Fault-Tolerant Electronic Devices Using Artificial Neural Network Paradigm For Safety-Critical Devices And Nanoelectronic Devices

Microelectronic Systems Laboratory
Swiss Federal Institute of Technology (EPFL)

Student: Neil Joye
Project Supervisors: Prof. Y. Leblebici, A. Schmid and Prof. T. Asai (Hokkaido University, Japan)
Acknowledgement

I would like to take this opportunity to thank Professor Asai, Professor Leblebici and the scholarship of the EPFL for the financial support without which this project could not have been carried out in Hokkaido University.

I would also like to thank all the members of the Laboratory of Advanced LSI Engineering in Hokkaido University for their hospitality and support during my stay in Sapporo. And finally, I would like to thank Alexandre Schmid for his very useful technical advice.
# Table of Contents

**Introduction** ....................................................................................................................... 1

**Part 1: Development of a Fault-tolerant Boolean** .............................................................. 2
1. Introduction......................................................................................................................... 3
2. Theory ............................................................................................................................ 4
   2.1. Fault-tolerance using Artificial Neural Networks ..................................................... 4
   2.2. Learning algorithm ............................................................................................... 6
   2.3. Physical implementation ....................................................................................... 7
      2.3.1. Spiking neuron .............................................................................................. 7
      2.3.2. Synapses ...................................................................................................... 8
3. Development of the spiking neuron and the synapses ...................................................... 10
   3.1. Minimum gate length of the transistors ................................................................ 10
   3.2. Spiking neuron .................................................................................................... 10
   3.3. Synapses ............................................................................................................. 12
      3.3.1. Excitatory synapses with one input .......................................................... 12
      3.3.2. Variation of the transistors threshold voltage .......................................... 16
      3.3.3. Influence of the input transistor of the integrate-and-fire neuron .................. 16
      3.3.4. Excitatory synapses with two inputs ...................................................... 18
      3.3.5. Inhibitory synapses with one input ........................................................ 21
      3.3.6. Inhibitory synapses with two inputs ....................................................... 24
      3.3.7. Excitatory and inhibitory synapses with one input .................................. 24
      3.3.8. Surface of the different synapses ............................................................ 25
4. Fault-tolerant Boolean gates ............................................................................................. 27
   4.1. NOT gate ............................................................................................................ 28
   4.2. NOR gate ............................................................................................................ 30
   4.3. NAND gate ........................................................................................................ 31
5. Conclusion ...................................................................................................................... 33

**Part 2: Depressing Synapses and their Application to Synchrony Detection and**
**Contrast-Invariant Pattern Classification** ........................................................................... 34
1. Introduction .................................................................................................................... 35
2. Depressing synapse ...................................................................................................... 36
3. Depressing synapse with application to synchrony detection ....................................... 39
4. Depressing synapse with application to contrast-invariant pattern classification ........... 43
5. Conclusion .................................................................................................................. 47

**Conclusion** .................................................................................................................... 48

**References** ..................................................................................................................... 49

**Appendix 1**
Initial indications of the project: administrative data, content and foreseen milestones

**Appendix 2**
Paper (abstract) submitted for possible publication in the Ninth International Conference
on Cognitive and Neural Systems (ICCNS) 2005
Table of Figures
Part 1 – Development of a Fault-tolerant Boolean Gates Library

Figure 1: Logic gate built from a feed-forward artificial neural network ..................................4
Figure 2: Representation of a library of fault-tolerant XOR gates, each one having a
different degree of degree of defect density it can support .............................................5
Figure 3: Implementation of the learning algorithm with a software program ......................6
Figure 4: Implementation of the learning algorithm with SPICE ..........................................6
Figure 5: Circuit of the integrate-and-fire neuron [3] ............................................................7
Figure 6: Excitatory and inhibitory synapses ...................................................................8
Figure 7: Circuits of both excitatory and inhibitory synapses with one input each ..........8
Figure 8: Excitatory and inhibitory synapses without thee complementary input signal
been available ..................................................................................................................9
Figure 9: Excitatory and inhibitory synapses using a differential pair ..............................9
Figure 10: Architecture of the integrate-and-fire neuron .................................................10
Figure 11: Output signals of the three inverters of the integrate-and-fire neuron ............11
Figure 12: Types of synapses that will be developed ......................................................12
Figure 13: Architecture of the excitatory synapse with one input ...................................12
Figure 14: Maximum drain current of the input transistor
(maximum frequency and L_IN = 2 µm) ...........................................................................14
Figure 15: Architecture of the circuit in order to measure the properties of the excitatory
synapse with one input ..................................................................................................14
Figure 16: Input pulse .....................................................................................................15
Figure 17: (Measured output – Theoretical output) of an excitatory synapse with
one input .............................................................................................................................15
Figure 18: Architecture of the circuit in order to measure if the output pulses of a spiking
neuron are still decoded correctly by other synapses and neurons ...........................15
Figure 19: Proposed layout technique to minimise the variance of the threshold voltage ....16
Figure 20: Output current of the excitatory synapse for an input equal to 1 and
a synaptic weight equal to 1 .........................................................................................16
Figure 21: Output current of the excitatory synapse for an input equal to 1 and
a synaptic weight equal to 0.755 ..................................................................................17
Figure 22: Architecture of the excitatory synapse with two inputs .................................18
Figure 23: Architecture of the circuit in order to measure the properties of the excitatory
synapse with two inputs ...............................................................................................18
Figure 24: (Measured output – Theoretical output) of an excitatory synapse with
two synchronous input signals equal to 1 .................................................................19
Figure 25: (Measured output – Theoretical output) of an excitatory synapse with
two asynchronous input signals equal to 1 ...............................................................19
Figure 26: Vdd-EPSP with synchronous and asynchronous inputs ..............................20
Figure 27: EPSC with synchronous input signals.
The inputs and the weights are equal to 1 .................................................................20
Figure 28: EPSC with synchronous input signals.
The inputs and the weights are equal to 1 .................................................................21
Figure 29: Architecture of the inhibitory synapse with one input ..................................21
Figure 30: (Measured output – Theoretical output) of an inhibitory synapse with
one input ...........................................................................................................................22
Figure 31: Output current of the inhibitory synapse with a gate width of the spiking
neuron input transistor equal to 4.6 µm .................................................................23
Figure 32: Output current of the inhibitory synapse with a gate width of the spiking
neuron input transistor equal to 50 µm ..................................................23
Figure 33: Architecture of the inhibitory synapse with two inputs .........................24
Figure 34: Excitatory synapse with one input. (a) Distribution of the surface between the neuron and the synapse. (b) Distribution of the surface between the different components of the synapse.................................................................25
Figure 35: Excitatory synapse with two inputs. (a) Distribution of the surface between the neuron and the synapse. (b) Distribution of the surface between the different components of the synapse.................................................................25
Figure 36: Inhibitory synapse with one input. (a) Distribution of the surface between the neuron and the synapse. (b) Distribution of the surface between the different components of the synapse.................................................................25
Figure 37: Inhibitory synapse with two inputs. (a) Distribution of the surface between the neuron and the synapse. (b) Distribution of the surface between the different components of the synapse.................................................................26
Figure 38: The proposed fault-tolerant architecture based on multiple layers ..............27
Figure 39: (a) Excitatory synapse connected to a spiking neuron
(b) Inhibitory synapse connected to a spiking neuron......................................28
Figure 40: Architecture of the NOT gate using a minimum number of synapses and spiking neurons .................................................................28
Figure 41: Fault-tolerant architecture of the NOT gate with two redundant units ........28
Figure 42: Fault-tolerant architecture of the NOT gate with three redundant units .........29
Figure 43: Probability of correct operation of the NOT gate against the number of faulty devices with two to five redundant units.........................................................29
Figure 44: Probability of correct operation of the NOT gate against the density of faulty devices with two to five redundant units.........................................................30
Figure 45: Architecture of the NOR gate using a minimum number of synapses and spiking neurons .................................................................30
Figure 46: Fault-tolerant architecture of the NOR gate with two redundant units ..........30
Figure 47: Simplest architecture of the NAND gate ................................................31
Figure 48: Architecture of the NAND gate ..........................................................31
Figure 49: Fault-tolerant architecture of the NAND gate with two redundant units ........31
Figure 50: Probability of correct operation of the NAND gate against the number of faulty devices with two to five redundant units.........................................................32
Figure 51: Probability of correct operation of the NAND gate against the density of faulty devices with two to five redundant units.........................................................32
Part 2 – Depressing Synapses and their Application to Synchrony Detection and Contrast-Invariant Pattern Classification

Figure 52: Architecture of the depressing synapse………………………………………….36
Figure 53: Architecture for the circuit in order to simulate the properties of the depressing synapse…………………………………………………………………….37
Figure 54: Experimental results of the depressing synapse circuit; (a) successive spike inputs, (b) the degree of synaptic depression, and (c) its outputs………………37
Figure 55: Changes in amplitude of the output of the depressing synapse circuit against the firing rate of the presynaptic neuron……………………………………38
Figure 56: Representation of a burst input signal…………………………………………..39
Figure 57: Response of the EPSP for single burst input via nondepressed synapse circuit ….40
Figure 58: Response of the EPSP for single burst input via depressed synapse circuit …..40
Figure 59: Responses of the EPSP for asynchronous burst inputs via nondepressed synapse circuit………………………………………………………………….41
Figure 60: Responses of the EPSP for asynchronous burst inputs via depressed synapse circuit………………………………………………………………….41
Figure 61: Responses of the EPSP for synchronous burst inputs via nondepressed synapse circuit………………………………………………………………….41
Figure 62: Responses of the EPSP for synchronous burst inputs via depressed synapse circuit………………………………………………………………….42
Figure 63: Changes in the EPSP against the number of active presynaptic neuron and their firing rates using nondepressing synapses ……………………………….44
Figure 64: Changes in the EPSP against the number of active presynaptic neuron and their firing rates using depresssing synapses ……………………………….44
Figure 65: Changes in the EPSP against the number of active presynaptic neuron and their firing rates. (a) 1 active presynaptic neuron, (b) 2 active presynaptic neurons, (c) 3 active presynaptic neurons and (d) 4 active presynaptic neurons…45
Figure 66: Results for dependence of the integrate-and-fire neuron on their firing rate of presynaptic neurones (4 neurones) ……………………………………………………45
Figure 67: Large-scale simulation results (100 neurons) for the same experiments shown in Figure 66 ……………………………………………………………………….46
Figure 68: The leaky integrate-and-fire neuron proposed by Bugmann for pattern recognition [8]………………………………………………………………….46
Introduction

Results from neuroscience research confirm the stupefying computational power of our brain. For example, complex tasks like face recognition are carried out in real time using minimal power, thus beating any available digital computer. Besides, the human brain is extremely robust and immune to component failure. In fact, where a single component failure could result in a complete system breakdown in a digital processor, the neurobiological system can re-adapt itself and continue to operate correctly with no apparent loss of performance. Therefore, using neurobiological models to develop microelectronic devices could offer original solutions to problems like power consumption or fault-tolerance that are becoming central issues for current and future circuit fabrication technologies.

These advantageous properties also underline the fact that neurobiological systems process information in a completely different way compared to standard digital processors. First of all, the information is represented by relative values of analog signals rather than by absolute values of digital signals. Secondly, the computation in neurobiological systems takes place in real and continuous time. This means that no clock is needed. Typically, the signals that are used are spikes and the information is coded in their frequencies and timings for example. Finally, processing at the neural level seems very slow compared to the operating frequencies of digital processors. However, massive parallelism of the architecture compensates this lack of speed. Considering these facts, a new approach for developing electronic devices based on neurobiological models must be quite different from the usual digital approach.

In this context, this project consists in implementing some of the advantageous properties of neurobiological systems into microelectronic devices. Furthermore, spiking neurons will be used in this project since they are able to encode information in their spikes frequencies and timings. The project is divided into two parts. The first one is a study on the implementation of fault-tolerance into electronic devices. The goal is to develop a library of fault-tolerant Boolean gates using artificial neural network paradigm. The second part studies an application of the neurobiological model that is synchrony detection using depressing synapses.
Part 1

Development of a Fault-Tolerant Boolean Gates Library
1. Introduction

The ITRS technology roadmap predicts that semiconductor device dimensions will shrink to lower than 30 nm by the end of this decade. While technological feasibility of scaling-down the device dimensions is not disputed from the manufacturing point of view, such scaling is expected to result in a multitude of serious challenges at the circuit and system level, especially in terms of increased leakage currents, reduced power supply voltages, degraded reliability and increased power density. Therefore actual fault-tolerant design methods of VLSI circuits, which have traditionally been addressed at system level, involving algorithmic adaptation, block-level redundancy and majority voting as the main tools, will not be sufficient for future very-deep submicron CMOS devices where serious degradation of reliability is expected. Therefore, to implement robustness and fault-tolerance into these devices, new design approaches will need to be considered at low level of abstraction.

According to this fact, it was already shown in previous work [1] that compensation of faults such as process variations or device mismatch can be achieved by implementing multi-layer feed-forward artificial neural networks. The goal in this project is to implement this concept at hardware circuit-level using spiking neurons. Indeed, this type of neurons can easily encode analog signals used by neural networks as mean frequencies of impulses. Besides, excitatory and inhibitory synapses will be used. This means that the artificial neural network model that is used is a closer representation of the neurobiological system compared to mathematical models.
2. Theory

2.1. Fault-tolerance using Artificial Neural Networks

In a standard CMOS Boolean gate, if only one transistor doesn’t operate correctly, the gate doesn’t work anymore. As already said before, this problem can be solved by using artificial neural networks. Indeed, these networks can be trained to be tolerant to faults which may occur as a result of transient artefacts or permanent damage by choosing the correct synaptic weights previously calculated by a learning algorithm. Therefore, the idea is to construct an artificial neural network that acts like a Boolean gate as shown in Figure 1. Then if the correct synaptic weights have been chosen, the network can support an error and still operate as the desired logic function. It has to be noticed that the more neurons a network has, the more errors it can support. It is this inherent neuron redundancy and appropriate training that give the fault-absorbing property to the neural network. So a compromise has to be done between the maximum defect density a Boolean gate can support and its size.

![Logic gate built from a feed-forward artificial neural network.](image)

*Figure 1. Logic gate built from a feed-forward artificial neural network.*

Spiking neurons are employed because they can easily encode analog signals used by neural networks as mean frequencies of impulses. Furthermore, a maximum and minimum frequency respectively encodes a logic 1 and logic 0. These two values will be determined after having performed some simulations of the neuron. It also has to be noticed that the synaptic weights have values between zero and one.

Using this approach, the final goal is to build a library of logic gates, each one with a different maximum degree of defect density allowing correct operation. Then this library of fault-tolerant Boolean gates could allow the synthesis of complex systems using the well-established conventional design-flow applied for digital CMOS architectures while keeping all device-level development to enhance robustness and fault-tolerance largely transparent to the designer. Figure 2 shows an example of a library of fault-tolerant XOR gates, each one having a different degree of defect density that it can successfully absorb.
Two main disadvantages arise from utilizing artificial neural network to build logic gates:

- **Surface** – As it will be shown in Section 3.3.8, an integrate-and-fire neuron has nine transistors and a synapse has three transistors. Building logic gates using these elements will lead to circuits with many more transistors compared to the classical CMOS logic gates.

- **Power** – The fact that a large number of transistors are used to develop logic gates means that the power consumption will also increase.

Power consumption and silicon surface are central issues that cannot be neglected when conceiving electronic devices. However, for a safety-critical application, the most important aspect that has to be considered is the robustness of a circuit. For example, a device that has to be installed on a satellite should be very reliable. If it fails to operate because of a transient artefact or permanent damage, it is impossible to replace it, thus leading the satellite to be unusable. Therefore, it is preferable to employ larger devices that consume more power if the final circuit is able to absorb some faults.

Figure 2. Representation of a library of fault-tolerant XOR gates, each one having a different degree of defect density it can support.
2.2. Learning algorithm

The spiking neuron model that will be applied in this work has a closer representation of the neurobiological system compared to mathematical models which are typically implemented in computer programs. This means that learning algorithms like backpropagation cannot be applied to adapt weight values. However, a learning algorithm like weight perturbation [2] may be adapted for excitatory and inhibitory synapses which are employed in this project.

Then to implement this learning algorithm which calculates the correct synaptic weights, two solutions can be used. The first one is to apply a software program, for example in C or Matlab, as shown in Figure 3. The critical point is to find a mapping between the synaptic weights calculated by the learning algorithm and a physical parameter in the circuit, for example the size of a transistor. This step will be explained in Section 3.3.1.

Then to implement this learning algorithm which calculates the correct synaptic weights, two solutions can be used. The first one is to apply a software program, for example in C or Matlab, as shown in Figure 3. The critical point is to find a mapping between the synaptic weights calculated by the learning algorithm and a physical parameter in the circuit, for example the size of a transistor. This step will be explained in Section 3.3.1.

The second solution is to use SPICE as shown in Figure 4. First the extraction of some useful parameters of the circuit has to be done. Then a small software program executes one step of the learning algorithm. The parameters that were extracted have new some new values that can be used for a new SPICE simulation. The cycle is then carried out until the error of the output is negligible. This solution is more precise than the first one, but it is also slower.

**Figure 3. Implementation of the learning algorithm with a software program.**

**Figure 4. Implementation of the learning algorithm with SPICE.**
2.3. Physical implementation

2.3.1. Spiking neuron

A classical spiking neuron circuit based on a model referred to as the integrate-and-fire neuron is used in this project. As the name implies, this circuit integrates the input current on a capacitor and when a threshold level is reached, the neuron “fires” and resets its input. The circuit is taken from [3] and is shown in Figure 5. The main reason for this choice among several other spiking neuron circuits is the surface of this circuit which is limited.

![Image of the integrate-and-fire neuron circuit](image)

Figure 5. Circuit of the integrate-and-fire neuron [3].

To understand circuit operation, consider the circuit condition after an output pulse has completed. In this state, $V_o$ is at ground potential and $V_c$ is lower than the switching threshold of the oscillator. The discharge path of the state capacitor $C$ is closed and the charging path of $C$ is open. The circuit remains in this state until the input current $I_i$ increases $V_c$ to the switching threshold of the oscillator which leads $V_o$ to switch to Vdd. The new value of $V_c$ is above the switching threshold of the oscillator and depends on the relative values of $C_f$ and $C$. Once a pulse begins, the discharge path of $C$ is open and the charging path of $C$ is closed. $V_p$ sets the discharge rate of $C$ and thus the width of the pulse. The circuit remains in this state until $V_c$ decreases to the switching threshold of the oscillator. Then $V_o$ switches to ground potential and the pulse is completed. $V_c$ is smaller than the switching threshold of the oscillator and depends on the relative values of $C_f$ and $C$.

It has to be noticed that the oscillator is made of two inverters connected in series.
2.3.2. Synapses

Two types of synapses will be used, namely excitatory and inhibitory synapses. These synapses are shown in Figure 6.

![Figure 6. Excitatory and inhibitory synapses.

EPSP = Excitatory postsynaptic potential. IPSP = Inhibitory postsynaptic potential. EPSC = Excitatory postsynaptic current. IPSC = Inhibitory postsynaptic current.](image)

The goal of these synapses is to control the input current of a spiking neuron (see section 2.3.1). Excitatory synapses are responsible for its increase and inhibitory synapses for its decrease. Furthermore, the inputs of these synapses are the outputs of other spiking neurons. These inputs are therefore represented as voltages. Consequently, the synapses have to undertake a voltage to current conversion.

The circuit implementation that was chosen (Figure 7) is adopted from [4] principally because the number of components is small.

![Figure 7. Circuits of both excitatory and inhibitory synapses with one input each.](image)
Each synapse (excitatory and inhibitory) is composed of a current mirror with a capacitor which prevents the output transistor to leave the saturation mode. Besides, to be able to convert the input voltage to the current output, an input transistor has to be added compared to the circuit from [4]. Therefore, the β factor of this transistor encodes the synaptic weight (this will be explained more in details in section 3.3.1).

To be able to use the architecture shown in Figure 7, the complementary input signal has to be available for the inhibitory synapse. Therefore, a third inverter as been added to the oscillator of the spiking neuron (Figure 5). If this signal is not available, a current mirror has to be added to the inhibitory synapse as shown in Figure 8.

![Figure 8. Excitatory and inhibitory synapses without the complementary input signal been available.](image)

The synaptic physical implementation of the synapses can also be obtained by using a differential pair as shown in Figure 9. This solution has not been selected because of the large size of its transistors.

![Figure 9. Excitatory and inhibitory synapses using a differential pair.](image)
3. Development of the spiking neuron and the synapses

The circuits showed in the previous Section will be implemented in AMS 0.35 \( \mu \text{m} \) CMOS technology. The power supply voltage Vdd is equal to 3.3 V.

3.1. Minimum gate length of the transistors

In order to keep the output resistance of transistors as high as possible (so that the drain current doesn’t depend too much on the source and drain voltages) and to reduce the effects of channel length and mobility modulation, the gate length of the transistor must be large enough. Therefore, a general design rule is to set the gate length in analog applications to two to five times the minimum gate length (0.35 \( \mu \text{m} \) in this case).

After having measured the drain current against the drain voltage with different gate length and gate voltages, the minimum gate length will be set to 1.60 \( \mu \text{m} \) in order to prevent the undesirable effects mentioned above to happen.

3.2. Spiking neuron

![Figure 10. Architecture of the integrate-and-fire neuron.](image)

The complementary signal of the output spike is needed for the inhibitory synapse. Although this signal is available at the output of the first inverter (\( T_{4p} \) and \( T_{4n} \)), a third inverter (\( T_{6p} \) and \( T_{6n} \)) has been added to the oscillator as shown in Figure 10. This is due to the fact that the rise and fall time of the output signal of the first inverter are to long. This property is shown in Figure 11 which represents the output signals of the three inverters.
The maximum frequency will be set to 100 MHz (period of 10 ns) which represents logic 1, and the width of a pulse will be 2 ns. This is equivalent to a maximum duty cycle of 20%. Above this value, the relation between the input current of the spiking neuron and the output frequency is not linear. The width of the pulses can be assigned with the value of the voltage $V_p$. In this project, $V_p$ will be fixed to $V_{dd}$ (which corresponds to a minimum pulse width).

The sizes of the components in the integrate-and-fire neuron are displayed on Table 1. The gate lengths are equal to 1.60 $\mu$m.

<table>
<thead>
<tr>
<th>Components</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>50 fF ($10 \mu$m x 5.50 $\mu$m)</td>
</tr>
<tr>
<td>$C_f$</td>
<td>10 fF ($10 \mu$m x 0.95 $\mu$m)</td>
</tr>
<tr>
<td>$T_1$</td>
<td>$W_1 = 4.6 \mu$m</td>
</tr>
<tr>
<td>$T_2$</td>
<td>$W_2 = 4.6 \mu$m</td>
</tr>
<tr>
<td>$T_3$</td>
<td>$W_3 = 4.6 \mu$m</td>
</tr>
<tr>
<td>$T_{4p}$</td>
<td>$W_{4p} = 3 \mu$m</td>
</tr>
<tr>
<td>$T_{4n}$</td>
<td>$W_{4n} = 3 \mu$m</td>
</tr>
<tr>
<td>$T_{5p}$</td>
<td>$W_{5p} = 10 \mu$m</td>
</tr>
<tr>
<td>$T_{5n}$</td>
<td>$W_{5n} = 1 \mu$m</td>
</tr>
<tr>
<td>$T_{6p}$</td>
<td>$W_{6p} = 8 \mu$m</td>
</tr>
<tr>
<td>$T_{6n}$</td>
<td>$W_{6n} = 1 \mu$m</td>
</tr>
</tbody>
</table>

Table 1. Sizes of the components in the integrate-and-fire neuron.

The values of the capacitors have been chosen very small in order to have a circuit with an acceptable size. As explained in following Sections, the sizes of the capacitors can become very large. Therefore, in order to have a small circuit, it is principally the size of the capacitors that must be decreased. As for the transistors $T_1$, $T_2$ and $T_3$, their sizes were settled in order to have a correct operation of the spiking neuron. The size of the oscillator will be explained in following sections.

Fan-out problems could appear. In this case the solution is to increase the width gate of the transistors of the oscillator.
3.3. Synapses

The goal is to develop the three types of synapses shown in Figure 12, each with at least one to three inputs.

![Figure 12. Types of synapses that will be developed.](image)

3.3.1. Excitatory synapses with one input

On Table 2 is displayed the values of the different components of the synapse circuit shown in Figure 13. The gate lengths are equal to 1.60 µm.

<table>
<thead>
<tr>
<th>Components</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_2$</td>
<td>100 fF (10 µm x 11.2 µm)</td>
</tr>
<tr>
<td>$T_a$</td>
<td>$W_a = 10 \mu m$</td>
</tr>
<tr>
<td>$T_b$</td>
<td>$W_b = 9.4 \mu m$</td>
</tr>
<tr>
<td>$T_{IN}$</td>
<td>$W_{IN} = 1 \mu m$</td>
</tr>
</tbody>
</table>

*Table 2. Sizes of the components in the excitatory synapse with one input.*
Coding of the synaptic weight

A very important aspect of this circuit is the fact that the size of the input transistor $T_{IN}$, which controls the value of the current injected in the synapse, determines the synaptic weight. To understand this point, let us consider the equation of the drain current of a transistor in the saturation mode.

$$I_D = \frac{\beta}{2n} (V_G - V_{TO} - nV_S)^2$$

(1)

$$\beta = \frac{W}{L} \mu C_{ox}$$

(2)

The drain current of a transistor depends on the gate and source voltage and the factor $\beta$ of the transistor. In the case of $T_{IN}$, the source voltage is at ground potential and the gate voltage is equal to $V_{dd}$ during an input spike (ground potential when there is not any input pulse). This means that the only way to control the drain current of the input transistor, and by the same way the current injected in the synapse, is to vary the length or the width of $T_{IN}$. Therefore in this project, it as has been decided to use the gate length to encode the synaptic weight. This choice has been made principally because it is easier to keep the current mirror in the saturation mode. It is also very important to notice that the input transistor must not leave the saturation mode in order to have the drain current independent from the drain voltage.

Considering these facts, the minimum gate length of the input transistor has been chosen to be equal to 2 $\mu$m. This value encodes a maximum weight equal to one. Then, for example, a weight of 0.5 is obtained by increasing the gate length in order to have the drain current of $T_{IN}$ two times smaller. Considering equations (1) and (2), this property can be obtained with a gate length equal to 4 $\mu$m. Therefore a relation between the synaptic weight and the gate length of the input transistor can be established.

$$L_{MIN} = 2 \mu m$$

$$L = \frac{L_{MIN}}{\omega}$$

(3)

This is one of the main aspects of this project. A direct mapping between the synaptic weight and a physical parameter of the circuit, in this case the gate length of a transistor, is accomplished.

On the other hand, a disadvantage of this approach is that the gate length of $T_{IN}$ can become very big if the synaptic weight is very small. To overcome this problem, the maximum gate length is set to 20 $\mu$m. To obtain weights smaller than 0.1, the width of the input transistor has to be decreased in order to code the correct synaptic weight.

Finally, despite the fact that transistors dissipate less power in weak inversion, it has been decided to operate the circuits in strong inversion. This choice has been made because in weak inversion the drain current depends strongly on the threshold voltage and therefore is not well controlled.
**Brief explanation of how the sizes of the transistors are chosen**

In order to acquire the correct performances of the synapse, all transistors must operate in the saturation mode. However, with the transistors sizes that are displayed on Table 1, the input transistor operates in the conduction mode if the input current (drain current of T_{IN}) is too large. Therefore this current depends on the EPSP and is not constant. The worst case, which is when the input current is maximum, is obtained with a maximum input frequency and a minimum gate length (input and synaptic weight equal to 1). In this case, the input current varies as shown in Figure 14. The difference between these two values is about 1% of the mean current. Therefore, the error that is made is negligible.

This problem could be solved if the value of C_2 is increased in order to decrease the variance of the EPSP. However, in this case the silicon area of the synapse can be very large.

![Figure 14. Maximum drain current of the input transistor (maximum frequency and \( L_{IN} = 2 \mu m \)).](image)

**Simulations of the excitatory synapse with one input**

The sizes of the transistors that are displayed in the previous Sections have been specified after having performed some simulations of these circuits. One excitatory synapse connected to a spiking neuron, as shown in Figure 15, has been simulated.

![Figure 15. Architecture of the circuit in order to measure the properties of the excitatory synapse with one input.](image)

The amplitude of the input pulse is equal to Vdd. The rise time and the fall time are equal to 0.1 ns and \( \Delta t \) equals 1.35 ns (see Figure 16). The width of the input pulse is smaller than the width of the output pulse (about 2 ns) in order to have both input and output signals with the same integrated current in a spiking neuron.
Figure 16. Input pulse.

Figure 17 shows the results that have been obtained.

![Graph showing error vs synaptic weight for different input currents](image)

Figure 17. (Measured output – Theoretical output) of an excitatory synapse with one input.

A few conclusions can be made:

- The more input current is injected in the synapse, the larger is the error due to the fact that the relative error is pretty much constant. This explains why the maximum error is nearly always obtained with an input equal to logic 1 (maximum input frequency).
- The maximum error is approximately equivalent to 7% of the maximum logic value.
- For weights between 0.5 and 1, the output is too small. However, it is the opposite for weights inferior to 0.5. These properties are due to two phenomena that will be explained in the following two Sections.

Some simulations have also been made with two spiking neurons connected in series as shown in Figure 18. This has been done to check if the output signal of the first neuron is still correctly decoded by the second neuron, which is the case because the results are the same as shown in Figure 17. It is with these simulations that the width of the input spikes has been specified.

Figure 18. Architecture of the circuit in order to measure if the output pulses of a spiking neuron are still decoded correctly by other synapses and neurons.
3.3.2. Variation of the transistors threshold voltage

The threshold voltage of a transistor slightly decreases with the increase of the gate length. Considering equation (1) on page 13, it implies that the relation between the synaptic weight and the gate length of the synapse input transistor as shown in equation (3) is not quite linear anymore. For example, a small synaptic weight will involve a slightly smaller threshold voltage of the synapse input transistor and therefore a larger injected current in the synapse as it is dictated by equation (1). This explains why the measured outputs in Section 3.3.1 are larger than they should be for small synaptic weights.

This problem could be partially solved by using the following layout technique (Figure 19). If two transistors have the same gate length and that there are in the same local bulk, the two transistors together act like only one transistor with a gate length twice as big. This technique could be used for small values of the synaptic weight.

![Figure 19. Proposed layout technique to minimise the variance of the threshold voltage.](image)

3.3.3. Influence of the input transistor of the integrate-and-fire neuron

While a spiking neuron “fires” an impulse, the charging path of the capacitor C\textsubscript{1} is closed (see Figure 10). This implies that no current is integrated during the firing of an output pulse and that the EPSC is equal to zero. This property is the main reason why the output of a neuron inclines to be too small for large synaptic weights (see Figure 17).

Figure 20 and 21 show the output current of the excitatory synapse obtained by simulations performed in section 3.3.2 (see Figure 15 to 17).

![Figure 20. Output current of the excitatory synapse for an input equal to 1 (input frequency of 100 MHz) and a synaptic weight equal to 1.](image)
Figure 21. Output current of the excitatory synapse for an input equal to 1 (input frequency of 100 MHz) and a synaptic weight equal to 0.755.

For a synaptic weight equal to 1 (Figure 20), the spiking neuron does not integrate the output current of the excitatory synapse while this current is relatively small. This property explains why the output frequency of the neuron is exactly the same as the input frequency of the synapse if the synaptic weight is equal to one. This implies that the spiking neuron adjusts the closer of its charging path in order to have an output equal to the input.

However, with a synaptic weight equal to 0.755 (Figure 21), the spiking neuron doesn’t integrate the output current of the excitatory synapse while this one could be large. This is due to the fact that the output frequency of the spiking neuron, and by the same way the frequency of the closer of its charging path, is not equal to the input frequency of the excitatory synapse. This means that the overall current that is not integrated is relatively large and explains why the measured output of the spiking neuron is too small (see Figure 17).

A solution to this problem would be to have a much larger capacitor in order to have a negligible variance of the EPSP. In this case, the output current of the synapse would be constant. This means that when $T_1$ (Figure 10) is in an off state, it is always the same amount of current that is not integrated by the spiking neuron regardless of the synaptic weight. Unfortunately, this solution has a big disadvantage. The size of the capacitor can be huge.
3.3.4. Excitatory synapses with two inputs

![Figure 22. Architecture of the excitatory synapse with two inputs.](image)

Table 3 displays the values of the different components of the synapse circuit shown in Figure 22. The gate lengths of $T_a$ and $T_b$ are equal to 1.60 $\mu$m and the synaptic weights are encoded as described in equation (3) on page 13.

<table>
<thead>
<tr>
<th>Components</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_2$</td>
<td>300 fF (10 $\mu$m x 34 $\mu$m)</td>
</tr>
<tr>
<td>$T_a$</td>
<td>$W_a = 10 \mu$m</td>
</tr>
<tr>
<td>$T_b$</td>
<td>$W_b = 9.4 \mu$m</td>
</tr>
<tr>
<td>$T_{IN_1}$ and $T_{IN_2}$</td>
<td>$W_{IN_1} = W_{IN_2} = 1 \mu$m</td>
</tr>
</tbody>
</table>

*Table 3. Sizes of the components in the excitatory synapse with two inputs.*

Compared to the excitatory synapse with one input, all components have the same size except for the capacitor that has a larger value. This property will be explained in following paragraphs.

**Simulations of the excitatory synapse with two inputs**

The architecture shown in Figure 23 will be used to simulate the behaviour of the excitatory synapse with two inputs.

![Figure 23. Architecture of the circuit in order to measure the properties of the excitatory synapse with two inputs.](image)
Many simulations have been done with different input values and synaptic weights. Therefore, in order to present the main results and problems, only two figures that exhibit the main results and problems incurred will be displayed. Figures 24 and 25 have both been obtained with input values equal to one. The difference between the two figures is the phase difference between the input signals. Figure 24 has been acquired with synchronous input signals and Figure 25 with a phase difference of $\pi$.

**Figure 24.** (Measured output – Theoretical output) of an excitatory synapse with two synchronous input signals equal to 1. $w_a$ and $w_b$ are the synaptic weights of the two inputs.

**Figure 25.** (Measured output – Theoretical output) of an excitatory synapse with two asynchronous input signals equal to 1. $w_a$ and $w_b$ are the synaptic weights of the two inputs.
Considering Figures 24 and 25, an important conclusion can be made. The output frequency of the integrate-and-fire neuron depends on the phase difference between the input signals. This is due to the fact that the EPSP has not the same maximum and minimum value depending on the phase difference as shown in Figure 26. This implies that the drain current of the synapse input transistor is not quite the same because it depends on the drain voltage as it was already explained on page 14. A second reason is that drain current of T_b (output transistor of the synapse) also depends of the variance of the EPSP (which is also its gate voltage). Therefore, the influence of the closer of the charging path of the spiking neuron (section 3.3.3) also leads to this difference between the asynchronous and synchronous inputs.

![Figure 26. Vdd-EPSP with synchronous (t_delay = 0) and asynchronous (t_delay = 5 ns) inputs.](image)

Figures 25 also shows that the output obtained with asynchronous input signals can be larger than 1 (output frequency higher than 100 MHz). This is due to the fact that too much current is injected in the excitatory synapses as it was already explained on page 14. This implies that the adjustment of the closer of the charging path of C_1 in order to have an output frequency equal to 100 MHz can not be obtained anymore. This property is shown in Figure 27 and 28.

![Figure 27. EPSC with synchronous input signals. The inputs and the weights are equal to 1.](image)
As explained for the excitatory synapse with one input, the solution is to increase the value of C_2 (see Figure 22) in order to have a negligible variance of the EPSP. Therefore, the influence of the closer of the charging path of the spiking neuron capacitor would be minimized. However, here again the disadvantage of this solution is the silicon area of the C_2.

To conclude, the precision of the excitatory synapse with two inputs is worst compared to the synapse with only one input. The limitations that have been discussed in this section become even worse if the number of inputs is increased. Therefore, a serious problem to the approach used in this project is the limitation of the number of synaptic inputs.

### 3.3.5. Inhibitory synapses with one input

Table 4 displays the values of the different components of the synapse circuit shown in Figure 29. The gate lengths of T_a and T_b are equal to 1.60 \( \mu \)m and the synaptic weights are coded in the gate length as described in equation (3) on page 13.

**Figure 28.** EPSC with asynchronous input signals. The inputs and the weights are equal to 1.

**Figure 29.** Architecture of the inhibitory synapse with one input.
A current source $I_o$ has been added. Its value is 28 $\mu$A which is the current needed to have an output frequency of 100 MHz when the IPSC is equal to zero.

The main problem of this circuit is to have the transistor $T_b$ operating in the saturation mode. This problem appears when the IPSP is large, which means that the drain current of the input transistor is large too. The only solution to solve this problem is to increase the value of the capacitor $C_2$ in order to decrease the variation of the IPSP. But in this case, much larger values of $C_2$ are needed compared to the excitatory synapse.

Another solution would be to decrease the gate width of the input transistor and increase $W_a$ in order to decrease the value of the IPSP. However, $W_b$ should also be increased in order to still have a correct output current. This would imply that the problem of keeping $T_b$ in the saturation mode would not be solved. This also explains why an increase of the power supply voltage, for example to 5 V, does not solve the problem either.

To conclude, the best performances seems to be reached when using the sizes that are displayed on Table 4.

### Simulation of the inhibitory synapse with one input

The same experimental conditions as for the excitatory synapse (see Figure 15) have been used to obtain the simulation results shown in Figure 30.

As it can be seen, the output error is acceptable except for an input and a synaptic weight equal to 1. In this case the transistor $T_b$ operates in the conduction mode and its drain current IPSC is equal to $I_o$ which leads to an output current equal to zero. This problem is explained in the following paragraphs.

<table>
<thead>
<tr>
<th>Components</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_2$</td>
<td>400 fF (10 $\mu$m x 45.4 $\mu$m)</td>
</tr>
<tr>
<td>$T_a$</td>
<td>$W_a = 2 \mu$m</td>
</tr>
<tr>
<td>$T_b$</td>
<td>$W_b = 3 \mu$m</td>
</tr>
<tr>
<td>$T_{IN}$</td>
<td>$W_{IN} = 2.95 \mu$m</td>
</tr>
</tbody>
</table>

*Table 4. Sizes of the components in the inhibitory synapse with one input.*
Figure 31. Output current of the inhibitory synapse with a gate width of the spiking neuron input transistor \( (T_1) \) equal to 4.6 \( \mu \text{m} \).

A negative current is a current that is injected in the spiking neuron.

The output voltage of the synapse, which is also the drain voltage of the transistor \( T_b \), is too low. This leads the output transistor of the current mirror to operate in the conduction mode. In this case, the current \( I_o \) flows entirely through \( T_b \) and the output current of the synapse is equal to zero as shown in Figure 31. This implies that the inhibitory synapse doesn’t discharge the capacitor of the integrate-and-fire neuron through \( T_b \) when it should. Therefore too much current is injected in the spiking neuron which leads to an output that is too large.

A solution would be to increase the width gate of the input transistor \( T_1 \) of the integrate-and-fire neuron. This would imply that the voltage between the source and the drain of this transistor is decreased and that by the same way the output voltage of the synapse would be increased. Figure 32 shows the output current of the synapse with \( W_1 \) (Figure 10) equal to 50 \( \mu \text{m} \). As it can be seen, the problem is not completely resolved and the width of the input transistor of the spiking neuron has become very large. Therefore this solution has not been selected.

Figure 32. Output current of the inhibitory synapse with a gate width of the spiking neuron input transistor equal to 50 \( \mu \text{m} \).
The influence of the closer of the charging path of the integrate-and-fire neuron capacitor (see section 3.3.3) is not a main problem as it is for the excitatory synapse. This is due to the fact that the current $I_o$ is constant and that the output frequency, which is also the frequency of the closer of the charging path of $C_1$, is small for large values of the IPSC.

### 3.3.6. Inhibitory synapses with two inputs

![Figure 33. Architecture of the inhibitory synapse with two inputs.](image)

Table 5 displays the values of the different components of the synapse circuit shown in Figure 33. The gate lengths of $T_a$ and $T_b$ are equal to 1.60 $\mu$m and the synaptic weights are coded in the gate length as described in equation (3) on page 13.

<table>
<thead>
<tr>
<th>Components</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_2$</td>
<td>400 fF $(10 \mu$m x 45.4 $\mu$m)</td>
</tr>
<tr>
<td>$T_a$</td>
<td>$W_a = 2 \mu$m</td>
</tr>
<tr>
<td>$T_b$</td>
<td>$W_b = 3 \mu$m</td>
</tr>
<tr>
<td>$T_{IN_1}$ and $T_{IN_2}$</td>
<td>$W_{IN_{1,2}} = 2.95 \mu$m</td>
</tr>
</tbody>
</table>

*Table 5. Sizes of the components in the inhibitory synapse with two inputs.*

All components have the same size except for $C_2$ which has a larger value compared to the inhibitory synapse with one input (Table 4) in order to minimize the variance of the EPSP.

The final results from simulations show that the maximum error is equal to 0.25. This report will not display these results because the problems are exactly the same as for the inhibitory synapse with one input.

### 3.3.7. Excitatory and inhibitory synapses with one input

The circuit shown in Figure 7 hasn’t been implemented at hardware circuit-level. The main reason is that the output transistor of the inhibitory synapse is very hard to operate all the time in the saturation mode. A solution would be to use capacitors with huge values, but then the total surface of the circuit would be much too big. Therefore, it has been decided not to develop this type of synapse.
3.3.8. Surface of the different synapses

The goal of this section is to show which components have the largest silicon surface in order to know which one must be specially reduced.

Figure 34. Excitatory synapse with one input.
(a) Distribution of the surface between the neuron and the synapse.
(b) Distribution of the surface between the different components of the synapse.

Figure 35. Excitatory synapse with two inputs.
(a) Distribution of the surface between the neuron and the synapse.
(b) Distribution of the surface between the different components of the synapse.

Figure 36. Inhibitory synapse with one input.
(a) Distribution of the surface between the neuron and the synapse.
(b) Distribution of the surface between the different components of the synapse.
Figure 34 to 37 demonstrate that the capacitor $C_2$ is the component which has the largest silicon area. This is specially the case of the inhibitory synapses.
4. Fault-tolerant Boolean gates

The goal has been to built the NOT, NOR and NAND gates using the integrate-and-fire neuron and the synapses that have been developed in section 3. With these Boolean gates, any digital function can be built. Then the same logic gates, but supporting various amount of defects, have been developed.

No learning algorithm has been used as described in section 2.2 due to the limited time allocated to this project. However, the synaptic weights can be determined very easily in our case because the Boolean gates that have been built are simple logic functions.

In this context, the architecture that has been used to develop fault-tolerant logic gates is inspired from [5]. As seen from Figure 38 (which is taken from [5]), this architecture consists of four layers in which the data is strictly processed in a feed-forward manner. The first layer is an input layer which accepts binary signal level (0 or 1). The core operation is performed in the second layer which is composed of N identical logic blocs. The third layer operates an average of the different outputs of the second layer and the last layer decides if the output is equal to 1 or 0 depending on the output of the averaging layer. It has been demonstrated in [5] that this architecture can successfully absorb a certain amount of defect density in the second layer depending on the number of redundant units.

![Figure 38. The proposed fault-tolerant architecture based on multiple layers.](image)

The errors that have been dealt with in this project are stuck-at-one and stuck-at-zero applied to the outputs of the second layer. The “zero” and “one” represent the logic levels which are encoded into the mean frequencies of the signals and not the ground and power supply voltages. Thus only faulty components from the synapses are considered because the ones from the integrate-and-fire neuron can lead to errors such as suck-at-Vdd. Therefore, from this point of view, coding the logic levels into mean frequencies of the signals is a disadvantage.

In order to obtain the probability of correct operation with respect to the number of faulty devices, Monte Carlo simulations have been performed using process parameters variations. Matching variations have not been considered because they lead to much smaller variations of the output level. For each number of faulty devices in the logic layer, only the errors that lead to the worst outputs have been considered for the determination of the probability of correct operation.
In Figure 40 to 49, the representation shown in Figure 39 is used:

![Figure 39. (a) Excitatory synapse connected to a spiking neuron. (b) Inhibitory synapse connected to a spiking neuron.](image)

4.1. NOT gate

The NOT gate is smallest gate to built. It consists of an inhibitory synapse connected to a spiking neuron as shown in Figure 40. This architecture uses a minimum number of synapses and spiking neurons. Therefore it can not support any faulty components.

![Figure 40. Architecture of the NOT gate using a minimum number of synapses and spiking neurons. The number is the synaptic weight.](image)

The NOT gate shown in Figure 40 is used as the logic block in the second layer of the fault-tolerant architecture displayed in Figure 38. Figure 41 and 42 are two examples of this architecture for the NOT gate with respectively two and three redundant units. The averaging layer is composed of an excitatory synapse having N inputs, with N being the number of redundant units. The synaptic weights of this excitatory synapse are all identical and equal to 1/N in order to perform an average of the logic layer outputs as shown in equation (4).

\[
y = \sum_{i=1}^{N} \omega_i x_i = \sum_{i=1}^{N} \frac{1}{N} x_i = \frac{1}{N} \sum_{i=1}^{N} x_i
\]  

(4)

![Figure 41. Fault-tolerant architecture of the NOT gate with two redundant units.](image)
No decision layer has been built as shown in Figure 38. However, in this project a threshold equal to 0.5 has been considered.

Finally, a disadvantage of this circuit is that only faulty devices from the logic layer are considered. Therefore, if the unit of the averaging layer has a faulty component, the output is incorrect.

**Simulation of the NOT gate**

Figure 43 and 44 display the main results obtained after having performed Monte Carlo simulations of the different circuits.

![Figure 43. Probability of correct operation of the NOT gate against the number of faulty devices with two to five redundant units.](image)

**Figure 42. Fault-tolerant architecture of the NOT gate with three redundant units.**
Figure 44. Probability of correct operation of the NOT gate against the density of faulty devices with two to five of redundant units.

Figure 43 shows that the fault immunity increases with the number of redundant units. However, Figure 44 also demonstrates that the density of faulty devices in the logic layer must be inferior to 50% in order to have a correct operation of the Boolean gate.

4.2. NOR gate

The NOR gate is an inhibitory synapse with two inputs each one having a synaptic weight equal to one as shown in Figure 45.

Figure 45. Architecture of the NOR gate using a minimum number of synapses and spiking neurons.

Figure 46 displays the fault-tolerant architecture used for the NOR gate with two redundant units. The averaging layer operates exactly the same way as for the NOT gate (see equation (4) on page 28).

Figure 46. Fault-tolerant architecture of the NOR gate with two redundant units.

The results obtained after the Monte Carlo simulations are not displayed because they are exactly the same as for the NOT gate.
4.3. NAND gate

Figure 47 displays the simplest architecture of the NAND gate. Unfortunately the output can be larger than one if the two outputs of the logic layer are also equal to one (a and b equal to zero). This problem was explained in section 3.3.4. This would be a serious problem if the output of this NAND gate had to be used as an input of another Boolean gate.

The solution consists in using inhibitory synapses instead of the excitatory synapse with two inputs. The final architecture of the units that are used in the logic layer (Figure 38) is shown in Figure 48.

Figure 48 display an example of the fault-tolerant architecture used for the NAND gate with two redundant units. The averaging layer operates exactly the same way as for the NOT and NOR gates (see equation (4) on page 28).

The architecture of the redundant units is shown in Figure 48.
Simulation of the NAND gate

Figure 50 and 51 display the main results obtained after having performed Monte Carlo simulations of the different circuits.

Figure 50. Probability of correct operation of the NAND gate against the number of faulty devices with two to five redundant units.

Figure 51. Probability of correct operation of the NAND gate against the density of faulty devices with two to five redundant units.

The results are nearly the same as for the NOT and NOR gates except that the probability of correct operation is a little bit worse due to the fact that more spiking neurons and synapses are used.
5. Conclusion

In order to compensate the serious degradation of reliability of future very-deep submicron CMOS devices, a new design approach consisting in the implementation of a multi-layer feed-forward artificial neural network in hardware using spiking neurons has been developed. Indeed, this type of neuron can easily encode the analog signals used by neural networks as mean frequencies of impulses.

First of all, the development of an integrate-and-fire neuron with excitatory and inhibitory synapses has been achieved. The most important property of the synapses is the fact that the synaptic weights are encoded into the gate length of a transistor. Therefore, if the correct synaptic weights have been previously calculated with a learning algorithm, the neural network can absorb a certain amount of defect density. Some limitations of the physical implementation of these synapses, such as the variation of the transistors threshold voltage and the difficulty to have all the transistors of the synapses operating in the saturation mode, have also been demonstrated.

Then a library of fault-tolerant Boolean gates (NOT, NOR and NAND) has been built using an architecture based on redundant layers arranged in a feed-forward manner. Considering logic errors such as stuck-at-one or stuck-at-zero, it was proved that the more redundant units are available, the more faulty devices can a neural network support and still operate as the desired Boolean gate.

However, the disadvantage of using this design approach is that high performance will be lost. Especially, these circuits will consume more power and will have a larger silicon surface compared to standard devices. However, for a safety-critical application, the most important aspect that has to be considered is the robustness of a circuit. Therefore, it might be preferable in some cases to consume more power and to use more silicon area if the electronic device is able to absorb so faults.

Finally future works will first consist in building the layout of these fault-tolerant Boolean gates. Then, the implementation of a learning algorithm (see section 2.2) will also have to be accomplished to be able to develop larger digital functions.
Part 2

Depressing Synapses and their Application to Synchrony Detection and Contrast-Invariant Pattern Classification
1. Introduction

In the first part of the project “Development of a fault-tolerant Boolean gates library”, the concept of mean firing rate has been applied for the coding of the signals used by the artificial neural networks. However, this temporal average neglects all the information contained in the exact timing of spikes. For example, the human brain can recognize and respond to visual scenes in less than a second. Recognition and reaction involve several processing steps from the retinal input to the finger movement at the output. If at each processing step, neurons had to wait and perform a temporal average in order to read the message of the presynaptic neurons, the reaction time would be much longer. Therefore, synapse whose conductivity changes based on the firing rate or spike timing of presynaptic neurons should be used. This type of synapse is called dynamic synapse.

In this project, some applications of the depressing synapse have been studied. In previous work [6], two applications of this synapse which are the synchrony detection and the contrast-invariant pattern classification have been implemented at hardware circuit-level using AMIS 1.5 µm CMOS technology. In this part of the project, identical work has been done but using AMS 0.35 µm CMOS technology.

This part of the project is divided into three parts. The first one is a study of the properties of the depressing synapse. Then the second part deals with the hardware circuit-level implementation of the depressing synapse application to the synchrony detection. In Senn’s previous work [7], it was shown that an easy way to extract coherence information among cortical neurons by projecting spike trains through depressing synapses onto a postsynaptic neuron. Finally, the last part shows a hardware circuit-level implementation of a depressing synapse application to is contrast-invariant pattern classification. In Bugmann’s previous work [8], it was shown that the strength of a time-averaged current injected into the soma by using a spike train is independent of its frequency, which implies that the response strength of a target neuron depends only on the number of active inputs.
2. Depressing synapse

The architecture of the depressing synapse circuit as shown in Figure 52 is the same as the one used in [6]. This circuit is made of a current mirror (T3 and T5) and a pMOS common-source amplifier (T2 and T4). An input transistor is also added in order to convert an input voltage $V_{ctr}$ to a current $I_{in}$. Finally, $C_e$ is a parasite capacitance.

Figure 52. Architecture of the depressing synapse.

To understand circuit operation, let us consider that the circuit has no input spike ($I_{in} = 0$). Therefore, in this state, the output current is also equal to zero and the capacitance $C_e$ is discharged through T2. This implies that $V_e$, which is the gate voltage of transistor T1, is at ground potential. The output transistor of the synapse (T1) is in this case in an on state.

Then if an input is applied to this depressing synapse, the input current $I_{in}$ is mirrored to the output current $I_{out}$ through T1. In this state, the input current also charges the capacitance $C_e$ which leads to an increase of $V_e$. Therefore the transistor T1 enters an off state. Because of the parasitic capacitance, the increase of $V_e$ has a short time delay. Therefore, before T1 enters the off state, the synapse has a short time to “fire” an impulse. When the input current becomes zero again, T2 discharges the capacitance $C_e$ and $V_e$ returns to ground potential.

Now, assume that a second input pulse is applied at the input of the synapse a short time after the first one. If this spike enters before $V_e$ returns to the ground potential, the output spike $I_{out}$ will be smaller than the first one. It can be concluded that the amplitude of the output spikes decreases when $V_e$ increases. Therefore, it is possible to control the time of the depression by adjusting the voltage $V_{bias}$. For example if this voltage is increased, the discharge current of the capacitance $C_e$, which is also the drain current of T2, will increase. Therefore the voltage $V_e$ will decrease faster. It has to be noticed that if $V_{bias}$ is set at Vdd, the circuit behave as a nondepressed synapse because $C_e$ is always discharged ($V_e = 0$).

An explanation of the operation of this synapse is shown in Figure 54 and 55.
**Simulation results**

In order to simulate the properties of the depressing synapse, the output of a depressing synapse is connected to the input of a leaky integrate-and-fire as shown in Figure 53.

![Depressing synapse](image)

*Figure 53. Architecture for the circuit in order to simulate the properties of the depressing synapse. EPSP = Excitatory postsynaptic potential.*

All transistors have the same gate length of 1.60 µm as explained in section 3.1 of the first part of the report. They also have the same gate width of 1 µm except for \( W_2 = 10 \) µm and \( W_{in} = 20 \) µm in order to be able to adjust \( V_{bias} \) and \( V_{ctr} \) more easily. Finally, the parasite capacitance \( C_e \) and \( C_1 \) have been set to 10 fF.

Figure 54 shows time courses of the output of the synapse circuit for increasing input-spike intervals. Table 6 displays the experimental conditions.

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{leak} )</td>
<td>300 mV</td>
</tr>
<tr>
<td>( V_{bias} )</td>
<td>240 mV</td>
</tr>
<tr>
<td>Input spike width</td>
<td>1 µs</td>
</tr>
<tr>
<td>Spike amplitude</td>
<td>100 nA</td>
</tr>
</tbody>
</table>

*Table 6. Experimental conditions for the depressing synapse connected to a leaky integrate-and-fire neuron.*

As shown in Figure 54, the amplitude of the output increases with the decrease of the voltage \( V_e \).

![Experimental results](image)

*Figure 54. Experimental results of the depressing synapse circuit; (a) successive spike inputs, (b) the degree of synaptic depression, and (c) its outputs.*
Figure 55 displays the change in amplitude of the output spike against the input firing rate. The normalized spike amplitude is equal to the measured amplitude divided by the maximum amplitude of the output signal (when $V_e = 0$). Figure 55 demonstrates that if $V_{bias}$ is increased, the cut-off frequency of the depressing synapse is shifted toward higher frequencies. This is due to the fact that the discharge current of the parasite capacitance, which is the drain current $T_2$, increases with the increase of $V_{bias}$.

![Figure 55. Changes in amplitude of the output of depressing synapse circuit against the firing rate of presynaptic neuron.](image)

To conclude the depressing synapse can be considered as a low-pass filter.
3. Depressing synapse with application to synchrony detection

Senn showed that an easy way to extract coherence information among cortical neurons is to project spike trains through depressing synapses onto a postsynaptic neuron [7]. This section demonstrates this property using the depressing synapse circuit that has been developed in section 2.

Simulation results

The same leaky integrate-and-fire neuron as shown in Figure 53 has been used except that the value of the capacitance $C_1$ is this time equal to 100 fF. The inputs of the depressing synapses are burst inputs as in Senn’s original work [7]. During a burst input, the output current of the depressing synapse rapidly decreases for successive spikes due to the increase of $V_e$ and its slow recovery. But during a nonbursting period, the parasite capacitance $C_e$ has enough time to be completely discharged ($V_e = 0$) which implies that the value of the EPSP at the onset of the next burst will be maximal. If this dynamic response is compared with one from a nondepressed synapse having the same mean value of the EPSP, the depressed synapse will have a larger response at the burst onset and a smaller response toward the end of the burst. Figure 57 and 58 show two examples of these properties with depressing and nondepressing synapses.

Table 7 and Figure 56 display the experimental conditions.

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{\text{leak}}$</td>
<td>215 mV</td>
</tr>
<tr>
<td>$V_{\text{bias}}$ (for depressing synapses)</td>
<td>240 mV</td>
</tr>
<tr>
<td>$V_{\text{bias}}$ (for nondepressing synapses)</td>
<td>3.3 V</td>
</tr>
<tr>
<td>Input spike width</td>
<td>1 $\mu$s</td>
</tr>
<tr>
<td>$t_1$</td>
<td>100 $\mu$s</td>
</tr>
<tr>
<td>$t_2$</td>
<td>4 ms</td>
</tr>
</tbody>
</table>

*Table 7. Experimental conditions for the depressing synapse with application to the synchrony detection.*

$V_{\text{bias}}$ has been adjusted in order to have the cut-off frequency between $t_2^{-1}$ and $t_1^{-1}$. In this case, $V_e$ can’t recover between two spikes of the same burst but has enough time to be set at ground potential between two bursts. The burst input signal is shown in Figure 56.

Figure 56. Representation of a burst input signal.

Figure 57 and 58 show the response of the EPSP with burst inputs for respectively a nondepressed and a depressed synapse. Amplitudes of burst inputs have been set at 100 nA for depressing synapses and 6 nA for nondepressing synapses in order to have the same mean value of the EPSP.
As it has been previously explained, the EPSP has a larger response at the burst onset if depressing synapses are used instead of nondepressing synapses.

Figure 57. Response of the EPSP for single burst input via nondepressed synapse circuit.

Figure 58. Response of the EPSP for single burst input via depressed synapse circuit.

The fact that using depressing synapses the synchrony between different burst inputs can be detected will now be demonstrated. The circuit that is used is the same as in Figure 53 but with two depressing synapses and the experimental conditions are displayed on Table 7.

When the burst inputs are not synchronized (Figure 59 and 60), the maximum value of the EPSP with depressing synapses is equal to 275 mV and 229 mV with nondepressing synapses. The difference between these two values is not very large. However, when burst inputs are synchronized (Figure 61 and 62), the maximum value of the EPSP is equal to 536 mV with depressing synapses and 454 mV with nondepressing synapses. These two values have increased, especially with the depressing synapses, compared to the one measured with asynchronous burst inputs. If a much larger network is simulated, the increase of the maximum value of the EPSP would be significant with depressing synapses and would be quite smaller for nondepressing synapses. Therefore, the synchrony detection using depressing synapses could be clearly demonstrated. However, this property can also be shown with the results that have been obtained here. If the appropriate threshold of the integrate-and-fire neuron is defined, for example 500 mV, the neuron with depressing synapse circuits can fire when burst inputs are synchronized while the one with nondepressing synapse circuits can not. Therefore it is possible to detect the synchrony of burst inputs using dynamic synapses.
Figure 59. Responses of the EPSP for asynchronous burst inputs via nondepressed synapse circuit.

Figure 60. Responses of the EPSP for asynchronous burst inputs via depressed synapse circuit.

Figure 61. Responses of the EPSP for synchronous burst inputs via nondepressed synapse circuit.
Figure 62. Responses of the EPSP for synchronous burst inputs via depressed synapse circuit.
4. Depressing synapse with application to contrast-invariant pattern classification

Bugmann showed that the strength of a time-average current injected into the soma by using a spike train inclines to be independent of its frequency, which implies that the response strength of a target neuron depends only on the number of active inputs [8]. This section demonstrates this property using the depressing synapse circuit that has been developed in section 2.

Simulation results

The same leaky integrate-and-fire neuron (LIFN) as employed for the synchrony detection (section 3) is used in this project. The EPSP of this circuit depends on the input current of the LIFN which depends on the number of active presynaptic neurons. Therefore, by setting the appropriate threshold of the LIFN corresponding to a certain number of active neurons, the LIFN can detect if the number of active neurons is inferior or superior to this number. On the other hand, the EPSP also increases in proportion to firing rates of presynaptic neurons. Therefore, the performance to discriminate the number of presynaptic active neurons largely deteriorates if the firing rate is not a constant value.

However, it has been shown in [8] that this problem can be solved by using depressing synapses. If input spikes are given to the depressing synapse successively in a short period, the efficiency to increase the EPSP per spikes drops. Even if the number of input spikes increases with the increase on firing rate, the value of EPSP does not change greatly because the efficiency per spike is lowered by the synaptic depression. Therefore, the discrimination performance of the network inclines to be independent of the firing frequencies of presynaptic neurons. This property has been demonstrated in Figure 63 to 66 by using the circuits that have been developed previously in section 2.

Table 8 displays the experimental conditions.

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{\text{leak}}$</td>
<td>215 mV</td>
</tr>
<tr>
<td>$V_{\text{bias}}$ (for depressing synapses)</td>
<td>240 mV</td>
</tr>
<tr>
<td>$V_{\text{bias}}$ (for nondepressing synapses)</td>
<td>3.3 V</td>
</tr>
<tr>
<td>Input spike width</td>
<td>1 µs</td>
</tr>
</tbody>
</table>

*Table 8. Experimental conditions for the depressing synapse with application to contrast-invariant pattern classification detection.*

First, the operation of the neuron circuit with nondepressed and depressed synapses as the number of active presynaptic neuron increases has been simulated (Figure 63 to 65). In case of nondepressed synapses, average values of the EPSP increases monotonically as the firing rate of postsynaptic neurons and the number of presynaptic neuron increases (Figure 63). On the other hand, in case of the depressed synapse, the EPSP doesn’t increases monotonically as the firing rate increases (Figure 64). This is due to the fact that the depressing synapse acts like a low-pass filter. For frequencies higher than the cut-off frequency of the depressing synapse, the EPSP decreases with the increase of the firing rate.
Figure 63. Changes in the EPSP against the number of active presynaptic neuron and their firing rates using nondepressing synapses.

Figure 64. Changes in the EPSP against the number of active presynaptic neuron and their firing rates using depressing synapses.
Figure 65. Changes in the EPSP against the number of active presynaptic neuron and their firing rates. (a) 1 active presynaptic neuron, (b) 2 active presynaptic neurons, (c) 3 active presynaptic neurons and (d) 4 active presynaptic neurons. 

DS = depressing synapses. NDS = nondepressing synapses.

Now, the firing threshold of the leaky integrate-and-fire neuron is set to 1.5 V. The firing rates when the EPSP exceeds this threshold to the number of active neurons are plotted in Figure 66 for both depressed (DS) and nondepressed (NDS) synapses. In this figure, the independence between the firing rate and the number of active neurons that are detected by the LIFN isn’t apparent. This is due to the fact that the networks which have been simulated are too small.

Figure 66. Results for dependence of the integrate-and-fire neuron on the firing rate of presynaptic neurons (4 neurons).
However, if a large-scale network is simulated, the properties of the depressing synapse can be demonstrated. Figure 67, which is taken from [6], was obtained by simulating a network with 100 synapses. As it can be seen, the value of the firing rate doesn’t influence the number of active neurons that is detected by the LIFN if this frequency is larger than 6 kHz. This simulation hasn’t been done due to the limited time allocated to the project.

Figure 67. Large-scale simulation results (100 neurons) for the same experiments shown in Figure 66.

Now, let’s assume that the presynaptic neurons are arranged on a 2D rectangular grid which forms some patterns like “E”, “L” or “-”, and that “E” is made of 90 active neurons, “L” of 50 neurons and “-” of 10. Figure 68 is taken from [8] and shows the example for the “E” pattern. Then, the firing rates of the active neurons are assumed to represent the “contrast” strength of these patterns. If there is little dependence on the presynaptic firing rates, the neuron can classify these patterns independently of their contrast strength. The result displayed in Figure 67 indicates that with depressing synapses, correct classification can be achieved for all patterns.

Figure 68. The leaky integrate-and-fire neuron proposed by Bugmann for pattern recognition [8].
5. Conclusion

In order to use the information encoded in the exact timing of the spikes, dynamic synapses have to be considered. In this approach, a previous work [6] has implemented at hardware circuit-level depressing synapses and their application to synchrony detection and contrast-invariant pattern classification. In this project, the same work but using AMS 0.35 μm CMOS technology has been performed.

The basic operations of these two applications have been demonstrated with small-scale (up to four synapses) networks. However, future work will have to perform large-scale simulations in order to show the real mechanism of these applications.
Conclusion

Neurobiological systems carry out computational tasks in a fundamentally different way as standard digital processors. Based on badly conditioned data, complex matching operations are carried out in real time with extremely limited computational elements with slow and noisy interconnections. Therefore, neurobiological systems have very advantageous properties that could be used in microelectronic devices. In this context, this project has studied the implementation of some of these concepts at hardware circuit-level.

In this context, it was already shown in previous work [1] that compensation of faults such as process variations or device mismatch can be achieved by implementing multi-layer feed-forward artificial neural networks. Thus in this project, this concept has been implemented at hardware circuit-level using spiking neurons. Indeed, this type of neurons can easily encode the analog signals used by neural networks as mean frequencies of impulses. Two types of synapses have also been used, namely excitatory and inhibitory synapses. Using this approach, NOT, NAND and NOR Boolean gates have been built in AMS 0.35 µm CMOS technology. The most important aspect of this work is that the neural weight value is encoded into the gate length of the input transistor of the synapse. Therefore a very straightforward mapping between the value of the weight and a physical parameter of the circuit is achieved. This library of fault-tolerant Boolean gates can in the future be used for designing safety-critical electronic devices using the well-established conventional design-flow applied for digital CMOS architectures.

In the second part of the project, depressing synapses and their applications to synchrony detection and contrast-invariant pattern classification has been implemented at hardware circuit-level. The two applications have been demonstrated using small-scale simulations. Therefore, future work will have to be done for large-scale simulations in order to correctly show the properties of these applications.

To conclude, future submicron electronic devices will face important fabrication limitations that will partly result in degrading reliability and severe increased power density. Therefore, using electronic devices based on a neurobiological approach will overcome some of the basic obstacles of technology scaling.

Sapporo, February 17, 2005

Neil Joye
References


