Ephaptic Dependence Detectors: Theory and Simulations

Aman Chawla

***Abstract*—This paper documents a focused investigation into whether an ephaptic-perceptron-inspired framework can be used as a detector of statistical dependence between two input streams. The project began with an information-theoretic hypothesis grounded in entropy inequalities for mixed random variables and progressed through a sequence of simulation studies in MATLAB. We first examined whether disagreement between entropy-based channel preference and the current ephaptic perceptron implementation could serve as a dependence signal. That hypothesis was evaluated on synthetic data with known ground truth and on real wearable acceleration recordings. We then tested a direct theorem-driven criterion based on the inequality $H(Z) \geq \max\{H(A), H(B)\}$ for a mixed output Z formed from two binary streams. The direct criterion produced no violations on either independent or dependent synthetic data and no violations on real data, demonstrating that it is a one-way consistency check rather than a practical detector. Finally, we generalized the mixing rule from a one-parameter convex form to a two-parameter form $Z = \lambda_1 A + \lambda_2 B$ and evaluated features derived from the entropy gap surface. The minimum gap feature achieved modest but statistically significant above-chance detection on synthetic data, with balanced accuracy 0.584 and permutation p-value 0.0017, while mean and variance features remained at chance. Across all experiments, mutual information remained the strongest and most interpretable dependence indicator, while ephaptic gap features emerged as weak but potentially useful auxiliary cues.**

***Index Terms*—Ephaptic coupling, entropy, dependence detection, mutual information, perceptron, simulation.**

I. Introduction

The central objective of this investigation is practical and specific: to identify a defensible use for the current ephaptic perceptron implementation without rewriting its core logic. The starting point was a set of MATLAB scripts that estimate binary entropies, choose high-entropy channels, and update perceptron-like weights. Early experiments on real accelerometer data suggested that the model often favored the vertical axis z, while entropy ranking frequently favored acceleration magnitude. This mismatch raised a natural question. Could disagreement between entropy-based ranking and model preference itself indicate non-independence among inputs?

The motivating theoretical context is a short but influential argument from ephaptic information selection: in certain coupling models, the output stream may preserve or amplify informational salience associated with high-entropy stimuli. If independence assumptions hold and the output is formed by non-degenerate mixing, entropy inequalities impose strong constraints. If those constraints are violated, assumptions are invalid. The practical challenge is that not every valid inequality is discriminative for detection. A detector must separate independent from dependent cases in finite-sample conditions, not merely satisfy a theorem under idealized assumptions.

Accordingly, this work proceeds in stages. First, we test whether the existing ephaptic perceptron disagreement signal is predictive of dependence. Second, we evaluate a direct theorem-based inequality test in synthetic and real settings. Third, we expand the mixing model from a single parameter to two parameters and test whether richer entropy-gap geometry yields better separation. The result is a grounded summary of what works, what fails, and what remains promising.

II. Theoretical Background

Consider two discrete random variables A and B. Under ordinary arithmetic, define

$$Z = \lambda A + (1 - \lambda)B, \quad 0 < \lambda < 1. \tag{1}$$

For independent A and B, a useful argument relies on conditional entropy and injective transforms. Since Z is an injective affine transform of A when conditioning on $B = b$ with $\lambda \neq 0$, one has

$$H(Z \mid B = b) = H(A), \quad H(Z \mid B) = H(A). \tag{2}$$

Then

$$H(Z) \geq H(Z \mid B) = H(A). \tag{3}$$

By symmetry, $H(Z) \geq H(B)$, yielding

$$H(Z) \geq \max\{H(A), H(B)\}. \tag{4}$$

This implication supports the contrapositive diagnostic statement: if $H(Z) < \max\{H(A), H(B)\}$, then the assumptions used above are broken. However, that statement is one-way. Absence of violation does not establish independence. In finite data, this distinction is crucial.

Another important distinction concerns mixture distributions versus linear combinations of random variables. Entropy concavity guarantees

$$H(\lambda P_A + (1 - \lambda)P_B) \geq \lambda H(P_A) + (1 - \lambda)H(P_B), \tag{5}$$

but this does not imply entropy exceeds the larger component entropy in all cases. Our experiments use explicit variable-level combinations, not only distribution mixtures, and therefore rely on mapping structure and collision behavior in addition to concavity.

Finally, when using real signals, structural coupling can invalidate independence assumptions by construction. For example, if magnitude is computed as $\sqrt{x^2 + y^2 + z^2}$, it necessarily depends on z. Any detector claim must acknowledge such deterministic linkage before interpreting model behavior.

III. Data and Experimental Framework

The project used two data sources. The first source was synthetic binary and continuous streams generated in MATLAB, where dependence structure could be controlled exactly. The second source was real acceleration recordings in CSV format from wearable-like experiments, including files such as `TotalAccelerationfirstbig.csv` and later additions for a total of ten recordings. Each CSV contained at least `x`, `y`, and `z`, and magnitude channels were derived as needed.

Binary streams were created through median thresholding of each real-valued channel. Entropies were computed with a corrected binary entropy implementation that safely handles zero probabilities. Dependence baselines were estimated using absolute Pearson correlation and empirical binary mutual information. These baselines served as ground-truth-oriented comparators for all ephaptic-derived heuristics.

The core scripts built in the investigation subfolder were `run_investigation.m`, `run_theorem_test.m`, `run_lambda_sweep.m`, and `run_lambda12_sweep.m`. Each script writes outputs to CSV and visualizes distributions to ensure that conclusions are tied to both numeric summaries and behavior across trials.

IV. Experiment 1: Mismatch as a Dependence Signal

The first test asked whether disagreement between entropy winner and ephaptic-perceptron winner indicates dependence. Synthetic experiments used two groups with 200 trials each. Independent trials used unrelated generated streams. Dependent trials used linearly coupled streams with additive noise.

The key summary was the mismatch rate

$$\mathbb{P}\left(\hat{c}_{\text{entropy}} \neq \hat{c}_{\text{model}}\right), \tag{6}$$

evaluated separately per group. Results were unambiguous. Independent mismatch rate was 0.455, while dependent mismatch rate was 0.390. If mismatch were a strong dependence signature, the dependent rate should be distinctly larger, which was not observed.

Two heuristic detectors were evaluated. The first classified a sample as dependent when mismatch equals one. This gave true positive rate 0.390, true negative rate 0.545, and balanced accuracy 0.468, which is below chance. The second classified dependence when mismatch equals zero. This gave true positive rate 0.610, true negative rate 0.455, and balanced accuracy 0.532, only marginally above chance. These outcomes indicate that mismatch carries weak signal at best and can invert depending on coding direction.

Meanwhile, baseline dependence metrics behaved as expected. Independent synthetic pairs yielded mean $|corr| = 0.035$ and mean mutual information 0.001 bits, while dependent pairs yielded mean $|corr| = 0.688$ and mean mutual information 0.377 bits. Therefore the synthetic construction was valid, and the weakness lies in the mismatch heuristic rather than in dataset quality.

V. Experiment 2: Direct Theorem Test with One-Parameter Mixing

The second test directly evaluated the inequality $H(Z) \geq \max\{H(A), H(B)\}$ using

$$Z = 0.5A + 0.5B \tag{7}$$

with discrete entropy estimation from empirical PMFs. For synthetic data, 400 trials were generated per group. Violation rates were exactly zero for both independent and dependent groups. On real data, ten files were evaluated for `x_vs_y` and `z_vs_mag` binary pairs. Violation rates were again zero in both cases.

The real-data mean entropy gaps were positive and substantial: approximately 0.3590 for `x_vs_y` and 0.5545 for `z_vs_mag`. Yet mutual information remained nonzero, with means 0.0784 and 0.0368 bits, respectively. Thus non-violation coexists with measurable dependence. This confirms the theoretical warning that the test is one-way. Violation can falsify assumptions, but non-violation does not prove independence.

This result also clarifies why this criterion cannot directly serve as a detector in realistic data analysis. A detector requires differential response across classes, while a consistency inequality may hold broadly across both classes due to mapping structure and finite alphabet behavior.

VI. Experiment 3: One-Parameter Lambda Sweep

The third experiment swept λ over $[0.05, 0.95]$ and examined the gap curve

$$g(\lambda) = H(\lambda A + (1-\lambda)B) - \max\{H(A), H(B)\}. \tag{8}$$

Summary features included minimum, mean, and standard deviation of $g(\lambda)$. Synthetic balanced accuracies were 0.500 for minimum and mean, and 0.518 for standard deviation. These are effectively chance-level outcomes.

The real-data summary exhibited a revealing pattern: mean minimum gap equaled mean gap, and standard deviation was approximately zero across files for both pairs. This indicates the gap curve was essentially flat in this binary setup. The flatness is consistent with near-injective mappings for many λ values, where entropy of transformed symbols changes little except at collision points.

The conclusion from this stage is direct. Varying one convex parameter does not create useful discrimination for dependence in these binary streams.

VII. Experiment 4: Two-Parameter Lambda Surface

To increase flexibility, mixing was generalized to

$$Z = \lambda_1 A + \lambda_2 B, \quad \lambda_1, \lambda_2 \in [0, 1], \tag{9}$$

excluding the degenerate point $(0, 0)$. An 11×11 grid generated a two-dimensional gap surface. Features from this surface were then used for classification.

Synthetic results improved for one feature only. Minimum gap over the λ_1–λ_2 grid achieved balanced accuracy 0.584

at threshold approximately -0.0012. Mean and standard deviation of the gap surface remained at chance with balanced accuracy 0.500.

Because 0.584 is modest, a permutation test was added with 3000 random label permutations. The observed balanced accuracy was 0.584, the empirical null 95% interval was $[0.502, 0.560]$, and the permutation p-value was 0.0017. This establishes that the effect is statistically significant in synthetic conditions, even though practical effect size is limited.

Real-data summaries for ten files showed mean minimum gaps near zero for both pairs, with positive mean gap and nontrivial surface standard deviation. Specifically, `x_vs_y` had mean minimum -0.0001, mean gap 0.7211, and mean standard deviation 0.3613; `z_vs_mag` had mean minimum -0.0000, mean gap 0.7686, and mean standard deviation 0.3632. Mutual information remained 0.0784 and 0.0368 bits, respectively. These values suggest that minimum-gap behavior may carry subtle information but is far weaker and less interpretable than direct dependence metrics.

VIII. Implications for the Current Ephaptic Perceptron

The guiding practical question was not how to redesign the model from scratch, but how to use it as it currently stands. The evidence supports three conclusions.

First, the current perceptron behavior is not a reliable dependence detector by itself. Disagreement with entropy ranking can occur for reasons unrelated to dependence, including model-output mechanics and finite-sample dynamics.

Second, theorem-based entropy inequalities are informative as consistency checks but weak as detectors. They are suitable for falsification in presence of violation, not for confirmation in absence of violation.

Third, ephaptic-style entropy-gap geometry provides a weak but real auxiliary signal when expanded to two-parameter mixing and summarized by minimum gap. This signal becomes statistically credible in synthetic data but remains modest in effect size and should not be used as a standalone decision rule.

Therefore, the most defensible immediate use of the present ephaptic code is as an exploratory channel-prioritization and auxiliary scoring module. It can indicate where informative structure might lie, while dependence decisions are anchored by stronger measures such as mutual information and correlation.

IX. Limitations

This study has several limits that should be stated explicitly. The entropy estimates are empirical and finite-sample, so small numerical differences near zero can affect thresholded features such as minimum gap. Binary thresholding, while convenient and aligned with current scripts, discards amplitude detail and may obscure nonlinear dependencies present in raw continuous data. Real data were limited to ten files from a specific motion context, and conclusions may shift with broader tasks or sensors.

The current perceptron implementations also mix modeling assumptions. Some outputs are vector-valued signs rather than single scalar decisions, and some update rules embed implicit biases that complicate interpretation as pure entropy selectors. These properties are acceptable for exploratory analysis but limit strict theoretical mapping.

X. Conclusion

This investigation mapped a clear trajectory from theoretical intuition to empirical testing. The original hope that entropy-perceptron mismatch could directly indicate dependence was not supported in strong form. Direct inequality checks confirmed theoretical consistency but lacked discriminative power. Lambda sweeps in one parameter were effectively uninformative in the binary setting. A two-parameter mixing surface introduced a meaningful improvement: minimum entropy gap produced modest but statistically significant above-chance separation on synthetic data, with permutation evidence against random fluctuation.

The net outcome is pragmatic. Ephaptic-inspired entropy constructs can contribute useful auxiliary structure, especially through extremal gap features on richer mixing families. However, robust dependence detection still relies on primary metrics such as mutual information and correlation, with ephaptic features serving as secondary signals. This framing provides a concrete and realistic use for the present ephaptic perceptron codebase: exploratory dependence-aware channel scoring rather than standalone dependence inference.

Acknowledgment

The simulations and scripts referenced in this paper were developed in the `01March2026Investigations` workflow and evaluated on both synthetic streams and real acceleration recordings in the local AGI dataset.

References

[1] A. Chawla and S. D. Morgera, “Ephaptic synchronization as a mechanism for selective amplification of stimuli,” *BMC Neuroscience*, vol. 15, suppl. 1, P87, 2014.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.