Neuronal Circuits for Robust Online Fixed-Point Detection

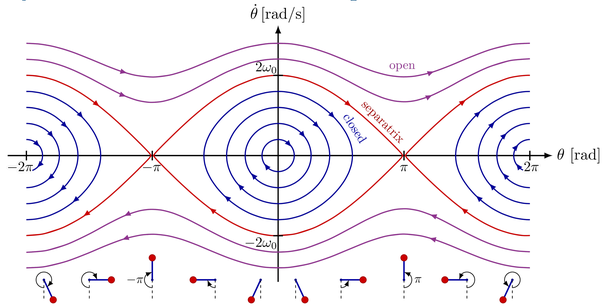
"Tony" Runzhe Yang[1], David Lipshutz[2], Tiberiu Tesileanu[2], Johannes Friedrich[2], Dmitri Chklovskii[2,3]

[1] Computer Science Dept. & Neuroscience Institute, Princeton University
[2] Center of Computational Neuroscience, Flatiron Institute [3] Neuroscience Institute, NYU Langone Health

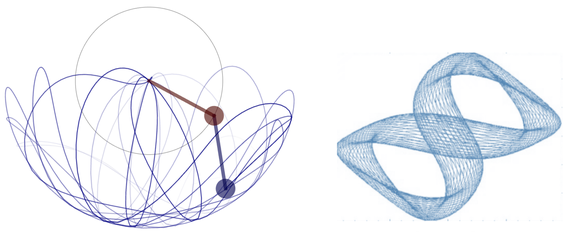
FLATIRON INSTITUTE

COSYNE 2023

How the brain learns the non-linear dynamics of the complex world?



Phase Portrait of the Single Pendulum



Double Pendulum

θ_1 vs θ_2

Learning topological skeleton of dynamics

- Data-driven learning in high-dimensional spaces requires lifting the ***curse of dimensionality***.
- It's more efficient to learn from observed trajectories the underlying **topological skeleton**, i.e., to learn the ***fixed points*** connected by *invariant manifolds* of a dynamical system.
- Detecting ***stable*** and ***unstable fixed points*** is critical for balance or postural maintenance.
- To be implemented by a *neuronal circuit*, the learning algorithm must be ***online*** and ***local***.

Offline fixed-point estimation

time-invariant system w/ Gaussian noise
$$\mathbf{x}_t = \mathbf{F}(\mathbf{x}_{t-1}) + \epsilon_t$$

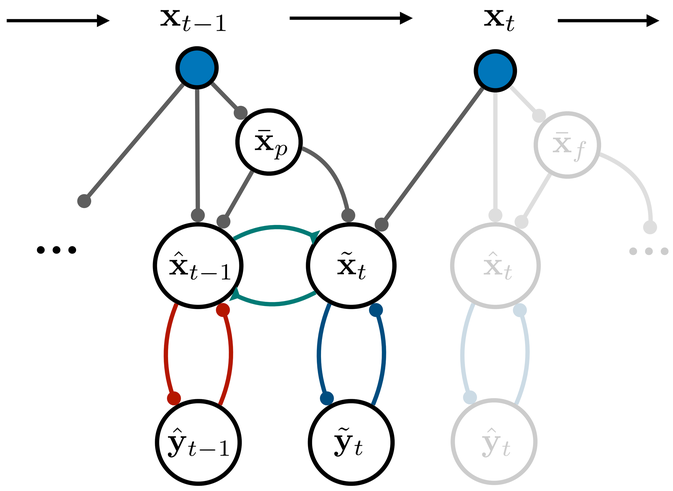
linearization around a fixed point *b*
$$\mathbf{b} = \mathbf{F}(\mathbf{b})$$
$$\mathbf{x}_t \approx \mathbf{b} + \tilde{\mathbf{A}}(\mathbf{x}_{t-1} - \mathbf{b}) + \epsilon_t$$

constructing "past" and "future" from a single trajectory
$$\underbrace{\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_{T-1}}_{\mathbf{X}_p}, \underbrace{\mathbf{x}_T}_{\mathbf{X}_f}$$

unbiased offline estimate of the Jacobian matrix
$$\tilde{\mathbf{A}} \approx \mathbf{C}_{fp}\mathbf{C}_{pp}^{-1} \quad C_{fp} := \mathrm{Cov}(X_f, X_p),\ C_{pp} := \mathrm{Cov}(X_p, X_p)$$

least-square estimate fixed point *b*
$$\mathbf{b} \approx \left(\mathbf{I} - \tilde{\mathbf{A}}\right)^{+} \left(\mathbf{x}_t - \tilde{\mathbf{A}}\mathbf{x}_{t-1}\right)$$

Phase I: online learning of forward dynamics



$$\tilde{\mathbf{A}} = \mathbf{C}_{fp}\mathbf{C}_{pp}^{-1} \approx \mathbf{M}_{fp}\mathbf{W}_{pp}$$

Step 1. maintain trajectory mean
$$\bar{\mathbf{x}}_p \leftarrow (1-\eta_p)\cdot\mathbf{x}_p + \eta_p \cdot \mathbf{x}_{t-1}$$

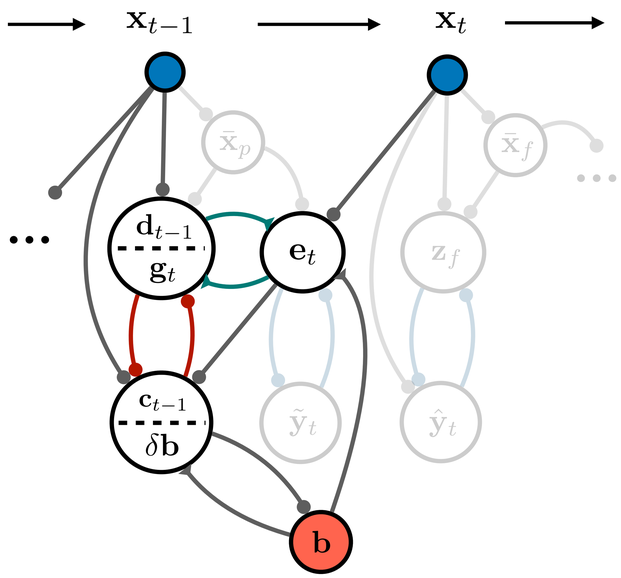
Step 2. deviation from mean
$$\hat{\mathbf{x}}_{t-1} \leftarrow \mathbf{x}_{t-1} - \bar{\mathbf{x}}_p$$
$$\tilde{\mathbf{x}}_t \leftarrow \mathbf{x}_t - \bar{\mathbf{x}}_p$$

Step 3. approximate $\mathbf{C}_{fp}$ and $\mathbf{C}_{pf}$ with ***Hebbian plasticity***
$$\mathbf{M}_{fp} \leftarrow (1-\eta_m)\cdot\mathbf{M}_{fp} + \eta_m \cdot \tilde{\mathbf{x}}_t\hat{\mathbf{x}}_{t-1}^\top$$
$$\mathbf{M}_{pf} \leftarrow (1-\eta_m)\cdot\mathbf{M}_{pf} + \eta_m \cdot \hat{\mathbf{x}}_{t-1}\tilde{\mathbf{x}}_t^\top$$

Step 4. approximate $\mathbf{C}_{pp}^{-1}$ with *Sherman–Morrison* formula
$$\hat{\mathbf{y}}_{t-1} \leftarrow \mathbf{W}_{pp}\hat{\mathbf{x}}_{t-1}$$
$$\mathbf{W}_{pp} \leftarrow \frac{1}{1-\eta_w}\mathbf{W}_{pp} - \frac{\eta_w(1-\eta_w)^{-1}}{1-\eta_w+\eta_w\cdot\hat{\mathbf{x}}_{t-1}^\top\hat{\mathbf{y}}_{t-1}}\cdot\hat{\mathbf{y}}_{t-1}\hat{\mathbf{y}}_{t-1}^\top$$

* an unbiased estimation can be obtained by replacing one $\hat{\mathbf{y}}_{t-1}$ with $\tilde{\mathbf{y}}_{t-1} = \mathbf{W}_{pp}\tilde{\mathbf{x}}_{t-1}$

Phase II: fixed-point Inference with the learned recurrent circuit



iterate

Step 1. subtract **b** from inputs
$$\mathbf{c}_{t-1} \leftarrow \mathbf{x}_{t-1} - \mathbf{b}$$

Step 2. calculate forward errors
$$\mathbf{d}_{t-1} \leftarrow \mathbf{W}_{pp}\mathbf{c}_{t-1}$$
$$\mathbf{e}_t \leftarrow (\mathbf{x}_t - \mathbf{b}) - \mathbf{M}_{fp}\mathbf{d}_{t-1}$$

Step 3. calculate -gradient w.r.t. **b**
$$\mathbf{g}_t \leftarrow -\mathbf{M}_{pf}\mathbf{e}_t$$
$$\delta\mathbf{b} \leftarrow \mathbf{e}_t + \mathbf{W}_{pp}\mathbf{g}_t$$

Step 4. update **b**
$$\mathbf{b} \leftarrow \mathbf{b} + \delta\mathbf{b}$$

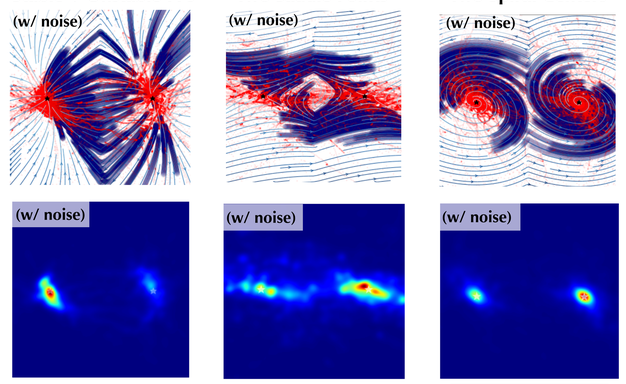
$$\min_{\mathbf{b}} \|(\mathbf{x}_t - \mathbf{b}) - \mathbf{M}_{fp}\mathbf{W}_{pp}(\mathbf{x}_{t-1} - \mathbf{b})\|_F^2$$

Robust fixed-point predictions

Piecewise Linear Systems

• fixed-point predictions
• trajectories

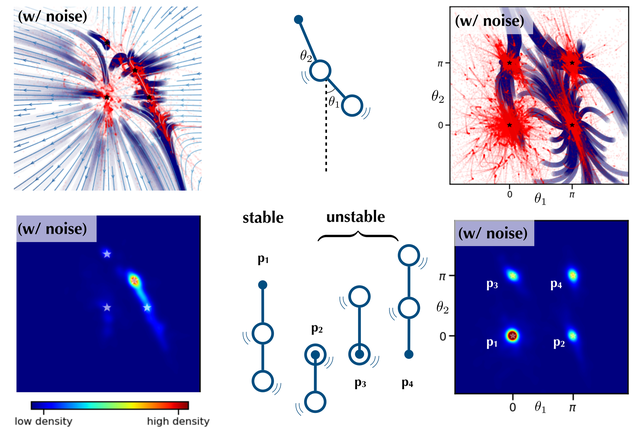
Stable + Unstable FPs | Two Saddle Points | Two Spiral Centers



Nonlinear Systems

• fixed-point predictions
• trajectories

Stable/Unstable/Saddle FPs | Double Pendulum



A building block of bio-plausible neural circuits for control

- We design a ***neural circuit*** that can ***efficiently*** *and* ***robustly*** predict fixed point from a ***single trajectory***. The learning rule is ***mostly local***.
- Applying a bio-plausible **online *K*-means** [1] on predicted fixed-points (can be further filtered by error $|\mathbf{e}_t|$) would recover all fixed-points.
- The learned weights approximate the **forward (or backward) dynamics**, which is also useful for **dynamical mode decomposition** [2] and estimating the **long-term Lyapunov exponent** with bio-plausible algorithms [3].

[1] Pehlevan, Chklovskii, A Hebbian/anti-Hebbian network derived from online non-negative matrix factorization can cluster and discover sparse features. ACSSC 2014
[2] Kutz et al. Dynamic mode decomposition: data-driven modeling of complex systems. SIAM, 2016.
[3] Lipshutz, et al. A biologically plausible neural network for multichannel canonical correlation analysis. Neural computation 2021