A low power silicon synapse with tunable reversal potential.

Ashish Gautam
Graduate School of Engineering, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan

Takashi Kohno
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan
E-mail: asgautam@sat.t.u-tokto.ac.jp, kohno@sat.t.u-tokyo.ac.jp
www.u-tokyo.ac.jp

Abstract

Synapses are a building block of signal processing and computation in neuronal cell's networks. We present concept and simulation results of a pseudo five-bit, low power silicon synapse circuit capable of emulating both excitatory (NMDA/AMPA type) and inhibitory (GABA type) responses. The post synaptic current generated in this circuit is proportional to the difference between the post synaptic membrane potential and a tunable synaptic reversal potential.

Keywords: Synapse, neuromorphic, reversal potential, synaptic weights.

1. Introduction

Research interest in the field of neuromorphic engineering has grown significantly in the last decade. The driving force behind the research is an attempt to arrive at a novel computing architecture which surpasses the limitations of the currently prevalent von-Neumann architecture by taking advantage of a highly parallelized network structure of localized memory and processing units to emulate the information processing in neuronal cell's networks. Neuronal and synaptic circuits that emulate the dynamics of neuronal cells and synapses form the building blocks of this nascent computing architecture. Over the years a wide variety of neuronal and synaptic circuits with a varied degree of detail have been proposed. Earlier models[1] simplified post-synaptic currents as pulses and others replicated the exponentially rising and decaying characteristic of these currents. Recently proposed models[2] focused on reducing the hardware footprint by sharing of synaptic circuits having a single synapse circuit emulate multiple synapses. These circuits are compact and power-efficient but do not reproduce the synaptic reversal potential which is thought to be playing an important role in the information processing. In addition, their static power consumption was not evaluated even though it has a large impact on the power consumption of large-scale networks.

We propose a low-power synaptic circuit with synaptic reversal potential that has pseudo five-bit of synaptic weights, implementing a familiar kinetic model of synaptic conductance. Our design places its emphasis on ultra-low power consumption and the exponential

©The 2018 International Conference on Artificial Life and Robotics (ICAROB2018), Feb. 1-4, B-Con Plaza, Beppu, Oita, Japan

profile of synaptic current emulating the currents mediated by various neurotransmitters (AMPA, NMDA, GABAa, and GABAb) in the brain. In addition, the synaptic current in our circuit is proportional to the difference between the membrane potential and the synaptic reversal potential.

2. Synaptic Model

It was shown that the kinetic models of the chemical synapses which describe the neurotransmitter kinetics can describe and reproduce the synaptic activities precisely[3]. Following the analysis discussed in Ref.3, the post-synaptic current in our circuit is given by the simplest version of the kinetic models.

$$\boldsymbol{I_{syn}(t) = g_{syn} * r(t) * (V_{syn}(t) - E_{syn})}, \quad (1)$$

where r represents the fraction of bound post-synaptic receptors, g_{syn} is the maximal value of synaptic conductance, V_{syn} is the post-synaptic potential and E_{syn} is the synaptic reversal potential.

3. Proposed Synaptic Circuit

3.1. *Description and analysis*

The schematic diagram of the proposed synaptic circuit is shown in Fig. 1. It consists of three major blocks, the input stage (devices M1 to M10), the integrator (C_{syn}, M11, and M12) and the transconductance amplifier (M13 to M19). The input stage is functionally similar to log domain integrator synapse circuit proposed by Merolla et al.[4] with roles of PMOS and NMOS interchanged. The PMOS transistor M1 acts as a switch, transistor M2 and the inverter I0 together provide the necessary charge to be injected during the rising and falling phase of the input pulse so as to avoid the distortion of the current waveform at the switching points of the input pulse due to the effect parasitic capacitors inherent in the Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) device. Transistors M3, M5, M7, and M9 act as switches and determine the state of corresponding branches shown in the figure. These branches are for implementing synaptic weights discussed later. For the present discussion, we assume only the branch consisting M3 is turned ON leaving M5, M7, and M9 in the OFF state. An input pulse at the gate of transistor M1 pulls the drain terminal of NMOS M4

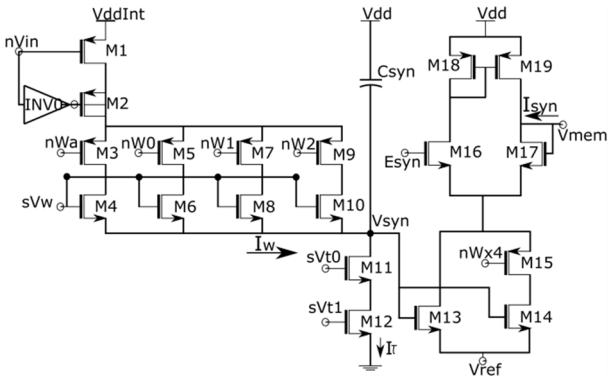

Figure. 1: Schematic diagram of the proposed synaptic circuit. Dimensions of the transistors: M4=0.3758*(w/l), M6=w/l, M8= 2*(w/l) and M10=4*(w/l).Nets sVt0 and sVt1 are shorted and connected to a common node Vτ. Tail transistor M14 is three times wider than M13.

to VddInt and current that flows through the device is given by (ignoring the short channel effect)

$$\boldsymbol{I_w = I_{wo} e^{\frac{(k*sV_w - V_{syn})}{U_T}}}, \quad (2)$$

where I_{w0} represents the leakage current, sV_w is the voltage applied to the gate of M4 and V_{syn} represents the source voltage of the transistor M4. This current is then integrated by the RC network formed by capacitor C_{syn} and the NMOS transistors M11 and M12 (as their gates are shorted they act as a single transistor). During the charging phase, these transistors are not in saturation as V_{syn} is initially 0. When the input pulse turns off, the gate of the transistor M1 goes high, the current I_w shuts off and the capacitor begins to discharge linearly through the transistors M11-M12. When these transistors are saturated the current I_τ is given by

$$\boldsymbol{I_\tau = I_{\tau o} e^{\frac{(kV_\tau)}{U_T}}}, \quad (3)$$

where $I_{\tau 0}$ represents the leakage current, V_τ controls the time constant of the synaptic circuit. Voltage V_{syn} is fed into the tail transistor of the transconductance amplifier (M13 and M14), which generates a current proportional to the difference of the voltage applied at the inputs of the differential pair (E_{syn}-V_{mem}). The generated current I_{syn} is given by

$$\boldsymbol{I_{syn} = I_o e^{\frac{(kV_{syn} - sV_{ref})}{U_T}} * tanh(\frac{k}{2U_T}(E_{syn} - V_{mem}))}. \quad (4)$$

The source of the tail transistors is connected to V_{ref}, which minimizes the leakage current when V_{syn} is set to zero during the inactive state of the synapse

circuit and turns off the tail transistor as soon as possible during the discharging phase when the transistors M11-M12 come out of saturation. By differentiating Eq. (4) we get

$$\frac{dI_{\mathrm{syn}}}{dt} = I_{\mathrm{syn}} * \frac{k}{U_T} * \frac{dV_{\mathrm{syn}}}{dt}, \tag{5}$$

and the dynamics of the node V_{syn} is given by

$$C_{\mathrm{syn}} * \frac{dV_{\mathrm{syn}}}{dt} = (I_w - I_\tau), \tag{6}$$

combining Eq. (5) and Eq. (6), we get

$$\frac{C_{\mathrm{syn}}}{I_\tau} * \frac{U_T}{k} * \frac{dI_{\mathrm{syn}}}{dt} + I_{\mathrm{syn}} = \frac{I_{\mathrm{syn}} * I_w}{I_\tau}. \tag{7}$$

In the log domain integrator synapse described in Ref. 4, the right-hand side of Eq. (7) comes out to be a constant due to the inverse relation between I_w and I_{syn}. This is not exactly true in our circuit due to the body effect of the NMOS transistor M4, but to avoid complexity in our analysis we neglect the body effect and approximate the dynamics of our circuit as that of the log domain integrator. By this Eq. (7) reduces to the equation of a first order low pass filter and its response to a spike arriving at t_i^- and ending at t_i^+ during the charging phase is given by

$$I_{\mathrm{syn}}(t) = I_{\mathrm{const}}\left(1 - exp\left(-\frac{t - t_i^-}{\tau}\right)\right), \tag{8}$$

and during the discharge phase

$$I_{\mathrm{syn}}(t) = I_{\mathrm{syn}}^+ exp\left(-\frac{t - t_i^+}{\tau}\right), \tag{9}$$

where I_{const} is the constant term on the right-hand side of Eq. (9), I_{syn}^+ is the initial condition at t_i^+ and τ is the time constant given by

$$\tau = \frac{C_{\mathrm{syn}} * U_T}{k * I_\tau}. \tag{10}$$

During the discharge phase, a small current flows out of the V_{syn} node back into the synaptic input stage due to the parasitic capacitance of the NMOS transistors M4, M6, M8, and M10 which is comparable to I_τ for higher values of time constant (100ms to 200ms), and in this range to calculate the time constant this additional current flowing out of the V_{syn} node must be added to I_τ in Eq. (10).

3.2 *Pseudo five-bit synaptic weights.*

Wang et al. proposed a programmable five-bit synaptic weight circuit[5], where the weights are set by an on-chip DAC. Because the number of synaptic circuits in silicon neuronal networks is very large, their footprint size has to be minimized. We propose a more compact circuit whose resolution is almost five-bit. In Fig.1 the four branches consisting of transistors M3 to M10, with the NMOS transistors sized appropriately provide 4 bit of synaptic weight. The fifth bit is represented by node nWx4 which when active turns ON the switch M15 and incorporates transistor M14 along with M13 as the tail device thus providing four times increase in the magnitude of synaptic current. Table 1 below lists down the synaptic weights given by the five-bit input, nWa, nW0, nW1, nW2, and nWx4. We call it pseudo five-bit as the total number of weight values we get are 27 instead of 32. In our circuit, instead of using 32 transistors to realize full five-bit, we used seven full-sized transistors for M6, M8 and M10 and one half size transistor for M4. This shrinks not only the footprint of these transistors but also that of M2 by reducing the charge injection phenomenon.

Table 1. Pseudo five-bit synaptic weights.

nW2	nW1	nW0	nWa	nWx4	Weight Value
0	0	0	0	0	0
0	0	0	1	0	1
⋮	⋮	⋮	⋮	⋮	⋮
1	1	1	1	0	15
0	1	0	1	1	16
0	1	1	0	1	17
⋮	⋮	⋮	⋮	⋮	⋮
1	1	1	1	1	26

4. Simulation Results

We designed our synaptic circuit to be connected to the silicon neuron circuit in Ref. 6. The scale of synaptic current in our circuit was designed so as to interact with this silicon neuron circuit, whose membrane capacitance is about 900fF. Since the membrane capacitance of the neuronal cells is about several hundreds of picofarads and the scale of synaptic currents is about several hundreds of picoamperes, the synaptic current scale of our silicon synapse circuit is set to about 10pA. This range is far smaller than that in other silicon synapse circuits in Ref.2 and Ref.5. The static power consumption of our circuit was less than 2pW. To

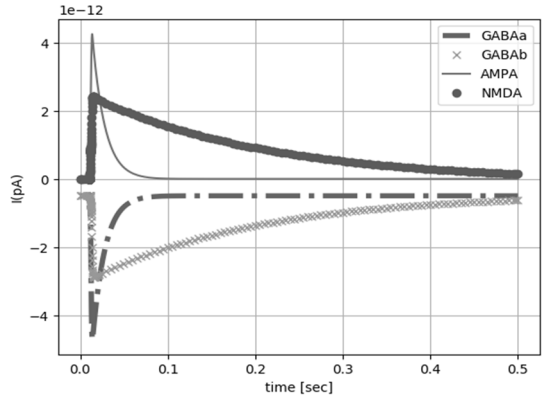

Figure 2: Synaptic currents emulating the response of (a) AMPA (b) NMDA (c) GABAA (d) GABAB neurotransmitters. *sVw* was set to 80mV for NMDA and $GABA_b$ mode and to 110mV for AMPA and $GABA_a$ mode.

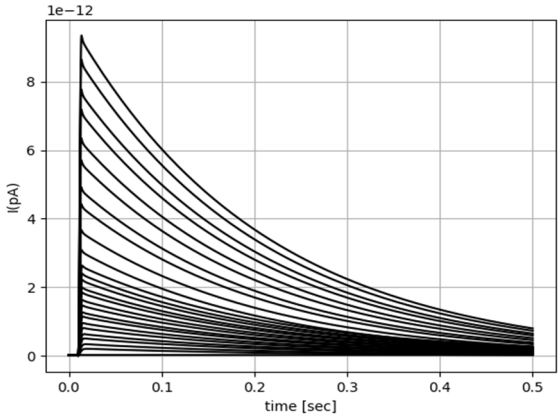

Figure 3: Synaptic current output over the dynamic range of the synaptic weights.

process an input spike the dynamic energy consumption was calculated to be about 500fJ. This value was arrived at by subtracting the value of static energy dissipated from the total energy consumed to process an input spike, with the circuit configured to peak current value of 3pA and time constant of 200ms. For the maximum current of 10pA and maximum time constant of 200ms, this value was calculated to be about 1.5pJ. Figure 2 shows the waveform of the synaptic currents emulating the response of various neurotransmitters and Fig. 3 shows the plot of synaptic currents over the dynamic range of synaptic weights described in Table 1. These results were obtained by circuit simulation using Spectre. All the traces are in typical shape of the synaptic current waveforms. The offset observed in the inhibitory setting ($GABA_a$ and $GABA_b$ in Fig. 2) will be compensated by silicon neuron circuit's parameters.

5. Conclusion

We designed a silicon synapse circuit for a low power silicon neuron circuit using Taiwan Semiconductor Manufacturing Company (TSMC) 250nm CMOS process and presented the simulation results. Our circuit is designed to consume lower power compared to 2.8pJ[7] which implements a DPI synapse discussed in Ref.2. We were unable to find static power consumption value for previously described synaptic circuits in Ref.2 and Ref.5. The area occupied by a single synapse is 26.02x140.2 um^2. In our future work, we plan to reduce the area of synaptic circuits by significantly reducing the size of the integrator stage and sharing one transconductance amplifier with many input stages, as well as improve the bit resolution of the synapse circuit.

6. Acknowledgements

This study was supported by JST PRESTO and CREST, and VDEC, The University of Tokyo in collaboration with Cadence Design Systems, Inc.

References

1. Mead, C. (1989). *Analog VLSI and neural systems*. Reading, MA: Addison-Wesley.
2. Bartolozzi,C; Indiveri,G (2007). Synaptic dynamics in analog VLSI. *Neural Computation*, 19(10):2581-2603.
3. Destexhe.et.al. (1998). *Kinetic models of synaptic transmission*. In Methods in neuronal modelling, from ions to networks (pp. 1–25). Cambridge, MA: MIT Press.
4. Merolla P and Boahen K. 2004. A recurrent model of orientation maps with simple and complex cells. In: *Advances in Neural Information Processing Systems* 16 (NIPS) MIT Press, Cambridge, MA. pp. 995-1002.
5. Wang, Y.,& Liu, S. (n.d.). Programmable synaptic weights for an aVLSI network of spiking neurons. 2006 *IEEE International Symposium on Circuits &Systems*.doi:10.1109/iscas.2006.1693637.
6. Kohno, T., & Aihara, K. (n.d.). A three variable ultralow-power analog silicon neuron circuit. *NOLTA 2016* (Yugawara, Japan, November 27-30, 2016).
7. Giacomo Indiveri, Federico Corradi, and Ning Qiao. Neuromorphic architectures for spiking deep neural networks. In *IEEE International Electron Devices Meeting (IEDM), 2015*.