

# Analog Low-Power Hardware Implementation of a Laguerre-Volterra Model of Intracellular Subthreshold Neuronal Activity

Viviane S. Ghaderi, *Student Member, IEEE*, Shane Roach, Dong Song, *Member, IEEE*, Vasilis Z. Marmarelis, *Fellow, IEEE*, John Choma, Jr., *Fellow, IEEE* and Theodore W. Berger, *Fellow, IEEE*

**Abstract**— The right level of abstraction for a model mimicking a neural function is often difficult to determine. There are trade-offs between capturing biological complexities on one hand and the scalability and efficiency of the model on the other. In this work, we describe a nonlinear Laguerre-Volterra model of the synaptic temporal integration of input spikes to postsynaptic potentials. This model is then efficiently implemented using analog subthreshold circuits and can serve as a foundation for future large-scale hardware systems that can emulate multi-input multi-output (MIMO) spike transformations in populations of neurons. The normalized mean square error in estimating real data using the circuit implementation of this model is less than 15%. The model components are modular and its parameters are adjustable for modeling temporal integration by neurons in other brain regions. The total power consumption of this nonlinear Laguerre-Volterra system is less than 5nW.

## I. INTRODUCTION

Understanding high-level brain functions, such as memory, learning, and cognition has long been a goal of the neuroscience community. Over the past decades, the quest for creating artificial systems that mimic these important functions has become a major challenge for the engineering community. The motivations for creating such systems are manifold; from implementing cognitive hardware systems which interact with their environments and solve real-time tasks to designing implantable prostheses that replace cognitive brain functions. However, hardware implementations are yet to accomplish the robustness and parallel processing capabilities of the brain together with its compactness and power efficiency.

Synapses are the fundamental signal processing units in the brain that transform temporal spiking patterns into analog postsynaptic potentials. Many complex processes such as neurotransmitter release and diffusion, receptor kinetics, and plasticity contribute to this transformation. Despite our knowledge of these mechanisms and the role they play in the neuron-to-neuron communication, it is still unknown how the collective activity of populations of neurons and synapses results in higher level brain functions [1]. Therefore, the level of details that have to be included in a model is difficult to

assess. There is a tradeoff between the biological complexity included in a model and its scalability and efficiency.

In this work, we want to capture the nonlinear transformation of spike inputs in the hippocampus from CA3 neurons to postsynaptic potentials by CA1 neurons. The hippocampus has been shown to be essential for long-term memory formation and capturing the temporal integration performed by the CA1 neuron is an important nonlinear mechanism in this process [1]. This mechanism is modeled using the Laguerre Expansion of Volterra Kernel method which is employed to convert spike inputs into subthreshold postsynaptic potentials [2].

This modeling approach was chosen for various reasons. First of all, this model allows us to replicate and predict the transformation of the spike inputs to the postsynaptic potential outputs in real-time and for any brain region, simply by adjusting its parameters. Even though the complex biological mechanisms underlying this transformation are not modeled in detail, the nonlinear interactions performed by the neuron and its synapses can still be captured. Furthermore, the system is efficiently implemented in hardware using filters, multipliers and gain cells. Because of its modularity, the implementation can easily be extended to a large scale system. The model parameters are programmable and adaptive processing techniques can be applied to train them [3].

In the following sections, we outline the steps taken to efficiently implement a second order nonlinear Laguerre Volterra model in hardware using subthreshold CMOS analog circuit design.

## II. MODEL DESCRIPTION

### A. Experimental Procedure

The postsynaptic potential (PSP) data used for training and testing of the proposed Laguerre-Volterra model was obtained from whole-cell patch-clamp recordings [4]. CA1 pyramidal neurons were synaptically stimulated with Poisson random interval trains (RITs) through the Schaffer collaterals to mimic the spiking behavior observed in CA3 hippocampal neurons. The mean frequency of stimulation was 2Hz and the stimulation intensity was adjusted so that no action potentials were induced. The PSPs were then recorded at the soma of the cell.

V.S. Ghaderi (email: vghaderi@usc.edu) and J. Choma (jchoma@usc.edu) are with the Department of Electrical Engineering, University of Southern California, Los Angeles, USA.

S. Roach (e-mail: shaneroa@usc.edu) is with the Department Neuroscience, University of Southern California, Los Angeles, USA.

D. Song (dsong@usc.edu), V.Z. Marmarelis (vzm@bmsr.usc.edu) and T.W. Berger (email: berger@bmsr.usc.edu) are with the Department of Biomedical Engineering, University of Southern California, Los Angeles, USA.

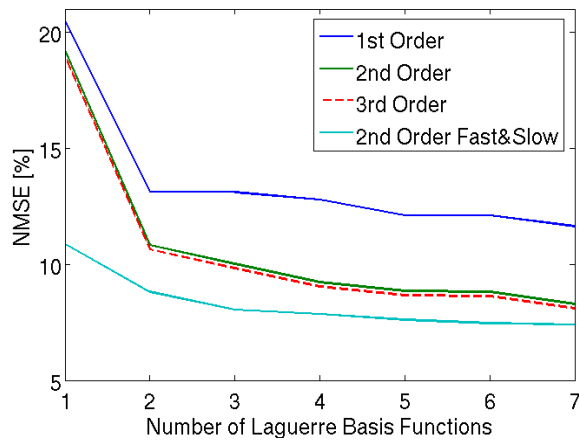


Figure 1: Percent NMSE vs. the Number of Laguerre Basis Functions for Different Orders of Nonlinearity

### B. PSP Estimation and Prediction Using LEV

The modeling approach used here to estimate and predict the neuron's postsynaptic response is similar to the one in [1] and [4]. In this work, the goal is to use a similar model to describe the nonlinear transformation from presynaptic spike to postsynaptic potential that is optimized for hardware implementation. In [5], the continuous-time Laguerre polynomial expansion method in the time domain was discussed. A weighted sum of orthonormal Laguerre basis functions (LFs) can be used to "approximate causal signals that asymptotically decay in an exponential manner for large  $t$ " [5]. The Laplace transform of a LF of order  $n$  is:

$$H(s) = H_{LP}(s)H_{AP}^n(s) \quad (1a)$$

$$H_{LP}(s) = \frac{1}{p+s} \quad (1b)$$

$$H_{AP}(s) = \frac{p-s}{(p+s)} \quad (1c)$$

Therefore, this suggests a frequency domain implementation of  $H(s)$  as shown in Fig. 1, in which the impulse response at each node is a LF and  $p$  determines the LF decay rate.

As mentioned above the PSP has a nonlinear dependence on the input action potential. The linear system of Fig. 1 will be insufficient in capturing the nonlinearities that occur on presence of paired pulses, triplets, etc. Therefore, the model has to be expanded to an LEV (Laguerre Expansion of Volterra Kernels) scheme [1] in which a time-domain estimation is described by:

$$y(t) = c_0 + \sum_{n=1}^L c_1(n)v_n(t) + \sum_{n_1=1}^L \sum_{n_2=1}^{n_1} c_2(n_1, n_2)v_{n_1}(t)v_{n_2}(t) + \sum_{n_1=1}^L \sum_{n_2=1}^{n_1} \sum_{n_3=1}^{n_2} c_3(n_1, n_2, n_3)v_{n_1}(t)v_{n_2}(t)v_{n_3}(t) + \dots \quad (2a)$$

$$\text{where: } v_n(t) = \int_0^{\infty} l_n(\tau)x(t-\tau)d\tau \quad (2b)$$

is the convolution of the  $n$ th order LF with the input  $x(t)$  (here, the random interval input spike train),  $c_m$  are weighting coefficients that can be found using linear estimation methods.

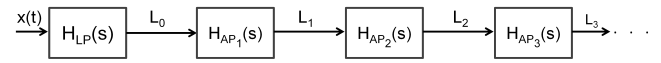


Figure 2: Block Diagram of Frequency Domain Laguerre Functions

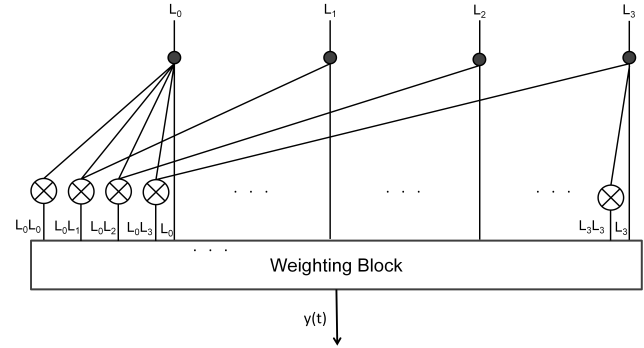


Figure 3: Block Diagram of a 2<sup>nd</sup> Order LEV Model with 4 LFs

In this work the Moore-Penrose Pseudo-Inverse is used to find  $c_m$ , which provides a least square fit of the output of the LEV to the training data set of PSPs [5].  $L$  is the number of LFs employed in the model. Equation (2) can be implemented in hardware using the linear system of Fig. 1 and the multipliers and weighting adders in Fig. 2. It is also possible to use the sum of two LEVs with different values of  $p$  to estimate the response of a system that exhibits a fast as well as a slow decaying component [6].

### C. Optimization of the LEV Model for Hardware

Since the complexity of the system grows with the number of LFs employed and the degree of nonlinearities included, there are practical limitations for implementing an  $n$ th order LEV system with a large number of Laguerre basis functions. Fig. 3 shows the variation of normalized mean square error (NMSE) in percent versus the number of Laguerre basis functions employed to estimate PSP output data. The top curve is a linear combination of LFs (first order LEV). The second one also includes the cross products (2<sup>nd</sup> order). The third curve which almost coincides with the second one includes third order products. This clearly shows that while including the second order nonlinear terms has an advantage over first order linear estimation, the third order nonlinear terms provide negligible improvement. Finally, the bottom curve shows the second order estimation when both slowly and rapidly decaying time constants are used ( $p_{fast}=50s^{-1}$  and  $p_{slow}=3s^{-1}$ ). Including two sets of LFs can significantly improve model performance. This result is expected considering the underlying biological mechanisms. The transformation performed by glutamatergic synapses from action potential to post-synaptic potentials includes fast-acting ionotropic receptors, such as AMPA and NMDA, and slow-acting metabotropic receptors, such as mGluR. In this work we implement the second order estimation with one time constant ( $p=50s^{-1}$ ). The extension to two time constants is straightforward.

### D. Hardware Implementation of the LEV Model

To implement the model described above, the first step is to build the set of LFs using the least power and area for future expansion to large-scale multi-input/output spike systems. In this work,  $H_{LP}$  and  $H_{AP}$  of (1) were implemented using OTA-C filters. The low-pass section is shown in Fig. 4.

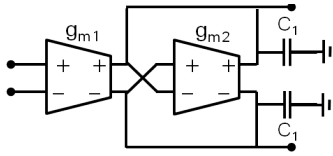


Figure 4: Low-Pass Filter Implementation

The capacitors can be connected across the output differentially to save area if a common mode feedback loop is not required. The transfer function for the low-pass filter is given by:

$$\frac{v_{out}}{v_{in}} = \frac{\frac{gm_1}{C_1}}{s + \frac{gm_2}{C_1}} \quad (3)$$

Comparing this transfer function to (1b), it can be inferred that  $p = gm_2/C_1$ . For unity dc gain  $gm_1 = gm_2$ . The identical all-pass sections are implemented using the filter structure in Fig. 5. The transfer function of the all-pass filter can be derived as:

$$\frac{v_{out}}{v_{in}} = \frac{(\frac{gm_1 gm_3}{gm_5 C_1} - \frac{gm_2}{C_1}) - s}{s + \frac{gm_2}{C_1}} \quad (4)$$

While more compact all-pass topologies exist, this particular configuration provides fully differential outputs as well as a large input impedance to minimize the loading effect on the preceding stage. To match this transfer function to (1b) we choose  $gm_2/C_1 = p$ ,  $gm_4 = -gm_5$  and  $gm_1 = gm_2$ .

### E. Design Considerations for Laguerre Filters

The  $p$ -parameter of the LFs is determined based on the decay time of the PSP signals, which is on the order of tens of milliseconds. For a given  $p = gm/C$  it is important to minimize  $C$  to reduce the on-chip area. Using high-density capacitors as well as active impedance converters can accomplish area reduction. Scaling down the  $gm$  cells with the capacitors can also shrink the size. This means a smaller current level for the OTAs. The lower limit for this current however is set by the reverse conduction current of the source/drain diode to the bulk and to a much smaller extent the leakage currents through the gate. This leakage can be reduced by using low-leakage processes or circuit design techniques that allow femto-Amp current-mode designs [7].

Ideal LFs have initial values that are reached immediately after a pulse is applied to the system. In a physiological system, however, there is a delay between the time an action potential reaches the pre-synaptic bouton and the time a post-synaptic potential can be recorded at the soma.

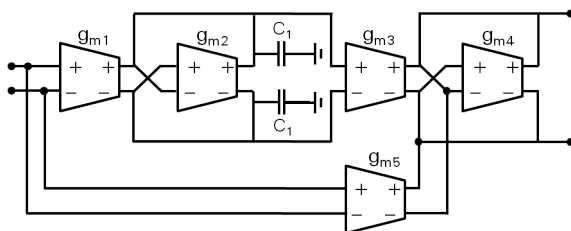


Figure 5: All-Pass Filter Implementation

This is due to several mechanisms, such as the distance that the input signal has to travel on the Schaffer collaterals to reach the CA1 synapses, the neurotransmitter diffusion through the cleft, and the summation of synaptic currents in the soma. Any hardware implementation will also have a finite rise time which mimics this effect. The only challenge here is to make sure that the rise times for all Laguerre filter sections are matched. This is difficult to achieve since LFs are generated by propagating the input through multiple filter sections and each section has parasitic poles and zeros that will limit the rise and fall time. Therefore, when designing the transistor level implementation of Fig. 4 and 5 it is important to minimize the number of parasitic poles for each transconductance cell and properly delay the output of each Laguerre section so that their initial values are reached congruently.

### E. Transistor Level Implementation

In this work, all circuit blocks were designed using CMOS transistors operating in the subthreshold region. Since the transconductance for a given current level is maximized in subthreshold it is the most power efficient region of operation. The circuit components were designed and simulated in 0.13 $\mu$ m Cypress Semiconductor CMOS technology.

The transistor level implementation of the low-pass filter is shown on the left in Fig. 6. This circuit implements (3) without creating any parasitic poles. Since the all-pass of Fig.5 includes a low-pass section, the low-pass circuit on the left in Fig. 6 can be appended by the one on the right to yield  $H_{AP}$ . This design minimizes the number of components and optimizes power consumption in each section. Bias was chosen to be 35pA for the low-pass sections along with a capacitor value of 10pF to achieve the desired  $p$  value. In this technology, the area occupied by one metal-to-metal 10pF capacitor is around 3000 $\mu$ m<sup>2</sup>. Bias was set to be sufficiently larger than the leakage levels which are in the order of a few pA plus some margin for fine tuning the time constant. This also keeps the capacitor values needed to achieve the desired time constants at a reasonable size. All input transistors are PMOS to minimize flicker noise, which is important to consider especially for low frequency biomimetic systems. Since designs in subthreshold, especially those with low current levels, are prone to great variability, this has to be mitigated with mismatch calibration techniques, which will require some additional area.

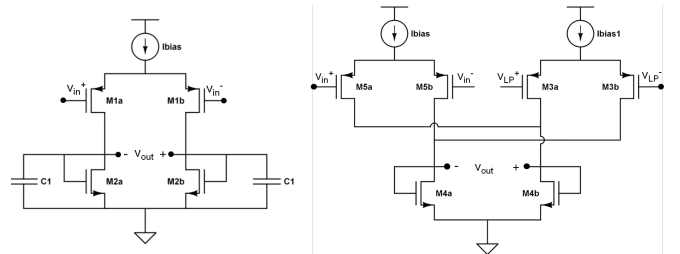


Figure 6: Transistor Level Implementation of the Low-Pass and All-Pass Filters

These sections are then cascaded as shown in Fig. 2 to achieve the first order LEV model. Each Laguerre filter output is properly delayed in order to align their rising edges. To implement the second order terms of (2), the outputs of

the Laguerre filters were cross-multiplied using ten Gilbert multipliers, each with a bias current of 100pA. In order to obtain the full second order LEV model, the outputs of the first and second order PSD systems have to be weighted and added, which can be achieved using programmable weighting multipliers whose output currents are tied together.

### III. RESULTS

The simulated impulse response of the circuit with matched rise times of each Laguerre filter is shown in Fig. 7. The input is a series of pulses of 200 $\mu$ s duration and amplitude of 800mV applied differentially, resembling ideal impulses. The rms error between the signals generated by the circuit and ideal LFs is less than 5%. The weighted sum of these LFs along with its cross-products provides a second order LEV system. Fig. 8 shows the performance of the ideal hardware LEV model in estimating the PSP recordings. The top part of the graph shows the recorded PSP data due to a random impulse interval. The second graph shows the LEV hardware response to the same random input using a second order LEV implementation, and as a comparison, the third graph shows the ideal second order LEV approximation. The bottom graph shows the absolute value of the difference between the recording and the circuit output. The normalized mean square error between circuit estimation and data is 14.98%. The total power consumption is 5.3nW. It can be inferred that the approximation does not yet completely capture all the effects of synaptic temporal integration. As mentioned above, the system performance can be improved significantly by including a second set of LFs with a slow time constant. In a practical implementation an automatic calibration system should be included to eliminate the sensitivity to excessive offsets and mismatches.

### IV. CONCLUSION

A second order LEV model was implemented in low power hardware to approximate and predict intracellular subthreshold potentials evoked by random interval data. The estimation was evaluated through comparison to real data using normalized mean square error. The system was designed in a modular way to allow for easy scaling and programming. The model parameters can be adjusted by changing the bias voltages and currents in the circuit.

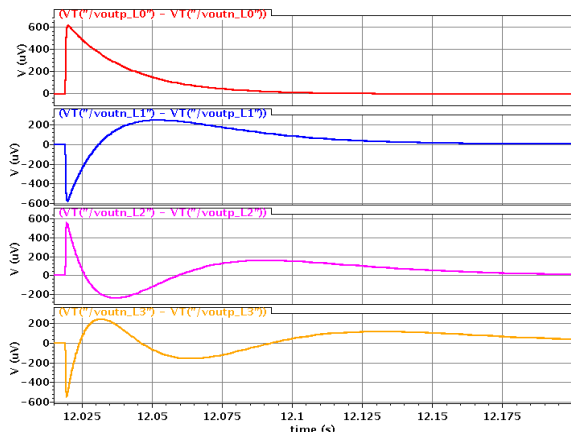


Figure 7: Impulse Response of Laguerre Filter Circuits

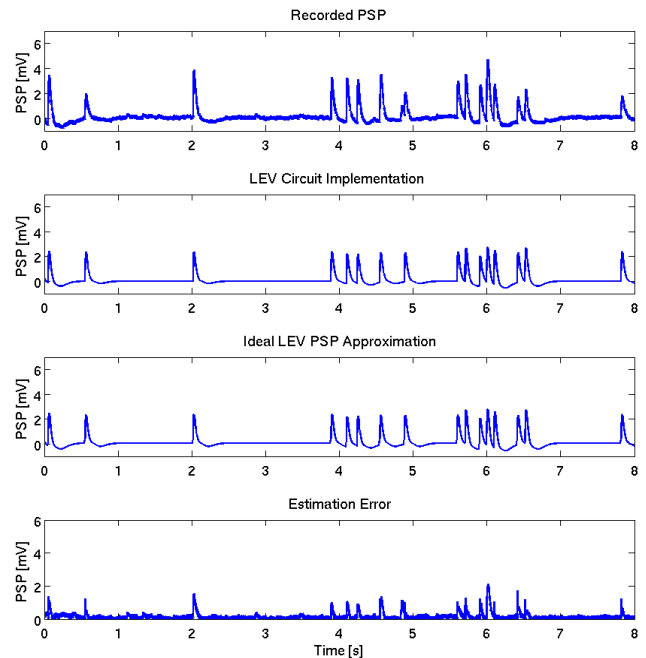


Figure 8: Comparison between Recorded Data and the Second Order LEV Hardware System Estimation

Therefore, the system can be retrained for various types of data and application. The compactness of the modeling approach in addition to the nano-Amp implementation of the system allows for effective scaling to large-scale neuron models. Improvement in error can be achieved by including another set of LFs with a different time constant.

### ACKNOWLEDGMENT

The author would like to thank Dr. Hasler of the Georgia Institute of Technology for valuable discussions and suggestions on subthreshold circuit design and Cypress semiconductor for access to their design kit.

### REFERENCES

- [1] T.W. Berger, D. Song, R.H. Chan, and V.Z. Marmarelis, "The Neurobiological Basis of Cognition: Identification by Multi-Input, Multi-Output Nonlinear Dynamic Modeling," *Proc IEEE Inst Electr Electron Eng*, vol. 98, (no. 3), pp. 356-374, Mar 2010.
- [2] V.Z. Marmarelis, "Identification of nonlinear biological systems using Laguerre expansions of kernels," *Ann Biomed Eng*, vol. 21, (no. 6), pp. 573-89, 1993 Nov-Dec 1993.
- [3] R.H. Chan, D. Song, A.V. Goonawardena, S. Bough, J. Sesay, R.E. Hampson, S.A. Deadwyler, and T.W. Berger, "Tracking the changes of hippocampal population nonlinear dynamics in rats learning a memory-dependent task," *Conf Proc IEEE EMBS*, pp. 3326-9, 2011.
- [4] U. Lu, D. Song, and T.W. Berger, "Nonlinear dynamic modeling of synaptically driven single hippocampal neuron intracellular activity," *IEEE Trans Biomed Eng*, vol. 58, (no. 5), pp. 1303-13, May 2011.
- [5] V.S. Ghaderi, S.L. Allam, N. Ambert, J.M. Bouteiller, J. Choma, and T.W. Berger, "Modeling neuron-glia interactions: from parametric model to neuromorphic hardware," *Conf Proc IEEE EMBS*, pp. 3581-4, 2011.
- [6] G.D. Mitsis, R. Zhang, B.D. Levine, and V.Z. Marmarelis, "Modeling of nonlinear physiological systems with fast and slow dynamics. II. Application to cerebral autoregulation," *Ann Biomed Eng*, vol. 30, (no. 4), pp. 555-65, Apr 2002.
- [7] B. Linares-Barranco and T. Serrano-Gotarredona, "On the design and characterization of femtoampere current-mode circuits," *IEEE JSSC*, vol. 38, (no.8 ), pp.1353-1363, Aug 2003.
- [8] R. Serrano-Gotarredona, L. Camuas-Mesa, T. Serrano-Gotarredona, J.A. Leero-Bardallo, and B. Linares-Barranco, "The Stochastic I-Pot: A Circuit Block for Programming Bias Currents," *IEEE Trans Circuits II*, vol.54, (no.9), pp.760-764, Sept 2009.