Integrated Optical Analog-to-Digital Converter

by

Anatol Khilo

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2008

© Massachusetts Institute of Technology 2008. All rights reserved.

Author...
Department of Electrical Engineering and Computer Science
February 1, 2008

Certified by...........   ....
Franz X. Kärtner
Professor of Electrical Engineering
Thesis Supervisor

Accepted by   ..../...
Terry P. Orlando
Chairman, Department Committee on Graduate Students
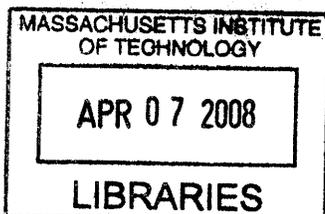
INTEGRATED OPTICAL ANALOG-TO-DIGITAL CONVERTER

by

ANATOL KHILO

Submitted to the Department of Electrical Engineering and Computer Science
on February 1, 2008 in partial fulfillment of the
requirements for the degree of Master of Science in
Electrical Engineering and Computer Science

Abstract

An optically-sampled frequency-demultiplexed wideband analog-to-digital converter (ADC) which has potential to exceed the performance of electronic ADCs by orders of magnitude is studied analytically and numerically. The accuracy of the ADC as a function of its parameters is analyzed and impact of various imperfections of ADC components on its operation is evaluated. A universal error compensation algorithm for improving the conversion accuracy is proposed. On the way to implementation of the integrated optical ADC, two of its critical components – ring resonator filter bank and fiber-to-chip coupler – are designed. A novel coupler from a standard single mode fiber to a strongly confining silicon waveguide is proposed. The results of characterization of the filter bank and fiber-to-chip coupler fabricated on the silicon-on-insulator platform are presented and analyzed.

Thesis Supervisor: Franz X. Kärtner
Title: Professor of Electrical Engineering

CONTENTS

ACKNOWLEDGEMENTS

Support

I. INTRODUCTION

1. Using Optics for Analog-to-Digital Conversion

Analog-to-digital converters (ADCs) are the window that connects the world of analog signals, which is the only kind of signals existing in nature, and the digital world of modern electronics. ADCs are used to digitize speech, communications and imaging signals, radar signals, for following processing in digital domain. The constantly increasing performance of digital electronics opens new exciting possibilities of signal processing and require increasingly fast and accurate ADCs. The performance of electronic ADCs is constantly improving too, but not as fast as the performance of digital electronics [1], at the rate of approximately 1.5 bits in 8 years [2].

The review of electronic analog-to-digital converters by R. Walden examines the factors limiting the performance of electronic ADCs [2]. Fig. 1 shows the plot of effective number of bits (ENOB) achieved by electronic ADCs as a function of frequency of the input analog signal. The accuracy of the conversion decreases at the rate of 1 bit/octave due to sample-to-sample variation of the instant in time at which sampling occurs. This variation is called the aperture jitter and is now the primary factor limiting the performance of the electronic ADCs. The state-of-the art electronic jitter is about 100fs. The idea to use optics for analog-to-digital conversion comes from availability of remarkably stable optical pulse sources generated by mode-locked lasers with timing jitter an order of magnitude smaller than the electronic limit.

The research in the field of optical ADCs has been going on for the last 30 years, starting with the papers from Taylor in the late seventies [3]. However, up to now optical ADCs did not find application in practical systems. The renewed interest to optical conversion technology [1, 4] is caused by the progress in mode-locked laser technologies and the advent of silicon photonics which enables integration of optical devices together with electronics..
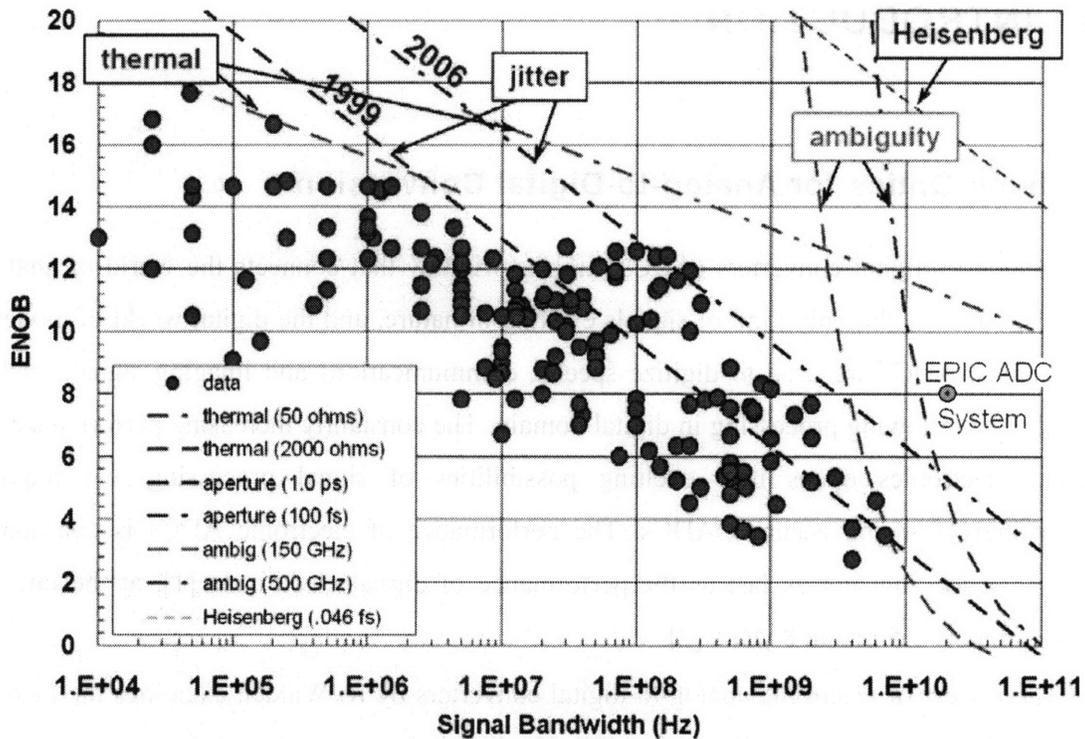
Figure 1. Effective number of bits (ENOB) of electronic ADCs as a function frequency of the input signal. The red circles represent ADCs reported in the literature. The lines show the limits imposed on conversion accuracy by thermal noise, aperture jitter, comparator ambiguity and Heisenberg principle. The plot is taken from R. Walden's review of electronic ADC converters [1,2]. The green circle represent the target parameters of the EPIC ADC system.

There is a large variety of the optical ADCs proposed in the literature. One can distinguish four classes of photonic ADCs [1]:

1) photonic assisted ADC, in which a photonic device is used to improve performance of an electronic ADC;

2) photonic sampling and electronic quantizing ADC, where the optics is used to sample the analog electric signal which is then quantized by an electronic ADC;

3) electronic sampling and photonic quantizing ADC;

4) photonic sampling and quantizing ADC.

In this thesis ADCs of the second type are considered, where the analog electronic signal is optically sampled and electronically quantized. The idea behind optical sampling is to modulate the amplitude of the low-jitter optical pulse train with the analog RF signal we want to sample; if the optical pulses are sufficiently short, each of them "cuts out" a very small

portion of the electric signal [3, 5]. The operation of this kind of ADCs is schematically shown in Fig. 2 [1]. Each optical pulse at the output of the modulator has energy proportional to the electric signal at the corresponding time moment and thus represent a sample of the electrical signal. Because the optical pulse train is very stable, very low jitter sampling can be achieved, overcoming the 100fs electronic jitter barrier. The obtained pulse train can then be detected by a photodetector and sampled by an electronic ADC.
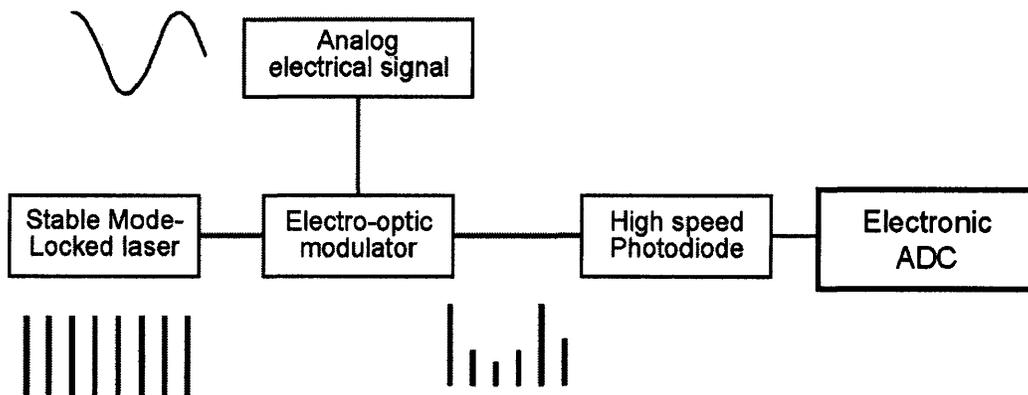
Figure 2. The principle of an optically sampled, electronically quantized ADC (taken from [1]).

Optical sampling we just described helps to reduce the aperture jitter, but still requires fast electronic ADC for quantization of the modulated pulse train. Two ways to reduce the required speed of an electronic ADC are to use time- or frequency-domain demultiplexing. In time domain demultiplexing, the individual pulses are separated into several pulse streams with smaller rate, each of which is then quantized by a slower electronic ADC [6, 7]. In frequency domain demultiplexing, the pulses in the optical pulse train have different frequencies, so that they can be separated using a set of filters and then detected individually [8, 9, 10, 11]. The frequency-differentiated pulse train can be obtained either using delay lines to time interleave pulses at different carrier frequencies, or by introducing time-frequency mapping within each pulse by chirping the pulse train with a dispersive element. The latter approach is adopted in this thesis; the detailed description of the operation of such ADC can be found in chapter 4.

The electronic photonic integrated circuit (EPIC) ADC project being carried out at MIT has the goal of demonstrating analog-to-digital conversion of signals with bandwidths up to

20 GHz with effective resolution of 8 bits, at it is shown with a green circle in Fig. 1, with the final goal of improving the resolution to 10 bits. Note that the stated performance goals are not at the limit of the optical ADC technology: the low jitter of the mode-locked lasers allows to improve the performance even further.

The unique feature of the EPIC ADC project is that the optical ADC is being implemented using integrated optical components on a single chip. The advances in silicon microphotonics and nanofabrication technologies make this goal feasible. At the first step, and external mode-locked laser is used, and the photodiode output is quantized by external electrical ADC. However, efforts are under way to implement the mode-locked laser on the same chip with the rest of the ADC system.

2. Silicon Photonics

Integrated optics has been around for many decades, with mainly low-index-contrast structures being the object of research and application. Recent years witnessed an explosion of interest to integrated optics because of the shift towards high-index-contrast (HIC) structures, such as those implemented on the silicon-on-insulator platform. High-index-contrast structures has the following advantages over their low-index-contrast counterparts.

• A waveguide with high index contrast between the core and overcladding has much smaller cross-section. For example, a typical silicon waveguide with an index contrast of about $\Delta n = 2.0$ has a cross-section of about 0.45 x 0.2μm, about three orders of magnitude smaller than that of a standard fiber with index contrast of $\Delta n = 0.003$ and core radius of about 9μm. In addition to being small, HIC waveguides support a much smaller bend radii without radiation losses. This allows miniaturization of photonic devices and packing numerous devices on the same chip, potentially enabling creation of complex integrated optical with sophisticated functionality.

• An important consequence of the ability of HIC waveguide to support small bend radii is the possibility to create filters with large free spectral range (FSR). The filters implied here are the ring resonator filters, which in the simplest case are just a single ring which is evanescently coupler to two waveguides (see Fig. 3). In principle of operation, ring filters are similar to Fabry-Perot resonators. For resonant frequencies, i.e. the frequencies with the phase shift per round trip being an integer number of 2π, the light will accumulate in the

resonator; the light exiting the ring into the through port with interfere destructively with the input light passing by the ring, thus suppressing the field in the through port. The light at resonant frequency accumulated in the ring will exit it through the drop port, and the nonresonant light will pass the ring and go to the through port, thus making the ring a frequency-selective element. The free spectral range, i.e. the frequency spacing between the adjacent resonant frequencies, is an inverse of the ring round trip time. The smaller the ring, the shorter is the round-trip time, and the larger is FSR. Ring filters implemented on silicon-on-insulator platform can have FSR of up to 5 THz (40nm) and more. This makes it possible to create useful devices such as optical add-drop demultiplexers and reconfigurable switches, which was hardly possible with low-index contrast waveguides because of the limited useful frequency range of the filters.

drop port

input   through port
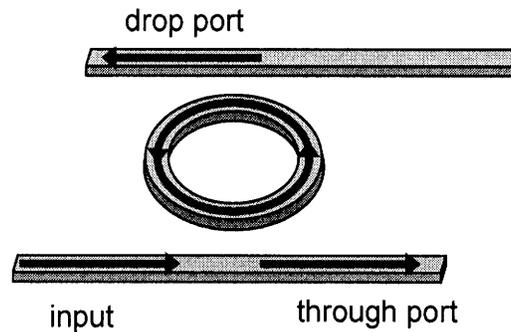
Figure 3. Single ring filter. The input light is transmitted into the drop port at resonant frequencies and into the though port otherwise.

• Another critical advantage of the silicon photonics is the ability to use the same fabrication techniques which are successfully used in integrated electronics industry, which is also based on silicon. This makes potential wide-spread application of the integrated optics much more realistic because the same infrastructure and low-cost mass-production fabrication facilities available without additional investments.

• In addition to economic benefit of compatibility with semiconductor electronics, silicon photonics has potential to produce also important functional benefits. The idea is to use electronics and photonics on the same circuit. The electronic-photonic integrated circuits can implement the functions at which the light is best at with photonic elements, and the functions where the electronics has advantage with electronic elements. For example, light can be used for data transport between the memory and the processors or between the cores

of the same processor, potentially eliminating the interconnection bottleneck which hampers the rate of performance enhancement of the modern processors.

There are also several challenges on the way to application of HIC microphotonics.

• The microphotonic devices are very sensitive to fabrication errors because of their small size. For example, the resonant frequency of a ring filter can shift by a THz per 1 nm of error in Si waveguide height. Development of fabrication technologies [12] and error-tolerant device designs [13, 14] are needed to overcome this problem.

• The HIC structures are strongly polarization-dependent because of large effective index difference between the TE and TM modes. On the other hand, the light transmitted through an optical fiber can have an arbitrary state of polarization. To be useful in fiber optics links, the HIC devices must be made polarization-independent. One way to achieve it is to use polarization-diversity scheme where the TE and TM polarization are split, the TM polarization is transformed into TE, the two are processed by identical photonic structures optimized for TE polarization [15].

• Miniaturization of HIC structures makes it hard to couple light into them from a fiber. The reason is the mode mismatch between the fiber and HIC waveguide which can reach three orders of magnitude. The problem is illustrated in Fig. 4, which compares the size of the core of a standard single-mode fiber to the core of a Si waveguide. Inefficient coupling can be the main source of power loss in a HIC device such as optical add-drop multiplexer, therefore an efficient fiber-to-chip coupler is a critical milestone on the way to practical application of HIC photonic circuits. Section IV of the thesis is devoted to the coupling problem.
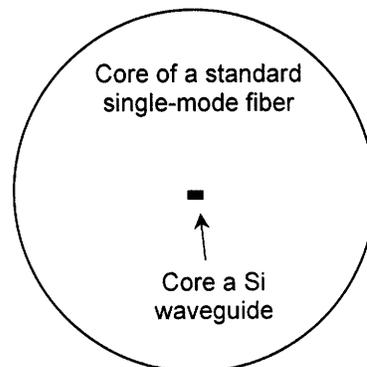
Figure 4. The core of a standard single-mode fiber (circle) and a silicon waveguide (rectangle), drawn to scale.

3. Outline of the thesis

The thesis comprises three main parts. The first is devoted to system aspects of the photonic ADC system, the and the other two – to the design of two individual components for EPIC ADC system, a filter bank and a fiber-to-chip coupler.

The photonic ADC section starts with a detailed description of the operation of the optically-sampled frequency-demultiplexed ADC system which is a subject of this thesis. The basic system as well as some of its variations are considered. The analysis of the ADC performance starts in chapter 5 with derivation of an analytical expression describing the ADC output in the ideal case when the components of the ADC do not have any deviations from the required performance. The obtained expressions are analyzed in chapter 6 in order to understand fundamental limits on accuracy of the ADC. It is shown that even in an ideal case direct reconstruction of the analog RF signal from the ADC output is problematic. A numerical iterative error compensation algorithm is proposed in chapter 7. This algorithm allows to improve the accuracy of the ADC at the post-processing step provided that ADC parameters are properly characterized. Memory and computation time requirements of the algorithm are discussed and several ways to improve them are suggested. Some aspects of characterization of the ADC system in order to obtain the parameters necessary to run the error compensation algorithm are studied in chapter 8. Chapter 9 carries out an extensive study of imperfections in different ADC components such as modulator or the filter bank. The impact on the accuracy of the ADC is examined; the results establish guidelines for requirements to ADC system components and precision with which they are characterized to allow error compensation. The results also help to choose the parameters of the ADC system to make it less sensitive to imperfections of its components. Although the EPIC ADC project that the author was a part of was concerned with development of a particular ADC, and attempt was made throughout this section to generalize the description of the ADC and understand the inherent tradeoffs in selection of its parameters.

Section III is devoted to design of the filter band for EPIC ADC system. The procedure of the design of a 20-channel second- and third-order filter banks is explained in chapter 10. The experimental results are presented and interpreted in chapter 11.

Section III described the design of horizontal fiber-to-chip couplers and starts with an overview of existing couplers from a fiber to a HIC waveguide. The theory behind the modeling and optimization of adiabatic couplers is presented in chapter 13. The next chapter

applies these methods to design a fiber-to-chip coupler which works with small-core input fibers. Chapter 0 presents a novel two-stage adiabatic coupler designed to work with standard single-mode fibers with large mode size. The results of the characterization of the coupler designed in chapter 14 are presented and discussed in the final chapter.

The appendices contain the description of the approach of calculation of effective number of bits used in Section II and a list of notation used in this thesis.

II. PHOTONIC ANALOG-TO-DIGITAL CONVERTER

4. Optically-Sampled Frequency-Demultiplexed ADC System

1. General concept of ADC

This chapter describes the operation of the ADC system which is the subject of the present work. This ADC uses sampling in optical domain and applies frequency demultiplexing to achieve high sampling rate, as it is explained below.

The layout of the ADC system is shown in Fig. 5 [16, 17]. A low-jitter pulse train with repetition rate f_R generated by a mode-locked laser passes through a dispersive fiber with length L and dispersion coefficient β_2. The fiber imposes chirp on the pulse train so that a frequency component ω gets delayed by

$$\tau(\omega) = \beta_2(\omega - \omega_{ref})L, \qquad (4.1)$$

where ω_{ref} is the reference frequency at which the group delay is assumed to be zero. The impact of the dispersive fiber on a pulse is illustrated in Fig. 6. The initial broadband pulse can be represented as a superposition of sub-pulses corresponding to different frequency bands of the whole pulse, as shown in Fig. 6(a). Each of these sub-pulses travels with different group velocity in the dispersive fiber, so at the output these sub-pulses will spread across some time interval, as shown in Fig. 6(b).
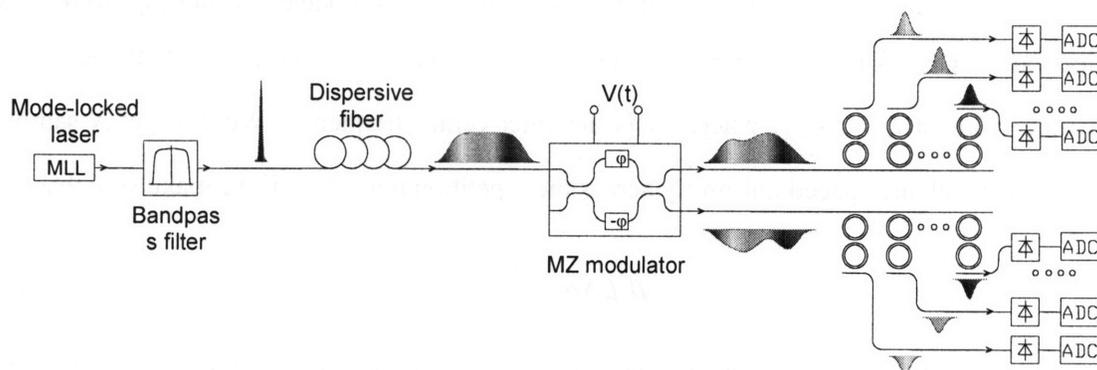
Figure 5. Layout of the optically sampled frequency-demultiplexed ADC.
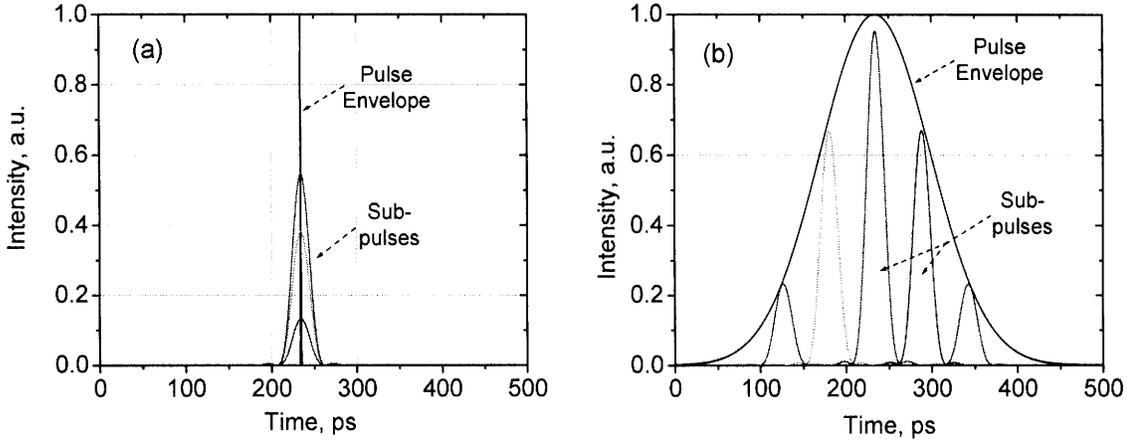
Figure 6. A pulse from the mode-locked laser and its sub-pulses as selected by optical filter bank (a) at the input, (b) at the output of the dispersive fiber.

The chirped pulse train is then modulated by a Mach-Zehnder (MZ) modulator whose RF driving voltage $v(t)$ is the signal to be sampled. Because of time-frequency mapping expressed by Eq. (4.1), the modulator effectively imprints the time dependence of $v(t)$ onto the optical spectrum; each of the sub-pulses gets multiplied by $v(t)$ at corresponding time moment. After this the optical signal is demultiplexed into N channels by a filter bank; each filter of this bank selects a sub-pulse at corresponding frequency. The filter bank shown in Fig. 5 contains second-order ring filters for implementation in an EPIC ADC; the filter bank can be implemented in any other way such as arrayed waveguide grating (AWG).

After the filter bank, each sub-pulse is detected by a photodetector and digitized by an electronic ADC taking one sample per pulse. This sample, $v_{ADC}(t_n)$, represents the RF signal at time moment $t_n = \tau(\omega_n)$, where ω_n is the filter center frequency. We will get N samples $v_{ADC}(t_n)$ which are spaced uniformly across the repetition period T_R if we make sure that

$$\beta_2 L \Delta \omega = \frac{T_R}{N}, \tag{4.2}$$

where $\Delta\omega$ is the channel frequency spacing[1]. We thus get an ADC system with sample rate $f_s = f_R N$. The benefit of this approach is that we are solving the problem of sampling at a

---

[1] In our notation, $\Delta\omega$ is frequency spacing in terms of angular frequency, which is related to frequency spacing Δf as $\Delta\omega = 2\pi \Delta f$.

high rate f_s by replacing it with N less challenging problems of sampling at N times smaller rate. For example, a very high sampling rate of high rate $f_s = 40$ GHz can be obtained using 20 channel system with 20 electronic ADCs sampling at 2 GHz; such electronic ADCs are currently commercially available.

The required length of the dispersive fiber can be found from (4.2). Let us rewrite this equation in terms of the dispersion coefficient D, which is normally specified by manufacturers in data sheets. Dispersion coefficient is related to β_2 with $\beta_2 = \dfrac{c}{2\pi f_{ref}^2} D$, where f_{ref} is the dispersion reference frequency at which the group delay is assumed to be zero (in our case this is the center frequency of the filter bank). We can then rewrite (4.2) as

$$\frac{c}{2\pi f_{ref}^2} DL\Delta\omega = \frac{T_R}{N},$$

or, finally,

$$L = \frac{1}{f_R N} \frac{f_{ref}^2}{c} \frac{1}{D\Delta f}, \tag{4.3}$$

where $\Delta f = \Delta\omega/(2\pi)$.

2. MZ linearization using complimentary outputs

After explaining the general principles of ADC operation, let us describe some modifications of the scheme. The first is the use of both complimentary outputs of the MZ modulator (as shown in Fig. 5). The problem with an ADC which uses just one output is that it is susceptible to random pulse-to-pulse energy variations in the mode-locked laser. Such variations cannot be distinguished from the RF signal variations, so the accuracy of the ADC will suffer. Another issue is the nonlinearity of the MZ modulator response for the case of large modulation. If the MZ bias voltage creates a π phase shift between the two arms and the RF voltage produces phase shift φ in the upper arm and $-\varphi$ in the lower arm of the MZ proportional to the drive RF voltage v, the optical fields at the upper and lower output of the MZ are

$$e^{up} = e_0 \cdot \sin\left(\frac{\pi}{4} + v\right),$$

$$e^{down} = e_0 \cdot \cos\left(\frac{\pi}{4} + v\right),$$

where e_0 is the pulse amplitude at the input of the MZ modulator; the proportionality factor between the phase shift φ and voltage v was omitted. When these signals are detected by photodiodes and digitized by electronic ADCs, we obtain

$$v_{ADC}^{up} = |e_0|^2 \cdot \sin^2\left(\frac{\pi}{4} + v\right),$$

$$v_{ADC}^{down} = |e_0|^2 \cdot \cos^2\left(\frac{\pi}{4} + v\right),$$

at the upper and lower outputs, respectively. We see that the ADC output depends on the applied RF signal in a nonlinear way. To linearize the MZ response, we can construct the following expression

$$\frac{v_{ADC}^{up} - v_{ADC}^{down}}{v_{ADC}^{up} + v_{ADC}^{down}}.$$

Substituting the expressions for v_{ADC}^{up} and v_{ADC}^{down} from above, we see that the denominator will be equal to the total pulse of power, $v_{ADC}^{up} + v_{ADC}^{down} = |e_0|^2$, and numerator

$$v_{ADC}^{up} - v_{ADC}^{down} = |e_0|^2 \sin(2v),$$

so that the pulse power will cancel. Therefore, the original RF signal can be recovered using [18, 8]

$$v_{ADC} = \frac{1}{2}\text{asin}\left(\frac{v_{ADC}^{up} - v_{ADC}^{down}}{v_{ADC}^{up} + v_{ADC}^{down}}\right). \tag{4.4}$$

This approach allows to linearize the MZ response and make the ADC output insensitive to pulse-to-pulse amplitude noise. The price we pay for it is the increase in the number of componets: we need to use *dual* filter bank, i.e. the filter bank which has two identical sets of filters, one for the upper and one for the lower MZ output, the number of photodetectors and electronic ADCs will double as well. In the further discussion we will normally assume that ADC uses this linearization scheme. More thorough treatment of the ADC operation and study of the limitations of this linearization approach will be performed in chapter 5.

In an alternative approach, the MZ nonlinearity is cancelled by inverting the sine or cosine modulator transfer function at the post-processing stage [19]. In this case only one MZ output is needed. ADC output for zero RF signal must be accurately measured to be used as a reference level in this method. This method can be efficient in suppressing the nonlinearity of the MZ transfer function. However, it does not reduce susceptibility to pulse-to-pulse amplitude variation of the laser source.

3. Bandlimiting optical filter

One more component of the ADC system which was not mentioned yet is the bandlimiting optical filter which might be necessary to apply if optical filters of the filter bank have finite free spectral range (FSR), which is the case for example for ring or other resonant filters. To understand the problem, let us consider a filter with center frequency ω_n which produces a sample corresponding to time moment $t_n = \tau(\omega_n)$ according to the time-frequency mapping (4.1). However, the filter passband is repeated every FSR, so the filter will also transmit the frequency components $\omega_n \pm FSR$ which are proportional to RF signal at time moments $\tau(\omega_n \pm FSR)$. If the optical spectrum has non-negligible power at these frequencies, the power at these frequency components will add to the primary component at ω_n, leading to crosstalk and ENOB degradation. The bandlimiting optical filter is intended to prevent this from happening by suppressing the optical spectrum outside one FSR of the filter bank. The filter in Fig. 5 is shown before the fiber; note, however, that if nonlinearity of the fiber is significant leading to spectral broadening, is beneficial to use a bandlimiting filter both before the fiber (to reduce optical power and thus nonlinearity) and after the fiber (to get rid of spectral components created by nonlinearity).

4. Integrator

The photodetector in the ADC scheme might be followed by an integrator which is reset every repetition period of the laser. The integrator calculates total power of each pulse which is then quantized by an electronic ADC. The benefit of an integrator is that it makes the ADC system not sensitive to the aperture jitter of the electronic ADC and also simplifies the analysis of the system (see chapter 5 for more details). A precise integrator, however, might be difficult to implement; the ADC system can also work without an integrator, in which case the electronic ADCs should be locked to the laser.

5. Integrated ADC System pursued at MIT

The work for this thesis was done within the MIT EPIC project, where EPIC stands for electronic-photonic integrated circuits. The goal of this project is to demonstrate possibility of building an optically-sampled wavelength-multiplexed system described above with integrated optical components, ideally building the whole ADC system on the same chip [17, 20, 21]. The system has the following components:

• The mode-locked laser, which is initially implemented externally. A mode-locked laser with 200MHz repetition rate, pulse duration of about 200fs and jitter below 16fs is built by Jeff Chen at MIT. The repetition rate is multiplied by 10 in an external repetition rate multiplication cavity to yield 2GHz pulse train. The work on the mode-locked laser integrated on a chip is also underway [22].

• The MZ modulator is made in silicon using carrier injection to change the refractive index of the waveguide [23]. Modulators with almost 40GHz bandwidth have been designed at MIT and fabricated and characterized at Lincoln Lab.

• The filters in the filter bank are implemented as two-ring filters; the material of choice was silicon-rich nitride (see Section III for details). The channel spacing was chosen to be 80GHz, and filter FWHM bandwidth – 25 GHz. The two-ring filters suppress neighboring channels by more than 30 dB according to the design, so the crosstalk between the channels should not be limiting the ENOB to less than about 9.5 bits without any error compensation.

• The photodetectors are being implemented in silicon (using defect state absorption) and silicon germanium.

• Electronic ADCs not implemented on the chip; commercially-available ADCs are used instead for the demonstration.

• The components of ADC system are initially fabricated on independent chips, which will be connected with a fiber. To get the light from the fiber to the chip, a horizontal adiabatic coupler is build, as it is described in detail in Section IV.

The number of channels chosen for the system was 20, so given the 2GHz pulse repetition rate, our goal is an ADC system sampling at 40 Gsamples/s; the bandwidth of the ADC is 20GHz. The target number of bits is 8. The ADC with such target parameters will be referred to as "EPIC ADC" in the rest of the text.

Let us now proceed to the analysis of the operation of the optically-sampled ADC and analyzing its accuracy limits.

5. The Output of an Ideal ADC System

This chapter is devoted to derivation of the expression for the output samples of the optically-sampled ADC described in chapter 4. We will assume that all components of the ADC are ideal e.g. filter center frequencies are precisely spaced, path losses for different channels are exactly balanced, the modulator is biased exactly at the quadrature point, the mode-locked laser produces pulse train without timing jitter, and so on. The goal is to understand limitations inherent to this ADC even in the absence of any component imperfections.

Let us go along the link of the ADC and find the signal at all conversion steps; the ADC architecture and notation for the signals at different stages are illustrated in Fig. 7. The mode-locked laser produces pulse train with amplitude $e_{laser}(t)$ in time domain. The spectrum of the laser $E_{laser}(\omega)$ is connected to $e_{laser}(t)$ by Fourier transform

$$E_{laser}(\omega) = \int e_{laser}(t)\exp(-i\omega t)dt,$$

$$e_{laser}(t) = \frac{1}{2\pi}\int E_{laser}(\omega)\exp(i\omega t)d\omega.$$

We adopt the notation that lowercase letters correspond to time domain and uppercase – to frequency domain. If a bandwidth limiting filter is used at the laser output, we assume for simplicity that its impact is included into $e_{laser}(t)$.
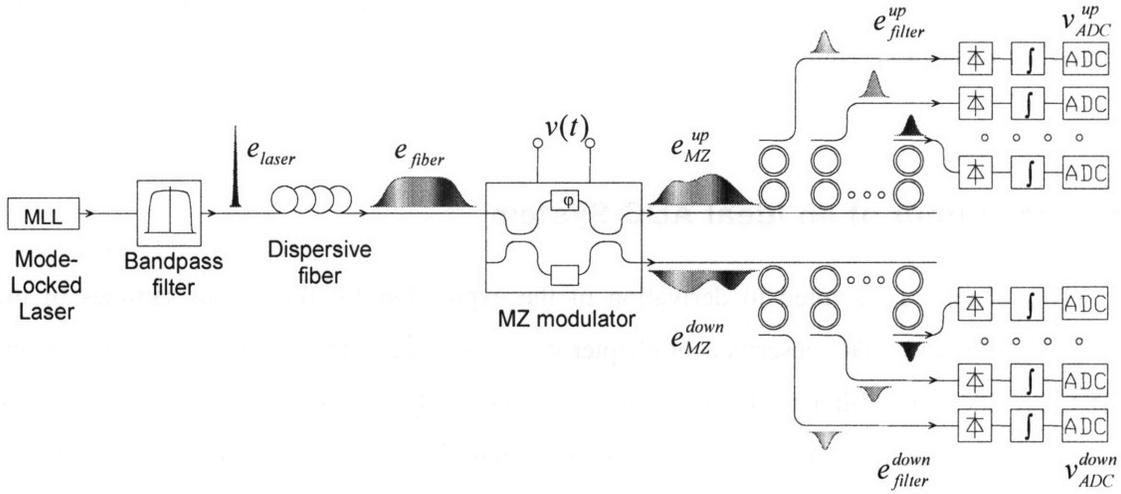
Figure 7. Layout of the optically-sampled ADC system with notation for signals at different conversion stages.

The pulse after the dispersive fiber is given by

$$E_{fiber}(\omega) = E_{laser}(\omega)\exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2\right],\tag{5.1}$$

where L is the length of the fiber and β_2 is its second-order dispersion coefficient. We assume that the fiber has no dispersion slope. The frequency ω_{ref} is the reference frequency at which the group delay is assumed to be zero.

The electrical signal being sampled $v(t)$ is applied to the RF input of the MZ modulator. We assume that the modulator biased at quadrature so that the differential phase shift between two arms is π. We also assume that the modulator is driven in push-pull configuration, i.e. the phase shifts in its upper and lower arms are

$$\phi_{MZ}^{up}(t) = \frac{\pi}{4} + v(t),$$

$$\phi_{MZ}^{down}(t) = -\frac{\pi}{4} - v(t).$$

These expressions assume that the phase shifts changes linearly with the applied voltage; the proportionality factor between them is omitted because it is not essential in the following derivations. Note that push-pull configuration of the modulator is necessary to avoid phase

modulation of the output optical signal (the output phase is $(\phi_{MZ}^{up} + \phi_{MZ}^{down})/2$, which equals zero for push-pull configuration).

At the output of MZ modulator, the phase modulation is transformed into amplitude modulation so that the output optical amplitude is

$$e_{MZ}^{up}(t) = e_{fiber}(t) \cdot \sin\left(\frac{\pi}{4} + v(t)\right),$$ (5.2)

$$e_{MZ}^{down}(t) = e_{fiber}(t) \cdot \cos\left(\frac{\pi}{4} + v(t)\right).$$ (5.3)

Introducing

$$v_{MZ}^{up}(t) = \sin\left(\frac{\pi}{4} + v(t)\right),$$ (5.4)

$$v_{MZ}^{down}(t) = \cos\left(\frac{\pi}{4} + v(t)\right),$$ (5.5)

we can rewrite Eqs. 5.2, 5.3 as

$$e_{MZ}^{up}(t) = e_{fiber}(t) \cdot v_{MZ}^{up}(t),$$ (5.6)

$$e_{MZ}^{down}(t) = e_{fiber}(t) \cdot v_{MZ}^{down}(t).$$ (5.7)

The derivation will be continued for the upper output of the ADC; the analysis for the lower output is identical.

Multiplication in time domain in Eq. (5.6) corresponds to convolution in frequency domain:

$$E_{MZ}^{up}(\omega) = \int E_{fiber}(\omega - \omega') V_{MZ}^{up}(\omega') d\omega' =$$

$$= \int E_{laser}(\omega - \omega') \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega' - \omega_{ref})^2\right] V_{MZ}^{up}(\omega') d\omega'.$$ (5.8)

If the bandwidth of the modulating signal $V_{MZ}^{up}(\omega)$ is much narrower than the laser spectrum $E_{laser}(\omega)$, we can make an approximation that the laser spectrum is constant within the bandwidth of the modulating signal so that we take it out from integration in Eq. 5.8

$$E_{MZ}^{up}(\omega) \approx E_{laser}(\omega) \int \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega' - \omega_{ref})^2\right] V_{MZ}^{up}(\omega') d\omega'.$$ (5.9)

For the EPIC ADC system (see page 22), the laser bandwidth is more than 1.6THz, while the RF signal bandwidth is limited to 20 GHz, so this approximation should be valid with good precision. Note that the bandwidth of $V_{MZ}^{up}(\omega)$ is actually larger than the bandwidth of the RF signal $V_{MZ}^{up}(\omega)$ because the former signal contains harmonic multiple harmonics of 20 GHz according to Eqs. (5.4), (5.5), especially if the modulation index is large. Another requirement for validity of this approximation is that $E_{laser}(\omega)$ is a smooth function of frequency, i.e. it doesn't have rapid oscillations within the bandwidth of $V_{MZ}^{up}(\omega)$.

The approximate equation (5.9) can be transformed as

$$E_{MZ}^{up}(\omega) \approx E_{laser}(\omega) \int V_{MZ}^{up}(\omega') \exp\left[i\frac{\beta_2 L}{2}(\omega^2 + \omega'^2 + \omega_{ref}^2 - 2\omega\omega' - 2\omega\omega_{ref} + 2\omega'\omega_{ref}) \right] d\omega'$$

$$= E_{laser}(\omega) \exp\left[i\frac{\beta_2 L}{2}(\omega^2 + \omega_{ref}^2 - 2\omega\omega_{ref}) \right] \cdot \int V_{MZ}^{up}(\omega') \exp\left[i\frac{\beta_2 L}{2}(\omega'^2 - 2\omega\omega' + 2\omega'\omega_{ref}) \right] d\omega' =$$

$$= E_{laser}(\omega) \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2 \right] \cdot \int V_{MZ}^{up}(\omega') \exp\left[i\frac{\beta_2 L}{2}\omega'^2 \right] \exp[i\beta_2 L(\omega_{ref} - \omega)\omega'] d\omega'.$$

Let us denote

$$V_{eff}^{up}(\omega) \equiv V_{MZ}^{up}(\omega) \exp\left[i\frac{\beta_2 L}{2}\omega^2 \right]. \qquad (5.10)$$

This "effective" voltage is the dispersed version of the signals defined by Eqs. (5.4), (5.5); as we see, the dispersion is transferred from the chirped optical pulses to the RF signal. With this definition, we can rewrite the spectrum at the output of MZ modulator as

$$E_{MZ}^{up}(\omega) \approx E_{laser}(\omega) \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2 \right] \cdot \int V_{eff}^{up}(\omega') \exp[i\beta_2 L(\omega_{ref} - \omega)\omega'] d\omega'.$$

The integral in the last equation is the inverse Fourier transform of $V_{eff}^{up}(\omega)$, so

$$E_{MZ}^{up}(\omega) \approx E_{laser}(\omega) \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2 \right] \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref}) \right).$$

Note that in this equation v_{eff} is the function defined in time domain into which we substitute $\beta_2 L(\omega - \omega_{ref})$ instead of time. Finally, using Eq. (5.1) we get

$$E_{MZ}^{up}(\omega) \approx E_{fiber}(\omega) \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref}) \right).$$

We see that the effective time-dependent voltage $v_{eff}^{up}(t)$ is imprinted onto the spectrum of the signal. Also note that according to this equation, the spectrum is modulated not by the original RF signal $v(t)$, but by the dispersed sine or cosine of this signal, $v_{eff}^{up}(t)$. The equations up to this point were previously obtained by Jonathan Birge [24].

The signal then passes through the filter bank which demultiplexes it into separate channels. Let us denote the amplitude transmission of n-th filter as $T(\omega - \omega_n)$, where ω_n is the filter center frequency; the function $T(\omega)$ is centered around zero. The output of n-th filter is then

$$E_{filter,n}^{up}(\omega) \approx E_{MZ}^{up}(\omega) \cdot T(\omega - \omega_n)$$

$$= E_{laser}(\omega) \cdot T(\omega - \omega_n) \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2 \right] \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref})\right).$$

The filter transfer function $T(\omega - \omega_n)$ should be much narrower than $E_{laser}(\omega)$, so we can approximately replace $E_{laser}(\omega)$ by a its value at ω_n

$$E_{filter,n}^{up}(\omega) \approx E_{laser}(\omega_n) \cdot T(\omega - \omega_n) \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref})\right) \cdot \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2 \right]. \qquad (5.11)$$

Each channel is then detected by a photodetector and integrated by an integrator with integrates the detected current within the pulse repetition period T_R. The output of the integrator is quantized by an electronic ADC to obtain one sample; the samples obtained from different channels must be interleaved to finally obtain the output of the upper branch of our photonic ADC system, $v_{ADC}^{up}(t_n)$.

The signal at the output of the integrator for one time slot T_R can be written as

$$v_{ADC}^{up}(t_n) = R_d \int_0^{T_R} \left| e_{filter,n}^{up}(t) \right|^2 dt, \qquad (5.12)$$

where $e_{filter,n}^{up}(t)$ is the time-dependent optical amplitude of the n-th channel and R_d is the photodetector responsivity. The time t_n is the time within the time slot T_R corresponding to n-th channel; it is related to the filter center frequency as $t_n = \beta_2 L(\omega_n - \omega_{ref})$. The system parameters must be chosen in such a way that t_n of different channels fill the time slot T_R uniformly. Note that actually the photodetector response is frequency dependent; we do not

include this into the analysis because this frequency dependence can be compensated at the post-processing step.

If the photodetector is not fast enough, the tails of its response function can extend into the next time slot, leading to crosstalk between adjacent pulses. In the analysis of this chapter we assume that this does not happen, i.e. the photodiode current excited by one pulse decays to zero when the next pulse arrives. For EPIC ADC system with 2 GHz repetition rate, the photodiode impulse response must decay to negligible value within the pulse repetition period $T_R = 500\text{fs}$. The integration limits in Eq. (5.12) can then be extended to infinity, (implying integration over only one pulse):

$$v_{ADC}^{up}(t_n) = R_d \int_{-\infty}^{\infty} \left| e_{filter,n}^{up}(t) \right|^2 dt \ .$$

Using Parseval's relation,

$$\int_{-\infty}^{\infty} \left| x(t) \right|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| X(\omega) \right|^2 d\omega \ ,$$

we can replace integration over time with integration over frequency

$$v_{ADC}^{up}(t_n) = \frac{R_d}{2\pi} \int_{-\infty}^{\infty} \left| E_{filter,n}^{up}(\omega) \right|^2 d\omega \ .$$

Substituting signal spectrum (5.11) into this expression we get

$$v_{ADC}^{up}(t_n) \approx \int \left| E_{laser}(\omega_n) \cdot T(\omega - \omega_n) \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref})\right) \cdot \exp\left[i\frac{\beta_2 L}{2}(\omega - \omega_{ref})^2 \right] \right|^2 d\omega$$

$$= \left| E_{laser}(\omega_n) \right|^2 \int \left| T(\omega - \omega_n) \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref})\right) \right|^2 d\omega \ . \tag{5.13}$$

Let us now substitute ω_n with time t_n according to the time-frequency mapping

$$\omega_n = \omega_{ref} + \frac{t_n}{\beta_2 L} \ .$$

Then (5.13) can be rewritten as

$$v_{ADC}^{up}(t_n) \approx \left| E_{laser}(\omega_n) \right|^2 \cdot \int \left| T\left(\omega - \omega_{ref} - \frac{t_n}{\beta_2 L}\right) \cdot v_{eff}^{up}\left(\beta_2 L(\omega - \omega_{ref})\right) \right|^2 d\omega \tag{5.14}$$

Introducing a new variable t' defined as

$$t' \equiv \beta_2 L(\omega - \omega_{ref})$$

so that

$$\omega - \omega_{ref} = \frac{t'}{\beta_2 L}, \qquad d\omega = \frac{dt'}{\beta_2 L},$$

we can rewrite (5.14) as

$$v_{ADC}^{up}(t_n) \approx \left|E_{laser}(\omega_n)\right|^2 \cdot \int \left|T\left(\frac{t'}{\beta_2 L} - \frac{t_n}{\beta_2 L}\right) \cdot v_{eff}^{up}(t')\right|^2 dt'. \qquad (5.15)$$

Let us denote the laser envelope spectrum density at frequency ω_n corresponding to t_n as

$$e_{laser}^{\frown}(t_n) \equiv \left|E_{laser}\left(\omega_n = \omega_{ref} + \frac{t_n}{\beta_2 L}\right)\right|.$$

The superscript $\frown$ symbolically represents "envelope" and is added for distinction from the laser amplitude at time moment t_n. Eq. (5.15) then becomes

$$v_{ADC}^{up}(t_n) \approx \left|e_{laser}^{\frown}(t_n)\right|^2 \cdot \int \left|v_{eff}^{up}(t')\right|^2 \cdot \left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'. \qquad (5.16)$$

Similar expression is valid for the samples obtained from the lower ADC branch:

$$v_{ADC}^{down}(t_n) \approx \left|e_{laser}^{\frown}(t_n)\right|^2 \cdot \int \left|v_{eff}^{down}(t')\right|^2 \cdot \left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'. \qquad (5.17)$$

Let us recall that these expressions are approximate because we assumed that the laser spectrum envelope is constant within the spectrum of the RF signal and within the passband of the optical filter.

We see that even if all components of the ADC are ideal and noise is absent, the obtained samples are not equal to the samples of the original signal $v(t_n)$. The next chapter continues the analysis of an ideal ADC system by studying the obtained equations and analyzing the accuracy of the analog-to-digital conversion.

6. Accuracy of an Ideal ADC System

1. Analysis of expressions for ADC output

In the previous chapter we derived expressions (5.16), (5.17) for the output of the upper and lower branches of the optically-sampled ADC. We will now analyze what these expressions mean and what is the accuracy that can be obtained with this ADC system. Similar to the previous chapter, we assume here an ideal ADC system, i.e. ADC system free of errors in its parameters. We will restrict our discussion to the ADC systems with practical parameters, specifically, with limited optical power and bandwidth of the mode locked laser, so that power and spectral efficiency are important.

The output of the ADC output is related to the applied RF signal according to Eqs. (5.16), (5.17) with transformations illustrated in Fig. 8. We will consider the upper branch of the system; the analysis for the lower branch is similar. First the sine function is applied to the original RF signal according to Eq. (5.4) due to nonlinearity of the MZ modulator. Then the obtained signal is dispersed according to (5.10). The dispersed signal is then convolved with the filter transfer function according to (5.16) and multiplied by the laser envelope according to the same equation.
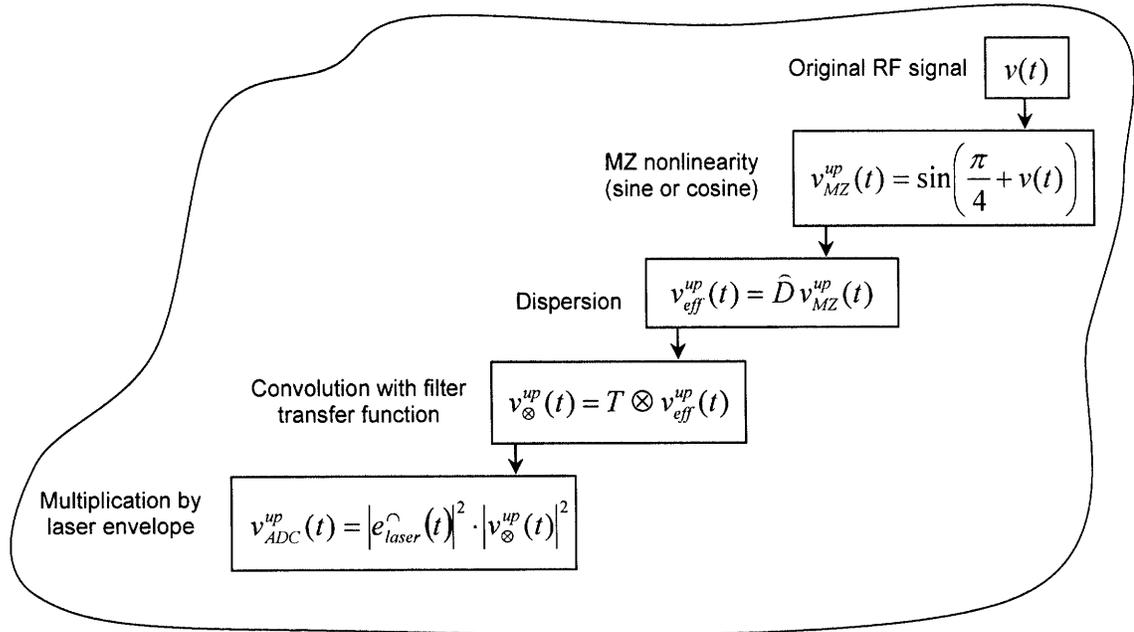
Figure 8. Signal transformations in the optically-sampled ADC.

Each of the transformations described above contribute to the difference between the original RF signal $v(t)$ and the ADC output $v_{ADC}^{up}(t)$, thus reducing the effective number of bits (ENOB) produced by this ADC. The modulator nonlinearity can be mitigated according to the linearization scheme according to Eq. (4.4); however, we will now show that this cannot compensate the nonlinearity precisely. Substitution of ADC outputs (5.16), (5.17) into the linearization equation (4.4) gives

$$
v_{ADC}(t_n) = \frac{1}{2}\mathrm{asin}\left(\frac{v_{ADC}^{up}(t_n) - v_{ADC}^{down}(t_n)}{v_{ADC}^{up}(t_n) + v_{ADC}^{down}(t_n)}\right) = \frac{1}{2}\mathrm{asin}\left(\frac{\int\left[\left|v_{eff}^{up}(t')\right|^2 - \left|v_{eff}^{down}(t')\right|^2\right]\cdot\left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'}{\int\left[\left|v_{eff}^{up}(t')\right|^2 + \left|v_{eff}^{down}(t')\right|^2\right]\cdot\left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'}\right) \quad (6.1)
$$

We see that the laser envelope indeed cancels in this expression, so according to this equation the ADC output will be independent on pulse-to-pulse intensity variations and the exact laser spectrum. However, let us recall that this equation was obtained in the approximation that the laser envelope is constant within the signal bandwidth, so in reality non-flatness of the laser envelope lead to ENOB degradation.

The previous equation tells us that the MZ linearization does not work precisely because of the effect of dispersion and convolution with the filter shape. Because of the impact of dispersion, the numerator of Eq. (6.1) does not contain $\sin^2(\phi) - \cos^2(\phi)$ and thus the output will not be proportional to $\sin(2\phi)$; the denominator does not contain $\sin^2(\phi) + \cos^2(\phi)$ and thus does not give us the total optical power. Let us prove that the linearization will be precise only if the dispersion and convolution are neglected.

If the dispersion impact of RF signal is negligible, we have

$$
v_{eff}^{up}(t) = v_{MZ}^{up}(t) = \sin\left(\frac{\pi}{4} + v(t)\right), \quad v_{eff}^{down}(t) = v_{MZ}^{down}(t) = \cos\left(\frac{\pi}{4} + v(t)\right).
$$

The ADC output is then

$$
v_{ADC}(t_n) = \frac{1}{2}\mathrm{asin}\left(\frac{\int\left[\sin^2\left(\frac{\pi}{4} + v(t')\right) - \cos^2\left(\frac{\pi}{4} + v(t')\right)\right]\cdot\left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'}{\int\left[\sin^2\left(\frac{\pi}{4} + v(t')\right) + \cos^2\left(\frac{\pi}{4} + v(t')\right)\right]\cdot\left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'}\right),
$$

$$v_{ADC}(t_n) = \frac{1}{2}\operatorname{asin}\left(\frac{\int \sin(2v(t')) \cdot \left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'}{\int \left|T\left(\frac{t'-t_n}{\beta_2 L}\right)\right|^2 dt'}\right).$$

If the filter is very narrow so that it has transfer function in frequency domain close to the delta-function, the integrals in the last expression disappear and we finally get

$$v_{ADC}(t_n) = \frac{1}{2}\operatorname{asin}\bigl(\sin(2v(t_n))\bigr) = v(t_n),$$

i.e. the linearization will work perfectly and ADC will be sampling the RF signal with no error.

Let us now analyze in more details all factors which limit the ENOB of the ideal photonic ADC.

2. The impact of dispersion

As shown above, one of the reasons for ENOB degradation is the dispersion transferred from the chirped optical pulses to the RF signal. The accumulated dispersion is defined by the product $\beta_2 L$, which is related to the channel spacing $\Delta\omega$ and sampling rate f_s according to Eq. (4.2):

$$\beta_2 L \Delta\omega = 1/f_s. \tag{6.2}$$

It follows from this equation that if we want to decrease the impact of dispersion in an ADC with given sample rate, we need to increase the channel spacing $\Delta\omega$ and thus the total laser bandwidth $\Delta\omega\, N$. The bandwidth of the mode-locked laser is limited, so the dispersion impact cannot be made negligibly small in a practical wideband ADC system.

The signal degradation is expected to be proportional not only to the accumulated dispersion, but also to the square of the RF signal bandwidth (because dispersion length is inversely proportional to the signal bandwidth). If the signal bandwidth is limited to Nyquist frequency $f_{max} = f_s/2$, the signal degradation is expected to be proportional to $\beta_2 L f_s^2 \sim f_s / \Delta\omega$. This means that ENOB will decrease if we are trying to build an ADC with high sampling rate with densely packed channels. This is illustrated in Fig. 9, which shows ENOB as a function of channel spacing for several sampling rates. The ENOB was obtained

from ADC system simulation; to make sure that ENOB was reduced only by dispersion, all other factors were made negligibly small: (a) very small modulation index was used to avoid MZ nonlinearity, (b) very narrow filter to avoid distortion due to convolution, (c) the laser was assumed to have very broadband so that its spectrum was flat over the RF bandwidth. The method used for calculation of ENOB of the ADC system is described in the Appendix. Let us underline that it is implied in Fig. 9 that the signal frequency is changing together with the sampling rate so that $f_{max} = f_s / 2$. Note that the impact of dispersion will not depend on the number of channels N or repetition rate f_R separately, but rather on the aggregate sampling rate $f_s = N f_R$.
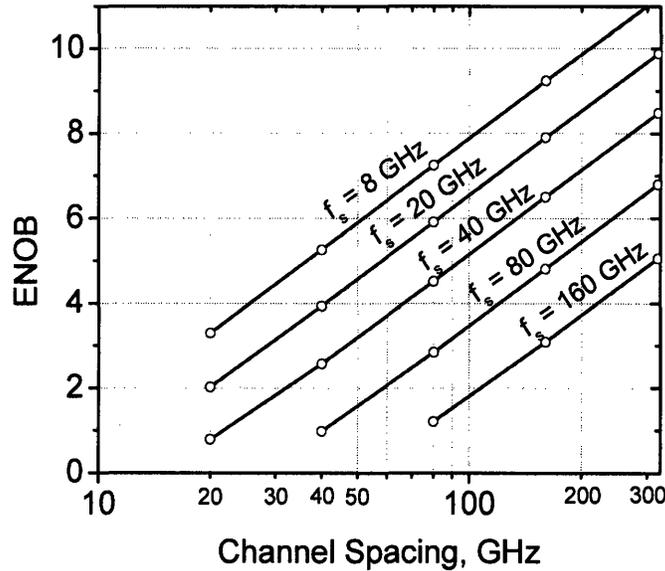
Figure 9. Dispersion-limited ENOB as a function of frequency spacing between the channels for different sampling rates.

3. The impact of convolution with the filter transfer function

The accuracy of an ideal ADC is also limited because the ADC output is the convolution of the signal with the filter transfer function. Qualitatively speaking, the optical filter selects not a single frequency component but a range of frequencies, so the values of the signal $v(t)$ at time moments corresponding to all these frequencies will contribute to the sample produced by the ADC $v_{ADC}(t)$. The narrower the optical filter, the smaller is the ENOB degradation; however, narrowing down the optical filter reduces the power efficiency of the

system because only a small fraction of the input optical power is used. The power at the photodetector must be sufficiently large to avoid ENOB degradation due to photodetector noise, so narrowing down the optical filter means that the optical power at the input of the ADC system must be increased. Having high input power is not only impractical but is not always possible, because, for example, in silicon waveguides the maximum optical power is limited by two-photon absorption. Therefore, in a practical power-efficient ADC system the filter bandwidth will not be small and the corresponding ENOB degradation will not be negligible.

Fig. 10 shows the calculated ENOB as a function of the spectral fill factor, which is the ratio of the optical filter FWMH $\Delta\omega_{FWHM}$ to channel spacing $\Delta\omega$. Different curves in this figure correspond to different sampling rates of the ADC (assuming that the RF spectrum is changed accordingly). To make sure that dispersion does not contribute to ENOB degradation, the channel spacing $\Delta\omega$ was made very large (640GHz). We see that ENOB depends strongly on the optical filter bandwidth; as we try to improve power efficiency or scale the ADC to a larger number of channels by increasing the spectral fill factor, the ENOB will be reduced. We see that the ENOB does not depend on the sampling rate of the ADC, which can be understood from the following considerations. Because of the time-frequency mapping, the filter bandwidth defines the "convolution time interval" around the sampling point which contribute to the given sample in our ADC; the larger is this interval, the larger is the error. If the sampling rate is increased, the given convolution time interval will produce larger error because there's more signal variation within it; however, the length of this interval is inversely proportional to the sampling rate, therefore the net error is independent on the sampling rate. Some difference in ENOB for small fill factors in Fig. 10 was due to the contribution of dispersion, which becomes noticeable when the impact of convolution becomes very small. Small difference in ENOB for large fill factors is due to statistical nature of the simulation.
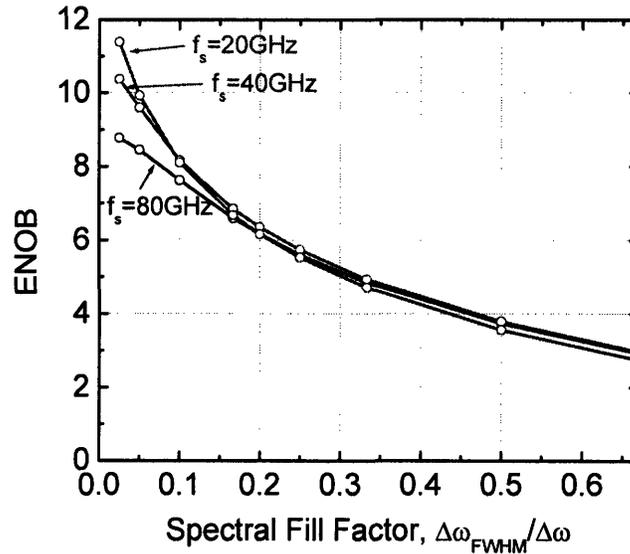
Figure 10. ENOB limited by convolution with the optical filter transfer function as a function of spectral fill factor for several ADC sampling rates.

4. Variation of laser spectrum within RF bandwidth

ADC output sampled described by Eq. (5.9) is obtained in approximation that the laser spectrum stays constant within the RF bandwidth. Therefore, variation of the laser spectrum on the scale of RF bandwidth will reduce ENOB. In time domain non-flatness of the laser spectrum corresponds to the envelope of the dispersed signal changing at rate comparable to the rate of change of RF signal; the impact of changing optical envelope is not distinguishable from the impact of RF signal varying in time, so the measured signal will be distorted and ENOB will be degraded. This distortion is the weaker the flatter is the optical spectrum. However, it's hard to produce optical spectrum which is flat within the filter bank bandwidth; moreover, even if we produced such a spectrum, all its power except the center part will not be used, so such a system will not be power-efficient. Therefore, in a practical system the spectrum will not be constant over RF bandwidth and ENOB will be reduced.

The degree of signal degradation due to non-flat optical spectrum will depend on the width of the filter bank bandwidth relative to the laser optical bandwidth, as well as on the RF bandwidth relative to the laser bandwidth. Therefore, we should expect that ENOB will decrease when the channel spacing increases or when the ADC sample rate increases. Fig. 11 shows calculated ENOB as a function of channel spacing. Different curves correspond to different pulse durations, and therefore to different values of optical bandwidth, assuming

transform-limited Gaussian pulses. Different plots correspond to different number of channels with laser repetition rate being 2GHz in all cases. We see that ENOB can be severely limited of the optical bandwidth of the laser is insufficient to accommodate all channels. In Fig. 11, ENOB is also limited by dispersion and convolution, especially for small values of channel spacing. The curve with the highest ENOB correspond to delta pulse and perfectly flat spectrum; the difference from this curve is the penalty due to non-flatness of the optical spectrum. We see that for ADC systems with small channel spacing and sufficient optical bandwidth ENOB is mostly limited other factors rather than non-flat optical spectrum. Therefore the approximation made in derivation of Eqs. (5.16, 5.17) is valid for practical ADC systems. However, it is important to keep in mind that ENOB can be severely reduced if the optical bandwidth is not adequate. The outermost samples at the edges of the spectrum will have large errors because the spectrum there varies very quickly. If the detection noise is taken into account, the error for the outermost samples will be even larger because the corresponding channels will have low optical power and will thus be degraded by thermal and shot noise.
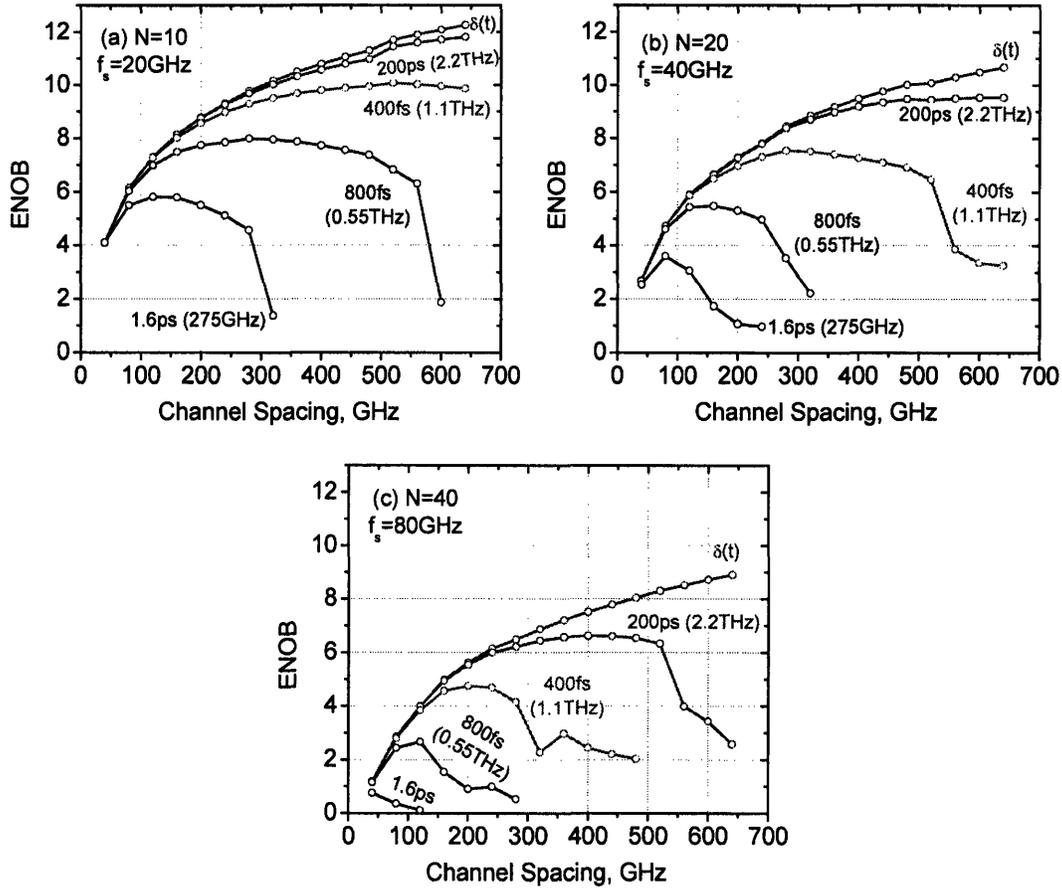
Figure 11. ENOB as a function of channel spacing. For large values of channel spacing ENOB is limited mostly by non-flat optical spectrum, for small channel spacing – by dispersion and convolution. The curve parameter is the FWHM duration of a Gaussian pulses generated by the mode-locked laser; the corresponding FWHM bandwidth is given in brackets. Number of channels is (a) N=10, (b) N=20, (c) N=40; laser repetition rate is 2GHz.

5. Nonlinearity of MZ modulator transfer function

The ENOB will also be reduced due to nonlinearity of the MZ modulator which cannot be compensated exactly because of dispersion and convolution with the filter shape, as it was explained in the analysis above. This distortion will be smaller if the modulation index is reduced, however, a practical ADC system should have a large modulation index because it requires less optical power so it is more power efficient. The reduced optical power at photodetectors also helps to avoid nonlinearity of the photodetector response which takes place at several mW of optical power for a typical photodetector [19]. Therefore, the

modulation index of the MZ modulator should be high for a practical system, so the question about the impact of uncompensated nonlinearity needs to be studies.

The impact of MZ nonlinearity which is uncompensated due to dispersion and convolution is shown in Figs. 12, 13 respectively. For comparison, the ENOB produced by an ADC without linearization is shown on the right. We can see that the linearization really helps to improve accuracy for high modulation indices[1]. The amount of uncompensated nonlinearity turns out to be relatively small so that ENOB penalty does not exceed 1 bit for modulation indices up to 0.95. The ENOB penalty due to uncompensated nonlinearity is smaller than due to dispersion and convolution, so uncompensated nonlinearity turns out to be a secondary factor when the MZ linearization is used.
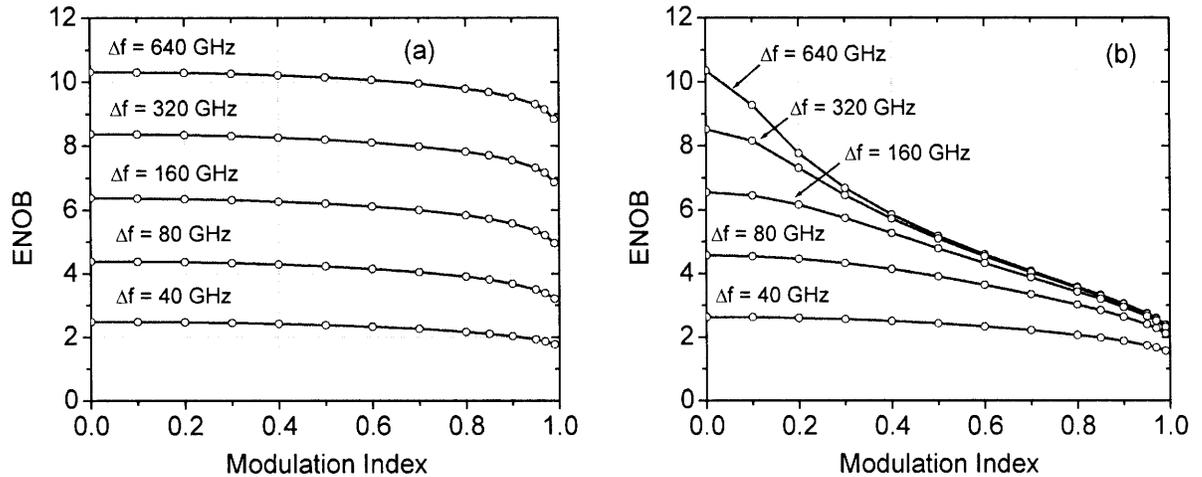
Figure 12. ENOB versus the modulation index as limited by dispersion with (a) linearization scheme applied, and (b) without linearization. Different curves correspond to different values of channel spacing Δf.

---

[1] The modulation index of the MZ modulator is defined as the maximum variation of the output power relative to the average level, $m = (P_{max} - P_0)/P_0$, where P_{max} is the maximum output power and P_0 is the average output power. The average power in our ADC system corresponds to zero driving voltage, $P_0 \sim \sin^2(\pi/4) = 0.5$. If φ_{max} is the maximum phase shift in single arm of MZ modulator driven in push-pull configuration, the maximum output power is $P_{max} \sim \sin^2(\pi/4 + \varphi_{max})$. The modulation index is then $m = \sin(2\varphi_{max})$. If v_{max} is the maximum RS signal, the modulation index is $m = \sin(2\pi v/v_\pi)$, where v_π is the voltage producing π phase shift.
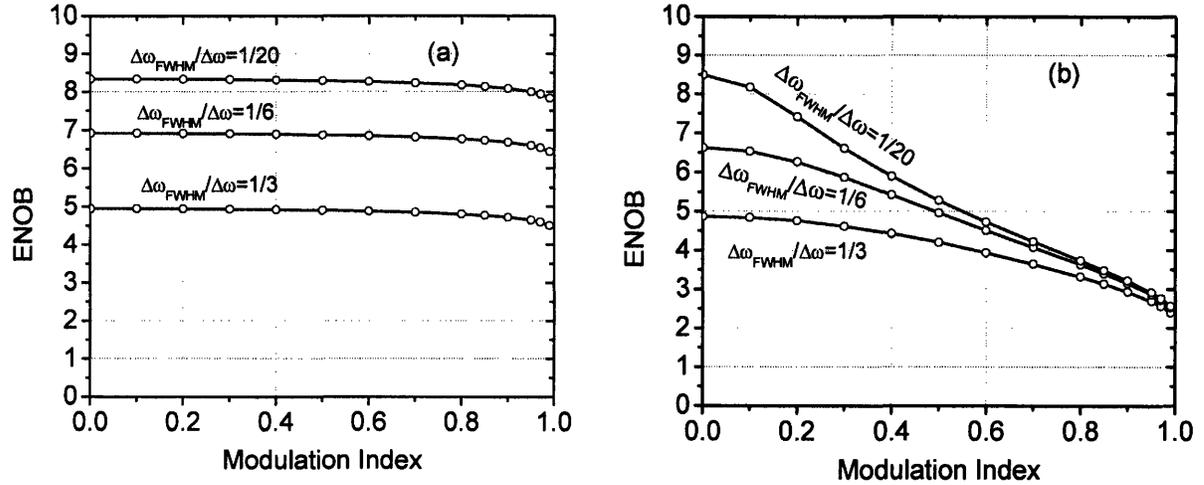
Figure 13. ENOB versus modulation index due to convolution with (a) linearization scheme applied, and (b) without linearization.

6. Difficulties with straightforward signal recovery

Is there is a straightforward way to recover the original signal from the ADC output? According to Eqs. (5.16) (5.17), the ADC output is connected to the original RF signal by sine or cosine, dispersion and convolution with the filter shape. Is it possible to recover the original signal by deconvolving the ADC output with the filter transfer function, compensating the dispersion and then inverting the sine or cosine? It turns out that there are two things that prevent this algorithm from working.

• Deconvolution can in principle be implemented in frequency domain as division of the ADC output spectrum by the filter transmission function spectrum. This, however, requires the knowledge of the spectrum of the ADC signal at the output of the ADC. If the signal we are sampling is bandlimited to some maximum frequency f_{max}, the modulator nonlinearity will create multiple harmonics of this frequency, for instance harmonics at $3f_{max}$, $5f_{max}$, etc. The frequency of all these harmonics need to be smaller than the Nyquist rate of the ADC to avoid aliasing during deconvolution in frequency domain. Thus the deconvolution will be precise only for signals with f_{max} several times smaller than the Nyquist rate. However, we are interested in building a wideband ADC system which will produce high ENOB for signals with frequencies up to its Nyquist rate. We see that the deconvolution for such a wideband system is not an option.

For example, consider a 20-channel ADC system with 2GHz laser repetition rate. Its sampling rate is 40GHz and Nyquist rate is 20GHz. If we are sampling an RF signal with $f_{max} = 20GHz$, the MZ modulator will create harmonics at $3f_{max} = 60GHz$, $5f_{max} = 100GHz$, etc. The ADC is not sampling fast enough to capture these harmonics, so we cannot determine the spectrum of $v_{ADC}(t)$ accurately because of aliasing and deconvolution will not work. If RF signal is bandlimited to $f_{max} = 4GHz$, the ADC will capture its 5th harmonics and deconvolution will be reasonably accurate; however, large reduction of the nominal ADC bandwidth defies our goal of building a wideband ADC and must be avoided.

• The dispersion compensation for this ADC system also turns out to be problematic. If we try to do it by inverting the (5.10),

$$V_{MZ}^{up}(\omega) = V_{eff}^{up}(\omega) \exp\left[-i \frac{\beta_2 L}{2} \omega^2 \right],$$ (6.3)

we again run into the problem that the spectrum of the ADC output $V_{eff}^{up}(\omega)$ is not known for a wideband ADC, as it was just explained. Moreover, $V_{eff}^{up}(\omega)$ cannot be found because phase of the dispersed signal $v_{eff}^{up}(t)$ is lost during detection. The function $v_{eff}(t)$ is complex for the following reason. The function $v_{MZ}(t)$ is real, so the phase of its spectrum is odd, $V_{MZ}(\omega) = V_{MZ}^*(-\omega)$. After quadratic phase is added to it due to dispersion according to Eq. (5.10), the phase is no longer odd, so $v_{eff}(t)$ is complex. Measurements gives us only the absolute value of $v_{eff}(t)$, so the phase of $v_{eff}(t)$ is lost, its spectrum cannot be found, and the dispersion cannot be compensated according to (6.3).

The deleterious impact of dispersion can be avoided if the ADC system is modified in such a way that it does not use dispersive fiber. The delays between different sub-pulses of the pulse train can be created by replacing the fiber with frequency demultiplexer, delay lines introducing necessary delays for each channel, and then multiplexing the pulses into one signal which then goes to the input of the modulator. Such a scheme was proposed in [10] and demonstrated in [11] and [25]. The benefit of this architecture is the absence of signal distortion due to dispersion; if the system is built of discrete components, the ADC system can be modified to work with laser with any repetition rate because the delays in the delay lines can be adjusted. The drawback of this approach seems to be the increased complexity of the system and potential problems with stability of the delay lines if they are implemented e.g.

as fibers of different lengths. It also not straightforward to implement this concept when the ADC is built on a single chip because of the large length of required delay lines; for example, ~2.5cm of waveguide with refractive index = 3 is required to delay a pulse by 250ps.

7. Conclusion

The accuracy of the optically-sampled ADC is limited even if all components of the system are ideal. The main factors reducing system ENOB is dispersion and convolution with optical filter shape, which become the more severe the more densely the channels are packed in frequency domain and the more efficient is the spectrum usage, and the higher is sampling rate. For the EPIC ADC system the accuracy is limited by less than 4 bits. It is necessary to keep in mind, however, that in the analysis of this section did not include the effect of detector noise. For example, ENOB goes down in Fig. 10 as the spectral fill factor is increased; in reality, ENOB might go up because larger fill factor means larger optical power at the photodetector and thus less noisy signal

The incomplete dispersion compensation also reduces the accuracy for high modulation indices, but this reduction is small. Non-flat optical spectrum reduces ENOB as well and can be an issue for insufficiently wide or oscillating laser spectra.

The direct signal reconstruction by backcalculation was shown to be problematic even for an ideal ADC system because of loss of phase information and spectral aliasing. A novel method of reconstruction of the original RF signal described in the next section can compensate the errors caused by dispersion and convolution, as well as errors caused by ADC components imperfections. This method allows to overcome limitations described in this chapter.

7. Iterative Error Compensation Algorithm

1. Description of the algorithm

In this chapter we present a novel approach which allows to compensate various ADC errors numerically at the post-processing stage provided that ADC parameters are properly characterized [26]. The problem we are trying to solve can be formulated as follows. We have a system which transforms the input signal (applied RF voltage) into the output (measured samples) in some nonlinear way. We know how to calculate the system's output

from its input, but not the other way around. Our system has the property that its output is *approximately* equal to its input, but the two are not equal exactly. Our goal is, given the system's output, find the system's input which produced this input as precisely as possible.

Let us denote the input signal as $v(t)$ and the measured signal at the output of the ADC as $v_{ADC}(t)$; the notation we use is illustrated in Fig. 14. Let us also suppose we have an approximate value of the input signal, $v^{(1)}(t)$. We can run the ADC simulation to find the output of the system $v_{ADC}^{(1)}(t)$. If our guess $v^{(1)}(t)$ is close to the real signal $v(t)$, the calculated output signal $v_{ADC}^{(1)}(t)$ will be close to the measured output $v_{ADC}(t)$. The idea of the error correction algorithm is that the difference between the measured and simulated outputs $v_{ADC}^{(1)}(t) - v_{ADC}(t)$ is approximately equal to the difference between the guess and true input signals,

$$v_{ADC}^{(1)}(t) - v_{ADC}(t) \approx v^{(1)}(t) - v(t) . \tag{7.1}$$

From this equation, the true input signal is approximately

$$v(t) \approx v^{(1)}(t) - \left[v_{ADC}^{(1)}(t) - v_{ADC}(t) \right].$$

We can use this as an improved guess for the input signal:

$$v^{(2)}(t) = v^{(1)}(t) - \left[v_{ADC}^{(1)}(t) - v_{ADC}(t) \right]. \tag{7.2}$$

We can improve this new guess using the same approach, i.e. at n-th step the guess $v^{(n)}(t)$ can be improved according to

$$v^{(n+1)}(t) = v^{(n)}(t) - \left[v_{ADC}^{(n)}(t) - v_{ADC}(t) \right]. \tag{7.3}$$

The iterations can be terminated when get a solution which remains steady from iteration to iteration with desired precision. The measured signal itself $v_{ADC}(t)$ can be used as an initial guess $v^{(1)}(t)$.

The described error compensation algorithm can also be explained in a more mathematical way. The relation between the input RF signal $v(t)$ and the output samples $v_{ADC}(t)$ can be written as:

$$v_{ADC}(t) = (1 + \hat{\varepsilon}) v(t) , \tag{7.4}$$

where $\hat{\varepsilon}$ is some nonlinear operator describing the distortion introduced by the system (the used notation is shown in Fig. 14). For a reasonably accurate ADC system, $\hat{\varepsilon}v(t)$ should be much smaller than $v(t)$. Let us suppose we have a guess for the input signal, $v^{(1)}(t)$,

$$v^{(1)}(t) = v(t) + \Delta v(t), \tag{7.5}$$

where $\Delta v(t)$ is the error which should be small if our guess is good. If we run the ADC simulation with this guess as the input, we get

$$v^{(1)}_{ADC}(t) = (1+\hat{\varepsilon})v^{(1)}(t) = (1+\hat{\varepsilon})\big(v(t)+\Delta v(t)\big) = v(t) + \hat{\varepsilon}v(t) + \Delta v(t) + \hat{\varepsilon}\Delta v(t). \tag{7.6}$$

The difference between the calculated and measured ADC outputs, (7.6) and (7.4) is then

$$v^{(1)}_{ADC}(t) - v_{ADC}(t) = \Delta v(t) + \hat{\varepsilon}\Delta v(t) \tag{7.7}$$

which is the error in our guess signal $\Delta v(t)$ plus a second-order term $\hat{\varepsilon}\Delta v(t)$, which should be small because both $\hat{\varepsilon}$ and $\Delta v(t)$ are small. We have thus proved that Eq. (7.1) holds up to the terms of second order of magnitude. Substituting this into Eq. (7.2) for iterative guess improvement we get a new guess

$$v^{(2)}(t) = v^{(1)}(t) - \hat{\varepsilon}\Delta v(t).$$

Comparing this to our initial guess (7.5) we see that the error in the new guess has been reduced from $\Delta v(t)$ to $\hat{\varepsilon}\Delta v(t)$. The error can be further reduced by running more iterations of the algorithm.

Real input signal $v(t)$	$\Longrightarrow$	ADC	$\Longrightarrow$	Measured output signal $v_{ADC}(t)$
Guess input signal $v^{(1)}(t)$	$\Longrightarrow$	ADC	$\Longrightarrow$	Calculated output signal $v^{(1)}_{ADC}(t)$
$\circ$ $\circ$ $\circ$ $\circ$ $\circ$				
Guess input signal $v^{(n)}(t)$	$\Longrightarrow$	ADC	$\Longrightarrow$	Calculated output signal $v^{(n)}_{ADC}(t)$
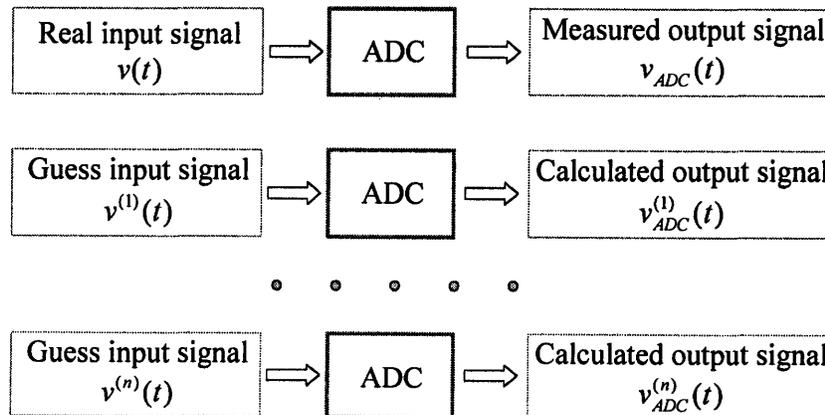
Figure 14. Notation used in description of the iterative error compensation algorithm.

The proposed algorithm was verified to be effective against all error sources described in chapters 6 and 9, such as dispersion, channel-to-channel loss variations, bias voltage error, crosstalk between time slots, and many others. The errors which cannot be compensated for are the random errors due to laser jitter and photodetection noise. For sufficiently low jitter and high optical power on the photodetectors, the ENOB achievable with the proposed algorithm is limited only by the precision with which the ADC components are characterized. Because some ADC parameters may drift in time, e.g. filter center frequency can drift with temperature, periodic calibration by sampling of appropriate test signals derived from the modelocked laser itself is needed (see chapter 8 for more details).

Note also that the proposed approach is generic and does not depend on a particular implementation of the ADC system. The algorithm is therefore expected to be effective for other photonic ADC configurations as well.

2. Example of error compensation

As an example we consider the EPIC ADC system (see page 22). For this system, the ENOB is limited to about 4 by dispersion and filter bandwidth even if all system components are perfect (see more details in chapter 6). As a test case for the error compensation algorithm, we also assume rather poor ADC components: photodetectors with 1 GHz bandwidth, leading to overlap between the detected pulses, filter banks with randomly spaced center frequencies with a variance of 15GHz, and an MZ modulator with 30/70 splitting ratio in the output coupler. The signal obtained directly from of this ADC have ENOB ~ 1. The error compensation algorithm allows to improve ENOB to 10 bits in 11 iterations. Fig. 15 illustrates how the error is reduced on initial iterations and Fig. 16 shows how the ENOB as a function of the iteration number. Note that if the ADC produces a better initial signal, it requires fewer iterations to achieve high ENOB; the iterations can also converge faster.
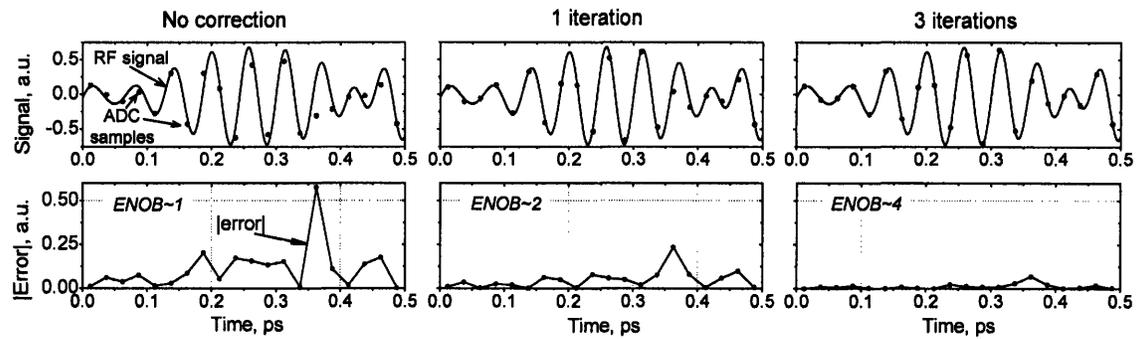
Figure 15. Example of error compensation by the iterative algorithm. In the upper plots, the curves correspond to the input RF input signal, and red points – to the samples obtained by the ADC. The difference between the two is the error, $v_{ADC}(t) - v(t)$, which is shown in the bottom plots.
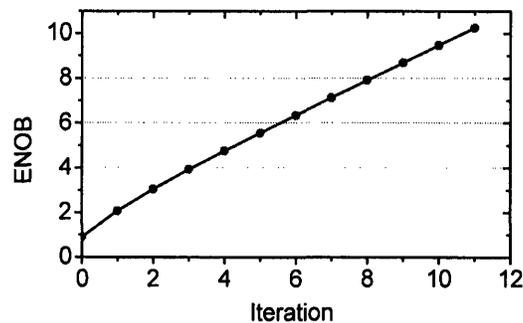
Figure 16. ENOB obtained with the error compensation algorithm as a function of the number of iterations.

3. Memory requirements and segment-by-segment simulation

The error compensation algorithm requires running simulation of the whole ADC system, which can be time and memory-consuming. If the ADC simulation is not implemented efficiently, time and memory requirements can be too high for practical applications of the system. Let us analyze this issue in more detail and describe ways to improve efficiency of the algorithm.

The first idea for improving computation efficiency is to run ADC simulation not for the whole signal at once, but split the signal into segments and run the simulation one segment at a time. The problem with running ADC simulation with the full signal is the following. The signal produces by the ADC can be very large; for example, EPIC ADC would produce $40 \cdot 10^6$ samples per millisecond. ADC simulation requires sampling rates which are much

higher than the ADC sampling rate. For the EPIC ADC, the sampling rate is 40 Gsamples/sec, but the optical signal at the input of the MZ modulator is about 2THz in bandwidth, so the sample rate of about 2 THz must be used in simulation, which is 50 times larger than the ADC sampling rate. Therefore, to run error compensation for a 1ms segment of ADC output, we need to work with about $2 \cdot 10^9$ samples. It takes a lot of memory just to store that much data; even more memory is needed to run Fourier transforms for this data. The idea for reducing the memory requirements is to divide the whole signal into multiple parts and run the simulation part-by-part, as it is pictorially illustrated in Fig. 17. For example, we can divide the 1 ms-log signal into $2 \cdot 10^6$ segments, each with duration equal to the laser repetition period $T_R = 500\,ps$. ADC simulation with 2THz bandwidth for each segment would then need just 1000 points. The amount of memory required is thus greatly reduced and is made independent on the duration of the time interval. The problem with high memory usage is therefore solved.
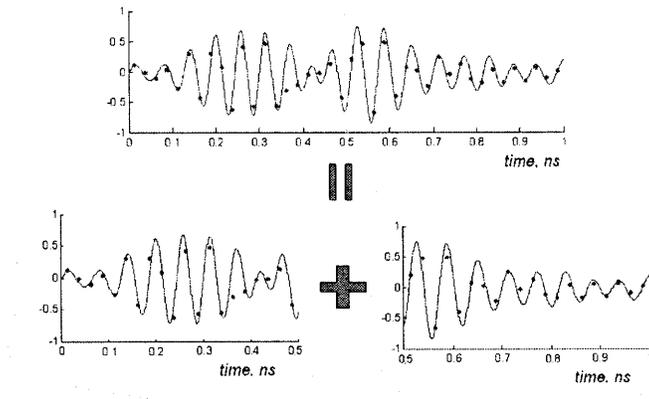
Figure 17. Illustration of a principle that running ADC simulation for whole signal (plot on the top) is equivalent to running simulations for segments of the signal separately (plots at the bottom) and then concatenating the obtained results.

There is one important issue with segment-by-segment data processing described above. Let us call the part of the signal with duration equal to repetition period as one time slot. The problem is that it is possible to have crosstalk between different time slots, i.e. the output in one time slot depends on the input not only in the same time slot but also in the neighboring time slots. This can happen, for example, when there is non-negligible crosstalk between optical filters, or when the photodiodes are not fast enough so that their response continues

into the next time slot. In this way running the simulation segment-by-segment and concatenating of the obtained results as shown in Fig. 17 will not give precise results because it will disregard the crosstalk between segments. The solution of this problem is to pad the output signal segment with additional slots on the left and on the right, so that the "spilling" the signal into neighboring time slots will be included into the simulation. For example, we can run the ADC simulation with segment length equal to one time slot. The input signal will therefore always be one slot long. For the output time slot, we can add two additional slots before and two slots after it, so that the output segment will be 5 slots long, as it is illustrated in Fig. 18. In presence of crosstalk, each input time slot will contribute to 5 output time slots. To obtain the output of the ADC, contributions of all input time slots into each output time slot must be summed up. For example, referring to Fig. 18, the signal at the output time slot 3 is obtained by summing contribution from input slots 1-5,

$$\text{out } 3 = (\text{in } 1 \rightarrow \text{out } 3) + (\text{in } 1 \rightarrow \text{out } 3) + (\text{in } 1 \rightarrow \text{out } 3) + (\text{in } 1 \rightarrow \text{out } 3) + (\text{in } 1 \rightarrow \text{out } 3)$$

In this way the ADC simulation can be run in a slot-by-slot fashion with low memory usage and the effect of crosstalk included. Depending on the strength of crosstalk, the number of padding slots can be increased or reduced. Note that an additional benefit of the segment-by-segment approach is that the simulation can be easily run on several processors in parallel, which can greatly increase the performance.
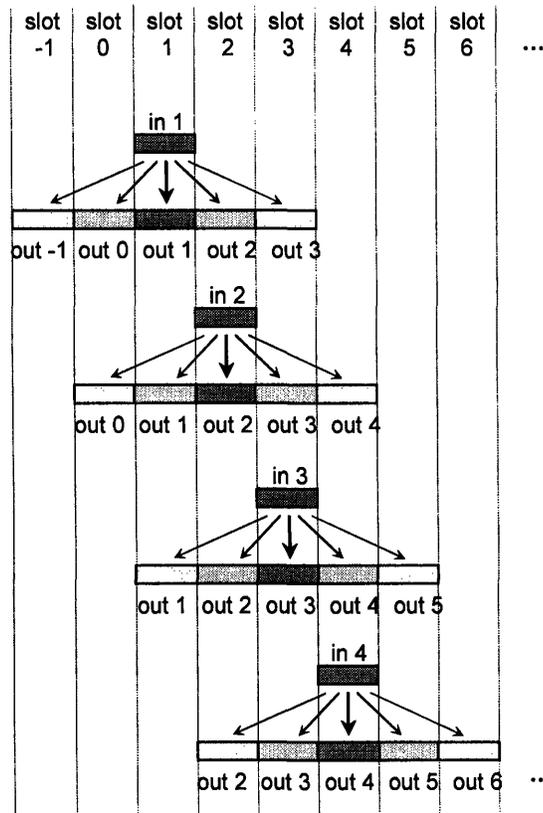
Figure 18. Running ADC simulation one time slot at a time. To capture the effect of crosstalk between time slots, the output slot is padded with two slots in the beginning and two slots in the end. The overall output is obtained by combining contributions of all input time slots into a given output time slot.

The price we have to pay for including the effect crosstalk into slot-by-slot simulation is the increased size of the output signal for each simulation and thus increased computation time. For example, if we pad each output slot with 5 slots before and 5 slots after, the output signal will be 11 slots long, so the computation time will increase by about an order of magnitude. To solve this problem, we can make the input signal segment longer, e.g. 100 slots long. The output segment will be 100 slots long, plus 5 slots of padding before and after it, giving in total 110 time slots. The computation time increase will be mere 10%, which can further be reduced by increasing the length if the input signal segment.

The final issue with the segment-by-segment simulation worth mentioning is the upsampling of the input RF signal. The upsampling is necessary because the RF sample rate must match the sample rate of the optical signal for doing calculations. The upsampling is commonly done by converting the signal into the frequency domain, padding the spectrum

with zeros, and converting it back to the time domain. However, we do not want to do upsampling in such a way for the whole RF signal at once, because it might require too much memory. The solution can be to do upsampling for one segment at a time. To avoid signal distortion at the ends of the segment, the RF signal segment being upsampled should include several additional time slots of the RF signal at the beginning and the end.

All of the above suggestions were implemented and were confirmed to be effective for reducing memory usage and simulation speed of the signal reconstruction algorithm.

4. Variable sample rate

One way to speed up the ADC simulation is to use variable sample rate at different stages along the ADC. The laser pulse train must be sampled at a high rate (several THz) because of its large bandwidth. However, at the output of the filter bank there is no need to keep same large sample rate because the bandwidth of the signal is comparable to the FWHM bandwidth of the optical filter; sample rate on the order of 100GHz should be sufficient. Furthermore, after the optical signal is detected by a photodetector, sample rate can be reduced once again because of limited bandwidth of the photodetector. Using variable sample rate in ADC simulation as described above allows to achieve about an order-of-magnitude performance improvement compared to straightforward implementation when the same large sample rate is used at all stages of the ADC simulation.

Writing an efficient ADC simulation code is still a challenging task even with optimization suggested above. For example, the author's code written in Matlab spends about 1-2 minutes on one iteration of the error compensation algorithm for a signal with $200 \cdot 10^3$ samples; the simulation is run on a desktop with a single-core 3GHz Pentium IV processor. A wideband ADC system producing data at a huge rate is likely to be connected to an equally powerful multi-core computer, so the simulation will run much faster, but writing an efficient code is still expected to be very important.

8. Calibration of the ADC System

1. Mach-Zehnder Modulator Characterization

The Mach-Zhehnder interferometer can have several kinds of imperfections leading to degradation ADC accuracy, as described in chapter 9. If we want to use error compensation

algorithm to improve ENOB of ADC with imperfect MZ modulator, we need to characterize the MZ interferometer to find all its relevant parameters. It turns out, however, that finding all parameters of the MZ is not a trivial task; for example, it is not immediately clear how to find the splitting ratio of the input coupler. We now present an approach to characterize an imperfect MZ modulator to find its properties sufficient for using the error reconstruction algorithm. We will see that some parameters are important in combinations, so we don't have to measure every single parameter of the MZ.

Let us introduce the notation for the MZ parameters which can lead to error in ADC performance (see Fig. 19).
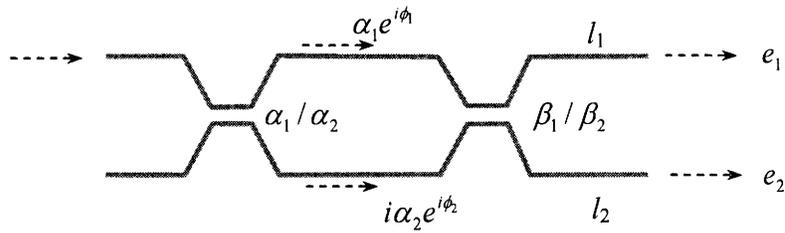
Figure 19. Mach-Zehnder interferometer with non-ideal input and output splitting ratios and difference in internal and output loss between the upper and lower arms.

- α_1 and α_2 are the amplitude splitting coefficients of the input coupler. An ideal 50/50 splitter has $\alpha_1 = \alpha_2 = \sqrt{0.5}$; we are interested in the case when $\alpha_1 \neq \alpha_2$. The power conservation requires that $\alpha_1^2 + \alpha_2^2 = 1$;

- β_1 and β_2 are the amplitude splitting coefficients of the output coupler;

- ϕ_1 and ϕ_2 are phase shifts in the upper and lower arms of the MZ modulator; this includes both the phase shift due to the driving RF signal and the bias phase shift;

- l_1 and l_2 are amplitude transmission coefficients of the upper and lower branches *after* the output MZ splitter. In an ADC system, difference between l_1 and l_2 can be caused, for example, by different propagation lengths of light in the upper and lower arms of the ADC, or by difference in coupling efficiencies from MZ arms to the output fiber (if the ADC is not realized on a single chip);

- e_1 and e_1 are the optical amplitudes at the output of the upper and lower arms, respectively.

In out ADC system the MZ has light only at one of its inputs. In this case a possible difference in losses which occur in the upper and lower arms of the modulator between the two splitters can be included into α_1, α_2 (see Eqs. (9.1), (9.2)). For simplicity, let us assume that the signal with amplitude 1 is at the input of the upper arm of the MZ. The amplitudes after the input α_1/α_2 coupler and the phase shifting sections of the upper and lower arms are then $\alpha_1 e^{i\phi_1}$ and $\alpha_2 e^{i\phi_2}$, respectively. After the output β_1/β_2 coupler and output losses l_1, l_2 the field amplitudes become

$$e_1 = l_1\left[\alpha_1\beta_1 e^{i\phi_1} - \alpha_2\beta_2 e^{i\phi_2}\right],$$

$$e_2 = l_2 i\left[\alpha_1\beta_2 e^{i\phi_1} + \alpha_2\beta_1 e^{i\phi_2}\right].$$

This can be rewritten as

$$e_1 = l_1\left[\alpha_1\beta_1 e^{i\frac{\phi_1-\phi_2}{2}} - \alpha_2\beta_2 e^{-i\frac{\phi_1-\phi_2}{2}}\right]e^{i\frac{\phi_1+\phi_2}{2}},$$

$$e_2 = l_2 i\left[\alpha_1\beta_2 e^{i\frac{\phi_1-\phi_2}{2}} + \alpha_2\beta_1 e^{-i\frac{\phi_1-\phi_2}{2}}\right]e^{i\frac{\phi_1+\phi_2}{2}}.$$

Introducing the average phase shift ϕ and differential phase shift $\Delta\phi$

$$\phi \equiv \frac{\phi_1 + \phi_2}{2}, \qquad \Delta\phi \equiv \frac{\phi_1 - \phi_2}{2},$$

we can rewrite the fields at the upper and lower outputs as

$$e_1 = l_1\left[\alpha_1\beta_1 e^{i\Delta\phi} - \alpha_2\beta_2 e^{-i\Delta\phi}\right]e^{i\phi}, \tag{8.1}$$

$$e_2 = l_2 i\left[\alpha_1\beta_2 e^{i\Delta\phi} + \alpha_2\beta_1 e^{-i\Delta\phi}\right]e^{i\phi}. \tag{8.2}$$

We see that the average phase will determine the phase of the output light; for push-pull modulator configuration the average phase shift is constant so the phase modulation is avoided. The differential phase shift between the arms leads to purely amplitude modulation if the couplers are perfectly 50/50; if errors in coupling coefficients exist, the light will show some phase modulation as well.

The values we can directly measure are the output power which are proportional to $|e_1|^2$, $|e_2|^2$. Let us now find a way to determine the unknown MZ parameters from the measurements of the output powers for varying phase shifts in the modulator arms. Let us

assume that the modulator is overdriven, so that the phase shift between the arms is changed by more than π. The power at the upper and lower outputs will oscillate between some minimum and maximum values which can be measured. From Eqs. (8.1), (8.2) the minimum and maximum power at the upper output $P_{1,\min}$, $P_{1,\max}$ and the lower output $P_{2,\min}$, $P_{2,\max}$ can be expressed as

$$\sqrt{P_{1,\min}} = l_1(\alpha_1\beta_1 - \alpha_2\beta_2),\tag{8.3}$$

$$\sqrt{P_{1,\max}} = l_1(\alpha_1\beta_1 + \alpha_2\beta_2),\tag{8.4}$$

$$\sqrt{P_{2,\min}} = l_2(\alpha_1\beta_2 - \alpha_2\beta_1),\tag{8.5}$$

$$\sqrt{P_{2,\max}} = l_2(\alpha_1\beta_2 + \alpha_2\beta_1).\tag{8.6}$$

The sum and difference of the minimum and maximum powers are

$$\sqrt{P_{1,\min}} + \sqrt{P_{1,\max}} = l_1(\alpha_1\beta_1 - \alpha_2\beta_2 + \alpha_1\beta_1 + \alpha_2\beta_2) = 2l_1\alpha_1\beta_1,$$

$$\sqrt{P_{1,\max}} - \sqrt{P_{1,\min}} = l_1(\alpha_1\beta_1 + \alpha_2\beta_2 - \alpha_1\beta_1 + \alpha_2\beta_2) = 2l_1\alpha_2\beta_2$$

for upper arm, and

$$\sqrt{P_{2,\max}} + \sqrt{P_{2,\min}} = l_2(\alpha_1\beta_2 + \alpha_2\beta_1 + \alpha_1\beta_2 - \alpha_2\beta_1) = 2l_2\alpha_1\beta_2,$$

$$\sqrt{P_{2,\max}} - \sqrt{P_{2,\min}} = l_2(\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_1\beta_2 + \alpha_2\beta_1) = 2l_2\alpha_2\beta_1$$

for the lower arm. The ratios of these sums and differences are

$$\frac{\sqrt{P_{1,\min}} + \sqrt{P_{1,\max}}}{\sqrt{P_{1,\max}} - \sqrt{P_{1,\min}}} = \frac{\alpha_1\beta_1}{\alpha_2\beta_2},\tag{8.7}$$

$$\frac{\sqrt{P_{2,\max}} + \sqrt{P_{2,\min}}}{\sqrt{P_{2,\max}} - \sqrt{P_{2,\min}}} = \frac{\alpha_1\beta_2}{\alpha_2\beta_1}.\tag{8.8}$$

Let us introduce the following notation:

$$M_1 \equiv \frac{\sqrt{P_{1,\min}} + \sqrt{P_{1,\max}}}{\sqrt{P_{1,\max}} - \sqrt{P_{1,\min}}},\tag{8.9}$$

$$M_2 \equiv \frac{\sqrt{P_{2,\max}} + \sqrt{P_{2,\min}}}{\sqrt{P_{2,\max}} - \sqrt{P_{2,\min}}},\tag{8.10}$$

$$k_1 \equiv \frac{\alpha_1}{\alpha_2}, \qquad k_2 = \frac{\beta_1}{\beta_2}.$$

The values of M_1 and M_2 are known from the measurements, and the values of k_1 and k_2 are to be found. Eqs. (8.7) and (8.8) can then be rewritten as

$$M_1 = k_1 k_2,$$

$$M_2 = k_1 / k_2.$$

The unknown parameters k_1, k_2 can then be found as

$$k_1 = \sqrt{M_1 M_2}, \qquad (8.11)$$

$$k_2 = \sqrt{M_1 / M_2}. \qquad (8.12)$$

Equations (8.3)-(8.6) can be rewritten in terms of k_1, k_2 as

$$\sqrt{P_{1,min}} = l_1 \alpha_2 \beta_2 (k_1 k_2 - 1), \qquad (8.13)$$

$$\sqrt{P_{1,max}} = l_1 \alpha_2 \beta_2 (k_1 k_2 + 1), \qquad (8.14)$$

$$\sqrt{P_{2,min}} = l_2 \alpha_2 \beta_2 (k_1 - k_2), \qquad (8.15)$$

$$\sqrt{P_{2,max}} = l_2 \alpha_2 \beta_2 (k_1 + k_2). \qquad (8.16)$$

Introducing

$$\mu_1 \equiv l_1 \alpha_2 \beta_2,$$

$$\mu_2 \equiv l_2 \alpha_2 \beta_2,$$

and subtracting (8.13) from (8.14) and (8.15) from (8.16), we find the unknown parameters μ_1 and μ_2:

$$\mu_1 = \frac{\sqrt{P_{1,max}} - \sqrt{P_{1,min}}}{2}, \qquad (8.17)$$

$$\mu_2 = \frac{\sqrt{P_{2,max}} - \sqrt{P_{2,min}}}{2k_2}. \qquad (8.18)$$

Equations (8.1), (8.2) for MZ output amplitude can be rewritten as

$$e_1 = l_1 \alpha_2 \beta_2 \left[\frac{\alpha_1}{\alpha_2} \frac{\beta_1}{\beta_2} e^{i\Delta\phi} - e^{-i\Delta\phi} \right] e^{i\phi},$$

$$e_2 = il_2\alpha_2\beta_2 \left[\frac{\alpha_1}{\alpha_2} e^{i\Delta\phi} + \frac{\beta_1}{\beta_2} e^{-i\Delta\phi} \right] e^{i\phi},$$

or, substituting the MZ parameters k_1, k_2, μ_1, and μ_2,

$$e_1 = \mu_1 \left[k_1 k_2 e^{i\Delta\phi} - e^{-i\Delta\phi} \right] e^{i\phi},$$

$$e_2 = i\mu_2 \left[k_1 e^{i\Delta\phi} + k_2 e^{-i\Delta\phi} \right] e^{i\phi}.$$

We thus found a way to characterize MZ performance based on parameters extracted from measuring the minimum and maximum output power in both output arms of the overdriven modulator according to (8.11), (8.12), (8.9) (8.10), (8.17), (8.18). With this approach, we can find the MZ output signal without having to determine each MZ parameter individually.

2. Link loss and laser envelope characterization

To measure the losses of the ADC link for different channel, zero RF signal can be applied and detected power at all outputs can be measured. Then, during the ADC operation, the measured outputs can be normalized by the power measured at the calibration step. This approach promises to significantly increase the tolerance of the ADC system towards such imperfections as channel-to-channel loss variations, filter bandwidth variations, and photodetector sensitivity variations.

3. Determining ring center frequencies and dispersion slope

To characterize timing errors caused by fluctuations of the center frequencies of the filters and dispersion slope of the fiber, the following approach can be used. A calibration run of the ADC can be periodically performed e.g. every $1\mu s$. At this run, a Nyquist frequency signal can be sampled by the ADC. If the timing errors are absent, the output of every channel will be zero. With timing errors, some of the outputs will be nonzero, as it is illustrated in Fig. 20. By measuring the power in these outputs, timing errors of the system can be characterized.
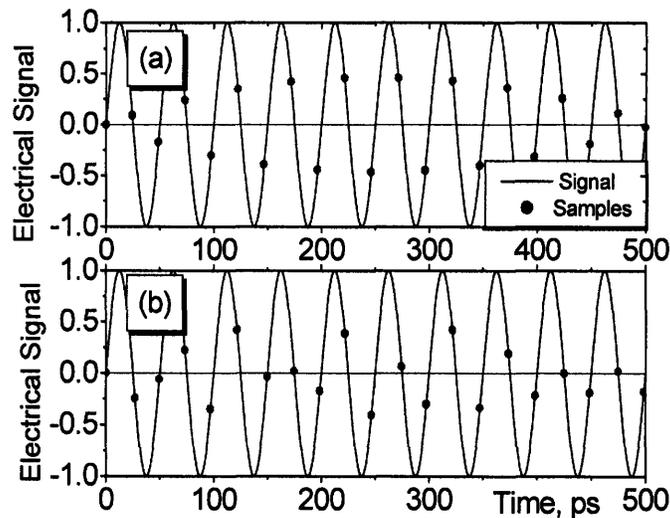
Figure 20. An example of characterization of the timing error of the system with Nyquist-rate signal. (a) timing errors are caused by dispersion slope, (b) in addition to dispersion slope, there is 15GHz random error in center frequency of the filters.

9. Accuracy of a Non-Ideal ADC System

In this chapter we analyze the sensitivity of the ADC to errors in different components of the system. As we've seen before, these errors can be eliminated at post-processing stage if the system is properly characterized. Why then do we need then to study these errors if they can be eliminated anyway? There are two answers to this question

(a) This error analysis establishes guidelines to how accurate the system components must be; it is preferable to do as little as possible at post-processing step to save computation efforts. Moreover, if the system is highly inaccurate, the error compensation algorithm might be unable to reconstruct the true signal; a rule of thumb can be adopted that the ADC should be able to produce about 3 effective bits without any error compensation.

(b) This error analysis determines the accuracy with which the system must be characterized. For example, 1degree error in bias phase without error compensation leads to approximately the same ENOB penalty as 5 degree error in bias phase measured as being 6 degrees error with error compensation algorithm. Therefore, the analysis of this section gives both the requirements for system components for ADC without reconstruction, and the requirements for characterization of ADC components with reconstruction.

The errors in the following ADC system components are analyzed below.

- Bias voltage of the MZ modulator;

- Splitting ratios of the input and output couplers of MZ modulator; loss difference in two arms of the MZ modulator;

- Voltage-dependent loss in the MZ modulator;

- Different loss in upper and lower branches of the ADC and channel-to-channel loss variation. this includes also difference in photodetector and electronic ADC sensitivities;

- Variations of center frequencies of filters in the filter bank;

- Variations in filter bandwidth in the filter bank;

- The length of the dispersive fiber;

- Dispersion slope of the dispersive fiber;

- Thermal and shot noise in the photodetectors;

- Timing jitter of the laser;

- Pulse-to-pulse amplitude variations of the laser.

This rest of this chapter analyzes the impact of each of these nonidealities on ADC performance.

1. Bias phase of the MZ modulator

The bias phase of the MZ modulator should be $\pi/2$ when MZ linearization using both complimentary outputs is used. If $\Delta\varphi^*$ is the error in bias phase, the output of the upper arm of the MZ modulator is described by

$$i(t) \sim \left(1 - \sin\Delta\varphi^* \cos v(t) - \cos\Delta\varphi^* \sin v(t)\right).$$

This is something different from what is obtained in correctly biased modulator, an attempt to invert this using our linearization scheme will not work precisely. In spectral domain, an error in bias phase will lead to second-order distortion in each of the modulator outputs which will be present also after the linearization.

Fig. 21 shows how the effective number of bits is degraded by the bias phase error. For plot (a) the iterative error compensation was performed so that all signal distortion are

removed except the distortion due to bias phase error. We are also interested in ADC performance without error compensation, and the question one might ask is whether plot (a) is of any use in this case? It turns out that the ENOB shown in (a) can be understood as the highest ENOB obtained in ADC without error compensation. This is illustrated in plot (b), which shows ENOB the system with no compensation (MIT EPIC ADC system parameters were assumed for both plots). We see that for small bias phase error the ENOB is limited to ~4.5 bits by other factors such as dispersion; the ENOB starts going down for higher bias phase errors exactly at points predicted by plot (a). Therefore (a) indeed establishes the limit on ENOB with given bias phase error. In the rest of this chapter we will be calculating only ENOB with error compensation as we did in plot (a).

We see that the impact of the bias phase error depends a lot on the modulation depth. It was to be expected, because variation in bias phase is not distinguishable from variation in phase caused by the applied RF signal. If the RF phase variation is weak, even small bias phase error will be comparable to the RF phase leading to low ENOB. The same bias phase error will be smaller relative to large RF phase variation taking place for high modulation indices, hence the sensitivity to bias phase is lower for strong modulation. For low modulation depths, the ADC accuracy is extremely sensitive to this error, with bias phase needed to be controlled or monitored within 0.01 degrees to get 4 bits for 1% modulation; such phase accuracy is very hard to achieve. Therefore, using higher modulation indices is necessary. For 95% modulation (which we will refer to as "full modulation" in the following discussion), the bias phase control should be better than 0.05 degrees to get 8 bits. Such precise phase control is not trivial to achieve and might require monitoring of optical power and using feedback loop to control the bias voltage [19]. Using the error compensation algorithm relaxes these requirements: for full modulation, the bias phase should be controlled to within 1 degree to get 4 bits and continuously measured with accuracy 0.05deg. to restore ENOB up to 8 bits at the post-processing step.
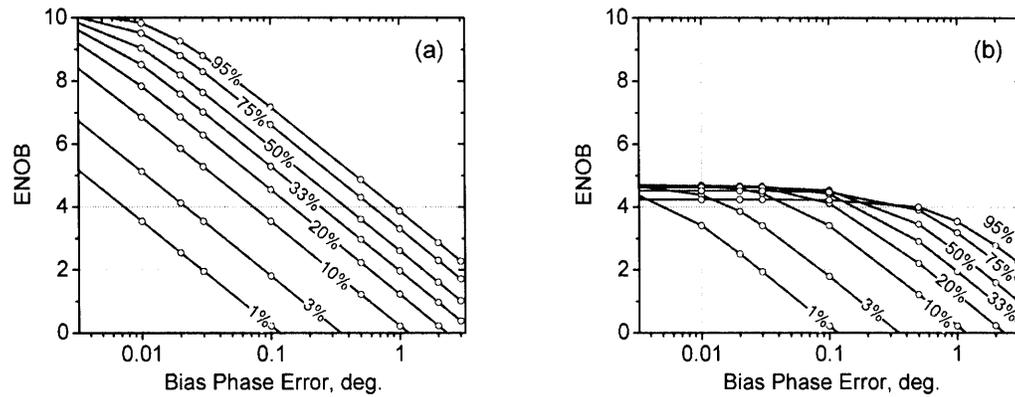
Figure 21. ENOB as a function of error in bias phase of the MZ modulator. The curve parameters is modulation depth. (a) Error compensation algorithm in used to cancel all signal distortions except the distortion due to error in bias voltage; (b) No error compensation is applied.

2. Splitting ratios of the input and output couplers of the MZ modulator

The analysis of an ideal ADC operation assumed that the MZ interferometer has perfect 50/50 input and output splitters. In reality, however, the splitters will not be perfect because of fabrication errors and the MZ output will be proportional to the sine of the applied RF signal, but to some different function (more details on transfer function of a non-ideal MZ interferometer can be found in chapter 8). Let us analyze the impact of errors in the splitting ratios on the ENOB achievable with our ADC.

Fig. 22 shows color map of the effective number of bits versus the error in input and output coupling coefficient. One percent error in these plot means that the power splitting ration is 49/51 rather than perfect 50/50. Figures (a)-(d) correspond to modulation indices of 1%, 10%, 50%, and 95% respectively. The contour lines with number on them represent the lines of constant ENOB. The plots show that the ADC becomes more sensitive to coupling errors when the modulation index is decreased, because in this case the signal level is small to that any disbalance between modulator arms creates distortion which is comparable to this signal. For example, 1% error in input and output couplers can reduce ENOB to 2-3 bits for 1% modulation; if 95% percent modulation is used, the same MZ can produce solid 7 bits of precision.

The MZ interferometer can also have losses in its arms between the two couplers. These losses can be different for two arms if these arms are not identical, e.g. when π bias phase shift is created in one arm but not in another. The differential loss between the two arms

inside the modulator has the same effect as non-perfect splitting ratio in the input coupler. If the error in the splitting ratio is ε %, so that the coupling ratio is $(50 + \varepsilon)/(50 - \varepsilon)$, the result will be the same as if the lower arm has loss α dB less than the upper arm, where

$$\alpha(\text{dB}) = 10 \log\left(\frac{50 + \varepsilon}{50 - \varepsilon}\right). \tag{9.1}$$

For splitting ratio errors less than 10% the differential loss can be approximated as

$$\alpha(\text{dB}) \approx 0.174 \cdot \varepsilon. \tag{9.2}$$

This equation together with Fig. 22 allows to find ENOB degradation due to unequal losses in two arms of the MZ modulator.
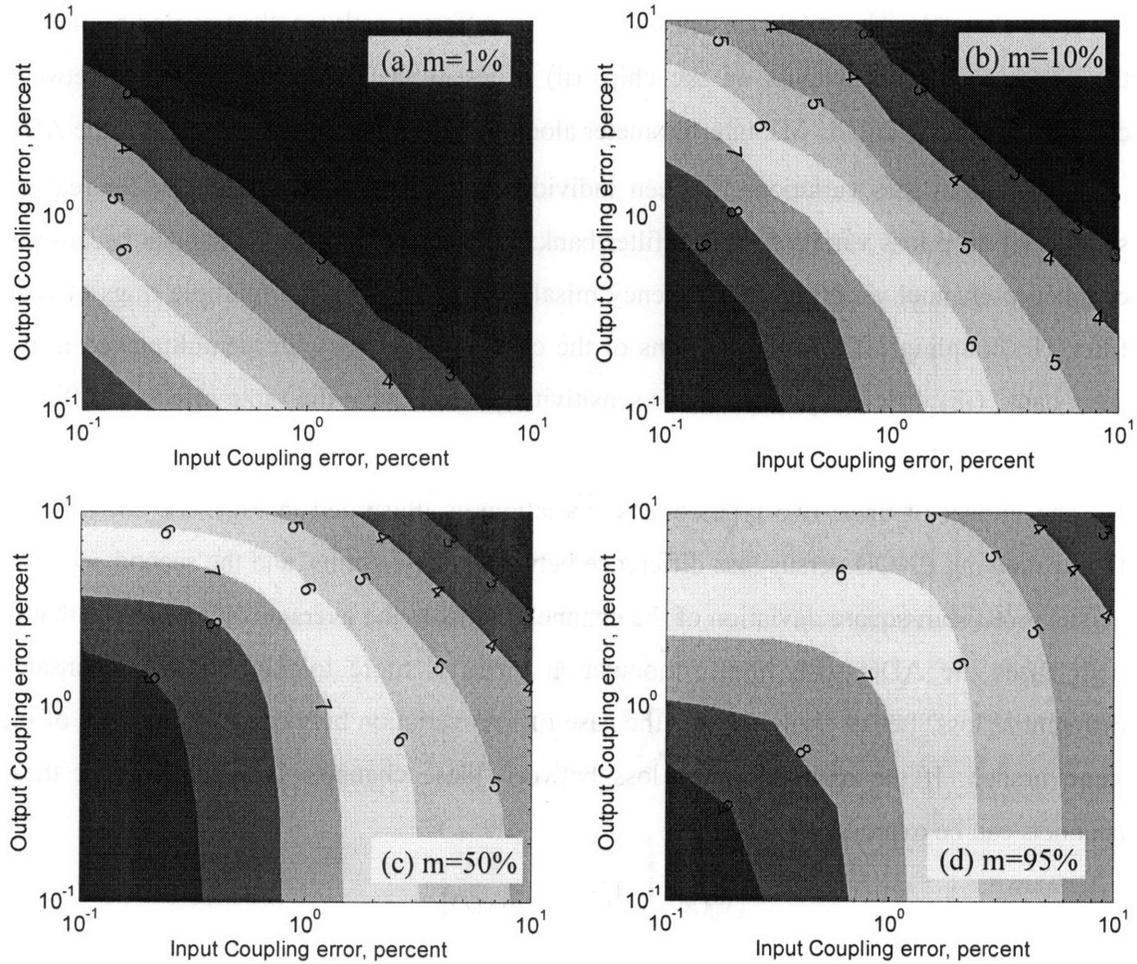
Figure 22. ENOB as a function of error in input and output coupling coefficients of the MZ modulator for modulation indices (a) m = 1%, (b) m = 10%, (c) m = 50%, (d) m = 95%. The numbers along the contour lines are the number of bits, with one bit step between the contours.

3. Channel-to-channel loss variations

In the optically-sampled ADC system we are studying, high sample rate is achieved by splitting fast RF signal into multiple slower channels and then interleaving these channels to obtain the original signal. If the losses or sensitivities in these channels are not precisely matched, the signal will be distorted. Calibration of multi-channel ADC is known to be a challenging task in electronic ADCs. Let us see how sensitive is our system to channel-to-channel loss variations.

Two types of loss difference between channels can be distinguished:

(a) loss difference between two complementary outputs of the MZ modulator. For example, for a 20-channel system, the upper 20 channels might have larger loss than the lower 20 channels. This might be caused by e.g. (i) different path lengths traveled by light in the upper and lower outputs on the chip, (ii) different in coupling coefficients between components following the MZ interferometer along the upper and lower branches of the ADC.

(b) random loss variations between individual channels. This can happen for reasons such as (i) drop loss variations in the filter bank; for ring filters, this can happen because of channel-to-channel variations in frequency misalignment between the multiple rings of each filter, (ii) coupling efficiency variations of the channels after they are demultiplexed in the filter bank, (iii) different photodetector sensitivities, which have the same effect as different optical losses, etc.

The impact of these two types of loss variations is illustrated in Figs. 23, 24, the first figure showing ENOB versus loss difference between the two arms, and the second – ENOB versus root-mean-square deviation of the channel loss from the average value. We see that in both cases the ADC with higher modulation index is more tolerant to uncompensated differential loss. Let us show this for the case of loss variation between two channels of the same branch. If the relative power loss between these channels is ε, the output these channels can be expressed as:

$$v_1(t) = 1 \cdot [1 + m \cdot \sin v(t)],$$

$$v_2(t) = (1 + \varepsilon) \cdot [1 + m \cdot \sin v(t)],$$

where m is the modulation index, $v(t)$ is the applied RF signal. Subtracting these equations we obtain that the error is given by $\varepsilon[1 + m \cdot \sin v(t)] \approx \varepsilon$ for small modulation indices. The absolute error level therefore does not depend on the modulation depth, while the useful

signal $m \cdot \sin v(t)$ does. Thus the error ε will be the larger relative to the signal $m \cdot \sin v(t)$, the smaller is the modulation index. The modulation index in fact defines the "enhancement factor" of any channel-to-channel loss mismatch. For example, if we have an ADC with 10% modulation depth, the precision of the channel loss equalization or measurement must be about one order of magnitude better than the target accuracy of our ADC. For example, if we want to have an ADC with 8 bit resolution, the channel losses must be measured with precision of about 11 bits because of the "error enhancement" due to incomplete modulation.
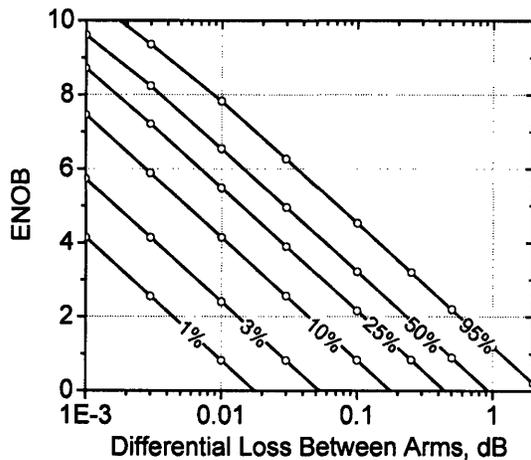
Figure 23. ENOB limited by uncompensated loss difference between the upper and lower branches of the ADC. Different curves correspond to different modulation indices, shown on the curves.
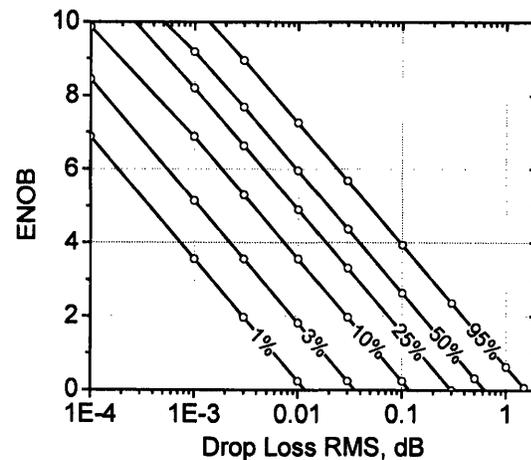
Figure 24. ENOB limited by uncompensated channel-to-channel loss variations. Different curves correspond to different modulation indices, shown on the curves.

The contour plots of ENOB as a function of both the differential loss between the arms and channel-to-channel loss variation are shown in Fig. 25.
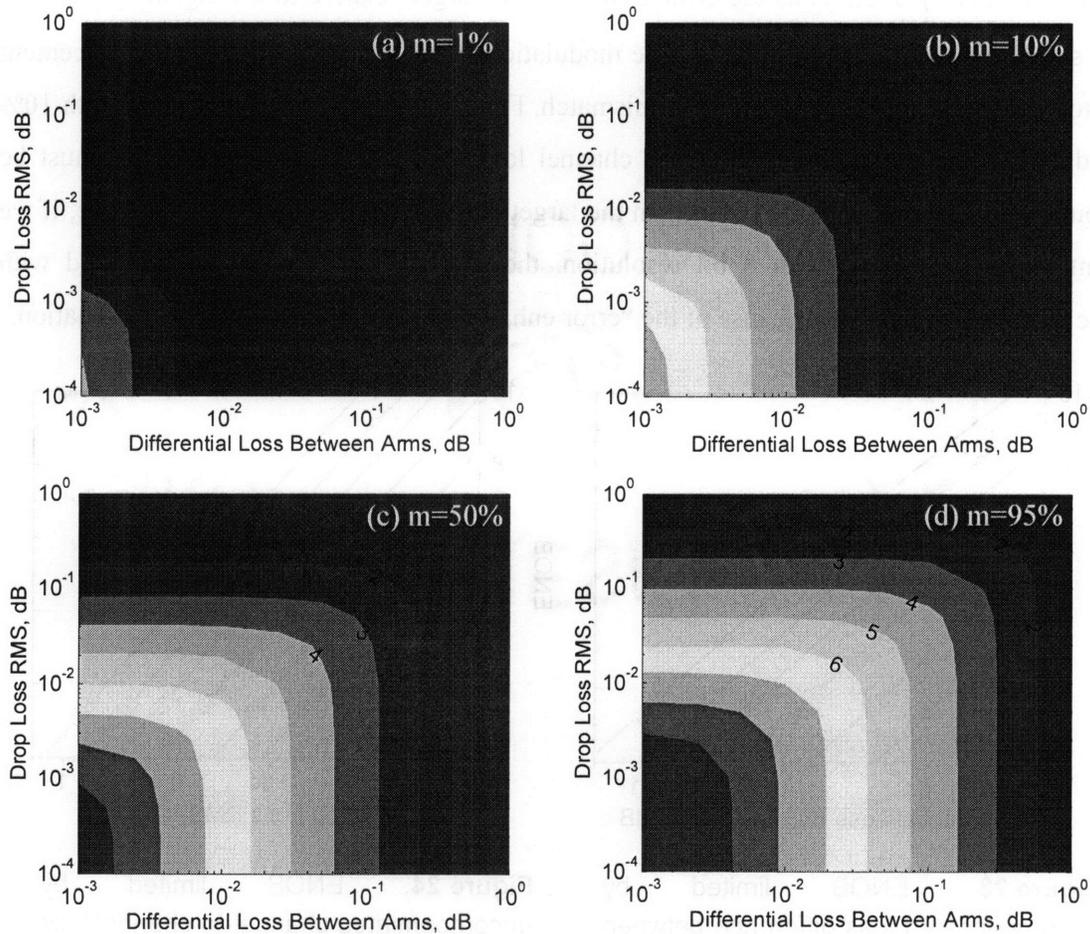
Figure 25. ENOB as a function of loss difference between two complementary outputs (horizontal axis) and root-mean-square deviation of random channel-to-channel loss variation. The modulation indices are (a) m = 1%, (b) m = 10%, (c) m = 50%, (d) m = 95%. The numbers shown along the contour lines is the effective number of bits, with 1 bit step between the contours.

4. Timing jitter of the laser

The timing jitter of the laser source will lead to the RF signal being sampled at wrong time moments and thus ENOB degradation. Timing jitter is especially important because it cannot be measured on a per-pulse basis and therefore the error it causes cannot be compensated at post-processing stage.

The equation for the effective number of bits produced by an ADC with timing jitter was derived in many sources [1, 4, 2, 27]:

$$ENOB = \log_2\left(\frac{1}{\sqrt{3}\,\pi\sigma_t f_s}\right), \qquad (9.3)$$

where σ_t is the RMS timing jitter and f_s is the sampling frequency. This equation is derived assuming that the signal being sampled is at Nyquist frequency $f_s/2$.

Fig. 26 shows plots of ENOB versus RMS liming jitter for several values of sampling rate, as calculated with Eq. (9.3). We see that jitter of less than 100 fs (current state-of-the-art in electronics) is required for obtaining a high-resolution wideband ADC system. For example, the jitter needs to be below 16fs to obtain 8 effective bits at 40Gsample/s. That is why photonic ADCs using very stable mode-locked lasers with jitter of tens of femtoseconds and less can perform much better than electronic ADCs.
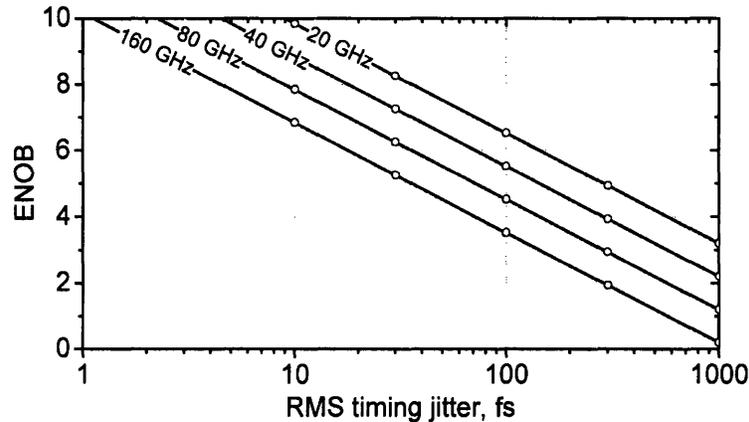
Figure 26. ENOB limited by the timing jitter of the laser for different values of the sampling rate.

5. Filter center frequency variations

Center frequencies of the filters in the filter bank might be misaligned so that the frequency spacing between the channels in not uniform. Because of time-frequency mapping, nonuniform frequency spacing between the channels will translate into nonuniform time intervals between samples and thus reduced ENOB.

The ENOB penalty due to frequency variations depends on the channel spacing but not on the sample rate. This can be understood by converting frequency variations into equivalent timing jitter according to the time-frequency mapping. If σ_f is the standard deviation of frequency from its nominal value, this will be equivalent to RMS timing jitter σ_t

$$\sigma_t = 2\pi \sigma_f \cdot \beta_2 L.$$

From uniform sampling condition (4.2)

$$\beta_2 L = \frac{2\pi}{\Delta \omega f_s} = \frac{1}{\Delta f \, f_s},$$

where Δf is channel frequency spacing. The equivalent timing jitter is then

$$\sigma_t = \frac{\sigma_f}{\Delta f} \frac{1}{f_s}. \tag{9.4}$$

We see that the timing jitter is inversely proportional to channel spacing, so the system with larger channel spacing is more tolerant to channel frequency variations. Substituting (9.4) into (9.3) we get a simple expression for the effective number of bits limited by frequency variations:

$$ENOB = \log_2\left(\frac{\Delta f}{\sqrt{3}\,\pi\sigma_f}\right). \tag{9.5}$$

Although impact of timing jitter is proportional to sampling rate, the impact of frequency variations is not, because the "equivalent timing jitter" corresponding to frequency variations is inversely proportional to sampling rate.

Fig. 27 shows ENOB as a function of σ_f for several values of cannel spacing Δf. The values shown in the plot are about 1 bit higher than those predicted by (9.5). This happens for the following reasons.

• In the simulation, the frequencies of upper and lower parts of the dual filter bank were varied independently, while the equation implies that the frequencies of the upper and lower parts are the same. When the frequencies are different, the jitter in the upper and lower parts partially cancels each other so the ENOB is higher. From statistical point of view, if we have two random variables f_1 and f_2 with standard deviation σ_f, the average of these variables $(f_1 + f_2)/2$ will have standard deviation $\sigma_f / \sqrt{2}$.

• Another reason for the difference between simulation and theory is that the simulation assumes that the frequency spectrum of the signal being sampled contains range of frequencies $[0.55...0.95] \cdot f_s / 2$ (see the Appendix), while the theoretical equation (9.5) was

obtained for signal at $f_s/2$. Because the frequencies in the simulation were somewhat smaller, the obtained ENOB was larger (by about 0.3dB).

Note that in contrast to errors caused by the timing jitter, errors caused by filter frequency variations can be eliminated using the iterative algorithm if these frequencies are accurately measured.
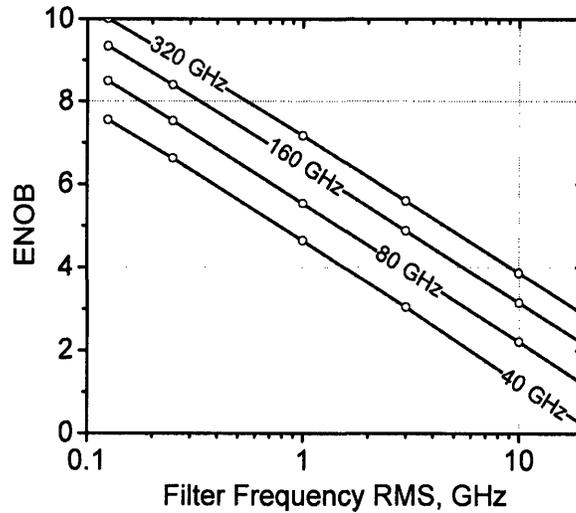
Figure 27. Impact of filter frequency variations on the accuracy of ADC. The curve parameter is channel spacing.

6. Error in fiber length

Errors in fiber length can be divided into two types: (a) fluctuations in fiber length due to thermal expansion and mechanical strain, and (b) systematic error in fiber length.

Fluctuating fiber length will lead to fluctuations in group delay experienced by pulses and therefore in sampling times. Even small variations in length can result in jitter comparable to the jitter of the laser. For example, 10μm variation of length in a 2km fiber will lead to 50fs variation of the pulse arrival time. The fluctuations in fiber length will therefore lead to distortion of the signal produced by our ADC. However, fluctuations of pulse arrival times are slow and happen over many repetition periods of the laser; the pulses stay approximately uniformly spaced within characteristic time of these fluctuations. Therefore, this will produce only low-frequency errors in the ADC, and such errors are not important in practice. If necessary, fiber length variations can be compensated as described in [28], where the authors

used a moving stage with a feedback loop to obtain timing stabilization of 200m fiber link to within 10fs over many hours.

Systematic error in fiber length as compared to the length (4.3) defined by uniform sampling condition will result in incorrect relative time delay between sub-pulses corresponding to different channels and thus non-uniform sampling by our ADC. Fig. 28 shows ENOB limited by incorrect fiber length. The accuracy depends only on the relative error in fiber length and the number of channels in the system, as it follows from the following argument. If the fiber of length L produces relative time delay between channels equal to $\dfrac{T_R}{2}$, where T_R is the repetition period of the laser, error ΔL in fiber length will lead to commensurate timing error $\Delta\tau = \dfrac{T_R}{2}\dfrac{\Delta L}{L}$. The impact on ENOB is defined by $\dfrac{1}{f_s\Delta\tau}$, where f_s is the sampling rate. Substituting $\Delta\tau$ and $f_s = \dfrac{N}{T_R}$ we conclude that the ENOB will be defined by

$$ENOB = \log\left(\frac{1}{f_s\Delta\tau}\right) + C = \log\left(\frac{1}{N}\frac{L}{\Delta L}\right) + C.$$

We see that ENOB depends only the number of channels and relative error in fiber length.
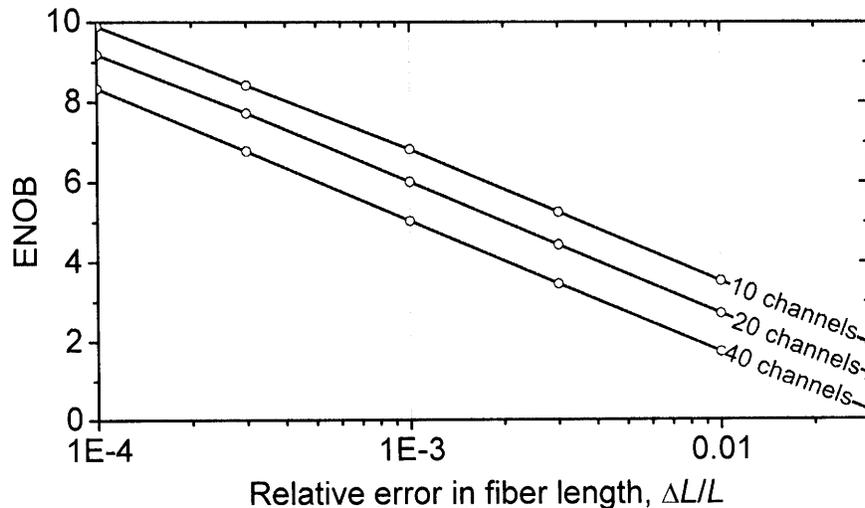
Figure 28. ENOB versus relative error in fiber length; the curve parameter is the number of channels in the system.

7. Polarization fluctuations in the fiber

If the fiber is not polarization maintaining, the pulse train will experience PMD effect. This will lead to fluctuations in pulse arrival times because of difference in group velocities between the two polarization. For example, if a 2km fiber is used with PMD coefficient of $0.1\,ps/\sqrt{km}$, the average group delay between the two polarizations will be 140 fs. This will lead to sampling time error. However, variations in pulse arrival times are slow so they will not be important for real ADC system (similar to the case of fiber length variation).

PMD will also lead to time-dependent polarization of the output signal. If the ADC components following the fiber (i.e. modulator and filter bank) work for just one polarization, the accuracy of ADC can be degraded if the input polarization is wrong. For example, the ring filter bank described in chapter 10 works only for TE polarization. If at some time point polarization fluctuates so that the polarization at the output of the fiber is TM, ADC will not work. Therefore, in this case we must use either a polarization-maintaining fiber, or a polarization controller which makes sure that the output polarization is always TE. Alternatively, a polarization diversity scheme can in principle be implemented on the integrated optical chip [15]. However, it might be better to avoid complicating the ADC system by doubling the number of modulators and filters as it is necessary for implementing the polarization diversity scheme, and use single polarization and polarization-maintaining fiber instead.

8. Dispersion slope of the fiber

In addition to second-order dispersion, the dispersive fiber also has the third-order dispersion described by dispersion coefficient β_3. The third-order dispersion will introduce additional delay between channel which is proportional to square of the frequency spacing. Let us study how it will reduce the accuracy of the ADC.

The timing error for the outermost channels $\Delta\tau$ is proportional to square of frequency spacing between them, $\Delta\omega N$:

$$\Delta\tau \sim \beta_3 L(\Delta\omega N)^2, \qquad (9.6)$$

where L is the length of the fiber, $\Delta\omega$ is the channel frequency spacing, N is the number of channels, and β_3 is the third-order dispersion coefficient which is related to dispersion slope S by

$$\beta_3 = \frac{c^2}{(2\pi)^2 f_{ref}^4} S,$$

where f_{ref} is the reference frequency at which zero group delay is assumed. The effective number of bits will be defined by the inverse of the timing error and sampling rate,

$$ENOB = \log_2\left(\frac{1}{f_s \Delta \tau}\right) + C. \tag{9.7}$$

From the uniform sampling condition (4.2) we get

$$L\Delta\omega = \frac{1}{\beta_2 f_s}.$$

Substituting this into expression (9.6) for $\Delta\tau$ and then $\Delta\tau$ into (9.7) we get

$$ENOB = \log_2\left(\frac{1}{\Delta\omega N^2 \beta_3/\beta_2}\right) + C. \tag{9.8}$$

We see that the accuracy is inversely proportional to channel spacing and square of the number of channels. The accuracy is also be inversely proportional to the ratio of third and second order dispersion coefficients, which is a property of the dispersive fiber.

Fig. 29 shows ENOB as a function of channel spacing for several values of the number of channels. The parameters of the fiber were close to parameters of Corning SMF-28 fiber. We see that the impact of dispersion slope is important for this type of fiber. For example, for a system with 20 channels with 80GHz channel spacing and 2GHz repetition rate, ENOB is limited to only about 2.5. The timing error for the outermost channels can be calculated as $SL\Delta\lambda/2$, with $\Delta\lambda$ being 10 times wavelength channel spacing. This gives the timing error of about 4 ps, which is a large error for a system like ours (compare to Fig. 26), and therefore ENOB is significantly degraded. The error compensation algorithm can take this error out once the dispersion slope is properly characterized, however, even in this case using a fiber with better D/S is highly desirable in order to improve the quality of the signal with which iterations start. Using a fiber with much larger ratio of D/S is necessary for getting high ENOB in this ADC system; according to (9.8), ENOB will improve by 1 for each factor of 2 increase in D/S. If a fiber with sufficiently high D/S is not available, dispersion slope compensation scheme might be implemented, i.e. two fibers with opposite signs of dispersion slope can be concatenated so that the overall dispersion slope is zero. Alternatively, some

other dispersive element with smaller dispersion slope can be used instead of a fiber, such as a pair of reflection gratings.
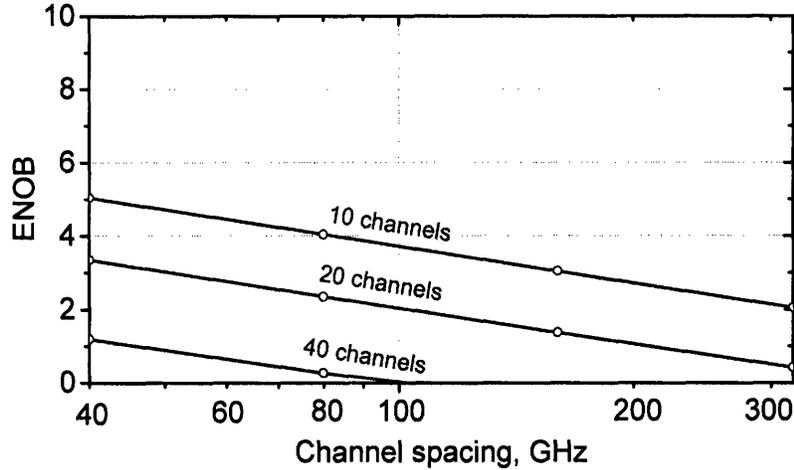
Figure 29. ENOB limited by fiber dispersion slope versus channel spacing for different number of channels. A fiber with parameters close to Corning SMF-28 was assumed, with $D = 16\,ps/(nm \cdot km)$ and $S = 0.08\,ps/(nm^2 \cdot km)$.

9. Nonlinearity of the fiber

When ultrashort pulse train from the laser passes the fiber, it will experience nonlinear effects because their peak power is high. The nonlinear phase will distort the time-frequency mapping on which the operation of our ADC relies and therefore reduce the obtained ENOB.

Fig. 30 illustrates ENOB degradation due to Kerr nonlinearity in the fiber. When channel spacing is increased, the fiber needs to be shorter, but the nonlinearity does not really go down because most of the nonlinear phase shift happen in the beginning of the fiber where the pulses are not yet broadened by dispersion and their peak power is high. When the pulse duration is increased, its peak power decreased; however, it turns out that this also does not decrease nonlinearity because a longer pulse has longer dispersion length so it retains high peak power over a longer part of the fiber than a shorter pulse. The simulation for Fig. 30 was done using parameters close to those of SMF-28 fiber: core area ~80μm^2, $n_2 = 2.6 \cdot 10^{-20}\,m^2/W$, $D = 16\,ps/(nm \cdot km)$, and ADC parameters were $f_R = 2GHz$ and $N = 20$. Simulation was done with split-step Fourier algorithm with adaptive step size control which uses high resolution at the beginning of the fiber where the peak power high large and increases step size as the peak power and thus nonlinearity is decreased.
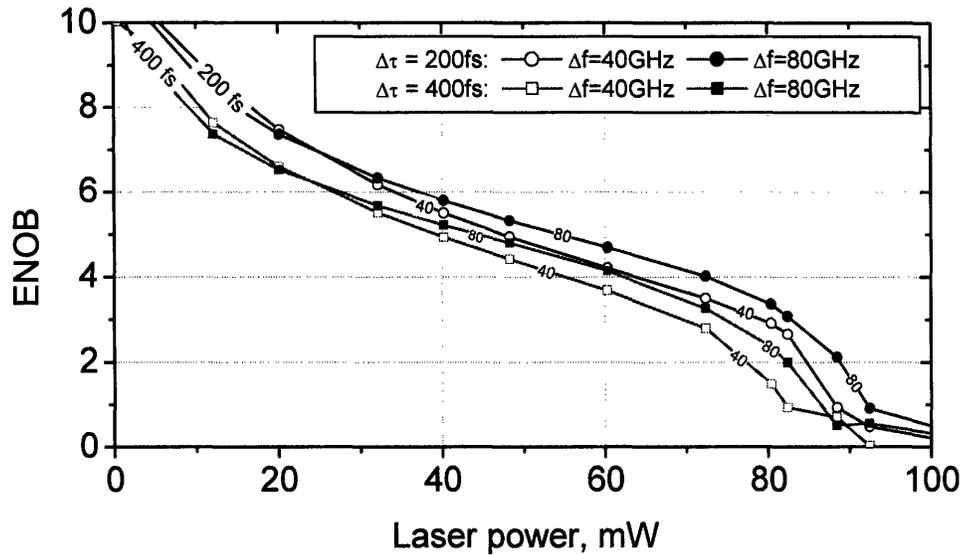
Figure 30. ENOB limited by fiber nonlinearity versus average power at the input of the fiber. The pulse duration $\Delta\tau$ is 200 and 400fs and the channel spacing Δf is 40 and 80 GHz. SMF-28 fiber parameters were used in the simulation.

Simulation results show that fiber nonlinearity can be an important issue in a system which is sufficiently energy efficient and requires high optical power. One way to mitigate nonlinearity can be to use small optical power at the input of the fiber and then amplify it with an optical amplifier. Another solution can be to use a different method to introduce dispersion, such as using a pair of diffraction gratings.

10. Filter shape variations

The filter in the filter bank might have non-identical shapes of the transmission function, which will lead to different transmission characteristics for different channels and reduction in ADC accuracy. Filter shapes might vary in many different ways, e.g. in bandwidth, peak transmission, shape of the passband, etc. For the purpose of understanding the impact of filter variations on ADC performance, we consider just two types of variations:

- variations of filter bandwidth with all other characteristics being identical for all filters;

- variations of shapes of the filters implemented as double ring resonators. Filter shapes can vary from filter to filter because of random mismatch of resonant frequencies of the two

rings. Frequency mismatch will to an increased filter bandwidth and also an increased drop loss at the same time.

First we consider the case when only the bandwidth of the filter is varied. Fig. 31 shows ENOB as a function of relative standard deviation of filter FWHM bandwidth for several values of modulation index. For example, if the filter FWHM bandwidth is 20GHz, the value of 0.01 along the x-axis corresponds to standard deviation of 20 MHz. Note that variations of the filter bandwidth are in many respects similar to variations of drop loss considered above, because larger bandwidth means more transmitted power. Similar to that case, the impact on ENOB strongly depends on modulation index.
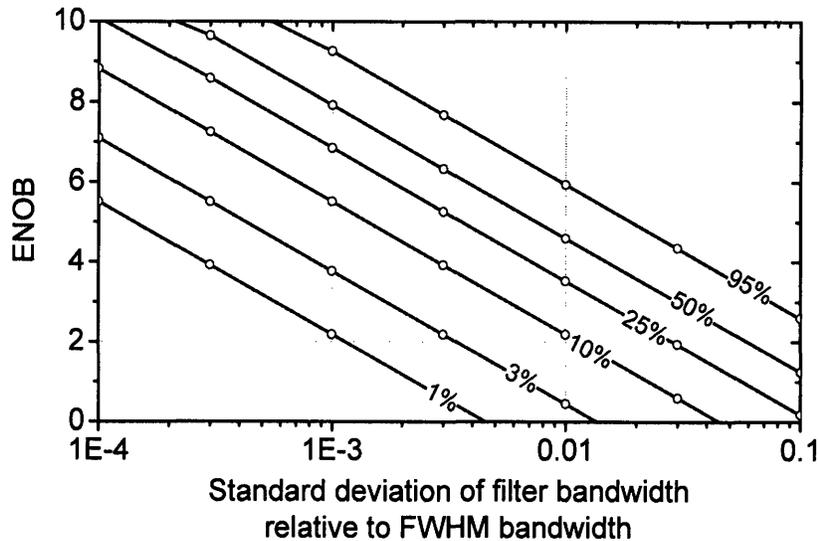
Figure 31. Impact of filter-to-filter bandwidth variation on accuracy of the ADC. The curve parameter is the modulation depth. The peak transmission is the same for all filters.

Now we consider two-ring filters with random frequency mismatch between the two rings, which changes both filter bandwidth and drop loss. Fig. 32 shows ENOB as a function of the standard deviation of the frequency mismatch between rings[1] for several modulation indices. It is interesting to compare these results with a plot of error caused by variations in center frequencies of the filters, Fig. 27. The ENOB penalty due to misalignment of the two rings is

[1] Ring filter parameters described in chapter 10 were used: bus-ring and ring-ring power coupling coefficients were 3.35e-2 and 7.60e-4, respectively; the radius of the ring 10μm and propagation loss 10 dB/cm were assumed. Such a filter has FWHM bandwidth of about 25 GHz and FSR of 2.06 THz

smaller than the penalty due to shift in center frequency of the whole filter by the same amount if the modulation index is large. This can be seen in Fig. 33 which puts some curves from Figs. 27 and 32 on the same plot. The fact that ENOB is less sensitive to frequency misalignment can be explained by the fact that the dependence of filter bandwidth and drop loss on frequency mismatch is quadratic, as it is clear from Fig. 34 which plots them in logarithmic scale. For small values of frequency mismatch variations (which should be the case for a good ADC system), the bandwidth and drop loss variations are very small so they do not have much impact on the system performance.
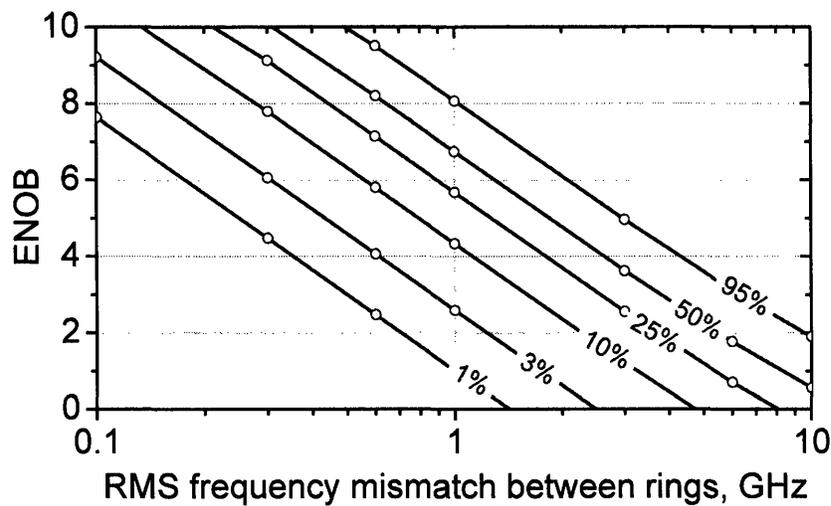
Figure 32. Effective number of bits limited by random frequency mismatch between the rings of the two-ring filter. Different curves correspond to different modulation depths.
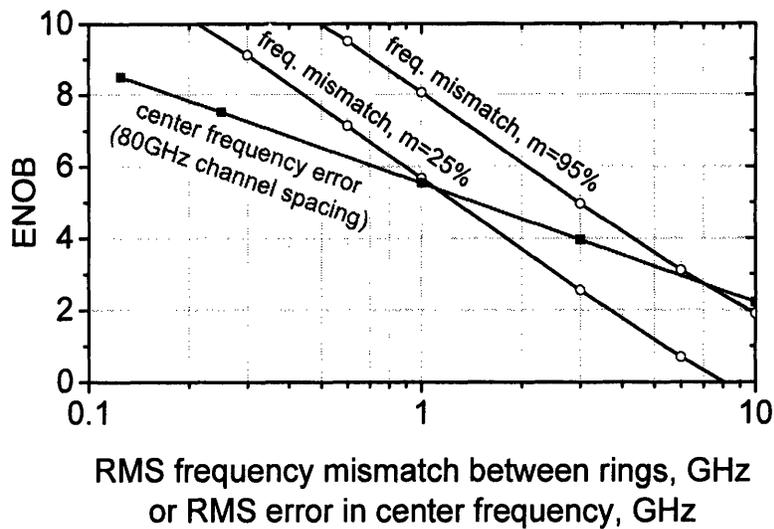
Figure 33. Comparison of ENOB penalty caused by errors in center frequencies of the filters (Fig. 27) and caused by the same amount of frequency mismatch between the two rings (Fig. 32). The curves for frequency mismatch correspond to modulation indices m=25% and m=95%, and the curve for center frequency error corresponds to 80 GHz channel spacing.
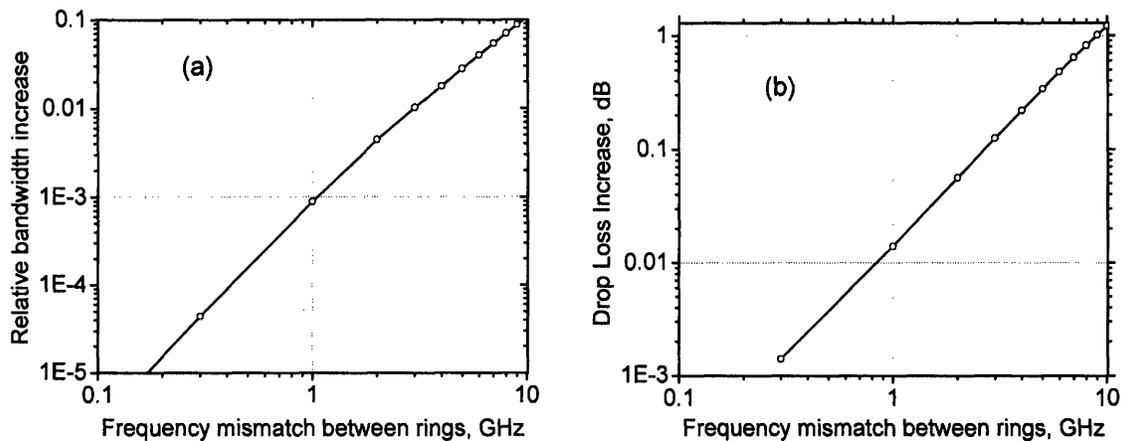
Figure 34. (a) Relative increase in filter bandwidth and (b) increase in drop loss as a function of frequency mismatch between the rings of the two-ring filter.

11. Photodetector noise and power budget

Our analysis so far did not include the impact of detection noise. Detection noise limits the minimum energy in a pulse hitting the photodetector. Let us derive an expression for the SNR of our ADC system limited by photodetector noise.

The total power of the detected signal is

$$P_{det} = R\bar{I}^2,$$

where R is resistance, and $\bar{I}$ is the average current given by

$$\bar{I} = \frac{\eta q}{hf} P_{in},$$

where η is the quantum efficiency of the photodetector, f is the frequency of the signal, h is Planck's constant, and P_{in} is the optical power incident on the photodetector. The power of the shot noise is

$$P_{shot} = 2q\bar{I}\,\Delta f_{det}R,$$

where q is the electron charge and Δf_{det} is photodetector bandwidth. The total detected signal contains a constant part and a varying part with amplitude determined by the modulation index m; only the latter contains useful signal. SNR of our ADC system can thus be expressed as

$$SNR = m^2 \frac{P_{det}}{P_{shot}}.$$

Substituting expressions for P_{det} and P_{shot} we get

$$SNR = m^2 \frac{\bar{I}}{2q\,\Delta f_{det}} = m^2 \frac{\eta P_{in}}{2\,hf\,\Delta f_{det}}. \qquad (9.9)$$

The optical power incident at photodetector P_{in} can be expressed as

$$P_{in} = \alpha \frac{P_{laser}}{2N}, \qquad (9.10)$$

where P_{laser} is the output power of the laser which is into two parts by MZ modulator and then into N channels by the filter bank. Parameter α defines power loss in all components of the ADC, which includes

- spectrum fill factor, i.e. the fraction of the spectrum actually passed through by the filters; this factor depends on the ratio of filter bandwidth to channel spacing;

- drop loss of filters;

- modulator insertion loss;

- loss during coupling between different components of the system, etc.

Substituting the input power P_{in} (9.10) into the expression for SNR (9.9) we finally obtain

$$SNR = \alpha m^2 \frac{\eta}{4 \, hf \, \Delta f_{det}} \frac{P_{laser}}{N}. \tag{9.11}$$

Fig. 35 shows ENOB obtained from the simulation of the ADC system with the presence of shot noise. In this simulation, the photodetector bandwidth was $\Delta f_{det} = 5 GHz$, quantum efficiency of the photodetectors $\eta = 1$, spectrum fill factor = 1/5 (corresponding to e.g 25 GHz FWHM channel bandwidth and 80 GHz channel spacing). It was also assumed that losses in the system are absent, i.e. $\alpha = 1$. The laser power shown along x-axis is the power after the bandlimiting filter was applied; the required power from the laser is actually higher, depending on how much of its spectrum is not used and is thus rejected by the bandlimiting filter. The simulation results shown in Fig. 35 agree well with predictions of Eq. (9.11). For other ADC parameters, the results shown in Fig. 35 can be scaled according to Eq. (9.11).
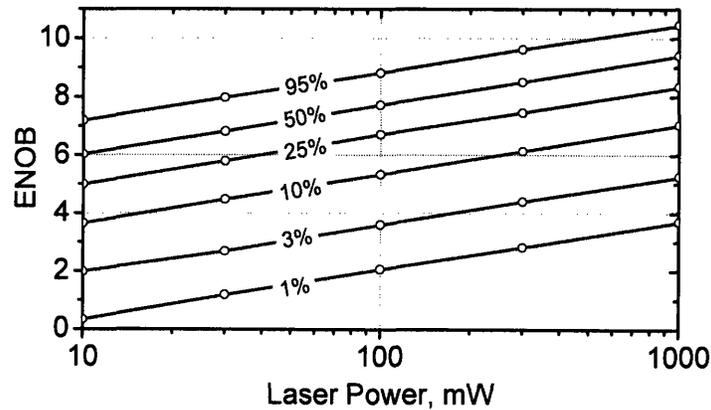
Figure 35. ENOB limited by photodetector shot noise versus laser power for different values of modulation index. ADC parameters used in simulation were N=20, spectrum fill factor = 1/5, photodetector bandwidth $\Delta f_{det} = 5 GHz$ and quantum efficiency η=1, absence of optical power loss (α=1).

The impact of detection noise is one of the main factors influencing the design of the ADC system. Unlike many other errors, error due to detection noise cannot be taken out by the error compensation algorithm. Therefore, we need to have sufficient optical power for accurate operation of the ADC. If the optical losses in the system are substantial, the required optical power will be high, which is undesirable not only from the power efficiency point of view, but sometimes also from physical constraints. For example in Si waveguides where the two-photon absorption and free carrier absorption will lead to large losses if the optical power is high[1]. High optical power can also lead to nonlinear effects in the fiber (Fig. 30) and waveguides. In this respect it is important to note that the impact of noise strongly depends of the modulation index, so using high modulation indices is essential for reducing power requirements. Increasing filter bandwidth to improve the spectrum fill factor is also important; this will of course lead to errors caused by filter convolution (chapter 6), but this kind of error can be eliminated by error compensation algorithm.

---

[1] The largest allowed optical power is on the order of 250mW, as estimated by Milos Popovic.

III. RING RESONATOR FILTER BANK

10. Ring Resonator Filter Bank Design

1. Ring resonators

Filters based on ring resonators [29, 30] are now widely used in integrated optics and enable many useful applications. The direct application of the ring filters include optical filtering and demultiplexing; reconfigurable add-drop demultiplexer can be created by tuning the resonant frequency of the filters. The ring filters also find application in nonlinear optics because of significant field enhancement achievable inside a ring [31]. The sensitivity of the ring frequency response to variations in the ring characteristics allows to use them for optical sensing [32] and efficient high-speed modulation [33].

The design of a ring filter consists of choosing the waveguide dimensions and coupling coefficients [34]. The choice of waveguide dimensions must consider multiple factors, such as high sensitivity to fabrication errors, scattering loss dependence on waveguide cross-section, low bend losses, etc. [13, 14, 35]. The polarization dependence of the resonant frequency can be worked around using a polarization diversity scheme [15].

2. Target filter bank parameters

This chapter describes step-by-step the procedure of design of the filter bank for EPIC ADC system described on page 22. Although our calculations are done keeping in mind the specific requirements and constraints of that system, the approach is general and fully applicable to other cases.

The target characteristics of the EPIC ADC filter bank are:

- 20 filters with 80GHz channel spacing and 25GHz FWHM bandwidth;

- crosstalk between adjacent channels must not exceed -30dB (which makes ENOB of up to 10 achievable without iterative error compensation);

- possible materials for the waveguide core are stoichiometric silicon nitride with refractive index up to n = 2.0 and silicon-rich nitride with n = 2.2. The former has potentially lower loss.

- free spectral range (FSR) must exceed the total bandwidth of the filter bank, i.e. $20 \cdot 80\,\text{GHz} = 1.6\,\text{THz}$. The requirement that FSR must be at least 2.0 THz was adopted to allow a bandlimiting filter to be used to prevent crosstalk due to transmission at multiple FSRs;

- the waveguides are overcladded with hydrogen silsesquioxane (HSQ) with 1.5 μm thickness (see chapter 14 for more details); the undercladding is silicon oxide with $n = 1.445$;

- it is desirable that the height of the waveguide does not exceed 400nm because fabrication of SiN waveguides taller than 400nm is problematic and requires additional research.

3. Choosing waveguide material and dimensions

We start the filter bank design with choosing the material for the waveguide core among stoichiometric and silicon-rich SiN with refractive indices 2.0 and 2.2, respectively. The procedure was developed in the authoritative work of Milos Popovic [13, 30].

Let us check if the required FSR is achievable with rings made of both available materials. The FSR is determined by the group index of the mode n_{gr} and the bend radius of the ring R; the smaller is the bend radius, the higher is FSR:

$$FSR = \frac{c}{2\pi R} n_{gr}.$$

The smallest bend radius is limited by radiation loss in the ring. The radiation loss depends on the waveguide dimensions, with general rule being that the radiation loss decreases as the mode becomes more confined in the waveguide, which happens when the waveguide cross-section becomes bigger. However, the waveguide cannot be made arbitrarily large because of the requirement that it must stay single-mode; if the ring waveguide is multi-mode, power coupling between the modes will occur in the ring, which will prevent its proper operation. Therefore, we are interested in choosing a waveguide with the largest possible cross-section provided that the waveguide remains single-mode.

Fig. 36(a) shows a curve defining waveguide cross-sections at cut-off condition for the second TE mode of stoichiometric nitride waveguide[1]. The x-axis of the plot is the width and

---

[1] All results described in this chapter were obtained with the mode solver created by Milos Popovic.

y-axis is the height of the waveguide. The waveguide has only one TE mode for all cross-sections below this curve and is multi-mode for all cross-sections above it. We are interested in waveguides with cross-sections lying at this curve because they provide the highest confinement and thus smallest bend radius and highest FSR while staying single-mode.
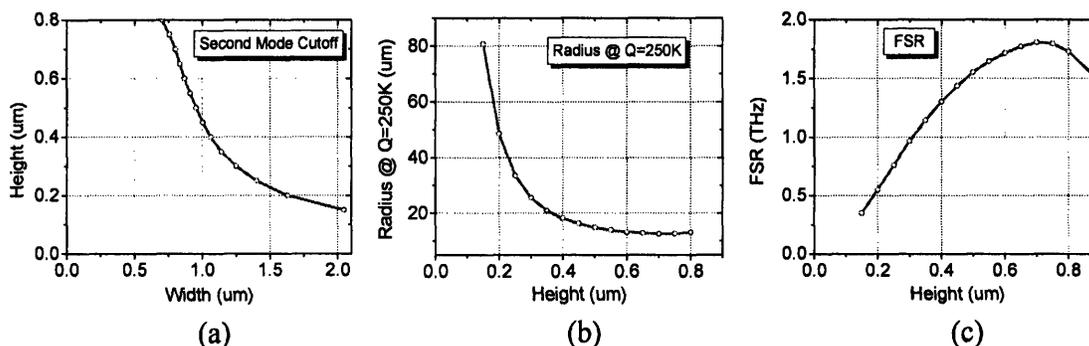
Figure 36. Selecting cross-section of the waveguide made of stoichiometric SiN (n=2.0). (a) the curve defining the cutoff condition for the second TE mode; the axes are width and height of the waveguide; (b) bend radius providing radiation Q=250k versus height of the single-mode waveguide with width from plot (a); and (c) FSR versus height of the single-mode waveguide with width from plot (a) and bend radius from plot (b).

Fig. 36(b) shows the center radius of the ring which gives bend loss-limited Q-factor of 250,000. The height of the waveguide is defined by the x-axis and the width – by the single-mode curve in Fig. 36(a). The number 250,000 for the Q-factor is somewhat arbitrary and is chosen to be high enough so that the power loss per round trip due to bend loss can be neglected compared to coupling to the bus waveguide. Fig. 36(c) shows FSR versus waveguide height; the waveguide width is the one defined by single-mode curve in Fig. 36(a) and bend radius is defined from Q=250,000 condition, as shown Fig. 36(b). We see that the smallest bend radius is limited to about 12 μm and largest FSR does not exceed 1.8 THz for stoichiometric silicon nitride. The largest possible FSR of 1.8 THz is achieved for a waveguide which is about 700 nm tall, which is more than the maximum height that can be fabricated without additional research, so even FSR = 1.8 THz is not realistic. Therefore, stoichiometric nitride is not a suitable material for making filter bank for EPIC ADC system.

Fig. 37 shows similar plots for silicon-rich nitride (n=2.2). The smallest radius in this case is about 9 μm and the largest FSR is about 2.4 THz corresponding to 600nm tall waveguide. For 400nm tall waveguide which is easier to fabricate the FSR is about 2.1THz, which

exceeds the required 2THz. Therefore we choose silicon-rich nitride as the material for the waveguides.

(a)   (b)   (c)
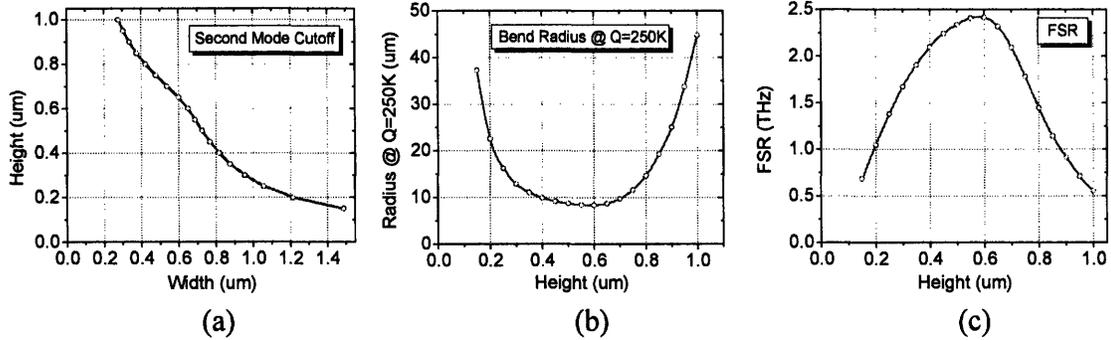
Figure 37. Selecting cross-section of the waveguide made of silicon-rich SiN (n=2.2). Plots similar to those of Fig. 36: (a) the curve defining the cutoff condition for the second TE mode; the axes are width and height of the waveguide; (b) bend radius providing radiation Q=250k versus height of the single-mode waveguide with width from plot (a); and (c) FSR versus height of the single-mode waveguide with width from plot (a) and bend radius from plot (b).

Apart from fabrication constraints, an additional reason to choose the height of the waveguide to be 400nm rather than some larger value is that we want the waveguide to be thinner and wider so that the design is less susceptible to variations in waveguide width due to fabrication errors and sidewall roughness plays smaller role [13, 36]. A waveguide with even smaller height and larger width will be even less sensitive to fabrication errors, but it will not allow to obtain the target FSR of 2 THz according to Fig. 37(c). Therefore we fix the waveguide height to be 400 nm.

We now need to choose the width of the waveguide. For the 400-nm-thick waveguide to be single-mode, its width should be below 815nm according to Fig. 37(a). The single-mode curve, however, was obtained for straight waveguide, and as it was found out in [35], a bent waveguide might still support a second mode; the light can couple into this second mode, leading to loss. It was found that for SiN waveguides this issue is less important than for Si waveguides considered in [35], but nevertheless the width of the ring waveguide should be decreased from 815 to 800nm to make it single-mode.

The second important consideration for choosing waveguide width described in [35] is the following. In the bus-ring coupling region, the two waveguides are close together and it turns out that they can locally support a third mode. Some fraction of the power will be coupled into the this mode in the coupling region, and this power will be radiated after the

light exits the coupling region. This will lead to some loss which is not so important for the light passing through the ring, but is critical for the light resonating inside the ring because it will experience this loss each time it passes through the coupling region. In order to reduce power loss, we can reduce the waveguide width in order to avoid having the third mode in the coupling region. Making the ring waveguide narrower is not a good solution because it will reduce field confinement and thus the FSR; therefore, we should reduce the width of the bus waveguide. Fig. 38 shows the intensity distribution of the third mode of two straight waveguides separated by a small gap; one waveguide (ring) is 800nm wide, and the other (bus) is 800, 700, and 600nm wide. The effective index of the third mode is shown too. The third mode becomes cut off when its effective index goes below the index of the undercladding, i.e. below 1.445. This happens for the width of about 600 nm, as it can be seen from the rightmost part of Fig. 38; we also see from the intensity distribution that the third mode becomes unconfined. Therefore, we choose 600 nm for the width of the bus waveguide.

$w = 800$ nm: $n_{\text{eff}} = 1.460$ $w = 700$ nm: $n_{\text{eff}} = 1.447$ $w = 600$ nm: $n_{\text{eff}} = 1.4436$
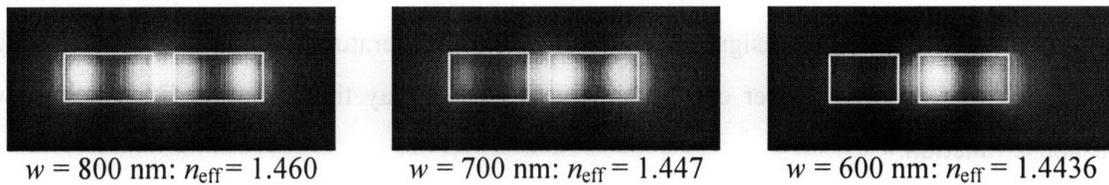
Figure 38. The intensity distribution and effective index of the 3rd mode in the bus-ring coupling region for several values of the bus waveguide width w. The third mode is cut-off for $w = 600$ nm.

Once the cross-section of the waveguides is chosen, we can fix the bend radii of the rings. Table 1 lists the radii of all 20 rings of the filter bank. The bend loss-limited Q factor of the rings changes from 340K to 410K, and the FSR from 2.04 to 2.06 THz. The radius step for 80GHz channel spacing between the filters is 5.7nm. In fabrication, it is hard to ensure that the step is exactly equal to the specified value so that 80 GHz channel spacing cannot be expected; furthermore, filter-to-filter variation of waveguide height and width can lead to even larger error in channel frequencies. To overcome this problem, a precise fabrication technique was developed by C. W. Holzwarth *et. al.* [12]; a calibration chip is first fabricated and characterized, and then the device chip is fabricated with precise dose control allowing sum-nanometer accuracy in waveguide width. Channel spacing uncertainty was reduced to

about 6GHz with this method without using heaters or similar post-fabrication tuning techniques [12].

No.	Frequency, THz	Center Radius, μm	No.	Frequency, THz	Center Radius, μm
1.	193.30	10.113	11.	194.10	10.055
2.	193.38	10.107	12.	194.18	10.049
3.	193.46	10.101	13.	194.26	10.044
4.	193.54	10.095	14.	194.34	10.038
5.	193.62	10.090	15.	194.42	10.032
6.	193.70	10.084	16.	194.50	10.026
7.	193.78	10.078	17.	194.58	10.021
8.	193.86	10.072	18.	194.66	10.015
9.	193.94	10.067	19.	194.74	10.009
10.	194.02	10.061	20.	194.82	10.004

Table 1. Center radii of the rings of the 20-channel filter bank.

The theory of the filter design is well-described in the literature [30, 29, 34]. The coupling coefficients of an n-ring filter can be chosen in such a way that the filter has Chebyshev response function:

$$T(f) = \frac{1}{1 + \varepsilon^2 T_n^2(f/\Delta f)},$$

where T_n is Chebyshev polynomial of n-th order, Δf is the bandwidth of the Chebyshev function, and ε is the ripple factor. The order of the function determines how quickly the function rolls off. We need to have a filter with 25 GHz bandwidth rolling off by at least 30dB at 80 GHs from the center frequency. Second order filter is sufficient to meet these specifications. We will also design a third-order filter because it rolls off much faster and thus leads to a much smaller crosstalk with adjacent channels. Having fixed the Chebyshev bandwidth Δf and the ripple factor ε, we can find the ring-bus and ring-ring coupling coefficients [34]. The coupling coefficients and characteristics of the obtained response functions of two- and three-ring filters are given in Table 2. The "lossless performance" data corresponds zero propagation loss in the rings. In practice, SiN waveguides will have some loss; it was estimated to be about 10 dB/cm from previous fabrication experience[1]. The

---

[1] According to Charles W. Holzwarth.

performance of the filters assuming 10 dB/cm propagation loss is also given in Table 2. The transfer functions of the filters are shown in Fig. 39. Note that the through port extinction ratio is not very large. This would be a problem in such applications as optical add-drop multiplexers; in our case, however, we are using only the drop port, the extinction of the through port is not important.

2-ring filter design	3-ring filter design
Power coupling coefficients: bus-ring $3.35 \cdot 10^{-2}$ ring-ring $7.60 \cdot 10^{-4}$	Power coupling coefficients: bus-ring $5.32 \cdot 10^{-2}$ ring-ring $7.62 \cdot 10^{-4}$
Lossless performance: Chebyshev bandwidth: 20 GHz 3-dB bandwidth: 24 GHz Adjacent channel (center): -36 dB Ripple: 0.97 dB	Lossless performance: Chebyshev bandwidth: 21 GHz 3-dB bandwidth: 27 GHz Adjacent channel (center): -50 dB Ripple: 0.17 dB
Performance for 10dB/cm loss: 3-dB bandwidth: 25 GHz Adjacent channel (center): -33 dB Ripple: 0.07dB Drop loss: 3.1 dB	Performance for 10dB/cm loss: 3-dB bandwidth: 25 GHz Adjacent channel (center): -47 dB Ripple: none Drop loss: 3.4 dB

Table 2. The coupling coefficients and some parameters of the response function of of two- and three-ring filters.
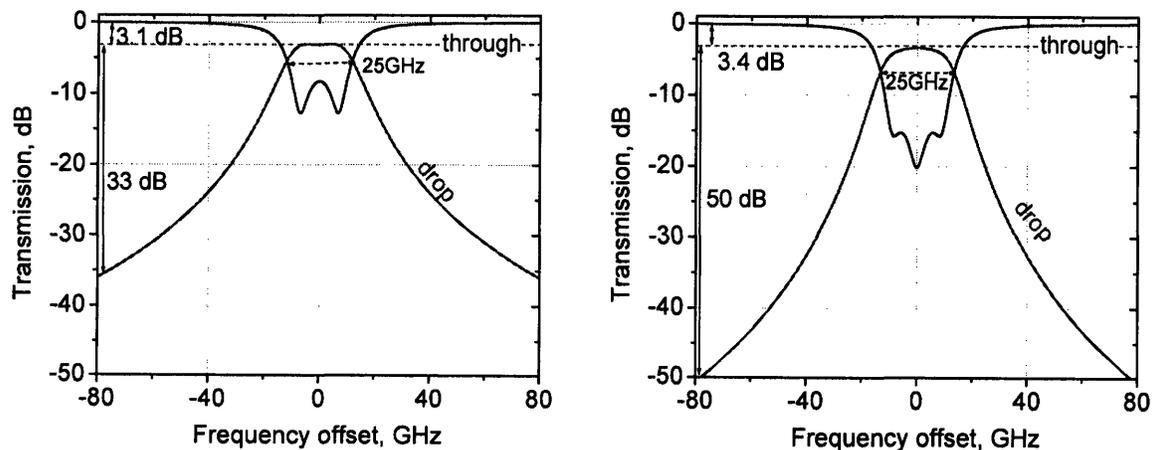
Figure 39. Simulated drop and through port response of the designed (a) second order ring filter, (b) third order ring filter.

The final step in the filter design is determining the ring-bus and ring-ring gaps corresponding to the desired coupling coefficients. The coupling coefficients as a function of gap width are computed using three-dimensional finite-difference time-domain (FTDT) simulation[1] which solves Maxwell's equations in space and time numerically without any approximation other than discretization of time- and space-domains [37]. The input field is the waveguide mode calculated with the mode solver; the coupling coefficients are found by finding overlaps of the fields at the output of the bus and the ring with corresponding modes[2].

The calculated coupling coefficients as a function of gap size are shown in Fig. 40. The gap sizes necessary to obtain necessary bus-ring and ring-ring coupling coefficients from Table 2 are listed in Table 3 for all 20 filters. The coupling coefficients change from filter to filter because of some dependence of the coupling coefficients on wavelength. The values of gaps in Table 3 are given so precisely more for consistency rather than for any practical reason, because in fabrication the gaps could not be controlled with precision better than 5nm.
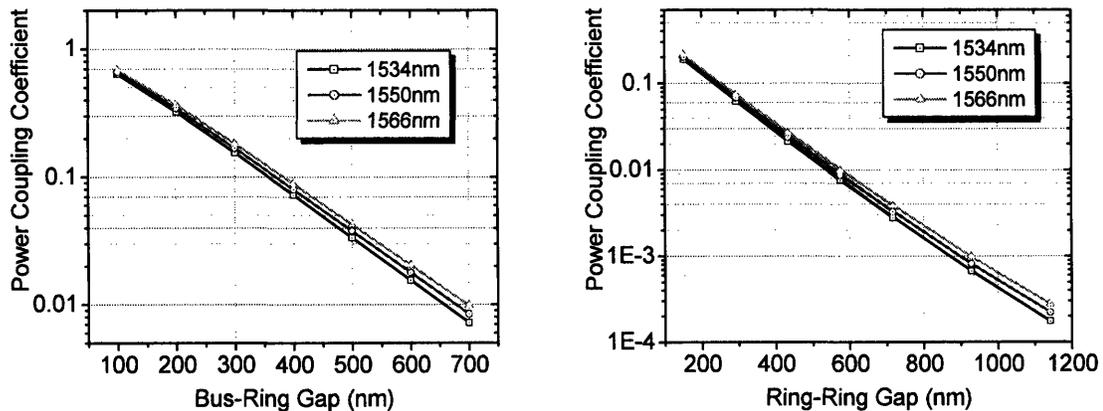
Figure 40. (a) Ring-bus, (b) ring-ring coupling coefficient versus coupling gap for several wavelengths obtained from 3d FDTD simulation.

The final design parameters for the two- and three-ring filters are depicted in Fig. 41.

---

[1] FDTD codes were written in Fortran by the former and present members of the group: Christina Manolatou [38, 39], Mike Watts, and Milos Popovic [30]. Simulations were run on an SGI supercomputer with 8 Ithanium processors and 32Gb of memory.

[2] The Matlab code providing FDTD input field and calculating coupling coefficients was written by Milos Popovic.

Two-ring filter design				Three-ring filter design		
No.	f_{center}, THz	ring-bus gap, nm	ring-ring gap, nm	No.	f_{center}, THz	ring-bus gap, nm	ring-ring gap, nm
1.	193.30	517.3	942.4	1.	193.30	455.7	942.0
2.	193.38	516.6	941.1	2.	193.38	455.1	940.7
3.	193.46	516.0	939.9	3.	193.46	454.5	939.4
4.	193.54	515.3	938.7	4.	193.54	453.9	938.2
5.	193.62	514.7	937.5	5.	193.62	453.3	937.0
6.	193.70	514.0	936.3	6.	193.70	452.7	935.9
7.	193.78	513.4	935.1	7.	193.78	452.1	934.6
8.	193.86	512.7	933.9	8.	193.86	451.5	933.5
9.	193.94	512.1	932.7	9.	193.94	450.9	932.3
10.	194.02	511.4	931.5	10.	194.02	450.4	931.1
11.	194.10	510.8	930.3	11.	194.10	449.8	929.9
12.	194.18	510.1	929.1	12.	194.18	449.2	928.7
13.	194.26	509.5	927.9	13.	194.26	448.6	927.5
14.	194.34	508.8	926.7	14.	194.34	448.0	926.3
15.	194.42	508.2	925.5	15.	194.42	447.4	925.1
16.	194.50	507.5	924.3	16.	194.50	446.8	923.9
17.	194.58	506.9	923.1	17.	194.58	446.2	922.7
18.	194.66	506.2	922.0	18.	194.66	445.6	921.5
19.	194.74	505.6	920.8	19.	194.74	445.1	920.3
20.	194.82	504.9	919.6	20.	194.82	444.5	919.1

Table 3. Bus-ring and ring-ring gaps for two and three ring filter designs.
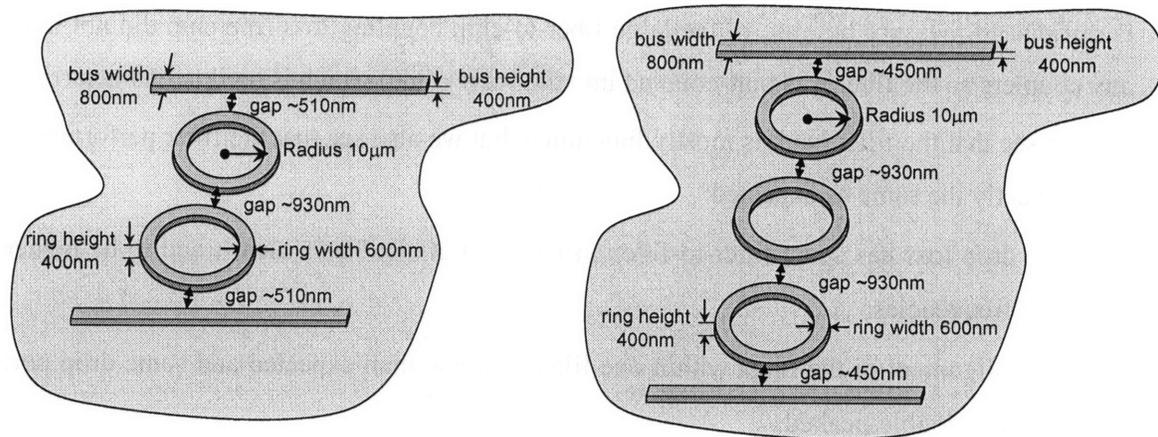
Figure 41. Final parameters of the two- and three-ring filter designs.

11. Experimental Results for the EPIC Filter Bank

The dual 20-channel two-ring filter bank was fabricated by Charles W. Holzwarth of MIT Microsystems Laboratory. The image of the fabricated chip is shown in Fig. 42 (the line in the middle of the image is because it was obtained by concatenating two images taken independently using a microscope). There is one heater on top of each filter, so that the center

frequencies of the filters can be adjusted. The two rings of the same filter did not contain individual heaters because the resonant frequencies of these rings were expected to be well-aligned [12].
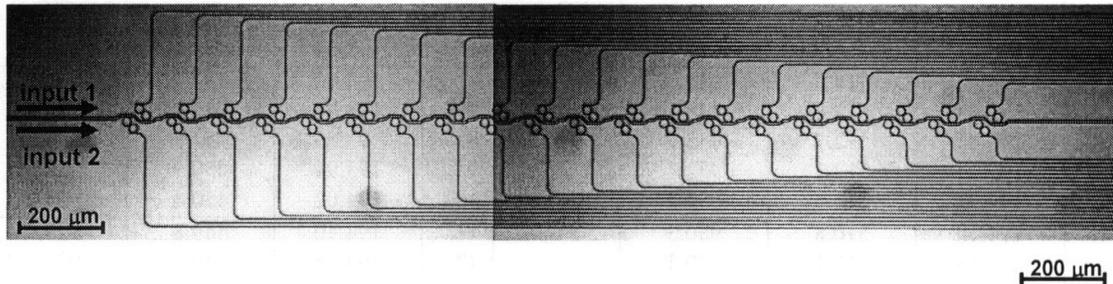
Figure 42. Image of the fabricated 20-channel filter bank (before the heaters were added)[1].

The drop responses of the 20 filters of the bottom filter bank are shown in Fig. 43[2]. The grey line on top represent several measurements of the through port transmission. Note that the drop loss shown in Fig. 43 (and other figures of this chapter) is relative; the absolute loss is difficult to measure because of the large fiber-to-chip coupling loss (the chip did not have any couplers so the fiber was butt-coupled into the waveguide) which is hard to characterize.

We see that the filter bank is mostly functional, but we also see that the filter performance is not exactly the same as expected:

• the drop loss has some filter-to-filter variations, but also has a clear trend to be higher for lower frequencies;

• the alignment of the rings within one filter is worse than expected and some drop port responses are double-peaked;

• the bandwidth of the filters and crosstalk between adjacent channels is higher than expected and varies from filter to filter.

---

[1] The image has been created by Charles W. Holzwarth.

[2] The measurements were done by Marcus Dahlem and Charles W. Holzwarth.
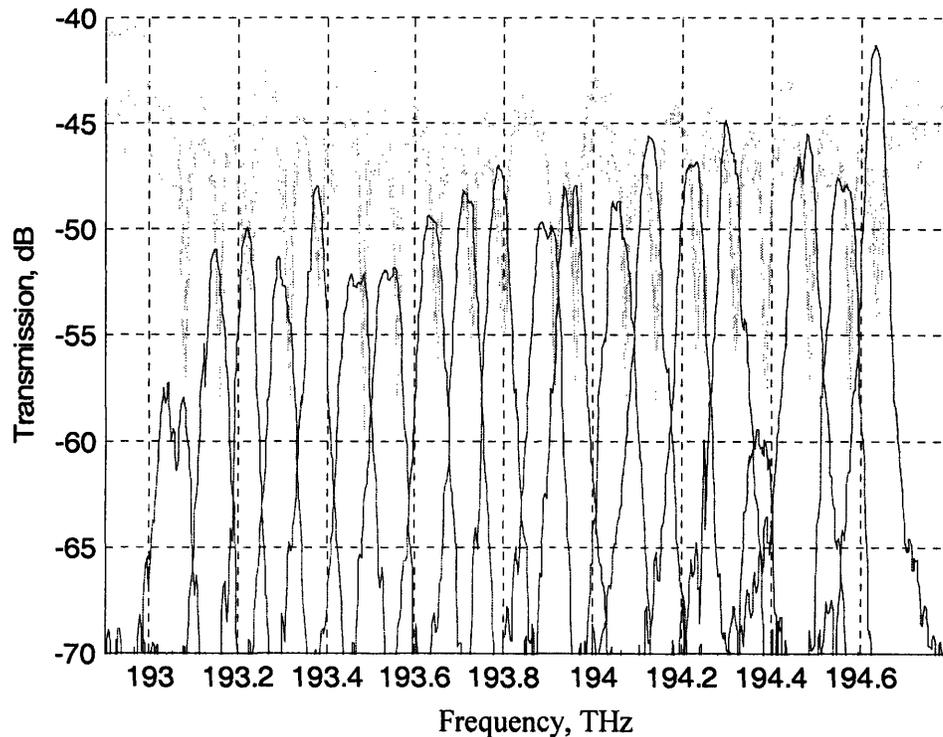
Figure 43. Measured relative transmission of the drop and through port of the 20-channel filter bank.

Let us try to understand how the observed deviations from the designed performance can be explained. First of all, let us plot the drop loss versus the channel number for the top and bottom filter banks. Fig. 44 shows the measured data points and a linear fit; the line has approximately 0.5 dB/filter slope both for the top and the bottom filter bank. The there is one channel in the top and one in the bottom filter bank with drop loss much higher than in the other channels; these high losses can be caused by bad quality of the facets of output waveguides after the chop cleaving, which leads to very inefficient coupling. These points were not considered in drop loss fitting.

The 0.5dB/filter trend can be explained by the losses at the junction of straight and bent waveguide sections. As it can be seen from Fig. 42, the waveguide going through all filters have many "wiggles", which were created in order to have each pair of filters from the top and bottom banks in the same 100μm x 100μm write field so that the fabricated rings are more similar to each other. The loss at the junction of straight and bent sections of the waveguide occurs because of the mode mismatch between the modes in the two sections. The

through waveguide has 12 such junctions per filter, and the loss was estimated to be 0.035-0.04dB/junction with 3d FDTD simulation, corresponding 0.42-0.5 dB of loss per filter. Before being dropped, the light in the n-th channel must pass (n - 1) filters which precede it along the through waveguide, therefore the 0.5 dB/filter trend we observe in the drop port. The solution of this problem is to increase the bend radius of the through waveguide or to avoid creating so many bends in it.

(a)   (b)
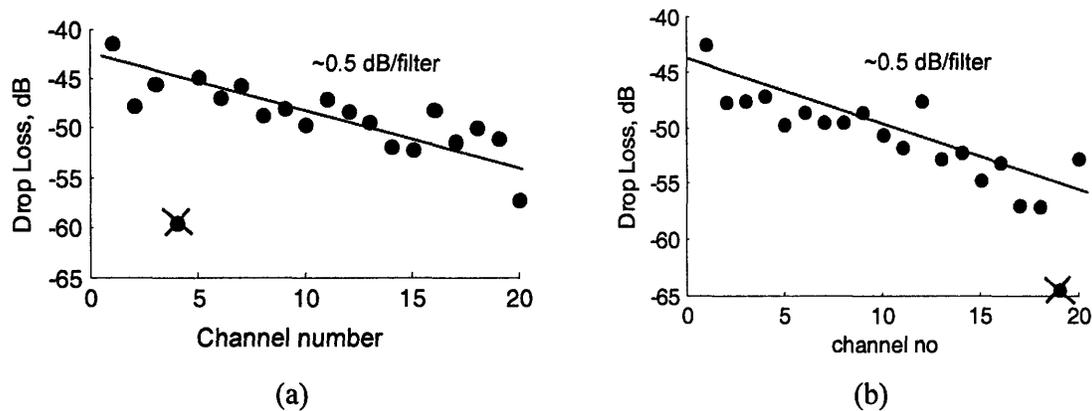
Figure 44. Measured drop loss versus channel number for (a) top and (b) bottom filter banks. The lines shows the linear fit with 0.5dB/filter slope. The crossed out points are thought to be caused by bad output facets and were not considered in fitting.

The next thing that can be noticed in the measured drop port responses is that the channel spacing is less uniform than expected. Fig. 45 shows histograms of channel spacing distribution for the top and the bottom filter banks. The average channel spacing is about 83GHz, close to the design value of 80GHz. However, there are also considerable spread of channel spacing around the average value, with channel spacing ranging from 56 to 107 GHz. Although the channels are not aligned as precisely as was expected [12], the obtained results are still excellent because the amount of frequency tuning needed to make the spacing uniform is small and thus the required tuning power is low.

Average = 83 GHz, RMS = 8.4GHz   Average = 83 GHz, RMS = 8.5GHz

Min. = 56GHz, Max. = 107GHz   Min. = 63GHz, Max. = 103GHz
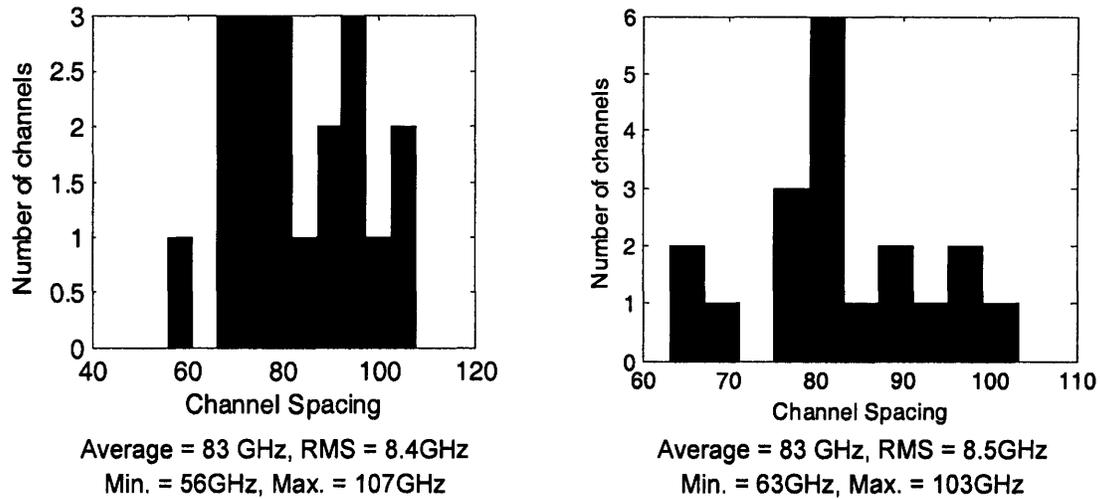
Figure 45. Histograms of channel spacing for (a) top and (b) bottom filter bank. The target channel spacing is 80GHz.

There is also some frequency misalignment between the rings of individual filters, as it can be seen from the broadening and sometimes splitting of the drop port responses. Fig. 46 shows the drop and through port responses of each filter in the bottom filter bank. The multiple lines for the through port response were obtained with different measurements of the same device. The x-axis is the frequency relative to the center of the drop port, and the y-axis is the relative transmission in dB. We see that the dips in the through port response are misaligned with respect to the center of the drop port response by some value $\Delta f_{misalign.}$, which is a sign that the there is frequency mismatch between the rings. The frequency mismatch be estimated as $2 \cdot \Delta f_{misalign.}$ and is shown in the middle of each plot. The average frequency mismatch is about 30 GHz.
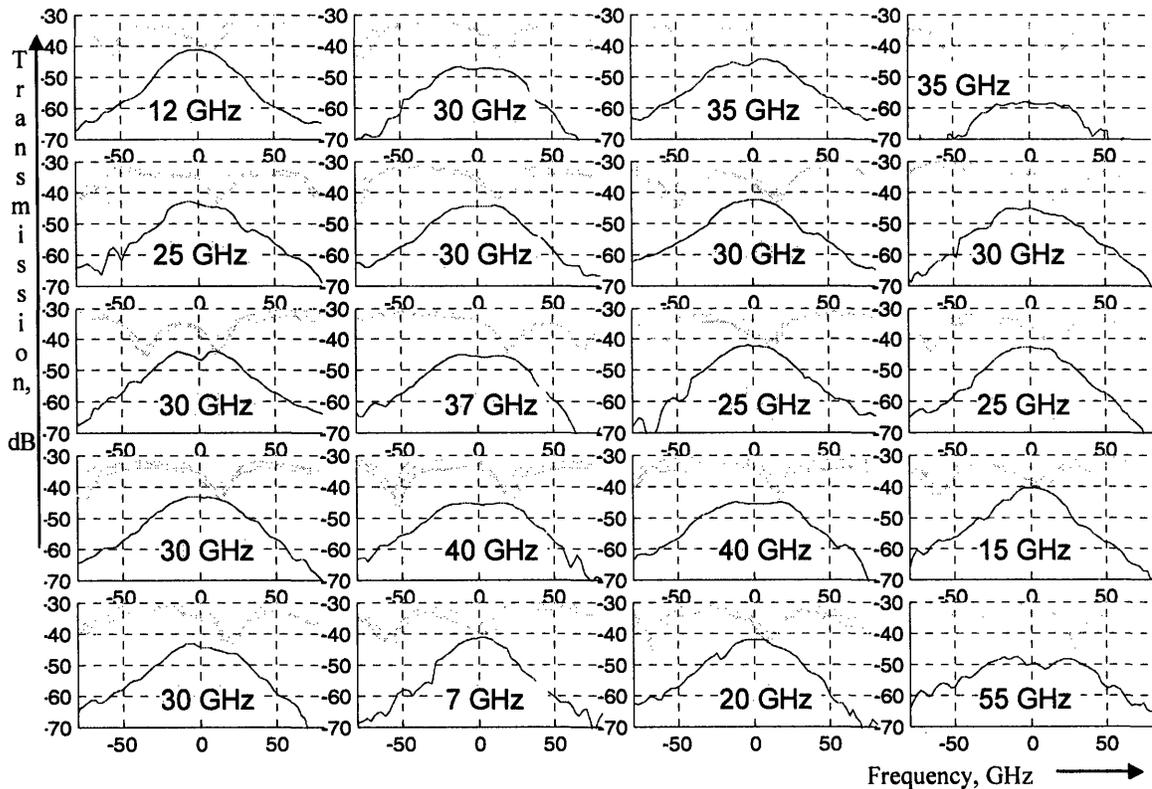
Figure 46. Drop and through port responses of the individual filters. Several lines for through port response correspond to different measurements. The horizontal axis is frequency, in GHz, and the vertical is relative transmission, in dB. The numbers on the middle of the plots show frequency mismatch between the rings of two-ring filters. The 0.5dB/filter slope was taken out from the measured drop port responses.

Note that the frequency mismatch between the rings can be eliminated by thermal tuning of one ring with respect to the other. The 20-channel filter bank, however, did not have separate heaters for each filter, so it was not possible to demonstrate it. To demonstrate the compensation of the frequency mismatch, an additional 2-channel filter bank was fabricated with separate heaters on to of each ring. Fig. 47 shows the measured drop and through port responses as the frequency misalignment is being reduced from about 40 GHz to below 0.5 GHz.
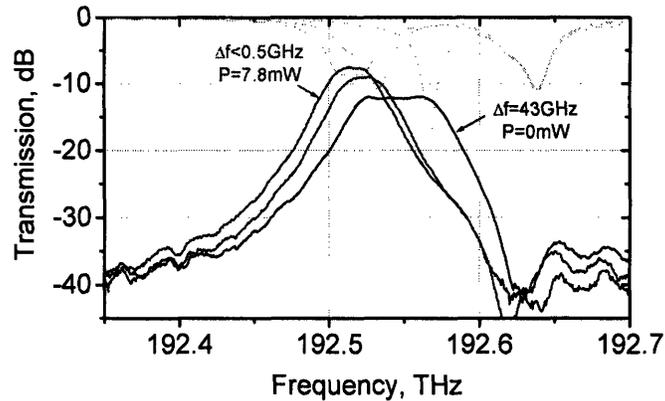
Figure 47. Elimination of the frequency mismatch between the rings of a two-ring filter by thermal tuning. Δf is the fitted frequency mismatch and P is the tuning power.

Let us now return to interpretation of measurement results for the 20-channel filter bank. It turns out that loss in the "wiggles" of the through waveguide and frequency mismatch between the rings is not sufficient to explain the observed responses shown in Fig. 46. The bandwidth of the filters is larger than expected, which can be caused either by high loss or by larger than expected coupling coefficients. To understand what is happening we can examine the depth of the dip in the through port response, illustrated in Fig. 48(a), as a function of the frequency mismatch between rings, Fig. 48(b). If the waveguide loss is about 10 dB/cm as it was expected, the dip depth versus frequency mismatch exhibits a completely different dependence than that observed in experiment. To fit the measured data it is sufficient to assume that the loss is 35-40 dB/cm and the coupling coefficients are about 20% higher than expected. Fig. 48 shows the simulated dip width if the propagation loss of 38dB/cm is assumed. Fig. 49 presents the individual drop and through port responses fitted

- 38dB/cm propagation loss;

- frequency mismatch between the rings as shown in Fig. 46;

- 20% larger than designed coupling coefficients.

The absolute transmission of the fitted drop port response was adjusted for each filter individually because of the unknown channel-to-channel variations in fiber coupling efficiency. The fitted through port, however was not adjusted for each filter. We see that the quality of the fits are good, which confirms the conclusion about high propagation loss.
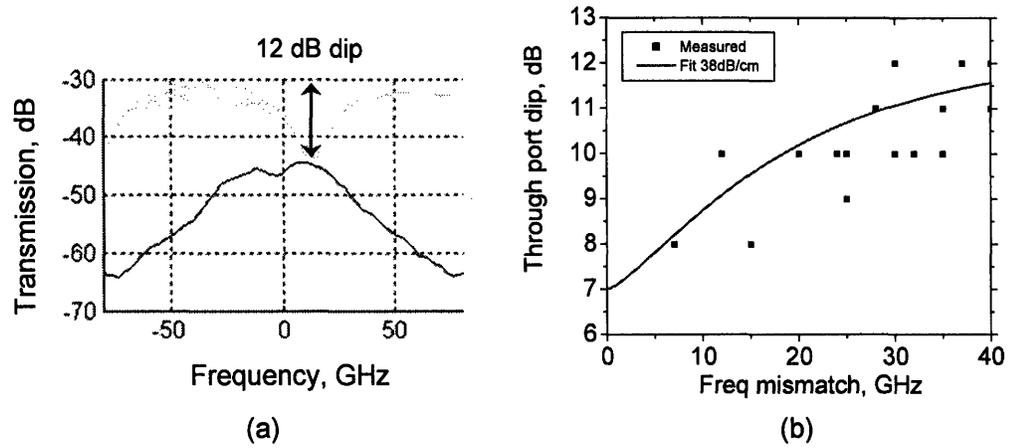
Figure 48. (a) An example of the filter through port response with 12 dB dip depth; (b) through port dip depths versus frequency mismatch between rings, points correspond to the measured data and the line obtained from simulation with 38dB/cm waveguide propagation loss.
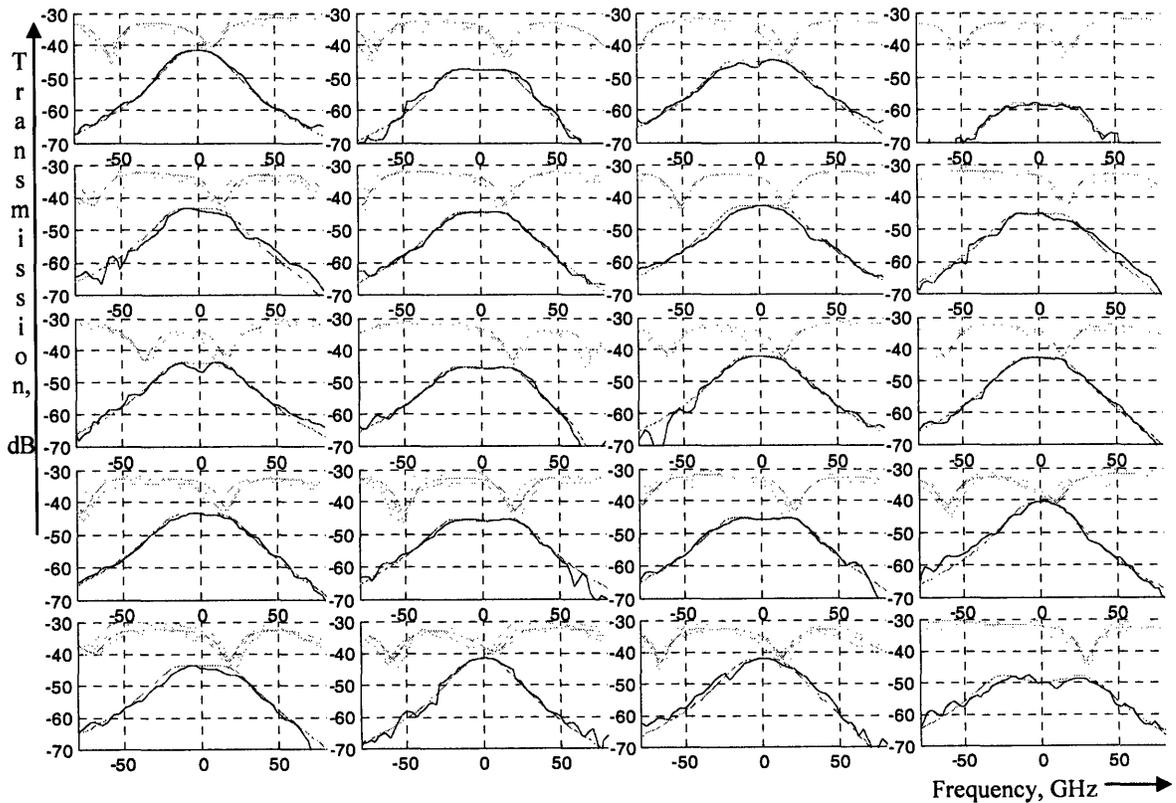
Figure 49. Drop and through port responses of 20 filters of the bottom filter bank. Blue lines represent measured data and red lines – fitted responses assuming 38dB/cm propagation loss.

IV. HORIZONTAL FIBER-TO-CHIP COUPLERS

12. Overview of Horizontal Fiber-to-Chip Couplers

The problem of coupling of light from a fiber to a high-index-contrast waveguide stems from the large mode mismatch between the fiber and waveguide mode, as it was outlined in chapter 2. Several solutions to the coupling problem were proposed in the literature. The most practical of them can be divided into two categories: inverse taper-based horizontal couplers and grating-based vertical couplers.

The idea to use inversely tapered waveguide was first proposed by Shani *et. al.* in 1989 in the context of coupling light generated by a semiconductor laser into a fiber [40]. In this approach, the waveguide is gradually narrow down in so that its mode becomes less and less confined and adiabatically expands until it matches the size of the fiber mode. The concept of adiabatic transitions is explained in more detail in chapter 13. The coupler we design for the EPIC ADC system is based in the inverse taper, therefore more details on the operation of such a coupler can be found in chapter 14 (and particularly in Fig. 59 which illustrates the mode evolution along the tapered waveguide). The overcladding shown in Fig. 50 serves to confine the mode when it expands from the narrowed SiN waveguide. The size of the overlcadding should match the size of the fiber to obtain efficient coupling.
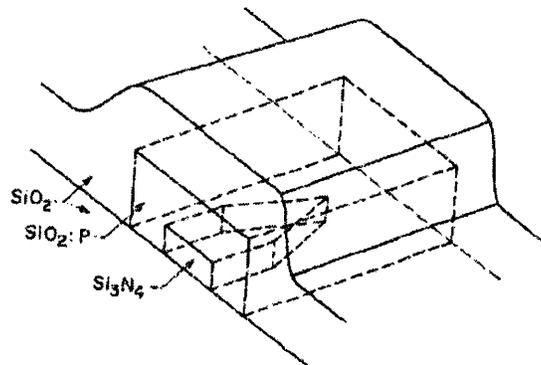
Figure 50. Inverse taper used by Shani *et. al.* for coupling light from a semiconductor laser into a fiber [40].

In Shani's work, the size of the overcladding was 4x7μm and the size of the waveguide was 0.1x2μm, which is higher than the size of a typical waveguide we are interested in. The total coupling loss was measured to be 3.1dB, consisting of 0.5 dB of propagation loss in the taper, 0.8 dB mode mismatch loss between the overcladding and the fiber, 0.8 dB due to the separation between the laser and the waveguide, and 0.8dB due to sidewall scattering in the laser.

The idea was picked up by the researchers at NTT [41, 42, 43] and IBM [44] and applied for coupling light into submicron Si waveguides. Fig. 51 shows a 200μm-long inverse coupler demonstrated by Shoji et. al. in 2002 [41]. The coupling loss was 0.8dB with polarization-dependent loss of 0.5dB. The coupling loss was found by subtracting the propagation loss of 2.0 dB from the total loss of 3.5dB. The propagation loss was measured with low-coherence reflectometry technique. The coupler works with small-core fibers; the mode field diameter of the fiber used in the experiment was 4.3μm. The claimed result of 0.8dB conversion loss is impressive, although it is necessary to keep in mind that this loss does not include the scattering losses due to roughness of the taper. These losses are expected to be higher in the taper than in a normal waveguide because the field intensity of a weakly confined mode of the taper at sidewalls is higher than for a strongly-confined wide waveguide. Therefore, the taper mode "sees' more of the sidewall roughness and is expected to have increased loss as compared to normal waveguide. Another comment to this result is that it is not clear how such a good mode matching was achieved between the 4.3μm fiber and 3μm polymer waveguide.
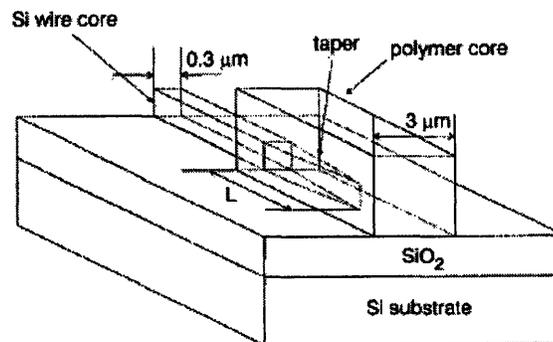
Figure 51. Horizontal coupler by Shoji *et. al.* [41].

Another result from the same group was published in [42, 43]. The device is the same, except an additional low-index polymer covering was added on top of the polymer waveguide, as shown in Fig. 52; the other parameters were the same as in the previous work. The loss for coupling into a 4.3μm fiber was 0.5dB, and for coupling into a 9μm SMF fiber – 2.5 dB.
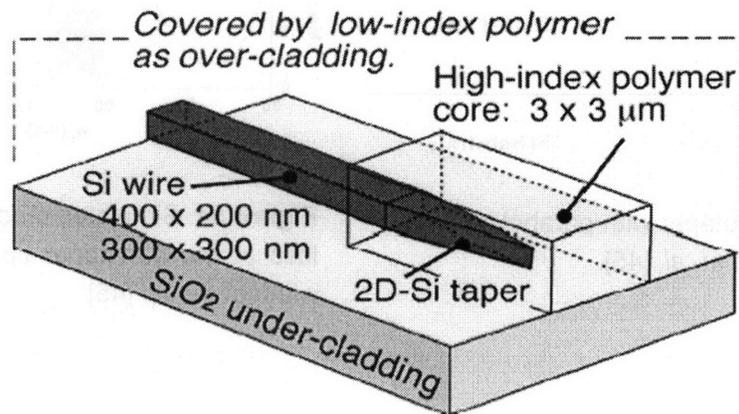
Figure 52. Another design of horizontal coupler from NTT researchers [42, 43].

The coupler demonstrated at IBM uses the same idea of an inverse taper overcladded with a polymer waveguide [44]. The polymer size was 2x2μm and the spot diameter at the focus of the lensed fiber was 2.1μm. The experimental results give 0.5±0.4dB per coupler for TE mode. The loss for TM mode was about 1μm even though the coupler was optimized for TE mode only. Note that the cited coupler loss was calculated by subtracting the waveguide loss from the total loss. The waveguide loss was estimated to be 3.5±2 dB/cm, so the precision of coupler loss estimation is not very high.

An coupler based on inverse taper without an overlaying polymer waveguide was implemented by M. Lipson group at Cornell [45], see Fig. 53. The profile of the taper was parabolic, which allowed to reduce taper length to about 40 μm. The simulated loss at the tip of the taper as a function of taper width is shown in Fig. 54. It can be seen that the region of low loss are different for TE and TM modes, which means that the coupler will be efficient for just one of them. Apart from loss at the tip, the loss occurs during the mode conversion in the taper; it was calculated to be 0.25dB. The measurement results gave 3.3dB coupling loss for TM and 6.0dB loss for TE modes for coupling from a lensed fiber with mode field diameter of about 5μm.
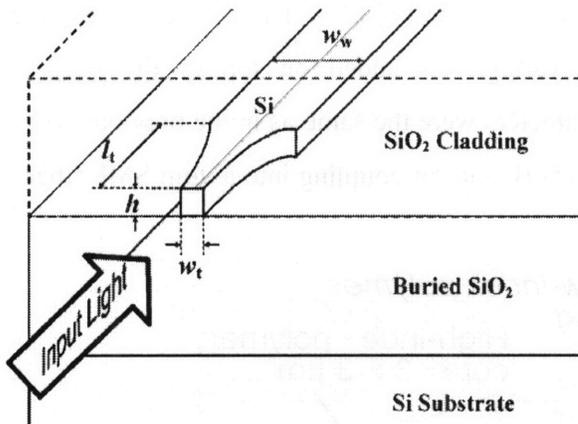
Figure 53. Nanotaper with parabolic profile from V. Almeida *et. al.* [45].
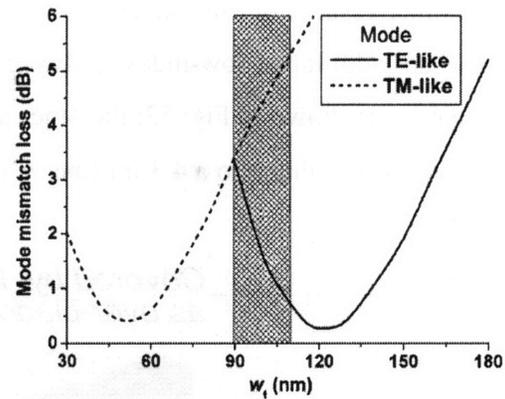
Figure 54. Simulated mode mismatch loss at the nanotaper tip versus the width of the tip [45].

Another important kind of fiber-to-chip couplers are vertical grating-based couplers, with important works in this field produced by the group at Ghent university [46, 47] and, recently, by M. Fan and M. Popovic at MIT [48].
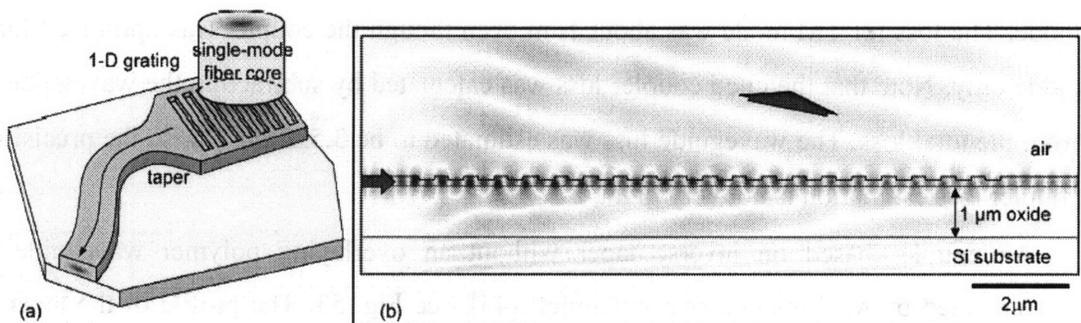
Figure 55. Vertical fiber-to-chip coupler by W. Bogaerts *et. al.* [46]

Fig. 55 illustrates a grating-based vertical coupler design [46]. The calculated efficiency is quoted to be 30%. The calculated efficiency increases to about 70% if an optimized multi-layer SOI stack with an additional first-order grating is used. The measured losses were 6.8 dB. Another way proposed to increase the efficiency coupling was to add a gold mirror at the bottom of the grating region [47]. Due to symmetry of the structure, the light entering the grating region will radiate both upward and downward leading to large loss; the golden

mirror is intended to break the symmetry and prevent downward radiation. This allowed to increase the coupling efficiency to 69% (1.6dB loss).

An alternative approach to break the symmetry of the structure was proposed by M. Fan and M. Popovic [48]. They created a two-layer grating which is inherently asymmetric and achieved suppression of downward radiation. The calculated coupling efficiency was 74% (1.3dB loss)
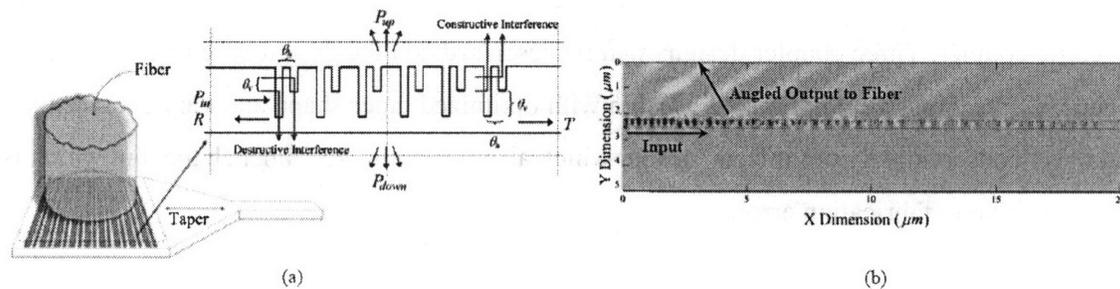
Figure 56. (a) The design of a two-layer grating and (b) field distribution showing scattering in upward direction only [48].

The last mode conversion approach we would like to mention uses tapered ridge waveguide to reduce the vertical extent of the mode [49, 50]. This is illustrated in Fig. 57(a) which shows the rib width reduction which "pushes" the mode downward into the wider section of the waveguide with a much thinner rib. We will use a similar idea when constructing a two-stage fiber-to-chip coupler in chapter 0.
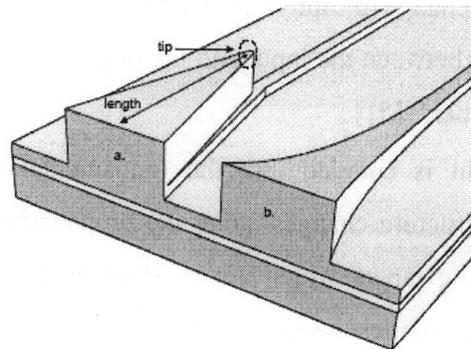
Figure 57. Tapered rib waveguide for reducing vertical extent of the mode [49].

13. Modeling and Optimization of Adiabatic Couplers

In ADC being build within the EPIC project, the light from an external mode-locked laser must be coupled to and from an integrated optical chip. It very important to have an efficient broadband fiber-to-chip coupler because of the power budget constraints. The coupler idea based on the inverse taper overcladded with the polymer was chosen because it potentially offers high coupling efficiency with large bandwidth. This section describes the design of such a coupler, starting with theoretical background and proceeding to optimization of our specific coupler. Three coupler designs will be described: (1) the design based on linear taper shape, (2) "aggressively optimized" design with optimized taper shape and very short length, and (3) "conservatively optimized" design which also has optimized taper shape, but which is less sensitive to fabrication errors.

1. Theory: Optimization of taper shape

The cross-section of a tapered waveguide is changing along the light propagation direction, therefore such a structure does not support optical modes. However, it is useful to consider the local modes, which are defined for some position z along the taper as the modes which would exist if the waveguide cross-section at position z would stay the same along all the taper. The local modes do not exist in reality, but they are a useful tool for understanding and simulating slowly-varying optical structures.

The adiabaticity principle can be stated as:

"If the medium characteristics of a tapered electromagnetic waveguide are such that the local normal modes do not change abruptly throughout the taper, adiabatic transformation of each local mode is possible between the input and the output of the tapered waveguide in the limit of an infinitely long taper." [51]

In other words, if light is coupled into the fundamental mode, it will stay in the fundamental mode if the structure changes along the propagation direction slowly enough. We are interested in making tapers which are as short as possible, so the goal of taper optimization is to find the smallest taper length over which the mode transformation is approximately adiabatic so that most of the power stays in the same mode and coupling to other modes is minimized. We are also interested in optimizing the taper shape to on order to obtain lowest power losses in smallest length.

.

The theory for calculating the mode conversion loss was developed by Snyder in [52] and applied to taper shape optimization in [51]. Mode conversion loss of i-th mode $\overline{\eta}_i$, which is the amount of power which does not stay in the original mode but is coupled to other modes along the taper can be written in first-order approximation as

$$\overline{\eta} \cong \sum_{j \neq i} \left| \int_0^1 \hat{C}_{ji}(s) \exp\left\{ jL \int_0^s [\beta_j(s') - \beta_i(s')]ds' \right\} ds \right|^2 , \tag{13.1}$$

where the summation is over all modes except i-th mode we are considering, L is the length of the taper, β_j is the position-dependent propagation constant of the j-th local mode, s is the relative position along the taper so that $z = sL$, and $\hat{C}_{ji}(s)$ is the local coupling coefficient between the modes caused by structure variation. The local coupling coefficient can be expressed as

$$\hat{C}_{ji}(s) = \frac{\omega \varepsilon_0}{4[\beta_j(s) - \beta_i(s)]} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dn^2(x,y,sL)}{ds} \vec{E}_j^*(s) \cdot \vec{E}_i(s) dxdy , \tag{13.2}$$

where $n(x,y,sL)$ is the cross-section index distribution which depends on relative position s along the taper and $\vec{E}_j$ is the electrical field distribution in j-th local modes. Equations (13.1), (13.2) can be rewritten in the following form [51]:

$$\overline{\eta} \cong \sum_{j \neq i} \left| F_{ji}(L) \right|^2 , \tag{13.3}$$

$$F_{ji}(L) \equiv \int_0^1 T_{ji}(u) \exp(juL) du , \tag{13.4}$$

$$T_{ji}(u) \equiv \left[\frac{\hat{C}_{ji}(s)}{\beta_j(s) - \beta_i(s)} \right]_{s = f_{ji}^{-1}(u)} , \tag{13.5}$$

$$u = f_{ji}(s) \equiv \int_0^s [\beta_j(s') - \beta_1(s')]ds' . \tag{13.6}$$

Equation (13.4) states that the coupling coefficient to a higher mode j is defines by a Fourier transform of the characteristic taper function $T_{ji}(u)$. This function is proportional to the rate of change of the waveguide cross-section at a given position and inversely proportional to the mismatch of propagation constants of the two modes. Variable u defined by Eq. (13.6) describes the phase difference between the two coupled modes accumulated starting from the

beginning of the taper until the position s. Equation (13.3) simply says that the total loss in i-th mode is a sum of losses into all other modes.

Equations (13.3)-(13.6) provide not only a way to find the mode conversion loss within given taper of finite length and shape, but also a way to optimize the taper shape in order to increase conversion efficiency. We are interested in minimizing $F_{ji}(L)$, the coupling coefficient to other modes (13.4), as a function of length. For doing this, we can optimize $T_{ji}(u)$, the taper characteristic function (13.5), which we can control by choosing a taper with slower or faster rate of shape variation at given position L. For linear taper shape, the characteristic function usually has a large peak at some position u and is relatively small elsewhere. The Fourier transform of such function will decay slowly with L, requiring a long taper for small conversion to higher-order modes. The common choice for optimized taper shape is to set the characteristic function $T_{ji}(u)$ to constant, then its Fourier transform will decay much faster and we can expect low conversion loss in shorter length. However, we can do even better by choosing other taper functions, such as super-Gaussian and Hamming functions, whose Fourier transform decays faster than that of the constant. Fig. 58 illustrates the mode conversion loss in an example of taper for several choices of the characteristic windowing functions. The linear taper shape is clearly worse than the others, and the taper with Hamming characteristic function outperforms tapers with constant and super-Gaussian characteristic functions and allows to achieve mode conversion losses below 10^{-4} dB for much shorter taper lengths than the rest. It is necessary, however, to note the following.

• In reality it doesn't make much sense to design a taper with 10^{-4} dB mode conversion loss, because the losses due to other factors such as fabrication errors will be larger anyway. If we are interested in losses of about 1%, the taper with super-gaussian shape actually outperforms the one based on Hamming function. Therefore it makes sense to look not for tapers with loss $F_{ji}(L)$ decaying to the smallest values, but for tapers with $F_{ji}(L)$ decaying to the loss value of interest (such as 1%) in the shortest length L.

• The integral of the taper characteristic function remains constant for different taper shapes because the input and output waveguide cross-sections of the taper are fixed. Therefore, if we are choosing a characteristic function which smoothly decays to zero at the edges of the interval, such a function must be larger in the middle of the interval to conserve the total area under the curve. This is different from the traditional optimization of the

windowing function, where the area under the curve is not constrained. Therefore, in our case we should expect smaller impact from using different characteristic function than in the theory of windowing.
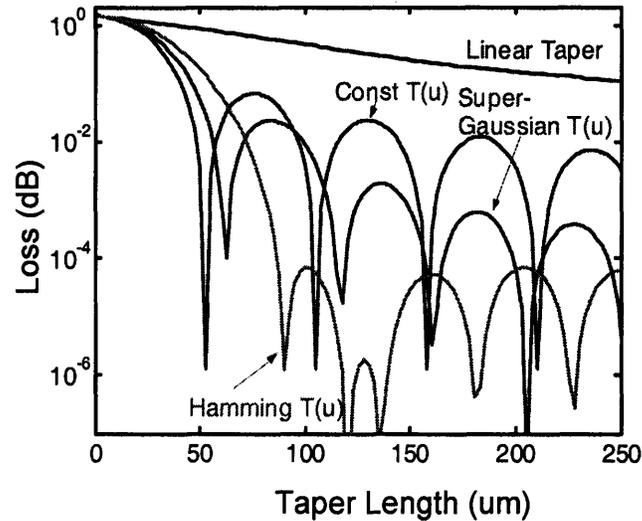
Figure 58. Mode conversion loss versus taper length for taper shapes corresponding to different characteristic functions calculated with Eqs. (13.3)-(13.6).

The taper shape optimization algorithm described above allows to design short and efficient tapers. This approach, however, has some limitations.

• Taper shape can be optimized based on coupling into just one of the higher-order modes rather than on coupling into all modes simultaneously, as it takes place in reality. As mentioned in [51], it is common that coupling to just one of the modes dominates, in which case this approach is expected to provide good results. In reality, however, coupling to other higher-order modes will also take place especially for short tapers, and the loss curves will look different than shown in Fig. 58. For example, the first sharp minimum of the loss $F(L)$ occurring at about 55μm for constant characteristic function relies on destructive interference from light scattered into the higher-order mode at different positions along the taper. If other higher-order modes are taken into account, the phase cancellation for them will not take place at the same length, so in reality the sharp dip in loss at 55μm will not be there because of coupling to other modes. Therefore, it doesn't make much sense to thoroughly optimize the taper shape based on coupling to just one mode because coupling to other modes will change

this result anyway. The performance of the taper optimized for one mode should be verified with methods which consider coupling to other higher-order modes as well.

• If the taper performance is improved by optimizing it for TE mode, it does not necessarily mean that the taper performance will be equally improved for TM mode as well. In strongly confined photonic structures the TE and TM modes usually have very different propagation constants, so optimization for one of them does not mean equally good optimization for another. The optimized taper can also become less broadband compared to a longer linear taper.

• The performance of very short optimized tapers relies on exact phase cancellation of light scattered at different positions along the taper. It becomes therefore sensitive to fabrication errors and variation in material parameters. For example, if the refractive index of some material turns out to be different from the design value by 0.01, this will change the phase distribution along the taper and might eliminate the phase cancellation and thus severely degrade the taper performance.

2. Theory: Simulation of light propagation along a tapered waveguide

Once the taper is designed, it is necessary to be able to evaluate its performance numerically. Although the FDTD method is the most straightforward and accurate approach, its use in practice is problematic because the simulation region is too large for FDTD. An alternative approach is the mode overlap method [53].

In mode overlap method, the taper is split into multiple cross-sections and local modes are found for each of them. At the input of the taper we usually have either one excited mode, or the excitation coefficients of the modes can be calculated from mode overlaps. Once we know the mode excitation coefficient in one cross-section, we can propagate this modes until the next cross-section by multiplying each mode coefficient by its corresponding phase term. At the new cross-section, the field can be expanded in terms of the new modes; the excitation coefficients of the modes in the new cross-sections are found from the boundary conditions for electric and magnetic fields. We can step this way through all the taper and find the total transfer matrix of the taper is a product of transfer matrices at each step. This approach is

implemented in commercially available software FimmProp from Photon Design[1], which was used for carrying out taper simulations.

In the next chapter we describe the application of the algorithms described above to the design of a coupler for the EPIC ADC system.

14. Horizontal Coupler Designs for EPIC ADC

We now proceed to applying the design algorithm described in the beginning of this chapter to building an efficient fiber-to-chip coupler for the EPIC ADC system. The coupler is based on the inverse taper concept [41, 42, 43, 44] and is shown in Fig. 59. As explained in chapter 12, the light from the fiber is first coupled into the polymer waveguide, and then transferred into the silicon nitride (SiN) waveguide with an inverse taper. The mode evolution along the taper is shown on the right of Fig. 59.

1. Thickness of overcladding layer

Let us start with optimizing the dimensions of the waveguide layers taking into account specific project and fabrication constraints.
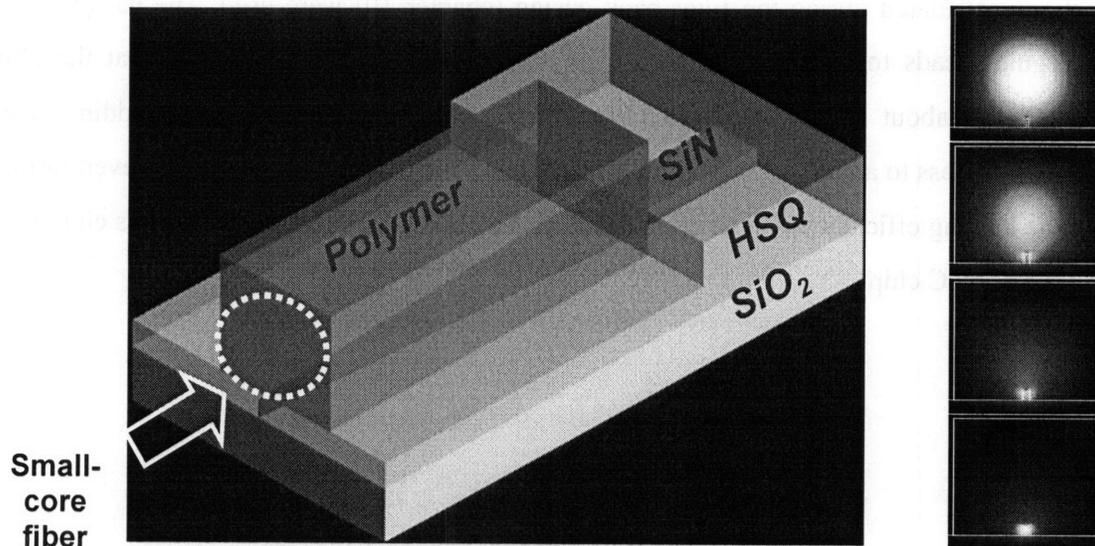
Figure 59. Fiber-to-chip coupler based on inverse taper. The mode evolution along the taper is illustrated on the right.

---

[1] http://www.photond.com/

The couplers are intended to be used for the ADC chip which also contains ring filter bank. Metallic heaters are used for ring filters in order to fine-tune and stabilize the ring center frequency. In order for the metallic heaters not to cause a lot of optical loss, they must be spatially separated from the waveguide core by an additional layer, as shown in Figs. 59, 60. This layer is made of hydrogen silsesquioxane (HSQ), which has refractive index of about 1.44. Let us find the thickness of the HSQ layer which is sufficient for isolating the optical mode from the metal heaters.

Metal heater

H_{HSQ} HSQ, n = 1.44

SiN, n = 2.2

SiO$_2$, n = 1.45
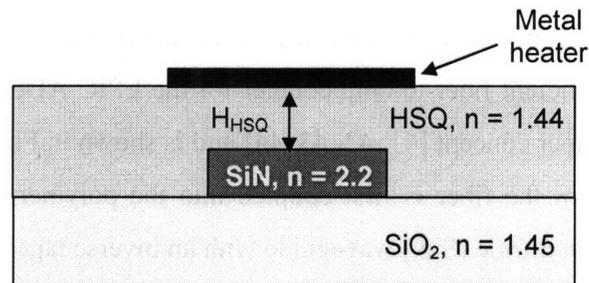
Figure 60. Cross-section of the waveguide layers with metal heaters on top.

Fig. 61 shows the calculated loss in dB/cm due to absorption in the metal layer versus HSQ thickness. The waveguide core is made of silicon-rich nitride and the core cross-sections calculated during the filter bank design (chapter 10) were used. The narrower bus waveguide leads to higher loss because its mode is less confined. We see that the HSQ thickness of about 1.5μm (which is 1.9μm if measured from the SiO$_2$ underladding layer) reduces the loss to an almost negligible value. Thicker HSQ will reduce the loss even further, but the heating efficiency will be compromised. The HSQ thickness of 1.5μm was chosen for the EPIC ADC chip.
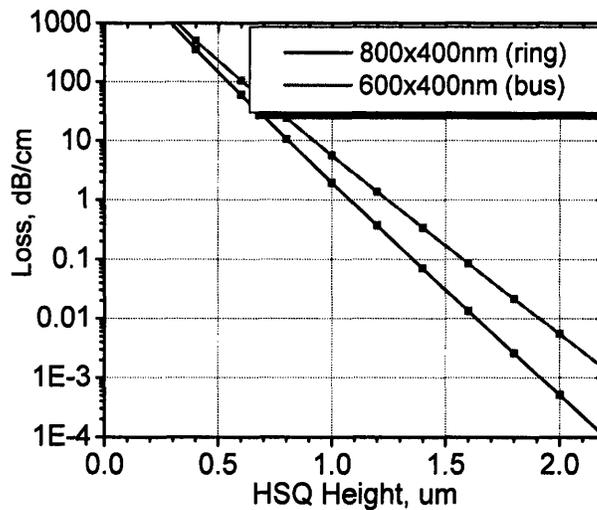
Figure 61. Loss due to absorption on metal versus the thickness of HSQ overcladding layer on to of the SiN waveguide.

2. Cross-section of the polymer waveguide

The light from the fiber is first coupled into a big waveguide made of polymer. The polymer was chosen as the material because of convenience in fabrication and wide variety of available polymers. The polymer chosen for the EPIC coupler was cyclotene with refractive index of 1.53.

To ensure efficient coupling, the width and height of the polymer waveguide should be chosen in such a way as to minimize mismatch between the fiber and polymer waveguide mode. We will assume that the light is coupled from a small-core high NA fiber because it requires smaller polymer waveguide for mode matching. Smaller polymer waveguide makes it easier to convert the mode into the SiN waveguide; thinner polymer layer is also easier to fabricate. The problem of coupling light from a standard single-mode fiber is considered in chapter 0.

The loss at the interface between the fiber and the polymer waveguide as a function of the waveguide width and height is shown in Fig. 62. The loss was calculated with the mode overlap method. The fiber was assumed to have mode field diameter of 4μm. We see that the optimum waveguide cross-section is about 5.3x4.5μm, which is larger than the diameter of the fiber mode. This can be explained by the fact that the polymer waveguide/air index contrast is substantial, and the light is well-confined in the polymer not extending much into the air. Therefore the mode diameter of the polymer mode is somewhat smaller than the

polymer cross-section. Note also that the optimum polymer waveguide height is smaller than the width; the reason for this is that the polymer mode extends into the undercladding because the index contrast between the polymer and SiO_2 is relatively small. The dimensions of the polymer waveguide were chosen to be 5.2x4.3μm. The selected height was chosen to be a little bit smaller than the optimum value to facilitate subsequent mode conversion; the penalty due to the increased mode mismatch is very low (less than 0.05dB).
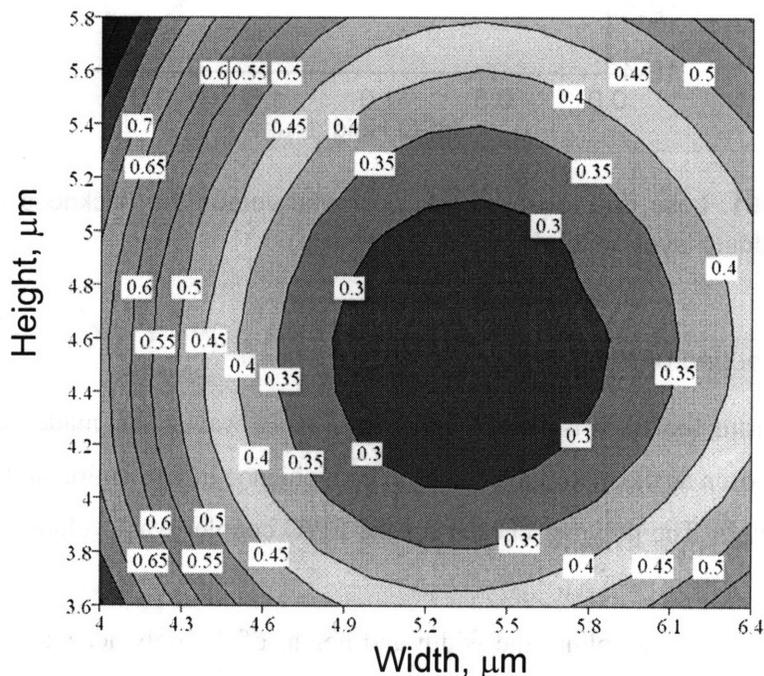
Figure 62. Coupling loss from the small-core fiber with 4 μm mode field diameter into the polymer waveguide versus width and height of the polymer waveguide. The numbers along the contour lines represent loss in dB.

3. Thickness of HSQ layer in the coupling region

To get into the SiN waveguide from the polymer, the light must be coupled through the HSQ layer. HSQ has a refractive index of about 1.44, which is less than the index of the cyclotene polymer, 1.53. Therefore, the mode is evanescently decaying in the HSQ layer, which serves as a barrier between the polymer and the SiN guide. Let us now estimate how serious a problem the 1.5μm layer of HSQ will present.

Fig. 63(a) shows mode conversion loss as a function of taper length when HSQ layer is 1μm thick. The taper shape was assumed to be linear; the loss was calculated with mode

overlap method using commercial software Fimmwave. We see that the taper must be about 1cm long to obtain less than 1 dB loss. Such a long taper is clearly not practical: in addition to occupying very large area on the chip, it will also lead to increased propagation losses in the waveguides, the net result being a huge and very lossy coupler. The coupler for 1.5μm overcladding will be even longer. Taper shape optimization can help reduce the required length, but still the this coupler will be impractical because of very large sensitivity to fabrication errors.

Thinning down the HSQ layer in the coupling region while leaving it intact over other parts of the chip helps a lot to solve this problem. Fig. 63(b) illustrates the loss dependence on taper length in case when the HSQ is thinned down to 100nm (which is 500nm if counted from the SiO₂ undercladding). The coupler efficiency increased dramatically: 500μm long taper with linear shape has only about 0.1dB of loss. In EPIC chip, the HSQ layer was thinned down to about 100nm; this was assumed in all following simulations. For the linear coupler design, the length was chosen to be 500μm.
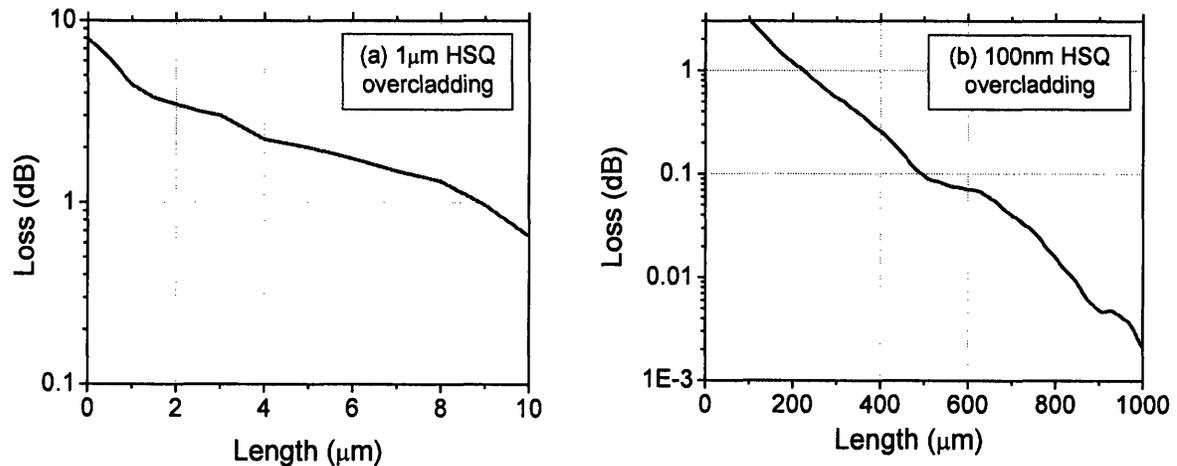
Figure 63. Mode conversion loss as a function of taper length for HSQ overcladding thickness of (a) 1μm, (b) 100nm.

4. Width of the SiN taper tip

After light is coupled into the polymer waveguide, it needs to be converted into SiN waveguide with an inverse taper. When the light encounters the tip of the SiN waveguide, it can experience some scattering due to discontinuity of refractive index distribution at this point. The loss calculated by mode overlap method is plotted in Fig. 64 versus the width of

the tip. We see that the loss is negligible as long as the tip width is below 200 μm. The tip width was chosen to be 200μm for "aggressively optimized" design and 100μm for "conservatively optimized" and linear coupler designs.
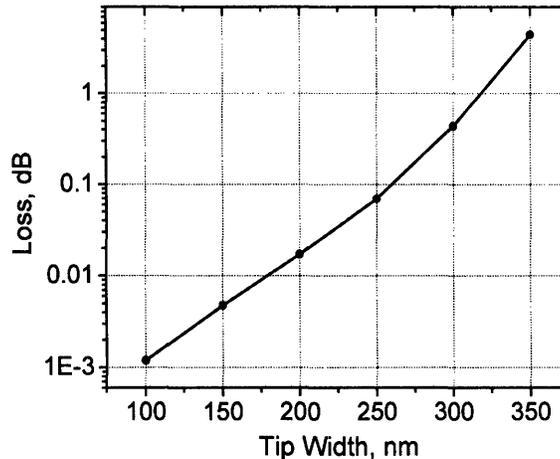
Figure 64. Loss at the SiN taper tip versus the width of the tip.

5. Width of the SiN waveguide at the end of polymer waveguide

When the light is transformed into the SiN waveguide, the polymer waveguide on top of it must be terminated. Due to fabrications limitations, it is likely that there will be a gap between the end of the polymer waveguide and the beginning of the thick HSQ layer, as it is shown in Fig. 65. There will be some loss due to mode mismatch between the SiN with and without polymer on top. The loss will depend on the width of the SiN waveguide at the point when the polymer end: the wider is the SiN guide, the more its mode is confined in the core and the less it sees of the index discontinuity, therefore the loss will be lower. Fig. 66 shows the mode mismatch loss as a function of the SiN waveguide width. We see that the loss should be about 0.1dB for the 600nm wide waveguide (which is our design width for the bus waveguide). We choose 600nm to be the width of the SiN waveguide at end of polymer for our "aggressively optimized" coupler design. This loss will be experienced by the mode twice: first time when it leaves the polymer waveguide, and the second time when it enters the region with thick HSQ overcladding. The total loss for the "aggressively optimized" design will therefore be 0.2dB. For the "conservatively optimized" design, we choose 800nm as the output SiN waveguide width; the total loss in this case is expected to be 0.05dB x 2 = 0.1dB.
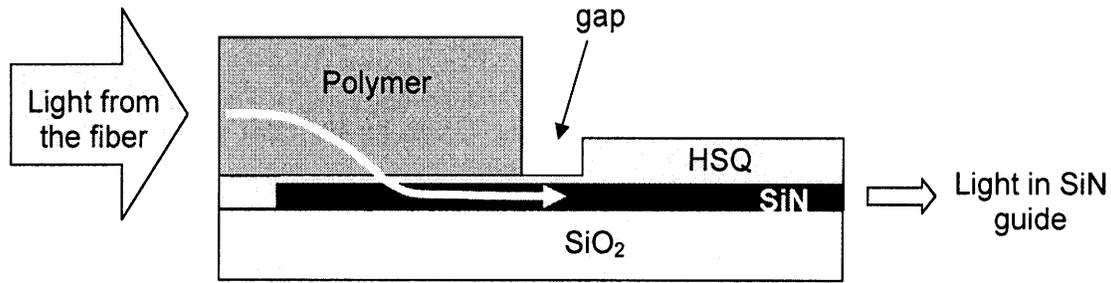
Figure 65. Cross-section of the coupler along the light propagation direction.
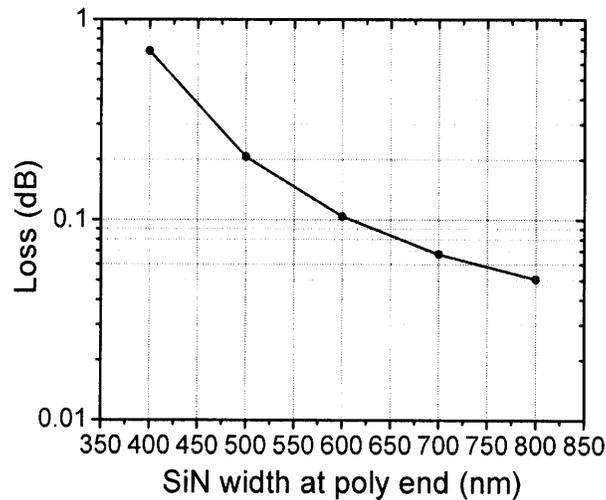
Figure 66. Loss experienced by SiN waveguide mode due to mode mismatch at the end of the polymer waveguide as a function of width of the SiN waveguide.

6. Taper shape optimization

Let us now proceed to optimization of the taper shape using the theory described in chapter 13. As we have decided previously, the taper in the "aggressively optimized" design starts from 200nm tip and widens to 600nm, and in "conservatively optimized" design – from 100nm to 800nm. Fig. 67 show loss dependence on taper length for taper shapes given by a super-Gaussian, Hamming, and constant characteristic taper functions for the taper going from 200 to 600nm. We see that the mode conversion loss as low as 0.1 dB can be achieved with 100μm taper. The taper based on constant characteristic function performs worse than the others, and tapers based on super-Gaussian and Hamming functions lead to approximately the same efficiency. We choose 100μm taper with shape based on super-Gaussian

characteristic function for our "aggressively optimized" design. Fig. 68 shows the actual evolution of SiN waveguide width along the taper.

Fig. 69 shows the loss dependence on the taper length for the taper going from 100 to 800nm. We choose 250μm long taper with Hamming characteristic function for our "conservatively optimized" design. Theoretically this should ensure mode conversion loss of 0.01-0.02 dB. The actual taper shape is shown in Fig. 70.
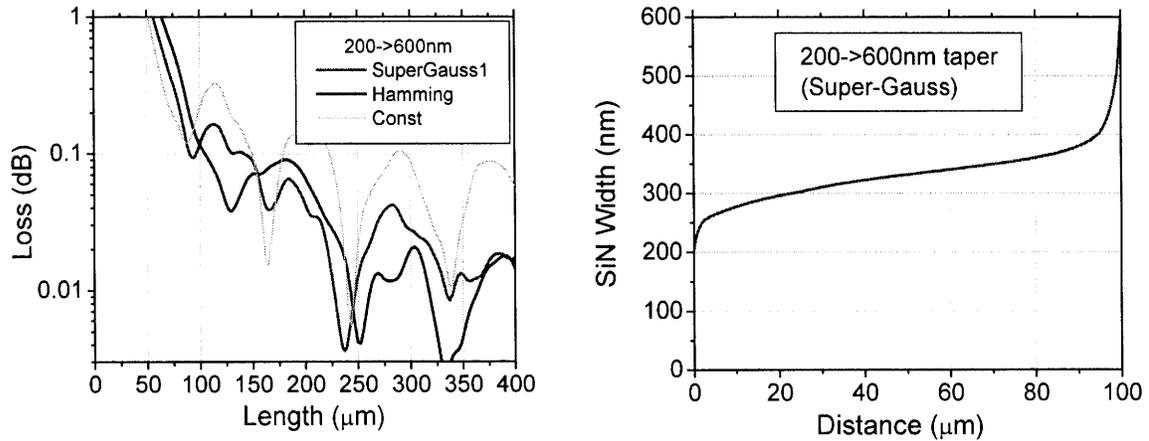
Figure 67. Mode conversion loss as a function of taper length for tapers with different characteristic functions. The SiN waveguide width is 200nm at the beginning of the taper and 600nm at its end.

Figure 68. Shape of the "aggressively optimized" taper with super-Gaussian characteristic function corresponding to Fig. 67.
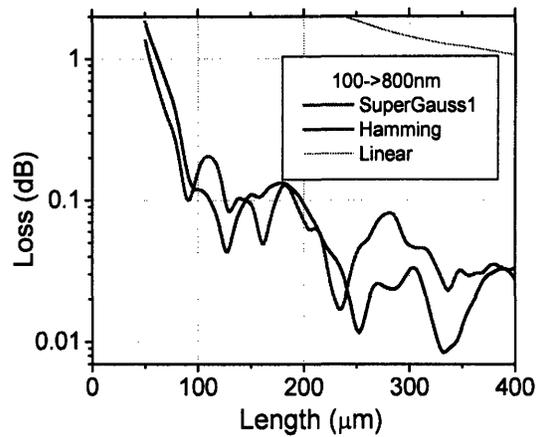
Figure 69. Mode conversion loss as a function of taper length for tapers with different characteristic functions. The SiN waveguide width is 100nm at the beginning of the taper and 800nm at its end.
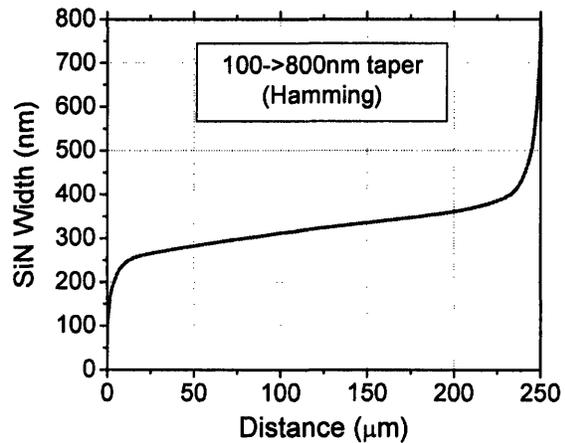
Figure 70. Shape of the "aggressively optimized" taper with super-Gaussian characteristic function corresponding to Fig. 69.

7. Summary: three coupler designs

In this chapter we found the parameters for three coupler designs: (1) the coupler with linear taper, (2) "aggressively optimized" coupler, and (3) "conservatively optimized" coupler. The parameters and theoretical performance of these couplers are summarized in the Table 4 below. The propagation loss in SiN waveguide was assumed to be 10dB/cm.

112

	(1) Coupler with linear taper	(2) "aggressively optimized" design	(3) "conservatively optimized" design
Design Parameters:			
• SiN tip width	100nm	200nm	100nm
• SiN width at taper end	600nm	600nm	800nm
• Coupler length	500µm	100µm	250µm
• Taper shape	Linear	Fig. 68	Fig. 70
Theoretical Loss:			
• SMF-poly mode mismatch	0.28 dB	0.28 dB	0.28 dB
• Taper mode conversion	0.10 dB	0.10 dB	0.01 dB
• SiN propagation loss	0.30 dB	0.05 dB	0.13 dB
• End of poly air gap	0.20 dB	0.20 dB	0.10 dB
Total Loss:	0.88 dB	0.63 dB	0.52 dB

Table 4. Final parameters and expected performance of the three coupler designs.

15. Two-Stage Fiber-to-Chip Coupler Design

The couplers described in the previous chapter are designed for coupling light from a special small-core fiber with mode field diameter of 4µm. All efficient horizontal fiber-to-chip couplers published in literature are intended for working with small-core fibers as well. However, commonly used fibers are standard single-mode fibers (SMF) with mode field diameter of about 9µm. It turns out that using the inverse taper concept for coupling light from standard fibers is problematic. This chapter describes a novel horizontal coupler which is designed to work with standard fibers. The waveguide material for the present design was silicon which has even smaller mode size than SiN, so that the proposed coupler is reducing the optical mode area by about 3 orders of magnitude.

1. Two-stage coupler concept

The proposed coupler is based on two-stage adiabatic mode evolution and is illustrated in Fig. 71 [54]. The light is first coupled into a polymer waveguide of approximately the same dimensions as the SMF fiber. At the first stage of the coupler, a rib is introduced into the rectangular polymer waveguide and this rib width is adiabatically narrowed until most of the

field gets "squeezed" into the lower part of the waveguide. At this point the rib can be terminated with very little power loss. As a result of stage I, the light is transferred from a large polymer waveguide into a smaller one. At the stage II, the field is transferred into a high-index contrast waveguide by using an inverse taper described in chapter 14. In the beginning of the inverse taper the silicon is very narrow so the fundamental mode is not confined to Si waveguide core and matches well the mode of the polymer guide. The silicon waveguide is then adiabatically widened until the fundamental mode of the structure becomes well-confined in the silicon core. The mode field evolution in both stages is illustrated at the bottom of Fig. 71.
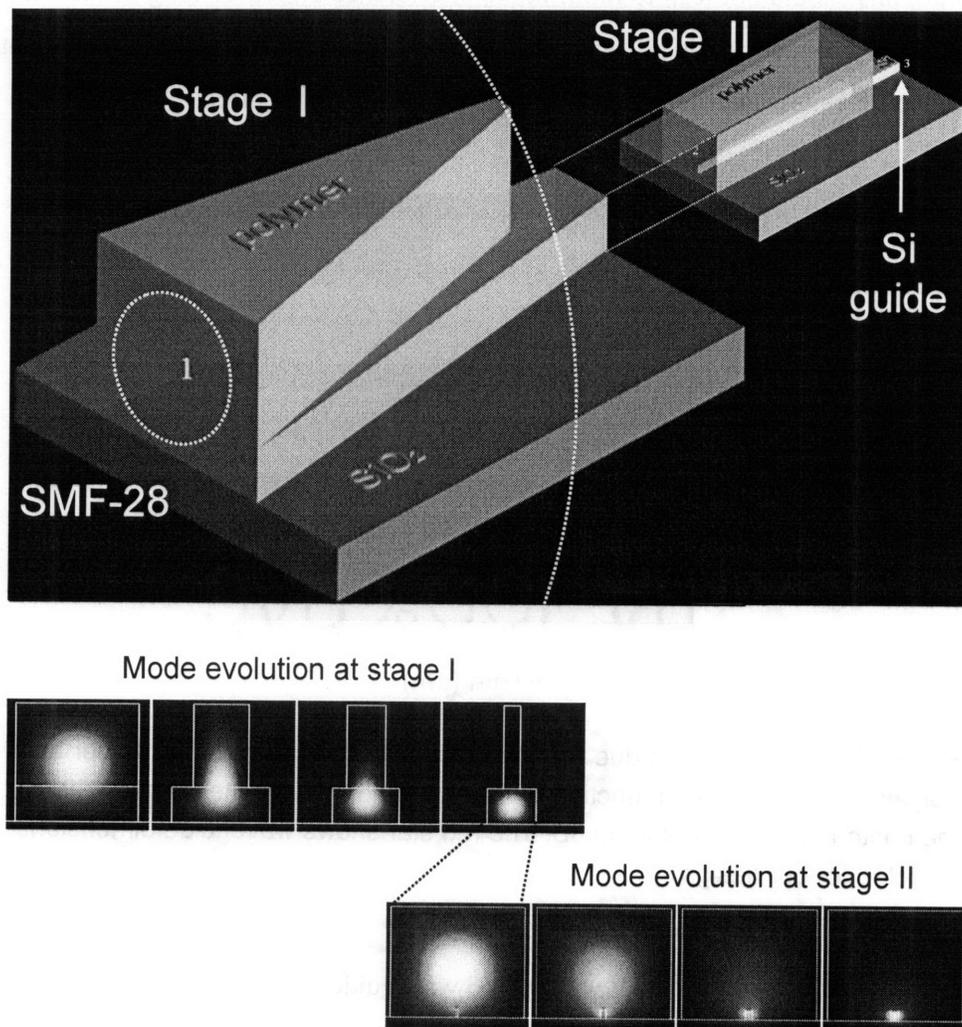
Figure 71. Two-stage coupler from standard fiber into to a high index contrast waveguide (to) and mode evolution along the two-stage coupler (bottom).

2. Optimization of coupler parameters

The coupler parameters to be optimized include dimensions of the polymer waveguide at the beginning of stage I, dimensions of the polymer waveguide at stage II, and taper shapes for both stages. For this specific design, we assumed that the output Si waveguide cross-section is 400x250nm and the polymer of choice is again cyclotene (n=1.53).

The dimensions of the input polymer waveguide were chosen to maximize overlap between the waveguide and the fiber mode. For SMF-28 fiber with mode field diameter of about 10μm; the polymer waveguide with width 12μm and height 11μm was found to give the coupling loss of 0.28dB (see Fig. 72). Note that the waveguide with such dimensions is multi-mode; this should not lead to multi-mode interference effects as long as the waveguide cross-section is changed gradually enough along the taper and the coupler does not have much surface roughness and slanted sidewalls introduced in fabrication.
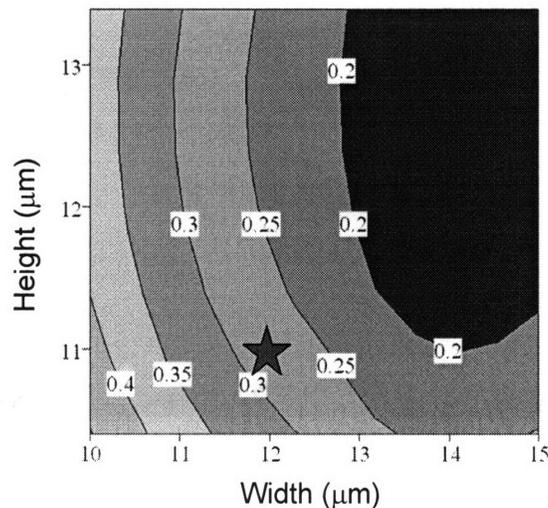
Figure 72. Coupling loss due to mode mismatch between SMF-28 fiber and polymer waveguide as a function of its width and height. The numbers along the contour lines are losses in dB. The red star shows waveguide dimensions selected for our design.

The plot of the mode mismatch loss at the Si waveguide tip at stage II is shown in Fig. 73. We choose the tip width of 50nm, the loss for such a narrow tip is negligible. The Si waveguide width in our coupler design is thus changing from 50 to 400nm along the inverse taper of stage II.
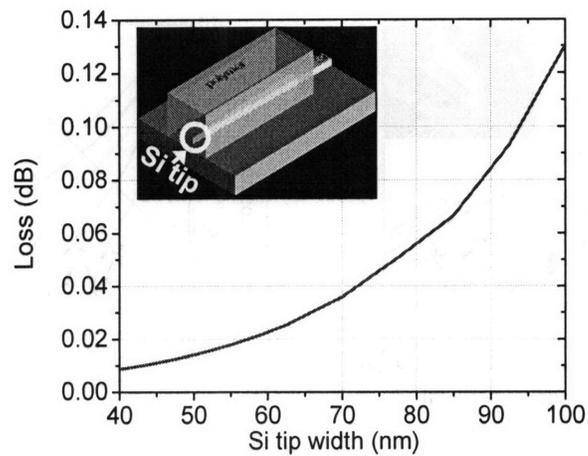
Figure 73. Loss at the Si waveguide tip as a function of tip width.

Let us now proceed to optimizing dimensions of the polymer waveguide at stage II. It turns out that the height of this waveguide H_{II} is an important factor determining the overall efficiency of the coupler. It the stage II waveguide is too tall, the length of inverse taper in stage II required for efficient mode conversion becomes too large. On the other hand, as the polymer height at stage II is decreasing, the first stage must become longer because it must reduce the vertical mode extent by a larger fraction. This is illustrated in Fig. 74, which shows mode conversion loss in stage I and II as a function of their length for different values of height of the polymer waveguide in stage II. The taper shapes were assumed to be linear in these simulations. It can be said that the value of H_{II} "divides" work between the two sections: for small H_{II} the first section needs to "work harder" than the second, and vice versa. One should expect that the optimal value of H_{II} must exist which optimizes the overall efficiency achievable with the coupler. This is confirmed in Fig. 75, which shows the overall mode conversion loss versus H_{II} for different total lengths of both stages. The ratio of lengths of two stages was optimized independently for each point of the plot so that the conversion efficiency is maximized. We can see that the optimum value for our coupler parameters is around 3.5-4.5µm. The overall length of two stages of 1.2mm allows to achieve 0.2dB mode conversion loss.
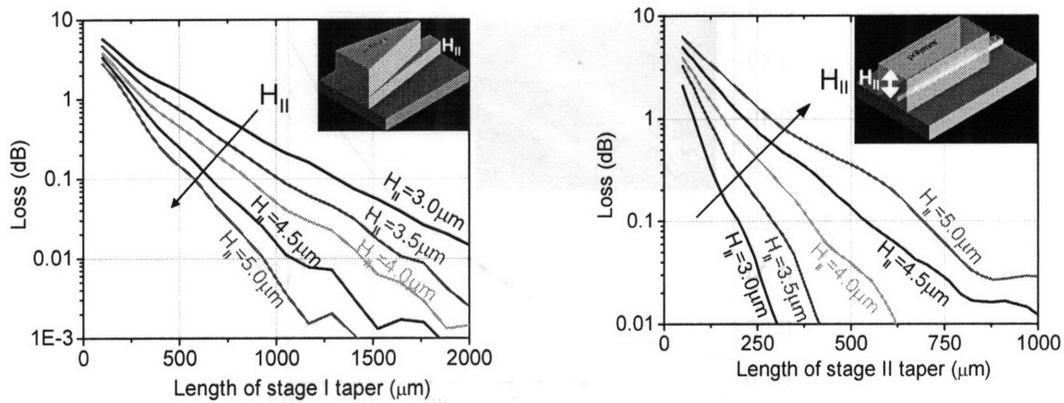
Figure 74. Mode conversion loss in stage I (left) and stage II (right) versus length of these stages. The curve parameter is the height H_{II} of the polymer in stage II.
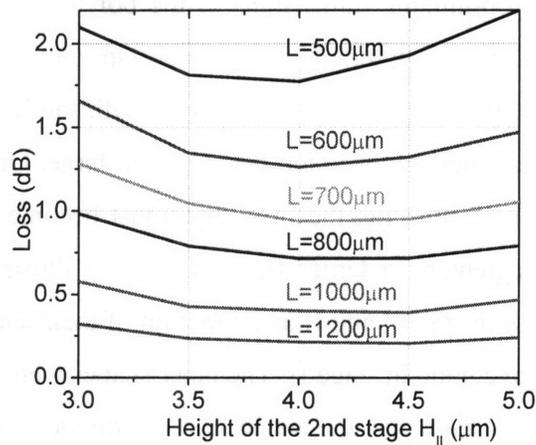
Figure 75. Loss versus height of stage II for different values of total length of both stages.

It is interesting to note that if the height of stage II is 11 μm, that is the first stage is absent, the required taper length becomes on the order of centimeters, which is too much for practical applications. This is the reason why the inverse taper alone without the first stage cannot be used for coupling light from standard single-mode fibers.

The taper lengths can be considerably reduced by optimizing their shapes. The mode conversion loss in each stage versus its length is shown in Fig. 3 for linear taper shape and optimized shapes based on constant, super-Gaussian and Hamming characteristic functions.

We choose super-Gaussian function for the first stage and Hamming function for the second; the corresponding taper shapes are shown in Fig. 77.

The total loss versus the total length of both section $L = L_I + L_{II}$ is shown in Fig. 78. The relative length of the sections L_I / L_{II} was again optimized individually for each point of the plot. We see that the mode conversion loss of 0.2dB can be achieved for total coupler length to 500µm (335µm stage I + 165µm stage II). The overall coupler loss of the coupler, including the mode mismatch loss with the fiber, is therefore 0.5dB. The taper parameters and performance is summarized in Fig. 79.
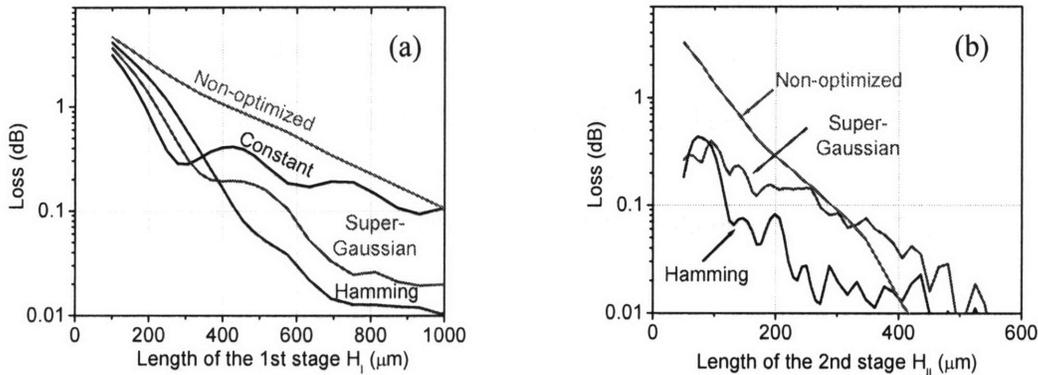
Figure 76. Mode conversion loss in (a) stage I and (b) stage II versus length of the stage for different taper shapes.
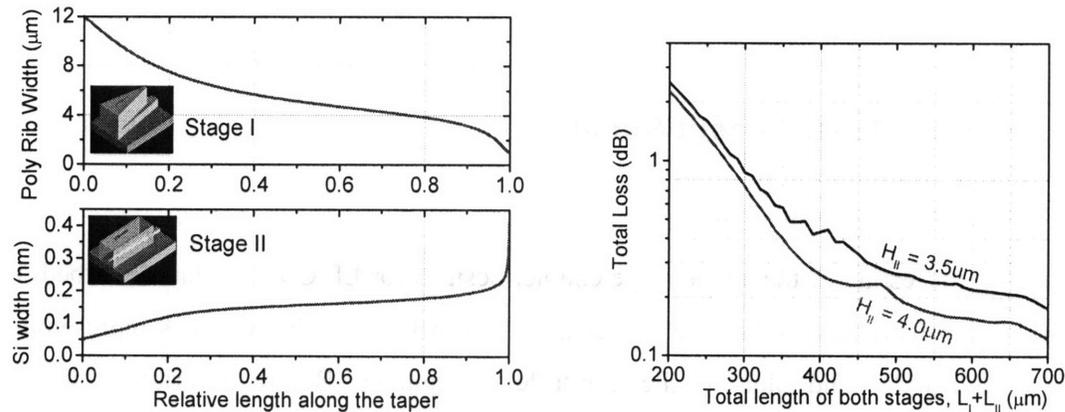
Figure 77. Taper shape for stage I (top) and stage II (bottom).
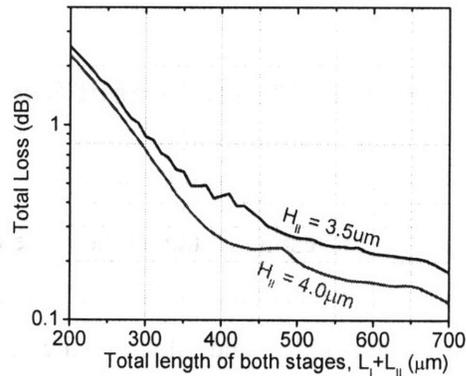
Figure 78. Total loss as a function of total length of the coupler; the curve parameter is the height of the polymer at stage III.

Fiber → Polymer Waveguide Mode Mismatch Loss = 0.28dB

Stage I: 335μm rib taper, mode conversion loss = 0.18dB

Stage II: 165μm inverse Si taper, mode conversion loss = 0.02dB

Total Length = 500μm   Total Loss = 0.5 dB
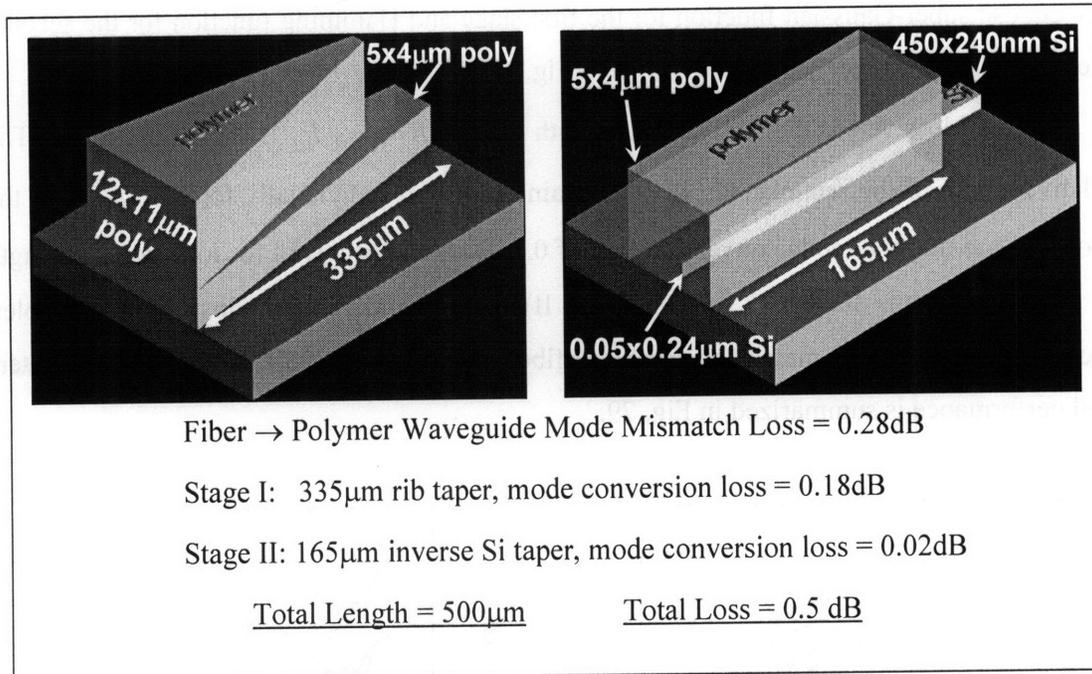
Figure 79. Final parameters and simulated performance of the two-stage coupler.

The taper parameters given above were obtained by optimizing it for TE mode only. The taper shapes found for TE mode are not automatically the best ones for TM mode. Therefore, if efficient operation is required for both polarizations, the coupler length must be increased; in our case 800μm coupler gives about 0.8dB loss for TM mode.

16.   Coupler Measurement Results

1. Layout of the coupler test chip

The chip for testing of the single-stage coupler designs for EPIC ADC chip described in chapter 14 was fabricated[1]. The chip layout is shown in Fig. 80, the red lines representing polymer waveguides and blue – silicon nitride waveguides. Before describing the test

---

[1] Coupler fabrication was done by Mohammad Araghchini and Charles W. Holzwarth. Raul Barreto and Tymon Barwicz fabricated an initial version of the coupler and initially developed the fabrication process. The chip layout was mostly designed by Charles W. Holzwarth with contributions from the author.

structures included on the chip, let us explain the problems with loss measurements we are trying to solve:

- if the couplers are efficient, the loss will be small and it becomes small to measure is reliably;

- the loss experienced by light going across the chip might be caused not only by coupler loss but also by waveguide propagation loss, which can exceed the coupler loss;

- waveguide facets after cleaving can be of different quality; this might cause waveguide-to-waveguide loss variations exceeding the expected coupler losses.
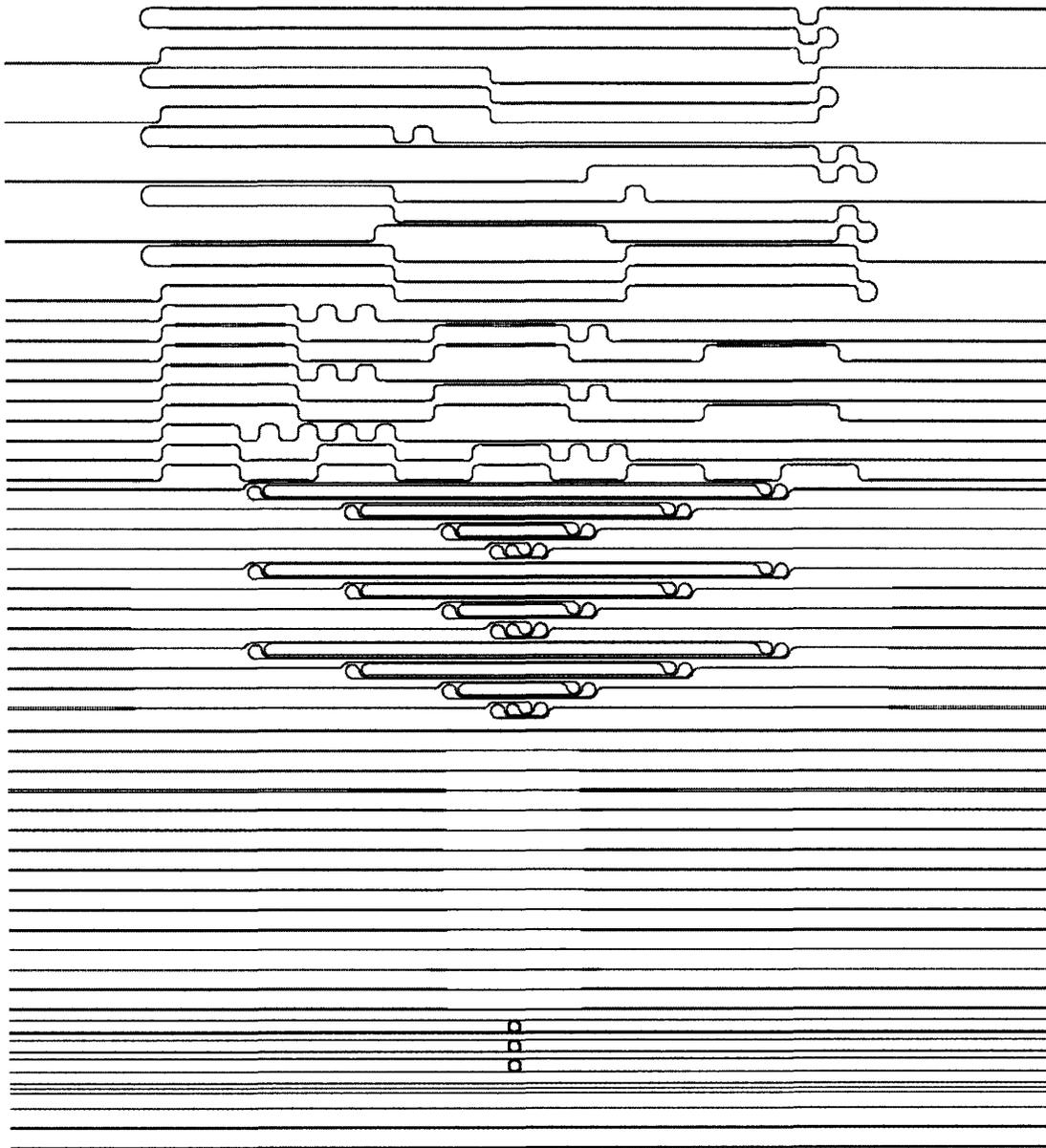
Figure 80. Layout of the coupler test chip. Polymer waveguides are red and SiN waveguides are blue.
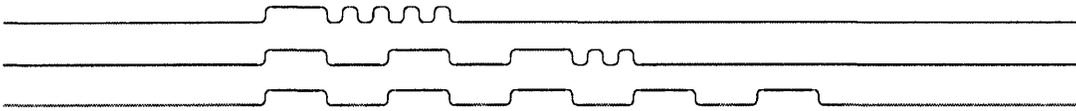
Figure 81. Test structures for 100μm coupler. The structures are identical in all respects except the number of tapers, which equals 4, 12, and 20 starting from top. The "wiggles" are created to shift polymer waveguide transversally away from each other, so if there is light at the end of one polymer wavegude, it will not be coupled into the next waveguide by propagating in the overcladding (or air).

The test structures fabricated on the chip can be divided into the following groups.

- For each taper design, there is a set of waveguide structures with several mode converters to and from SiN waveguides. The structures differ in the number of mode converters. For example, for the 100μm coupler design ("aggressively optimized" coupler), there are structures with 4, 12, and 20 couplers, as shown in Fig. 81. All properties of the structures are the same, including the total length of SiN and polymer waveguides and the number of bends, only the number of tapers is different. If measurements show that the loss is different between the structures, this difference can only be explained by different number of tapers. The measured losses can be plotted as a function of number of tapers and loss in one taper can be obtained from fitting the plotted points with a line. This approach helps to solve the problem with low coupler loss by using many couplers and the problem of other loss sources by making sure that the only difference between the structures is in number of couplers. However, the problem of coupler characterization is not solved completely: with this method, we are measuring the mode conversion loss in the taper, which is only first part of the coupler loss, the second being the mode mismatch loss. However, the mode conversion loss is also valuable information which distinguishes one coupler design from another.

- For each coupler design there is also a set of structures with just one input and one output coupler and different lengths of SiN waveguide section. Plotting the power at the output of each paperclip as a function of SiN length we can estimate the SiN waveguide propagation loss; the coupler loss can also be estimated from the intersection of the obtained curve with the y-axis.

- For each coupler design there are just straight sections of polymer waveguide with one section of SiN waveguide in between. The structures are also different in transversal shift of the SiN waveguide from the center of the polymer waveguide. These shifts were introduced

so that at least one of the shift values will compensate shifts resulting from possible fabrication errors. These structures are not very effective for measuring loss because of the problem mentioned above, but they can shine some light on coupler behavior.

• There are also auxiliary structures, such as straight polymer waveguides and ring filter for measuring SiN propagation loss.

2. Coupler measurement results

Now we present some preliminary measurements of coupler performance[1]. Fig. 82 shows the measured transmission spectrum of the 100μm coupler test structures shown in Fig. 81. The transmission in this figure is not normalized, i.e. the absolute transmission is not meaningful and only relative difference in transmission should be considered. The transmission curves do not behave as expected: the curve for 4 tapers has strong oscillatory behavior with dip of about 20dB at 1536nm. The alignment of fiber was optimized to obtain maximum transmission at 1550 nm.
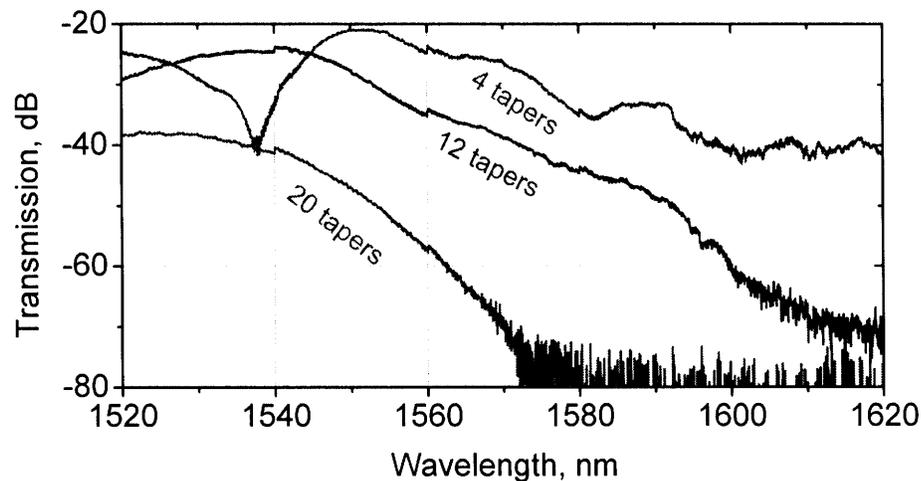
Figure 82. Relative transmission of 100μm coupler test structures with 4, 12, and 20 tapers (see Fig. 81).

The oscillatory frequency dependence of the transmission curve is characteristic of multi-mode interference effects, which can take place in the polymer waveguide which is multimode. To confirm that the transmission minimum at 1536 nm is not caused by the taper,

---

[1] The measurements were done by Marcus Dahlem.

the measurement was repeated for a different alignment of the fiber with respect to the polymer waveguide, see Fig. 83. One of the curves in this plot corresponds to the fiber alignment which maximizes transmission at 1550 nm (i.e. the same as in Fig. 82), and for another curve, the fiber was aligned to maximize transmission at 1536nm, the minimum of the first curve. We see that the dip in transmission shifts to other wavelength when the fiber is moved, which supports the hypothesis that the oscillations in the transmission are caused by multi-mode interference.
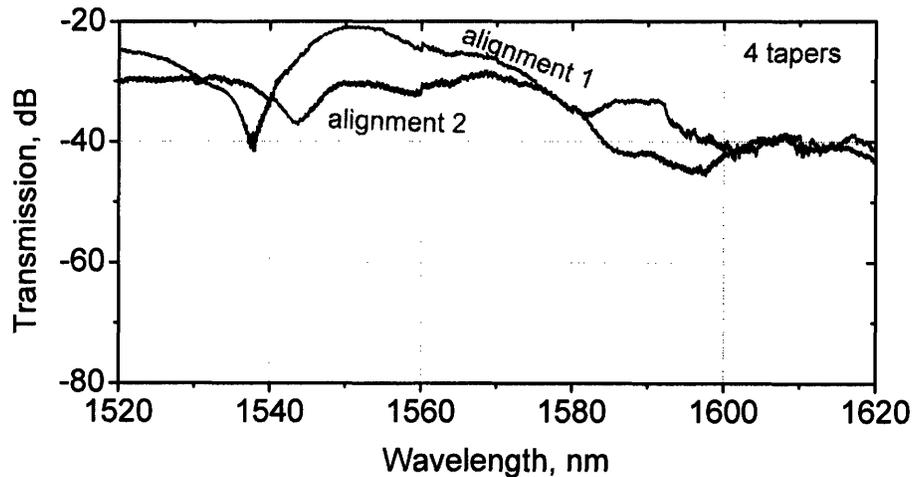
Figure 83. Dependence of transmission of four 100μm "aggressively optimized" couplers on the alignment of the fiber.

Note also that the transmission goes down a lot for longer wavelengths in Fig. 82, especially for structures with 12 and 20 tapers. This is explained by the fact that the 100μm coupler is the "aggressively optimized" design intended to work at 1550 nm and is not expected to have very large bandwidth.

Fig. 84 presents transmission spectra of the test structures for 500μm linear tapers. The layout of the test structures is given in Fig. 85 (these are the same as in Fig. 80, just zoomed in). The oscillations in the spectrum are still present, especially for the structure with 8 tapers.
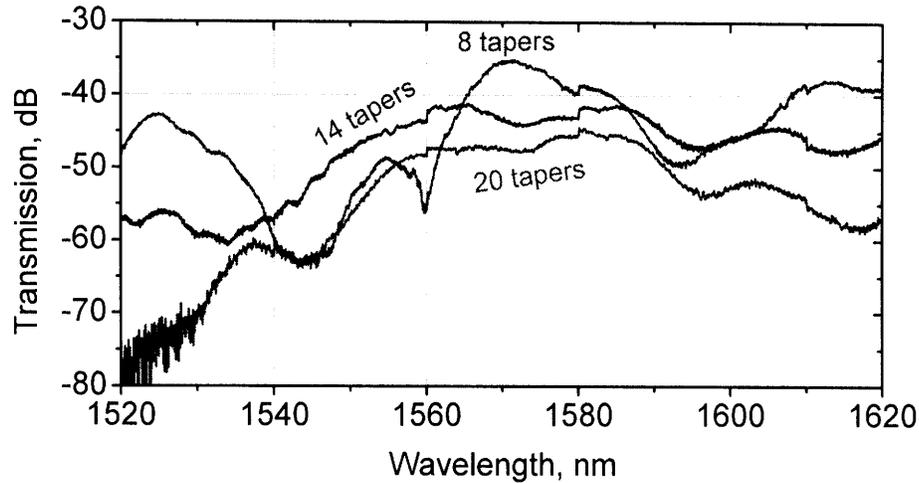
Figure 84. Relative transmission of test structures for 500μm linear coupler with 8, 14, and 20 tapers.
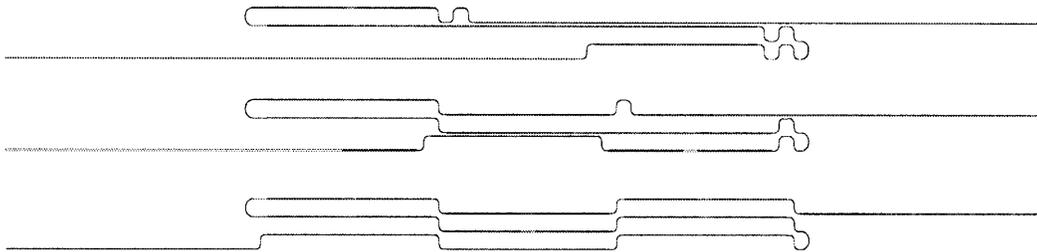
Figure 85. Test structures for 500μm linear couplers with 8, 12, and 20 tapers.

The chip also contains test structures with 750μm linear tapers, which turned out to have less oscillatory transmission spectrum, see Fig. 86. The sensitivity to the alignment of the input and output fibers is low, as it is shown in Fig. 87. The layout of the test structures is shown in Fig. 88. The transmission spectra for 8 and 14 tapers are almost parallel, as it should be the case without multi-mode effects. In this case we can actually estimate the mode conversion loss by subtracting transmission for 8 and 14 tapers and dividing it by 6; the result is shown in Fig. 89. The loss per taper is within 2–2.5 dB, which includes also the waveguide propagation loss. Our SiN waveguides have a larger-than-expected loss estimated to be 40dB/cm, which gives about 1.5dB waveguide propagation loss per taper. The mode conversion loss is therefore 0.5-1dB per taper.
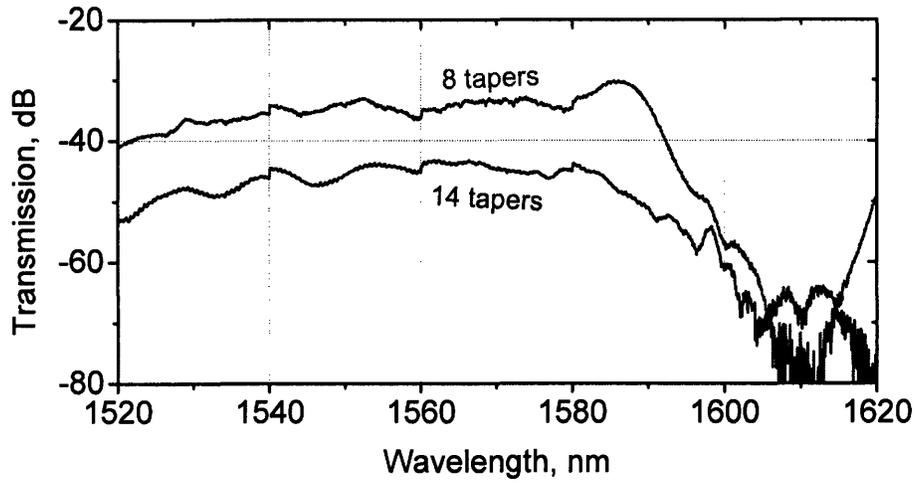
Figure 86. Relative transmission of test structures for 750μm linear coupler with 8, and 14 tapers.
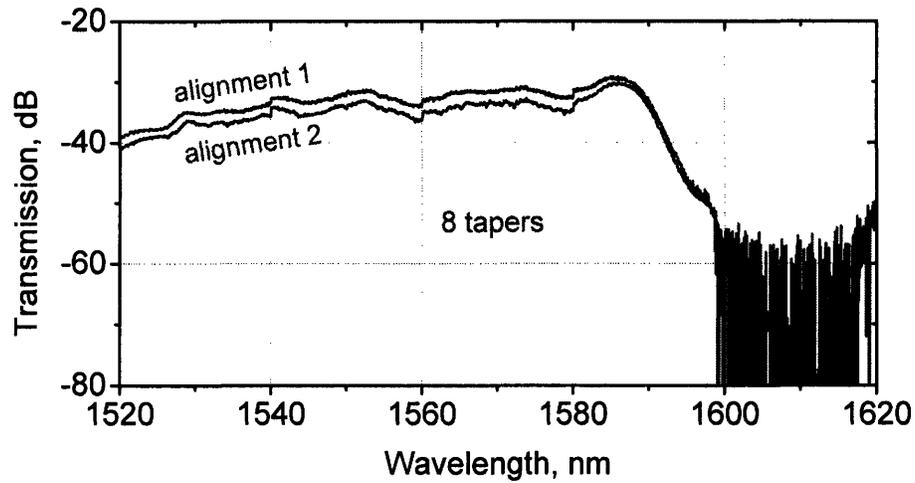
Figure 87. Effect of fiber alignment on transmission of the 8-taper test structure for 750μm linear coupler.
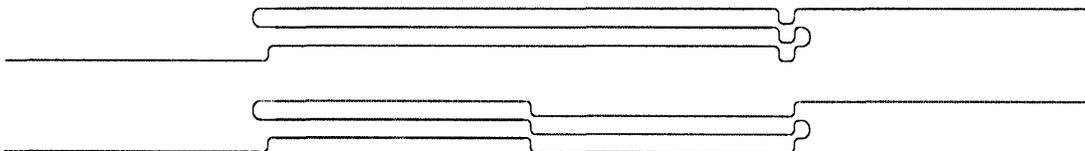
Figure 88. Test structures for 750μm linear couplers with 8, 12, and 20 tapers.
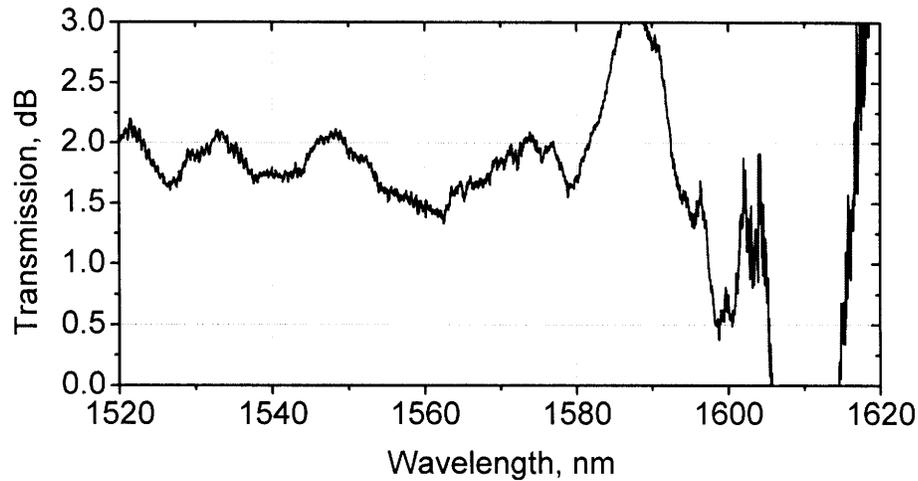
Figure 89. Mode conversion loss in a 750μm linear taper, as obtained from Fig. 86. The loss includes the conversion loss itself as well as waveguide propagation loss.

3. Discussion of measurement results

As we have seen above, the transmission spectra of the coupler test structures show oscillatory behavior which can be explained by multi-mode interference. Another confirmation that what we observe is the result of multi-mode interference is that the calculated period of beating between the 1st and the 3rd modes is $\Delta\lambda_{13} \approx 42\text{nm}$, which agrees well with the oscillation period observed in Fig. 83.

However, the question remains open about how exactly the multi-mode interference occurs. If the input fiber is not perfectly aligned or has a wrong mode field diameter, coupling into higher-order modes will take place. However, the tapers are adiabatic and only the fundamental mode must get coupled into the SiN waveguide while the higher order modes should be radiated; imperfect coupling should lead only to power loss but not to spectral oscillations we observed.

The explanation to the experimental data can be the coupling between modes which occurs as light propagates along the polymer waveguide. The ACM image of the waveguide, see Fig. 90, shows some sidewall tilt and roughness which can result in coupling between the modes. At it is known from coupled mode theory, almost complete power transfer from one mode to another can occur if the propagation constants of the two modes are close. If the light is completely transformed into a higher-order mode in the taper region, the light will not

make it into the SiN waveguide and will be radiated away at the end of the polymer, resulting in a dip in the transmission spectrum. If the taper region is at the point when most energy is transferred from a higher-order mode to the fundamental mode, the light will get coupled into the SiN waveguide resulting in transmission peak.

The dependence of the wavelength of the transmission minimum on the alignment of the fiber agrees with out explanation. When the input fiber is moved, the mode excitation coefficients are changed. This changes the initial conditions for the process of field oscillation between the modes, and thus shifts the transmission minima.

Another factor leading to spectral dependence of the transmission is the dependence of efficiency of light coupling into the output fiber on mode excitation coefficients. Multi-mode interference at the output of the coupler will create some intensity pattern which will change with frequency. The coupling efficiency into the output single-mode fiber will depend on this intensity pattern and therefore on wavelength. To reduce the oscillations because of this phenomenon we tried to do the transmission measurements with a multi-mode fiber at the output. The spectral oscillations were indeed reduced, but not to the extent sufficient for reliable characterization of the couplers. The measurements of brightness of the light scattered at the output facet of polymer waveguide using an infrared (IR) camera confirmed the presence of wavelength dependence of the output pattern and reduction of oscillations when the total scattered power is estimated[1].
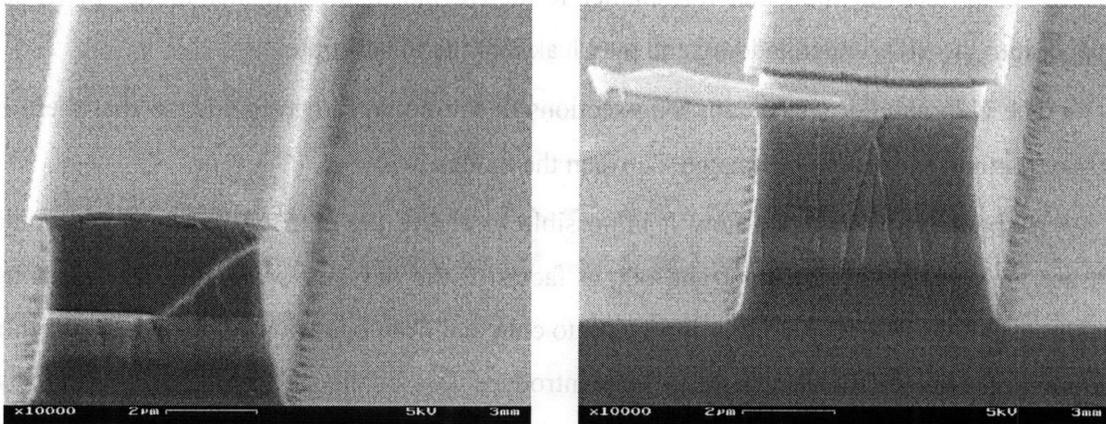
Figure 90. SEM images of the cross-section of polymer waveguide[2].

---

[1] IR camera measurements were done my Milos Popovic and Marcus Dahlem.

[2] SEM images of the polymer waveguides were taken by Mohammad Araghchini.

The fact of mode coupling occurring along the polymer waveguide is indirectly confirmed by spectral oscillation being the largest in the structures with the longest sections of the polymer waveguide. If we look at the results of 100μm coupler measurements in Fig. 82, we see that the oscillations are largest for the 4-taper structure, which has longer polymer section than 12 and 20-coupler structures (Fig. 81). The structures for testing the 750μm couplers have relatively short polymer sections (compare Fig. 90 to Figs. 81 and 85), therefore the oscillations are smaller than for other tapers so that we could get a meaningful loss estimate for 750μm taper mode conversion loss. The transmission spectrum of 750μm test structures is also insensitive to alignment of the input and output fibers (Fig. 87), another indicators that the multi-mode effects in these structures are smaller.

4. Future directions

The following improvements in coupler design and fabrication process can be considered.

• The fabrication process should be optimized to obtain the sidewalls of the polymer waveguide which are smoother and less slanted. This should reduce the mode transfer between the modes.

• A polymer with smaller refractive index should be considered so that the polymer waveguide becomes single-mode or with fewer number of higher-order modes. However, it is necessary to keep in mind that for smaller refractive index, the polymer modes penetrate deeper into the SiO_2 undercladding and may leak into the Si substrate.

• The chip should not contain long sections of the polymer waveguide, so that there is less opportunity for energy exchange between the modes.

• It should be considered how it is possible to cleave the chip in a way which would ensure high quality of the input and output facets of the polymer wavegude. The cleaving quality can now change randomly from chip to chip and from one waveguide on a single chip to another. Nonperfect cleaving can now introduce loss on the order of several decibels, which is larger than the expected loss of the coupler.

APPENDICES

Appendix A. ENOB calculation

ADC performance is usually characterized with such metrics as the amount of second- and third-order distortions and noise-limited ENOB [1, 2, 4, 55]. Although these metrics are useful in practice, they are not universal in the sense that the values of ENOB estimated using each of these metrics can be different from each other and the results can depend on the test signal. For example, pulse-to-pulse crosstalk due to slow photodetectors cannot be notices if a sinusoidal test signal is used.

In the thesis we adopted a different approach to ENOB calculation. Because our studies are mostly theoretical, we have the benefit of knowing the precise analog RF signal with arbitrary accuracy. We thus calculate the ENOB directly by comparing the input analog RF signal to the output of our ADC according to [1]:

$$SNR(dB) = 20 \cdot \log\left(\frac{v_{RMS}}{\Delta v_{RMS}}\right),$$

$$ENOB = \frac{SNR(dB) - 1.76}{6.02},$$

where v_{RMS} and Δv_{RMS} are root mean square values of the signal and the ADC error.

The ENOB calculated with this approach depends on the type of the analog signal, namely on its spectrum. ENOB will often be much higher for a signal rich in lower frequencies. To obtain an objective value for ENOB, a random test signal was generated and then the lower frequencies were filtered out so that only the higher half of the spectrum with frequencies in the range $[0.55...0.95] \cdot f_s / 2$ remains, where f_s is the sampling rate of the ADC and $f_s / 2$ is its Nyquist frequency. This is a reasonably hard signal which allows to test the ADC performance in various cases without the danger of missing some of its properties. The results will statistically depend on the specific implementation of the signal; to get a reliable ENOB estimate, a reasonably long test signal was used in simulations so that the resultant ENOB is averaged over long time window and shows little variations from one implementation of the signal to another.

The effects of signal quantization by the electronic ADCs were not considered in ENOB calculation.

Appendix B. Notation used in the thesis

The following is the list of the most important notation used in the thesis.

f_R repetition rate of the mode-locked laser;

T_R repetition period of the laser; $T_R = 1/f_R$;

f frequency of the optical signal

ω angular frequency of the optical signal

Δf channel frequency spacing

$\Delta \omega$ channel frequency spacing in terms of angular frequency, $\Delta \omega = 2\pi f$;

Δf_{FWHM} full-width at half maximum bandwidth of an optical filter;

$\Delta \omega_{FWHM}$ full-width at half maximum bandwidth of an optical filter in terms of angular frequency;

Δf_{det} bandwidth of the photodetector;

ω_{ref} reference frequency, that is the frequency of zero group delay; this is assumed to be the center frequency of the filter bank;

N number of channels;

f_s sampling rate of the ADC system; $f_s = N f_R$ for an ADC system with N channels;

β_2 dispersion coefficient of the fiber; $\beta_2 = \dfrac{c}{2\pi f_{ref}^2} D$

L length of the dispersive fiber;

m modulation index of the MZ modulator;

e_0 optical field of the pulse train at the input of MZ modulator;

e^{up}, e^{down} optical fields at the output of the upper and lower arms of the MZ modulator;

$v, v(t)$ analog RF signal at the input of the ADC being quantized;

v^{up}_{ADC}, v^{down}_{ADC} the signal obtained at the output of the upper and lower parts of the ADC system using complimentary outputs of the MZ modulator;

v_{ADC}, $v_{ADC}(t)$ the signal obtained at the output of the ADC; our goal is to make this signal as close to the applied analog RF signal as possible;

$e_{laser}(t)$ optical field of the pulse train at the output of the mode-locked laser; $E_{laser}(\omega)$ is the same field in frequency domain; throughout the thesis, capital letter correspond to frequency and lowercase letters – to time domain;

$e_{fiber}(t)$ optical field at the output of the dispersive fiber;

$\phi^{up}_{MZ}(t)$, $\phi^{down}_{MZ}(t)$ phase shift in the upper and the lower arms of the MZ modulator;

$v^{up}_{MZ}(t)$ defined as $v^{up}_{MZ}(t) = \sin\left(\dfrac{\pi}{4} + v(t)\right)$;

$v^{down}_{MZ}(t)$ defined as $v^{down}_{MZ}(t) = \cos\left(\dfrac{\pi}{4} + v(t)\right)$;

$v^{up}_{eff}(t)$ the "dispersed" version of $v^{up}_{MZ}(t)$ defined by $V^{up}_{eff}(\omega) \equiv V^{up}_{MZ}(\omega)\exp\left[i\dfrac{\beta_2 L}{2}\omega^2\right]$;

$E^{up}_{filter,n}(\omega)$ optical field at the output of n-th filter of the upper filter bank;

$e^{\frown}_{laser}(t)$ envelope   of   the   laser   pulse   broadened   by   dispersion;

$$e^{\frown}_{laser}(t_n) \equiv \left| E_{laser}\left(\omega_n = \omega_{ref} + \dfrac{t_n}{\beta_2 L}\right) \right|;$$

f_{max} maximum frequency in the spectrum of the analog RF signal.

REFERENCES

1. G. C. Valley, "*Photonic analog-to-digital converters,*" Opt. Express **15**, pp. 1955-1982 (2007).

2. R. H. Walden, "*Analog-to-digital converter survey and analysis,*" IEEE J. Sel. Areas Comm. **17**, pp. 539-550 (1999).

3. H. F. Taylor, "*An optical analog-to-digital converter - design and analysis,*" IEEE J. Quantum Electron. 15, 210-216 (1979).

4. B. L. Shoop, "*Photonic Analog-to-Digital Conversion,*" Springer, New York, 2000.

5. T. Shibata, M. Yoneyama, "*A novel sample and hold system using an optical modulator,*" IEEE Photon. Technol. Lett. **4**, pp. 588-591 (1992).

6. J. A. Bell, M. C. Hamilton, D. A. Leep, H. F. Taylor, Y.-H. Lee, "*A/D Conversion of microwave signals using a hybrid optical/electronic technique,*" Proc. SPIE **1476**, 326-329 (1991).

7. J. A. Bell, M. C. Hamilton, D. A. Leep, "*Optical sampling and demultiplexing applied to A/D conversion,*" Proc. SPIE **1562**, 276-280 (1991).

8. P. W. Juodawlkis, J. C. Twichell, G. E. Betts, J. J. Hargreaves, R. D. Younger, J. L. Wasserman, F. J. O'Donnell, K. G. Ray, R. C. Williamson, "*Optically sampled analog-to-digital converters,*" IEEE Trans. Microwave Theory Tech. **49**, pp. 1840-1853 (2001).

9. J. C. Twichell, J. L. Wasserman, P. W. Juodawlkis, G. E. Betts, R. C. Williamson, "*High-linearity 208-MS/s photonic analog-to-digital converter using 1-to-4 optical time-division demultiplexers,*" IEEE Photonic Technol. Lett. **13**, pp. 714-716 (2001).

10. J. U. Kang, R. D. Esman, "*Demonstration of time interweaved photonic four-channel WDM sampler for hybrid analogue-digital converter,*" Electron. Lett. **35**, pp. 60-61 (1999).

11. T. R. Clark, J. U. Kang, R. D. Esman, "*Performance of a Time- and Wavelength-Interleaved Photonic Sampler for Analog–Digital Conversion,*" IEEE Photon. Technol. Lett. **11**, pp. 1168-1170 (1999).

12. C. W. Holzwarth, T. Barwicz, M. A. Popović, P.T. Rakich, E.P. Ippen, F.X. Kärtner, H.I. Smith, "*Accurate resonant frequency spacing of microring filters without postfabrication trimming,*" J. Vac. Sci. Technol. B 24, 3244-3247, 2006.

13. M. Popovic, T. Barwicz, E. P. Ippen, F. X. Kärtner, "*Global Design Rules for Silicon Microphotonic Waveguides: Sensitivity, Polarization and Resonance Tunability,*" CLEO 2006, paper CTuCC1.

14. M. Popović, T. Barwicz, "*Fabrication tolerant silicon waveguides and resonators,*" U.S. Provisional Patent App., MIT case 12410, filed Dec. 2006, pending.

15. T. Barwicz, M.R. Watts, M.A. Popović et.al., "*Polarization-transparent microphotonic devices in the strong confinement limit,*" Nature Photonics 1, 57 (Jan. 2007)

16. M. Y. Frankel, J. U. Kang, R. D. Esman, "*High-performance photonic analogue-digital converter,*" Electron. Lett. 33, pp. 2096–2097 (1997).

17. F. X. Kärtner, et. al., "*Electronic photonic integrated circuits for high speed, high resolution, analog to digital conversion*", in Proc. SPIE 6125, paper 612503 (2006).

18. J. C. Twichell, R. Helkey, "*Phase-encoded optical sampling for analog-to-digital converters,*" IEEE Phot. Tech. Lett. 12, pp. 1237-1239 (2000).

19. T. R. Clark, M. Currie, P. J. Matthews, "*Digitally Linearized Wide-Band Photonic Link,*" J. Lightwave Technol. 19, pp. 172-179 (2001).

20. M. E. Grein, S. J. Spector, T. M. Lyszczarz et. al., "*Demonstration of a photonic analog-to-digital converter scalable to 40 GS/s with 8-bit resolution*", submitted to Conference on Lasers and Electro-Optics 2008.

21. S. J. Spector et al., "*Integrated optical components in silicon for high speed analog-to-digital conversion,*" Proc. SPIE 6477, paper 647700 (2007).

22. Hyunil Byun, "*Silicon-Germanium Saturable Absorbers and Erbium-doped Waveguides for Integrated Mode-locked Lasers*", M.S. thesis, MIT 2006.

23. F. Gan, F. Kartner, "*High-speed silicon electrooptic modulator design,*" IEEE Photon. Technol. Lett. 17, pp. 1007-1009 (2005).

24. Jonathan R. Birge, "*Chirped Pulse Sampling: Analysis Of Time-To-Frequency Coupling And Its Limits,*" internal document (2006).

25. A. S. Bhushan, F. Coppinger, S. Yegnanarayanan, B. Jalali, "*Nondispersive wavelength-division sampling,*" Opt. Lett. **24**, pp. 738-740 (1999).

26. A. Khilo, J. R. Birge, F. X. Kärtner, "*Adaptive Error Compensation for Photonic Analog-to-Digital Converters,*" submitted to Conference on Lasers and Electro-Optics 2008.

27. G. C. Valley, J. P. Hurrell, G. A. Sefler, "*Photonic analog-to-digital converters: fundamental and practical limits,*" Proc. SPIE **5618**, 96-106 (2004).

28. J. A. Cox, J. Kim, F. X. Kärtner, "*10-femtosecond Precision, Long-term Stable Timing Distribution*", submitted to XVI International Conference on Ultrafast Phenomena, Stresa, Italy 2008.

29. B. E. Little, S. T. Chu, H. A. Haus, J. Foresi, J.-P. Laine, "*Microring resonator channel dropping filters,*" J. Lightwave Technol. **15**, pp. 998-1005 (1997).

30. M.A. Popović, "*Theory and Design of High-Index-Contrast Microphotonic Circuits,*" Ph.D. Thesis, MIT 2007.

31. Vilson R. Almeida Carlos A. Barrios Roberto R. Panepucci, Michal Lipson, "*All-optical control of light on a silicon chip*", Nature **431**, pp.1081-1084 (2004).

32 A. Yalçin, K. C. Popat, J. C. Aldridge, T. A. Desai, J. Hryniewicz, N. Chbouki, B. Little, O. King, V. Van, S. Chu, D. Gill, M. Anthes-Washburn, M. S. Ünlü, "*Optical sensing of biomolecules using microring resonators,*" J. Sel. Top. Quantum. Electron. **12**, 148155 (2006).

33. Qianfan Xu, Bradley Schmidt, Sameer Pradhan, and Michal Lipson, "*Micrometre-scale silicon electro-optic modulator,*" Nature **435**, 325-327 (2005).

34. A. Melloni, M. Martinelli, "*Synthesis of direct-coupled-resonators bandpass filters for WDM systems,*" J. Lightwave Technol. **20**, pp. 296–303 (2002).

35. M. Popovic, M. Watts, T. Barwicz, P. Rakich, L. Socci, E. Ippen, F. Kartner, H. Smith, "*High-index-contrast, wide-FSR microring-resonator filter design and realization with frequency-shift compensation*", OFC 2005, paper OFKI.

36. T. Barwicz, "*Accurate Nanofabrication Techniques for High-Index-Contrast Microphotonic Devices,*" Ph.D. Thesis, MIT, September 2005.

37. A. Taflove, "*Computational Electrodynamics: The Finite-Difference Time-Domain Method*," Artech House Publishers, 2005.

38. Christina Manolatou, "*Passive Components for Dense Optical Integration Based on High Index-Contrast*," PhD Thesis, MIT, 2002.

39. C. Manolatou, H. A. Haus, "*Passive components for dense optical integration*," Boston, MA: Kluwer Academic Publishers, 2002.

40. Y. Shani, C. H. Henry, R. C. Kistler, K. J. Orlowsky, D. A. Ackerman, "*Efficient coupling of a semiconductor laser to an optical fiber by means of a tapered waveguide on silicon*," Applied Physics Letters, pp. 2389-2391 (1989).

41. T. Shoji, T. Tsuchizawa, T. Watanabe, K. Yamada, H. Morita, "*Low loss mode size converter from 0.3um square Si wire waveguides to singlemode fibres*," Electronics Letters **38** (2002).

42. S. Itabashi, H. Fukuda, T. Tsuchizawa, T. Watanabe, J. Takahashi, T. Shoji, K. Yamada, "*Silicon wire waveguide fabrication and microphotonics devices*," OFC 2005, paper OFL1.

43. T. Tsuchizawa, K. Yamada, H. Fukuda et. al., "*Microphotonics Devices Based on Silicon Microfabrication Technology*," IEEE Journal Of Selected Topics In Quantum Electronics **11**, pp. 232 – 240 (2005).

44. S. J. McNab, N. Moll, Y. A. Vlasov, "*Ultra-low loss photonic integrated circuit with membrane-type photonic crystal waveguides*," Opt. Express 11, 2927 (2003).

45. Vilson R. Almeida, Roberto R. Panepucci, and Michal Lipson, "*Nanotaper for compact mode conversion*," Opt. Letters **28**, (2003).

46. W. Bogaerts, D. Taillaert, B. Luyssaert, P. Dumon, J. Van Campenhout, P. Bienstman, D. Van Thourhout and R. Baets, V. Wiaux, S. Beckx, "*Basic structures for photonic integrated circuits in Silicon-on-insulator*", Opt. Express **12**, pp. 1583–1591 (2004).

47. F. Van Laere, G. Roelkens, J. Schrauwen, D. Taillaert, P. Dumon, W. Bogaerts, D. Van Thourhout, R. Baets, "*Compact grating couplers between optical fibers and Silicon-on-Insulator photonic wire waveguides with 69% coupling efficiency*," presented at Optical Fiber Communication Conference 2006, paper PDP15.

48. M. Fan, M. A. Popović, F. X. Kartner, "*High Directivity, Vertical Fiber-to-Chip Coupler with Anisotropically Radiating Grating Teeth,*" in Conference on Lasers and Electro-Optics (CLEO) 2007, paper CTuDD3.

49. "*Silicon Photonics*", published in Intel Technology Journal, May 10, 2004 (http://developer.intel.com/technology/itj/2004/volume08issue02/art06_siliconphoto/p06_coupling.htm)

50. Day, I.; Evans, I.; Knights, A.; Hopper, F. et al, "*Tapered silicon waveguides for low insertion loss highly-efficient high-speed electronic variable optical attenuators,*" OFC 2003, paper TuM5.

51. G. H. Song, W. J. Tomlinson, "*Fourier Analysis and Synthesis of Adiabatic Tapers in Integrated Optics*", JOSA A 9, pp.1289-1300 (1992).

52. A. W. Snyder, J. D. Love, "*Optical Waveguide Theory*" (Chapman & Hall, New York, 1983).

53. P. Bienstman, "*Rigorous and Efficient Modeling of Wavelength Scale Photonic Components,*" Ph.D thesis, Ghent University, 2001.

54. A. Khilo, M. A. Popović, F. X. Kärtner, "*Efficient Planar Fiber-to-Chip Coupler based on Two-Stage Adiabatic Evolution,*" presented at ICONO/LAT Conference, Minsk, Belarus, paper IO2/VIII-1 (2007).

55. C. H. Cox, III, "*Analog Optical Links,*" Cambridge University Press, Cambridge UK (2004).