A Subthreshold Second-Order Integration Circuit for Versatile Synaptic Alpha Kernel and Trace Generation

Ole Richter
Hugh Greatorex
o.j.richter@rug.nl
University of Groningen
Groningen, The Netherlands

Benjamín Hučko
University of Groningen
Groningen, The Netherlands

Madison Cotteret
Technische Universität Ilmenau
Ilmenau, Germany
University of Groningen
Groningen, The Netherlands

Willian Soares Girão
University of Groningen
Groningen, The Netherlands

Ella Janotte
Istituto Italiano di Tecnologia
Genova, Italy
University of Groningen
Groningen, The Netherlands

Michele Mastella
Elisabetta Chicca
University of Groningen
Groningen, The Netherlands

ABSTRACT

In neuromorphic hardware the choice of synaptic kernels and trace dynamics are key for the correct system abstraction and representation of information. This paper presents a novel second-order integration circuit for the implementation of traces and kernels, the Second-order Differential Pair Integrator (SoDPI). It provides smooth alpha-kernel shaped responses to spike input in analog subthreshold complementary Metal Oxide Semiconductor (MOS) technology. Our approach utilises two Differential Pair Integrator (DPI) circuits in series to implement an effective current-mode second-order translinear low-pass filter. Theoretical analysis and experimental measurements demonstrate the improved reliability of this design, which offers a promising approach for modelling biological synaptic and neural responses in neuromorphic hardware, as well as improving the stability of integrated on-chip learning systems.

CCS CONCEPTS

• **Hardware** → **Analog and mixed-signal circuits**; Integrated circuits; • **Computing methodologies** → **Neural networks**; **Bio-inspired approaches**.

KEYWORDS

Spiking Neuronal Network, Neuromorphic Engineering, CMOS, Synaptic Kernel, Synaptic Trace.

Author Contribution: Conceptualisation, Methodology - O.R.; Software/Hardware - E.J., H.G., M.C., M.M., O.R., W.S.G.; Investigation - B.H., O.R.; Formal analysis - H.G.; Writing - original draft - B.H., H.G., M.C., O.R.; Writing – review and editing - E.C., E.J, M.M, W.S.G; Visualisation - B.H., H.G., M.C., O.R.; Supervision - E.C. (CRediT, authors in alphabetical order).

ICONS '23, August 1–3, 2023, Santa Fe, NM, USA

ACM ISBN 979-8-4007-0175-7/23/08.
https://doi.org/10.1145/3589737.3606008

ACM Reference Format:
Ole Richter, Hugh Greatorex, Benjamín Hučko, Madison Cotteret, Willian Soares Girão, Ella Janotte, Michele Mastella, and Elisabetta Chicca. 2023. A Subthreshold Second-Order Integration Circuit for Versatile Synaptic Alpha Kernel and Trace Generation. In *International Conference on Neuromorphic Systems (ICONS '23), August 1–3, 2023, Santa Fe, NM, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3589737.3606008

1 INTRODUCTION

Research in the field of neuromorphic engineering strives to build computational systems which draw inspiration from biology to develop efficient and intelligent information processing systems and to gain a deeper understanding of the computational principles governing their biological counterparts [8]. An immediate and reoccurring question is the level of abstraction required to replicate the most important behaviour of biological observations. Within Spiking Neural Network (SNN) and their hardware implementations, the myriad of chemical processes governing the movement of ions and neurotransmitters within synapses [11] are often abstracted to

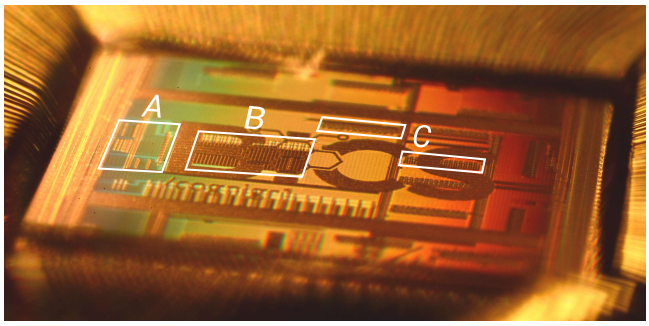

Figure 1: Photograph of the realised "Cognigr1" Application Specific Integrated Circuit (ASIC). The box labelled B indicates the area covered by the array of 40 SoDPI circuits with 520 synaptic branches, as presented in this work. The structures labelled by C are the on-chip Digital to Analog Converters (DAC) and A indicates the location of asynchronous Adress Event Representation (AER) and Input and Output (I/O) circuitry.

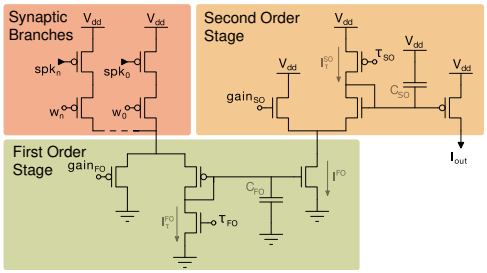

Figure 2: The SoDPI circuit in its compact variant, with relevant currents and biases labelled. It consists of the synaptic branches or tails, and two DPI circuits for implementing both stages of the second order filter. Upon arrival of a presynaptic spike, a fixed charge packet with its size modulated by the individual weight is supplied to the first DPI. The main blocks are two DPI low-pass filters, here seen as first order stage in n-type MOS (nMOS) variant and second order stage in p-type MOS (pMOS) variant.

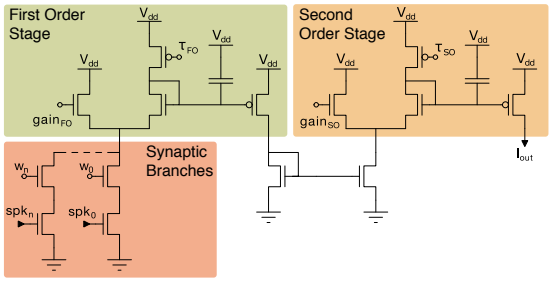

Figure 3: The SoDPI circuit in its pMOS variant. The advantage compared to Fig. 2 is improved usability as the parameters $\tau_{\mathrm{FO/SO}}$ and $\mathrm{gain}_{\mathrm{FO/SO}}$ match and behave exactly the same for both stages. The cost of this is larger area and higher power consumption as an additional current mirror is required.

a reduced set of dynamical variables which are influenced by pre- and postsynaptic spike activity. These so-called traces often play a crucial role in the various mechanisms of inference, adaptation, plasticity and learning [5, 7].

How faithfully biological synaptic dynamics should be emulated is thus key to finding optimal representations of the information available to the synapses and therefore to building more efficient integrated neuromorphic hardware systems.

In on-chip learning systems in particular, traces are heavily relied upon to inform online adaptation and weight update circuitry [7]. Choosing the incorrect trace type for the use case at hand degrades performance and usability significantly. Nevertheless, this is a common mistake which can be avoided thanks to the availability of a variety of filters. In the field of neuromorphic subthreshold analog Complementary MOS (CMOS) design, different traces are available for different use cases: capped traces implemented by capacitive discharge circuits or pulse extenders are useful for indicating the time passed since last spike, while first-order integration circuits such as [1] are useful for indicating the short-term history of spike activity or determining credit assignment in the form of eligibility traces [2]. For models that require information about the average neuron activity [13, 4] there is the need for higher-order integration circuits, as filtering the short-term fluctuations due to single spikes is vital to minimise the variance of the estimated mean spike rate. The response kernels for such filtering circuits are the so-called alpha kernels. They provide a good approximation of both excitatory and inhibitory Post Synaptic Currents (PSCs) resulting from complex ion channel activation [11]. Consequently, multiple designs for alpha kernel synaptic integrator circuits have been proposed and successfully implemented in analog neuromorphic hardware [3, 14, 16].

We present the SoDPI circuit, which improves upon the reliability and flexibility of previous designs. We discuss the increased applicability of our design for implementing integrative traces, superposing multiple synaptic traces, and the stabilising properties offered to neuromorphic systems with online learning capabilities.

2 METHODS

The proposed SoDPI circuit in Fig. 2 is composed of two DPI circuits originally proposed in [1]. The first DPI acts as a First Order (FO) translinear low-pass filter. Its response to an ideal digital input spike on the node $\mathrm{spk}_{\mathrm{n}}$ of the n-th synaptic branch is described by the following equation [1]:

$$I^{\mathrm{FO}}(t) = I^{\mathrm{FO}+} e^{-\frac{(t-t^+)}{\tau_{\mathrm{FO}}}} \tag{1}$$

where $I^{\mathrm{FO}}(t)$ is the output current, $\tau_{\mathrm{FO}} = \frac{U_T C_{\mathrm{FO}}}{\kappa I_\tau^{\mathrm{FO}}}$ is the time constant, t^+ is the time of the falling edge of the input spike and $I^{\mathrm{FO}+}$ is the output current I^{FO} at time t^+.

The equation above can be derived [1] by assuming transistor saturation and applying the translinear principle [9]. Similarly, the output of the SO stage is described by the following differential equation:

$$\tau_{\mathrm{SO}} \frac{\mathrm{d}I_{\mathrm{out}}}{\mathrm{d}t} + I_{\mathrm{out}} = \frac{I_{\mathrm{gain}}^{\mathrm{SO}}}{I_\tau^{\mathrm{SO}}} I^{\mathrm{FO}+} e^{-\frac{t}{\tau_{\mathrm{FO}}}} \tag{2}$$

where $\tau_{\mathrm{SO}} = \frac{U_T C_{\mathrm{SO}}}{\kappa I_\tau^{\mathrm{SO}}}$ is the time constant, $I_{\mathrm{gain}}^{\mathrm{SO}}$ is the virtual p-type subthreshold current biased by $\mathrm{gain}_{\mathrm{SO}}$ [1] and all other variables are defined in Fig. 2. The equation can be solved to derive the following expression for the output current:

$$I_{\mathrm{out}} = \frac{I_{\mathrm{gain}}^{\mathrm{SO}} I^{\mathrm{FO}+}}{I_\tau^{\mathrm{SO}}} \frac{\tau_{\mathrm{FO}}}{(\tau_{\mathrm{SO}} - \tau_{\mathrm{FO}})} \left(e^{-\frac{t}{\tau_{\mathrm{SO}}}} - e^{-\frac{t}{\tau_{\mathrm{FO}}}} \right) \tag{3}$$

where we have set $t^+ = 0$ such that we consider the dynamics of I_{out} at the falling edge of an incoming spike. In the limit of $\tau_{\mathrm{FO}} = \tau_{\mathrm{SO}} = \tau$ we recover the alpha kernel form:

$$I_{\mathrm{out}} = \frac{I_{\mathrm{gain}}^{\mathrm{SO}} I^{\mathrm{FO}+}}{I_\tau^{\mathrm{SO}}} \frac{t e^{-\frac{t}{\tau}}}{\tau} \tag{4}$$

Multiple synaptic branches can share the same synaptic kernel circuit, as commonly done in [3, 12]. This is possible thanks to the superposition principle of linear circuits. To ensure such linearity, transistors in the DPI should operate in saturation ($V_{\mathrm{ds}} >> U_T$),

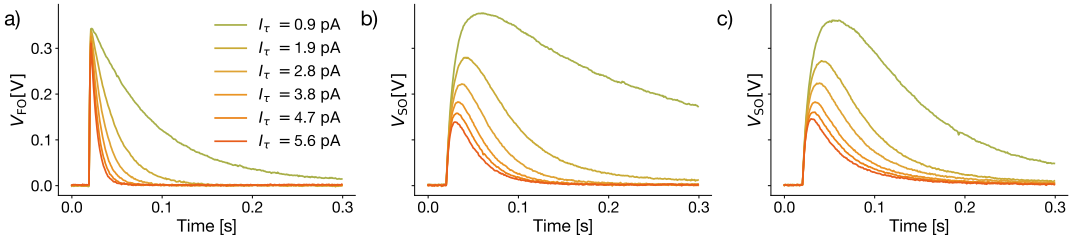

Figure 4: Measured kernel responses to a single input spike for different time constants (see legend, valid for all plots). a) Response of the DPI block (FO stage). b) Response of the full SoDPI circuit with the parameter sweep in the FO stage. c) Response of the full SoDPI circuit with the parameter sweep in the Second Order (SO) stage. While we are showing data for the voltage across the output capacitor of the circuits, the next stage uses the current produced by the filter. This is given by the exponential of the plotted voltage and it is affected by the relevant transistor variables (e.g., U_T, κ) [8].

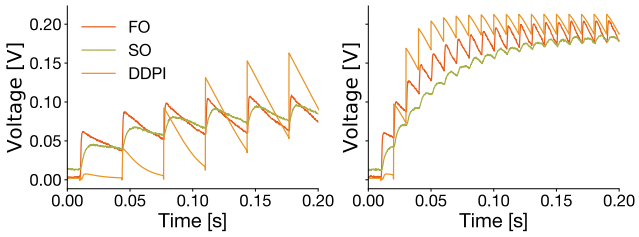

Figure 5: The integration behaviour from onset to steady state for a 30 Hz (left) and 100 Hz (right) input train of pulses for the DPI (FO), SoDPI (SO) and Double DPI (DDPI) circuits (with colour code defined in the legend). The advantage of the SoDPI is clearly visible as the response retains its alpha kernel shape for all input spikes. The DDPI data points are obtained through Simulation Program with Integrated Circuit Emphasis (SPICE) simulations, while SoDPI and DPI plots show ASIC measurements.

the conditions $I_{\text{gain}} << I_{\text{out}}$ and $I_\tau << I_{\text{in}}$ should be valid for each stage.

The circuit shown in Fig. 3 is replicated 40 times on the ASIC Cognigr1 shown in Fig. 1 with a variable number of branches per circuit $(1-32)$ that sum up to a total of 520 synaptic branches. The circuit has been implemented in XFAB 180nm technology as a cost-effective proof of concept of subthreshold analog design. A single circuit without synaptic branches has an area of $740\mu m^2$ in our implementation, but can be size-optimised to $275\mu m^2$ and lower by using the circuit shown in Fig. 2 with same transistor sizes and Back-End Of Line (BEOL) Metal Insulator Metal (MIM) or Metal Oxide Metal (MOM) capacitors as a replacement for MOS Capacitor (MOSCAP). Note that the MOSCAPs are operated in their constant capacitance range, and Fig. 2 and 3 are drawn with capacitor equivalents for clarity and syntactic consistency to previous works. The dynamic current consumption of the circuit presented is highly parameter-specific, but can be well approximated by $I_{\text{total}} = I_{\text{out}} + I_{\text{in}} + I_{\text{FO}}$ for the SoDPI of Fig. 2 and $I_{\text{total}} = I_{\text{out}} + I_{\text{in}} + 2 \cdot I_{\text{FO}}$ for its pMOS variant of Fig. 3. The circuit operation regime is set by an on-chip DAC, designed based on [15]. The DAC is configured by a micro controller via a First In First Out Register (FIFO). Synaptic branches can be stimulated by an asynchronous Quasi Delay Insensitive (QDI) AER decoder infrastructure. The chip has a monitor system able to probe both V_{FO} and V_{SO} of all 40 instances on a limited number of I/O. The monitoring is done via current mirroring and therefore the voltages are measured as $V_{\text{measument}} = V_{\text{dd}} - V_{\text{FO/SO}}$. As comparison we simulated the DDPI circuit [14] in a SPICE based simulator with the ASIC production Process Development Kit (PDK) using matching sizes and similar circuit parameters. We swapped the second current mirror from [14] to a pMOS version for correct circuit operation.

3 RESULTS AND DISCUSSION

In Fig. 4 we present and compare the synaptic response kernel voltage curves recorded from the widely used DPI circuit and the SoDPI circuit proposed in this work, with different time constants. The linearity assumption required for equation 3 is not fulfilled for single and sparse events (i.e. events that are separated in time such that they always occur when there is no charge on the capacitor, therefore accumulation of the effect of subsequent spikes does not occur). During the discharge phase, the non linearity can be clearly observed in the traces in Fig. 4 for $V \lessapprox 100\,\text{mV}$. This is due to the transistor responsible for generating the discharge current I_τ leaving the saturation and entering the ohmic regime. Given the integration requirement (i.e. accumulation of the effect of subsequent spikes) for operating in the linear regime, implementing these circuits with several synaptic branches increases the probability of operating them in the proper regime, satisfying the linearity assumption. Furthermore, using a single filter for several synapses reduces the area overhead per synapse.

One key improvement of this work is that the alpha kernel is retained both in the single or sparse response regime and in the integration regime as seen in Fig. 5, making it easily configurable and user friendly. The previous DDPI can not retain its alpha kernel, as seen in Fig. 5, as it is not robust to the non linearity presented above and in Fig. 6. Furthermore, the requirement to tune parameters to obtain reasonable integration behaviours produces an onset response with unreasonable latencies and lack of response for frequencies $< 20Hz$ (see Fig. 6b).

Another key improvement seen in Fig. 6 is the low signal fluctuation in steady state operation leading to a more accurate representation of the mean rate of the input compared to previous work. The

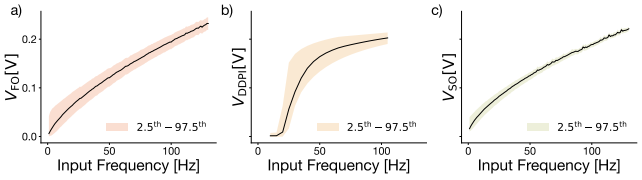

Figure 6: The steady state for a sweep of input frequency for the DPI (a), DDPI (b), and SoDPI (c). The 2.5-97.5 percentile shows the signal fluctuations with regard to the mean of the steady state response. The increase in signal fluctuations and change of the first derivative of the output voltage with respect to the input frequency observed for the DDPI is due to an intrinsic limitation of the circuit which does not comprise a parameter regime in which a proper kernel shape is produced for both single spikes and high frequency inputs. The DDPI data points are obtained through SPICE simulations, while SoDPI and DPI plots show ASIC measurements.

lower fluctuations can be employed for smother responses in plasticity circuitry like homeostasis, calcium and mean rate traces [6, 10], gating and conductance circuitry [14] making the responses less sensitive to input spike timing.

4 CONCLUSION

We presented a robust second-order integrator circuit, which can replace standard first-order synaptic integration circuits for mimicking alpha shaped biological Inhibitory PSCs (IPSCs) and Excitatory PSCs (EPSCs) with minimal added overhead. Furthermore, the circuit is able to produce mean rate trace representations with low steady state fluctuations. We provided analytical solutions for the response kernel in the saturation regime, and showed its functioning in a 180 nm CMOS technology. The circuit improves the previously presented solutions in terms of spike kernels with a stable response across input regimes, from isolated spikes to sparse and integration regime. Furthermore, given the translinear design principle, the impact of device mismatch is limited to the linear range [9] as opposed to the exponential impact on voltage based subthreshold circuits. The process variation is reduced given the equation containing predominantly ratio between currents.

The lower signal-to-fluctuation-ratio of the output promises improvements in stability and precision for emulated learning algorithms implementations relying on accurate mean rate information. The proposed circuit provides a novel basic building block for the implementation of learning systems. It can be used to port a variety of algorithms (e.g. calcium based stop-learning, frequency based learning rules, gating, homeostasis) into full-custom low-power hardware systems. The properties of the circuits demonstrated in this paper support its application to learning algorithms can guarantee stability, consistency and precision of synaptic updates.

ACKNOWLEDGMENTS

The authors would like to thank Giacomo Indiveri for the discussions that led to this work, Philipp Klein for soldering the PCBs and Matteo Cartiglia for assisting with microcontroller programming. Additional picture credits Thorben Schoepe. The authors acknowledge the financial support of the CogniGron research center and the Ubbo Emmius Funds (Univ. of Groningen). This work has been supported by EU H2020 projects NeuTouch (813713), BeFerroSynaptic (871737) and MANIC (861153). Additional funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): Project MemTDE Project number 441959088 (SPP 2262 MemrisTec: 422738993); Project NMVAC Project number 432009531.

REFERENCES

[1] Chiara Bartolozzi and Giacomo Indiveri. 2007. Synaptic Dynamics in Analog VLSI. en. *Neural Computation*, 19, 10, (Oct. 2007), 2581–2603. DOI: 10.1162/neco.2007.19.10.2581.

[2] Guillaume Bellec, Franz Scherr, Anand Subramoney, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. 2020. A solution to the learning dilemma for recurrent networks of spiking neurons. en. *Nature Communications*, 11, 1, (July 2020), 3625. DOI: 10.1038/s41467-020-17236-y.

[3] Ben Varkey Benjamin, John V. Arthur, Peiran Gao, Paul A. Merolla, and Kwabena Boahen. 2012. A superposable silicon synapse with programmable reversal potential. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, San Diego, CA, (Aug. 2012), 771–774. DOI: 10.1109/EMBC.2012.6346045.

[4] Joseph M. Brader, Walter Senn, and Stefano Fusi. 2007. Learning Real-World Stimuli in a Neural Network with Spike-Driven Synaptic Dynamics. en. *Neural Computation*, 19, 11, (Nov. 2007), 2881–2912. DOI: 10.1162/neco.2007.19.11.2881.

[5] Wulfram Gerstner, Marco Lehmann, Vasiliki Liakoni, Dane Corneil, and Johanni Brea. 2018. Eligibility Traces and Plasticity on Behavioral Time Scales: Experimental Support of NeoHebbian Three-Factor Learning Rules. *Frontiers in Neural Circuits*, 12, (July 2018), 53. DOI: 10.3389/fncir.2018.00053.

[6] Giacomo Indiveri and Elisabetta Chicca. 2011. A VLSI neuromorphic device for implementing spike-based neural networks. *Neural Nets WIRN11*. Frontiers in Artificial Intelligence and Applications 234, 305–316. DOI: 10.3233/978-1-60750-972-1-305.

[7] Lyes Khacef, Philipp Klein, Matteo Cartiglia, Arianna Rubino, Giacomo Indiveri, and Elisabetta Chicca. 2022. Spike-based local synaptic plasticity: A survey of computational models and neuromorphic circuits. arXiv:2209.15536 [cs]. (Nov. 2022).

[8] Carver Mead. 1989. *Analog VLSI and Neural Systems*. Addison-Wesley Longman Publishing Co., Inc., USA. ISBN: 0-201-05992-4.

[9] Jan Mulder, Wouter A. Serdijn, Albert C. Woerd, and Arthur H. M. Roermund. 1999. *Dynamic Translinear and Log-Domain Circuits*. Springer US, Boston, MA. DOI: 10.1007/978-1-4615-4955-0.

[10] Melika Payvand and Giacomo Indiveri. 2019. Spike-Based Plasticity Circuits for Always-on On-Line Learning in Neuromorphic Systems. In *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, Sapporo, Japan, (May 2019), 1–5. DOI: 10.1109/ISCAS.2019.8702497.

[11] Dale Purves, George J. Augustine, David Fitzpatrick, William Hall, Anthony-Samuel LaMantia, and Leonard White, (Eds.) 2018. *Neuroscience*. (Sixth edition ed.). Oxford University Press, New York. ISBN: 978-1-60535-380-7.

[12] Ning Qiao, Hesham Mostafa, Federico Corradi, Marc Osswald, Fabio Stefanini, Dora Sumislawska, and Giacomo Indiveri. 2015. A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128K synapses. *Frontiers in Neuroscience*, 9, (Apr. 2015). DOI: 10.3389/fnins.2015.00141.

[13] João Sacramento, Rui Ponte Costa, Yoshua Bengio, and Walter Senn. 2018. Dendritic cortical microcircuits approximate the backpropagation algorithm. In *Advances in Neural Information Processing Systems*. S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, (Eds.) Vol. 31. Curran Associates, Inc.

[14] Dora Sumislawska, Ning Qiao, Michael Pfeiffer, and Giacomo Indiveri. 2016. Wide dynamic range weights and biologically realistic synaptic dynamics for spike-based learning circuits. In *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, Montréal, QC, Canada, (May 2016), 2491–2494. DOI: 10.1109/ISCAS.2016.7539098.

[15] Minhao Yang, Shih-Chii Liu, Chenghan Li, and Tobi Delbruck. 2012. Addressable current reference array with 170dB dynamic range. In *2012 IEEE International Symposium on Circuits and Systems*. IEEE, Seoul, Korea (South), (May 2012), 3110–3113. DOI: 10.1109/ISCAS.2012.6271979.

[16] Theodore Yu and Gert Cauwenberghs. 2010. Analog VLSI Biophysical Neurons and Synapses With Programmable Membrane Channel Kinetics. *IEEE Transactions on Biomedical Circuits and Systems*, 4, 3, (June 2010), 139–148. DOI: 10.1109/TBCAS.2010.2048566.