Integer Quadratic Integrate-and-Fire (IQIF): *A Neuron Model for Digital Neuromorphic Systems*

Wen-Chieh Wu[1#], Chen-Fu Yeh[2#], Alexander James White[2], Cheng-Te Wang[2], Zuo-Wei Yeh[3], Chih-Cheng Hsieh[3], Ren-Shuo Liu[3], Kea-Tiong Tang[3], Chung-Chuan Lo[2,4*]*

#: contributed equally. *: corresponding authors
[1] Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan
[2] Institute of Systems Neuroscience, National Tsing Hua University, Hsinchu 30013, Taiwan
[3] Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan
[4] Brain Research Center, National Tsing Hua University, Hsinchu 30013, Taiwan
Contact information: cclo@mx.nthu.edu.tw, cheng-te@lolab-nthu.org

Abstract—**Simulation of a spiking neural network involves solving a large number of differential equations. This is a challenge even for modern computer systems, especially when simulating large-scale neural networks. To address this challenge, we design a neuron model: the Integer Quadratic Integrate-and-Fire (IQIF) neuron. Instead of computing on floating point numbers, as is typical with other spiking neuron models, the IQIF model is computed purely on integers. The IQIF model is a quantized and linearized version of the classic quadratic integrate-and-fire (QIF) model. The IQIF model retains all dynamic characteristics of the QIF model with much lower computation complexity, at the cost of a limited dynamic range of the membrane potential and the synaptic current. We compare IQIF to other spiking neuron models based on their simulation speeds and the number of neuronal behaviors they can perform. We further compare the performance of IQIF with the leaky integrate-and-fire model in a classical decision-making network that exhibits nonlinear attractor dynamics. Our results show that the IQIF neurons are capable of performing computation that other spiking neuron models can do while having the advantages of speed. Moreover, the IQIF model is digital hardware friendly due to its pure integer operation and is therefore easily to be implemented in custom-built neuromorphic systems.**

Keywords—SNN, Neuron model, integer, styling

I. Introduction

In computational neuroscience, dozens of spiking neuronal models have been proposed to account for experimental observations. Examples include leaky integrate-and-fire (LIF) model, quadratic integrate-and-fire (QIF) model, exponential integrate-and-fire model, Izhikevich model and Hodgkin-Huxley model [1]–[6]. In order to accurately reproduce the complex neuronal behavior, a model is usually described by ordinary differential equations and solving these equations are computationally intensive.

By contrast, in artificial intelligence, efficient computation is much more important than reproducing neuronal behavior. Therefore, the neuron models used in the artificial neural networks (ANN) are rather simple and usually non-spiking. Nevertheless, the advantages of computing with spiking neurons are still extensively discussed [7]–[9]. Owning to the potential of spiking neurons in ANN, several projects [10]–[14] on neuromorphic chips have also been carried out but with limited success primarily due to the complexity and cost of these systems.

One step toward solving this issue is to develop a neuron model which is endowed with complex behavioral characteristics while being digital-system friendly so that the development of specialized hardware are much easier. To this end, we design a neuronal model called integer quadratic integrate-and-fire (IQIF). The IQIF model is a quantized and linearized version of the classic QIF model. We choose the QIF model because it is the simplest spiking model that exhibits nonlinear dynamics, which is key to many complex neuronal behavior. Our IQIF model exhibits six neuronal behaviors and can be computed purely based on 8-bit integer. In this paper, we demonstrate the characteristics and efficiency of the model, and show that a IQIF network is able to carry out the decision-making task, one of the most fundamental functions of a biological neural network.

II. The Neuron Model and Its Basic Dynamics

The IQIF model is developed based on the QIF model, which is described by

$$C\frac{dV}{dt} = f(V) = K(V_r - V)(V - V_t) + I(t) \qquad (1)$$

Instead of having a simple linear $f(V)$ function as in the LIF model (Figure 1 top-left), the QIF has a quadratic $f(V)$ function (Figure 1 top-right). The QIF model is only slightly more complex than the LIF model but can exhibit more number of distinct behaviors due to its quadratic $f(V)$ function. We design the IQIF model by quantize and linearize the quadratic function as follows

$$\frac{\Delta V}{\Delta t} = f(V) = \begin{cases} a(V_r - V) + I(t), & V < \dfrac{bV_t + aV_r}{a+b} \\ b(V - V_t) + I(t), & V \geq \dfrac{bV_t + aV_r}{a+b} \end{cases} \qquad (2)$$

$$V \to V_{reset} \ when \ V > V_{max}$$

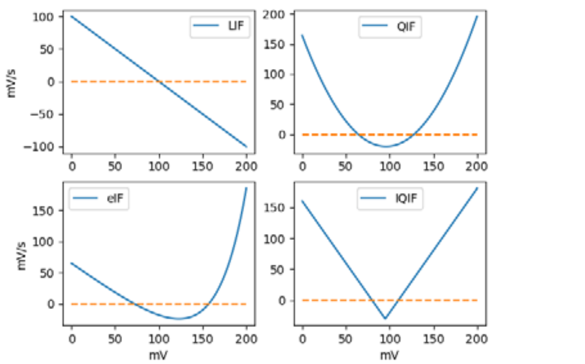

Figure 1. The dV/dt vs. V plots for four neuron models: LIF, QIF, EIF and IQIF. The intersections of the curve and the $y = 0$ line represent the equilibrium points of the models and the sign of the slope at each intersection indicate its stability. The LIF model has only one stable equilibrium point presenting the resting state, while the other three models have two equilibrium points with one stable and one unstable which represents the spike threshold.

The model has four main parameters a, b, V_r and V_t. The parameters a and b determine the slopes of the descending and rising part of $f(V)$(Figure 1 bottom-right), respectively. Due to the ability to independently adjust the two slopes, the behavior of IQIF is also similar to the exponential integrate-and-fire model (Figure 1 bottom-left). The parameter V_r is the resting potential that determines the stable equilibrium point of the membrane potential. The parameter V_t is the threshold potential that determines the unstable equilibrium point of the potential. The IQIF neuron fires an action potential when V_t is exceeded. After the maximum potential V_{max} is reached, the potential is reset to the reset potential V_{reset}. In the current version, we set the reset potential same with the resting potential V_r.

The parameter a plays the role similar to the leaky time constant in the LIF model. Specifically, a determines how fast the potential relaxes back to the resting potential (Figure 2). In contrast, a does not have a significant effect on the response curve of the model (Figure 3 top-left). The slope of the response curve, or the gain, can be adjusted by the parameter b (Figure 3 top-right). This is due to the positive effect of b on the rising speed of the membrane potential once the firing threshold is exceeded. This effect is stronger for larger input and hence determines the slope of the response curve. The parameters V_r and V_t have a similar effect on the response curve. They determine how much input is required to fire a neuron and therefore affect the value of the rheobase of the response curve (Figure 3 bottom-left and bottom-right). These two parameters do not significantly affect the slope of the response curve.

By understanding the effects of each parameter on the dynamics, we are able to generate a variety of neuronal behavior using IQIF. We discover that a single IQIF neuron is capable of producing six different behaviors, which include class 1 spiking

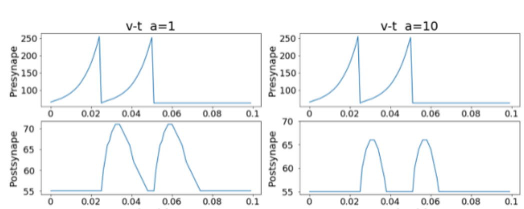

Figure 2. The membrane potential of a postsynaptic IQIF neuron (bottom) in response to the presynaptic input (top) for different values of the parameter a (left versus right). The membrane potential relaxes back to the resting state faster if a is larger.

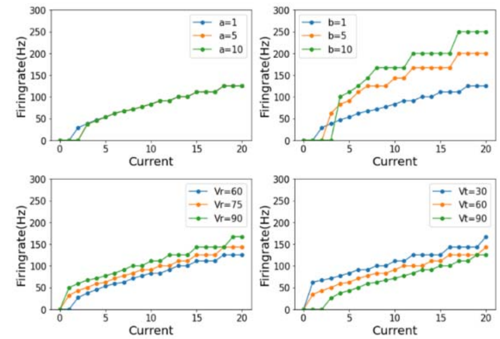

Figure 3. The response curve (the f-I curve) of the IQIF neurons. Top-left: the curve is relatively insensitive to the parameter a. Top-right: The gain, or the slope of the response curve, of the IQIF neuron can be changed by the parameter b. Bottom-left: the resting potential V_r does not affect the gain but can change the rheobase of the response curve. Bottom-right: the spike threshold V_t has an effect on the response curve similar to V_r

excitation, bistability, tonic spiking, burst spiking and delayed and integrator (Figure 4). By contrast, the classical LIF model exhibits only three behaviors. As a comparison, the Izhikevich model exhibits as many as 21 behaviors but with a cost of 13 floating point operations per timestep [4], [5] (Table 1). By contrast, IQIF model only needs six integer operations.

The dynamics of a spiking neural network is highly fluctuated due to the discrete spike events. This can be resolved by adopting the biologically realistic synaptic current, which typically exhibits an exponentially decayed form with a long time constant. This slow synaptic current is also crucial for forming working memory, a fundamental function of a neural network. Therefore, instead of using the commonly used short pulse, we implement the following form of the synaptic current I:

$$I(t) = g_s e^{\frac{t-t'}{\tau}}, \tag{3}$$

where g_s represents the synaptic weight, t' the presynaptic spike time and τ the time constant of the synaptic current. Due to the integer characteristic of IQIF, the exponential decay performed by timestep-wise division drops much faster than it should be. To address the issue, the division is performed every $\log(7/8) \div \log((\tau-1)/\tau)$ steps, and this produces a current decay close to the correct decay rate.

Next, we compare the simulation efficiency between IQIF and other neuron models. In non-parallel simulations, the run time grows in an $O(N^2)$ scale with N neurons in an all-to-all network. In parallel simulations using OpenMP, there is a small overhead when N is small but the run time improves dramatically with a large N. Comparing to other neuron models, IQIF runs significantly faster. Under single thread mode with N=100, IQIF (0.24s) is 7 times faster than the Izhikevich model (1.68s) and 5 times faster than float point LIF (1.21s) (Figure 5). In multi-thread parallelization, IQIF is still much faster than other float point models (Figure 5). We have also implemented a quantized version of the LIF model (ILIF) and compared it with the IQIF model. We find that IQIF runs as fast as ILIF despite that IQIF is capable of exhibition more complex behavior.

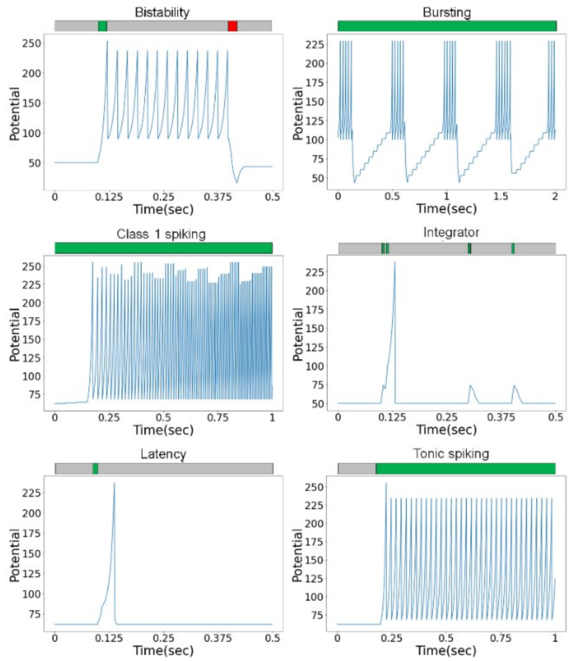

Figure 4. The six distinct behaviors exhibited by IQIF. The upper bar in each panel indicates the stimulus protocol with the blue and red bars for the onset of excitatory and inhibitory stimuli, respectively.

There is a drawback of the IQIF model the users should be aware of. Due to the nature of integer operation, the response curve forms staircases when short integers are used. This problem limits the usable range of the input strength and output firing rate (Figure 6). Our tests show that in order to obtain a smooth response curve, operating on 16-bit integers is preferable and an 8-bit system is the minimum. The working range for the firing frequency of the IQIF neurons is between 0 and 300 Hz.

III. Tests on Decision Network

To test the behavior of IQIF in neural networks, we construct a classical decision-making network, which consists of two excitatory populations (A and B) and one inhibitory population (C)(Figure 7). The function of this network is to make a decision on which of the two inputs (to the population A or B) is stronger. The classical decision network is characterized by attractor dynamics and exhibits winner-take-all competition, in which the two excitatory populations compete against each other through feedback and feedforward inhibition when receiving inputs. The decision is made when the firing rate of either A or B reaches a preset value (decision threshold). This type of models has been used in a wide range of neural networks for a variety of cognitive functions [15]–

Table 1. The number of behaviors performed by four neuron models, and the required number of operations (integer for IQIF and floating point for others) per iteration. LIF: leaky integrate-and-fire, EIF: exponential integrate-and-fire. IZH: Izhikevich. The numbers for LIF and IZH models are from [5].

Neuron Model	Firing pattern number	Number of operations per iteration
LIF	3	5
IQIF	6	6
EIF	6	10
IZH	21	13

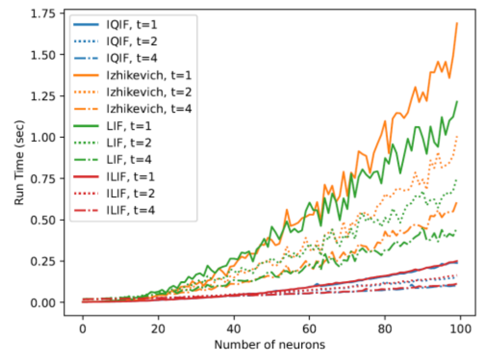

Figure 5. Run time of IQIF and other neuron models as a function of the number of neurons in an all-to-all network. The tests were performed with single thread (t=1), two threads (t=2) and four threads (t=4). The IQIF is faster than Izhikevich and LIF models and is comparable to quantized LIF (ILIF) in all conditions.

[18]. Moreover, the decision-making models reproduce empirical observations in many neurobiology studies and are thought to describe the underlying neural mechanisms of perceptual decision [19]–[24].

In the present study we only use one single excitatory neuron in each of A and B, and use three inhibitory neurons in C. The decision network constructed with IQIF neurons does successfully make decisions about the input strength (Figure 7 middle). For comparison, we perform simulations with a decision network made of LIF neurons with 50 neurons in each population (A, B and C) (Figure 7 bottom). Note that in order to increase the signal-to-noise ratio, we usually include dozens or hundreds neurons in each population. However, due to the limited dynamical range of IQIF neurons, a single neuron cannot receive too many inputs from highly activated neurons. Therefore, we limit the number of neurons in the decision network constructed by IQIF neurons. As show below, the IQIF neurons are still capable of exhibiting decent performance that is comparable to that constructed by a 150 LIF neurons.

We further evaluate the decision performance, which is usually done by plotting the psychometric functions: the percentage of correct choices and the reaction time versus the difference between the strengths of inputs (Figure 8). The IQIF decision network performs comparably with the LIF one. Although the IQIF network has slightly lower percentage of correct than the LIF network does, the reaction time of the IQIF network is much faster than the LIF network at the condition of equal input.

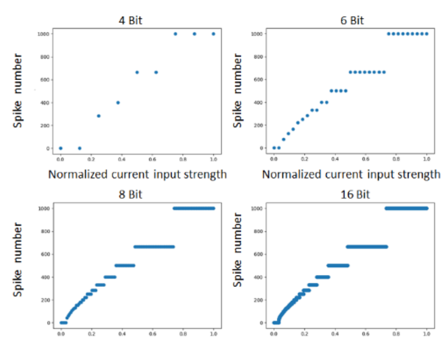

Figure 6. The response curve of the IQIF neurons with different integer bits. The IQIF neurons work well with integers of 8-bit or larger with a firing rate below 300 Hz

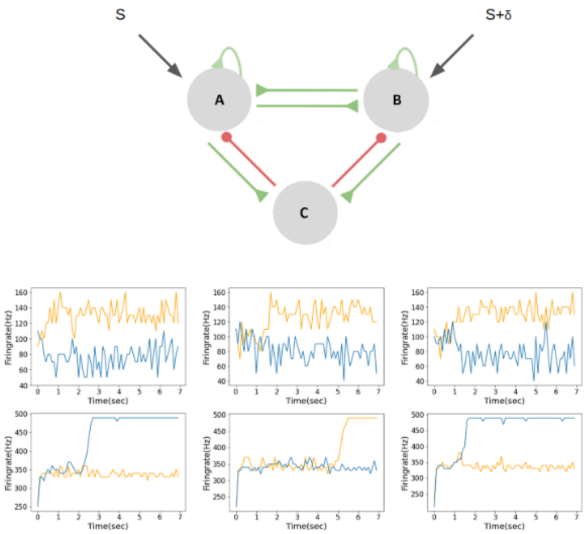

Figure 7. The decision-making task. Top: Schematics of the decision network. Green lines represent excitatory connections and red lines are inhibitory connections. The population A and B receive inputs of strength S and S+δ, respectively. Middle: The firing rates of neurons A (blue) and B (orange) in response to the inputs in three representative trials. The separation of the firing rates of A and B indicates that the decision is made. Bottom: Same as in the middle row but for the LIF network. There are 50 neurons in each population of A, B and C. Therefore, the firing rates appear smoother than the IQIF network, which has only one neuron in each of A and B, and three in C.

IV. Discussion

In this paper, we demonstrate that the IQIF model exhibits a number of distinct behaviors while does not require a large number of operations per timestep. The model operates purely on integers and therefore runs much faster than other spiking neuron models. One advantage of the IQIF model is its rich dynamics. One IQIF neuron can exhibit behavior that requires 2 or 3 LIF to perform. Therefore, we may be able to save the number of neurons in a network constructed by IQIF neurons. We will demonstrate this feature in follow-up studies. Furthermore, the IQIF model is digital hardware friendly and we have designed a prototype of neuromorphic chips that support the IQIF model [25]. The chip consists of two populations of IQIF neurons with configurable recurrent connections. The overall architecture comprise I/O devices (input/output stream buffers), a hierarchy-population scheduler, a top controller, and two neuromorphic processing units (NPUs), which contain 32 neurons and 128 IQIF neurons separately [25].

Further analysis will be conducted in follow-up studies to optimize the IQIF model and to compare its performance with other spiking neuron networks in various tasks.

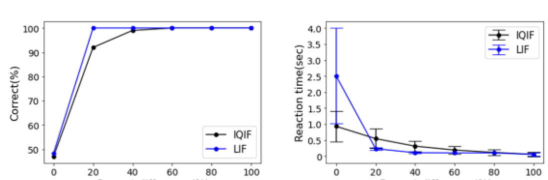

Figure 8. The psychometric functions of the decision task performed by decision networks made of IQIF or LIF neurons. Top: Accuracy (percentage of correct) as a function of the difference between the input currents. Bottom: Reaction time as a function of the difference between the input currents. IQIF neurons perform comparably with the LIF neurons

[1] W. Gerstner and W. M. Kistler, *Spiking neuron models: single neurons, populations, plasticity*. Cambridge, U.K.; New York: Cambridge University Press, 2002.

[2] "5.3 Quadratic Integrate and Fire | Neuronal Dynamics online book." https://neuronaldynamics.epfl.ch/online/Ch5.S3.html (accessed Jan. 22, 2021).

[3] N. Fourcaud-Trocmé, D. Hansel, C. van Vreeswijk, and N. Brunel, "How Spike Generation Mechanisms Determine the Neuronal Response to Fluctuating Inputs," *J. Neurosci.*, vol. 23, no. 37, pp. 11628–11640, Dec. 2003, doi: 10.1523/JNEUROSCI.23-37-11628.2003.

[4] E. M. Izhikevich, "Which model to use for cortical spiking neurons?," *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1063–1070, Sep. 2004, doi: 10.1109/TNN.2004.832719.

[5] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003, doi: 10.1109/TNN.2003.820440.

[6] 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: https://doi.org/10.1113/jphysiol.1952.sp004764.

[7] H. Paugam-Moisy and S. Bohte, "Computing with Spiking Neuron Networks," in *Handbook of Natural Computing*, G. Rozenberg, T. Bäck, and J. N. Kok, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 335–376.

[8] R. Brette, "Philosophy of the Spike: Rate-Based vs. Spike-Based Theories of the Brain," *Front. Syst. Neurosci.*, p. 151, 2015, doi: 10.3389/fnsys.2015.00151.

[9] M. Pfeiffer and T. Pfeil, "Deep Learning With Spiking Neurons: Opportunities and Challenges," *Front. Neurosci.*, vol. 12, 2018, doi: 10.3389/fnins.2018.00774.

[10] E. Farquhar, C. Gordon, and P. Hasler, "A field programmable neural array," in *2006 IEEE International Symposium on Circuits and Systems*, May 2006, p. 4 pp. – 4117, doi: 10.1109/ISCAS.2006.1693534.

[11] C.-S. Poon and K. Zhou, "Neuromorphic Silicon Neurons and Large-Scale Neural Networks: Challenges and Opportunities," *Front. Neurosci.*, vol. 5, 2011, doi: 10.3389/fnins.2011.00108.

[12] B. V. Benjamin *et al.*, "Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations," *Proc. IEEE*, vol. 102, no. 5, pp. 699–716, May 2014, doi: 10.1109/JPROC.2014.2313565.

[13] P. A. Merolla *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, Aug. 2014, doi: 10.1126/science.1254642.

[14] M. Davies *et al.*, "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan. 2018, doi: 10.1109/MM.2018.112130359.

[15] C.-T. Wang, C.-T. Lee, X.-J. Wang, and C.-C. Lo, "Top-Down Modulation on Perceptual Decision with Balanced Inhibition through Feedforward and Feedback Inhibitory Neurons," *PLoS ONE*, vol. 8, no. 4, p. e62379, Apr. 2013, doi: 10.1371/journal.pone.0062379.

[16] C.-C. Lo and X.-J. Wang, "Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks," *Nat Neurosci*, vol. 9, no. 7, pp. 956–963, 2006, doi: 10.1038/nn1722.

[17] X.-J. Wang, "Decision Making in Recurrent Neuronal Circuits," *Neuron*, vol. 60, no. 2, pp. 215–234, Oct. 2008, doi: 10.1016/j.neuron.2008.09.034.

[18] C.-C. Lo and X.-J. Wang, "Conflict Resolution as Near-Threshold Decision-Making: A Spiking Neural Circuit Model with Two-Stage Competition for Antisaccadic Task," *PLOS Comput Biol*, vol. 12, no. 8, p. e1005081, Aug. 2016, doi: 10.1371/journal.pcbi.1005081.

[19] P. Miller and D. B. Katz, "Accuracy and response-time distributions for decision-making: linear perfect integrators versus nonlinear attractor-based neural circuits," *J. Comput. Neurosci.*, pp. 1–34, 2013, doi: 10.1007/s10827-013-0452-x.

[20] P. Cassey, A. Heathcote, and S. D. Brown, "Brain and Behavior in Decision-Making," *PLoS Comput Biol*, vol. 10, no. 7, p. e1003700, Jul. 2014, doi: 10.1371/journal.pcbi.1003700.

[21] A. K. Churchland, R. Kiani, and M. N. Shadlen, "Decision-making with multiple alternatives," *Nat. Neurosci.*, vol. 11, no. 6, pp. 693–702, Jun. 2008, doi: 10.1038/nn.2123.

[22] R. Kiani, C. J. Cueva, J. B. Reppas, and W. T. Newsome, "Dynamics of Neural Population Responses in Prefrontal Cortex Indicate Changes of Mind on Single Trials," *Curr. Biol.*, vol. 24, no. 13, pp. 1542–1547, Jul. 2014, doi: 10.1016/j.cub.2014.05.049.

[23] J. Bahuguna, A. Aertsen, and A. Kumar, "Existence and Control of Go/No-Go Decision Transition Threshold in the Striatum," *PLoS Comput Biol*, vol. 11, no. 4, p. e1004233, Apr. 2015, doi: 10.1371/journal.pcbi.1004233.

[24] A. Pastor-Bernier and P. Cisek, "Making Choices between Rules or between Actions," *Neuron*, vol. 70, no. 3, pp. 382–384, May 2011, doi: 16/j.neuron.2011.04.017.

[25] Z.-W. Yeh *et al.*, "POPPINS : A Population-Based Digital Spiking Neuromorphic Processor with Integer Quadratic Integrate-and-Fire Neurons," *ISCAS 2021 (accepted)*.