When does the leading eigenvector agree with cluster means? A fixed-point analysis of Oja's and SoftHebb's streaming rules

Author: Ovidiu-Alexandru Argherie | **Supervisors**: Stephanie Tan, Yaqi Guo

1. Introduction

Backpropagation is poorly suited for neuromorphic and energy constrained hardware, for the reasons as follows:

1. **biologically implausible**, as the brain does not send errors signals backwards along the same forward connection [1];
2. **update locked**, as layers must wait for the full forward and backward pass, thus no local, asynchronous updates [1];
3. **memory bottleneck**, as all intermediate activations must be stored for the backward pass, which is very costly for large models [1].

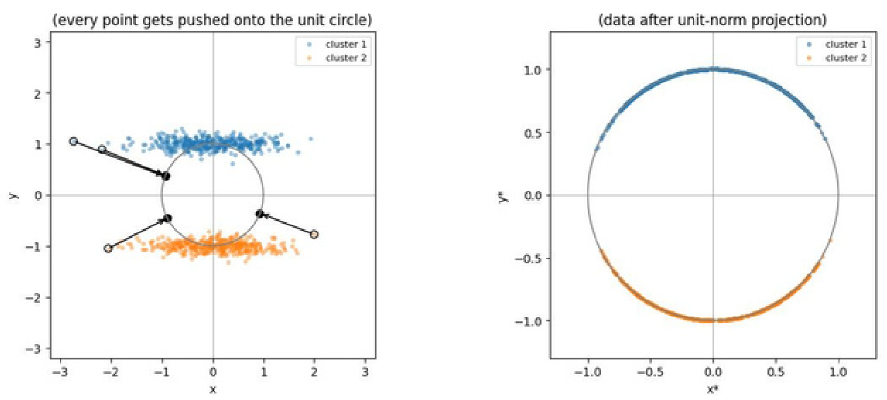
This motivates backpropagation-free, local learning rules, one of which is Hebbian learning, the oldest and the most biologically grounded.

2. Background

Hebbian rule. Neurons that fire together, they wire together. It is local and biologically plausible, but unstable, as weights can go without bound.

Oja [2]. It adds a normalizing decay to the hebbian rule. It converges to the top eigenvector of the data covariance and computes the first principal component, from purely local information [3].

SoftHebb [4]. It extends Oja's rule to a soft winner-take-all network of K neurons. Its weights converge to the normalized component means, making it behave analogously to K-means.



$$\Delta \boldsymbol{w}^{(\mathrm{Oja})} = \eta\, y\, (\boldsymbol{x} - y\, \boldsymbol{w}) \qquad \Delta \boldsymbol{w}_k^{(\mathrm{SoftHebb})} = \eta\, y_k\, (\boldsymbol{x} - u_k\, \boldsymbol{w}_k)$$

Gap. Oja's rule and SoftHebb are built from the same Hebbian parts, yet they compute different objectives. In the batch settings [5], PCA and K-means are tightly connected. Does this also survive in the streaming rules?

3. Methodology

Controlled synthetic data. Balanced two cluster with 2D Gaussian mixture, shared covariance $diag\left(\sigma_1^2, \sigma_2^2\right)$, and centers placed symmetrically about the origin. Thus, we can vary $r = \sigma_1/d$ and fix everything else.

Two configurations. First, isotropic, with center at (±d,0) and separation along the x-axis. Second, anisotropic, with centers at (0, ±d) and $\sigma_1 > \sigma_2$. That is, separation along the y-axis and the elongation along the x-axis.

Key	σ_1	σ_2	d	Sep. axis	$r = \sigma_1/d$	N
Isotropic	1.0	1.0	3.0	x	0.33	2500
Anisotropic	$r \cdot d$	0.1	1.0	y	$0.7 \to 2.2$	1200

4. Theoretical Analysis

Shared lens. A fixed-point analysis recipe (i.e., ODE method of stochastic approximation [6]) that reduces each rule to a stable fixed point.

Isotropic two clusters. Total covariance is diag(10,1), thus Oja lands on x-axis. SoftHebb's normalized component means lie on x-axis, hence agreement.

Naive anisotropic prediction. Normalized generative's cluster means yield SoftHebb on y-axis, whilst Oja on x-axis for any $r > 1.005$, hence divergence.

Why does it fail? Large orthogonal spread collapeses the clusters into overlapping arcs near the x-axis, thus the best way split becomes y-axis.

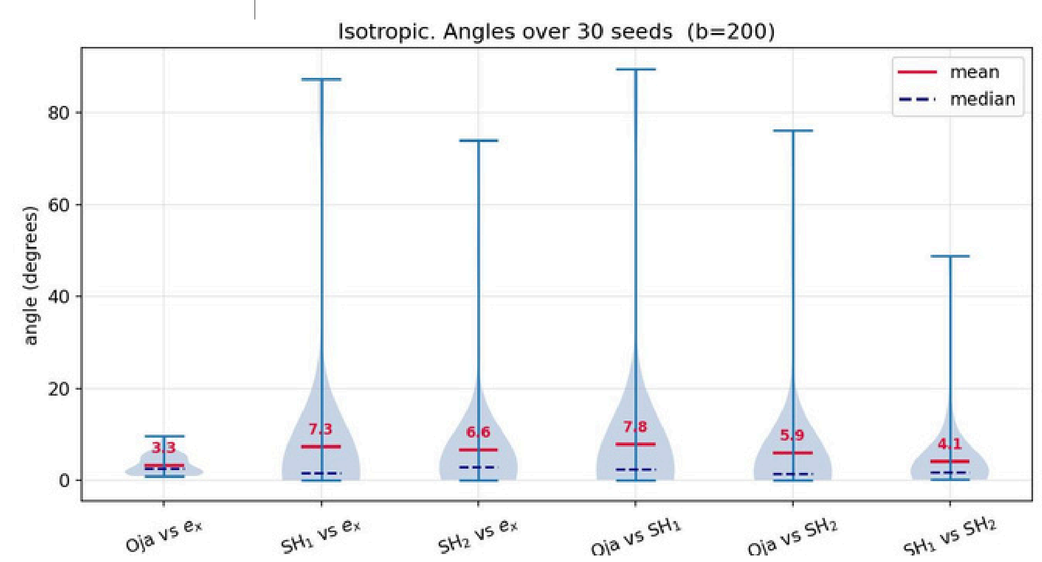
Reduced objective. In hard WTA limit, SoftHebb maximizes $S(\theta) = \mathbb{E}[\,|\cos\theta\, x^* + \sin\theta\, y^*|\,]$ leaving only two candidate splits defined by $\theta \in \{0, \frac{\pi}{2}\}$, as the data has reflection symmetry across both axes.

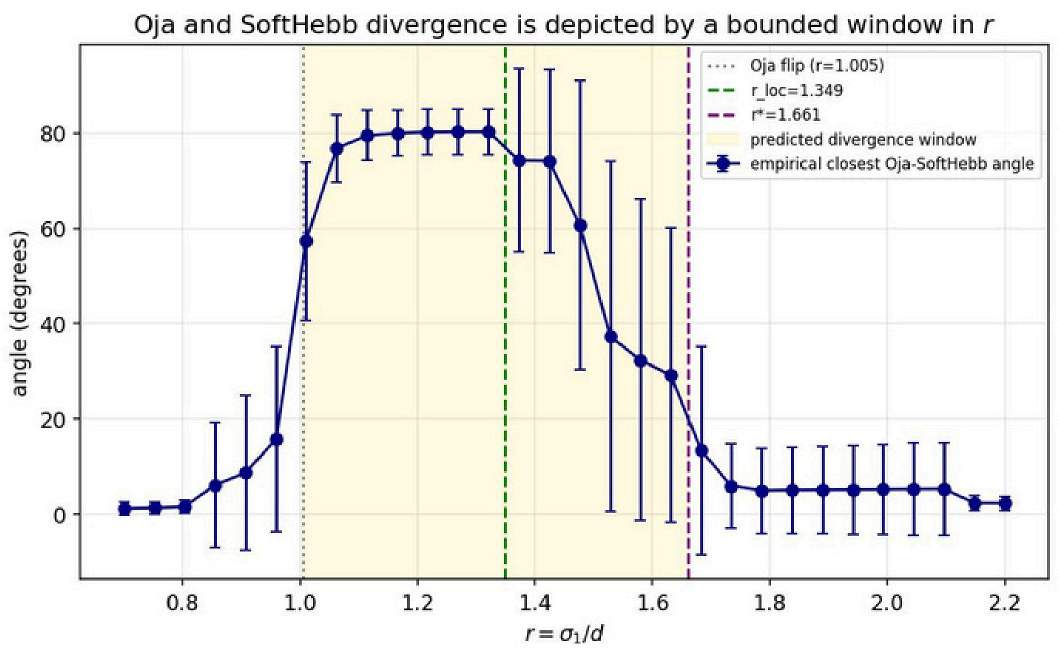
The ratio $r = \frac{\sigma_1}{d}$. The cluster split wins iff $\mathbb{E}[\,|y^*|\,] \geq \mathbb{E}[\,|x^*|\,]$ and the closed form yield a unique crossover at approximately 1.687 on the r-axis.

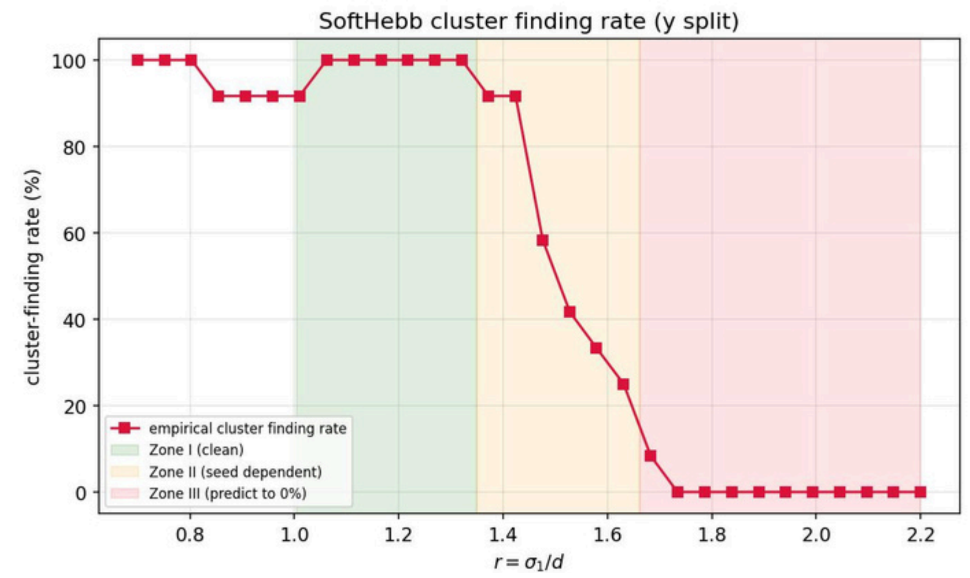
Stability. Curvature of the objective function shows the y split is a local maximum, whilst the x split switches from a valley to a peak at around 1.330.

Three Zones. Zone I (1.005, 1.35) shows clean y split. Zone II (1.35, 1.66) is seed dependent. Zone III (above 1.66) SoftHebb collapses to the x-axis [b=200]. Oja flips at 1.005 because the total covariance is $diag\left(\sigma_1^2, \sigma_2^2 + d^2\right)$.

5. Experimental Results







6. Conclusions

Isotropic clusters. When separation aligns with the variance axis, Oja and SoftHebb converge to the same direction.

Naive right angle divergence is false. SoftHebb acts on unit normalized inputs, so large orthogonal spread collapses the clusters into overlapping arcs and its split flips back towards Oja.

One governing ratio $r = \frac{\sigma_1}{d}$. Agreement at small and large values of r, diverge only inside a bounded interval. Thresholds were derived analytically and matched by the sweep.

Mechanism. The two rules read different geometries of the same data (i.e., euclidean and directional geometry), so divergence is bounded in r.

7. Future work

As our analysis was controlled, it naturally marks the next steps:

1. extending the analysis to **higher dimensions** or more clusters;
2. as **Zone II** is **seed dependent** and zone edges behave as **soft crossovers**, it makes worth describing the transition analytically;
3. extending Oja's to **Sanger's rule** [3] implies a more in-depth PCA and K-means correspondance analysis.

8. References

[1] R. Ye, C. Ye, C. Huang, M. Tang, and Y. Liu, "Beyond-backpropagation training: Methods, applications, and perspectives," TechRxiv, vol. 2026, no. 0103, 2026. [Online]. Available: https://www.techrxiv.org/doi/abs/10.36227/techrxiv.176740426.63642005/v1

[2] E. Oja, "Simplified neuron model as a principal component analyzer," Journal of mathematical biology, vol. 15, no. 3, pp. 267–273, 1982.

[3] B. A. Olshausen, "Linear hebbian learning and pca," 1998.

[4] T. Moraitis, D. Toichkin, A. Journé, Y. Chua, and Q. Guo, "Softhebb: Bayesian inference in unsupervised hebbian soft winner-take-all networks," Neuromorphic computing and engineering, vol. 2, no. 4, p. 044017, 2022.

[5] C. Ding and X. He, "K-means clustering via principal component analysis," in Proceedings of the twenty-first international conference on Machine learning, 2004, p. 29.

[6] L. Ljung, "Analysis of recursive stochastic algorithms," IEEE transactions on automatic control, vol. 22, no. 4, pp. 551–575, 1977.