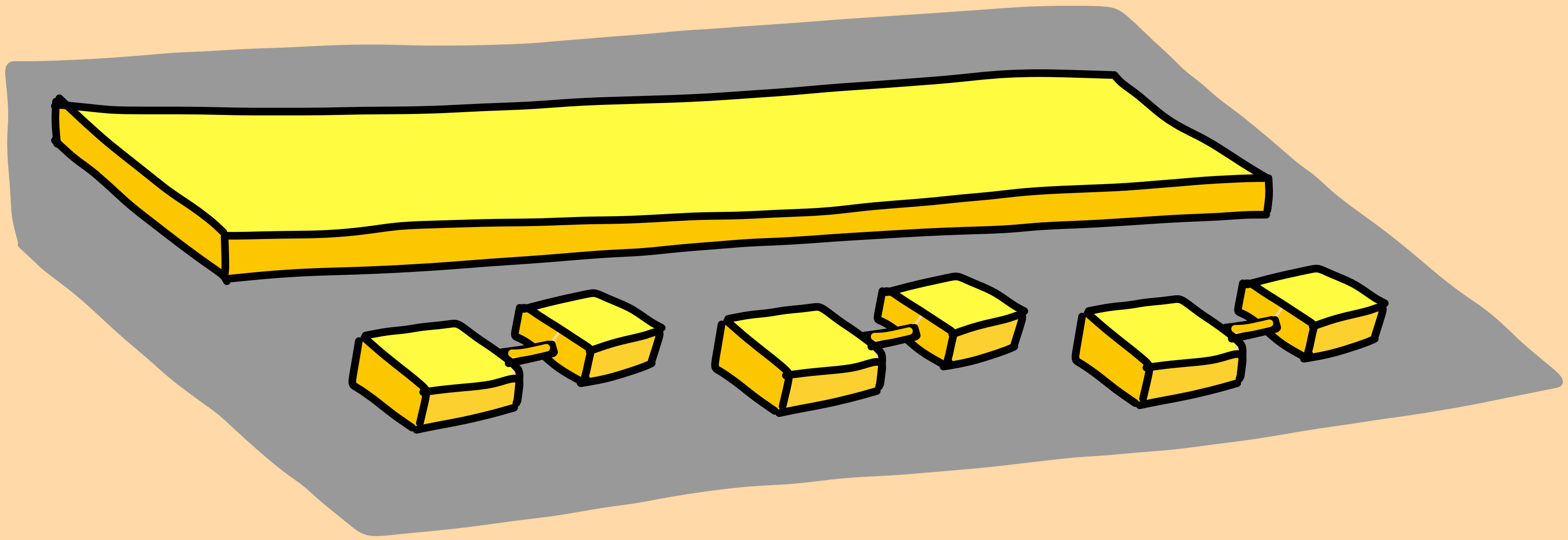
WAVEGUIDES & NONLINEAR RESONATORS



BUILD ON OPTICAL NEURAL NETWORKS:

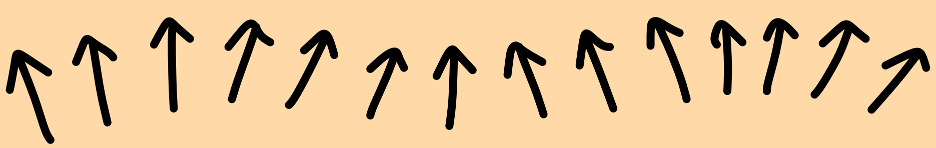
WAGNER & PSALTIS (1987ff.), SKINNER ET AL (1995), SHEN,..., ENGLUND, SOLJACIC (2017), HUGHES,...,FAN (2018), GUO,..., LVOVSKY (2019), FELDMAN,..., PERNICE (2019), ...

REVIEW: WETZSTEIN ET AL (2020)



SUPERCONDUCTING
MICROWAVE CIRCUITS

NONLINEAR
MATTER WAVES



SOME SPIN WAVES
(TIME-REVERSAL SYMMETRY!)

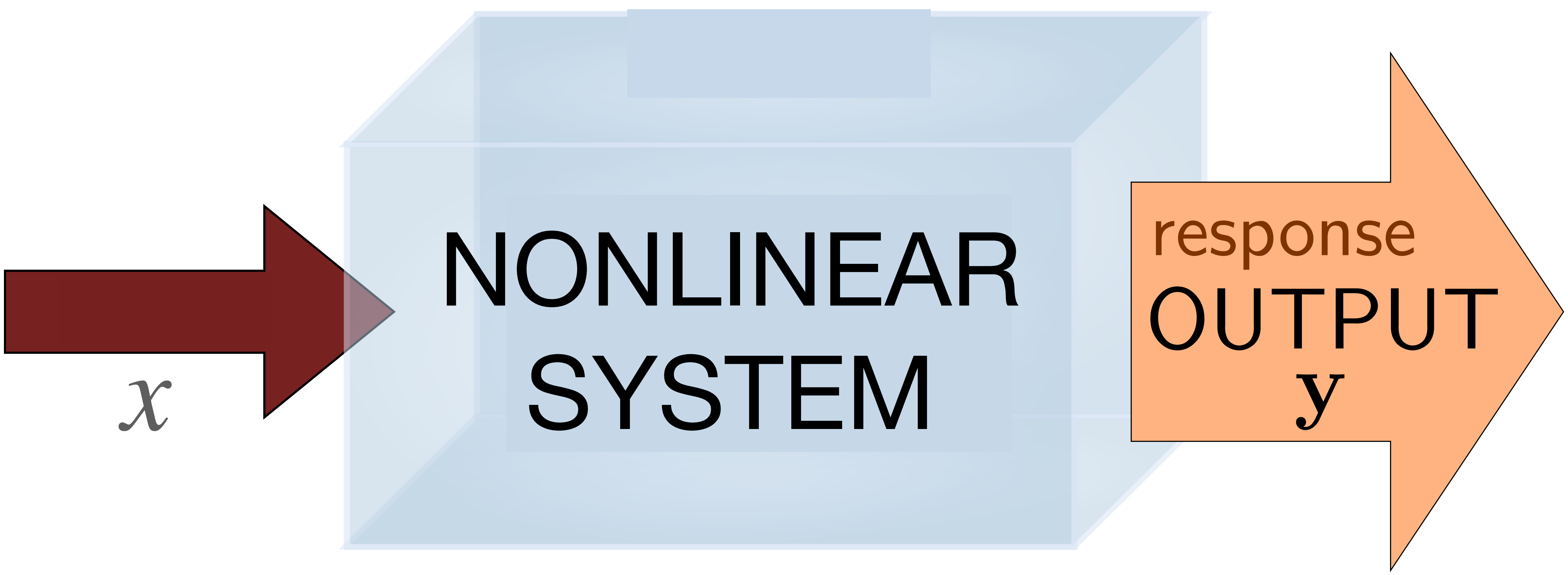
Victor Lopez-Pastor & F.M. , Phys. Rev. X 13, 031020 (2023)

Nonlinear neuromorphic system from linear wave scattering

Clara Wanjura & F.M. arXiv 2308.16181

typical optical neuromorphic system

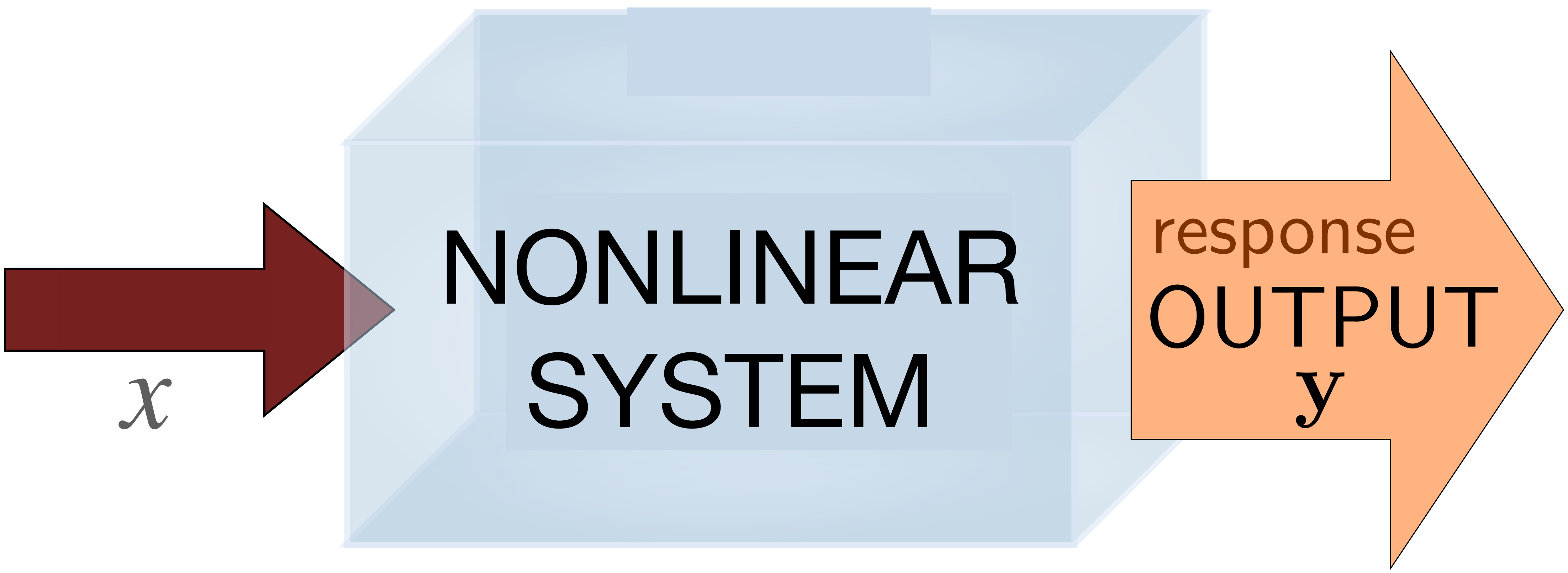
nonlinearity for **expressivity**



$$y = f(x)$$

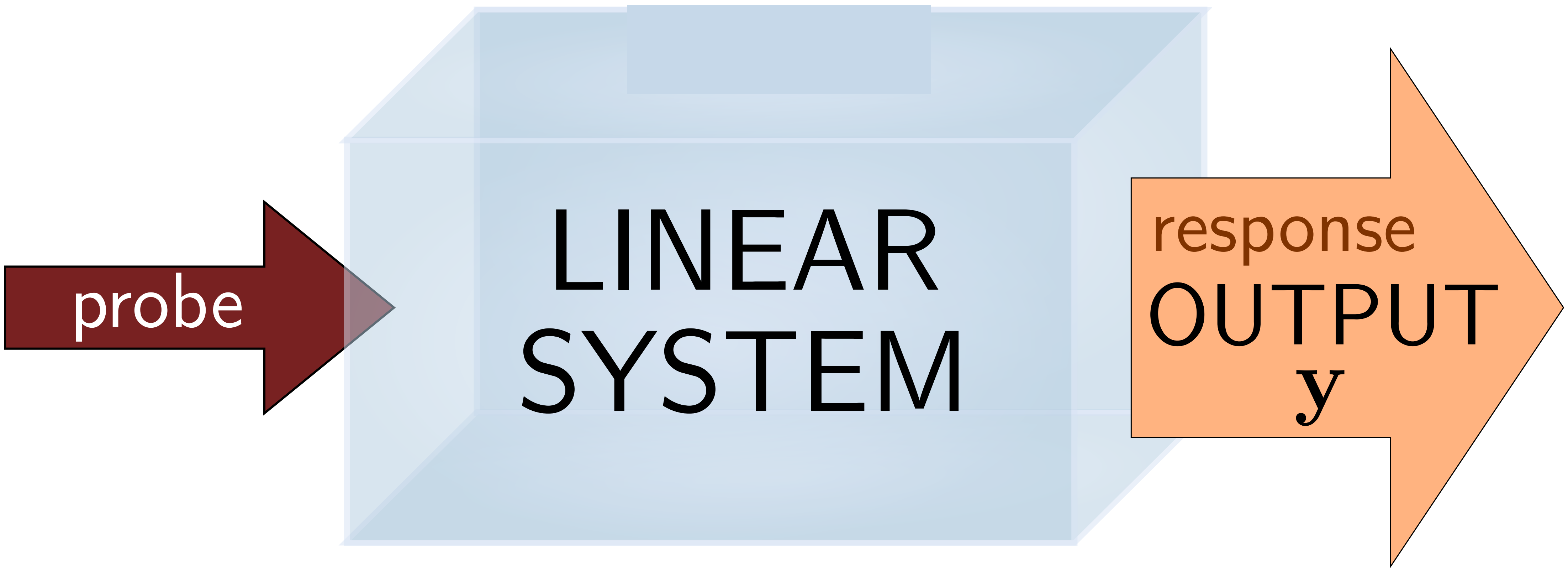
typical optical neuromorphic system

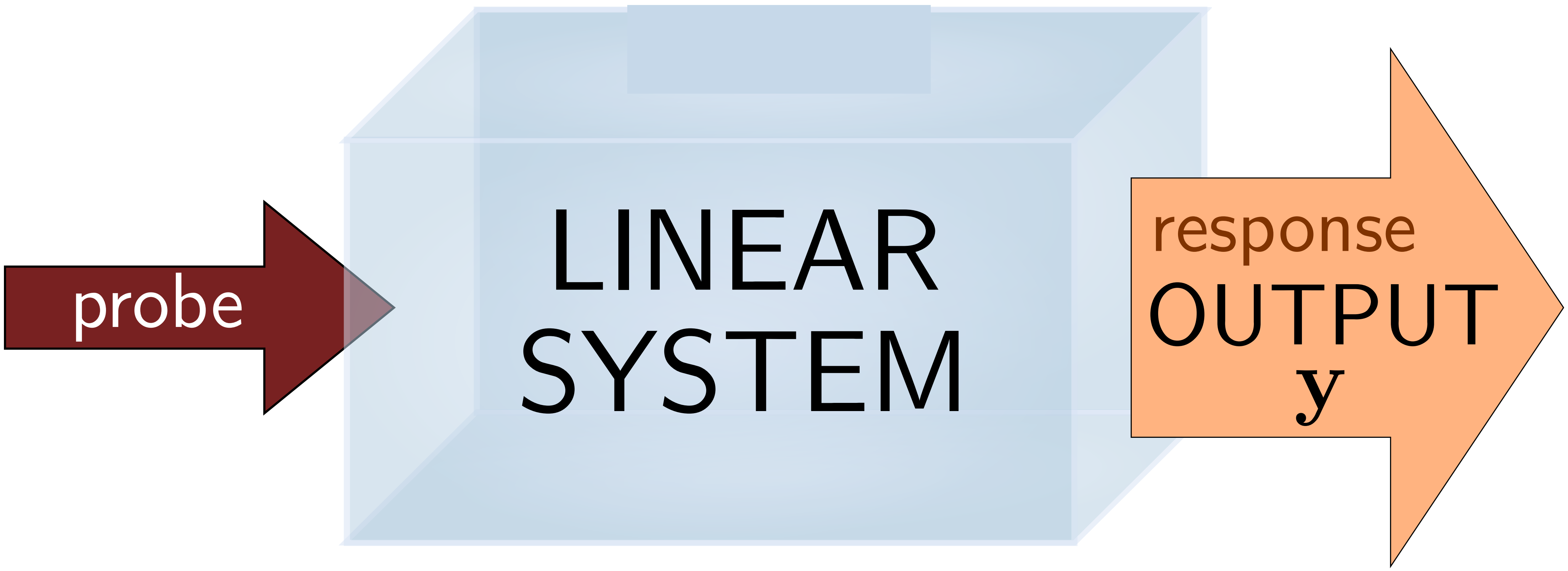
nonlinearity for **expressivity**



$$y = f(x)$$

optical nonlinearities (but: power levels)
optoelectronics (but: delays, power)

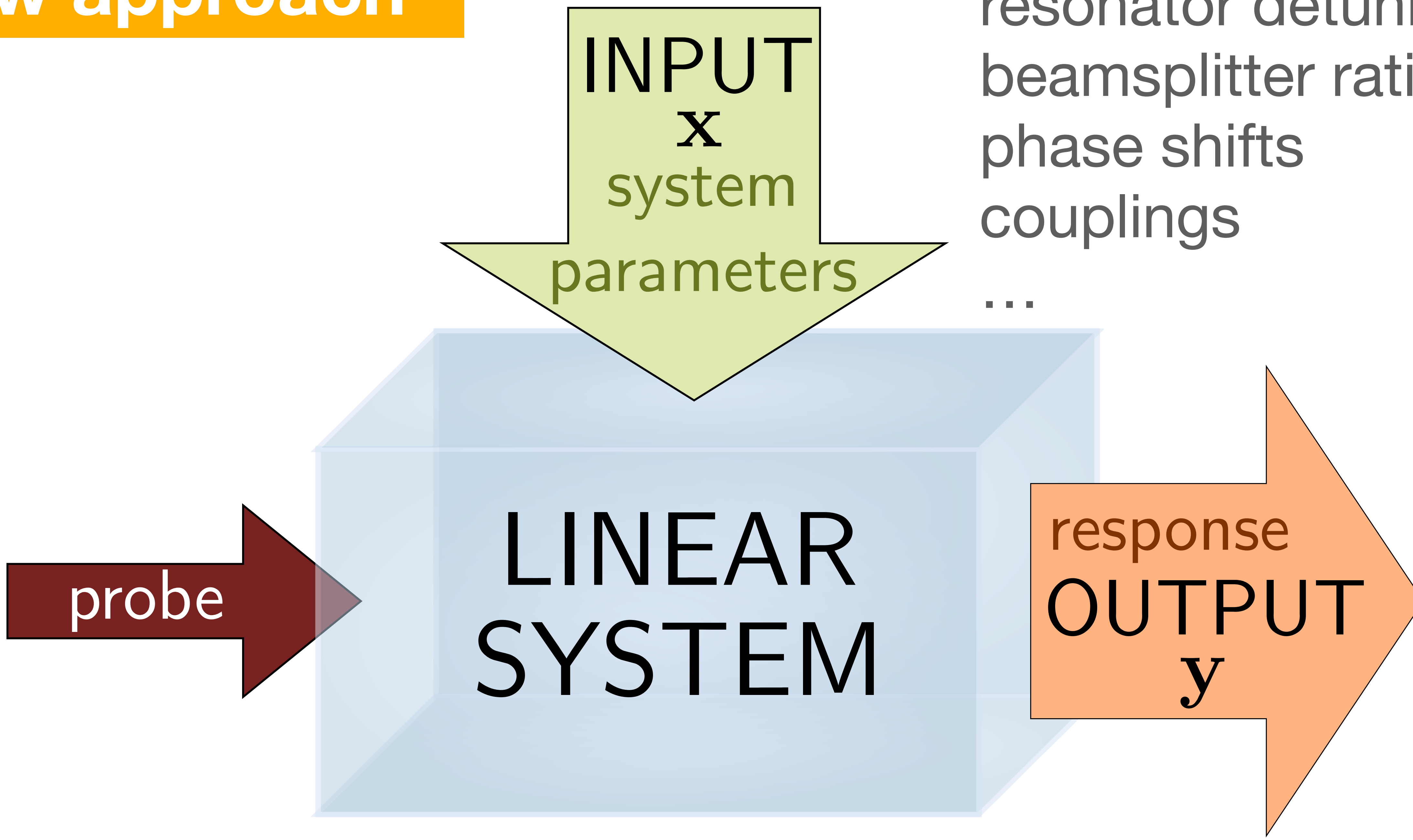

probe
LINEAR
SYSTEM
response
OUTPUT
$\mathbf{y}$



$$y = S(\omega)a_{\mathrm{probe}}$$

Scattering Matrix

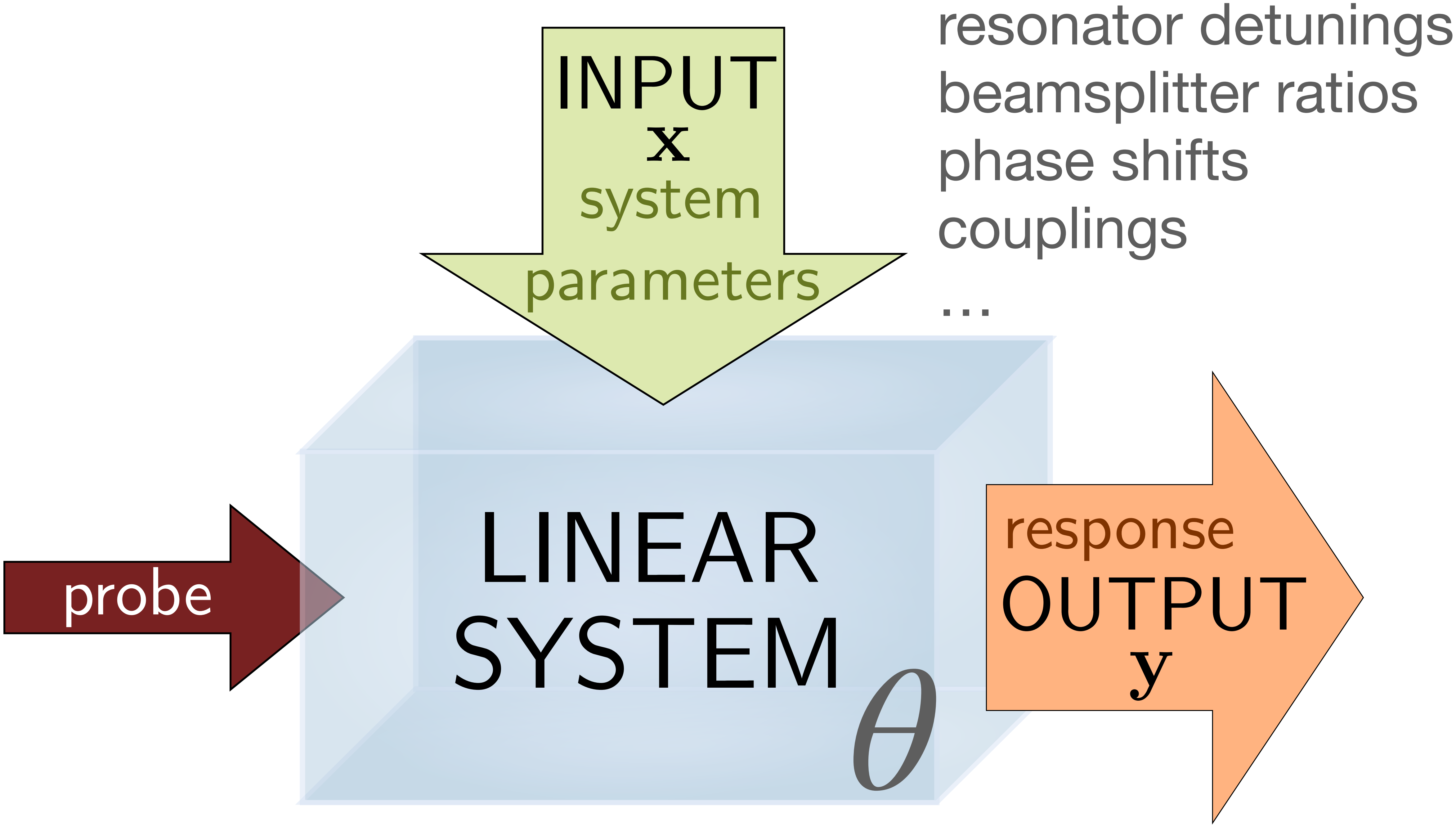
new approach



resonator detunings
beamsplitter ratios
phase shifts
couplings
...

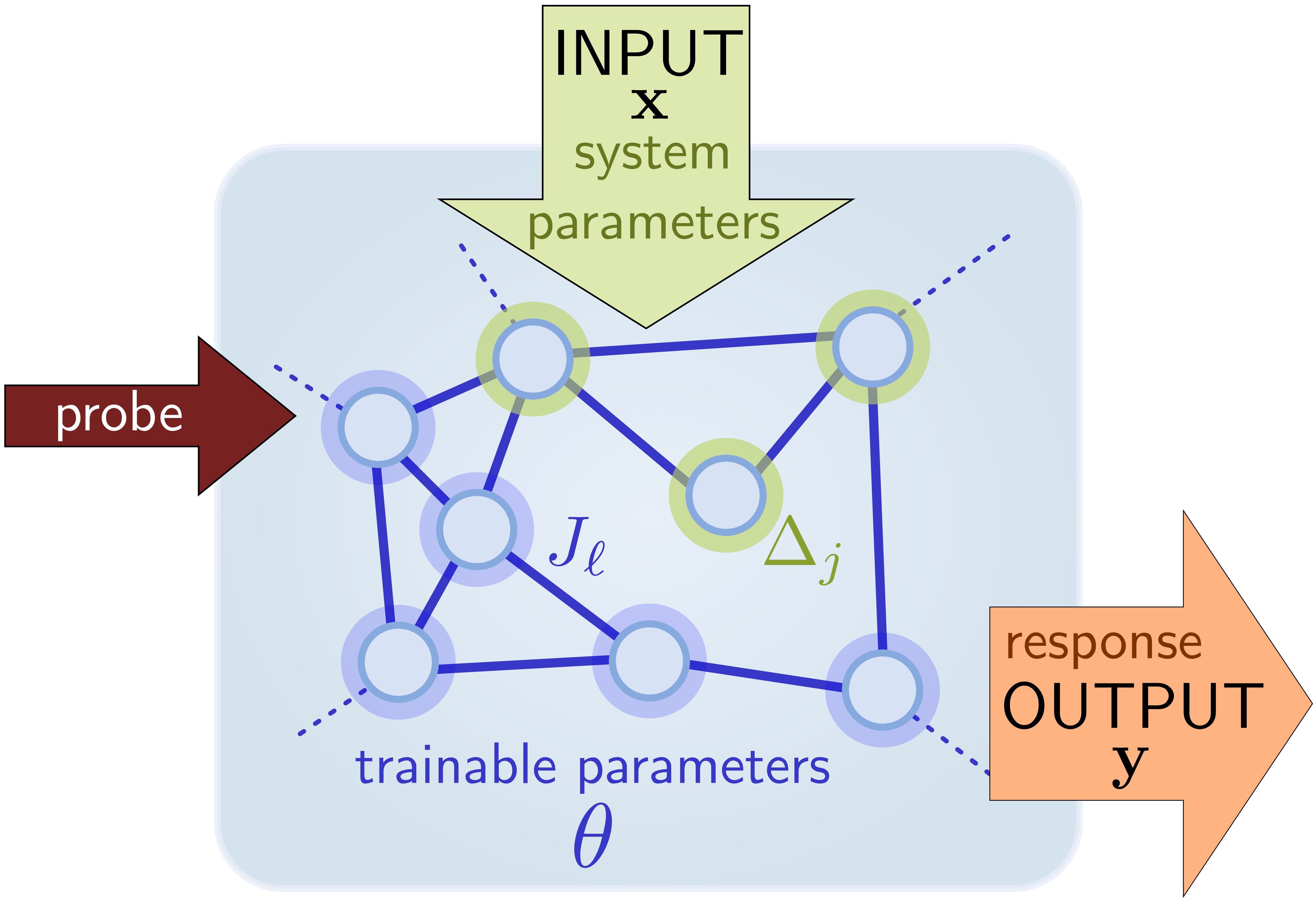
$$y = S(x, \omega) a_{\text{probe}}$$

nonlinear function of **input x**



$$y = S(x, \theta, \omega) a_{\text{probe}}$$

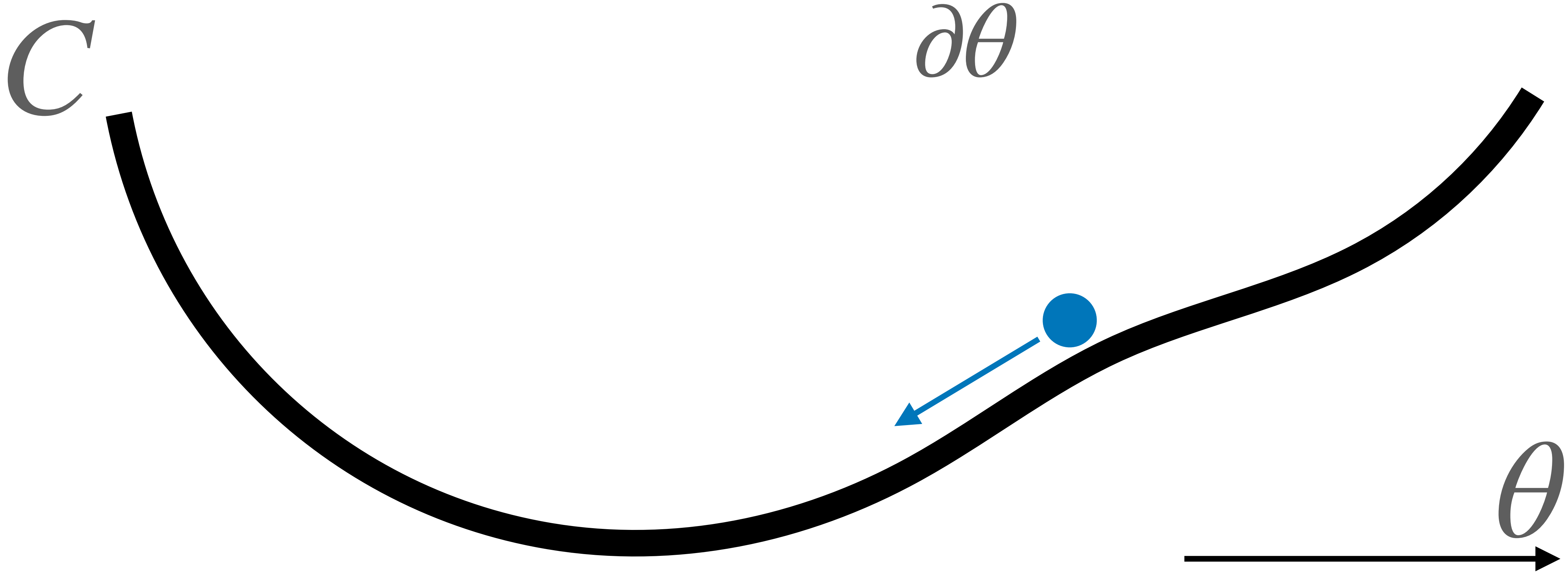
nonlinear trainable function of **input x**


INPUT
$\mathbf{x}$
system
parameters
probe
J_ℓ
Δ_j
response
OUTPUT
$\mathbf{y}$
trainable parameters
θ

Training

Gradient descent on cost function

$$\delta\theta = -\frac{\partial C}{\partial \theta}$$



Training

Gradient descent on cost function

$$\delta\theta = -\frac{\partial C}{\partial \theta}$$

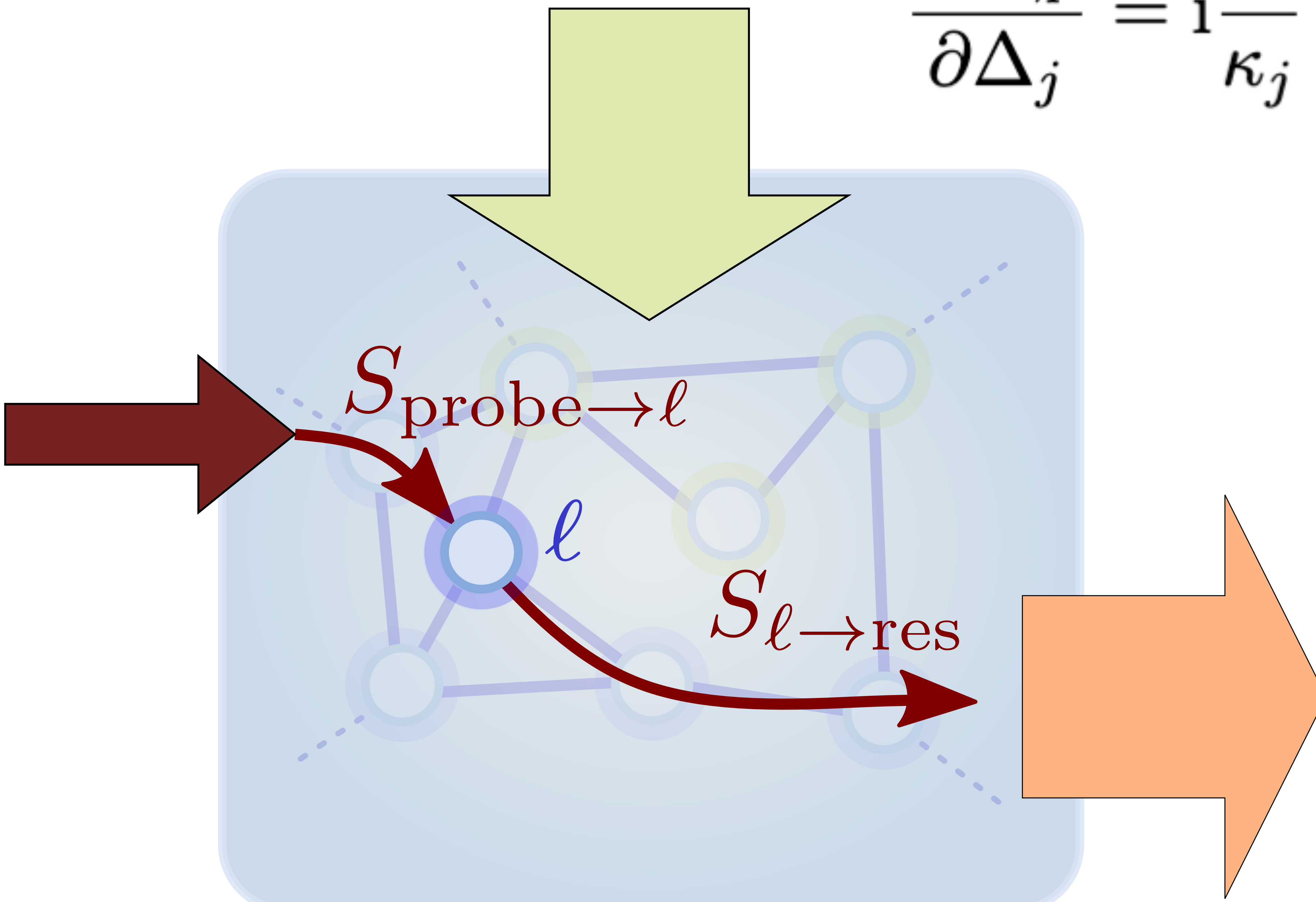
Challenge: obtain gradients efficiently for a physical system!

Backpropagation on a model (but: model?)
Hamiltonian Echo Backpropagation (time-reversal)
Equilibrium propagation (relaxation system)

Training

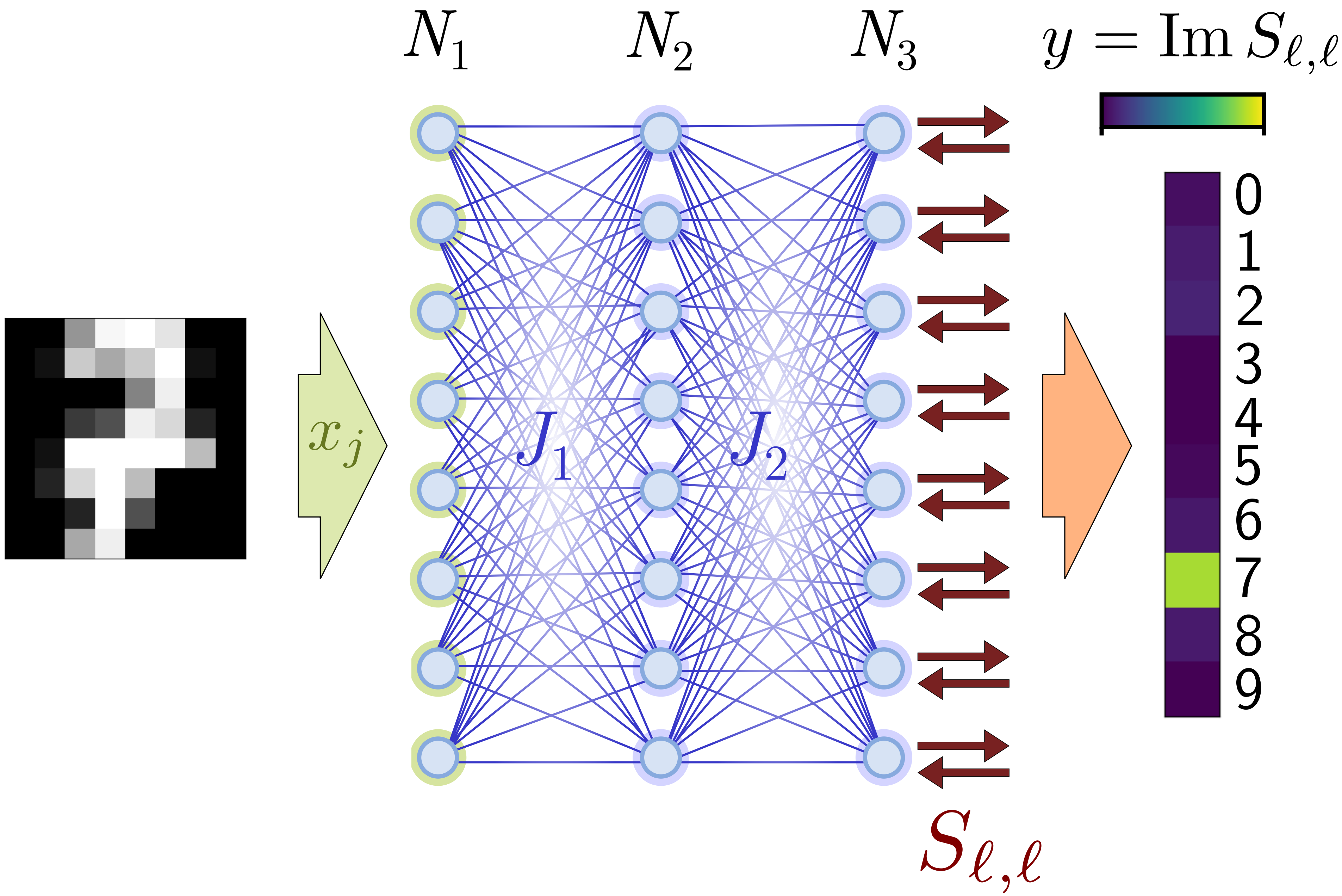
Here: **Gradients from simple scattering matrix measurements!**

$$\frac{\partial S_{r,p}}{\partial \Delta_j} = \mathrm{i}\frac{1}{\kappa_j} G_{j,p} G_{r,j}$$



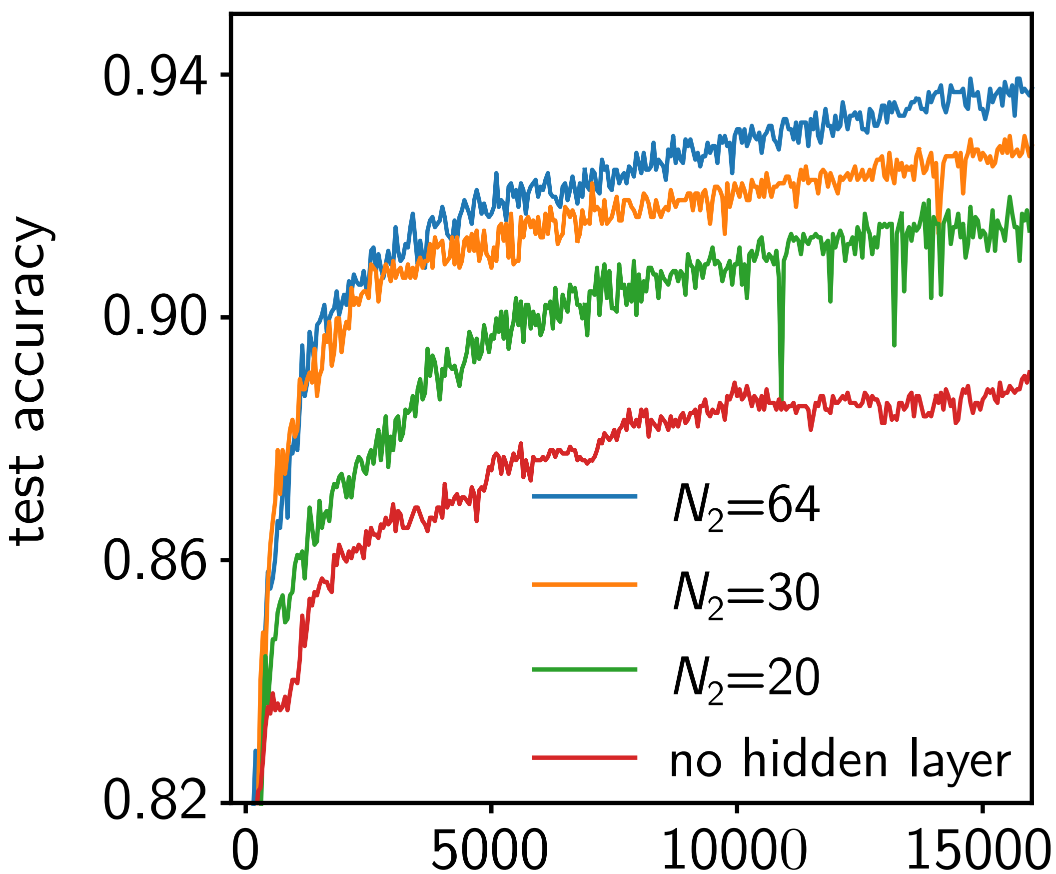
C. Wanjura, F. Marquardt arXiv: 2308.16181

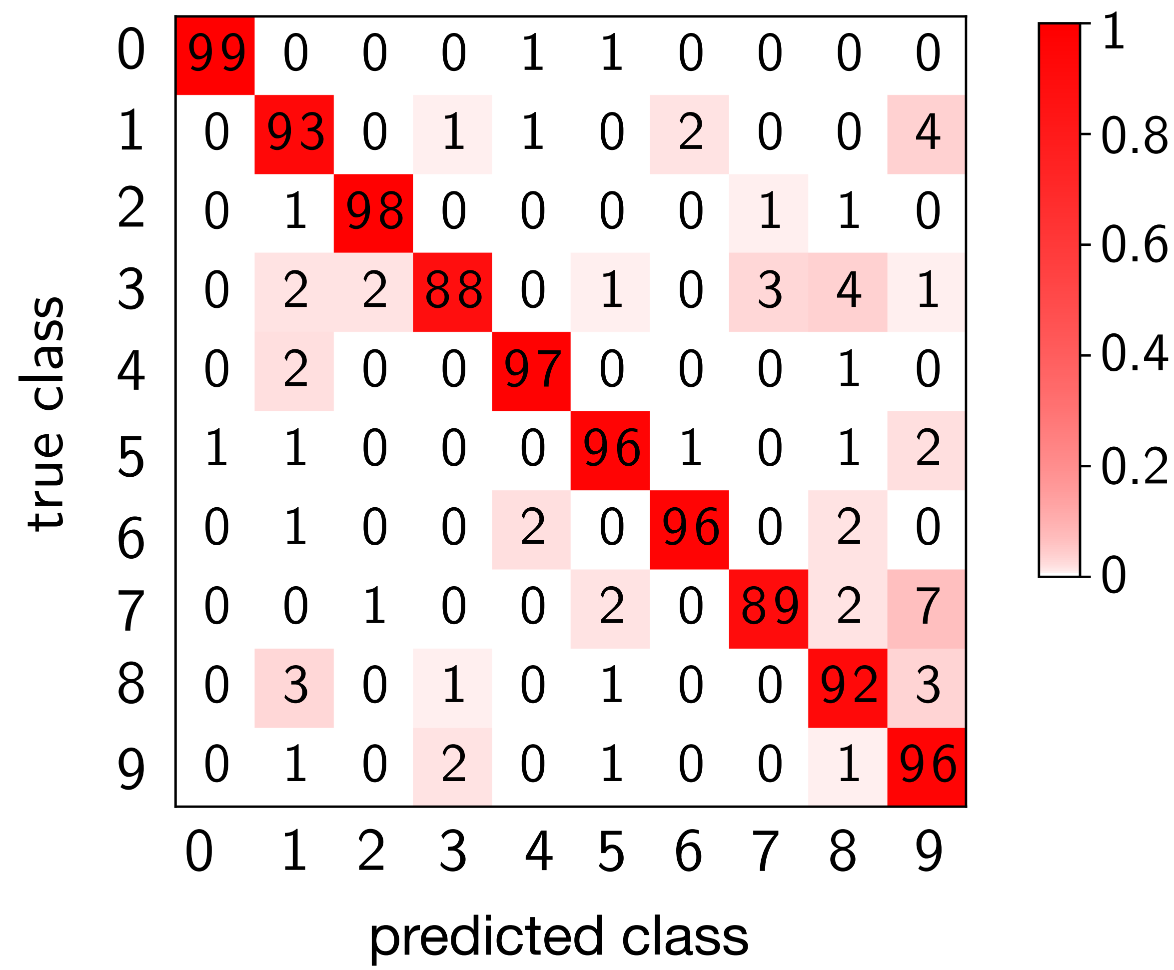
Simple tight-binding model



Simple tight-binding model

Training on handwritten digits

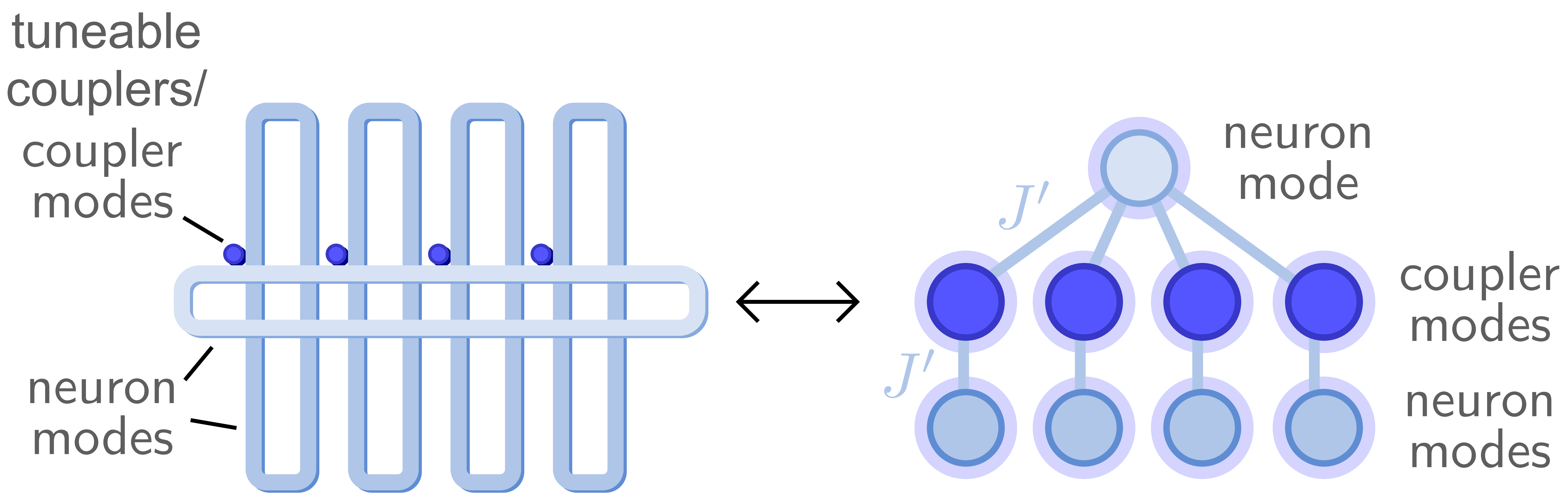




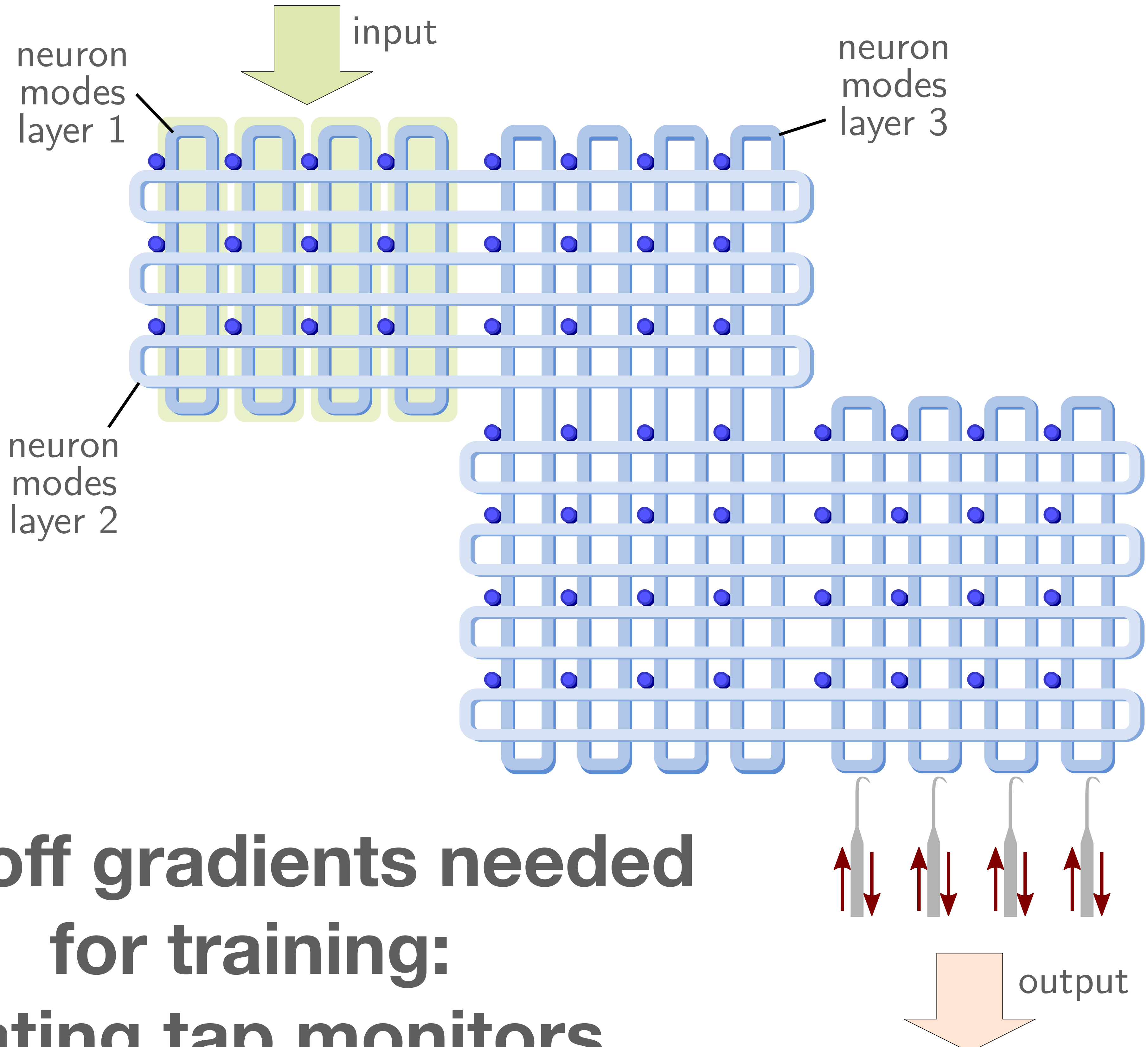
Better accuracy than purely linear neural network

Possible optical implementation

racetrack resonators



Possible optical implementation



read off gradients needed for training: grating tap monitors

Fully nonlinear neuromorphic learning machine based on linear wave scattering

...should work in many platforms
simple training, simple inference

C. Wanjura, F. Marquardt arXiv: 2308.16181

similar ideas & free-space experiments:

M. Yildirim, N. U. Dinc, I. Oguz, D. Psaltis, and C. Moser, arXiv:2307.08533

F. Xia, K. Kim, Y. Eliezer, L. Shaughnessy, S. Gigan, and H. Cao, arXiv:2307.08558

Physical self-learning machines as new tools for machine learning

Hamiltonian Echo Backpropagation

General physical training procedure

Victor Lopez-Pastor & F.M.

Phys. Rev. X 13, 031020

Nonlinear neuromorphic system via linear waves

Suitable for any linear platform

Clara Wanjura & F.M. arXiv 2308.16181