

Final report of the Degree in Physics Engineering

CMOS photonic devices characterization and quantum computing architecture design

Darío de la Fuente García

Supervised by:

Prof. Rajeev J. Ram

Prof. José Antonio Lázaro Villa

Physical Optics and Electronics Group, Research Laboratory of Electronics
Massachusetts Institute of Technology

Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona
Universitat Politècnica de Catalunya

January 2018



Table of contents

Abstract.....	4
Acknowledgements.....	5
1. Introduction.....	6
1.1. Motivation.....	6
1.2. Quantum computing fundamentals.....	7
1.2.1. Reversible computing.....	7
1.2.2. Qubits and quantum gates.....	8
1.2.3. Sources of error in quantum computing.....	10
1.2.4. Quantum computing algorithms.....	12
1.3. Ring modulator fundamentals.....	13
1.3.1. Physics of the P-N junction.....	13
1.3.2. Physics of the micro ring modulator.....	14
1.4. Energy consumption for modulation.....	18
1.5. Thesis outline.....	19
2. Quantum computing architecture.....	20
2.1. Current challenges in quantum computing.....	20
2.2. Distributed architecture.....	20
2.3. Bandwidth requirements.....	24
2.4. Possible bottlenecks.....	26
3. Photovoltaic optical modulator.....	28
3.1. Photovoltaic effect in P-N junctions.....	28
3.2. DC Measurements.....	29
4. Frequency response of PV modulator.....	34
4.1. Transistor-switched modulator.....	34
4.2. Frequency measurements setup.....	35
4.3. Results.....	37
4.4. Interpretation of the results.....	41
5. Conclusions and future work.....	46
5.1. Conclusions.....	46
5.2. Future work.....	46
References.....	48

Abstract

Quantum computing is an exciting new technology that will allow to solve some problems that current computers cannot do. However, most of the technologies being researched for quantum computing need ultra-low temperatures on the order of tens of mK, at which digital computers are extremely inefficient to run, because of the cost of dissipating the heat produced. In order to interface a quantum computer and a digital one, cables cannot be used, as they conduct too much heat for the cooler to dissipate. Optical fibers do not have this problem, but then low energy conversion between optical and electrical signals are necessary.

This work has two areas: on the one hand, the architecture of a distributed quantum computer and its interfacing with a digital computer is described, together with some bandwidth requirements; on the other, a CMOS ring modulator that uses its photocurrent to switch is characterized.

Acknowledgements

First I would like to give special thanks to Professor Rajeev J. Ram for accepting me in his lab and for giving me guidance throughout the whole stay. His great insight and advice were of utmost importance in the development of this research.

I would also like to thank Dr. Amir H. Atabaki. His assistance with experiments, theory and all sorts of troubleshooting. His help and expertise were vital for being able to perform the practical part of this research successfully.

I thank as well all the people in the Physical Optics and Electronics group at MIT. They were an amazing group of incredibly talented people with very diverse interests.

From the UPC, I would like to thank Professor Jose Antonio Lázaro for being the co-director of this thesis and helping me during its development. Also thanks to the CFIS, both for making my stay at MIT possible through the grant, and for providing the opportunity of pursuing the double degree I am now finishing.

I also want to mention all my friends, both old ones from Oviedo and Barcelona and the new ones I met in Boston, for making life more enjoyable.

Last, but not least important, a million thanks for my family, who has been supporting and caring for me all these years.

1. Introduction

1.1. Motivation

Quantum computers are expected to be able to do in a reasonable time some tasks that a classical computer cannot do efficiently. One of the most promising technologies underlying quantum computation, superconducting qubits, requires temperatures of about 20 mK in order to work. As quantum speedup is only known to be possible for some specific algorithms ([1], [2]), a solution would be a quantum coprocessor that would offload work from a main classical processor for said algorithms.

This setup raises one problem – classical computers produce heat which has to be dissipated, even more so in an ultra-low temperature environment. Therefore, the most energy efficient solution would be to have the classical computer at room temperature and interface it with the quantum computer. However, conventional metallic cables are not fit for this purpose: they conduct well electricity (and thus, information in the shape of electric signals), but they conduct well heat, and dissipating this heat becomes prohibitively expensive.

A solution to this is to send the information between the coprocessor and the central processor with an optical fiber, as fiber is not a good conductor of heat. This approach, however, needs optical modulators that can work at room temperature and at low temperature.

In this work, a low-energy photovoltaic ring modulator that can convert between electrical and optical signals with a single transistor is described and characterized. The energy savings of this approach over a regular ring modulator are twofold: on one hand, the transistor makes switching between the on and off states require much less peak-to-peak voltage; and on the other hand, the light that is absorbed into the ring generates the photovoltaic energy that allows the modulator to work, and thus less energy is wasted.

Another issue with quantum computers is that dilution refrigerators, which are currently used to cool down quantum computers, have very limited capacity, and thus a useful quantum computer would not be able to fit in a single one. A distributed architecture [3] can solve this problem, but introduces new ones, such as inter-node connection and coordination.

In this work, a distributed quantum computer architecture is described where a specialized classical processor serves as an intermediary between the CPU and the quantum processor.

In the remaining part of this chapter, the theoretical fundamentals for quantum computing and microring modulators are described, as well as a comparison of energy cost per bit of information transmitted for both electrical links and optical links.

1.2. Quantum computing fundamentals

In this section I will first introduce reversible computing, as quantum computing is an extension of it. Then I will explain how quantum states and quantum gates are represented. Afterwards I will describe the sources of error in actual quantum computers and how they are characterized. Finally I will talk about two quantum algorithms that are of special interest.

1.2.1. Reversible computing

Due to the laws of thermodynamics, there is a minimum energy cost to erase a single bit of data [4], called the Landauer energy, equal to $kT \ln 2$.

For example, let us consider a XOR gate. The XOR gate has two inputs and one output. The output is 1 if the inputs have different values and 0 if they do not, as shown below.

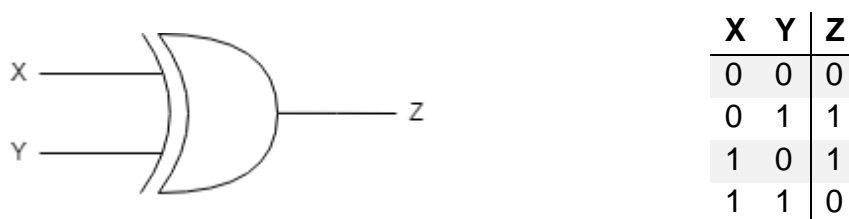


Figure 1.1: Icon and truth table for the XOR gate.

This gate is irreversible: there is no way to guess the value of X and Y if only Z is known. However, with another additional bit of data, for example, the value of X, we could reconstruct the inputs. Therefore, we can say that this gate destroys one bit of information, and therefore cannot be executed without spending at least $kT \ln 2$ joules of energy. With current technology, energy consumption per gate is several orders of magnitude above the Landauer energy and thus irreversible gates are used for computations, as they are simpler to build and use.

However, if a computation only used gates that do not erase information (and are therefore reversible) the theoretically minimal energy cost can be 0.

Reversible computing is a model of computing where the computations are performed by reversible gates, that do not destroy information and thus their energy cost can be arbitrarily small.

As with standard irreversible computing, reversible computing can be decomposed in a set of elemental gates. Gates in reversible computing must have the same number of inputs and outputs, as to not destroy information.

Moreover, since they must be reversible, different inputs have to result in different outputs.

In irreversible computing, there are two elemental gates (NAND and NOR) that are called universal, because any one of the two is sufficient to express any computation. These gates are not reversible, as they have two inputs and only one output, but there are reversible equivalents, the FREDKIN and TOFFOLI gates ([5], [6]).

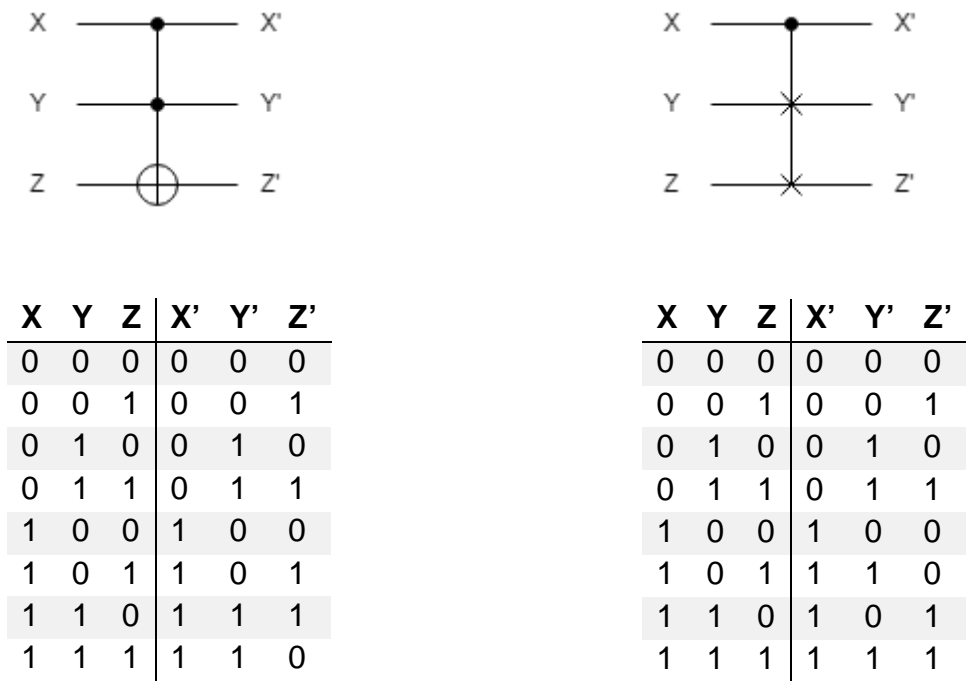


Figure 1.2: Icon and truth table for the TOFFOLI gate (left) and for the FREDKIN gate (right).

Another important concept for reversible logic are ancilla bits. Ancilla bits are bits with a constant value that are not inputs of the computation. They are used to store intermediate results and some computations do require a minimum number of them [5], but usage of a higher amount can lead to fewer gates.

1.2.2. Qubits and quantum gates

A digital bit can only have two different states: 0 and 1. On the other hand, a quantum bit or qubit has infinite states, any state $\alpha|0\rangle + \beta|1\rangle$ with α and β complex numbers and $\alpha^2 + \beta^2 = 1$ is a possible state (although, as with any quantum state, global phase cannot be measured). All the possible states of a single qubit can be corresponded to the points in the surface of a sphere called the Bloch sphere:

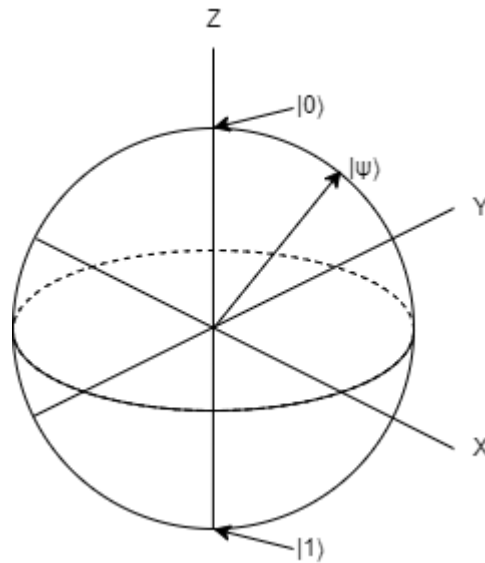


Figure 1.3: Bloch sphere.

$|0\rangle$ and $|1\rangle$ represent basis states, so we can also represent the state of a qubit as $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$. This idea can be extended for several qubits: the basis states would

be $|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $|01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $|10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, $|11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ for two qubits, for example.

Quantum gates operate on the state of one or more qubits. We can represent a quantum gate as a 2^N by 2^N matrix, where N is the number of qubits it acts on. Since quantum information cannot be destroyed, all quantum gates have to be reversible. Additionally, the matrix for a quantum gate has to be unitary, that is, that its inverse is equal to its conjugate transpose. This required property is an extension of the property that digital reversible gates have.

There is also a universal base for quantum gates. Any quantum operation can be executed to arbitrary precision with only the following gates [7]:

$$\text{Hadamard gate: } H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\text{Pauli X rotation gate: } X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\text{Pauli Z rotation gate: } Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\pi/8 \text{ gate: } T = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix}$$

$$\text{Controlled-NOT gate: } CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

1-qubit gates are equivalent to a rotation on the Bloch sphere. For example, the Pauli X rotation gate is equivalent to a 180° rotation along the \hat{x} axis of the Bloch sphere, and a Hadamard gate is equivalent to a 180° rotation along the $\frac{1}{\sqrt{2}}(\hat{x} + \hat{z})$ axis.

When measured, a qubit on a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ will collapse to 0 with $|\alpha|^2$ probability and to 1 with $|\beta|^2$ probability.

As with classical reversible computing, ancilla qubits (prepared beforehand in a known state, usually $|0\rangle$ or $|1\rangle$) are necessary for performing most computations. Moreover, ancilla qubits introduce a space-time tradeoff: additional ancilla qubits can reduce the number of necessary gates for a computation).

1.2.3. Sources of error in quantum computing

There are three sources of error for quantum computers: gate noise, decoherence, and state preparation and measurement errors. All three will be described in this section.

- Gate noise

The average fidelity of a quantum gate U , defined as how well it approximates a quantum gate V is defined by the following equation [8]:

$$\bar{F}(U, V) = \int d\psi \langle \psi | V^\dagger U | \psi \rangle \langle \psi | U^\dagger V | \psi \rangle \quad (1.1)$$

The fidelity will be equal to 1 if $U = V$, and lower otherwise. Therefore, a possibility would be to consider gate error as 1 minus the fidelity. However, the fidelity is hard to measure and gate noise is usually modeled in a simplified manner. For example [9], the gate noise can be determined by a single parameter ϵ . In this model, gates are performed perfectly with probability $1 - \epsilon$ and there's a probability ϵ of applying the gate and a 180° rotation along the \hat{x} , \hat{y} or \hat{z} axis.

- **Decoherence**

If we have a quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, with $|\alpha|^2 + |\beta|^2 = 1$, we can obtain its density matrix:

$$\rho = |\psi\rangle\langle\psi| = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix} \quad (1.2)$$

The diagonal components are the populations, and the off-diagonal ones are the coherences. We can use the density matrix to compute the expectation value of an operator \hat{A} :

$$\langle\psi|\hat{A}|\psi\rangle = Tr(\rho\hat{A}) \quad (1.3)$$

However, a quantum system interacts with its environment, so there is a time evolution, as the states of the system entangle with the corresponding states of the environment:

$$|\psi(t)\rangle = \alpha|0\rangle \otimes |e_0\rangle + \beta|1\rangle \otimes |e_1\rangle \quad (1.4)$$

$|e_0\rangle$ and $|e_1\rangle$ are not necessarily perpendicular:

$$\langle e_0|e_1\rangle = \cos\theta \quad (1.5)$$

The reduced density matrix $\rho(t)$ becomes:

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \cos\theta \\ \alpha^*\beta \cos\theta & |\beta|^2 \end{bmatrix} \quad (1.6)$$

As the environment evolves, $\cos\theta \rightarrow 0$ and the coherences disappear. In the limit, the density matrix is diagonal and corresponds to $|\alpha|^2$ probability of state $|0\rangle$ and $|\beta|^2$ probability of state $|1\rangle$: the superposition of states is lost. This process has a lifetime which varies depending on the technology of the quantum computer. For example, in the case of superconducting qubits, the lifetime is about 60 μ s [10].

- **State preparation and measurement errors**

Quantum algorithms require a specific state to be set to a qubit (usually $|0\rangle$) before being used. There is a probability of the qubit not being applied that state,

and that is a state preparation error. Measurement errors occur when the measurement, instead of returning the correct bit, returns the other.

1.2.4. Quantum computing algorithms

In this section I will talk about two quantum algorithms: quantum simulation and Shor's algorithm. These two algorithms have special interest because they need a relatively low amount of qubits to surpass classical algorithms. However, there are many more, such as Grover's search algorithm [2].

- Quantum simulation

Quantum simulation refers to a set of algorithms to solve physical and chemical problems with a quantum computer. This was first suggested by Feynman [11], arguing that since nature is quantum, it should be simulated with a quantum computer, not with a digital one.

The problem classical computers have simulating quantum systems is that a quantum system is represented by a Hilbert space of a number of dimensions that scales exponentially with the number of particles. Therefore, exact simulation gets prohibitively expensive in time and resources quickly, so that even supercomputers cannot run them in reasonable time, and classical quantum simulation algorithms have to rely on approximations.

The most common problem in this category is calculating the ground state of a molecule [12], because it is expected to require little quantum resources to surpass current supercomputers. This can be done, given the Hamiltonian \hat{H} of the molecule, by applying the time evolution operator $e^{-\frac{i\hat{H}t}{\hbar}}$ and then applying a phase estimation algorithm to find the eigenvalues of the Hamiltonian, which are the energies of the states, and thus the lowest eigenvalue will be the ground state of the molecule.

The best result with an actual quantum computer so far has been determining the ground state energy of BeH_2 , with a seven-qubit quantum computer manufactured by IBM [13].

- Shor's algorithm

Shor's algorithm [1] is a quantum algorithm for factoring integers. It offers exponential speedup over the best known classical algorithms, which scale exponentially [14], and is of special interest because one of the most used methods of public key cryptography, RSA [15], relies on the fact that factoring very large numbers is impossible for even the largest classical computers.

Shor's algorithm transforms the problem of factoring a number N into finding the period of the sequence $f(x) = a^x \bmod N$, being a an arbitrary number coprime with N . Using a Quantum Fourier Transform, this can be solved in polynomial time by a quantum computer.

So far, only the number 15 has been factorized by an actual quantum computer [16].

As explained in Section 1.1, the necessity of using dilution refrigerators imposes restrictions on the maximum number of qubits that can be put together. This can be solved by having a distributed architecture of several nodes, each consisting in a dilution fridge with some number of qubits on it which can communicate among themselves and with the classical computer. In Chapter 2, a quantum computing distributed architecture that attempts to solve these communication problems is proposed.

1.3. Ring modulator fundamentals

In this section, the physics of the ring modulator are explained. In Chapters 3 and 4, it will be shown how to turn this device into a PV modulator, and how this allows for lower energy dissipation.

1.3.1. Physics of the P-N junction

A P-N junction is the union of a P-type semiconductor (a semiconductor doped with holes, with density N_A) and an N-type semiconductor (a semiconductor doped with electrons, with density N_D).

In the absence of an external voltage, due to diffusion, electrons will flow into the P-type semiconductor and holes will flow into the N-type semiconductor. Therefore, the regions close to the boundary become electrically charged, and thus generates an electrical field which moves the electrons and holes in the opposite direction until it compensates the diffusion effect and the system arrives to an equilibrium.

We call the region where the semiconductors have charge the depletion layer region. The width of this region can be calculated by the following equation:

$$W = \sqrt{\frac{2\epsilon\phi_{eq}}{q} \cdot \frac{N_A + N_D}{N_A \cdot N_D}} \quad (1.7)$$

Where ϕ_{eq} is the potential of the junction in equilibrium.

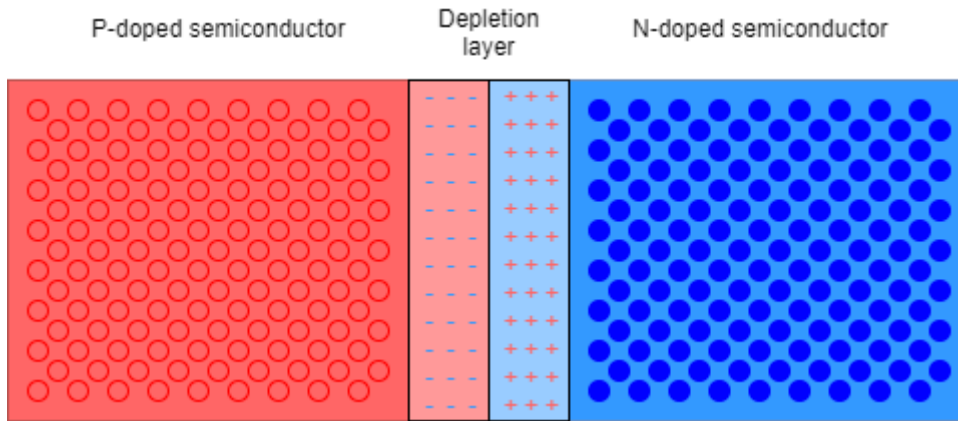


Figure 1.4: Schematic of the P-N junction.

When a bias voltage V_b is applied, the width of the region changes to:

$$W = \sqrt{\frac{2\epsilon \cdot (\phi_{eq} - V_b)}{q} \cdot \frac{N_A + N_D}{N_A \cdot N_D}} \quad (1.8)$$

The bias voltage also changes the concentration of carriers in the semiconductor, because it changes the rate of diffusion of carriers. Under low-level injection (that is, low N_A and N_D), the concentration of minority carriers at the extremes of the depletion layer increases by a factor of $e^{\frac{q \cdot V_b}{kT}}$. This effect is used in the ring modulator.

1.3.2. Physics of the micro ring modulator

A ring modulator consists in an optical waveguide next to a waveguide in the shape of a ring.

Part of the light that goes through the waveguide enters into the ring, travels through the ring and part of it goes back into the waveguide. The output of the modulator depends on the resonance of the light that has travelled through the ring with the light that enters the ring, as we will see later.

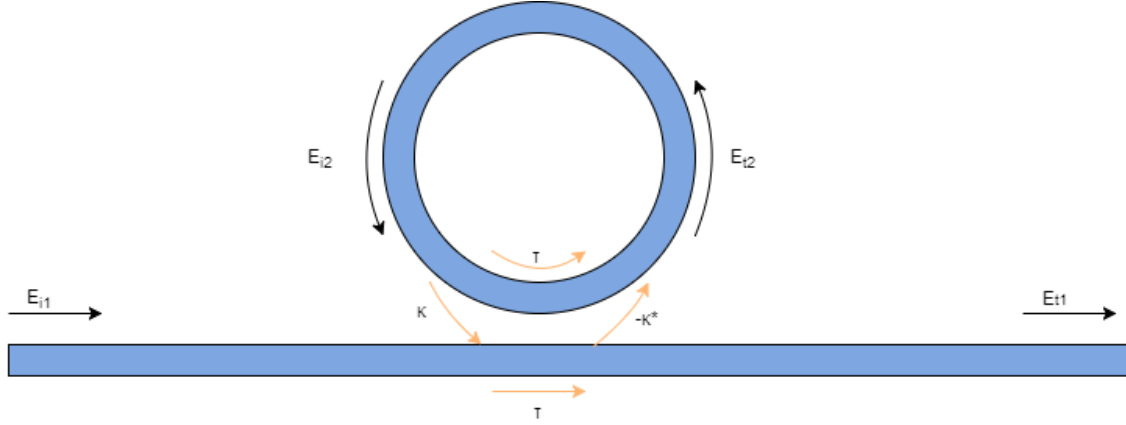


Figure 1.5: Optical ring modulator diagram.

In this diagram, τ is the transmittance of the waveguide and κ is the coupling coefficient between the waveguide and the ring. We can express the relations between the amplitudes of the fields in matrix form:

$$\begin{bmatrix} E_{t1} \\ E_{t2} \end{bmatrix} = \begin{bmatrix} \tau & \kappa \\ -\kappa^* & \tau^* \end{bmatrix} \cdot \begin{bmatrix} E_{i1} \\ E_{i2} \end{bmatrix} \quad (1.9)$$

We can obtain an expression [17] of the output power, P_{t1} :

$$P_{t1} = |E_{t1}|^2 = \frac{\alpha^2 + |\tau|^2 - 2\alpha|\tau| \cdot \cos\left(\frac{4\pi^2 nr}{\lambda} + \phi_t\right)}{1 + \alpha^2|\tau|^2 - 2\alpha|\tau| \cdot \cos\left(\frac{4\pi^2 nr}{\lambda} + \phi_t\right)} P_{i1} \quad (1.10)$$

In this expression, α is the loss coefficient of the ring, ϕ_t the phase introduced by the coupler, r the radius of the ring, λ the wavelength of the wave and n the refractive index of the ring. This equation shows that there is a peak of resonance when $\cos\left(\frac{4\pi^2 nr}{\lambda} + \phi_t\right) = 1$, which occurs when E_{i2} and E_{t2} form a constructive interference, and results in a minimum of output power. The minimum transmission is:

$$\left(\frac{P_{t1}}{P_{i1}}\right)_{min} = \frac{\alpha^2 + |\tau|^2 - 2\alpha|\tau|}{1 + \alpha^2|\tau|^2 - 2\alpha|\tau|} = \frac{(\alpha - |\tau|)^2}{(1 - \alpha|\tau|)^2} \quad (1.11)$$

Our objective is that the minimum transmission is as low as possible, because it sets a limit on how good the resonator can be. For example, if the difference between the maximum and minimum possible transmission is 3 dB, then in the

best case the on and off states of our resonator will have 3 dB of difference between them. In order to achieve a good extinction ratio (defined as $\frac{P_{t1_{max}}}{P_{t1_{min}}}$), the minimum power should be as low as possible, and for that, $\alpha = |\tau|$, which we call the critically-coupling condition.

Experimentally, we can observe the total quality factor Q of the ring with the following equation (FWHM = Full Width at Half Maximum):

$$Q = \frac{\lambda}{\delta\lambda_{FWHM}} \quad (1.12)$$

This quality factor measures the ratio of stored / lost energy per unit time. There are other two Q factors that help with the characterization of a ring modulator: the coupling Q , Q_c , which determines the amount of energy that goes into the ring; and the intrinsic Q , Q_i , which determines the amount of energy lost by the ring. The expressions are:

$$Q_c = \frac{4\pi^2 nr}{\lambda \kappa^2} \quad (1.13)$$

$$Q_i = \frac{2\pi n}{\lambda \alpha} \quad (1.14)$$

With these definitions, we can express the quotient between the output and input power near the resonance peak as:

$$\frac{P_{t1}}{P_{i1}} = \frac{\frac{1}{2}\left(\frac{1}{Q_i} - \frac{1}{Q_c}\right) - j\frac{\Delta\omega}{\omega}}{\frac{1}{2}\left(\frac{1}{Q_i} + \frac{1}{Q_c}\right) - j\frac{\Delta\omega}{\omega}} \quad (1.15)$$

Where $\Delta\omega$ is the distance from resonance. If $\omega = \omega_{resonance}$, the equation reduces to:

$$\frac{P_{t1}}{P_{i1}} = \frac{\frac{1}{Q_i} - \frac{1}{Q_c}}{\frac{1}{Q_i} + \frac{1}{Q_c}} \quad (1.16)$$

This expression of the output power in the resonance frequency helps us to distinguish between two different causes for low extinction ratio: undercoupling, where $\frac{1}{Q_i} < \frac{1}{Q_c}$ and the coupling loss is bigger than the intrinsic loss; and overcoupling, $\frac{1}{Q_i} > \frac{1}{Q_c}$.

The coupling factor Q_c can be modified by changing the distance between the waveguide and the ring, and the intrinsic factor Q_i by varying the input power of the laser.

The ring is a P-N junction. There are three different ways the P-N junction can be arranged in a ring: lateral, vertical and interleaved.

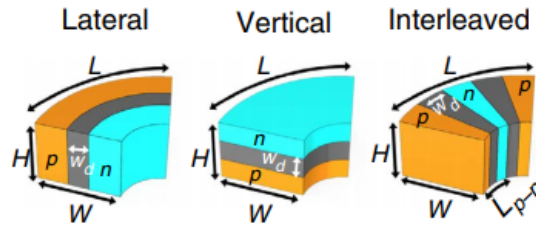


Figure 1.6: Schematic of the three different types of junction in a ring modulator [18].

Due to restrictions in the CMOS fabrication process, vertical junctions cannot be made. In our case, the P-N junction of the ring is lateral.

If we are transmitting light with a fixed wavelength, we can change the resonance wavelength of the device by modifying the refractive index of the ring. This results in a variance in output power, further amplified with a good extinction ratio. In order to modify the refractive index, we apply a voltage on the P-N junction of the ring, so that the width of its depletion layer region changes, which has an impact on the carrier concentration of the ring, which in turn changes its refractive index. This last relation was experimentally characterized by Soref et al. in [19] and for a wavelength of 1550 nm at 300 K the equation is:

$$\Delta n = -(2.1 \cdot 10^{-22} \cdot N_e^{1.04} + 3.4 \cdot 10^{-18} \cdot N_h^{0.82}) \quad (1.17)$$

Where N_e and N_h are the density of electrons and holes, respectively, in cm^{-3} .

To sum up, in order to transform an electrical signal into an optical signal with a ring modulator, the electrical signal is applied to the P-N junction of the ring, and this changes the depletion width, which modifies the overall concentration of carriers, which in turn changes the refractive index of the ring, and the resonance peak shifts. If the frequency of the laser is close to the resonance frequency of the ring, we get a shift of output power between the two different voltages.

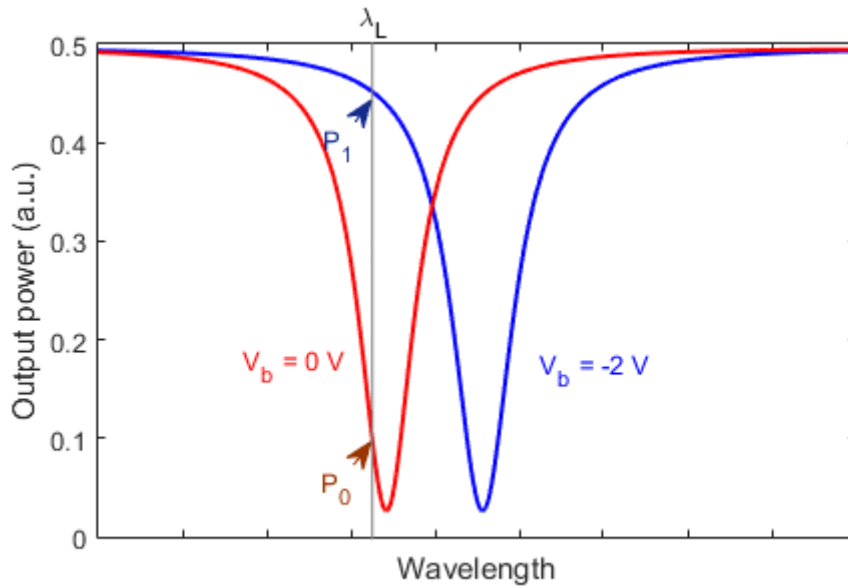


Figure 1.7: Example of ring resonator modulation.

1.4. Energy consumption for modulation

Dilution refrigerators [20] are capable of dissipating a maximum amount of heat, which is known as its cooling power. The cooling power, apart from the specifics of the refrigerator, depends on the target temperature. For example, a state-of-the-art dilution refrigerator, the BF-XLD1000, has 30 μW of cooling power at a temperature of 20 mK and 1000 μW at 100 mK. This sets a maximum on how much power the energy modulation can consume.

This is a very low amount of power for computing purposes. To put an example, the commercial device that currently offers the most floating point operations per joule is the Virtex-7 690T, an FPGA manufactured by Xilinx. It offers 78 GFLOP/J. At cryogenic temperature, power efficiency will increase by about 50%, because leakage currents, which constitute about 30-40% of the energy consumption of CMOS chips, decreases dramatically with temperature (the other factors of energy consumption will likely not be affected) [21]. Therefore, inside a dilution fridge, current state-of-the-art would be 117 GFLOP/J, which is equivalent to 117 KFLOP/ μJ . That is to say, one million floating point operations per second require at least 8.55 μW with current technology.

If we were to assume that we connected the room-temperature classical computer with a low-temperature quantum computer using a 1 meter long copper cable of 0.5 mm^2 section, we can estimate the heat transfer by conduction with the following equation:

$$H = kA \frac{\Delta T}{L} \quad (1.18)$$

Here H is the heat transfer in watts, k is the thermal conductivity of the material (assumed constant), A the area of the medium, ΔT is the temperature difference and L is the length of the cable. The thermal conductivity of copper is around $400 \text{ W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$, so the result is that the heat transfer is about 60 mW, which is several orders of magnitude outside the possibilities of dilution refrigerators, without even taking into account the energy cost of the data transmission itself.

For modulation of optical signals, it has to be taken into account that there are two sources of energy consumption: the light that is absorbed by the microring and transformed into heat, and the electrical energy consumed by the circuit. The electrical energy consumption increases more or less linearly with the frequency. However, the optical power consumption does not increase with bandwidth to first order, although in order to maintain signal-to-noise ratio, higher laser power is needed for very high frequencies. This means that in terms of efficiency it is best to operate a photonic modulator as fast as it can work, but since in this case the energetic constraints are so tight, it may make sense to operate a device slower to reduce the total consumption even though the cost per bit is higher.

Our energy consumption per bit target depends of the specifics of the refrigerator and the desired bandwidth. If the power that can be dissipated is $10 \mu\text{W}$, and the desired bandwidth is 10 Gbps, energy consumption must be 1 fJ/bit, which is two orders of magnitude better than the current state-of-the-art [22], which achieves a data rate of 3.5 Gbps and 70 fJ/bit for the transmitter, as well as 2.5 Gbps and 220 fJ/bit.

1.5. Thesis outline

The remainder of this thesis is organized as follows:

Chapter 2 presents the current technological obstacles for quantum computing scalability and with that base it presents a distributed architecture that aims to solve the scalability issues.

An optical ring modulator that can be switched with a single transistor by taking advantage of the open circuit voltage that is generated by the light entering the ring is described in Chapters 3 and 4. Chapter 3 is dedicated to explaining the physics of this device and for the DC measurements, whereas chapter 4 contains the information about the necessary setup to test the frequency response of the device and the AC measurements.

Finally, in Chapter 5, conclusions of the work and possible future work are discussed.

2. Quantum computing architecture

2.1. Current challenges in quantum computing

There are a number of as of yet unresolved challenges for practical quantum computing to be achieved.

- 1- Scaling the number of qubits: increasing the number of qubits results in lower ability to target them individually.
- 2- Quantum gates have a lot of noise (at best around 3% gate error [23]).

Problem number 2 can be solved, as long as gate error is under a threshold, with quantum error correction (QEC) algorithms. The needed threshold varies with the algorithm, for example, $1.5 \cdot 10^{-3}$ for Steane error correcting codes ([24], [25]) and 1% for C4/C6 architecture [26]. QEC algorithms use several physical qubits to represent a single qubit (which is called a logical qubit) which has less gate error than the original physical qubits. In order to obtain arbitrarily small gate error, this method can be concatenated, that is, several logical qubits forming a single higher-order logical qubit.

However, this solution spawns two new, different problems:

- 3- Some QEC algorithms cannot perform all the necessary fundamental qubit gates on the logical qubits directly.
- 4- All qubits cannot fit in a single chip.

Problem 3 means that for some QEC algorithms [27] a percentage of the available logical qubits have to be dedicated to generating specific states with high precision, as using these states the missing gates can be implemented. However, this requires even more qubits and thus problem 4 is aggravated.

Problem 4 stems from the fact that the qubits must be separated about $10 \mu\text{m}$ to prevent undesired interactions between neighboring qubits, and the capacity of dilution refrigerators is small, which puts a cap on how many qubits there can be in a dilution refrigerator.

2.2. Distributed architecture

To solve the problem of fitting all the qubits necessary for QEC (which will vary depending on the specific algorithm used), a distributed architecture can be used. This approach has already been suggested before: in one paper there were a set of logic units that hold qubits and can operate with them connected by an optical cross-connect switch that provides all-to-all connectivity [28], and another paper proposes links between nodes using quantum teleportation [29]. Here I present some improvements for the second option.

The high-level view of this architecture is a set of nodes, where the qubits and gates to manipulate them are, connected in a 2-D mesh architecture. In order to be able to operate between logical qubits in different nodes, in the middle of each link there must be a device that generates EPR pairs (a pair of qubits that are in an specific entangled state called a Bell state) [30] in order to be able to teleport qubits between nodes.

The nodes interchange digital information with the exterior through optical fibers. Thanks to frequency multiplexing, a single optical fiber can transmit the information of several nodes.

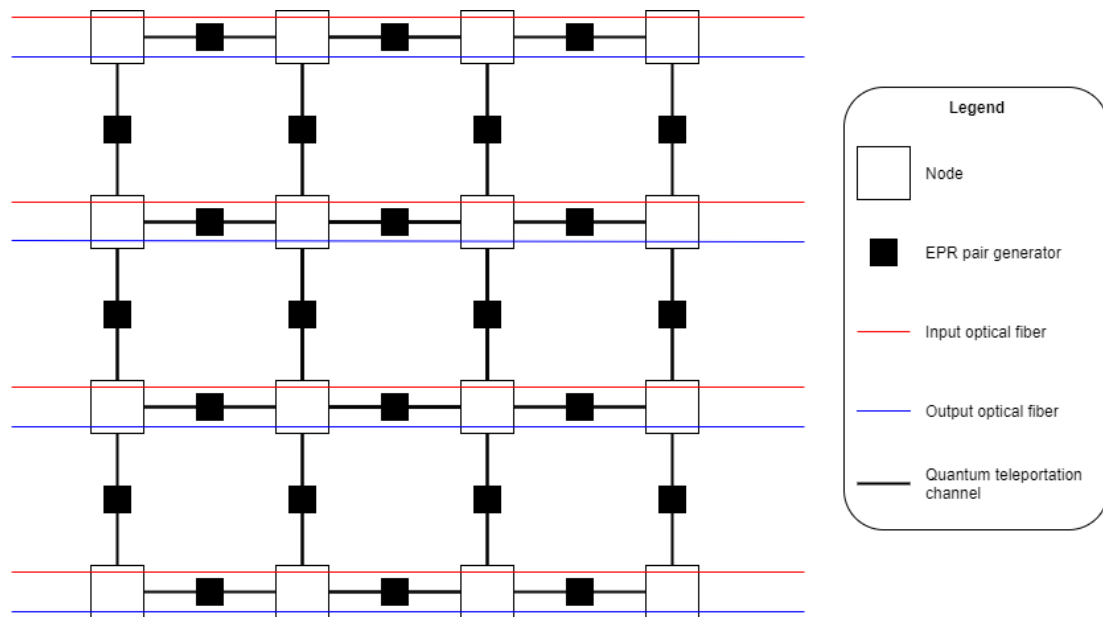


Figure 2.1: Schematic of the proposed architecture. The black lines represent communication between nodes by quantum teleportation.

Each node would have the following structure:

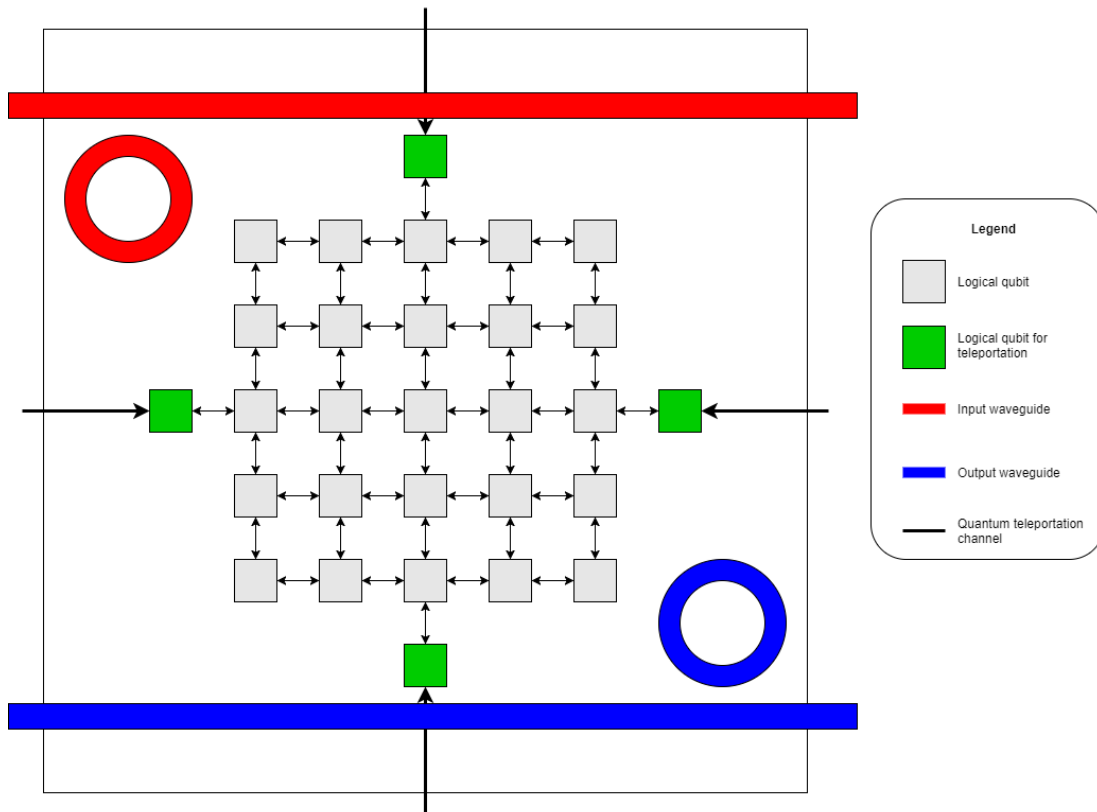


Figure 2.2: Schematic of the structure of a node.

The qubits are arranged in a 2-D mesh with an extra qubit on each side dedicated for quantum teleportation. 2-D mesh is a configuration that makes sense for qubits that are setup on a plane, like superconducting qubits. Moreover, the number of links required for all-to-all interconnect scales quadratically and is thus hard to do when the number of qubits gets large. 2-D mesh has a maximum number of steps between any two qubits on the node of $2\sqrt{N}$, where N is the number of qubits in the mesh. Operations on logical qubits should be parallelized in order to boost performance. QEC operations on the physical qubits need to be parallelized in order to maintain coherence on the qubits.

Since there is a possibility of error when applying QEC, a response has to be sent back to the processor whenever an operation finishes. This means latency between the nodes and the processor is of great importance for the performance of the device. For that purpose, I propose a specialized processor that can control the quantum processors. This processor (Quantum Control Processor), would be connected to the CPUs and memory with the system bus:

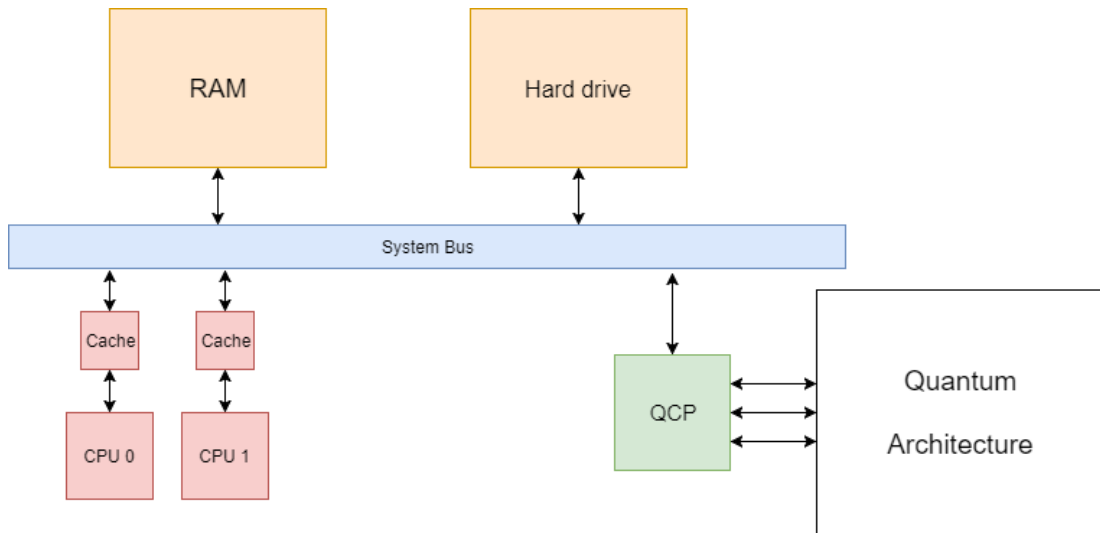


Figure 2.3: Connection with the classical computer.

The QCP would have the following high-level structure:

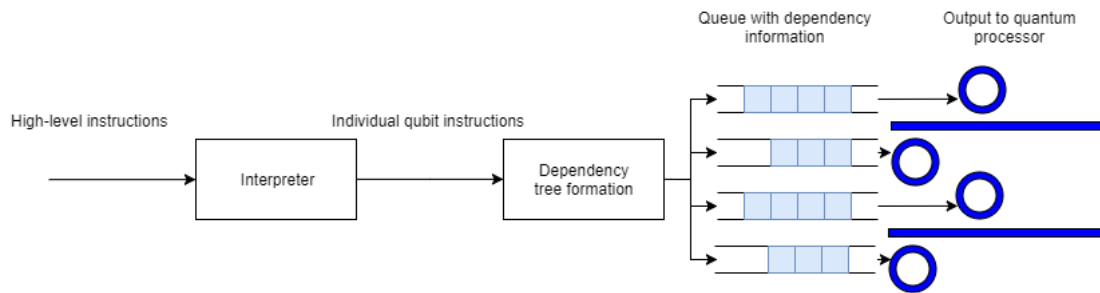


Figure 2.4: QCP pipeline for processing input from the CPU.

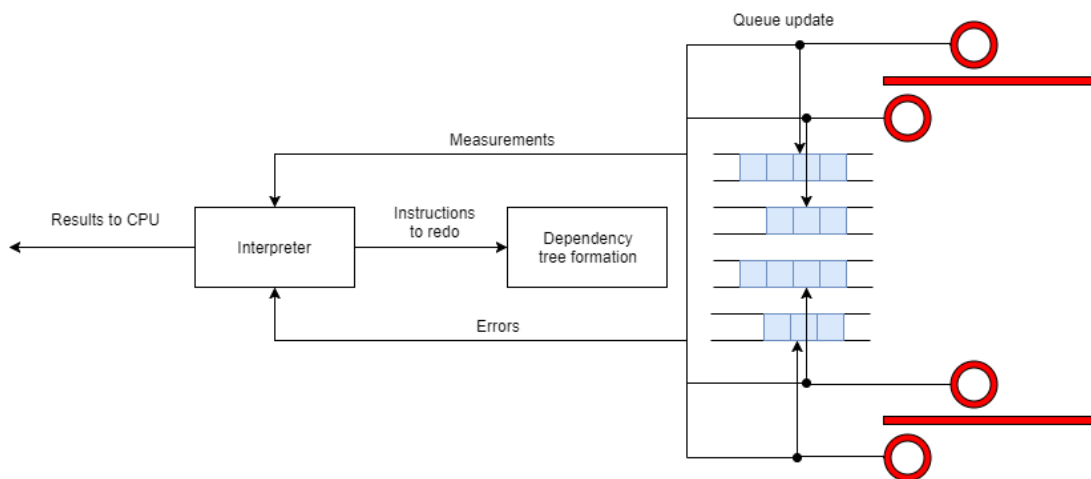


Figure 2.5: QCP pipeline for processing input from the quantum processor.

From the CPUs there would come high-level instructions, such as “Do a 50-qubit QFT”, which a first stage, the interpreter, would break down into individual qubit instructions such as “Apply X gate on qubit #134”, taking into account the layout of the quantum processor and minimizing inter-node communications. These instructions would be separated depending on which node they have to be applied at and they would be arranged into a dependency tree so that the instructions can be put into a queue and be sent optimally to the quantum processor.

From the quantum processor there would come confirmations that the instruction that was being done on a qubit has ended, and the results from measurements. If the instruction finished successfully, the queue would be updated and the next instructions sent. On the other hand, if there were an unrecoverable error, it should be communicated to the interpreter in order to redo the execution of the affected instructions. Measurement results will go back to the interpreter to discern if they have to be sent to the CPU, or used internally.

2.3. Bandwidth requirements

In this section I will estimate the amount of bandwidth this architecture would need with the main classical processor. The necessary bandwidth will vary mainly on the number of qubits of the node and whether the quantum processor can handle error correction by itself and thus no instructions for error correction are needed. Both cases will be presented.

- Case 1: No error correction instructions needed

This case is the simplest. We can assume a single-qubit gate time of t_{1q} , a CNOT gate time of t_{2q} and a measurement time of t_m . We will consider a 2-D grid of logical qubits of $W \cdot L$, so the total number of logical qubits will be $W \cdot L + 4$, as we have 4 qubits on the sides solely for quantum teleportation.

We don't have to consider the effect of there being several nodes here because the data for different nodes will be transmitted through different channels: either a different optical fiber or a different frequency mode in the same fiber.

Each instruction sent to the node has to encode two things:

- Which qubit the gate will be acted on.
- Which gate will be applied.

For the first one, the number of bits needed to encode a number between 1 and $W \cdot L + 4$ is $\lceil \log_2 W \cdot L + 4 \rceil$, whereas for the second we need to count the number of possible operations. There are 10 different possibilities: Measurement of the qubit, Hadamard gate, X gate, Z gate, T gate, CNOT with 4 different

neighboring qubits, and in the case of the qubits for teleportation, quantum teleportation. With 10 possibilities, 4 bits are needed. Since 4 bits allow for up to 16 choices, we could include in the possible operations some non-essential but nonetheless useful gates, such as the inverse of the T gate, the SWAP gate, or the phase gate ($S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$).

The final minimum number of bits for encoding an instruction is:

$$N_{bits} = 4 + \lceil \log_2(W \cdot L + 4) \rceil \quad (2.1)$$

In the worst case, every single qubit is doing a 1 qubit gate operation at the same time, as 1 qubit gates are the fastest. The necessary bandwidth would be:

$$BW_{max} = \frac{N_{bits} \cdot (W \cdot L + 4)}{t_{1q}} \quad (2.2)$$

For example, for a number of qubits of 1000, and a 1-qubit gate time of 1000 ns, the necessary bandwidth would be 14 Gbit/s. This is high, but it's also the worst possible case, and if it isn't met it would just mean the coprocessor is not operating at full capacity.

These calculations are for inbound data rates. Outbound data rates will be smaller, since there are less possibilities. There are only 4 possible outputs for each operation: operation unsuccessful, operation successful and not a measurement, measurement successful, result 0 and measurement successful, result 1.

We can encode that information with only 2 bits, so the responses will be $2 + \lceil \log_2(W \cdot L + 4) \rceil$ bits long, shorter than the instructions. Since the node can only respond to the instructions the central processor has asked for, the outbound bandwidth will always be smaller than the inbound bandwidth.

- **Case 2: Error correction instructions needed**

If QEC cannot be managed by the chip itself, then it is necessary to be able to address every single physical qubit. Let's name the number of physical qubits per each logical qubit as K.

The possible operations will be the same as before, but the number of bits to address the specific qubit will increase.

$$N_{bits_2} = 4 + \lceil \log_2(K \cdot (W \cdot L + 4)) \rceil \quad (2.3)$$

However, this time we do not have to consider the case of all the qubits operating at once. In a paper by A. Petznick and B. W. Reichardt [31] it is described how to prepare $|0\rangle$ states in the $[[7,1,3]]$ and $[[23,1,7]]$ QEC algorithms. For the 7-qubit one, there are 8 CNOT gates in 3 steps, so 2.67 CNOT gates at once on average; whereas for the 23-qubit one there are 57 CNOT gates operating in 7 stages for an average parallelism of 8.14 CNOT gates operating at once. Assuming the load stays similar for all operations, if we were to concatenate the two, in order to achieve a good effective gate error [29], we would have $K = 23 \cdot 7 = 161$ and assuming a similar CNOT gates/physical qubits ratio as before, there would be 57 CNOT gates executing at once per logical qubit.

Assuming a $W \cdot L + 4$ of 100, and a CNOT gate time of 100 ns:

$$BW_{err_{corr}} = \frac{N_{bits_2} \cdot N_{gatesperlogicalqubit} \cdot (W \cdot L + 4)}{t_{2q}} = 798 \text{ Gbps} \quad (2.4)$$

This is too high of a bandwidth per node, considering the node only has 100 logical qubits. The bandwidth can be lowered, with a proportional performance cost, as long as QEC is done before the qubits decohere.

2.4. Possible bottlenecks

The most likely bottleneck in this architecture is the inter-node communication. The node diagram in section 4.3.2 only had qubits in the node can perform quantum teleportation. This can pose an issue if the number of logical qubits in a single node increases. However, it is perfectly possible to add more qubits for quantum teleportation. If it truly were to pose a problem, there could even be quantum teleportation links in every qubit in the side, like this:

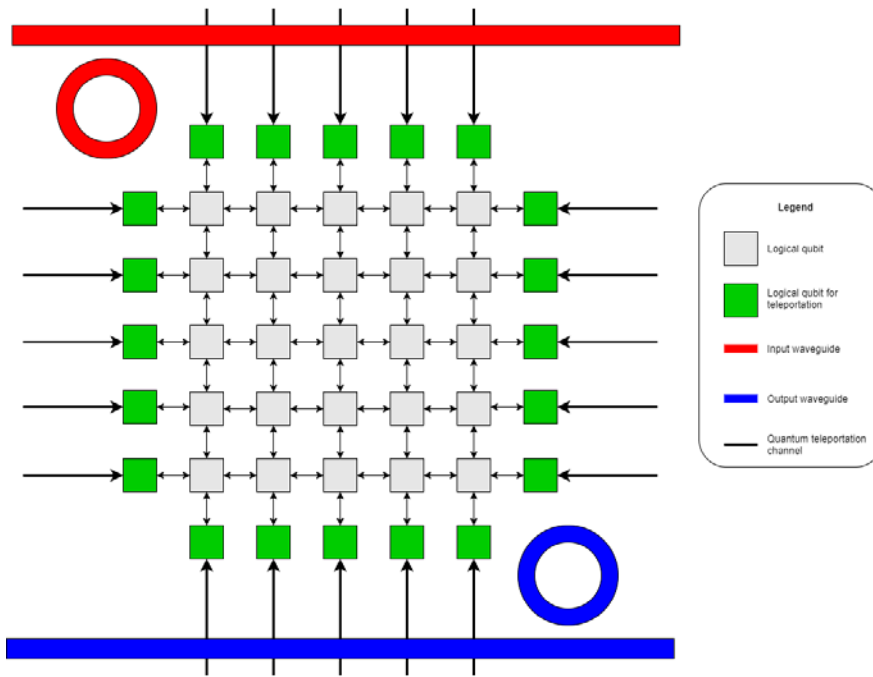


Figure 2.6: Schematic of a node with all the side qubits being dedicated to quantum teleportation.

Nevertheless, even in the last case, the compiler and QCP should minimize the links between nodes, as they probably will end up having higher latency than intra-node communication.

3. Photovoltaic optical modulator

In this section an optical ring modulator is described. First the physics of the device are explained. Then a mechanism that takes advantage of the photocurrent of the device to modulate a signal is discussed and finally it is experimentally measured.

3.1. Photovoltaic effect in P-N junctions

In our case, we do not want to switch the ring modulator with a bias voltage, but rather by short circuiting and open circuiting the ring. The motivation is twofold: it has not been done before, and it could lead to lower power consumption per bit transmitted. In this section the physics on how this works will be explained.

The current that goes through a P-N junction in the ideal case can be expressed as:

$$I(V) = I_0 \cdot \left(e^{\frac{qV}{kT}} - 1 \right) \quad (3.1)$$

Where k is the Boltzmann constant and q the elemental charge.

If the P-N junction is illuminated, electron-hole pairs are generated and cause an additional current source:

$$I(V) = I_0 \cdot \left(e^{\frac{qV}{kT}} - 1 \right) - I_{ph} \quad (3.2)$$

Here, the P-N junction has an open-circuit voltage, which equals:

$$V_{oc} = \frac{\eta kT}{q} \cdot \ln \left(\frac{I_{ph}}{I_0} + 1 \right) \quad (3.3)$$

The unknown variables are I_{ph} , the photonic current; I_0 , the reverse bias saturation current; and η , the ideality factor. I_{ph} depends on the input optical power P_{in} and the responsivity R , measured in A/W:

$$I_{ph} = R \cdot P_{in} \quad (3.4)$$

Silicon does not absorb much IR light (which is why silicon waveguides are possible), so the responsivity will be small. On a similar device [32] a responsivity of 12 mA/W was reported, which would result in a photocurrent of 12 μ A for a P_{in} of 1 mW. For a similar polysilicon P-N junction [33] η was 2.1. With this data, if we assume a reverse bias saturation current of $1 \cdot 10^{-10}$ A, the open circuit voltage would be 636 mV, which seems reasonable. In the next section we will check our assumptions in this area with the experimental results.

3.2. DC Measurements

The device we have tested is a polysilicon ring modulator. The diameter of the ring is 9 μ m, its width is 500 nm and the distance to the waveguide is 250 nm.

To measure the wavelength shift between open-circuit and short-circuit, a 1550 nm tunable laser (HP 8164A) was used. In order to obtain higher optical power, an Erbium-doped fiber amplifier (FITEL ErFA 11000) is added. In order to control the polarization of the light going into the chip, a polarization controller is used. After that, in order to measure the output power, there is a 90/10 coupler. The coupler is necessary because the power of the laser changes in function of the frequency. Therefore, in order to accurately compute the transmission of the device, two measurements are needed: one from the 10% fiber to know the power of the laser, and another one of the device. The tunable laser doubles as photodetector.

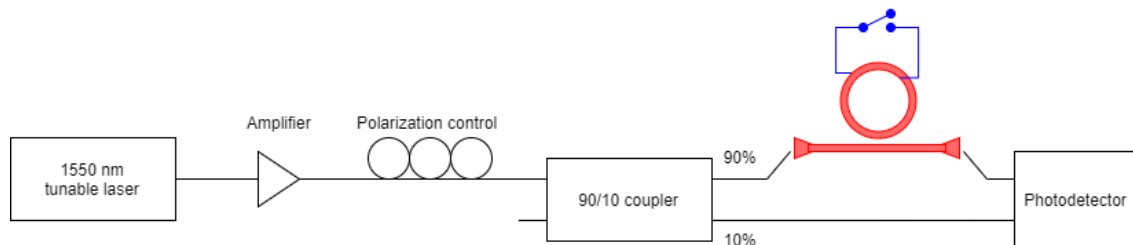


Figure 3.1: Experimental setup for DC measurements. Electrical links are in blue, on-chip waveguides in red and optical links in black.

The device has three metal contacts for a probe to be connected to the circuit. In order to switch between open and short circuit a SMA socket is connected to a probe, which can be short circuited by putting a small metal piece between the contacts.

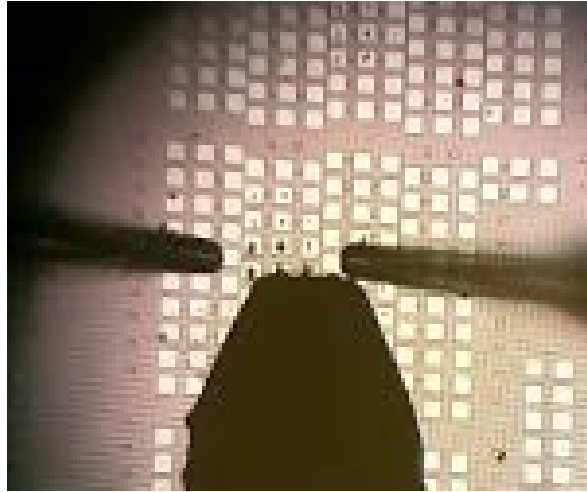


Figure 3.2: Surface of the chip observed through a microscope. Optical power enters the waveguide through the left optical fiber and is collected in the one on the right.

Figure 3.3 shows the wavelength shift between open and short circuit, for 3 and 10 mW of input power. The plotted data is the average of three measurements for each curve.

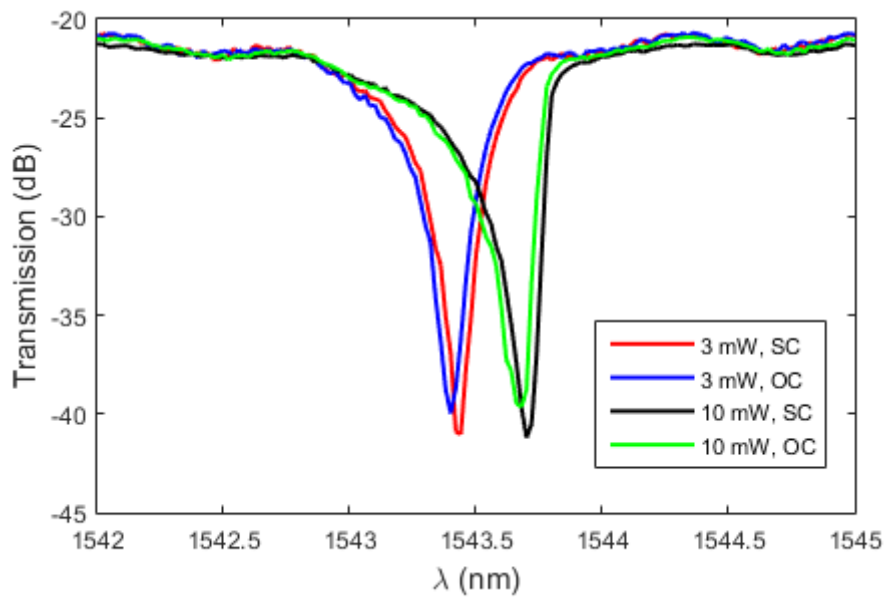


Figure 3.3: Wavelength shift between open and short circuit, for 3 and 10 mW of optical power.

In both cases, there is a wavelength shift of 30 pm, which is the same shift that is observed between bias voltages of 0 and 0.6 V, as predicted by Section 3.1. This is equivalent to a 50 pm/V resonance shift, which is a good figure for ring modulators operating at positive voltages [34]. The Q factor is 3400 both for 3 mW and for 10 mW of input power. We can also see that with 10 mW the resonator acts nonlinearly and the transmission curve is no longer symmetric. For

3 mW of input power, a shift of output power of about 2 dB can be obtained. For 10 mW the plot shows that the shift in transmitted power is about 3 dB, considering that only wavelengths lower than the resonance peak should be used for modulation, in order to prevent instability due to temperature effects.

To compute the parameters of the P-N junction, we do an I-V sweep of the junction with no light.

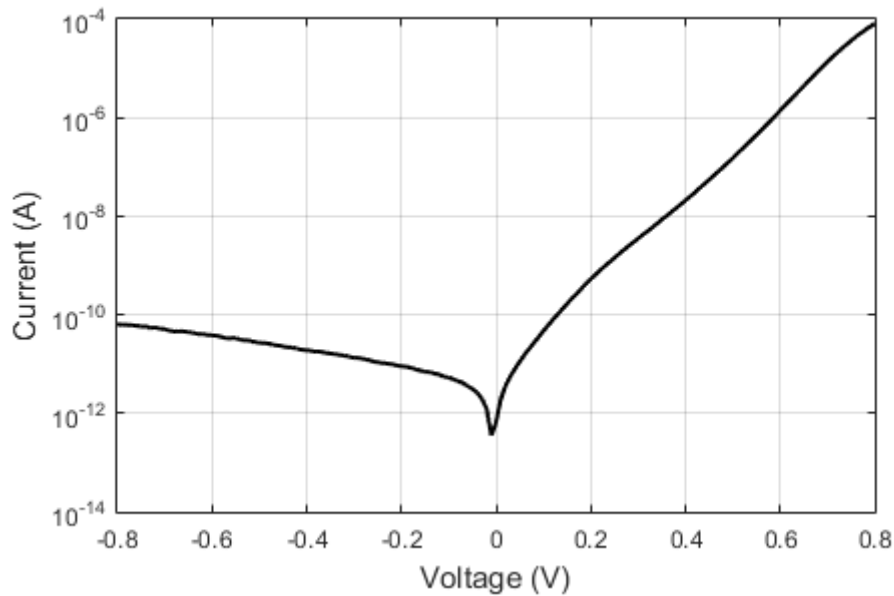


Figure 3.4: I-V curve of the modulator, for no input power. Current is represented in absolute value.

If the ring were an ideal P-N junction, the current for negative voltages should be constant and equal to I_0 , however that is not the case here, and it makes it somewhat harder to compute I_0 . After doing a least-squares fitting, the parameters turned out to be: $I_0 = 5.5 \cdot 10^{-12}$ A, $\eta = 1.86$. The ideality factor is reasonably similar to what we predicted (18% difference), but the reverse bias saturation current is 18 times smaller than the estimation.

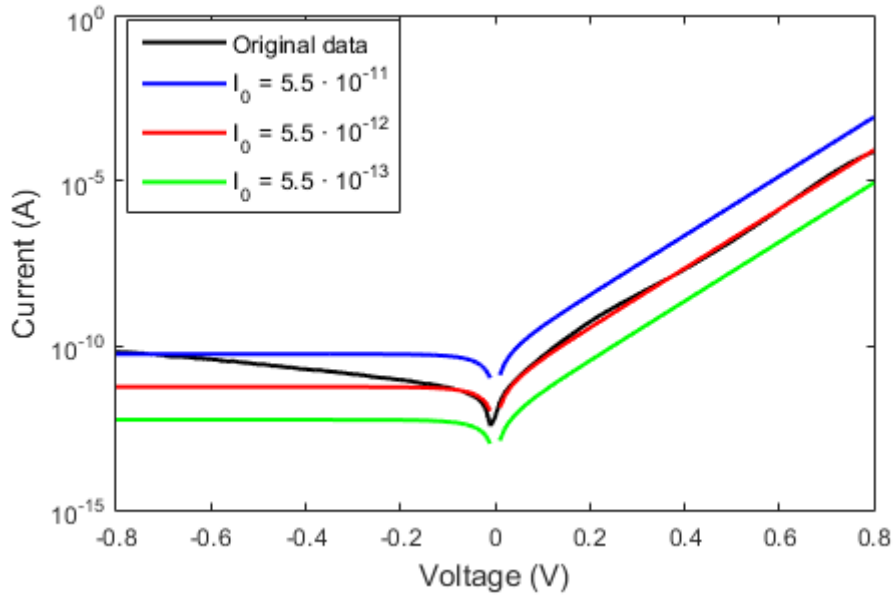


Figure 3.5: Fitting of different curves with $\eta = 1.86$ and several values of I_0 .

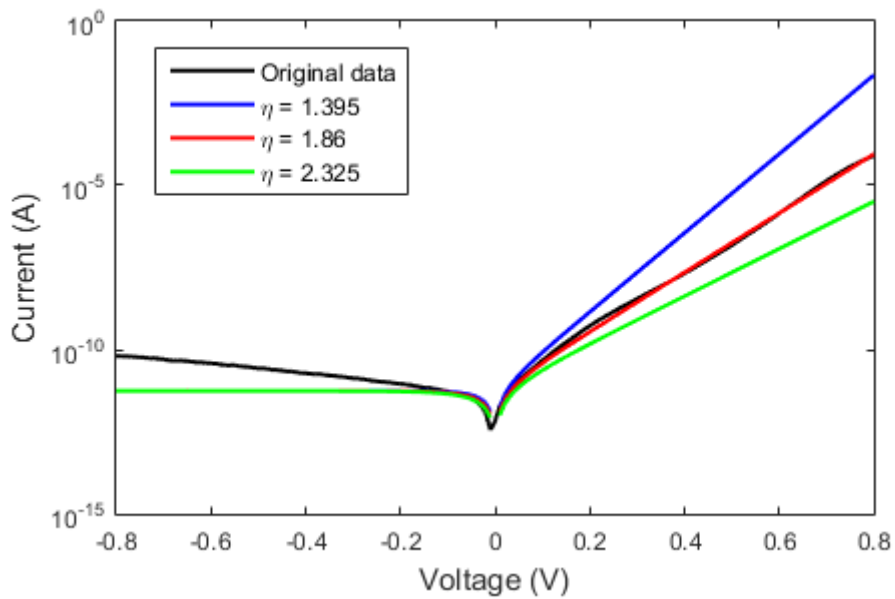


Figure 3.6: Fitting of different curves with $I_0 = 5.5 \cdot 10^{-12} A$ and several values of η .

In both Figures 3.5 and 3.6, it can easily be seen that although the behavior for positive voltages follows closely the equation, that is not the case for negative V.

In order to obtain the responsivity of the cavity, we measure the I-V curves for both of the input powers.

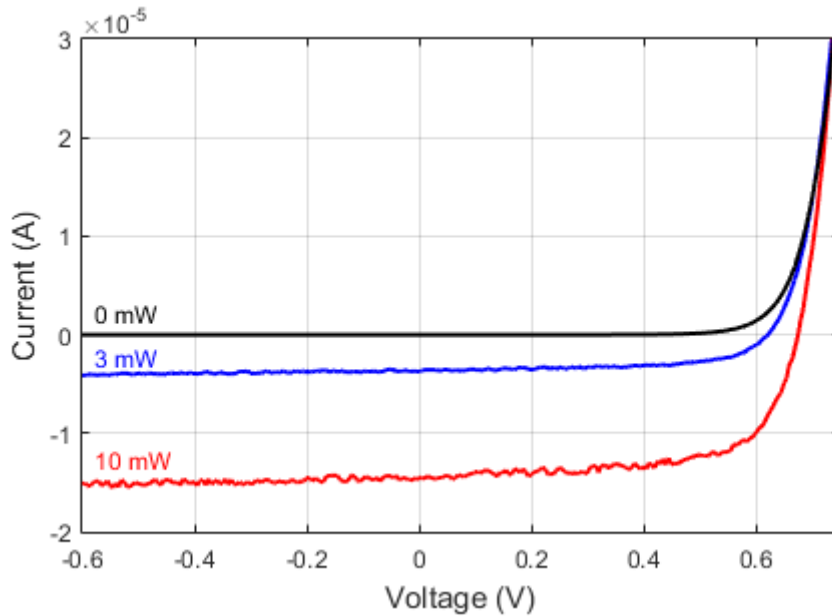


Figure 3.7: IV curve of the modulator, for 0, 3 and 10 mW of optical power, at the respective resonance peaks.

We must take here into consideration that the power that goes through the input optical fiber is not the same P_{in} in our equations in Section 3.1. This is because there is an insertion loss of about 10 dB between the optical fiber and the waveguide on-chip. Therefore, if the input power is 10 mW, P_{in} will be about 1 mW.

As we can see in Figure 3.5, the open circuit voltage of the modulator is 0.62 V for an input power of 3 mW and 0.67 V for 10 mW. We observe a responsivity of 12.3 mA/W for 3 mW of optical power and 14.5 mA/W for 10 mW after taking into account the insertion loss of the waveguide, which are close to the predicted number. Overall, the estimations of the variables of Equation 3.3 balance out and the experimental open circuit voltage is pretty close to what was predicted.

4. Frequency response of PV modulator

The modulator described and measured in the last section must be switched manually between the on and off states. In order to be able to switch using an electrical signal, it is necessary to attach a transistor so that the signal can be directly applied. However, since we are interested in high-frequency modulation, the experimental setup must be designed to have as little parasitic capacitance as possible.

4.1. Transistor-switched modulator

In the DC measurements, the on and off switching was done manually, connecting and disconnecting two terminals. For an actual modulator, this cannot be the case. For this purpose, we will use a transistor as a high-frequency switch. The best one for this approach would be one that has high switching frequency, low activation voltage and low energy consumption. In order to choose the most adequate one, an LTspice simulation was made.

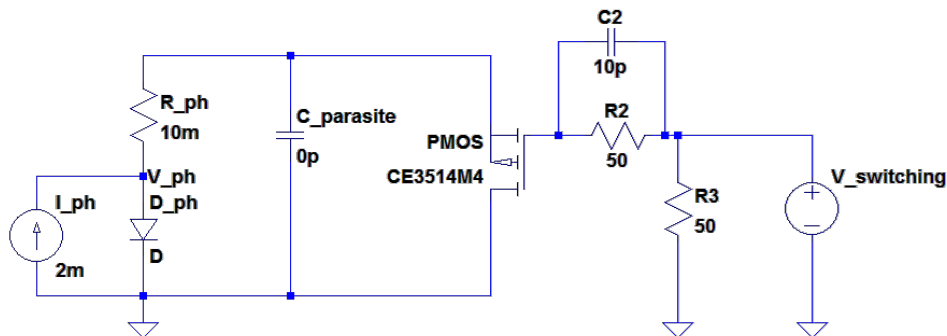


Figure 4.1: LTspice simulation of the transistor-switched resonator.

In this diagram, I_{ph} , D_{ph} and R_{ph} constitute a simplified model for the modulator: the PN junction of the ring works as a diode and has, due to pair generation caused by light absorption, a current source. A small series resistance is added. $C_{parasite}$ is used to study the effects of parasitic capacitance. $R2$, $R3$ and $C2$ are used to couple the AC voltage source $V_{switching}$. In order to choose the best transistor for the task, the internal parameters of PMOS were changed to test different existing models of transistors.

In the simulation, the objective is to alternate V_{ph} between 0 and the open circuit voltage. In the case that the transistor operates as an open circuit, V_{ph} will equal the open circuit voltage. If the transistor is operating as a small resistor, V_{ph} will be close to zero.

Several types and models of transistor were tested, and the transistor chosen in the end was a MOSFET, the CE3514M4, manufactured by California Eastern Laboratories. This transistor was picked because of its IV curves.

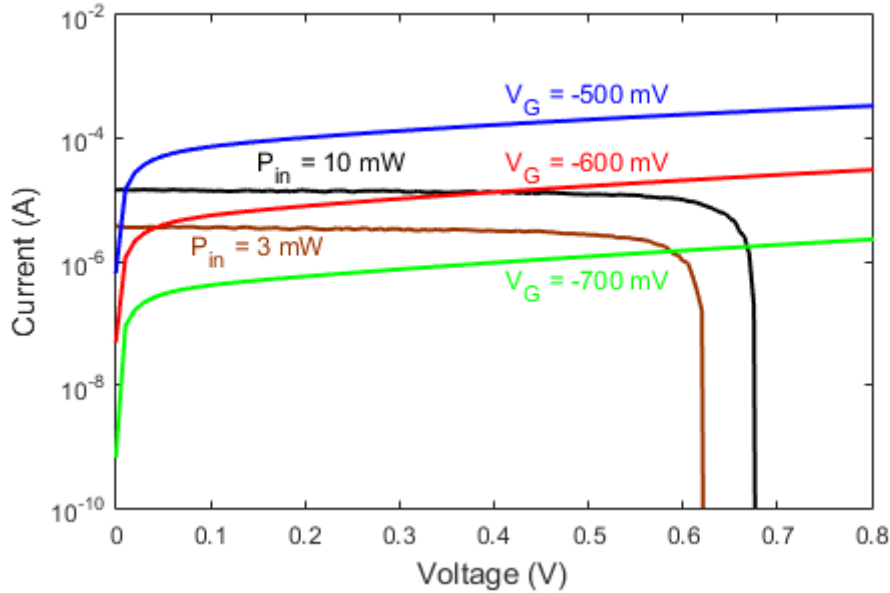


Figure 4.2: Experimental measurements of the IV curves of the CE3514M4 transistor for several gate voltages and of the ring modulator for different amounts of optical power.

As we can see in Figure 4.2, the open circuit / short circuit state of the modulator can be changed with only a peak-to-peak voltage of 0.2 V (Between -0.5 and -0.7 for 10 mW, and between -0.6 and -0.8 for 3 mW).

4.2. Frequency measurements setup

The circuit designed in the previous section was put into a PCB to achieve minimal parasitic capacitance:



Figure 4.3: EAGLE schematic and photo of the PCB.

In the schematic, X_IN and X_OUT are SMA connectors. The size of the board is 25.4 x 8.89 mm.



Figure 4.4: PCB with components mounted on it.

The experimental setup for the AC measurements is similar to the one for the DC measurements, with some added components. Since only one specific frequency is being tested at a time, and the optical amplifier introduces some unwanted noise, a tunable frequency filter is added. The PCB is connected on one end to the probe and on the other end to a wave generator (Agilent 81180A), which is set to generate square waves with arbitrary frequency, bias voltage and peak-to-peak voltage. In order to verify that V_{DS} is varying, a probe is connected on the PCB and connected to the oscilloscope (LeCroy WaveSurfer 422). In order to verify that the output power of the microring is switching, the output of the chip is amplified again and sent into the oscilloscope with the aid of a photodetector (u²t XPDV2320R).

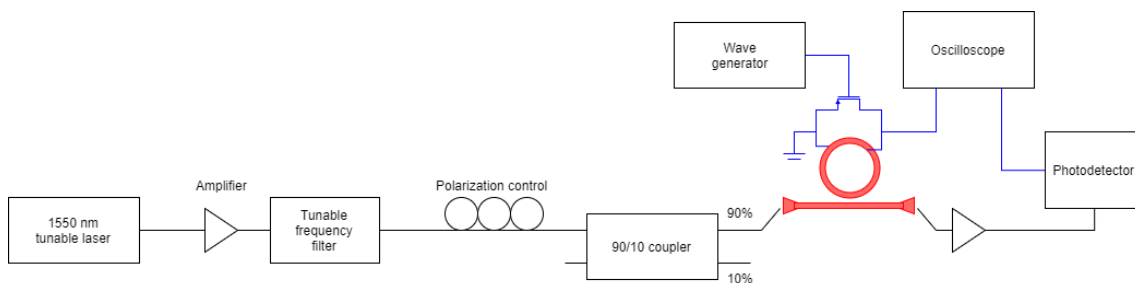


Figure 4.5: Experimental setup for AC measurements. Electrical links are in blue, on-chip waveguides in red and optical links in black.

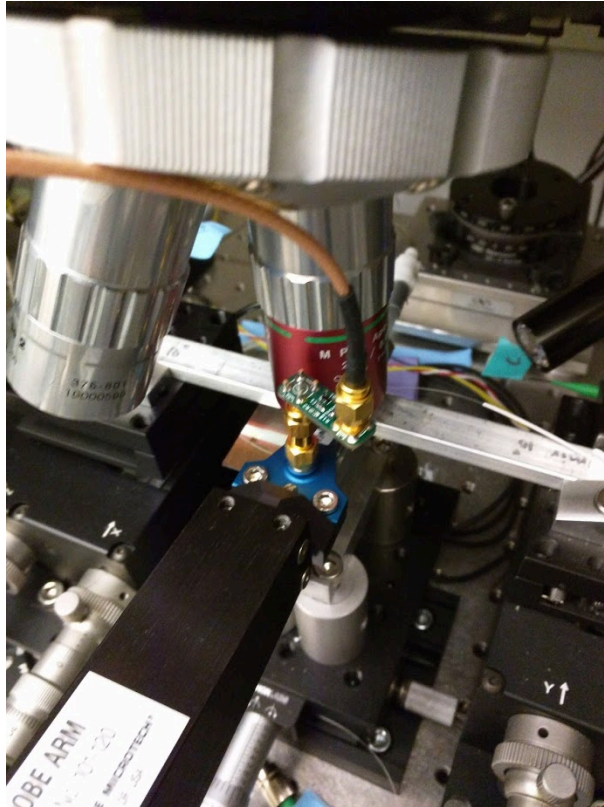


Figure 4.6: Connection of the PCB and the probe on the chip. The probe that measures V_{DS} is not present.

4.3. Results

First the V_{DS} evolution was tested for 10 mW of input power at a low frequency:

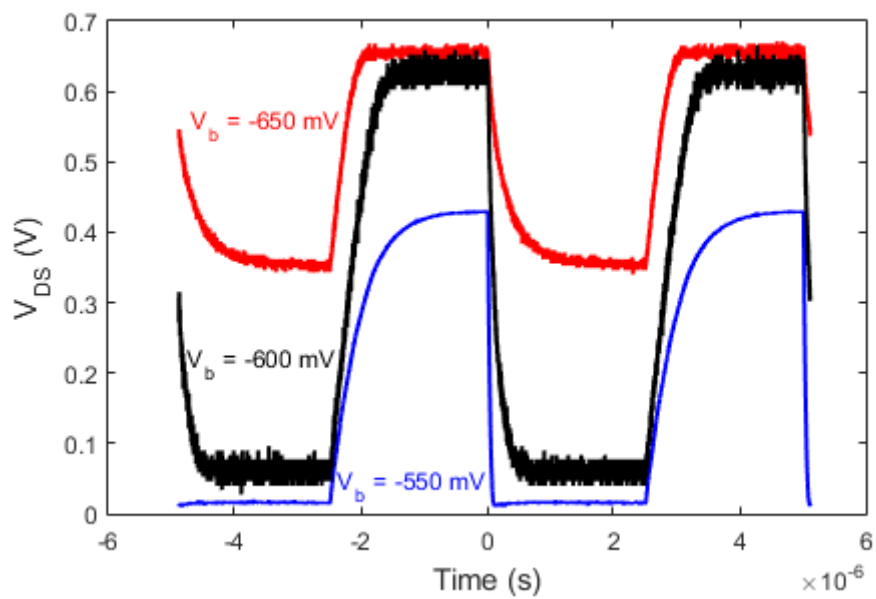


Figure 4.7: V_{DS} evolution for square waves of 200 KHz and peak-to-peak voltage of 100 mV, for different bias voltages.

As we can see in Figure 4.7, offset voltage makes a great difference in the voltage swing. This is a good thing, as this allows us to reduce the peak-to-peak voltage of the input electric signal, and since energy consumed by a digital circuit depends quadratically on the voltage, the consumed energy significantly decreases. We can also notice a big difference between rise times and fall times.

Next we measure the output power. The probe that measured V_{DS} introduced noticeable parasitic capacitance and for that reason it was removed during the output power measurements. We seek to optimize the extinction ratio ($ER = 10 \cdot \log_{10} \frac{P_{max}}{P_{min}}$ dB) for each peak-to-peak voltage chosen.

Extinction ratios of over 3 dB were achieved for peak-to-peak voltages of 200 mV and 100 mV:

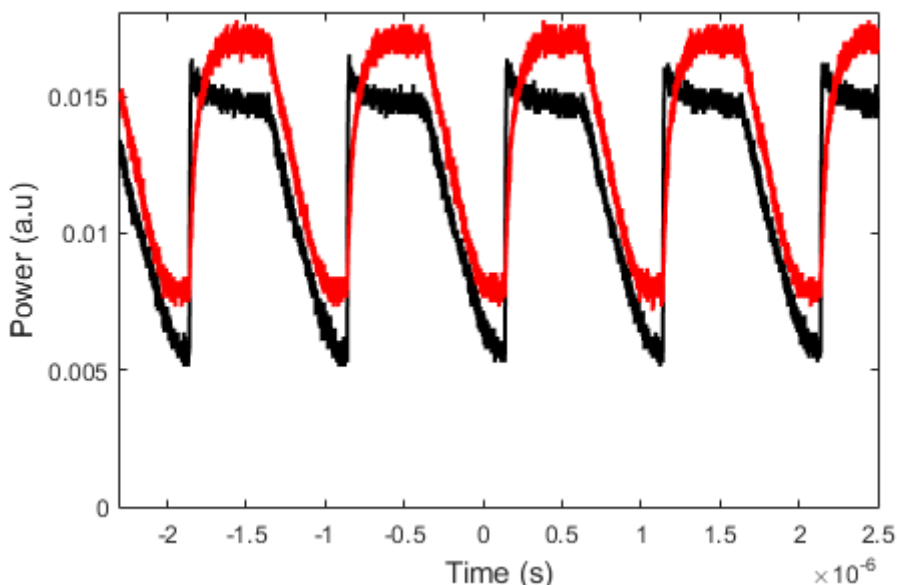


Figure 4.8: Best extinction ratio found for 200 mV peak-to-peak (black) and for 100 mV peak-to-peak (red) at 1 MHz frequency.

The results found were 4.0 dB extinction ratio for 200 mV peak-to-peak, for an optical power of 10 mW, 1543.65 nm frequency and -540 mV bias voltage. For the 100 mV peak-to-peak sample, 3.3 dB extinction ratio was found, with the same parameters as before except that the bias voltage was -610 mV.

In both measurements, we can see a high discrepancy between rise time and fall time. For the 200 mV peak-to-peak one, we can also observe some overshoot.

Reducing the peak-to-peak voltage further gets the extinction ratio below 3 dB, but the results are still interesting:

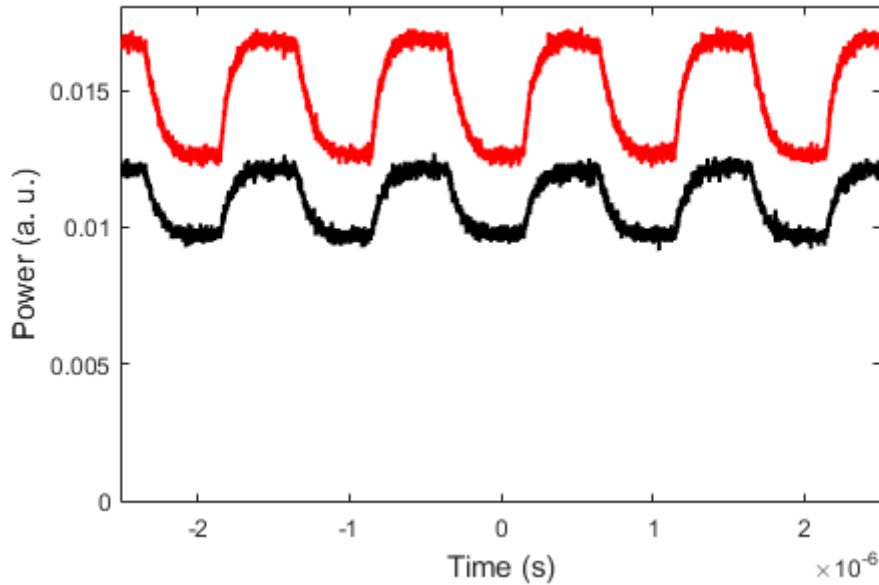


Figure 4.9: Best extinction ratio found for 30 mV peak-to-peak (red) and for 21.9 mV peak-to-peak (black) at 1 MHz frequency.

We achieve 1.2 dB of ER for a peak-to-peak voltage of 30 mV, bias voltage of -583 mV, 14.36 mW of optical power and at a wavelength of 1543.85 nm. For a peak-to-peak voltage of 21.9 mV, an ER of 1.0 dB was observed, with a bias voltage of -586 mV, 15.3 mW of optical power and a wavelength of 1543.8 nm.

Another measurement that was done was directly applying voltages between 0 and 0.7 V to the ring modulator, without using a transistor. This was done for two reasons: to measure the frequency response of the ring modulator itself, and to measure the amplifying effect that the transistor has on the peak-to-peak voltage.

In order to discover the maximum frequency at which the modulator can operate, eye diagrams are done for several frequencies. For this measurement, a different wave generator (Picosecond Pulse Labs SDG Model 12072) and oscilloscope (Agilent DCA-X 86100D) are needed, since the ones used for the previous measurements had a maximum bandwidth of 200 MHz. Figure 4.10 shows the eye diagram for 8 Gbps, and an input signal of 250 mV of bias voltage and 500 mV of peak-to-peak voltage:

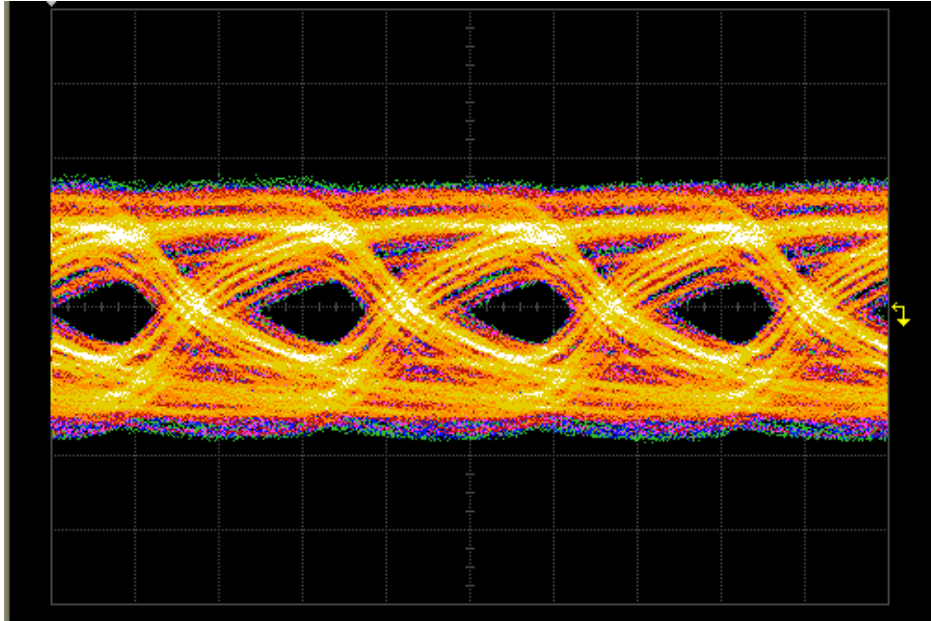


Figure 4.10: Eye diagram of the ring modulator without transistor, at 8 Gbps, modulated with 0.5 V of amplitude.

As we can see in this figure, the 'eye' is barely open, so 8 GHz is (almost) the fastest the device can operate. Taking into account this behavior without the transistor, the frequency bottleneck of the PV modulator must be the parasitic capacitance of the setup.

For 10 mW of input power, extinction ratios for several peak-to-peak voltages have been measured and organized into the following graph:

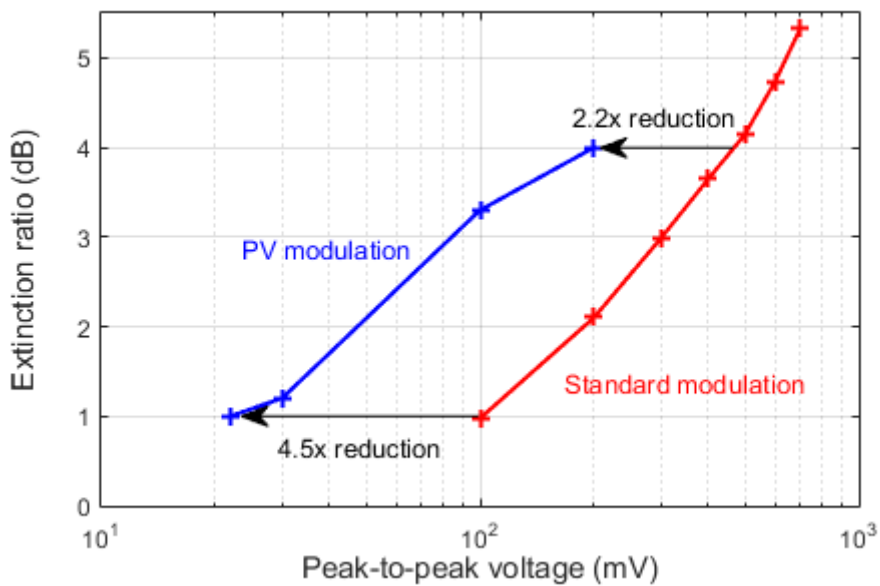


Figure 4.11: Extinction ratio for different peak-to-peak voltages and two modulation methods, for 10 mW of input power.

As we can see, PV modulation achieves similar extinction ratio with peak-to-peak voltages up to 4 times lower than by modulating directly, although it has diminishing returns for higher peak-to-peak voltages.

4.4. Interpretation of the results

The first thing that must be explained about the results is the poor frequency response of the PV modulation, especially compared to the frequency response of just applying the bias voltage directly into the ring without any transistor. In order to explain this, we model the circuit as follows:

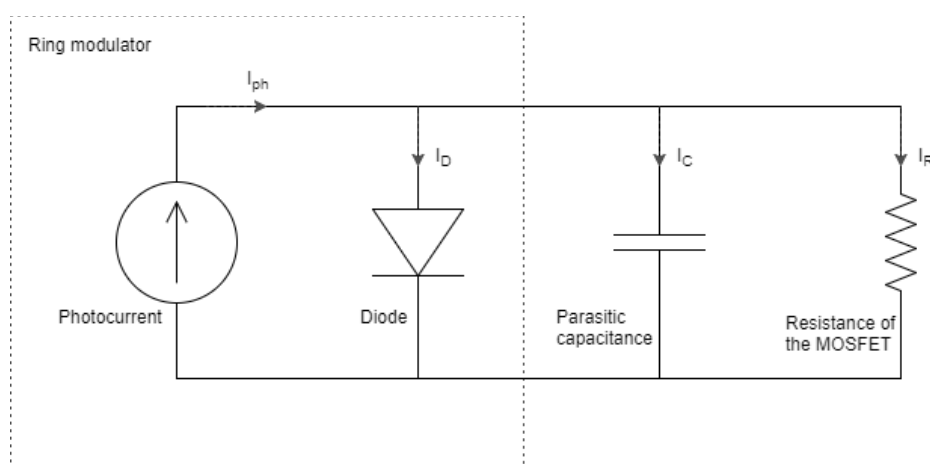


Figure 4.12: Circuit model of the ring resonator and transistor

Here in this model C is the total parasitic capacitance of both the transistor, P-N junction and connections, and R is the resistance of the MOSFET for a certain V_G and V_{DS} , and is calculated as $R = \frac{V_{DS}}{I_{DS}}$. The photocurrent is also assumed constant.

In this model, the voltage between the two terminals V is constant until V_G changes and thus the resistance of the MOSFET changes. From the diagram, we can see that:

$$I_{ph} = I_D + I_C + I_R \quad (4.1)$$

Since the equations that give the current that goes through a diode, capacitor and resistor are known, we can express this as:

$$I_{ph} = I_0 \cdot e^{\frac{V}{\eta \cdot V_T}} + C \frac{dV}{dt} + \frac{V}{R} \quad (4.2)$$

Attempting to solve the resulting ODE, we get:

$$\left(I_{ph} - I_0 \cdot e^{\frac{V}{\eta \cdot V_T}} - \frac{V}{R} \right) \cdot dt = C \cdot dV \quad (4.3)$$

$$t + K = \int \frac{C \cdot dV}{I_{ph} - I_0 \cdot e^{\frac{V}{\eta \cdot V_T}} - \frac{V}{R}} \quad (4.4)$$

Where K is the integration constant. Unfortunately, the integral in Equation 4.4 has no analytical solution, and there is no exact solution. However, some reasonable approximations can be applied here. Up to approximately 0.5 V, as shown in Chapter 3, the IV curve of the ring is mostly flat, which means that I_D is small compared to I_C and I_R . Therefore, for low voltage:

$$t + K = \int \frac{C \cdot dV}{I_{ph} - \frac{V}{R}} = -RC \cdot \ln \left(I_{ph} - \frac{V}{R} \right) \quad (4.5)$$

Solving for V:

$$I_{ph} - \frac{V}{R} = K_2 e^{-\frac{t}{RC}} \rightarrow V(t) = R \cdot I_{ph} - (R \cdot I_{ph} - V(0)) e^{-\frac{t}{RC}} \quad (4.6)$$

Therefore, the time constant of the circuit is $\tau = RC$, the final voltage is $R \cdot I_{ph}$, which is just the new V_{DS} , and the circuit is fundamentally an RC circuit.

This last equation is valid as long as the current that flows through the diode is negligible. For higher voltage values, it will be necessary to model the diode. We can approximate it roughly as a piecewise linear model. For voltages over a activation voltage V_{ON} ,

$$I_D \cong \frac{V - V_{ON}}{r} \quad (4.7)$$

This is equivalent to a V_{ON} voltage source in series with a resistor of resistance r . With this approximation, the ODE is:

$$t + K = \int \frac{C \cdot dV}{I_{ph} - \frac{V - V_{ON}}{r} - \frac{V}{R}} = \int \frac{C \cdot dV}{I_{ph} + \frac{V_{ON}}{r} - \frac{r + R}{rR} V} \quad (4.8)$$

This integral has a similar solution to Equation 4.5:

$$t + K = -\frac{RrC}{R + r} \cdot \ln \left(I_{ph} + \frac{V_{ON}}{r} - \frac{r + R}{rR} V \right) \quad (4.9)$$

$$V(t) = \frac{R \cdot r \cdot I_{ph}}{R + r} + \frac{R \cdot V_{ON}}{R + r} - \left(\frac{R \cdot r \cdot I_{ph}}{R + r} + \frac{R \cdot V_{ON}}{R + r} - V(0) \right) e^{-\frac{R+r}{RrC}t} \quad (4.10)$$

This equation describes another RC circuit, but this time with two resistors in parallel. We can see that for both $V < V_{ON}$ and $V > V_{ON}$ the time constants depend linearly on the parasitic capacitance, which is good news: reducing the parasitic capacitance of the modulation setup by, for example, integrating the transistor into the chip will improve the frequency response by orders of magnitude.

This model is very approximated, since the equivalent resistance of the MOSFET changes at the same time the voltage changes. The most precise way to compute this would be to integrate along the IV curves of both the ring modulator and the MOSFET, but that would not allow us to have a simplified expression for the voltage evolution.

We can test this model with the data from Figure 4.7:

Bias voltage (mV)	V_{low} (mV)	V_{high} (mV)	t_{rise} (ns)	t_{fall} (ns)
-550	15	430	1270	70
-600	60	630	800	360
-650	350	656	450	1100

Figure 4.13: Table with the data from Figure 4.7.

In Figure 4.13 t_{rise} and t_{fall} are computed as the time it takes to reach 95% of the final value, which is equivalent to three time constants. In theory, the V_{high} of the signal with bias voltage of -550 mV and the V_{low} of the signal with bias voltage should be equal, as the gate voltage under those two conditions should be the same. This is not the case here, probably because of some imprecision in the wave generator. The resistance used to compute the t_{fall} of the signal with bias

voltage of -650 mV had to be adjusted for the model to converge to the actual final voltage.

We also consider $V_{ON} = 0.5$ V and that I_D reaches I_{ph} for $V = 0.7$ V, which makes $r = \frac{0.2 V}{I_{ph}}$. I_{ph} is $1.45 \cdot 10^{-5}$ A, as shown in Figure 3.5, considering that the measurements were taken for 10 mW of input power. The resistances of the transistor are calculated from Figure 4.2 and the IV curves of the transistor for gate voltages of -0.55 and -0.65 V, which were not shown in Figure 4.2 for lack of space.

From the first row we can compute the parasitic capacitance of the setup, since the voltage evolution is always below V_{ON} and therefore is modeled as a single exponential. The results are 14 pF and 23 pF. This is a significant difference, so for the rest of the calculations we will take the average of the two, and assume that C is 19 pF.

Now that we have an estimation for the parasitic capacitance, we can check how accurate the model is:

Bias voltage (mV)	t_{rise} (Exp.)	t_{rise} (Model)	% Error	t_{fall} (Exp.)	t_{fall} (Model)	% Error
-550	1270	1716	35%	70	58	-17%
-600	800	1163	45%	360	270	-25%
-650	450	581	29%	1100	1359	24%

Figure 4.14: Comparison of results provided by the model and experimental results.

Overall, for such a simple approximation of the behavior of the P-N junction, the results the model provides are satisfactory. For all the experimental results, the model was able to estimate with up to 45% of error the rise and fall times.

We also want to check whether 19 pF is a reasonable estimation of the parasitic frequency. When measuring V_{DS} , most of the parasitic capacitance comes from the probe. We can know this because measurements with a probe had to be done at 200 KHz, like in Figure 4.7; whereas measurements without a probe could have a frequency of 1 MHz, like in Figures 4.8 and 4.9. According to our model, this implies that the parasitic capacitance due to the probe is 80% of the total parasitic capacitance.

The probe was connected to a coaxial cable of about 1 m of longitude. These kind of cables have a parasitic capacitance of between 50 and 100 pF per meter. Since the probe represents 80% of the capacitance, this means that the total parasitic capacitance of the circuit is between 62.5 pF and 125 pF. Therefore, the capacitance estimate of our model is too low, but within an order of magnitude of the real value.

The most positive result from the experiments is the low peak-to-peak voltage: 100 mV of peak-to-peak voltage on the gate of the transistor is amplified and results in 3.3 dB of extinction ratio, whereas applying 300 mV of peak-to-peak voltage directly results in only 3 dB of ER. For even lower voltage differences we still observe a noticeable extinction ratio. This is thanks to the IV curve of the chosen transistor, which amplifies a lot the signal coming through V_{GS} .

From these results we can estimate the consumed energy per bit, assuming that the minimum ER needed is 3 dB. The consumed power was the 10 mW of optical power used (since the wavelength of the laser was very close to the resonance peak, meaning that the output power was several orders of magnitude smaller than the input power, and thus negligible) and $C \cdot \Delta V^2 \cdot \nu$ of electrical power. Since the frequency of the signal was 200 KHz, the power per bit will be:

$$E_{1 \text{ bit}} = \frac{P_{\text{optical}} + P_{\text{electrical}}}{\nu} = \frac{0.01}{200 \cdot 10^3} + 100 \cdot 10^{-12} \cdot 0.1^2 = 50 \frac{\text{nJ}}{\text{bit}} \quad (4.11)$$

This is not close to the state-of-the-art. However, if we considered that the transistor was on-chip and therefore its parasitic capacitance were on the order of 10 fF, then the maximum frequency, as explained in this section, would increase by a factor of 10000 to 2 GHz. Assuming as well that the loss entering the waveguide is only 1 dB instead of 10 dB, we would get:

$$E_{1 \text{ bit}} = \frac{P_{\text{optical}} + P_{\text{electrical}}}{\nu} = \frac{0.00126}{2 \cdot 10^9} + 10 \cdot 10^{-15} \cdot 0.1^2 = 630 \frac{\text{fJ}}{\text{bit}} \quad (4.12)$$

This is a much better result, although still inferior to the state-of-the-art. The limiting factor is the optical power, which constitutes more than 99.9% of the total.

5. Conclusions and future work

5.1. Conclusions

In Chapter 2, the current technical challenges for quantum computation were presented. Aiming to partially solve this issues, a distributed quantum computing architecture was presented, together with the interfacing with a classical computer. Finally, bandwidth requirements per node were discussed and estimated. It was shown that if the quantum coprocessors can perform QEC by themselves without needing communication with the classical processor, it could reduce the necessary bandwidth between quantum coprocessors and classical processor by approximately a factor of 50.

In Chapter 3 it was shown that the modulator behaves as an ideal diode which is generating photocurrent due to absorption of light inside the microring. This photocurrent produces a positive voltage. With this voltage, short-circuiting and open-circuiting the ring can produce a noticeable resonance shift. We also have observed non-linearities for high input power, which make the transmission curve asymmetrical. These non-linearities are probably due to the dissipation of heat inside the ring, which increases its temperature and shifts its resonance peak to the right.

In Chapter 4, the frequency response of the device under standard modulation and under PV modulator was measured. High extinction ratios with low peak-to-peak voltage were found: with only 100 mV of peak-to-peak voltage, 3.3 dB of extinction ratio was observed with PV modulation, which compares to 1.0 dB of ER for the same peak-to-peak voltage with standard modulation. Moreover, a circuit model for PV modulation was presented, showing that the frequency response is inversely proportional to the parasitic capacitance. This model was tested using the experimental data and showed a maximum deviation of 45%.

5.2. Future work

There are several directions in which the quantum computing architecture work can be expanded. For example, models of communication overhead can be built, and with those models estimate the effect of different network topologies and QEC algorithms. Another interesting result would be to describe an efficient algorithm for resolving the dependencies of quantum operations that the QCP could implement in hardware and with it estimate the average latency between receiving confirmation of the end of one instruction, and sending the next one.

There are three more experiments that should be done with the PV modulator described in this work for it to be a viable solution for low-energy data modulation and for quantum computation.

First, in order to reduce the parasitic capacitance, an on-chip transistor could be used instead of a PCB. This should decrease the capacitance by orders of magnitude. As seen in Section 4.4, this should in turn decrease the rise and fall times proportionally.

Second, low-temperature measurements should be performed to make sure that the PV modulator can work at cryogenic temperatures. Under such conditions, it would also be interesting to measure how much thermal energy is released by the device.

Third, this PV modulator has not been tested as a photodetector. In order to verify that this device can be used to transform optical signals into electrical signals, the photocurrent for different input power and wavelength should be measured, as well as the frequency response.

Another possible improvement would be to use a different material for the modulator. Although the modulator used had a good wavelength shift per volt, the Q factor is comparatively low. A modulator with a similar wavelength shift but a 10 times larger Q factor would be able to achieve much higher extinction ratios while using both lower peak-to-peak voltage to modulate the signal, and lower laser power.

References

- [1] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, Oct. 1997.
- [2] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*, 1996, pp. 212–219.
- [3] R. Van Meter, T. D. Ladd, A. G. Fowler, and Y. Yamamoto, "Distributed Quantum Computation Architecture Using Semiconductor Nanophotonics," 2009.
- [4] R. W. Keyes and R. Landauer, "Minimal Energy Dissipation in Logic," *IBM Journal of Research and Development*, vol. 14, no. 2, pp. 152–157, Mar. 1970.
- [5] T. Toffoli, "Reversible computing," in *Automata, Languages and Programming*, 1980, pp. 632–644.
- [6] E. Fredkin and T. Toffoli, "Conservative logic," *International Journal of Theoretical Physics*, vol. 21, no. 3, pp. 219–253, 1982.
- [7] C. M. Dawson and M. A. Nielsen, "The Solovay-Kitaev algorithm," *arXiv preprint*, 2005.
- [8] M. A. Nielsen, "A simple formula for the average gate fidelity of a quantum dynamical operation," *Physics Letters A*, vol. 303, no. 4, pp. 249–252, Oct. 2002.
- [9] A. M. Steane, "Overhead and noise threshold of fault-tolerant quantum error correction," *Physical Review A*, vol. 68, no. 4, p. 42322, Oct. 2003.
- [10] N. M. Linke *et al.*, "Experimental comparison of two quantum computing architectures," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3305–3310, Mar. 2017.
- [11] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6–7, pp. 467–488, Jun. 1982.
- [12] B. P. Lanyon *et al.*, "Towards quantum chemistry on a quantum computer," *Nature Chemistry*, vol. 2, no. 2, pp. 106–111, Feb. 2010.
- [13] A. Kandala *et al.*, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, vol. 549, no. 7671, pp. 242–246, Sep. 2017.
- [14] L. M. Adleman, C. Pomerance, and R. S. Rumely, "On Distinguishing Prime Numbers from Composite Numbers," *The Annals of Mathematics*, vol. 117, no. 1, p. 173, Jan. 1983.
- [15] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, Feb. 1978.
- [16] E. Lucero *et al.*, "Computing prime factors with a Josephson phase qubit

- quantum processor,” *Nature Physics*, vol. 8, no. 10, pp. 719–723, Aug. 2012.
- [17] D. G. Rabus, “Ring Resonators: Theory and Modeling,” in *Integrated Ring Resonators*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 3–40.
- [18] E. Timurdogan, C. M. Sorace-Agaskar, J. Sun, E. Shah Hosseini, A. Biberman, and M. R. Watts, “An ultralow power athermal silicon modulator,” *Nature Communications*, vol. 5, pp. 1–11, Jun. 2014.
- [19] R. Soref and B. Bennett, “Electrooptical effects in silicon,” *IEEE Journal of Quantum Electronics*, vol. 23, no. 1, pp. 123–129, Jan. 1987.
- [20] O. V. Lounasmaa, “Dilution refrigeration,” *Journal of Physics E: Scientific Instruments*, vol. 12, no. 8, pp. 668–675, Aug. 1979.
- [21] A. P. Chandrakasan and R. W. Brodersen, “Minimizing Power Consumption in Digital CMOS Circuits,” *Proceedings of the IEEE*, vol. 83, no. 4, Apr. 1995.
- [22] M. Georgas *et al.*, “A monolithically-integrated optical transmitter and receiver in a zero-change 45nm SOI process,” in *2014 Symposium on VLSI Circuits Digest of Technical Papers*, 2014, pp. 1–2.
- [23] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, “Demonstration of a small programmable quantum computer with atomic qubits,” *Nature*, vol. 536, no. 7614, pp. 63–66, 2016.
- [24] A. Steane, “Multiple-Particle Interference and Quantum Error Correction,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 452, no. 1954, pp. 2551–2577, Nov. 1996.
- [25] J. Preskill, “Reliable quantum computers,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1969, pp. 385–410, Jan. 1998.
- [26] E. Knill, “Quantum computing with realistically noisy devices,” *Nature*, vol. 434, no. 7029, pp. 39–44, Mar. 2005.
- [27] N. C. Jones *et al.*, “Layered Architecture for Quantum Computing,” *Physical Review X*, vol. 2, no. 3, p. 31007, Jul. 2012.
- [28] C. Monroe *et al.*, “Large-scale modular quantum-computer architecture with atomic memory and photonic interconnects,” *Physical Review A*, vol. 89, no. 2, p. 22317, Feb. 2014.
- [29] R. D. Van Meter, “Architecture of a Quantum Multicomputer Optimized for Shor’s Factoring Algorithm,” *arXiv preprint*, 2006.
- [30] T. P. Spiller, K. Nemoto, S. L. Braunstein, W. J. Munro, P. van Loock, and G. J. Milburn, “Quantum computation by communication,” *New Journal of Physics*, vol. 8, no. 2, pp. 30–30, Feb. 2006.
- [31] A. Paetznick and B. W. Reichardt, “Fault-tolerant ancilla preparation and noise threshold lower bounds for the 23-qubit Golay code,” *Quantum Information and Computation*, vol. 12, no. 11–12, pp. 1034–1080, 2011.
- [32] M. Yang *et al.*, “Boosting Hot-Electron Extraction Through Deep Groove

- Perfect Absorber for Si-Based Photodetector," *IEEE Photonics Technology Letters*, vol. 29, no. 21, pp. 1884–1887, Nov. 2017.
- [33] E. C. Garnett and P. Yang, "Silicon Nanowire Radial p–n Junction Solar Cells," *Journal of the American Chemical Society*, vol. 130, no. 29, pp. 9224–9225, Jul. 2008.
- [34] X. Xiao *et al.*, "25 Gbit/s silicon microring modulator based on misalignment-tolerant interleaved PN junctions," *Optics Express*, vol. 20, no. 3, p. 2507, Jan. 2012.