All-optical computing architectures in silicon photonics

Thesis by

Rahul Trivedi

In partial fulfilment for the requirements of the degree of

Bachelor of Technology in **Electrical Engineering**

Department of Electrical Engineering

Indian Institute of Technology Delhi

April, 2016

Abstract

All-optical computing architectures can potentially outperform conventional electronic computing architectures due to the significantly higher speeds of propagation of optical signals ($\sim 10^8$ m $\cdot$ s^{-1}) as compared to those of electrical signals ($\sim 10^5$ m $\cdot$ s^{-1}). One of the primary challenges in realizing optical information processing systems is that of achieving on-chip optical nonlinearity at low power. A number of systems, including cavity enhanced $\chi^{(2)}$ and $\chi^{(3)}$ nonlinear systems, have been proposed to achieve on-chip nonlinearity and have subsequently been used to design all-optical logic circuits. However, experimentally realization such systems that can perform any reasonably complex information processing task is technologically challenging as it involves the fabrication of a large number of high quality factor cavities (in case of $\chi^{(2)}$ nonlinearity, the systems need to be fabricated on III-V materials).

Silicon photonic systems, which can be fabricated using standard CMOS processes, are technologically more amenable to implementing large information processors. In this thesis, we present the design of all-optical computing architectures using silicon photonics components such as ring resonators, waveguides, integrated photo-detectors etc. The central nonlinear block used in our design is a self-electrooptic cavity, which comprises of a ring resonator with an integrated photo-detector with the nonlinearity being induced due to an electrical feedback into the optical system. We derive a time-domain simulation model to compute the characteristics of this nonlinear device. We also present the design and analysis of all-optical logic gates (AND gate and NOT gate) and digital memory elements (latch). Moreover, we also investigate the design of an all-optical neuron, which can potentially be employed for the implementation of fast optical neural networks and use it to solve a classification problem.

Acknowledgements

I would like to thank my advisers, Dr. Arka Majumdar and Dr. Uday Khankhoje, for providing me with the opportunity to investigate quantum optics through this project. This project would not have been successful and conclusive without their guidance and support. I would also like to thank the Electrical Engineering department of IIT Delhi for allowing me to pursue this project as my undergraduate thesis. Finally, I would like to thank the committee members who evaluated my thesis, their encouragement, observations and constructive feedback has helped me improve the quality of work presented in this thesis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction and past work

Ever increasing demand for fast information processing systems and recent advances in nano fabrication [1,2] have inspired research into alternative computing paradigms (such as quantum computing [3, 4] and neuromorphic computing [5]) and alternative physical platforms for implementing these system. Amongst the different physical platforms, using integrated optical systems [6–8] have been widely investigated and used due to the extremely high speeds of optical waves ($\sim 10^8$ m $\cdot$ s^{-1}). In terms of propagation delays, optical information processing systems are orders of magnitude faster than the more commonly employed electronic information processing systems [9], which are limited by the saturation velocity of the electrons ($\sim 10^5$ m $\cdot$ s^{-1}).

However, optical systems are typically linear systems, with a very weak nonlinear response. Linear systems alone cannot be used to build versatile information processors – for instance, even basic logic gates operate nonlinearly on the inputs. It has been shown theoretically and experimentally that use of resonators, which are dielectric structures that can confine optical fields to small volumes, can enhance the optical nonlinearity thereby reducing the power consumption of the designed system [10, 11]. Several implementations of optical information processing systems (including digital information processing) based on cavity enhanced second ($\chi^{(2)}$) [11] and third ($\chi^{(3)}$) [12, 13] order nonlinearity have been proposed. Apart from the inherent dielectric nonlinearity (e.g. $\chi^{(2)}$ and $\chi^{(3)}$ nonlinearity), it has also been shown that coupling quantum emitters,

such as quantum dots [14], colour centres [15] etc., to optical resonators can make the composite system strongly nonlinear. These advances have indeed made the prospect of building an all-optical computer significantly achievable.

Most nonlinear systems, however, are not silicon compatible. For instance, strong $\chi^{(2)}$ nonlinearities are often found in III-V systems (such as GaAs or GaP) [16]. Even quantum emitters are fabricated out of non-silicon systems. Silicon does have a $\chi^{(3)}$ nonlinearity, but this nonlinear effect is very weak and using this nonlinearity at sufficiently low input powers requires the fabrication of ultra-high quality factor cavities [11, 17]. Although nonlinear devices in non-silicon platforms have been fabricated and demonstrated [18, 19], silicon based systems are more amenable to large-scale on-chip integration. This is partly due to the availability of well developed large scale silicon-based fabrication processes. Silicon photonics fabrication processes can enable easy on-chip fabrication of photo-detectors, static beam splitters and phase shifters. Integrating on-chip light sources is hindered by the indirect bandgap of silicon, however it is possible to route laser signals into the system through coupled optical fibers. Designing computing systems on such a silicon photonics platform would thus enable a speedy and large-scale realization of an optical computer.

In this thesis, we investigate the problem of designing computing architectures in silicon photonics. The fundamental nonlinear block used in our design is the self-electrooptic cavity [20]. This device uses an electrically modulated ring-resonator which can be typically constructed by integrating a reverse-biased p-n junction in the ring of the resonator. The depletion width of this junction can be modulated by changing the reverse-biased voltage – this change in the depletion region changes the cavity refractive index profile and hence achieves a modulation of the resonant frequency of the ring resonator with the reverse-biased voltage. By employing the p-n junction as a photo-detector in addition to a refractive index modulator and giving an electrical feedback to the reverse bias, the overall optical characteristics of the cavity becomes nonlinear. Unlike $\chi^{(2)}$ and $\chi^{(3)}$ nonlinear systems, the nonlinearity of this electrooptic

system can be tuned by changing the electrical circuit – this allows for a more versatile network design. With this element as the central building block, we design all-optical combinational and sequential digital circuit elements that are a prerequisite for building complex digital information processing systems.

The integrated optics platform is also suited for the design of a number of specialised computing systems. These include neural networks and neuromorphic learning machines [21], which attempt to mimic the functionality of a biological brain, and typically adjust their function on the basis of a set of training data samples. Although their usability is limited only to certain specific applications, using optics can potentially speed up such systems by several orders of magnitude. Using the self-electrooptic cavity, we also present the design of a single optical neuron which can form the basis for implementing high-speed large-scale neural networks.

This thesis is organised as follows – chapter 2 is devoted to the analysis of the self-electrooptic cavity and other static linear devices (phase shifters, beam splitters) used in the designs. A time-domain simulation model for describing this cavity is derived and the steady state input-output characteristics of this cavity are computed using this simulation model. Chapter 3 shows how to design digital circuit elements (AND gate, NOT gate and latch) using the self-eletrooptic cavity. Chapters 4 and 5 are devoted to design and optimisation of the optical neuron.

Chapter 2

Optical circuit elements and their simulation models

This chapter is devoted to describing the different optical circuit elements that we use in designing computing architectures in the subsequent chapters. We also develop time-domain simulation models for the nonlinear and linear circuit elements used in our designs. The circuit elements detailed in this chapter include the self-electrooptic cavity and linear circuit elements such as static directional couplers and phase shifters. Section 2.1 elucidates the simulation model for the self-electrooptic cavity and its input-output characteristics are presented in section 2.2. The simulation models for linear circuit elements (phase shifters and directional couplers based on channel waveguides) are described in section 2.3.

2.1 Simulation model for the nonlinear cavity

The basic-setup that we use to develop the simulation model is shown in Fig. 2.1. We assume that the cavity is coupled to only a single waveguide – it thus has only one input and one output port. The model developed can easily be extended to more complicated systems such as systems with multiple coupled waveguides or electrically coupled cavities – some of these systems are analyzed in detail as and when they are

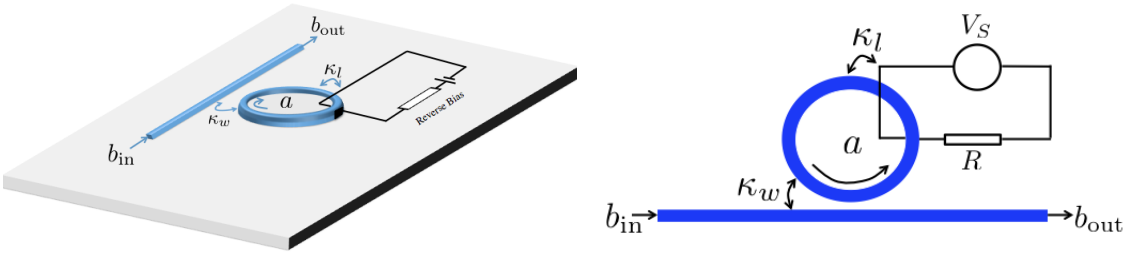


Figure 2.1: Schematic of the self-electrooptic cavity coupled to a single waveguide with a reverse biased p-n junction in the ring.

required in the subsequent designs. The cavity has a photo-detector integrated within the ring, which is reverse biased through an external voltage source. We assume that the cavity input is modulated onto a reference laser frequency ω_L, which is sufficiently close to the cavity resonant frequency – i.e. it is of the form $b_{\text{in}}(t)\exp(-i\omega_L t)$. The cavity dynamics is described by a complex function of time, $a(t)$ which is proportional to the slowly varying amplitude of the cavity electric field around frequency ω_L (i.e. the cavity electric field will have a time dependance of $a(t)\exp(-i\omega_L t)$). It follows from standard coupled mode theory that $a(t)$ satisfies the following dynamical equation [20, 22]:

$$\frac{\mathrm{d}a(t)}{\mathrm{d}t} = -\left(i\Delta + \frac{\kappa}{2}\right)a(t) - \sqrt{\kappa_w}b_{\text{in}}(t) \tag{2.1}$$

where $\Delta = \omega - \omega_L$ is the detuning of the reference laser frequency ω_L from the cavity resonant frequency ω, and would depend on $a(t)$ due to electrical circuit coupled to the optical cavity. $\kappa = \kappa_l + \kappa_w$ captures the losses in the cavity, and can be considered as a sum of two different sources of loss in the cavity – κ_w corresponding to the coupling between the cavity and the input-output waveguide and κ_l corresponding to the power absorbed by the integrated p-n junction, thereby resulting in the generation of electron-hole pairs and hence a photo-current in the photodetector. The waveguide output $b_{\text{out}}(t)\exp(-i\omega_L t)$ is a sum of the waveguide input and radiation by the cavity: $b_{\text{out}}(t) = b_{\text{in}}(t) + \sqrt{\kappa_w}a(t)$. In order to complete our description of the simulation model, we need to analyze two different aspects of the nonlinear cavity:

1. The dependance of ω or Δ on the p-n junction depletion width.

2. The dynamics of the photo-generation process.

2.1.1 Modulation of ω by the p-n depletion width

Throughout our analysis, we assume a linearly graded junction to model a realistically fabricable device. The doping density of the cavity can be expressed as [23]:

$$N_D(r) - N_A(r) = \begin{cases} 2N_0(r-R)/d & |r-R| < d/2 \\ 0 & |r-R| > d/2 \end{cases} \tag{2.2}$$

where d is the ring width and N_0 is the maximum net doping density in the p-n junction. The depletion width (w_D) depends on the reverse-bias voltage (V_D) across the cavity through the following well-known equation [23]:

$$w_D = \left[\frac{6\epsilon_0 n_{\text{Si}}^2 d}{N_0 e}(V_D + V_{\text{bi}})\right]^{1/3} \tag{2.3}$$

where n_{Si} is the refractive index of silicon and $V_{\text{bi}} \sim 2kT\log(N_0/n_i)/e$ is the diode built-in voltage. Schematic plots of the doping density and the electron and hole carrier densities ($\mu_n(r)$ and $\mu_p(r)$ respectively) under the abrupt junction approximation are shown in Figs. 2.2a and 2.2b. The dependance of the cavity resonant frequency on the

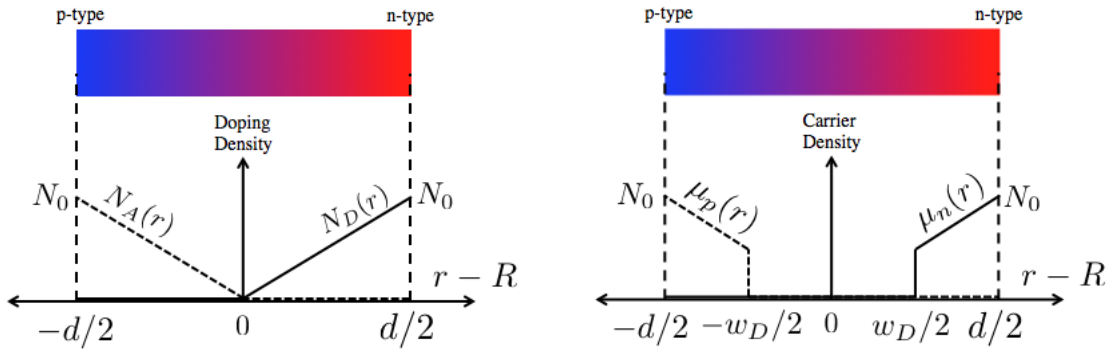


Figure 2.2: Schematic plots of (a) the doping densities ($N_D(r)$ and $N_A(r)$) and (b) the carrier densities ($\mu_p(r)$ and $\mu_n(r)$) across the graded p-n junction integrated into the ring of the ring-resonator.

depletion width is intimately linked with the dependance of the refractive index of Si on the number of free carriers per unit volume. While this dependance is generally very complex and can only be empirically determined [24], the Drude model allows us to approximate it by a linear relationship between the Si permittivity (n_{si}^2) and the carrier densities $\mu_n(r)$ and $\mu_p(r)$ [25, 26]:

$$n^2(r) = n_{\text{Si}}^2 - \alpha_p \mu_p(r) - \alpha_n \mu_n(r) \tag{2.4}$$

with α_p and α_n being (frequency dependant) material parameters. To obtain a simple analytical estimate of the dependance of ω_{res} on w_D, we assume a scalar gaussian approximation $\mathcal{E}(\mathbf{r})$ to modal field profile:

$$\mathcal{E}(\mathbf{r}) = \mathcal{E}_0 \exp\left[-\frac{1}{2}\left(\frac{(r-R)^2}{\sigma_r^2} + \frac{z^2}{\sigma_z^2} \right) \right] \exp(im\phi) \tag{2.5}$$

where σ_r and σ_z are measures of field confinement in the radial and z directions respectively and the index m determines the mode order and would usually depend on the ratio of the ring circumference and the wavelength of light in Si. Since $\mathcal{E}(\mathbf{r})$ is an approximation to the actual field profile, it only approximately satisfies the Helmholtz equation:

$$\nabla^2 \mathcal{E}(\mathbf{r}) + \frac{\omega^2}{c^2} n^2(\mathbf{r}) \mathcal{E}(\mathbf{r}) \approx 0 \tag{2.6}$$

which is also equivalent to:

$$\omega^2 \approx c^2 \frac{\int |\nabla \mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r}}{\int n^2(\mathbf{r}) |\mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r}} \tag{2.7}$$

Denoting the refractive index profile in the absence of a depletion region by $n_0(r)$, and defining $\Delta n^2(r) = n^2(r) - n_0^2(r)$, Eq. 2.7 can be rewritten as:

$$\omega^2 = c^2 \frac{\int |\nabla \mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r}}{\int n_0^2(\mathbf{r}) |\mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r} + \int \Delta n^2(\mathbf{r}) |\mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r}} = \frac{\omega_0^2}{1 + \gamma_{w_D}} \tag{2.8}$$

where

$$\gamma_{w_D} = \frac{\int \Delta n^2(\mathbf{r})|\mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r}}{\int n_0^2(\mathbf{r})|\mathcal{E}(\mathbf{r})|^2 \mathrm{d}^3\mathbf{r}} \tag{2.9}$$

and ω_0 is the resonant frequency in the absence of a depletion region. Using Eq. 2.4, $\Delta n^2(r)$ can be expressed as:

$$\Delta n^2(r) = \begin{cases} 2\alpha_n N_0(r-R)/d & R < r < R + w_D/2 \\ 2\alpha_p N_0(R-r)/d & R - w_D/2 < r < R \\ 0 & |r-R| > w_D/2 \end{cases} \tag{2.10}$$

A closed form expression for γ_{w_D} can be obtained by substituting Eqs. 2.10 and 2.5 into Eq. 2.9:

$$\gamma_{w_D} = \frac{N_0\sigma_r^2(\alpha_n+\alpha_p)(1-\exp(-w_D^2/4\sigma_r^2))}{\sqrt{\pi}n_{\mathrm{Si}}^2\sigma_r d - N_0\sigma_r^2(\alpha_n+\alpha_p)(1-\exp(-d^2/4\sigma_r^2))} \approx \frac{N_0(\alpha_n+\alpha_p)}{4\sqrt{\pi}n_{\mathrm{Si}}^2\sigma_r d} w_D^2 \tag{2.11}$$

Using this expression for γ_{w_D}, the detuning $\Delta = \omega - \omega_L$ can be related to V_D through:

$$\Delta \approx \omega_0 - \omega_L - \frac{1}{2}\omega_0\gamma_{w_D} = \omega_0 - \omega_L - \Gamma_0(V_D + V_{\mathrm{bi}})^{2/3} \tag{2.12}$$

where the coefficient Γ_0 quantifies the extent of modulation and is given by:

$$\Gamma_0 = \frac{N_0(\alpha_n+\alpha_p)}{4\sqrt{\pi}n_{\mathrm{Si}}^2\sigma_r d}\left(\frac{6\epsilon_0 n_{\mathrm{Si}}^2 d}{N_0 e}\right)^{2/3} \tag{2.13}$$

Assuming the p-n junction to be reverse biased through a voltage source with EMF V_S, we can define $\Delta_0 (= \omega_0 - \omega_L + \Gamma_0(V_S + V_{\mathrm{bi}})^{1/3})$ as the detuning in the absence of any photo-current, with which Δ can be written as:

$$\Delta = \Delta_0 - \Gamma_0((V_D + V_{\mathrm{bi}})^{1/3} - (V_S + V_{\mathrm{bi}})^{1/3}) \tag{2.14}$$

Eq. 2.14 is the final equation relating the cavity detuning Δ to the reverse biased voltage V_D and is used in all the simulations in this thesis.

2.1.2 Modelling the photo-generation process

The amount of power absorbed by the p-n junction integrated in the cavity and converted in to a photo-current is given by $\kappa_l|a(t)|^2$. The photo-current produced not only depends on this absorbed power, but also on the reverse-biased voltage across the p-n junction since the reverse biased voltage determines the strength of the electric field in the cavity, which in turn determines the drift velocity of the carriers. Moreover, the electrical circuit providing feedback into the cavity has two intrinsic delays associated with it – the time taken by the generated carriers to move across the depletion region and the RC time constant arising due to the depletion capacitance of the photo-detector. To accurately predict the speed of the optical device, these delays need to be included in the simulation model.

The circuit model for the electrical circuit is shown in Fig. 2.3. The photo-detector is modelled by a parallel combination of a current source corresponding to the photo-current (I_{ph}), a current source corresponding to the dark current I_{dark} and the depletion capacitance C_{dep}. Both I_{dark} and C_{dep} are functions of the reverse bias voltage V_D, and

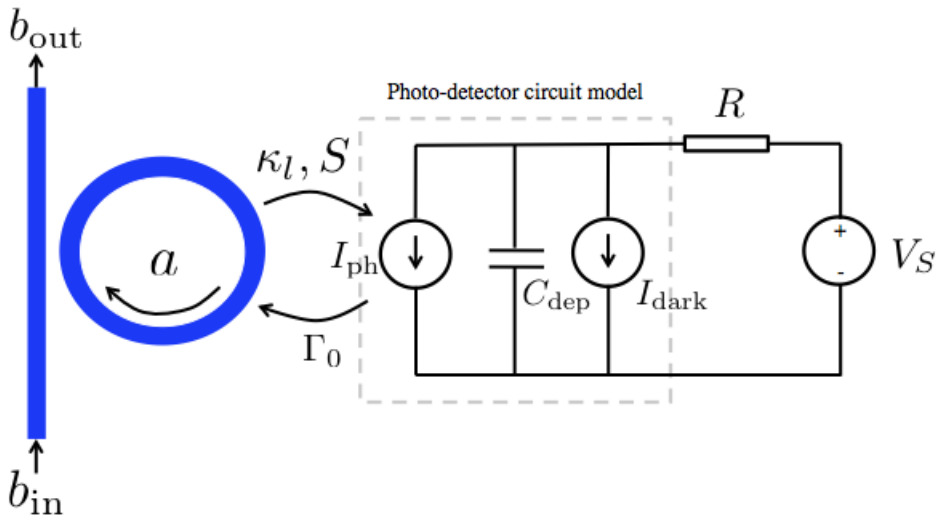


Figure 2.3: Equivalent circuit model for the electrical simulation of the photo-detector. The optical to electrical coupling occurs is captured through κ_l, which governs the power absorbed by the photo-detector ($\kappa_l|a|^2$).

are given by [23]:

$$C_{\text{dep}}(V_D) = \frac{C_0}{(1 + V_D/V_{\text{bi}})^{1/3}} \tag{2.15}$$

$$I_{\text{dark}}(V_D) = I_S(1 - \exp(-V_D/V_T)) \tag{2.16}$$

where C_0 is the open-circuit depletion capacitance, I_S is the reverse saturation current and $V_T = kT/e$ is the thermal voltage.

Modelling the generated photo-current (I_{ph}) is not so straightforward and the model has to incorporate the nonlinear dependance of the drift velocity of the carriers on the junction electric field and the delays associated with the finite drift velocity of the carriers. Since the photo-current is a combination of both electron and hole currents, I_{ph} can be expressed as a sum of electron ($I_{\text{ph},n}$) and hole ($I_{\text{ph},p}$) currents. It is clear that both $I_{\text{ph},n}$ and $I_{\text{ph},n}$ would be proportional to $\kappa_l|a|^2$, which is the optical power absorbed by the cavity. The two currents would also be proportional to the drift velocity of the corresponding carriers. The drift velocities (denoted by v_n for electrons and v_p for holes), however, depend nonlinearly on the electric field in the junction ($\mathcal{E}_{\text{jun}}$) [27]:

$$v_n = \frac{\mu_n \mathcal{E}_{\text{jun}}}{1 + \mu_n|\mathcal{E}_{\text{jun}}|/v_{\text{sat},n}} \tag{2.17a}$$

$$v_p = \frac{\mu_p \mathcal{E}_{\text{jun}}}{\sqrt{1 + \mu_p^2 \mathcal{E}_{\text{jun}}^2/v_{\text{sat},p}^2}} \tag{2.17b}$$

where $v_{\text{sat},n}$ and $v_{\text{sat},p}$ are the electron and hole saturation velocities ($v_n, v_p \to v_{\text{sat},n}, v_{\text{sat},p}$ as $\mathcal{E}_{\text{jun}} \to \infty$). The junction electric field would in general depend nonlinearly on the reverse biased voltage. However, as an approximate estimate, $\mathcal{E}_{\text{jun}}$ can be expressed as:

$$\mathcal{E}_{\text{jun}} \approx \frac{V_D}{d} \tag{2.18}$$

Putting together Eqs. 2.17 and 2.18, and using the fact that the electron and hole photo-currents are proportional to the absorbed power ($\kappa_l|a|^2$), we obtain:

$$I_{\text{ph},n} = \frac{S\kappa_l|a|^2}{2}\frac{V_D/V_{\text{sat},n}}{1+|V_D/V_{\text{sat},n}|} \tag{2.19a}$$

$$I_{\text{ph},p} = \frac{S\kappa_l|a|^2}{2}\frac{V_D/V_{\text{sat},p}}{\sqrt{1+V_D^2/V_{\text{sat},p}^2}} \tag{2.19b}$$

where $V_{\text{sat},n} = v_{\text{sat},n}d/\mu_n$ and $V_{\text{sat},p} = v_{\text{sat},p}d/\mu_p$ are the electron and hole saturation voltages, and can be thought of measures of the reverse bias voltage at which the respective carrier's velocity would saturate. S is the photo-detector sensitivity – it measures the extent of conversion of the absorbed power into a photo-current. In general, even S depends on both the absorbed power and the reverse bias voltage, we ignore these non-ideal effects and assume S to be a constant.

While the above analysis incorporates the effect of the reverse bias voltage on the generate photo-current, it still does not model the non-zero time taken by the carriers to move across the depletion region. To include these effects in our simulation model, we make use of the *Ramo's Theorem* [28, 29], according to which the photocurrent due to the generation of a charge Q uniformly across the junction is given by:

$$I_{\text{ph},n} = \frac{2Q}{\tau_n}\left(1-\frac{t}{\tau_n}\right) \tag{2.20a}$$

$$I_{\text{ph},p} = \frac{2Q}{\tau_p}\left(1-\frac{t}{\tau_p}\right) \tag{2.20b}$$

Where τ_n and τ_p are measures of the time taken by the electrons and holes to drift across the junction. Note that in the limit of $\tau_n, \tau_p \to 0$, the above expression reduce to a delta function, which conforms with our intuition of an infinitely fast transfer of charge across the junction. Eqs. 2.20 can be thought of as the impulse response of the photo-current with respect to the generation rate, and the finite speed of charge transfer across the junction can be modeled by convolving Eq. 2.19 with this impulse

response:

$$I_{\mathrm{ph},n} = \frac{S\kappa_l}{\tau_n}\int_0^{\tau_n}\left(1-\frac{\tau}{\tau_n}\right)\frac{V_D(t-\tau)/V_{\mathrm{sat},n}}{1+|V_D(t-\tau)/V_{\mathrm{sat},n}|}|a(t-\tau)|^2\mathrm{d}\tau \tag{2.21a}$$

$$I_{\mathrm{ph},p} = \frac{S\kappa_l}{\tau_p}\int_0^{\tau_p}\left(1-\frac{\tau}{\tau_p}\right)\frac{V_D(t-\tau)/V_{\mathrm{sat},p}}{\sqrt{1+V_D^2(t-\tau)/V_{\mathrm{sat},p}^2}}|a(t-\tau)|^2\mathrm{d}\tau \tag{2.21b}$$

Finally, the net photo-current is simply a sum of the electron and hole currents: $I_{\mathrm{ph}} = I_{\mathrm{ph},n} + I_{\mathrm{ph},p}$. We note here that since the electron and hole drift velocities are functions of the junction electric field and hence the reverse biased voltage, τ_n and τ_p are not constants. However, we approximate them as constants since a reverse biases p-n junction usually has a junction electric field high enough to push the electrons and holes to saturation, in which case the drift velocities become constant.

The electrical simulation equation can now be easily obtained by applying the Kirchoff's current law on the equivalent circuit in Fig. 2.3:

$$\frac{\mathrm{d}V_D(t)}{\mathrm{d}t} = -\frac{V_D(t)}{RC_{\mathrm{dep}}(V_D)} + \frac{V_S}{RC_{\mathrm{dep}}(V_D)} - \frac{I_{\mathrm{ph}}(t)+I_{\mathrm{diode}}(V_D)}{C_{\mathrm{dep}}(V_D)} \tag{2.22}$$

This put together with the optical equation (Eq. 2.1) complete the description of the time domain simulation model for the self-electrooptic cavity.

While developing the above model, we have assumed coupling to only a single waveguide, the simulation model can easily be generalised to a cavity coupled to N waveguides with coupling constants $\kappa_{w,n}$ $n \in \{1,2,3\ldots N\}$, inputs $b_{\mathrm{in},n}$ $n \in \{1,2,3\ldots N\}$ and outputs $b_{\mathrm{out},n}$ $n \in \{1,2,3\ldots N\}$ by modifying the optical equation (Eq. 2.1) to:

$$\frac{\mathrm{d}a(t)}{\mathrm{d}t} = -\left(i\Delta+\frac{\kappa}{2}\right)a(t) - \sum_{n=1}^{N}\sqrt{\kappa_{w,n}}b_{\mathrm{in},n}(t) \tag{2.23a}$$

$$\kappa = \sum_{n=1}^{N}\kappa_{w,n} + \kappa_l \tag{2.23b}$$

$$b_{\mathrm{out},n}(t) = b_{\mathrm{in},n}(t) + \sqrt{\kappa_{w,n}}a(t) \tag{2.23c}$$

We summarise the final simulation equations below for easy referencing:

$$\frac{\mathrm{d}a(t)}{\mathrm{d}t} = -\left(i\Delta + \frac{\kappa}{2}\right)a(t) - \sum_{n=1}^{N}\sqrt{\kappa_{w,n}}b_{\mathrm{in},n}(t) \tag{2.24a}$$

$$\frac{\mathrm{d}V_D(t)}{\mathrm{d}t} = -\frac{V_D(t)}{RC_{\mathrm{dep}}(V_D)} + \frac{V_S}{RC_{\mathrm{dep}}(V_D)} - \frac{I_{\mathrm{ph}}(t) + I_{\mathrm{diode}}(V_D)}{C_{\mathrm{dep}}(V_D)} \tag{2.24b}$$

$$\kappa = \sum_{n=1}^{N}\kappa_{w,n} + \kappa_l \tag{2.24c}$$

$$\Delta = \Delta_0 - \Gamma_0((V_D + V_{\mathrm{bi}})^{1/3} - (V_S + V_{\mathrm{bi}})^{1/3}) \tag{2.24d}$$

$$C_{\mathrm{dep}}(V_D) = \frac{C_0}{(1 + V_D/V_{\mathrm{bi}})^{1/3}} \tag{2.24e}$$

$$I_{\mathrm{dark}}(V_D) = I_S(1 - \exp(-V_D/V_T)) \tag{2.24f}$$

$$\begin{aligned} I_{\mathrm{ph}} = {} & \frac{S\kappa_l}{\tau_n}\int_0^{\tau_n}\left(1 - \frac{\tau}{\tau_n}\right)\frac{V_D(t-\tau)/V_{\mathrm{sat},n}}{1 + |V_D(t-\tau)/V_{\mathrm{sat},n}|}|a(t-\tau)|^2\mathrm{d}\tau \\ & + \frac{S\kappa_l}{\tau_p}\int_0^{\tau_p}\left(1 - \frac{\tau}{\tau_p}\right)\frac{V_D(t-\tau)/V_{\mathrm{sat},p}}{\sqrt{1 + V_D^2(t-\tau)/V_{\mathrm{sat},p}^2}}|a(t-\tau)|^2\mathrm{d}\tau \end{aligned} \tag{2.24g}$$

$$b_{\mathrm{out},n}(t) = b_{\mathrm{in},n}(t) + \sqrt{\kappa_{w,n}}a(t) \tag{2.24h}$$

There are a number of electrical and optical device parameters that appear in the simulation equations. Some of these parameters are either fixed material properties or difficult to design during fabrication and can therefore not be used to tune the characteristics of the self-electrooptic cavity. Table. 2.1 divides the parameters into two categories – fixed parameters and variable parameters. In all the subsequent circuit designs, only the variable parameters will be used to tune the properties of different self-electrooptic cavities.

Type	Symbol
Fixed	$C_0, I_S, V_S, \tau_n, \tau_p, V_{\mathrm{sat},n}, V_{\mathrm{sat},p}, V_{\mathrm{bi}}, V_T$
Variable	$S, R, \kappa_l, \kappa_w, \Gamma_0, \Delta_0$

Table 2.1: Division of the electrooptic nonlinear device's parameters into fixed (i.e. parameters that cannot be changed while fabrication or operation of the device) and variable (i.e. parameters that can be changed during fabrication).

2.2 Steady state characteristics of the self-electrooptic cavity

Steady state characteristics of the self-electrooptic cavity can be simulated by solving Eq. 2.24 for a constant input b_{in}. These steady state characteristics are crucial to the designs that follow in the subsequent chapters. The modulation coefficient Γ_0 can be estimated using Eq. 2.13. The device parameters used in Eq. 2.13 for a typical silicon ring resonator are listed in Table. 2.2a and with these parameters, $\Gamma_0/2\pi \sim 100$ GHz/V$^{2/3}$. The simulation parameters used in computing the steady state characteristics of the cavity are listed in Table. 2.2b [26, 29, 30].

The input-output characteristics of the self-electrooptic cavity can be quantified by defining a transmission coefficient. For an input $b_{\text{in}} = B_{\text{in}} \exp(i\phi_{\text{in}})$, and output $b_{\text{out}} = B_{\text{out}} \exp(i\phi_{\text{out}})$, the transmission coefficient $\tau(B_{\text{in}}) \exp(i\Phi(B_{\text{in}}))$ can be defined by:

$$\tau(B_{\text{in}}) = \frac{B_{\text{out}}}{B_{\text{in}}} \tag{2.25a}$$

$$\Phi(B_{\text{in}}) = \phi_{\text{out}} - \phi_{\text{in}} \tag{2.25b}$$

Fig. 2.4 shows the steady state cavity energy ($|a|^2$) as a function of the input power ($|b_{\text{in}}|^2 = B_{\text{in}}^2$) for a cavity coupled to a single waveguide. Both a forward and reverse sweep of the input power is simulated for different detunings Δ_0. It can be easily seen

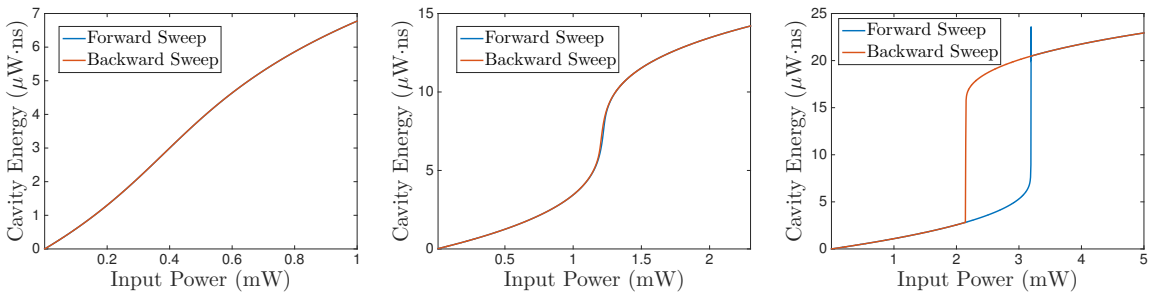
N_0	5×10^{24} m^{-3}
α_n (at 1550 nm)	9×10^{-27} m^{-3}
α_p (at 1550 nm)	3×10^{-27} m^{-3}
n_{Si}	3.4
d	1 μm
σ_r	0.3 μm

(a)

τ_n	60.0 ps	C_0	1.0 fF
τ_p	20.0 ps	I_S	10^{-19} A
$V_{\text{sat},n}$	2.0 V	V_S	7.0 V
$V_{\text{sat},p}$	0.8 V	S	1.0 A/W
V_{bi}	0.7 V	V_T	25 mV

(b)

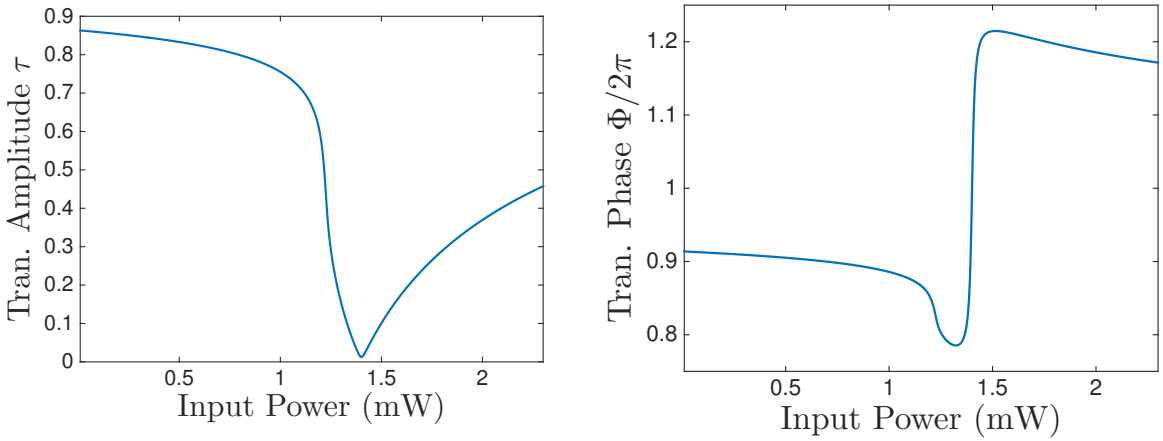
Table 2.2: Table listing (a) the parameters on which Γ_0 depends (Eq. 2.13) (b) the final simulation parameters used while solving Eq. 2.24.



(a) $\Delta_0 = -60$ GHz (b) $\Delta_0 = -80$ GHz (c) $\Delta_0 = -100$ GHz

Figure 2.4: Forward and backward sweep characteristics of the self-electrooptic cavity. In all the simulations, $\Gamma_0/2\pi = 100\text{GHz/V}^{2/3}$, $\kappa_l/2\pi = \kappa_w/2\pi = 20$ GHz and $R = 1$ kΩ.

that for small negative Δ_0, the cavity is monostable (i.e. it has a unique steady state output for a given input irrespective of the initial state of the cavity) whereas for large negative Δ_0, the cavity is bistable (i.e. the steady state depends on the initial cavity state). Moreover, for the assumed parameters, the cavity is at the edge of monostability and bistability at $\Delta_0/2\pi \sim -80$ GHz – this is the detuning at which the energy inside the cavity sharply transits from a low magnitude to a large magnitude. Such a sharp transition is important from the point of view of designing optical logic gates wherein this transition demarcates the high and low signal levels.



(a) (b)

Figure 2.5: Steady state transmission (a) amplitude (b) phase of the self-electrooptic cavity. In all the simulations, $\Gamma_0/2\pi = 100\text{GHz/V}^{2/3}$, $\kappa_l/2\pi = \kappa_w/2\pi = 20$ GHz and $R = 1$ kΩ.

Fig. 2.5 show the transmission characteristics (both amplitude τ and phase Φ) as a function of the input power at $\Delta_0/2\pi \sim -80$ GHz. Clearly, the derivative of τ with respect to B_{in} becomes large near the input power where the cavity energy transits from a low to high magnitude – the neighbourhood of this transition point can thus be used for amplifying a small signal input biased at the transition point.

The transition input power, henceforth called the *threshold power*, is an important measure of the power required to observe nonlinearity in the cavity response. At input powers much smaller than this threshold power, the cavity response would be similar to that of a linear cavity with the same parameters. While designing large optical networks, it is often essential to tune this threshold power. Although it is possible to optimize the cavity threshold through a number of parameters (such as $\Gamma_0, S, V_S, R, \kappa_l$ etc.), not all parameters are experimentally accessible. We investigate the impact of two parameters, namely the cavity quality factor Q and the external resistance R, that can easily be controlled in experiment. The cavity quality factor is a direct measure of the loss in the cavity and hence strongly effects the cavity threshold. The external

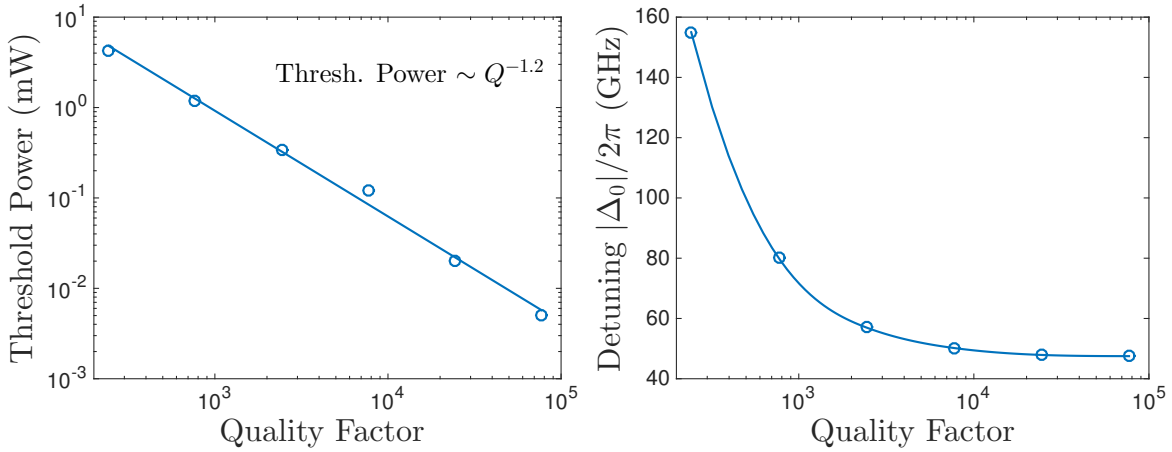


Figure 2.6: Variation of (a) input threshold power of the self-electrooptic cavity (b) detuning required to operate the cavity at the edge of monostable and bistable regimes ($\Delta_0 < 0$) as a function of the cavity quality factor. In all the simulations, $\Gamma_0/2\pi = 100\text{GHz/V}^{2/3}$, $\kappa_l = \kappa_w = \omega_0/4\pi Q$ and $R = 1$ kΩ.

resistance R is a measure of the effect of the photo-current (and hence the input power) on the reverse biased voltage – it can thus be also used to tune the cavity threshold.

Fig. 2.6a shows how the threshold power scales with the quality factor $Q \sim \omega_0/2\pi(\kappa_w + \kappa_l)$. We assume the critical coupling condition in our simulations ($\kappa_l = \kappa_w$) – a condition that is typically satisfied by fabricated resonators. The detuning parameter Δ_0 is chosen so as to operate the cavity at the edge of the monostable and bistable regimes – Fig. 2.6b shows the Δ_0 required to achieve this condition as a function of the quality factor. Clearly, an increase in Q reduces the threshold power – our simulations indicate the the threshold power approximately scales inversely with $Q^{1.2}$. Recent advances in fabrication technologies have enabled the on-chip integration of silicon cavities with quality factors as high as $\sim 10^5$ [31, 32]. Fig. 2.6a clearly shows that it is possible to achieve threshold power of the order of 1 μW at these high quality factors. However, it should also be noted that the speed of the optical cavity varies inversely with $\kappa = \kappa_l + \kappa_w$ (Eq. 2.1) and therefore the cavity becomes slower on increasing Q. Speeds $\sim 10 - 100$ GHz can be achieved with $Q \sim 10^3 - 10^4$ and the self-electrooptic cavity have threshold powers $\sim 100\ \mu\text{W} - 1$ mW at these quality factors.

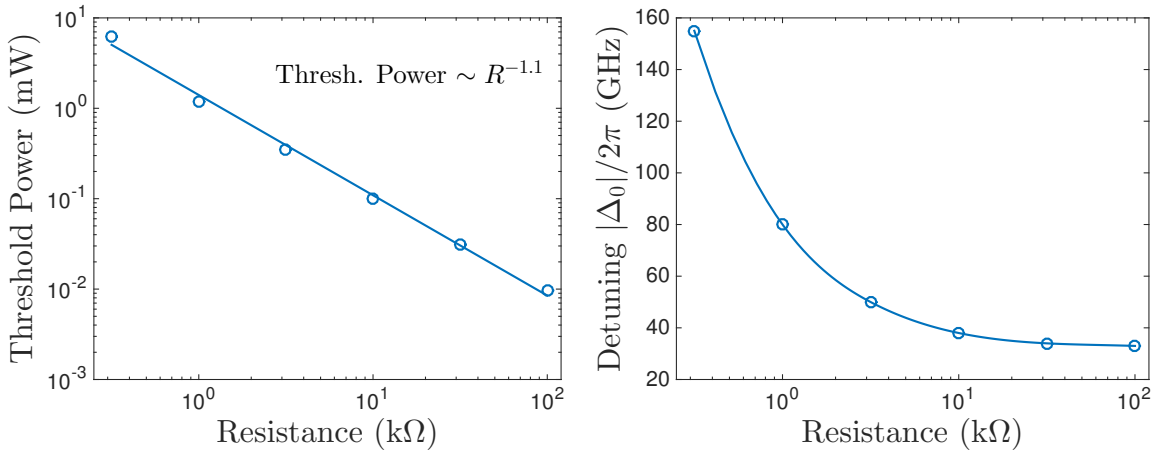
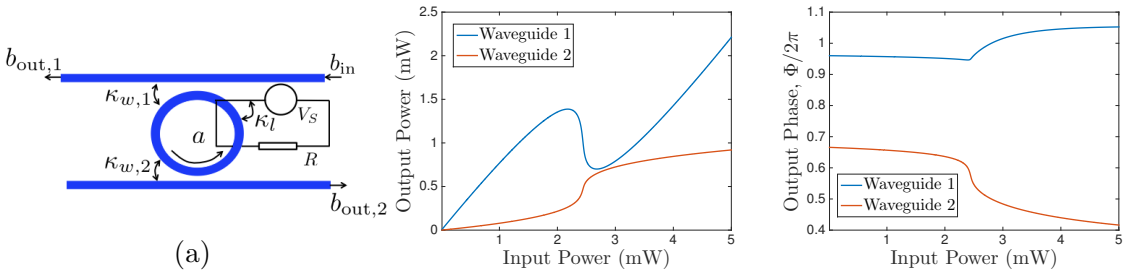


Figure 2.7: Variation of (a) input threshold power of the self-electrooptic cavity (b) detuning required to operate the cavity at the edge of monostable and bistable regimes ($\Delta_0 < 0$) as a function of the external resistance R. In all the simulations, $\Gamma_0/2\pi = 100\text{GHz/V}^{2/3}$ and $\kappa_l/2\pi = \kappa_w/2\pi = 20$ GHz.

Increasing the resistance R also reduces the threshold power (Fig. 2.7a) with the threshold power scaling inversely with $R^{1.1}$. This is expected since for large R, even a small photo-current (corresponding to a small input power) can produce a large change in the reverse-biased voltage and hence the cavity resonance. However, it should also be noted that increasing resistance also increases the power dissipation in the electrical circuit, even though it makes the resonator more optically power efficient. Again, in our simulations we have chosen Δ_0 so as to operate the cavity at the edge on the monostable and bistable regimes, and the required Δ_0 as a function of R is shown in Fig. 2.7b

Finally, we consider the transmission characteristics of a self-electrooptic cavity coupled to two waveguides (Fig. 2.8a). The cavity has one input $b_{\text{in}}(t)$ which is fed through one of the coupled waveguides and two outputs, $b_{\text{out},1}(t)$ and $b_{\text{out},2}$, corresponding to the two waveguides. Since the input is being fed into only one waveguide, $b_{\text{out},1} = b_{\text{in}} + \sqrt{\kappa_{w,1}}a$ and $b_{\text{out},2} = \sqrt{\kappa_{w,2}}a$ corresponding to the two waveguide outputs. The variation of the output power and phase corresponding to the two waveguides with the input power is shown in Figs. 2.8b and 2.8c. It can clearly be seen that this structure behaves like a thresholder between the input and the output of the waveguide complementary to the input waveguide.



(a) (b) (c)

Figure 2.8: (a) Schematic of cavity circuit with two waveguides (b) Steady state output power as a function of input power (c) Steady state output phase as a function of input power. In all simulations, $\Gamma_0/2\pi = 100$ GHz/V$^{2/3}$, $\kappa_{w,1}/2\pi = \kappa_{w,2}/2\pi = 20$ GHz and $\Delta_0/2\pi = -80$ GHz.

2.3 Linear circuit elements: directional couplers and phase shifters

A number of versatile linear operations can be performed on a silicon photonics platform by using isolated and coupled structures based on channel waveguides [33]. Circuit designs presented in the subsequent chapters require the use of phase-shifter and directional couplers. Realisation of these circuit elements in silicon photonics is described below.

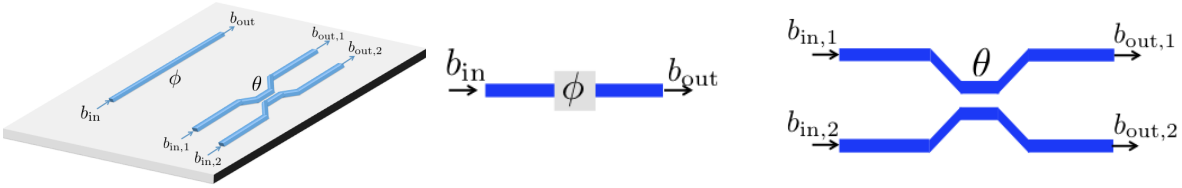
A phase-shifter can be realised by an appropriately chosen length L of a channel waveguide (Fig. 2.9a). For an input signal $b_{\text{in}}(t)$, the output of the phase shifter $b_{\text{out}}(t)$ is given by:

$$\tilde{b}_{\text{out}}(\omega) = \exp(i\phi(\omega))b_{\text{in}}(\omega) \tag{2.26}$$

where $\tilde{b}_{\text{in}}(\omega)$ and $\text{b}_{\text{out}}(\omega)$ are the fourier transforms of $b_{\text{in}}(t)$ and $b_{\text{out}}(t)$ respectively. The phase $\phi(\omega)$ can be designed by appropriately changing the length of the channel waveguide:

$$\phi(\omega) = \frac{\omega + \omega_L}{c} L n_{\text{eff}}(\omega + \omega_L) \tag{2.27}$$

$n_{\text{eff}}(\omega)$ being the effective refractive index of the waveguide. Bandwidths of the input signals would typically be ~ 10 GHz while the laser frequency is many order of magnitudes higher $\omega_L/2\pi \sim 10^6$ GHz. It is therefore reasonable to neglect the frequency dependance of $\phi(\omega)$ (since $\omega + \omega_L \sim \omega_L$ in Eq. 2.27) and the input-output relationship



(a) (b) (c)

Figure 2.9: (a) Schematic of an on-chip phase shifter and directional coupler realized using channel waveguides (b) Circuit symbol for the phase shifter (c) Circuit symbol for the directional coupler.

can be simply expressed as:

$$b_{\text{out}}(t) = \exp(i\phi) b_{\text{in}}(t) \tag{2.28}$$

It should however be noted that the frequency dependance of $\phi(\omega)$ has to be accounted for an accurate calculation of the effect of waveguide dispersion on the characteristics of the phase-shifter. The circuit symbol used for a phase-shifter is shown in Fig. 2.9b.

Directional couplers (Fig. 2.9a) can be constructed by fabricating two channel waveguides in close proximity to each other. The two waveguides then couple through their evanescent fields resulting in the output of one waveguide also being dependant on the input to the other waveguide. Mathematically, the directional coupler can be described by:

$$\begin{bmatrix} \tilde{b}_{\text{out},1}(\omega) \\ \tilde{b}_{\text{out},2}(\omega) \end{bmatrix} = \begin{bmatrix} \cos\theta(\omega) & -i\sin\theta(\omega) \\ -i\sin\theta(\omega) & \cos\theta(\omega) \end{bmatrix} \begin{bmatrix} \tilde{b}_{\text{in},1}(\omega) \\ \tilde{b}_{\text{in},2}(\omega) \end{bmatrix} \tag{2.29}$$

where the $b_{\text{in},i}(t),\ i \in \{1,2\}$ are the directional coupler inputs and $b_{\text{out},i}(t),\ i \in \{1,2\}$ are the directional coupler outputs (Eq. 2.29) is expressed in terms of the fourier transforms of the input and output signals.The parameter $\theta(\omega)$ depends on the length L of the coupled region, and overlap integrals between the modal field profiles of the two channel waveguide:

$$\theta(\omega) = \frac{(\omega+\omega_L)L}{cn_{\text{eff}}(\omega+\omega_L)} \frac{\int \Delta n^2(\boldsymbol{\rho})\mathcal{E}(\boldsymbol{\rho})\mathcal{E}(\boldsymbol{\rho}-\boldsymbol{\rho}_c)\ \mathrm{d}^2\boldsymbol{\rho}}{\int \mathcal{E}^2(\boldsymbol{\rho})\ \mathrm{d}^2\boldsymbol{\rho}} \tag{2.30}$$

where $\boldsymbol{\rho} = (x, y)$ are the transverse coordinates (i.e. coordinates in plane perpendicular to the propagation axis) of the point in question, $\boldsymbol{\rho}_c$ is the centre-centre displacement between the two channel waveguides and $\Delta n^2(\boldsymbol{\rho})$ is the difference between the refractive index profile of the coupled structure and the refractive index profile of an isolated waveguide. $\theta(\omega)$ can be designed by appropriately choosing the length of the coupled region. Again, the frequency dependance of $\theta(\omega)$ may be neglected under the assumption of a small signal bandwidth, and the input-output relation for the directional coupler

can be expressed as:

$$\begin{bmatrix} b_{\text{out},1}(t) \\ b_{\text{out},2}(t) \end{bmatrix} = \begin{bmatrix} \cos\theta & -i\sin\theta \\ -i\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} b_{\text{in},1}(t) \\ b_{\text{in},2}(t) \end{bmatrix} \tag{2.31}$$

The circuit symbol used for a directional coupler is shown in Fig. 2.9c.

It is worthwhile to note that the linear operations achieved by the phase-shifter and the directional coupler are unitary operations – a physical implication of this mathematical property is the conservation of energy between the input and output signals. Moreover, any unitary (and by extension energy conserving) operation can be realized with an appropriate combination of directional couplers and phase shifters.

Chapter 3

Optical logic gates

The thresholding characteristics of the self electro-optic nonlinear cavity can be used for realising digital logic gates and memory elements. In this chapter, we present circuits that implement a NOT gate, an AND gate (and by extension any combinational logic architecture) and an optical latch (thereby opening up the possibility of implementing sequential logic). In the implementation presented in chapter, the binary data is encoded into the amplitude of the optical signals in the circuit.

The implementation of the optical NOT gate is presented in section 3.1, optical AND gate is presented in section 3.2 and optical latch in section 3.3

3.1 Optical NOT gate

The optical circuit implementing the NOT gate is shown in Fig. 3.1 [12, 13]. A combination of phase shifters and a directional coupler is used interfere the NOT gate input b_{in} with a constant signal b_0 which is also equal to the input high level to produce the cavity input b_{cav}:

$$b_{\text{cav}} = \frac{b_0 - b_{\text{in}}}{\sqrt{2}} \tag{3.1}$$

Clearly, when the gate input is high, the cavity input is 0 and when the gate input is low, the cavity input is $b_0/\sqrt{2}$. If the cavity threshold is designed so as to lie between

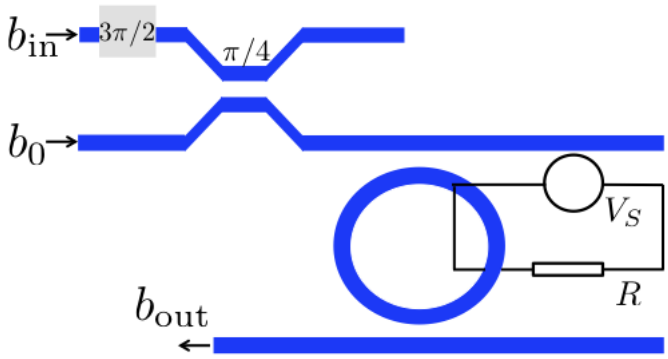
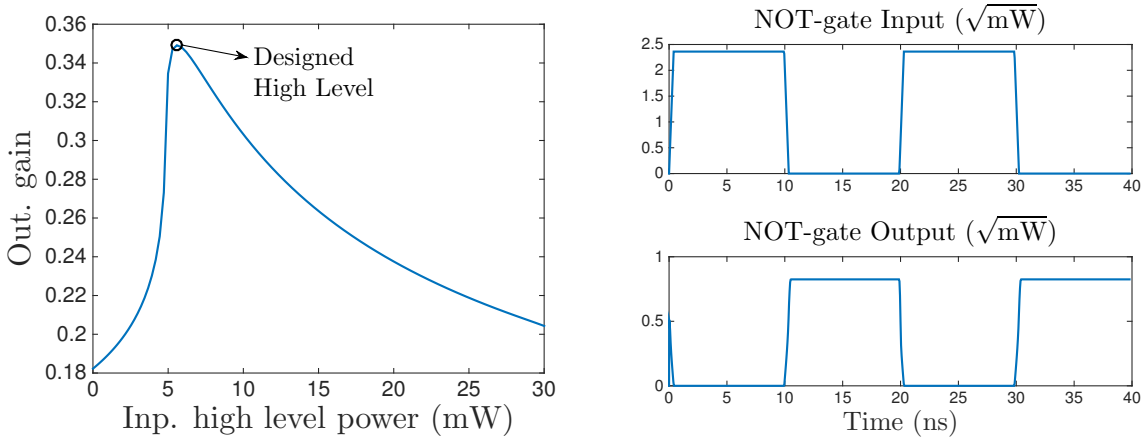


Figure 3.1: Schematic of the circuit implementing an optical NOT gate. The constant signal b_0 is also equal to the high level amplitude of the input signal.

0 and $b_0/\sqrt{2}$, then this circuit performs a NOT operation on the cavity input. An important consideration in gate design is for the circuit to be cascadable i.e. the gate output should be a valid input for other logic gates. Assuming all the logic gates to be designed for the same input high level, this is equivalent to requiring equal output and input high levels. However, since the cavity is lossy and since the gate output corresponds to the output of only one of the waveguides, the output high level of the NOT gate is typically smaller than the input high level. This necessitates the insertion of level restorers at different points within a large network. So as to minimize the number of such level restorers, we design the input high level b_0 so as to maximize



(a) (b)

Figure 3.2: (a) Variation of the output gain of the NOT gate with the input high level power (b_0^2) – the circled point indicates the designed input high level (b) Transient simulation of the designed NOT gate.

the output gain (i.e. ratio of the output high level to the input high level) as shown in Fig 3.2a. Fig. 3.2b shows a transient simulation illustrating the inversion operation performed by this circuit.

3.2 Optical AND gate

The optical AND gate can be realized using the circuit shown in Fig. 3.3 [11–13]. The phase shifters and a directional coupler interfere the AND gate inputs $b_{\text{in},1}$ and $b_{\text{in},2}$ so as to produce the following cavity input b_{cav}:

$$b_{\text{cav}} = \frac{b_{\text{in},1} + b_{\text{in},2}}{\sqrt{2}} \tag{3.2}$$

Feeding both low inputs into the AND gate results in $b_{\text{cav}} = 0$, and thus both the cavity outputs are equal to 0. This circuit therefore trivially satisfies one entry of the AND gate truth table (low·low = low). For the input combination with one high and one low, $b_{\text{cav}} = b_0/\sqrt{2}$ and for the input combination with both high, $b_{\text{cav}} = \sqrt{2}b_0$ where b_0 is the input high level. To satisfy the three remaining entries of the AND gate truth table, we design the input high level so as to ensure that the cavity threshold lies between $b_0/\sqrt{2}$ and $\sqrt{2}b_0$.

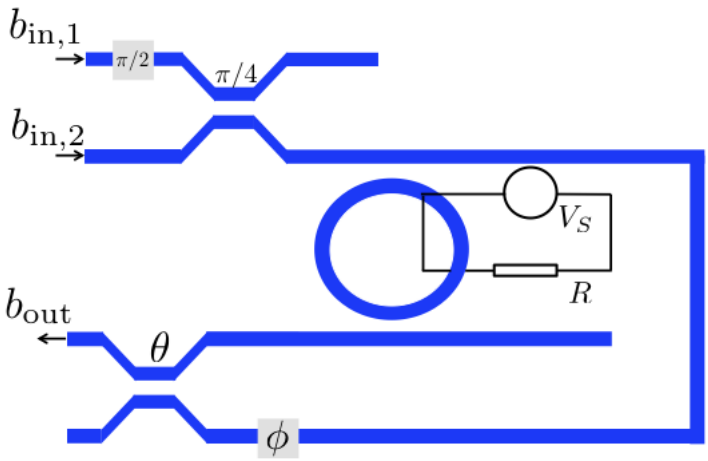


Figure 3.3: Schematic of the circuit implementing an optical AND gate.

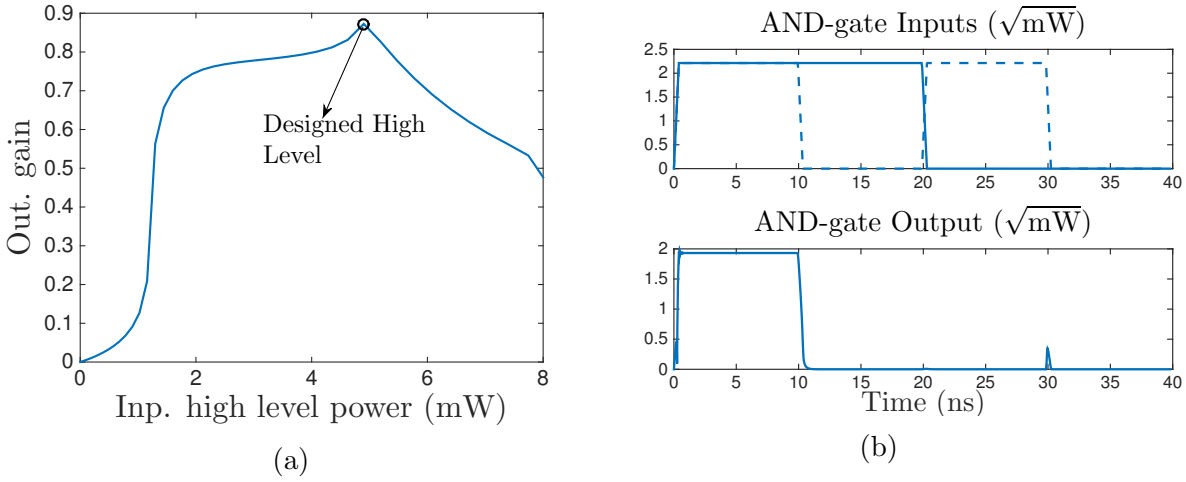


Figure 3.4: (a) Variation of the output gain of the AND gate with the input high level power (b_0^2) – the circled point indicates the designed input high level (b) Transient simulation of the designed AND gate.

However, since the cavity thresholding is not perfect, the cavity output will not be exactly 0 for the input combination with one high input and one low input. It is possible to cascade another thresholder with the AND gate so as to decrease this amplitude, a simpler approach is to interfere the outputs of the two waveguides destructively through a combination of a beam splitter and a phase shifter as shown in Fig. 3.3. The parameters θ and ϕ are chosen so as to ensure that the output amplitude for an input combination of one high and one low is 0. We point out that since the cavity is a nonlinear device, destructively interfering the waveguide outputs for this input combination does not imply a destructive interference for the input combination of high and high. Again, for the gate design to be cascadable, it is necessary for the output high level to be as close to the input high level as possible. We therefore design the input high level b_0 so as to maximize the ratio of output high level to input high level as shown in Fig. 3.4a. A transient simulation illustrating the operation of an AND gate is shown in Fig. 3.4b.

3.3 Optical latch

The optical circuit implementing the latch is shown in Fig. 3.5a [11–13]. The two cavities are provide feedback to each other and this results in the overall circuit behaving bistably. The input high level b_0 is designed to lie near the cavity threshold (both the cavities are operated monostably). The two inputs to the latch either set it to a high output ($b_{\text{set}} = b_0, b_{\text{reset}} = 0$), reset it to a low output ($b_{\text{set}} = 0, b_{\text{reset}} = b_0$) or hold the output at its previous state ($b_{\text{set}} = b_0, b_{\text{reset}} = b_0$).

To see that this circuit indeed behaves like a latch, consider its operation during a set or a reset phase. For instance, during a set phase, the upper cavity in Fig. 3.5a operates above threshold. The phase ϕ_b is designed so as to ensure that the constant input b_c interferes destructively with the output of the upper cavity so as to give a low input to the lower cavity and push it below the threshold. The circuit is in a similar state during a reset phase, with the upper cavity being below threshold and the lower cavity being above threshold. The feedback phase ϕ_f is designed so as to ensure that the cavity states remain the same for a hold following a set (or reset) phase. This is equivalent to ensuring that the feedback phase ϕ_f imparted to the feedback signal

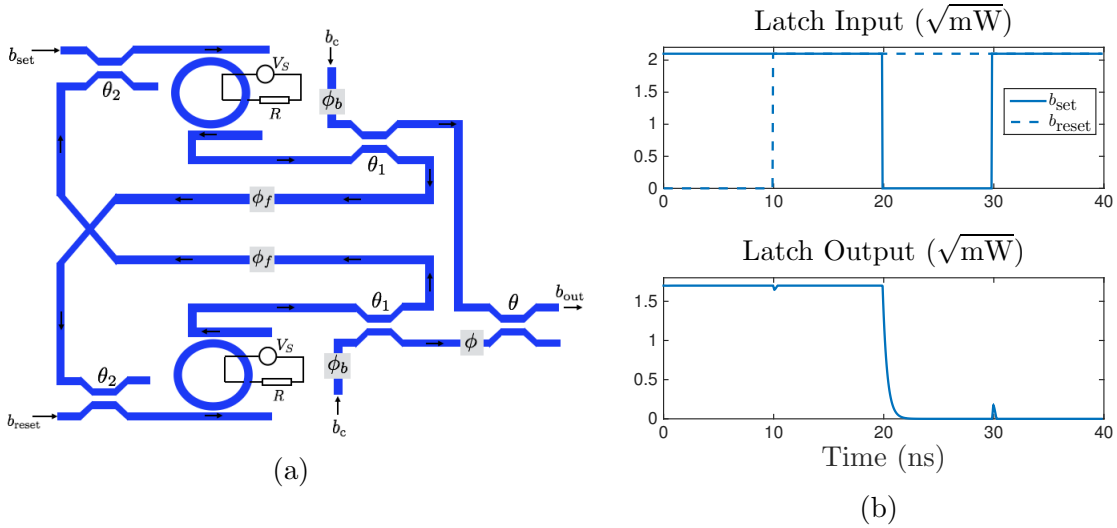


Figure 3.5: (a) Schematic of the circuit implementing an optical latch (b) Transient simulation of the designed latch.

ensures that it interferes destructively with the set (or reset) input to the cavity below threshold, thereby preserving the previous state of the circuit.

The phase shifter ϕ and the beam splitter θ is designed so as to ensure that the final output b_{out} has the same amplitude during a set and the following hold, and a reset and the following hold. A transient simulation of this circuit is shown in Fig. 3.5b .

Chapter 4

Optical neuron

The neuron is a bio-inspired nonlinear system that can potentially be used as a building block for more complex neural networks and learning machines. Mathematically, a neuron is equivalent to a mapping from N inputs to one output, with an input-output relation given by (Fig. 4.1a):

$$b^{\text{out}} = \mathcal{F}\left(\sum_{n=1}^{N} w_n b_n^{\text{in}}\right) \tag{4.1}$$

where $\mathcal{F}(z)$ is the neuron 'activation' function, and is typically chosen according to the specific application at hand. Common choices for the activation function include the sigmoid activation function $\mathcal{F}(z) = 1 + \tanh(\beta z)$ (β being a design parameter) and the step activation function $\mathcal{F}(z) = 1$ if $z > 0$ and 0 if $z \leq 0$ (Fig. 4.1b). The weights w_n are termed as the neuron 'synaptic' weights – an important property of these weights is their adaptability i.e. they need to be designed in a way that they can be tuned using an external control signal. This tunability is an essential requirement if the neuron is to perform a learning task, wherein it adapts its transfer function according to a given training data set.

This chapter is devoted to the design and simulation of the synaptic weights and the activation function for an all-optical neuron using the self-electrooptic nonlinear device as the fundamental building block. The designed neuron is then used to solve a

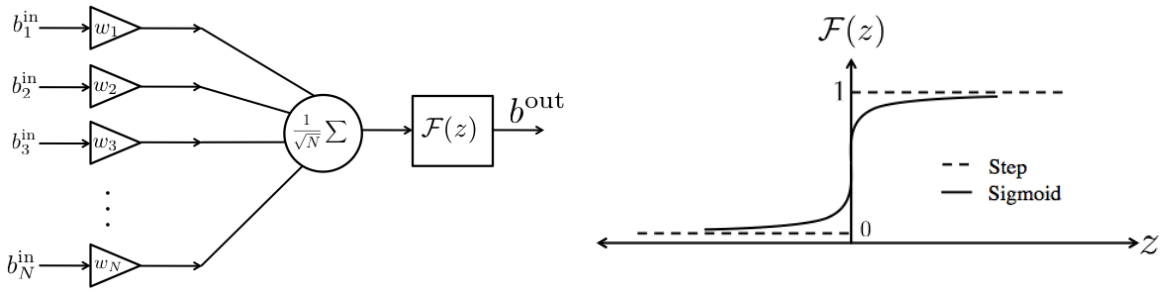


Figure 4.1: (a) Block diagram of a neuron with inputs b_n^{in} and output b^{out} (b) Schematic plots of the sigmoid and step activation functions

simple classification problem, and its performance is compared with a similarly trained Support Vector Machine (SVM) [34].

4.1 Design of synaptic weights

A tunable amplifier can easily be constructed by biasing the input signal to the transition region of the self-electrooptic cavity. The bias can then be used to control the gain of the amplifier. One issue with this design is the fact that the self-electrooptic cavity not only scales the amplitude of the input signal when used as an amplifier, but also introduces a non-zero bias dependant phase shift in the amplifier output. By definition, the neuron synapse should have a real output for a real input. So as to remove the undesired imaginary part introduced by the amplifier, we cascade the amplifier with a quadrature filter i.e. a circuit that takes in a complex signal and filters out the imaginary part of the complex signal.

The amplifier is implemented using a differential structure as shown in Fig. 4.2 [35]. The differential structure not only provides for an automatic bias cancellation, but also increases the linearity of the output. The amplitude of the bias B_{bias} can be used to control the gain of the amplifier. A general definition of the amplifier gain needs to account for the fact that the input and output signals are complex. We define an

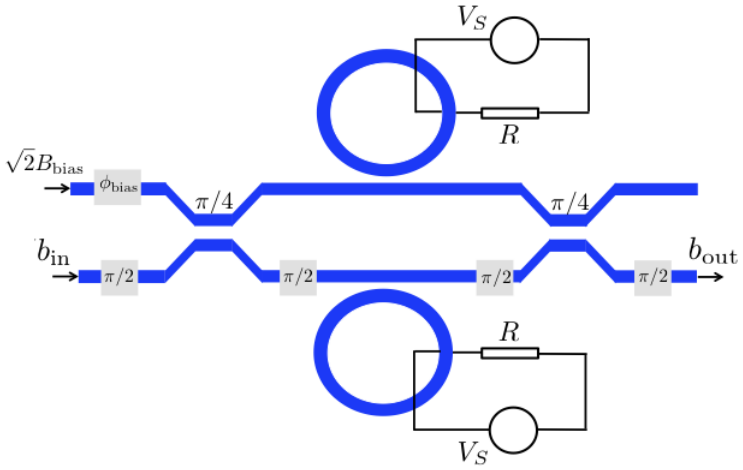


Figure 4.2: Schematic of the circuit implementing the bias-controlled differential amplifier. The gain of this structure can be tuned by changing the bias amplitude B_{bias}.

in-phase gain g_+ and a quadrature-phase gain g_- via:

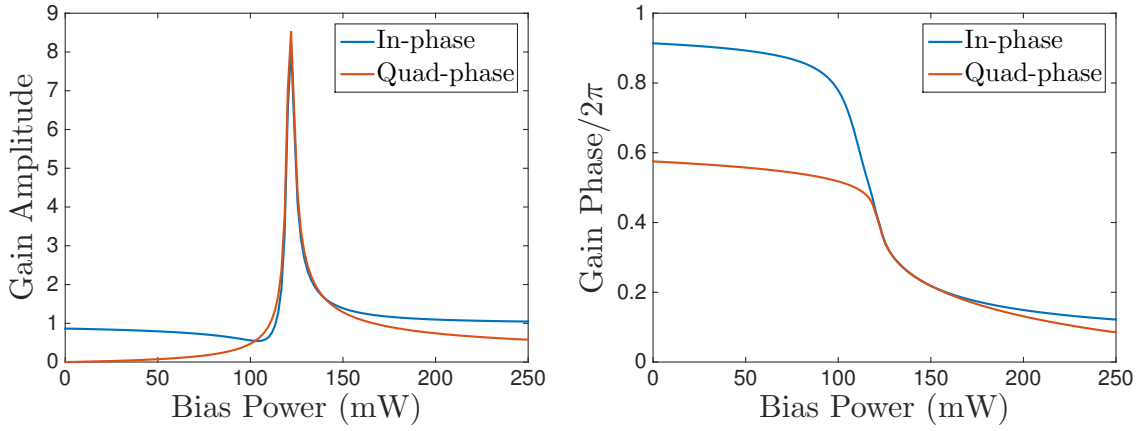
$$\begin{bmatrix} b_{\text{in}} \\ b_{\text{in}}^* \end{bmatrix} = \begin{bmatrix} g_+ & g_- \\ g_+^* & g_-^* \end{bmatrix} \begin{bmatrix} b_{\text{out}} \\ b_{\text{out}}^* \end{bmatrix} \tag{4.2}$$

The gains g_+ and g_- can be related to the nonlinear transmission coefficient $\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))$ of a self-electrooptic cavity coupled to a single waveguide (Eq. 2.25) via (refer to appendix A for derivation):

$$g_+ = \tau(B_{\text{bias}})\exp(i\Phi(B_{\text{bias}})) + \frac{B_{\text{bias}}}{2}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}\Big[\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))\Big]_{B_{\text{in}}=B_{\text{bias}}} \tag{4.3a}$$

$$g_- = \frac{B_{\text{bias}}\exp(2i\phi_{\text{bias}})}{2}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}\Big[\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))\Big]_{B_{\text{in}}=B_{\text{bias}}} \tag{4.3b}$$

Fig. 4.3, shows the variation of the magnitude and phase of the in-phase and quadrature-phase gains with the bias power $|B_{\text{bias}}|^2$ assuming $\phi_{\text{bias}} = 0$. Clearly, the magnitude of these gains becomes large near the threshold of the SEO cavity. Moreover, the phase of these gains also changes sharply near the cavity threshold. The range of the gains that can be achieved by varying the bias power can also be controlled using the detuning Δ_0. From Fig. 4.4, it can be seen that as the detuning is increased, the self electro-optic cavity tends to the bistable regime, and the range



(a) (b)

Figure 4.3: Variation of (a) amplitude (b) phase of the in-phase and quadrature-phase gains for the differential amplifier with the bias power (B^2_{bias}). In all the simulations, $S = 0.01$ A/W, $\Gamma_0/2\pi = 100$ GHz/V$^{2/3}$, $\kappa_l/2\pi = \kappa_w/2\pi = 20$ GHz, $\Delta_0/2\pi$ = -80 GHz and $\phi_{\text{bias}} = 0$

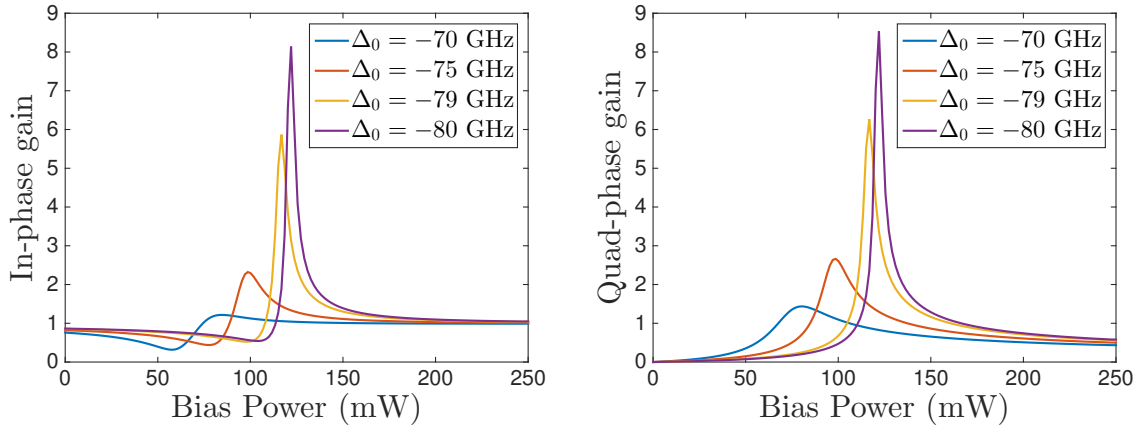
.

of achievable gains increases significantly. In conclusion, the gains themselves can be controlled using the bias power, and the range of achievable gains can be controlled via the detuning.

The amplifier structure described above can also be employed to implement a quadrature filter by appropriately choosing the bias amplitude in phase. According to the gain definitions, the complex output of the amplifier b_{out} is related to the input b_{in} through the following equation:

$$b_{\text{out}} = g_+ b_{\text{in}} + g_- b^*_{\text{in}} \tag{4.4}$$

where g_+ and g_- depend on B_{bias} and ϕ_{bias}. If the bias is chosen so as to ensure $g_+ = g^*_-$, then it is clear that the amplifier output would be purely real irrespective of the input phase. This particular choice of the bias thus achieves quadrature filtering of the input signal. The procedure for appropriately choosing the bias amplitude B_{bias} and phase is straightforward, since $|g_+|$ and $|g_-|$ are independent of the bias phase, the bias amplitude can be chosen so as to ensure $|g_+| = |g_-|$. The bias phase can then be chosen to ensure that $\angle g_+ + \angle g_- = 0$. The synapse can then be designed by cascading

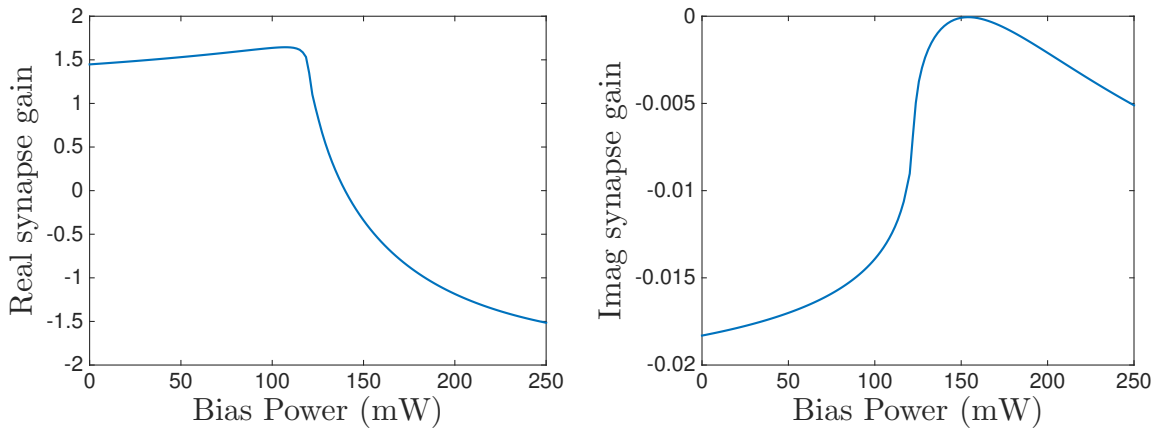


(a) (b)

Figure 4.4: Variation of (a) in-phase gain amplitude (b) quadrature-phase gain amplitude the differential amplifier with the bias power (B^2_{bias}) for different detuning Δ_0. In all the simulations, $S = 0.01$ A/W, $\Gamma_0/2\pi = 100$ GHz/V$^{2/3}$, $\kappa_l/2\pi = \kappa_w/2\pi = 20$ GHz.

the an amplifier with a quadrature filter. Note that in our design, the quadrature filter's bias is fixed and cannot be used to tune the synapse gain – the synapse gain is entirely controlled by the amplifier bias. Additionally, even though the amplifier can be controlled using both the bias amplitude and phase, changing the phase is usually not as straightforward as changing the amplitude. Since the weights w_n in the neuron transfer function that are being implemented by the controlled amplifier can be both negative and positive, the amplifier bias phase is chosen so as to ensure that the overall synapse gain has equal positive and negative swings over the tunable range of bias amplitudes. Fig. 4.5a shows the synapse gain as a function of the amplifier bias power for a synapse designed according to the guidelines outline above. Fig. 4.5b shows the maximum gain in the imaginary part of the output for a small signal input with power upto 0.1 mW. Clearly, the quadrature filter ensures that the synapse output is almost purely real with an imaginary part that is nearly 80 dB smaller than the real part.

It can also be noted that the synapse implementation described above is only valid for input powers much smaller than the bias powers – increasing the input power makes the nonlinear dependance of the output on the input more significant. This requirement thus limits the magnitude of the inputs that can be fed into the neuron.



(a) (b)

Figure 4.5: Variation of (a) real synapse gain (b) maximum imaginary synapse gain (for maximum input amplitude of 0.1 mW) with the bias power (B^2_{bias}) for different detuning Δ_0. In all the simulations, $S = 0.01$ A/W, $\Gamma_0/2\pi = 100$ GHz/V$^{2/3}$, $\kappa_l/2\pi = \kappa_w/2\pi = 20$ GHz and $\Delta_0 = -80$ GHz.

4.2 Design of activation function

The typical activation function used in most neural networks is the sigmoid activation function - $\mathcal{F}(z) = 1 + \tanh(\beta z)$. Certain classes of problems, called the classification problems, require the step activation function: $\mathcal{F}(z) = 1$ if $z > 0$ and 0 if $z \leq 0$. Hardware implementations of these activation functions can only approximate them. However, the one essential property that the hardware implementation should satisfy is that the activation function output saturates as the input becomes infinitely large. Failure to conform to this property might lead to undesirable instabilities in the neural network. The circuit implementation of the activation function is shown in Fig. 4.6a. An optically controlled phase shifter is wrapped in a Mach-zender interferometer with a constant input being fed into one of its arm – depending on the phase of this phase shifter, this constant input is switched in and out of the output port. The structure implementing the optically controlled phase-shifter is shown in Fig. 4.6b. Two self-electrooptic cavities are electrically coupled to each other i.e. their junctions are connected in parallel. The input to one of the cavities therefore affects the detuning of the second cavity. From Eq. 2.1, the steady state output corresponding to any one of

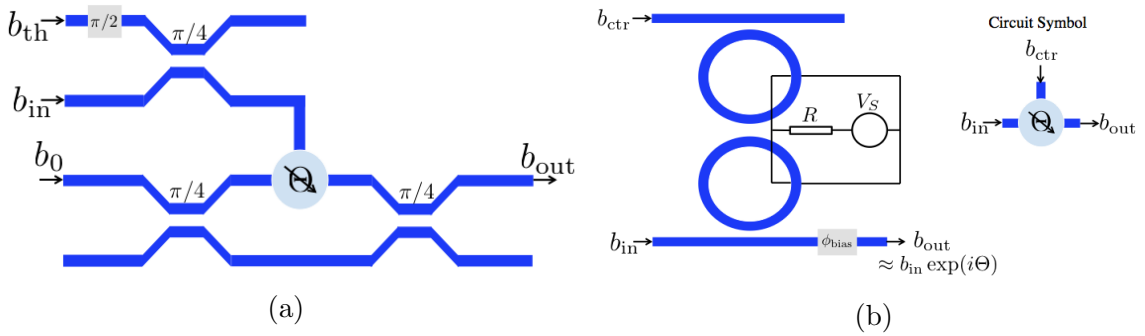


Figure 4.6: (a) Schematic of the circuit implementing the sigmoid or step activation function (b) Implementation of the controlled phase shifter by electrically coupling to self-electrooptic cavities.

the cavities is related to its input and the cavity detuning via:

$$b_{\text{out}} = \left(\frac{i\Delta + (\kappa - 2\kappa_w)/2}{i\Delta + (\kappa)/2}\right) b_{\text{in}} = \left(\frac{i\Delta + (\kappa_l - \kappa_w)/2}{i\Delta + (\kappa_l + \kappa_w)/2}\right) b_{\text{in}} \tag{4.5}$$

Clearly, if $\kappa_l << \kappa_w$, then $b_{\text{out}} \approx b_{\text{in}} \exp(i\Theta)$ with $\Theta = \pi - 2\tan^{-1}(2\Delta/\kappa_w)$. Since the two cavities are electrically coupled, Δ and hence the transmission phase Θ corresponding to one waveguide depends on the input b_{ctr} to the complementary (other) waveguide. Fig. 4.7a shows this transmission phase as a function of the input power corresponding to the complementary waveguide. Moreover, it can also be seen from

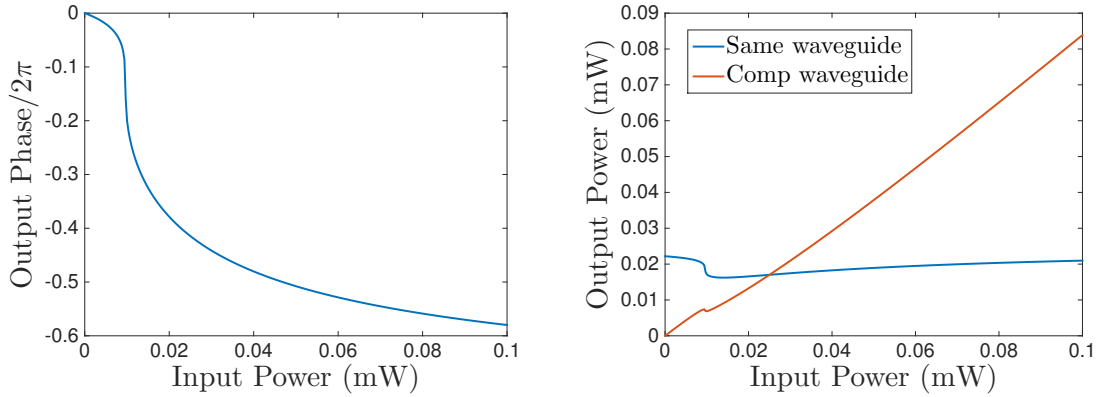


Figure 4.7: Variation of the (a) transmission phase of the output waveguide (b) output amplitudes of both the waveguides as a function of the input power to the waveguide complementary to the output waveguide. In all simulations, $S = 1$ A/W, $\Gamma_0/2\pi = 500$ GHz/V$^{2/3}$, $\kappa_l/2\pi = 2$ GHz, $\kappa_w/2\pi = 20$ GHz, $\Delta_0/2\pi = -32$ GHz, $\phi_{\text{bias}}/2\pi = 0.03$ and the power input to the output waveguide is maintained at 0.025 mW.

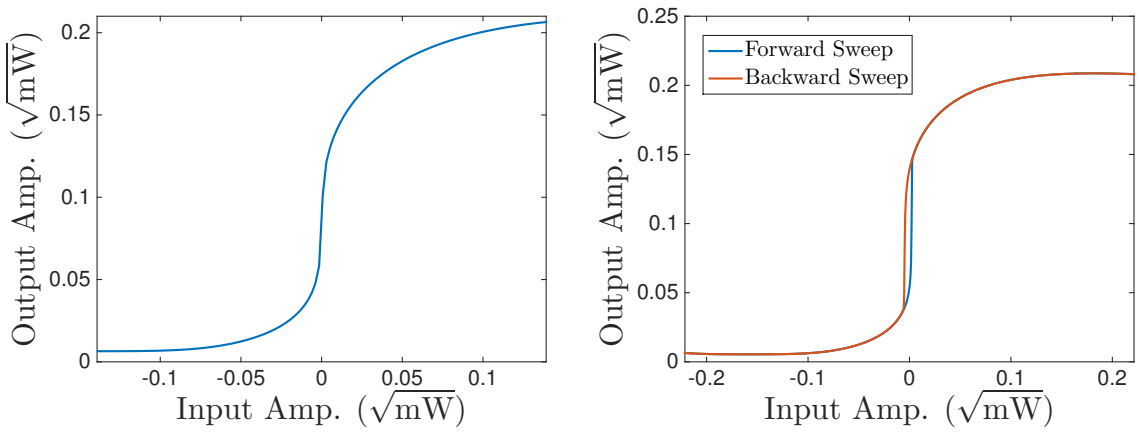
Fig. 4.7a that the output amplitudes in both the waveguides are nearly equal to the input amplitudes – this structure therefore only induces a phase-shift in the waveguide output relative to the waveguide input. It can be noted that the detuning and hence the transmission phase is only a function of the input amplitude and not the input phase since the electrical coupling is affected through the generated photo-current which depends on the power absorbed in the cavity p-n junction. From a design perspective, it is often convenient to cascade a static phase-shifter ϕ_{bias} with this structure so as to control the range of phases that can be achieved by changing the input amplitude.

It is now straightforward to see that the output b_{out} of the circuit implementing the activation function is given by:

$$b_{\text{out}} \approx \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \exp(i\Theta) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} b_0 \\ 0 \end{bmatrix} = \frac{b_0}{\sqrt{2}}(1 - \exp(i\Theta)) \tag{4.6}$$

The input b_{in} to this circuit is linearly combined with a constant signal through a directional coupler and $(b_{\text{in}} + b_{\text{th}})/\sqrt{2}$ fed into the phase-shifter. This allows the control of the activation function threshold – for instance, adjusting b_{th} allows us to design the input threshold to be equal to 0 so as to achieve an approximation of the sigmoid activation function. Fig. 4.8a shows the steady state characteristics of an approximate sigmoid neuron implemented with this structure.

It is often desired to use the neuron as a classifier – i.e. design its activation function to be equal to the step activation function ($\phi(z) = 1$ if $z > 0$ and 0 if $z \leq 0$). Specific applications of such neuron include hardware implemention the support vector machine [34] and reservoir computing systems [36]. Increasing the detuning so as to make the self-electrooptic cavity bistable is favourable in this design, since the thresholding characteristics of a bistable structure becomes much sharper than a monostable structure. Fig. 4.8b shows the forward and reverse sweep characteristics of an optically implemented classification neuron. It can clearly be seen that the transition region for the classification neuron is much smaller than the transition region for the sigmoid



(a) (b)

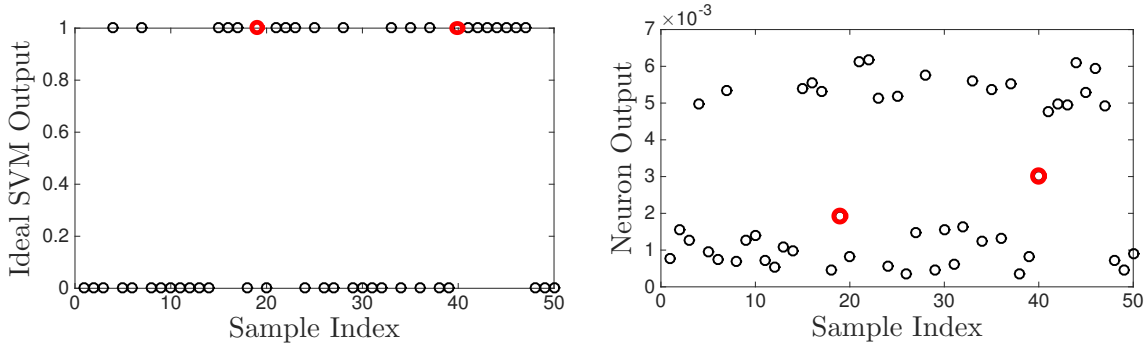
Figure 4.8: Hardware approximations of the (a) Sigmoid activation function (b) Step activation function obtained using the optical neuron. In all simulations, $S = 1$ A/W, $\Gamma_0/2\pi = 500$ GHz/V$^{2/3}$, $\kappa_l/2\pi = 2$ GHz, $\kappa_w/2\pi = 20$ GHz. $\Delta_0/2\pi = -32$ GHz is used for the sigmoid activation function and $\Delta_0/2\pi = -33$ GHz is used for the step activation function.

neuron. However, it should also be noted that this neuron can misclassify the input if it lies within the hysteresis loop of the activation function, but the probability of producing an invalid output (i.e. an output that cannot be interpreted as high or low), is very small due to its bistable characteristics.

4.3 Performance of the optical neuron

So as to benchmark the performance of the designed optical neuron, we use it to solve a simple linear classification problem. The neuron is trained off-chip using the SVM algorithm to compute the weights w_n from the training test. The synapse bias corresponding to the required w_n can be calculated using Fig. 4.5b. Since we are solving a classification problem, it is desirable to operate the activation function bistably (i.e. use large detuning). Fig. 4.9 compares the classification performed by the optical neuron to an ideal SVM. It can clearly be seen that the optical neuron has a performance similar to an ideal SVM trained on the same dataset. The misclassification performed by the optical neuron (shown in red in Fig. 4.9) can be attributed to the nonlinearity in the

synapse, and the indeterminacy in the activation function threshold due to its bistable operation.



(a) (b)

Figure 4.9: Performance of (a) an ideal SVM (b) optical neuron trained on the same training dataset. The red dots indicate a misclassification performed by the optical neuron as compared to the ideal SVM.

Chapter 5

Optimizing the synapse design

The synapse design described in chapter 4 requires the use of 4 self-electrooptic cavities per synapse. In a large neural network, the number of synapses can be huge – as an example, consider a feedforward neural network (Fig. 5.1) – one of the most commonly used learning machines. The number of synapses between two neuron layers is equal to the product of the number of neurons in these layers. Typical feedforward neural networks have as many as 100 neurons per layer, which is equivalent to approximately 10^4 synapse connections between two layers. Integrating these many cavities on-chip

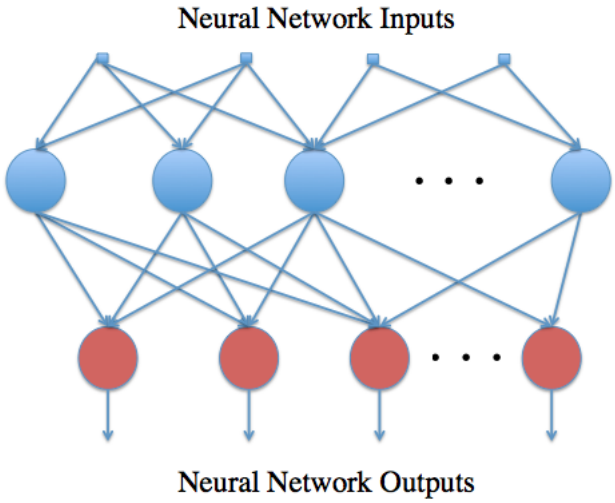


Figure 5.1: Schematic of a feedforward neural network showing different neuron layers and synapses interconnecting those layers.

is experimentally difficult and the chip layout optimisation may itself be an extremely complicated task.

In this chapter, we describe an alternative design of the synapse between two neuron layer. The proposed design uses a silicon based linear structure integrated on top of a planar waveguide to achieve a matrix multiplication equivalent to the synapse operation. The tunability of the synaptic weights is achieved by changing the refractive index of the linear structure by employing the thermooptic property of silicon.

5.1 Image propagation through a tunable linear medium as a vector multiplication

All the designs presented so far used single mode channel waveguides as interconnects to transmit optical signals. Single mode channel waveguides cannot be used transmit spatially multiplexed information (i.e. an image), since only the modal field profile (which has a fixed spatial dependance) can propagate through the waveguide. A planar waveguide, on the other hand, guides electromagnetic fields only in one direction, and

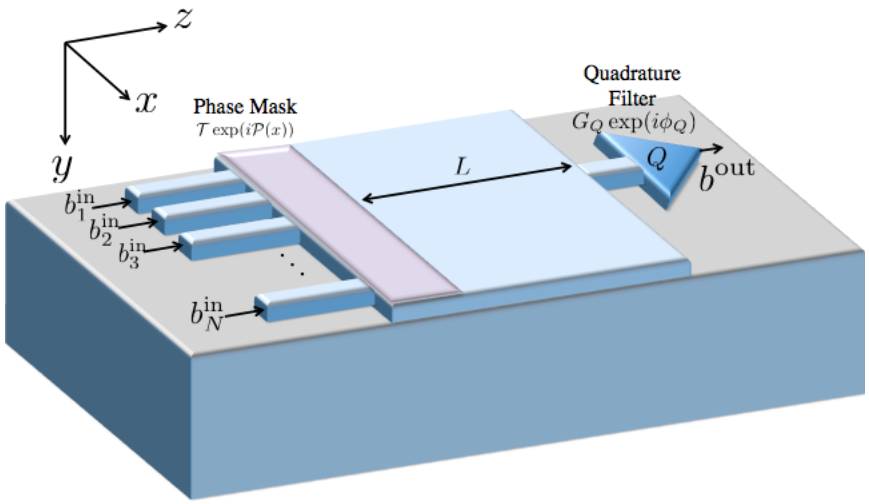


Figure 5.2: Three-dimensional schematic of the modified synapse – the inputs taken from a channel waveguide are fed into a planar waveguide. They then propagate through a phase mask which imparts a x dependant phase profile. After propagation through length L, the imaginary part of the electric field is filtered through a quadrature filter to obtain the synapse output.

can thus be used to transmit a one-dimensional image. We make use of this property of planar waveguides in the subsequent design.

Fig. 5.2 shows a schematic of the modified synapse. The synapse inputs, $b_1^{\text{in}}, b_2^{\text{in}} \ldots b_N^{\text{in}}$, are fed into a planar waveguide as a one dimensional image. This image then propagates through a phase-mask which imparts a different phase to each pixel of the one dimensional image. This phase-mask is designed so as to thermo-optically tunable – the phase profile imparted by this phase-mask can be changed by varying its temperature non-uniformly across the structure. The output of the phase mask propagates through a length L before reaching the quadrature filter. The input to the quadrature filter is thus a linear transformation of the synapse inputs, with the phase-mask introducing tunability in this linear transformation. The quadrature filter then removes the unwanted imaginary part of the input, and the resulting output is a purely real linear combination of the synapse inputs. In the remainder of this section, we show how to design the phase mask corresponding to a given set of synaptic weights w_n.

Assuming the synapse inputs to be fed in through channel waveguides, with modal field profile $\boldsymbol{\mathcal{E}}_c(\boldsymbol{\rho})$, the field incident onto the planar waveguide is given by (Note $\boldsymbol{\rho} \equiv (x, y)$):

$$\mathbf{E}_{\text{inc}}(\boldsymbol{\rho}) = \sum_{n=1}^{N} b_n^{\text{in}} \boldsymbol{\mathcal{E}}_c(\boldsymbol{\rho}_n) \tag{5.1}$$

where $\boldsymbol{\rho}_n = \boldsymbol{\rho} - (n - N/2)\boldsymbol{\rho}_c$ is the transverse coordinate with respect to the centre of the n^{th} channel waveguide, and $\boldsymbol{\rho}_c$ is the centre-centre displacement between two adjacent channel waveguides. Assuming the planar waveguide to be single mode, the electric field transmitted into the planar waveguide can be expressed as:

$$\mathbf{E}_{\text{tran}}(\boldsymbol{\rho}) = \boldsymbol{\mathcal{E}}_p(y) f(x) \tag{5.2}$$

where $\boldsymbol{\mathcal{E}}_p(y)$ is the modal field profile of the planar waveguide and

$$f(x) \approx \sum_{n=1}^{N} b_n^{\text{in}} \int_{-\infty}^{\infty} \boldsymbol{\mathcal{E}}_p^*(y) \cdot \boldsymbol{\mathcal{E}}_c(\boldsymbol{\rho}_n)\, \mathrm{d}y \tag{5.3}$$

This electric field then propagates through a phase mask, at the output of which the electric field is given by:

$$\mathbf{E}_{\text{ph}}(\boldsymbol{\rho}) = \mathcal{T}\boldsymbol{\mathcal{E}}_p(y) f(x) \exp(i\mathcal{P}(x)) \tag{5.4}$$

where $\mathcal{T}$ is the (approximately uniform) transmission amplitude of the phase-mask and $\mathcal{P}(x)$ is its transmission phase. The electric field at the at the input plane corresponding to the quadrature filter can now be computed from $\mathbf{E}_{\text{ph}}$ using (refer to appendix B for derivation):

$$\mathbf{E}_{\text{qf}}(\boldsymbol{\rho}) = \left(\frac{in_{\text{eff}}}{\lambda_0 L}\right)^{1/2} \exp(-ik_0 n_{\text{eff}} L)\mathcal{T}\boldsymbol{\mathcal{E}}_p(y) \int_{-\infty}^{\infty} f(x') \exp(i\mathcal{P}(x')) \times \exp\left(-\frac{ik_0 n_{\text{eff}}}{2L}(x-x')^2\right) \mathrm{d}x' \tag{5.5}$$

Assuming the channel waveguide feeding input to the quadrature filter is centred at $x = 0$, the input to the quadrature filter (b_Q^{in}) can be computed from:

$$b_Q^{\text{in}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \boldsymbol{\mathcal{E}}_c^*(\boldsymbol{\rho}) \cdot \mathbf{E}_{\text{qf}}(\boldsymbol{\rho}) \mathrm{d}^2\boldsymbol{\rho} \tag{5.6}$$

The synapse output $b_{\text{sun}}^{\text{out}}$, which is also the activation function input, is thus given by:

$$b_{\text{syn}}^{\text{out}} = \text{Re}(G_Q \exp(i\phi_Q) b_Q^{\text{in}}) \tag{5.7}$$

where $g_+ = G_Q \exp(i\phi_Q)(= g_-^*)$ is the in-phase gain of the quadrature filter. Using Eqs. 5.3, 5.5, 5.6 and 5.7, the synapse output is given by:

$$b_{\text{syn}}^{\text{out}} = \sum_{n=1}^{N} w_n b_n^{\text{in}} \tag{5.8}$$

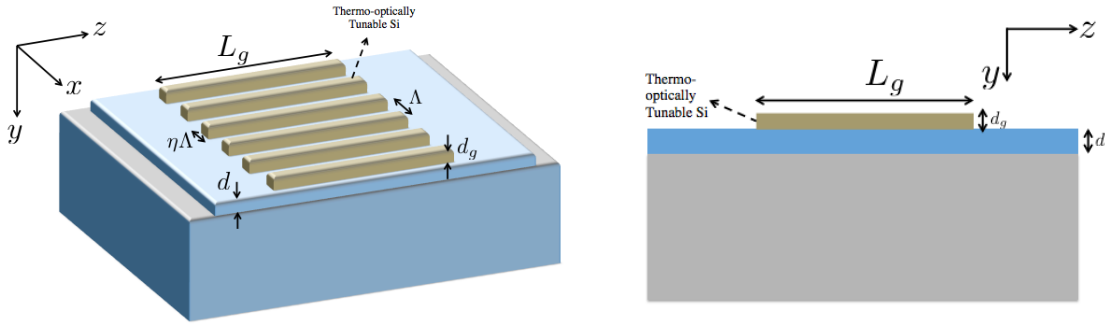
where the weights w_n are given by:

$$w_n = G_Q \left(\frac{n_{\text{eff}}}{\lambda_0 L}\right)^{1/2} \mathcal{T} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\boldsymbol{\mathcal{E}}_p(y) \cdot \boldsymbol{\mathcal{E}}_c^*(\boldsymbol{\rho}))(\boldsymbol{\mathcal{E}}_p^*(y') \cdot \boldsymbol{\mathcal{E}}_c(\boldsymbol{\rho}_n'))\times$$
$$\cos\left(\frac{\pi}{4} + \phi_Q - k_0 n_{\text{eff}} L - \frac{k_0 n_{\text{eff}}}{2L}(x - x')^2 + \mathcal{P}(x')\right) \mathrm{d}x\mathrm{d}y\mathrm{d}x'\mathrm{d}y' \tag{5.9}$$

Eq. 5.9 is the required relationship between the phase mask $\mathcal{P}(x)$ and the synaptic weights w_n, and maybe numerically solved to approximate $\mathcal{P}(x)$ for a given set of synaptic weights.

5.2 Design of on-chip tunable phase mask

The on-chip tunable phase mask can be realized by fabricating a silicon grating on top of the planar waveguide used for transmitting the one-dimensional image as described in the previous section. Three-dimensional and two-dimensional schematics of the proposed structure are shown in Fig. 5.3. The presence of the grating on top of the planar waveguide locally changes the effective refractive index of the planar waveguide – this effective refractive index can be controlled by thermo-optically controlling the refractive index of the grating. Since the phase accumulated by the waveguide mode on propagating through the grating depends on this effective refractive index, this structure provides a thermo-optic control over the imparted phase profile. Silicon's thermo-optic effect [37,38] induces an approximate change of 0.2 in its refractive index over an achievable temperature range (50 K – 300 K). The amount of phase-change this change in refractive index translates to depends on the grating length L_g and the grating thick-

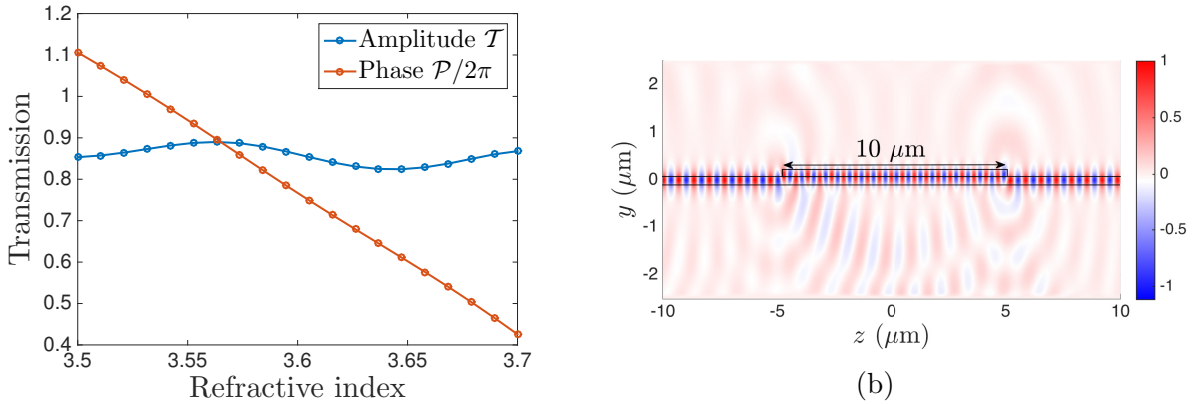


(a) (b)

Figure 5.3: (a) Three-dimensional schematic of the on-chip grating used as a tunable phase mask (b) Two dimensional schematic of the same structure. The two dimensional structure is used for the finite difference frequency domain simulations.

ness d_g. While a larger grating thickness d_g would increase the variation in the grating effective index for a fixed change in the silicon refractive index, it is often desired to retain a single-mode structure through out and increasing d_g indefinitely might result in multiple modes propagating through the grating region. d_g is thus bounded by this requirement.

Increasing the grating length L_g increases the range of phases that can be achieved by this design for a fixed change in silicon refractive index. We verified the design's performance by simulating a two-dimensional structure (Fig. 5.3b) in using the finite difference frequency domain simulator [39, 40]. It was observed that for $L_g \sim 10$ μm,



(a) (b)

Figure 5.4: (a) Variation of the transmission amplitude and phase of the two dimensional structure shown in Fig. 5.3b (b) Electric field profile for the same structure. In all the simulations, $d = d_g = 120$ nm, $L_g = 10$ μm and $n_{\text{Si}} = 3.5$.

all phases from 0 to π can be achieved over a change of 0.2 in the silicon refractive index (Fig. 5.4a). The transmission amplitude of this structure remained approximately constant ($\mathcal{T} \sim 0.8$). Fig. 5.4b shows the electric field profile obtained on simulating this structure. It can clearly be seen that within the grating region, the mode maxima shifts to lie in the grating layer rather than the original waveguiding layer – this is a consequence of the grating layer having a refractive index similar to that of the guiding layer. The coupling between the grating layer and the planar waveguide is therefore stronger than an evanescent coupling (which would have occurred if the grating layer had a refractive index much smaller than the guiding layer). A large grating length is thus not required to achieve a desired variation in the output phase of this structure.

Chapter 6

Conclusion

In this thesis, we presented the self-electrooptic cavity as a fundamental nonlinear optical element that can be used to design silicon based optical information processing systems. The self-electrooptic cavity was analyzed and an approximate time-domain simulation model to compute its characteristics was derived. The self-electrooptic cavity was then used to design optical logic gates (NOT gate, AND gate, latch) and an optical neuron. The design methodology described in this thesis can form the basis of building advanced computing systems, including large-scale digital systems and neuromorphic learning machines.

A significant part of this thesis was devoted to the problem of designing an optical neuron. In an integrated optics setting, using the synaptic weights based on the self-electrooptic cavity would require a very large number of cavities for implementing a neural network. As an alternative, synapse based on propagation of one-dimensional image through a planar waveguide was proposed. The fundamental building block of this alternative synapse was a tunable phase-shifter – using simple two dimensional simulations, it was shown that an on-chip tunable phase-shifter could be realized by fabricating thermooptically controlled silicon structures on top of the planar waveguide used for image propagation.

The designs developed and presented in this thesis are expected to be of relevance to the implementation of silicon based information processing systems.

Appendices

Appendix A

Derivation of Equation 4.3

To analyze the differential amplifier (Fig. 4.2), we first compute the small signal gain of a self-electrooptic cavity coupled to a single waveguide. From section 2.2, the transmission amplitude $\tau(B_{\text{in}})$ and phase $\Phi(B_{\text{in}})$ are defined as: for a constant input $b_{\text{in}} = B_{\text{in}} \exp(i\phi_{\text{in}})$, the steady state output is given by:

$$b_{\text{out}} = \tau(B_{\text{in}}) \exp(i\Phi(B_{\text{in}}))b_{\text{in}} = \tau(B_{\text{in}})B_{\text{in}} \exp(i(\Phi(B_{\text{in}}) + \phi_{\text{in}})) \tag{A.1}$$

Consider an input of the form $b_{\text{in}} = B_0 \exp(i\phi_0) + \delta b_{\text{in}}$, where δb_{in} is a small (complex) signal. The output would consequently be of the form $b_{\text{out}} \approx \tau(B_0)B_0 \exp(i(\phi_0 + \Phi(B_0))) + \delta b_{\text{out}}$ where

$$\delta b_{\text{out}} = \left.\frac{\partial b_{\text{out}}}{\partial b_{\text{in}}}\right|_{B_0 \exp(i\phi_0)} \delta b_{\text{in}} + \left.\frac{\partial b_{\text{out}}}{\partial b^*_{\text{in}}}\right|_{B_0 \exp(i\phi_0)} \delta b^*_{\text{in}} \tag{A.2}$$

The partial derivatives can easily be computed by changing variables from $(b_{\text{in}}, b^*_{\text{in}})$ to $(B_{\text{in}}, \phi_{\text{in}})$ through:

$$B_{\text{in}} = \sqrt{b^*_{\text{in}} b_{\text{in}}} \text{ and } \phi_{\text{in}} = -\frac{i}{2}\log\left(\frac{b_{\text{in}}}{b^*_{\text{in}}}\right) \tag{A.3}$$

Thus,

$$\begin{aligned}
\frac{\partial b_{\text{out}}}{\partial b_{\text{in}}} &= \frac{\partial b_{\text{out}}}{\partial B_{\text{in}}}\frac{\partial B_{\text{in}}}{\partial b_{\text{in}}} + \frac{\partial b_{\text{out}}}{\partial \phi_{\text{in}}}\frac{\partial \phi_{\text{in}}}{\partial b_{\text{in}}} \\
&= \frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}(B_{\text{in}}\tau(B_{\text{in}})\exp(i(\Phi(B_{\text{in}})+\phi_{\text{in}})))\left(\frac{b_{\text{in}}^*}{4b_{\text{in}}}\right)^{1/2} - \frac{iB_{\text{in}}}{2b_{\text{in}}}(i\tau(B_{\text{in}})\exp(i(\Phi(B_{\text{in}})+\phi_{\text{in}}))) \\
&= \frac{1}{2B_{\text{in}}}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}(B_{\text{in}}\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))) + \frac{1}{2}\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}})) \\
&= \tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}})) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}(\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))) \qquad \text{(A.4)}
\end{aligned}$$

and similarly,

$$\begin{aligned}
\frac{\partial b_{\text{out}}}{\partial b_{\text{in}}^*} &= \frac{\partial b_{\text{out}}}{\partial B_{\text{in}}}\frac{\partial B_{\text{in}}}{\partial b_{\text{in}}^*} + \frac{\partial b_{\text{out}}}{\partial \phi_{\text{in}}}\frac{\partial \phi_{\text{in}}}{\partial b_{\text{in}}^*} \\
&= \frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}(B_{\text{in}}\tau(B_{\text{in}})\exp(i(\Phi(B_{\text{in}})+\phi_{\text{in}})))\left(\frac{b_{\text{in}}}{4b_{\text{in}}^*}\right)^{1/2} + \frac{iB_{\text{in}}}{2b_{\text{in}}}(i\tau(B_{\text{in}})\exp(i(\Phi(B_{\text{in}})+\phi_{\text{in}}))) \\
&= \frac{\exp(2i\phi_{\text{in}})}{2B_{\text{in}}}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}(B_{\text{in}}\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))) + \frac{\exp(2i\phi_{\text{in}})}{2}\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}})) \\
&= \frac{\exp(2i\phi_{\text{in}})}{2}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}(\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))) \qquad \text{(A.5)}
\end{aligned}$$

Clearly from Fig. 4.2, the inputs to the two cavities are given by $B_{\text{bias}}\exp(i\phi_{\text{bias}}) \pm b_{\text{in}}/\sqrt{2}$. The output of the amplifier is simply a difference of the two cavity outputs, scaled down by a factor of $\sqrt{2}$. To the first order, b_{out} can thus be expressed as:

$$b_{\text{out}} = g_+ b_{\text{in}} + g_- b_{\text{in}} \qquad \text{(A.6)}$$

where from Eqs. A.4 and A.5:

$$g_+ = \tau(B_{\text{bias}})\exp(i\Phi(B_{\text{bias}})) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}\left[\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))\right]_{B_{\text{bias}}} \qquad \text{(A.7a)}$$

$$g_- = \frac{\exp(2i\phi_{\text{bias}})}{2}\frac{\mathrm{d}}{\mathrm{d}B_{\text{in}}}\left[\tau(B_{\text{in}})\exp(i\Phi(B_{\text{in}}))\right]_{B_{\text{bias}}} \qquad \text{(A.7b)}$$

Appendix B

Derivation of Equation 5.5

Consider the problem of computing the electric field profile at $z = L$ given the electric field profile at $z = 0$ in a planar waveguide (with z axis as the propagation constant and y axis as the confinement direction). Let the field at $z = 0$ be given by:

$$\mathbf{E}(x, y, z = 0) = \boldsymbol{\mathcal{E}}_p(y) f(x) \tag{B.1}$$

wherein we have assumed that the fields are sufficiently paraxial so as to ignore difference in the polarisation of modes not propagating along the z-axis. This electric field can be expressed as a superposition of planar waveguide modes propagating along directions not parallel to the z-axis. To see this, we re-express Eq. B.1 as:

$$\mathbf{E}(x, y, z = 0) = \int_{-\infty}^{\infty} \boldsymbol{\mathcal{E}}_p(y) \tilde{f}(k_x) \exp(ik_x x)\ \mathrm{d}k_x \tag{B.2}$$

where $\tilde{f}(k_x)$ is the fourier transform of $f(x)$:

$$\tilde{f}(k_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \exp(-ik_x x)\ \mathrm{d}x \tag{B.3}$$

Each of the components $\boldsymbol{\mathcal{E}}(y)\tilde{f}(k_x)$ corresponds to a waveguide mode propagating along an angle $\sin^{-1}(k_x/k_0 n_{\text{eff}})$ with $z-$axis (n_{eff} is the effective index of the waveguide). The

electric field at $z = L$ can be computed by multiplying each of these components by an appropriate phase shift according to their propagation direction:

$$\mathbf{E}(x, y, z = L) = \int_{-\infty}^{\infty} \boldsymbol{\mathcal{E}}_p(y) \tilde{f}(k_x) \exp(i(k_x x + k_z L))\ \mathrm{d}k_x \tag{B.4}$$

where $k_z = \sqrt{k_0^2 n_{\text{eff}}^2 - k_x^2}$. For paraxial fields, $k_x << k_0 n_{\text{eff}}$, and hence $k_z \approx k_0 n_{\text{eff}} - k_x^2/2k_0 n_{\text{eff}}$. Using this approximation and Eq. B.3, Eq. B.4 can be expressed as:

$$\begin{aligned}
\mathbf{E}(x, y, z = L) &= \frac{\boldsymbol{\mathcal{E}}_p(y) \exp(ik_0 n_{\text{eff}} L)}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x') \exp(-ik_x x') \exp\left\{ i\left(k_x x + \frac{k_x^2 L}{2k_0 n_{\text{eff}}} \right) \right\} \mathrm{d}k_x \mathrm{d}x' \\
&= \frac{\boldsymbol{\mathcal{E}}_p(y) \exp(ik_0 n_{\text{eff}} L)}{2\pi} \int_{-\infty}^{\infty} f(x') \left[\int_{-\infty}^{\infty} \exp\left\{ i\left(k_x(x - x') + \frac{k_x^2 L}{2k_0 n_{\text{eff}}} \right) \right\}\ \mathrm{d}k_x \right] \mathrm{d}x'
\end{aligned} \tag{B.5}$$

The integral with respect to k_x in Eq. B.5 can easily be evaluated to obtain:

$$\mathbf{E}(x, y, z = L) = \left(\frac{in_{\text{eff}}}{\lambda_0 L} \right)^{1/2} \exp(-ik_0 n_{\text{eff}} L) \boldsymbol{\mathcal{E}}_p(y) \int_{-\infty}^{\infty} f(x') \exp\left(-\frac{ik_0 n_{\text{eff}} (x - x')^2}{2L} \right)\ \mathrm{d}x' \tag{B.6}$$

which is the desired result.

Bibliography

[1] Yifang Chen. Nanofabrication by electron beam lithography and its applications: A review. *Microelectronic Engineering*, 135:57–72, 2015.

[2] Ivo Utke, Stanislav Moshkalev, and Phillip Russell. *Nanofabrication using focused ion and electron beams: principles and applications*. Oxford University Press, 2012.

[3] Phillip Kaye, Raymond Laflamme, and Michele Mosca. An introduction to quantum computing. 2006.

[4] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[5] Robert Kozma, Robinson E Pino, and Giovanni E Pazienza. *Advances in neuromorphic memristor science and applications*, volume 4. Springer Science & Business Media, 2012.

[6] Christian Koos, Philipp Vorreau, Thomas Vallaitis, Pieter Dumon, Wim Bogaerts, Roel Baets, Bweh Esembeson, Ivan Biaggio, Tsuyoshi Michinobu, François Diederich, et al. All-optical high-speed signal processing with silicon–organic hybrid slot waveguides. *Nature Photonics*, 3(4):216–219, 2009.

[7] Carsten Langrock, Saurabh Kumar, John E McGeehan, AE Willner, and MM Fejer. All-optical signal processing using χ (2) nonlinearities in guided-wave devices. *Lightwave Technology, Journal of*, 24(7):2579–2592, 2006.

[8] YA Zaghloul and ARM Zaghloul. Complete all-optical processing polarization-based binary logic gates and optical processors. *Optics Express*, 14(21):9879–9895, 2006.

[9] D Cotter, RJ Manning, KJ Blow, AD Ellis, AE Kelly, D Nesset, ID Phillips, AJ Poustie, and DC Rogers. Nonlinear optics for high-speed digital information processing. *Science*, 286(5444):1523–1528, 1999.

[10] Taylor K Fryett, Christopher M Dodson, and Arka Majumdar. Cavity enhanced nonlinear optics for few photon optical bistability. *Optics express*, 23(12):16246–16255, 2015.

[11] Rahul Trivedi, Uday K. Khankhoje, and Arka Majumdar. Cavity-enhanced second-order nonlinear photonic logic circuits. *Phys. Rev. Applied*, 5:054001, May 2016.

[12] Hideo Mabuchi. Nonlinear interferometry approach to photonic sequential logic. *Applied Physics Letters*, 99(15):153103, 2011.

[13] Charles Santori, Jason S Pelc, Raymond G Beausoleil, Nikolas Tezak, Ryan Hamerly, and Hideo Mabuchi. Quantum noise in large-scale coherent nonlinear photonic circuits. *Physical Review Applied*, 1(5):054005, 2014.

[14] Arka Majumdar, Michal Bajcsy, Dirk Englund, and Jelena Vuckovic. All optical switching with a single quantum dot strongly coupled to a photonic crystal cavity. *Selected Topics in Quantum Electronics, IEEE Journal of*, 18(6):1812–1817, 2012.

[15] Paul E Barclay, Kai-Mei Fu, Fedor Jelezko, and Marko Loncar. Diamond photonics: introduction. *JOSA B*, 33(4):DP1–DP1, 2016.

[16] Ichiro Shoji, Takashi Kondo, Ayako Kitamoto, Masayuki Shirane, and Ryoichi Ito. Absolute scale of second-order nonlinear-optical coefficients. *JOSA B*, 14(9):2268–2294, 1997.

[17] M Dinu, F Quochi, and H Garcia. Third-order nonlinearities in silicon at telecom wavelengths. *Applied Physics Letters*, 82(18):2954–2956, 2003.

[18] J Sweet, BC Richards, JD Olitzky, J Hendrickson, G Khitrova, HM Gibbs, D Litvinov, D Gerthsen, DZ Hu, DM Schaadt, et al. Gaas photonic crystal slab nanocavities: Growth, fabrication, and quality factor. *Photonics and Nanostructures-Fundamentals and Applications*, 8(1):1–6, 2010.

[19] Arka Majumdar and Dario Gerace. Single-photon blockade in doubly resonant nanocavities with second-order nonlinearity. *Physical Review B*, 87(23):235319, 2013.

[20] Arka Majumdar and Armand Rundquist. Cavity-enabled self-electro-optic bistability in silicon photonics. *Optics letters*, 39(13):3864–3867, 2014.

[21] Raúl Rojas. *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.

[22] A Afroozeh, IS Amiri, K Chaudhary, J Ali, and PP Yupapin. Analysis of optical ring resonator. *Advances in Laser and Optics Research*, pages 1–16, 2014.

[23] Simon M Sze and Kwok K Ng. *Physics of semiconductor devices*. John wiley & sons, 2006.

[24] Graham T Reed, G Mashanovich, FY Gardes, and DJ Thomson. Silicon optical modulators. *Nature photonics*, 4(8):518–526, 2010.

[25] Stefan Alexander Maier. *Plasmonics: fundamentals and applications*. Springer Science & Business Media, 2007.

[26] BJ Lee and ZM Zhang. Temperature and doping dependence of the radiative properties of silicon: Drude model revisited. In *Advanced Thermal Processing of Semiconductors, 2005. RTP 2005. 13th IEEE International Conference on*, pages 10–pp. IEEE, 2005.

[27] CK Sun, PKL Yu, Ching-Ten Chang, and DJ Albares. Simulation and interpretation of fast rise time light-activated pin diode switches. *Electron Devices, IEEE Transactions on*, 39(10):2240–2247, 1992.

[28] M Martini and G Ottaviani. Ramo's theorem and the energy balance equations in evaluating the current pulse from semiconductor detectors. *Nuclear Instruments and Methods*, 67(1):177–178, 1969.

[29] Wladyslaw Dabrowski. Transport equations and ramo's theorem: Applications to the impulse response of a semiconductor detector and to the generation-recombination noise in a semiconductor junction. *Progress in quantum electronics*, 13(3):233–266, 1989.

[30] Martin Dentan and BAUDOUIN de Cremoux. Numerical simulation of the nonlinear response of a pin photodiode under high illumination. *Lightwave Technology, Journal of*, 8(8):1137–1144, 1990.

[31] Ahmad Rifqi Md Zain, Nigel P Johnson, Marc Sorel, and M Richard. Ultra high quality factor one dimensional photonic crystal/photonic wire micro-cavities in silicon-on-insulator (soi). *Optics express*, 16(16):12084–12089, 2008.

[32] Parag B Deotare, Murray W McCutcheon, Ian W Frank, Mughees Khan, and Marko Lončar. High quality factor photonic crystal nanobeam cavities. *Applied Physics Letters*, 94(12):121106, 2009.

[33] Amnon Yariv and Pochi Yeh. *Photonics: optical electronics in modern communications (the oxford series in electrical and computer engineering)*. Oxford University Press, Inc., 2006.

[34] Marti A. Hearst, Susan T Dumais, Edgar Osman, John Platt, and Bernhard Scholkopf. Support vector machines. *Intelligent Systems and their Applications, IEEE*, 13(4):18–28, 1998.

[35] Nikolas Tezak and Hideo Mabuchi. A coherent perceptron for all-optical learning. *EPJ Quantum Technology*, 2(1):1–22, 2015.

[36] Benjamin Schrauwen, David Verstraeten, and Jan Van Campenhout. An overview of reservoir computing: theory, applications and implementations. In *Proceedings of the 15th European Symposium on Artificial Neural Networks. p. 471-482 2007*, pages 471–482, 2007.

[37] J Komma, C Schwarz, G Hofmann, D Heinert, and R Nawrodt. Thermo-optic coefficient of silicon at 1550 nm and cryogenic temperatures. *Applied Physics Letters*, 101(4):1905, 2012.

[38] Bradley J Frey, Douglas B Leviton, and Timothy J Madison. Temperature-dependent refractive index of silicon and germanium. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 62732J–62732J. International Society for Optics and Photonics, 2006.

[39] Wonseok Shin and Shanhui Fan. Choice of the perfectly matched layer boundary condition for frequency-domain maxwell?s equations solvers. *Journal of Computational Physics*, 231(8):3406–3431, 2012.

[40] Finite difference frequency domain. https://github.com/wsshin/maxwellfdfd, 2015.