Department of Physics and Astronomy
University of Heidelberg

Bachelor Thesis
in Physics
submitted by
Gerd Kiene
born in Nienburg (Weser), Germany

2014

Evaluating the Synaptic Input of a Neuromorphic Circuit

This Bachelor Thesis has been carried out by Gerd Kiene at the
Kirchhoff Institute for Physics
Ruprecht-Karls-Universität Heidelberg
under the supervision of
Prof. Dr. Karlheinz Meier

Evaluating the Synaptic Input of a Neuromorphic Circuit

This thesis presents an analysis of the synaptic input circuit found on the High-Input Count Analog Neural Network Chip (HICANN). The HICANN is a neuromorphic chip that includes analog synapse and neuron circuits. The synaptic input circuit, a sub-circuit of the neuron circuit, translates synaptic events from the synapse circuit to exponentially shaped conductance signals influencing the neuron circuit. Consecutive events are integrated to a summation of the exponential shapes. The strength of the signal and the time constant of the exponential decay should be adjustable within the synaptic circuit.

In measurements on the chip the synaptic input circuit shows deviations from the expected behavior. To understand the causes for these deviations, the properties of the synaptic input are analyzed in simulation. It is shown that the synaptic input saturates quickly when exposed to input events. This effect is caused by an undersized capacitor that is used in the circuit for the integration of events. Furthermore it is difficult to set a time constant for the exponential decay. In the circuit this constant is tuned by a resistive element. The resistive property of this element shows a strong dependence on the voltage drop across it resulting in highly non ideal behavior. In addition, an offset which is varying between instances of the circuit is found in the integrator circuit that can not be compensated for within the synaptic input circuit. The extent of these deviations is quantified in simulations. Improvements for the next generation of the chip are suggested in the outlook of this thesis.

Untersuchung des Synaptischen Eingangs einer Neuromorphen Schaltung

In dieser Arbeit wurde die Schaltung des synaptischen Eingangs, die Teil des High-Input Count Analog Neural Network Chips (HICANN) ist, analysiert. Der HICANN ist ein neuromorpher Chip, der analoge Neuronen- und Synapsenschaltungen enthält. Der synaptische Eingang ist ein Teil der Neuronenschaltung und wandelt synaptische Ereignisse in exponentiell abfallende Konduktanzen, welche die Neuronenschaltung stimulieren, um. Aufeinander folgende synaptische Ereignisse werden zu einer Summe der exponentiellen Kurven integriert. Weiterhin sollten die Signalstärke und die Zeitkonstante in der Schaltung einstellbar sein.

Dieser Arbeit vorausgegangenen Messungen zeigte der synaptische Eingang Abweichungen vom erwarteten Verhalten. Ziel der Arbeit war es, die Gründe für diese Abweichungen aufzudecken. Zunächst ist hier die schnelle Sättigung der Schaltung bei Anregung durch synaptische Ereignisse zu nennen. Der Grund für diese liegt in dem unterdimensionierten Kondensator, welcher im synaptischen Eingang genutzt wird. Weiterhin wurden Schwierigkeiten bei der Einstellung der Zeitkonstante des exponentiellen Abfalls festgestellt. Diese Schwierigkeiten werden durch den Widerstand, mit welchem die Zeitkonstante eingestellt wird hervorgerufen. Dieser zeigt ein stark nichtlineares Verhalten. Zuletzt ist auch ein nicht kalibierbarer, zwischen den Instanzen der Schaltung variierender Offset im synaptischen Eingang festzustellen. Im Ausblick der Arbeit werden Vorschläge für mögliche Verbesserungen der Schaltung gemacht.

Contents

1 Introduction

The Electronic Visions group in Heidelberg develops neuromorphic electronic systems in the context of the 'Brainscales' [3] and the 'Human Brain Project' [1] research projects. Neuromorphic electronic systems are systems that provide the user with options to emulate neural networks. They consist of circuitry that is built to mimic the behavior of neurons and their connections. The benefit of using these systems for emulation of neural networks, rather than employing conventional computing systems for network simulation is the scalability of the neuromorphic systems due to comparatively low power consumption per neural event, for details see [17]. The neuromorphic system developed by the Electronic Visions group aims to reach this scalability by using very large scale integration (VLSI) for the development of the neuromorphic circuits. These VLSI circuits are employed in a wafer scale system [21]. Currently the system is scaled up to a multi-wafer-scale system described in [11]. The basic unit of the wafer system is the HICANN. The wafer is divided in reticles, on each of the 48 reticles 8 HICANN chips are found. The biggest building block of the HICANN is formed by the Analog Neural Network Core (ANNCORE) circuits [11]. This block consists of analog neuron and synapse circuits that provide low power models of their biological paragons. Additionally all processes in the neuron and synapse circuits are accelerated by a factor of 10^4 compared to biology. Up to 512 neuron and 114k synapse circuits are available for emulation of neural network models per HICANN. The neuron model implemented on the hardware is an Adaptive Exponential Integrate and Fire neuron model (AdEx), as described in [4]. The design of the neuron circuit design is described in [18].

The presented thesis analyses one of the sub-circuits of the neuron circuit: the synaptic input. The purpose of the sub-circuit is to convert event signals the synapse circuit sends to currents onto the membrane capacitance. Synaptic events are emitted by the synapse circuit as rectangular current pulses. These need to be converted into signals of exponentially decaying shape that are integrated.

In experiments with the wafer-scale system the synaptic input has shown problematic behavior. Visible are fast saturation, problems to measure a synaptic time constant and problematic leakage calibration. The simulations in this thesis show the causes for these problems. Following a better understanding of the current hardware behavior can be achieved and suggestions for future implementations are made.

The chapter 2 describes the model that is used in the synaptic input and ranges for the model parameters. It explains the circuit schematics and uses ideal components to verify the principle functioning of the circuit in theory and in simulation using ideal circuit elements.

Chapter 3 describes simulations of the synaptic input circuit and on individual components of that circuit as they are realized on the chip. For simulation the original transistor

level schematics of the synaptic input are used with the transistor models given in [13]. First the individual components are discussed. The deviations from the ideal circuit behavior (see Chapter 2), can be studied for the individual circuits. In the latter part of the simulation chapter the benchmarks and single circuit results are employed to test the full circuit implementation. Here the full schematic of the synaptic input circuit is tested in settings that correspond closely to the on chip environment. In simulation DC sweeps, AC- and transient-analysis are employed to facilitate the investigation. Special focus in all simulations is put on Monte Carlo analysis using mismatch data for the respective circuits. By simulations using missmatch data differences between the individual realizations of the circuits can be quantified and effects that are visible on chip can be traced back to their origins.

2 Ideal Circuit Operation

This chapter introduces the model used for the synaptic conductances in the hardware implementation and states ranges for time constants associated with this model. The chapter discusses the principle of operation of the circuit: starting with ideal components, the circuits reaction to stimulations is calculated in theory. It concludes with simulations of the full synaptic input circuit using self programmed ideal elements.

2.1 Model and Time Constants

The neuron circuit realized on the wafer scale hardware system, in the following simply called the hardware, should emulate an instance of the AdEx neuron model [4, 11]. The main AdEx model equation is reproduced here:

$$\begin{aligned} C_{mem}\frac{dV_{mem}}{dt} &= -g_l(V_{mem}-E_l) + g_l\Delta_t exp\left(\frac{V_{mem}-V_t}{\Delta_t}\right) \\ &\quad - g^x_{syn}(t)(V_{mem}-E^x_{syn}) - g^i_{syn}(V_{mem}-E^i_{syn}) - w \end{aligned} \tag{2.1}$$

where V_{mem} and C_{mem} are the membrane voltage and capacitance, g_l is the leakage conductance, E_l is the leakage reversal potential, Δ_t the slope factor, V_t the threshold potential, g^i_{syn} and g^x_{syn} are the inhibitory and excitatory conductances, E^i_{syn} and E^x_{syn} are the corresponding reversal potentials, finally w is the adaptation current. A equation for the adaption current and a reset mechanism are also part of the model, they can be found in [4].

The above neuron model consists of various terms that contribute currents onto a membrane capacitance. The voltage across this membrane capacitance V_{mem} represents the result of an integration process over the summed currents. The hardware implementation of the model achieves this by separate circuits for each of the terms, all connected to generate currents onto an capacitor on the chip. The integration process happens on the capacitor: the accumulated charge Q on it is given by the currents flowing onto it over time. The voltage V_{mem} across the capacitor is given by $V_{mem} = Q/C$. The circuitry mimics the equation (2.1) with a physical model [18].

An overview over the complete schematic and a discussion of the implementation of all terms found in the AdEx neuron model is given in [18]. Within the scope of this thesis the synaptic input circuit is tested exclusively, the structure of the term for this circuit out of equation (2.1) is:

$$I_{syn} = g_{syn}(t)(V_{mem} - E_{syn}) \tag{2.2}$$

where I_{syn} is the current put out by the term onto the membrane capacitance, $g_{syn}(t)$ is the synaptic conductance, V_{mem} and E_{syn} are the membrane and synaptic reversal potentials. Here and in the following the discrimination between excitatory and inhibitory synapses is not made anymore, as they are identical from a schematic point of view and are only discriminated by the choice of reversal potential (they can differ in weight and time constant, but this does not make them excitatory or inhibitory). Lower synaptic reversal potentials than the leakage potential E_l results in inhibitory, higher in excitatory terms. Two of the circuits are integrated into each of the neuron circuits, see [18]. The model for the synaptic conductance that is implemented on the hardware can be written as:

$$g_{syn}(t) = \sum_f \bar{g}_{syn} e^{-(t-t^{(f)})/\tau_{syn}} \Theta(t - t^{(f)}) \tag{2.3}$$

where t^f denotes the arrival times of the synaptic events, τ_{syn} is the membrane time constant, $\bar{g}_{syn}$ the synaptic weight and $\Theta(x)$ the Heaviside function. This equation is taken from [8]. Basically this is an exponentially decaying conductance shape that is incremented by a constant step for each synaptic event. The synaptic input should be able to emulate this model with the synaptic time constants available in literature. Important to note in this context is the hardware speedup: all time constants in the biological model are scaled down by a factor of 10^4 [11, 18]. This speedup defines two different time frames: on the one hand *biological time* were processes usually happen on a scale of milliseconds, on the other the *hardware time*, where events on the neuron capacitance happen on a scale of micro seconds. The biological time represents the model time, the hardware time the real time of the experiment. Ranges for the synaptic time constant can be found in various literature sources, see table 2.1. The range of the found time constants is 1.5 to 100 ms. This results in 0.15 to 10 μs in hardware time, assuming the mentioned hardware speedup factor of 10^4.

$\tau_{exitatory}$	$\tau_{inhibitory}$	reference
10 ms	10 ms	[19]
5 ms	10 ms	[23]
2 ms	10 - 100 ms	[16]
1.5 ms	5 ms	[8]

Table 2.1: Synaptic time constants found in literature. Ideally the circuit would cover biological time constants of 1.5 - 100 ms. With the speedup factor of 10^4 this results in hardware time constants of 0.15 - 10 μs

2.2 On Chip Environment of the Synaptic Input

In the ANNCORE circuit on the neuromorphic wafer system synaptic events are represented by current pulses that the synapse circuit sends to the synaptic input circuit. The synaptic input circuit has to turn these current pulses of adjustable height into conductance shapes according to the model stated in equation (2.3). These current pulses have a length of 4 to 10 ns. This length is determined by a full clock circle of the system clock [11]. Currently used a system clock frequency of 100 MHz is used [9], which fixes the pulse width to 10 ns. The synaptic input was not designed for such long pulses, but for half the pulse length according to [18].

The maximal amplitude of the pulses is variable and ranges from 400 nA to 160 μA [18]. Though higher currents are therefore principally possible, the synapse is used with a maximum current setting of 8 μA [20]. This current can be scaled down in discrete steps of 4 bit resolution by the synaptic weight. There is a current offset in the synapse circuit found in simulations described in [12]. For the case of 8 μA maximum current, this results in an offset of approximately 800 nA. When setting the synapse weight to 0 which should turn off the synapse completely, this offset current is still flowing each time a synaptic event is generated. Setting the synapse to lower values probably needs further investigation, especially concerning the possibility to scale the offset with the synaptic weight. As the synapse circuit produces the current pulses by a NMOS-transistor sink configuration, an appropriate voltage is needed at the output of the synapse to ensure correct operation. For the HICANN synapse the output voltage needs to be set to a value higher than 0.8 V, ideal is an operating point of 1 V [20]. This is important for the synaptic input circuit as the output potential is the voltage level at the terminal I_{syn}, for the naming see Figure 2.1.

The synaptic input circuit is tuned by a number of voltages and currents to adjust and bias the amplifier, operational transconductance amplifier (OTA) and resistor found in the circuit. These parameters are set by floating gate circuits discussed in [18]. All currents can be scaled between 0 and 2.5 μA, all voltages between 0 to 1.6 V.

2.3 Ideal Circuit Behavior

In this section the principle of operation for this circuit is discussed. The model equation (2.3) is to be emulated by the circuit. It should give an exponentially decaying conductance that is increased by a constant value each time a synaptic event arrives. The synaptic weight $\bar{g}_{syn}$, the time constant τ_{syn} and the synaptic reversal potential E_{syn} need to be adjustable.

These requirements can principally be met by the circuit shown in Figure 2.1. The current pulses form the synapse array are integrated by a leaky integrator. The leaky integrator in this circuit is built from an operational amplifier with negative feedback over a parallel combination of a resistor and a capacitor. This forms voltage signals of exponentially decaying shape, as discussed below. The rest of the circuit, consisting of two OTAs, converts this shape to a corresponding conductance signal. Ideal circuit

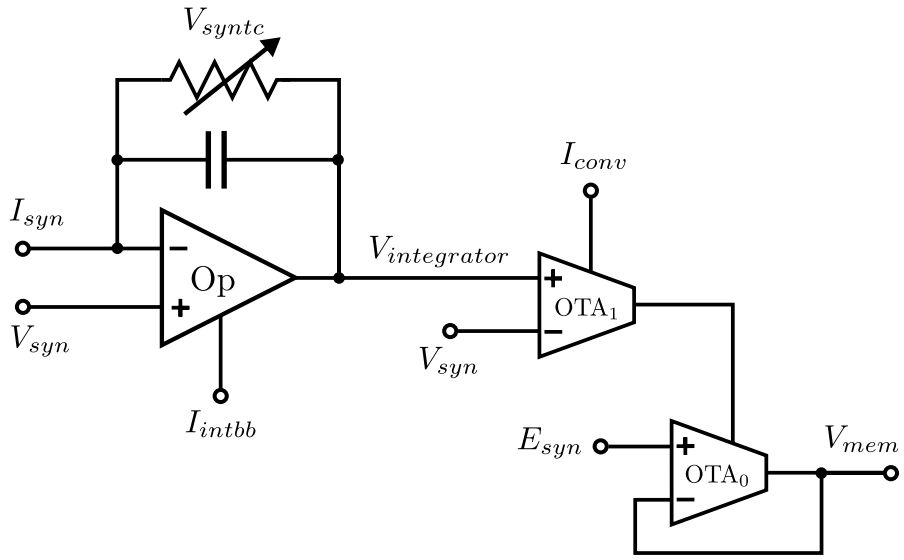


Figure 2.1: Schematic of the synaptic input. The leaky integrator is formed by an operational amplifier (OP), a capacitor on which the voltage is integrated and a resistor that implements the leakage. Here V_{syn} gives a reference voltage, I_{intbb} is the amplifiers bias current, V_{syntc} controls the resistance value and I_{syn} is the terminal that receives the current excitations. The integrated voltage $V_{integrator}$ is converted to a current by OTA_1. The output current of OTA_1 is scaled with a bias I_{conv}. This output current sets the conductance of OTA_0, which is leakage connected with the synaptic reversal potential E_{syn}.

elements for the operational amplifier, resistor and capacitor are assumed in the following calculation. Considering only the leaky integrator with a single input pulse, one can deduce for a single synaptic event starting the integrator at rest:

$$V_{integrator} = \begin{cases} V_{syn} - I_{syn}R\left(1 - e^{\frac{t-t_{start}}{RC}}\right) & t_{start} \leq \mathrm{t} \leq t_{start} \\ V_{syn} - I_{syn}R\left(1 - e^{\frac{t_{syn}}{RC}}\right)e^{-\frac{t-t_{start}-t_{syn}}{RC}} & \mathrm{t} > t_{start} + t_{syn} \end{cases} \tag{2.4}$$

where $V_{integrator}$ and V_{syn} are the voltage over the capacitor and the reference voltage, I_{syn} is the height of the current excitation (and has a negative value), t_{start} and t_{syn} are the start time and the duration of the excitation current. Important to note is that the voltage at terminal I_{syn} is held equal to the reference voltage V_{syn} by the amplifier. This is due to the negative feedback loop over resistor and capacitor. In this circuit the approximation $t_{syn} \ll \tau = RC$ holds, as the minimum synaptic time constant found in literature is 1.5 ms (see table 2.1), corresponding to 150 ns hardware time constant, which is much greater that the 10 ns pulse length that is generated by the synapses circuit. This allows for a Taylor expansion of the first term, simplifying equation (2.4) to:

$$V_{integrator} = \begin{cases} V_{syn} - \frac{I_{syn}}{C}(t - t_{start}) & t_{start} \leq \mathrm{t} \leq t_{start} + t_{syn} \\ V_{syn} - \frac{I_{syn}t_{syn}}{C}e^{-\frac{t}{RC}} & \mathrm{t} > t_{start} + t_{syn} \end{cases} \tag{2.5}$$

The above equations were partially reproduced and extended from [18], all following equations represent own considerations.

A pictorial explanation of the approximation made above is that the charge flowing onto the capacitor does not decay by a relevant amount while the excitation happens. This results in equal voltage steps for each of the excitations. These steps can be seen in a simulation using ideal circuit elements, see Figure 2.2a. Here it can be seen that the height of the excitations is the same for all voltage levels. Later in section 3.2 we will discuss how good this approximation holds in the used cases.

The above equations consider only a single excitation. Neglecting the finite ramp up time during the excitation and assuming it to be a voltage step one can write (because of $t_{syn} \ll \tau_{syn}$) for the membrane voltage under many excitations:

$$V_{integrator} = V_{syn} - \frac{I_{syn} t_{syn}}{C} \sum_f e^{-\frac{t-t^{(f)}}{RC}} \Theta(t-t^{(f)}) \tag{2.6}$$

where $t^{(f)}$ are the excitation times.

The voltage across the capacitor is limited by an equilibrium point. This is set by two counteracting currents that equilibrate in a dynamic equilibrium: one is the average input current given by the synaptic current I_{syn} times the pulse length t_{syn} divided by the average stimulation rate $f_{stimuation,ave}$. The other is given by the current that leaks over the resistor. For small excitations against the overall voltage level this equilibrium point can be calculated to be:

$$V_{integrator} = f_{stimulation,ave} t_{syn} I_{syn} R. \tag{2.7}$$

For larger excitations this still holds approximately. Generally, the equilibrium integrated voltage rises with larger pulse height, larger stimulation frequency and larger resistance values. The stable state behavior and its scaling with the resistance value can be seen in Figure 2.2a.

This voltage is connected to the positive input of OTA_1, the negative input is set to V_{syn}. The difference between the inputs is converted to a current by OTA_1. This conversion would ideally be governed by the equation given in [7]:

$$I_{\mathrm{OTA}_1} = g(I_{conv}) \cdot (V_{integrator} - V_{syn}) \tag{2.8}$$

with $g(x) = x \cdot h_{\mathrm{OTA}}$, where h is a constant characteristic for the OTA implementation. This current signal is finally converted to a conductance shape by OTA_0.

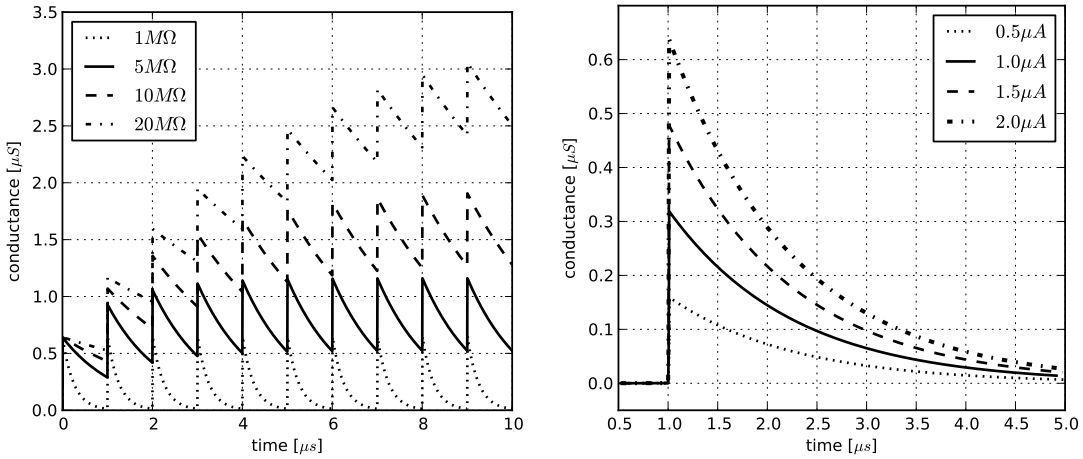
$$I_{\mathrm{OTA}_0} = g(I_{\mathrm{OTA}_1}) \cdot (V_{E_{syn}} - V_{mem}) \tag{2.9}$$

$$g_{\mathrm{OTA}_0}(I_{\mathrm{OTA}_1}) = h_{\mathrm{OTA}_0} I_{\mathrm{OTA}_1} \tag{2.10}$$

Using equations (2.6), (2.8) and (2.10) one can deduce for the conductance of OTA_0:

$$g_{\mathrm{OTA}_0} = -h_{\mathrm{OTA}_0} g(I_{conv}) \frac{I_{syn} t_{syn}}{C} \sum_f e^{-\frac{t-t^{(f)}}{RC}} \Theta(t-t^{(f)}) \tag{2.11}$$

This exactly resembles the theoretical model in equation (2.3), with the substitutions: $\bar{g}_{syn} = -h_{\mathrm{OTA}_0} g(I_{conv}) \frac{I_{syn} t_0}{C}$ and $\tau_{syn} = RC$. This is exact resemblance is not found in the publications reviewed here and represents own work. Assuming all elements are linear and ideal, this circuit is therefore able to perform the task it is meant for. For comparison with realistic circuit simulations the synaptic input circuit is simulated using exclusively ideal circuit elements. The result is shown in Figure 2.2a.



(a) The resistance value is stepped. Time constant and equilibrium point are scaled up with the resistance value. The height of the excitations is no function of the resistance, as derived in equation (2.5)

(b) Scaling the conductance by I_{conv}. By setting the gain of the OTA by giving h and the bias current I_{conv} any conductance can be reached.

Figure 2.2: The synaptic input seen in Figure 2.1 is simulated using ideal elements. The conductance of OTA_0, as seen in equation (2.9) is shown.

It can be seen, that the equations derived above provide a good estimate for the circuits behavior: The height of the reaction of the conductance shapes is independent of the resistance. This is predicted in equation (2.5). Also the equilibrium point due to leakage over the integrator is visible, as derived in equation (2.7). The principle of the scaling by I_{conv} can be seen in Figure 2.2b. Here a single synaptic event is given to the circuit. The height of the conductance shape is directly proportional to the bias current I_{conv}, in correspondence with equation (2.11). Here the approximation does not hold exactly, as the pulses height is not negligible compared to the absolute level, but the principle can be seen.

3 Simulation

This chapter presents simulations of the synaptic input circuit that was designed in [18]. The simulations where all done on the synaptic input and the periphery circuits around it. For an overview see the schematic of the synaptic input in Figure 2.1. Starting with simulations on the single elements of the synaptic input, the chapter moves on to analysis of the complete circuit. The chapter concludes with simulations of the synaptic input circuit in the context of other parts of the neuron circuit to give results that are directly comparable with hardware measurements.

For the simulation ideal sources were used to set the voltages and currents. These were chosen to separate the analysis of the circuit elements from imperfections of the sources on chip.

A number of different analysis is employed in simulation, these are briefly mentioned here. The DC-operating point analysis calculates all voltages and currents in the circuit for a single setting of all biases. An extension to this is the DC-sweep. In a DC-sweep a single variable that is given by a DC-source is scanned over a given range. These DC-sweeps can be parameterized to do sweeps for steps in the value of another variable. Many of the following plots are produced by parameterized DC-sweeps. In transient analysis the circuit is analysed under time dependent excitations, voltages and currents are recorded against time, this is used in the full circuit analysis.

A very important in the course of this thesis was the Monte Carlo analysis of the circuits. The Monte Carlo analysis draws samples out of distributions of transistor parameters that are provided by the chip vendor for every transistor in the circuit. The distributions of transistor parameters are called mismatch data, as they mainly represent the varying width and length of the transistors. The respective analysis (one out of the list given above) is performed on the complete circuit with the altered parameters. This technique allows for simulations that are very close to the actual hardware implementation, as each of the mismatch sets corresponds to a possible instance of the circuit as it is found on chip. Comparatively easy is the Monte Carlo analysis for the DC operating point. Here a large number of simulation runs can be used for generating and recording data for single operating point quantities like a node voltage or the current out of a terminal. For all simulation apart from the DC operating point analysis it is possible to work with Monte Carlo corner cases. These corner cases can be chosen out of a distribution of the operating point analysis and be used for different kinds of analysis like sweeps, transient analysis and others. This way different simulation techniques can be used with the Monte Carlo analysis enabling testing of the transistor level circuits that correspond closely to circuits found on chip.

3.1 Single Circuit Elements

Here the circuit elements found in Figure 2.1 are analyzed. Transistor level models of the single circuit elements as they are realized on chip are used for simulation. Where necessary to enable simulation adaptations are made. The most important adaptation is to replace the low V_T-transistors that are found in the input stage of the OTA and the amplifier with normal transistors of the same width and length for Monte Carlo simulations. Low V_T-transistors differ from normal transistors by their lower threshold voltage. This replacement is necessary, as no statistical data is available on those transistors. The Monte Carlo simulation is not possible if the data is not available for all transistors in the circuit. It was carefully checked that this adaptation does not change the overall behavior, but small deviations are expected. As the conclusions are mostly of qualitative nature this does not affect the results severely. Apart form the Monte Carlo analysis of the amplifier and OTA all simulations are carried out for transistor models that correspond to the transistors that are realized on the chip. All current and voltage sources used in the simulations are ideal sources to isolate the problems the real sources have from the analyzed circuitry.

3.1.1 Resistor

The resistive element in the circuit plays a central role: the only way to adjust the synaptic time constant is to change the resistance value of the tunable resistor. The source for control and nonlinearities in the synaptic time constant is to be found in the resistive element, as the capacitor can not be changed and shows no significant nonlinearities section 3.1.2. According to table 2.1 that gave a range of 0.15 - 10 μs as time constants in hardware time, we see that resistances of 0.6 - 40 MΩ are desirable (the circuits capacitance is fixed at C = 249 fF). The implementation chosen in [18] for this is shown in Figure 3.1. This element was originally designed in [15].

The resistor was tested by fixing one terminal to V_{syn} and scanning a voltage at the other terminal $V_{integrated}$. The naming scheme is chosen to resemble the voltages that are to be found at the terminals of the resistor as found in Figure 2.1. The default settings used in these simulations where: $V_{syntc} = 1.4\,\mathrm{V}$ and $V_{syn} = 1\,\mathrm{V}$, deviations are recorded in the respective figures legend. As we will see in the discussion of Figure 3.4a the absolute voltage level V_{syn} is important for the circuits function, this motivates the usage of the absolute voltage level in the plots.

To visualize the dependence of the resistance value on the control voltage V_{syntc} the simulations shown in Figure 3.2 are employed. The resulting current is shown in Figure 3.2a, the resistance value in Figure 3.2b is the large signal resistance given by Ohms law. The family of curves shown here does not resemble ideal behavior: for an ideal resistor the resistance does not depend on the voltage across the resistor. This resistor implementation shows in contrast more complicated behavior: for low values of V_{syntc} the resistance value is almost constant, but the resistance value below 0.1 MΩ. As argued above the lower limit for appropriate time constants is 0.6 MΩ, the synapse would be therefore to fast for normal operation. For higher settings of V_{syntc} appropriate resis-

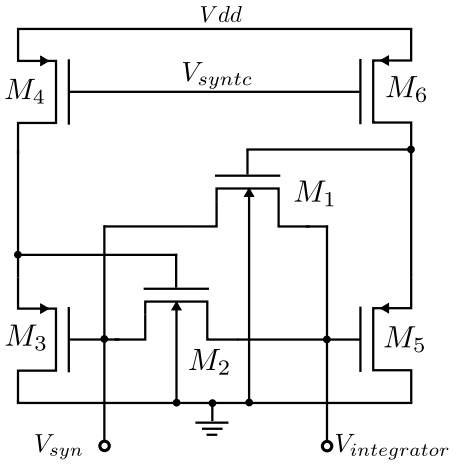
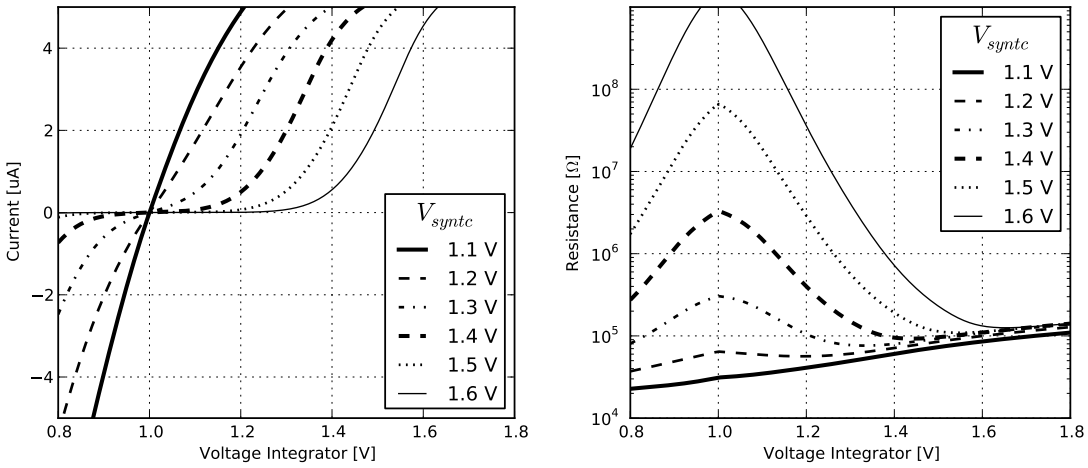


Figure 3.1: Schematics of the resistive element. It represents a floating resistor with two source-follower biasing stages. The transistors M1 and M2 are kept in the triode region. For details of the function see [15]. The resistor was tested by keeping the V_{syn} terminal at a fixed voltage level and scanning the voltage $V_{integrator}$ at the other terminal.

tances are reached, but linearity is lost. In this range of operation the resistance value drops exponentially with higher differential voltage and converges to a value in the range of 0.1 MΩ.

In Figure 3.2 the peaks dependence on V_{syntc} is to be noted: this value depends strongly on V_{syntc} and is reached for a low differential voltage. In Figure 3.3 we extract the maximum value of resistance for each setting and use this as a measure for the total resistance value. In the logarithmic plot a linear dependence is visible: the peak value depends exponentially on V_{syntc}, 6 orders of magnitude can be reached for the maximum resistance. This only holds for small differential voltages, for higher differential voltages the resistor is almost indifferent to the V_{syntc} value: for all settings the curves converge to the mentioned almost constant resistance value at $\approx 10^5\,\Omega$, see Figure 3.2b. As a result the synapse is to fast for all possible settings of V_{syntc}, as soon as a certain voltage is integrated on the capacitor. As the time constants functional dependence is given by the resistor as discussed above, the time constant varies as a function of the integrated voltage by several orders of magnitude for a single set of parameters. The resistances highly nonlinear behavior makes it impossible to set a certain synaptic time constant in the current implementation. We can only speak of a certain time constant function, that can be selected by setting V_{syntc}.

The resistance is dependent on the reference voltage V_{syn} as well, see Figure 3.4a. The complete range of resistances is available for all V_{syn} between 0.4 V and 1.4 V, but V_{syntc} needs to be scaled accordingly. There are no advantages concerning sensibility of the circuit to variations of V_{syntc}, so that the best value for V_{syn} is chosen to make the maximum linear input range of the OTA available while still satisfying the operating range of the synapse. As mentioned the dependence on V_{syn} is the reason for denoting



(a) Current flowing through resistor.

(b) Value of the large signal resistance.

Figure 3.2: Resistor tested for different values of V_{syntc}. The terminal V_{syn} is held at $V_{syn} = 1\,\mathrm{V}$, the voltage at the voltage integrator terminal is scanned. In the current plot, an ideal resistor would produce a straight line with 0 current at 1 V, in the resistance plot an ideal resistor would be represented by a horizontal line. The simulations show in contrast to the ideal behavior a strong dependency on the differential voltage: the resistance value varies over several orders of magnitude.

of the full voltage and not just the differential voltage on the x-axis.

Another parameter is the temperature dependence of the circuit: the current wafer system temperature is at approximately $35\,°C$ and drifts $\pm 3\,°C$ due to time of the day dependent room temperature variation [24]. The result of the resistive element being simulated for different temperature values can be seen in Figure 3.4b. While the time of the day variations are to small to change the systems behavior according to simulation, an influence could be visible when the system temperature rises: the chips temperature could be as high as $65\,°C$ medium [10]. For the resistor the temperature dependence can be calibrated for.

Next the device variations considering transistor mismatch will be discussed. An example distribution of the resistance value for a single combination of V_{syn} and $V_{integrated}$ is shown in Figure 3.5a. The distribution is recorded for a number of different operating points, the resulting means and the corresponding standard deviations are shown in Figure 3.5b. The resistors dependence on mismatch is significant, resistors in different neuron circuits therefore can show significantly different time constant functions for the same V_{syntc} and V_{syn} settings. The functional dependence of the synaptic time constant needs to be calibrated for by scaling V_{syntc} accordingly. Due to the extreme sensitivity of the resistance on V_{syntc} this is difficult to accomplish. The sensitivity of the resistive element was mentioned in [18] and noted circuit property that needs improvement.

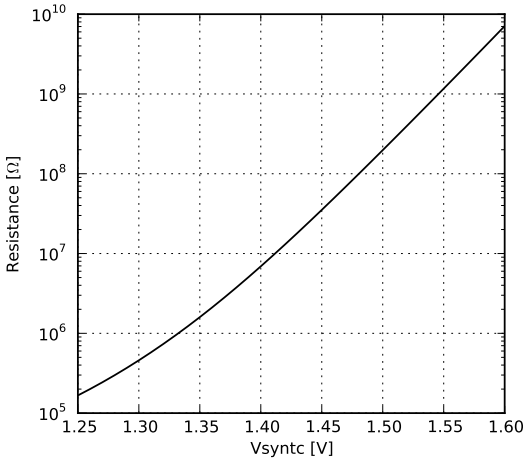
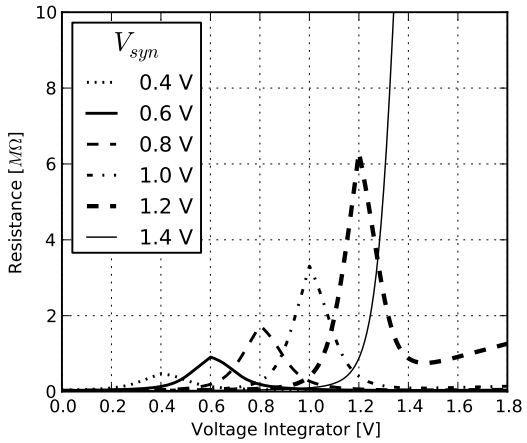


Figure 3.3: Here the maximum resistance is plotted as a function of V_{syntc}. The maximum value was extracted out of Figure 3.2b by recording the maximum for that curve. An approximate linear dependency can be seen in the logarithmic plot, the peak resistance value is exponentially dependent on V_{syntc}.

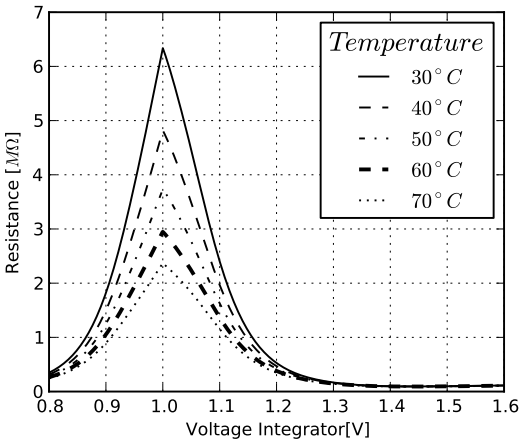
3.1.2 Capacitor

The capacitor in the integrator of the synaptic input circuit is fixed at $250fF$ and realized as a MIMCAP. An An ideal capacitor would show a constant capacitance, independent of the common mode voltage or differential voltage. The MIMCAP model available in the simulation software showed voltage dependent variation of the capacitance below $1/1000$ of the total value over a voltage range of 0 to 1.8 V. As MIMCAPS are very stable in consideration to DC voltage level no large deviations from linearity due to voltage differences were expected, see for example the deviations in [5]. We can therefore assume the capacitor to be constant, especially in comparison to the resistor.

Important for the capacitor is not only the linearity, but also mismatch considerations. As there is no mismatch data available for this capacitor, this effect can not be simulated here. The missmatch for the capacitors is process dependent, but general considerations can be found in [6] suggests variances below 1%. This would give a small error on a critical quantity, as the capacitor value, according to equation (2.5) limits the height of the voltage rise to a single excitation current. As the capacitance can not be calibrated directly, the only option to calibrate for this variation would be to change the amplitude of the current pulses put out by the synapse.

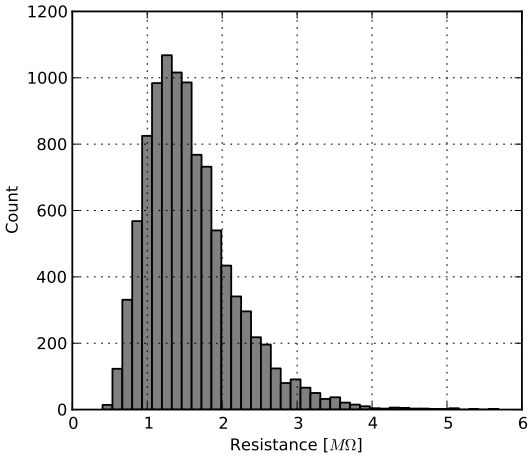


(a) Resistor for different reference voltages V_{syn}. The resistance value depends on the reference voltage, small reference voltages result in small resistances for the same V_{syntc}.

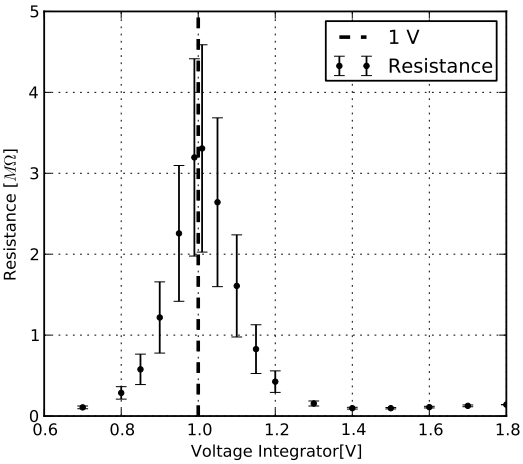


(b) Resistor circuit in different temperature regimes. Day to day variations are on a scale of single °C and therefore negligible.

Figure 3.4: Both simulations using $V_{syntc} = 1.4V$.



(a) Distribution of resistance for $V_{integrated} = 1.1V$, using N = 10000 points.



(b) Variation of resistance for different voltage levels: each of the points is simulated as a distribution like shown in (a).

Figure 3.5: Monte Carlo simulation of the resistance value. $V_{syn} = 1V$ and $V_{syntc} = 1.4V$ are fixed for these figures. The resistive elements behavior varies significantly with mismatch. These variations can be calibrated for.

3.1.3 Amplifier

The amplifier is the final part of the integration stage. It holds its negative terminal (I_{syn}) at the same voltage level that is provided as reference at the positive terminal (V_{syn}). It restores the voltage at the I_{syn} terminal by providing the current that was drawn by the synaptic event. This current is integrated on the capacitor.

The bias current for the amplifier is by default set to -2 μA. This current enhances the circuits transient performance under load: for larger offset currents larger loads can be driven.

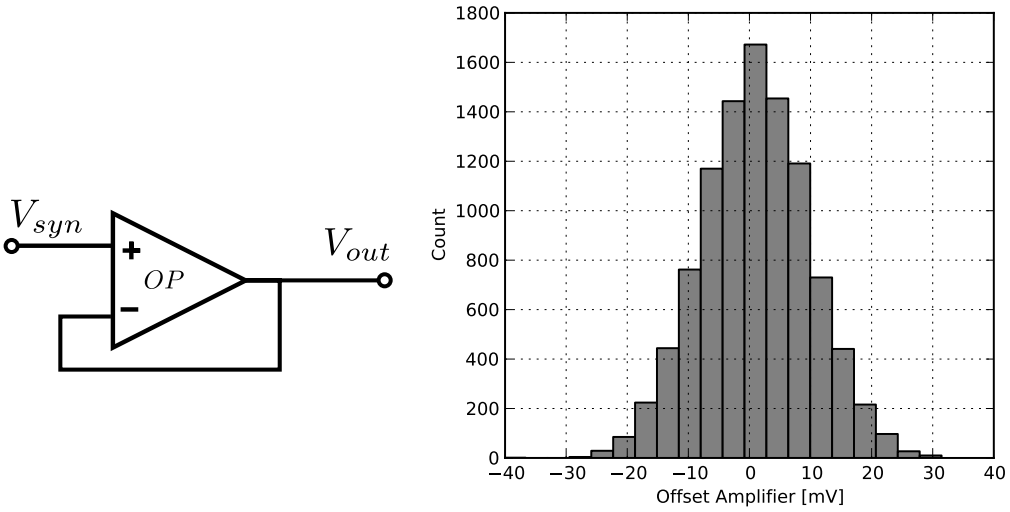
Important for the function of the synaptic input is the amplifiers offset. The offset voltage is the voltage at which the amplifier provides 0 current at its output terminal. This is tested by giving direct negative feedback to the input on a chosen voltage level and simulating the voltage difference between the two terminals, see Figure 3.6a for the schematic. The simulation is performed for many (N = 10000) different samples of mismatch for all transistors in the amplifier. The resulting offset distribution can be seen in Figure 3.6b. The result of the Monte Carlo simulation is an offset of:

$$V_{offset,amplifier} = 0.7 \pm 8.7\,\mathrm{mV} \tag{3.1}$$

This offset is like the other deviations caused by mismatch fixed for every single instance of the amplifier that is realized on the chip. The offset voltage is stable over the range of possible V_{syn}, with variations of ± 2 mV for the medium value, no preferred value for the reference voltage can be given here. There is also no possibility to calibrate for this offset, as the integrator works always relative to the given V_{syn}, see equation (2.4). This results in an offset voltage at the OTAs inputs that is per instance of the circuit constant and different between the instances, resulting in a nonzero conductance signal in absence of excitations. The OTA is another source of offset, that will be presented in the following.

3.1.4 Operational Transconductance Amplifier

Here the operational transconductance amplifier (OTA) used in the circuit is analysed, for a symbol see Figure 3.7a, for the ideal behavior equation (2.8). The two consecutive OTA in the circuit perform a voltage to conductance conversion. For a full circuit analysis of this converter see section 3.2.1. The OTAs are connected differently: OTA_1 has two diode connected transistors as load, one of them connected as voltage drop, the other forms one end of the input current mirror for the biasing stage of OTA_0. For OTA_0 the output and the negative input terminal are connected to the voltage across the membrane capacitance. These two have to be tested separately due to the load voltage dependency of the output that can be seen in Figure 3.7b. Depending on the load voltage the current range is limited: for very high/low loads the current can only flow into/out of the OTA. The behavior is qualitatively similar for the two different load configurations, but quantitatively the circuits response depends severely on the load voltage. This makes separate analysis of the two OTAs advisable, which is performed here. As load voltage sources are used: for the diode load the input characteristic is quite steep so that a



(a) Schematic of the amplifier offset test. The voltage V_{syn} is given, the offset is given by $V_{syn} - V_{out}$

(b) Monte Carlo analysis of the amplifiers offset with N = 10000 points. Using the settings $I_{intbb} = 2\,\mu\text{A}$ and $V_{syn} = 1\,\text{V}$. Mean: 0.7 mV, standard deviation: 8.7 mV.

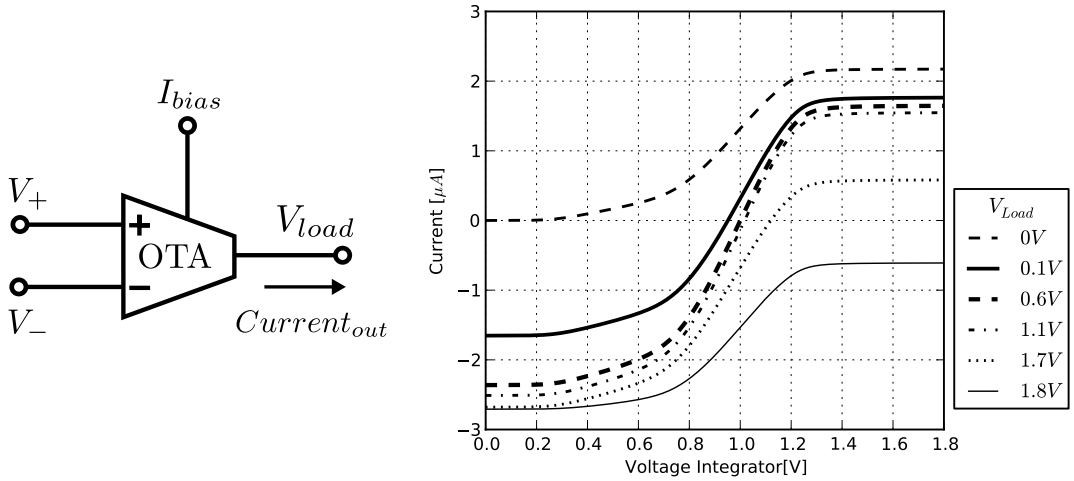
Figure 3.6: The offset of the amplifier is tested, the test principle and the result are shown. This offset can not be calibrated for, as the integrator always integrates relative to the reference voltage V_{syn} with which the integrated voltage is differentiated at the input of OTA_1.

constant source provides a good estimate of the used case. For OTA_0 the constant load voltage mimics a infinite capacitance.

The effect of the OTAs bias current is shown in Figure 3.8. It can be seen that the maximum output current is limited by the bias current in both cases. This limit is given because the input current to the OTA is mirrored to the outside, the direction of the mirroring and what part of the current is mirrored is given by the input voltages at the positive and negative terminal. The assumption that the maximum output current is given by the bias current is not true for bias currents greater than $2\,\mu\text{A}$: here saturation effects are visible, larger biases do not result in higher output currents.

For the effects of the reference voltages in the synaptic input, V_{syn} at the integrator with OTA_1 and E_{syn} as the synaptic reversal potential connected to OTA_0 see Figure 3.9. The reachable current and the linearity are functions of the reference voltage, as shown in [14] and visible here. For values next to ground or the voltage rail large offsets are visible for the leakage connected OTA: for example in Figure 3.9a the curve with $V_{ref} = 0.5\,\text{V}$ shows zero output current at 0.4 V. For the open loop OTA these effects are not visible to the same extend, due to no biasing by changes of the load. The ideal operating range will be subject of discussion in section 3.2.

The OTAs offset is simulated the same way the amplifiers offset was tested: by giving a voltage offset as positive signal and negative feedback, for the schematic see Figure 3.10a.



(a) Symbol of an OTA. For OTA_1 V_+ is connected to the output voltage of the integrator, V_- is held at V_{syn}. For OTA_1 V_+ is connected to E_{syn}, V_- is connected to the output.

(b) The voltage at the output of the OTA is swept. Current differences due to the load voltage of $\approx 1\,\mu$A in the range of interest (between 0.1 and 1.7 V) are recorded. For measurement the OTAs with different loads have to be considered separately to gain reliable results.

Figure 3.7

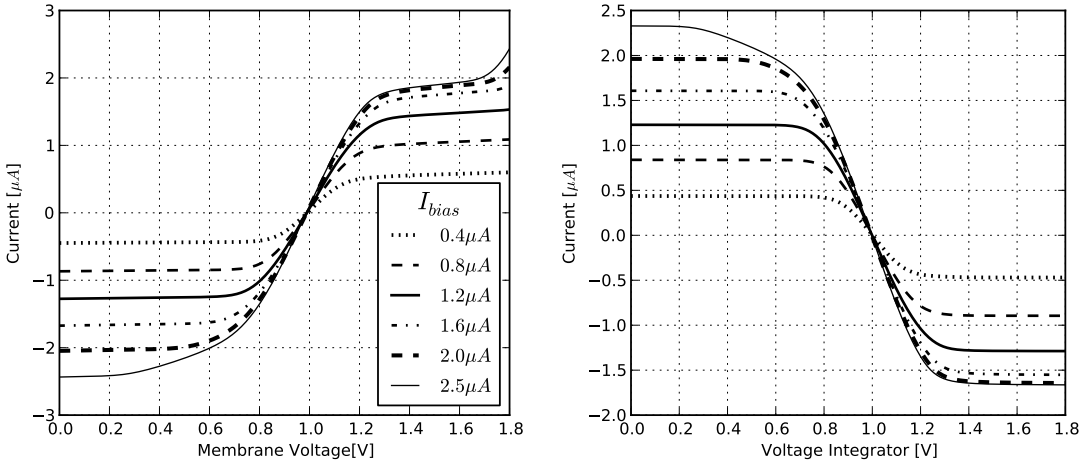
This results in Figure 3.10b, the value of the offset is:

$$V_{offset,OTA} = -12 \pm 23\,\text{mV} \tag{3.2}$$

Result of this is a large nonzero medium and a standard deviation that is much larger than for the amplifier. This offset is no problem when it can be calibrated for like in the case of OTA_0. As the input terminals are independent form one another the reference voltage can be set to cancel the offset voltage. For OTA_1 this is not possible: here the integrator and the integrators reference are connected to the positive and negative input terminals, no independent level is possible and therefore no calibration. The amplifiers offset that was discussed before is included in the integrator signal, the OTAs offset is additional. Both can not be calibrated for, as the integration process is, as mentioned, relative to the voltage reference. As both offsets are distributed gaussian, they can be added quadratically. This culminates in an estimate for the offset that can not be calibrated for in the circuit that is an input signal for OTA_1 even when no synaptic events are present:

$$V_{offset,total} = -11 \pm 25\,\text{mV} \tag{3.3}$$

To verify this result both circuits were simulated together in. This simulation, with $I_{\text{OTA},bias} = 2.5\,\mu$A, $I_{\text{OP},bias} = 2\,\mu$A yields an offset of: $-13 \pm 26\,$mV. There are no significant differences between the quadratic addition and the combined simulation observed.



(a) OTA_0, the load is equal to the Membrane voltage. The reference potential was set at 700 mV.

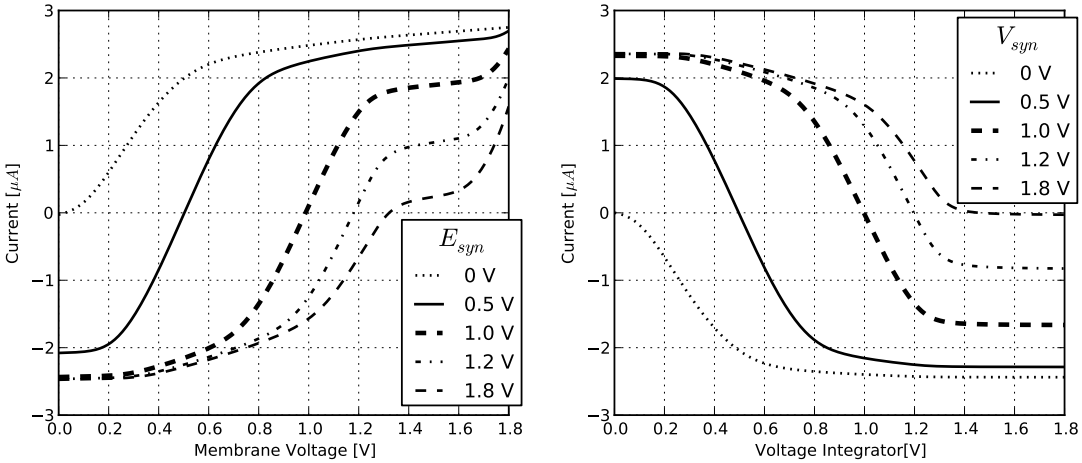
(b) OTA_1, using a fixed load voltage of 500 mV. For the legend see (a).

Figure 3.8: OTA output current for different bias currents in the leakage connected and open loop configuration. Two saturation effects can be noted: for all settings of I_{bias} the linear input range is about 200 mV voltage difference between the terminals and above 400 mV the input changes very little. The other saturation effect is the dependence on the bias current: here the maximum output current in not enlarged substantially for currents above 2 μV

3.1.5 Current Mirror

The next circuit element that is considered here is the current mirror used to mirror the bias current for the leakage circuit. The leakage circuit is part of the neuron circuit and consists of a single OTA that is connected like OTA_1, that was tested above. For more discussions see [18]. This is no direct part of the synaptic input, but has severe impact on the circuits behavior and is important for the simplified neuron simulations shown in section 3.3. Parts of this measurements in were presented in [18] but they are extended here to give a better idea of uncertainties that arise when setting values for I_{gl}. The floating gate cell that provides the current can output 0 to 2500 μA. The current mirror is adjusted by two bits: a 'fast' and a 'slow' bit. These can set the transformation ratio to 1:1, 1:3, 1:9 and 1:27, the current is multiplied with the respective factor. The lower the ratio, the slower the membrane, the default ration is 1:3 [18]. It can be seen in Figure 3.11 that the behavior of the current mirror may deviate severely from the wanted behavior. The distributions show both a medium deviation from the ideal transformation ratio and a large spread. The result is an uncertainty concerning the bias current I_{gl} and a limited range for some neurons. It will be difficult in calibration to separate the effects due to OTA mismatch from the current mirror faults. Both will add up to a number of neurons with limited range or precision.

For low currents the transformation ratio is current dependent as well, see Figure 3.12.

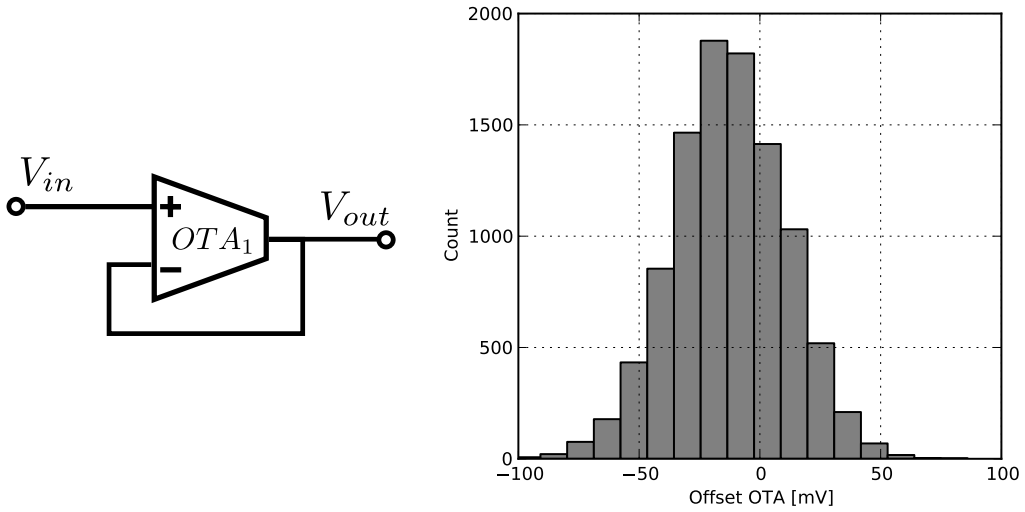


(a) OTA_0, the load is equal to the Membrane voltage.

(b) OTA_1, using a fixed load voltage of 500 mV.

Figure 3.9: The reference voltage is stepped. For very high/low reference voltage, the current is biased towards negative/positive values. OTA_0 shows very nonlinear behavior for synaptic reversal potentials of $E_{syn} > 1\,\mathrm{V}$. Apart from these deviations the two different configurations show similar properties.

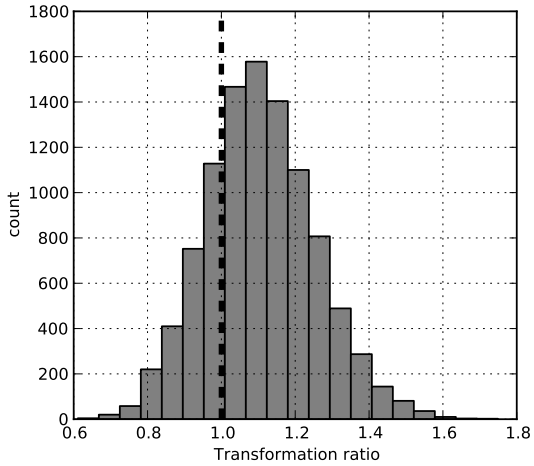
This results in significant uncertainties when setting large time constants (that correspond to small values for I_{gl}). As the ratio is biased towards higher averages, for most neurons the full range given by the 1:3 setting will be available, but the precision might be smaller than expected.



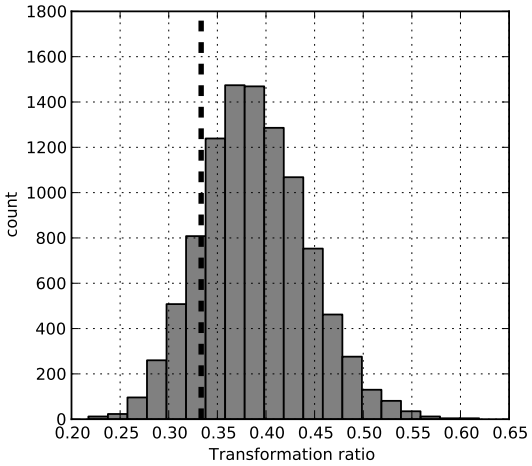
(a) Schematic of the OTA offset test. The difference between the terminals $V_{in} - V_{out}$ in DC operation is a measure for the offset of the amplifier.

(b) Monte Carlo analysis of the OTAs offset with N = 10000 points. Using the settings $I_{bias} = 2\mu A$ and $V_{syn} = 1V$. Mean: $12.7mV$, standard deviation: $23.4mV$

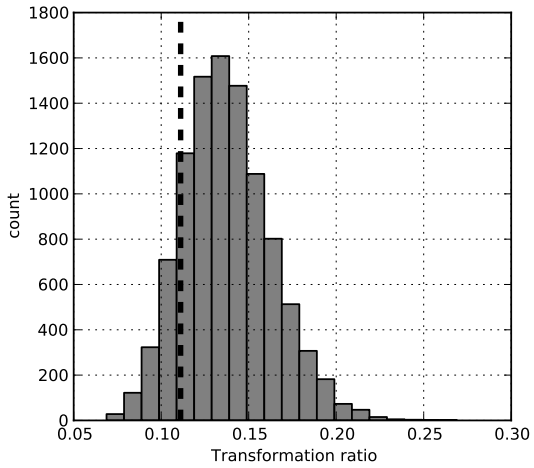
Figure 3.10



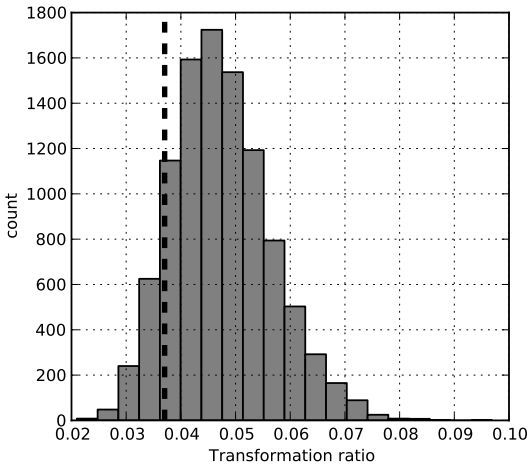
(a) Transformation ratio 1:1. $\mu = 1.1$, $\sigma = 0.1$



(b) Transformation ratio 1:3. $\mu = 0.38$, $\sigma = 0.05$



(c) Transformation ratio 1:9. $\mu = 0.14$, $\sigma = 0.03$



(d) Transformation ratio 1:27. $\mu = 0.047$, $\sigma = 0.009$

Figure 3.11: Transformation ratio of the current mirror for the different settings of the "fast" and "slow" option. The projected transformation ration is shown in the respective caption and as dotted line in the histogram. The mean is denoted as μ, the standard deviation as σ. For these recordings, the maximum input current, $2.5\mu A$ was used and N = 10000 samples drawn each time.

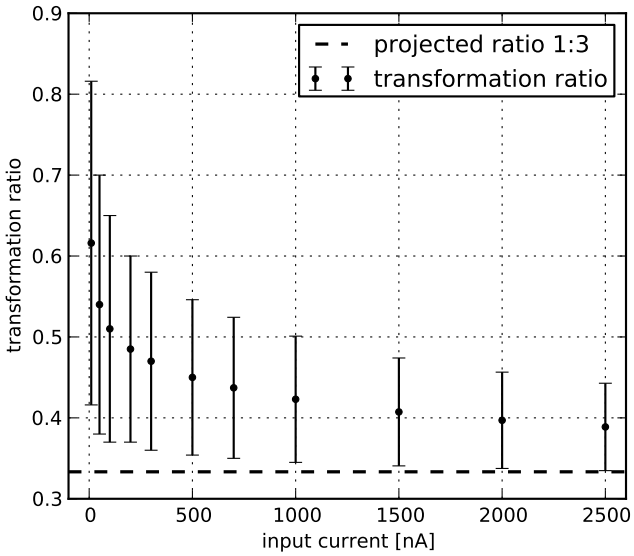


Figure 3.12: Transformation ratio of the current mirror as function of the current to be translated, the error-bars indicate the standard deviation of the resulting distribution. Especially for low currents the transformation ratio is quite different from the wanted ratio and object of large uncertainties.

3.2 Complete Circuit Behavior

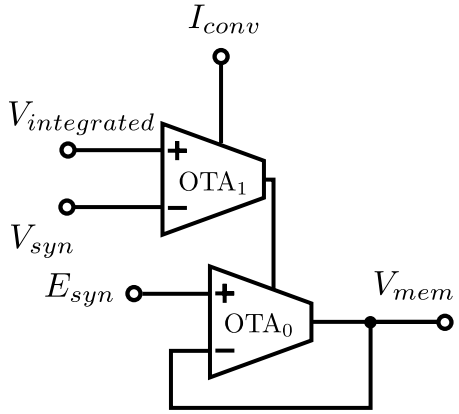
Here the single circuit elements of the synaptic input that were discussed in section 3.1 are combined to realize higher level circuits. This allows for composite analysis of the elementary transistor level circuits and provides results very close to measurements on chip. Three steps are taken in this analysis: First two macro modules, the voltage to conductance conversion stage formed by OTA_0 and OTA_1, as well as the integrator with OTA_1, are simulated and analyzed. Second the benchmarks of Chapter 2 meet with the analysis that where done here so far by using the full circuit schematic as seen in Figure 2.1 with all real transistor level models of the single circuit elements in simulation. As a final step, a simplified neuron setup is used to produce results that are closely comparable with hardware measurements. To ensure this the default settings are all synchronized with the settings that are common in hardware tests [22].

3.2.1 Voltage to Conductance Conversion

As a first macro module the voltage to conductance conversion stage, formed by the two OTA connected like they are found in the synaptic input, was simulated. For the sub-circuits schematic see Figure 3.13a. This macro circuit converts the integrated voltage to a conductance shape. The settings for this test are recorded in Figure 3.13b.

In this test the voltage at the positive terminal, $V_{integrated}$ is swept. The result of doing this for different values of the reference voltage V_{syn} is shown in Figure 3.14b. Several conclusions can be drawn from these simulation results: The maximum positive conductance is available when using low reference voltages V_{syn}. Reference voltages below 500 mV are generally not well suited for operation of the OTAs in the neuron, as the OTAs output characteristic is highly asymmetric for these values, see Figure 3.9b. For 500 mV a balance concerning conductance and linearity of the voltage to current conversion stage is reached. But this value for V_{syn} is not ideal for the surrounding circuitry: the synaptic circuit needs a minimum line voltage, that is given by 0.8 V. As the ideal voltage for the operation of the synapse is given by 1 V [20] and the difference in conductance between the two reference voltage values is not large, no new operating point is suggested. An important fact visible in Figure 3.14b is that only integrated voltages of above the reference voltage V_{syn} are resulting in conductance values different from zero. This is because the OTA can only be biased with positive currents (the mirroring diode is connected to ground). Also visible is the linear range of the OTA: for the used case of $V_{syn} = 1\,V$ the input is approximately linear for integrated voltages of up to 1.1 V. The voltages up to 1.4 V do still change the output conductance, all higher values have no measurable effect. This limits the maximum integrated voltage that influences the membrane to 1.4 V, for larger integration ranges the linear range of OTA_1 would need to be extended.

The maximum conductance the voltage to conductance conversion can reach is scaled with the bias current I_{conv} that is given into OTA_1. Higher bias currents result in larger conductances, as explained in section 3.1.4. The influence the bias currents have can be seen in Figure 3.14a. It can be seen that further increasing the bias current has



(a) Schematic of the voltage to current conversion stage of the synaptic input.

variable	default value
E_{syn}	800 mV
V_{mem}	700 mA
V_{syn}	1 V
I_{conv}	2.5 μA

(b) Standard settings for the test of the voltage to conductance conversion stage.

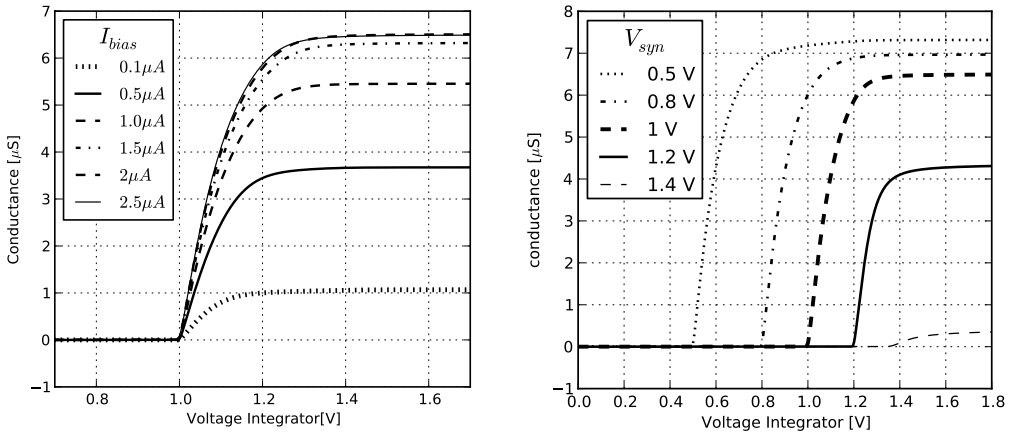
Figure 3.13

only little effect as soon as the saturation effects found in section 3.1.4 are reached for bias currents larger than 2 μA. The maximum conductance that can be reached with the synaptic input for these settings is 6.5 μS, the reachable conductance is therefore comparable with the conductance of the leakage circuit (For more on the leakage term see below). Hence the synaptic term should be able to influence the membrane potential strongly. Even for maximum settings of I_{gl} (assuming the default 1:3 current mirror setting, see section 3.1.5) the membrane potential can be pulled up to the mean value between synaptic and leakage reversal potential for maximum settings of I_{conv}.

3.2.2 Synaptic Event to Current Shape Conversion

In Figure 3.15 another macro module was tested within the synaptic input: the integrator with the voltage to current conversion OTA for the current offset the OTA was generating at its output. These currents were connected to the biasing stage of OTA$_0$. To achieve realistic behavior, the settings of Figure 3.18b were used, that are similar to what hardware users employ for their measurements on chip. The current offset was recorded in a Monte Carlo simulation using mismatch.

The current is exclusively positive: the mirror transistors that form the load of OTA$_1$, that is converting the offset of the integrating capacitor and its own to a bias current for OTA$_0$ are connected to ground. There is no lower potential then ground in this circuit, the current can only flow into the OTA (by convention being positive), not out of it. As the medium offset of the stage is negative ($V_{offset,total} = -11 \pm 25$ mV section 3.1.4) most of the samples therefore do not contribute to the membrane in the default setting, but are biased towards not being excited.



(a) Scaling of the voltage to current conversion with I_{conv}. Saturation is visible for very high currents.

(b) The output conductance for different reference voltages V_{syn}.

Figure 3.14: The standard case for the reference voltage is 1 V, here the linear range reaches up to 1.2 V, integrated voltages above 1.4 V have no effect on the membrane.

3.2.3 Complete Circuit

Next the complete synaptic circuit as it is shown in Figure 2.1 is used in simulation. The settings given in Figure 3.16a are used as default. When different settings are used the deviations from the default settings are reported.

The results of a run with default settings is shown in Figure 3.16b. The top picture shows the resulting conductance, the middle one the integrated voltage and the bottom picture the excitation current. Due to the input excitations the circuit reaches the steady state described in section 2.3 after a few input excitations. The input events cause a almost instantaneous voltage step followed by a decay of exponential-like shape, following the excitation. This proves that the circuit is able to perform the function it was designed for in this setting: a substantial current is put onto the membrane that is given by a rate of excitation that is integrated in the circuit. But it is visible in Figure 3.16b that this state does not test the dynamic range that should be ideally covered by the circuit: the pulse height given out by the synapse circuit is only 1/8 of the possible height (actually the default setting is the lowest possible setting, see section 2.2) and the integrated voltage is just $\approx 150\,\text{mV}$ above the reference voltage, while the integrator could support voltages up to the voltage rail. Still we see already: the input reaches the steady state after about 5 synaptic events, which foreshadows a possible fast saturation.

As we discussed above, the dynamic range is very limited, in Figure 3.17 two factors limiting the performance are shown. In Figure 3.17a the pulse height is set to the maximum value ($8\mu A$), which is 8 times the default. The result are excitations that are 8 times higher, one can derive from equation (2.5), that the voltage step is always given by

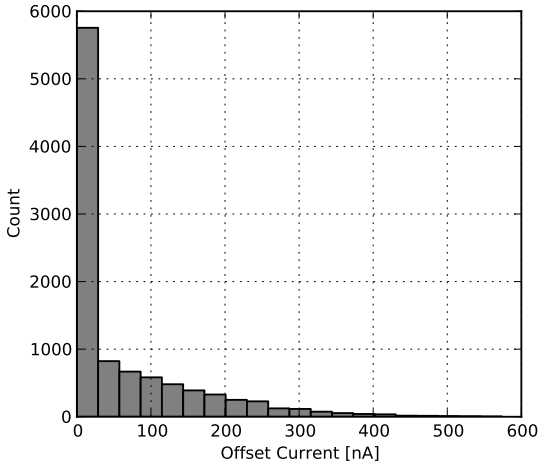


Figure 3.15: DC current due to offsets of amplifier and OTA that flows as biascurrent for leakage connected OTA. The distribution should be symmetric following the above discussion of the offsets of the devices. But as this is measured as part of the complete setup, the negative currents can not flow (would require a potential lower than ground, that is not available).

$$\Delta V_{syn} = \frac{I_{syn}}{C} t_{excitation} \tag{3.4}$$

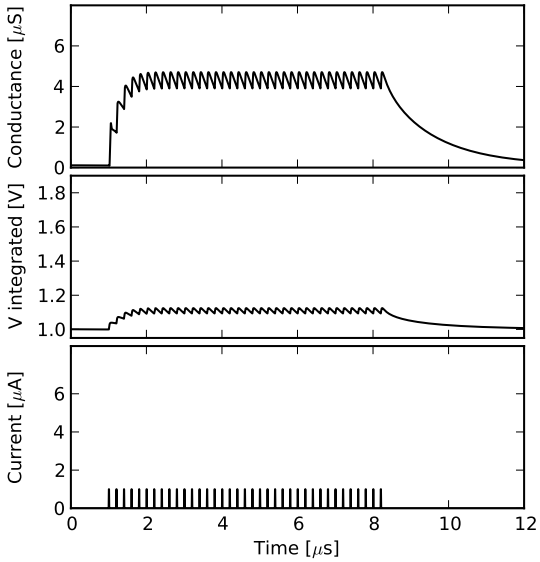
For this circuit we have $C = 249\,\mathrm{fF}$, $I_{syn} = 8\,\mu\mathrm{A}$ and $t = 10\,\mathrm{ns}$. This results in $\Delta V_{integrator} = 320\,\mathrm{mV}$ for a single spike. As the voltage over the integrating capacitor equals the voltage across the resistor this raises the voltage level across that element by $\Delta V_{integrator}$ as well. The higher voltage decays with a much faster time constant, see Figure 3.2: the voltage change induces a drop of the time constant by 2 orders of magnitude. This results in an stable state behavior for the integrated voltage that is visible in Figure 3.17a, the voltage steps decay very fast, although the current was increased by a factor of 8, the mean voltage is just increased by a factor of 2. Another effect visible here is the limited linear input range of the OTA. This is seen in the conductance peaks: they do not show the same sharp edge as the integrated voltage, as the OTA does not react linear above 200 mV difference between the terminals, see section 3.2.1.

This limitation of the output by the linear input range is well visible in Figure 3.17b. Here the resistor is set to a very high value ($V_{syntc} = 1.6\,\mathrm{V}$), resulting in integration up to the voltage rail. It is visible that the input is still integrating the incoming events, while the output conductance is already saturated and does not change at all after the integration process has reached 1.4 V.

We see two major difficulties in these tests: one is the small capacitor that can not handle the large current pulses that are put out by the synapse. The integrated voltage rises over the linear input range of the OTA after just two synaptic events. The other is the extreme nonlinearity seen in the resistive element, see also section 3.1.1. Due to

variable	default value
E_{syn}	800 mV
I_{conv}	2.5 μA
I_{intbb}	-2 μA
V_{mem}	700 mV
V_{syn}	1 V
V_{syntc}	1.4 V
C_{input}	1 pF
$F_{modulation}$	60 KHz
$F_{stimulus}$	5 MHz
I_{syn}	1 μA
t_{syn}	10 ns

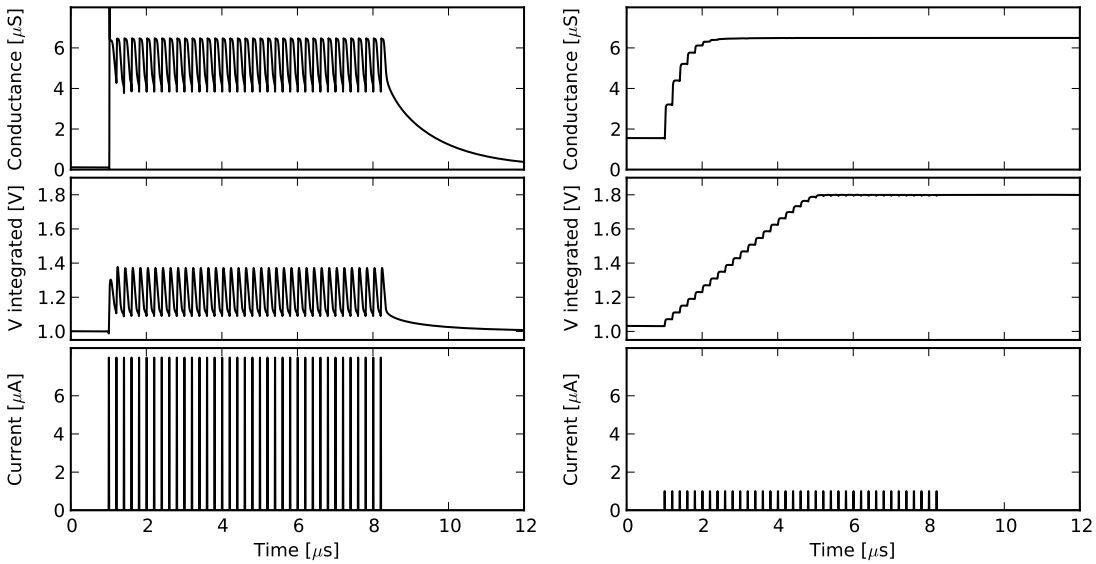
(a) Standard settings for the transient simulation of the synaptic input. Deviations from these standard settings are recorded in the respective figures caption.



(b) Run with the default settings given in Figure 3.16a. The top figure shows the output conductance $g = I_{out}/(V_{mem} - E_{syn})$, the middle one the voltage on the integrating capacitor and the bottom one the excitatory current.

Figure 3.16: Default run for the full circuit analysis.

the extremely voltage dependent time constants the synaptic input is not characterizable with a single time constant and significant nonlinearities can be observed.



(a) $I_{stimulus} = 8\,\mu\text{A}$ **(b)** $V_{syntc} = 1.6\,\text{V}$.

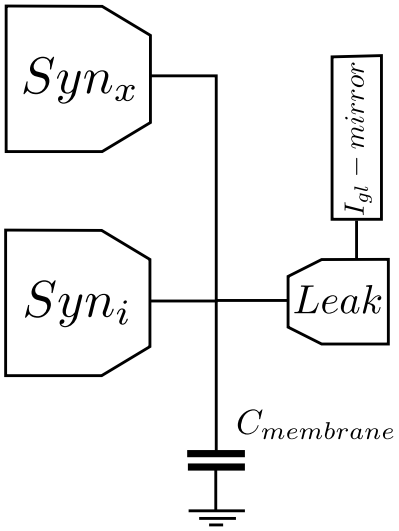
Figure 3.17: Transient analysis, for explanation of the sequence see Figure 3.16b In (a) a very high input current is used, here it is visible, that the membrane voltage rises sharply due to the input current and decays extremely fast due to these excitation. A sharp artifact at the start is visible because the amplifier can not follow the steep voltage slope. In (b) the limited input range of the OTA can be seen: the conductance output saturates, when the integrated voltages reaches 1.4 V, resulting in 400 mV differential voltage.

3.3 Simplified Neuron Circuit Test

In this section tests on a simplified Neuron model will be used to test the circuit in a similar environment as the one that is used on the chip. The simplified neuron model uses a leakage term (see [18] and equation (2.1)) with the current mirror used on chip, see section 3.1.5. The synaptic terms share their schematics and are only distinguishable by their reversal potentials: for the inhibitory term the synaptic reversal potential is below, for the excitatory term the potential is above the leakage potential. The full neuron schematic is not used to simplify the troubleshooting and exclude influences of not fully understood terms. Most of the tests were done due to questions of hardware users. The aim is to reproduce behavior that was measured to make sure that these measurements are no artifacts, as they show serious malfunctions of the circuit. Especially important to simulate here is the membrane voltage V_{mem}, as this is the only voltage in the neuron that can be read out in measurements [18]. The simulations done here should aid the calibration of the circuits and show how the parts that where analyzed before interact, especially by using Monte Carlo analysis to get realistic performance.

variable	value
E_l	700 mV
E^i_{syn}	600 mV
E^x_{syn}	800 mV
V_{syn}	1 V
V_{syntc}	1.4 V
I_{intbb}	-2 μA
I_{gl}	2.5 μA
I_{conv}	2.5 μA
I_{syn}	0 μA

(a) Default settings for the simplified neuron circuit test.



(b) Schematic for the simplified neuron test.

Figure 3.18: In the simplified neuron setup two synaptic terms (Syn_i, Syn_x) alongside a leakage term ($Leak$) are connected to a membrane capacitance of the same size that is available in the neuron. For realistic biasing of the leakage term the current mirror realized on chip is included in the simulation.

3.3.1 Uncalibrated Membrane Voltage

Here the DC-operating point is calculated for a multitude of different Monte Carlo samples considering mismatch. This measurement is done because in hardware measurements it proves to be difficult to set the resting potential for some neurons directly by setting the leakage potential E_l. Here the reason for these deviations are to be visualized. The settings used for this test are reproduced in Figure 3.18a.

Note the very low setting for I_{gl} that is used to illustrate the effect due to input offsets directly. The distribution gets narrower for higher values of I_{gl}, but by using large values of I_{gl} one omits the possibility to set large time constants.

This results in a behavior that is dependent on the I_{conv} settings, shown in Figure 3.19. With I_{conv} the terms can be completely shut off, resulting in next to no influence on the membrane potential. The result is as anticipated: for setting both I_{conv} to 0 a single peak was obtained. By setting the respective I_{conv} to the maximum the other plots show additional peaks where the reversal potential for the respective plot is seen.

Out of the multitude of Monte Carlo samples a subset was chosen from Figure 3.19d that was approximately equally distributed over the histogram to encounter for all kinds of mismatch combinations. The corners were selected by equidistant steps out of Figure 3.19d to cater for all possible cases. The selected corners were then used to produce Figure 3.20a. Here one can see, that the leakage term can, in some cases never control the membrane potential fully. This is due to the different strength the nonzero offset for a number of neurons in the integration stage. For some neurons there is a large offset

current that is comparable with the maximum leakage bias current ($I_{gl,max} = 833\,\mathrm{nA}$), see Figure 3.15.

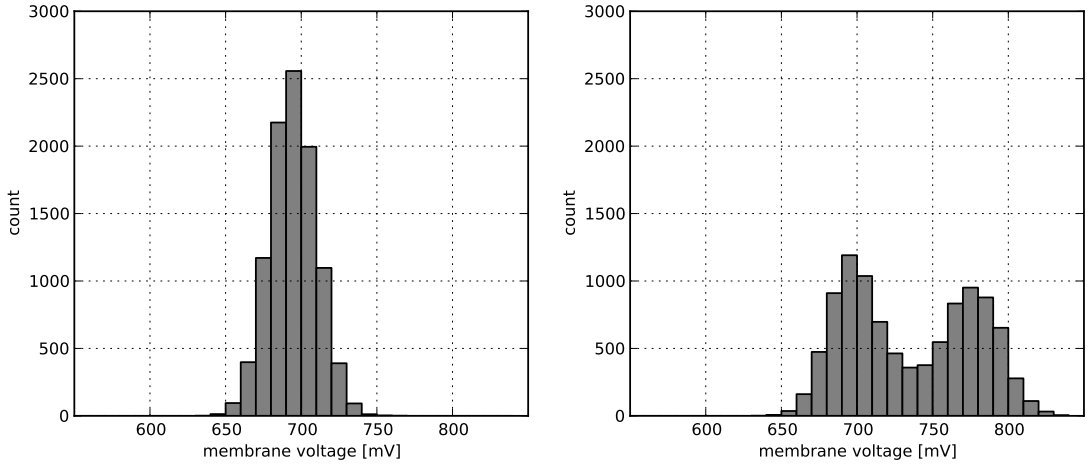
3.3.2 Excitations of the Neuron Model

For comparison with measurements the membranes reaction to step current stimuli was simulated for the different Monte Carlo corners. By giving a step current onto the membrane like it is possible in the hardware setup the voltage across the membrane capacitor V_{mem} is simulated in the simplified Neuron setup. The result can be seen in Figure 3.20b. The uncalibrated membrane potential exhibits different time constants for the different Monte Carlo corners. Also visible are the different voltage levels on which the excitation happened, they correspond to the uncalibrated membrane.

As final simulation the measurable effect the synaptic input has on the simplified neuron setup was simulated. This is possibly the most important measurement for comparison with measurements on the real hardware, all the hardware effects discussed before can be seen here in an indirect way by simulating the membrane potential. The membrane potential is the only quantity in the current neuron that can be read out.

For this simulation it is interesting to quantify how large the excitation the synaptic term can exert on the membrane can be for single pulses of different height or time constant. Figure 3.21 shows the effect the pulse height has on a single post synaptic potential. The default settings shown in Figure 3.16a are used, apart from the stimulus frequency, that is lowered to 10kHz resulting in a single excitation being simulated and the two different synaptic potentials $E_{syni} = 600\,\mathrm{mV}$ and $E_{synx} = 800\,\mathrm{mV}$. Excitatory and Inhibitory terms are shown in Figure 3.21a and Figure 3.21b. The effects discussed before are visible: between 4 and $8\,\mu A$ the membrane voltage shows no real difference to the excitation, as OTA_1 is already saturated. Reason for this is the already discussed capacitor that is to small to cater for the large currents, given the limited linear range of the OTA.

The influence of the resistor control voltage V_{syntc} is shown in Figure 3.22. As stated before, the synaptic time constants functional dependence (there cannot be spoken of a single time constant in the current implementation) is sensitive to changes of V_{syntc}. For settings below 1.3 V the influence on the membrane potential is negligible, as the synaptic excitation decays so fast that no measurable current can flow onto the capacitor. For V_{syntc} greater than 1.5 V the time constant is so large that the voltage decays on a inappropriate timescale (see exponential dependence of the time constant in Figure 3.3).



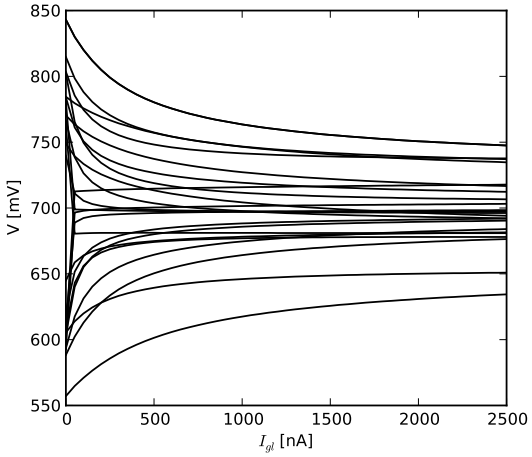
(a) $I_{convi} = 0\,\mathrm{nA}$, $I_{convx} = 0\,\mathrm{nA}$

(b) $I_{convi} = 0\,\mathrm{nA}$, $I_{convx} = 2500\,\mathrm{nA}$

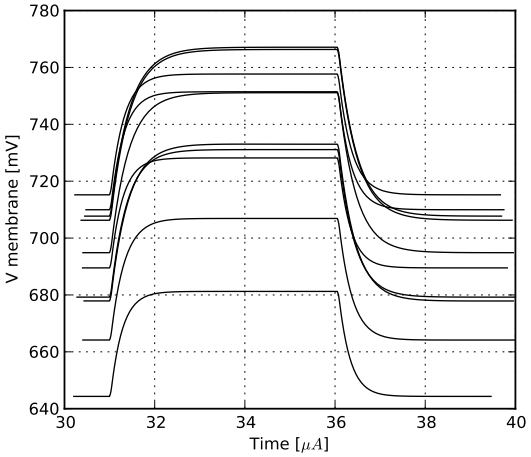
(c) $I_{convi} = 2500\,\mathrm{nA}$, $I_{convx} = 0\,\mathrm{nA}$

(d) $I_{convi} = 2500\,\mathrm{nA}$, $I_{convx} = 2500\,\mathrm{nA}$

Figure 3.19: DC operating point of the membrane voltage for the different I_{conv} settings given below the figures. Each time N = 10000 Monte Carlo samples are drawn. The influence of the synaptic reversal potentials that is visible here is reduced for large I_{gl}, these plots illustrate the strength of the offset current that is directly plotted in Figure 3.15

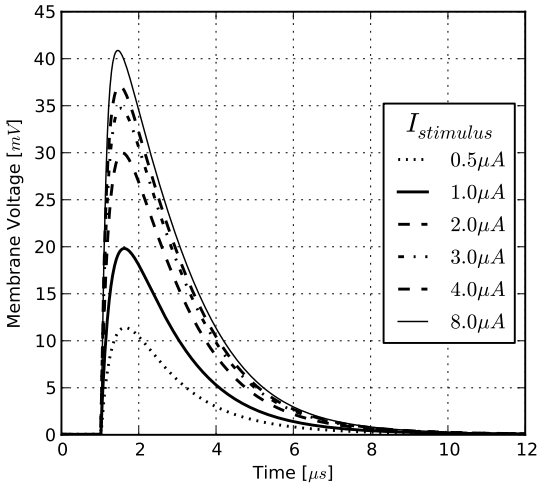


(a) Membrane voltage plotted versus I_{gl}. It can be seen that some of the neurons do not converge to $E_{syn} = 700mV$, but remain closer to the synaptic reversal potentials that are set to 800 and 600 mV

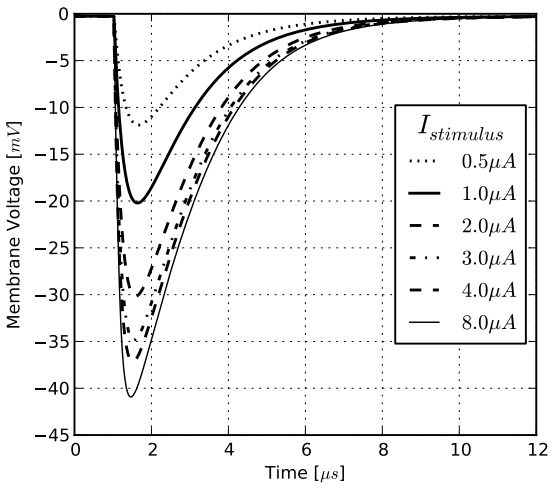


(b) Reaction of the membrane potential to current pulses in the simplified neuron simulation.

Figure 3.20: Monte Carlo Corner cases of the simplified neuron setup. Each of the traces corresponds to a single instance of the neuron on chip.



(a) Excitatory



(b) Inhibitory

Figure 3.21: Excitatory (a) and inhibitory (b) synaptic terms are stimulated with a single synaptic current event of variable height $I_{stimulus}$. The Voltage trace is normed to a resting potential of 696 mV. Saturation of the height sensitivity is seen for currents of 4 to 8 μA. The terms are very much symmetric for these single events, as the voltage swings are not large enough to catapult them out of the linear range seen in Figure 3.8a.

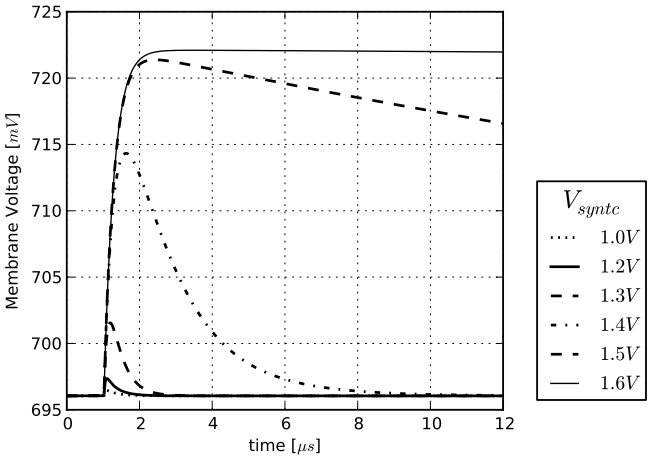


Figure 3.22: Membrane voltage trace for different settings of V_{syntc} using a single synaptic event with $I_{stimulus} = 1\,\mu$A. The small range for V_{syntc} is visible: only settings between 1.3 and 1.5 V produce measurable post synaptic potentials.

4 Discussion

In this thesis the synaptic input circuit of the neuron circuit was analyzed in detail. Initially the behavior of the synaptic input circuit was evaluated on a theoretical basis using ideal elements. Ranges the circuit should be able to cover in terms of time constants were specified alongside the model the circuit should emulate. An equivalence between the function of the circuit and the model equations it was built to emulate was successfully shown in theory and simulation, again using ideal elements. Following on these general considerations the circuit was simulated on a transistor level. In the various simulations three major problems of the synaptic input were found using transistor level simulations of the synaptic input circuit. First to mention is the limited dynamic range of the synaptic input circuit concerning the weight of the synaptic events. This was simulated in section 3.2.3. The voltage step caused by a single synaptic event, using the maximum weight setting (tuning the output current as high as possible) for the output of the synapse, is enough to saturate the synaptic input circuit. Even when using the minimum available weight of the output current the synaptic input saturates after a few spikes in the current setup. The second major problem of the circuit is the highly sensitive time constant that is not constant in the operation of the circuit. Deviations of the time constant from ideal behavior are caused by the resistive element, as the capacitor is comparatively ideal. The variability of the time constant can be seen in Figure 3.2. There the resistive element exhibited variations of the synaptic time constant over two orders of magnitude within 400 mV differential voltage. This is the range in which the OTA proves to be sensitive to the integrated voltage. Combined with the first imperfection of the limited range, a single event with maximum amplitude is able to saturate the synaptic input and lower its time constant by two orders of magnitude. It can therefore not be spoken of a single time constant, but rather of a time constant function, for further detail, see section 3.1.1. As the third major point there is a permanent current out of the synaptic input onto the membrane for some neurons, see section 3.3. This current is caused by an offset in the synaptic input that can not be calibrated for. Individual offsets of amplifier and OTA are the reason for this combined offset, see sections 3.1.3 and 3.1.4.

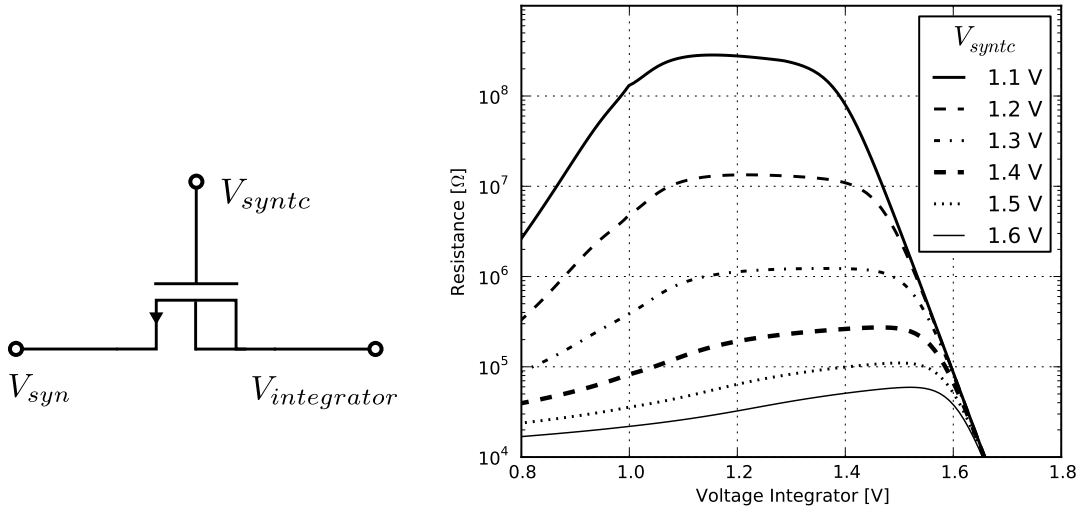
Especially the impossibility to set a certain synaptic time constant (see section 3.1) and the limited dynamic range of the circuit (see section 3.2) are severe problems. These restrict the range of models that can be emulated by the wafer scale system, they produce surprising behaviour due to saturation and abrupt changes in the time constant, see section 3.2.3. These problems could possibly be overcome by a revision of the synaptic input circuit, for proposed changes see the outlook in chapter 5. The revised circuit would need to feature a bigger capacitor and a more linear resistor. Due to changes in these two elements, a revision of the amplifier seems to be advisable as well. As tests for the

proposed changes are missing, they can not be treated as solutions. For the resistor there is concern about its implementability as the models for the transistors may not cover the special mode of operation correctly, for the capacitor the effects on the amplifiers behavior need to be simulated carefully. Further investigation of the suggestions is needed.

5 Outlook

Here suggestions for possible improvements of the membrane circuit are made and a roadmap for their implementation is proposed. The fast saturation of the synaptic input circuit can directly be countered by using a bigger capacitor in the integrator circuit. The capacitors possible size is given by the available area to implement it. To make this area available, one or two sections of 4 synapse rows each could be deleted [20]. This deletion could make additional capacitances of 1.3 or 2.6 pF available. Initial simulation results using these bigger capacitances show, that the amplifier would need to be adjusted to be able to drive the bigger capacitor. Increasing the bias current of the amplifier to 10 μA (the current maximum possible setting is 2.5 μA) counters the problem only partially. Starting from the enlargement of the capacitor, different ranges for the resistor would be needed to satisfy the benchmarks for synaptic time constants given in chapter 2. As the capacitor would be one order of magnitude larger, the resistances would need to be smaller by the same factor. This results in needed resistances of 0.06 - 4 MΩ. For the current implementation this would result in a more linear behavior with respect to the differential voltage over the integrator, see Figure 3.2b. Different implementations were tested to provide a more linear resistor. A simplistic derivate of the concept presented in [2] seems to provide a very linear resistor. The concept uses two transistors with a bulk drain connection in series to achieve resistance values on the scale of gigaohms. Here we would employ only one resistor and bias it to be in the non ultra-high-resistance regime, see the low resistivity region in [2, figure 1]. To achieve this the source needs to be at a lower potential than the drain in the NMOS implementation shown in Figure 5.1a. Simulation results are shown in Figure 5.1b. It can be seen that the resistor is tune able over a range just as wide as the implementation analyzed in section 3.1.1 while being much more linear. The fact that the resistance breaks down at differential voltages above 400 mV is irrelevant for the circuit, as OTA_1 of the synaptic input is in saturation for these voltages anyway, see section 3.1.4. To varify the correct operation, the resistor would need additional testing. For this the tests done on the current implementation in this thesis could be employed. The simulation results shown in Figure 5.1 have to be questioned, as the mode of operation is not a typical one (bulk and source are usually never connected). Therefore the concept will need further investigation.

For the offset cancellation an additional voltage available in the bias voltage generation on the HICANN (two floating gate biases where verified to be unconnected per neuron, one is available per synaptic input circuit). The additional voltage could be used to use separate voltages V_{syn1} and V_{syn2} for generating the reference voltages of amplifier and OTA. This scheme allows for offset compensation to the level of the floating gates precision. A better solution would be to use a voltage common to both terminals and adjust this voltage for one of the terminals. At one terminal the voltage would remain



(a) Schematic of the proposed alternative resistor

(b) Large signal resistance for different settings of V_{syntc}. Dimensions of transistor: W = 2.5 μm, L = 4 μm.

Figure 5.1: Suggested alternative resistor. The basic tests of the functionality shows an improved linear range compared to Figure 3.2. This implementation would still be very sensitive to changes in V_{syntc}

unchanged at the level V_{syn}, at the other terminal it would be changed to $V_{syn} + \delta_{syn}$, with δ_{syn} on the scale of 50 mV. With this scheme the offset could be canceled to a better than the floating gates precision. No implementable way of cancelling the offset in this elaborate way is found to the present state. But this cancellation scheme of the offset would be dangerous to implement as well: it adds complexity to the circuit and requires a further calibration step to make the circuit usable. As the current that the synaptic input circuit generates can not be monitored directly on the chip, calibration would be a challenge. This redesign option is not suggested, as the risks of this are estimated to be more severe than the possible profits, as long as no carefully tested option is available.

Bibliography

[1] The Human Brain Project. URL: https://www.humanbrainproject.eu/. [online; accessed: 2014-07-09].

[2] E.J. Brauer A. Tjalli, Y. Leblebici. Implementating ultra-high-value floating tunable cmos resistors. *ELECTONIC LETTERS*, 44(5), 2008.

[3] BrainScaleS. Research. `http://brainscales.kip.uni-heidelberg.de/public/index.html`, 2012.

[4] R. Brette and W. Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.*, 94:3637 – 3642, 2005.

[5] J.A. Babcock et al. Analog characteristics of metal-insulator-metal capacitors using pecvd nitride dielectics. *IEEE Electron device letters*, 22(5), 2001.

[6] M. Marin et al. Modeling the mismatch of high-k mim capacitors. *IEEE International Conference on Microelectronic Test Structures*, 2007.

[7] R.L. Geriger and E. Sánchez-Sinencio. Actice Filter Design Using Operational Transconductance Amplifiers: A Tutorial. *IEEE Circuits and Devices Magazine*, 1:20 – 32, March 1985.

[8] Wulfram Gerstner and Werner Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, 2002.

[9] A. Grübl. Personal communication, 2014.

[10] M. Güttler and D. Husmann de Olivera. personal communication, 2014.

[11] HBP SP9 partners. *Neuromorphic Platform Specification*. Human Brain Project, March 2014.

[12] M. Hock. *Modern Semiconductor Technologies for Neuromorphic Hardware*. PhD thesis, Ruprecht-Karls University Heidelberg, July 2014.

[13] Chenming Hu, Weidong Liu, and Xiaodong Jin. *The BSIM3v3.2 MOSFET Model*, Dec 1998.

[14] G. Kiene. Transistor-level simulations of the adex neuron model. Internship report, University of Heidelberg, 2014.

[15] Y. Tsividis M. Banu. Floating Voltage-Contolled Resistors in COMS Technology. *ELECRONIC LETTERS*, 18(15):678–679, July 1982.

[16] Timothée Masquelier and Gustavo Deco. Network bursting dynamics in excitatory cortical neuron cultures results from the combination of different adaptive mechanism. *PloS one*, 8(10):e75824, 2013.

[17] C. A. Mead. Neuromorphic electronic systems. *Proceedings of the IEEE*, 78:1629–1636, 1990.

[18] Sebastian Millner. *Development of a Multi-Compartment Neuron Model Emulation*. PhD thesis, Ruprecht-Karls University Heidelberg, November 2012.

[19] Mihai A Petrovici, Johannes Bill, Ilja Bytschok, Johannes Schemmel, and Karlheinz Meier. Stochastic inference with deterministic spiking neurons. *arXiv preprint arXiv:1311.3211*, November 11th, 2013.

[20] J. Schemmel. Personal communication, 2014.

[21] J. Schemmel, J. Fieres, and K. Meier. Wafer-scale integration of analog neural networks. In *Proceedings of the 2008 International Joint Conference on Neural Networks (IJCNN)*, 2008.

[22] D. Schmidt. Personal communication, 2014.

[23] Tim P Vogels and L F Abbott. Signal propagation and logic gating in networks of integrate-and-fire neurons. *J Neurosci*, 25(46):10786–95, Nov 2005.

[24] BrainScales Website. Internal Section, Wafer Status, Temperature Info. [online; accessed: 2014-07-15].

Statement of Originality (Erklärung):

I certify that this thesis, and the research to which it refers, are the product of my own work. Any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, August 1, 2014

.....
(signature)