

A Digital Neural Array IC for Real-Time Neural Network Replication from Spike Activities

Donghyun Park*, Hajung Mun[†], Minhyeok Jeong*, Dahee Kang[‡], Jongmin Lee[§], Yoonmyung Lee*

*Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Republic of Korea

[†]Memory Division, Samsung Electronics, Hwaseong, Republic of Korea

[‡]Department of Semiconductor and Display Engineering, Sungkyunkwan University, Suwon, Republic of Korea

[§]Department of Intelligence Semiconductor Engineering, Ajou University, Suwon, Republic of Korea
ehdqus0131@g.skku.edu, hiway2.mun@samsung.com, jongmin@ajou.ac.kr, yoonmyung@skku.edu

Abstract—Growing demand to deepen understanding of the human brain has accelerated efforts to identify the structure of biological neural networks from neuronal activities. This paper presents a fully digital, tile-able neural array integrated circuit (IC) that, to our knowledge, is the first hardware platform for network reconstruction—inferring synaptic connectivity directly from spike-train data generated by a biological (ground-truth) network. Designed with the overarching goal of emulation of biological networks, the architecture employs repeatable digital neuron-module tiles to ensure the scalability, flexibility and verifiability. Two chip-level run-time interfaces are integrated: a writable spike-forcing path for injecting biological spike pulses into selected IC neurons for synchronized co-firing, and a dedicated monitoring path for streaming spike events, synaptic weights, and membrane potentials. Scalability is further enabled by single-timer Δt capture and a piecewise-linear STDP Δw generator shared per neuron, avoiding complex LUTs/multipliers while preserving biological plausibility. The platform is realized in silicon and tested with an FPGA-based setup. Using only spike activity, the system reconstructs synaptic connectivity with high fidelity across diverse ground-truth networks ranging from simple 2-layer topologies to biologically derived *C.elegans* head network, as well as networks with bimodal and trimodal weight distributions. Accuracy was comprehensively quantified from multiple perspectives for both connectivity and spike-train similarity, confirming faithful replication. These results demonstrate that the proposed platform can recover the synaptic structure from spikes alone and provide a practical tool for predicting network learning responses under varied stimuli, advancing biological neural network research and real-time neuromorphic experimentation.

Index Terms—Network replication, Neuromorphic IC, digital tile-able neuron module, SNN, STDP, biological-plausibility

I. INTRODUCTION

In recent years, demand for AI inspired by the structure and operating principles of neurons and synapses in biological neural networks (BNNs) has surged. However, our understanding of BNNs remains limited, primarily due to challenges in accurately capturing signals from individual neurons. To address these limitations, recent studies have explored methods for recording electrical signals from BNNs [1]. Among them, a

This work was supported in part by Samsung Electronics, and in part by the National Research Foundation (NRF) of Korea (RS-2024-00408040, RS-2025-00562391, RS-2025-10032968) and IITP Korea (IITP-2025-RS-2020-I201821). The EDA tool was supported by the IC Design Education Center (IDEC), Korea. Jongmin Lee (jongmin@ajou.ac.kr) and Yoonmyung Lee (yoonmyung@skku.edu) are corresponding authors.

Harvard research team proposed a novel neural signal recording method using a nanoelectrode array [2]. They cultured BNNs on integrated circuits (ICs) to capture intracellular signals from multiple neurons, enabling detailed analysis of neuronal interconnections. Replicating this neural structure at the IC level would facilitate simulating neural circuit responses to new stimuli and serve as a high-dimensional analysis tool.

Although BNNs have not yet been replicated in hardware (e.g., as IC), there has been active work on software algorithms that infer neural network (NN) structures from spike-train data. Cross-correlograms (CCGs) have long served as a foundational method for such efforts, with one representative variant—the generalized linear-model-based CCG (GLMCC) [3]—designed to absorb large-scale background fluctuations. More recently, acknowledging that even GLMCC exhibits considerable performance variations depending on neuronal dynamics, the ensemble artificial neural network (eANN) method has been proposed [4]. By learning from validated outputs of multiple established algorithms, eANN predicts a multiclass connectivity and improves robustness, particularly for inhibitory synapses. Despite these advances, these algorithmic approaches remain limited to binary classification of connectivity (i.e., presence or absence of a synapse). They do not evaluate or quantify how accurately the synaptic weights of the reconstructed NN replicate (or infer) those of the ground-truth NN at the individual synapse level.

In contrast, at the hardware level, bio-inspired neuromorphic chips such as TrueNorth [5] and Loihi [6] have been actively studied. Platforms like TrueNorth and Neurogrid [7] deliver high-throughput spike processing but lack any intrinsic learning algorithms. Similarly, a recently reported system [8] emphasizes faithful emulation of biological features—such as Excitatory/Inhibitory postsynaptic potentials (EPSPs/ IPSPs), refractory periods, and lateral inhibition—yet omits any embedded learning engine. While Loihi incorporates spike-timing-dependent plasticity (STDP) to support on-chip learning, no prior research has directly targeted the replication or inference of the synaptic connectivity of NNs at the hardware level.

In this work, we effectively replicate ground-truth NNs, or BNNs, without any external teaching signals, by employing unit neuron modules that incorporate various biologically plau-

sible mechanisms, including leaky integrate-and-fire(LIF) with EPSPs, IPSPs and STDP. To ultimately enable the real-time replication of BNNs from spiking activity alone, the proposed neural array architecture was implemented as a silicon IC, enabling direct connection with cultured biological neurons. However, current technological limitations hinder the acquisition of complete synaptic weight and spike information from real BNNs, necessitating alternative benchmarking strategies. Following approaches adopted in prior NN reconstruction algorithm studies [4] [3] [9], we constructed pseudo-biological NNs to serve as ground-truth references, treating them as proxies for actual BNNs. Using spike activity as the sole input, we assessed the feasibility of replicating these reference NNs. The fidelity of reconstruction was evaluated by comparing the synaptic weights and spike-train activities of the replicated (destination) NN against those of the ground-truth NN.

Consequently, we demonstrated that the proposed circuit successfully replicates not only simple two-layer NNs exhibiting bimodal synaptic weight distributions [10] [11], but also more complex cases such as trimodal distributions. Furthermore, we replicated a more complex network based on the *C. elegans* head NN [12], converted into a bimodal distribution through thresholding. Replication accuracy was quantitatively assessed through microscopic verification of synaptic connectivity using normalized mean absolute error (NMAE) scores, which confirmed high similarity to the ground-truth. Additionally, a comprehensive evaluation was conducted from multiple perspectives, including macroscopic connectivity metrics as well as microscopic and macroscopic analysis of spike-activity similarity. Collectively, these results thoroughly validate the accuracy of the NN replication.

The proposed neural array architecture is hierarchically organized into three primary components: (1) a tile-able unit neuron module integrating key biological features with on-chip learning capability; (2) a scalable top-level array; and (3) a dedicated interface supporting real-time spike forcing and NN state monitoring. The subsequent sections detail the unit neuron module (Section II), followed by the chip-level architecture and its interfaces (Section III).

II. TILE-ABLE UNIT NEURON MODULE DESIGN

Fig. 1 illustrates the concept and simplified block diagram of the unit neuron module comprising our neural array. Each neuron module is capable of forming synaptic connections with multiple neighboring neurons and integrates essential neuronal primitives - LIF, STDP, refractory controls, lateral inhibition, and synaptic delays.

Because biological spikes occur on a millisecond scale, this work employs two distinct clock domains as shown in Fig. 2. The spike clock (CLK_SP) runs at 8kHz, providing 0.125ms spike-pulse resolution, while the internal operation clock (CLK_OP) runs at 50MHz. This dual clock domain approach lets CLK_OP process every presynaptic event within a single spike window and time-share hardware, yielding an area-efficient design. In the block diagram of Fig. 1, dark-gray sections (in-/out-spike buffers, spike generator and STDP timer)

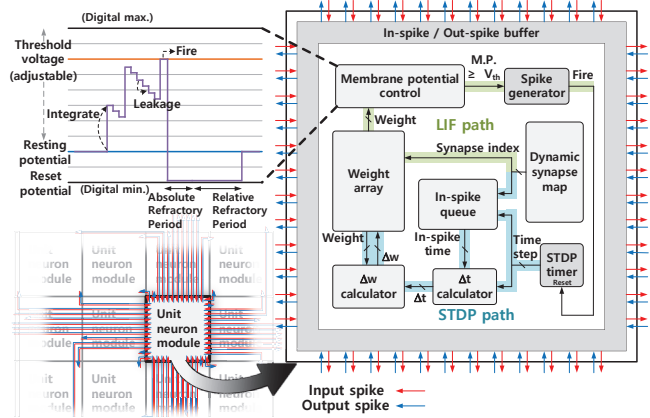


Fig. 1: Local connection and Block diagram of the unit neuron module with digital membrane potential control

operate at the spike timescale, whereas the light-gray sections execute rapid computations within a single spike timestep.

A. Membrane Potential Control

The proposed digital neuron module implements biologically plausible membrane potential dynamics using a digital LIF neuron model, as illustrated in Fig. 1 top left. The membrane potential (MP) is digitally represented using a signed 10-bit resolution, capturing levels for resting, threshold, and inhibitory potentials. When presynaptic spikes arrive, their corresponding synaptic weights are integrated into the MP. The module also supports lateral inhibition, wherein the membrane potential is temporarily suppressed upon receiving inputs at dendrites pre-selected for this function, preventing laterally inhibited neuronal firing [13].

When the MP exceeds a configurable threshold, the neuron generates a spike event, resets its MP to a predefined level, and transmits the spike after an adjustable axonal delay to model biological propagation delays [14]. Furthermore, the neuron module incorporates biological refractory periods, separated into absolute and relative phases. During the absolute refractory period (ARP), incoming spikes have no influence on the MP. Conversely, spikes arriving during the relative refractory period (RRP) affect the MP only if their synaptic weight surpasses an adjustable threshold [15] [16].

Combining digital LIF behavior, selective integration, axonal delay, lateral inhibition, and dual-phase refractory periods, this neuron module provides a hardware-efficient and biologically accurate emulation of neuronal membrane potential dynamics.

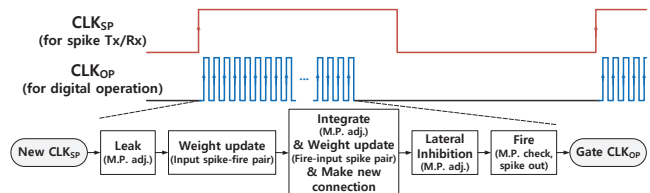


Fig. 2: Operation flow and clock gating during a single timestep

B. On-chip synaptic weight STDP learning with a single timer and piecewise linear curve generator

We implemented on-chip STDP to replicate BNNs using spiking activity data [17]. STDP learning requires two key computational steps: calculating the time interval between pre- and post-synaptic spikes (Δt) and subsequently determining the corresponding change in synaptic weight (Δw). To maintain both biological plausibility and hardware efficiency, Δw is determined through logic-based operations that reflect existing synaptic strengths, rather than using resource-intensive lookup tables or multipliers. This approach allows efficient hardware implementation while accurately mimicking biological synaptic plasticity.

For efficient hardware implementation of STDP, each neuron module incorporates only a single 8-bit timer to measure spike timing differences (Δt). This timer resets and then increments at every neuron firing event, immediately providing $\Delta t_{post-to-pre}$ upon input spike arrival. To obtain $\Delta t_{pre-to-post}$, input spike timestamps are stored in a queue at the timer value marked by red circles in Fig. 3, and $\Delta t_{pre-to-post}$ is subsequently computed upon neuron firing. If no input spike arrives within a time window, the timer is clock gated and resumes operation once a new input spike or forced spike event, i.e. locally emulated spike to represent BNN’s spiking activity, occurs. This single-timer approach significantly reduces circuit complexity and hardware overhead, enabling scalable neural array implementations.

To calculate Δw from Δt , the most common and intuitive approach is using a Look-Up Table (LUT) [18] [19]. However, LUT-based implementations require extensive hardware resources to store all STDP curve points, especially considering the biological diversity in STDP curves across neuronal regions and connection types [20] [21]. Additionally, biologically plausible STDP demands Δw values to be computed based on the current synaptic weight, necessitating hardware multipliers and thus further increasing circuit complexity and area overhead.

To overcome these limitations, we propose using a single, shared Δw calculation block per neuron, featuring a piecewise linear curve generator as depicted in Fig. 4. This significantly reduces hardware area and allows direct accounting of existing weight values. It also allows optional introduction of stochastic variations via a Linear Feedback Shift Register (LFSR), enhancing biological plausibility. This noise component provides stochastic behavior in weight updates, closely resembling

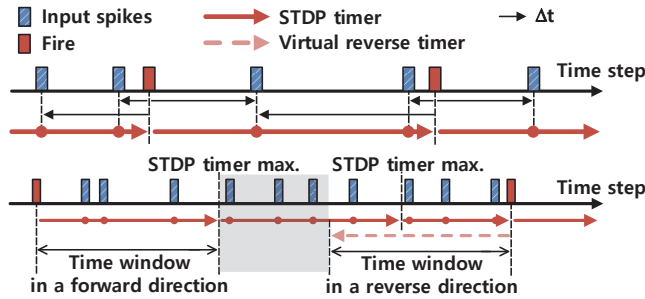


Fig. 3: Illustration of the STDP timer, which is used to calculate the interval between spikes

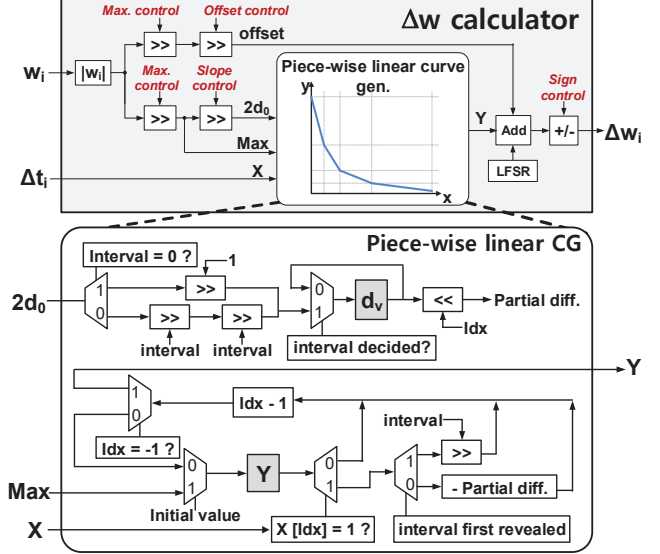


Fig. 4: Δw calculator for the i -th synapse. d_0 represents the decrement per timestep in the first segment of the piecewise linear function.

synaptic fluctuations observed in biology. As shown in Fig. 5, the proposed design can generate up to 16k distinct STDP curves through four tunable hyperparameters (max value, slope, offset, and sign), facilitating independent curves for excitatory and inhibitory synapses. This approach significantly improves both the flexibility and biological fidelity of the neural module.

In summary, our configurable STDP design effectively balances biological fidelity and hardware efficiency. By leveraging shared computation logic and hyperparameter-driven STDP curve generation, the system achieves versatile synaptic learning with minimal hardware resource demands. This architecture not only preserves biological plausibility but also supports scalable deployment in large neuromorphic systems.

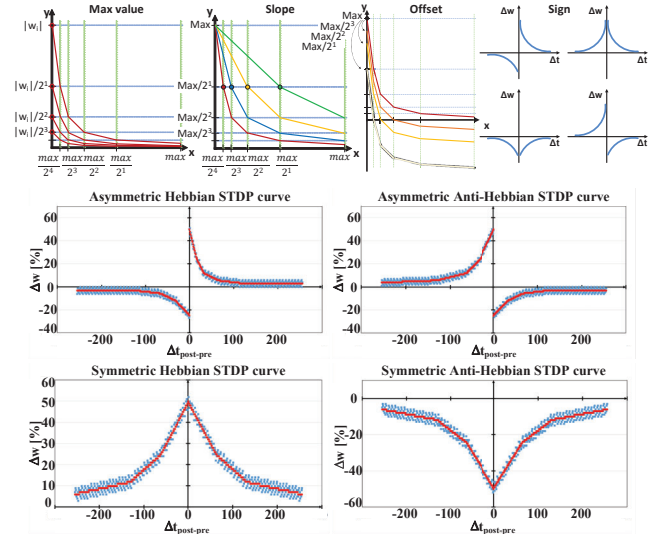


Fig. 5: STDP curve parameters(top) and Examples of various STDP curves generated by the Δw generator block; red indicates curves without noise, and blue with noise(bottom).

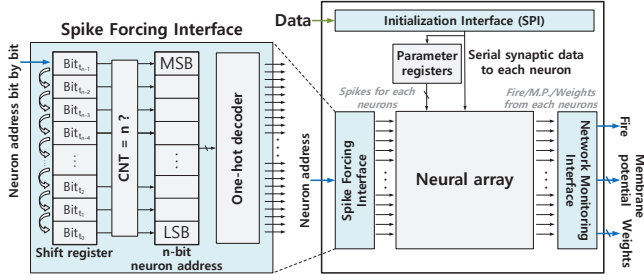


Fig. 6: Top level block diagram including Spike forcing interface for mapping spike activity data from biological neural network

III. CHIP LEVEL ARCHITECTURE

The proposed top-level neural array architecture consists of tiled digital neuron modules, enabling scalable implementations that can be tailored to various neural network configurations. As depicted in Fig. 6, the architecture integrates multiple neuron modules and dedicated interfaces that support both initialization and run-time operations.

During the initialization phase, common parameters and neuron-specific synaptic connections are configured using a Serial Peripheral Interface (SPI). These parameters contain various switching signals for STDP, integration, refractory period, and other functionalities, as well as basic parameters such as leakage and threshold values for LIF models, weight thresholds for RRP, and tracking windows for STDP.

In the subsequent operation phase, neuron modules can perform spike-driven learning and real-time neural processing. To replicate BNNs precisely, real-time spike injection into individual neuron modules is supported, as illustrated in Fig. 6 left. A serial input interface efficiently manages forced spike events using address decoding and minimal hardware pins, ensuring accurate synchronization with the neural network timing.

Additionally, dedicated monitoring interface is implemented to allow continuous monitoring of neuron activities by observing membrane potentials, synaptic weights, and spike events. This real-time monitoring capability facilitates comprehensive network analysis and debugging during runtime. The combined functionalities of scalable neuron modules, efficient spike interfaces, and comprehensive monitoring make the proposed architecture highly effective for real-time BNN replication, and this top-level architecture has been realized in a silicon chip and measured via FPGA as shown in Fig. 7.

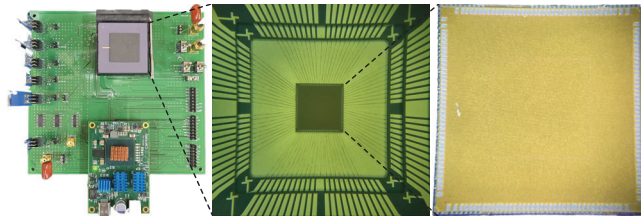


Fig. 7: Chip measurement setup with FPGA(left) and chip packaging(middle) with die photo (right)

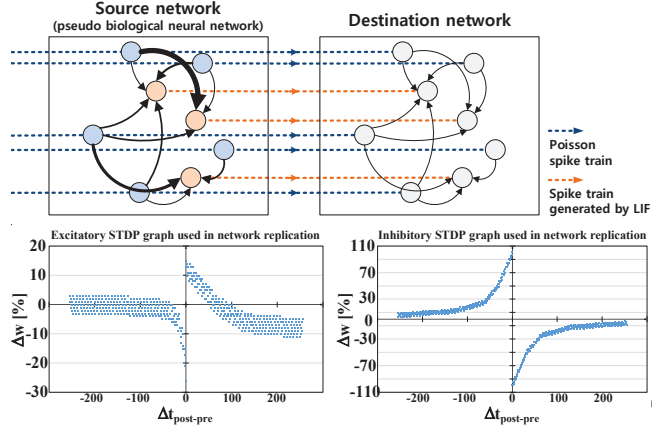


Fig. 8: Verification Strategy(top) and STDP graph used for training destination network(bottom)

IV. CHIP MEASUREMENT RESULT

A. Verification Strategy

Fig. 8, illustrates our verification strategy employed to validate the replication capability of our neural array. The source network generates postsynaptic spikes using a LIF mechanism driven by Poisson-distributed input spikes; these spikes then serve as input for STDP-based learning in the destination network, whose weights were randomly initialized within a narrow mid-range. After learning over 400k timesteps, the synaptic weight distributions of the destination and source networks are statistically compared to evaluate replication accuracy.

B. Reconstructed Networks

We first performed validation using a simplified 2-layer neural network consisting of 1,210 pre-neurons and 250 post-neurons, without intra-layer connections. For this initial validation, the synaptic weights in the source network were configured using two distinct distributions: a bimodal distribution, where half the weights were set at the minimum value and half at the maximum value; and a trimodal distribution, where weights were assigned equally among three distinct levels: minimum (1), midpoint (256), and maximum (511). The STDP curve utilized for network reconstruction is presented at the bottom of Fig. 8, and the resulting reconstructed 2-layer networks for both bimodal and trimodal weight distributions are shown on the left and middle of Fig. 9.

To further validate our neural array, we replicated the head neural network of *C.elegans*, known for its relatively simple yet functionally significant structure. Since the pivotal study by [22] in 1986, various works have anatomically analyzed the network and inferred synaptic weights based on synapse counts [23] [24]. However, a recent study provided functional signal propagation measurements that differed from predictions based solely on anatomical synapse counts. Utilizing the calcium response data and statistically significant connections (q -values < 0.05) reported by [12], we established synaptic weights by separately thresholding excitatory and inhibitory synapses into bimodal distributions. Consequently, we constructed a *C.elegans* head network featuring four distinct synap-

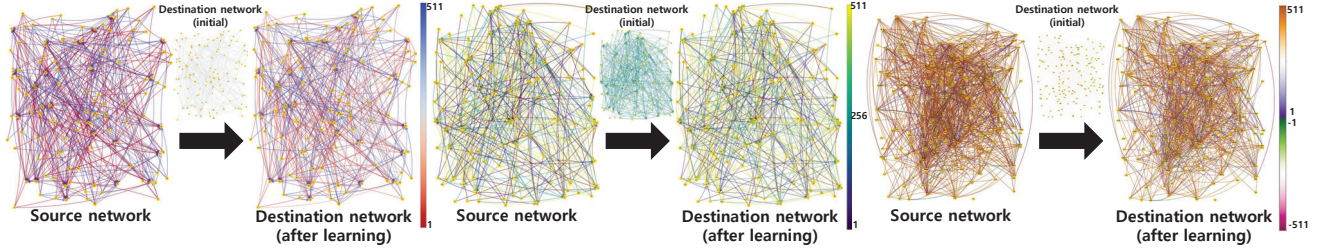


Fig. 9: Network reconstruction result of 2-layer network with bimodal(left) and trimodal(middle) and C.elegans Head network(right)

tic weight distributions (namely -512, -1, 1, 511), accurately reflecting functional connectivity characteristics. The utilized eSTDP and iSTDP curves are presented in Fig. 8 bottom, and the resulting reconstructed C.elegans network is shown in Fig. 9 right.

Fig. 10 presents box plots of the post-training weight distributions for the 2-layer and the C.elegans network after 400k timesteps. The central orange line within each box represents the median value, and shorter boxes indicate a more concentrated distribution. Despite a few outliers, synaptic weights consistently diverge from their initial midpoint values (256 for excitatory synapses and -256 for inhibitory synapses) toward their respective targets, indicating effective reconstruction.

C. Reconstruction Accuracy Evaluation

In Fig. 9-10, we visually demonstrated that our digital neural array chip successfully infers the synaptic connectivity of a pseudo-biological network using only spike pulse data, without direct access to its internal structure. To further quantitatively verify the accuracy of our network reconstruction, additional evaluations were conducted from multiple perspectives.

Firstly, following previous network inference studies [3] [4], we evaluated synaptic connectivity accuracy using the averaged precision score (APS) and Matthews correlation coefficient (MCC) [25]. The confusion matrix for calculating APS and MCC was constructed by defining synapses with weights of ± 1 as negative(absence) and ± 511 as positive(presence). As APS and MCC inherently assess binary connectivity classification, they were applicable only to the bimodal weight distribution in the 2-layer network. Although the overall weight distribution in the C. elegans head network is quad-modal, we treated excitatory and inhibitory synapses separately as bimodal distributions to compute individual APS and MCC values. By subsequently combining these values using weight averaging, we achieved an overall APS of 97.83% (excitatory: 98.89%,

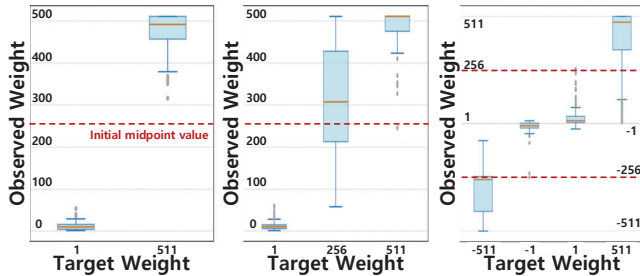


Fig. 10: Box plot for all networks(left for 2 layer bimodal, middle for 2 layer trimodal, right for C.elegans Head network)

inhibitory: 84.58%) and MCC of 73.79% (excitatory: 72.91%, inhibitory: 84.76%) for the C. elegans head network.

Since our primary objective was precise replication of ground-truth synaptic weights rather than mere binary connectivity matching, we introduced the Normalized Mean Absolute Error Score (NMAE-score) as a microscopic verification metric. For each synapse i , the score is defined as:

$$\text{score}_i = 1 - \frac{|w_i^{\text{pred}} - w_i^{\text{true}}|}{W_{\text{max}} - W_{\text{min}}} \quad (1)$$

where w_i^{pred} and w_i^{true} represent the synaptic weights of the reconstructed network and ground-truth network, respectively. An overall NMAE-score was calculated as the average of the individual scores across all synapses. Fig. 11 (left) shows the NMAE score progression for the three networks over the training time. After 400k timesteps, the networks achieved high final NMAE scores of 95.64%, 89.66%, and 89.56%, respectively.

With APS, MCC and NMAE-score, we have evaluated how accurately the destination network replicates the synaptic weights of the source network from both macroscopic and microscopic perspectives. Beyond synaptic connectivity, we further assessed the similarity of spike activity between the source and reconstructed networks. Fig. 11 (right) illustrates raster plots of 5 randomly selected post-neurons from both the source and reconstructed networks, where green dots represent spikes simultaneously generated by both networks at identical timesteps (CLK_SP cycles). As with connectivity, spike activity was evaluated using both macroscopic and microscopic methods. For macroscopic evaluation, we quantified the overall temporal alignment between reference and reconstructed spike trains by computing the Pearson correlation coefficient. We discretized time into M bins of timestep and computed the total spike count in each bin for both spike trains. This yielded two sequences, $\{x_i\}$ and $\{y_i\}$, for $i = 1, \dots, M$. The Pearson correlation coefficient (r) was then calculated as:

$$r = \frac{\sum_{i=1}^M (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^M (x_i - \bar{x})^2 \sum_{i=1}^M (y_i - \bar{y})^2}}, \quad (2)$$

where \bar{x} and \bar{y} represent the mean values of $\{x_i\}$ and $\{y_i\}$, respectively. This coefficient ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no correlation. Despite the conservative bin size of $\Delta t = 1$ timestep, 2-layer network demonstrated high correlation coefficients of 87.44% for bimodal and 74.08% for trimodal distributions.

TABLE I: Comparison with prior HW and SW studies

		This work	HW implemented but no NW inferencing				NW inferencing but no HW implemented			
			FNINS 2025 [8]	Loihi [6]	Neurogrid [7]	JSSC 2019 [18]	PLOS C. Bio. 2024 [4]	Nat. Com. 2019 [3]	J. Comp. Ns 2015 [9]	
HW implementation		O (Silicon)	O (Silicon)	O (Silicon)	O (Silicon)	O (Silicon)	X (Simulated)	X (Simulated)	X (Simulated)	
HW Methodology		Tile-able Digital	MRAM Mixed	SRAM Digital	Multichip Mxed	SRAM Digital	-	-	-	
Technology		28nm	28nm	14nm FinFET	180nm	10nm FinFET	-	-	-	
Neuron model		LIF	LIF	LIF	Adaptive Quad IF	LIF	LIF	HH, LIF	GLM, LIF	
Learning Algorithm		STDP	X	STDP	X	STDP	Ensemble ANN (eANN)	GLMCC	MLE(Gradient ascent)	
Iterative learning Required		Not required	X	Not required	X	Not required	Not required	Required	Required	
Network reconstruction		O	X	X	X	X	O	O	O	
Connectivity type		Exc, Inh	Exc, Inh	Exc, Inh	Exc, Inh	Exc	Exc, Inh	Exc, Inh	Exc, Inh	
Reconstructed network scale (# neurons / synapses)		2-layer (1210 / 3750) & C.elegans head (168 / 658)	-	-	-	-	Random LIF	HH-based sim., Rat hippocampal	Randomly balanced, Synfire Chain	
# neurons / synapses		225 / 3375	256 / 4k	128k / 128M	65k / 100M	4k / 1M	**300 / 4.5k	**1000 / 150k	**1000 / 200k	
Area [mm ²]		4	7.74	60	168	1.72	-	-	-	
*Area per neuron [28nm] normalized [mm ²]		0.018	0.03	0.002	0.000062	0.0033	-	-	-	
Clock frequency		GCLK : 8k LCLK : 50M	GCLK : 1k-50k LCLK : N.A.	Event-driven async	Real-time	105M-506M	-	-	-	
Clock frequency adaptability		Real-time acc.	Real-time acc.	Real-time acc.	Real-time	Accelerated	-	-	-	
Energy [fJ] normalized [pJ/SOP]		***85.3	N.A.	41.9(min)	119	13.8	-	-	-	
Real-time interaction with bio-network		O	X	X	X	X	-	-	-	
***Monitor-able data		SA, MP, SW	SA, MP	SA, SW	SA, MP, GV	SA, SW	-	-	-	
SW metrics		Recon. Accuracy (Synapse connectivity)			NMAE score (Microscopic verification)			Binary classification (Macroscopic verification)		
		2 yr bi			2 yr tri			C.elegans		
		95.64%			89.66%			89.56%		
		99.7% (APS)			97.83% (APS)			80-93% (APS)		
99.5% (MCC)			73.79% (MCC)			75-87% (MCC)				
99.5% (MCC)			73.79% (MCC)			~70% (MCC)				
Misclassification Error Rate < 1% (ideal conditions)										
SW metrics		Recon. Accuracy (Spike activity)			ISI-distance (Microscopic verification)			Pearson Correlation Coefficient (Macroscopic verification)		
		2 yr bi			2 yr tri			2 yr bi		
		0.0876			0.1912			0.0876		
		2 yr bi			2 yr tri			2 yr bi		
87.44%			74.08%			87.44%				
Biologically plausible Characteristics		EPSP, IPSP, Refractory period, Axonal delay, STDP, Lateral inhibition	EPSP, IPSP, Refractory period, Lateral inhibition	EPSP, IPSP, Refractory period, Axonal delay, STDP, Lateral inhibition	EPSP, IPSP, Refractory period, Synaptic delay, Lateral inhibition, ion-channel dynamics	EPSP, STDP, Lateral inhibition	EPSP, IPSP	EPSP, IPSP, Synaptic delay	EPSP, IPSP, Refractory period, Synaptic delay	

*Estimated from reported die area divided by neuron count. **Synapse counts were inferred from reported connection probabilities or sparsity levels. ***Dynamic energy. ****SA : Spiking Activity, MP : Membrane Potential, SW : Synaptic Weight, GV : Gating Variable

However, since the Pearson correlation coefficient assesses linear similarity at the macroscopic level by considering overall firing rate and synchronicity, an additional metric for microscopic evaluation was adopted for detailed spike-train evaluation: parameter-free inter-spike-interval (ISI) distance metric [26]. For each post neuron n , let $ISI_{ref}^{(n)}(t)$ and $ISI_{meas}^{(n)}(t)$ represent the instantaneous inter-spike intervals of the two spike trains at time t . The ISI-distance for neuron n is defined as:

$$D_{ISI}^{(n)} = \frac{1}{T} \int_0^T \frac{|ISI_{ref}^{(n)}(t) - ISI_{meas}^{(n)}(t)|}{\max\left(ISI_{ref}^{(n)}(t), ISI_{meas}^{(n)}(t) \right)} dt \quad (3)$$

where T is the total observation time. The Mean ISI-distance is calculated by averaging the ISI-distances of each neuron across the entire network. An ISI-distance value of 0 indicates perfect similarity in spike timing, while values closer to 1 indicate greater dissimilarity. In our experiments, the 2-layer

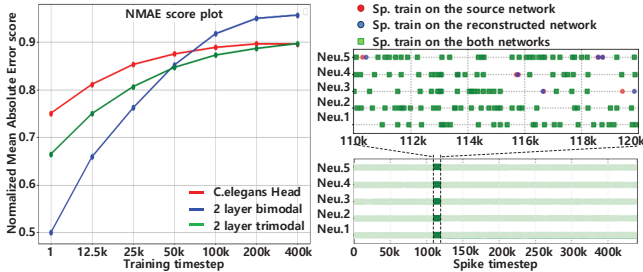


Fig. 11: NMAE score plot for connectivity verification(left) and spike raster plot used for spike activity verification(right).

network achieved mean ISI-distances of 0.0876 (bimodal) and 0.1912 (trimodal), confirming high temporal similarity [26].

Table I compares our digital neural array architecture with prior studies. Because this is, to our knowledge, the first silicon implementation of network reconstruction, a like-for-like comparison was not possible. Therefore, we compared against (1) neuromorphic hardware studies lacking network-inference capability and real-time spike injection/monitoring interfaces, and (2) algorithmic network-inference studies evaluated solely in simulations without hardware realization.

Unlike previous memory-centric hardware optimized for integration, our architecture adopts a tile-able structure, ensuring each neuron module internally stores all data required for its operations. Although incorporating multiple biologically plausible features-absent in prior hardware study-can increase area, we successfully maintained comparable area efficiency by area-optimized strategies such as time-division processing and a piecewise linear STDP curve generation. Critically, the chip uniquely provides a real-time interface to BNNs for spike-pulse injection-enabling synchronized firing-and runtime monitoring of internal neuron module states.

V. CONCLUSION

In this work, we presented a fully digital neural array realized in silicon that enables real-time emulation and replication of BNNs. The scalable tile-able architecture incorporates a configurable piecewise-linear STDP engine and reconstructs network connectivity from spike-train data with high, comprehensively verified accuracy. This platform provides a practical foundation for BNN studies and real-time neural processing.

REFERENCES

- [1] J.-H. Cha, J.-H. Park, Y. Park, H. Shin, K. S. Hwang, I.-J. Cho, and S.-J. Kim, "A cmos microelectrode array system with reconfigurable sub-array multiplexing architecture integrating 24,320 electrodes and 380 readout channels," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 16, no. 6, pp. 1044–1056, 2022.
- [2] J. Abbott, T. Ye, K. Krennek, R. S. Gertner, S. Ban, Y. Kim, L. Qin, W. Wu, H. Park, and D. Ham, "A nanoelectrode array for obtaining intracellular recordings from thousands of connected neurons," *Nature biomedical engineering*, vol. 4, no. 2, pp. 232–241, 2020.
- [3] R. Kobayashi, S. Kurita, A. Kurth, K. Kitano, K. Mizuseki, M. Diesmann, B. J. Richmond, and S. Shinomoto, "Reconstructing neuronal circuitry from parallel spike trains," *Nature communications*, vol. 10, no. 1, p. 4468, 2019.
- [4] C. Donner, J. Bartram, P. Hornauer, T. Kim, D. Roqueiro, A. Hierlemann, G. Obozinski, and M. Schröter, "Ensemble learning and ground-truth validation of synaptic connectivity inferred from spike trains," *PLoS Computational Biology*, vol. 20, no. 4, p. e1011964, 2024.
- [5] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, B. Brezoz, I. Vo, S. K. Esser, R. Appuswamy, B. Taba, A. Amir, M. D. Flickner, W. P. Risk, R. Manohar, and D. S. Modha, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014. [Online]. Available: <https://www.science.org/doi/abs/10.1126/science.1254642>
- [6] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. H. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, Y. Liao, C.-K. Lin, A. Lines, R. Liu, D. Mathakutty, S. McCoy, A. Paul, J. Tse, G. Venkataramanan, Y.-H. Weng, A. Wild, Y. Yang, and H. Wang, "Loihi: A neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, 2018.
- [7] B. V. Benjamin, P. Gao, E. McQuinn, S. Choudhary, A. R. Chandrasekaran, J.-M. Bussat, R. Alvarez-Icaza, J. V. Arthur, P. A. Merolla, and K. Boahen, "Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 699–716, 2014.
- [8] J. Lee, J. Song, H. S. Im, J. Kim, W. Lee, W. Yi, S. Kwon, B. Jung, J. Kim, Y. Lee *et al.*, "A scalable neural network emulator with mram-based mixed-signal circuits," *Frontiers in Neuroscience*, vol. 19, p. 1599144, 2025.
- [9] Y. V. Zaytsev, A. Morrison, and M. Deger, "Reconstruction of recurrent synaptic connectivity of thousands of neurons from simulated spiking activity," *Journal of computational neuroscience*, vol. 39, no. 1, pp. 77–103, 2015.
- [10] P. A. De Abreu Urbizagastegui, A. van Schaik, and R. M. Wang, "Memory-efficient neurons and synapses for spike-timing-dependent-plasticity in large-scale spiking networks," *Frontiers in Neuroscience*, vol. 18, p. 1450640, 2024.
- [11] S. Song, K. D. Miller, and L. F. Abbott, "Competitive hebbian learning through spike-timing-dependent synaptic plasticity," *Nature neuroscience*, vol. 3, no. 9, pp. 919–926, 2000.
- [12] F. Randi, A. K. Sharma, S. Dvali, and A. M. Leifer, "Neural signal propagation atlas of caenorhabditis elegans," *Nature*, vol. 623, no. 7986, pp. 406–414, 2023.
- [13] V. Egger, K. Svoboda, and Z. F. Mainen, "Mechanisms of lateral inhibition in the olfactory bulb: efficiency and modulation of spike-evoked calcium influx into granule cells," *Journal of Neuroscience*, vol. 23, no. 20, pp. 7551–7558, 2003.
- [14] M. Madadi Asl, A. Valizadeh, and P. A. Tass, "Dendritic and axonal propagation delays may shape neuronal networks with plastic synapses," *Frontiers in physiology*, vol. 9, p. 1849, 2018.
- [15] M. J. Berry and M. Meister, "Refractoriness and neural precision," *Journal of Neuroscience*, vol. 18, no. 6, pp. 2200–2211, 1998. [Online]. Available: <https://www.jneurosci.org/content/18/6/2200>
- [16] E. N. Marieb and K. Hoehn, *Human anatomy & physiology*. Pearson education, 2007.
- [17] G.-q. Bi and M.-m. Poo, "Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type," *Journal of neuroscience*, vol. 18, no. 24, pp. 10464–10472, 1998.
- [18] G. K. Chen, R. Kumar, H. E. Sumbul, P. C. Knag, and R. K. Krishnamurthy, "A 4096-neuron 1m-synapse 3.8-pj/sop spiking neural network with on-chip stdp learning and sparse weights in 10-nm finfet cmos," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 4, pp. 992–1002, 2018.
- [19] Y. Zhong, X. Cui, Y. Kuang, K. Liu, Y. Wang, and R. Huang, "A spike-event-based neuromorphic processor with enhanced on-chip stdp learning in 28nm cmos," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2021, pp. 1–5.
- [20] N. Caporale and Y. Dan, "Spike timing-dependent plasticity: a hebbian learning rule," *Annual review of neuroscience*, vol. 31, pp. 25–46, 2008.
- [21] T. P. Vogels, R. C. Froemke, N. Doyon, M. Gilson, J. S. Haas, R. Liu, A. Maffei, P. Miller, C. Wierenga, M. A. Woodin *et al.*, "Inhibitory synaptic plasticity: spike timing-dependence and putative network function," *Frontiers in neural circuits*, vol. 7, p. 119, 2013.
- [22] J. G. White, E. Southgate, J. N. Thomson, S. Brenner *et al.*, "The structure of the nervous system of the nematode caenorhabditis elegans," *Philos Trans R Soc Lond B Biol Sci*, vol. 314, no. 1165, pp. 1–340, 1986.
- [23] T. A. Jarrell, Y. Wang, A. E. Bloniarz, C. A. Brittin, M. Xu, J. N. Thomson, D. G. Albertson, D. H. Hall, and S. W. Emmons, "The connectome of a decision-making neural network," *science*, vol. 337, no. 6093, pp. 437–444, 2012.
- [24] S. J. Cook, T. A. Jarrell, C. A. Brittin, Y. Wang, A. E. Bloniarz, M. A. Yakovlev, K. C. Nguyen, L. T.-H. Tang, E. A. Bayer, J. S. Duerr *et al.*, "Whole-animal connectomes of both caenorhabditis elegans sexes," *Nature*, vol. 571, no. 7763, pp. 63–71, 2019.
- [25] B. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) - Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0005279575901099>
- [26] T. Kreuz, J. S. Haas, A. Morelli, H. D. Abarbanel, and A. Politi, "Measuring spike train synchrony," *Journal of Neuroscience Methods*, vol. 165, no. 1, pp. 151–161, 2007. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0165027007002671>