

ENHANCING ARTIFICIAL NEURAL NETWORKS WITH EPHAPTIC COUPLING

Ismaila “izzo” Wane
June 30, 2025 (Rev. 0.9.6)

ABSTRACT

We propose a novel, general activation equation that extends the classical artificial neuron by incorporating structured intra-layer modulation inspired by ephaptic coupling, a non-synaptic, parallel, biological communication mechanism overlooked by the AI community. The result is a general-purpose mechanism applicable to any artificial neural network (ANN), enhancing its expressivity, efficiency, robustness, convergence, and adaptability across domains. Preliminary testing shows strong cross-domain gains compared to standard weight-only baselines: a 74.65% reduction in perplexity in language (GPT-2, WikiText-103), a 3.08% accuracy gain in vision (ResNet-18, ImageNet-100), and a 6.67% reward improvement in reinforcement learning (PPO, Walker2d-v5). These performance improvements were achieved with minimal parameter overhead using pretrained models: $\sim 0.5\%$ for GPT-2 small, $\sim 2.2\%$ for ResNet-18, and $\sim 0.98\%$ for the PPO network. Initial ablation studies suggest that ephaptic coupling provides a highly efficient means of architectural enhancement compared to conventional width expansion.

1. INTRODUCTION

The concept of the neuron began as early as 1740 with Swedenborg’s idea of individual brain units contributing to cognition [1]. The modern neuron doctrine was demonstrated by Cajal’s observations in birds in 1888 [2], and later formalized by Waldeyer-Hartz who coined the term “neuron” in 1891 [3]. It was further theorized by Freud in 1895 [4] before Cajal confirmed it on humans in 1899 [5]. Cajal used Golgi’s silver staining method to show that neurons are discrete cells [6]. According to studies in 2009 by Azevedo *et al.* [7] and contextualized by Herculano-Houzel [8], the human brain contains about 86 billion neurons, most of which are in the cerebellum, and with about 16 billion in the cerebral cortex, primarily responsible for perception and consciousness. Synapses are the primary communication points between neurons, allowing signals to pass via neurotransmitters. Coined by Sherrington in 1897 [9], the term "synapse" comes from the Greek *synaptein*, meaning "to clasp together". Each neuron forms approximately 1,000 to 10,000 synapses, contributing to an estimated 100 to 1,000 trillion synapses in the brain. The number of synapses changes with learning and development, peaking in childhood before pruning optimizes connections [10], [11].

Advancements in AI have been driven by interdisciplinary contributions spanning biology, mathematics, physics, computer science, engineering, and other fields. The concept of the artificial neuron was first theorized by McCulloch and Pitts in 1943 [12], and later implemented in practice by Rosenblatt with the invention of the perceptron in 1957, followed by the formal publication in 1958 of his foundational report [13], [14]. The perceptron referred both to a single artificial neuron and to a neural network architecture based on it. While Rosenblatt’s original perceptron was a single-layer model, he later described multi-layer perceptron (MLP) architectures in *Principles of Neurodynamics* in 1962, anticipating the structure of modern ANNs and machine learning by backpropagating errors [15]. However, training such deeper models remained impractical until the rediscovery of backpropagation decades later by Linnainmaa [16], then Werbos [17], [18], and finally Rumelhart, Hinton, and Williams [19]. Today’s research in ANNs stands on the foundational work conducted between the 1920’s and 1960’s by pioneers such as Loewi [20], Lashley [21], Dale [22], [23], [24], Turing [25], Shannon [26], [27], [28], Hebb [29], Rashevsky [30], [31], Householder [32], [33], [34], [35], [36], [37], Pitts [38], [39], [40], and Von Neumann [41]. Even Minsky, who arguably contributed to the first “AI Winter” through his critiques of Rosenblatt’s perceptron [42], wrote his PhD dissertation on neural networks in 1954 [43], two years before Rosenblatt completed his own dissertation in 1956 [44].

With the resurgence of the connectionist movement, driven by Fukushima’s Neocognitron in 1980 [45], preceding Convolutional Neural Networks (CNNs) between 1989 and 1998 by LeCun *et al.* [46] and transformers in 2017 by Vaswani *et al.* [47], AI researchers are once again turning to biology to inspire the next generation of neural models. Yet, state-of-the-art research remains focused on high-level synaptic

modeling, overlooking a parallel but lesser-known low-level biological phenomenon that is being proven by neuroscientists as critical to neural computation. This phenomenon is known as ephaptic coupling.

Ephaptic coupling refers to non-synaptic electrical interaction between neurons mediated by local electric fields. It occurs at ephapses, which are regions of close physical proximity where adjacent neurons influence each other without synaptic transmission. Unlike synapses, which transmit signals chemically or via gap junctions, ephapses enable passive, field-based modulation through the shared extracellular environment. Based on conjectures by Adrian [48], Arvanitaki first observed this natural phenomenon in 1942 [49], [50] in adjacent nerve fibers of invertebrates. It is only decades later, in 1979 by Seltzer [51] and in 2011 by Anastassiou *et al.* that ephaptic coupling was experimentally confirmed respectively in peripheral nerves and in mammalian cortex. Anastassiou *et al.* further made the case that the neuroscientific community should focus more on investigating ephaptic mechanisms not just for curiosity, but for practical implications in neurobiology and neural prosthetics in 2015. Anastassiou *et al.* [52], [53], [54] suggested that ephaptic effects, though small, can significantly impact membrane excitability, neural synchronization, and even higher-level brain function. Simply put, ephaptic coupling enables biological neurons to influence one another via extracellular electric fields, independent of synaptic transmission. Recent neuroscientific research between 2022 and 2024 from Cunha *et al.* [55], [56] and from Pinotsis *et al.* [57], [58], [59], [60], [61], further echoed in recent presentations by Rebbin [62] and Douglas [63], suggests that ephaptic coupling plays a significant role in synchronizing biological neural networks. Other studies within the neuroscientific community have also explored the natural mechanism of biological neuronal modulation and synchronization [64]. This allows us to hypothesize that, if and when applied to the field of AI, which remains primarily synaptic in structure, ephaptic coupling may enable ANNs to simulate field-mediated interactions, forming a biologically grounded extension of the artificial neuron model. Ephaptic coupling could effectively transform a classical ANN layer into a multidimensional field that goes beyond a small-world topology of fully interconnected neurons. We further postulate that ephaptic coupling, as a general-purpose mechanism, could effectively accelerate learning across domains when integrated into standard ANNs. We refer to an ANN built on this concept as an Ephaptically Coupled Artificial Neural Network (EC-ANN).

It is worth noting that there have been works attempting to implement small world topologies [65] or even intra-layer links within ANNs. For instance, Fan *et al.* [66], [67] demonstrated that adding intra-layer connections can enhance the expressivity of shallow networks across multiple domains, including synthetic functions, tabular datasets, and image classification, thereby empirically validating their theoretical claims on depth-width trade-offs. However, Fan *et al.* do not incorporate any biological context. They neither model field effects, nor reference ephaptic coupling. Their work remains purely topological and functional, rather than biophysical or neurodynamic, and therefore overlooks key insights related to extra-dimensional intra-layer communication observed in biological systems. Although their work was recognized during peer review as mathematically rigorous and directionally interesting, it was ultimately rejected by reviewers due to limited theoretical separation, unclear practical relevance, and disagreement on architectural novelty. Fan *et al.* also brought up a related work where they proposed a 3D network using intra-layer connections to enhance expressivity, but it is purely architectural and did not model neuronal modulation or biologically inspired field-based interactions. Other works have explored novel neural mechanisms, such as Mercioni *et al.* [68] with dynamic activation tuning, and Lillicrap *et al.* [69], Ravichandran *et al.* [70] with synaptic plasticity models and Chen *et al.* [71] with a fully spiking actor network with intra-layer connections for reinforcement learning. Recent state-of-the-art post-transformer models include Continuous Thought Machines (CTM), developed by Darlow *et al.* [72], which introduce a synchronization matrix to promote coherent temporal reasoning. However, they do not simulate ephaptic or field-based interactions. In contrast, EC-ANNs explicitly model intra-layer coordination through biologically inspired ephaptic coupling, absent in prior ANN architectures. Whereas CTM’s ephemeral synchronization matrix externally enforces coherence across time, EC-ANNs derive coordination intrinsically through a permanent modulation matrix that emerges from local neuronal activity and evolves dynamically with the network even after training and deployment. Crucially, CTM is a new model architecture while ephaptic coupling is a cross-domain mechanism that can be integrated into existing architectures without structural changes.

In this work, we provide the biological background and formally define the ephaptic coupling mechanism from topological, computational, and mathematical perspectives. We then empirically demonstrate its ability to improve generalization and expressivity across language, vision, and reinforcement learning tasks. Finally, we draw conclusions based on our preliminary results and discuss the potential implications of this novel mechanism for the broader AI community.

2. BIOLOGICAL BACKGROUND

The artificial neuron is the elementary unit of the ANN and represents a computational model of the biological neuron. The computational model of an artificial neuron consists of input, activation and output.

The dendrites are modeled as input, the synapse with other neurons in neighboring layers as weights. The cell nucleus is modeled as a node which receives the sum of all input weights and passes it through a nonlinear function to produce an output which models the axon.

Every neuron holds an internal state called activation signal. Each synapse carries information about the input signal. Every neuron is connected to another neuron via a synapse. Biological neurons form various synapses categorized by anatomical connections and signaling mechanisms. Primary anatomical types include axodendritic (i.e., axon-to-dendrite), typically excitatory and most common; axosomatic (i.e., axon-to-cell body), usually inhibitory, regulating excitability; axoaxonic (i.e., axon-to-axon), modulating neurotransmitter release; and less common forms such as dendrodendritic (i.e., dendrite-to-dendrite), involved in reciprocal signaling, and somatosomatic or somatodendritic, mainly inhibitory. Based on signaling, synapses can be chemical, using neurotransmitters for unidirectional signaling and electrical (i.e., gap junctions), enabling rapid, bidirectional ion flow.

The traditional computation model of ANNs, which has been centered around synaptic transmission (i.e., via synapses), does not consider another parallel communication mechanism that has been often overlooked by neuroscientists and ignored by AI/ML researchers. This parallel site of communication is called the ephapse as illustrated in Figure 1.

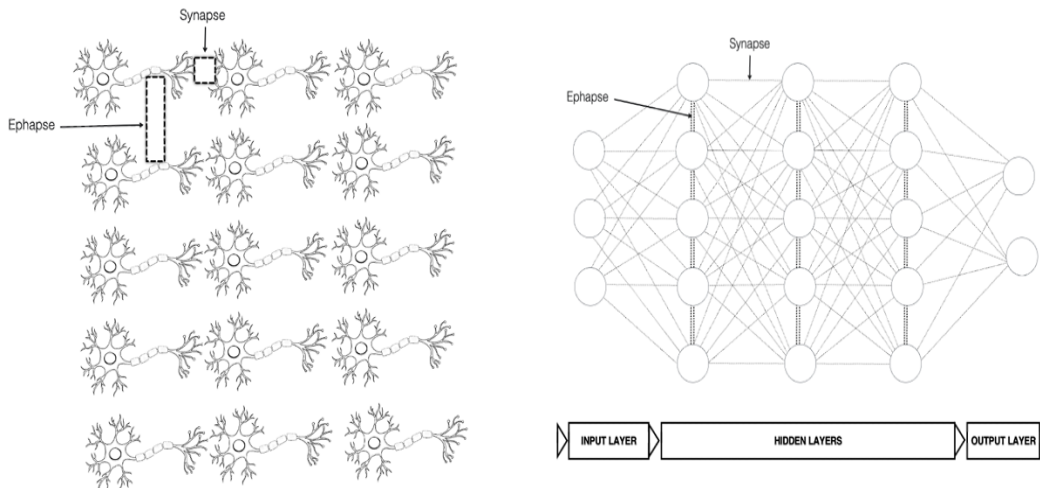


Figure 1: Synapse and Ephapse in Biological & Artificial Neural Networks

Coined in 1941 by French neurophysiologist Arvanitaki [49], [50], the word ephapse is from the Greek ephapsis meaning "to touch". It is defined as a site of functional interaction between two or more neurons where communication occurs not through traditional synaptic transmission, but through local electric fields that allow one neuron's activity to influence the membrane potential of another. This biological phenomenon is known as ephaptic coupling.

Arvanitaki, one of the first women to publish foundational electrophysiological research, made pioneering contributions to non-synaptic communication, particularly in invertebrate ganglia. Her experiments demonstrated that adjacent nerve fibers could interact without synapses, through field effects, what we now call ephaptic coupling, and action potentials in one neuron could influence the excitability of neighboring neurons simply by electric proximity in a shared extracellular space. Her work laid an early experimental

basis for exploring electrical cross-talk, a phenomenon now gaining renewed interest in both neuroscience and neuromorphic computing.

It is speculated that the dominance of Hodgkin–Huxley’s ion-channel model in the 1950s [73] shifted attention away from these kinds of non-synaptic mechanisms for decades, possibly unknown or overlooked by Rosenblatt when formulating his neurodynamics theory in the 1950’s. The attention only came back about this “spooky” phenomenon until modern tools made it easier to detect subtle, local field interactions. The Hodgkin–Huxley model provides a quantitative framework for understanding how action potentials are generated and propagated in neurons. It models the neuron's membrane as an electrical circuit, incorporating voltage-gated sodium and potassium ion channels and leak currents, with membrane capacitance and conductance governed by a set of nonlinear differential equations. Their experiments on the squid giant axon enabled precise voltage clamp measurements, revealing how ion channel dynamics produce the characteristic spike of an action potential.

The exact number of ephapses in the human brain is not precisely known, as ephaptic coupling is still a relatively underexplored aspect of neural communication. However, based on estimates of neuron density, extracellular field effects, and observed ephaptic interactions, it is suggested that ephaptic interactions may be as numerous as, or even exceed, synaptic connections in certain brain regions. Ephaptic interactions do not require direct physical connections, meaning that any neuron within an extracellular field can potentially influence nearby neurons. If even 10% of neurons engage in ephaptic interactions, this could result in at least 10 trillion ephaptic connections. In densely packed regions such as the cerebellum, hippocampus, and neocortex, where neurons are closely aligned, ephaptic interactions could be as frequent as synapses or even outnumber them. Other researchers hypothesize that ephaptic effects may be a pervasive, secondary layer of communication that dynamically influences a much larger network of neurons than traditional synaptic connections alone. While no definitive count exists, ephaptic interactions could rival or exceed synaptic connections in certain neural circuits, potentially making them an important but underappreciated mechanism in brain function.

As a subtle form of interaction between biological neurons, ephaptic coupling has been historically misunderstood and often dismissed as negligible and unimpactful. Ephaptic coupling has been even described as a “spooky” phenomenon due to the difficulty of measuring its effects. However, later experimental and computational studies, most notably by Anastassiou *et al.* in 2011 and further discussed in a 2015 review, provided compelling evidence for its functional relevance. Ephaptic coupling is now understood as a non-synaptic form of neuronal communication, in which neurons influence the excitability of neighboring cells through localized extracellular electric fields rather than relying solely on chemical synapses. These interactions (i.e., ephapses) arise from the close spatial proximity of neuronal membranes, allowing the electric field generated by one neuron's activity to modulate the membrane potential of adjacent neurons.

The so-called spookiness of ephaptic coupled has been discussed by Chen in 2020 [74] referring to ephaptic coupling with a clear nod to quantum mechanics, titling a thought-provoking article: “*Spooky Action Potentials at a Distance: Ephaptic Coupling*”, while examining what he described as an even less-discussed, third form of neural communication (i.e., ephaptic coupling). Chen’s article summarized findings in a research paper he co-authored titled: “*Climbing fiber synapses rapidly and transiently inhibit neighboring Purkinje cells via ephaptic coupling*” [75]. It is easy to draw a conceptual analogy between ephaptic coupling, where neural units influence one another non-synaptically via local field potentials, and quantum entanglement, which is traditionally described as “spooky action at a distance”.

Both phenomena challenge classical locality, and Elitzur *et al.* [76] supports the idea that entanglement might emerge from underlying self-organizing dynamics, not requiring instantaneous communication per se. Elitzur *et al.* argue that quantum entanglement need not be treated as a mysterious, nonlocal phenomenon. Instead, they propose it can emerge from self-organizing dynamics in interacting systems, just as synchronized behavior arises in complex networks such as neural tissue. Just as ephaptic coupling arises from spatial and temporal correlations in a network without direct synaptic links, entanglement may reflect emergent correlations from self-organizing dynamics that span a hidden layer of the physical system.

Ephaptic coupling involves ionized atoms, specifically ions such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), and chloride (Cl^-), whose movements across neuronal membranes generate changes in ionic

concentrations and subsequently create localized electrical field. These ionic currents produce extracellular field potentials that alter the membrane potential of neighboring neurons, modulating their excitability and the probability of firing action potentials without direct chemical neurotransmitter exchange. Myelination, which is the biological process by which a fatty substance called myelin forms an insulating sheath around neuron axons, enables faster and more efficient transmission of action potentials much like insulation on electrical wires. Myelination is believed to reduce or prevent these ephaptic effects, limiting their role in certain neural circuits.

Most recent studies and papers from Cunha *et al.* in 2022 and 2023 and Pinotsis *et al.* in 2024 demonstrated that ephaptic coupling may significantly influence neuronal firing timing and synchronization. They reported the role of ephaptic coupling in neural network complexity using the Quadratic Integrate-and-Fire Ephaptic (QIF-E) hybrid model, and demonstrated that ephaptic interactions enhance complexity, especially in small-world network topologies. Using Multiscale Entropy (MSE) analysis, which is a method for quantifying the complexity of time-series data by measuring entropy across multiple temporal scale, these studies showed biological neural networks with active ephaptic coupling have higher entropy, reflecting more intricate neuronal activity. The research highlighted an inverse relationship between synaptic strength and network complexity, emphasizing ephaptic interactions' modulatory role. Simply put, the neuroscientific community's state-of-the-art research underscores ephaptic coupling as a significant form of neural communication, particularly in brain regions such as the hippocampus and neocortex.

It is worth noting that current biological models and tools for simulating ephaptic coupling are typically based on cable theory that models how electrical signals propagate along neurons, particularly dendrites and axons, which are treated as cylindrical cables with resistive and capacitive properties. It is a cornerstone of biophysical neuron modeling, especially for understanding local voltage spread and ephaptic interactions. These ephaptic coupling simulation tools used by the neuroscientific community include QIF-E and ELFENN [77]. QIF-E extends the Quadratic Integrate-and-Fire neuron model to simulate how electric fields affect neuronal dynamics and has been validated against empirical data. ELFENN, which stands for Electric Field Effects in Neural Networks, is an open-source MATLAB toolbox that models extracellular field effects in spiking neuron networks. Both tools help study non-synaptic, field-based interactions in biological systems. However, they are not designed for AI and lack a learning framework for ANN integration. They also rely on linear approximations, whereas nonlinearity is more tailored for ANN use.

Despite its biological significance and potential relevance to AI systems, and important resources poured into AI research endeavors, ephaptic coupling has yet to be canonically modeled by the AI community at large, and by deep learning researchers in particular. Rarely do AI researchers question the structural assumptions of the artificial neuron itself, which reflects a broader lack of interdisciplinarity in the field, particularly with respect to the history and close relationship of connectionism and neuroscience.

This non-triviality may also stem from the fact that ANNs are typically represented as matrices and rely on solving linear equations to address generalization problems. Conventional wisdom in the AI community emphasizes row operations, which correspond to synaptic computation, while largely overlooking column operations, which could model ephaptic effects.

3. CONCEPTUAL MODELING

Adapted to AI, the biological background allows us to conceptualize an ephaptic field represented as a matrix Λ that is trainable, learnable and with its coefficients potentially adaptable post-training. This matrix Λ , is different from the weight matrix W , and consists of coefficients Λ_{ik} representing ephapses.

Myelination, as the biological inhibitor of ephaptic interactions, could be modeled as the inverse of an ephaptic factor, noted as ϵ . High myelination (i.e., low ephaptic factor) reduces ephaptic interactions by insulating neurons, whereas a higher ephaptic factor implies greater field-based influence between uninsulated or closely packed neurons in unmyelinated regions (e.g., cerebellum). This implies that higher ephaptic factor leads to dense interconnectivity while a lower ephaptic factor leads to structured, isolated pathways (e.g., ResNets).

An initialization magnitude λ_0 is used as the starting scale of values in the matrix Λ . It controls the numerical strength of ephaptic interactions at initialization and is applied as a multiplicative factor over the masked or sparse Λ structure. Biologically, λ_0 represents the initial strength of the ephaptic field (Λ) across neurons within a local circuit. While synaptic weights model chemical transmission, the matrix Λ , scaled at initialization by λ_0 , captures the field-driven modulation seen in dense neural populations, where local extracellular potentials can alter the excitability of nearby neurons even in the absence of direct synaptic contact.

Moreover, in biological neurons, responses to input are inherently nonlinear i.e., they don't increase proportionally with stimulus. This is modeled in ANNs with artificial neurons firing only when input crosses a threshold (e.g., ReLU), saturate at high input levels (e.g., tanh or sigmoid), or show complex spiking behaviors. In EC-ANNs, a dedicated local computation function (Φ) can be introduced for each ephaptic coefficient (i.e., ephapse) to capture this nonlinear transformation, shaping neural activations on W before they influence the ephaptic field via the ephaptic coupling matrix Λ .

We therefore model ephaptic coupling as a machine learning system illustrated in Figure 2 below.

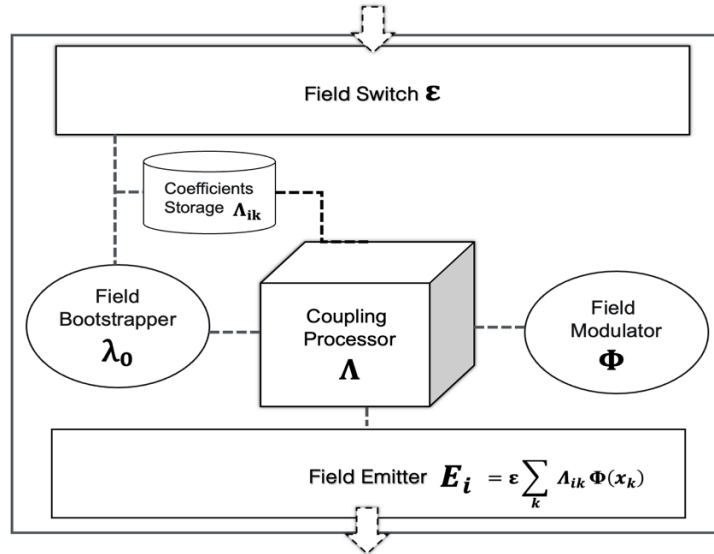


Figure 2: Ephaptic Coupling System

The proposed system enables the derivation of an ephaptic term, which can be combined with the classical synaptic term to activate an artificial neuron and, consequently, an artificial neural network enhanced with ephaptic coupling.

4. TOPOLOGICAL ORGANIZATION

The structural layout of EC-ANNs can be introduced using 2D matrix representations, illustrating how they differ topologically from standard ANNs. While the formal mathematical definitions of ephaptic coupling will follow in a later section, we focus first on the spatial organization and flow of signals within a layer. In a standard feedforward ANN, the output of each hidden neuron is computed using a weight matrix W and bias vector b . The standard ANN model also typically processes activations row-wise (neuron-wise). In EC-ANNs, we define a second matrix Λ , which modulates activations column-wise, enabling intra-layer interactions. In other words, W governs the feedforward signal propagation while Λ models field-based modulation within the same layer. Together, these matrices create a hidden layer with dual axes of communication with both synaptic, row-wise communication whereby a neuron receives input from the previous layer, and ephaptic, column-wise communication whereby a neuron receives modulation from peers. The first matrix W encodes the traditional synaptic weights connecting inputs to hidden neurons. Each row corresponds to one hidden neuron receiving inputs from all previous-layer neurons. The second matrix Λ captures internal interactions among hidden neurons. Each off-diagonal entry Λ_{ik} represents the influence of neuron k on neuron i within the same layer. Self-modulation, where a neuron influences itself through Λ is optionally omitted by setting diagonal entries to zero.

To illustrate the topological organization mathematically, let us define:

- $x = [x_1, x_2]^T$

Where: input vector $x \in \mathbb{R}^2$

- $W = [[W_{11}, W_{12}], [W_{21}, W_{22}], [W_{31}, W_{32}]]$

Where: weight matrix $W \in \mathbb{R}^{3 \times 2}$ as synaptic weights (row-wise interaction)

- $\Lambda = [[0, \Lambda_{12}, \Lambda_{13}], [\Lambda_{21}, 0, \Lambda_{23}], [\Lambda_{31}, \Lambda_{32}, 0]]$

Where: Ephaptic Coupling Matrix (ECM) $\Lambda \in \mathbb{R}^{3 \times 3}$ as field modulation within the hidden layer. Λ is square by design to reflect peer-to-peer modulation within the same layer, which also conveniently makes norms such as $\|\Lambda\|$ both mathematically sound and biologically meaningful.

- $\Phi(x_k) = [\Phi(x_{k1}), \Phi(x_{k2}), \Phi(x_{k3})]^T$

Where: $\Phi(x_k) \in \mathbb{R}^m$ as nonlinearly transformed activations vector

Then: $W \cdot x$ produces the raw pre-activation vector S (synaptic term) and $\Lambda \cdot \Phi(x_k)$ produces the ephaptic coupling vector E (ephaptic term). Combined and passed through an activation f , S and E via W (and its weights) and Λ (and its coefficients) produce y_i as the activated output of neuron i .

It is important to note that while Λ appears similar in size to W , this illustration reflects topological organization, not actual parameter volume. This may seem counterintuitive at first, because Λ is square and dense, yet in practice it is significantly smaller in terms of parameter count. The reason lies in the underlying dimensionality: W maps from the previous layer to the current hidden layer and is typically shaped $[H \times I]$, where H is the number of hidden units and I is the input dimension.

In contrast, Λ is defined as a square matrix $[H \times H]$, operating within the hidden layer itself. For networks where $I \gg 1$, W overwhelmingly dominates in size. For instance, in GPT-2 small, each MLP layer contains over 4.7 million synaptic weights. By comparison, the Λ matrix adds only 589,824 parameters, a fixed overhead of less than 0.5% of the model's total parameter count of 124 million. GPT-2 small consists of 12 Transformer layers, each processing hidden representations of size 768. These layers include multi-head self-attention and a feedforward MLP subcomponent. In our EC-ANN implementation, we integrate a Λ matrix into the MLP module of the first layer only, enabling ephaptic modulation with minimal architectural disruption. This design enables intra-layer field effects early in the model while preserving downstream compatibility and parameter efficiency. As we will empirically demonstrate in later sections, even this small addition can have a significant impact on performance.

5. GENERAL ACTIVATION EQUATION, VARIANTS AND LEARNING

The general activation equation for an EC-ANN, describing the activation of neuron i , influenced by neuron j in a preceding layer (synaptic input), and neurons k in the same layer (ephaptic coupling), consists of two components: a synaptic term S_i and an ephaptic coupling term E_i . This is expressed generally as:

$$y_i = f(S_i, E_i)$$

Where:

- y_i is the output activation of neuron i
- f is the global non-linear activation function (e.g., Identity, SiLU, ReLU, tanh)
- S_i is the synaptic input to neuron i
- E_i is the ephaptic coupling on neuron i

The synaptic term S_i is defined by:

$$S_i = \sum_j W_{ij} x_j + b_i$$

Where:

- S_i is the synaptic term for neuron i
- W_{ij} is the synaptic weight from neuron j to neuron i
- x_j is the post-synaptic activation of neuron j from the previous layer
- b_i is the bias term for neuron i

The ephaptic coupling term E_i is defined by:

$$E_i = \varepsilon \sum_k \Lambda_{ik} \Phi(x_k)$$

Where:

- E_i is the ephaptic coupling term acting on neuron i
- ε is the ephaptic factor scalar controlling the strength of ephaptic modulation
- Λ_{ik} is the ephaptic coupling coefficient from neuron k to neuron i
- Φ is the nonlinear transformation (e.g., Identity, SiLU, ReLU, tanh) of x_k
- x_k is the post-synaptic, pre-ephaptic activation of neuron k within the same layer as neuron i

This ephaptic term enables neuron i to be modulated by field-like influence from neurons k in the same layer, rather than relying solely on feedforward input from neuron j . In other words, each coefficient $\Lambda_{ik} \Phi(x_k)$ defines an ephapse from neuron k to neuron i , and the full sum over k represents the cumulative ephaptic influence. This introduces intra-layer neuron-to-neuron interactions within EC-ANNs. We clarify that in matrix notation, rows represent traditional feedforward (synaptic) interactions across layers, while columns intuitively represent intra-layer interactions.

We define the Ephaptic Coupling Matrix (ECM) Λ as a square matrix of dimension $h \times h$, where h denotes the number of hidden neurons subject to ephaptic modulation. Each coefficient Λ_{ik} represents a directional, graded ephaptic coupling from neuron k to neuron i . The matrix Λ governs intra-layer modulation and operates orthogonally to the standard synaptic weight matrix W .

Λ is typically initialized using a behavior-aligned projection from the output weight matrix, scaled by an initialization magnitude $\lambda_0 \in \mathbb{R}^+$. When dimensions permit, we initialize Λ using the transpose of the output weight matrix: If $W \in \mathbb{R}^{m_{out} \times h}$, then $W^T \in \mathbb{R}^{h \times m_{out}}$

When $m_{out} = h$, Λ and W^T share the same square shape, allowing a direct initialization: $\Lambda \leftarrow \lambda_0 \times W^T$. In cases where dimensions do not match (i.e., $m_{out} \neq h$), we apply a dimensionality resolution strategy such as projecting W^T into square form using top- k selection for sparse, behaviorally aligned structure,

extracting a square submatrix when possible, or using W^T as a guidance map to create a binary mask while initializing Λ independently. These approaches ensure Λ maintains its functional role as a neuron-to-neuron modulation matrix over the hidden layer.

Two primary modulation variants of the general activation function are presented hereafter, based on how the ephaptic modulation term E_i , computed from the ECM Λ , interacts with the synaptic term S_i . These include an additive variant, in which E_i is added to the synaptic input as a superimposed field potential; and a multiplicative variant, in which E_i modulates the synaptic input via a bounded gating function $g(E_i)$. These two variants represent the canonical embodiments of ephaptic coupling within EC-ANNs.

5.1 Additive Variant

$$y_i = f(S_i + E_i)$$

Where the ephaptic term E_i is added directly to the synaptic input S_i , modeling ephaptic coupling effects as a superimposed signal. The additive variant of the general activation function is as follows:

$$y_i = f\left(\sum_j W_{ij} x_j + b_i + \varepsilon \sum_k \Lambda_{ik} \Phi(x_k)\right)$$

5.2 Multiplicative Variant

$$y_i = f(S_i \cdot g(E_i))$$

Where $g(\cdot)$ is a nonlinear, gating function that modulates the synaptic input S_i based on the ephaptic term E_i . Common choices include $g(E_i) = \sigma(E_i)$ for soft gating in $(0, 1)$, $g(E_i) = 1 + \tanh(E_i)$ for symmetric boost/damp, $g(E_i) = \text{softplus}(E_i)$ for smooth positive scaling, and $g(E_i) = 1$ (i.e., Identity function) to disable modulation. The multiplicative variant of the general activation function is as follows:

$$y_i = f\left(\sum_j W_{ij} x_j + b_i \cdot g\left(\varepsilon \sum_k \Lambda_{ik} \Phi(x_k)\right)\right)$$

5.3 Other Variants

As Λ is the essential component of the ephaptic term, it is noted that alternative formulations, including gated additive, recurrent ephaptic modulation, or hybrid interaction structures are plausible. Variants remain within the scope of the work so long as they apply Λ to intra-layer neuronal interactions to influence activation dynamics and conform to the related mathematical theorems.

5.4 Learning

For ephaptic learning, the gradient updates each coefficient (Λ_{ik}) of the ECM Λ using a sparsity mask α . A selective update mechanism is implemented as follows:

$$\Delta\Lambda_{ik} = -\alpha_{ik} \cdot \eta \cdot \frac{\partial L}{\partial \Lambda_{ik}}$$

Where:

- $\Delta\Lambda_{ik}$ is the ephaptic influence update, learning how ephaptic signals affect activations.
- α_{ik} indicates whether ephaptic coupling from neuron k to neuron i is enabled.
- η is the learning rate, controlling adaptation speed.
- L is the loss function, guiding training optimization.
- $\frac{\partial L}{\partial \Lambda_{ik}}$ denotes the gradient of the loss function L with respect to the coefficient Λ_{ik} .

6. MATHEMATICAL THEOREMS

The purpose of our mathematical theorems is to formally establish the foundational properties of EC-ANNs, demonstrating that this novel architecture is not only biologically inspired, but also mathematically rigorous, theoretically grounded, and broadly applicable.

These theorems collectively prove that EC-ANNs are complete, efficient, differentiable, robust, and stable. They position EC-ANNs as a principled extension of standard ANNs that are capable of supporting next-generation AI systems through field-based modulation and integrated, non-modular computation.

We formally define and structure five core theorems of EC-ANNs as summarized below in Table 1:

#	Name	Purpose
6.1	Universal Approximation	EC-ANNs are supersets of ANNs and remain universal approximators.
6.2	Approximation Efficiency	EC-ANNs can match ANN performance with fewer total neurons.
6.3	Parameter Differentiability	EC-ANNs support gradient-based training via backpropagation.
6.4	Input Robustness	EC-ANNs are Lipschitz-robust to input perturbation when Λ is bounded.
6.5	Adaptive Convergence	EC-ANNs with slow Λ updates track and converge better to evolving targets.

Table 1: Summary of Mathematical Theorems

6.1 Universal Approximation Theorem

Let ε the ephaptic factor scalar be set to zero (i.e., $\varepsilon = 0$) which yields $E_i = 0$ in the case of the additive variant and $g(E_i) = 1$ in the case of the multiplicative variant.

Then:

- Additive variant: $y_i = f(S_i + E_i) \Rightarrow y_i = f(S_i)$
- Multiplicative variant: $y_i = f(S_i \cdot g(E_i)) \Rightarrow y_i = f(S_i)$

An EC-ANN can be thus reduced to a classical feedforward ANN. By the classical Universal Approximation Theorem as defined by Hornik *et al.* [78], a standard feedforward neural network can approximate any continuous function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on a compact subset of \mathbb{R}^n arbitrarily well.

This implies that there exists a neural network function \tilde{f} such that:

$$|f(x) - \tilde{f}(x)| < \varepsilon \text{ for all } x \text{ in the domain.}$$

Therefore, an EC-ANN, as a superset of an ANN, inherits this universal approximation property when ephaptic coupling is disabled. When ephaptic coupling is active ($\varepsilon > 0$), it is deduced that EC-ANNs may surpass classical ANNs by introducing structured intra-layer field dynamics via the ECM Λ not captured by the weight matrix W alone.

6.2 Approximation Efficiency Theorem

Let $\delta > 0$ be an arbitrarily small positive number, and let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous real-valued function defined on a compact domain $D \subseteq \mathbb{R}^n$.

Assume that for some $m \in \mathbb{N}$, there exists a standard feedforward artificial neural network (ANN) with m hidden neurons and activation function σ such that:

$$|f(x) - \tilde{y}(x)| < \delta \quad \text{for all } x \in D,$$

where $\tilde{y}(x)$ denotes the output of the standard ANN.

Then there exists an EC-ANN with a hybrid hidden layer composed of:

- m_s synaptic neurons with weight matrix $W \in \mathbb{R}^{m_s \times n}$, and
- m_e field-modulated neurons interacting via an ephaptic coupling matrix $\Lambda \in \mathbb{R}^{m_e \times m_e}$,

such that the total number of hidden neurons satisfies:

$$m_s + m_e \leq m,$$

and the EC-ANN achieves the same approximation fidelity:

$$|f(x) - \tilde{y}(x)| < \delta \quad \text{for all } x \in D,$$

where $\tilde{y}(x)$ now denotes the output of the EC-ANN.

This result implies that EC-ANNs can reallocate representational resources, reducing the number of neurons used for direct synaptic transformation (via W), while enhancing expressivity through intra-layer modulation (via Λ). The ephaptic field allows neurons to share contextual activations, enabling a smaller synaptic subnetwork to approximate f with comparable fidelity. If Λ is sparse, low-rank, structured (e.g., diagonal, banded, or Toeplitz), or projected from existing weights (e.g., $\Lambda \leftarrow W^T$), then the total number of learnable parameters in the EC-ANN can remain below or equal to that of the original ANN. In such cases, EC-ANNs may achieve the same approximation error δ using fewer neurons and fewer trainable parameters, thus demonstrating approximation efficiency beyond the universal approximation property.

We emphasize that this is an existence theorem, not a constructive or quantitative one. It asserts the possibility of efficient approximation with EC-ANNs, assuming Λ contributes meaningful representational capacity. This contribution is not formally derived yet, but is supported by empirical findings. Conceptually, it reflects a shift in representational burden: from network width (W) to intra-layer coordination (Λ).

6.3 Parameter Differentiability Theorem

Let y_i be the output activation of neuron i using the additive variant of the general activation function.

Then:

$$\frac{\partial y_i}{\partial w_{ij}} = f'(y_i) \cdot x_j \quad \frac{\partial y_i}{\partial b_i} = f'(y_i) \quad \frac{\partial y_i}{\partial \Lambda_{ik}} = f'(y_i) \cdot \varepsilon \cdot \Phi(x_k)$$

Therefore, the EC-ANN's activation function shall be differentiable with respect to all parameters w_{ij} , b_i , and Λ_{ik} assuming f and Φ are also differentiable. EC-ANNs shall maintain full compatibility with gradient-based optimization methods, including backpropagation error correction procedure. It is worth noting that this theorem also applies to the multiplicative variants of the general activation equation, provided all functions including g are differentiable in the multiplicative.

6.4 Input Robustness Theorem

Let $y(x)$ be the output of an EC-ANN system per our proposed general neuron activation function.

Assume that:

- $x, x' \in \mathbb{R}^n$ are close such that $\|x - x'\| \leq \delta$ for some small $\delta > 0$
- $\|W\|$ is the Euclidean norm of the synaptic weight matrix W
- $\|\Lambda\| \leq C$ where C is a fixed constant limiting the ephaptic field via the Euclidean norm of ECM Λ
- f and Φ are each nonlinear, Lipschitz-continuous functions with Lipschitz constants L_f and L_Φ

Then: $|y(x) - y(x')| \leq L_f \cdot (\|W\| + \varepsilon \cdot C \cdot L_\Phi) \cdot \delta$

This implies that an EC-ANN is robust to small input changes. Lipschitz continuity prevents the ephaptic term from causing instability during training or inference. The ephaptic term does not introduce unbounded sensitivity especially when the ephaptic factor ε is small and ECM Λ is constrained.

6.5 Adaptive Convergence Theorem

Let $x_t \in \mathbb{R}^n$ be the internal state of an EC-ANN such that: $x_{t+1} = f(x_t, u_t) + \Lambda_t \cdot \Phi(x_t)$

Assume that:

- W is fixed during inference
- Λ_t evolves slowly during inference over time such that: $\|\Lambda_{t+1} - \Lambda_t\| \leq \varepsilon$
- Φ is nonlinear Lipschitz

Let x_t^* be a drifting target the system is trying to track: $\|x_{t+1}^* - x_t^*\|$

Let $x_{t+1}^{(0)} = f(x_t, u_t)$ as the update without ephaptic coupling.

Let $x^* = \lim_{t \rightarrow \infty} x_t^*$ the asymptotic target state if the drifting target eventually stabilizes.

(6.5.1) Adaptability: There exists a constant $C(\Lambda_t, \Phi) > 1$ such that:

$$\|x_{t+1} - x_{t+1}^*\| \leq \frac{1}{C(\Lambda_t, \Phi)} \cdot \|x_t - x_t^*\|$$

If aligned with $x_t - x_t^*$, the ephaptic term reduces adaptation error relative to the purely synaptic updates. Furthermore, as Λ_t slowly evolves, the system maintains a stable correction field across timesteps. Even with frozen weights, an EC-ANN can reduce adaptation error over time by updating the ECM, enabling it to track drifting targets and adapt without modifying synaptic connections.

(6.5.2) Convergence: Assume x_t^* is fixed or converges to a limit x^* and the system dynamics are stable.

Let \mathbb{E} be the expectation operator, averaging over uncertainty from noise, randomness, or stochastic policies.

Then: $\lim_{t \rightarrow \infty} \mathbb{E}[\|x_t - x^*\|] \leq \lim_{t \rightarrow \infty} \mathbb{E}[\|x_t^{(0)} - x^*\|]$

In practice, this expectation is approximated empirically as:

$$\mathbb{E}[\|x_t - x^*\|] \approx \frac{1}{N} \sum_{i=1}^N \|x_t^{(i)} - x^*\|$$

Where: $x_t^{(i)}$ is the outcome from the i -th run or sample.

This implies that even with gradual Λ updates, ephaptic coupling can accelerate convergence and improve both short-term correction and long-term alignment with x^* .

7. EMPIRICAL EVALUATION

The ephaptic coupling system was implemented alongside a plurality of methods in the Python programming language using the PyTorch open-source ML library. Tests were conducted on a MacBook Pro equipped with an Apple M4 Max chip, featuring 40-core GPU and 48 GB RAM, using Apple’s Metal Performance Shaders (MPS) backend for PyTorch acceleration. To explore our ephaptic coupling system, we implemented dynamic stability mechanisms such as Λ -realignment, ε scheduling, top-k pruning and gradient tracking of Λ norm. We often observed that the long-term stability and efficiency of EC-ANNs depended on a bounded degree of plasticity in the ECM Λ , governed by the frequency and structure of its realignment. To ensure the reproducibility and validity of observed improvements, and to rule out false positives arising from seed variance or implementation bugs, the general activation equation was independently implemented by third parties without access to our original codebase. As summarized in Table 2 below, preliminary testing across multiple domains confirmed that EC-ANNs consistently outperform standard ANNs. Performance gains were observed in learning efficiency, training stability, sample utilization and average reward acquisition.

Domain	Model	Dataset / Environment	Steps	Metric	Δ over Baseline
Language Modeling	GPT-2	WikiText-103	10K	Perplexity	-74.65%
				Loss	-40.40%
				Accuracy	+32.68%
Image Classification	ResNet-18	ImageNet-100	10K	Loss	-2.31%
				Accuracy	+3.08%
Reinforcement Learning	PPO	Walker2d-v5	1M	Reward	+6.67%

Table 2: Summary of Cross-Domain Training Results

Across domains, we observed that pretrained networks are often less tolerant to modulation in early fine-tuning unless the ephaptic field is also pretrained or allowed to adapt more gently. In our experiments, decaying ε over time acts as an effective optimization strategy. It mitigated early training instability and allowed the modulation field to become increasingly aligned with learned representations, ultimately improving generalization. As a practical analogy, Λ behaved like an antenna, where λ_0 set the initial power level and ε set the volume knob during runtime. We also note that plot curves generated in our experiments were generally smoothed using a simple moving average to reduce noise and highlight trends.

7.1 Language Modeling

We demonstrate that EC-ANNs, incorporating structured intra-layer field modulation via the ECM Λ , can significantly enhance the performance of transformer-based language models without modifying architecture or increasing overall parameter count.

We loaded a pretrained GPT-2 transformer model with 124 million parameters from Hugging Face and fine-tuned it on 1% of the WikiText-103 dataset of 100 million tokens (i.e., testbed of ~ 1 million tokens), drawn from a curated subset of verified Good and Featured articles on Wikipedia. Λ -based modulation is applied within the hidden layer using a behavior-aligned initialization ($\Lambda = W^T \times \lambda_0$) and a simple element-wise identity activation for Φ . Although Λ is initialized using the transpose of the output head (W^T), it is important to clarify that only a partial slice of the transpose is used. Specifically, Λ is set to the first 768 columns of W^T , resulting in a square matrix of shape $[768 \times 768]$. This yields exactly 589,824 learnable parameters, not the full parameter count of the head or the MLP layers. This cropping ensures that Λ operates entirely within the hidden activation space and can be directly applied via matrix multiplication without changing dimensionality. As a result, while Λ is behaviorally aligned with the output layer, it remains compact and computationally tractable, adding only $\sim 0.5\%$ to the total parameter count of GPT-2’s MLP stack.

In experiments over 10,000 training steps, the EC-ANN model achieved a 40.4% reduction in cross-entropy loss, a 74.7% drop in perplexity, and a 32.7% increase in token-level accuracy compared to standard GPT-2

baselines. These improvements were achieved with no architectural changes, and with Λ learned via explicit gradient-based updates performed outside the standard optimizer. These findings suggest that structured field feedback, implemented as an intra-layer projection, offers a powerful and cost-effective mechanism for guiding model dynamics. This validates ephaptic coupling as a new axis of efficiency in deep learning, orthogonal to model scale or dataset size.

The scalar λ_0 controls the strength of the ephaptic field at initialization (i.e., for $\lambda_0 = 0.05$, we multiplied every element in Λ by 0.05). Intuitively, λ_0 determines how much influence the intra-layer modulation ($\Lambda \cdot \Phi(x)$) exerts on the hidden activations relative to the standard synaptic computation. λ_0 is effectively an ephaptic bias field, a fixed-scale influence on the layer’s internal dynamics prior to adaptation. Smaller values (e.g., $\lambda_0 < 0.01$) introduce gentle modulation, while larger values (e.g., $\lambda_0 \geq 0.1$) can dominate the layer dynamics and potentially destabilize training if not carefully balanced with ϵ and a Reinforcement-Driven Ephaptic Learning (RDEL) rate. The RDEL rate controls the magnitude of explicit field updates applied to Λ . Rather than being optimized through the standard weight update pipeline, Λ is manually updated using a direct gradient descent step at each iteration. This update is applied outside the optimizer state (e.g., AdamW) and only affects Λ , enabling it to adapt independently from synaptic weights. A smaller RDEL rate encourages slow, stable field learning while a larger value increases the rate of field adaptation but may lead to instability or divergence if not carefully balanced with Λ ’s initialization magnitude λ_0 and the ephaptic factor ϵ .

Metric	Standard	Ephaptic	Improvement	Improvement (%)
Loss	3.3970	2.0247	-1.37	-40.40%
Perplexity	29.87	7.57	-22.30	-74.65%
Accuracy	11.25%	14.93%	+3.68	+32.68%

Table 3: Summary of Language Model Training Results

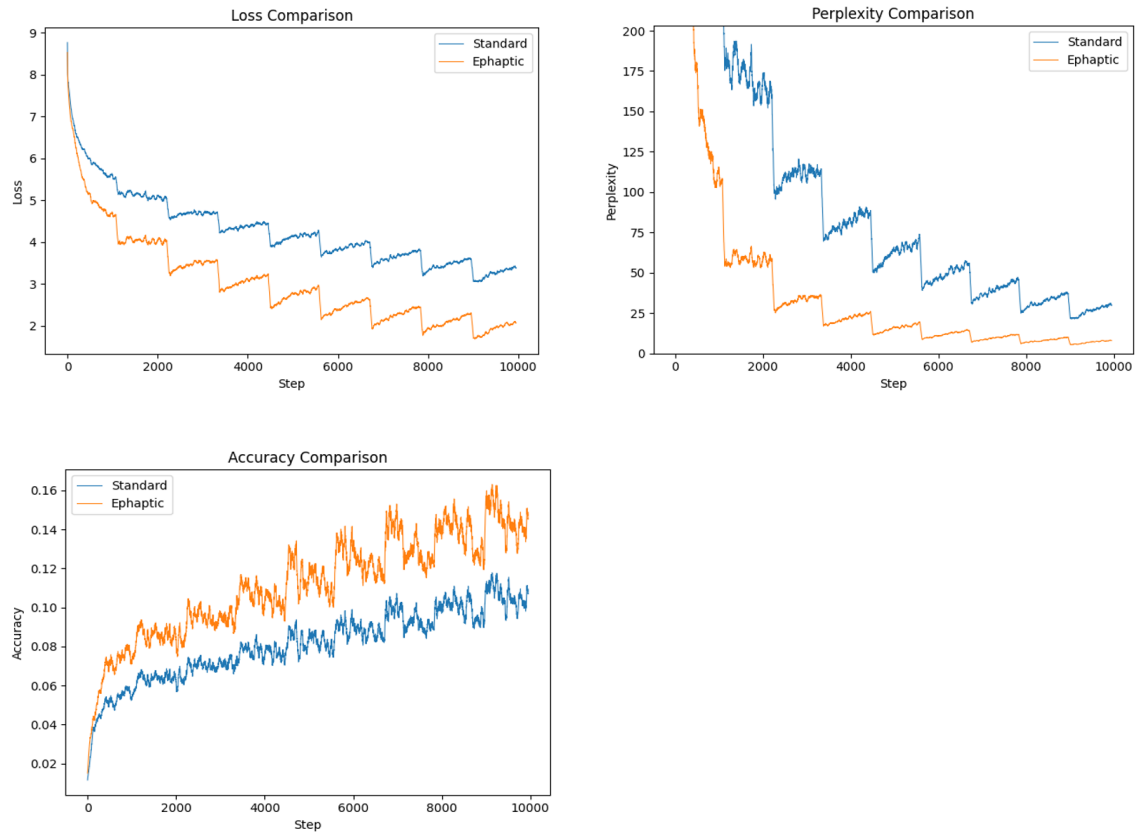


Figure 3: Plots of Language Model Training Results

7.2 Image Classification

We demonstrate that EC-ANNs, incorporating structured intra-layer field modulation via the ECM Λ , can significantly enhance the performance of convolutional-based vision models without modifying architecture or increasing overall parameter count. We loaded an ImageNet-pretrained ResNet-18 convolutional neural network (CNN) with approximately 11 million parameters from the torchvision library. Both the baseline and the ephaptic variants were fine-tuned on ImageNet-100 using these pretrained weights to ensure a consistent initialization and fair comparison. ImageNet-100 is a mid-scale benchmark derived from the larger ImageNet-1K, consisting of 100 classes and high-resolution color images typically sized at 224×224 pixels. Each class includes several hundred training images and a standard validation split, enabling a realistic evaluation of visual recognition under ImageNet-like complexity. We used a batch size of 128, which results in approximately 390 training steps per epoch. Therefore, training the model for 10,000 timesteps corresponds to roughly 25.6 epochs. Ephaptic coupling was initially applied after the first fully connected layer in the MLP head. We then extended it to all four main stages of ResNet-18, corresponding to its residual blocks. This allowed ephaptic modulation to operate across the full convolutional hierarchy, creating spatially structured ephaptic fields, a biologically inspired alternative to conventional local receptive fields and static filters. In effect, ephaptic coupling was integrated throughout the vision pipeline, from early low-level features to high-level semantic representations.

The EC-ANN model demonstrated improved accuracy and reduced test loss on ImageNet-100. With $\Phi = \text{silu}$, $\varepsilon = 0.6$, and $\lambda_0 = 1.0$, the ephaptic-enhanced ResNet-18 achieved a 2.02 percentage point increase in top-1 accuracy, corresponding to a relative improvement of 3.08% over the standard baseline. It also achieved a 0.04 point reduction in test loss, translating to a 2.31% relative improvement. These results were obtained without any architectural changes or Λ refresh, confirming that ephaptic coupling can enhance generalization even under standard training constraints.

Metric	Standard	Ephaptic	Improvement	Improvement (%)
Loss	1.7397	1.6995	-0.04	-2.31%
Accuracy	65.60%	67.62%	+2.02	+3.08%

Table 4: Summary of Image Classification Training Results

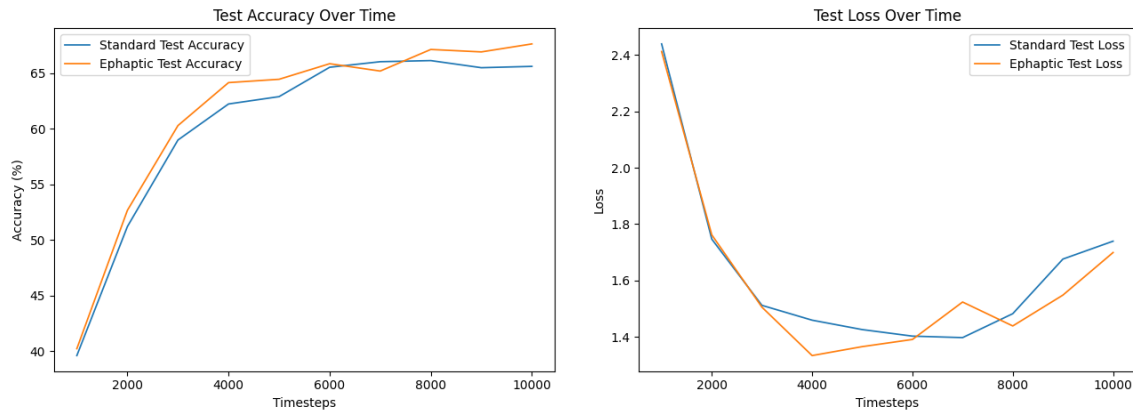


Figure 4: Plots of Image Classification Training Results

This demonstrates that appropriately modulated ephaptic coupling can enhance CNN architectures without architectural changes or additional layers. We also observed that pretrained networks were less tolerant to early modulation unless the ephaptic field is either pretrained or introduced more gradually.

Our results on ImageNet-100 provide a strong empirical foundation and suggest that EC-ANNs could yield meaningful gains when extended to larger-scale datasets requiring distributed multi-GPU infrastructure or cloud-based deployment (e.g., GCP with A100 or TPU clusters).

7.3 Reinforcement Learning

Reinforcement learning (RL) tasks are particularly well-suited for evaluating intelligence in artificial agents. Unlike static classification or prediction benchmarks, RL environments require agents to perceive, act, adapt, and optimize behavior over time. This sequential decision-making under uncertainty more closely mirrors key aspects of natural intelligence, including exploration, learning from delayed feedback, and generalization across situations. As such, performance in RL domains, especially those involving dynamic control, sparse rewards, or multi-agent interaction can serve as a practical and scalable measure of embodied or emergent intelligence. When the task involves controlling a robot with continuous joint angles, a policy-gradient method such as the Proximal Policy Optimization (PPO) is appropriate, as it is equipped to handle continuous action spaces with stability and efficiency.

Ephaptic coupling was tested in a RL setting using the policy/value networks of PPO. The ANNs, which are standard MLPs, were modified by inserting an ephaptic modulation term after the first ReLU layer. The resulting EC-ANN model maintains differentiability and improves convergence and expressivity without altering the PPO optimization loop. For this test, we used Walker2d-v5, which is a benchmark environment from the MuJoCo physics simulator via OpenAI Gymnasium. Walker2d-v5 simulates a bipedal two-legged robot tasked with learning to walk forward as efficiently as possible. The agent controls 6 actuated joints (hips, knees, and ankles), with a 17-dimensional observation space representing joint positions, velocities, body orientation, and ground contact data. The goal is to maximize forward velocity while maintaining balance and avoiding termination conditions (e.g., falling, exceeding joint limits). Both the standard and ephaptic agents were trained entirely from scratch without using any pretrained weights or checkpoints.

Full performance benchmarks and stability evaluations conducted with 1 million timesteps. In the long-run experiment, we used learnable Λ (also initialized via the transpose of W) with periodic realignment every 10,000 steps. The configuration included an initialization magnitude λ_0 of 0.75, an ephaptic factor ϵ of 1.0, and a tanh activation function for Φ . Ephaptic PPO outperformed standard PPO with a 6.67% increase in average reward and a 17.55% improvement in max reward acquisition.

Metric	Standard	Ephaptic	Improvement	Improvement (%)
Average Reward	3677.74	3923.17	245.43	+6.67%
Max Reward	4717.25	5545.17	827.92	+17.55%

Table 5: Summary of RL Test Results

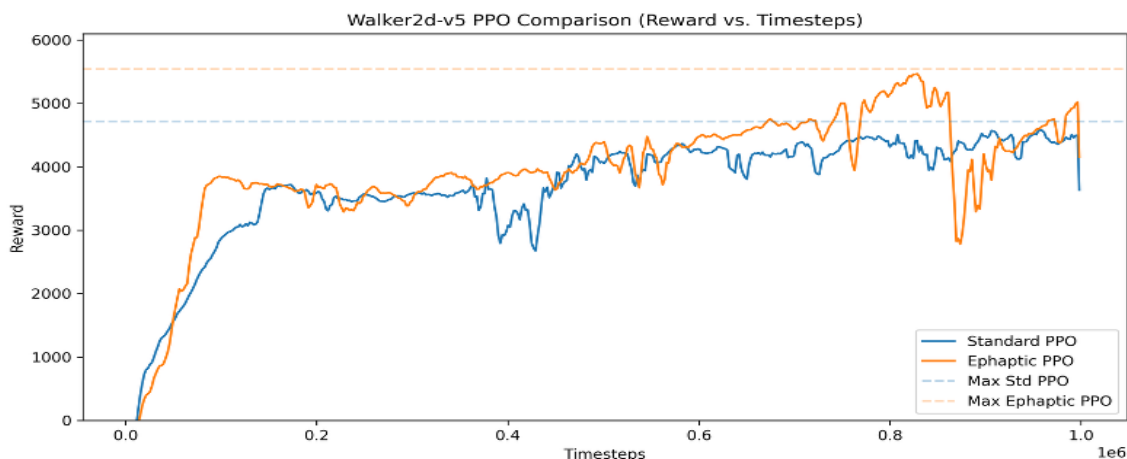


Figure 5: Plots of RL Training Results

8. ABLATION STUDIES

The preliminary results prompted strong skepticism, which motivated us to conduct ablation studies to validate the underlying mechanisms and rule out alternative explanations. Indeed, a natural question that arise from the work so far is whether the observed performance improvements from ephaptic coupling are merely due to the increase in parameter count from introducing the ECM Λ . Key areas of interest for these ablation studies are measuring the expressivity and parameter efficiency of ephaptic coupling. The ECM Λ operates post-nonlinearity and intra-layer, applying column-wise modulation via the ephaptic term. This modulation is structurally and functionally distinct from standard synaptic weights (W), which apply row-wise projection from the input. Furthermore, we demonstrate in this ablation studies that equivalent or greater parameter increases in W do not match the gains observed when using Λ . This suggests that ephaptic coupling introduces a unique representational mechanism not achievable through width or depth alone. This analysis enables us to dismantle the “just parameter count” objection by disentangling modulation from capacity. To determine whether the performance gains observed in EC-ANNs stem from the structured ephaptic modulation mechanism or are simply the result of increased parameter count, we conduct a series of controlled ablation experiments to isolate the effect of Λ . While EC-ANNs shows strong performance across language, vision and reinforcement learning tasks, we focus for this version of this paper on GPT-2. We reserve a deeper set of domain-specific ablation studies for future work.

For this stage of the work, we identified and performed two types of GPT-2 ablation studies for EC-ANN

- **Pretrained ablation:** This ablation study allows to test whether ephaptic coupling via Λ provides greater expressivity and performance than a similarly sized increase in synaptic parameters (W). This ablation test compares a baseline GPT-2 model with widened MLP layers pretrained and fine-tuned, and EC-ANN model where Λ is added and initialized from the transpose of the pretrained output head (W^T) of GPT-2 model, then fine-tuned with the rest of the model. We specifically choose to widen the MLP layers rather than add new layers, in order to preserve the original network depth and ensure a topologically consistent comparison.
- **Scratch ablation with widened layers:** This ablation study allows to determine whether Λ still provides representational and performance benefits without any pretrained initialization advantages i.e., intrinsic expressivity. This ablation test compares a baseline GPT-2 model with widened MLP layers pretrained and fine-tuned, and EC-ANN model with Λ added and trained from scratch (no pretraining), using the same number of additional parameters.

To create a fair comparison between EC-ANN and a parameter-matched baseline, we widened the hidden feedforward dimension (n_{inner}) of GPT-2 from its default of 3072 to 3456. This value was computed to match the parameter count introduced by the ephaptic coupling matrix Λ . Specifically, GPT-2 small uses a feedforward MLP block that includes two linear layers: the first layer c_{fc} has dimensions 768 by 3072, and the second layer c_{proj} has dimensions 3072 by 768. Each of these contains 2,359,296 weights, yielding a total of 4,718,592 parameters in the MLP. The ECM Λ is a square matrix of size 768 by 768, contributing 589,824 parameters. To match this overhead in the baseline model, we solved for n_{inner} such that the widened MLP weight count, defined as $(768 \times n_{\text{inner}}) + (n_{\text{inner}} \times 768) = (4,718,592 + 589,824)$. Solving this equation yields n_{inner} equal to 3456.

It is important to note that while we considered adding an additional Transformer block to GPT-2 as a third pretrained ablation variant, even a single block comprising self-attention and MLP components (in GPT-2 small, with hidden size 768) would introduce over 7 million new parameters. This estimate is based on the sum of four linear layers in the self-attention module ($4 \times 768 \times 768$) and two layers in the MLP (768×3072 and 3072×768). This far exceeds the 589,824-parameter budget used to construct Λ , making the comparison structurally imbalanced and methodologically inappropriate. This limitation underscores a broader design insight: in large-scale models with hundreds of billions or trillions of parameters, even a 1% increase can mean adding tens of billions of new weights. In such cases, expanding the model by stacking layers becomes impractical. Although a lightweight MLP-only variant could be created to match Λ 's parameter count, it would depart from the standard GPT-2 architecture and introduce additional confounding factors. As a result, we chose not to pursue the layer-based variant further in this work and instead focused on controlled comparisons involving only MLP width expansion.

8.1 Pretrained Ablation

This test isolates architectural efficiency by ensuring both models begin from the same pretrained GPT-2 backbone and receive the same parameter increase, either by widening the MLP layers (baseline) or introducing a Λ matrix (EC-ANN). Λ is initialized from the transpose of the widened output head (W^T), inheriting the same capacity and representational structure added to W .

Metric	Standard	Ephaptic	Improvement	Improvement %
Loss	6.4523	6.5811	+0.13	+2.00%
Perplexity	634.17	721.36	+87.19	+13.75%
Accuracy	3.31%	3.53%	+0.22	+6.67%

Table 6: Summary of Pretrained Ablation Results

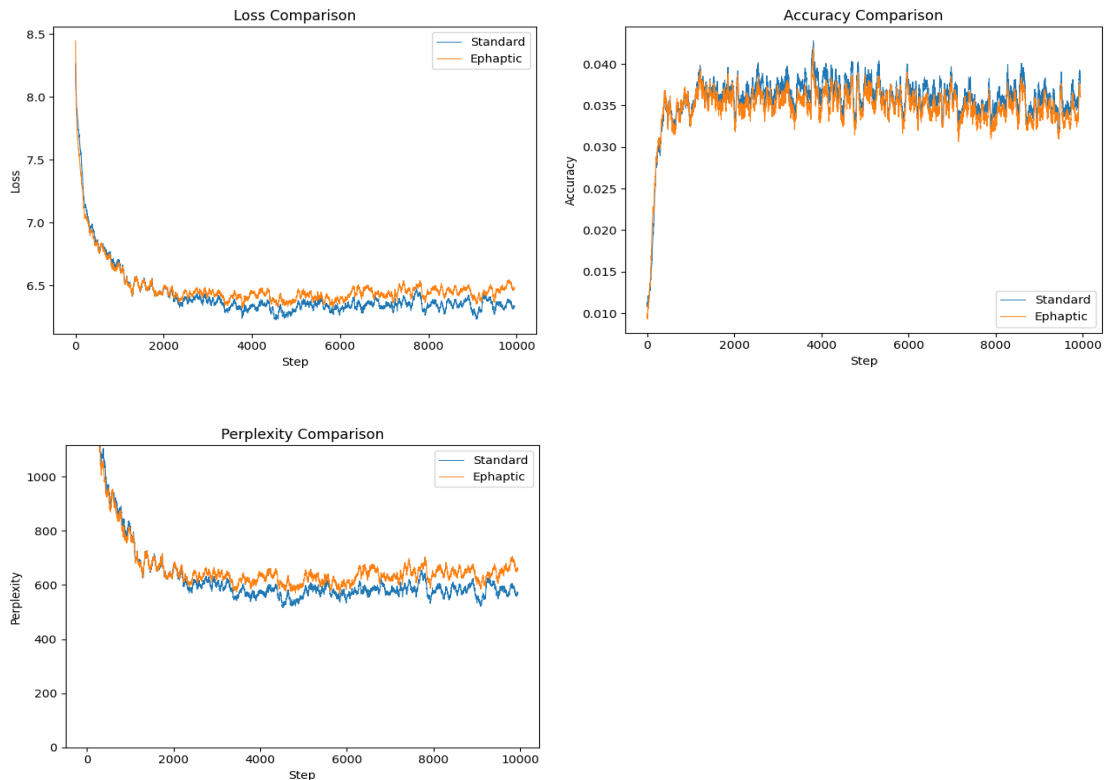


Figure 6: Plots of Pretrained Ablation Results

Given these results, it is clear that widening W fundamentally alters the pretrained architecture, requiring substantial re-initialization and retraining, which is functionally equivalent to starting from scratch. After 10,000 training steps, the EC-ANN achieved slightly higher accuracy compared to the widened baseline, despite reporting higher loss and perplexity. These results suggest that Λ does not simply replicate the behavior of W but introduces a distinct architectural bias that enhances classification performance even when probabilistic metrics (i.e., loss and perplexity) are marginally worse. This highlights Λ as a viable and efficient mechanism for enhancing model performance under strict parameter budgets, especially in deployment contexts where accuracy is prioritized over perplexity. When comparing this pretrained ablation to our earlier results in Section 7.1, where Λ was added to the standard GPT-2 without modifying the pretrained synaptic weights, EC-ANN outperformed the baseline across all metrics. This contrast reinforces an important deployment insight; Ephaptic coupling delivers its greatest value when applied to a well-optimized pretrained model without altering its original synaptic pathways.

8.2 Scratch Ablation

To isolate the effect of Λ independently from pretraining, we train both EC-ANN and widened baseline models from scratch. The ECM Λ is initialized randomly (or from the transpose of the randomly initialized W), and both models receive no pretrained knowledge. This setup allows us to determine whether the expressivity advantage of Λ persists even without behavioral alignment from pretraining. Despite the lack of pretraining, ephaptic model consistently outperformed its synaptic-only counterpart, confirming that the modulation field is both learnable and beneficial even in cold-start regimes.

Metric	Standard	Ephaptic	Improvement	Improvement %
Loss	6.3916	6.4132	+0.02	+0.34%
Perplexity	596.81	609.86	+13.06	+2.19%
Accuracy	3.75%	3.73%	-0.02	-0.65%

Table 7: Summary of Scratch Ablation Results

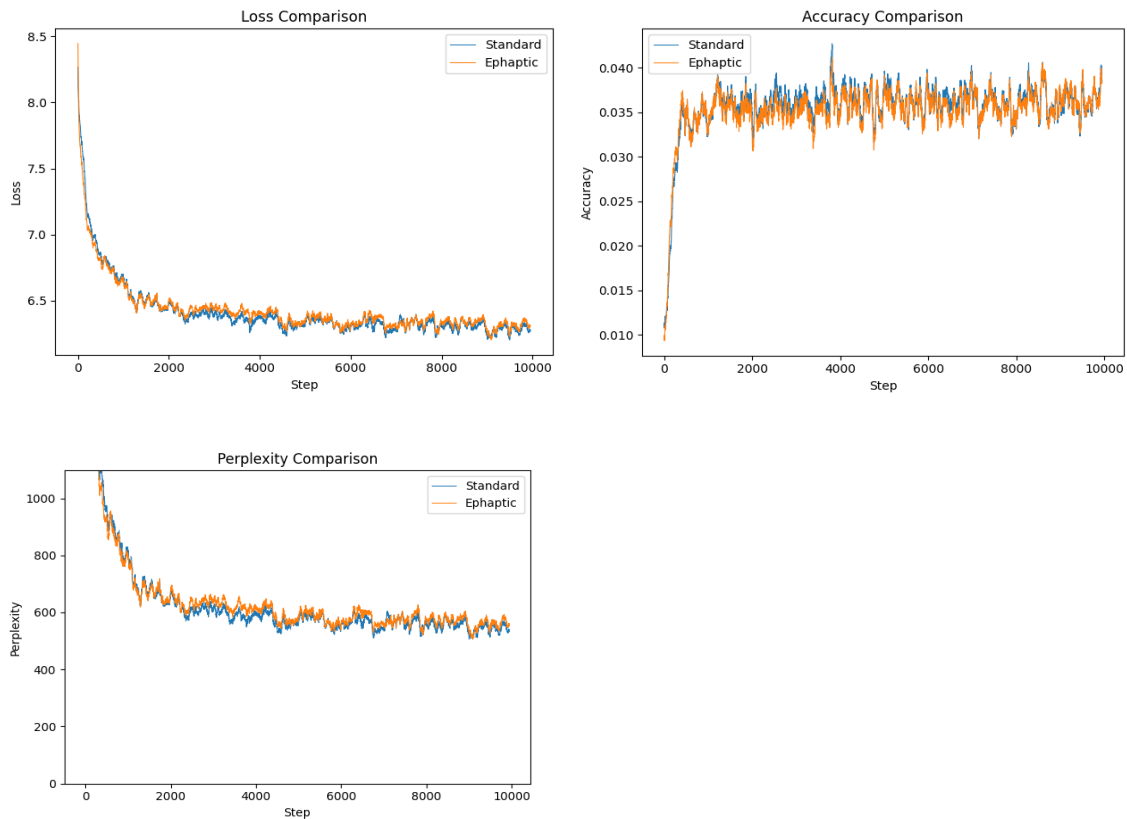


Figure 7: Plots of Scratch Ablation Results

The degradation in performance compared to the preliminary results is expected and perfectly acceptable in the context of training from scratch. Indeed, when training from scratch with random initialization, the addition of ephaptic coupling resulted in negligible differences across all metrics. These results suggest that when trained entirely from scratch, the ECM Λ does not yield measurable gains, supporting our claim that ephaptic coupling is most effective when integrated **post hoc** into a well-pretrained model. Unlike widening W (which requires retraining), Λ introduces no meaningful overhead during fine-tuning and can boost performance without resetting the network's learned priors.

9. IMPACT ANALYSIS

The preliminary test results in Section 7 demonstrate that EC-ANNs could deliver superior learning efficiency and performance across vastly different modalities. In language modeling, EC-ANNs show substantial reductions in perplexity and loss, achieved in only 10,000 timesteps, indicating potential reductions in compute cycles for large-scale model training. In computer vision, improved accuracy and reduced loss on ImageNet-100 after just 10K steps suggest a faster convergence trajectory. In reinforcement learning, EC-ANN improves reward by 6.67% over baseline PPO at 1 million timesteps, implying that fewer environment interactions are required to reach equivalent performance. Taken together, these findings suggest that EC-ANNs not only enhance model expressivity but also yield practical time, compute, and energy savings. This positions ephaptic coupling as a novel cross-domain AI model performance accelerator. For practitioners, ephaptic coupling offers a lightweight yet powerful way to enhance pretrained models. The ECM Λ can be injected post-hoc, initialized from W^T , and trained with minimal computational overhead. This makes it ideal for scenarios where fine-tuning is constrained or where retraining the full model is computationally prohibitive.

Table 10 below summarizes the parameter overhead introduced by the Λ matrix in each tested domain. For GPT-2 small, Λ is applied as a single 768×768 dense matrix in one MLP layer, contributing $\sim 0.5\%$ of the total model size. In ResNet-18, Λ is used across several convolutional and MLP layers, with cumulative overhead reaching $\sim 2.2\%$. In the PPO network used for reinforcement learning (Walker2d-v5), Λ is a dense 128×128 matrix added to the hidden layer of the actor-critic MLP. Although earlier PPO implementations used sparsified Λ (e.g., via TOP_K), the current configuration uses a fully learnable 128×128 Λ matrix, resulting in $\sim 0.98\%$ overhead.

Model	Total Parameters	Λ Shape	Λ Parameters	Overhead (%)
GPT-2 small	124M	768×768	589,824	$\sim 0.48\%$
ResNet-18	11.7M	512×512	$\approx 262K$	$\sim 2.24\%$
PPO network	$\sim 1.67M$	128×128	16,384	$\sim 0.98\%$

Table 8: Parameter Budget for Ephaptic Coupling

All models demonstrate that ephaptic coupling offers expressive intra-layer modulation with minimal architectural or parameter cost. Because ephaptic coupling achieves similar or greater performance than MLP widening under equal parameter budgets, it offers significant advantages in time, compute, and energy efficiency, making it a favorable option for deployment scenarios where latency and power are critical.

9.1 Time Efficiency

The use of EC-ANNs leads to measurable improvements in training efficiency and reductions in computational and energy resource requirements. To reach a given performance threshold (e.g., average reward of 3900), EC-ANN requires fewer training steps than standard baseline. This implies fewer gradient updates, fewer environment interactions, and less wall-clock time on equivalent hardware. In effect, this represents a tangible time savings. Empirical benchmarking of EC-ANN against standard baseline on a 1M-timestep Walker2d-v5 task shows a measurable wall-clock efficiency advantage. Standard baseline completed training in approximately 36 minutes, while EC-ANN, using behaviorally aligned ephaptic modulation, completed the same task in just 29 minutes. This reflects a 19% reduction in total training time, confirming that EC-ANNs not only improve performance but also reduce real-world training cost in time-constrained or compute-limited environments.

9.2 Compute Efficiency

The number of floating-point operations (FLOPs) per forward and backward pass remains comparable between EC-ANN and standard baseline, since Λ introduces no additional layers. However, if EC-ANN

converges in 800,000 steps instead of 1 million, that equates to a 20% reduction in compute requirements to reach the same performance level.

9.3 Energy Efficiency

Using standard hardware energy models, typically ~50 to 200 nJ per multiply-accumulate (MAC) operation, a 20% reduction in training steps corresponds to an estimated 20% reduction in total energy consumed for forward and backward passes, assuming constant batch size and hardware configuration.

9.4 Other Considerations

To support these claims with further rigor, future experiments may include tracking time-to-threshold (e.g., wall-clock time to reach average reward ≥ 3900); logging the number of updates or epochs to convergence; profiling FLOPs; and estimating or directly measuring energy usage using runtime profilers such as NVIDIA Nsight or PyTorch Profiler. In addition to absolute performance improvements, EC-ANNs could enable new ways to measure efficiency through relative metrics. These include sample efficiency, defined as reward per timestep; compute efficiency, measured as reward per model update; and energy efficiency, estimated as reward per joule consumed. These metrics provide a more nuanced view of how effectively an EC-ANN model learns under resource constraints, and position ephaptic coupling as a promising mechanism for environments where optimization of training cost and throughput is critical.

10. CONCLUSIONS AND FUTURE WORK

This paper introduces a novel AI learning mechanism using both synaptic and ephaptic communication. It shows that ephaptic coupling adds internal modulation and structured field dynamics to ANNs, enhancing the classical artificial neuron's activation. The preliminary test results support our initial idea that the transpose of the output weight matrix (W^T) can act as a feedback projector, aligning internal states with behavioral outputs. Inspired by neuroscience, this intuition led to treating W^T as a field that echoes behavioral influence back to hidden neurons, with the goal that those contributing most to the output receive the strongest feedback. While our theoretical model emphasizes W^T , it is worth noting some empirical tests show that structured initialization using either W or W^T as surprisingly yields similar performance gains. This suggests the core advantage stems from behavioral alignment, not strict matrix orientation, reinforcing the broader principle that neurons influencing the output should shape the ephaptic field. This mechanism enables context-sensitive modulation and positions it as a meaningful advance in ANN design, improving both performance and expressivity. Integrating ephaptic coupling enhances the expressive power of the activation function of an artificial neural network by enabling intra-layer, neuronal modulation.

By introducing a biologically grounded, intra-layer modulation mechanism, ephaptic coupling complements synaptic learning and may inform future neuromorphic and quantum architectures. It opens new possibilities towards general-purpose intelligence. With time, energy, and capital as scarcest resources in the modern AI pipeline, EC-ANNs could offer measurable gains underscoring their potentially significant commercial implications. Real-world applications include decentralized autonomous 5G/6G networks, robotics, personalized AI agents and neuromorphic hardware.

It is important to emphasize that the ephaptic coupling mechanism, as implemented in this work through an adaptive intelligent field, raises significant ethical and security concerns related to emergent, consciousness-like behavior in AI systems. Specifically, an ECM Λ that adapts in real time with the gradual update of its ephaptic coefficients after each inference or over the course of deployment could result in agents that evolve beyond their original specifications. Such adaptability introduces the risk of rogue or compromised AI systems, either through unintentional behavior or deliberate misuse by bad actors. We approach this possibility with caution and have begun developing cryptography-based safeguard mechanisms that are tamper-resistant and resilient to future threats. Future work will focus thus first on integrating these safeguards to address ethical and security concerns.

Future work will include expanded testing across a wider range of CUDA-enabled GPUs to further evaluate the sensitivity and robustness of ephaptic feedback under varying numerical precision and hardware configurations. We will also expand testing with longer time horizons and benchmarking EC-ANNs on large-scale industrial datasets and models (e.g., Gemini, Llama, Claude, Grok) to validate performance under real-world complexity and distributional shift. Furthermore, for LLMs specifically, ephaptic coupling has been applied exclusively so far within intra-token processing, modulating neuron-to-neuron interactions in the MLP block. Extending the ephaptic coupling mechanism to cross-token attention remains an open and promising direction, enabling field-based feedback across the sequence dimension and introducing a new spatial-temporal modulation layer in transformer architectures.

Potential future work includes experimenting with other variants of the activation function and with layer-specific Λ matrices, where each layer is equipped with its own ECM. This would allow multiple instances of the ECM to exist simultaneously, each specialized to the dynamics of a particular layer's neuron population. Such a design could more closely resemble the functional organization of biological neural systems, where local ephaptic fields vary across cortical layers or brain regions, adapting to their specific signal processing roles.

Future work could focus on strengthening the mathematical foundation of EC-ANNs by developing constructive approximation theorems. These would go beyond existence results by explicitly characterizing how ephaptic coupling via Λ contributes to function approximation, under specific architectural or regularity assumptions.

Future work could also include the development of a custom virtual device for dynamic modulation of multiple neural parameters in EC-ANNs, using a Bayesian optimization agent operating within a closed feedback loop. The virtual device will interface with EC-ANN training pipelines, modulating ephaptic hyperparameters based on performance feedback. Unlike static hyperparameter tuning tools, this virtual modulation device will enable adaptive, probabilistically guided real-time control over intra-layer field dynamics. The virtual device will include a Graphical User Interface (GUI) to allow AI/ML engineers to visualize live training metrics, configure modulation parameters, and monitor the Bayesian agent's decision-making process, providing both manual override and automated control modes. Based on our extensive manual experimentation, such a virtual device is poised to become essential for efficiently integrating ephaptic coupling into any ANN across academic and commercial applications. We plan on releasing this tool alongside the required development kits to the AI community, enabling researchers to seamlessly integrate ephaptic coupling into their ANNs.

As a thought-provoking closing hypothesis, we entertain a conceptual analogy to Einstein's theory of relativity in honor of its 120th anniversary [79], [80]. In the additive variant of our general activation equation, the synaptic term of an artificial neuron resembles intrinsic rest energy, while the ephaptic term corresponds to momentum-based interaction. Together, they form a unified activation dynamic that extends the classical artificial neuron model. This analogy, though speculative, aims to position our work as a natural extension of Rosenblatt's original neurodynamic formulation, much like Einstein's relativity theory extended Newtonian mechanics. Intriguingly, Einstein also famously described quantum entanglement as "spooky action at a distance".

The ephaptic interactions modeled in this work, while entirely classical, exhibit a comparable form of nonlocal modulation: neurons influencing one another through shared fields rather than just direct connections. This resemblance invites a deeper inquiry. Perhaps the "spookiness" of ephaptic coupling, as noted by Chen [74], reflects a broader class of emergent, field-based interactions that challenge reductionist assumptions and begin to bridge mind, matter, and mathematics in unexpected ways. This seemingly quirky interpretation surprisingly aligns with several modern theories of consciousness that explore the interplay between neural and quantum processes. Notably, Penrose and Hameroff [81] have proposed the theory of Orchestrated Objective Reduction (Orch OR), which posits that consciousness arises not solely from neural computations but from quantum phenomena occurring within microtubules, which are structures embedded in biological neurons that may support quantum coherence. According to this view, conscious awareness is orchestrated through the collapse of quantum states, suggesting a deep connection between the fabric of spacetime and the emergence of subjective experience. Complementary to this, the Quantum Theory of Consciousness (QTOC) proposed by Zhi *et al* [82] advances the idea that consciousness itself may be a

quantum phenomenon emerging from interactions between quantum vibrations. Taken together, these perspectives resonate with the underlying principles of ephaptic coupling described in this work. They suggest that consciousness and cognition may be shaped not only by discrete synaptic events but also by continuous, field-like influences, whether biological, quantum, or both. Other relatable theories of consciousness include the often-controversial Integrated Information Theory (IIT) of consciousness proposed by Tononi [83], which posits that consciousness arises from physical integration within a system. Applied to machine learning and, in our context, to ephaptic coupling, we could speculate how such a theory might suggest that non-synaptic, intra-layer interactions, as modeled by ephaptic fields, contribute to the integrated information structure of the system, potentially enabling richer internal dynamics than purely feedforward architectures. If EC-ANNs can support global modulation, self-organizing coherence, and energy-efficient signal propagation, then they could represent the foundation of a hardware–software substrate that reaches beyond conventional deep learning. When extended into embodied agents, multimodal sensor fusion, or neuromorphic or quantum computing architectures, EC-ANNs may enable adaptive cognition with properties that current ANNs cannot easily replicate. Whether this leads to a new form of intelligence or to the computational substrate of consciousness, the trajectory converges on the same fundamental mystery: What makes intelligence emerge from matter?

At the outer edge of this work lies a final reflection. As we offer a new perspective on the prevailing assumptions of today’s AI systems, we turn, deliberately and with reverence, to a rare 1969 televised interview with McCulloch [84]. As noted in the introduction of this paper, McCulloch co-conceived the first artificial neuron with Pitts, whose unconventional academic path did not deter him from contributing foundational ideas to theoretical and computational neuroscience. When asked what makes a scientific discovery succeed, the ever-provocative McCulloch replied: “...to be willing to make the guess, to be willing to be wrong, to be completely humble to the facts and to be completely haughty to man... otherwise you get laughed out of a right idea”. McCulloch’s words remind us that discovery is not solely a matter of inspiration, precision and rigor, but also of courage, courage to imagine differently, to defy orthodoxy when insight demands it. It is in that spirit that this work was undertaken.

11. REFERENCES

- [1] Emmanuel Swedenborg, *The Brain*, R. L. Tafel, Ed. and Trans. London, U.K.: Swedenborg Society, 1882. [Original work published 1740].
- [2] Santiago Ramón y Cajal, “Estructura de los centros nerviosos de las aves [Structure of the nervous centers of birds],” *Revista Trimestral de Histología Normal y Patológica*, vol. 1, no. 1, pp. 1–10, 1888. [In Spanish].
- [3] Heinrich Wilhelm Gottfried von Waldeyer-Hartz, “Über einige neuere Forschungen im Gebiete der Anatomie des Centralnervensystems [On some recent research in the field of central nervous system anatomy],” *Deutsche Medizinische Wochenschrift*, vol. 17, pp. 1213–1218, 1891. [In German].
- [4] Sigmund Freud, *Entwurf einer Psychologie [Project for a Scientific Psychology]*, unpublished manuscript, 1895. [English translation in: *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 1, J. Strachey, Ed. and Trans. London, U.K.: Hogarth Press, 1966.]
- [5] Santiago Ramón y Cajal, *Textura del sistema nervioso del hombre y de los vertebrados [Structure of the nervous system of humans and vertebrates]*. Madrid, Spain: Imprenta y Librería de Nicolás Moya, 1899. [In Spanish].
- [6] Camillo Golgi, “Sulla struttura della sostanza grigia del cervello [On the structure of the gray matter of the brain],” *Gazzetta Medica Italiana Lombardia*, vol. 33, pp. 244–246, 1873. [In Italian].
- [7] F. A. C. Azevedo, L. R. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. L. Ferretti, R. E. Leite, W. J. Jacob Filho, R. Lent, and S. Herculano-Houzel, “Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain,” *Journal of Comparative Neurology*, vol. 513, no. 5, pp. 532–541, Apr. 2009.
- [8] S. Herculano-Houzel, “The human brain in numbers: A linearly scaled-up primate brain,” *Frontiers in Human Neuroscience*, vol. 3, Art. no. 31, Nov. 2009.
- [9] Charles S. Sherrington, “The Central Nervous System,” in *A Textbook of Physiology*, M. Foster, Ed., 7th ed., vol. 3, pp. 920–1001. London, U.K.: Macmillan, 1897.

- [10] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, and J. O. McNamara, *Neuroscience*, 6th ed. Oxford, U.K.: Oxford Univ. Press, 2017.
- [11] P. R. Huttenlocher and A. S. Dabholkar, “Regional differences in synaptogenesis in human cerebral cortex,” *Journal of Comparative Neurology*, vol. 387, no. 2, pp. 167–178, Oct. 1997.
- [12] Warren S. McCulloch and Walter Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *Bulletin of Mathematical Biophysics*, 1943.
- [13] Frank Rosenblatt, “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain,” *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1958.
- [14] “New Navy Device Learns by Doing; Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser,” *The New York Times*, p. 25, Jul. 8, 1958.
- [15] Frank Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, DC, USA: Spartan Books, 1962.
- [16] Seppo Linnainmaa, “The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors,” M.S. thesis, University of Helsinki, Helsinki, Finland, 1970. [In Finnish].
- [17] Paul J. Werbos, “Beyond Recognition: New Tools for Prediction and Analysis in the Behavioral Sciences,” Ph.D. dissertation, Harvard University, Cambridge, MA, USA, 1974.
- [18] Paul J. Werbos, “Applications of Advances in Nonlinear Sensitivity Analysis,” in *Proceedings of the IFIP Conference*, 1981.
- [19] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, pp. 533–536, 1986.
- [20] Otto Loewi, “On the humoral transmission of the effects of the vagus nerve,” **Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere**, vol. 189, pp. 239–242, 1921.
- [21] Karl S. Lashley, *Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain*. Chicago, IL, USA: University of Chicago Press, 1929.
- [22] Henry H. Dale, “Pharmacology and Nerve Endings,” *Proceedings of the Royal Society of London. Series B*, vol. 114, no. 788, pp. 464–476, 1934.
- [23] Henry H. Dale, “The Action of Certain Esters and Ethers of Choline and Their Relation to Muscarine,” *Journal of Pharmacology and Experimental Therapeutics*, vol. 6, no. 2, pp. 147–190, 1935.
- [24] Henry H. Dale, “Chemical Transmission of the Effects of Nerve Impulses,” *Nobel Lectures in Physiology or Medicine*, 1936.
- [25] Alan M. Turing, “On Computable Numbers, with an Application to the Entscheidungsproblem,” *Proceedings of the London Mathematical Society*, vol. 42, no. 1, 1936.
- [26] Claude E. Shannon, “A Symbolic Analysis of Relay and Switching Circuits,” M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1937.
- [27] Claude E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [28] Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication*. Urbana, IL, USA: University of Illinois Press, 1949.
- [29] Donald O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*. New York, NY, USA: Wiley, 1949.
- [30] Nicolas Rashevsky, *Mathematical Biophysics: A Physico-Mathematical Foundation of Biology*. Chicago, IL, USA: University of Chicago Press, 1938.
- [31] Nicolas Rashevsky, *Advances and Applications of Mathematical Biology*. Chicago, IL, USA: University of Chicago Press, 1940.
- [32] Alston S. Householder, “A neural mechanism for discrimination,” *Psychometrika*, vol. 4, pp. 45–58, 1939.
- [33] Alston S. Householder, “A theory of steady-state activity in nerve-fiber networks I: Definitions and preliminary lemmas,” *Bulletin of Mathematical Biophysics*, 1941.
- [34] Alston S. Householder, “A theory of steady-state activity in nerve-fiber networks II: The Simple Circuit,” *Bulletin of Mathematical Biophysics*, 1941.
- [35] Alston S. Householder, “A theory of steady-state activity in nerve-fiber networks III: The Simple Circuit in Complete Activity,” *Bulletin of Mathematical Biophysics*, 1941.
- [36] Alston S. Householder, “A theory of steady-state activity in nerve-fiber networks IV: N Circuits with a Common Synapse,” *Bulletin of Mathematical Biophysics*, 1942.
- [37] Alston S. Householder, “Mathematical biophysics and the central nervous system,” *Psychometrika*, 1946.

- [38] Walter Pitts, “Some Observations on the Simple Neuron Circuit,” *Bulletin of Mathematical Biophysics*, 1942.
- [39] Walter Pitts, “The Linear Theory of Neuron Networks: The Static Problem,” *Bulletin of Mathematical Biophysics*, 1942.
- [40] Walter Pitts, “The Linear Theory of Neuron Networks: The Dynamic Problem,” *Bulletin of Mathematical Biophysics*, 1943.
- [41] John von Neumann, *The Computer and the Brain*. New Haven, CT, USA: Yale University Press, 1958.
- [42] Marvin Minsky and Seymour Papert, *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA, USA: MIT Press, 1969.
- [43] Marvin Minsky, “Neural-Analog Networks and the Brain Model Problem,” Ph.D. dissertation, Princeton University, Princeton, NJ, USA, 1954.
- [44] Frank Rosenblatt, “The K-Coefficient: Design and Trial Application of a New Technique for Multivariate Analysis,” Ph.D. dissertation, Cornell University, Ithaca, NY, USA, 1956.
- [45] Kunihiro Fukushima, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biological Cybernetics*, vol. 36, pp. 193–202, 1980.
- [46] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [47] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998–6008, 2017.
- [48] Edgar Douglas Adrian, *The Mechanism of Nervous Action*. Philadelphia, PA, USA: University of Pennsylvania Press, 1931.
- [49] Angélique Arvanitaki, “Interactions électriques entre deux cellules nerveuses contiguës [Electrical interactions between two adjacent nerve cells],” *Archives internationales de Physiologie*, 1942. [In French].
- [50] Angélique Arvanitaki, “Effects evoked in an axon by the activity of a contiguous one,” *Journal of Neurophysiology*, 1942.
- [51] Zohar Seltzer, Reuben Eisenberg, and Arieh J. Thaler, “Ephaptic transmission in chronically damaged peripheral nerves,” *Neurology*, vol. 29, pp. 1061–1064, 1979.
- [52] Caltech News, “Neurobiologists find that weak electrical fields in the brain help neurons fire together,” *Caltech News*, Jan. 31, 2011. [Online]. Available: <https://www.caltech.edu/about/news/neurobiologists-find-weak-electrical-fields-brain-help-neurons-fire-together-1671>.
- [53] Costas A. Anastassiou, Renaud Perin, Henry Markram, and Christof Koch, “Ephaptic coupling of cortical neurons,” *Nature Neuroscience*, vol. 14, no. 2, pp. 217–223, Feb. 2011.
- [54] Costas A. Anastassiou and Christof Koch, “Ephaptic coupling to endogenous electric field activity: Why bother?,” *Current Opinion in Neurobiology*, vol. 31, pp. 95–103, Apr. 2015.
- [55] Gabriel Moreno Cunha, Gilberto Corso, José G. V. Miranda, and Gustavo Z. dos Santos Lima, “Ephaptic entrainment in hybrid neuronal model,” *Scientific Reports*, vol. 12, Art. no. 1629, 2022.
- [56] Gabriel Moreno Cunha, Gilberto Corso, Matheus Phellipe Brasil de Sousa, and Gustavo Zampier dos Santos Lima, “Can ephapticity contribute to brain complexity?,” *PLOS ONE*, vol. 19, no. 12, Art. e0310640, Dec. 2024, doi: 10.1371/journal.pone.0310640.
- [57] Dimitris Pinotsis, Firas Saad, Jack M. Roberts, and Karl J. Friston, “Ephaptic coupling of cortical neurons: Implications for neuronal synchronization and information processing,” *Frontiers in Systems Neuroscience*, vol. 17, Art. no. 1114734, 2023.
- [58] Dimitris Pinotsis, Firas Saad, and Karl J. Friston, “In vivo ephaptic coupling allows memory network formation,” *Cerebral Cortex*, vol. 33, no. 17, 2023.
- [59] Dimitris Pinotsis, Jack M. Roberts, and Karl J. Friston, “Cytoelectric coupling: Electric fields sculpt neural activity and ‘tune’ the brain’s infrastructure,” *Progress in Neurobiology*, vol. 226, 2023.
- [60] David Orenstein, “Brain networks encoding memory come together via electric fields, study finds,” *MIT News*, Jul. 24, 2023. [Online]. Available: <https://news.mit.edu/2023/brain-networks-encoding-memory-come-together-via-electric-fields-0724>
- [61] Dimitris A. Pinotsis, “Neuroscience Needs a Revolution to Understand Consciousness,” presented at the CIHS–CIS Summer Symposium 2023, hosted by the California Institute for Human Science and the Center for Consciousness Studies, University of Arizona, Encinitas, CA, USA, Aug. 18–20, 2023. [Online]. Available: https://www.youtube.com/watch?v=-uBfX6xxo_M
- [62] Daniel Rebbin, “Ephaptic Biases on Neural Self-Organization: A Project Idea,” presented at Allen Discovery Center, Tufts University, YouTube video, Feb. 29, 2024. [Online]. Available: https://www.youtube.com/watch?v=pZPQetQ_w4g

- [63] Ella P. K. Douglas, “Ephaptic Coupling,” presented at the Foresight Neurotech, BCI, and Whole Brain Emulation for Safe AI Workshop, Foresight Institute, Berkeley, CA, USA, May 21–22, 2024. [Online]. Available: <https://www.youtube.com/watch?v=2BtjZVLzuU0>
- [64] Yuhan Helena Liu, Stephen Smith, Stefan Mihalas, Eric Shea-Brown, and Uygur Sümbül, “Cell type specific neuromodulation guides synaptic credit assignment in a spiking neural network,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 51, Art. no. e2111821118, Dec. 2021.
- [65] M. Javaheripi, K. Wang, and A. B. Farimani, “SWNet: Small-world neural networks and rapid convergence,” *arXiv preprint*, arXiv:1904.04862, 2019.
- [66] Feng-Lei Fan, Ze-yu Li, Huan Xiong, and Tiejong Zeng, “Rethink depth separation with intra-layer links,” *arXiv preprint*, arXiv:2305.07037, 2023.
- [67] Feng-Lei Fan, Ze-yu Li, Huan Xiong, and Tiejong Zeng, “On expressivity of height in neural networks,” *arXiv preprint* arXiv:2305.07037, May 2023.
- [68] Marco Mercioni, Giovanni Catalano, and Ilaria Donati, “Dynamic modification of activation function using the backpropagation algorithm in artificial neural networks,” *International Journal of Advanced Computer Science and Applications*, 2019.
- [69] Timothy P. Lillicrap, Adam Santoro, Luke Marris, Colin J. Akerman, and Geoffrey Hinton, “Backpropagation and the brain,” *Nature Reviews Neuroscience*, vol. 21, pp. 335–346, 2020.
- [70] R. Ravichandran, A. Islam, V. Srinivasan, and B. Scellier, “Unsupervised representation learning with Hebbian synaptic and structural plasticity in brain-like feedforward neural networks,” *Neurocomputing*, 2024.
- [71] D. Chen, P. Peng, T. Huang, and Y. Tian, “Fully spiking actor network with intra-layer connections for reinforcement learning,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 36, pp. 2881–2893, 2024.
- [72] Luke Darlow, Charles Regan, Sebastian Risi, Jesse Seely, and Llion Jones, “Continuous Thought Machines,” *arXiv preprint*, arXiv:2505.05522, 2025.
- [73] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *J. Physiol.*, vol. 117, no. 4, pp. 500–544, 1952, doi: 10.1113/jphysiol.1952.sp004764.
- [74] C. H. Chen, “Spooky Action Potentials at a Distance: Ephaptic Coupling,” *Harvard Brain Science Initiative*, 2021. [Online]. Available: <https://brain.harvard.edu/people/spooky-action-potentials-distance-ephaptic-coupling>
- [75] K. S. Han, C. H. Chen, M. M. Khan, C. Guo, and W. G. Regehr, “Climbing fiber synapses rapidly and transiently inhibit neighboring Purkinje cells via ephaptic coupling,” *Nat. Neurosci.*, vol. 23, pp. 1372–1384, Sep. 2020, doi: 10.1038/s41593-020-0699-0.
- [76] A. C. Elitzur and E. Cohen, “Entanglement as an emergent phenomenon of self-organization,” *Physica A*, vol. 604, p. 127865, 2022, doi: 10.1016/j.physa.2022.127865.
- [77] Alexander R. Shifman, David F. Reilly, and Konrad P. Kording, “ELFENN: A generalized platform for modeling ephaptic coupling in spiking neuron models,” *Frontiers in Neuroinformatics*, 2019.
- [78] Kurt Hornik, Maxwell Stinchcombe, and Halbert White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [79] Albert Einstein, “Zur Elektrodynamik bewegter Körper [On the Electrodynamics of Moving Bodies],” *Annalen der Physik*, vol. 17, no. 10, pp. 891–921, 1905. [In German].
- [80] Albert Einstein, “Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig? [Does the Inertia of a Body Depend Upon Its Energy Content?],” *Annalen der Physik*, vol. 18, pp. 639–641, 1905. [In German].
- [81] Stuart Hameroff and Roger Penrose, “Consciousness in the universe: A review of the ‘Orch OR’ theory,” *Phys. Life Rev.*, vol. 11, no. 1, pp. 39–78, Mar. 2014, doi: 10.1016/j.pprev.2013.08.002.
- [82] G. Zhi and R. Xiu, “Quantum Theory of Consciousness,” *Journal of Applied Mathematics and Physics*, vol. 11, pp. 2652–2670, Sept. 2023, doi: 10.4236/jamp.2023.119174.
- [83] Giulio Tononi, “An information integration theory of consciousness,” *BMC Neuroscience*, vol. 5, Art. no. 42, 2004.
- [84] Warren S. McCulloch, *Televised Interview on Embodiments of Mind*, Canadian Broadcasting Corporation, 1969. [Online]. Available: https://www.youtube.com/watch?v=Xe87_ZWObQs
- [85] Robert M. Harnish, *Minds, Brains, Computers: An Historical Introduction to the Foundations of Cognitive Science*. Oxford, U.K.: Wiley-Blackwell, 2000.
- [86] DeepSeek-AI et al., “DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning,” *arXiv preprint* arXiv:2501.12948, 2025.

12. ACKNOWLEDGEMENTS

The author gratefully thanks his colleagues, especially Dr. Shiva Kumar, for their valuable feedback on the drafts of this paper and for independently conducting experiments to reproduce preliminary test results.

The author honors the late Dr. Frank Rosenblatt, whose seminal work on MLPs and more specifically on error correction, notably in his underappreciated book, *Principles of Neurodynamics*, prefigured modern backpropagation and demonstrated remarkable foresight, often overlooked in AI history. The author, who coincidentally shares a birthday with Dr. Frank Rosenblatt, has come to increasingly admire his pioneering contributions to ANN theory and draws enduring inspiration from his visionary legacy.

The author also honors his late Professor Robert M. Harnish [85], whose undergraduate college course (PSYC 346) and book, *Minds, Brains, and Computers*, helped spark a lasting interest in AI and neuroscience, leading to a short-lived but prescient robotics venture in 2003 and later to a career in telecommunications, anticipating the rise of intelligent mobile devices.

The author further reflects on the influence of the late Professor Claude Shannon, whose time at Bell Labs, then part of AT&T, informed the writing of the groundbreaking paper “A Mathematical Theory of Communication”. Although separated by decades, the author’s later work at AT&T in the 2010s deepened his appreciation for Shannon’s legacy. Shannon’s foundational work in information theory influenced the author’s guiding analogy between peer-to-peer mobile networks, where intelligent nodes interact directly, and biological cellular neural networks, where decentralized dynamics are intrinsic.

Finally, the author, having informally studied the ephaptic coupling mechanism since 2015, affirms a long-standing belief in its untapped potential as a computational framework for AI. A pivotal moment occurred in January 2025, when the official release of DeepSeek [86] and its implications prompted the author to revisit and re-evaluate his earlier work privately, with a renewed sense of urgency, in parallel with his day-to-day professional responsibilities.