Superconducting Optoelectronic Networks

A Build-It-Up Tour: From a Single Josephson Junction to a Learning, Remembering, Cryogenic Brain

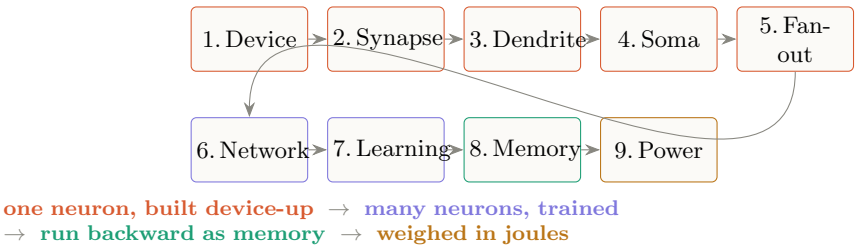
Lecture notes

June 15, 2026

Abstract

These notes assemble, layer by layer, a complete picture of *superconducting optoelectronic networks* (SOENs): brain-inspired hardware that communicates with single photons and computes with superconducting circuits. We start from one device — a Josephson-junction SQUID source — and successively add a synapse, a dendrite, a firing soma, and an optical fan-out, building a single spiking neuron. We then zoom out: many neurons reduce to a rate model that can be *trained* (equilibrium propagation) and can *remember* (Hopfield attractor dynamics). Finally we account for the *energy*: why the whole thing lives at 4 K and still wins on operations per watt. The governing dynamics throughout are the single first-order "leaky-integrator" equation $\dot{s} = \gamma\, g(\phi, s) - s/\tau$, derived for SOENs from a circuit Lagrangian by Shainline and co-workers [1, 2]. Each section corresponds to one stage of the companion interactive model.

Orientation: the nine layers



The one equation to keep in mind. Every active element — synapse, dendrite, soma — is a single first-order "loop" obeying

$$\boxed{\frac{\mathrm{d}s}{\mathrm{d}t} \;=\; \gamma\, g(\phi, s) \;-\; \frac{s}{\tau}} \tag{1}$$

where $s = I/I_c$ is the (dimensionless) circulating loop current, ϕ is the applied magnetic flux from upstream, g is the device *source function* (the SQUID nonlinearity), γ sets the drive strength, and τ the leak time. Read it as a leaky bucket: γg fills, s/τ drains. Everything below is this equation, copied and wired.

1 Stage 1 — The device: Josephson junction and SQUID source

The atom of a SOEN is the *Josephson junction* (JJ): two superconductors separated by a thin barrier. Its state is the phase difference φ of the macroscopic wavefunction across the barrier, obeying

$$I = I_c \sin\varphi, \qquad \frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{2\pi}{\Phi_0} V, \tag{2}$$

with $\Phi_0 = h/2e \approx 2.07 \times 10^{-15}\,\mathrm{Wb}$ the flux quantum. A current-biased, damped junction behaves exactly like a particle in a *tilted washboard potential*

$$U(\varphi) = -E_J \cos\varphi - \frac{\Phi_0 I}{2\pi}\,\varphi, \tag{3}$$

which is also the equation of a driven pendulum. Below a critical tilt the "phase particle" sits trapped in a well (zero voltage, superconducting). Past threshold it rolls — each 2π slip emits one flux quantum, a discrete, quantized event (Fig. 1a).

Two JJs in a loop form a *SQUID* (superconducting quantum interference device). Its output current is a steep, saturating function of the applied flux. In the phenomenological model this becomes the **source function** $g(\phi, s)$ — the single nonlinearity reused at every layer. A convenient idealization is a logistic turn-on whose threshold is set by a *bias current* i_b:

$$g(\phi, s) \;\approx\; \sigma\big(k(\phi - \phi_{\mathrm{th}})\big)\,(1 - s), \qquad \phi_{\mathrm{th}} = \tfrac{1}{2} - \alpha\, i_b, \tag{4}$$

where $\sigma(x) = 1/(1 + e^{-x})$ and the factor $(1 - s)$ enforces loop saturation (Fig. 1b). Raising i_b slides the curve left, making the element more excitable. *This is the only nonlinearity in the entire system.*

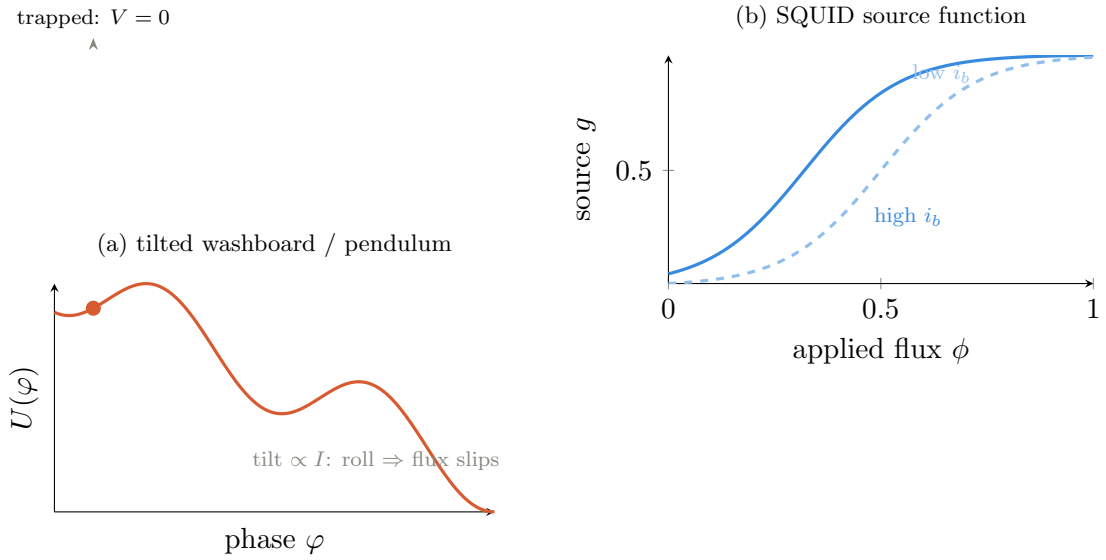


Figure 1: (a) A current-biased Josephson junction is a phase particle in a tilted washboard; equivalently a driven pendulum. (b) The two-junction SQUID gives a steep, saturating flux-to-drive transfer function $g(\phi)$; the bias current i_b sets where it turns on.

2 Stage 2 — The synapse: a photon becomes flux

Communication between SOEN neurons is *optical and discrete*: a presynaptic neuron emits photons. At the synapse a *single-photon detector* (SPD; in hardware a superconducting nanowire, SNSPD) absorbs a photon and, in doing so, injects a quantized packet of flux into a small superconducting *synaptic loop*. The synaptic weight w is set physically by how much flux each detection deposits — i.e. by a stored bias — and is therefore an analog, locally adjustable quantity.

The synaptic loop is itself an instance of Eq. (1): each detection is an impulse on ϕ, the loop integrates it into a signal that then leaks,

$$\frac{\mathrm{d}s_{\mathrm{syn}}}{\mathrm{d}t} = w \sum_k \delta(t - t_k) \;-\; \frac{s_{\mathrm{syn}}}{\tau_{\mathrm{syn}}}. \tag{5}$$

So a synapse is a leaky memory of recent presynaptic spikes, with a tunable gain w (Fig. 2).

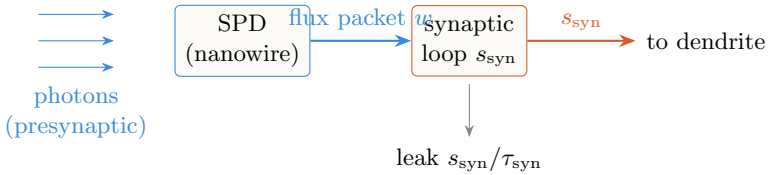


Figure 2: The synapse converts an optical spike into stored flux. A single-photon detector fires on absorption and deposits a weight-scaled flux packet w into a superconducting loop, whose signal then leaks away.

3 Stage 3 — The dendrite: a leaky integrator from a Lagrangian

A dendrite collects flux from one or more synapses and integrates it. Physically it is a SQUID feeding an output LRC (inductor–resistor–capacitor) branch. Writing the Lagrangian of that branch and applying least action, then keeping the over-damped (low-pass) limit, collapses the five coupled Josephson equations to the *single* first-order equation (1) [1]. With $s = I/I_c$ dimensionless,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \gamma\, g(\phi, s) - \frac{s}{\tau}, \qquad \phi = \sum_j J_{ij}\, s_j \;\text{ (flux coupled in from upstream loops).} \tag{6}$$

For a constant input the steady state is s_∞ with $g(\phi, s_\infty) = s_\infty/(\gamma\tau)$, approached with time constant τ. The leak τ is a genuine design knob: large τ makes the dendrite hold and *sum* inputs over a long window; small τ makes it a fast coincidence element (Fig. 3). This is exactly the leaky-integrator dendrite ubiquitous in computational neuroscience — the formal correspondence is the central result of Ref. [1].

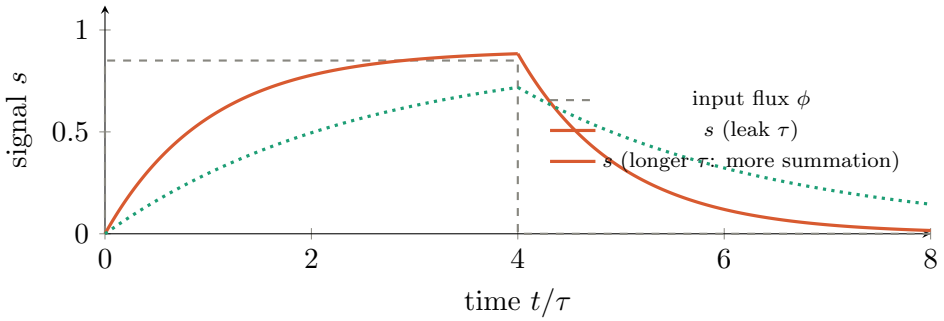


Figure 3: Dendritic step response. During a flux step the loop charges toward saturation; when input ends it leaks at rate $1/\tau$. A longer τ (dotted) integrates and holds longer, summing inputs across time.

4 Stage 4 — The soma: integrate, threshold, fire, refract

The soma is the decision element. It integrates its dendritic input exactly as in Eq. (1), but adds two new ingredients that turn a smooth integrator into a *spiking* neuron:

1. **Threshold + fire.** When the applied flux pushes the somatic signal across threshold θ, the soma's transmitter circuit emits a *photonic action potential* — a burst of photons sent downstream — and the loop partially resets.
2. **Refractory loop.** Firing also charges a second loop r whose signal is *subtracted* from the soma and which decays with its own time constant τ_r. While r is large the effective input cannot reach threshold, enforcing a refractory period.

A compact rule (the form used in the companion model) is

$$\dot{s} = \gamma g(\phi, s) - \tfrac{s}{\tau_s}, \quad \dot{r} = -\tfrac{r}{\tau_r}, \qquad \text{if } s - r \geq \theta: \;\; \text{spike}, \; r \mathrel{+}= \Delta_r, \; s \to \rho s. \tag{7}$$

The maximum sustainable firing rate is set almost entirely by τ_r: longer recovery throttles the rate (Fig. 4). This single feedback loop is the SOEN analogue of the refractory dynamics that, in biology, require dedicated ion channels.

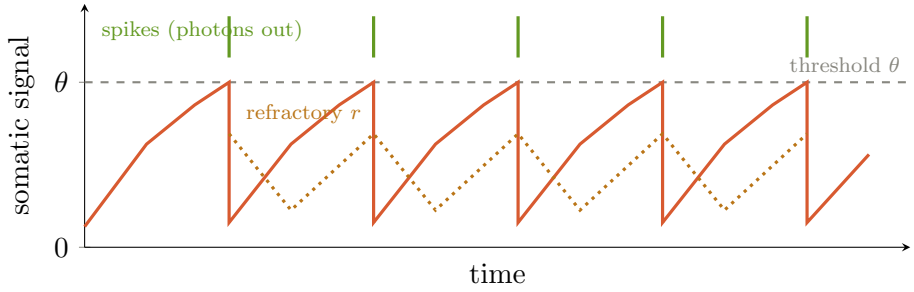


Figure 4: Integrate-and-fire with refractory feedback. The soma charges toward θ; on crossing it emits a spike and resets while the refractory variable r (amber) jumps and decays, suppressing immediate re-firing. The recovery time τ_r sets the maximum firing rate.

5 Stage 5 — Fan-out: why light wins

A firing soma emits light, and light *broadcasts*. A single photonic action potential is split by a passive optical arbor (waveguides) and delivered to thousands of downstream synapses essentially simultaneously (Fig. 5). The cost of adding another target is one more waveguide tap — negligible additional energy and latency. This is the structural advantage of optics over electrical interconnect, where driving a wire to many destinations costs charge $\propto$ capacitance $\propto$ fan-out. High fan-out ($\sim 10^3$–10^4) is what makes brain-like connectivity affordable, and it is the reason the energy argument of Stage 9 can come out in SOEN's favour despite the cooling overhead.

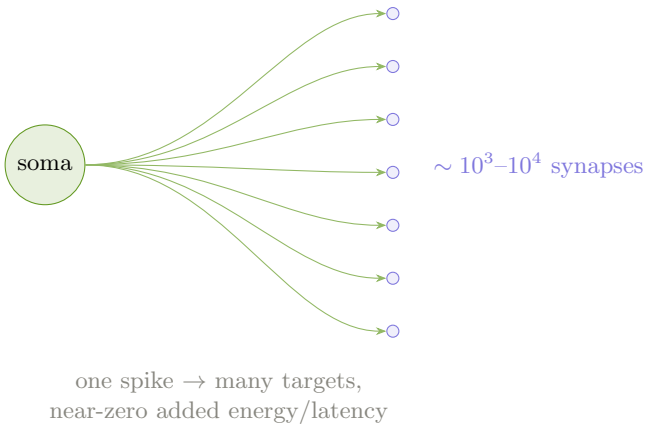


Figure 5: Optical fan-out. One photonic spike is split to a large number of downstream synapses at once. Because the carriers are photons in passive waveguides, broadcasting to many targets adds almost no energy — unlike charging an electrical bus.

Single-neuron summary. A SOEN neuron is: photons $\to$ SPD $\to$ synaptic loop (w) $\to$ dendritic leaky integrator (τ) $\to$ soma (threshold θ, refractory τ_r) $\to$ photonic spike $\to$ optical fan-out. Every loop obeys the same equation $\dot{s} = \gamma g(\phi, s) - s/\tau$; only the wiring and a few constants differ.

6 Stage 6 — From spikes to a network: the rate reduction

To reason about many neurons we abstract each one by its time-averaged *firing rate*. Repeating the Stage-4 experiment at many input levels traces out an f–I curve: silent below threshold, then a monotonic, saturating rise (the saturation set by τ_r). This justifies replacing each spiking neuron by a smooth rate unit $a = f(\text{input})$, with f essentially the source function g again (Fig. 6b). We then wire a layered network — here two inputs, six hidden units, one output — with synaptic weights on the edges (Fig. 6a). Activity propagates left to right; weight sign and magnitude are the only free parameters.

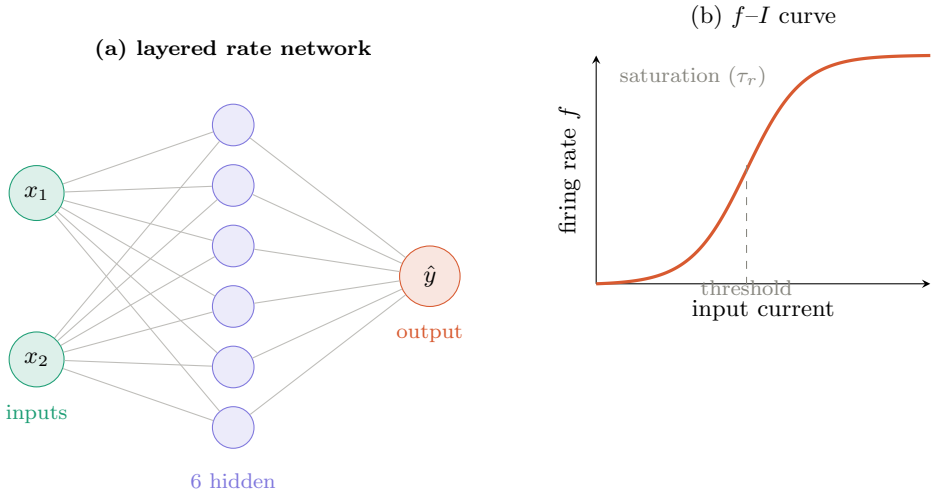


Figure 6: (a) Each neuron is summarized by a firing rate, and neurons are wired into a layered network. (b) The f–I curve: no firing below threshold, then a saturating rise. This smooth nonlinearity is what makes the network analytically tractable and trainable.

7 Stage 7 — Learning by equilibrium propagation

How do the weights get set? SOENs are naturally *energy-based*: the relaxation dynamics settle to a fixed point that minimizes an energy $E(\mathbf{s}; \mathbf{x})$. Equilibrium propagation (EqProp) [3] trains such systems using only *local* information and two relaxations:

1. **Free phase.** Clamp the input $\mathbf{x}$, let the network relax to equilibrium $\mathbf{s}^0$ (output free).
2. **Nudged phase.** Add a weak force pulling the output toward the target y (strength β); relax to $\mathbf{s}^\beta$.

The gradient of the task loss with respect to each weight is then the *difference of two local correlations* measured in the two phases:

$$\Delta W_{ij} \;\propto\; \frac{1}{\beta}\Big(s_i^\beta s_j^\beta \;-\; s_i^0 s_j^0 \Big). \tag{8}$$

No global backward pass, no separate adjoint circuitry: every weight update uses only the activities of the two units it connects, measured before and after a small nudge. In the companion model a 2–6–1 network learns XOR this way — the loss falls and the truth table turns correct as the weights reshape (Fig. 7). The very same relaxation that the dendrites and soma already perform *is* the inference engine; learning is just two settlings and a subtraction.

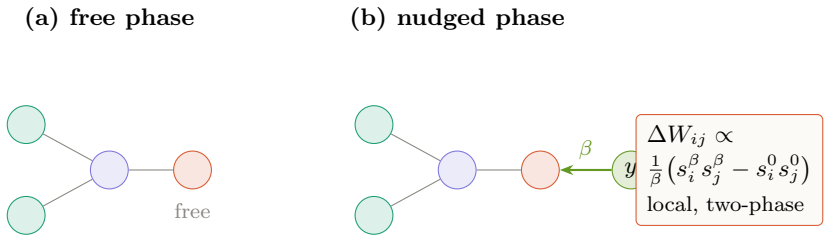


Figure 7: Equilibrium propagation. The network relaxes once with the output free and once with the output gently nudged toward the target. Each weight changes by the difference of the two local correlations — a purely local learning rule that needs no backward pass.

8 Stage 8 — Associative memory: the same physics, run backward

Forward, the network *classifies*. Made recurrent and symmetric, the very same energy dynamics *remember*. A Hopfield network of N units $s_i \in \{-1, +1\}$ with symmetric weights stores patterns $\{\boldsymbol{\xi}^\mu\}$ via the Hebbian prescription

$$W_{ij} = \frac{1}{N}\sum_\mu \xi_i^\mu\, \xi_j^\mu \quad (i \neq j), \qquad E(\mathbf{s}) = -\tfrac{1}{2}\sum_{i,j} W_{ij} s_i s_j. \tag{9}$$

The stored patterns are local minima (*attractors*) of E. Asynchronous updates $s_i \leftarrow \mathrm{sgn}\big(\sum_j W_{ij} s_j\big)$ never increase E, so a corrupted input rolls downhill into the nearest basin — the network *recalls* the clean pattern (Fig. 8). In the companion model a 5×5 network storing three shapes recovers the exact pattern from $\sim$ 5-bit corruption almost every time. Recall is just the relaxation of Stages 3–4 run on a recurrent graph; the descent of E is literally the loops shedding energy into their leaks.

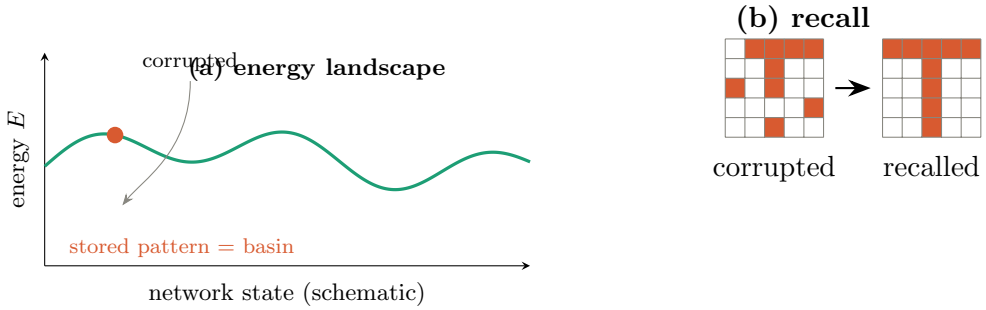


Figure 8: Hopfield associative memory. (a) Stored patterns are minima of an energy E; a corrupted state relaxes downhill into the nearest basin. (b) A 5×5 network recovers the exact stored shape from a corrupted input. The dynamics are identical to forward inference, run recurrently until E stops decreasing.

9 Stage 9 — The power budget: why 4 K still wins

Finally, the joules. SOENs run at $\sim$ 4 K, and cooling is expensive: removing one watt at 4 K costs $r \sim 10^3$ watts at the wall (Carnot plus inefficiency). So why bother? Because the *device* energy per synaptic operation is extraordinarily small — of order $E_s \sim 10^{-17}$ J (tens of attojoules, set by the optical and flux quanta involved) — versus $E_g \sim 10^{-12}$ J ($\sim$pJ) effective per synaptic op on a GPU. For a fixed wall-plug power P we compare deliverable synaptic ops per second:

$$P_{\text{cold}} = \frac{P}{(1+r)\,\eta}, \qquad S_{\text{SOEN}} = \frac{P_{\text{cold}}}{E_s}, \qquad S_{\text{GPU}} = \frac{P}{E_g}, \qquad \text{edge} = \frac{S_{\text{SOEN}}}{S_{\text{GPU}}}, \tag{10}$$

where $\eta \approx 1.4$ bundles support/PUE overhead and only the *cold* power does SOEN compute. With illustrative values $P = 1\,\mathrm{GW}$, $E_s = 20\,\mathrm{aJ}$, $r = 1000$, $E_g = 2\,\mathrm{pJ}$ one finds $S_{\mathrm{SOEN}} \approx 3.6 \times 10^{22}$ vs $S_{\mathrm{GPU}} \approx 5.0 \times 10^{20}$ ops/s — roughly a **70× efficiency edge**, even though only ∼0.07% of the budget reaches the cold compute and ∼71% is spent on cooling (Fig. 9). The device is so efficient that it wins despite paying a thousand-fold cooling tax. (Numbers are order-of-magnitude assumptions, not a vendor benchmark; the model is meant for exploring the trade-off.)

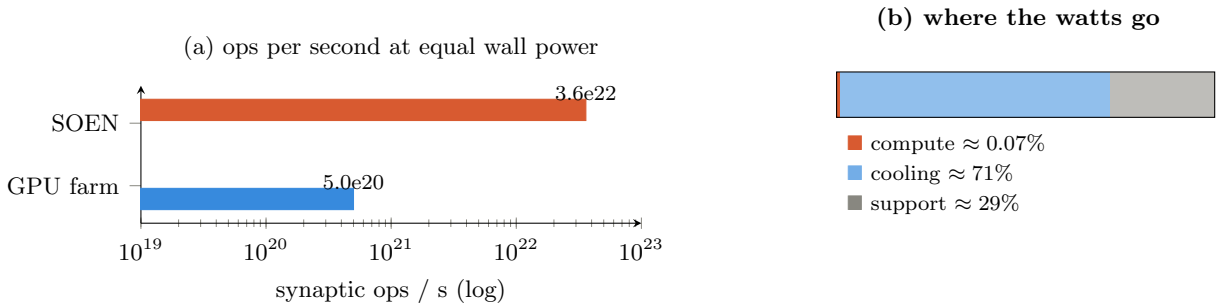


Figure 9: Energy accounting at a fixed wall-plug power. (a) Despite cooling overhead, the tiny per-op device energy gives SOEN a large lead in deliverable synaptic ops per second. (b) Almost the entire budget is spent on cooling and support; the cold compute is a sliver — yet it still out-delivers the GPU.

Synthesis

The arc is a single idea elaborated nine times. One nonlinear device (the SQUID source g) and one dynamical equation ($\dot{s} = \gamma g - s/\tau$) generate the synapse, the dendrite, and the soma. Wiring plus a threshold-and-refractory loop makes a spiking neuron; optics make its output broadcast cheaply. Averaging spikes to rates makes a network; an energy view makes that network both *trainable* (EqProp: two relaxations and a local subtraction) and a *memory* (Hopfield: relax until energy stops falling). The cryogenic cost is real but is bought back by an attojoule-scale per-operation energy and massive optical fan-out.

Stage	Element	Governing idea	Key knob / result
1	SQUID device	$g(\phi, s) = \sigma(k(\phi - \phi_{\mathrm{th}}))(1 - s)$	bias i_b sets excitability
2	Synapse	photon → flux packet, leak τ_{syn}	weight w (analog, stored)
3	Dendrite	$\dot{s} = \gamma g - s/\tau$ (from Lagrangian)	leak τ = integration window
4	Soma	integrate + threshold θ + refractory r	τ_r caps firing rate
5	Fan-out	optical broadcast in waveguides	10^3–10^4 targets, near-free
6	Network	rate reduction, f–I curve	layered 2–6–1
7	Learning	EqProp, $\Delta W \propto \frac{1}{\beta}(s^\beta s^\beta - s^0 s^0)$	local; learns XOR
8	Memory	Hopfield $E = -\frac{1}{2} s^\top W s$	attractor recall from noise
9	Power	$S = P_{\mathrm{cold}}/E_s$ vs P/E_g	∼70× edge at 4 K

References

[1] J. M. Shainline, B. A. Primavera, R. O'Loughlin, *Relating Superconducting Optoelectronic Networks to Classical Neurodynamics*, arXiv:2409.18016 (2024), NIST.

[2] J. M. Shainline *et al.*, *Superconducting optoelectronic loop neurons*, J. Appl. Phys. **126**, 044902 (2019).

[3] B. Scellier, Y. Bengio, *Equilibrium Propagation: Bridging the Gap between Energy-Based Models and Backpropagation*, Front. Comput. Neurosci. **11**, 24 (2017).

[4] J. J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, PNAS **79**, 2554 (1982).