Biologically Plausible Computing: Navigating Energy Landscapes

Francesco Bullo

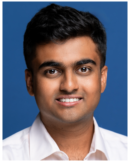
Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
https://fbullo.github.io

Department of Mechanical Engineering and
UCR Artificial Intelligence Research and Education Institute
UC Riverside, February 26, 2026

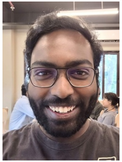
Acknowledgments

Simone Betteti
Italian Institute of Artificial Intelligence

Veronica Centorrino
ETH



Anand Gokhale
UC Santa Barbara



William Retnaraj
UC San Diego

Francesca Rossi
Scuola Sup Meridionale

Giacomo Baggio
Univ Padova

Jorge Cortes
UC San Diego

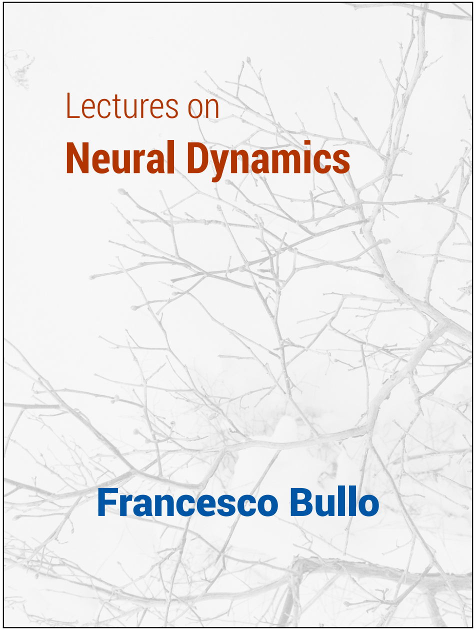
Alexander Davydov
Rice Univ

Giovanni Russo
Univ Salerno

Sandro Zampieri
Univ Padova

AFOSR ARO ONR DTRA/ERDC

Frederick Leve @AFOSR FA9550-22-1-0059
Marc Steinberg @ONR N00014-22-1-2813
Derya Cansever @ARO W911NF-24-1-0228



Lectures on Neural Dynamics, Francesco Bullo, May 2025. (144+xiv pages, 66 figures, 18 exercises)
Latest revision: Jan 23, 2026

1. Textbook with exercises and answers.
2. Content:
   - Neural circuit models based on firing rates and Hopfield networks: dynamics, interconnections, and local Hebbian adaptation rules
   - Stability in dynamic neural networks using Lyapunov methods, multistability, and energy functions
   - Optimization in neural networks through biologically inspired gradient dynamics and sparse representations.
   - Unsupervised learning via neural dynamics, linking Hebbian rules to tasks like PCA, clustering, and similarity-based representation learning.
3. PDF Freely available at:
   https://fbullo.github.io/lnd

"Continuous improvement is better than delayed perfection"
Mark Twain

Outline

§1. Chapter #1: Context and motivation for biologically-plausible neural circuits

§2. Chapter #2: Neural circuits for optimization

§3. Chapter #3: Neural circuits for multiplayer optimization

§4. Conclusion and ongoing research

Despite incredible achievements, deep learning models remain limited in

- **interpretability** (the "black box" problem)
- **computational efficiency** (the power-hungry nature of GPUs)
- **physical grounding** (gap between silicon and biological efficiency)

1. What are the *fundamental limits* of information processing, given the laws of physics?
2. What *architectures* and *strategies* enable the brain's extreme energy efficiency?
3. Can *analog, oscillator-based, and neuromorphic computing* translate these biological principles into silicon?

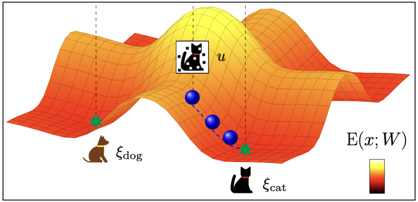
Despite incredible achievements, deep learning models remain limited in

- **interpretability** (the "black box" problem)
- **computational efficiency** (the power-hungry nature of GPUs)
- **physical grounding** (gap between silicon and biological efficiency)

1. What are the *fundamental limits* of information processing, given the laws of physics?
2. What *architectures* and *strategies* enable the brain's extreme energy efficiency?
3. Can *analog, oscillator-based, and neuromorphic computing* translate these biological principles into silicon?
4. To what extent do these questions reduce to cost minimization and energy landscapes?

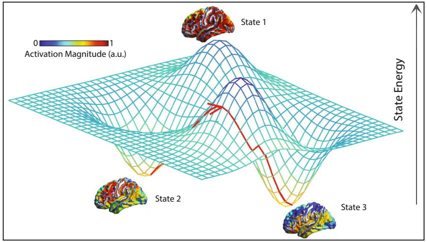
"The idea that the brain functions so as to minimize certain costs pervades theoretical neuroscience."

S. C. Surace, J.-P. Pfister, W. Gerstner, and J. Brea. On the choice of metric in gradient-based theories of brain function. *PLOS Computational Biology*, 16(4):e1007640, 2020.



Energy landscape for associative memory in Hopfield models

S. Betteti, G. Baggio, F. Bullo, and S. Zampieri. Input-driven dynamics for robust memory retrieval in Hopfield networks. *Science Advances*, 11(17), 2025a.



Energy of neurophysiological activity

S. Gu, M. Cieslak, B. Baird, S. F. Muldoon, S. T. Grafton, F. Pasqualetti, and D. S. Bassett. The energy landscape of neurophysiological activity implicit in brain network structure. *Scientific Reports*, 8(1), 2018.

Firing-rate network:

$$\dot{x} = \mathsf{F}_{\mathsf{FR}}(x) := -x + \Phi(Wx + Bu)$$

where W is *synaptic matrix*, Φ is *activation function*, and u is *stimulus*

Firing-rate network:

$$\dot{x} = \mathsf{F}_{\mathsf{FR}}(x) := -x + \Phi(Wx + Bu)$$

where W is *synaptic matrix*, Φ is *activation function*, and u is *stimulus*

1. What functionality does F_{FR} implement?
2. What energy does F_{FR} minimize?
3. Is there an optimization-based top-down framework for neural circuits?
   That is, a framework that derives neural circuits from a mathematical objective?

C. Pehlevan and D. B. Chklovskii. Neuroscience-inspired online unsupervised learning algorithms: Artificial neural networks. *IEEE Signal Processing Magazine*, 36(6):88–96, 2019.

Outline

Regularized optimization problem

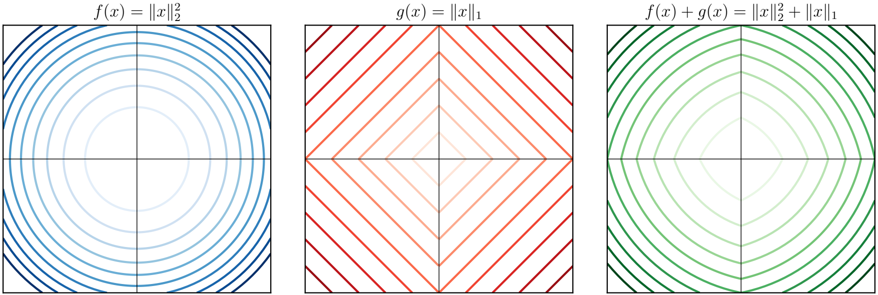
$$\min_{x \in \mathbb{R}^n} \quad \mathcal{E}_{\text{regularized}}(x, u) \quad = \quad f(x, u) \quad + \quad g(x)$$

- nominal cost $f(x, u)$ is well behaved
- regularizer $g(x)$ may be poor behaved

Regularized optimization problem

$$\min_{x\in\mathbb{R}^n} \quad \mathcal{E}_{\text{regularized}}(x, u) \quad = \quad f(x, u) \quad + \quad g(x)$$

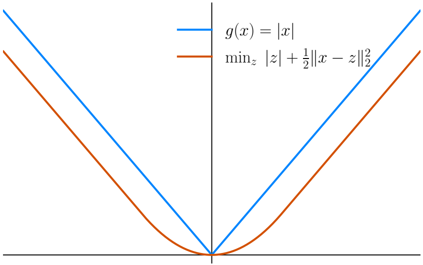
- nominal cost $f(x, u)$ is well behaved
- regularizer $g(x)$ may be poor behaved



proximal operator for the regularizer g

$$\mathrm{prox}_g(x) := \operatorname*{argmin}_{z \in \mathbb{R}^n} \quad g(z) + \frac{1}{2}\|x - z\|_2^2$$

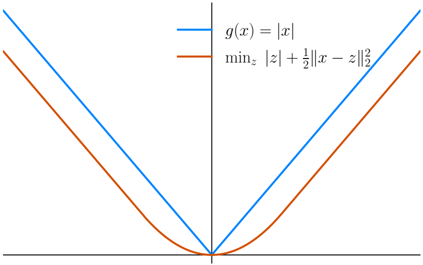
- simple regularized problem
- the quadratic term keeps optimal point close to input x
- the prox is a map that turns x into a "g-better" point



proximal operator for the regularizer g

$$\operatorname{prox}_g(x) \;:=\; \underset{z\in\mathbb{R}^n}{\operatorname{argmin}} \quad g(z) + \frac{1}{2}\|x - z\|_2^2$$

- simple regularized problem
- the quadratic term keeps optimal point close to input x
- the prox is a map that turns x into a "g-better" point



$f(\mathbf{x})$	$\mathrm{dom}(f)$	$\mathrm{prox}_f(\mathbf{x})$	Assumptions	Reference
$\frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$	$\mathbb{R}^n$	$(\mathbf{A}+\mathbf{I})^{-1}(\mathbf{x}-\mathbf{b})$	$\mathbf{A} \in \mathbb{S}^n_+$, $\mathbf{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$	Section 6.2.3
λx^3	$\mathbb{R}_+$	$\frac{-1+\sqrt{1+12\lambda[x]_+}}{6\lambda}$	$\lambda > 0$	Lemma 6.5
μx	$[0,\alpha] \cap \mathbb{R}$	$\min\{\max\{x-\mu,0\},\alpha\}$	$\mu \in \mathbb{R}$, $\alpha \in [0,\infty]$	Example 6.14
$\lambda\lVert\mathbf{x}\rVert$	$\mathbb{E}$	$\left(1 - \frac{\lambda}{\max\{\lVert\mathbf{x}\rVert,\lambda\}}\right)\mathbf{x}$	$\lVert\cdot\rVert$—Euclidean norm, $\lambda > 0$	Example 6.19
$-\lambda\lVert\mathbf{x}\rVert$	$\mathbb{E}$	$\left(1+\frac{\lambda}{\lVert\mathbf{x}\rVert}\right)\mathbf{x}, \quad \mathbf{x} \neq \mathbf{0}$, $\{\mathbf{u} : \lVert\mathbf{u}\rVert = \lambda\}, \quad \mathbf{x} = \mathbf{0}$.	$\lVert\cdot\rVert$—Euclidean norm, $\lambda > 0$	Example 6.21
$\lambda\lVert\mathbf{x}\rVert_1$	$\mathbb{R}^n$	$\mathcal{T}_\lambda(\mathbf{x}) = [\lvert\mathbf{x}\rvert - \lambda\mathbf{e}]_+ \odot \mathrm{sgn}(\mathbf{x})$	$\lambda > 0$	Example 6.8
$\lVert\boldsymbol{\omega} \odot \mathbf{x}\rVert_1$	$\mathrm{Box}[-\boldsymbol{\alpha},\boldsymbol{\alpha}]$	$\mathcal{S}_{\boldsymbol{\omega},\boldsymbol{\alpha}}(\mathbf{x})$	$\boldsymbol{\alpha} \in [0,\infty]^n$, $\boldsymbol{\omega} \in \mathbb{R}^n_+$	Example 6.23
$\lambda\lVert\mathbf{x}\rVert_\infty$	$\mathbb{R}^n$	$\mathbf{x} - \lambda P_{B_{\lVert\cdot\rVert_1}[\mathbf{0},1]}(\mathbf{x}/\lambda)$	$\lambda > 0$	Example 6.48
$\lambda\lVert\mathbf{x}\rVert_a$	$\mathbb{E}$	$\mathbf{x} - \lambda P_{B_{\lVert\cdot\rVert_{a,*}}[\mathbf{0},1]}(\mathbf{x}/\lambda)$	$\lVert\mathbf{x}\rVert_a$—arbitrary norm, $\lambda > 0$	Example 6.47
$\lambda\lVert\mathbf{x}\rVert_0$	$\mathbb{R}^n$	$\mathcal{H}_{\sqrt{2\lambda}}(x_1) \times \cdots \times \mathcal{H}_{\sqrt{2\lambda}}(x_n)$	$\lambda > 0$	Example 6.10
$\lambda\lVert\mathbf{x}\rVert^3$	$\mathbb{E}$	$\frac{2}{1+\sqrt{1+12\lambda\lVert\mathbf{x}\rVert}}\mathbf{x}$	$\lVert\cdot\rVert$—Euclidean norm, $\lambda > 0$,	Example 6.20
$-\lambda\sum_{j=1}^n \log x_j$	$\mathbb{R}^n_{++}$	$\left(\frac{x_j+\sqrt{x_j^2+4\lambda}}{2}\right)_{j=1}^n$	$\lambda > 0$	Example 6.9
$\delta_C(\mathbf{x})$	$\mathbb{E}$	$P_C(\mathbf{x})$	$\emptyset \neq C \subseteq \mathbb{E}$	Theorem 6.24
$\lambda\sigma_C(\mathbf{x})$	$\mathbb{E}$	$\mathbf{x} - \lambda P_C(\mathbf{x}/\lambda)$	$\lambda > 0$, $C \neq \emptyset$ closed convex	Theorem 6.46
$\lambda\max\{x_i\}$	$\mathbb{R}^n$	$\mathbf{x} - \lambda P_{\Delta_n}(\mathbf{x}/\lambda)$	$\lambda > 0$	Example 6.49
$\lambda\sum_{i=1}^k x_{[i]}$	$\mathbb{R}^n$	$\mathbf{x} - \lambda P_C(\mathbf{x}/\lambda)$, $C = H_{\mathbf{e},k} \cap \mathrm{Box}[\mathbf{0},\mathbf{e}]$	$\lambda > 0$	Example 6.50
$\lambda\sum_{i=1}^k \lvert x_{\langle i\rangle}\rvert$	$\mathbb{R}^n$	$\mathbf{x} - \lambda P_C(\mathbf{x}/\lambda)$, $C = B_{\lVert\cdot\rVert_1}[\mathbf{0},k] \cap \mathrm{Box}[-\mathbf{e},\mathbf{e}]$	$\lambda > 0$	Example 6.51
$\lambda M_f^\mu(\mathbf{x})$	$\mathbb{E}$	$\mathbf{x} + \frac{\lambda}{\mu+\lambda}\left(\mathrm{prox}_{(\mu+\lambda)f}(\mathbf{x}) - \mathbf{x}\right)$	$\lambda,\mu > 0$, f proper closed convex	Corollary 6.64
$\lambda d_C(\mathbf{x})$	$\mathbb{E}$	$\mathbf{x} + \min\left\{\frac{\lambda}{d_C(\mathbf{x})},1\right\}(P_C(\mathbf{x}) - \mathbf{x})$	$\emptyset \neq C$ closed convex, $\lambda > 0$	Lemma 6.43
$\frac{\lambda}{2}d_C^2(\mathbf{x})$	$\mathbb{E}$	$\frac{\lambda}{\lambda+1}P_C(\mathbf{x}) + \frac{1}{\lambda+1}\mathbf{x}$	$\emptyset \neq C$ closed convex, $\lambda > 0$	Example 6.65
$\lambda H_\mu(\mathbf{x})$	$\mathbb{E}$	$\left(1 - \frac{\lambda}{\max\{\lVert\mathbf{x}\rVert,\mu+\lambda\}}\right)\mathbf{x}$	$\lambda,\mu > 0$	Example 6.66
$\rho\lVert\mathbf{x}\rVert_1^2$	$\mathbb{R}^n$	$\left(\frac{v_i x_i}{v_i+2\rho}\right)_{i=1}^n$, $\mathbf{v} = \left[\sqrt{\frac{\rho}{\mu}}\lvert\mathbf{x}\rvert - 2\rho\right]_+$, $\mathbf{e}^T\mathbf{v} = 1$ ($\mathbf{0}$ when $\mathbf{x} = \mathbf{0}$)	$\rho > 0$	Lemma 6.70
$\lambda\lVert\mathbf{A}\mathbf{x}\rVert_2$	$\mathbb{R}^n$	$\mathbf{x} - \mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \alpha^*\mathbf{I})^{-1}\mathbf{A}\mathbf{x}$, $\alpha^* = 0$ if $\lVert\mathbf{v}_0\rVert_2 \leq \lambda$; otherwise, $\lVert\mathbf{v}_{\alpha^*}\rVert_2 = \lambda$; $\mathbf{v}_\alpha \equiv (\mathbf{A}\mathbf{A}^T + \alpha\mathbf{I})^{-1}\mathbf{A}\mathbf{x}$	$\mathbf{A} \in \mathbb{R}^{m\times n}$ with full row rank, $\lambda > 0$	Lemma 6.68

A. Beck. *First-Order Methods in Optimization*. SIAM, 2017. ISBN 978-1-61197-498-0

Transcribing regularized optimization into prox gradient descent

$$\min \quad \underbrace{f(x,u)}_{\text{nominal}} \quad + \quad \underbrace{g(x)}_{\text{regularizer}}$$

proximal gradient descent:

$$\dot{x} \;=\; -x \;+\; \mathrm{prox}_g\big(x - \nabla_x f(x,u)\big) \quad =: \quad \mathsf{F}_{\mathsf{ProxG}}(x,u)$$

Transcribing regularized optimization into prox gradient descent

$$\min \quad \underbrace{f(x,u)}_{\text{nominal}} \quad + \quad \underbrace{g(x)}_{\text{regularizer}}$$

proximal gradient descent:

$$\dot{x} \;=\; -x \;+\; \mathrm{prox}_g\big(x - \nabla_x f(x,u)\big) \quad =: \quad \mathsf{F}_{\mathsf{ProxG}}(x,u)$$

note: **energy system, determined by the energies f and g**

(just like gradient descent $\dot{x} = -\nabla_x f$ is determined by the energy f)

S. Hassan-Moghaddam and M. R. Jovanović. Proximal gradient flow and Douglas-Rachford splitting dynamics: Global exponential stability via integral quadratic constraints. *Automatica*, 123:109311, 2021.

A. Gokhale, A. Davydov, and F. Bullo. Proximal gradient dynamics: Monotonicity, exponential convergence, and applications. *IEEE Control Systems Letters*, 8:2853–2858, 2024.

End of the prologue: Result #1: proximal gradient descent = firing rate network

$$
\begin{aligned}
\dot{x} &= \mathsf{F}_{\mathsf{FR}}(x,u) &&:= -x + \Phi\left(Wx + Bu\right)\\
\dot{x} &= \mathsf{F}_{\mathsf{ProxG}}(x,u) &&:= -x + \operatorname{prox}_g\big(x - \nabla_x f(x,u)\big)
\end{aligned}
$$

If f is quadratic in (x,u) and $\Phi(x) = \operatorname{prox}_g(x)$,
then $\mathsf{F}_{\mathsf{ProxG}} = \mathsf{F}_{\mathsf{FR}}$

Result #2: the Hopfield energy is a regularized energy

The firing rate recurrent neural network

$$\dot{x} \;=\; \mathsf{F}_{\mathsf{FR}}(x,u) \;=\; -x \;+\; \Phi(Wx+Bu)$$

is the proximal gradient descent for **Hopfield energy = regularized energy**

$$\mathcal{E}_{\text{regularized}}(x,u) \;=\; \mathcal{E}_{\text{network}}(x,u) + \sum\nolimits_{i=1}^{n} \mathcal{E}_{\text{activation},i}(x_i),$$

- **network energy** captures interaction and effect of stimulus

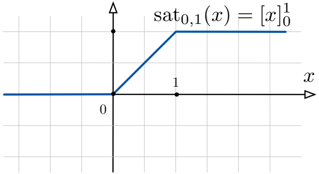
$$\mathcal{E}_{\text{network}}(x,u) \;=\; \tfrac{1}{2}x^\top(I_n - W)x - x^\top Bu$$

- **activation energy** determines activation function

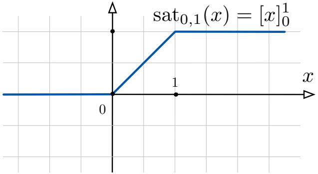
$$\Phi_i(y) \;=\; \operatorname{prox}_{\mathcal{E}_{\text{activation},i}}(y)$$

S. Betteti, W. Retnaraj, A. Davydov, J. Cortes, and F. Bullo. Competition, stability, and functionality in excitatory-inhibitory neural circuits. *Technical report*, 2025c. doi. arXiv:2512.05252

$$g(x) = \begin{cases} 0 & \text{if } 0 \le x \le 1 \\ +\infty & \text{otherwise} \end{cases} \quad \Longrightarrow \quad \operatorname{prox}_g(x) =$$

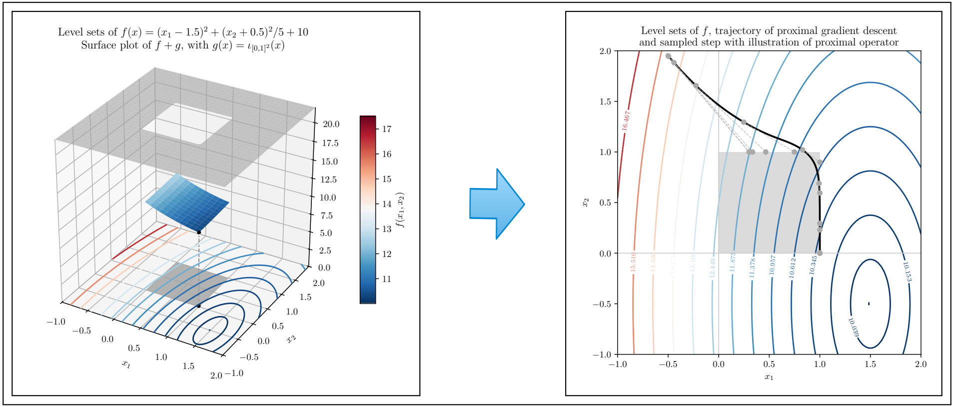


$$g(x) = \begin{cases} 0 & \text{if } 0 \le x \le 1 \\ +\infty & \text{otherwise} \end{cases} \quad \Longrightarrow \quad \mathrm{prox}_g(x) =$$



Firing rate network $=$ **linear threshold model**

$$\dot{x} = -x + [Wx + Bu]_0^1$$



Result #3: Dynamical systems analysis of proximal gradient descent

1. $\mathsf{F}_{\mathsf{ProxG}}$ is **well-posed**, **Lipschitz**, and **uniquely determined by f and g**

2. **equivalence:** x^* minimizes $f + g \iff x^*$ is an equilibrium of $\mathsf{F}_{\mathsf{ProxG}}$

3. **decreasing energy:**

   (when bounded) regularized cost $f + g$ non-increasing along flow

Result #3: Dynamical systems analysis of proximal gradient descent

1. $\mathsf{F}_{\mathsf{ProxG}}$ is **well-posed**, **Lipschitz**, and **uniquely determined by f and g**

2. **equivalence:** x^* minimizes $f + g \iff x^*$ is an equilibrium of $\mathsf{F}_{\mathsf{ProxG}}$

3. **decreasing energy:**

   (when bounded) regularized cost $f + g$ non-increasing along flow

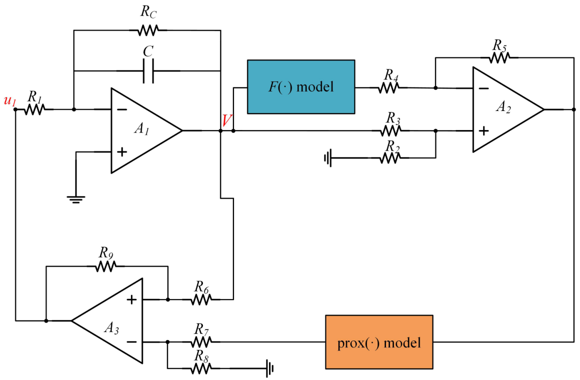
4. **contractivity:**

   $W \prec I_n \implies$ flow along $\mathsf{F}_{\mathsf{ProxG}}$ is a contraction

5. **proximal Polyak–Łojasiewicz condition**

A. Gokhale, A. Davydov, and F. Bullo. Proximal gradient dynamics: Monotonicity, exponential convergence, and applications. *IEEE Control Systems Letters*, 8:2853–2858, 2024.

Result #4: Analog circuit implementation



Analog circuit implementation: 3 amplifiers for each dimension, $F(\cdot)$ denotes ∇f.

J. Wu, X. He, Y. Niu, T. Huang, and J. Yu. Circuit implementation of proximal projection neural networks for composite optimization problems. *IEEE Transactions on Industrial Electronics*, 71(2):1948–1957, 2024.

From regularized energy to firing rate networks

$$\mathcal{E}_{\text{network}}(x,u) + \sum\nolimits_{i=1}^{n} \mathcal{E}_{\text{activation},i}(x_i) \qquad \Longrightarrow \qquad \dot{x} = -x + \Phi(Wx + Bu)$$

- network energy $\mathcal{E}_{\text{network}}$ describe interaction
- regularization terms $\mathcal{E}_{\text{activation},i}$ capture physical limitations

Comments

From regularized energy to firing rate networks

$$\mathcal{E}_{\text{network}}(x,u) + \sum_{i=1}^{n} \mathcal{E}_{\text{activation},i}(x_i) \qquad \Longrightarrow \qquad \dot{x} = -x + \Phi(Wx + Bu)$$

- network energy $\mathcal{E}_{\text{network}}$ describe interaction
- regularization terms $\mathcal{E}_{\text{activation},i}$ capture physical limitations

1. firing-rate dynamics re-interpreted as **proximal gradient dynamics** defined by **regularized energy**
2. **symmetric synapses**
3. **normative framework** $=$ optimization-based top–down framework that derives neural circuits from a mathematical objective

Outline

§1. Chapter #1: Context and motivation for biologically-plausible neural circuits

§2. Chapter #2: Neural circuits for optimization

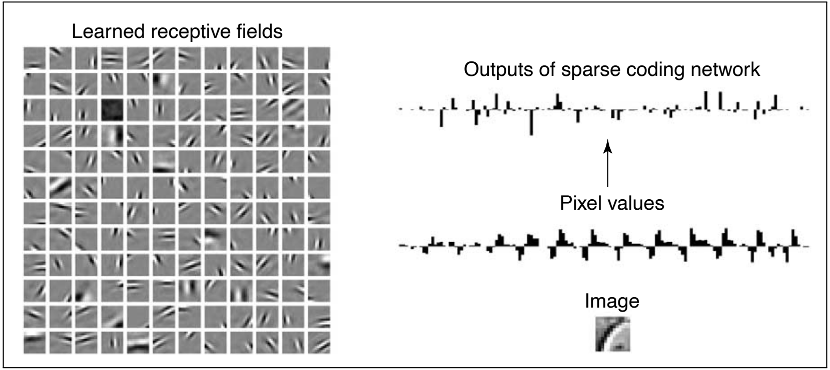
- Proximal gradient descent
- Case study #1: Sparse signal reconstruction
- Case study #2: Policy composition via free energy

§3. Chapter #3: Neural circuits for multiplayer optimization

- Proximal gradient play
- Case study #3: Contrast enhancement via excitatory-inhibitory networks

§4. Conclusion and ongoing research

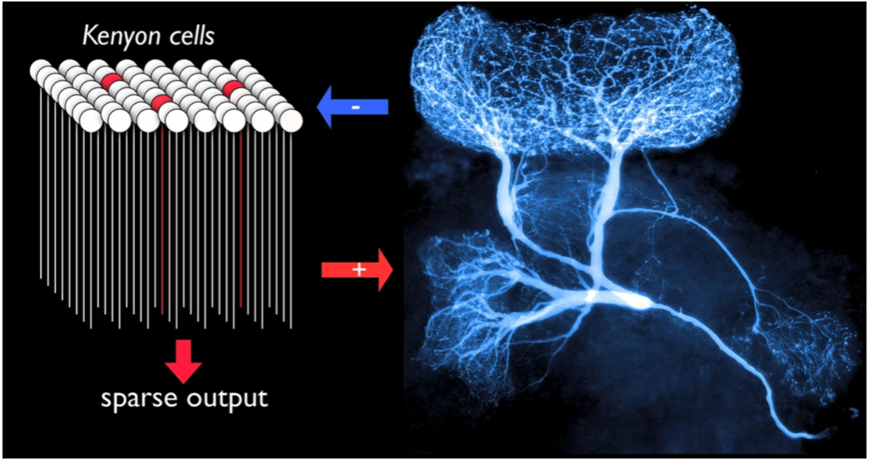
Case study #1: Sparse signal reconstruction



- primary visual area (V1) sparsifies signals
- receptive fields (≈ dictionary) are learned empirically

B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, 1996.

B. A. Olshausen and D. J. Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, 2004.



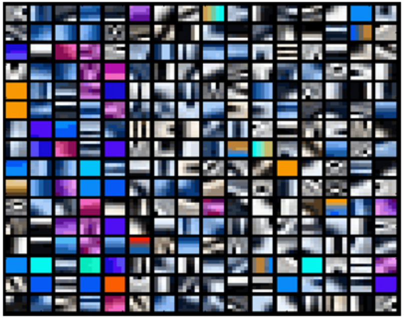
- mushroom body of a locust
- Kenyon cells and a giant (GABAergic) interneuron
- **each excitatory $\rightarrow$ inhibitory interneuron $\rightarrow$ all excitatory** : enables sparse coding

M. Papadopoulou, S. Cassenaer, T. Nowotny, and G. Laurent. Normalization for sparse encoding of odors by a wide-field interneuron. *Science*, 332(6030):721–725, 2011.

Sparse signal reconstruction in engineering

a noisy input image

denoised image, reconstructed from sparse approx



dictionary patches

- identify and exploit sparsity in signals
- dimensionality reduction in machine learning

E. J. Candes and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12):4203–4215, 2005

J. Wright and Y. Ma. *High-Dimensional Data Analysis with Low-Dimensional Models: Principles, Computation, and Applications*. Cambridge University Press, 2022

Positive lasso as a regularized objective

$$\min_{x\in\mathbb{R}^N,\quad x\geq 0} \quad \mathcal{E}_{\text{lasso}}(x) \;:=\; \underbrace{\|u-\Phi x\|_2^2}_{\text{quadratic reconstruction cost}} \;+\; \lambda \underbrace{\|x\|_1}_{\text{sparsity-promoting regularizer}}$$

where Φ *overcomplete dictionary matrix*, with $\|\Phi_i\| = 1$ and $\Phi_i \cdot \Phi_j$ = similarity between (i, j)

$$\underset{(M\times 1)}{\boxed{u}} \;\approx\; \underset{(M\times N)}{\boxed{\Phi}} \; \underset{(N\times 1)}{\boxed{x}} \;=\; \underset{(M\times N)}{\boxed{\Phi_1 \mid \Phi_2 \mid \cdots \mid \Phi_N}} \; \underset{(N\times 1)}{\boxed{x}}$$

where x is k-sparse and $k \ll M \ll N$

Firing rate network for sparse reconstruction

$$\min_{x\in\mathbb{R}^N,\quad x\geq 0} \quad \mathcal{E}_{\mathsf{lasso}}(x) := \|u - \Phi x\|_2^2 \ + \ \lambda\|x\|_1$$

proximal gradient dynamics is **positive competitive network**:

$$\dot{x} = -x + \mathrm{relu}\Big((I_n - \Phi^\top\Phi)x + \Phi^\top u - \lambda\mathbb{1}_n\Big)$$

C. J. Rozell, D. H. Johnson, R. G. Baraniuk, and B. A. Olshausen. Sparse coding via thresholding and local competition in neural circuits. *Neural Computation*, 20(10):2526–2563, 2008.

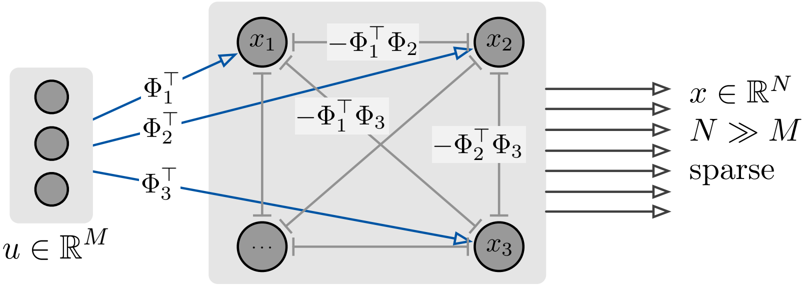
A. Balavoine, J. Romberg, and C. J. Rozell. Convergence and rate analysis of neural networks for sparse approximation. *IEEE Transactions on Neural Networks and Learning Systems*, 23(9):1377–1389, 2012.

V. Centorrino, A. Gokhale, A. Davydov, G. Russo, and F. Bullo. Positive competitive networks for sparse reconstruction. *Neural Computation*, 36(6):1163–1197, 2024.

Biological interpretation = competition via direct lateral inhibition

Nonnegative firing rates and non-negative dictionary elements Φ_i:

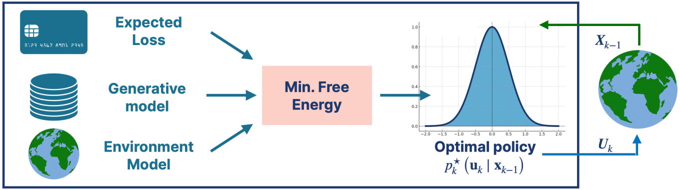
$$\dot{x}_i = -x_i + \mathrm{relu}\left(\sum_{j\neq i} \underbrace{(-\Phi_i^\top \Phi_j)}_{\leq 0,\ \textit{lateral inhibition}} x_j + \underbrace{\Phi_i^\top u}_{\textit{stimulus}} - \underbrace{\lambda}_{\textit{bias}}\right)$$



Outline

Case study #2: The free energy principle

Probabilistic mind theory: information as probabilities + Bayesian inference



Free energy principle: adaptive behaviors in natural/artificial agents arise from minimization of free energy (or "surprise")

- (perception:) adjust beliefs (variational Bayesian inference)
- (learning:) update generative models
- (decision:) change the sensory input (acting so the world matches predictions)

K. Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

A. Shafiei, H. Jesawada, K. Friston, and G. Russo. Distributionally robust free energy principle for decision-making. *Nature Communication*, 2025. . To appear

Optimal policy composition

$$\min_{\text{probabilities } w} \quad \underbrace{\text{surprise}(x, u)}_{\text{prior belief vs actual outcomes}} \;-\; \tau \; \underbrace{\text{entropy}(w)}_{\text{uncertainty}} \qquad \text{(free energy)}$$

$$\text{where} \qquad \text{policy}(u \,|\, x) \;=\; \sum\nolimits_{\alpha} w_a \; \text{primitive}_a(u \,|\, x) \qquad \text{(mixture of policies)}$$

Optimal policy composition

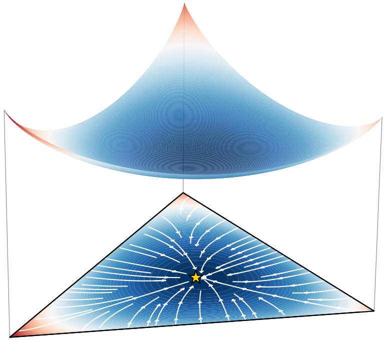
$$\min_{\text{probabilities } w} \quad \underbrace{\text{surprise}(x,u)}_{\text{prior belief vs actual outcomes}} \; - \; \tau \; \underbrace{\text{entropy}(w)}_{\text{uncertainty}} \qquad \text{(free energy)}$$

$$\text{where} \qquad \text{policy}(u \,|\, x) \; = \; \sum_{\alpha} w_a \; \text{primitive}_a(u \,|\, x) \qquad \text{(mixture of policies)}$$

resulting firing rate network =

softmax gradient descent

$$\dot{w} = -w + \mathrm{softmax}\big(-\tau^{-1} \nabla \text{surprise}(x, w)\big)$$



F. Rossi, V. Centorrino, F. Bullo, and G. Russo. Neural policy composition from free energy minimization. *Technical report*, 2025. arXiv:2512.04745

Outline

§1. Chapter #1: Context and motivation for biologically-plausible neural circuits

§2. Chapter #2: Neural circuits for optimization
- Proximal gradient descent
- Case study #1: Sparse signal reconstruction
- Case study #2: Policy composition via free energy

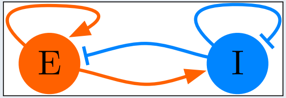
§3. Chapter #3: Neural circuits for multiplayer optimization
- Proximal gradient play
- Case study #3: Contrast enhancement via excitatory-inhibitory networks

§4. Conclusion and ongoing research

Dale's law: a neuron has the same type of effect, inhibitory or excitatory, on all its neighbors.

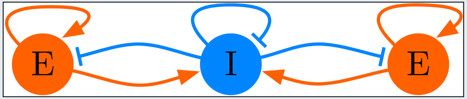
Dale's law: a neuron has the same type of effect, inhibitory or excitatory, on all its neighbors.

Classic motifs obeying Dale's law, with excitatory (E) and inhibitory (I) neurons



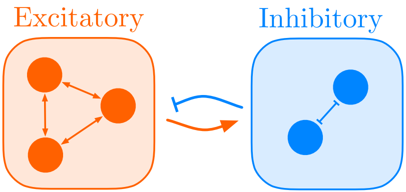
Wilson-Cowan model excitatory-inhibitory pair

H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12 (1):1–24, 1972.



Central inhibitory neuron mediates winner-take-all dynamic between two excitatory neurons

Dale's law (neuromodulator version): each neuron releases the same type of neuromodulator at all of its synapses.



For asymmetric/E-I networks,

- rich dynamic behavior is possible:
  global asymp. stability, multistability, limit cycles, chaotic behavior, etc
- lack of general analysis framework (stability and functionality)
- lack of general design framework (e.g., optimization-based, top-down)

Results on asymmetric networks

1. novel interpretation: **neurons are playing a game**
2. monostability
3. functionality

S. Betteti, W. Retnaraj, A. Davydov, J. Cortes, and F. Bullo. Competition, stability, and functionality in excitatory-inhibitory neural circuits. *Technical report*, 2025c. doi. arXiv:2512.05252

Symmetric networks: $\dot{x} = -x + \Phi(Wx + Bu)$ is **proximal gradient descent** for

$$\mathcal{E}_{\text{regularized}}(x,u) \;=\; \mathcal{E}_{\text{network}}(x,u) + \sum_{i=1}^{n} \mathcal{E}_{\text{activation},i}(x_i)$$

where $\mathcal{E}_{\text{network}}(x,u) \;=\; \frac{1}{2}x^\top(I_n - W)x - x^\top Bu \qquad \phi_i(y) \;=\; \mathrm{prox}_{\mathcal{E}_{\text{activation},i}}(y)$

Result #1: Neural circuits for multiplayer optimization

Symmetric networks: $\dot{x} = -x + \Phi(Wx + Bu)$ is **proximal gradient descent** for

$$\mathcal{E}_{\text{regularized}}(x, u) \quad = \quad \mathcal{E}_{\text{network}}(x, u) + \sum_{i=1}^{n} \mathcal{E}_{\text{activation},i}(x_i)$$

where $\mathcal{E}_{\text{network}}(x, u) \;=\; \frac{1}{2}x^\top (I_n - W)x - x^\top Bu \qquad \phi_i(y) \;=\; \mathrm{prox}_{\mathcal{E}_{\text{activation},i}}(y)$

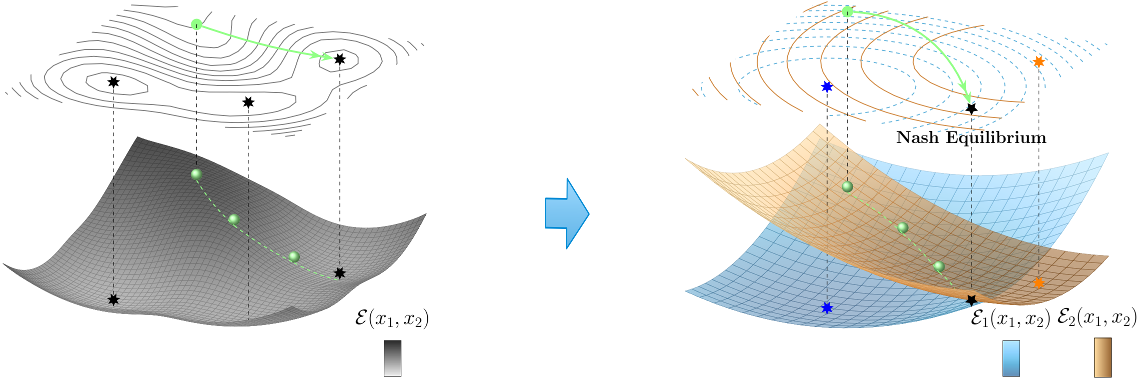
Asymmetric networks: $\dot{x} = -x + \Phi(Wx + Bu)$ is **proximal gradient play** for

$$\mathcal{E}_{\text{regularized},i}(x_i, x_{-i}, u) \quad = \quad \mathcal{E}_{\text{individual},i}(x_i, x_{-i}, u) + \mathcal{E}_{\text{activation},i}(x_i)$$

where

$$\mathcal{E}_{\text{individual},i}(x_i, x_{-i}, u) \;=\; \sum_{j=1}^{n} (\tfrac{1}{2}\delta_{ij} - 1)W_{ij}x_i x_j - x^\top Bu \qquad \phi_i(y) \;=\; \mathrm{prox}_{\mathcal{E}_{\text{activation},i}}(y)$$

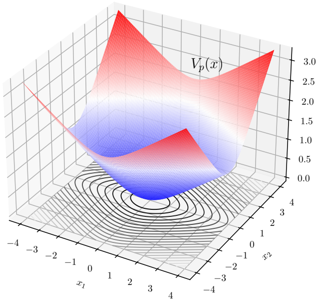
Result #1: Neural circuits for multiplayer optimization



S. Betteti, W. Retnaraj, A. Davydov, J. Cortes, and F. Bullo. Competition, stability, and functionality in excitatory-inhibitory neural circuits. *Technical report*, 2025c. doi. arXiv:2512.05252

Results on asymmetric networks

1. novel interpretation: neurons are playing a game
2. monostability: **monostability via constraints on synaptic weights**
3. functionality



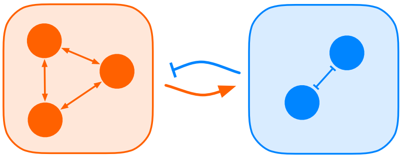
S. Betteti, W. Retnaraj, A. Davydov, J. Cortes, and F. Bullo. Competition, stability, and functionality in excitatory-inhibitory neural circuits. *Technical report*, 2025c. . arXiv:2512.05252

Result #2: Monostability for E-I networks

$\dot{x} = -x + \Phi(Wx + u),$ satisfying *Dale's law*: each neuron is either E or I

1. for each $i \in \mathrm{E}$ and $j \in \mathrm{I}$, *reciprocal connections*

$$(i, j) \text{ is an edge} \quad \Longleftrightarrow \quad (j, i) \text{ is an edge}$$



2. *synaptic weights are homogeneous*:
   - w_{EE} = weight of each E to E
   - w_{EI} = weight of each I to E
   - w_{IE} = weight of each E to I
   - w_{II} = weight of each I to I

Result #2: Monostability for E-I networks

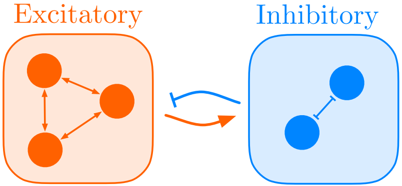
$\dot{x} = -x + \Phi(Wx + u)$, satisfying *Dale's law*: each neuron is either E or I

1. for each $i \in \mathrm{E}$ and $j \in \mathrm{I}$, *reciprocal connections*

   (i, j) is an edge $\iff$ (j, i) is an edge

2. *synaptic weights are homogeneous*:

   w_{EE} = weight of each E to E
   w_{EI} = weight of each I to E
   w_{IE} = weight of each E to I
   w_{II} = weight of each I to I



Monostability (single eq. point exists and is globally asymp stable) if

$$\left(\frac{\text{degree}_{\text{in}} + \text{degree}_{\text{out}}}{2}\right) w_{\mathrm{EE}} < 1 \quad \text{and} \quad \left(\frac{\text{degree}_{\text{in}} + \text{degree}_{\text{out}}}{2} - 2\right) w_{\mathrm{II}} < 1$$

Results on asymmetric networks

1. novel interpretation: neurons are playing a game
2. monostability: monostability via constrains on synaptic weights
3. functionality: **contrast enhancement via lateral inhibition**
    1. Lateral inhibition in E-I-E networks
    2. Winner-take-all in E^k-I networks
    3. Contrast enhancement in columns of E-I-E motifs

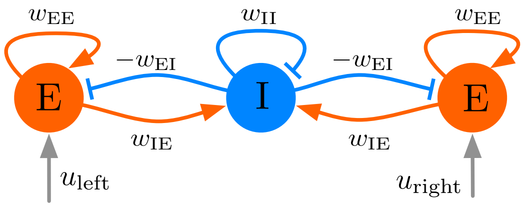
S. Betteti, W. Retnaraj, A. Davydov, J. Cortes, and F. Bullo. Competition, stability, and functionality in excitatory-inhibitory neural circuits. *Technical report*, 2025c. doi. arXiv:2512.05252

Result #3: Lateral inhibition in E-I-E networks

$$\dot{x} = -x + [Wx + Bu]_0^1$$

satisfying Dale's law with:

$$\begin{cases} w_{\mathrm{EE}} < 1 & \text{(monostability)} \\ w_{\mathrm{IE}} \geq 1 + w_{\mathrm{II}} & \text{(functionality)} \end{cases}$$

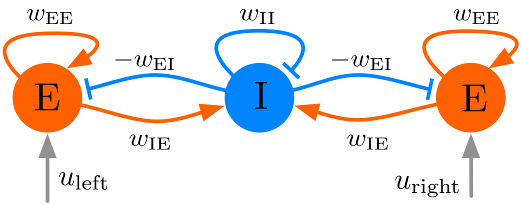


Result #3: Lateral inhibition in E-I-E networks

$$\dot{x} = -x + [Wx + Bu]_0^1$$

satisfying Dale's law with:

$$\begin{cases} w_{\mathrm{EE}} < 1 & \text{(monostability)} \\ w_{\mathrm{IE}} \geq 1 + w_{\mathrm{II}} & \text{(functionality)} \end{cases}$$



lateral inhibition leads to **binary decisions**:

when $u_{\text{left}} > u_{\text{right}} + \delta$, then (left E is high) and (right E is low)
when $u_{\text{right}} > u_{\text{left}} + \delta$, then vice-versa

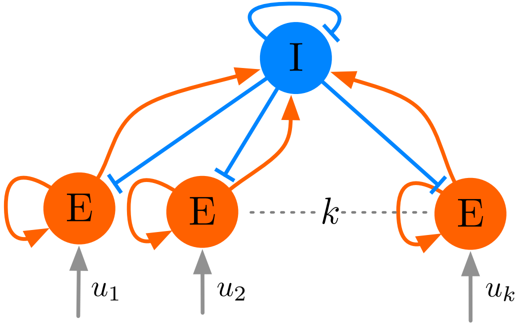
where $\delta := 1 - w_{\mathrm{EE}} + w_{\mathrm{EI}} > 0$

Result #3: Winner-take-all in E^k-I networks

$$\dot{x} = -x + [Wx + Bu]_0^1$$

satisfying Dale's law with:

$$\begin{cases} w_{\mathrm{EE}} < 1 & \text{(monostability)} \\ w_{\mathrm{IE}} \geq 1 + w_{\mathrm{II}} & \text{(functionality)} \end{cases}$$

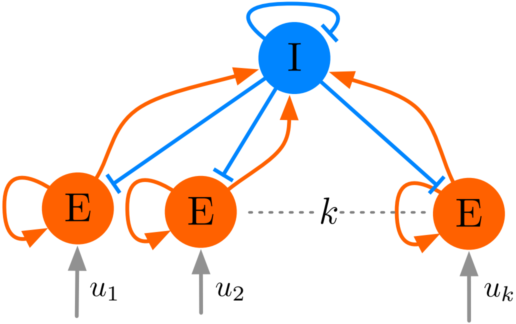


Result #3: Winner-take-all in E^k-I networks

$$\dot{x} = -x + [Wx + Bu]_0^1$$

satisfying Dale's law with:

$$\begin{cases} w_{\mathrm{EE}} < 1 & \text{(monostability)} \\ w_{\mathrm{IE}} \geq 1 + w_{\mathrm{II}} & \text{(functionality)} \end{cases}$$



mutual inhibition leads to **winner-take-all**:

when $u_i > u_j + 2\delta$, then (E_i is high) and (every other neuron j is low)

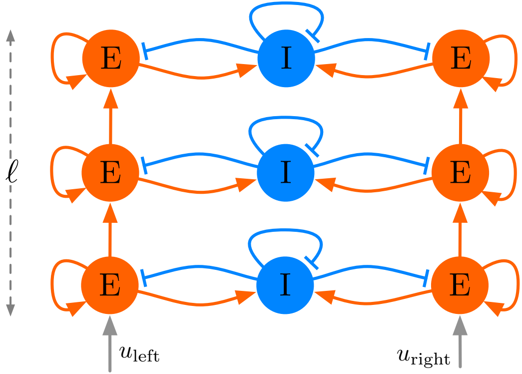
where $\delta := 1 - w_{\mathrm{EE}} + w_{\mathrm{EI}} > 0$

Result #3: Contrast enhancement in columns of E-I-E motifs

$$\dot{x} = -x + [Wx + Bu]_0^1$$

satisfying Dale's law with:

$$\begin{cases} w_{\mathrm{EE}} < 1/2 & \text{(monostability)} \\ w_{\mathrm{IE}} \geq 1 + w_{\mathrm{II}} & \text{(functionality)} \end{cases}$$

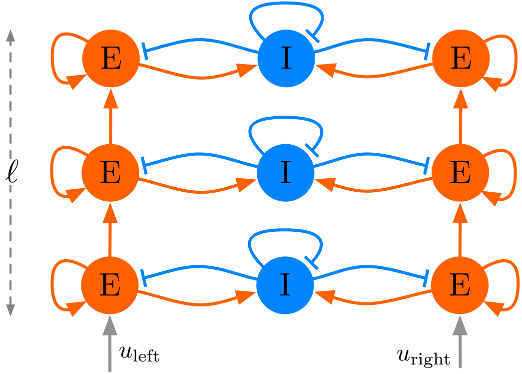


Result #3: Contrast enhancement in columns of E-I-E motifs

$$\dot{x} = -x + [Wx + Bu]_0^1$$

satisfying Dale's law with:

$$\begin{cases} w_{\mathrm{EE}} < 1/2 & \text{(monostability)} \\ w_{\mathrm{IE}} \geq 1 + w_{\mathrm{II}} & \text{(functionality)} \end{cases}$$



competition amount E pathways leads to **contrast enhancement**:

take $u_{\text{left}} > u_{\text{right}} + 2\epsilon$, for some small ϵ

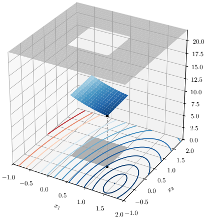
if (number of layers) $\ell \geq \ell_{\text{binary}} := 1 + \dfrac{\ln(\epsilon/\delta)}{\ln(1/w_{\mathrm{EE}} - 1)}$

then, at layer $\ell \geq \ell_{\text{binary}}$, $u_{\text{left}} > u_{\text{right}} + \delta$, and full contrast enhancement

Outline

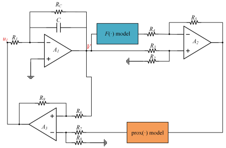
Future work at intersection of dynamical neuroscience & ML

1. **system-theoretic problems in neuroscience**
   - biologically-plausible learning and control
   - stimulus-driven cognitive phenomena
   - computational paradigms: astrocytes, dendritic computation, Hebbian learning, equilibrium propagation



2. **connections with ML**
   - unsupervised representation learning
   - self-attention dynamics and transformers
   - structured state space sequence models

3. **connections with nonconventional and analog computing**
   - analog implementation of prox gradient descent
   - analog implementation of proximal primal-dual gradient descent
   - oscillator-based computing



References

❶ biologically plausible optimization

V. Centorrino, A. Gokhale, A. Davydov, G. Russo, and F. Bullo. Positive competitive networks for sparse reconstruction. Neural Computation, 36(6):1163–1197, 2024.

A. Gokhale, A. Davydov, and F. Bullo. Proximal gradient dynamics: Monotonicity, exponential convergence, and applications. IEEE Control Systems Letters, 8:2853–2858, 2024.

V. Centorrino, F. Bullo, and G. Russo. Similarity matching networks: Hebbian learning and convergence over multiple time scales. Neural Computation, June 2025. . To appear

F. Rossi, V. Centorrino, F. Bullo, and G. Russo. Neural policy composition from free energy minimization. Technical report, 2025. . arXiv:2512.04745

S. Betteti, W. Retnaraj, A. Davydov, J. Cortes, and F. Bullo. Competition, stability, and functionality in excitatory-inhibitory neural circuits. Technical report, 2025c. . arXiv:2512.05252

❷ stimulus-driven energy models for associative memory

S. Betteti, G. Baggio, F. Bullo, and S. Zampieri. Input-driven dynamics for robust memory retrieval in Hopfield networks. Science Advances, 11(17), 2025a.

S. Betteti, G. Baggio, F. Bullo, and S. Zampieri. Firing rate models as associative memory: Excitatory-inhibitory balance for robust retrieval. Neural Computation, pages 1–32, 08 2025b.