POLITECNICO DI TORINO

Master of Science in Nanotechnologies for ICTs

Master Degree Thesis

TUNABLE SPIKE-FREQUENCY ADAPTATION IN ORGANIC ARTIFICIAL NEURONS

Supervisors:
Prof. Carlo RICCIARDI
Prof. Alberto SALLEO

Candidate:
Edoardo TENNA

ACADEMIC YEAR 2022-2023

Abstract

Neuromorphic computing is one of the most promising computing paradigms to build the next generation of energy-efficient computing systems. Many materials and devices are being investigated to develop these new architectures. Organic electrochemical transistors (OECTs) have emerged as candidate devices for the design of artificial synapses. Their ability to emulate a variety of neurological mechanisms, such as short-term and long-term plasticity, their operation relying on ionic transport and the possibility for global connectivity make them intrinsically similar to neurobiological membranes. Moreover, they present low switching energies and a wide range of tunability. Other neuromorphic components must be created with a similar technology to build an entirely organic neuromorphic system for spiking neural networks. Indeed, organic circuits that emulate biological neurons are crucial. Few examples of artificial neurons fabricated with OECT technology exist in literature. They emulate some spiking features of biological neurons with neurotransmitter or ion-based modulation. However, they all spike at a constant frequency for a constant applied stimulus and fixed operating conditions, thereby having limited neural encoding capability. Spike-frequency adaptation (SFA) is a fundamental neuronal mechanism that encodes information by modulating firing activities: adaptive neurons show an initial high frequency spiking activity at the onset of a constant stimulus, that gradually reduces with time to a steady-state response. In this thesis, a new circuit architecture to emulate spike-frequency adaptation in an all-organic artificial neuron is proposed. This circuit relies on the short-term plasticity property of OECTs and the possibility to fabricate OECTs with a secondary tuning gate, that acts on its transcharacteristic. The circuit is demonstrated first in a hybrid version, with combined OECT and standard CMOS technologies. Subsequently, an all-organic version of the circuit is proposed. The circuit operates at a supply voltage $< 1V$ and shows tunable spike-frequency adaptation. All the device modelling and circuit simulations have been carried out on LTSpice.

Contents

Abstract **i**

1 Introduction to neuromorphic engineering **1**
1.1 Origin and evolution of neuromorphic systems 1
1.2 Artificial neurons 4
1.2.1 Spike-frequency adaptation 6
1.2.2 Adaptive spiking neuron models 8
1.2.3 Adaptive spiking neuron circuits 12

2 OECTs **15**
2.1 Physics and operation 15
2.2 Steady-state response and scaling 17
2.3 Transient response 20
2.4 Double-gate architecture 29

3 Hybrid organic-inorganic adaptive neuron **31**
3.1 Adaptive double-gate OECT optimization 31
3.2 Circuit architecture 35
3.2.1 Axon-Hillock circuit 36
3.2.2 Adaptation block and integration 37
3.3 OECT fabrication 40
3.4 Double-gate OECT characterization 43
3.5 Double-gate OECT modeling 48
3.6 Circuit simulation 53

4 All-organic adaptive neuron **58**
4.1 Complementary OECTs modeling 58
4.2 Circuit simulation 61

5 Conclusion **66**
5.1 Future perspectives 66

List of acronyms

Abbreviation		Meaning
AdEx	=	Adaptive exponential integrate-and-fire
BiCMOS	=	Bipolar Complementary metal-oxide-semiconductor
e-beam	=	Electron beam
ELR	=	Extrapolation in the linear region method
FET	=	Field-effect transistor
IMC	=	In-memory computing
IF	=	Integrate-and-fire
ISI	=	Interspike interval
LIF	=	Leaky integrate-and-fire
LOL	=	Lift-off layer
LPF	=	Low-pass filter
LTP	=	Long-term plasticity
MOSFET	=	Metal-oxide-semiconductor field-effect transistor
ODE	=	Ordinary differential equation
OECT	=	Organic electrochemical transistor
PVT	=	Process, Voltage, Temperature
RIE	=	Reactive ion etching
SFA	=	Spike-frequency adaptation
SNN	=	Spiking neural network
SoA	=	State of the Art
STP	=	Short-term plasticity
TSMC	=	Taiwan Semiconductors Manufacturing Company
VLSI	=	Very large-scale integration

Chapter 1

Introduction to neuromorphic engineering

Neuromorphic engineering aims to develop computing hardware that mimics biological nervous systems.

In this Chapter, we give a brief overview on the field of neuromorphic engineering. In Sec. 1.1, we retrace the motivations that led to its creation in the 1980s. We describe the current scenario, where neuromorphic engineering has emerged as an interdisciplinary research field that includes a much broader scientific community. We point out how that can be attributed recently to simultaneous developments in materials and devices for ICs, the need for low power, sustainable AI and the approaching end of Moore's law.

In Sec. 1.2, we focus on a specific block of spiking neural networks (SNNs) circuits: artifical neurons. The discussion is aimed at introducing a biological, mathematical and technological context to the objective of this thesis, which is the development of an adaptive spiking architecture based on organic materials. We introduce the mechanism of spike-frequency adaptation (SFA) and discuss its computational significance in biological neural networks. We describe a series of available models from computational neuroscience that are able to reproduce spike-frequency adaptation and their level of complexity. We conclude by identifying some adaptive artificial neurons present in literature that have been built with different technologies.

With this introduction, we hope to impart the reasons why the development of adaptive spiking circuits is a defining step for emerging technologies towards the realization of more bio-realistic neuromorphic hardware. The introduction of an adaptive organic artificial neuron addresses this need for the field of organic electronics.

1.1 Origin and evolution of neuromorphic systems

The origin of neuromorphic engineering dates from the 1980s, by Carver Mead and his students at Caltech: the Physics of Computation group [1].

The field was created at a time when digital synchronous circuits were just beginning to dominate silicon production. Very large-scale integration (VLSI) technology enabled the growth of integrated circuits at a fast pace. Coherently with Moore's law, the number of transistors per integrated circuit was doubling every two years, as new technological nodes were developed to be smaller and smaller. The increasing demand for computational power was addressed by new product generations with smaller feature sizes, that did not require new conceptualization at the architecture level.

However, limits of technological scaling were already well understood. As transistors' sizes were being pushed down (to single digit μm scale at the time), silicon technology was getting closer to the fundamental physical limits. Much range was still to be exploited at that time, but it is now an issue more relevant than ever.

This was the premise to start rethinking the principles of computation, leading Mead and his group to look at the brain as source of inspiration.

The crucial consideration to start investigating neural computation was that biological information-processing systems are several orders of magnitude more efficient than any digital system ever built, at solving many classes of problems. Especially for those problems that require to process poorly conditioned input data, the difference in power consumption can be a factor of $10^7 - 10^9$. The origin of this efficiency gap is not to be searched at the device level. The energy per operation of a state-of-the-art (SoA) minimum-size MOSFET[1] is around $10^{-15} - 10^{-16} J$ nowadays, and between $10^{-13} - 10^{-16} J$ for a biological synapse or cell [2]. Rather, it is to be attributed at a more fundamental level: to the level of interconnection and even more to the representation of information [3].

The first point is addressed by making algorithms more local and designing hardware with memory function embedded inside the processing unit. This solution can be implemented either in analog or digital technology and is the main focus of In-memory computing[2] (IMC) [4, 5].

The representation of information is where the greater portion of efficiency is lost. In a conventional digital system information is confined into a binary representation of absolute values. Bits are computed through the elementary functions of Boolean algebra, which are digital gates. These elementary blocks are then combined together to create a great set of operations which are the ones used by algorithms.

[1] The energy per operation of a MOSFET is the energy it takes to fully charge or discharge its gate.

[2] In-memory computing is a broad field that includes all computational paradigms and architectures where data is processed inside the memory. Unlike conventional computer architectures, these systems avoid the throughput limitation and energy consumption caused by transferring data back and forth between memory and CPU. This limitation is known as "von Neumann bottleneck" or "memory wall". Neuromorphic engineering is a possible paradigm for in-memory computing.

This process is functioning but inefficient. First, because it forces a mathematical representation which is not the one of the device used: the entire operational capability of a transistor is reduced to a single bit value. Second, because the elementary operations that arise from this representation are not necessarily the ones used in the computation, which are then built artificially starting from these initial blocks. To overcome this limitation, new architectures have to be invented that exploit directly the set of computational primitives that arise naturally from the substrate.

The invention of neuromorphic engineering emerges from the analogy between the computational primitives of biological neural networks and the ones of analog (mixed-signal) VLSI [6, 7]. Both systems use state variables that are analog, represented by an electrical charge. Some neural primitives are also intrinsic to analog VLSI circuits: conservation of charge is naturally defined by Kirchoff's first law at zero cost, which implements a distributed addition; integration through time is the equation of a capacitor; thresholding, amplification, compression, all are distinctive both of neural systems and of the physics of silicon circuits. Above all, exponentiation of a current depending on a voltage is the physical result of populations of ion channels in the nerve membrane and of a transistor operated in weak inversion. Indeed, the main goal of the original neuromorphic engineering approach was to emulate neural computation by exploiting a medium whose physical behavior is analogous to that of biological nervous systems.

The development of neuromorphic systems was focused (and still is) to have both a scientific and a practical impact.

Neuromorphic circuits can be used as an emulation platform to study broad neural networks and understand the significance of their computational mechanisms. The choice over digital simulation (on a conventional computer architecture) is based on a number of factors. In general, digital simulation is precise, provides complete data that can be easily displayed, and is very flexible, as the network properties can be changed by just modifying and recompiling a code. Emulation is much more effective for very large networks, when simulation time, computational weight and energy consumption cannot be sustained by a digital computer. However, it is less accessible than simulation to the general research environment and requires specialized skills on the part of the designer.

The second objective was to build spatially dense, low power and highly parallel sensory processing and computing units. These systems are structurally predisposed to interface with the environment, being built on the same organizing principles of the brain. They eliminate noise at the single unit level by feedback from a collective signal at the next step of computation. They rely on adaptation and long-term learning mechanisms to extract information from noisy input data, that is unreliable in specific but meaningful overall. These same mechanisms enable them to perform computation in the presence of unit-to-unit variability and component failure.

This last property is what caused a renewed interest in neuromorphic engineering in more recent times, with the development of new memory technologies. These emerging memory devices exploit a broad range of materials and phenomena to store information as a relative resistance value (thus named resistive memories or memristors), instead of conventional charge-based memories [5]. They can be programmed in multiple resistance levels through the application of electrical SET and RESET pulses. However, they suffer from significant inter-device variability and variability across an array. This supports the integration in neuromorphic architectures, where these variations could be mitigated by adaptive collective processing.

At the same time, the rise of increasingly data-intensive applications, such as AI, demands for higher computational power and energy-efficiency hardware technology. It is foreseen that this demand will not be met by conventional von Neumann architectures [8].

In this social and technological context, neuromorphic engineering and computing is growing as a broad and diverse community aiming to build the next generation computer technology. The roadmap for its development is an interdisciplinary effort investigating new materials, devices, circuit architectures, algorithms supported by new advancements in the field of neuroscience [9, 10].

1.2 Artificial neurons

In spiking neural networks (SNNs), artificial neurons are hardware implementations that emulate the electro-physiological behavior of biological neurons [11]. They are usually designed to reproduce or take inspiration from the mathematical models of spiking neurons developed by computational neuroscience.

Spiking neuron circuits and models share the same classification criteria: they are divided between conductance-based circuits/models and phenomenological circuits/models. These two classes both emulate the evolution of the voltage across the cell membrane and the generation and transmission of action potentials. However, they differ in their level of detail, their ability to describe different populations of neurons and spiking behaviors, and the biological mechanisms they emphasize.

Conductance-based models/circuits are detailed biophysical representations that attempt to capture biophysical details of the neuron's membrane dynamics. They are based on the equivalent circuit representation of a cell membrane introduced by Hodgkin and Huxley in 1952 [12–16] (Fig. 1.1). In their simpler version, these models represent a neuron as a membrane capacitance in parallel to several conductances each with a series driving force (DC voltage generator): the membrane capacitance is the one of the lipid bilayer; the two nodes of the circuit are equivalent

to the intracellular and extracellular environments; the set of conductance-generator branches models the variety of ionic channels across the membrane (or the leakage conductance through the membrane); the driving potential for each branch is the Nernst potential[3] of that ion specie across the cell membrane. Each conductance is defined by a system of differential equations in mathematical models, and by a specific electronic block in spiking circuits. These representations are considered more physiologically realistic than their phenomenological counterparts because they incorporate the known biophysical properties of neurons. However, conductance-based models are usually more computationally expensive, more and more for an increasing degree of complexity. This complexity is usually transmitted from the model to the hardware as well, since more and more sophisticated differential equations usually require the design of a larger number of electronic blocks.

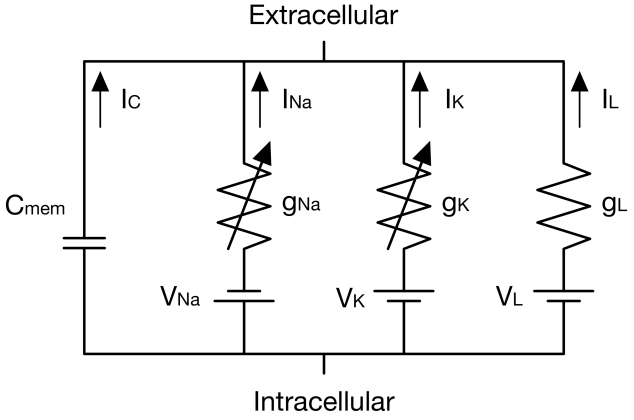


Figure 1.1: Equivalent circuit diagram of the Hodgkin-Huxley model.

Phenomenological models/circuits are simplified abstractions of neurons' behavior. They focus on capturing the overall transfer function of neuronal activity without explicitly representing the underlying biophysical mechanisms. Compared to conductance-based equivalents, phenomenological models involve fewer equations and parameters, making them suitable for large-scale networks simulation. For circuits, their size and power consumption is related to the specific design.

The choice between these models/circuits depends on the specific goals of the study and the computational/design-fabrication resources available.

In the next subsections we will focus on a specific spiking behavior, which is spike-frequency adaptation (SFA). We will introduce a series of phenomenological models that are able to reproduce it. We will then report examples of adaptive spiking circuits developed with different technologies.

[3]The Nernst potential is the electric potential that balances the movement of a specific ion across a biological membrane due to the concentration gradient of that ion.

1.2.1 Spike-frequency adaptation

Spike-frequency adaptation (SFA) is a biological mechanism where a neuron's response to a constant or repetitive input decays over time: the neuron's spiking frequency is maximum at the onset of the constant stimulus and then decreases at each spiking event towards a steady-state value.

SFA is a fundamental aspect of neural processing because it has important implications for how the nervous system encodes information: SFA enhances the ability of the nervous system to detect and prioritize novel and dynamic stimuli, that will cause higher firing rates, while filtering out redundant or static information. This mechanism is equivalent to a high-pass filter operation on the neuron's input [17,18]. This selectivity is essential for emphasizing dynamic sensory features in the environment. For example, it improves the ability of projection neurons in insects to encode the rate of change (gradient) of olfactory information [19]. Also, it was demonstrated how networks of simple Leaky integrate-and-fire (LIF) neurons with adaptation can be trained to process temporally-dispersed information and perform complex operations of sequence recognition [20]. Spike-frequency adaptation is also involved in maintaining stability in the firing rates across a neural network and contributes to energy efficiency of neural processing (by reducing the firing rate in response to sustained stimuli, neurons can conserve energy).

SFA is a property occurring with a slower dynamic compared to spike generation. A series of biophysical mechanisms can cause spike-frequency adaptation [21] and they all include a form of slow negative feedback to the excitability of the cell. In general, they can be due either to inactivation and slow recovery of Na channels after depolarization[4] (inactivation of depolarizing currents), or to activation of slow spike-dependent or voltage-dependent hyperpolarizing and shunting currents[5].

A universal model to analyze SFA was introduced in [21]. Adaptation is measured as the change in the instantaneous spiking frequency $f(t)$ of the neuron. As long as the stimulus is supplied, the instantaneous spiking frequency is defined as the reciprocal of the difference between the spiking times of the last two occurring spikes:

$$f(t) = \frac{1}{t_{spk}^{i} - t_{spk}^{i-1}} \quad , \quad t_{spk}^{i} \leq t < t_{spk}^{i+1} \quad , \quad i > 1 \tag{1.1}$$

The difference between two consecutive spiking times is called an interspike interval

[4]Na channels in the cell membrane are responsible for the depolarization of the cell and the generation of action potentials. Slowly recovering Na channels after a spike event cause a delayed generation of the following spike.

[5]Spike-dependence refers to the activation of a population of ion channels based on the spiking activity of the cell. Voltage-dependence is a dependence on the cell membrane voltage. Hyperpolarizing currents are current contributions that tend to bring the membrane potential at a lower value than its resting potential. They have a subtractive effect on the excitatory potential. Shunting currents also reduce the excitatory potential but they work by division.

(ISI).

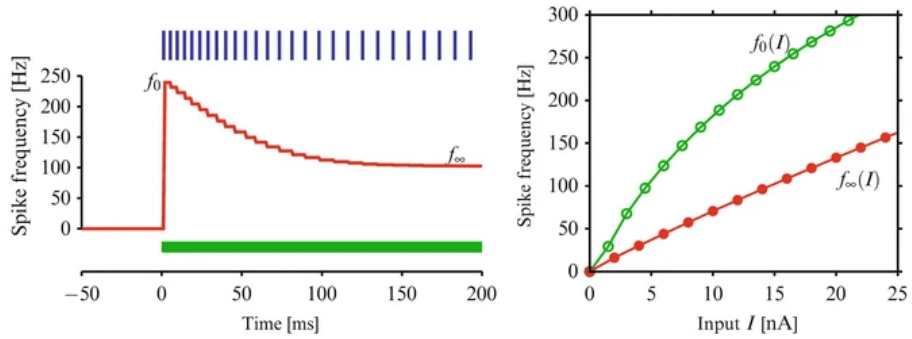


Figure 1.2: (left): Spike train (top) of an adapting neuron evoked by the onset of a constant stimulus (bottom, $I = 15nA$). The spike frequency drops from an initially high onset rate f_0 to a lower steady-state value f_∞ at an approximately exponential rate. (right): Onset frequency-current curve $f_0(I)$ and steady-state frequency-current curve $f_\infty(I)$ for the same neuron. The two plots report simulations of a single-compartment conductance-based model: the Ermentrout model [22]. Reproduced with permission from Springer Nature [23].

Fig. 1.2 (left) shows the increased interspike intervals of an adaptive spike train (top) and the variation of the instantaneous spiking frequency over time (bottom) for a simulated adaptive neuron with a constant current stimulus (shown in green). It is fundamental to highlight that the spiking response of a neuron depends on the intensity of the incoming stimulus. Therefore any plot of the instantaneous spiking frequency over time must be reported with the specific amplitude and time evolution of the applied stimulus.

The two spiking frequency values f_0 and f_∞ are of main importance. f_0 is the *Onset spiking frequency*, which is the instantaneous frequency measured between the first two spiking events, after the onset of the constant stimulus. The onset frequency is the maximum spiking frequency measurable for a specific amplitude of the input current. f_∞ is the *Steady-state spiking frequency*, which is the value the spiking frequency tends towards with time if the input is kept constant. To a first approximation, the decay of the firing rate from f_0 to f_∞ in response to a constant stimulus is exponential and can be described by an adaptation time constant.

Many measurements of exponential decay in the firing frequency can be summarized in a single plot, as the one in Fig. 1.2 (right). This diagram shows the values of the onset and steady-state frequencies of a neuron as a function of the constant input current's amplitude: $f_0(I)$ and $f_\infty(I)$. Frequency-current curves are valuable graphical representations used consistently in neuroscience to report the excitability and response properties of neurons [24, 25].

A series of quantities are of particular importance in this type of measurements.

First of all, the vertical distance between the two curves is a quantitative indicator of the maximum adaptation strength at a given stimulus intensity. Neurons can also have $f(I)$ curves that are shifted on the right compared to the one in Fig. 1.2. In these cases the neuron remains quiescent up to a minimum input current level and starts firing only with a stimulus above this threshold. This threshold value is the threshold current of the neuron. Finally, the firing activity of a neuron may reach a saturation point at high input currents. This occurs because the neuron's ability to generate action potentials is limited by factors such as its refractory period.

1.2.2 Adaptive spiking neuron models

Spike-frequency adaptation has been modeled in many extended and simplified models. In this subsection we report a few examples of these models. Some equations introduced here will be useful in later Chapters for the development of the organic adaptive spiking circuit.

The Adaptive exponential integrate-and-fire model (AdEx) [26] is one of the most used spiking neuron models in (computational) neuroscience. It is a phenomenological model capable of reproducing many neuronal firing patterns (e.g. adapting, bursting, delayed spike initiation, initial bursting, fast spiking, and regular spiking) with a simple system of two coupled ordinary differential equations (ODEs). The model defines the evolution of the membrane potential V for an injected postsynaptic (or external, or both) current I.

The two equations of the AdEx model are:

$$C_{mem}\frac{dV}{dt} = g_L(E_L - V) + g_L\Delta_T e^{(V-V_T)/\Delta_T} - w + I \tag{1.2}$$

$$\tau_w\frac{dw}{dt} = a(V - E_L) - w + b\tau_w \sum_{t=t_{spk}} \delta(t - t_{spk}) \tag{1.3}$$

with the reset condition:

$$\lim_{\delta t \to 0^+} V(t_{spk} + \delta t) = V_r \tag{1.4}$$

and the spiking times t_{spk} that are defined by crossing the spiking threshold voltage V_{cut} (or "cut" voltage because it cuts the increasing exponential):

$$t_{spk} : V(t_{spk}) = V_{cut} \tag{1.5}$$

In order: C_{mem} is the membrane capacitance; g_L is the total leak conductance; E_L is leak reverse potential, or effective rest potential; V_T is the threshold potential; Δ_T is the slope factor; w is the adaptation current; τ_w is the adaptation time constant; a is the adaptation coupling parameter; b is the spike-triggered adaptation parameter.

A complete description of the model is present in the original manuscript by Brette and Gerstner [26]. Also, its phase diagram describing the transition from one firing type to another by parameter tuning has been presented in [27].

The first equation describes the dynamics of the membrane potential as a conservation of currents across the membrane. The integration of the external current by the membrane capacitance occurs in presence of a leakage conductance and of an activation term with an exponential voltage dependence. The leak conductance is equivalent to the one of a Leaky integrate-and-fire (LIF) model. The exponential term is introduced to describe the upswing of the action potential in the spike generation process due to the voltage dependent activation of sodium channels. The upswing is stopped when the membrane potential reaches V_{cut} (e.g. $0mV$ or $+30mV$). V is reset at the value V_r at the next time instant. The adaptation current w, that is subtracted to the external current, is governed by the second equation.

The second equation is coupled to the evolution of V by the conductance parameter a. The adaptation current increases of a value b at every spiking event.

An example of spike-frequency adapting response of an AdEx model is reported in Fig. 1.3. The two plots show the evolution of the membrane voltage V and of the adaptation current w in response to a constant external current stimulus I.

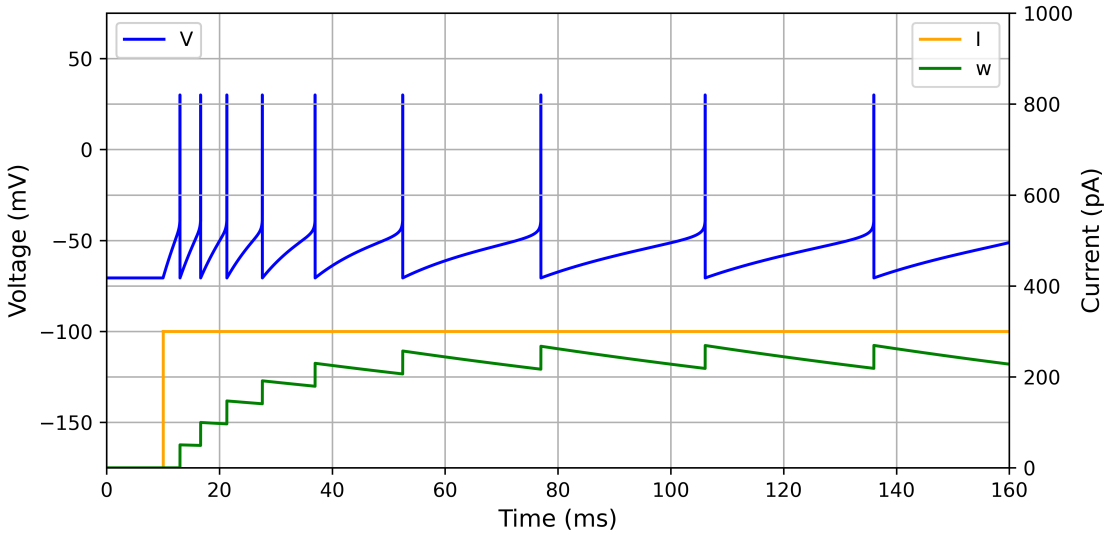


Figure 1.3: Dynamics of the membrane voltage V and of the adaptation current w of a simulated AdEx model showing spike-frequency adaptation in response to a step stimulus I. The neuron's parameters are: $C_{mem} \approx 28.27pF$, $g_L \approx 2.83nS$, $E_L = -70.6mV$, $\Delta_T = 2mV$, $V_T = -50.4mV$, $\tau_w = 144ms$, $a = 0.05nS$, $b = 50.5pA$, $V_r = -70.6mV$, $V_{cut} = -38.4mV$. The spikes are extended graphically up to $30mV$.

The AdEx model is a very versatile and compact model that can describe a variety of spiking patterns. For the case of spike-frequency adaptation, this behavior is reproducible by even simpler set of equations as well.

The Adaptive integrate-and-fire model (or Leaky integrate-and-fire with adaptation model) is a simplified version of the AdEx model that misses the exponential voltage dependent term for spike generation, in the first equation. It can be obtained from the AdEx by taking the limit $\Delta_T \to 0$. The second equation, governing the adaptation current, in unchanged. This simplified model can also reproduce SFA. In Fig. 1.4 we show the simulation of an Adaptive integrate-and-fire model with the same exact parameters as the AdEx reported previously (Fig. 1.3) and the same input stimulus.

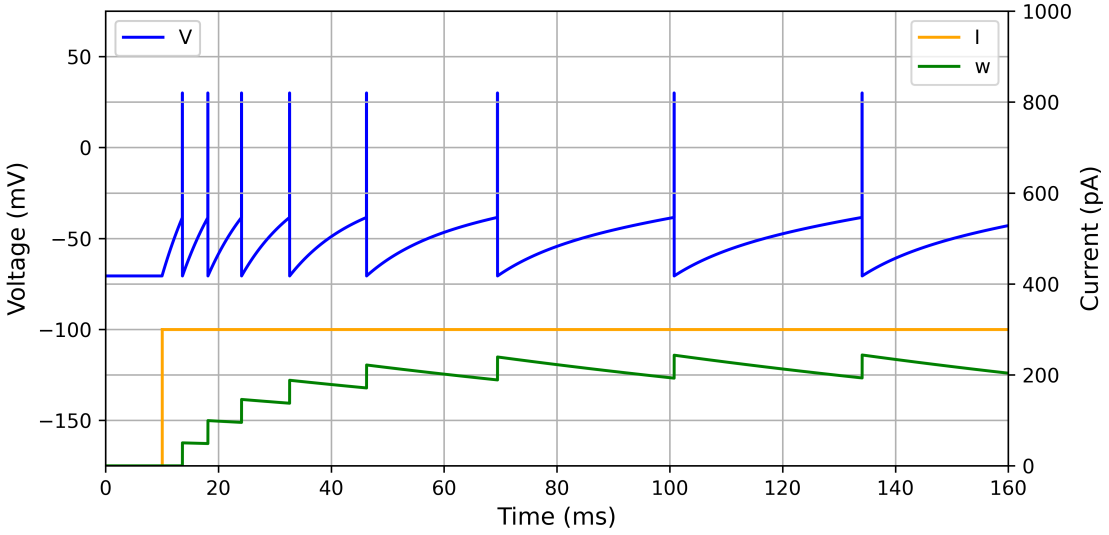


Figure 1.4: Dynamics of the membrane voltage V and of the adaptation current w of a simulated Adaptive integrate-and-fire model showing spike-frequency adaptation in response to a step stimulus I. The neuron's parameters are equivalent to the ones reported in Fig. 1.3 (except for the condition $\Delta_T \to 0$). As in the previous simulation, spikes are still generated when the membrane voltage reaches $V_{cut} = 38.4mV$. However, the equation defining the membrane voltage now misses the voltage-dependent exponential term. This causes a delay in the spiking times. More similar spiking times can be obtained by lowering the value of V_{cut}.

An even simpler (and also historically preceding) model is the one proposed by Treves in [28]. This was the first model to introduce a linear mechanism of adaptation in a spiking neuron model. It is obtained from the Adaptive integrate-and-fire model by setting the adaptation coupling parameter $a = 0$. The conductance a incorporates the fact that some adaptation currents are already activated by subthreshold membrane voltages: with $a = 0$ the adaptation has pure spike-triggered

coupling (only depends on the spiking events and not explicitly on the membrane voltage). Indeed, the model is a Leaky integrate-and-fire with spike-triggered adaptation.

For clarity, we write the equations of the model by Treves, that we obtained starting from the AdEx model and imposing the two restricting conditions expressed previously. The model by Treves writes:

$$C_{mem}\frac{dV}{dt} = g_L(E_L - V) - w + I \tag{1.6}$$

$$\tau_w\frac{dw}{dt} = -w + b\tau_w \sum_{t=t_{spk}} \delta(t - t_{spk}) \tag{1.7}$$

with the reset condition:

$$\lim_{\delta t \to 0^+} V(t_{spk} + \delta t) = V_r \tag{1.8}$$

and the spiking times defined by crossing the spiking threshold V_{cut}:

$$t_{spk} : V(t_{spk}) = V_{cut} \tag{1.9}$$

A simulation of this model is shown in Fig. 1.5, with the same neuron's parameters and input stimulus as the AdEx reported in Fig. 1.3.

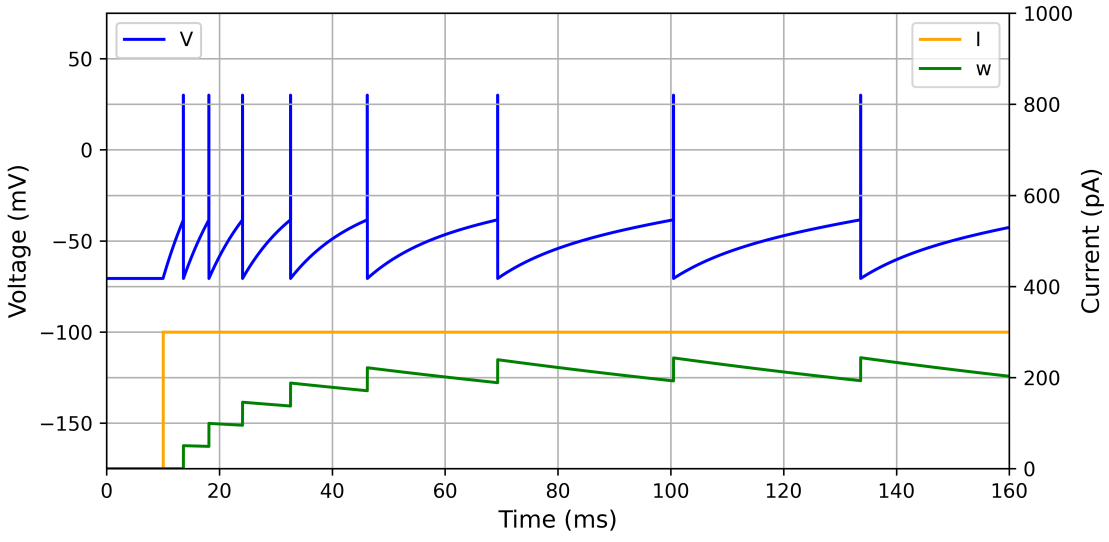


Figure 1.5: Dynamics of the membrane voltage V and of the adaptation current w of a simulated Leaky integrate-and-fire with spike-triggered adaptation model in response to a step stimulus I. The neuron's parameters are equivalent to the ones reported in Fig. 1.3 (except for conditions $\Delta_T \to 0$ and $a = 0$).

Integrate-and-fire neurons with adaptation variables are compact and efficient models that can reproduce a variety of neuronal firing patterns [29]. They are inadequate to describe complex neuronal behaviours (e.g. those related to the cell geometric structure or to dendritic compartments), but they are sufficient to reproduce the basic features of neuronal transduction. For example they can be used to predict the spiking times of cortical neurons under current injection [30, 31]. Emulation of these models on neuromorphic hardware enables large-scale neural network investigation, facilitating the exploration of network dynamics and the study of emergent properties.

1.2.3 Adaptive spiking neuron circuits

In the last thirty years, many examples of adaptive artificial neurons have been proposed in CMOS technology. These include circuits that reproduce with good accuracy the equations of biological spiking models, and circuits that are more loosely bound to mathematical models, but still emulate the characteristic firing patterns and are usually more compact.

Two main techniques are used to implement spike-frequency adaptation in spiking circuits.

The first one is to integrate the voltage spikes generated by the neuron itself and subtract the resulting slow dynamics current to the membrane potential. This can be achieved with different low-pass filtering (LPF) strategies and architectures that are well documented in [32, 33]. These adaptive blocks usually present tuning nodes (usually the gate of specific transistors in the circuit) that introduce degrees of freedom in the parameters of the adaptation current, e.g. tuning of the adaptation time constant τ_w or of the spike-triggered adaptation parameter b. For example, the "Tau-Cell" circuit, first reported in [34] as a BiCMOS[6] log-domain filter, was used to develop hardware representations of well-know spiking models, such as the Izhikevich model [35, 36].

A second mechanism to implent SFA is to introduce a dynamic spiking threshold. In this case the spiking threshold value is updated depending on the spiking activity of the neuron: it increases if spiking events occur too frequently, thereby increasing the next interspike intervals for a constant input, and viceversa. An example is given by the the Generalized linear integrate-and-fire model proposed by Mihalas and Niebur [37] and its hardware implementation [38].

This last strategy is frequently used also in artificial neurons developed with emerging resistive memory technologies. In these circuits, the output spikes of the neuron are used as SET or RESET pulses applied across a volatile or non-volatile memristive

[6]Bipolar Complementary Metal-Oxide-Semiconductor (BiCMOS) is a mixed technology that integrates CMOS and BJT on the same chip.

device, causing a resistance change, which is then converted into a spiking threshold variation by the circuit architecture. Artificial neurons with adaptive thresholds have been developed with different emerging technologies, including phase-change materials [39] and valence change memories [40, 41].

Memristive devices are also extremely versatile to build spiking neurons with architectures that closely resemble the circuit diagrams of conductance-based spiking models, such as the one for the Hodgkin-Huxley model reported in Fig. 1.1. For example, a variety of neuronal dynamics were emulated in [42] using similar circuit topologies and nanoscale vanadium dioxide (VO_2) active memristors, including spike-frequency adaptation. The main drawback to these type of circuits is that they integrate a significant number of passive elements, which may condition their scalability into large networks and also set many neuron properties at the time of fabrication.

In recent years, spiking neurons built with organic materials have also been developed. Organic electronics is a less mature technology than the ones previously mentioned, but it is promising for extended functionalities, such as biocompatibility [43] and for flexible electronics [44]: these two properties combined are crucial to develop a technology that can interface with biological tissue.

Organic electrochemical transitors (OECTs), in particular, are an emerging technology that intrinsically operates similarly to nerve membranes, as they use an electrolytic solution with ionic currents to charge and discharge an organic (polymeric) capacitance.

Both conductance-based and phenomenological circuits have been proposed with OECT technology. These artificial neurons introduce many interesting features for bio-realistic and possibly interfaceable neuromorphic systems. In many cases, they show modulation of the neuron's activity based on the concentration of ion species, neurotransmitters or aminoacids present in the electrolytic solution [45–48]. This modulation can occur as a modification of the frequency-current curve or as a variation in the response time after the application of a stimulus. The circuits can be tuned at the device level to be sensitive to common physiological ionic and biomolecular concentration variations of the surrounding environment. They can also present signal transduction from light stimulus to spiking frequency, additionally to ionic/biomolecular sensing [48]. Moreover, device engineering and organic chemistry can be exploited to fabricate devices that emulate individually the temporal dynamics of ionic channels in biological nerve membranes, for the development of compact Hodgkin-Huxley based circuits [47].

As we discussed in Sec. 1.1, the main goal of the original neuromorphic engineering approach was to emulate neural computation by exploiting a medium whose physical behavior is analogous to that of biological nervous systems. At its deepest

level, OECT-based neuromorphic engineering rests not on phenomenological analogy between media, but on the use of the media itself. The price to pay, within the current technological context, is limited scalability and stability compared to many solid-state technologies, even if progress is being made [44]. The development of compact circuits with an increasing number of functionalities is therefore fundamental at present for the development of the field, and will also be fruitful in the long run. As we hope to have justified with this introductory Chapter, the development of adaptive spiking circuits is an important step for emerging technologies towards the realization of more bio-realistic neuromorphic hardware. To the best of our knowledge, an organic artificial neuron that emulates spike-frequency adaptation has not been developed yet. In this thesis, we present a compact spiking architecture that exploits the synaptic plasticity property of OECTs and device engineering to implement tunable spike-frequency adaptation in an all-organic artificial neuron.

Chapter 2

OECTs

Organic electrochemical transistors (OECTs) are thin-film three-terminal transistors that exploit reversible ionic doping/de-doping of the semiconducting channel to actively modulate its conductivity. Since their first development by Wrighton et al. in the mid-1980s [49], they gained increasing interest due to their broad tunability, low-voltage operation, low-cost fabrication and biocompatibility.

More recently, they emerged as a promising technology for the design of artificial synapses and neurons. Their ability to emulate a variety of neurological mechanisms, such as short-term and long-term plasticity [50, 51], and their operation relying on ionic transport, make them intrinsically similar to neurobiological membranes.

In this Chapter, we discuss the architecture, the physics and the mechanisms of operation of OECTs. We also address some available models to describe their steady-state and transient response, depending on the applied voltages and their structural and electrochemical parameters. These models will be used to simulate the behavior of larger organic electrochemical circuits in the following Chapters.

All equations in this Chapter refer to hole transporting polymers. Indeed, p-type OECTs are the ones most frequently found in literature.

2.1 Physics and operation

An OECT consists of an organic semiconducting film connected to a source and a drain terminal and coupled to a gate electrode via an electrolytic solution (Fig. 2.1).

The semiconducting polymer has an initial conductivity $\sigma = q\mu p_0$, where q is the elementary charge, μ is the hole mobility; and p_0 is the initial hole density. Effectively, the conductivity of the OECT's channel is defined once the semiconducting polymer is immersed in the electrolyte and is permeated by water and ions (both anions and cations). By applying a voltage to the gate electrode, a voltage difference forms between the electrolyte and the semiconducting channel, and further ions can be injected into (subtracted from) the polymeric film. The sign and amplitude of this voltage difference determine the verse of ionic transport and the number of trans-

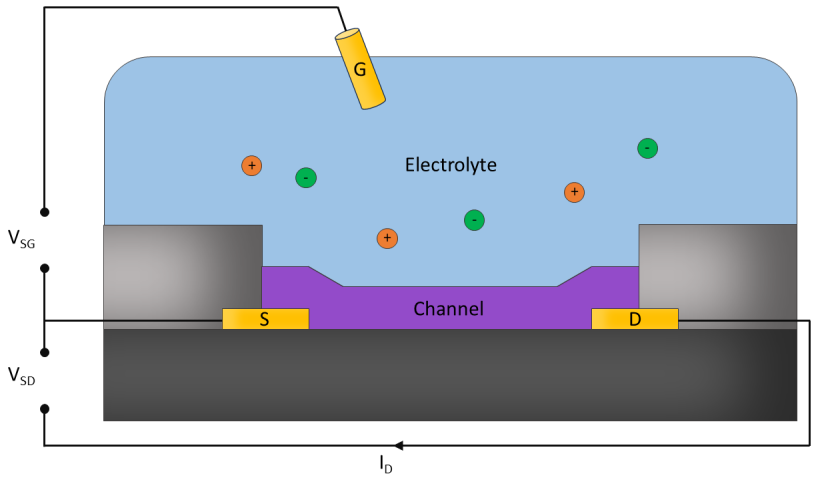


Figure 2.1: Architecture of an organic electrochemical transistor (OECT), showing source (S), drain (D) and gate (G) terminals. Current and voltages are consistent with the convention for a p-type transistor (more common type of OECT).

ported ions, therefore the conductivity variation. Two mechanisms are possible and they induce opposite modulation. In depletion-mode p-type OECTs, an increasing voltage applied to the gate causes injection of cations into the channel, which replace holes as positive charges to compensate for the anions present in the polymer: depletion-mode OECTs turn off with an increasing gate voltage, analogously to p-type MOSFETs. In accumulation-mode n-type OECTs, the cations injected by the increasing gate voltage induce accumulation of electrons inside the polymer: accumulation-mode OECTs turn on with an increasing gate voltage, analogously to n-type MOSFETs. Finally, the applied source-drain voltage induces a drain current proportional to the channel's conductivity.

OECTs rely on a completely different doping mechanism than field-effect transistors (FETs). A gate voltage in a MOSFET causes a 2D inversion layer to form at the interface with the oxide. The same voltage in an OECT causes a 3D bulk charge accumulation inside the channel (Fig. 2.2). This difference entails that the thickness d of the deposited polymer is an active parameter that influences the value of the gate-channel capacitance. This intrinsic property is responsible for the exceptionally large transconductance values of OECTs (on the order of millisiemens for micro-scale devices [52]) and the possibility to tune it without modifying the footprint ($W\backslash L$). OECTs also have only three terminals and miss a bulk terminal. However, they can still be described effectively by models analogous to the ones used for MOSFETs.

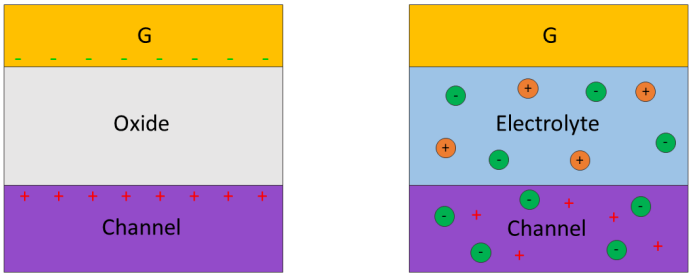


Figure 2.2: Different operating physics between MOSFETs and OECTs (p-type). A 2D inversion layer is formed at the insulator-semiconductor interface in a MOSFET due to the gate voltage. 3D bulk doping occurs inside the semiconducting channel of an OECT when a gate voltage is applied.

2.2 Steady-state response and scaling

The Bernards model, first reported by Daniel Bernards and George Malliaras [53], describes the electronic charge in an OECT's channel with the same set of equations used for long-channel MOSFETs, with a single modeling change: the oxide capacitance C_{ox} of MOSFETs is replaced by the product dC^* (thickness times volumetric capacitance), which considers for the modulation of conductivity in the entire volume of the semiconducting channel.

The Bernards model for a p-type OECT reads:

$$I_D = \begin{cases} 0 & V_{SG} \leq V_T \\ \mu C^* \frac{Wd}{L}[(V_{SG} - V_T)V_{SD} - \frac{1}{2}V_{SD}^2] & V_{SG} > V_T, V_{SD} \leq V_{SG} - V_T \\ \frac{1}{2}\mu C^* \frac{Wd}{L}(V_{SG} - V_T)^2 & V_{SG} > V_T, V_{SD} > V_{SG} - V_T \end{cases} \tag{2.1}$$

where I_D is the drain current; μ is the hole mobility; C^* is the volumetric capacitance of the channel's polymer; W, d and L are the channel width, thickness and length, respectively; V_{SG} and V_{SD} are the applied source-gate and source-drain voltages; and V_T is the threshold voltage.

The set of equations for the transconductance g_m can be derived from Eq. 2.1 and writes:

$$g_m = \begin{cases} \mu C^* \frac{Wd}{L} V_{SD} & V_{SG} > V_T, V_{SD} \leq V_{SG} - V_T \\ \mu C^* \frac{Wd}{L}(V_{SG} - V_T) & V_{SG} > V_T, V_{SD} > V_{SG} - V_T \end{cases} \tag{2.2}$$

The Bernards model does not report the OECT's behavior in weak inversion (here we reported a null current in the subthreshold region). However, it is important to clarify that OECTs can be operated in weak inversion and can be optimized to reach subthreshold swings close to the thermodynamic limit [54]. The OECT's model in

weak inversion is analogous to the equations written before and writes:

$$I_D = \Big\{ 2n\mu C^* \frac{Wd}{L} U_T^2 e^{\frac{V_{SG}-V_T}{nU_T}} (1 - e^{\frac{-V_{SD}}{U_T}}) \quad V_{SG} < V_T \tag{2.3}$$

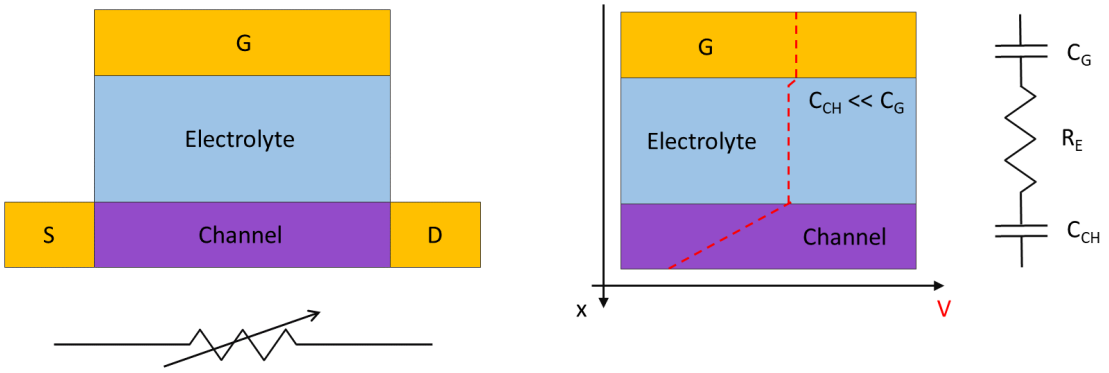
where n is the slope factor; and $U_T = \frac{K_b T}{q}$ is the thermal voltage.

The Bernards model also considers ionic transport inside the electrolyte, which correlates to the OECT's transient response and switching speed. The complete model combines an equivalent electronic circuit and an equivalent ionic circuit (Fig. 2.3): the former relies on the quasi-static approximation[1] for the charge distribution in the OECT's channel, defines the steady-state behavior of the device and is described by Eq.s 2.1 and 2.3; the latter models the drift-diffusion of ions in the gate-electrolyte-channel system and introduces temporal dependence to the model. This second component consists of two capacitors C_G and C_{CH}, that model the gate-electrolyte and channel-electrolyte interfaces, respectively, and a series resistor R_E, that corresponds to the ionic resistance of the electrolyte. For ease, the B-M model assumes that the permeation of negative ions has no effect on the organic semiconductor and that all charge densities are uniform across the channel's thickness (the model is limited to thin-film transistors). Moreover, it assumes that no electrochemical reactions occur at the gate electrode.

Fig. 2.3b also shows the potential distribution across the OECT. In a well-designed OECT, the gate capacitance C_G is much larger than the channel capacitance C_{CH}, so most of the applied gate voltage drops at the channel-electrolyte interface. This is equivalent in a well-designed MOSFET, where the poly-silicon gate is heavily doped in order to limit depletion inside of it. To achieve this result, two solutions are possible in OECTs. The first one is to deposit the same polymer used for the channel on a gate contact with a much larger footprint than the channel itself. This implementation exploits the volumetric ion intake of organic polymers: by increasing the footprint of the gate, more ions can be injected into the gate's polymer with respect to the channel, and the gate capacitance will be larger than the channel's one. Equivalently, one could increase the thickness of the gate's polymer. However, keeping a constant thickness for all polymers allows to deposit the channel and the gate in a single fabrication step. The second possibility is to use non-polarizable gate electrodes, such as Ag/AgCl pellets. Generally, a 10:1 ratio between C_G and C_{CH} is required for efficient gating.

The Bernards model predicts that the OECT's response is linked to the ionic RC time constant $\tau_i = R_E C_{CH}$. The two quantities R_E and C_{CH} depend on the dimensions of the channel, other than the type of semiconducting polymer and electrolytic medium. The channel capacitance is uniquely defined as $C_{CH} = C^* WdL$. On the other hand, the electrolyte resistance shows different proportionalities depending on

[1] The quasi-static approximation assumes that the drain current only depends on the instantaneous values of V_{SG} and V_{SD} and neglects the effect of their time-dependent variations.



(a) Electronic circuit.

(b) Ionic circuit.

Figure 2.3: The complete Bernards model. Electronic conduction inside the OECT's channel is modeled as a tunable resistor. The ionic system, made of gate electrode, electrolyte and polymeric channel, is modeled as a series of two capacitors (interfaces) and a resistor (ionic conduction through the electrolyte).

the OECT's architecture. Researchers demonstrated via electrochemical impedance spectroscopy that the electrolyte resistance for gold electrodes coated with PEDOT:PSS[2] surrounded by a much broader electrolyte volume scales as $R_E = r^* \frac{1}{\sqrt{WL}}$, where r^* has the unit of resistivity [55, 56]. In this case, the two scaling laws combined imply that the ionic time constant scales with the dimensions of the polymeric channel as $\tau_i \propto d\sqrt{WL}$. In the different case of a planar OECT fabricated with the electrolyte confined in a microfluidic path between gate and channel, with defined length h and cross-sectional area A, the resistance is given by $R_E = r^* \frac{h}{A}$. For this second architecture, the ionic time constant scales as $\tau_i \propto \frac{WdLh}{A}$, where the specific definition of A depends on the relative dimensions of the electrolyte path and the channel. In both cases $r^* = \frac{1}{\sigma}$, where σ is the electrolyte conductivity (or specific conductance). For sufficiently low ionic concentrations, $\sigma = q(c^+ z^+ u^+ + c^- z^- u^-)$, where c, z and u are the concentrations, ion charge and mobility, respectively, of the dissociated positive or negative ions.

In reality, the contribution of the gate capacitance C_G should also be considered when calculating the ionic time constant. The response speed is limited by the equivalent capacitance given by the two capacitors C_G and C_{CH} in series: $C_{eq} \equiv \frac{1}{1 \backslash C_G + 1 \backslash C_{CH}}$. However, this quantity can be approximated to $C_{eq} \approx C_{CH}$ if the condition defined earlier for efficient gating is valid.

As a note, OECTs can also be fabricated as all solid-state devices, by substituting the

[2]Poly(3,4-ethylenedioxythiophene)-poly(styrene sulfonate), abbreviated PEDOT:PSS, is the most common polymer used in OECTs. In PEDOT:PSS, the holes in the conductive backbones of PEDOT are compensated by the sulfonate ions (SO_3^-) of PSS, which acts as a stabilizing agent.

liquid electrolyte with an ion gel[3]. The tradeoff of not needing device encapsulation is that this type of ionic conductors have higher resistivities and slow down the operation of the OECT.

Bernards and Malliaras further develop their model and manage to obtain an expression for the time-dependent currents. As we will see in the next section, source and drain currents do not match in the transient response of an OECT. For a more intuitive explanation, we will introduce the OECT's transient response with a circuit schematic first proposed by Friedlein et al. [57] some years after the calculations of Bernards and Malliaras.

2.3 Transient response

The circuit diagram proposed by Friedlein et al. is depicted in Fig. 2.4. The schematic is equivalent to the electronic and ionic circuits in Fig. 2.3. However, the channel capacitance C_{CH} is now split into two components C_{CH_D} and C_{CH_S} that couple the gate to the drain and to the source, respectively. This expedient allows to draw a single equivalent circuit for the entire OECT.

We suppose to operate the OECT with a fixed source-drain voltage and a source-gate voltage that varies with time. The source and drain currents are now composed of two contributions: the channel current and the gate current.

The channel current i_{CH} is equivalent to the steady-state case. It represents the instantaneous value of the electronic conductance of the channel and it is described by Eq. 2.1 (and Eq. 2.3), except that V_{SG} must be substituted with v_{SE}:

$$i_{CH} = \begin{cases} 0 & v_{SE} \leq V_T \\ \mu C^* \frac{Wd}{L}[(v_{SE} - V_T)V_{SD} - \frac{1}{2}V_{SD}^2] & v_{SE} > V_T, V_{SD} \leq v_{SE} - V_T \\ \frac{1}{2}\mu C^* \frac{Wd}{L}(v_{SE} - V_T)^2 & v_{SE} > V_T, V_{SD} > v_{SE} - V_T \end{cases} \tag{2.4}$$

This new voltage v_{SE} represents the source-gate voltage after the propagation through the electrolyte and is the effective voltage difference that is sensed by the channel.

The gate current i_G is a transient current and is given by the sum of two contributions. The first one is the electronic leakage through the electrolyte i_{leak}, which is usually negligible and that will be neglected from now on. The second one is the ionic displacement current i_{ION}. The ionic displacement current represents the transient of majority carriers that leave/repopulate the channel due to the ionic doping/de-doping, caused by the time-varying gate voltage.

[3] An ion gel is a composite material consisting of an ionic liquid immobilized by an inorganic or a polymer matrix.

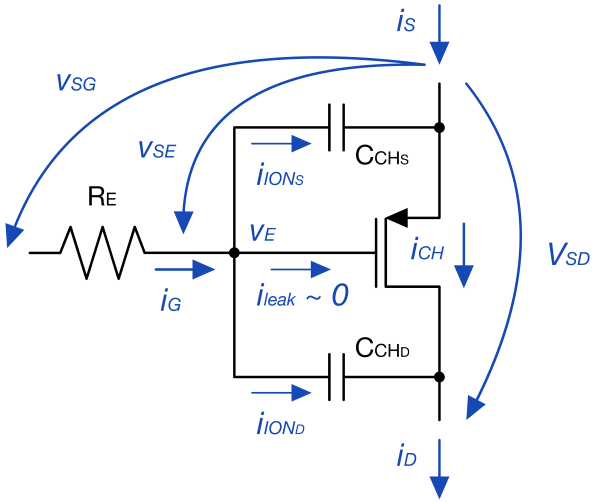


Figure 2.4: Circuit schematic of the Friedlein model. Compared to the Bernards model, the channel capacitance is split into two contributions C_{CH_D} and C_{CH_S}. The model assumes $C_G >> C_{CH}$, so the gate capacitance is neglected.

As an example, we take as a reference the p-type OECT in Fig. 2.1: the OECT has some initial applied voltages $V_{SG} > 0$ and $V_{SD} > 0$ with a defined value of channel conductance. If the gate voltage is increased, the gate electrode repels cations present in the electrolyte. The cations drift towards the polymeric channel and dope it. This motion occurs with the ionic time constant $\tau_i = R_E C_{CH}$ defined in Sec. 2.2. Accordingly, holes that were previously present in the channel, will be repelled by the newly absorbed cations and leave the polymer. At this point, the majority carrier concentration inside the polymer decreases, therefore its conductance, and this variation is considered by the channel current. Also, holes that leave the channel flow either towards the source, or the drain, with a certain ratio. This second effect is provided for by the ionic displacement current. Intuitively, holes will tend to move towards the lower voltage terminal, in this case the drain, and this is modeled by having $C_{CH_D} > C_{CH_S}$.

Finally, the drain current will be equal to the sum of the channel current and the portion of ionic displacement current which flows towards the drain; by contrast, the source current is given by the difference of the channel current and the portion of ionic displacement current which flows towards the source.

This phenomenology is described by the following set of equations, which can be directly derived from the circuit diagram in Fig. 2.4.

The total source and drain currents are:

$$i_D = i_{CH} + i_{ION_D} \tag{2.5}$$

$$i_S = i_{CH} - i_{ION_S} \tag{2.6}$$

where i_{ION_D} and i_{ION_S} are the portions of ionic displacement current that flow towards the drain and the source, respectively. These two contributions are given by:

$$i_{ION_D} = C_{CH_D} \frac{d}{dt}(v_E - V_D) \tag{2.7}$$

$$i_{ION_S} = C_{CH_S} \frac{d}{dt}(v_E - V_S) \tag{2.8}$$

where $C_{CH} = C_{CH_D} + C_{CH_S}$. Finally, we can write the ohm law through the electrolyte:

$$v_G - v_E = R_E i_G \tag{2.9}$$

where $i_G \approx i_{ION} = i_{ION_D} + i_{ION_S}$ since the electronic leakage in the electrolyte is negligible.

Eqs. 2.7 and 2.8 can be substituted into Eq. 2.9 to obtain the differential equation for the voltage v_E. Since V_D and V_S are constant in time, we obtain:

$$\frac{d}{dt} v_E = \frac{1}{R_E C_{CH}} (v_G - v_E) \tag{2.10}$$

R_E, C_{CH_D} and C_{CH_S} form a RC integrator ($\tau_i = R_E C_{CH}$) of the input gate voltage with transfer function

$$H(s) = \frac{1}{1 + \tau_i s} \tag{2.11}$$

This property of OECTs enables them to have plasticity behavior.

We can start by analyzing the OECT's response to a single pulse applied to the gate and then extend the discussion to a train of input pulses (these could represent incoming action potentials from an artificial neuron).

First, the RC integrator's *impulse response*, which is its response to a unit impulse input signal $v_G(t) = \delta(t)$, is:

$$v_E(t) = \frac{1}{\tau_i} e^{-\frac{t}{\tau_i}} \cdot u(t) \tag{2.12}$$

where the step function $u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & otherwise \end{cases}$; this is obtained by calculating the inverse Laplace transform of the transfer function ($v_G(s) = 1$). To know the impulse response of the system is an extremely important result. In fact, the response of a linear time-invariant system to any arbitrary signal $x(t)$ can be computed by performing the convolution of its impulse response with the signal itself [32].

In the practical case of a pulse input, we can derive the circuit's solution analytically by considering the pulse as the difference of two step functions. For a pulse $v_G(t)$, with amplitude v_{G_0}, that starts at $t = t_0$ and ends at $t = t_1$ ($t_1 \geq t_0$), we can write:

$$v_G(t) = v_{G_0} \cdot (u(t - t_0) - u(t - t_1)) = \begin{cases} v_{G_0} & t_0 \leq t < t_1 \\ 0 & otherwise \end{cases} \tag{2.13}$$

The integrator's *step response* (its response to a unit step input signal $v_G(t) = u(t)$) can be calculated analogously to the impulse response and yields:

$$v_E(t) = (1 - e^{-\frac{t}{\tau_i}}) \cdot u(t) \tag{2.14}$$

Therefore, the total solution for the voltage $v_E(t)$ is:

$$v_E(t) = \begin{cases} 0 & t < t_0 \\ v_{G_0}(1 - e^{-\frac{t-t_0}{\tau_i}}) & t_0 \leq t < t_1 \\ v_{G_0}(1 - e^{-\frac{t_1-t_0}{\tau_i}})e^{-\frac{t-t_1}{\tau_i}} & t \geq t_1 \end{cases} \tag{2.15}$$

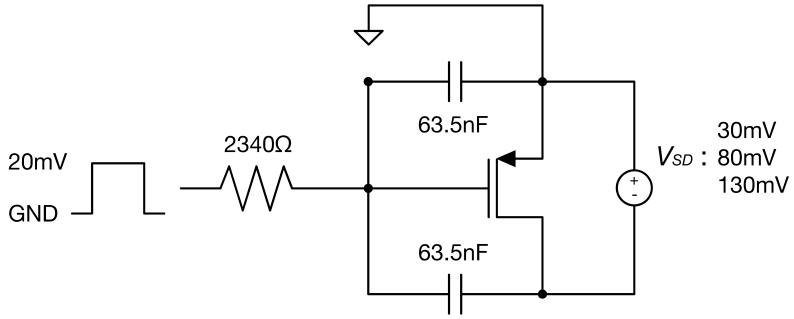
where we imposed continuity.

Fig. 2.5b shows the evolution of the voltage $v_E(t)$ on a simulated OECT (Fig. 2.5a), when a single upward pulse or a train of pulses are applied to the gate. The two cases also differ in the duration of the applied pulses (to give a complete overview on the possible behaviors of an OECT). The simulated circuit is matched to the p-type PEDOT:PSS[4] OECT reported in [57].

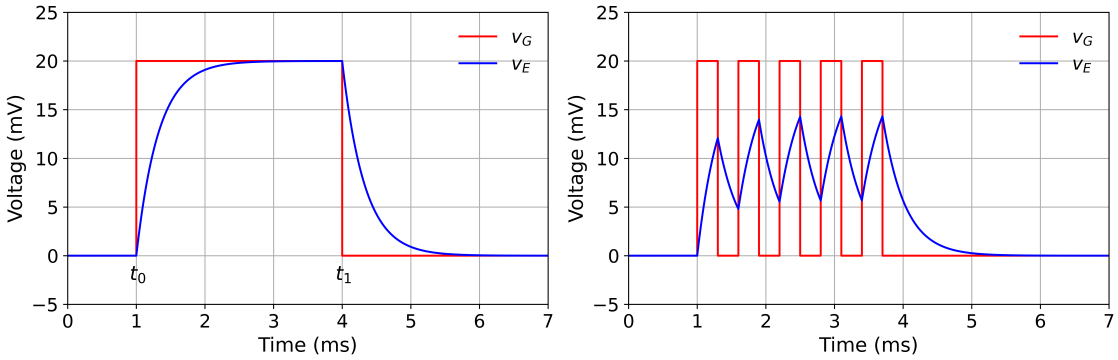
In the first case of Fig. 2.5b, the single pulse has a duration that is much longer than the ionic constant of the OECT: $t_1 - t_0 > 10\tau_i$. This allows the voltage $v_E(t)$ to reach steady-state (20mV) before decaying exponentially when the pulse ends. In this regime the OECT does not show plasticity but just a delayed response.

In the second case, the OECT is operated with very short gate pulses, that are also close in time: $t_1^i - t_0^i \approx \tau_i$, $t_0^{i+1} - t_1^i \approx \tau_i$ (index i specifies the pulse number in the pulse train). In this second condition, the RC rise of the voltage $v_E(t)$ is interrupted before it can reach steady-state (20mV), due to the short-duration of the input pulse. Also, the RC decay is stopped by a new incoming pulse, before $v_E(t)$ can reach 0mV. This entails that the new RC rise of $v_E(t)$ now starts with initial condition $v_E(t_0^{i+1}) \neq 0$. The result is that the maximum voltage v_E^{MAX} touched by the OECT keeps increasing with each pulse and saturates for a number of pulses $i \to \infty$, if the properties of the train of pulses do not change. v_E^{MAX} depends on v_{G_0}, τ_i, $t_1^i - t_0^i$ and $t_0^{i+1} - t_1^i$, as can be deduced from Eq. 2.15.

[4]The polymer PEDOT:PSS is conductive in its natural state. PEDOT:PSS OECTs can be operated effectively with an applied voltage difference $v_{SG} < 0$, as in the simulation of Fig. 2.5a.



(a) Simulated Friedlein schematic.



(b) Effective voltage v_E sensed by the OECT's channel for different pulsed inputs (v_G).

Figure 2.5: SPICE simulation of a p-type PEDOT:PSS OECT reported in [57]. The voltage v_E sensed by the channel is the output of a RC integrator with input the gate voltage v_G. The OECTs shows plasticity if the input pulses have duration and delay that are comparable to the OECT's ionic time constant τ_i or shorter. For the simulated OECT $\tau_i \approx 0.3ms$. Rise and fall times of the gate voltage are $100ns << \tau_i$, so the signals are rectangular pulses with good approximation.

We can write the generalized version of Eq. 2.15 for any pulse i in the pulse train, depending on the previous pulse $(i-1)$ and the time gap between the two $t_0^i - t_1^{i-1}$:

$$v_E(t) = \begin{cases} v_E(t_1^{i-1})e^{-\frac{t-t_1^{i-1}}{\tau_i}} & t_1^{i-1} \leq t \leq t_0^i \\ v_E(t_0^i) + (v_{G_0} - v_E(t_0^i))(1 - e^{-\frac{t-t_0^i}{\tau_i}}) & t_0^i \leq t \leq t_1^i \\ v_E(t_1^i)e^{-\frac{t-t_1^i}{\tau_i}} & t_1^i \leq t \leq t_0^{i+1} \end{cases} \tag{2.16}$$

where the values of $v_E(t_0^i)$ and $v_E(t_1^i)$ can be found by imposing continuity.

These are the fundamental formulas that describes OECT's short-term plasticity (STP) property, that is much used in literature [58–60].

At this point, we shall calculate the complete source and drain currents in the transient regime. To do that, we have to define how the channel capacitance C_{CH} splits between C_{CH_D} and C_{CH_D}, or equivalently how the displacement current splits between drain and source.

To account for this effect, Bernards and Malliaras first introduced a proportionality factor f that defines the fraction of displacement current that contributes to the drain current (symmetrically, $1 - f$ is the fraction that contributes to the source current). This factor considers for the spatial non-uniformity of the doping/de-doping process along the channel's length. Indeed, more cations will tend to dope the channel close to the lower voltage terminal, the drain, than close to the source. Therefore, the f-factor assumes values in the range $[\frac{1}{2}; 1]$: for instance, it reaches values close to $\frac{1}{2}$ when $V_G >> V_S > V_D$, while it approaches 1 when $V_S >> V_G > V_D$.

The treatment of this weighting factor distinguishes several models that are variations of the Bernards model [57, 61, 62].

As a first approximation, the Friedlein model [57] does not consider for non-homogeneous doping/de-doping of the channel and sets $C_{CH_D} = C_{CH_S} = C_{CH}/2$. Therefore, the ionic displacement current is equally distributed between source and drain: $i_{ION_D} = i_{ION_S} = i_{ION}/2$.

Fig. 2.6 (left) shows the channel and ionic current contributions for the OECT modelled with the Friedlein approximation of Fig. 2.5, for different values of V_{SD}.

The splitted ionic currents i_{ION_D} and i_{ION_S} are capacitive currents. They only depend on the time derivative of $v_G(t)$, since V_D and V_S are fixed in time (Eq.s 2.7 and 2.8). Therefore, they do not change between the three cases. The channel current i_{CH} depends on V_{SD} instead, as well as on $v_G(t)$ (Eq.s 2.4).

The sum/difference of these contributions gives the drain/source currents of the OECT, according to Eq.s 2.5 and 2.6. These are shown in Fig. 2.6 (right).

In particular, i_D is given by the sum of these two. In a p-type OECT, the variations of i_{CH} and i_{ION_D} due to a stepped or pulsed gate voltage are always opposite in sign. Therefore the drain current always shows a "spike and recovery" behavior, where the spike is the capacitive spike of the drain ionic displacement current i_{ION_D}. Note that this is equivalent for the drain current in a n-type OECT, where i_D is given by the difference of i_{CH} and i_{ION_D}, but these two now have same sign variations.

The source current shows different behaviors, instead. It evolves either in a "spike and recovery" way, or a "fast switching" way, or a "monotonic decay" way, depending on the ratio between the maximum variations of i_{CH} and i_{ION_S}.

We can better understand these three regimes by analyzing analytically these two

current variations. The former we define it as:

$$\Delta I_{CH} = I_{CH}(v_{SE}(t_0), V_{SD}) - I_{CH}(v_{SE}(t_1), V_{SD}) \tag{2.17}$$

which is the difference between the initial and final channel currents (right at the start and at the end of the applied gate voltage step); the latter is given by $i_{ION_S}(t_{0^+})$. From Eq. 2.8 with $v_E(t)$ given by Eq. 2.15, this can be calculated to be:

$$i_{ION_S}(t_{0^+}) = \frac{C_{CH_S} v_{G_0}}{\tau_i} = \frac{(1-f)C^* W d L v_{G_0}}{\tau_i} \tag{2.18}$$

If we consider an OECT operated in triode and that reaches steady-state, as the one in Fig. 2.6, we can write the variation of i_{CH} due to the first step of $v_G(t)$ as:

$$\Delta I_{CH} = \mu C^* \frac{Wd}{L} v_{G_0} V_{SD} \tag{2.19}$$

We can finally calculate the ratio between the two to be:

$$\frac{i_{ION_S}(t_{0^+})}{\Delta I_{CH}} = \frac{(1-f)L^2}{\mu V_{SD} \tau_i} \tag{2.20}$$

Eq. 2.20 clarifies the behaviors in Fig. 2.6. First of all, this ratio decreases as $\propto \frac{1}{V_{SD}}$. This entails that an increasing V_{SD} contributes to make the OECT's transition from the "spike and recovery" behavior to the "monotonic decay". The intermediate case of "fast switching" is obtained for a ratio close to 1, therefore for $V_{SD} \approx \frac{(1-f)L^2}{\mu \tau_i}$.

Also:

$$\frac{L^2}{\mu V_{SD}} = \tau_e \tag{2.21}$$

which is the electronic transit time across the channel.

The three behaviors are related to different ratios between the time scales of ionic and electronic transport. In the "spike and recovery" case, the OECT's switching time is limited by hole conduction through the channel and $(1-f)\tau_e > \tau_i$. In the "monotonic decay", it is limited by ionic conduction through the electrolyte, $(1-f)\tau_e < \tau_i$ and i_S evolves with an abrupt variation followed by an exponential relaxation. In the intermediate "fast switching" case, the two transport mechanisms have comparable time constants, $(1-f)\tau_e \approx \tau_i$ and the OECT's response is close to a square current step.

The model we presented up to now allows to engineer OECTs to behave in either of these regimes, for some specified operating conditions (V_{SG}, V_{SD}). This could be useful for some applications, such as ionic- or bio-sensing. For example, the capacitive spike signal could be maximized over the channel current by a series of design

choices, such as setting the OECT to the correct operating point, increasing the channel dimensions with a fixed W/L ratio or shortening the gap between channel and gate. Small variations of electrolyte resistance due to changes in the ionic concentration could be sensed by applying a know AC signal to the gate and measuring the drain current (or source current). From Sec. 2.2 we know that $R \propto \frac{1}{c_I}$. In this section we also calculated the maximum change in the source ionic displacement current (Eq. 2.18). Similarly, the maximum change in the drain ionic displacement current can be written as:

$$I_{ION_D}(t_{0^+}) = \frac{C_{CH_D} v_{G_0}}{\tau_i} = \frac{fC^* WdLv_{G_0}}{\tau_i} \tag{2.22}$$

Combining these two formulas we obtain:

$$I_{ION_D}(t_{0^+}) \propto v_{G_0} c_I f \tag{2.23}$$

This proportionality is independent from the channel capacitance. However, to increase the channel capacitance entails an increase in the ionic time constant. If the ratio W/L is unchanged, the channel current variation ΔI_{CH} of Eq. 2.17 is minimized because the cutoff frequency of the RC integrator (low-pass filter) is reduced. Also, $I_{ION_D}(t_{0^+})$ can be maximized by having a f-factor close to 1 and by increasing the amplitude v_{G_0} of the input gate voltage. These two conditions correspond to setting proper operating conditions of the OECT sensor.

In mixed-signal ICs however, having drain and source currents that do not match can be problematic in most situations. Therefore, the ionic displacement spikes should be minimized. Opposite considerations to the ones made for the OECT sensor hold in this case. Morevoer, the electrolyte resistance can be increased to reduce the amplitude of these spikes, at the expense of a slower OECT.

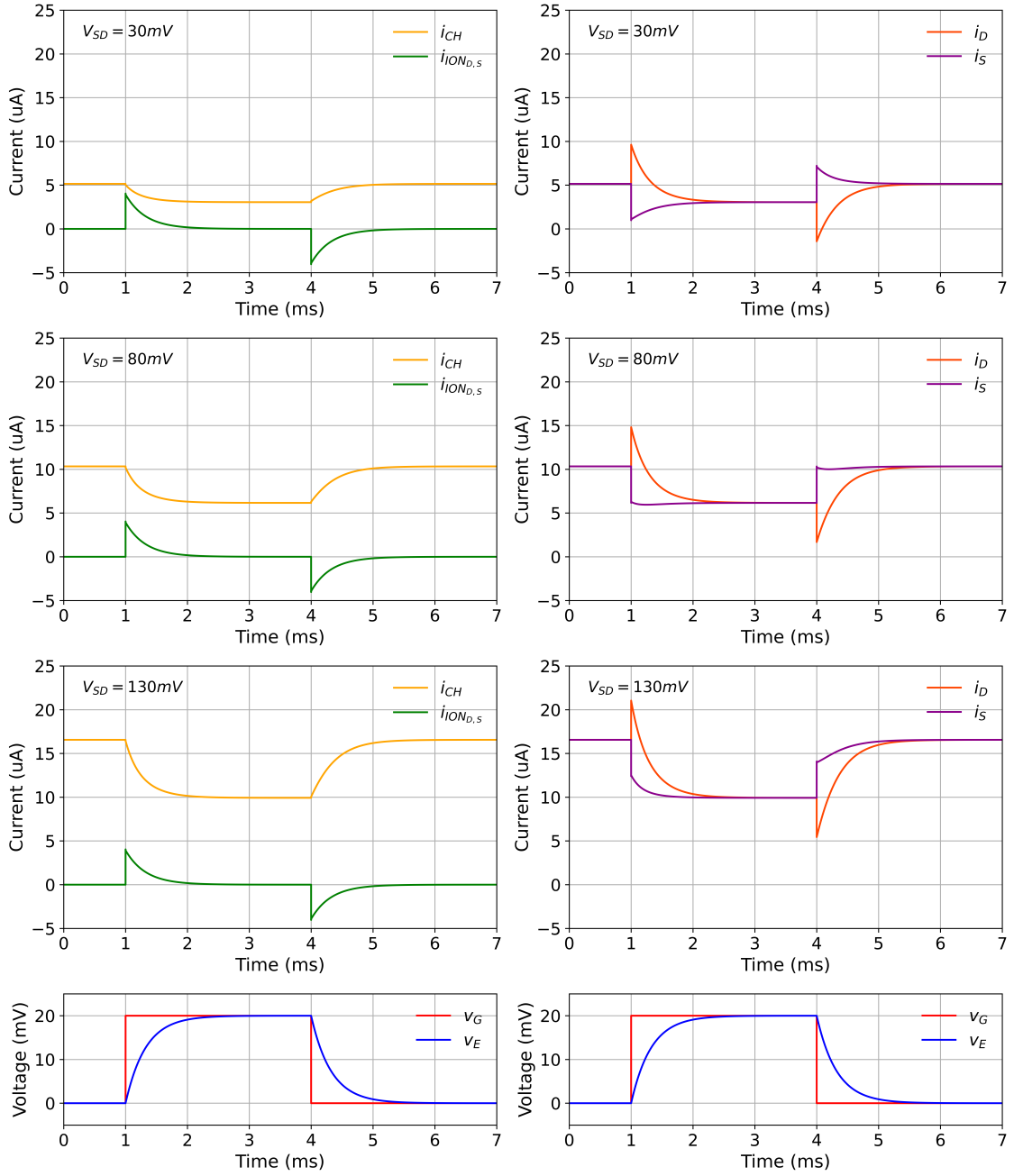


Figure 2.6: Channel current i_{CH} (yellow), split ionic displacement currents $i_{ION_{D,S}}$ (green), drain current i_D (orange) and source current i_S (purple) simulated for three different values of V_{SD} applied to the Friedlein modelled OECT. The applied gate voltage pulse $v_G(t)$ is the one reported previously. i_D always shows a "spike-and-recovery" behavior. i_S shows three possible behaviors: "spike and recovery" (upper case), "fast switching" (middle case), "monotonic decay" (lower case).

2.4 Double-gate architecture

As we saw in Sec. 2.2, OECTs can be fabricated either with a Ag/AgCl pellet gate electrode cantilevered on top of the channel, or with a lateral planar gate covered by organic polymer. The latter architecture in particular has been exploited in literature for the fabrication of OECTs with multiple gates.

These can be several and placed at different distances to the channel, to encode different spatiotemporal information [63,64]. In this case, the distance to the channel is related to the ionic resistance between each gate and the channel. A gate pulse exhibits then different ionic time constants depending on which gate it is applied to, and is sensed differently by the channel.

Another possibility is to have two gates, where the second one is used as a tuning gate [65]. A planar double-gate OECT architecture is shown in Fig. 2.7. In a double-gate device, the first gate is operated as the standard main gate (sweep gate). The second gate can be used as a control gate with an applied DC bias, to shape the OECT's transfer and output characteristics. This feature allows tunability of a single-device in situ, after this has been fabricated with a specific design.

The tuning of the double-gate characteristics by the voltage at the second gate can be expressed as a dependence of the threshold voltage, as reported in [65]. The threshold voltage V_T for a p-type double-gate OECT follows the formula:

$$V_T = V_T^0 \frac{(A_{gate_1} + A_{gate_2} + A_{channel})}{A_{gate_1}} - V_{SG_2} \frac{A_{gate_2}}{A_{gate_1} + A_{gate_2}} \tag{2.24}$$

where V_T^0 represents the threshold voltage without the presence of gate 2; V_{SG_2} is the voltage applied to the second gate (with respect to the source); A_{gate_1}, A_{gate_2} and $A_{channel}$ are the areas of gate 1, gate 2 and channel, respectively.

This dependence on V_{SG_2} allows to both decrease and increase V_T: if $V_{G_2} < V_S$ ($V_{SG_2} > 0$), then the threshold voltage decreases; conversely, if $V_{G_2} > V_S$ ($V_{SG_2} < 0$), the threshold voltage increases. This second case allows potentially to tune the OECT to always operate in weak inversion, for any biasing condition under a specified circuit voltage supply. For example, we can think of a circuit with a $0.6V$ supply, and standard OECTs spincoated with a polymer with threshold voltage of around $0.2V$. Clearly, any signal produced by the circuit is limited by the voltage supply, i.e. it has amplitude $< 0.6V$. Any standard OECT is biased either in weak, moderate or strong inversion depending on its biasing conditions. A double-gate OECT with a sufficiently increased threshold ($> 0.6V$) however, always works in weak (or moderate) inversion independently of its source to gate voltage V_{SG_1}, as long as this is limited to the voltage supply.

This feature allows simplification of circuit biasing by only changing the voltage at

an independent node. It is therefore of great significance.

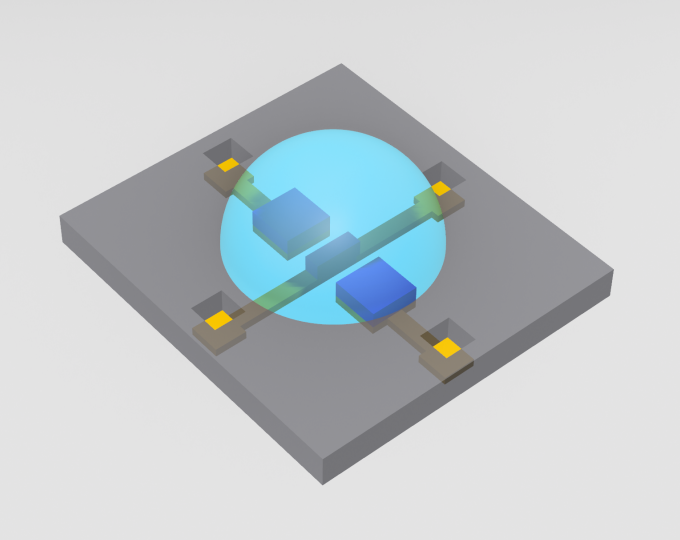


Figure 2.7: Schematic of a planar double-gate OECT with a drop of electrolyte connecting the two gates and the channel. The two gates have the same geometry, dimensions and distance to the channel, therefore they are equivalent.

Chapter 3

Hybrid organic-inorganic adaptive neuron

In Chapter 2 we described the operation of organic electrochemical transistors: we analyzed their steady-state and transient responses, their scaling properties and the possible device architectures. Above all, we introduced a mathematical model that predicts all these behaviors and features.

We gave a mathematical definition of two fundamental OECT's properties. We explained why and in which conditions OECTs show short-term plasticity (STP) by solving the Friedlein model for multiple pulses applied to the gate. We described in situ threshold voltage tuning in an OECT with a double-gate architecture, and we introduced a formula that could predict the V_T shift based on the source-gate voltage applied to the second gate.

In this Chapter, we make use of this complete mathematical model to develop a circuit architecture that emulates spike-frequency adaptation and that can be tuned in situ. We start by defining the analogy between the drain current of a short-term plastic OECT and the adaptation current equations that we introduced in Chapter 1. We describe how the OECT should be optimized to resemble as closely as possible the behavior of these equations. Also, we propose to tune the magnitude of this adaptation current by exploiting the double-gate architecture introduced previously. We then described the adaptive spiking circuit, its two blocks, its operation and biasing conditions. This first implementation of the artificial neuron is a hybrid version that combines OECT and CMOS technologies. To validate this design, we fabricated, characterized and then modelled on LTSpice the adaptive double-gate OECT and introduced it in a simulated circuit built with TSMC 0.18μm node. The results of this hybrid Spice simulated circuit are discussed in the last section.

3.1 Adaptive double-gate OECT optimization

In Sec. 1.2 we introduced a series of spiking neuron models from computational neuroscience that can be used to reproduce spike-frequency adaptation. In Sec. 2.3 we described the transient response of OECTs and the conditions in which they show short-term plasticity. Here we develop a brief mathematical discussion to show the

analogy between the adaptive current of spiking models and the plasticity of an OECT device. We start from Eq 1.7, defining the linear mechanism of adaptation introduced by Treves, and Eq 2.10, governing the effective voltage v_E sensed by the channel of an OECT with constant source and drain voltages. We report the two equations again here for clarity:

$$\tau_w \frac{d}{dt} w = -w + b\tau_w \sum_{t=t_{spk}} \delta(t - t_{spk})$$

$$\tau_i \frac{d}{dt} v_E = -v_E + v_G$$

The analogy between the two equations is evident: if the gate voltage signal v_G applied to the OECT is a series of Dirac delta functions, they are in fact the same equation.

This supports the thesis that the emulation of a bio-realistic adaptation current can be achieved in an artificial neuron circuit with a single OECT device. There are however a series of practical limitations and non-idealities that hinder this objective. These can be minimized with appropriate design choices.

Clearly, the first conclusion to the previous analogy is that the gate terminal of the OECT must be connected to the spiking output of the circuit. The circuit spikes must resemble Dirac deltas for the two equations to match. A Dirac delta can be approximated by a pulse (rectangular function) of duration ϵ and height $1/\epsilon$ with $\epsilon \to 0^+$, so that its integral over time is still equal to 1. In practice, this approximation holds if the circuit spikes have duration $\Delta t_{spk} << \tau_i$. Considering that bio-plausible adaptation time constants are $> 10ms$, this is not an issue for silicon circuits. However, it is a limitation for technologies that cannot reach high operating frequencies, such as electronics with OECTs. This point will be readdressed in Chapter 4 when dealing with the design of an all-organic adaptive spiking circuit. In any case, the tradeoff with decreasing the spike duration Δt_{spk} is that the response is attenuated of the same factor, since the amplitude of the input signal is fixed at v_{G_0} and does not increase as $1/\Delta t_{spk}$. The product $b\tau_w$ of the Treves adaptation equation is $v_{G_0}\Delta t_{spk}$ in the real case of the OECT equation. The spike duration has to be chosen carefully to have an adaptation that increases approximately linearly (with no saturation) for a sufficient number of spikes and a product $v_{G_0}\Delta t_{spk}$ that is sufficiently high to see the effect of spike-frequency adaptation on an appropriate time scale.

The second important note is that Eq 2.10 was derived imposing the conditions $V_S = const$ and $V_D = const$. This entails that the OECT must be introduced in a circuit architecture where the nodes corresponding to its source and drain are (approximately) constant over time. Oscillations in these two voltages would cause

variations of the ionic displacement currents i_{ION_D} and i_{ION_S}, according to Eq.s 2.7 and 2.8, that distort the short-term plasticity of the OECT. Looking at the circuit design, the source terminal can be simply connected to the voltage supply (V_{dd} for a p-type OECT and GND for a n-type). The connection to the drain terminal has to be chosen to allow extraction of the adaptive drain current without affecting the voltage of the node. This point is the main constraint behind the design of the Adaptation block, that will be described in the next section. As a preview, this block exploits the subthreshold swing of a current mirror in weak inversion to have the maximum range of current mirroring with the minimum change in gate voltage. If the gate voltage of the diode-connected transistor in the simple mirror does not change significantly, then this voltage is also the drain voltage of the adaptive OECT and the condition is respected.

Finally, the OECT transduces the voltage v_E connected to the spiking output of the circuit into the adaptive current. This introduces an additional equation $I_{adapt}(v_E)$, where v_E has the correct time evolution described by the adaptation equation. Depending on the region of operation of the OECT this relation can be linear (strong inversion, triode), quadratic (strong inversion, saturation), exponential (weak inversion). In our model of the OECT, this transduction produces the channel current i_{CH}, which follows the time evolution of v_E instantaneously. However, the total drain current of the OECT is composed by the sum of i_{CH} and i_{ION_D}, the portion of capacitive current flowing towards the drain. This additional contribution must be minimized to have a drain current that reproduces the dynamics of the adaptation current.

At the end of Sec. 2.3 we calculated the formula for the maximum amplitude of this capacitive spike. This was Eq 2.22, that we rewrite here:

$$I_{ION_D}(t_{0^+}) = \frac{C_{CH_D} v_{G_0}}{\tau_i} = \frac{f C^* W d L v_{G_0}}{\tau_i} = \frac{f v_{G_0}}{R_E}$$

where f is the f-factor defining the portion of capacitive current flowing towards the drain; v_{G_0} is the amplitude of the incoming gate pulse; R_E is the resistance of the electrolytic solution. We concluded that the electrolyte resistance has to be increased to reduce the amplitude of these spikes, at the expense of a slower OECT.

For our application the OECT has to be slow, in order to show plasticity. Its time constant τ_i (which is the adaptation time constant τ_w of the circuit) has to be in the order of $10ms$ to $100ms$, depending on the adaptation mechanism and the type of neuron that is emulated. For example it can range from $6ms$ to $2s$ for different types of cortical neurons [66–68]. The value set for the adaptation time constant must be obtained with a RC product with R that is maximized.

A first possibility is to decrease the electrolyte concentration, to linearly decrease the electrolyte resistivity and therefore the resistance. A decrease in the electrolyte

concentration can be also associated to a change in the conductivity of the channel and of the characteristics of the OECT though [69, 70]. Another option is to confine the electrolyte into a microfluidic channel or pattern a solid electrolyte so that the electrolyte is shaped into a long path with a small cross-sectional area, which is morphologically resistive. The effective resistance can also be increased by adding a series resistance to the electrolyte. We opted for this last option to have a degree of freedom after fabrication by which tuning the adaptation time constant. Also, if this series resistance is much larger than the electrolyte resistance, then the variation of R_E does not affect the dynamics of the OECT, with good approximation for a wide range of ionic concentrations. In this condition, only the dependence of the steady-state response remains.

This univocal dependence can be exploited to introduce a sensing property to the adaptive OECT: the transfer characteristic of the OECT changes depending on the electrolyte concentration, thereby causing a change in the magnitude of the adaptation current. This variation is then encoded by the neuron circuit into a different spiking frequency response, for the same input current stimulus.

To allow in situ tuning of the adaptation current we also introduce a second-gate to the adaptive OECT. Eq 2.24 showed that the effect of the voltage applied to this second gate is a shift in the threshold voltage of the OECT. The adaptation current then becomes a function of V_{SG_2}. Given the linear proportionality $V_T \propto -kV_{SG_2}$, we can write:

$$i_{adapt}(t) \approx i_{CH}(t) \propto \begin{cases} e^{\frac{v_{SE}+kV_{SG_2}}{nU_T}} & v_{SE} < V_T(V_{SG_2}) \\ v_{SE} + kV_{SG_2} & v_{SE} > V_T(V_{SG_2}), V_{SD} \leq v_{SE} - V_T(V_{SG_2}) \\ (v_{SE} + kV_{SG_2})^2 & v_{SE} > V_T(V_{SG_2}), V_{SD} > v_{SE} - V_T(V_{SG_2}) \end{cases} \tag{3.1}$$

where the first approximation neglects the capacitive ionic displacement current i_{ION_D}. These equations state the voltage V_{SG_2} can be used to shift horizontally (along the V_G axis) the transfer characteristic of the adaptive OECT. On a circuit perspective, as long as this shift does not cause the OECT to operate in a different region, then it can be viewed as a vertical shift (along the I_D axis) of the adaptation current I_{adapt}.

To conclude, this dependence can be also be used to balance a change in the transfer characteristic due to a change in the electrolyte concentration. For example, it can be used to tune the the adaptation current to be in a relevant range of magnitude for a chosen range of electrolyte concentration. This is equivalent to tuning of the ionic sensing properties of the OECT.

3.2 Circuit architecture

The circuit architecture for the hybrid adaptive neuron is reported in Fig. 3.1. It is made of two blocks.

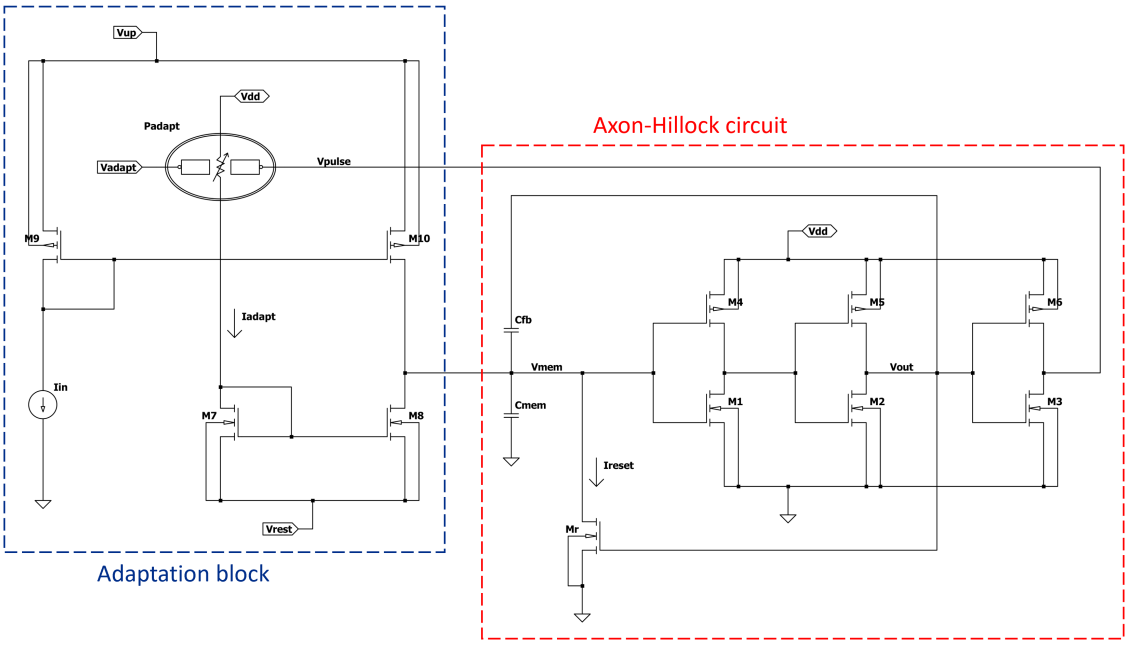


Figure 3.1: Schematic diagram of the hybrid adaptive artificial neuron. The circuit comprises an Axon-Hillock circuit (red, $M_{1-6,r}$) and an Adaptation block (blue, M_{7-10}, P_{adapt}).

The former is the Axon-Hillock circuit (red). This architecture was proposed by Mead in 1989 as one of the original circuits for generating discrete events, in the early-stages of VLSI silicon neuromorphic engineering [6]. Recently, it has been adopted by organic neuromorphics as a reference circuit for the fabrication of organic neurons, due to its simple design and excellent matching properties [45, 71, 72].

The latter is the proposed Adaptation block (blue). The Adaptation block is an analog current subtractor, which subtracts the tunable adaptation current I_{adapt}, obtained through the double-gate short-term plastic OECT, to the neuron's input I_{in}. The adaptation current increases at each spike event, therefore the effective current input to the Axon-Hillock decreases over time. This current variation results in a gradual reduction of the spiking frequency.

We will now analyse the single blocks independently in more detail.

3.2.1 Axon-Hillock circuit

The Axon-Hillock circuit (Fig. 3.2a) is a simple architecture that emulates the behavior of the Axon Hillock of a biological neuron[1]. It is a hardware implementation of an Integrate-and-fire (IF) neuron[2]. In this specific version of the circuit, a third inverter is added in series to the two that usually compose the non-inverting amplifying stage. The functionality of this third inverter is related to the integration with the Adaptation block and will be explained later in this section.

The evolution of voltages V_{out} and V_{mem} of a simulated Axon-Hillock circuit, for a constant current stimulus I_{in}, is reported in Fig. 3.2b.

The circuit integrates the current(s) I_{in} across its membrane capacitor C_{mem}. The actual capacitance value that integrates I_{in} is $C_{int} = C_{mem} + C_{fb}$, since the output node is initially shorted to GND and the two capacitors are therefore in parallel. The analog voltage V_{mem} increases linearly[3] until it reaches the inverter switching threshold. At this point, the two inverting stages cause the voltage V_{out} to quickly change from GND to V_{dd}. This sudden change has two effects: first, it activates a positive feedback through the capacitor divider implemented by C_{mem} and C_{fb}, which further increases the voltage V_mem very steeply; moreover, it switches on the reset transistor M_r. If the reset current I_{reset} is larger than the input I_{in}, the membrane capacitor is discharged, until it reaches the inverter's switching threshold again. V_{out} swings back to GND and the cycle repeats. The Axon-Hillock circuit, stimulated by a constant input, produces V_{out} voltage spikes (digital events) equally spaced in time, i.e. at a constant spiking frequency. The interspike interval (ISI) is inversely proportional to the neuron's input:

$$t_{low} = \frac{C_{fb} V_{dd}}{I_{in}} \tag{3.2}$$

The duration of a single spike depends on the reset current I_{reset}, as well as on I_{in}:

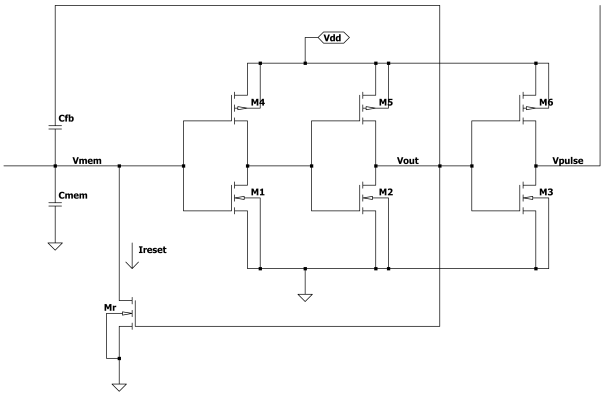
$$t_{hi} = \frac{C_{fb} V_{dd}}{I_{reset} - I_{in}} \tag{3.3}$$

From this last equation, we notice that the pulse duration is nearly independent of I_{in} if $I_{in} << I_{reset}$.

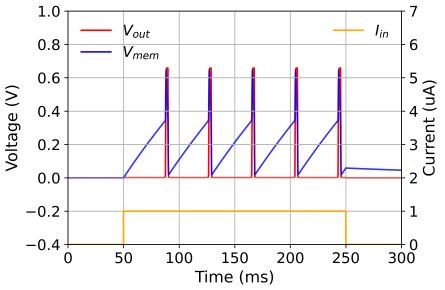
[1] The Axon Hillock is the portion of the neuron where action potentials (nerve pulses) are generated. It is located at the intersection between the first segment of an axon and the cell body. The Axon Hillock accumulates and stores inputs up to when the voltage across its membrane reaches a certain threshold value. At this point, nerve channels in the membrane engage in a cycle consisting of a positive feedback followed by delayed negative feedback, that generates a firing event and then causes the neuron to reset. A broad description of this mechanism is reported in [12–16] by Hodgkin, Huxley, and Katz, which represent the first quantitative model to describe the initiation of an action potentials, and their restoration as they propagate down the axon.

[2] In practice, it emulates a Leaky integrate-and-fire (LIF) neuron, since a non-infinite resistance is present in parallel to its membrane capacitor due to transistor leakage.

[3] V_{mem} increases with an RC time constant in the case of a LIF neuron.



(a) Schematic diagram of the Axon-Hillock circuit.



(b) V_{out} and V_{mem} traces over time in the simulated CMOS Axon-Hillock circuit for a constant input current stimulus.

Figure 3.2: Simulation of a CMOS Axon-Hillock circuit in TSMC 0.18μm node. $V_{dd} = 0.7V$, $C_{mem} = C_{fb} = 50nF$. A leakage parasitic resistor of $2M\Omega$ was added in parallel to the membrane capacitor for the purpose of the simulation.

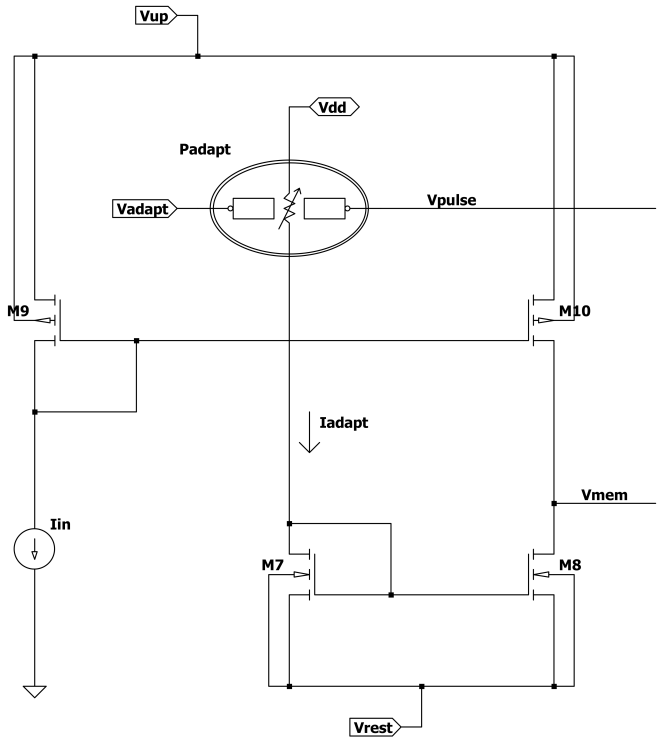
The main advantage of the Axon-Hillock architecture is that mismatch in the circuit mostly depends on the matching properties of the two capacitors, rather than any of its transistors. This feature is beneficial for CMOS implementation, but even more for other technologies that are affected by greater PVT (process voltage temperature) variations, such as organic electronics.

The tradeoff with this excellent matching is that the circuit dissipates significant power for slowly varying input signals: as the voltage V_{mem} gradually increases to cross the switching threshold, the first inverter stays in its fully conductive state (with both nFET and pFET conducting) for a considerable amount of time.

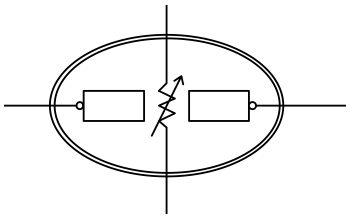
Finally, many spiking characteristics, e.g. the neuron's threshold voltage, are set at circuit design, depend on the technological process, and cannot be tuned after fabrication. However, the simple design and low number of transistors enable many emerging technologies to start testing neuromorphic circuitry.

3.2.2 Adaptation block and integration

The Adaptation block is the newly proposed architecture to implement tunable spike-frequency adaptation (Fig. 3.3a). It uses two current mirrors, one p-type and one n-type, to perform analog current subtraction between the neuron's input I_{in} and the adaptation contribution I_{adapt}, which is the drain current of the short-term plastic double-gate OECT (Fig. 3.3b). In this first version, the circuit is an organic-inorganic hybrid, since the current mirroring operation is performed with MOSFETs.



(a) Schematic diagram of the Adaptation block.



(b) Schematic symbol of the double-gate OECT.

Figure 3.3: Architecture of the proposed Adaptation block in its hybrid organic-inorganic version.

The connection to the Axon-Hillock circuit is as follows.

Gate1 of the double-gate OECT is connected to the output of the third inverting stage of the Axon-Hillock. This node provides voltage spikes that are inverted with respect to the spiking output V_{out}: Gate1 is at voltage V_{dd} whenever the neuron in not spiking and at GND during a spike, so that the drain current of the plastic OECT increases at each pulse.

The Axon-Hillock is usually designed with two inverters. The third stage is not necessary to produce the output spike. However, it is necessary in our case to provide a sufficiently high gain on the pulses that are sent to Gate1. If the output of the first stage were to be used instead of this additional node, the incoming gate pulses would have a non-negligible rise time. This is especially true because the input of the first stage is V_{mem}, which is a slow varying signal. This can cause the circuit to get stuck in an intermediate state where the first inverter's output is not low enough to make V_{out} spike and reset the circuit, but makes I_{adapt} increase sufficiently to have a net zero integrated current on the membrane capacitor. As a result V_{mem} remains stuck in between GND and V_{dd} until the neuron's input

changes.

The output of the Adaptation block is connected directly to V_{mem}. This allows the adaptation current to discharge the membrane capacitor to the rest potential even when there is no input external current I_{in}. Otherwise, the Axon-Hillock alone would discharge just by passive parasitic leakage.

The output swing of the current subtractor is given by the saturation conditions of the two current mirrors. For the two mirrors in weak inversion this is $V_{mem} \in [V_{up}-4U_T; V_{rest}+4U_T]$, where $4U_T \approx 100mV$. Since V_{mem} is limited by GND and V_{dd}, then the block has to be biased with $V_{up} \geq V_{dd}+100mV$ and $V_{rest} \leq GND-100mV$. In reality, V_mem does not ever reach V_{dd} during the capacitive feedback of the Axon-Hillock and we can choose to have $V_{up} = V_{dd}$.

In the single Axon-Hillock, the resting potential of the membrane voltage is GND and the reset potential is also GND. This is caused by the reset transistor, that resets V_{mem} to GND to end a spike and is a leak conductance to GND in its off state. When adding the Adaptation block, the properties of the membrane node change. This is because the transistor M_8 of the p-type mirror is effectively a resistance in parallel to the reset transistor. The two channel resistances share the voltage at their drain, but not the one at their source ($V_{rest} < GND$). The membrane voltage can rest between these two values, depending on the ratio between the two resistances.

We first think about the case where V_{mem} rests at GND. This is possible when the off current of the reset transistor at $V_{GS} = 0, V_{DS} = 0$ is $>>$ than the off current of the double-gate OECT at $V_{SG} = 0, V_{SD} > V_{SD,sat}$ (the n-type mirror is still in saturation so the OECT's current is mirrored exactly). During spiking activity, the adaptation current can increase above the reset off current. If this occurs, the adaptation current can effectively hyperpolarize the membrane voltage when the external stimulus ends, i.e. bring the membrane voltage below its resting potential (which is GND, in this case). The OECT drain current starts decaying exponentially at the end of the stimulus. Also, as V_{mem} decreases below GND, the mirror leaves the saturation condition and does not effectively copy the adaptive current. At this point the channel resistance of the reset transistor becomes smaller than the one transistor M_8 once again and the membrane voltage returns to GND.

The opposite condition is to have V_{mem} that rests at V_{rest}. In this case, the n-type mirror starts off of saturation but meets it as soon as the membrane voltage starts increasing. Also, saturation is met throughout the spiking activity because the reset transistor resets towards GND.

As we discussed in Sec. 3.1, the source and drain terminals of the adaptive OECT have to be approximately constant in voltage over time, to have the correct evolution of the adaptation current. The Adaptation block has to mirror the adaptation cur-

rent while meeting this constraint. A simple weak inversion current mirror provides effectively these operating conditions.

To understand why the double-gate OECT's drain voltage is approximately constant over time in this architecture, we can think at the two transduction steps between the generation of the adaptation current and its mirroring. $I_a dapt$ is caused by a spiking event at Gate1 of the OECT. This gate voltage pulse is attenuated by the RC integrator that describes the device, so that only a small portion of its amplitude is effectively sensed by the channel: this is voltage v_E. This small change causes an increase in $I_a dapt$ with a proportionality defined by the operating region of the OECT. The n-type diode-connected transistor matches this current increase by increasing its gate voltage (which is also its drain voltage), thereby increasing the voltage at the drain of the OECT. We suppose that this occurs instantaneously, i.e. the delay of the matching is negligible compared to the time constant of the OECT. If the OECT and the diode-connected transistors are perfectly balanced, both in saturation and weak inversion, then $\Delta v_D = -\Delta v_E$ and the condition is not met. However, if the OECT is operated in a different region, then the diode-connected transistor has a much greater amplification and the variation of v_D occurring to match the adaptation current is minimized. This also holds if the OECT and the mirror are both in weak inversion but the mirror has a much lower subthreshold swing. In this case, the total variation of v_D is attenuated by a factor $\frac{n_{OECT}}{n_{MIR}}$ compared to the variation of v_E, where n_{OECT} and n_{MIR} are the slope factors of the OECT and the diode-connected transistor, respectively. Subthreshold engineering can be exploited to fabricate an adaptive OECT with a small slope factor, to be used in the Adaptation block.

The Adaptation block provides a compact architecture to generate and mirror the adaptation current to the membrane node. If we do not consider the upper p-type mirror, which mirrors the external input, the block is made of just three transistors. The Adaptation block is also very versatile, since it can be integrated with any spiking circuit that has a downward spiking node and an integrating node that swings while preserving the saturation conditions of the two mirrors. Finally, the operation of the two mirrors in weak inversion is also an effective choice to limit power consumption, which is a necessary requirement of spiking circuit architectures.

3.3 OECT fabrication

At this point of the project, we want to validate the adaptive spiking architecture that we proposed. The first step is to fabricate the double-gate adaptive OECT, following the considerations and design rules discussed in Sec. 3.1 and 3.2. We report here the process flow used for the fabrication of the device, known as double-Parylene process. A schematic overview of the process steps is reported in Fig. 3.4.

The first part of the process concerns the deposition of the device metallic contacts and traces. The starting substrate is a silicon wafer with an upper thermal oxide layer of $300nm$, that ensures sufficient isolation. The wafer is cleaned of the organic residues with organic solvents and then undergoes a dehydration bake. A lift-off layer (LOL) (Microposit LOL-2000) is spin-coated on the substrate to achieve a thickness of around 200nm and is then baked for 30 minutes at $190°C$. It is crucial that this baking time is respected to ensure high resolution of the lift-off process. Photoresist SPR3612 is spin-coated on top of the LOL layer with a thickness of $1um$ and no primer[4]. The photoresist is then exposed with the design for the device contacts and traces and undergoes a post-exposure bake and a developing step. A $8nm$ layer of Titanium followed by a $55nm$ layer of Gold are evaporated with an e-beam evaporator. The wafer is soaked overnight in the lift-off remover bath to gently remove the photoresist and LOL layer[5]. At this point of the process, the wafer shows the patterned metallization layer (step 5 of Fig. 3.4).

The second part of the process is about the deposition and patterning of the double Parylene layer, that is used to spin-coat the polymer only on the active regions of the device. The wafer with the metallic contacts is coated with $1.5um$ of Parylene-C as the insulating layer. This layer has the same purpose of the field oxide in CMOS process flow, i.e. isolate the single devices. The Parylene-C layer is cross-linked with the adhesion promoter Silane A 174, to ensure good adhesion to the wafer. A dilute soap solution (3 % Micro-90 in water) is spin-cast on top. This intermediate layer prevents the second Parylene-C layer, deposited now with the same thickness, to adhere strongly to the surface underneath. The purpose of the double-Parylene layer with the soap solution in between is to peel-off the upper layer at the end of the process. The wafers are then coated with a hard mask layer made of $75nm$ of e-beam evaporated Titanium. Another lithography step is performed to define the active regions where the conductive polymer will be deposited, i.e. the channel and the gates of the OECT, and the contact pads. This step is composed once again of a SPR3612 photoresist spin-coating, an exposure step, that is now an aligned exposure, a post-bake and a developing step. To recapitulate, the wafer now presents the gold contacts and traces, a field-oxide layer that covers the entire wafer, a soap solution and another Parylene-C layer on top, the upper photoresist with the pattern for the conductive polymer. This is summarized by step 11 in Fig. 3.4. Finally, the Titanium hard mask is dry-etched with a reactive ion etching (RIE) step with the pattern of the upper photoresist layer. The same pattern is dug into the two Parylene layers by another RIE step, selective for oxides. The wafer is diced into the single chips that are spin-coated with the conductive polymer.

[4]A primer is a passivating layer that is spin-coated before a photoresist to promote its adhesion to the substrate. It works by preventing humidity absorption on the surface.

[5]To speed up the process it is possible to sonicate the bath but this may remove or damage portions of the metal design, especially finer features

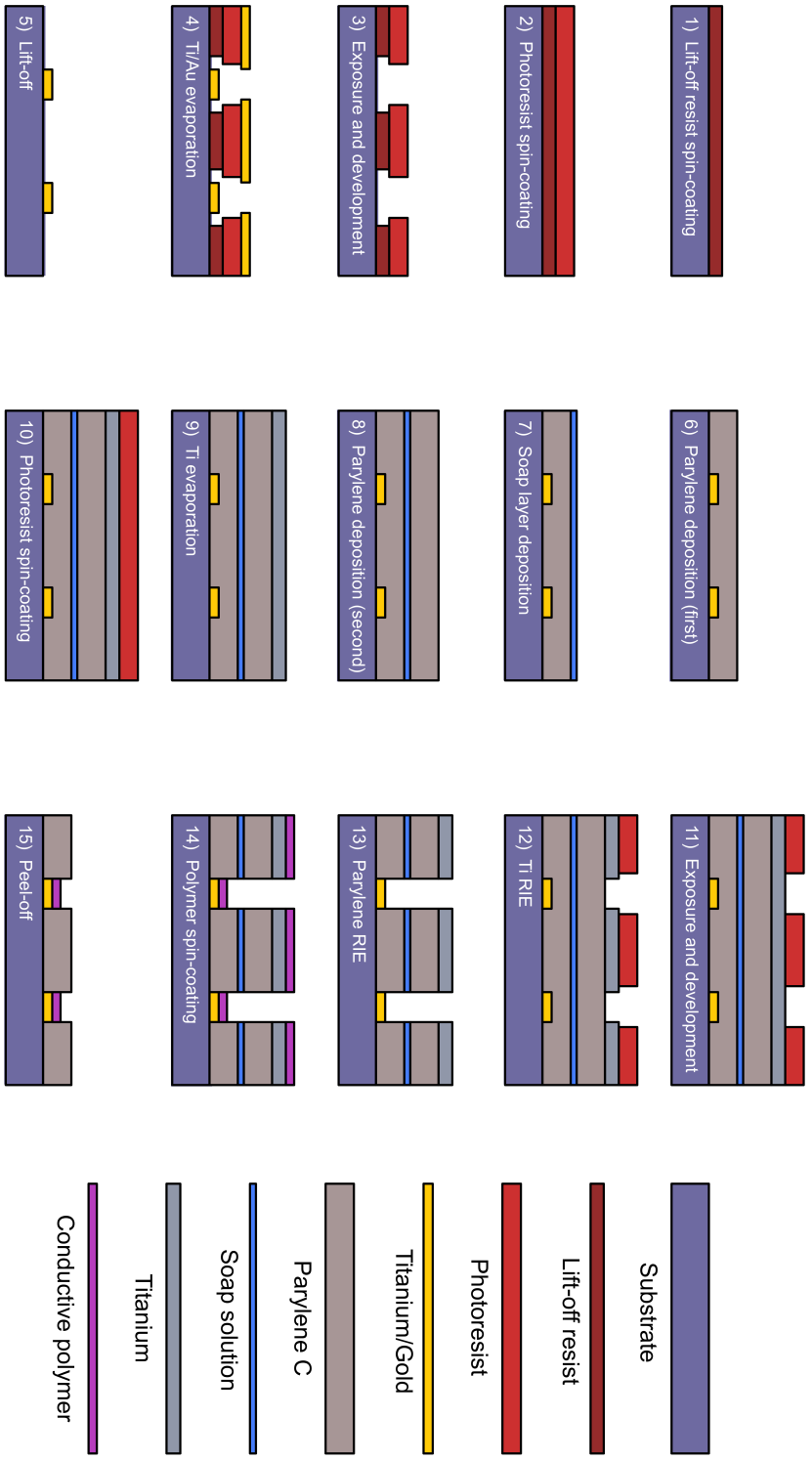


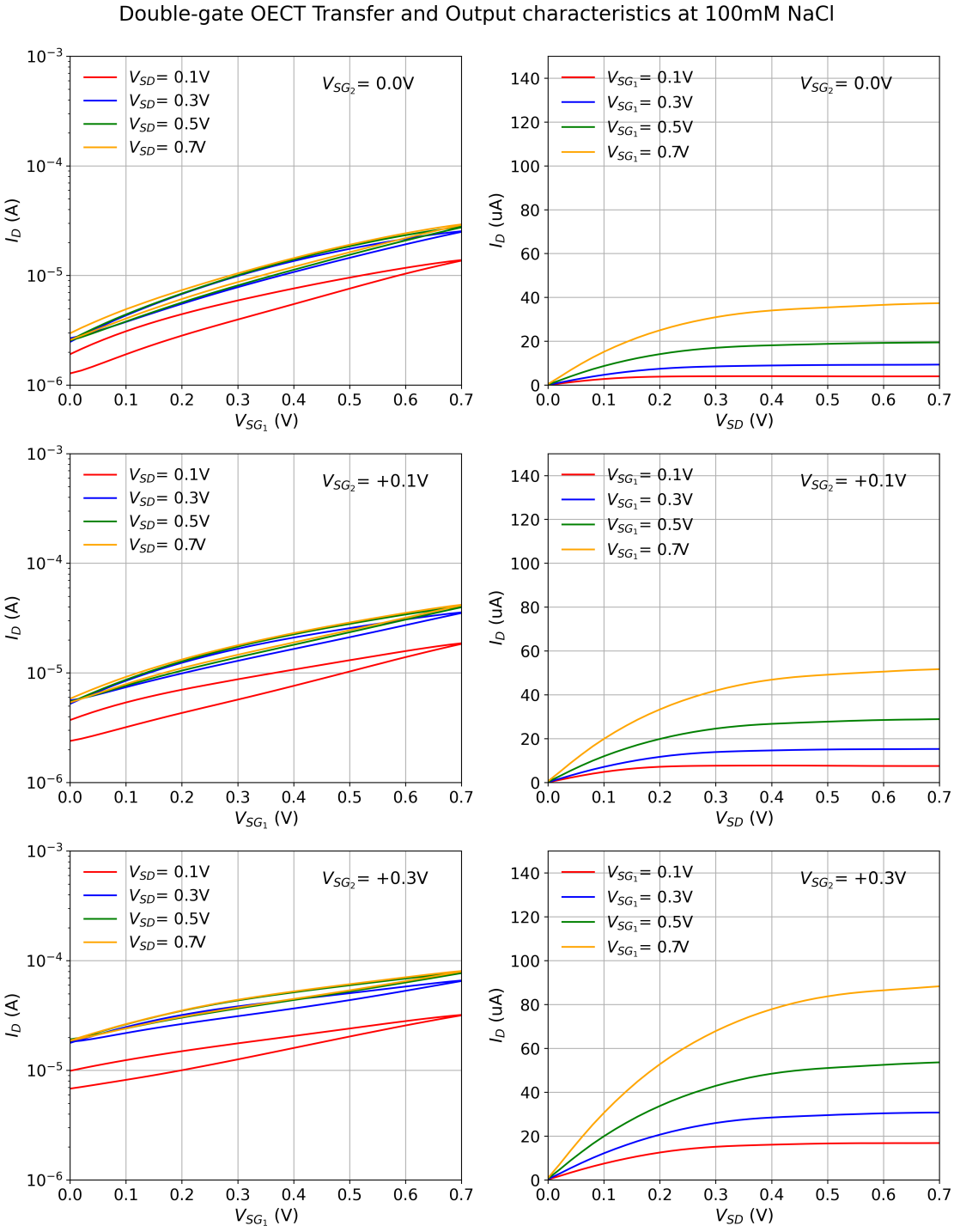
Figure 3.4: Schematic overview of the double-Parylene process used for the fabrication of the double-gate OECT.

For the double-gate OECT we used p(g2T-TT)[6], a hole conducting polymer with a high I_{ON}/I_{OFF} range and fast switching [73]. The dimensions of the channel are $150um$ x $50um$ (W x L) plus a $10um$ long contact to the source or drain trace at each side of the channel. The two gates are symmetric with respect to the channel and have dimensions $320um$ x $320um$, to respect the condition for efficient gating ($A_{gate} \geq 10A_{ch}$). p(g2T-TT) was dissolved in chloroform at a concentration of $2.5mg/mL$ and the solution was stirred at $100rpm$ for 2 hours. The polymer was spin-coated from hot ($50°C$) at $3000rpm$ with $1000rpm/s$ acceleration for 60s. The chip was then baked at $65°C$ for 20 min to let the solvent evaporate. The top Parylene-C layer was then peeled off to confine the organic semiconductor only in the lithographically defined channel and gate regions. The electrolytic solution was then drop-cast on top of the finished double-gate OECT.

3.4 Double-gate OECT characterization

We proceeded with the characterization of the fabricated device. Fig. 3.5 shows the transfer and output characteristics of the double-gate OECT for different values of V_{SG_2} applied between the source and the tuning Gate2. The transfer characteristics were obtained by sweeping the voltage difference V_{SG_1} from 0.0V to 0.7V and back. It is evident how the double-gate OECT presents hysteresis in the transfer curve. This phenomenon is characteristic of many OECTs and is reported often times in literature [54, 74, 75].

[6] poly(2-(3,3-bis(2-(2-(2-methoxyethoxy)ethoxy)ethoxy)-[2,2-bithiophen]-5-yl)thieno[3,2-b]thiophene)


Double-gate OECT Transfer and Output characteristics at 100mM NaCl
V_{SD}= 0.1V
V_{SD}= 0.3V
V_{SD}= 0.5V
V_{SD}= 0.7V
V_{SG_2}= 0.0V
I_D (A)
V_{SG_1} (V)
V_{SG_1}= 0.1V
V_{SG_1}= 0.3V
V_{SG_1}= 0.5V
V_{SG_1}= 0.7V
I_D (uA)
V_{SD} (V)
V_{SG_2}= +0.1V
V_{SG_2}= +0.3V

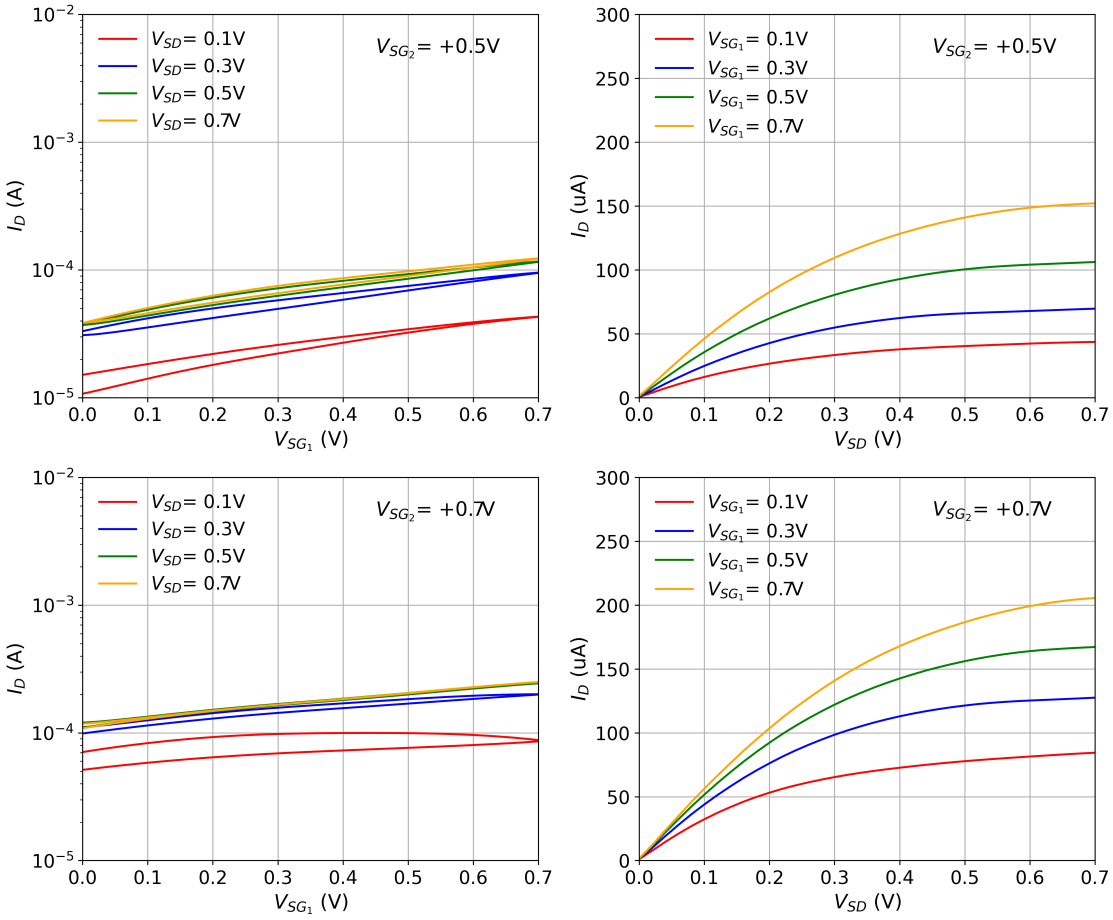


Figure 3.5: Steady-state characterization of the fabricated double-gate OECT for different values of V_{SG_2}. The transfer characteristics show an hysteretic loop, which is common in OECTs. The presence of the second gate reduces the modulation of I_D with V_{SG1}. However, it allows to tune the OECT's threshold voltage in-situ.

In Sec. 2.4 we discussed that the effect of the voltage applied to Gate2 of the OECT is to shift its threshold voltage. According to Eq. 2.24, the threshold decreases linearly with an increasing V_{SG_2}, with a proportionality factor given by the ratio between the area of Gate2 and the sum of the areas of Gate1 and Gate2. Our double-gate OECT has symmetric and equally dimensioned gates so we expect a proportionality factor of $\frac{1}{2}$. We extracted the threshold voltages values V_T from the transfer curves of Fig. 3.5 using the Extrapolation in the linear region method (ELR), which is one of the most commonly used approaches to extract this quantity in MOSFETs [76, 77]. The ELR method consists in finding the gate-voltage (V_{SG_1}) axis intercept ($I_D = 0$) of the linear extrapolation of the transfer curve at its maximum first derivative (slope) point (i.e. the point of maximum transconductance). The value of V_T is calculated by adding $\frac{V_{SD}}{2}$ to this intercept. In Fig. 3.6 we report graphically this extraction procedure and show the measured variation of V_T as a function of

V_{SG_2}. The threshold changes linearly with good approximation between $V_{SG_2} = 0V$ and $V_{SG_2} = 0.5V$. The shift in threshold is about -0.31V over the entire range. The proportionality factor for threshold modulation is therefore approximately 0.62, similar to the theoretical value that we expected. In Fig. 3.6 we also display the variation of the I_{ON} and I_{OFF} currents, taken at $V_{SD} = 0.7V$, $V_{SG_1} = 0.7V$ and $V_{SD} = 0.7V$, $V_{SG_1} = 0.0V$, respectively.

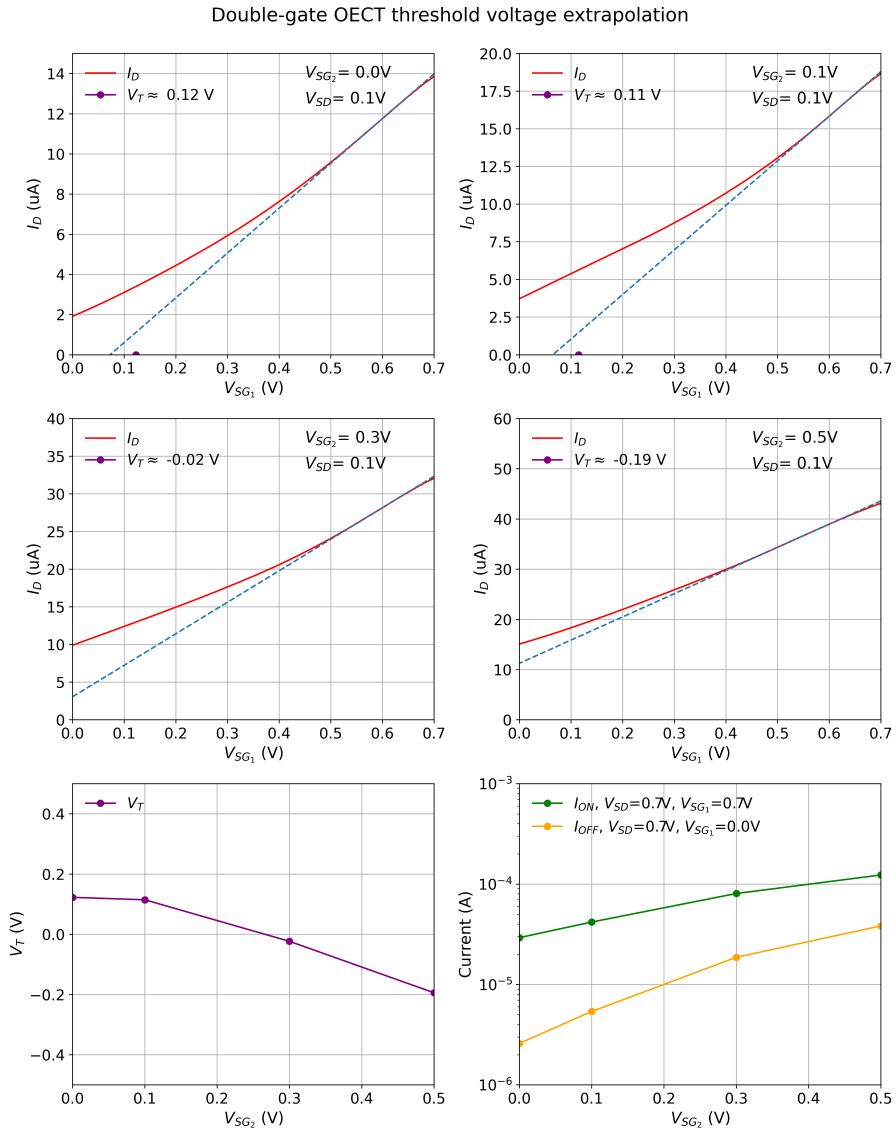
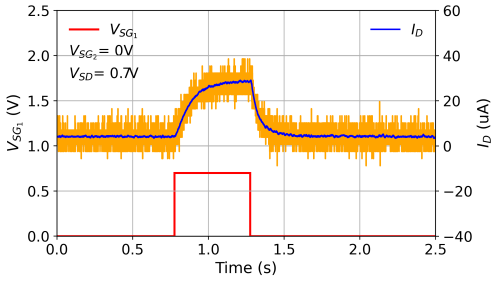


Figure 3.6: Extraction of the double-gate OECT threshold voltage using the ELR method, for different values of V_{SG_2}. Threshold modulation and the variation of the I_{ON} and I_{OFF} currents with V_{SG_2} are summarized in the last two plots.

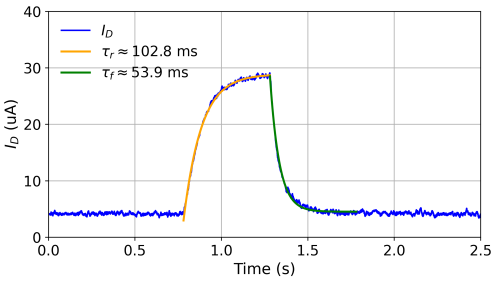
The transient response characterization was aimed at measuring the time constant of the OECT and verifying that its drain current behaves as an adapting current when

a train of pulses is applied to Gate1. To be able to tune the time constant of the OECT we introduced a limit resistor of $470k\Omega$ in series to the gate. This additional resistor is much larger than the resistance of the electrolytic solution and allowed us to achieve time constants in the order of 10-100ms, which is in the range discussed in Sec. 3.1 for biological adaptation mechanisms. Moreover, it limits the OECT's dependence on the electrolyte concentration just to its steady-state characteristics. As we mentioned, this univocal dependence could be exploited to develop a sensory neuron that encodes variations in the electrolyte concentration into different onset and steady-state spiking frequencies.

The response of the OECT to a single long (0.5s) gate pulse is reported in Fig. 3.7a. The current signal was measured as a voltage signal on a shunt probe resistor and then filtered using a moving average window on each data point. We extracted the OECT's time constants from this signal by applying a negative exponential fit, coherently with the model described, that is shown in Fig. 3.7a. The OECT shows two different rise and fall time constants, with $\tau_r \approx 2\tau_f$. This mismatch is common in semiconducting polymers, both p-type and n-type [78, 79].



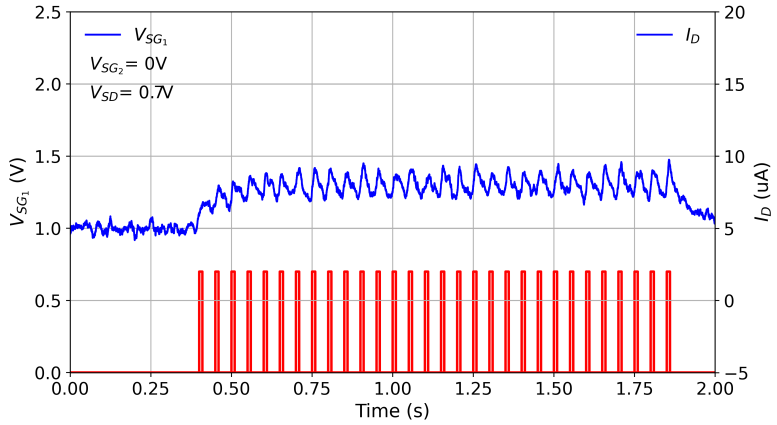
(a) Drain current response of the double-gate OECT to a prolonged single gate pulse of duration 0.5s (red). The plot shows the meaasured current signal (orange) and its filtered version (blue) obtained by applying a moving average window on each point of the dataset.



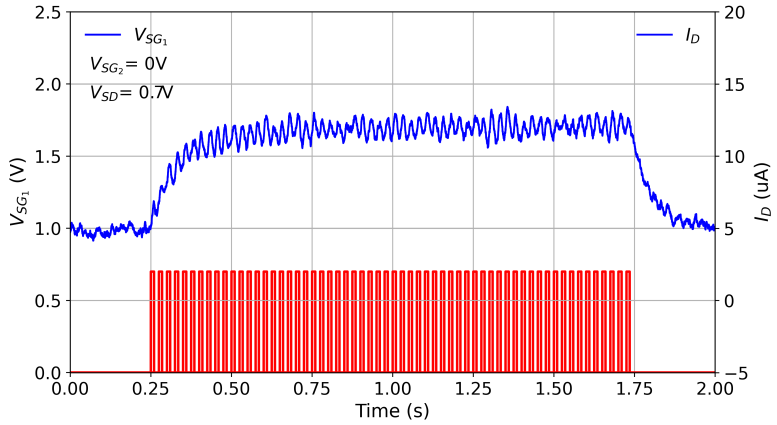
(b) Negative exponential fit on the filtered drain current signal. The OECTs shows different rise and fall time constants: $\tau_r > \tau_f$.

Figure 3.7: Extraction of the characteristic time constants of the OECT with a negative exponential fit.

To conclude our characterization, we tested the OECT's response to a series of Gate1 voltage pulses, to verify the presence of short-term plasticity. The OECT was tested with voltage spikes of duration $\Delta t_{spk} = 10ms$ at two constant spiking frequencies, 20Hz and 40Hz. The measured drain currents are shown in Fig. 3.8. An adapting current is evident in both cases. The increase is higher in the second case due to the higher frequency of the applied spikes. The amplitude of the gate pulses is kept constant between the two tests.



(a) Double-gate OECT response to a 20Hz gate spiking signal.



(b) Double-gate OECT response to a 40Hz gate spiking signal.

Figure 3.8: Adapting drain current of the double-gate OECT stimulated with a 20Hz and 40Hz spiking gate signal. The pulse duration as well as its amplitude are kept constant between the two cases: $\Delta t_{spk} = 10ms$, $V_{SG_1 max} = 0.7V$.

3.5 Double-gate OECT modeling

The next step was to build a spice model for the double-gate short-term plastic OECT that we just tested.

The transient response of the OECT and its plasticity can be reproduced via the Friedlein model, thaw we studied in Sec. 2.3. To be able to simulate the steady-state response of the OECT we have to develop an additional model that can show the threshold voltage modulation given by Gate2. For this purpose, we created an equivalent model that uses a non-inverting summing amplifier to sum the voltages at two terminals and apply the sum to the gate of a simple transistor model. The

two terminals are equivalent to Gate1 and Gate2 of the double-gate OECT.

The complete equivalent circuit to simulate the behavior of the double-gate OECT is shown in Fig. 3.9. The voltages at Gate1 and Gate2 are summed at the gate node of the Friedlein model, which then provides the low-pass filtering and the capacitive current contributions characteristic of the OECT. An important point about the design of the equivalent circuit is that the resistor R1 must be connected to the source in order for the two gates two provide a sum that is relative to the source, i.e. an equivalent V_{SG} to the gate of the Friedlein model.

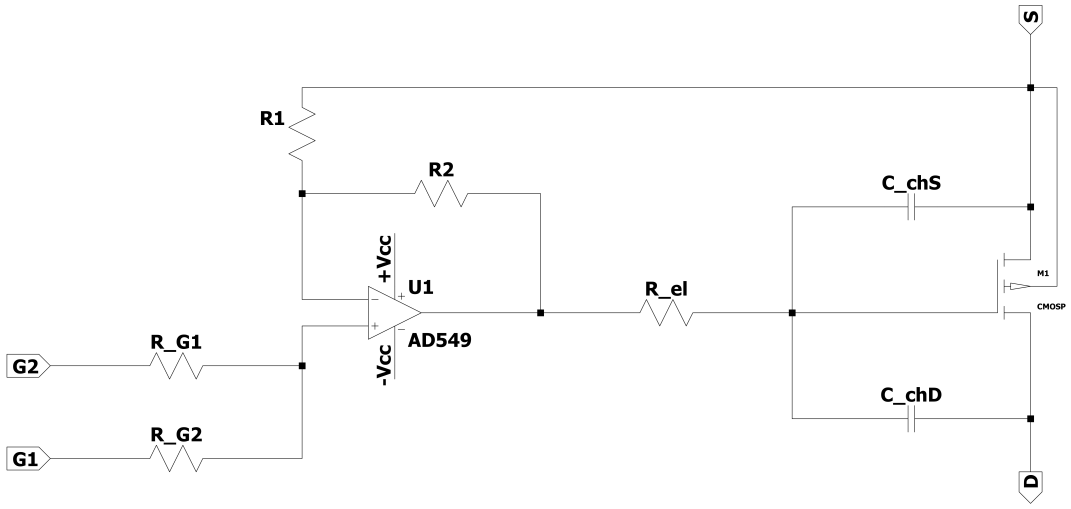
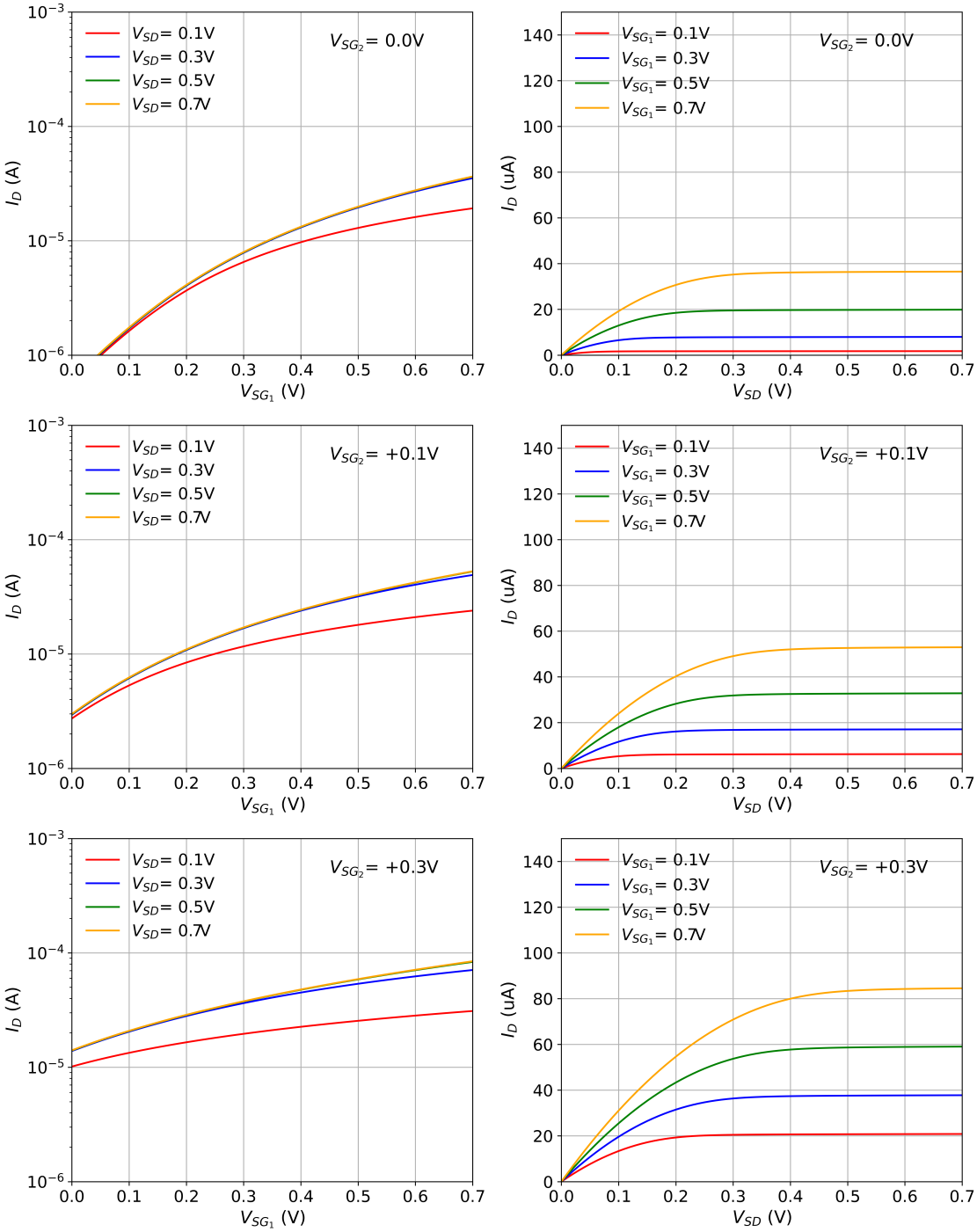


Figure 3.9: Complete equivalent model to simulate the behavior of the double-gate plastic OECT. The steady-state response, double-gate architecture and threshold voltage modulation are reproduced by the non-inverting summing amplifier. The transient response and short-term plasticity are modelled by the Friedlein equivalent circuit.

The complete equivalent circuit was tuned to fit the behavior of the characterized double-gate. We show here the transfer and output characteristics, the transient response to a prolonged gate voltage pulse and to multiple fast pulses of the final spice equivalent model. All the simulations are done with the same conditions and signal parameters as the measurements reported in the previous section for the fabricated double-gate OECT. The model is effective at describing with good precision both the steady-state and the transient response of the double-gate OECT. In particular about the transient response, it also replicates the mismatch of the rise and fall time constants reported earlier. For the spice equivalent model these two constants are: $\tau_r \approx 109.6ms$, $\tau_f \approx 41.3ms$.


V_{SD}= 0.1V
V_{SD}= 0.3V
V_{SD}= 0.5V
V_{SD}= 0.7V
V_{SG_2}= 0.0V
I_D (A)
V_{SG_1} (V)
V_{SG_1}= 0.1V
V_{SG_1}= 0.3V
V_{SG_1}= 0.5V
V_{SG_1}= 0.7V
I_D (uA)
V_{SD} (V)
V_{SG_2}= +0.1V
V_{SG_2}= +0.3V

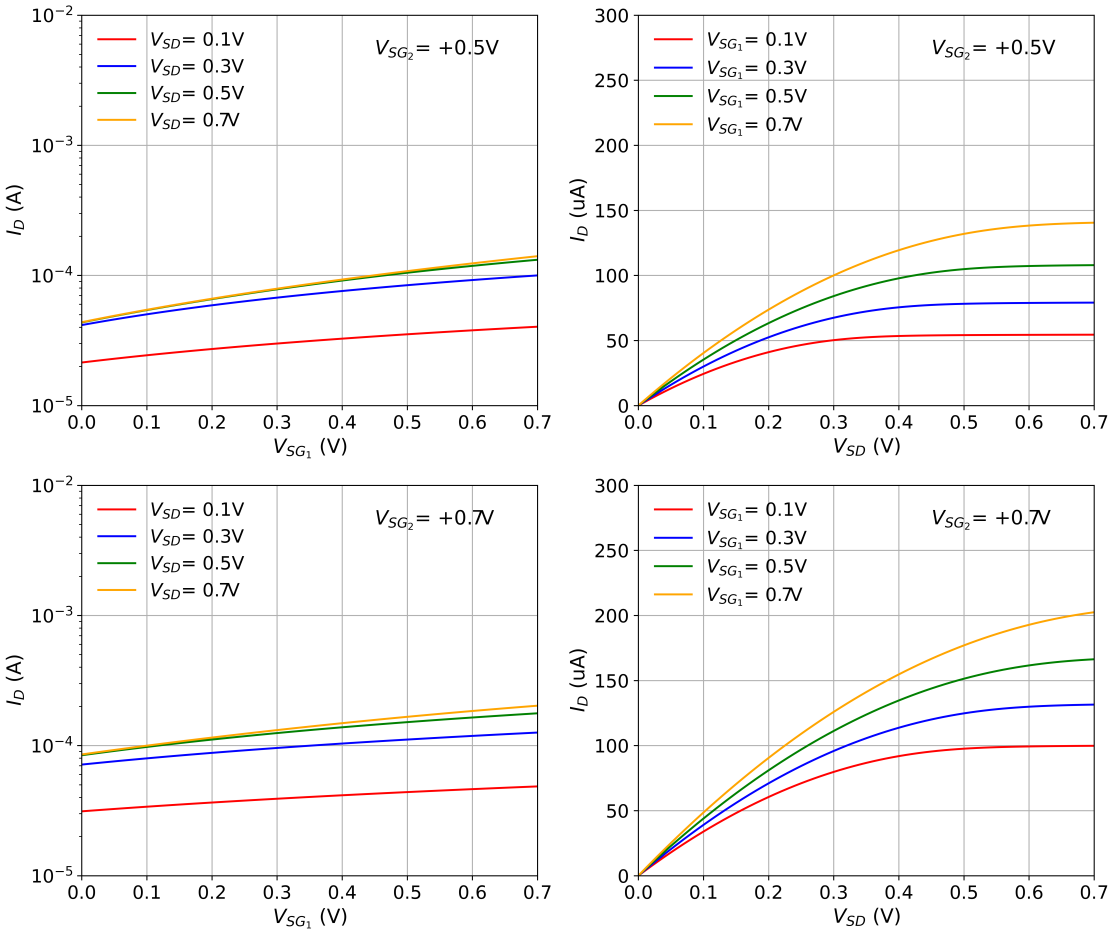
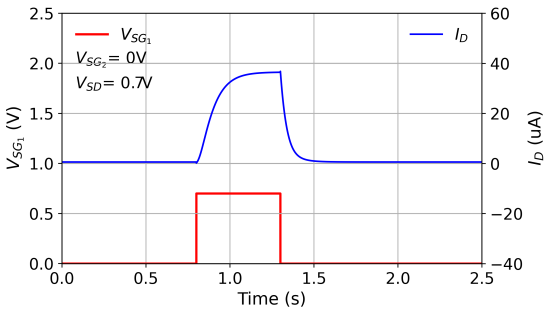
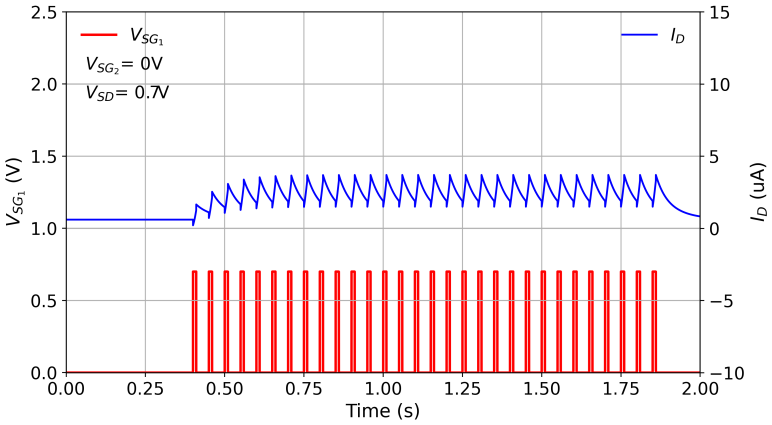


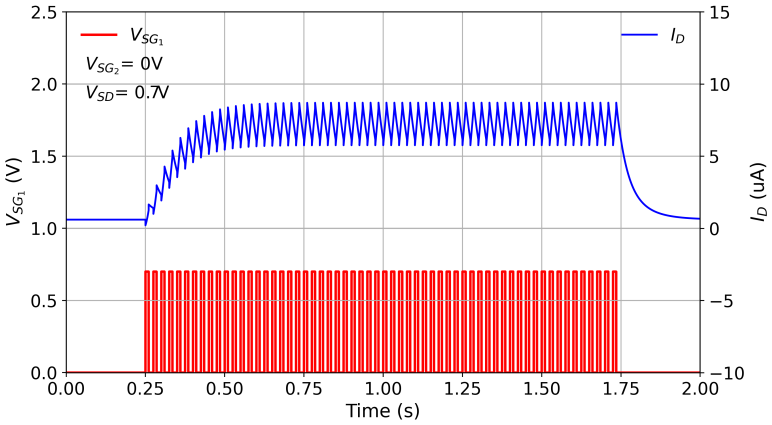
Figure 3.10: Steady-state modeling of the double-gate OECT at the measured values of V_{SG_2}.



(a) Response of the fitted spice equivalent model to a prolonged single gate pulse.



(b) Response of the fitted spice equivalent model to a 20Hz gate spiking signal.



(c) Response of the fitted spice equivalent model to a 40Hz gate spiking signal.

Figure 3.11: Transient behaviors of the fitted spice equivalent model.

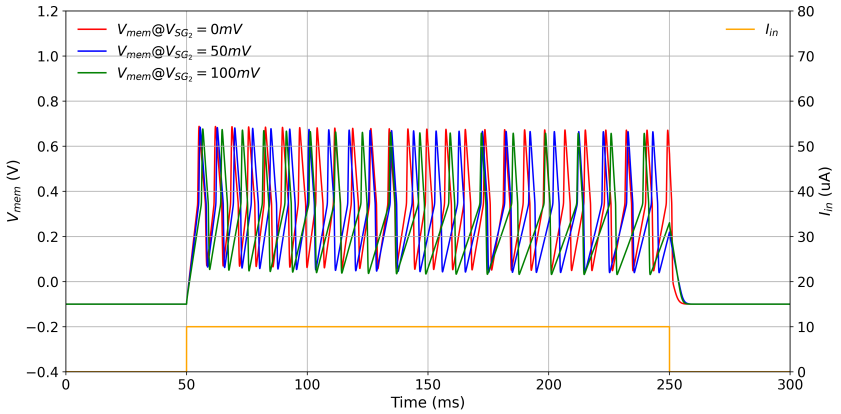
3.6 Circuit simulation

After modelling the adapting and tunable OECT, we proceeded with testing the hybrid spiking architecture that we described in Sec. 3.2 and that is reported in Fig. 3.1. In this first hybrid version of the circuit, all transistors except the double-gate OECT are MOSFET models of TSMC 0.18um node. These were properly-sized to obtain the desired spiking features. In particular, the W/L ratio of the reset transistor M_r was chosen small in order to have sufficiently wide spikes: by limiting the reset current of the Axon-Hillock, the reset of the membrane voltage takes longer and the spike duration t_{hi} increases, accordingly with Eq. 3.3. For the same purpose, C_{fb} and C_{mem} were set equal to 50nF. The tradeoff between increasing and decreasing the spike duration is related to the considerations made in Sec. 3.1, about the optimization of the adaptive OECT and matching it with its spiking gate stimulus. Shorter gate spikes are closer to resemble Dirac deltas if compared to the OECT ionic time constant, so the OECT reaches a steady-state response, i.e. saturation of its adaptation, in a larger number of spikes. However, since the amplitude of the gate spikes is fixed, the effective increase of drain current for each spike is very much decreased. This is equivalent to say that the spike-triggered adaptation factor b is much decreased. The right design sets conditions that are in between these two extremes, providing adaptation linearity for a sufficient number of spikes and a non-negligible spike-triggered adaptation parameter. The voltage supply V_{dd} is set at 0.7V while the two weak inversion current mirrors are biased with a V_{rest} of -0.1V and a V_{up} of 0.8V, all with respect to the ground.

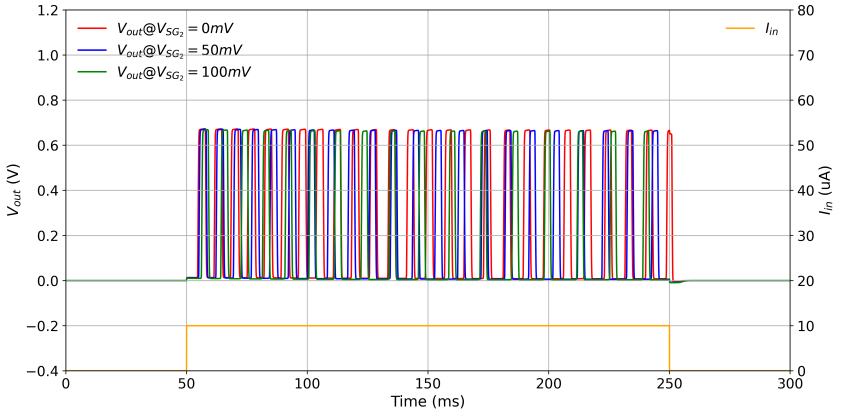
In this section, we want to study the behavior of our adaptive circuit with the same analysis tool that we introduced in Subsec. 1.2.1, that are commonly used in neuroscience.

The membrane voltages V_{mem}, output spiking voltages V_{out}, and adaptation currents I_{adapt} of the circuit simulated with an input current step of 10uA are shown in Fig. 3.12. The graphs show the response of the circuit for three different values of voltage V_{SG_2} applied to the second gate of the OECT.

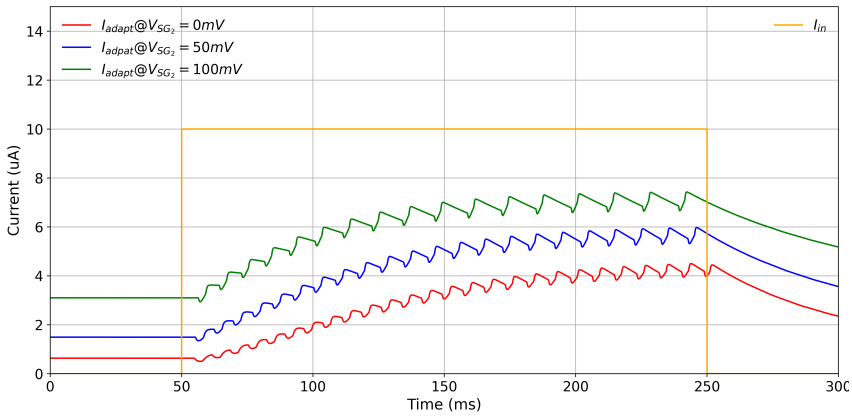
The effect of V_{SG_2} on the double-gate OECT is to shift its threshold voltage. As we can see in Fig. 3.12c, this shift produces a change in its I_{OFF} current, which is the adaptation current when the neuron is silent, i.e. before 50ms. This I_{OFF} current is subtracted at any time to the external input I_{in} of the circuit. The effective adaptation contribution is therefore the difference between this constant minimum and the signal I_{adapt}.



(a) Time evolution of the membrane voltage of the hybrid neuron in response to a 10uA constant current step.



(b) Time evolution of the output spiking voltage of the hybrid neuron in response to a 10uA constant current step.



(c) Time evolution of the adaptation current (the drain current of the double-gate adaptive OECT) of the hybrid neuron in response to a 10uA constant current step.

Figure 3.12: Simulation of the adaptive hybrid spiking circuit at three different values of voltage V_{SG_2} applied to the double-gate OECT.

A better understanding of the spiking activity of the circuit can be deduced from Fig. 3.13. This diagram shows the value of the instantaneous spiking frequency as a function of time. The definition of the instantaneous spiking frequency is the one given in Eq. 1.1. This definition is valid as long as the stimulus is supplied. When the stimulus ends and does not produce any more spikes, then the instantaneous frequency drops to 0. To this definition we added another frequency measurement to evaluate the instantaneous spiking frequency at the time of the first output spike[7]. This additional frequency is calculated as the reciprocal of the time interval between the first spiking time and the onset of the stimulus. From now on, we will define this instantaneous spiking frequency value as the *Onset spiking frequency* of the circuit, that will be labelled with the notation f_0. The *Steady-state spiking frequency* is still defined as the asymptotic frequency value the neuron tends to if the input is kept constant, and is labelled with the notation f_∞. As a general rule, we measure the spiking time when the membrane voltage reaches is maximum value due to the capacitive feedback.

From the diagram it is evident how the onset spiking frequencies between the three cases of applied V_{SG_2} are very different. This discrepancy is caused by the I_{OFF} current of the OECT. Intuitively, the effective current amplitude that charges the membrane capacitor of the circuit, before the first spike, is the difference between I_{in} = 10uA and the changing value of I_{OFF}. The case V_{SG_2}=100mV is the one with the higher OECT's off current, therefore the one that spikes later. The spiking activity of the three cases decays with time towards their steady-state frequency f_∞.

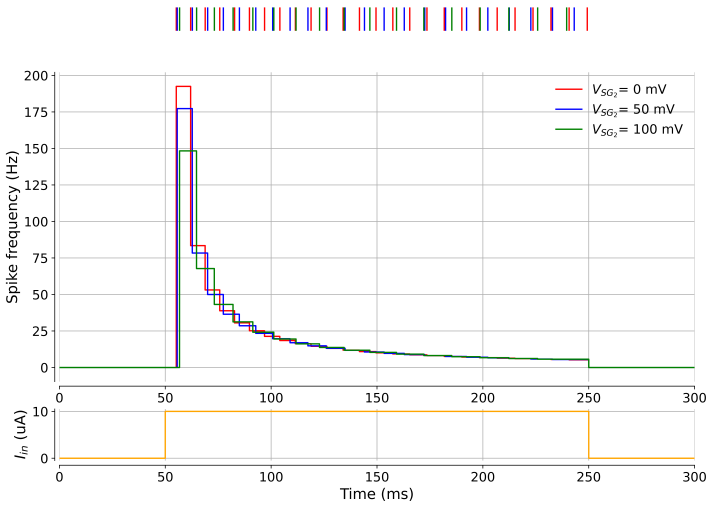


Figure 3.13: Instantaneous spiking frequency of the hybrid neuron as a function of time, relative to the spiking plots shown previously.

[7]The definition given in Eq. 1.1 defines the spiking frequency as the reciprocal of an interspike interval (ISI), therefore it can be calculated only when at least two spikes have been generated

A very broad understanding of the circuit's behavior is given by analyzing its frequency-current curves, that report the values of f_0 and f_∞ as a function of the input constant current's amplitude. These are shown in Fig. 3.14 for the three cases of V_{SG_2}. The first aspect that we note is that they all present a current threshold. The current threshold is the minimum current that has to be injected into the neuron to produce at least one spike. In our architecture, the current threshold is set by the I_{OFF} of the OECT, wihch is the input current that has to be injected just to break even. Therefore its value changes with V_{SG_2}. The double-gate OECT has indeed a double effect on the neuron's response: it shifts its threshold current and modulates the extent of the spiking adaptation.

The impact of spike-frequency adaptation depending on the applied V_{SG_2} can be be deduced by visual inspection from Fig. 3.14, by seeing that for high input currents the onset frequencies increase with the same slope, while the steady-state frequencies saturate at different pace. However, the change in the spiking threshold that occurs together with the adaptation modulation causes limited quantitative analysis.

To study individually the tuning of spike-frequency adaptation by the OECT's second gate voltage, we can plot a different type of f(I) curves, which are the curves $f(I-I_{th}$, where I_{th} is the threshold current of the neuron. This description considers the effective reduction of the input current due to adaptation, eliminating the initial dependence on I_{OFF}. These are shown in Fig. 3.15.

The onset frequency curves $f_0(I - I_{th})$ are now matched: the effect of I_{OFF} was eliminated so effectively the three neurons are now integrating the same input current on the same membrane capacitor. The $f_\infty(I - I_{th})$ curves do not match instead and indicate a different SFA impact. A higher source-gate2 voltage applied to the adaptive OECT induces higher adaptation to the circuit. This can be explained by thinking at the threshold modulation given by Gate2. If V_{SG_2} increases, V_T decreases and the OECT is more conductive for the same V_{SG_1}, V_{SD} applied. A more conductive OECT also reaches higher currents during its plasticity and these will be subtracted to the neuron's input. By plotting as a function of $I - I_{th}$ we get rid of the different I_{OFF}s for the double-gate tuned at different V_{SG_2}, but we do not get rid of the difference between this minimum and the maximum plastic current that can be reached by the OECT, for the same pulsing conditions. The result is a higher effective adaptive current contribution subtracted from the input.

With this study we demonstrated the operation of the adaptive architecture and the adaptation tuning via the double-gate OECT. We also addressed the change in threshold current of the neuron. Our goal is now to test the same architecture in a circuit made entirely with organic electrochemical transistors.

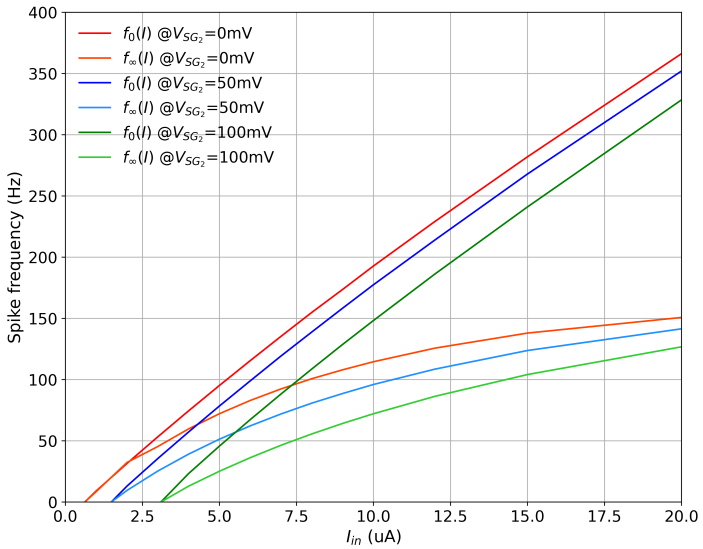


Figure 3.14: Frequency-current curves of the simulated hybrid neuron. Tuning of V_{SG_2} results in a shift of the threshold current and a variation of the adaptation impact.

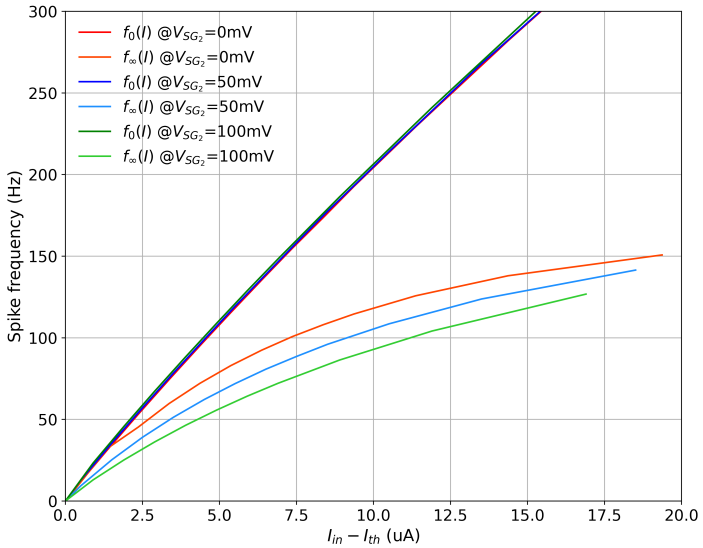


Figure 3.15: $f(I - I_{th})$ curves of the simulated hybrid neuron. This representation allows to estimate quantitatively the modulation of the spike-frequency adaptation by V_{SG_2} because it cancels the effect of V_{SG_2} on I_{OFF}.

Chapter 4

All-organic adaptive neuron

In the previous Chapter we validated the design of the adaptive spiking circuit in a hybrid version that combines OECT and CMOS technologies. In this last section we want to confirm that the same architecture can be used to build an artificial neuron that is entirely made with organic based transistors.

The use of organic transistors introduces some non-idealities and limitations compared to MOSFETs. For example, the mirroring operation is distorted by the capacitive spikes of the mirroring OECT. Moreover, organic transistors are orders of magnitude slower than silicon. To reduce the performance gap we are going to use State of the Art complementary OECTs.

As for the hybrid version, the validation of the all-organic neuron is performed by creating Spice models for the additional devices and simulating the complete circuit.

4.1 Complementary OECTs modeling

The initial purpose of the study is to investigate the optimal conditions for the operation of the organic circuit and choose the right materials that can provide those conditions.

In Sec. 3.1 we discussed how the circuit spikes must have duration $\Delta t_{spk} << \tau_i$ to produce an adaptation current that resembles the one by Treves. This condition ensures that the double-gate OECT does not reach a steady-state response for a sufficient number of spikes applied to the gate.

In Subsec. 3.2.1 we saw that the spike duration of the Axon-Hillock is given by Eq. 3.3, that we report again here:

$$t_{hi} = \frac{C_{fb} V_{dd}}{I_{reset} - I_{in}}$$

This formula assumes that the delay given by the inverting stages is negligible. This condition is valid for a silicon circuit, where the membrane and feedback capacitances are much larger than any transistor's capacitance. However, it does not hold

for circuits made with OECTs, since they exploit bulk doping and their capacitance is volumetric. Moreover, OECTs have mobilities that are much smaller than MOSFETs. Especially OECTs made with n-type semiconducting polymers struggle to reach switching times smaller than $5ms$.

State of the art for n-type organic polymers is high-molecular weight BBL, developed recently by Fabiano et al. [79]. This polymer can reach a switching time constant of $0.38ms$ and a ratio $I_{ON}/I_{OFF} > 10^5$ for an operating voltage of just 0.7V. In the same article, Fabiano and his group proposed a pair of balanced complementary OECTs, made with high molecular weight BBL_{152} and p(g2T-TT) (which is the same polymer we used for the double-gate OECT), that are among the best sub-1 V complementary inverters reported to date. To validate the operation of the adaptive organic neuron, we modelled on Spice these state of the art OECTs using the Friedlein model.

The output and transfer characteristics and the transient response of the two OECTs are reported in the supplementary material of [79]. The behavior of our Spice models is shown in Fig. 4.2 (the schematic symbols for the two OECTs are shown in Fig. 4.1).

An important aspect of the modeling is the matching of the capacitive spikes in the total drain current between the model and the measured data. These capacitive spikes are in fact the main source of non-idealities of the OECT and introduce rapidly changing currents that may corrupt the correct evolution of the voltage spikes or the double-gate adaptation current in the final circuit. To truly test the feasibility of our organic neuron and the architecture resilience to these noise factors it is mandatory to include accurate modeling of the ionic displacement currents.

Figure 4.1: Schematic symbols of the high molecular weight BBL n-type OECT (left) and p(g2T-TT) p-type OECT (right).

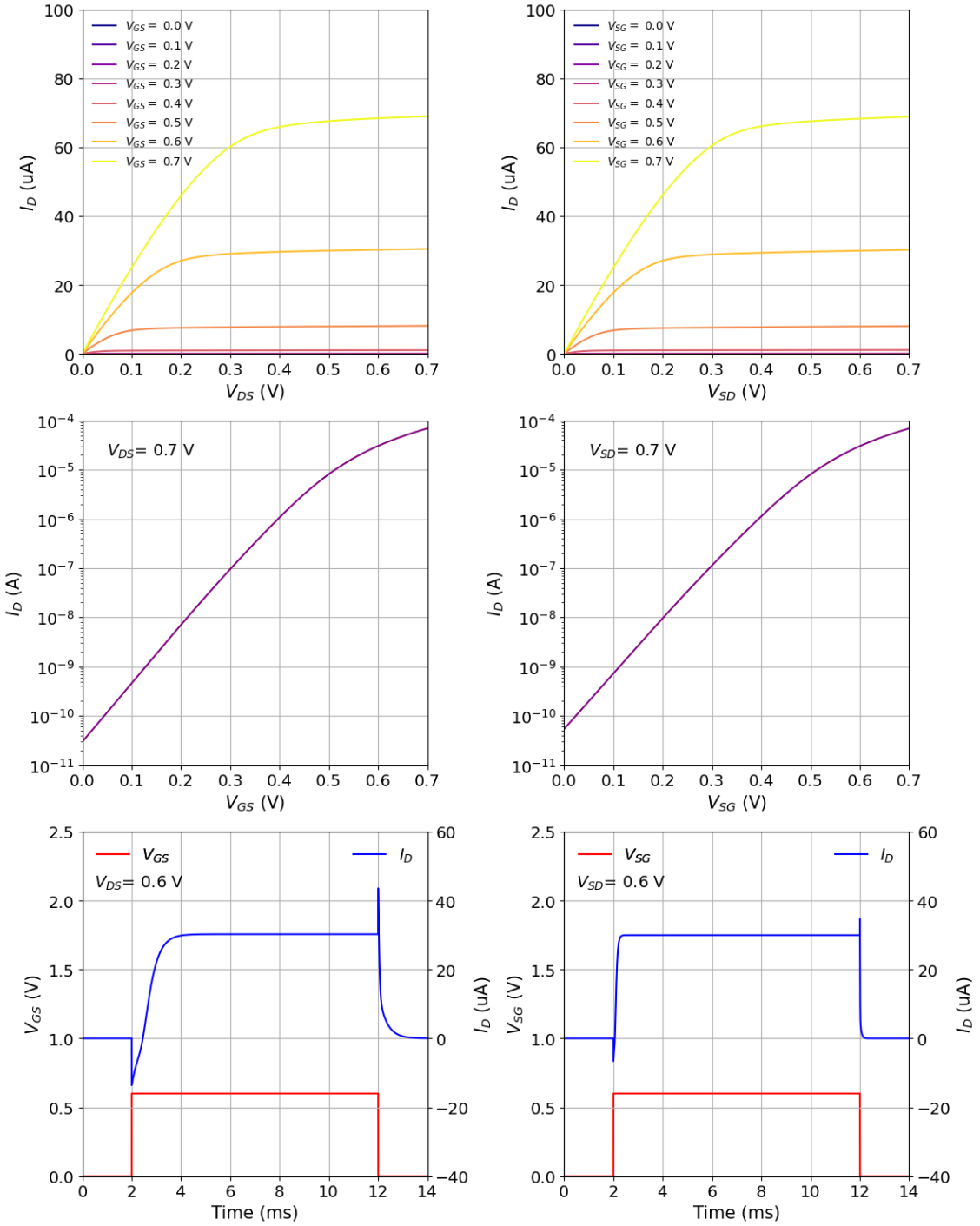


Figure 4.2: Spice modelling of the state of the art complementary OECTs reported in [79]. Output and transfer characteristics and transient response of the simulated OECTs: n-type (left) and p-type (right).

4.2 Circuit simulation

Now that we have a foundation for the optimal design of the organic circuit, and equivalent spice models that describe the devices employed, we want to test the performance of the organic adaptive neuron and compare it to the hybrid version. The complete all-organic neuron circuit is shown in Fig. 4.3.

The biasing conditions are equivalent to the ones of the simulated hybrid circuit: V_{dd}=0.7V, V_{rest}=-0.1V and V_{up}=0.8V. The main difference to the previous circuit is that the two capacitors are now $C_{mem} = (C_{fb} = 10nF$. The reason for this design choice is still related to the spike duration of the circuit and the optimal value to match to the double-gate OECT. Since the organic inverters and the reset transistor introduce longer delays compared to their CMOS equivalents, the spike duration is reduced by acting on the capacitive feedback.

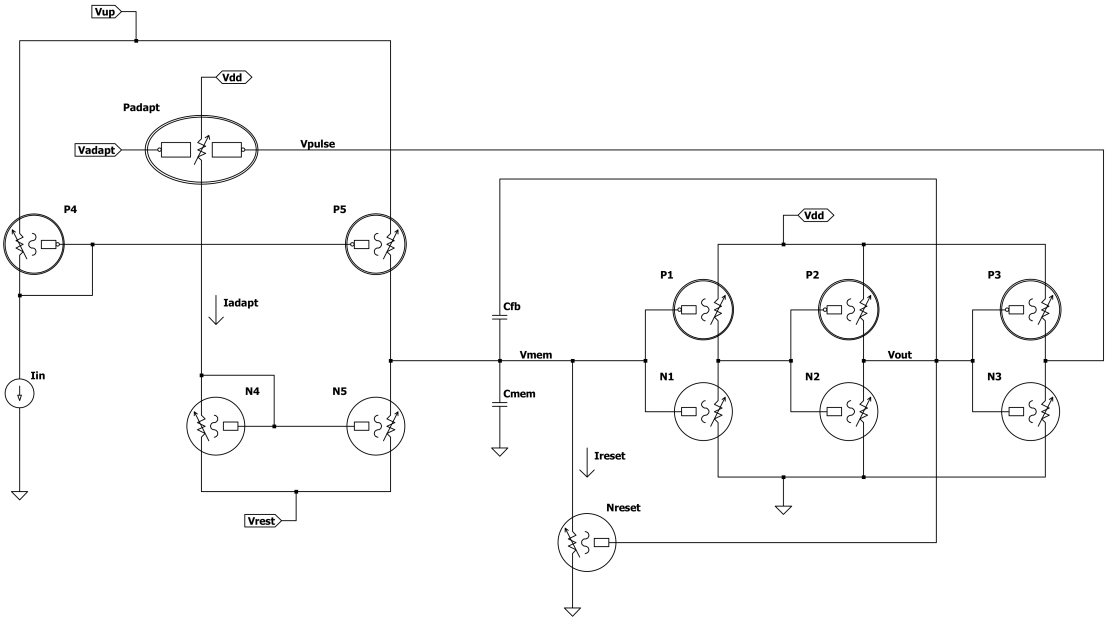
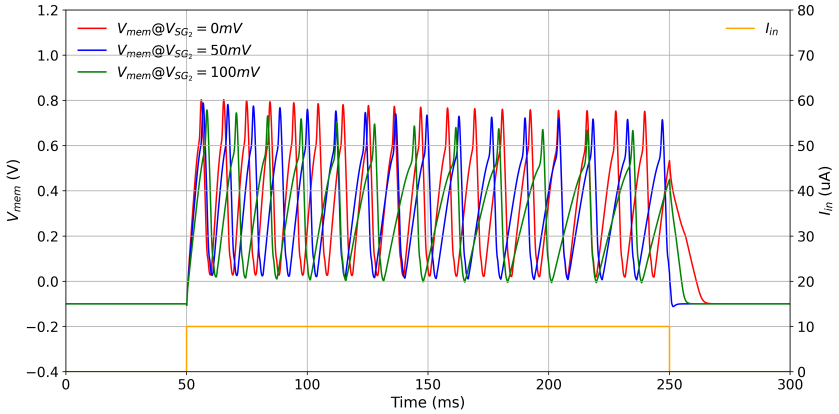


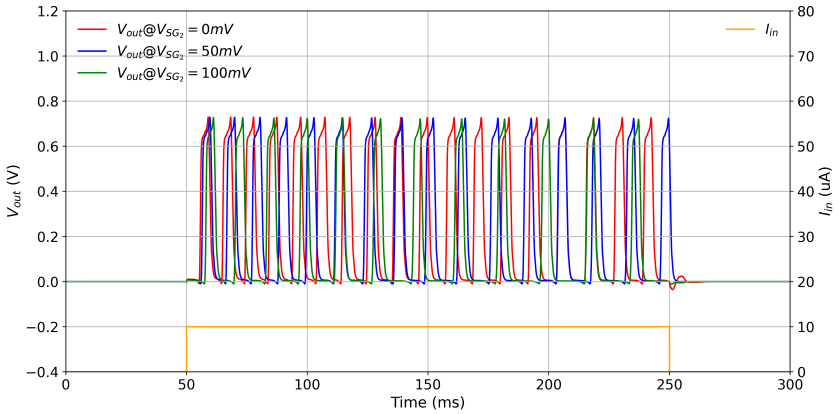
Figure 4.3: Schematic diagram of the all-organic adaptive artificial neuron. The analog current mirroring and the spike generation are now performed entirely by organic electronics.

In Fig. 4.4 we report the response of the membrane and spiking output nodes of the organic neuron to the same test cases analyzed for the hybrid version: $I_{in} = 10uA$ and the same values of the voltage V_{SG_2}. Some considerations can be done with respect to the hybrid circuit. The first one is that the membrane signal shows a non-negligible decay in its peak values, especially for the case $V_{SG_2} = 100mV$. This is in fact not a decay at values lower than 0.7V, which is V_{dd}, i.e. the voltage towards which the membrane voltage should tend, but from values higher than 0.7 towards

lower values that are confined in the range between the voltage supply and ground. This effect is due to the capacitive ionic displacement currents inside the OECTs and the spike-and-rebound behavior that they cause. This is a very important effect that our simulation is able to reproduce thanks to the capacitive spike modeling of the previous section. The second note is that the output spikes are longer in this case, as we expected. The output signal still shows a proper spike shape, that is also affected by the capacitive spike currents but that remains unchanged all throughout the simulation.



(a) Time evolution of the membrane voltage of the organic neuron in response to a 10uA constant current step.



(b) Time evolution of the output spiking voltage of the organic neuron in response to a 10uA constant current step.

Figure 4.4: Simulation of the adaptive all-organic spiking circuit at three different values of voltage V_{SG_2} applied to the double-gate OECT. All simulation's parameters are kept equal to the ones reported for the hybrid circuit.

The evolution of the instantaneous spiking frequency for the membrane and output spike signals is reported in Fig. 4.5. Due to the longer delays of the inverting stages

and of activation of the reset transistor, the onset spiking frequency is lower in all three cases (relatively to the hybrid simulation). The neuron still presents an exponentially decaying adaptation of its spiking frequency. The effect of V_{SG_2} is consistent with what we saw in the hybrid circuit.

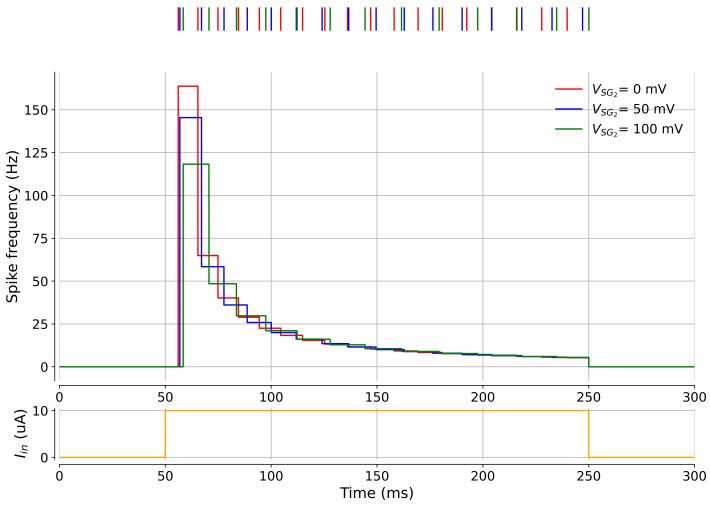


Figure 4.5: Instantaneous spiking frequency of the organic neuron as a function of time, relative to the spiking plots shown previously.

As for our previous SFA analysis, we proceeded by plotting the frequency-current curves of the organic neuron, which summarize the adapting behavior of the neuron and its properti tuning. These are plotted in Fig. 4.6. The behavior they show is similar to the one of the hybrid. One difference is that the onset curves have a deceleration that is much more evident than in the hybrid case. The slope they have at low currents reduces for increasing current values up to a fixed slope. To eliminate the shift in the threshold current we once again plot the $f(I - I_{th})$ curves, in Fig. 4.6. By plotting the $f(I - I_{th})$ curves we are able to see other differences with respect to the hybrid circuit. The onset curves do not increase at the same pace between the three cases of V_{SG_2}, while the steady-state ones still show different rates of saturation. This is a different adaptation behavior compared to the hybrid neuron. In the hybrid circuit, an increase in voltage V_{SG_2} causes an increase of the threshold current and lower steady-state spiking frequencies. The onset frequency remains unchanged. Therefore the voltage V_{SG_2} can be used to tune the two coupled quantities independently of the onset frequency. In the organic neuron, the onset frequencies also change with V_{SG_2} and they resemble the same dependence of the steady-state ones. In this second case, an increase in voltage V_{SG_2} causes an increase of the threshold current and a simultaneous decrease of the f(I) curves. The result is a neuron which is overall less responsive: it has a higher threshold current, a lower onset spiking frequency and a lower steady-state one. These considerations are useful when it comes to emulation of precise adaptive mechanisms and tuning

of thereof.

One final note is that for very low values of input currents, the general behavior of the circuit resembles once again the one of the hybrid: the voltage V_{SG_2} affects only the quantities I_{th} and f_∞, as it is shown in Fig. 4.8. This property may be taken into consideration when defining the correct operation range for the organic circuit, based on the wanted properties.

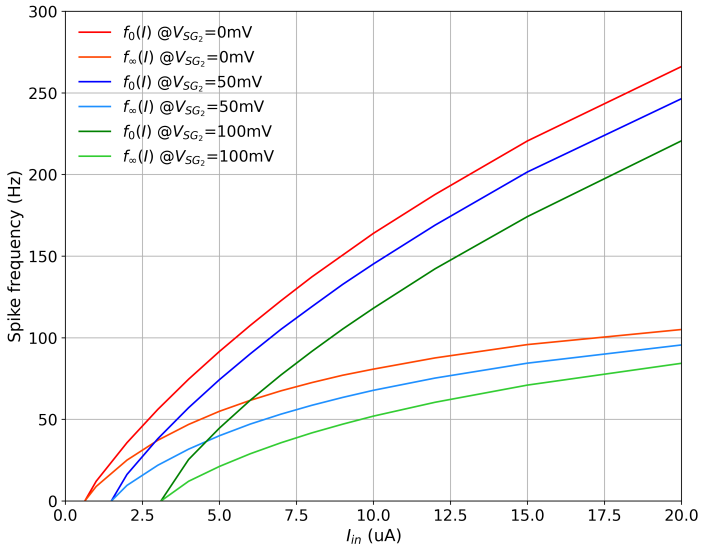


Figure 4.6: Frequency-current curves of the simulated organic neuron. Tuning of V_{SG_2} still results in a shift of the threshold current and a variation of the adaptation impact.

With this final study we characterized the operation of the adaptive spiking architecture on a simulated all-organic circuit. The circuit still shows adaptation, that is tunable, together with its threshold current. The additional $f(I - I_{th})$ analysis that we proposed allowed us to unbound the modulation of the adaptation from the shift in threshold current in the frequency-curves. With this method we were able to see secondary spiking effects that differ between the all-organic and the hybrid circuit, that may go unnoticed otherwise. The results we obtained promise successful implementation of an organic adaptive spiking circuit with the architecture and the design choices that we presented.

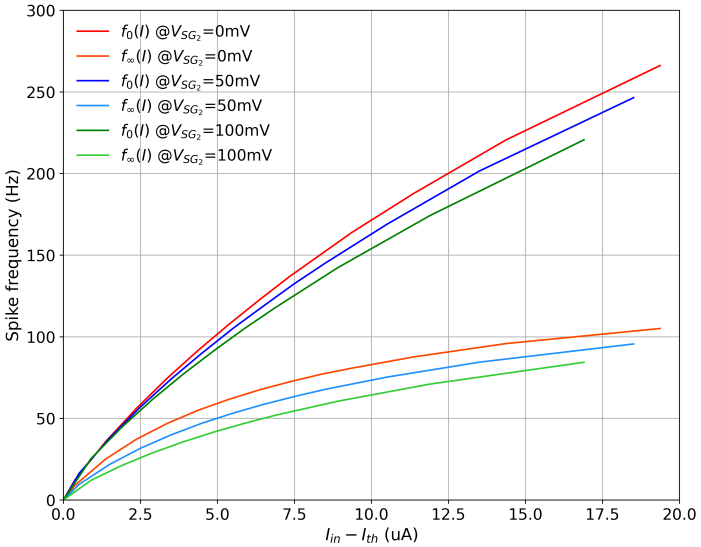


Figure 4.7: $f(I - I_{th})$ curves of the simulated organic neuron.

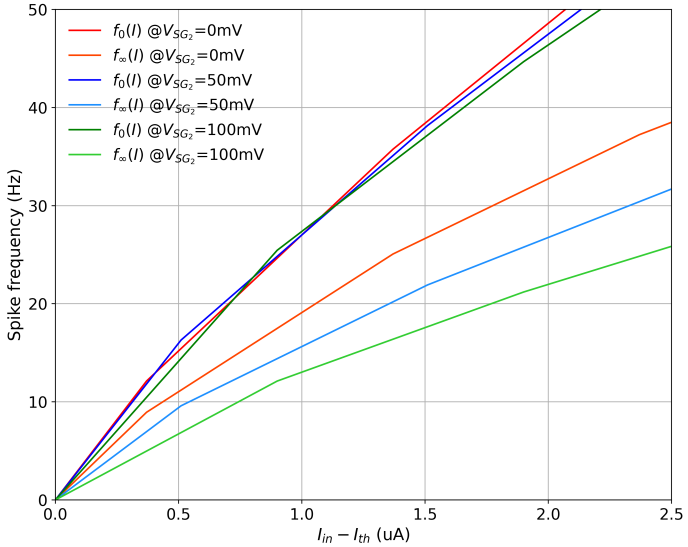


Figure 4.8: Zoom on the $f(I - I_{th})$ curves of the simulated organic neuron at low values of input current.

Chapter 5

Conclusion

5.1 Future perspectives

The spiking architecture we developed allows to emulate spike-frequency adaptation with an entirely organic matrix. However different optimizations can still be made.

Even though the circuit operates at a voltage supply ¡ 1V, power consumption is still much higher than silicon counterparts. Especially in neuromorphic applications, this is one of the most important evaluation parameter. The discussion in this thesis did not involve power consumption measurements. However, a few notes can be made in that regard.

First of all, the complete circuit architecture involves the use of an Axon-Hillock circuit. Although it has some great matching properties that are very useful for organic electronics, it consumes a fair amount of power due to its slowly changing membrane potential. The optimization of power consumption could start by integrating the adaptation block with a different more efficient spike generation block.

The integration with other neuromorphic blocks could also add tunability to the adaptation block itself. The adaptation block we proposed is very compact and versatile. It is made of just three transistors (if we do not consider the p-type mirror that is just used to mirror the input current to the neuron) and requires only an inverted spiking node. However it does not allow to tune the steady-state spiking frequency and the threshold current of the neuron independently, for example. The choice of proper integrated blocks could add degrees of freedom to the complete circuit to uncouple these two quantities.

The second point is that in order to have a functioning circuit and a correct adaptation signal, the double-gate OECT were to be designed to conduct currents in the range of 1s-100s of uA. This choice was necessary to have a channel current adaptive signal with higher amplitude than the parasitic capacitive spikes. Decreasing the ionic displacement currents is a difficult challenge in OECTs. These are minimized, as we saw, mainly by having a larger electrolytic resistance. To maintain an ionic time constant in the same order of magnitude, the OECT's channel capacitance has to be decreased, thereby forcing to make OECTs with a smaller design or with a

thinner polymeric layer.

Regarding future implementations, the functionality of the adaptation block as an ionic variation encoder may be studied. As we saw, by exploiting a limit resistor in series to the gate we were able to reduce the dependence of the OECT on the electrolyte concentration at just its steady-state response. A change in the electrolytic concentration would therefore have a similar impact on the adaptation current as the second-gate of the OECT. The second-gate could be used to tune the response of the adaptive neuron to have maximum response variation on a specific range of ionic concentration. Possibly, the double-gate OECT could be functionalized in order to be sensitive to specific bio-molecules.

Bibliography

[1] Carver Mead. How we created neuromorphic engineering. *Nature Electronics*, 3(7):434–435, July 2020.

[2] Simon B. Laughlin, Rob R. de Ruyter van Steveninck, and John C. Anderson. The metabolic cost of neural information. *Nature Neuroscience*, 1(1):36–41, May 1998.

[3] C. Mead. Neuromorphic electronic systems. *Proceedings of the IEEE*, 78(10):1629–1636, 1990.

[4] Naveen Verma, Hongyang Jia, Hossein Valavi, Yinqi Tang, Murat Ozatay, Lung-Yen Chen, Bonan Zhang, and Peter Deaville. In-memory computing: Advances and prospects. *IEEE Solid-State Circuits Magazine*, 11(3):43–55, 2019.

[5] Abu Sebastian, Manuel Le Gallo, Riduan Khaddam-Aljameh, and Evangelos Eleftheriou. Memory devices and applications for in-memory computing. *Nature Nanotechnology*, 15(7):529–544, March 2020.

[6] Carver Mead. Analog vlsi and neural systems. 1989.

[7] R Douglas, M Mahowald, and C Mead. Neuromorphic analogue VLSI. *Annual Review of Neuroscience*, 18(1):255–281, March 1995.

[8] Dragi Kimovski, Nishant Saurabh, Matthijs Jansen, Atakan Aral, Auday Al-Dulaimy, André Bondi, Antonino Galletta, Alessandro Papadopoulos, Alexandru Iosup, and Radu Prodan. Beyond von neumann in the computing continuum: Architectures, applications, and future directions. *IEEE Internet Computing*, PP:1–11, 08 2023.

[9] Giacomo Indiveri. Introducing 'neuromorphic computing and engineering'. *Neuromorphic Computing and Engineering*, 1(1):010401, jul 2021.

[10] Dennis V Christensen, Regina Dittmann, Bernabe Linares-Barranco, Abu Sebastian, Manuel Le Gallo, Andrea Redaelli, Stefan Slesazeck, Thomas Mikolajick, Sabina Spiga, Stephan Menzel, Ilia Valov, Gianluca Milano, Carlo Ricciardi, Shi-Jun Liang, Feng Miao, Mario Lanza, Tyler J Quill, Scott T Keene, Alberto Salleo, Julie Grollier, Danijela Marković, Alice Mizrahi, Peng Yao,

J Joshua Yang, Giacomo Indiveri, John Paul Strachan, Suman Datta, Elisa Vianello, Alexandre Valentian, Johannes Feldmann, Xuan Li, Wolfram H P Pernice, Harish Bhaskaran, Steve Furber, Emre Neftci, Franz Scherr, Wolfgang Maass, Srikanth Ramaswamy, Jonathan Tapson, Priyadarshini Panda, Youngeun Kim, Gouhei Tanaka, Simon Thorpe, Chiara Bartolozzi, Thomas A Cleland, Christoph Posch, ShihChii Liu, Gabriella Panuccio, Mufti Mahmud, Arnab Neelim Mazumder, Morteza Hosseini, Tinoosh Mohsenin, Elisa Donati, Silvia Tolu, Roberto Galeazzi, Martin Ejsing Christensen, Sune Holm, Daniele Ielmini, and N Pryds. 2022 roadmap on neuromorphic computing and engineering. *Neuromorphic Computing and Engineering*, 2(2):022501, may 2022.

[11] Giacomo Indiveri, Bernabé Linares-Barranco, Tara Julia Hamilton, André van Schaik, Ralph Etienne-Cummings, Tobi Delbruck, Shih-Chii Liu, Piotr Dudek, Philipp Häfliger, Sylvie Renaud, Johannes Schemmel, Gert Cauwenberghs, John Arthur, Kai Hynna, Fopefolu Folowosele, Sylvain Saighi, Teresa Serrano-Gotarredona, Jayawan Wijekoon, Yingxue Wang, and Kwabena Boahen. Neuromorphic silicon neuron circuits. *Frontiers in Neuroscience*, 5, 2011.

[12] A. L. Hodgkin, A. F. Huxley, and B. Katz. Measurement of current-voltage relations in the membrane of the giant axon of loligo. *The Journal of Physiology*, 116(4):424–448, 1952.

[13] A. L. Hodgkin and A. F. Huxley. Currents carried by sodium and potassium ions through the membrane of the giant axon of loligo. *The Journal of Physiology*, 116(4):449–472, 1952.

[14] A. L. Hodgkin and A. F. Huxley. The components of membrane conductance in the giant axon of loligo. *The Journal of Physiology*, 116(4):473–496, 1952.

[15] A. L. Hodgkin and A. F. Huxley. The dual effect of membrane potential on sodium conductance in the giant axon of loligo. *The Journal of Physiology*, 116(4):497–506, 1952.

[16] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, 1952.

[17] Jan Benda, André Longtin, and Len Maler. Spike-frequency adaptation separates transient communication signals from background oscillations. *The Journal of Neuroscience*, 25(9):2312–2321, March 2005.

[18] Jan Benda and R. Matthias Hennig. Spike-frequency adaptation generates intensity invariance in a primary auditory interneuron. *Journal of Computational Neuroscience*, 24(2):113–136, May 2007.

[19] Hayeong Lee, Lubomir Kostal, Ryohei Kanzaki, and Ryota Kobayashi. Spike frequency adaptation facilitates the encoding of input gradient in insect olfactory projection neurons. *Biosystems*, 223:104802, 2023.

[20] Darjan Salaj, Anand Subramoney, Ceca Kraisnikovic, Guillaume Bellec, Robert Legenstein, and Wolfgang Maass. Spike frequency adaptation supports network computations on temporally dispersed information. *eLife*, 10:e65459, jul 2021.

[21] Jan Benda and Andreas V. M. Herz. A Universal Model for Spike-Frequency Adaptation. *Neural Computation*, 15(11):2523–2564, 11 2003.

[22] Bard Ermentrout. Linearization of F-I Curves by Adaptation. *Neural Computation*, 10(7):1721–1729, 10 1998.

[23] Jan Benda and Joel Tabak. *Spike-Frequency Adaptation*, pages 1–12. Springer New York, New York, NY, 2013.

[24] T. Tateno, A. Harsch, and H. P. C. Robinson. Threshold firing frequency–current relationships of neurons in rat somatosensory cortex: Type 1 and type 2 dynamics. *Journal of Neurophysiology*, 92(4):2283–2294, October 2004.

[25] Fabrizio Gabbiani and Holger G. Krapp. Spike-frequency adaptation and intrinsic properties of an identified, looming-sensitive neuron. *Journal of Neurophysiology*, 96(6):2951–2962, December 2006.

[26] Romain Brette and Wulfram Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94(5):3637–3642, November 2005.

[27] Richard Naud, Nicolas Marcille, Claudia Clopath, and Wulfram Gerstner. Firing patterns in the adaptive exponential integrate-and-fire model. *Biological Cybernetics*, 99(4–5):335–347, November 2008.

[28] Alessandro Treves. Mean-field analysis of neuronal spike dynamics. *Network: Computation in Neural Systems*, 4(3):259–284, January 1993.

[29] Wulfram Gerstner, J. Leo van Hemmen, and Jack D. Cowan. What Matters in Neuronal Locking? *Neural Computation*, 8(8):1653–1676, 11 1996.

[30] Renaud Jolivet, Alexander Rauch, Hans-Rudolf Lüscher, and Wulfram Gerstner. Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience*, 21(1):35–49, April 2006.

[31] Christian Pozzorini, Richard Naud, Skander Mensi, and Wulfram Gerstner.

Temporal whitening by power-law adaptation in neocortical neurons. *Nature Neuroscience*, 16(7):942–948, June 2013.

[32] Shih-Chii Liu, Jörg Kramer, Giacomo Indiveri, Tobias Delbrück, and Rodney Douglas. *Analog VLSI: Circuits and Principles*. The MIT Press, 09 2002.

[33] Shih-Chii Liu, Tobi Delbruck, Giacomo Indiveri, Adrian Whatley, and Rodney Douglas. *Event-Based Neuromorphic Systems*. Wiley, December 2014.

[34] R. Edwards and Gert Cauwenberghs. Synthesis of log-domain filters from first-order building blocks. *Analog Integrated Circuits and Signal Processing*, 22:177–186, 03 2000.

[35] E.M. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6):1569–1572, November 2003.

[36] Venkat Rangan, Abhishek Ghosh, Vladimir Aparin, and Gert Cauwenberghs. A subthreshold avlsi implementation of the izhikevich simple neuron model. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, August 2010.

[37] Ştefan Mihalaş and Ernst Niebur. A generalized linear integrate-and-fire neural model produces diverse spiking behaviors. *Neural Computation*, 21(3):704–718, March 2009.

[38] Andre van Schaik, Craig Jin, Alistair McEwan, Tara Julia Hamilton, Stefan Mihalas, and Ernst Niebur. A log-domain implementation of the mihalas-niebur neuron model. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. IEEE, May 2010.

[39] Irene Muñoz-Martin, Stefano Bianchi, Shahin Hashemkhani, Giacomo Pedretti, Octavian Melnic, and Daniele Ielmini. A brain-inspired homeostatic neuron based on phase-change memories for efficient neuromorphic computing. *Frontiers in Neuroscience*, 15, August 2021.

[40] Xinxin Wang, Peng Huang, Zhen Dong, Zheng Zhou, Yuning Jiang, Runze Han, Lifeng Liu, Xiaoyan Liu, and Jinfeng Kang. A novel rram-based adaptive-threshold lif neuron circuit for high recognition accuracy. In *2018 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*. IEEE, April 2018.

[41] Ahmed Shaban, Sai Sukruth Bezugam, and Manan Suri. An adaptive threshold neuron for recurrent spiking neural networks with nanodevice hardware implementation. *Nature Communications*, 12(1), July 2021.

[42] Wei Yi, Kenneth K. Tsang, Stephen K. Lam, Xiwei Bai, Jack A. Crowell, and Elias A. Flores. Biological plausibility and stochasticity in scalable vo2 active memristor neurons. *Nature Communications*, 9(1), November 2018.

[43] Margarita R. Chetyrkina, Filipp S. Talalaev, Larisa V. Kameneva, Svetlana V. Kostyuk, and Pavel A. Troshin. Vat dyes: promising biocompatible organic semiconductors for wearable electronics applications. *Journal of Materials Chemistry C*, 10(8):3224–3231, 2022.

[44] Claudia Cea, Zifang Zhao, Duncan J. Wisniewski, George D. Spyropoulos, Anastasios Polyravas, Jennifer N. Gelinas, and Dion Khodagholy. Integrated internal ion-gated organic electrochemical transistors for stand-alone conformable bioelectronics. *Nature Materials*, 22(10):1227–1235, July 2023.

[45] Padinhare Cholakkal Harikesh, Chi-Yuan Yang, Deyu Tu, Jennifer Y. Gerasimov, Abdul Manan Dar, Adam Armada-Moreira, Matteo Massetti, Renee Kroon, David Bliman, Roger Olsson, Eleni Stavrinidou, Magnus Berggren, and Simone Fabiano. Organic electrochemical neurons and synapses with ion mediated spiking. *Nature Communications*, 13(1), February 2022.

[46] Tanmoy Sarkar, Katharina Lieberth, Aristea Pavlou, Thomas Frank, Volker Mailaender, Iain McCulloch, Paul W. M. Blom, Fabrizio Torricelli, and Paschalis Gkoupidenis. An organic artificial spiking neuron for in situ neuromorphic sensing and biointerfacing. *Nature Electronics*, 5(11):774–783, November 2022.

[47] Padinhare Cholakkal Harikesh, Chi-Yuan Yang, Han-Yan Wu, Silan Zhang, Mary J. Donahue, April S. Caravaca, Jun-Da Huang, Peder S. Olofsson, Magnus Berggren, Deyu Tu, and Simone Fabiano. Ion-tunable antiambipolarity in mixed ion–electron conducting polymers enables biorealistic organic electrochemical neurons. *Nature Materials*, 22(2):242–248, January 2023.

[48] Giovanni Maria Matrone, Eveline van Doremaele, Sophie Griggs, Gang Ye, Iain McCulloch, Francesca Santoro, and Yoeri van de Burgt. Organic neuromorphic spiking circuit for retina-inspired sensory coding and neurotransmitter-mediated neural pathways. December 2022.

[49] Henry S. White, Gregg P. Kittlesen, and Mark S. Wrighton. Chemical derivatization of an array of three gold microelectrodes with polypyrrole: fabrication of a molecule-based transistor. *Journal of the American Chemical Society*, 106(18):5375–5377, 1984.

[50] Paschalis Gkoupidenis, Nathan Schaefer, Benjamin Garlan, and George G. Malliaras. Neuromorphic functions in pedot:pss organic electrochemical transistors. *Advanced Materials*, 27(44):7176–7180, 2015.

[51] Paschalis Gkoupidenis, Nathan Schaefer, Xenofon Strakosas, Jessamyn A. Fairfield, and George G. Malliaras. Synaptic plasticity functions in an organic electrochemical transistor. *Applied Physics Letters*, 107(26), 12 2015. 263302.

[52] Dion Khodagholy, Jonathan Rivnay, Michele Sessolo, Moshe Gurfinkel, Pierre Leleux, Leslie H. Jimison, Eleni Stavrinidou, Thierry Herve, Sébastien Sanaur, Róisín M. Owens, and George G. Malliaras. High transconductance organic electrochemical transistors. *Nature Communications*, 4(1), July 2013.

[53] D.A. Bernards and G.G. Malliaras. Steady-state and transient behavior of organic electrochemical transistors. *Advanced Functional Materials*, 17(17):3538–3544, 2007.

[54] Anton Weissbach, Lukas M. Bongartz, Matteo Cucchi, Hsin Tseng, Karl Leo, and Hans Kleemann. Photopatternable solid electrolyte for integrable organic electrochemical transistors: operation and hysteresis. *Journal of Materials Chemistry C*, 10(7):2656–2662, 2022.

[55] Dimitrios A. Koutsouras, Paschalis Gkoupidenis, Clemens Stolz, Vivek Subramanian, George G. Malliaras, and David C. Martin. Impedance spectroscopy of spin-cast and electrochemically deposited PEDOT:PSS films on microfabricated electrodes with various areas. *ChemElectroChem*, 4(9):2321–2327, June 2017.

[56] Jacob T. Friedlein, Robert R. McLeod, and Jonathan Rivnay. Device physics of organic electrochemical transistors. *Organic Electronics*, 63:398–414, 2018.

[57] Jacob T. Friedlein, Mary J. Donahue, Sean E. Shaheen, George G. Malliaras, and Robert R. McLeod. Microsecond response in organic electrochemical transistors: Exceeding the ionic speed limit. *Advanced Materials*, 28(38):8398–8404, July 2016.

[58] Xudong Ji, Bryan D. Paulsen, Gary K. K. Chik, Ruiheng Wu, Yuyang Yin, Paddy K. L. Chan, and Jonathan Rivnay. Mimicking associative learning using an ion-trapping non-volatile synaptic organic electrochemical transistor. *Nature Communications*, 12(1), April 2021.

[59] Wentao Xu, Sung-Yong Min, Hyunsang Hwang, and Tae-Woo Lee. Organic core-sheath nanowire artificial synapses with femtojoule energy consumption. *Science Advances*, 2(6), June 2016.

[60] Scott T. Keene, Claudia Lubrano, Setareh Kazemzadeh, Armantas Melianas, Yaakov Tuchman, Giuseppina Polino, Paola Scognamiglio, Lucio Cinà, Alberto Salleo, Yoeri van de Burgt, and Francesca Santoro. A biohybrid synapse with neurotransmitter-mediated plasticity. *Nature Materials*, 19(9):969–973, June

2020.

[61] Gregório C. Faria, Duc T. Duong, and Alberto Salleo. On the transient response of organic electrochemical transistors. *Organic Electronics*, 45:215–221, 2017.

[62] Deyu Tu, Loïg Kergoat, Xavier Crispin, Magnus Berggren, and Robert Forchheimer. Transient analysis of electrolyte-gated organic field-effect transistors. In Zhenan Bao and Iain McCulloch, editors, *Organic Field-Effect Transistors XI*, volume 8478, page 84780L. International Society for Optics and Photonics, SPIE, 2012.

[63] Paschalis Gkoupidenis, Dimitrios A. Koutsouras, Thomas Lonjaret, Jessamyn A. Fairfield, and George G. Malliaras. Orientation selectivity in a multi-gated organic electrochemical transistor. *Scientific Reports*, 6(1), June 2016.

[64] Chuan Qian, Ling an Kong, Junliang Yang, Yongli Gao, and Jia Sun. Multi-gate organic neuron transistors for spatiotemporal information processing. *Applied Physics Letters*, 110(8), February 2017.

[65] Hsin Tseng, Anton Weissbach, Juzef Kucinski, Ali Solgi, Rakesh Nair, Lukas M. Bongartz, Giuseppe Ciccone, Matteo Cucchi, Karl Leo, and Hans Kleemann. Threshold voltage control in dual-gate organic electrochemical transistors. *Advanced Materials Interfaces*, page 2201914, October 2022.

[66] G Smith. A firing-rate model of spike-frequency adaptation in sinusoidally-driven thalamocortical relay neurons. *Thalamus amp; Related Systems*, 1(2):135–156, June 2001.

[67] Nachum Ulanovsky, Liora Las, Dina Farkas, and Israel Nelken. Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience*, 24(46):10440–10453, November 2004.

[68] Giancarlo La Camera, Alexander Rauch, David Thurbon, Hans-R. Lüscher, Walter Senn, and Stefano Fusi. Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. *Journal of Neurophysiology*, 96(6):3448–3464, December 2006.

[69] Achilleas Savva, Camila Cendra, Andrea Giugni, Bruno Torre, Jokubas Surgailis, David Ohayon, Alexander Giovannitti, Iain McCulloch, Enzo Di Fabrizio, Alberto Salleo, Jonathan Rivnay, and Sahika Inal. Influence of water on the performance of organic electrochemical transistors. *Chemistry of Materials*, 31(3):927–937, January 2019.

[70] Paolo Romele, Matteo Ghittorelli, Zsolt Miklós Kovács-Vajna, and Fabrizio Torricelli. Ion buffering and interface charge enable high performance electronics

with organic electrochemical transistors. *Nature Communications*, 10(1), July 2019.

[71] Mohammad Javad Mirshojaeian Hosseini, Elisa Donati, Tomoyuki Yokota, Sunghoon Lee, Giacomo Indiveri, Takao Someya, and Robert A Nawrocki. Organic electronics axon-hillock neuromorphic circuit: towards biologically compatible, and physically flexible, integrate-and-fire spiking neural networks. *Journal of Physics D: Applied Physics*, 54(10):104004, December 2020.

[72] Han-Yan Wu, Jun-Da Huang, Sang Young Jeong, Tiefeng Liu, Ziang Wu, Tom van der Pol, Qingqing Wang, Marc-Antoine Stoeckel, Qifan Li, Mats Fahlman, Deyu Tu, Han Young Woo, Chi-Yuan Yang, and Simone Fabiano. Stable organic electrochemical neurons based on p-type and n-type ladder polymers. *Materials Horizons*, 10(10):4213–4223, 2023.

[73] Alexander Giovannitti, Dan-Tiberiu Sbircea, Sahika Inal, Christian B. Nielsen, Enrico Bandiello, David A. Hanifi, Michele Sessolo, George G. Malliaras, Iain McCulloch, and Jonathan Rivnay. Controlling the mode of operation of organic transistors through side-chain engineering. *Proceedings of the National Academy of Sciences*, 113(43):12017–12022, October 2016.

[74] Raufar Shameem, Lukas M. Bongartz, Anton Weissbach, Hans Kleemann, and Karl Leo. Hysteresis in organic electrochemical transistors: Relation to the electrochemical properties of the semiconductor. *Applied Sciences*, 13(9):5754, May 2023.

[75] David Ohayon, Victor Druet, and Sahika Inal. A guide for the characterization of organic electrochemical transistors and channel materials. *Chemical Society Reviews*, 52(3):1001–1023, 2023.

[76] J. J. Liou, A. Ortiz-Conde, and F. Garcia-Sanchez. *Analysis and Design of Mosfets*. Springer US, 1998.

[77] Kazuo Terada, Katsuhiko Nishiyama, and Kei-Ichi Hatanaka. Comparison of mosfet-threshold-voltage extraction methods. *Solid-State Electronics*, 45(1):35–40, January 2001.

[78] Peiyun Li and Ting Lei. Molecular design strategies for ¡scp¿high-performance¡/scp¿ organic electrochemical transistors. *Journal of Polymer Science*, 60(3):377–392, September 2021.

[79] Han-Yan Wu, Chi-Yuan Yang, Qifan Li, Nagesh B. Kolhe, Xenofon Strakosas, Marc-Antoine Stoeckel, Ziang Wu, Wenlong Jin, Marios Savvakis, Renee Kroon, Deyu Tu, Han Young Woo, Magnus Berggren, Samson A. Jenekhe, and Simone Fabiano. Influence of molecular weight on the organic electrochemical transistor

performance of ladder-type conjugated polymers. *Advanced Materials*, 34(4), December 2021.