Wide dynamic range weights and biologically realistic synaptic dynamics for spike-based learning circuits

Dora Sumislawska, Ning Qiao, Michael Pfeiffer, and Giacomo Indiveri
Institute of Neuroinformatics
University of Zurich and ETH Zurich
Zurich, Switzerland
Email: [dora,ning,pfeiffer,giacomo]@ini.uzh.ch

Abstract—Spike-based neuromorphic learning circuits typically represent their synaptic weights as voltages, and convert them into post-synaptic currents so that they can be integrated by their afferent silicon neuron. This voltage-to-current conversion is often done using a single transistor. This results in an exponential (for weak-inversion) or quadratic (for strong inversion) non-linear transformation which severely restricts the type of learning algorithms that can be implemented. To overcome this problem we propose a range of solutions that perform a linear transformation from weight voltage to synaptic current, simplifying the implementation of a spike-based learning rules. We demonstrate the application of these conversion circuits using current-mode integrators that produce alpha-functions with biologically realistic temporal dynamics and amplitudes that are linearly proportional to the synaptic weights. The circuits proposed are low-power, and can be integrated in a wide range of spike-based learning frameworks that have been recently proposed. We describe the advantages and disadvantages of the various solutions proposed and validate them with circuit simulation results.

I. INTRODUCTION

Neuromorphic Very Large Scale Integration (VLSI) electronic circuits can be used to faithfully mimic the neural computation principles found in real neural systems [1, 2]. In recent years particular attention has been dedicated to the design and development of spike-based learning neuromorphic circuits [3–8]. In most of the solutions proposed the synaptic weight is represented as a voltage stored on a capacitor or floating gate, which is then converted into a current that is injected into the neuron membrane capacitance either directly, or indirectly, via a pre-processing stage such as a current-mode low-pass filter that implements synaptic temporal dynamics [9]. If the conversion from voltage to current is done using a single transistor (as is often the case [3–8]), this results in a non-linear transformation which is exponential if the transistor operates in weak-inversion or quadratic if the transistor operates in strong-inversion [10]. In both cases the non-linearity interferes with the learning process: the weight update operation typical of all learning mechanisms \( w(t + \Delta t) = w(t) \pm \Delta w \) produces a change of synaptic efficacy current which depends on the initial weight voltage: \( \Delta I_w(w) = f(V_w, \Delta w) \), such that the gain in current \( \Delta I_w(w) \) is no longer proportional to \( \Delta w \).

While there are examples of spike-based weight dependent learning mechanisms in the literature (e.g., with negative exponent exponential relationships [11]), the type of weight dependence produced by both subthreshold and above threshold conversion methods described above (with positive exponent or quadratic relationships) is severely limiting the usefulness of any learning mechanism: weight updates applied to low weight voltages have very little effect on the synaptic efficacy current, while the same weight updates applied to high weight voltages have a very strong and possibly dominant effect on the synaptic current produced. In order to overcome these limitations, we present a set of voltage-to-current conversion circuits that allow to apply the weight update rules over a wide dynamic range of weight voltages. Furthermore, to demonstrate how the linear weight-update current can be used in a neuromorphic learning application, we model the synaptic dynamics using a dual Differential Pair Integrator (DPI) circuit [9] configuration that produces biologically realistic alpha functions.

The overall block diagram describing the different circuits required to implement a spike-based learning neuromorphic neuron is shown in Fig. 1. Spike-based learning circuits produce pulses triggered by the pre-synaptic spike, the postsynaptic one, or both. These pulses produce weight updates either increasing or decreasing the synaptic weight voltage stored on a capacitor. This voltage is finally converted into a synaptic efficacy current. Currents produced by multiple copies of these blocks are summed spatially together and used as
input to the dual DPI synapse, which implements biologically plausible temporal dynamics. The total post-synaptic current is then fed into a neuron circuit which integrates the current and produces action potentials (e.g., see [8] for an example of a chip that implements this type of architecture).

In Section II we propose different circuits to perform a linear weight voltage to synaptic efficacy current conversion. In Section III we describe the operation of the dual DPI circuit, which produces the alpha-function-shaped outputs. In Section IV we present circuit simulation results, showing the effect of spike-based weight-updates on the current, and the response properties of the dual DPI synapse. Finally in Section V we highlight the potential benefits of the circuits proposed, discuss about their drawbacks, depending on the requirements, and draw concluding remarks.

II. LINEAR WEIGHT CONVERSION CIRCUITS

We present four analog circuits that perform a linear conversion from weight voltage to synaptic efficacy current. The requirements for all circuits are that they operate with minimum power consumption and they have a wide dynamic range for the input/output signals. Since the circuits are part of a negative feedback loop that involves learning, requirements on precision are less stringent. To ensure low power consumption, all circuits designed use transistors in the subthreshold domain [10].

The circuit of Fig. 2a represents a classical source follower. The standard use of this circuit is to read-out the source potential of the transistor $M_W$, which follows the changes in its gate voltage $V_W$. However, here we propose to use it in an unconventional way by reading as output the drain current of $M_W$, which should be linear with the voltage $V_W$. Figure 5 shows the transfer function of this circuit. There are three main regions of operation: for very low inputs (e.g. below the bias voltage $V_1$), $M_1$ operates in its ohmic region and the output current $I_W$ is close to zero; as $V_W$ rises beyond $V_1$, the $M_1$ transistor enters its saturation region, and the output current $I_W$ starts changing linearly with the input voltage, exploiting the transistor’s Early effect [10]. There are two distinct linear regions, with two different slopes because of the transition from weak- to strong-inversion regimes.

Figure 2b extends the basic source follower circuit (i.e., “SF” block) with an additional branch with two transistors. The $M_2$ transistor is used to produce a constant bias current, while $M_C$ is used as a cascode transistor which clamps the $V_C$ node and ensures that the current sourced by $M_2$ is constant. This extension shifts the operating region of the source follower and extends its linear range. Circuit simulation results of this circuit are shown in Fig. 5.

The circuit of Figure 3 makes use of a transconductance amplifier in negative feedback mode to modulate the drain-source voltage difference of $M_1$. The amplifier output is connected to the source-follower input, and the source-follower output is connected to the negative input of the amplifier. This configuration ensures that $V_W$ faithfully follows changes in $V_W$. As this imposes a $V_{DS}$ of $M_1$ equal to $V_W$, the circuit can operate correctly over a wide range of input voltages (see Fig. 5), by biasing $M_1$ in the saturation region, and exploiting the Early effect. To compensate for off currents and to shift the...
Fig. 5: Simulation results of the proposed weight voltage to current conversion circuits. The different traces represent the transfer function of the different circuits as a function of input voltage. The plot highlights the differences in dynamic range and linearity among the different circuits. Results were obtained for following bias settings: SF $V_1 = 460 \text{mV}$; SFC $V_1 = V_3 = 570 \text{mV}$, $V_2 = 1.3 \text{V}$; TSFC $V_1 = 540 \text{mV}$, $V_2 = 1.2 \text{V}$. Linearity and dynamic range properties of each circuit are quantified in Table 1.

To improve the transfer function of the circuit, we added a p-type transistor that sources a constant current set by $V_2$.

Figure 4 shows an alternative proposal based on coupled differential pair circuits. A single differential pair biased in subthreshold is characterized by a voltage-to-current transfer function that is sigmoidal, with a linear range that spans only a few hundred of milli volts [10]. The circuit proposed here extends this range by combining multiple differential pairs, biased with the same reference voltages $V_{bn}$ and $V_{bp}$. There are both n-type and p-type differential pairs to ensure that there will always be a differential pair with its transistors operating in the saturation region. We chose six differential pairs as a good compromise between linear range desired and circuit size. The threshold voltages $V_{1-6}$ are set in a way to equally distribute the centers of the differential-pair linear ranges (see Fig. 5 for circuit simulation results).

### III. ALPHA-FUNCTION SYNAPTIC CIRCUIT

The circuits proposed in the previous section can be useful for implementing instantaneous (linear) weight updates. However, to faithfully model the temporal dynamic properties of real neurons, additional circuits are necessary. Here we present an example of a circuit that can produce temporal dynamics with biologically plausible time constants (i.e., in the order of tens of ms) and realistic synapic alpha-function profiles, with an amplitude that scales linearly with the synapse weight voltage (see Fig. 6). The circuit, is based on the use of two coupled DPIs that subtract their output currents to produce a difference of exponential-like characteristic, which closely resembles an alpha function: $I_{\text{syn}} = I_{\text{DPPI}} - I_{\text{DPPII}}$, with

$$\tau [1.2] \frac{d}{dt} I_{\text{DPPI}[1.2]} + I_{\text{DPPI}[1.2]} = \frac{I_{w[\text{exc,inh}]} I_g}{I_{\text{tau}}},$$

where $I_g$ and $I_{\text{tau}}$ are two currents proportional to $V_{\text{thr}}$ and $V_{\text{tau}}$ respectively, and $\tau$ is a time constant directly proportional to $C$ and inversely proportional to $I_{\text{tau}}$ (see [9] for details).

Both DPI circuits share the same bias voltages and input signals. The only difference is the relative size of the capacitor, which translates to a difference in their time constants. Figure 7 shows the circuit’s impulse response, simulated for different values of synaptic weight voltage ($V_w$) and time-constant bias settings ($V_{\text{tau}}$). If the synapses of an afferent neuron share the same dynamics, it is sufficient to use one single dual DPI circuit that receives in input the sum of all synapse currents, exploiting the linear superposition principle.

### IV. WEIGHT UPDATES WITH TEMPORAL DYNAMICS

In order to test the combination of weight voltage to synaptic efficacy conversion circuits with the alpha-function dual DPI circuit we simulated an arbitrary learning rule that produces positive constant weight updates on the arrival of a pre-synaptic spike, for the first six spikes, and negative weight updates for the following six spikes. Figure 8 shows both the induced weight voltage changes and the synaptic responses over time of three different voltage-to-current conversion circuits: SFC, TSF, and MDP. All synaptic responses exhibit the same alpha function profile shown in Fig 6, but have different maximum amplitudes, due to their different linear conversion.

Fig. 6: Dual DPI circuit. Two differential pairs are coupled together to produce a difference of exponentials output current.

Fig. 7: Dual DPI impulse response as a function of different weight values (a) and time constant reference voltages (b). The amplitude of the alpha-function is linearly proportional to the weight values used. Changes in $V_{\text{tau}}$ modulate the circuit’s time constant.
properties. These maximum amplitudes are highlighted by the
dashed-dotted lines in the bottom quadrants. The figures in all
quadrants show also spurious peaks and glitches which are due
to simulation artifacts.

V. CONCLUSION AND DISCUSSION

In this work, we presented a set of circuits for implement-
ing linear synaptic weight voltage to synaptic efficacy current
conversion and for producing alpha-function post-synaptic
current response profiles with biologically plausible time-
constants. The proposed circuits range from more compact
ones to more accurate, area-expensive ones. Wider linear range
performance circuits, such as the MDP circuit of Fig. 4, require
larger area and power consumption.

We have shown that by exploiting the Early voltage in
the transistors even a basic source follower circuit (see SF
circuit in Fig. 2a) can produce linear voltage to current
transformation, albeit with a limited linear range and two
different slopes of its I-V characteristic. These non-idealities
however are less critical than the exponential or quadratic non-
linearities introduced by single-transistor conversion schemes.
This circuit could therefore prove to be very useful, if the
learning rule in which it is embedded in, is able to account for
its shortcomings. We showed, with the SFC circuit of Fig. 2b
how using two additional transistors allowed us to compensate
for some of these shortcomings and shift the operation of
the current source into the saturation region. The drawback
of using these additional transistors is that some fraction of
the consumed power is wasted, because it is used only for
shifting currents or canceling out leakage currents. The TSF
circuit of Fig. 3 provides the best compromise between area
consumption and linear range, which however requires a high-
gain active component (the amplifier) to improve its linear
output range.

Although the DPI and analogous circuits have been used in
the past to implement synaptic dynamics, this is the first time
that such circuits are combined to emulate the more realistic
alpha-function response profile of real synapses. We showed
how the current-conversion circuits can be directly interfac-
ted to the alpha-function circuit to implement synaptic dynamics
response properties with amplitudes that are linearly propor-
tional to the synaptic weight voltage. The circuits presented

<table>
<thead>
<tr>
<th>Circuit</th>
<th>RMS error</th>
<th>Area (µm²)</th>
<th>Power consumption/spike (µJ)</th>
<th>Static power consumption (µJ)</th>
<th>Dynamic range (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>2.11</td>
<td>494.9</td>
<td>0.388</td>
<td>0.007</td>
<td>0.17</td>
</tr>
<tr>
<td>SFC</td>
<td>3.12</td>
<td>614.3</td>
<td>1.124</td>
<td>0.985</td>
<td>1.13</td>
</tr>
<tr>
<td>TSFC</td>
<td>4.14</td>
<td>676.2</td>
<td>0.014</td>
<td>0.682</td>
<td>1.35</td>
</tr>
<tr>
<td>MDP</td>
<td>2.07</td>
<td>784.6</td>
<td>0.213</td>
<td>0.156</td>
<td>1.6</td>
</tr>
</tbody>
</table>

have been fabricated using a standard 180 nm Complementary
Metal-Oxide-Semiconductor (CMOS) process. We can now
measure the size versus power consumption trade-offs as well
as their linearity and robustness (e.g., to device mismatch
effect). A comparison of the different circuit features and
performance metrics obtained from the circuit simulations is
shown in Table I

ACKNOWLEDGMENT

We thank Chiara Bartolozzi, Simeon Bamford and the
CapoCaccia 2015 workshop participants for fruitful discus-
sions on the circuit architectures and collaborations. The
research was supported by the Swiss National Science Foun-
dation Grant 200021_146608, and the European Union ERC
Grant “Neurop” (257219).

REFERENCES

autonomous cognitive systems”. In: Proceedings of the IEEE
102.9 (Sept. 2014), pp. 1367–1388. ISSN: 0018-9219.
ings of the IEEE 78.10 (1990), pp. 1629–36.
in Silicon: Design, Implementation, Application, and Chal-
lenge”. In: Proceedings of the IEEE 102.5 (May 2014),
pp. 717–737. ISSN: 0018-9219.
[4] A. Basu et al. “Silicon spiking neurons for hardware imple-
mentation of extreme learning machines”. In: Neurocom-
- Constraints on enabling spike-timing dependent plasticity
in neuromorphic hardware”. In: Frontiers in Neuroscience 6
(2012). ISSN: 1662-453X.
Synapses With Spike-Time-Dependent Plasticity”. In: Biomed-
icial Circuits and Systems, IEEE Transactions on 5.3 (2011),
pp. 244–252. ISSN: 1932-4545.
configurable analog VLSI neural network of spiking neurons
and self-regulating plastic synapses”. In: Neural Computation
Neuromorphic Processor comprising 256 neurons and 128K
synapses”. In: Frontiers in Neuroscience 9.141 (2015). ISSN:
1662-453X.
VLSI”. In: Neural Computation 19.10 (Oct. 2007), pp. 2581–
2603.
Decision Making.” In: Neural Computation 22.6 (Feb. 2010),
pp. 1399–1444.