Design and implementation of a gigabit-rate optical receiver and a digital frequency-locked loop for phase-locked loop based applications

Francis Beaudoin

Department of Electrical and Computer Engineering
McGill University
Montréal, Canada
October 2003

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Master of Engineering

Your file Votre référence
ISBN: 0-612-98510-5
Our file Notre référence
ISBN: 0-612-98510-5

Table of Content

List of Figures

Abstract

The large demand for high-bandwidth communication systems has brought down the cost of optical system components. To be competitive in a crowded market, implementation of the different systems of an optical transceiver on a single chip has become mandatory.

CMOS technologies, especially state-of-the-art processes like the 0.18μm CMOS, permit integration of huge amounts of transistors per millimiter square. Furthermore, deep-submicron CMOS processes have similar RF performances to their tradionnal bipolar equivalent. It is therefore a small footstep to go to congregate high-speed analog circuits with digital cores on a single die.

This thesis adresses two of the building blocks found in an optical communication receiver, namely the analog front-end receiver and a digital frequency-acquisition based clock-and-data recovery circuit. The latter reduces the headcount of bulky passive components needed in the implementation of the loop filter by porting the analog loop to the digital domain. This circuit has been successfully fabricated and tested.

Finally, an optical front-end, comprising a transimpedance amplifier and a limiting amplifier is proposed and fabricated using a standard 0.18μm CMOS process. The speed of this circuit has been pushed up to 5Gb/s. Different techniques have been employed to increase the effective bandwidth of the input amplifier, namely the use of a constant-k filter.

Sommaire

Une demande sans cesse grandissante pour la bande passante fait tomber le prix des diverses composantes des systèmes optiques. Pour être compétitif dans un marché déjà saturé, l'intégration des différents systèmes électroniques des récepteurs optiques est maintenant obligatoire.

Les procédés à la fine pointe de la technologie, comme les transistors CMOS, permettent l'intégration de quantité impressionnante de transistors pour une surface donnée. Ces mêmes technologies commencent de plus en plus à avoir des caractéristiques similaires au traditionnel procédé bipolaire. Des lors, il est possible d'intégrer des systèmes analogiques à hautes vitesses ainsi que des circuits numériques sur une même puce.

Cette thèse présente deux blocs contenus dans les récepteurs optiques, plus précisément un pré-amplificateur et ainsi qu'un circuit numérique d'aide à l'acquisition de fréquence pour les systèmes de boucle asservie en phase. Ce dernier type de circuit réduit substantiellement le nombre de composantes passives requises. Cette réduction est attribuable à la conversion d'une partie de circuit analogique par son équivalent numérique.

Pour conclure, cette thèse discute d'un pré-amplificateur pour circuit optique, composé d'un amplificateur à transimpedance ainsi que d'un amplificateur limiteur. De plus, un prototype, utilisant la technologie CMOS avec une longueur de grille de 0.18um a été fabriqué. La vitesse d'opération maximale à être démontré pour ce circuit est de 5Gb/s. Différentes techniques pour augmenter la bande passante du circuit ont été utilisé, plus particulièrement l'utilisation d'un filtre a impédance constante.

Remerciements

À mes parents, Marie-Paule et Bernard, je leur dédie avec humilité cette thèse. À jamais je ne serai suffisamment reconnaissant de leur soutien et de leur amour.

À mon mentor Michel, pour sa guidance éclairée, son écoute attentive et son propos tempéré.

À tout mes amis d'avoir été présents.

Mon directeur de thèse Mourad El-Gamal, pour cette opportunité.

La Société canadienne de micro-électronique (CMC) et RESMIQ pour leurs supports financiers.

CHAPTER 1

Introduction

1.1 Motivation

With today's high speed networks, where concentrations of voice, video, text and other media is accelerating, capacity limit has never ceased of increasing. The major problem with electrical communication links is the capacity limit. To increase this capacity limit, the carrier frequency needs to be pushed higher in frequency. However, electrical circuitry can hardly cope with frequencies beyond few tens of gigahertz. With the recent invention of laser, the carrier has been changed from an electrical format to an optical format. Commercial communication systems using electrical carriers (RF systems) use frequencies up to 100GHz (mainly for satellite applications). In today's optical links, where the wavelength of the laser is 1.55μm, the carrier frequency is about 192 THz. This is approximately three orders of magnitude larger.

Network backbones are currently built from OC-192 links (10Gb/s), and soon OC-768 links (40Gb/s). With the use of wavelength-division multiplexing, a process of using different wavelengths for each channel, can boost the transmission capacity of the link [Keiser99]. Transfer rates in excess of 1Tb/s can be achieved on a single fibre [Chraplyvy98].

Before getting into the design of an optical receiver, it is important to know the requirements of the system. Optical communication systems are digital communication

systems working at very high speeds. In addition of using digital communication theory to characterize the link, knowledge of RF design, random signals, and transceiver architectures must also be mastered.

1.2 System Architecture

A typical optical system is shown in Figure 1.1. This system is used to convert an electrical stream of information to light. The modulated light is then propagated in a channel (i.e. an optical fibre). At the receiving end, a circuit converts back the light to the original electrical stream.

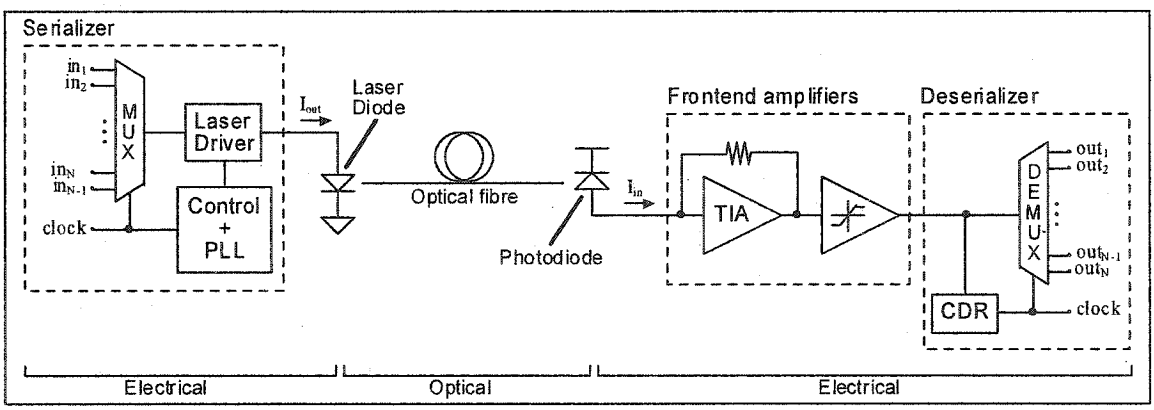


Figure 1.1 Generic optical system.

To produce a light equivalent to the electrical information, the electrical signal needs to be conditioned. The serializer converts a number of low speed parallel data to a high-speed serial stream. It is done using a multiplexer and a phase-locked loop (PLL). The PLL helps to synchronize the sampling of the incoming data by means of a reference clock. A laser driver then creates a current proportional to the serial bit stream. The control block is used to regulate the current conversion. The light produced by the laser diode is proportional to the current injected in it. The light is coupled to a propagation medium (i.e. an optical fibre).

To recover the information, the previous steps are done in a reversed order. The photodiode will create a current proportional to the light applied to it. This current is

converted to voltage using a transimpedance amplifier (TIA). Because the voltage produced by the TIA is still very small, a string of limiting amplifiers (LA) is needed to further amplify the voltage to digital level. By the help of a clock-and-data recovery circuit (CDR), the high-speed bit stream and the clock can be recovered. Finally, a demultiplexer will split the serial bit stream to a low speed parallel bit stream.

In today's modern IC CMOS processes, components characteristics variations can be as large as ±20% over process and temperature. Such high variations can change significantly the operation of a circuit. Good designs should be able to be as insensitive as possible to these extreme changes.

1.3 State-of-the-art frontend designs

In order to achieve low-cost fiber-optic links, CMOS technology is a very good candidate. Both digital and analog circuits can be implemented on the same die when implemented in CMOS. Cost reduction is possible, since the number of packages is reduced (e.g. 1 package for the complete system compared to two or more packages for the front-end and the digital circuitry). Power reduction is also reduced since power-hungry buffers, needed at the chip interface, are not necessary. Power consumption of buffers is particularly high for high bit rates.

CMOS optical receivers with bit-rates varying from 50MBits/s up to 10Gbits/s have already been reported in the literature [Pietruszynski88] [Nakamura98] [Ingels94] [Ingels99] [Tanabe98] [Mitran02] [Kiziloglu02]. CMOS front-ends are by their nature noisier than their bipolar or GaAs equivalent counterpart. Of interest is the transimpedance amplifier (TIA) proposed by Razavi [Razavi00]. Using capacitive feedback, instead of the traditional resistive shunt feedback, the noise generation has been brought to an equivalent level with respect to that of bipolar TIA.

CMOS front-ends are limited in speed because of the relatively large gate capacitances of the transistors. Different techniques can be employed to overcome this bandwidth

limitation. Mohan [Mohan99] proposes using inductors placed in series with the resistance used to produce the amplifier gain. Called "inductive peaking", this method can extend the bandwidth of the amplifier by an extra few percents. Limiting amplifiers also used the inductive peaking technique to broaden their bandwidths. However, the use of passive inductances is prohibited, because of the large number of inductors required. Also, the area of a passive inductance cannot be mitigated (e.g. 400μm x 400μm). Active inductors, built from a single transistor and a resistor, can be used in circuits with good results [Sackinger00]. It should be noted that active inductors are noisier than their passive equivalent. Their use in transimpedance amplifiers is not common, but by they can be used, with proper optimization, in limiting amplifiers.

Other techniques, such as applying a constant-k filter at the input [Beaudoin02], can be used to increase the bandwidth of the amplifier, or to reduce the noise generation of the TIA. A good analysis of the different topologies that can be used in the building a TIA is described in [Park97]. Common-source and common-gate input amplifiers are compared and noise analysis is done.

1.4 Thesis outline

This thesis consists of seven chapters. The first chapter is an introduction that discusses the motivations behind this research. The second chapter discusses the basic metrics used in optical communications. Chapter 3 presents the mathematical derivation of the equations used by a generic transimpedance amplifier. These equations account for the bandwidth and noise of the amplifier. The limiting amplifier is presented, as well as the issues associated with this type of amplifier. A circuit design is presented. Chapter 4 presents the measurement results and methodology. A comparison with the simulation results is done. A simple analysis of the package effect is presented. In chapter 5, the issues associated with the limited locking range of clock-and-data recovery circuits are discussed. A circuit is presented to broaden the acquisition range. Chapter 6 presents the measurement results of the circuit proposed in chapter 5. In addition, the measurement techniques are presented. Finally, chapter 7 concludes this research work and possible future work is proposed.

1.5 Thesis contributions

The following is a summary of the contributions of this thesis:

- The design, implementation and testing of a 5.0Gb/s optical front-end receiver [Beaudoin02].

  A. A method to extend the bandwidth of the amplifier was used. Instead of extending the frequency at the output employing techniques such as inductive peaking, the bandwidth was increased at the input with the help of a constant impedance filter.

  B. An investigation of the causes of the observed discrepancies between the simulation and measured performances was carried out. In addition, a method to easily perform the necessary measurements was presented in details.

- The design, implementation, and testing of a digital frequency locked-loop circuit used to broaden the acquisition range of PLL based circuits in general.

  A. A comparative study of the different types of loops available to simplify the frequency locking of PLL based circuits is presented. The theory and the governing equations of a digital frequency locked-loop (DFLL), based on its analog equivalent, are discussed.

  B. A DFLL implementation example is described. Each building block employed in the DFLL is discussed, with its associated requirements.

  C. The measurements results from a fabricated prototype are summarized, and the differences from simulations are thoroughly analyzed, along with a careful study of the on-chip loop dynamics.

1.6 References

[Keiser99] G. E. Keiser, A review of WDM technology and applications, Opt. Fiber Technol, vol. 5, pp.3-39, Jan 1999.

[Chraplyvy98] A. R. Chraplyvy, R. W. Tkach, "Terabit/second transmission experiments," IEEE J. Quantum Electron., vol. 34, pp. 2103-2108, Nov. 1998.

[Beaudoin02] F. Beaudoin, M. N. El-Gamal, "A 5-Gbit/s CMOS optical receiver frontend," The 2002 45th Midwest Symposium on Circuit and Systems, Volume 3, pp. 168-171, August 2002.

[Mitran02] P. Mitran, F. Beaudoin, and M. El-Gamal, "A 2.5 Gbit/s CMOS Optical Receiver Frontend," Proc. IEEE International Symposium on Circuits and Systems, Phoenix, May 2002.

[Pietruszynski88] M. Pietruszynski, J. M. Steininger, E. J. Swanson, "A 50-Mbit/s CMOS Monolithic Optical Receiver," IEEE J. Solid-State Circuits, vol. 23, pp. 1426-1432, Dec. 1988.

[Nakamura98] M. Nakamura, N. Ishihara, Y. Akazawa, "A 156-Mb/s Optical Receiver for Burst-Mode Transmission," IEEE J. Solid-State Circuits, vol. 33, pp. 1179-1187, August 1998.

[Ingels94] M. Ingels, G. van der Plash, J. Crols, and M. Steyaert, "A CMOS 18 THz Ohm 240Mb/s Transimpedance Amplifier and 155 Mb/s LED-driver for Low Cost Optical Fibre Links," IEEE J. Solid-State Circuits, vol. 29, pp. 1552-1559, Dec 1994.

[Ingels99] M. Ingels, M. Steyaert, "A 1Gb/s 0.7μm CMOS Optical Receiver with Full Rail-to-Rail Output Swing," IEEE J. Solid-State Circuits, vol. 34, pp. 971-977, July 1999.

[Tanabe98] A. Tanabe, M. Soda, Y. Nakahara, T. Tamura, K. Yoshida, A. Furukawa, "A Single-Chip 2.4-Gb/s CMOS Optical Receiver IC with Low Substrate Cross-Talk Preamplifier," IEEE J. Solid-State Circuits, vol. 33, pp. 2148-2153, Dec. 1998.

[Razavi00] B. Razavi, "A 622Mb/s 4.5pA/rt Hz CMOS Transimpedance Amplifier," ISSCC Digest of Technical Papers, pp. 162-163, Feb. 2000.

[Park97], S. M. Park, C. Toumazou, "Gigahertz low noise CMOS transimpedance amplifier," Proc. of IEEE International Symposium on Circuits and Systems, vol. 1 pp. 209-212, Jun. 1997.

[Mohan99] S. S. Mohan, T. H. Lee, "A 2.125 Gbaud 1.6 kΩ transimpedance preamplifier in 0.5 μm," Proc. of IEEE CMOS Custom Integrated Circuits, pp. 513 -516, May 1999.

[Sackinger00] E. Sackinger, W. C. Fischer, "A 3 GHz, 32 dB CMOS limiting amplifier for SONET OC-48 receivers," Tech. Digest of International Solid-State Circuits Conference, pp. 158 -159, Feb. 2000

[Kiziloglu02] A. K. Kiziloglu, K. Ty Yoon, F. Jr. Williams, M. R. Sandor, "Front-end CMOS chipset for 10 Gb/s communication," IEEE Radio Frequency Integrated Circuits (RFIC) Symposium, pp. 93 -96, Jun. 2002

CHAPTER 2

Optical Receivers Basics

This chapter covers the basics about channel-coding and how it relates to optical receivers. A derivation of the optimum bandwidth of the receiver amplifier is also performed. Some important metrics used in optical communication systems, such as bit-error rate and sensitivity are introduced. Finally, a brief overview of an optical link budget analysis and of a receiver performance plot are presented.

2.1 Channel coding

Formatting the information to be transmitted over an optical link is a critical consideration in system design. By using proper coding, the system speed and transmission length can be increased substantially. Optical communication systems transmits information mainly in a digital form. At the receiving end, a decision circuitry must be able to extract precise timing information from the received signal for data recovery. Such timing information is normally incorporated in the transmitted signal. It is generally achieved by rearranging or adding extra data in the stream. This process is called channel-coding.

From Shannon's channel-coding theory [Shannon48], it can be shown that channel-coding introduces coding gain. Coding gain is a measure of the increase in signal-to-noise ratio (decrease in bit-error rate) for different types of coding schemes. An increase in the

coding gain can be used to increase the length of a transmission link, or relax the design constraints for the receiver. A detailed mathematical analysis of channel-coding falls outside the scope of this thesis.

Three basic channel-coding methods for binary level signals are used in optical systems: non-return-to-zero (NRZ), return-to-zero (RZ), and phase-shift-keying (PSK). Both the NRZ and RZ coding schemes are simple extensions of the simplest form of encoding: on-off keying (OOK).

The simplest method to encode a binary stream is the NRZ coding. The signal representing a logic-one is on for the complete bit period T_b. Conversely, a logic-zero is off for the same period (Figure 2.1 (a)). Long strings of logic-zeros and ones need a highly stable clock in the clock-recovery circuit in order to be recovered properly. This results in an increase in both cost and complexity. However, by using scrambling or block coding (section 2.1.1), it is possible to reduce the number of consecutive zeros and ones. This is called the DC balance strategy, as explained in section 2.1.2.

The second type of line coding is the RZ, with a variety of code types existing. Figure 2.1 (b) shows the unipolar RZ code. A logic-one is represented by a half-period being at one, and the other half being off. A logic-zero is represented by no signal during the bit period. If long strings of zeros are sent, the same synchronization problem exists as for the NRZ coding. A type of coding not having this limitation is the biphase or Manchester coding.

Biphase or Manchester coding is a combination of RZ and PSK coding. From Figure 2.1 (c), one can see that a logic-zero is phase inverted compared to a logic-one. This coding has the advantage of having timing information (clock) embedded in the data stream, making timing recovery an easier task. However, the main drawback of RZ coding is the amount of bandwidth required. Compared to NRZ, RZ requires twice the amount of bandwidth. On the other hand, using a reduced bandwidth, NRZ requires very stable oscillators at the receiver. To improve the reliability of the communication system, extra

data need to be added to the stream: this includes block coding, DC balance strategies and forward-error correction, which are briefly described in the following subsections.

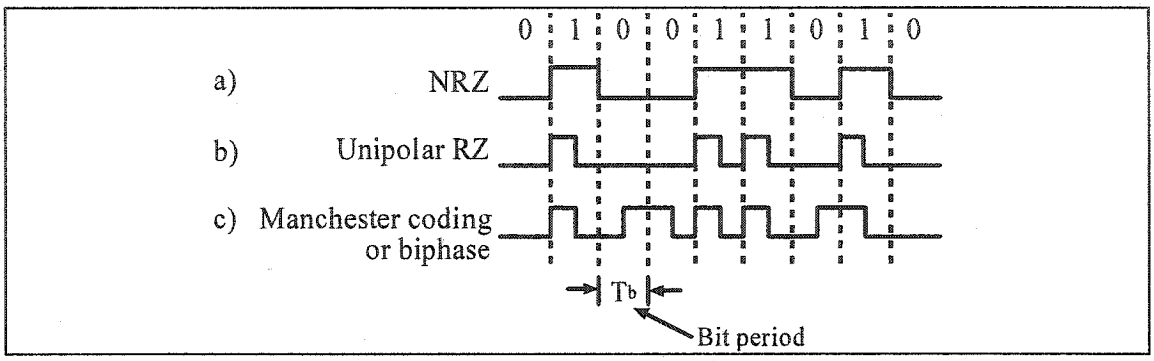


Figure 2.1 Different line coding formats.

2.1.1 Block coding

Block coding is based on using redundant binary codes. Such codes transform a block of m binary bits to a longer block of length n (n>m). Block coding also reduces the disparity between the number of zeros and ones (it tends to have the same number of zeros and ones). This is a desired feature, since long strings of ones and zeros are eliminated. At the expense of an increase in the required bandwidth, block coding provides an adequate way of introducing timing information in the data stream.

2.1.2 DC balance strategy

Block coding is a powerful way to change the power spectrum shape of the information stream, while transmitting approximately the same amount of information. By having a limited number of consecutive zeros or ones, the necessary complexity of the clock recovery circuit will be reduced. In addition, AC coupling between the different stages of the receiver is possible since the low frequency components are present. DC balance is a desired feature for digital transmission system.

2.1.3 Error correction

To improve the reliability in optical communication systems, forward error correction (FEC) is added to reduce the overall bit-error-rate. The SONET standard uses Reed-

Solomon codes to implements FEC. With a little overhead (normally 7.1% for RS (255,239) code), the BER can be improved by more than 10000 times [Sklar01].

2.2 Metrics used in optical communication systems

Many standards exist to quantify the performance of a system. The following paragraphs describe the main metrics used for optical communication receivers. Receivers are mostly quantified by the amount of noise they introduce, or by the amount of noise they are capable of dealing with. Time domain analysis is not critical for receiver amplifiers. However, time domain analysis (e.g. jitter) becomes a critical issue for clock-and-data recovery circuits.

2.3 Signal-to-noise ratio

One of the most important measures in analog systems in general is the signal-to-noise ratio (SNR). As the name suggests, it is a measure of how stronger is the signal power compared to the noise power at a node in a system. It can be expressed as a ratio of powers (in this case: optical powers), or also as a ratio of currents or voltages, depending on the system under analysis.

Because optical frontends amplify currents, it is more practical to define the SNR as a ratio of currents at the input node. The corresponding optical power actually received by the photodiode can then be estimated. All the noise introduced by the optical link and by the photodiode itself can be accounted for using a single equivalent current noise source.

2.4 Bit-error-rate (BER)

Several standard methods to measure the rate of error occurring in digital streams exist. A common one is to divide the number of errors occurring (N_e) by the number of received bits (N_b) each over a period of time (t) (equation 2.1). This gives the bit-error rate (BER). For example, for a BER of 10^{-9}, one bit will be erroneous for every one billion of bits sent.

$$BER = \frac{N_e}{N_b t}$$

Eq. 2.1

Since the BER is the probability of getting an erroneous bit, it is important to know the probability distribution of the received bits.

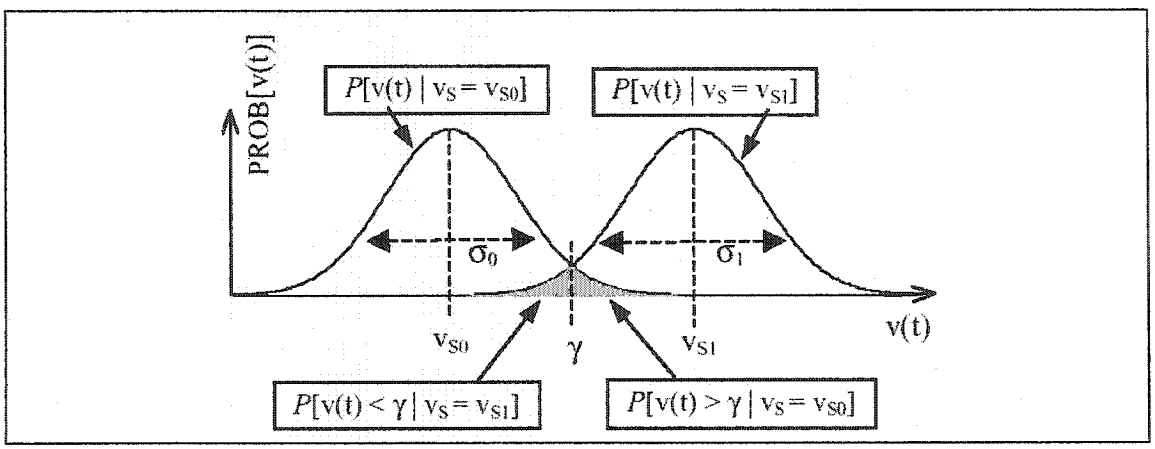


Figure 2.2 Probability distribution for a binary signal.

Figure 2.2 shows the probability distributions of a logic-zero and a logic-one at the output of a receiver. The threshold voltage γ is normally fixed at the middle of the intersection of the probability distribution functions in order to achieve optimum sampling. If the input voltage *v(t)* is greater than γ, then the system treats the received signal as a logic-one. On the other hand, if the input voltage *v(t)* is smaller than γ, the system treats the received signal as a logic-zero.

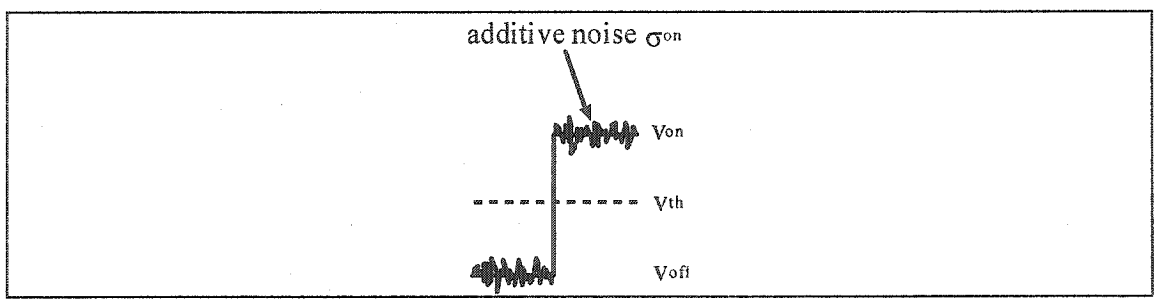


Figure 2.3 Signal with additive noise.

The BER is directly related to the signal-to-noise ratio of the receiving system. From Figure 2.3, we can define the parameter Q:

$$Q = \frac{v_{on} - v_{th}}{\sigma_{on}} \approx \frac{1}{2}\frac{S}{N}. \qquad \textbf{Eq. 2.2}$$

The variance σ_{on} is the RMS amount of noise present for a logic-one and v_{on}-v_{th} is the peak-to-threshold voltage for a logic-one. The same relationship can be defined from a logic-zero. Assuming a Gaussian probability density function, it is possible to compute the probability of error for which a logic-zero will be interpreted as a logic-one. In this case, the probability of error that noise will exceed the threshold voltage v_{th} is given by:

$$P_0(v_{th}) = \int_{v_{th}}^{\infty} p(y|0)dy = \frac{1}{\sqrt{2\pi}\sigma_{off}} \int_{v_{th}}^{\infty} \exp\left[-\frac{(v - v_{off})^2}{2\sigma_{off}^2}\right] dv, \qquad \textbf{Eq. 2.3}$$

where σ_{off}^2 is the noise variance for a logic-zero and v_{off} is the corresponding voltage for a logic-zero. Similarly, the probability of error of a logic-one is given by:

$$P_1(v_{th}) = \int_{-\infty}^{v_{th}} p(y|1)dy = \frac{1}{\sqrt{2\pi}\sigma_{on}} \int_{-\infty}^{v_{th}} \exp\left[-\frac{(v_{on} - v)^2}{2\sigma_{on}^2}\right] dv. \qquad \textbf{Eq. 2.4}$$

If the probabilities of a logic-zero and a logic-one are equally likely, then, using equations 2.3 and 2.4, the BER becomes [Keiser00]:

$$BER = P_e(Q) = \frac{1}{\sqrt{\pi}} \int_{Q/\sqrt{2}}^{\infty} e^{-x^2} dx. \qquad \textbf{Eq. 2.5}$$

Performing the integration, equation 2.5 can be approximated by 2.6 [Keiser00]:

$$BER = \frac{1}{2}\left[1 - erf\left(\frac{Q}{\sqrt{2}}\right)\right] \approx \frac{1}{\sqrt{2\pi}}\frac{e^{-Q^2/2}}{Q}. \qquad \textbf{Eq. 2.6}$$

Figure 2.4 shows the dependence of the BER given by equation 2.6 on the parameter Q. For example, a Q of 7.035 is required to get a BER of 10^{-12}. The corresponding SNR is twice the value of Q, i.e. 14.07 or 11.48dB. So as predicted, in order to have a lower bit-error rate, the required system SNR needs to be increased.

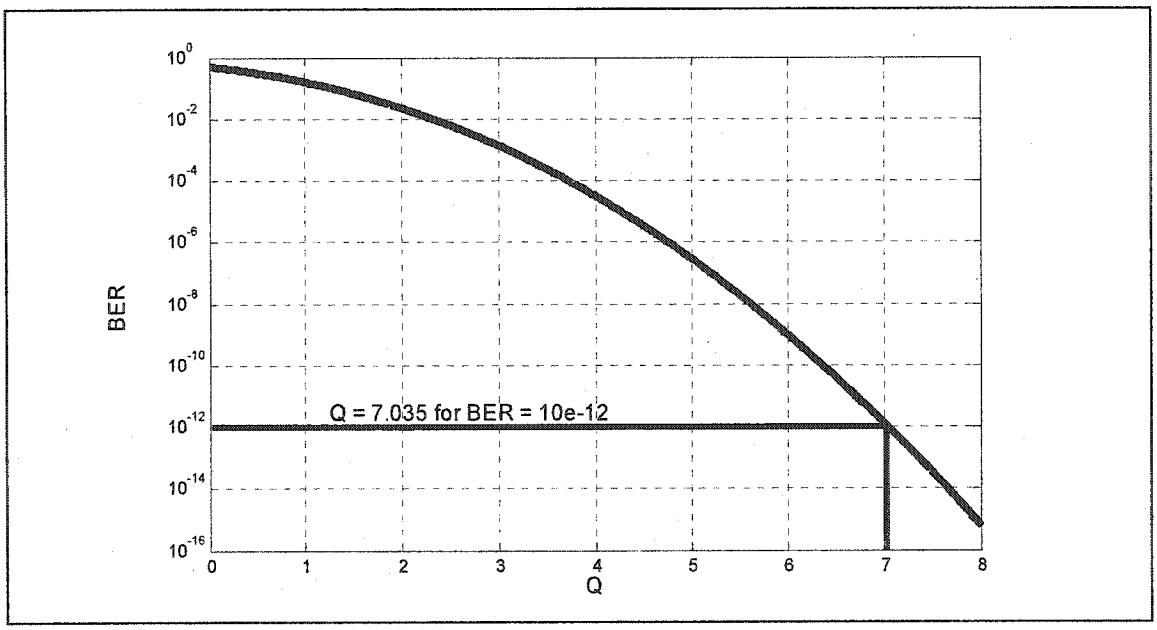


Figure 2.4 Bit-error-rate versus Q factor.

2.5 *The extinction ratio*

The extinction ratio is an effect introduced by laser diodes at the transmitter side, that are biased close to their lasing threshold (Figure 2.5) to reduce chirping. The extinction ratio has a significant impact on the receiver sensitivity [Keiser00]. The extinction ratio is defined as the ratio of a logic-one power level (P_1) relative to a logic-zero power level (P_0):

$$e_r = \frac{P_1}{P_0}. \qquad \text{Eq. 2.7}$$

Instead of hard driving the laser diode i.e. $I_{laser} = 0$ for logic-zero and $I_{laser} = I_{max}$ for logic-one, it is preferable to chose the biasing current to be close to the lasing threshold (Figure 2.5).

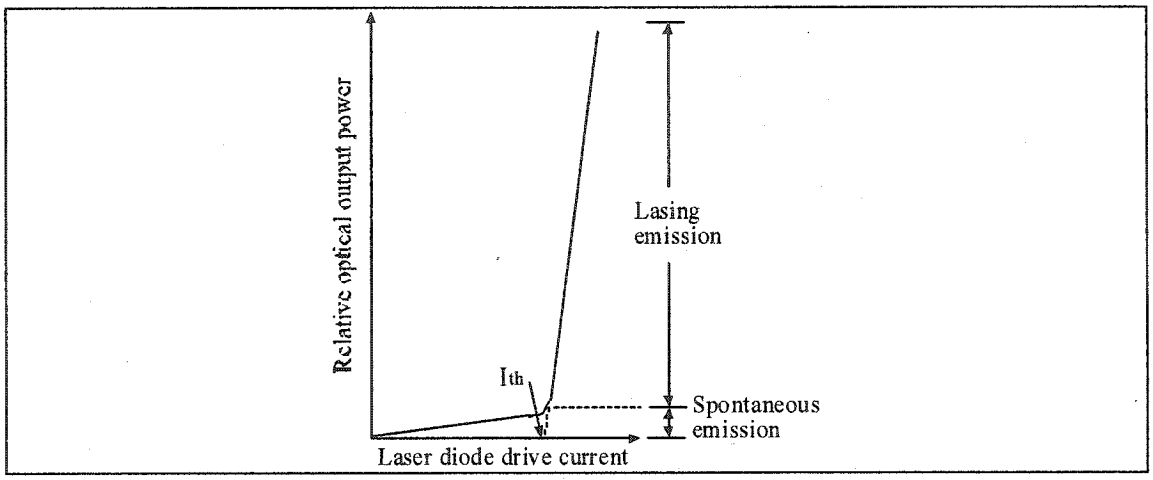


Figure 2.5 Optical power versus the injected current in a laser diode.

2.5.1 Chirping

Laser diodes may experience dynamic change of frequency when the injection current is directly modulated. This leads to significant dispersion effects in the optical fibre and an increase in the eye closure at the receiver [Keiser00]. The effect can be more easily seen in the time domain as it introduces jitter. If the laser diode is biased above the lasing threshold, the injection current becomes much less susceptible to direct modulation.

Having the laser diode always biased, some optical power is now emitted for a logic-zero. Because the signal now available excursion is reduced, a power penalty is introduced due to the smaller signal-to-noise ratio available at the receiver. This penalty is defined by:

$$P_{loss-e_r} = \frac{e_r + 1}{e_r - 1}. \qquad \text{Eq. 2.8}$$

For a typical power penalty of 1dB, the extinction ratio needs to be 8.7 or 9.4dB. The power penalty introduced by the extinction ratio can be reduced at the price of an increase in chirping. However, the designer does not have control over this parameter since this is a system design issue.

2.6 Sensitivity

In optical communication systems, sensitivity is a measure of the average power needed to achieve a specified bit-error rate. For example, the SONET standard for OC-48 requires a sensitivity of -18dBm for the receiver for links shorter than 15km. The objective here is to achieve a BER of less than 10^{-10} with the specified minimum power.

Sensitivity can be expressed as an average power, or as an optical modulation amplitude (OMA) in W_{pp}. Measuring the peak-to-peak power amplitude of a laser diode can be quite difficult, and requires expensive equipments. However, measuring the average power can be done more easily. The sensitivity is given by [MAXIM01]:

$$Sensivity = 10\log\left(\frac{\overline{i_n^2}SNR(r_e+1)}{2\rho(r_e-1)}1000\right)dBm, \qquad \textbf{Eq. 2.9}$$

where $\overline{i_n^2}$ is the average integrated noise, SNR is the required SNR for a given BER, ρ is the photodiode responsitivity in (A/W), and r_e is the extinction ratio. The sensitivity can also be expressed in terms of the OMA:

$$Sensivity = 10\log\left(OMA\frac{(r_e+1)}{2(r_e-1)}1000\right)dBm \qquad \textbf{Eq. 2.10}$$

Sometimes, sensitivity is expressed as the smallest current that can be detected and amplified by a transimpedance amplifier. Rewriting equation 2.9 and expanding the SNR, the minimum current corresponding to a given sensitivity is:

$$I_{signal} = \frac{2\rho(r_e - 1)}{\overline{i_n^2}(r_e + 1)1000} 10^{Sensitivity/10}$$ Eq. 2.11

2.6.1 Noise equivalent bandwidth

Before computing the required receiver sensitivity, it is important to be able to estimate the correct amount of integrated noise. The concept of noise equivalent bandwidth (NEB) is explained in Figure 2.6 [Johns97].

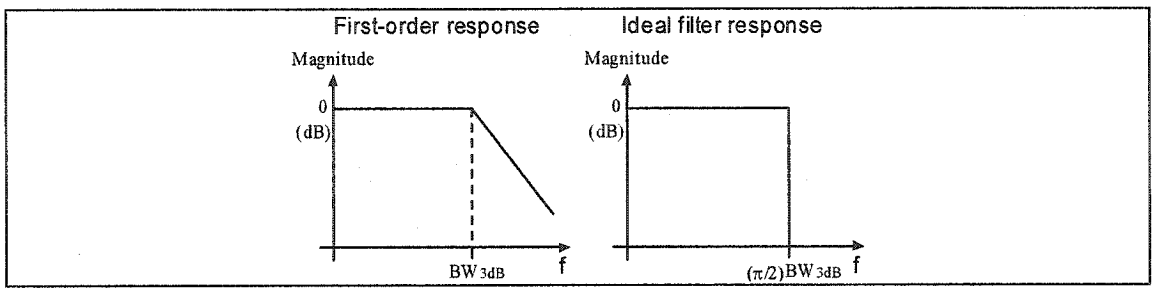


Figure 2.6 First-order response and noise equivalent bandwidth

The noise contribution of an amplifier at high frequency is low because of its decreasing gain response. The NEB is the bandwidth of an ideal rectangular low-pass filter that would produce the same amount of noise as the original amplifier. The NEB is found by integrating the magnitude response of an amplifier over the complete band of frequency. For a first-order system or a system with a dominant single pole, the magnitude response is described by:

$$|A(jf)| = \frac{1}{\left(1 + \left(\frac{f}{BW_{3dB}}\right)^2\right)^{1/2}}$$ Eq. 2.12

where BW_{3dB} is the 3 dB bandwidth of the system. Applying a white source of noise V_{nw} to equation 2.12, the total output noise RMS value of the system would be:

$$V_{no(rms)}^2 = \int_0^\infty \frac{V_{nw}^2}{1 + \left(\frac{f}{BW_{3dB}}\right)^2} df = V_{nw}^2 BW_{3dB} \arctan\left(\frac{f}{BW_{3dB}}\right)\Bigg|_0^\infty = \frac{V_{nw}^2 BW_{3dB}\pi}{2}$$ Eq. 2.13

If this same noise signal is applied to an ideal rectangular low-pass filter, the total output noise would equal:

$$V_{ideal}^2 = \int_0^{BW_{ideal}} V_{nw}^2 df = V_{nw}^2 BW_{ideal} = V_{nw}^2 NEB. \qquad \textbf{Eq. 2.14}$$

Finally, equating 2.13 and 2.14, the noise equivalent bandwidth of a first-order system is therefore given by:

$$NEB = \frac{\pi}{2} BW_{3dB}. \qquad \textbf{Eq. 2.15}$$

Thus, the total amount of noise present at the output of an amplifier having a first-order frequency response is equivalent to the noise of an ideal rectangular low-pass filter with a 3 dB bandwidth 1.57 times higher than the amplifier signal bandwidth.

The photodiode responsitivity can change significantly the sensitivity of a system. For example, a typical PIN photodiode has a responsitivity of 0.5 A/W to 0.95 A/W. For avalanche photodiodes, responsitivity is in the range of 5-10 A/W. By simply using an avalanche photodiode instead of a PIN photodiode, the sensitivity can be improved by more than 10 dB (10x). However, the noise produced by the avalanche photodiode is substantially higher and needs to be added to the input-referred current noise term ($\overline{i_n^2}$).

Finally, the power loss introduced by the extinction ratio also plays an important role in sensitivity computation. However, the designer does not have control over its value.

2.7 Signal spectra

Prior to the receiver design, it is mandatory to know the power spectrum of the signal to be amplified. By knowing the signal power spectrum, the designer will be able to carefully optimize the design for low noise operation, while respecting the bandwidth

requirements. The receiver bandwidth has a significant impact on the sensitivity, and on the bit-error-rate.

The most common line-coding used in optical communication systems is the NRZ coding, because it requires low bandwidth (half of the RZ coding). Because NRZ coding normally includes DC balancing, the low frequency of the spectrum is shaped differently than the regular NRZ coding. We can still work with the unmodified NRZ spectrum since we are mainly interested by the high frequency component. Before finding the optimum receiver bandwidth, we need to have a mathematical representation for the NRZ random stream data.

2.7.1 Autocorrelation of a random NRZ data stream

The Wiener-Khinchin theorem states that, for a random process that is at least wide-sense stationary, the power spectral density of the process is equal to the Fourier transform of its autocorrelation function. Using this theorem, we can compute the power spectral density of a random NRZ data stream by first determining the autocorrelation function, and then calculating its Fourier transform.

The autocorrelation of a function can be computed by multiplying it by a time-shifted version of itself. The later result is then integrated over all values of time shifts. Autocorrelation is defined by:

$$R_{XX}(\tau) = \int_{-\infty}^{\infty} X(t)X(t-\tau)d\tau \qquad \text{Eq. 2.16}$$

For random data, the limit of integration for the autocorrelation function can be changed from $+\infty$ and $-\infty$ to $+T_b$ and $-T_b$, where T_b is the period of a single bit [MAXIM01]. Using this fact, the autocorrelation computation is greatly simplified, leading to the function shown graphically in Figure 2.7.

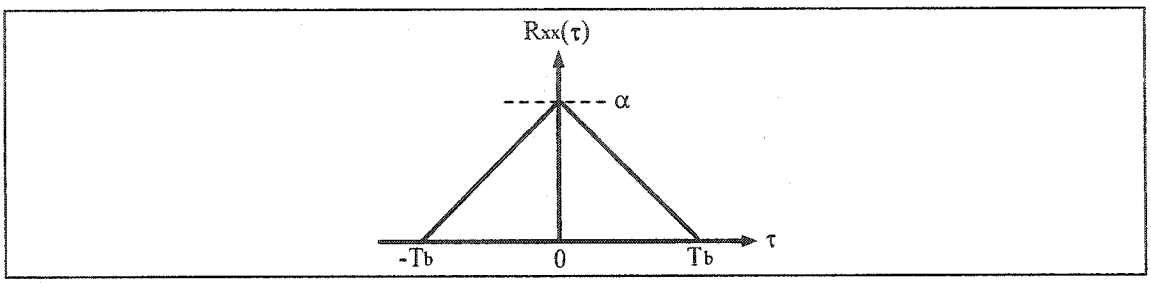


Figure 2.7 Autocorrelation of a random NRZ data stream.

The power spectral density is then found by applying the Fourier transform on the triangular function of Figure 2.7. The result is given by:

$$PSD = \alpha T_b \left[\frac{\sin(\pi T_b f)}{\pi T_b f} \right]^2, \qquad \text{Eq. 2.17}$$

where f is the frequency, and α is a normalization coefficient.

A graphical representation of the resulting power spectral density is shown by Figure 2.8. The y-axis is the normalized power, while the x-axis is the frequency normalized to the data rate (frequency/data rate). It is worth noting that no power exists at any integer multiple of the data rate frequency. In addition, the content of frequency extends to $+\infty$, but has no significant impact above 2 times the bit rate frequency.

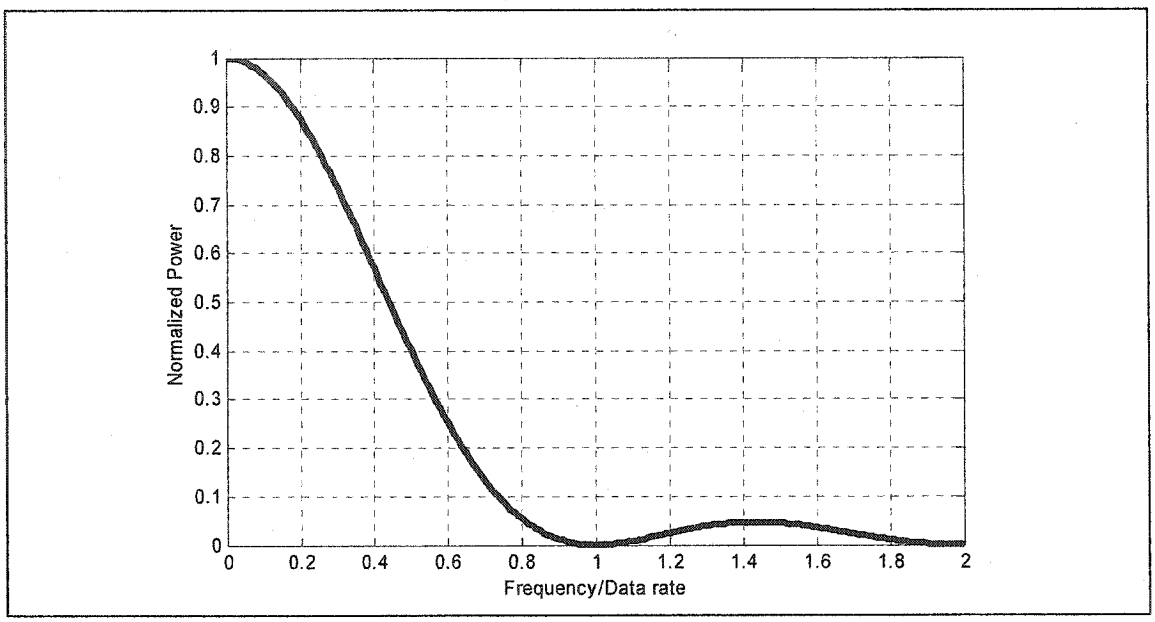


Figure 2.8 Power spectrum of random NRZ signal.

Knowing the power spectrum of a NRZ data stream, it is now possible to find the optimum bandwidth necessary for the system receiver.

2.8 Optimum bandwidth

As described in section 2.7, the power spectrum of a NRZ data stream extends to infinity. However, real systems have a limited bandwidth. Therefore, we have to choose a cutoff frequency to limit the bandwidth of an optical receiver.

The lowest possible cutoff frequency equals half the bit rate. This comes from the fact that the fastest transition in a NRZ stream is a repeating zero-one pattern, creating a square wave with half the bit rate frequency. However, to limit inter-symbol interference and account for possible variations, the cutoff frequency needs to be higher. On the other hand, a cutoff frequency being too high will lead to a large amount of integrated noise at the receiver, degrading its performance.

2.8.1 Inter-symbol interference

Inter-symbol interference (ISI) is introduced by the dispersion effect in the optical fibre, and by the transfer function of the receiver amplifier. Dispersion comes from the non-constant group delay of a transmission medium or a filter. ISI introduces significant amount of jitter at the receiving end. Dispersion introduced by the optical fibre link can be compensated using dispersion-shifted fibres, or dispersion-compensated fibres. Again, the receiver designer does not have control over this since this is a system design issue. As will be described in section 3.3, the transfer characteristic of the receiving amplifier has a significant impact on dispersion. Also, if the cutoff frequency of the amplifier is made too small, significant ISI will be produced since not enough power will be present to accurately reconstruct the original signal.

It is interesting to observe the cumulative (integrated) power spectrum of a random NRZ data stream, as that shown in Figure 2.9.

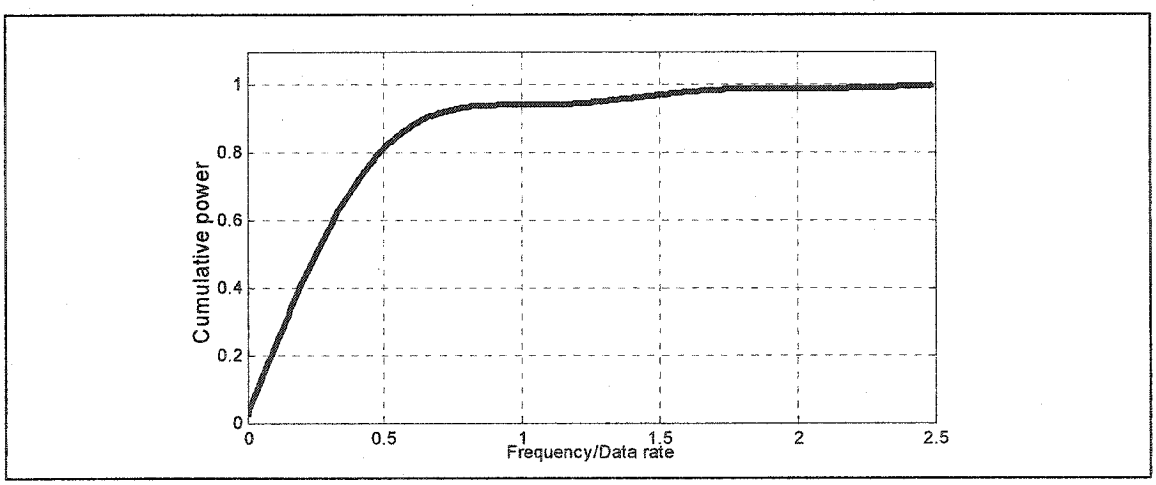


Figure 2.9 Cumulative power spectrum.

The majority of the signal power is included in frequencies lying between 0 to 0.9 times the data rate. At a frequency around 1 times the data rate, the cumulative signal power levels off. It then starts to increase by a small amount at frequencies between 1.2 to 1.8 times the data rate. After frequencies which are 2 times the data rate, 99.9% of the signal power is already recovered. Table 2.1 illustrate the gain signal power values versus the bandwidth.

Table 2.1 Tabulated values for cumulative power in a random NRZ signal.

Frequency / Data rate	Cumulative % of Total power
0.5	81.4%
0.6	87.8%
0.7	91.7%
0.8	94.3%
0.9	95.0%
1.0	95.1%
1.1	95.2%
1.2	95.5%
1.3	96.2%
1.4	97.1%
1.5	98.1%

The optimum bandwidth to be used with a NRZ signal is a trade-off between inter-symbol interference and SNR. If the bandwidth is too low, the signal will be distorted and ISI will be severe. However, the noise contribution will be reduced. Alternatively, if the bandwidth is excessively large, no ISI will appear at the output of the amplifier. However, the noise contribution will be high due to the wide bandwidth. There exists an optimum bandwidth that guarantees to have the lowest noise and minimum ISI.

The optimum bandwidth can be found by optimizing the SNR as a function of frequency. As explained earlier, the minimum bandwidth allowed is 0.5 times the data rate. As the cutoff frequency increases, the signal power shift increases non-linearly, then changes by negligible amounts for high frequencies. On the other hand, the noise power (assumed white) increases linearly as the frequency increases. For example, by increasing the bandwidth by 33%, the signal power increases only by 1.5% while the noise power would increase by 33%. Thus, for an increase of 33% in bandwidth, the SNR becomes 0.76 times the original SNR, resulting in a reduction of the BER by a factor of more than 7000.

The maximum SNR is found by differentiating the SNR function with respect to frequency, then equating to zero.

$$SNR|_{MAX} = \frac{d}{df}\frac{\alpha T_b\left[\frac{\sin(\pi T_b f)}{\pi T_b f}\right]^2}{\overline{i_n^2} f} = 0 \qquad \text{Eq. 2.18}$$

From equation 2.18, it can be shown that the optimum bandwidth is **0.75 times the data rate frequency**. Since the BER is directly related to the SNR, it is interesting to plot the relationship between the BER versus the bandwidth of the system (Figure 2.10).

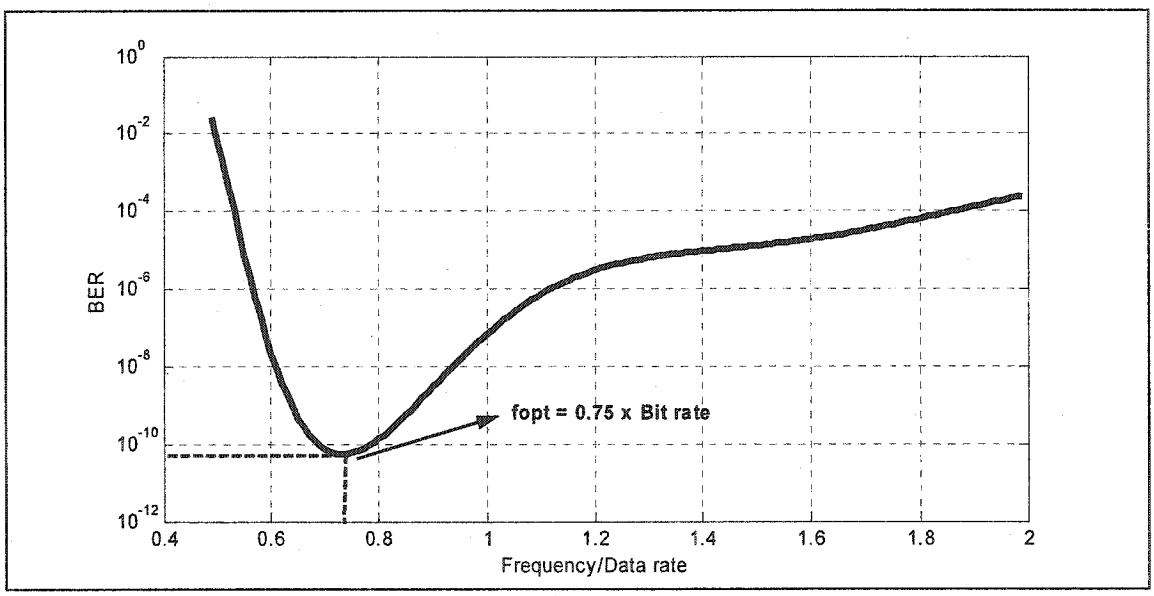


Figure 2.10 Optimum bandwidth for a NRZ system.

2.9 Link budget

The system level performance of an optical communication system can be easily quantified in terms of a link budget. A link budget being either tabulated or plotted provides a quick summary of the system constraints. All link budgets have three entries: one for the transmitter, one for the propagation path, and one for the receiver. A simple graphical link budget for optical power is shown in Figure 2.11. The transmitter power is 0 dBm or 1 milliwatt. Over the optical fibre, 20 dB of power is lost (fibre attenuation,

fibre splices, couplers, etc.). The power arriving at the receiver is -20 dBm. Assuming a receiver sensitivity of -25 dBm, the link margin will be +5 dBm. The link margin is the amount of headroom available at the receiver to meet the system BER. For example, if some fibre splices introduce more loss than predicted, the system will still be able to operate properly and within specification.

Link budgets are normally more complex than this simple example. They usually account for all the second order effects that can affect the operation of the link. A link budget can also be written for noise power. These budgets are helpful for determining if a specific component in a system would introduce too much noise and degrade the BER. Noise budget for optical communication systems are similar to noise budgets in RF systems [Alexander97].

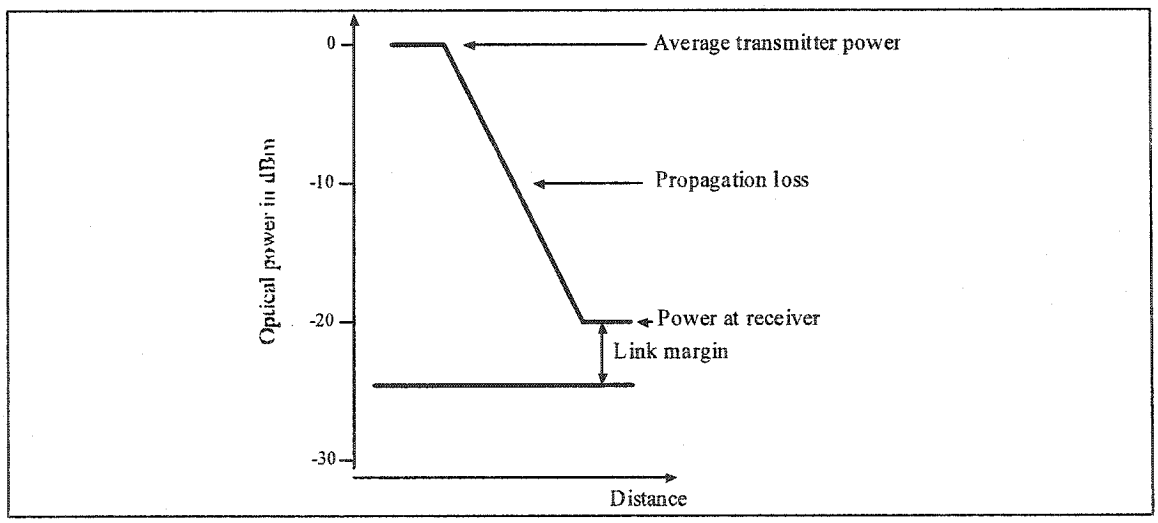


Figure 2.11 Example of a simple optical link budget.

2.10 Receiver performance plot

Digital systems are characterized by the amount of erroneous bits sent in a given time frame (BER). Similar to analog receivers, a plot showing the performance for an optical communication receiver can be generated. The frontend components of receivers, such as transimpedance amplifiers, do suffer from dynamic range overload. Such overloads can be quantified using the same analysis used in RF communication systems. For example, a

large input power applied to a transimpedance amplifier will saturate and distort the output signal.

BER is to digital systems what SNR is to analog systems. Figure 2.12 shows the BER of a receiver as a function of the optical input power. For very low input power, the BER is limited because of the low SNR due to insufficient power at the input. To meet the user quality-of-service (QoS) requirements, the receiver needs to be operated within its dynamic range. Operation in this region will produce very low BER, and sometime error-free operation of the receiver. For large input powers exceeding the dynamic range, the BER will increase. This is owed to the overload and signal compression in the transimpedance amplifier.

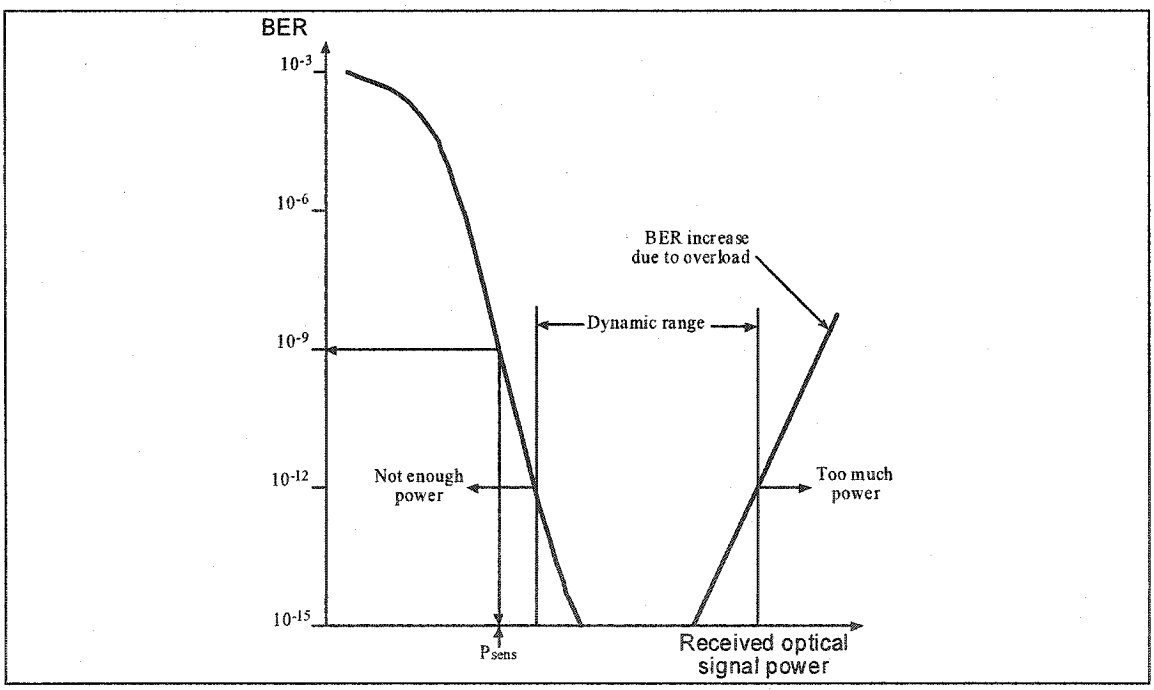


Figure 2.12 Typical receiver performance plot.

Consequently, the optical power received by the receiver should always remain within its dynamic range. Optical amplifiers or attenuators can be added at the input of the photodiode to either amplify or attenuate the optical power. However, if the receiver is to operate with wide input current variations, the receiver can be modified to cope with this

by changing the gain of the transimpedance amplifier [Alexander97]. However, this requires extra circuitry and is more complex to design.

2.11 Conclusion

Metrics used to characterize the receiver of an optical communication system were defined in this chapter. These metrics relate the signal power to the noise power in order to quantify the reliability of a link. From the link budget, one can learn many things about the overall system design and performance. It gives an overall system insight on whether the system will meet comfortably, marginally, or not the requirements. In addition, optimum bandwidth for the receiver amplifier resulting in the best operation was demonstrated to be at 0.75 times the bit rate frequency. This is an important starting point in designing the transimpedance and limiting amplifiers, both being at the front of an optical receiver.

2.12 References

[Shannon48] C. E. Shannon, "A Mathematical Theory of Communication," BSTJ, vol 27, pp.379-423, 623-657, 1948.

[Sklar01] B. Sklar, "Digital Communication, Fundamental and Applications," Second Edition, Upper Saddle River: Prentice Hall, 2001.

[Keiser00] G. E. Keiser, "Optical Fiber Communications," Third Edition, New York: McGraw-Hill, 2000.

[Johns97] D. A. Johns, K. Martin, Analog Integrated Circuit Design. New York: Wiley, 1997.

[Alexander97] S. B. Alexander, Optical Communication Receiver Design. IEE Telecommunications Series, Volume 37, 1997.

[MAXIM01] Maxim, HFAN-03.0.0: "Accurately Estimating Optical Receiver Sensitivity," www.maxim-ic.com, Application Note 607, Oct. 2001.

CHAPTER 3

Transimpedance and Limiting Amplifiers Design

3.1 Receiver front-end

Even if the optical communication receivers treat digital information, analog techniques are needed to amplify the weak signal (around -20dBm of optical power) of the photodiode. Low noise analog techniques, similar to the one used in RF designs, are needed to optimize the amplifiers design.

The basic receiver front-end is composed of a transimpedance amplifier (TIA) followed by a limiting amplifier (LA) (Figure 3.1). The role of the TIA is to transform the weak photodiode current to a voltage. One might think that using a simple resistor to perform the current-to-voltage conversion could be simpler than using a complex amplifier. This approach would not work well. The transimpedance gain needs to be stabilized over frequency. This is done by using an inverting amplifier and closing the loop with a resistive network.

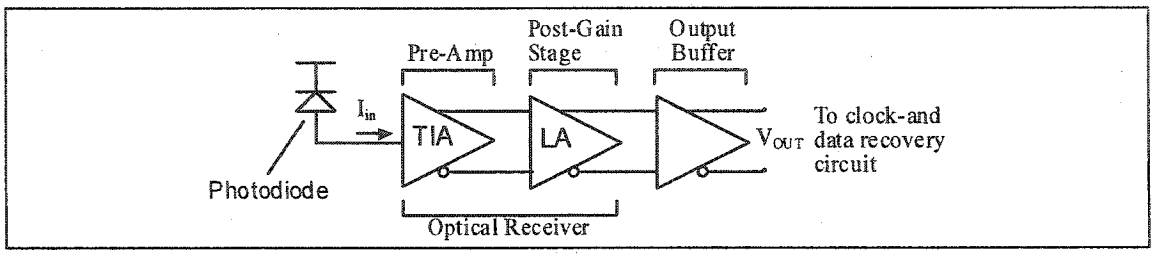


Figure 3.1 Typical front-end amplifier.

A clock-and-data recovery circuit normally follows the receiver front-end. Since this block accepts digital level voltages, the receiver front-end should also output digital voltage levels. One solution would be to have a TIA with a very large transimpedance gain. This solution is not viable because the designer will quickly reach the limit of the technology. As described in section 3.3.1, the bandwidth of a TIA is inversely proportional to the transimpedance gain. As a result, it is not possible to optimize both the gain and the bandwidth. To meet the TIA bandwidth requirements, the transimpedance gain can be reduced. In order to make a TIA electrically compatible with a clock-and-data recovery/demultiplexer circuit, its output needs to be amplified to digital levels. In addition, the output must be independent of the input voltage and it has to operate over a wide dynamic range. The two most common approaches to achieve this function is by employing (1) an automatic gain control (AGC) amplifier or using (2) a string of limiting amplifiers (LA) after the TIA.

Controlled by a peak detector, the variable gain amplifier will reduce or enlarge its gain to amplify the TIA signal to a digital level suitable for the CDR. The peak detector time constant needs to be very large in order to catch only the envelope of the amplitude of the input signal. Choosing a time constant too small will result in gain modulation. This kind of amplifier is not very suitable for on-chip integration since it requires off-chip capacitances to achieve the large time constant required.

On the other hand, the non-linearity of a simple differential amplifier can be used to perform the limiting function. By cascading a number of limiting amplifiers, the output voltage will eventually reach digital levels, independently of the input voltage. Limiting

amplifiers are by their nature easy to design, since no automatic-gain control circuit is needed. Also, their power consumption is low, and no external components are needed.

3.2 Type of preamplifier

Preamplifiers in optical communication receiver can be classified into two categories. These categories are not distinct and intermediate designs are possible. The two categories are high input impedance and transimpedance amplifiers.

3.2.1 High-impedance amplifiers

The goal of the high-impedance amplifier is to reduce all sources of noise to an absolute minimum. This is accomplished by reducing the input capacitance. Such amplifiers are normally built using GaAs or JFET transistors. The same techniques used for LNA design are applied here. Since the input impedance is high, the input RC time constant is large as well as the low frequency gain. Therefore, the front-end bandwidth is less than the signal bandwidth. To overcome this limitation, equalization techniques need to be employed. Figure 3.2 shows a typical high impedance bandwidth with the typical transfer functions.

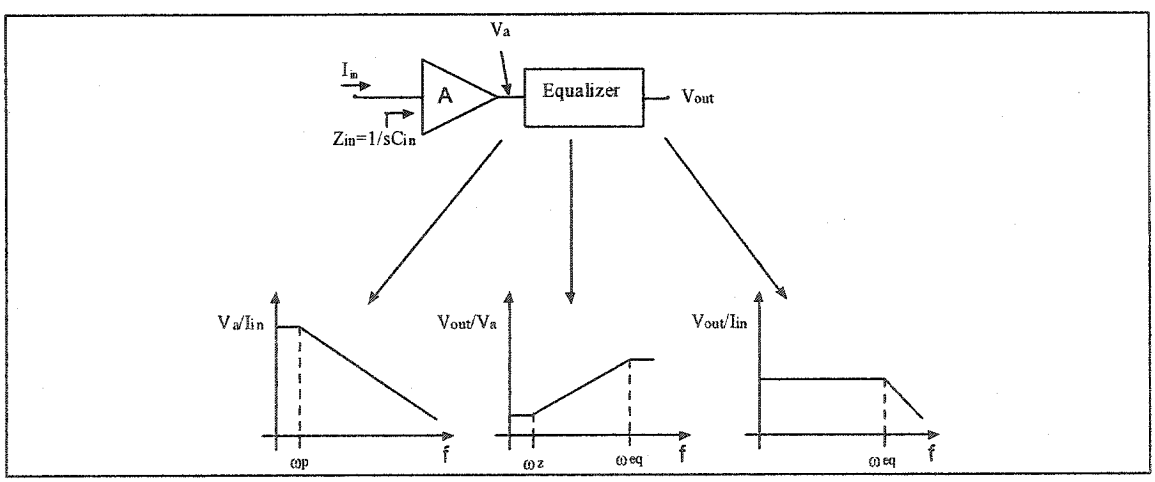


Figure 3.2 High impedance front-end

Equalization is normally done by inserting a zero at the high frequency amplifier pole location. Exact compensation of the pole by the zero is hard to achieve, and external adjustment is generally needed.

3.2.2 Transimpedance amplifiers

Transimpedance amplifiers (TIA's) provide the low noise characteristics of high-impedance amplifiers when used in conjunction with a resistive shunt feedback (current feedback) [Alexander97] [Sedra98]. This shunt feedback has the advantage of reducing the input impedance of the receiver, while stabilizing the gain of the overall amplifier. In addition, low noise and a large dynamic range are possible. The next section describes the mathematical details for this type amplifiers.

3.3 TIA Design

A generic transimpedance amplifier is shown in Figure 3.3.a. It consists of a high-gain, high-impedance amplifier, shunted with a resistive feedback. To properly complete the model, the impedance of the photodiode should be included. Some TIAs use a circuit to compensate for the DC current created by the photodiode. This current is present in systems where the extinction ratio is non-zero (section 2.5). The input impedance of this type of circuit should also be modeled. It yields the model shown in Figure 3.3.b. This section presents the mathematical development of the TIA transfer function.

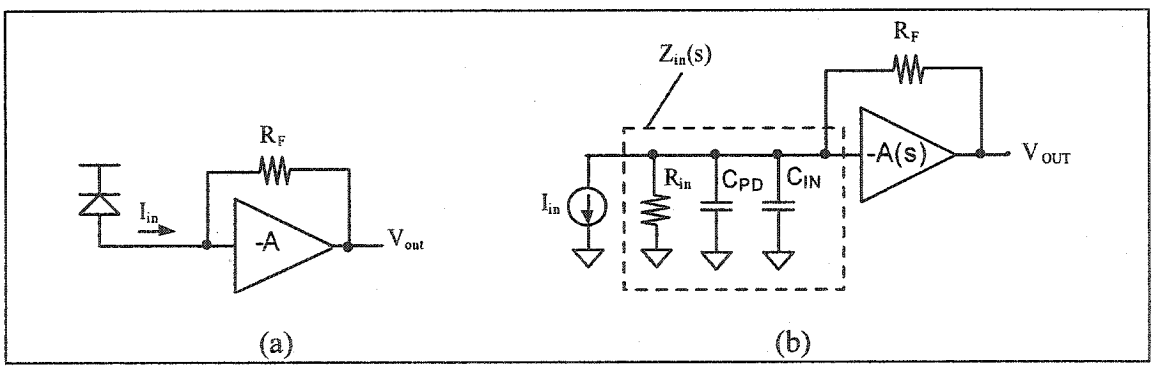


Figure 3.3 Basic transimpedance amplifier structure.

3.3.1 Basic

Using the small signal equivalent model, the transfer function of the loop presented in Figure 3.3.b is given by equation 3.1.

$$Z_t(s) = \frac{v_{out}(s)}{i_{in}(s)} = \frac{-R_f}{1 + \frac{1}{A(s)}\left[1 + \frac{R_f}{Z_{in}(s)}\right]}, \qquad \text{Eq. 3.1}$$

where R_f is the feedback resistor, A(s) is the gain transfer function of the amplifier and $Z_{in}(s)$ is the impedance seen at the input of the amplifier (the combination of the photodiode output resistance, the parasitic capacitance and the TIA input capacitance).

Suppose that the amplifier gain A is constant over the complete range of frequency, the transfer function of 3.1 can be simplified to:

$$Z_t(s) = \frac{-1}{1 + s\frac{R_f C_{in}}{A}}. \qquad \text{Eq. 3.2}$$

One can conclude that the pole created by the capacitance present at the input and the feedback resistor can be moved to higher frequency when the gain A is increased, thus increasing the bandwidth of the overall amplifier.

The bandwidth of the TIA is set by the first pole of its closed-loop transfer function and is given by:

$$BW \approx \frac{A+1}{2\pi R_F C_{in}}, \qquad \text{Eq. 3.3}$$

where A is the small-signal gain, R_F the feedback resistance and C_{in} the equivalent input capacitance of the amplifier. The feedback resistance and amplifier gain A should be selected such that the bandwidth is 0.75 times the bit rate. However, optimization needs to be done to reduce the noise created by the amplifier.

The input impedance of a TIA is given by the following equation:

$$Z_{in}(s) = \frac{R_f}{1+A} \| Z_{in}(s), \qquad \text{Eq. 3.4}$$

If the gain A could be increased to infinity, the input impedance of the amplifier would be 0. In practice, the gain A is limited by the technology used.

An ideal TIA used in high-speed applications would have high transimpedance gain and wide bandwidth. However, it is not possible to maximize both transimpedance gain and bandwidth at the same time. If one increases the gain, the bandwidth will decrease. On the other hand, if the bandwidth needs to be increased, the gain needs to be reduced. The tradeoff can be better explained by the following equation [Sanduleanu01]:

$$R_f * BW = \frac{A}{2\pi C_{in}}. \qquad \text{Eq. 3.5}$$

Because negative feedback is used, the amplifier must be kept stable. The TIA is built around a high-impedance amplifier. This amplifier can be modeled by equation 3.6:

$$A(s) = \frac{A}{1+\frac{s}{\omega_p}}, \qquad \text{Eq. 3.6}$$

where ω_p is the location of the dominant amplifier pole. This is a simplified model of a real amplifier. Real amplifiers have complex transfer functions with many poles located higher in frequency. To model an amplifier as a first order system, the second dominant pole must be around 2 to 3 times higher in frequency than the dominant pole.

From [Sanduleanu01], a limit on the transimpedance gain can be set. The transimpedance gain should be lower than:

$$R_f = \frac{A\omega_p}{C_{in} BW^2}, \qquad \text{Eq. 3.7}$$

in order to keep the amplifier stable. As it will be explained in section 3.4.1, it is better to have a large feedback resistance to reduce the noise transfer characteristic of the TIA. At the same time, the sensitivity of the amplifier is increased.

The next section describes the sources of noise present in a transimpedance amplifier.

3.4 TIA noise analysis

Every electronic circuit is by its nature noisy. The amount of noise tolerable in a system is defined by the application specifications. For radio-frequency communication system, the front-end amplifier has tight noise specifications [Alexander97]. In optical communication system, the front-end circuit is generally a TIA. The amount of noise generated by the TIA will set a lower limit on the sensitivity (smallest detectable signal) achieved by the system. Other stages, such as the limiting amplifier (LA) are also subject to noise degradation. This is most important when the input signal is small and comparable to the noise amplitude.

This section only present a quick overview of the noise sources present in optical front-ends. Excellent presentations/chapters of noise present in electronic systems are available throughout the literature. Of great interest is the chapter of Gray/Meyer [Gray01]. This chapter presents all the basic noise sources contributors in modern IC processes. Also, thorough examples of noise analysis are presented in other interesting books [Razavi98].

To analyse the noise generated by a TIA, it is important to define the main contributors. Such contributors are the channel noise of the transistors and the thermal noise of the feedback resistor. As the transistor size is shrinking (due to advances in process development), the channel noise is becoming more significant. Figure 3.4 illustrated the resistor and MOS transistor noise models used for mathematical analysis.

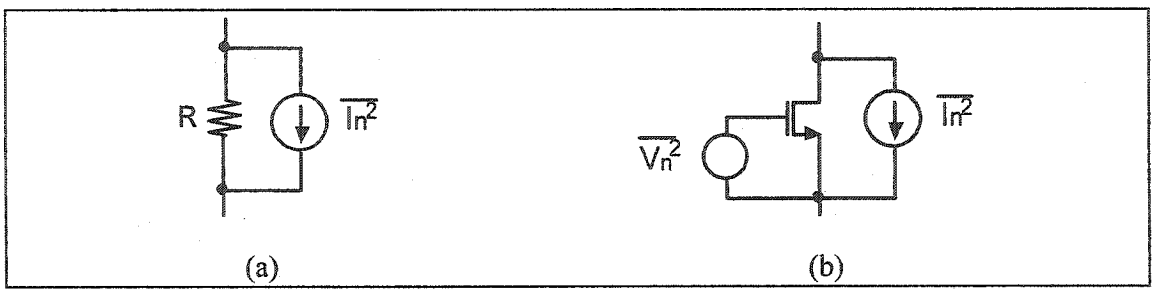


Figure 3.4 Noise representations in a resistor and a MOS device.

Only thermal noise is generated by a resistor. Thermal noise can be considered as a white noise source [Gray01]. It can be represented by a current source in parallel with the actual resistance. The magnitude of the current source is given by:

$$\overline{I_n^2} = 4kT\frac{1}{R}\Delta f \quad \text{Eq. 3.8}$$

where k is the Boltzmann's constant, T the absolute temperation, R the value of resistance and Δf is the bandwidth used for integration. From this equation, a large resistance will produce less current noise. This is important for TIA design, since the input is in the form of current.

MOS transistors have different noise contributors [Razavi98]. The channel of a transistor exhibits thermal noise. It can be defined as:

$$\overline{I_n^2} = \frac{8}{3}kTg_m, \quad \text{Eq. 3.9}$$

where g_m is the transconductance of the device. Flicker noise, also called the 1/f noise, can be an important noise contributor. It is normally represented by a voltage source connected in series with the gate. However, for circuits such as a TIA, where the frequencies of interest are above DC, it can be omitted in the analysis.

3.4.1 Generic TIA noise derivation

Many analysis of the noise present in a TIA have been reported [Ingels99] [Kim01] [Ingels94] [Tanabe98]. This section summarizes these studies. Before presenting the

mathematical derivation, one assumption has to be made: the noise of the amplifier A (Figure 3.3) is mainly due to one transistor, the input transistor. Other noise sources present in the amplifier should be smaller, and therefore neglected.

The equivalent input-referred current noise spectral power density for the TIA presented in Figure 3.2 is given by:

$$I_{n,eq}^2(s) \cong \frac{8}{3}\frac{kT}{g_m R_f^2}\left(1+sR_fC_{in}\right)^2 + 4\frac{kT}{R_f}, \qquad \textbf{Eq. 3.10}$$

where g_m is the transconductance of the input transistor, R_f is the feedback resistor, C_{in} is the equivalent input capacitance, T is the absolute temperature in Kelvin, and k is the Boltzmann's constant. As stated previously, equation 3.10 does not include the flicker noise because it is located below the frequency range of interest. The first term of the equation is the input transistor noise contribution, where as the second term is the feedback resistor's noise contribution.

At medium frequency, the main source of noise is the feedback resistor noise and the MOS channel thermal noise. The spectral noise density starts increasing at higher frequencies because of the zero present in the transfer function. Assuming that the g_mR_f product is greater than 1 (the channel noise being smaller than the feedback resistor noise), equation 3.10 can be simplified to:

$$I_{n,eq}^2(\omega) \cong \frac{8}{3}\frac{kT\omega^2C_{in}^2}{g_m} + 4\frac{kT}{R_f}. \qquad \textbf{Eq. 3.11}$$

The total integrated input-referred current noise power for a TIA is calculated by integrating 3.11. For the complete range of frequencies:

$$I_{n,eq}^2 \cong \int_{\omega=0}^{\omega=\infty}\left(\frac{8}{3}\frac{kT\omega^2C_{in}^2}{g_m} + 4\frac{kT}{R_f}\right)d\omega. \qquad \textbf{Eq. 3.12}$$

The upper integration limit, set to infinity, can be replaced by its noise equivalent bandwidth, as defined in section 2.6.1. The integrated input-referred current noise becomes:

$$I_{n,eq}^2 \approx \frac{1}{9}\frac{kTC_{in}^2\left(\frac{\pi}{2}BW\right)^3}{g_m} + 2\frac{kT\left(\frac{\pi}{2}BW\right)}{R_f}, \qquad \text{Eq. 3.13}$$

where BW is the 3dB bandwidth of the amplifier. This is valid for an amplifier approximated with a first order transfer function.

The equations derived in this section can be applied, with some variants, to almost any transimpedance amplifier with resistive feedback.

3.5 5Gb/s TIA with a resonant network

This section describes the design of a 5Gb/s transimpedance amplifier used as part of an optical front-end.

The preamplifier is based on a common-source configuration, combined with a feedback resistance R_F (Figure 3.5). To increase the bandwidth, transistors in cascode configuration (M_C) are used. Compared to active loads, resistive loading (R_C) only contributes with thermal noise. In addition, active loads are not suited for high-speed circuit since the parasitic source and drain capacitances add an extra load to the output node. Transistors M_B are used for buffering as well as level shifting. The output level is set by either changing the W/L ratio, or by controlling the drain current. To minimize the noise, a constant-k filter is used at the input [Kim01].

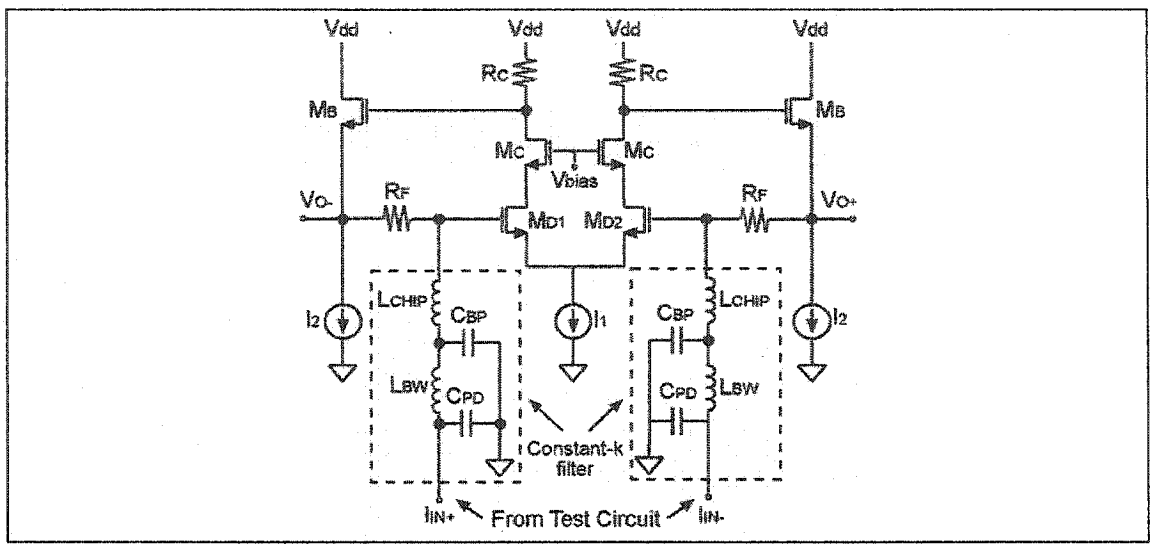


Figure 3.5 Transimpedance amplifier circuit.

The small signal gain of the amplifier A, formed by the differential pair M_{D1}-M_{D2}, cascode transistor M_C, resistive loads R_C and transistor buffers M_B is given by:

$$A_d \approx g_{m1} R_c, \qquad \textbf{Eq. 3.14}$$

where g_{m1} is the transconductance of the input pair transistor. The bandwidth of the TIA is set by the first pole of its closed-loop transfer function and is given by:

$$BW \approx \frac{A+1}{2\pi R_F \left(C_{in} + C_{pd}\right)}, \qquad \textbf{Eq. 3.15}$$

where A is the small-signal gain of the amplifier, R_F is the feedback resistance, C_{in} is the input capacitance of the amplifier, and C_{pd} is the photodiode capacitance, which is the dominant capacitance at the input. By multiplying the bandwidth by 2π, the input-pole location is readily available. The output pole location is defined by the capacitive loading of the next stage and the output resistance of the level shifter:

$$p_{out} \approx \left(R_{out-I2} \| \frac{1}{g_{m-MB}} \right) C_{out}, \qquad \textbf{Eq. 3.16}$$

where $R_{out\text{-}I2}$ is the output resistance of the current source, and $g_{m\text{-}MB}$ is the transconductance of the level-shifting transistor. The output capacitance C_{out} is the parallel combination of the gate capacitance of the next stage, in parallel with the drain capacitance of the current source I_2 and the gate-source capacitance of the level-shifting transistor M_B.

The bandwidth of the TIA can be extended in two ways: one can increase the intrinsic gain of the preamplifier while maintaining stability, or the feedback resistor R_F could be decreased, which comes at the expense of sensitivity and noise.

In a good design, the amplifier's major noise source would be concentrated in a single transistor [Ingels99]. The equivalent input-referred current noise power spectral density is given by:

$$\overline{I_n^2} \approx 2\left[\frac{4kT}{R_f} + \frac{8kT}{3g_m}\left(\frac{1+sR_fC_T}{R_f}\right)^2\right]. \qquad \textbf{Eq. 3.17}$$

This equation was based on the derivation done in the previous section. Note the factor of two in this equation due to the differential nature of the amplifier implemented here.

To minimize noise, R_F could be increased, but this would come at the expense of bandwidth and speed. By applying the classical network theory of Bode, it is possible to increase R_F while compensating for the decrease in bandwidth [Bode45] [Levy64]. This is done by using a constant-k filter. As illustrated in Figure 3.6, the filter can, in theory, double the gain-bandwidth product for an active device. This is the theoretical limit when using an infinite number of LC elements. However, using a single L-C section can provide a 20% increase. The mathematical derivation can be found in [Bode45].

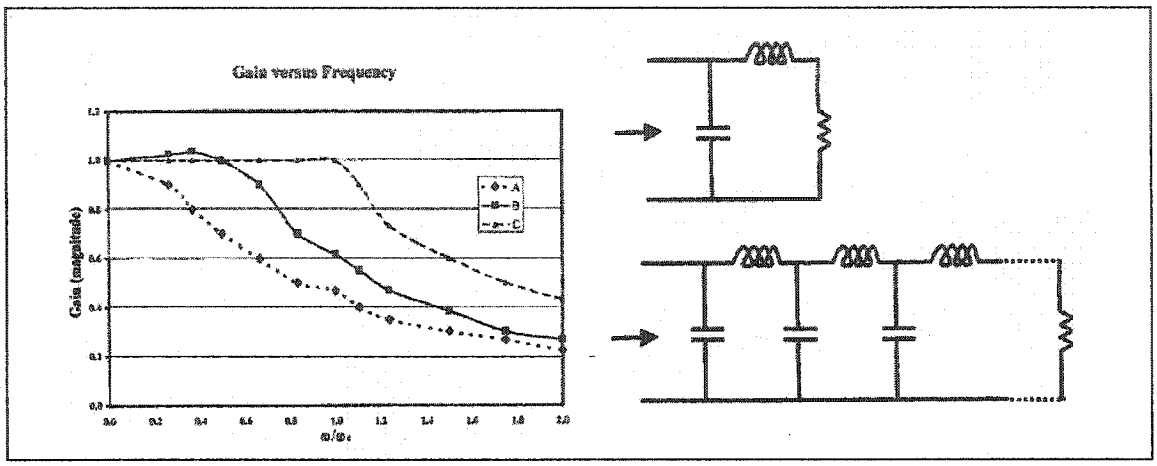


Figure 3.6 Frequency response using a constant-K filter.

For the prototype implementation in this thesis, a two-section filter was used. Half of the filter is built from the parasitic elements of the photodiode. It is composed of the photodiode capacitance C_{PD}, the bondwire inductance L_{BW} present between the photodiode and the TIA. The second section of the filter is composed of the bondpad capacitance C_{BP} and an on-chip inductor L_{CHIP} (Figure 4.13). The bandwidth being increased, the feedback resistor R_F can then be made larger to minimize the noise until restoring the bandwidth to its desired value. For a bandwidth increase of say 30%, the feedback resistor R_f can also be increased by 30%. However, the noise reduction is not linear.

The next section will present the next building block of an optical receiver front-end, the limiting amplifier.

3.6 Limiting amplifier design

Transimpedance amplifiers are normally followed by a post-amplifier, which amplifies the weak voltage signal of the TIA. Different techniques can be used to achieve this function. Limiting amplifiers, which rely on their non-linear transfer functions are well suited for on-chip integration. Typical LAs are built from a string of 3 to 5 amplifier cells [Beaudoin02] [Sackinger00]. Also, LAs have less stringent requirements regarding their noise characteristics.

Limiting amplifiers require having 1) a large gain and 2) a wide bandwidth. Because of the CMOS technology limits, it is better to split the gain over a number of smaller gain/larger bandwidth amplifiers. AC coupling between stages cannot be used for integrated-amplifiers because of the large capacitors required (which obviously depends on the bandwidth required). Therefore, DC coupled amplifiers were used. Since the DC open-loop gain is large, any offset at the input will be multiplied by the string gain and can saturate the output. To overcome this problem, a feedback loop needs to be used to stabilize the DC point at the input.

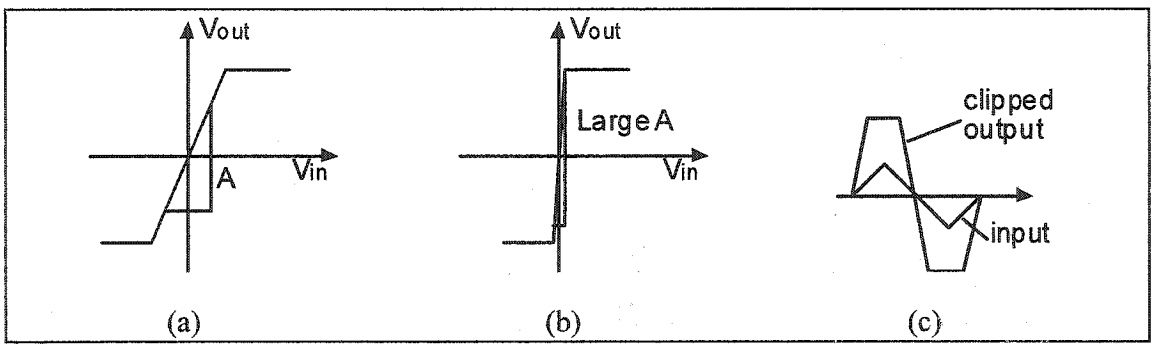


Figure 3.7 Limiting amplifier transfer function.

Cascading many stages of amplifiers with the transfer function defined in Figure 3.7 (a) will lead to a large gain amplifier with the transfer function shown in Figure 3.7 (b). The time response of such an amplifier is depicted by Figure 3.7 (c). An ideal limiting amplifier would have an infinite gain. Thus, any input below or above a certain threshold will be limited to a value set by the limit of the transistor version of the amplifier.

3.6.1 Offsets

One major problem with using open loop limiting amplifiers is the sensitivity to DC offsets. Since the stages are DC coupled, a large gain is present at DC, typically around 30 to 40dB. Any small offset present at the input can make the output saturate. Offsets have many sources; they can be extrinsic to the amplifier or intrinsic.

Extrinsic sources of offset are generally due to the previous stage, in this case, the transimpedance amplifier. When the TIA uses a fully differential structure, one input is

normally AC grounded. This means that no DC current is flowing in this input. However, photodiodes can exhibit large DC currents, because of the non-zero extinction ratio (section 2.5). This creates an inherent offset at the output of the TIA. One way to overcome the TIA offset is to correct the DC current at the input of the amplifier. This can be done in different ways [Razavi00].

Intrinsic offsets are due to the nature of the transistors themselves. Every differential pair of transistors exhibits offsets, when operated under the same biasing conditions (same I_d, same V_{ds}, same V_{gs}, ...). These offsets are due to local variations, i.e. the fluctuation in the length/width of the device, and other process parameters. Matched devices tend to reduce this phenomenon, but two devices can never be exactly the same.

Offsets between two devices (differential pair, current mirror) are normally expressed with respect to the variation in the drain current I_d [Drennan03]:

$$\sigma_{I_d}^2 = \frac{\sigma_\beta^2 + 4\sigma_{V_t}^2}{(V_{gs} - V_t)^2}, \qquad \text{Eq. 3.18}$$

where σ_β is the variation in the gain factor of the transistor, and σ_{V_t} is the variation of the threshold voltage. The threshold voltage variation is inversely proportional to the size of the device. This means that small devices will exhibit large offsets while large devices will have low offsets. Equation 3.18 can be referred back to the input and expressed as a function of the input voltage V_{gs}:

$$\sigma_{V_{gs}} = \frac{\sigma_{I_d}}{g_n}. \qquad \text{Eq. 3.19}$$

Offsets can only be estimates. Actual values used in equation 3.18 are measured from test chips, provided by the manufacturer. As guidelines, a designer should be aware that offsets are present in any circuit where two or more devices need to be matched. Also, the larger the devices are, the smaller the offset will be. Offsets analysis is comparable to

noise analysis in the sense that the offset of the input stage will have the most influence for stages cascaded.

Typically, an amplifier built from a differential pair can have an input-referred offset of few millivolts. Adding the offset due to the TIA, the overall offset present at the input can be around 10mV. The effect of an offset present at the input of a string of amplifiers can be graphically understood by looking at Figure 3.8.

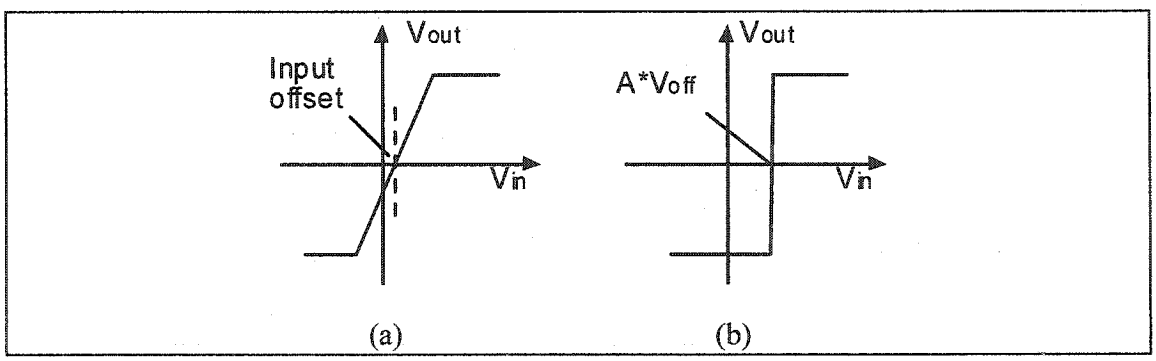


Figure 3.8 Effect of input offset on an amplifier string.

For a single-stage amplifier operated in open-loop, the effect of the offset is negligible as long as the input signal is greater than the offset magnitude. The transfer function curve is not centered at zero anymore (Figure 3.8 (a)). For a large gain amplifier, the offset is also amplified. This can push the amplifier to saturation if the gain is large enough. For example, if an amplifier with a gain of 100V/V has a 5mV offset at the input, the output will already be 0.5V without any input signal. For low voltage technologies where the voltage swing is limited, this means that the amplifier cannot be operated properly.

3.6.2 Offset-induced jitter

By changing the location of the toggling point, offsets will induce jitter. This is a very undesirable effect since modern systems rely on very low BER.

Jitter is created when the duty-cycle of the data stream is changed. Figure 3.9 shows this phenomenon.

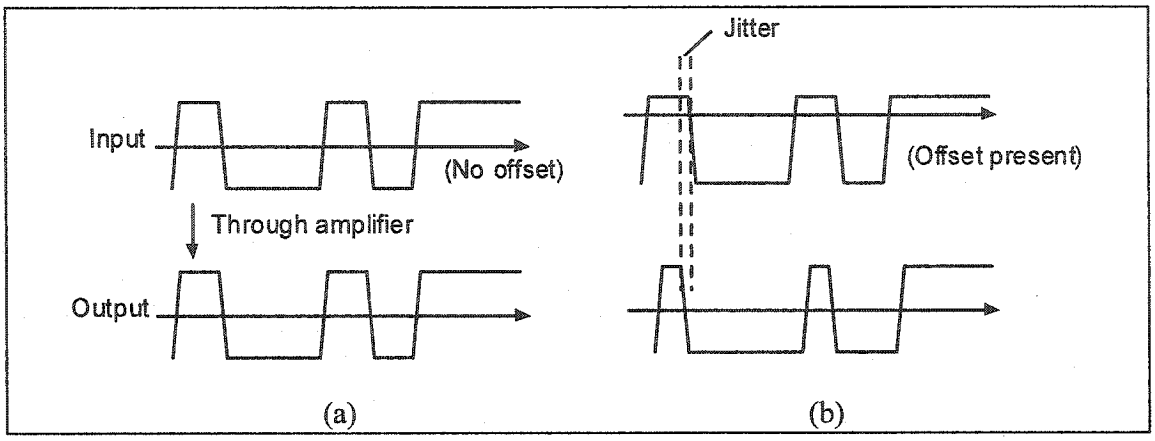


Figure 3.9 Offset-induced jitter.

When no offset is present at the input, the decision point stays centered. Thus the edges of the output data stream are at the same location of the edges of the input data stream. When an offset is present, the decision point is shifted. Because the edges of the signal have a finite slew rate, the variation in voltage can be translated to a variation in time, also know as jitter:

$$\Delta t = \frac{\Delta V}{slew\_rate}. \qquad \text{Eq. 3.20}$$

Bounds on this jitter type of sources are known and can be quantified. They are specified by equation 3.20. Different techniques exist to compensate for the effect of the offsets [Ingels99].

3.6.3 5.0Gb/s Limiting amplifier

Because of the wide-band nature of the application, AC coupling between the amplifier stages cannot be used. As a result, the biasing at the input of one stage is provided by the previous stage. To maintain an optimal biasing of the string, and reduce the effect of offsets present at the input of each sub-amplifier, a first-order low pass feedback loop is used (Figure 3.10). The corner frequency of the passband should be low enough to avoid inter-symbol interference (ISI).

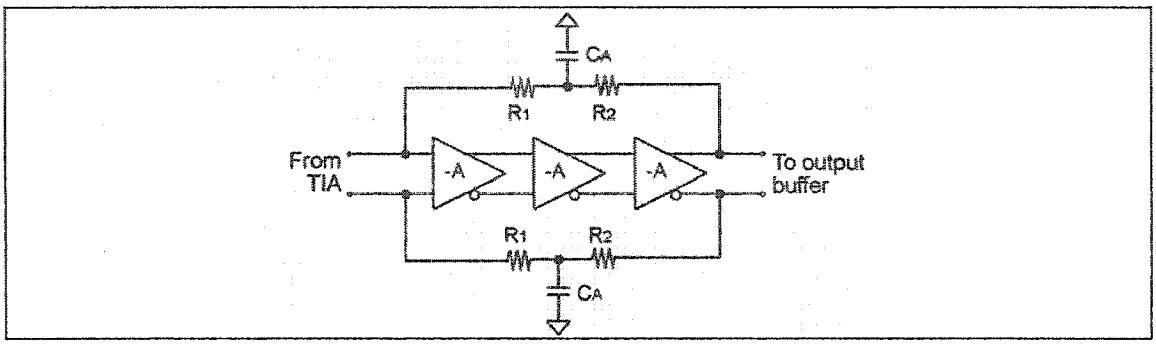


Figure 3.10 Block diagram of the limiting amplifier with passive DC bias control.

Capacitive loading at the input and output of the limiting amplifier is prohibited in order to ensure high-speed operation. For that reason, resistors $R_{1\text{-}2}$ serve two functions: (1) provide a buffering action between the input/output of the amplifier and the capacitor and (2), implement a filter with a low frequency pole at the output. Specifically, the filters are composed of resistors R_2 and capacitors C_A [the source of this circuit]. With a small-signal gain A per stage, the low corner frequency is given by:

$$f_P \approx \frac{A^3}{2\pi R_2 C_A}. \qquad \text{Eq. 3.21}$$

In order to achieve a low cut-off frequency and avoid capacitive loading at the input, the resistors are required to be large. Because their absolute values are not important, the resistors were implemented in polysilicon with an approximate value of 200KΩ. Metal-insulator-metal (MIM) capacitors of 11pF were used. The circuit has a frequency cut-off of only 72KHz, which is low enough to avoid significant ISI.

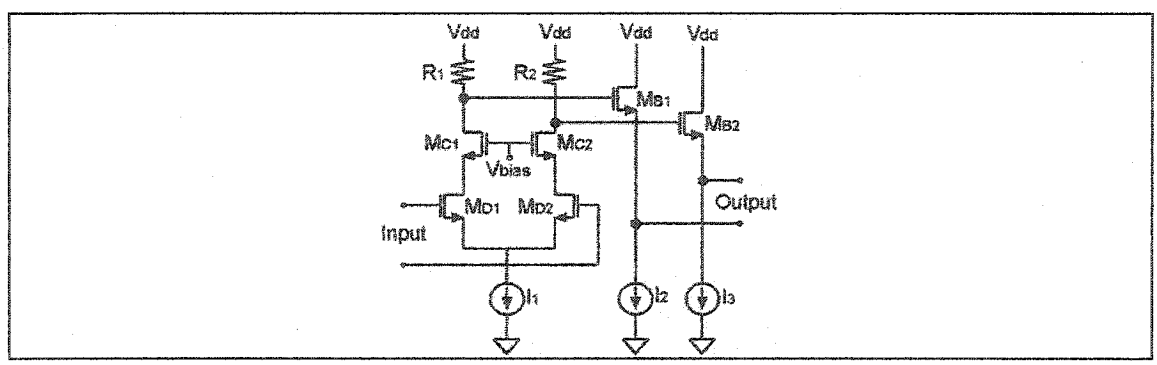


Figure 3.11 Limiting amplifier cell.

The transistor level details of the differential amplifier used are shown in Figure 3.11. It consists of an input differential transistor pair (M_{D1}, M_{D2}) and two transistors in cascode (M_{C1}, M_{C2}) in order to increase the bandwidth. Resistors R_1 and R_2 set the gain, and two source follower transistors (M_{B1}, M_{B2}) are used for buffering as well as for level shifting. No extra bandwidth extension techniques were used (e.g. peaking) since the achieved bandwidth and gain were judged to be sufficient.

The major limitation of this topology is the large area required by the capacitors and resistors. This could be mitigated for the polysilicon resistors by replacing them with the channel resistances of a long transistor.

Another technique that can be used to optimally bias the string of amplifiers is using an active-circuit instead of the simple low-pass filter. By using the Miller effect, the low-frequency pole of the circuit can be pushed down simply by having a gain greater than one. It has the advantages of having smaller passive elements. On the other hand, the design is more complex. Such a circuit has been implemented in papers such as in [Mitran02] [Ingels99].

3.7 Noise analysis (NF) of a limiting amplifier

Noise in limiting amplifiers is an issue that is not much discussed in the current literature. Even if the design specifications for these LA are less stringent than those of the transimpedance amplifier, it is necessary to understand the effect of the noise present in the limiting amplifier chain.

The noise analysis of a string of LA's is somewhat similar to the noise analysis of cascaded radio-frequency circuits. Analyzing the first stages using a small-signal equivalent, the Friis equation can be used. The first stage of the amplifier chain has the most effect on the total noise figure of the amplifier:

$$NF_{tot} = 1 + (NF_1 - 1) + \frac{NF_2 - 1}{A_1} + \ldots + \frac{NF_m - 1}{A_1 \cdots A_{m-1}}. \qquad \text{Eq. 3.22}$$

This is valid only for the first stages since the non-linear limiting action is not present due to the weak amplitude of the signal. Two conditions can be derived from equation 3.22: 1) the first amplifier of the chain should generate the least amount of noise as possible, and 2) the first amplifier of the chain should have a relatively large gain to reduce the effect of the noise of the subsequent amplifiers in the chain.

3.8 Conclusion

This chapter presented the building blocks of an optical front-end receiver. Theory and generic equations guiding the design of a transimpedance amplifier were discussed. The design of a prototype TIA with bandwidth-enhancement techniques was discussed. The second part of the chapter presented a discussion about the issues present when designing a limiting amplifier. Sources of extrinsic and intrinsic offsets were identified. A design example showed how to reduce the offset effects.

The following chapter discusses about the measurement results of a 5.0Gb/s optical front-end fabricated prototype.

3.9 References

[Alexander97] S. B. Alexander, Optical Communication Receiver Design. IEE Telecommunications Series, Volume 37, 1997.

[Sedra98] A. S. Sedra, K. C. Smith Microelectronic Circuit, Fourth Edition. New York: Oxford, 1998.

[Sanduleanu01] M. A. T. Sanduleanu, "A Low Noise, Wide Dynamic Range, Transimpedance Amplifier with Automatic Gain Control for SDH/SONET (STM16/OC48) in a 30GHz fT BiCMOS Process," ESSCIRC Digest of Technical Papers, Villach, Sep. 2001

[Gray01] by P. R. Gray, P. J. Hurst, S. H. Lewis, R. G. Meyer, Analysis and Design of Analog Integrated Circuits, Fourth Edition. New York: Jon Wiley & Sons, 2001.

[Razavi98] B. Razavi, RF Microelectronics, First Edition. Upper Saddle River: Prentice Hall, 1998.

[Ingels99] M. Ingels, and M. Steyaert, "A 1Gb/s 0.7um CMOS Optical Receiver with Full Rail-to-Rail Output Swing," IEEE J. Solid-State Circuits, vol. 34, pp. 971-977, July 1999.

[Kim01] H. Kim, S. Chandrasekhar, C. Burrus, and J. Mauman, "A Si BiCMOS Transimpedance Amplifier for 10-Gb/s SONET Receiver," IEEE J. Solid-State Circuits, vol. 36, pp.769-776, May 2001.

[Ingels94] M. Ingels, G. van der Plash, J. Crols, and M. Steyaert, "A CMOS 18 THz Ohm 240Mb/s Transimpedance Amplifier and 155 Mb/s LED-driver for Low Cost Optical Fibre Links," IEEE J. Solid-State Circuits, vol. 29, pp. 1552-1559, Dec 1994.

[Tanabe98] A. Tanabe, M. Soda, Y. Nakahara, T. Tamura, K. Yoshida, A. Furukawa, "A Single-Chip 2.4-Gb/s CMOS Optical Receiver IC with Low Substrate Cross-Talk Preamplifier," IEEE J. Solid-State Circuits, vol. 33, pp. 2148-2153, Dec. 1998.

[Razavi00] B. Razavi, "A 622Mb/s 4.5pA/rt Hz CMOS Transimpedance Amplifier," ISSCC Digest of Technical Papers, pp. 162-163, Feb. 2000.

[Bode45] H. W. Bode, Network Analysis and Feedback Amplifier Design. New York: Van Nostrand, 1945.

[Levy64] R. Levy, "Explicit Formulas for Chebyshev Impedance-Matching Networks, Filers, and Interstages," Proc. Inst. Elect. Eng., vol. 111, no. 6, pp. 1099-1106, June 1964.

[Beaudoin02] F. Beaudoin, M. N. El-Gamal, "A 5-Gbit/s CMOS optical receiver frontend," The 2002 45th Midwest Symposium on Circuit and Systems, Volume 3, pp. 168-171, August 2002.

[Sackinger00] E. Sackinger, W.C. Fischer, "A 3 GHz, 32 dB CMOS limiting amplifier for SONET OC-48 receivers," Digest of Technical Papers, ISSCC 2000, pp. 158-159, Feb. 2000.

[Drennan03] P. G. Drennan, C. C. McAndrew, , "Understanding MOSFET mismatch for analog design," IEEE Journal of Digest Solid-State Circuits, Volume38, Issue 3, pp. 450-456, Mar. 2003.

[Mitran02] P. Mitran, F. Beaudoin, and M. El-Gamal, "A 2.5 Gbit/s CMOS Optical Receiver Frontend," to appear in Proc. IEEE International Symposium on Circuits and Systems, Phoenix, May 2002.

CHAPTER 4

TIA and LA Simulations and Measurement Results

This chapter presents the simulation and measurement results for the 5Gb/s optical front-end described in the previous chapter. The presentation of a test circuit, used to emulate the photodiode characteristics is discussed. The results of different measurements (bandwidth, eye diagram, noise) are analyzed and compared to their simulation equivalent.

The testing of high-frequency circuits requires careful planning. As early as the specification phase, the designer should know the testing strategy that will be used. A description of the strategies used is described for each test performed.

4.1 Test circuit

In a typical optical receiver system, transimpedance amplifiers are often driven by a photodiode. To facilitate testing in this work, the photodiode is replaced by its electrical equivalent circuit. This has the advantage of allowing the use of standard equipment for testing.

The photodiode electrical model is shown in Figure 4.1. To reflect as much as possible the parasitic effect of an interconnect between a photodiode and the receiver, the

bondwire L_{BW} needs also to be included. The photodiode itself can easily be emulated by an active circuit having approximately the same characteristics. Such circuit was proposed in [Kim01].

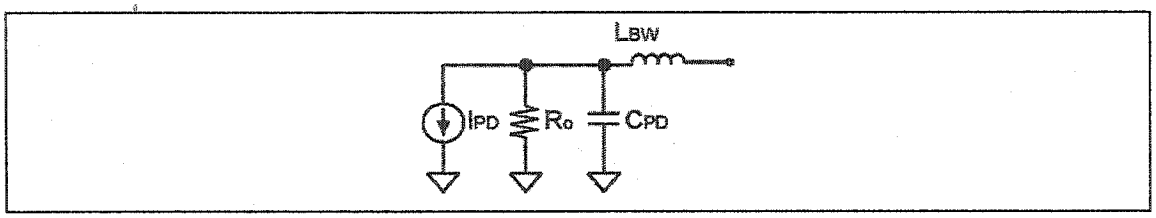


Figure 4.1 Photodiode model.

To apply the test signal to the TIA input, transistor M_9 in Figure 4.2 converts the input voltage to current. A second transistor (M_8) matched to M_9 is used to generate a reference voltage signal for calibration. Current source I_4 sets a limit on the maximum current that can be applied to the TIA. It is important that the output impedance of the cascode transistor M_{10} and of the current source I_4 do not load the TIA. The equivalent photodiode capacitance (0.2pF) is implemented as a pad, also used to connect the *simulated* bonding wire to the TIA. The same circuit with no input was connected to the second terminal of the differential TIA in order to keep the system balanced.

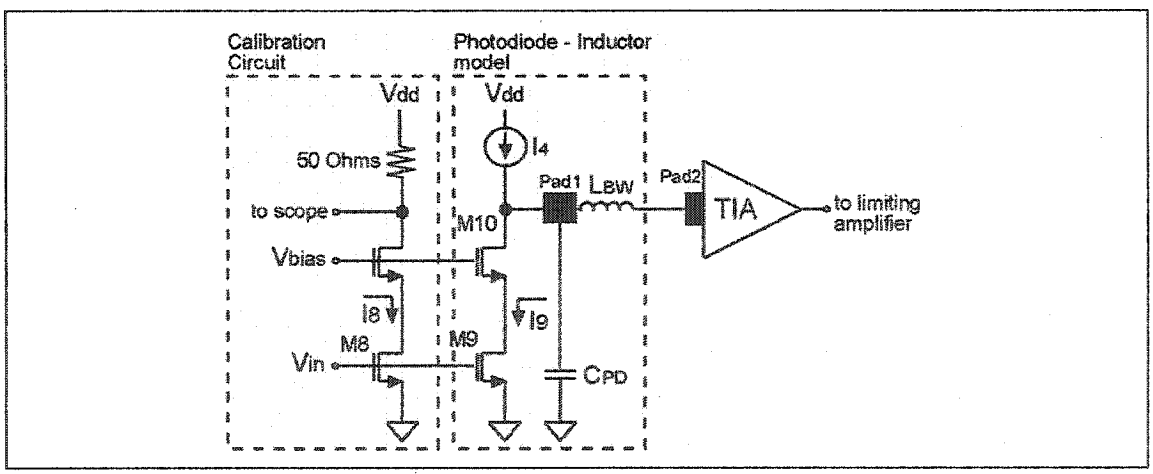


Figure 4.2 The input test circuit.

As described in section 3.5, the photodiode capacitance C_{PD} and the bondwire inductance L_{BW} are used to form half of the constant-k filter, implemented to increase the input

bandwidth of the amplifier. The second section of the filter is formed by the second bondpad (Pad2) and an on-chip inductance. Because of the differential nature of the TIA, the input test circuit has a fully differential input. Also, since the photodiode will source the current into one input of the TIA only, the other input was left AC floating, but correctly DC biased for proper operation.

4.2 Bandwidth

The 3dB bandwidth of the amplifier is important in order to know the maximum bit rate of the amplifier. Figure 4.3 shows the measured 3dB bandwidth of the transimpedance amplifier. The transimpedance gain is 58.7dBΩ with a 3dB bandwidth of 2.6GHz. Using the derivation of section 2.8, the maximum bit rate achieved by the amplifier is 3.5Gb/s. However, if the input power is increased, a bit rate of 5Gb/s can be achieved.

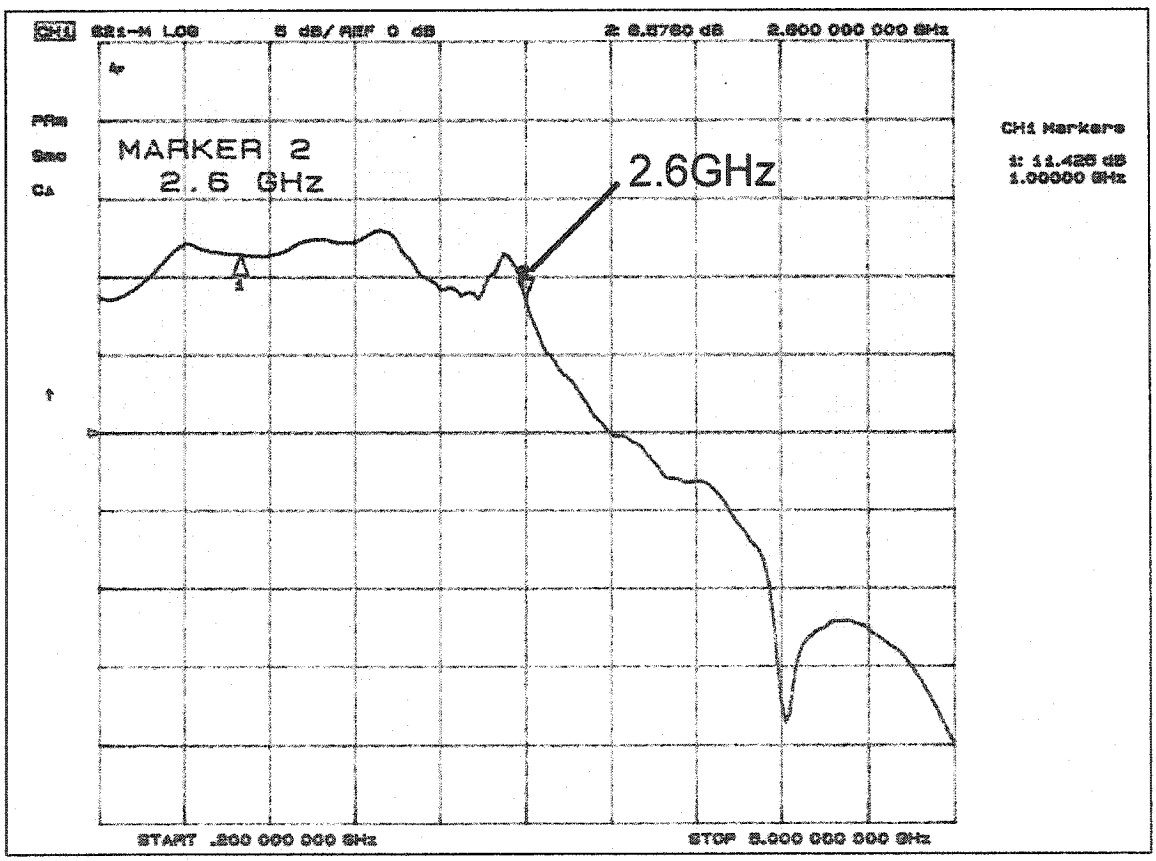


Figure 4.3 Measured TIA transimpedance transfer function.

To measure the bandwidth of the transimpedance amplfier, two consecutive measurements were done with the 8720ES Agilent Vector-Network Analyzer. The first measurement is used as a reference, to remove the frequency dependence of the test circuit. The forward transfer function (from the input to the reference output) of the test circuit is measured. The test function reference measured is removed from the TIA forward transfer function (subtracting the two curves when using the dB scale). This results in the curve observed in Figure 4.3.

To measure the transimpedance gain, conversion between the input power (in dBm), the gain of the test circuit and the TIA output power needs to be done. This is given by:

$$Z_{TIA} = \left(10^{\frac{P_2 - P_1}{20}}\right) C, \qquad \textbf{Eq. 4.1}$$

where P_1 is the power measured at the reference point, P_2 is the measured TIA output power, and C is the internal conversion gain of the test circuit (V/A).

The following figure shows the simulated transimpedance frequency response of the TIA. Its 3dB bandwidth is 3.1GHz with a transimpedance gain of 55.3dBΩ.

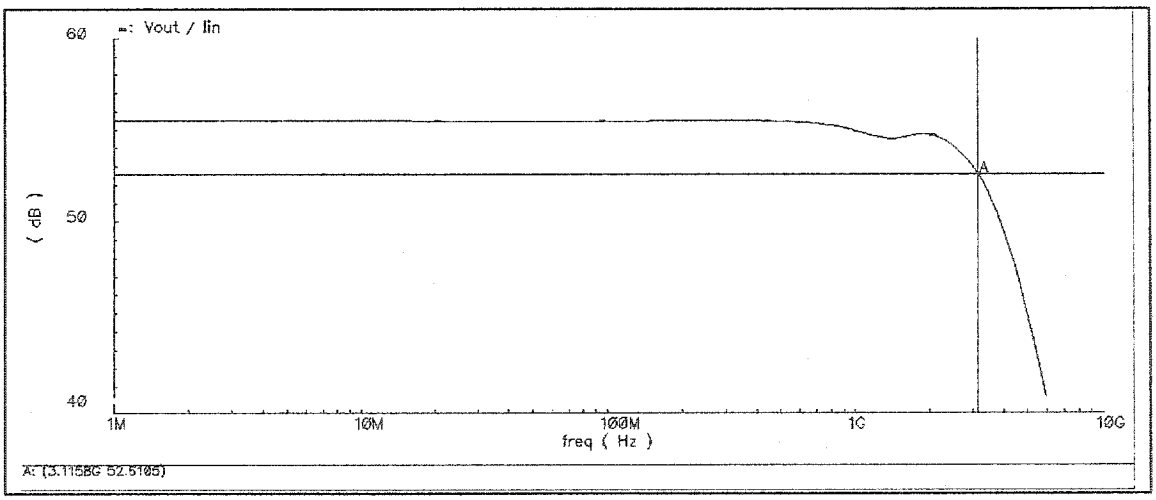


Figure 4.4 Simulated TIA transimpedance prototype transfer function.

4.2.1 Results analysis

Because the gain and bandwidth of the measured and the simulated results differ, it is better to do the comparison using the gain-bandwidth product. The gain-bandwidth of the measured amplifier is 152.6GHz x dBΩ. On the other hand, the simulated amplifier is 170.5 GHz x dBΩ. This corresponds to a 10.5% decrease over the designed value.

Different factors can explain this decrease in gain bandwidth. First, it was found that the output of the test circuit provides a very small amplitude signal, which can easily be corrupted by noise. Such noise comes from electro-magnetic interferences. Therefore, the reference value used in the power calculation can be off by few dBs. Also, the frequency dependence of the test circuit output is non-linear. This is explained partly by the ringing caused by the parasitic inductances of the package and the parasitic capacitances present everywhere in the circuit and on the test PCB.

Secondly, the gain of the amplifier relies mainly on the absolute value of the feedback resistance. Resistors built in CMOS processes can exhibit large deviations compared to their nominal values. Because the transimpedance gain is higher, we can advance that the sheet resistance of the feedback resistor is higher.

A higher sheet resistance can also be used to explain the decrease in bandwidth. The TIA is built around a simple g_mR amplifier. Increasing the resistance will produce higher gain. However, using equation 3.3 from section 3.3.1, the bandwidth should remain the same. Because of the very high frequency of operation, the dominant pole of the g_mR amplifier can affect significantly the behaviour of the circuit. Equation 3.3 holds true for a dominant pole placed at 2 to 3 times the bandwidth. However, the increase in resistance moves the dominant pole to lower frequencies, reducing the effective bandwidth of the amplifier.

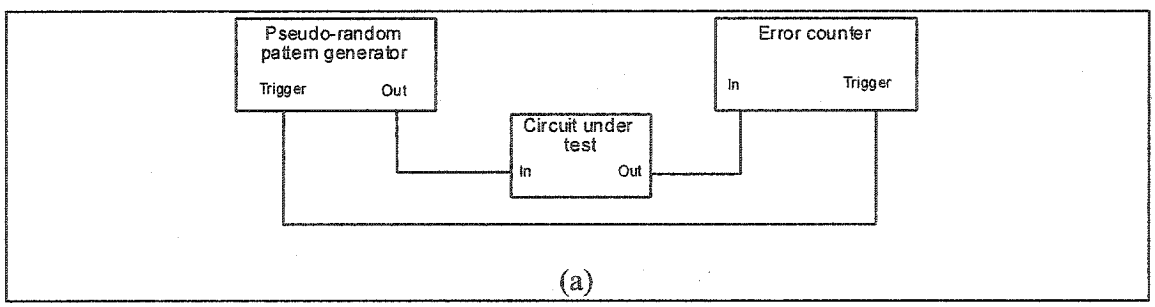
Another factor is the limited frequency response of the package. CFP24 are rated up to 1.5GHz. This can have a significant impact on the return losses at both input and output.

4.3 Eye diagram

Eye diagrams are a powerful and simple method to characterize certain components of a communication data link. This method is used extensively when measuring the characteristics of receivers, transmitters, channels, etc. Eye diagrams provide in an easy and readable way information about the timing and amplitude of a particular signal.

The eye diagram is measured with a bit-error rate tester (BERT). The BERT is composed of a generator, which generate a pseudo-random sequence. This sequence can be of different lengths and is normally built using a linear feedback shift register (LFSR). An LFSR, when correctly initialized, can generate sequences that are closely similar to random sequences. The sequence repeats itself every 2^N-1 of bit. N is the number of registers used. Circuits are normally tested with 2^7-1 to 2^{23}-1 length pseudo-random patterns. Pseudo-random bit sequences of 2^7-1 can correctly estimates the frequency content of 8B10B encoding, which is the mostly used type of encoding.

The second portion of the BERT setup is two folded. To examine the eye diagram, a high-speed sampling oscilloscope is normally used (Figure 4.5b). However, if one is interested by the numbers of errors introduced by the system, an error counter, which counts the discrepancies between the received data and an internal data stream reference, can be used (Figure 4.5a). Counting the numbers of errors is used to determine the bit-error rate (BER).



(a)

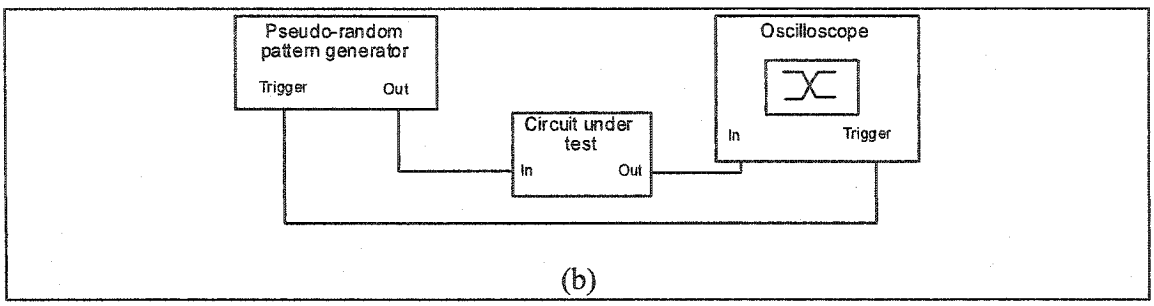


Figure 4.5 Different test setup when using a BERT.

Eye patterns are generated by superposing small section of the pseudo-random sequence, as illustrated in Figure 4.6. The pseudo-random sequence generates mostly all the different combinations of alternating zeros and ones.

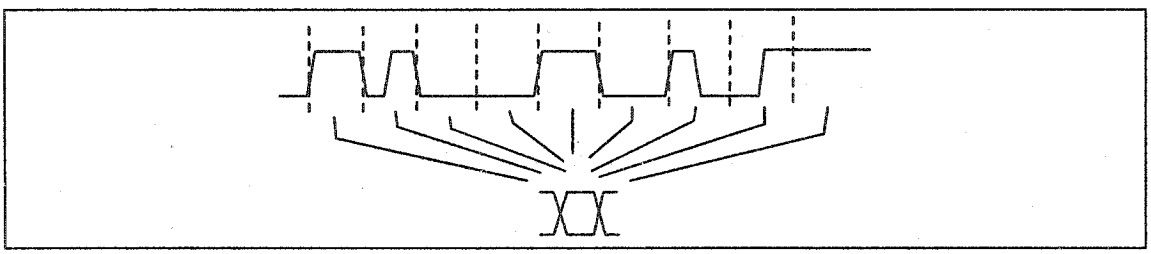


Figure 4.6 Creation of an eye diagram.

The longest length of consecutive bits is defined by the number of shift registers present in the LFSR. A bound can be set when defining the frequencies present in the data stream:

$$\frac{Bit\_rate}{2LFSR_{length}} \leq F_{PRBS} \leq \frac{Bit\_rate}{2}, \qquad \text{Eq. 4.2}$$

where $LFSR_{length}$ corresponds to the number of shift registers.

4.3.1 Simulated eye diagram

Simulation and plotting of eye diagrams can be done in Cadence. The setup is actually simple. A piece-wise linear source reading its input from a file is used. Section 8.2 is the Matlab code used to create a file containing a pseudo-random sequence. Another Matlab code is used to plot the eye diagram, as shown in Figure 4.7. As explained above, the

output stream of data is split in sub-windows, which are all superimposed on the same graph.

Figure 4.7 is the actual eye diagram of the front-end amplifier, taken differentially at the output of the limiting amplifier. The bit rate is 5.0Gb/s. It should be noted that creating an eye diagram with a simulator gives only an estimate of the actual expected eye diagram. Effects such as noise, which contribute for random jitter, is not shown. However, this method provides a good insight of the *deterministic* jitter created by the system.

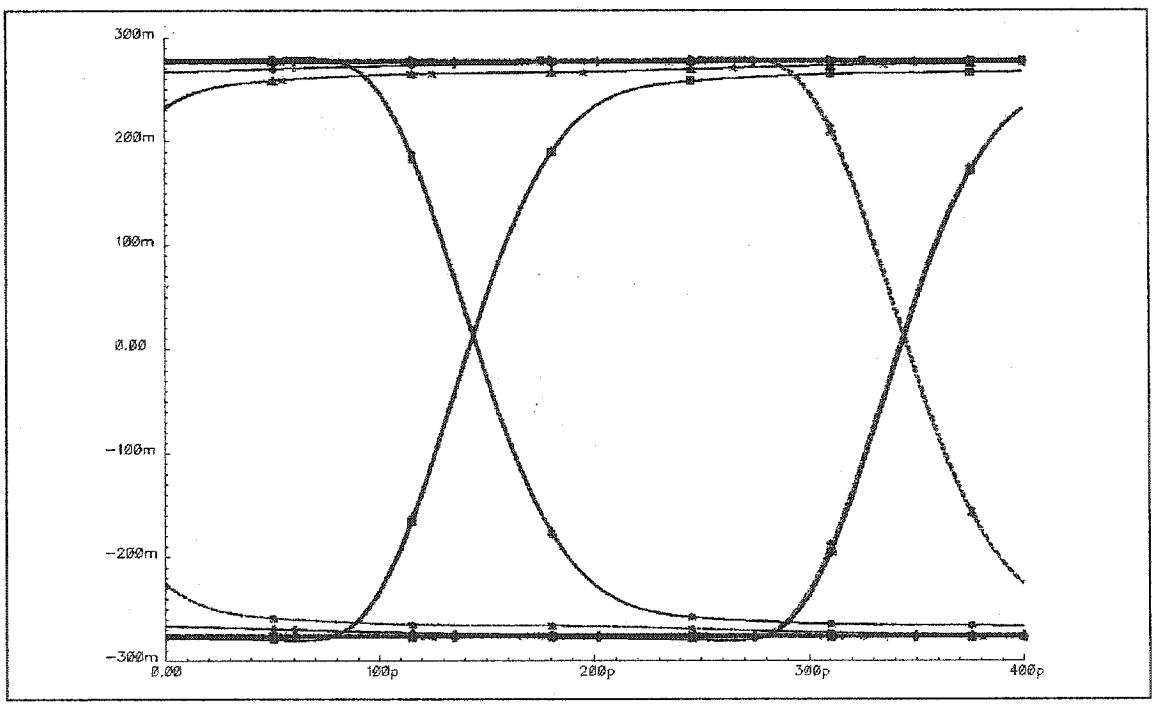


Figure 4.7 Simulated eye diagram at the output.

4.3.2 Measured eye diagram

The following next three figures show the measured eye diagrams, at different point in the system and different bit rates. Figure 4.8 is the eye diagram showing each one of the differential outputs, running at 5Gb/s. One can see that the eye is still open.

Figure 4.9 shows an eye diagram at the output of the transimpedance amplifier, running at 2.5Gb/s. A test buffer was added to probe the signal at the TIA output. One can see that

pattern dependent jitter is created (double crossing at zero). This can be due to the limited bandwidth of the buffer. Because the buffer is attached directly at the output of the TIA (loading its output), the buffer size needs to be relatively small.

Figure 4.10 shows again an eye diagram at the output of the transimpedance amplifier, running this time at 5.0Gb/s. The eye opening is getting smaller because of the high operation speed (bandwidth limited) and the small current injected at the input (circuit noise). The reader can see that the effect of the noise is more present compared to Figure 4.8.

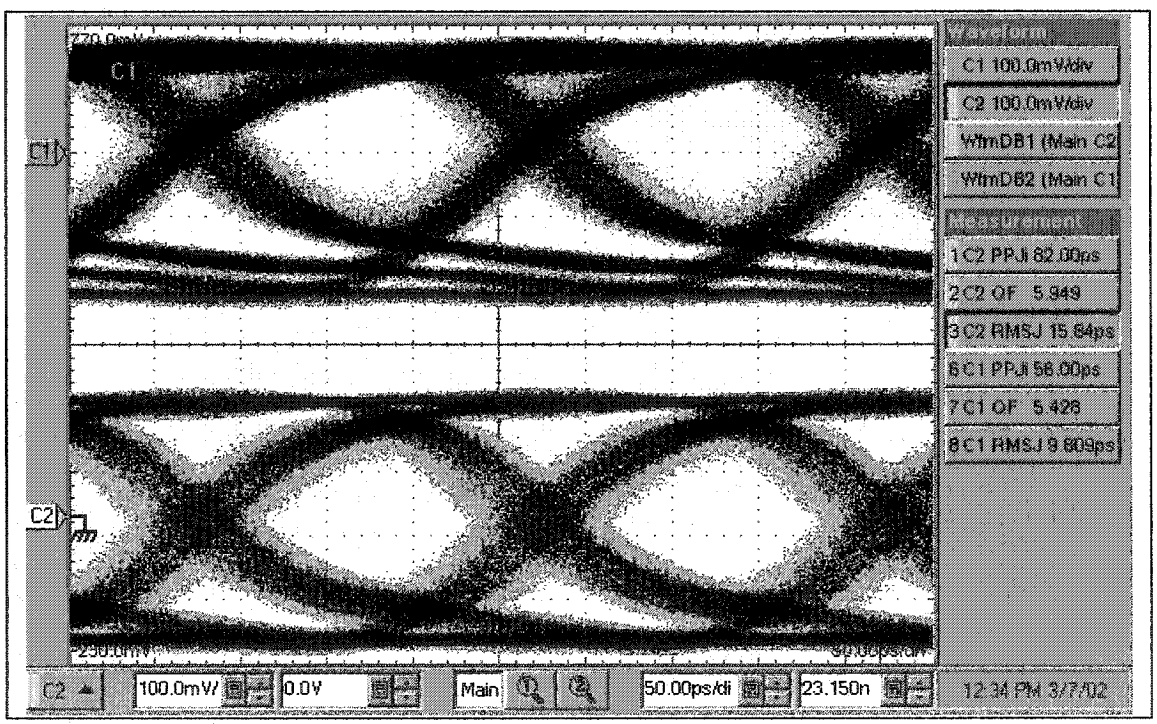


Figure 4.8 Complete receiver eye diagram at 5Gb/s.

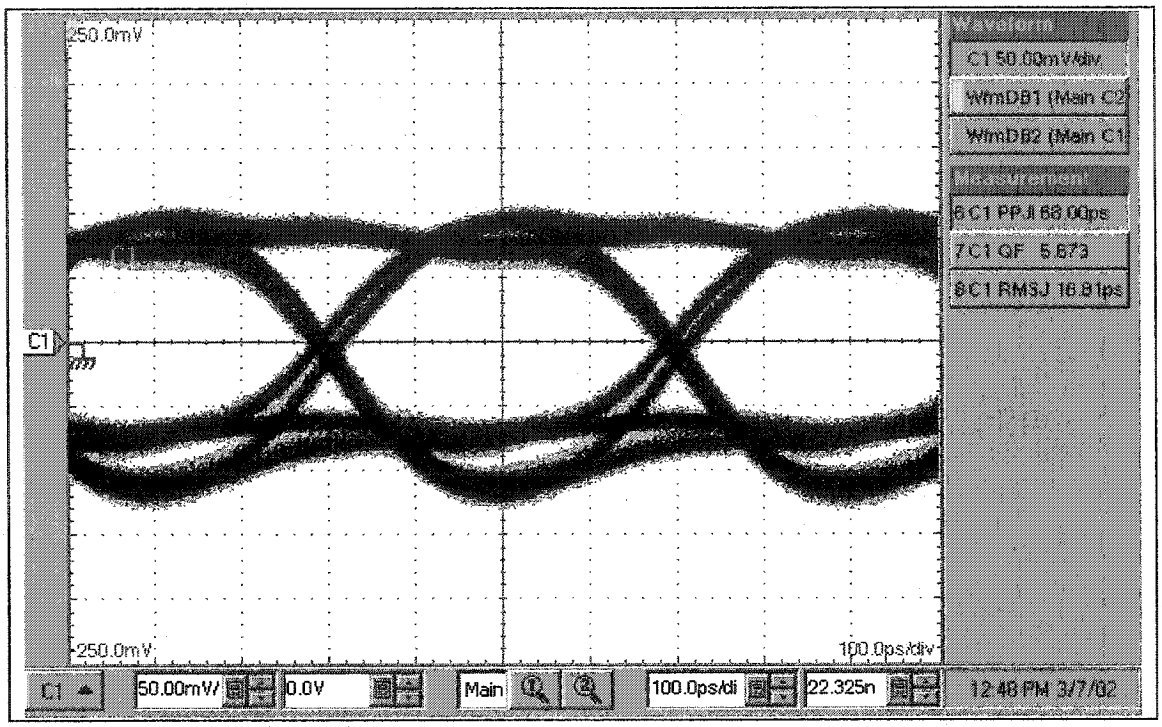


Figure 4.9 2.5G TIA output.

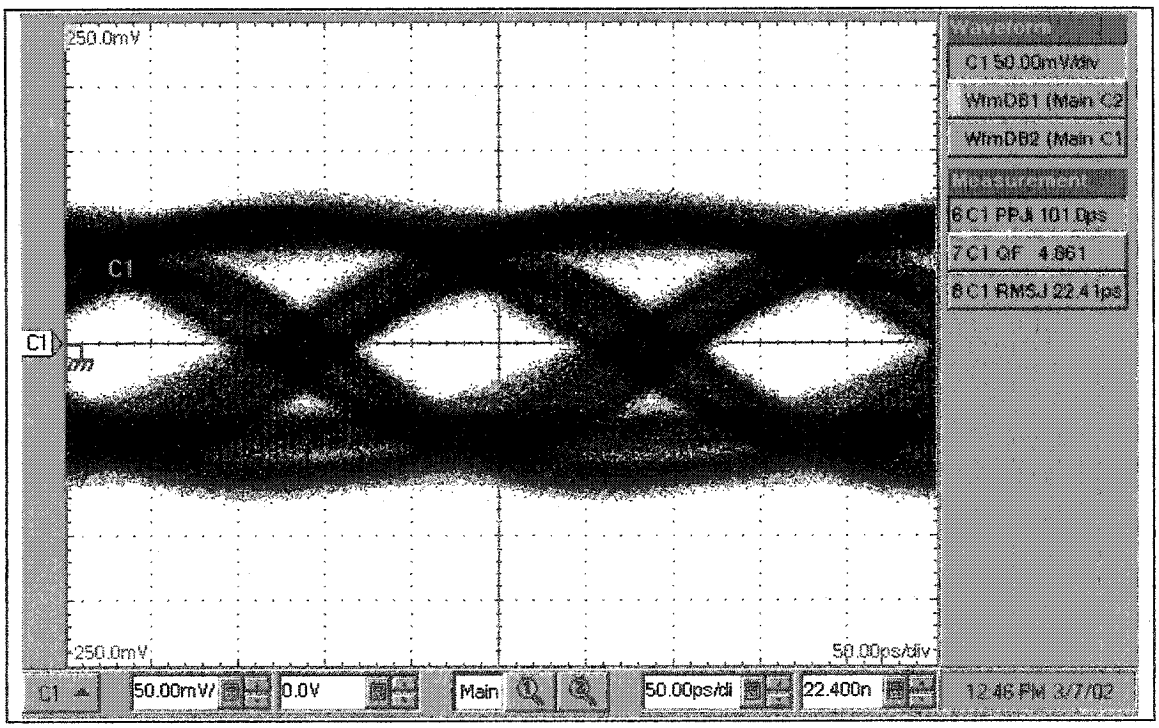


Figure 4.10 5.0Gb/s TIA output.

4.4 Noise

Noise is an important metric when characterizing analog systems. In the case of transimpedance amplifiers, the input referred current noise gives a good indication of the noise produced by the amplifier and its performance in optical systems.

Figure 4.11 shows the simulated input current referred noise versus frequency for the 5.0Gb/s transimpedance amplifier prototype. The lowest amount of current generation is located at 2.5GHz. The average input current referred noise is 13 $pA/\sqrt{Hz}$, when using the noise-equivalent bandwidth (NEB). This value is similar to the input current referred noise observed in bipolar applications [Lu01].

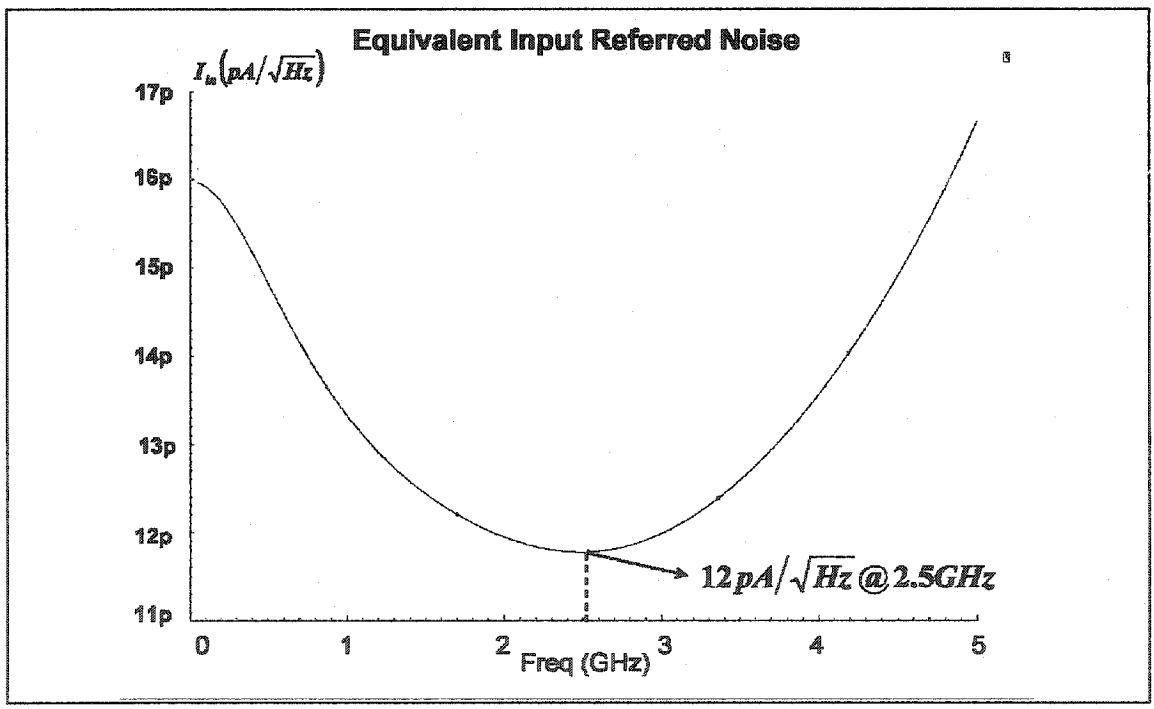


Figure 4.11 Simulated input referred current noise.

Figure 4.12 is the actual measurement of the input current referred noise. The average value of the input-referred noise is greater than 100 $pA/\sqrt{Hz}$. This represents a ten-fold difference between the simulation results and the measured value.

To explain this large discrepancy, it was found that the reference output of the test circuit is somewhat close to instability around 1.35GHz and 2.45GHz. This is clearly shown on Figure 4.12. Such peaking increases the effect of the noise, thus reducing the amplifier input sensitivity. This instability is attributed to ill-modelled parasitic capacitances and the bonding wire inductance.

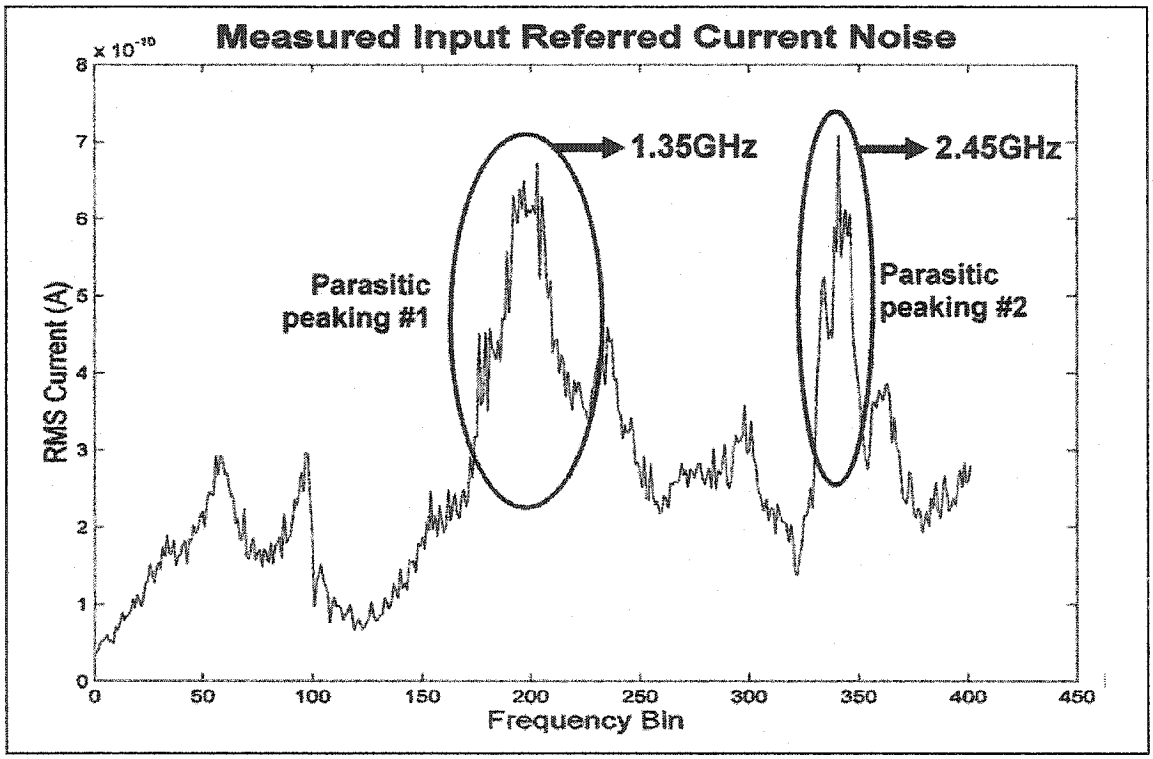


Figure 4.12 Measured input referred current noise.

4.5 *Sensitivity calculation*

The sensitivity of an optical receiver is an important metric in the design a of communication link. The receiver's sensitivity will dictate the maximum length of the link, which is inversely proportional to the optical power present at the photodiode. Without going into the derivation details, the sensitivity of a transimpedance amplifier is given by [Maxim01]:

$$S = 10\log\left(\frac{\overline{I_{in}^2}SNR}{2\rho}\frac{(r_e+1)}{(r_e-1)}1000\right)dBm, \qquad \text{Eq. 4.3}$$

where I_{in} is the average input referred current noise of the TIA, SNR is the signal-to-noise ratio required to achieve a particular bit rate, ρ is the photodiode responsitivity (in A/W), and r_e is the extinction ratio. Using standard values for the photodiode responsitivity and extinction ratio, it was found that the simulated sensitivity of the 5.0Gb/s prototype TIA is -20.4dBm. This is similar to values reported in the literature [Kim01].

4.6 Layout, packaging, and PCB

A microphotograph of the 5Gbit/s receiver is shown in Figure 4.13. The chip was housed in a standard CFP24 package. The complete chip area is 2 X 0.95 mm^2. The TIA and LA consume about half of the space; the other half being taken by the input test circuit and the bandwidth extension constant-k filter circuitry employing two on-chip octagonal inductors (L_{CHIP}).

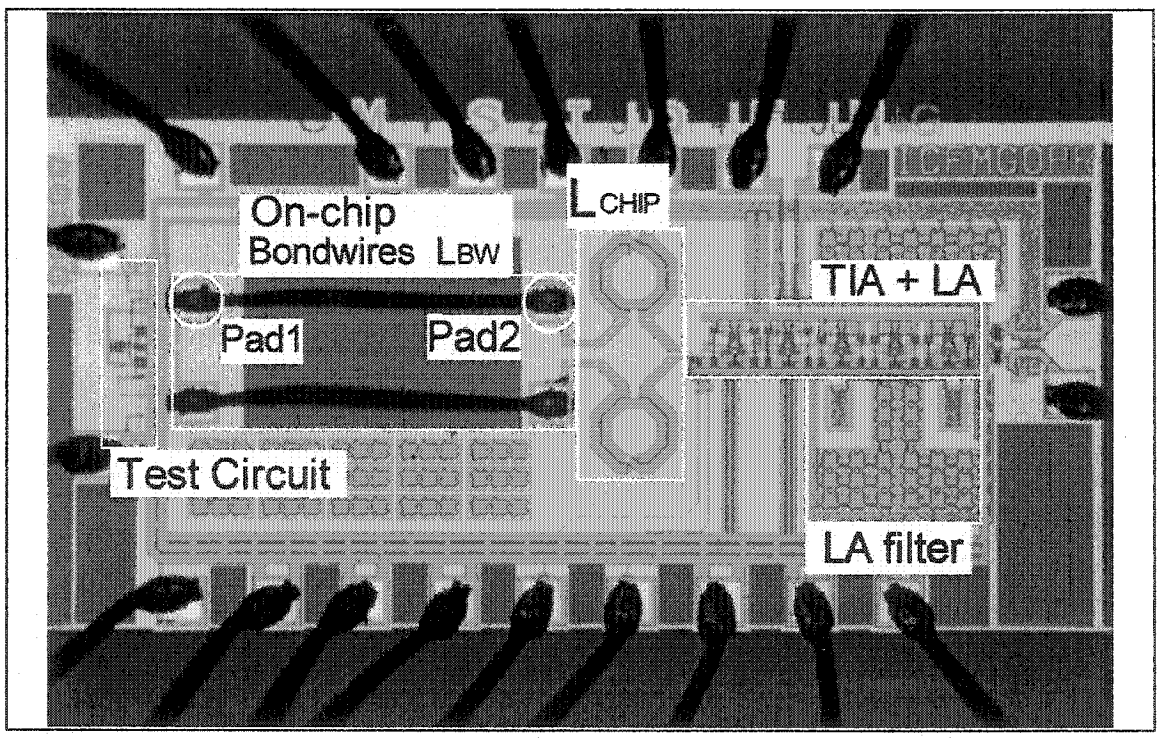


Figure 4.13 The 5Gb/s optical front-end chip microphotograph.

Models of the package and of all bondwires were incorporated early in the design phase. Large MIM capacitors are used for on-chip supply decoupling. These capacitors are of primary importance to ensure a good AC ground at V_{DD}. To help minimize ground bouncing, multiple parallel pins connected to ground were used. The same was done for V_{DD}.

To minimize reflections at the input, a 50Ω polysilicon resistor was connected between the input line and ground. The output buffers were designed to have 50Ω output impedances as well.

The 0.8pH on-chip octagonal inductors (L_{CHIP}) were built using two metal layers in order to decrease their series resistances. To reduce the parasitic capacitances to substrate, all high-frequency pads used the top three metal layers only.

Great importance was accorded to the layout of every amplifier. Because of its differential nature, the circuit was lay out symmetrically. A double guard ring is used around every transistor and circuit, in order to reduce substrate cross-talk. This protects to some extent the sensitive pre-amplifier and input test circuitry.

The packaged die was mounted on a double-sided FR-4 PCB (Figure 4.14). 50Ω transmission lines were kept as short as possible in order to minimize dielectric losses as well as mode dispersion. Decoupling capacitors (1nF and 100nF) used for power supply and biasing voltages were soldered as close as possible to the package in order to reduce the inductive effects of the PCB traces [Johnson93] [Guidon00].

As a future guideline, 50Ω transmission lines could be replaced with coplanar waveguides. This type of transmission line is built using smaller dimension (width) traces. This is more suitable for chips having a tight pitch between each pin. One disadvantage of using such lines is the complexity of the design, which can be alleviated using a field solver.

Nonetheless, it is up to the designer to chose, at the specification phase, all the testing strategies; this includes the placement of the test buffers, the type of off-chip loads (typically 50Ω for high-speed systems), the pin placement (to reduce the cross-talk between adjacent pins), etc. Choosing carefully the testing strategies at the specification phase will enhance the chances of success of a chip, and ease the testing of such a circuit.

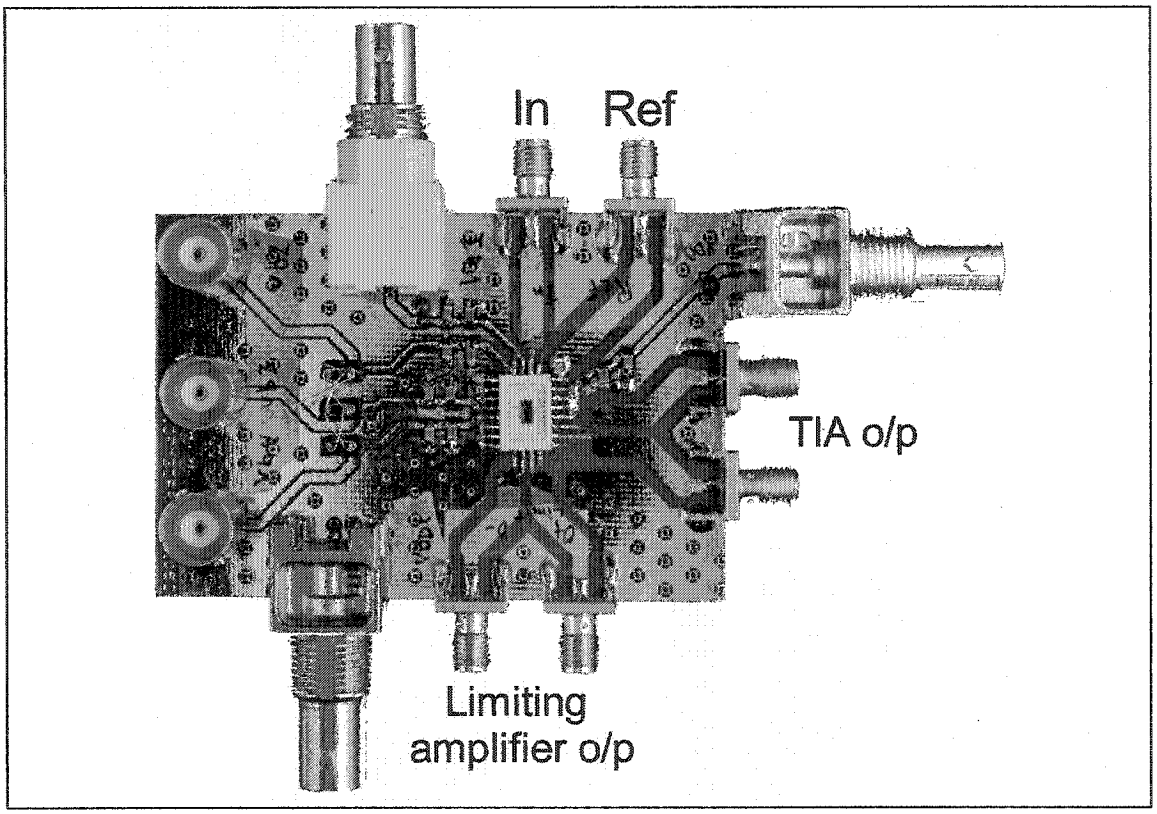


Figure 4.14 The PCB used for testing the chip.

4.7 Conclusion

A comparison between different transimpedance amplifiers is presented in Table 4.1. The first two TIAs are built in bipolar technologies and have bit rates of 10Gb/s. The last two columns are CMOS based amplifiers. This work and the work [Ingels99] have, in addition to the TIA, a post-amplifier.

Table 4.1 TIA performance.

	[Kim01]	[Lu01]	[Ingels99]	This work [Beaudoin02]
Technology	0.25 µm Si BiCMOS	Bipolar (f_t=25GHz)	0.7µm CMOS	0.18µm CMOS
3dB bandwidth (GHz)	9	10 (simulated)	0.5	2.6
Gain (dBΩ)	55	43	60	58.4
Average input noise ($pA/\sqrt{Hz}$)	-	8	7	13 (simulated)
Sensitivity (dBm)	-17	-	-	-20.4 (simulated)
Supply voltage (V)	5	2.3	5	1.8
Power dissipation (mW)	140	9	125	47 (no buffers) 97 (with buffers)

This chapter presented the simulated and measured results of a 5Gb/s front-end analog amplifier. Discrepancies in bandwidth, gain, and average input referred current noise where found, but a valid hypothesis was given in each case to explain the differences.

4.8 References

[MAXIM01] Maxim, HFAN-03.0.0: "Accurately Estimating Optical Receiver Sensitivity," www.maxim-ic.com, Application Note 607, Oct. 2001.

[Kim01] H. Kim, S. Chandrasekhar, C. Burrus, and J. Mauman, "A Si BiCMOS Transimpedance Amplifier for 10-Gb/s SONET Receiver," IEEE J. Solid-State Circuits, vol. 36, pp.769-776, May 2001.

[Guindon00] M. Guindon, "L'environnement industriel," COMALOG inc., Rosemère, 1997.

[Johnson93] H. Johnson, M. Graham, "High-speed digital design, a handbook of black magic," Prentice-Hall, New-York, 1993.

[Lu01] Y. Lu, M. N. El-Gamal, "A 2.3 V low noise, low power, 10 GHz bandwidth Si-bipolar transimpedance preamplifier for optical receiver front-ends," 2001 IEEE International Symposium on Circuits and Systems, Vol. 4, pp.834-837, May 2001.

[Ingels99] M. Ingels, M. Steyaert, "A 1Gb/s 0.7µm CMOS Optical Receiver with Full Rail-to-Rail Output Swing," IEEE J. Solid-State Circuits, vol. 34, pp. 971-977, July 1999.

[Beaudoin02] F. Beaudoin, M. N. El-Gamal, "A 5-Gbit/s CMOS optical receiver frontend," The 2002 45th Midwest Symposium on Circuit and Systems, Volume 3, pp. 168-171, August 2002.

CHAPTER 5

Digital Frequency-Locked Loop

An approach extensively described in the literature for clock-and-data recovery is the use of a specialized PLL [Salama99] [Greshishchev00]. Such circuits have been used with success in optical communication link receivers.

This chapter describes the basic operation of a clock-and-data recovery circuit related to optical receivers. Limitations of modern CDR architectures with respect to the acquisition range and period are presented. A solution to increase the locking range and minimizing the locking time is proposed. Finally, the theory and a circuit implementation are described in more details.

5.1 CDR basics

Clock-and-data recovery (CDR) circuits are members of the phase-locked loop (PLL) circuit family. It is one of the most important blocks in digital communications. The principal role of a CDR is to recover both the clock and raw data of a serial link, while providing the smallest number of data errors. The errors are produced when the CDR is making a wrong decision when recovering a bit.

5.1.1 CDR basic operation

A generic CDR is composed, as depicted in Figure 5.1, of a phase detector, a charge pump, a loop filter, a voltage-controlled oscillator (VCO), and a retimer element which is typically a D flip-flop. The serial data stream, generally coming from a limiting amplifier, is fed to a phase detector. The phase detector compares the phase error between the incoming data and a phase controlled clock generated by a VCO. The goal is to sample each bit of the incoming string of data at the mid-point of a bit. The mid period of a bit has the property of having the greatest distance, vector wise, to the previous and the next bit. Therefore, the probability of error should be at a minimum. Sampling at other moments will significantly increase the number of invalid data (errors).

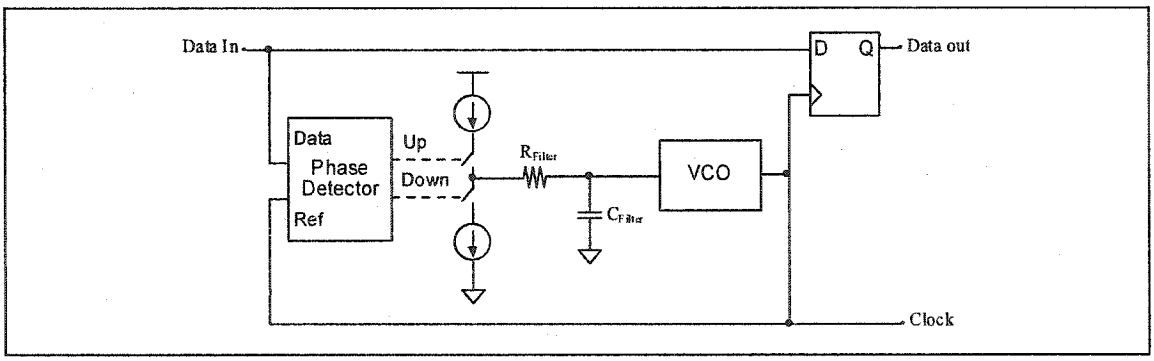


Figure 5.1 Generic CDR structure.

To find this mid point, a phase comparison is performed by a decision circuit. Bringing this back to a CDR circuit, a phase detector will generate an error signal corresponding to the phase difference between the incoming signal and the system clock. For example, if the data leads the clock, the phase detector will generate more DOWN pulses to bring down the VCO frequency. Modifying the VCO frequency will change the phase at the input of the phase detector. On the other hand, if the incoming data lags the clock, the phase detector will generate UP pulses to fasten the VCO. The UP and DOWN signals controls a charge pump, which is a bidirectional switchable current source. The current is then filtered thought R_{filter} and C_{filter} and produce a filtered voltage to control the VCO. Figure 5.2 illustrates the operation of such a loop.

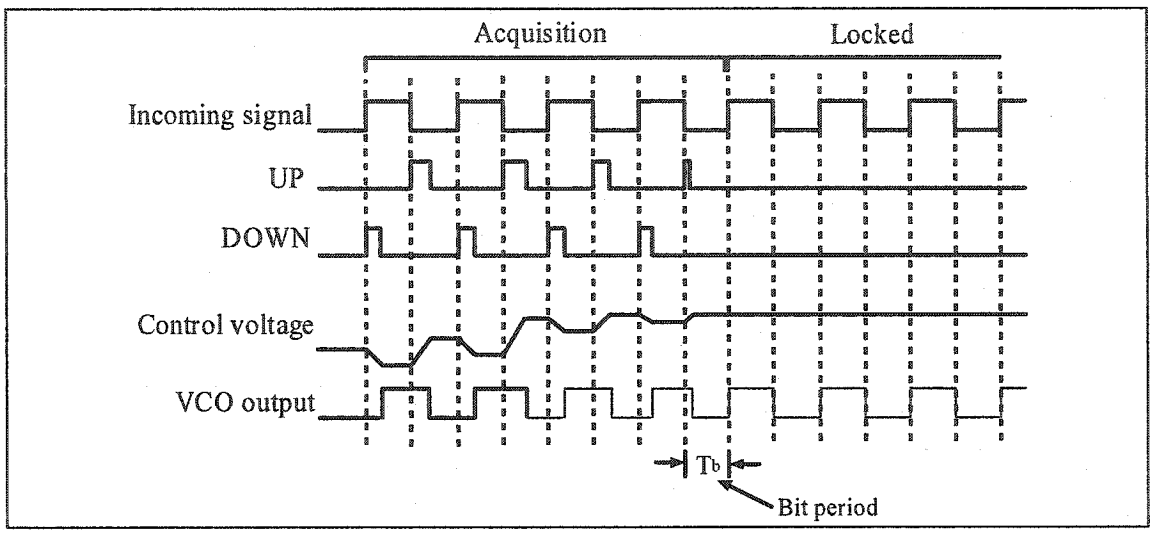


Figure 5.2 Acquisition example.

When the VCO is locked to the incoming signal, a rising edge-triggered D flip-flop samples the signal at the mid period of the incoming bit. This generates a re-centered data, which has a known phase relation with the VCO clock.

When the sampling is done on the data transition, the BER is greatly increased since there is little information about the actual value of the bit. In an ideal system with infinite rise/fall times and no jitter, sampling at the edge of a bit would give excellent BER. However, in a real system, the bit-to-bit transition is done in a finite time (rise/fall times). In addition, jitter, which is a time variation between the bit-to-bit transitions, also increases the BER. The optimum sampling point is right at the middle of the eye, where the eye opening is the largest, as depicted in Figure 5.3.

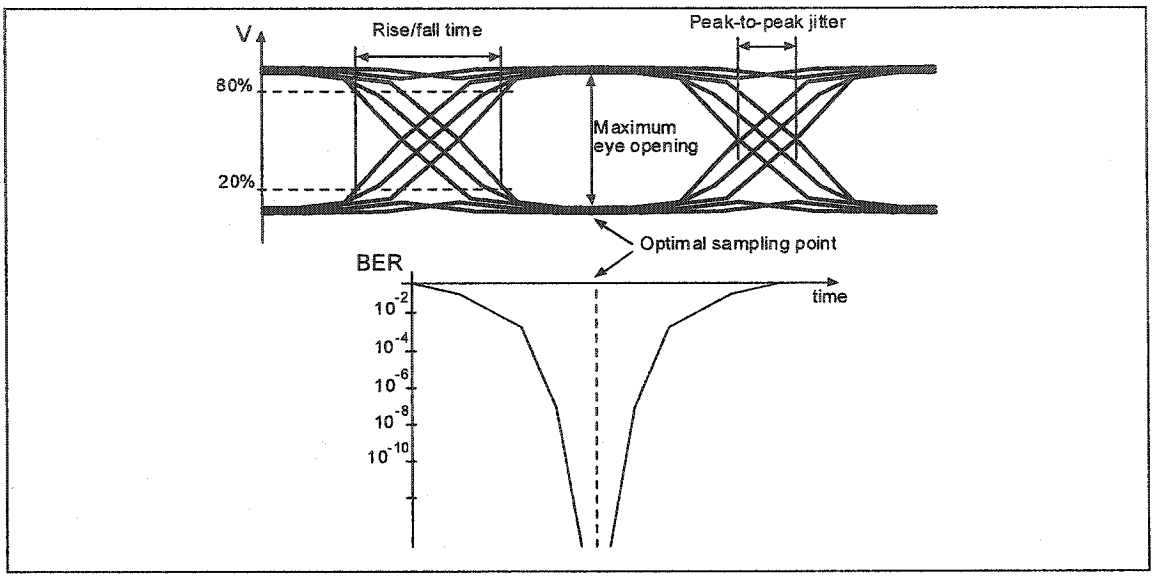


Figure 5.3 Optimal sampling point.

To achieve locking, the designer has to be sure that the CDR VCO will operate at the bit-rate frequency (or half the frequency for half-rate CDR). Failing to do so will let the CDR in continuous acquisition mode, while never being able to lock onto the desired frequency. The VCO center frequency is an important concern because the phase detector is normally limited in its frequency acquisition range.

5.1.2 Analog phase detector limitations

Analog phase detectors are by their nature limited in frequency acquisition. Frequency acquisition is defined as the range for which a PLL can lock to an incoming frequency and keep track of it, as shown in Figure 5.4. If the incoming frequency is too low or too high, the loop will never be able to lock. Typically, CDRs employing a simple phase detector have a locking range of less than 5MHz.

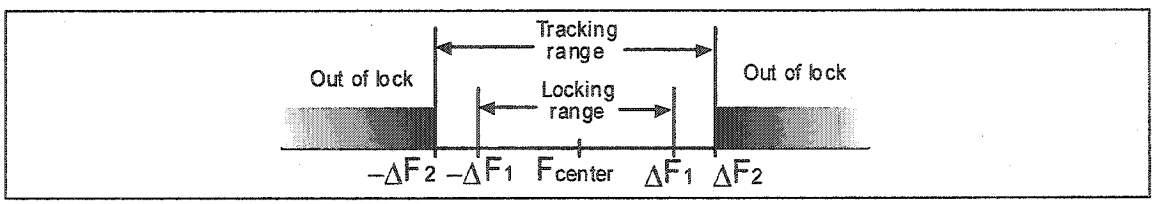


Figure 5.4 Locking range.

Another concern of the phase detector is its tracking range. When the loop is locked, the phase detector needs to be able to track any changes in phase/frequency at the input. If the incoming frequency jumps outside the tracking range, the CDR will be brought out of lock and will restart acquisition. To lock the loop again, the incoming frequency has to meet the condition described above.

5.2 System-level CDR design

This section presents four CDR architectures. Limitations relative to the locking range, locking speed, and jitter generation are described. The last architecture presents a solution to correct the limitation of the 3 previous designs.

5.2.1 Single loop CDR

The simplest type of CDR architectures is the one illustrated in Figure 5.1. It consists of a simple analog phase detector, a charge-pump/loop filter, a VCO, and a retimer element (D flip-flop). The operation of this circuit was described in section 5.1.1. The major limitation of the phase detector is its narrow locking range.

As defined in [Best93], a phase detector, combined with a second-order low-pass filter has a frequency acquisition range of:

$$\Delta\omega_{acq} = \left[\frac{\omega_{LPF}^2}{2}\left(-1+\sqrt{1+\frac{1}{4\xi^2}}\right)\right]^{1/2}, \qquad \text{Eq. 5.1}$$

where ω_{LPF} is the cutoff frequency of the filter and ξ is its damping factor. If the filter is critically damped, i.e. ξ = 0.707, then the acquisition range $\Delta\omega_{acq}$ is approximately $0.46\omega_{LPF}$. It should be noted that this equation underestimates the locking range. For example, to respect the tight jitter specification of the SONET standard, the low-pass filter cutoff frequencies is set to about 2MHz for the OC48 links [ANSI]. Using equation 5.1, this leads to a locking range of approximately 1MHz. If the CDR uses an on-chip VCO, the variation of its centre frequency (nominally at 2.5GHz) can be as large as 500MHz (2.25GHz – 2.75GHz) due to process and temperature variations. Clearly, a

simple phase detector cannot bring the loop nearly close to lock. Even if the CDR can lock to the incoming data stream, it is very unlikely that the CDR will stay locked due to tracking limitation issues.

5.2.2 Dual loop CDR with shared VCO control

One solution to increase the locking range of a CDR is to split the single loop into two loops: one loop to control the center frequency and a second loop to control the phase. This way, both the frequency and phase detection loops can be can be optimized separately. Figure 5.5 shows the block diagram of a dual loop shared VCO control CDR.

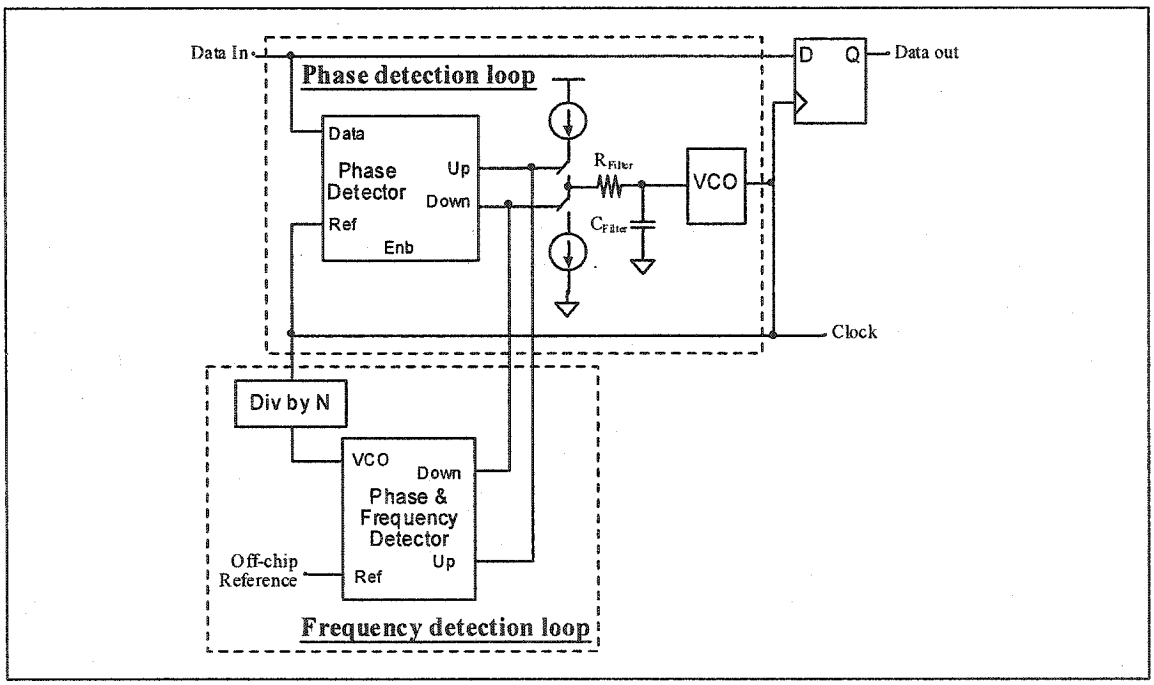


Figure 5.5 Dual loop CDR with shared VCO control.

The major design criteria for a dual loop CDR with shared control is the wide VCO frequency tuning range required (which also requires a large VCO gain K_{vco}). The VCO is very sensitive to noise on its control lines because of the large VCO gain. Because the charge-pump has a finite output impedance, switching noise will be introduced. Noise on the control line is modulated by the VCO, and appears at the output spectrum. From [Razavi98], the noise power is defined as:

$$P_{noise} = \left(\frac{K_{VCO}}{\omega_m}\right)^2 \frac{V_m^2}{4}, \qquad \text{Eq. 5.2}$$

where K_{VCO} is the VCO gain in rad/s/V, ω_m is the frequency of the noise and V_m is the noise in voltage on the control line. The frequency modulation due to the noise creates jitter in the output clock. If the jitter is too high, the decision circuit will incorrectly retime the data.

A better approach to reduce the VCO sensitivity to the charge pump noise is to use two independent control lines, while using the concept of the two loops.

5.2.3 Dual loop CDR with spliced VCO control

Because large-gain VCOs are very sensitive to noise on their control lines, one solution is to split the control lines in two. The first control line is used in the phase detection loop. The VCO gain of this line is small and corrects only for the phase error. The second control line is used for frequency detection. The VCO gain needs to be large to correct the offset in center frequency. Figure 5.6 shows the implementation of such a system.

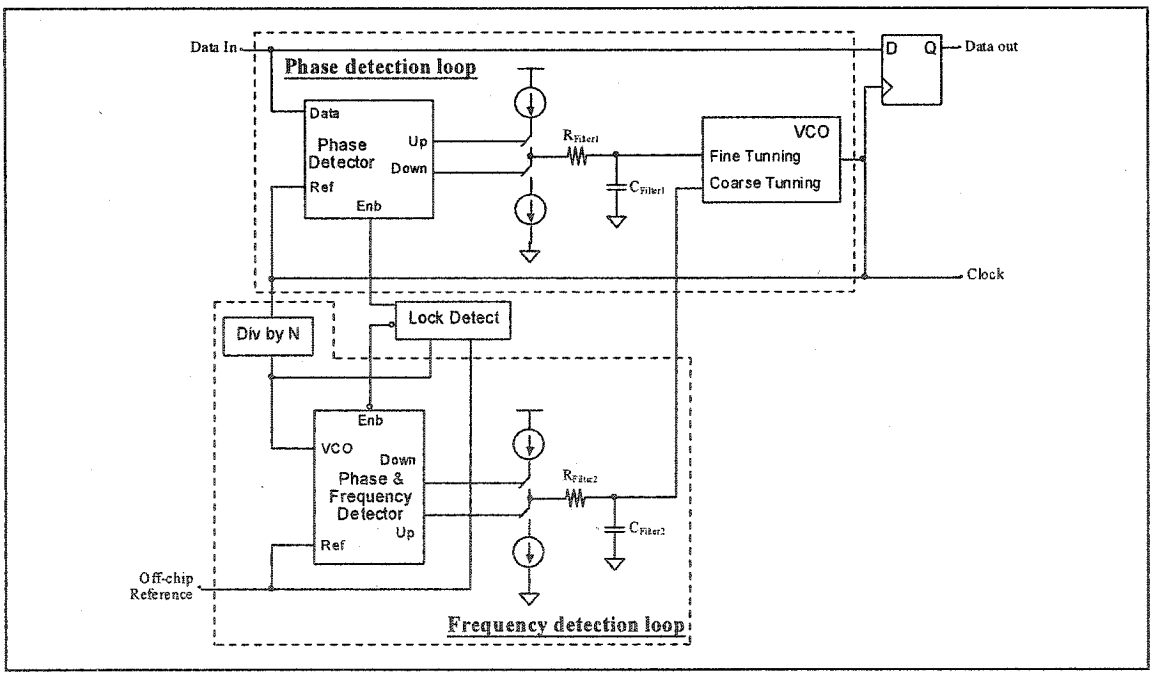


Figure 5.6 Dual loop with split VCO control.

The VCO output frequency can be described by the following equation:

$$\omega_{out} = \omega_{FR} + K_{VCO\_SMALL} \cdot V_{cont1} + K_{VCO\_LARGE} \cdot V_{cont2}, \qquad \text{Eq. 5.3}$$

where ω_{FR} is the free running frequency, K_{VCO_SMALL}, and K_{VCO_LARGE} are the two VCO gains and V_{cont1}, and V_{cont2} are the VCO control voltages. Normally, the frequency detection loop would have a large time constant. This reflects into a low frequency cutoff filter on the voltage control line. Low frequency cutoff filters realized with passive elements are difficult to integrate on-chip because of the large size of the capacitor normally required. The high frequency content of the noise created by the charge pump on the large gain control line gets attenuated, thus reducing the spectral power of the noise at the output of the VCO. Because the frequency detection is done at the reference rate, most of the spectral density of the noise is concentrated at the reference frequency. Even if the noise produced by the charge pump is very small, it appears as an upconverted signal at the output of the VCO ($F_{VCO\_CENTER} \pm F_{ref}$). These are called the reference spurs. The noise power of this system is described by equation 5.4.

$$P_{noise} = \left(\frac{K_{VCO\_LRAGE}}{\omega_{m1}}\right)^2 \frac{V_{m1}^2}{4} + \left(\frac{K_{VCO\_SMALL}}{\omega_{m2}}\right)^2 \frac{V_{m2}^2}{4}, \qquad \text{Eq. 5.4}$$

The solution proposed in this chapter is to completely remove the switching noise on the large gain control line of the VCO. This can be accomplished using a digital-to-analog converter.

5.2.4 Dual loop CDR with digital frequency detector

Because the reference frequency is only a small fraction of the VCO frequency (few 10's of MHz), the complete frequency detection loop can be replaced by a digital equivalent circuit [Hwang01]. The digital circuit can perform the same exact function as the analog detection, with some noticeable advantages. Figure 5.7 shows the block diagram of such an architecture.

The greatest advantage of having a digital loop is the complete elimination of the switching noise on the control lines. The only noise remaining that can affect the VCO is the *1/f* noise produced by the DAC. When the digital frequency loop is locked, the coarse tuning VCO control voltage is fixed to a constant DC value by the DAC. To reduce the noise furthermore, the digital portion of the circuit can be disabled to reduce substrate noise coupling. This can only be done if the assumption that the VCO frequency will not drift over time is valid. In the case where the VCO centre frequency changes or the rate of the input data stream is modified, the digital frequency detection loop can be brought back online to set a new control voltage. Such a digital frequency detection loop can be used not only in CDR applications, but in any PLL application. An implementation example as well as the theory of operation of this architecture is described in section 5.3.

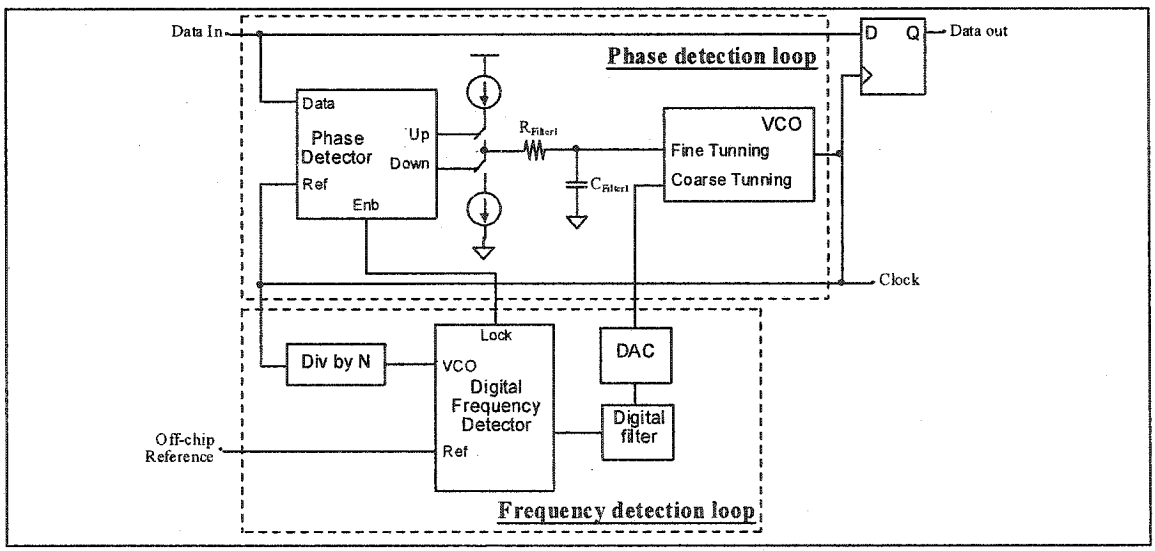


Figure 5.7 Dual loop with digital frequency detection.

Table 5.1 summarizes the characteristics of each CDR architecture described previously. One can see that a CDR using a dual-loop architecture, where the second loop controls the frequency with the help of a digital frequency locked-loop, has advantages over the other architectures. First, the locking range can be large (split VCO controls). When the digital loop is locked, it can disable itself and no reference spurs are present in the output frequency spectrum. In addition, only the DAC noise is present on the control line.

Table 5.1 Summary of CDR architectures properties.

	Single-loop	Dual loop with shared VCO control	Dual loop with split VCO control	Dual loop with digital frequency locked-loop
Locking range	Very small	Large	Large	Large
Locking time	Medium	Slow	Slow	Fast to very fast
Reference spurs	No	Yes	Yes	No
Sensitivity to noise in control path	Low	High	Medium	Low
Number of on-chip filters	1 filter (phase loop)	1 complex filter	2 filters (phase, frequency loop)	1 analog filter (phase) + 1 digital filter (frequency)

5.3 DFLL design

A digital frequency locked-loop (DFLL) can be use to increase the locking range of a CDR without significantly increasing the power consumption and at the cost of no added jitter. Also, the DFLL, by its construction, can quickly bring the CDR loop close to the locking frequency. Before designing a DFLL, it is worth the effort to characterize the loop dynamics in order to understand the tradeoffs associated with this design.

5.3.1 Theory

The DFLL is built around a digital frequency detector. Combined with a first order digital filter and a digital-to-analog converter, the system can control, within some quantization error, the center frequency of a VCO. The digital frequency detector simply counts the difference of VCO cycles with respect to a digitally programmed value within one reference cycle. The reference is provided externally for better precision. The speed at which the counter operates defines the quantization error. The quantization error can be reduced by increasing the speed of operation of the frequency detector counter. Figure 5.8 shows the state flow diagram of the operation of the DFLL.

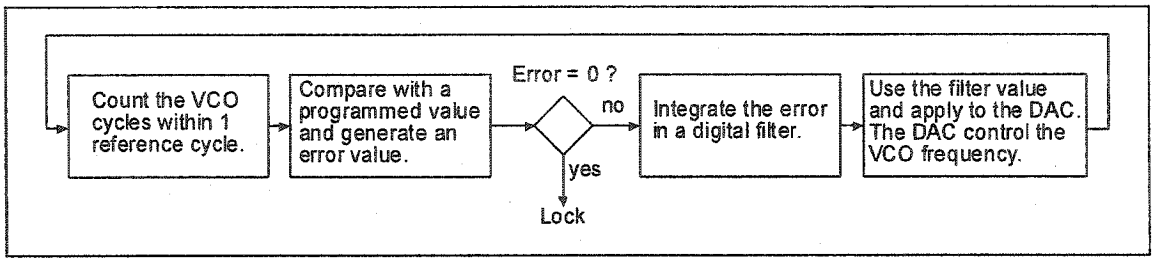


Figure 5.8 State flow operation of the DFLL.

The first step in characterizing the loop is to build a linear model of the system. Figure 5.9 a) illustrates the basic operating principle of the frequency detector. The x-axis represents the time and the y-axis represents the time normalized to T_{vck}, which is the period of one VCO cycle. The target of the loop is to reduce the difference between the reference (T_{ref}) and the VCO signal (T_{vck}). To increase the sensitivity of the phase detector, T_{out} should be much smaller than T_{ref}. It is possible to represent the period T_{ref} and T_{out} as integer multiples of T_{vck}.

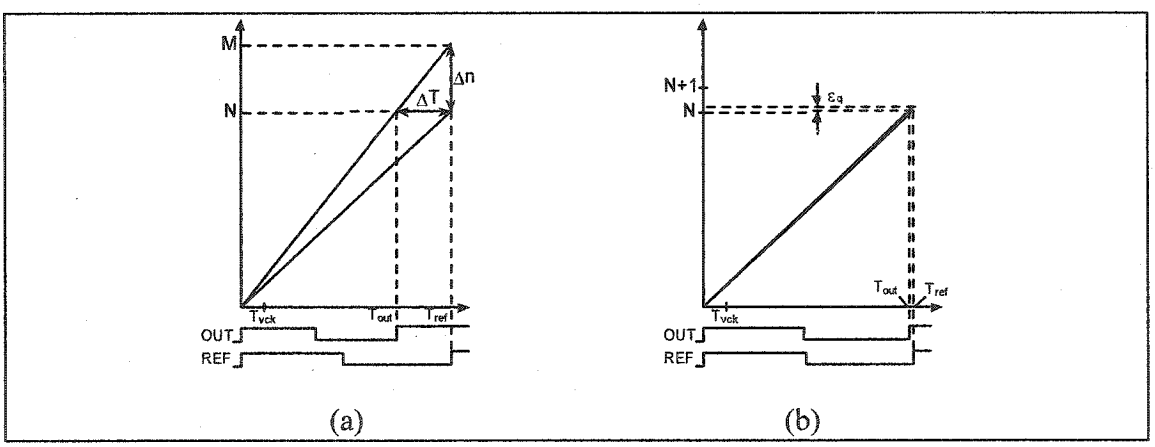


Figure 5.9 (a) Frequency difference measurement. (b) Quantization error.

Two equations can be derived from Figure 5.9:

$$T_{out} = MT_{vck}, \qquad \textbf{Eq. 5.5}$$

$$T_{ref} = NT_{vck}. \qquad \textbf{Eq. 5.6}$$

where M is the actual count of VCO_divided cycles and N is the number of targeted VCO_divided cycles. T_{vck} is the VCO_divided period targeted, and Tout is the actual VCO_divided period. The difference between M and N can be written using equations 5.5 and 5.6:

$$N - M = \frac{T_{ref} - T_{out}}{T_{vck}} + \varepsilon_q \quad , \qquad \textbf{Eq. 5.7}$$

where ε_q is a quantization error due to the digital nature of the system. Since we are interested by the frequency of the signal, equation 5.7 can be rewritten into a frequency equivalent representation:

$$N - M = \frac{N}{\omega_{ref}}\left(\omega_{ref} - \omega_{out}\right) + \varepsilon_q \,, \qquad \textbf{Eq. 5.8}$$

where ω_{ref} and ω_{out} are radian frequencies.

The loop is said to be frequency locked when the difference between N and M is 0. However, there is still a quantization error that cannot be corrected by the loop (Figure 5.9 b) since M and N can only be integer values. The quantization error should be smaller than the locking range of an analog phase detector if it is to be used in a CDR application. The quantization error is bounded between:

$$-1 < T_{vck} < 1 \,. \qquad \textbf{Eq. 5.9}$$

In terms of frequency resolution, it can be translated to:

$$\frac{1}{T_{ref} - T_{vck}} < F_{out} < \frac{1}{T_{ref} + T_{vck}} \,. \qquad \textbf{Eq. 5.10}$$

Normally, because of technology limitations, the reference period would be 16 to 32 times greater than the divided VCO period. In addition, the output of the VCO needs to be

divided because of its high operating frequency. Typically, for a 2.5GHz VCO, division by 8 is required to generate T_{vck} from the VCO output.

This system is very versatile. If the VCO centre frequency is much different (say half) of its targeted value, one can decrease the preset digital value of N for a given reference frequency. Another solution is to keep N constant and change the reference frequency.

A linear model of the loop is shown in Figure 5.10. Note that this model is analyzed in the frequency domain, while PLLs are normally analyzed in the phase domain.

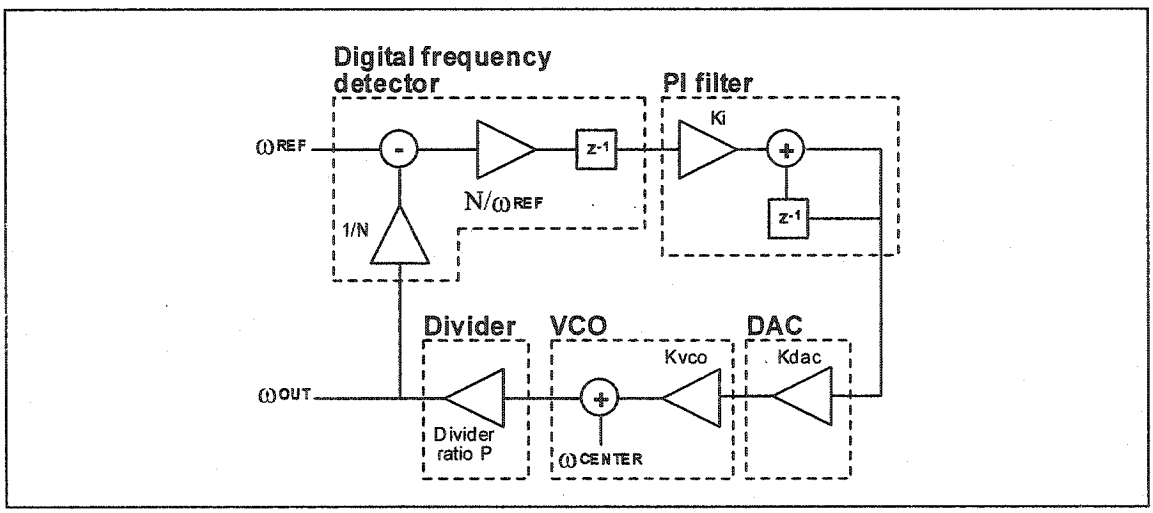


Figure 5.10 DFLL linear model.

Analyzing the loop of Figure 5.10 in the discrete time domain leads to equation 5.11:

$$H(z)=\frac{\omega_{out}}{\omega_{ref}}=\frac{Kz^{-1}}{1-(1-K)z^{-1}}, \qquad \text{Eq. 5.11}$$

and

$$K=\frac{K_i K_{DAC} K_{VCO}}{P\omega_{ref}}, \qquad \text{Eq. 5.12}$$

where K_i is the PI controller gain, K_{DAC} is the DAC conversion gain (digital word/V), K_{VCO} is the VCO gain (MHz/V), P is the divider ratio and ω_{ref} is the reference frequency. The system is stable if K is bounded between:

$$0 < K < 2. \qquad \text{Eq. 5.13}$$

Figure 5.11 shows the time response of the system when K is varied. If K is greater than one, the loop is under-damped. When K is smaller than 1, the loop is over-damped and the loop is critically damped when K equals to 1.

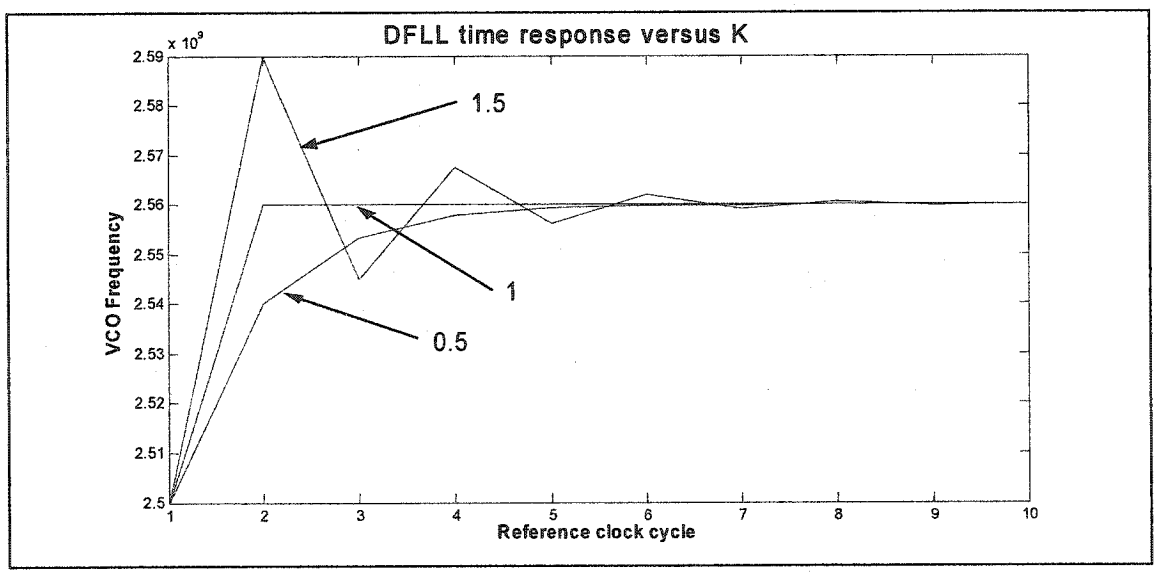


Figure 5.11 DFLL time response vs. K.

An interesting feature of this model appears when K is set to 1. The loop equation then becomes:

$$H(z) = \frac{\omega_{out}}{\omega_{ref}} = z^{-1} \qquad \text{Eq. 5.14}$$

This corresponds to a locking time of 1 clock cycle. This is a wanted feature in CDR applications. Therefore, the loop should be designed such that K equals to 1, in order to achieve the smallest locking time. However, in practice, it is not possible to tightly control every gain factor (e.g. K_{vco}, K_{DAC}). A locking time of 2 to 3 reference cycles is

more probable in real applications. In addition, the use of a multiplier can make the implementation difficult and cumbersome. Gains that are implemented by digital circuitry should be limited to powers of 2 (i.e. 2, 4, 8, 16, ...).

Because of implementation limitations, K_i is set to 1. For a given VCO with a gain of K_{vco} and a fixed divider ratio P, the only variable left for optimization is K_{DAC}. However, it should be note that having a large K_{DAC} can lead to implementation problems; the output range of the DAC might not be compatible with the power supply requirements.

5.3.2 Implementation

Because the design process of a DFLL requires a large amount of time and effort, design reusability is an interesting alternative for cutting down the design cycle time. By using standard logic gates, porting a design from one technology to another can easily be done without investing too much effort and time. Also, the circuit should be customisable to accommodate for various system specifications (in this particular case: VCO frequency, resolution, and locking time).

The circuit presented in this work was implemented in a standard 0.18μm CMOS technology. The digital portion of the loop is built using hand made static CMOS cells, even though standard cells from a library could be used. The high frequency part of the system (e.g. the VCO and high speed dividers) are implemented with current-steering circuits, known as current-mode logic (CML) [Tanabe01]. This choice was made because of the high switching speed attainable by this family of logic gate. The digital-to-analog converter is designed specifically for this circuit. The design of this component is not very critical (e.g. in terms of linearity and speed) because of the feedback loop of the system which minimizes its non-idealities.

All the circuits in this design operate from a power supply of 1.5V. Assuming that the current consumption stays the same, the reduction in the voltage supply from the recommended 1.8V to 1.5V decreases the overall power consumption by 16.6%. This reduction in voltage supply is possible since CML gates do not require large supply

voltage and because only a medium performance is required from the standard digital CMOS gates used. Finally, all of the analog circuitry (DAC) was designed for low voltage operation.

The DFLL implementation is shown in Figure 5.12. It consists of five blocks: the digital-frequency detector (DFD), the loop filter, a DAC, a VCO, and a high-speed frequency divider. It was chosen that the only variable left for optimization was the DAC control gain (K_{DAC}). By carefully choosing its value, it is possible to get the loop lock in one reference clock period.

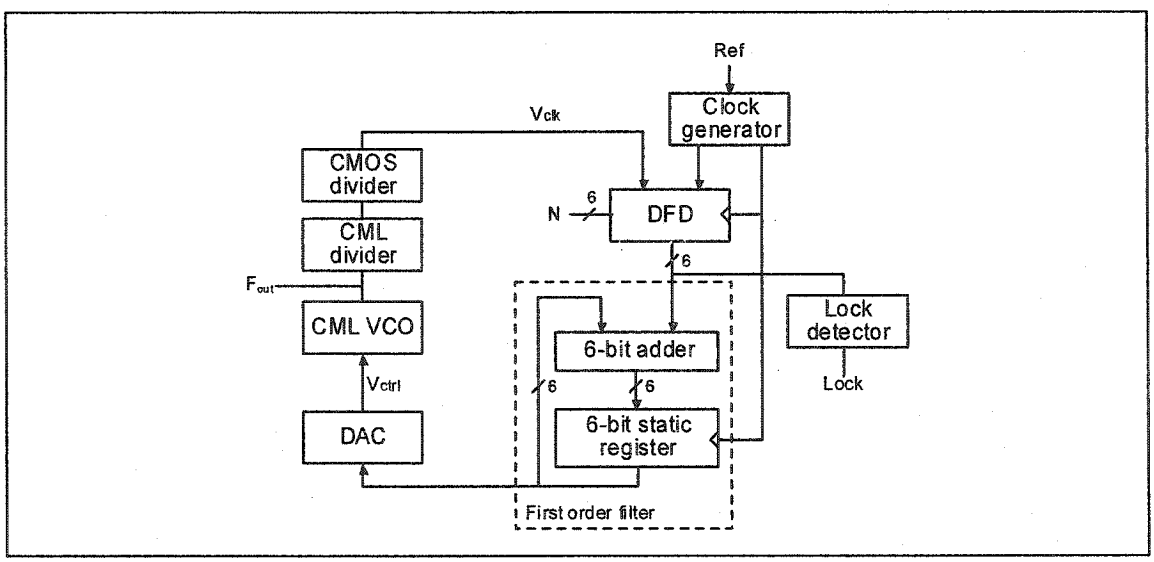


Figure 5.12 Block diagram of the DFLL.

Figure 5.13 illustrates the various signals within the block diagram of Figure 5.12 during the operation of the DFLL. For higher precision, the reference clock is a signal external to the chip (i.e. a crystal oscillator). The clock generator creates two signals; V_{ref} and Reset. V_{ref} serves the purpose of controlling the gating in the digital frequency detector (the counting period) and the refreshing of the filter registers. The Reset signal is used to reload a preset value in the digital frequency detector (which corresponds to M in equation 5.8). During the first half cycle of V_{ref}, a digital counter counts the number of VCO cycles. At the mid period of every V_{ref} period, the DFD outputs the difference between the pulse count and the preset value. This value goes directly to the digital filter

were it gets integrated. At the end of the second half cycle of V_{ref}, the digital counter inside the DFD is reseted. The DAC, VCO and frequency divider are continuous time elements. When the output value of the filter changes, it is immediately reflected by the DAC which changes the VCO frequency. The DAC and VCO have half a V_{ref} period to stabilize before another loop iteration is started. A lock detection circuit will assert a lock signal when the error value at the DFD output is equal to 0.

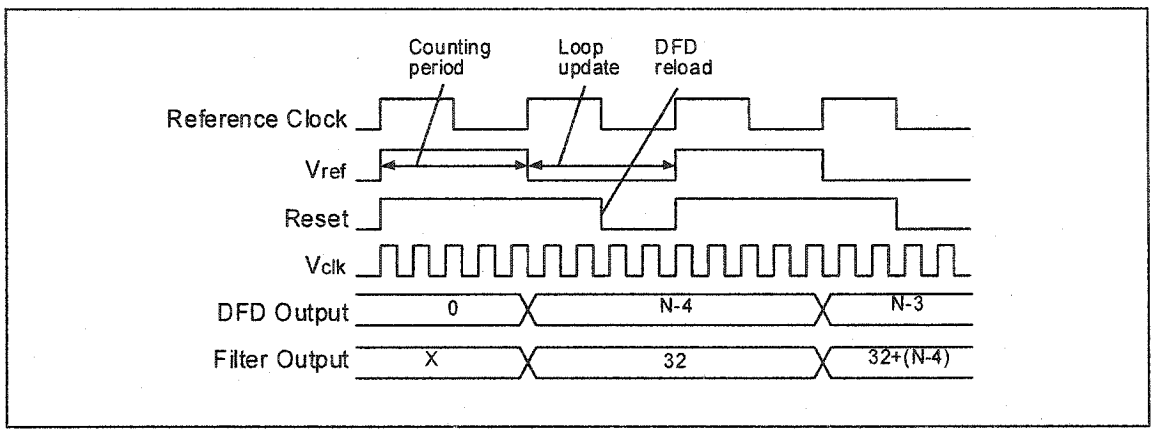


Figure 5.13 Time domain operation of the DFLL.

5.3.3 Digital frequency detector

The digital frequency detector (DFD) is the heart of the frequency-locked loop. Its function is to count the number of divided VCO (V_{clk}) periods within a time frame specified by the reference clock (V_{ref}). During the first half period of V_{ref}, V_{clk} is fed to a down counter through a tri-state buffer, initialized to a preset value N. Referring to Figure 5.14, latch L1 has the property of preventing metastability when Ref goes to low, disabling the buffer. The down-counter is re-initialized to the preset value N at the end of every V_{ref} cycle. The error value is available at the end of the first half V_{ref} period. The output of the down-counter is a signed value representing the difference between V_{clk} and the expected clock count (given by the initial value N). For example, if the VCO is running at a higher frequency, the output of the down counter will be negative, indicating the need to decrease its operating frequency. The difference is sent to the loop filter for integration.

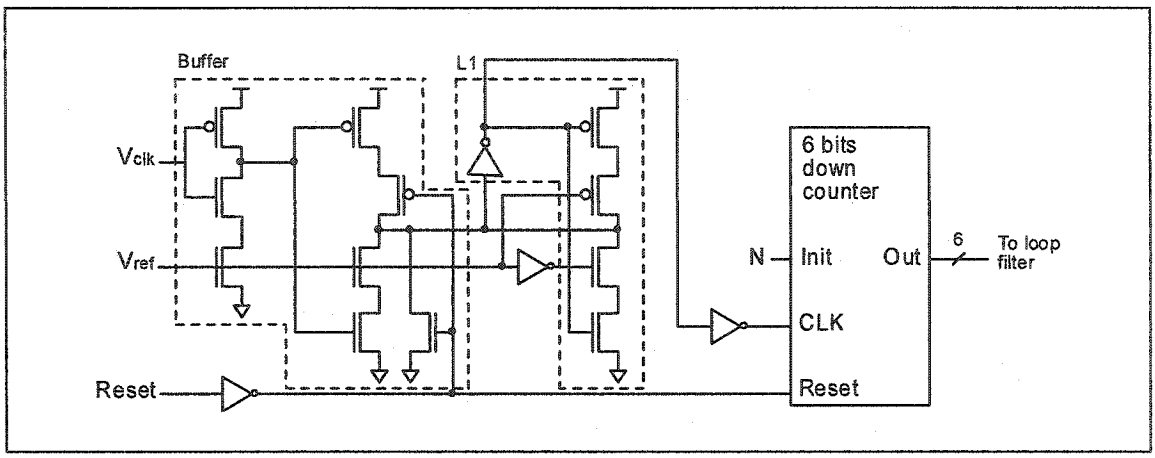


Figure 5.14 Digital frequency detector.

The maximum operating speed of the DFD is limited by the intrinsic speed of the down-counter. To decrease the granularity (i.e. increase the resolution), the down-counter speed should be increased (reducing the divider ratio P). Different approaches are available to increase the speed of operation. For example, the counter can be implemented using a carry-bypass or a carry-select method.

5.3.4 Loop filter

The loop filter is a simple first-order proportional-integrate (PI) filter. It is built using full-adders and static registers. Dynamic latches cannot be used in this application because of the leakage present at the input. When the loop is in locking condition, the register have to hold its value for an undetermined time (until the loop goes out of lock for any reason). If the register value is inadvertently changed while holding a value, the output frequency will change and it might bring the loop out of lock. Therefore, a static D flip-flop was design based on the 74HC74 D flip-flop. Because this architecture of flip-flops is not sensitive to leakage, it can hold a value as long as it is powered up.

Figure 5.15 shows a block diagram of the filter. To simplify the design, the gain of the filter was chosen to be one ($K_i = 1$). Because the DAC accepts unsigned numbers (i.e. positive numbers only), conversion from signed to unsigned numbers is needed. The unsigned value is created by adding an offset at the output of the filter. Adder2 in Figure

5.15 does that function. The offset, which is represented by the center value of the range of the DAC ($2^N/2$), is added to the value of the loop filter.

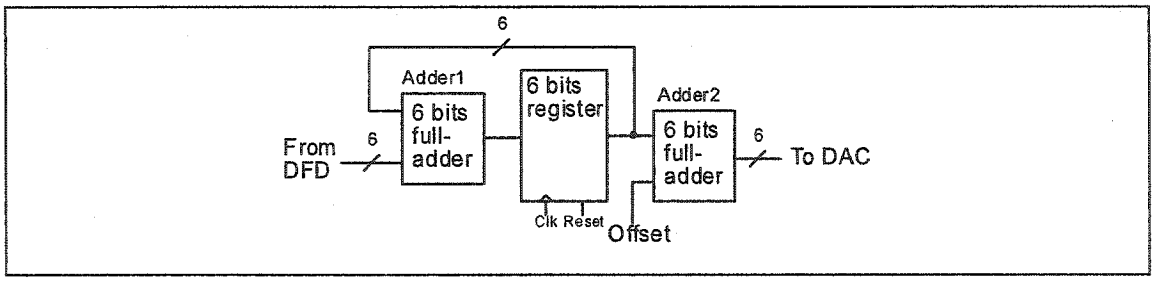


Figure 5.15 Digital first order filter.

5.3.5 Digital-to-analog converter

A digital-to-analog converter (DAC) is used as an interface between the digital filter and the VCO control. By changing the DAC output voltage, the center frequency of the VCO is changed.

Mixed-signal circuits need to be immunized against substrate coupling noise. This is especially important since the analog circuits are located closely to the digital ones. To reduce the sensitivity of the analog circuits substrate noise coupling, a fully-differential approach is used.

The DAC used in this circuit consists of a segmented thermometer-code current-mode digital-to-analog converter [Johns97]. The digital input is first converted to its equivalent current weight. A differential output is easily obtained using current-steering techniques. Implementing a non-segmented thermometer-code DAC would require 2^N cells to achieve the desired resolution. Because linearity is not a concern in this particular application, segmentation can be done to reduce the size of the DAC, and to reduce the complexity of the digital circuitry. Figure 5.16 shows the block diagram of the DAC.

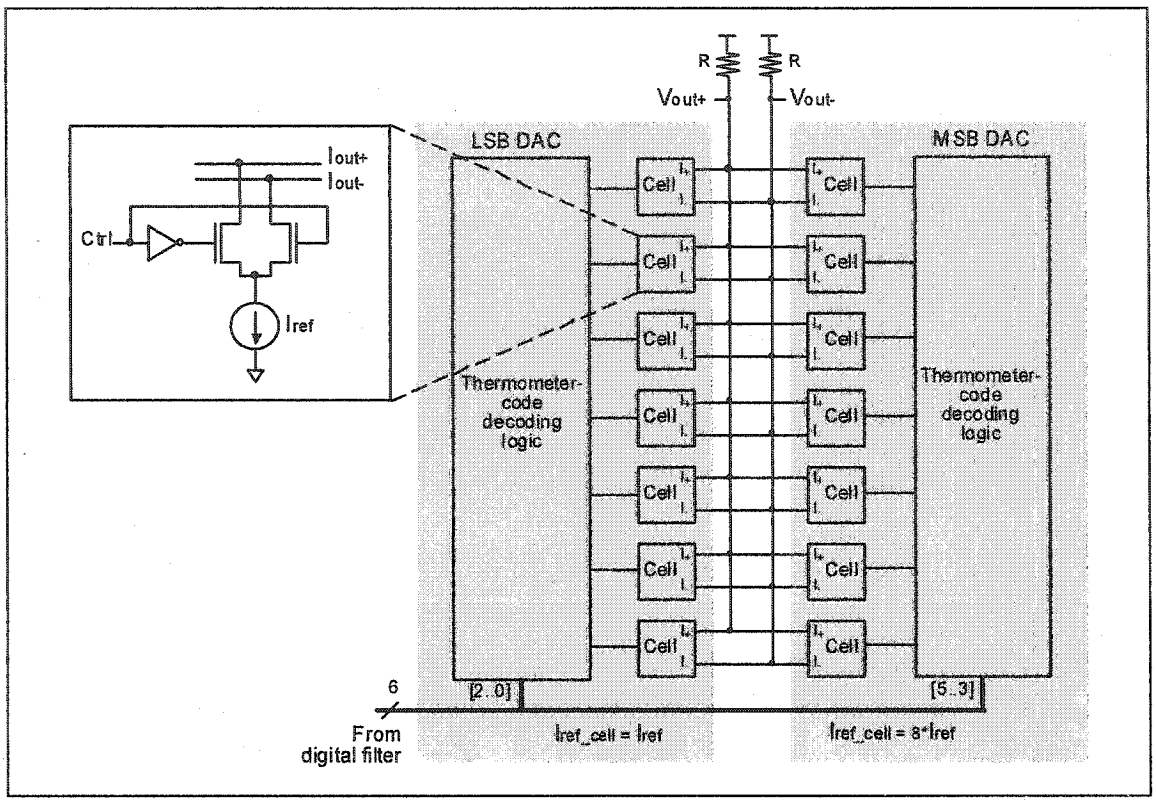


Figure 5.16 Block diagram of the DAC with cell details.

The digital filter outputs a 6-bit unsigned value, which needs to be converter by the DAC. Using segmentation, the three less significant bits are grouped together and form the LSB DAC portion. The MSB DAC portion, that has a current drive 8 times bigger than the LSB DAC, converts the 3 most significant bits. Since the outputs of both DACs are currents, the outputs can be summed together by simply connecting them together. The cells used in the DAC consist of a constant current source (I_{ref} or $8*I_{ref}$, depending if it is used in the LSB or the MSB DAC). The output current is steered between one branch to another by changing Ctrl signal. Cells from both the LSB and MSB DAC have the same architecture.

Because the VCO is controlled by voltage and not current, a current-to-voltage conversion is done using a simple resistor. The output voltage can be written as follow:

$$V_{out} = \left(I_{ref} In[2:0] + 8 I_{ref} In[5:3]\right) R, \qquad \text{Eq. 5.15}$$

where R is the resistor value, In[5:0] are the digital inputs and I_{ref} is the nominal current source value. The DAC gain defined in equation 5.15 can be changed by either varying the reference current a the resistor value. Figure 5.17 shows the operation of the DAC over its complete range of operation.

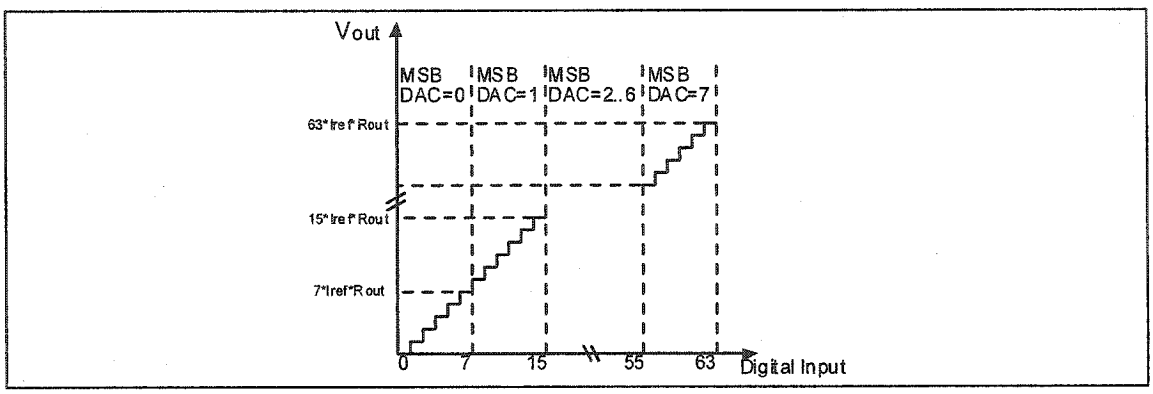


Figure 5.17 Operation of the DAC.

5.3.6 Voltage-controlled oscillator

Different VCO architectures can be used in a clock-and-data recovery system. LC oscillators can operate at very high speeds while having good phase noise characteristics. However, this type of oscillators is bulky, and its tuning range is limited. One way to increase the tuning range of an LC VCO is by the use of an array of switchable capacitances, to modify the center frequency. For low to moderate speeds, ring oscillators are attractive since they require a small amount of space and can be built with standard digital gates.

The VCO used is this design is an interpolation-ring CML voltage-controlled oscillator. It is implemented using standard CML gates. As described in [Chan97], interpolative VCO's are members of the ring oscillators VCO family. One advantage of using an interpolative VCO is the good control over the center frequency. By increasing the number of buffers in the delayed line (i.e. increasing the delay), the tuning range can be greatly broadened.

Using the Barkhausen's criteria, the condition to get an oscillatory circuit is to get a 180° phase shift, and a gain greater or equal to one at the wanted oscillation frequency. CML gates are good candidates to achieve this operation, since their gains are larger than one at high frequencies and inverted signals (180° phase shifted) are readily available due to their differential topology.

The heart of an interpolative VCO is the interpolator, which is simply a mixer. The mixer interpolates an output value from a signal and its delayed version. As illustrated in Figure 5.18, the mixing rate is simply the control voltage. Such a cell can easily be built in CML (section Chapter 8). The resulting output is given by:

$$Out = (Ctrl)In_1 + (1 - Ctrl)In_2, \qquad \textbf{Eq. 5.16}$$

where Ctrl is the mixing control signal, which is bounded between 0 and 1, and In_1 and In_2 are the input signals of the multiplexer. Because the multiplexer has a gain larger than one, the output signal will be regenerated. It means that the rate of the edges comes back to its original value.

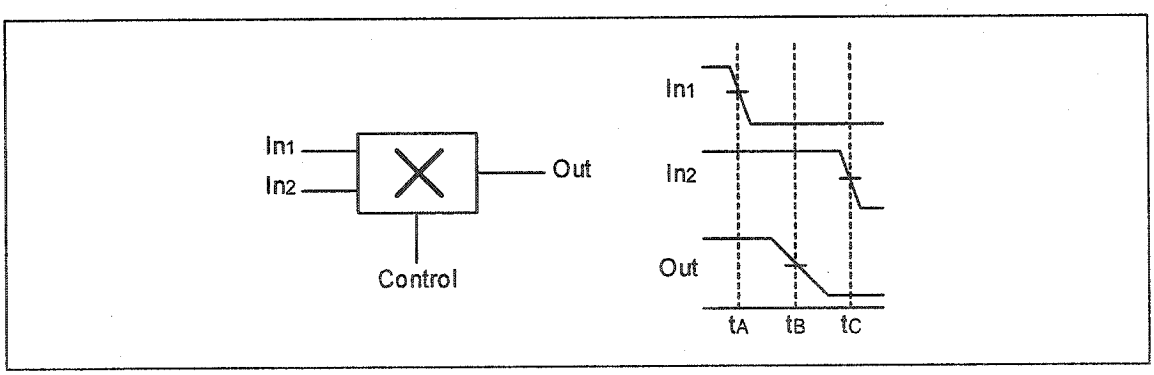


Figure 5.18 Edge mixing.

For a number N of stages, with a multiplexer delay of t_{muxdel} and a buffer delay of t_{bufdel}, it is possible to compute the oscillating frequency of the VCO:

$$T_{osc} = 2N \cdot t_{muxdel} + Ctrl \cdot t_{bufdel}. \qquad \textbf{Eq. 5.17}$$

The mixing control *Ctrl* signal is bounded between 0 and 1. Figure 5.19 is an example of a two stage interpolative VCO.

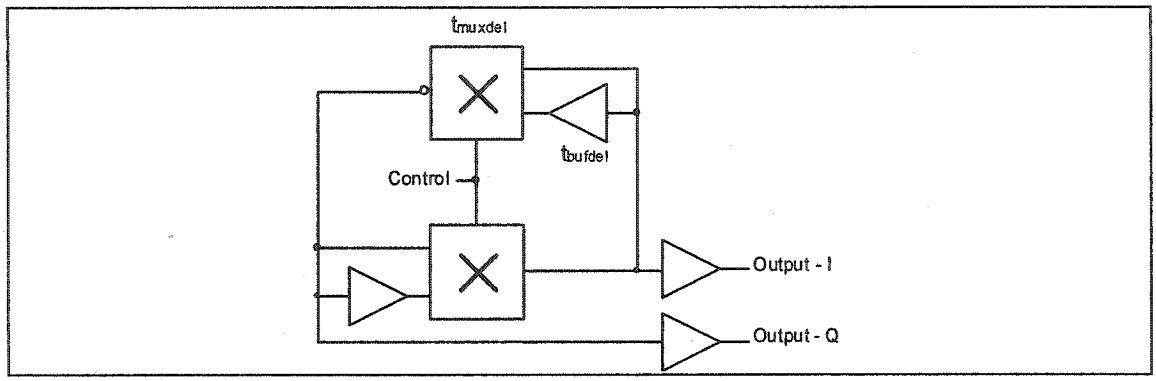


Figure 5.19 A two-stage interpolative VCO.

For an even number of stages, quadrature outputs are readily available because of the architecture. Depending on the number of stages used, it is possible to have a variety of phase shifted outputs. However, increasing the number of stages will reduce the operating frequency, as described by equation 5.17.

5.4 Conclusion

This chapter presented the motivation and the theory of using a digital frequency locked-loop to simplify and relax the design constraints on clock-and-data recovery circuits. The next chapter presents the simulation and measurement results of a fabricated prototype.

5.5 References

[Salama99] H. Djahanshahi and C. A. T. Salama, "Differential 0.35um CMOS circuits for 622MHz/933MHz monolithic clock and data recovery applications," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), vol. 2, Orlando, FL, pp. 93–96, June 1999.

[Greshishchev00] Y. M. Greshishchev, P. Schvan, M. Xu, J. Showell, J. Ohja, and J. E. Rogers, "A fully integrated SiGe receiver IC for 10-Gb/s data rate," IEEE J. Solid-State Circuits, vol. 35, pp. 1949–1957, Dec. 2000.

[Best93] R. E. Best, Phase-Locked Loops, Second Edition. New York: McGraw-Hill, 1993.

[ANSI] ANSI T1.105.03: SONET - Jitter at Network Interfaces.

[Razavi98] B. Razavi, RF Microelectronics, First Edition. Upper Saddle River: Prentice Hall, 1998.

[Hwang01] I. Hwang, S. Song, and S. Kim, "A Digitally Controlled Phase-Locked Loop With a Digital Phase-Frequency Detector for Fast Acquisition," IEEE J. Solid-State Circuits, vol. 36, pp. 1574–1581, Oct. 2001.

[Tanabe01] A. Tanabe, M. Umetani, I. Fujiwara, T. Ogura, K. Kataoka, M. Okihara, H. Sakuraba, T. Endoh, F. Masuoka, "0.18-μm CMOS 10-Gb/s multiplexer/demultiplexer ICs using current mode logic with tolerance to threshold voltage fluctuation," IEEE J. Solid-State Circuits, vol. 36, pp. 988–996, Jun. 2001.

[Johns97] D. A. Johns, K. Martin, Analog Integrated Circuit Design. New York: Wiley, 1997.

[Chan97] W. Chan, J. Lau,A. Buchwald, "A 622-MHz interpolating ring VCO with temperature compensation and jitter analysis," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), vol. 1, Orlando, FL, pp. 25–28, June 1997.

CHAPTER 6
DFLL Simulation and Measurement Results

This chapter presents the measurement results of two circuits: a ring-interpolative voltage-controlled oscillator and a digital-frequency-locked loop. The ring interpolative voltage controlled oscillator has a structure similar to the one used in the DFLL but running at twice the frequency. The VCO was built on a separate chip for characterisation. A discussion of the test setup and the measurement results are presented.

6.1 Ring-interpolative voltage-controlled oscillator

VCO's are characterized by their frequency of oscillation, tuning range, phase-noise and power dissipation [Mostafa01]. Most of these parameters can be simulated with a conventional simulator.

Figure 6.1 is a plot of the simulated response versus the tuning voltage of the VCO. As one can see, the linear tuning range of the VCO is quite large, from 2.7GHz to 3.3GHz, or 20% of the center frequency. The operation of the VCO is limited by the valid voltage that can be applied at the input of the multiplexer. This valid range is designed to be between 1.0V and 1.5V, where 1.25V is the setting of the center frequency.

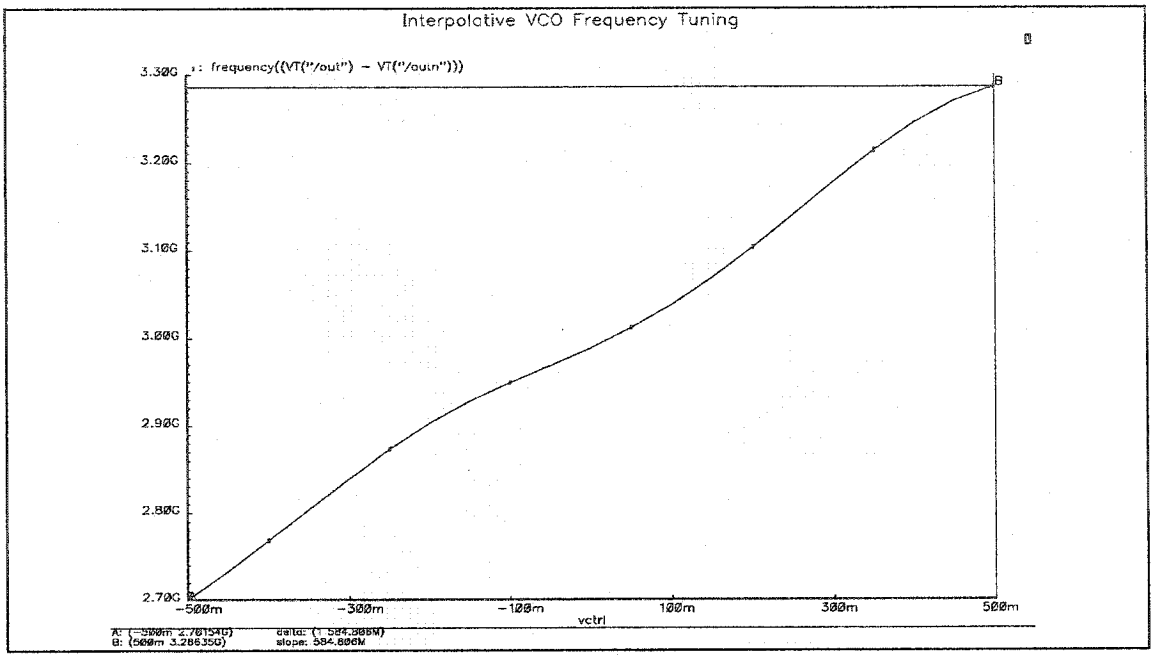


Figure 6.1 Stand-alone VCO tuning (simulated).

Figure 6.2 shows the center frequency measurement of the manufactured VCO. The linear range of operation is located between -0.4V and 0.4V differential voltage input. This is translated to absolute levels of 1.05V to 1.45V, (accounting for the CML gate and power supply constraints).

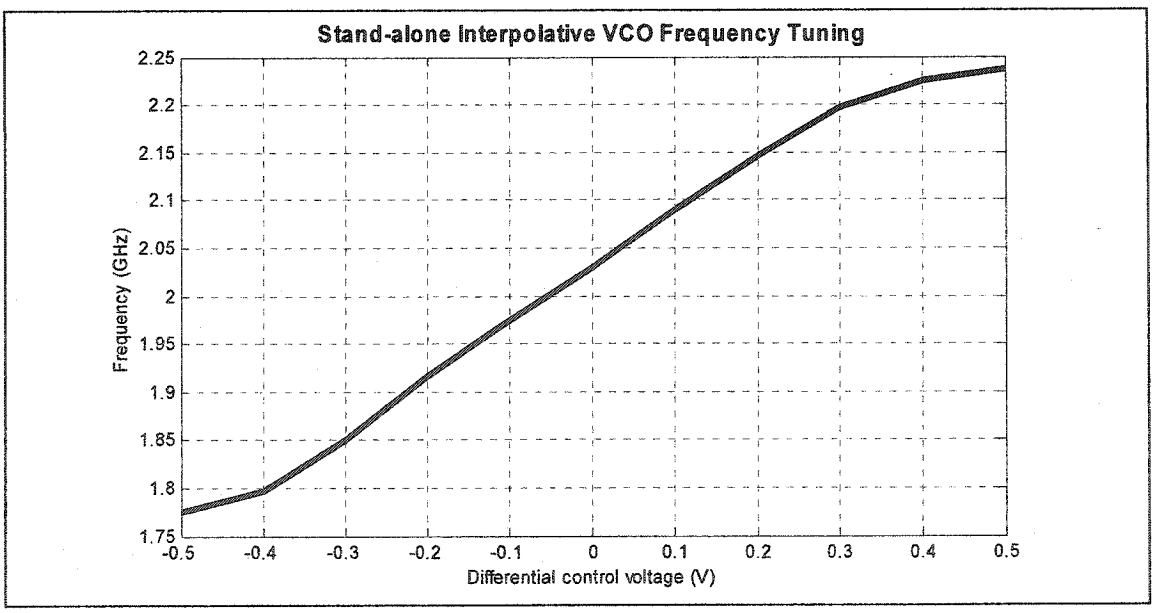


Figure 6.2 Stand-alone VCO tuning (measured).

Table 6.1 is a summary of the characteristics of the interpolative ring oscillator used in the digital frequency-locked loop circuit. Simulation results are from the layout of the oscillator with the parasitic capacitances extracted. Measurement results are taken from the actual manufactured circuit.

	Simulation (extracted)	Measured
Center frequency	2.99GHz	2.03GHz
Tuning range	584MHz (19.54%)	462MHz (22.7%)
VCO gain (K_{VCO})	564MHz/V	571MHz/V
Current	35.9mA	35.3mA
Power consumption	53.9mW	53.0mW
Output amplitude	$715mV_{ppd}$ (LVDS levels)	$700mV_{ppd}$ (LVDS levels)
Phase noise	Not simulated	-64dBc @ 1MHz

Table 6.1 Summary of VCO characteristics.

The large discrepancy between the measured and simulated center frequencies comes from the parasitic capacitances present on the internal nodes of the ring oscillator. As stated in 5.3.6, the propagation delay of a CML gate is directly proportional to the time constant of the output node of the gate. The time constant is defined as the product of the CML load, which is normally a resistance, and a parasitic capacitance. These parasitic capacitances are the gate capacitances of the next stage, coupled with the trace-to-substrate capacitance. Another parasitic capacitance that has to be taken into account is the side wall capacitance between two traces. In modern CMOS processes, metal traces can be placed down to 0.30μm apart. However, these traces can be up to 1μm thick. If the layout is done using the top metal layers in order to reduce to a minimum the capacitance to the substrate, the dominant capacitance would be that of the side wall coupling effect between two adjacent traces. At very high frequencies, this problem is easily observable, as experienced in with the results of the oscillator presented earlier.

Different techniques are available to correct this issue. Placing differential traces wide apart, e.g. using 2 to 3μm spacing, will help reducing considerably the side wall capacitances. This will also reduce the fringing capacitance which are non-negligeable for small dimension layout structures. To reduce the sensitivity of other circuits to the interference caused by the high speed transitions, shielding can be used [Greshishchev00]. Figure 6.3 illustrates this process. Metal traces are run on the high metal layers while the shield is created from a low metal layer. Because the shield is normally DC grounded and offers a low impedance, isolation to the substrate is created. In addition, this offers a low inductance path for high-frequency currents. Adding a protective shield has to be done very carefully. A bad layout example is if the shield is implemented in metal M_{N-1}, while the differential traces are ran on M_N. This will create a large C_{shield}, which adds an unnecessary load to the circuit nodes!

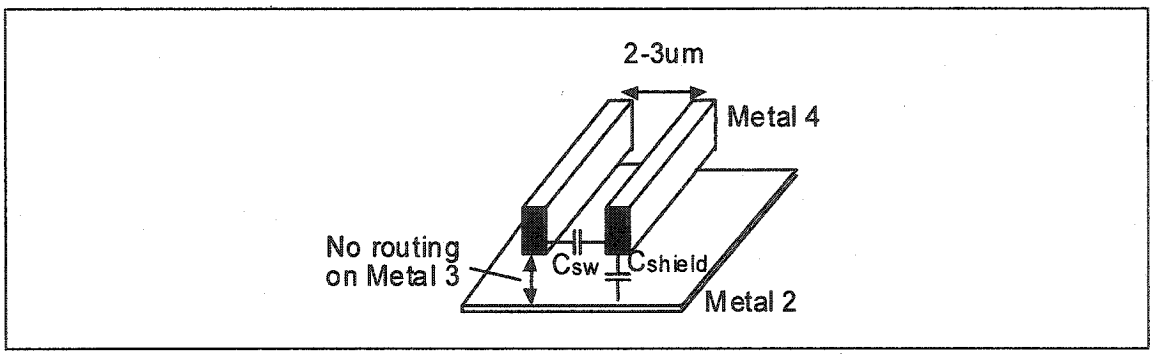


Figure 6.3 Proper shielding example of differential signals.

Another solution to counteract the capacitance increase on the output nodes is to increase the driving strength of the CML gates to improve the rise and fall times. The driving strength is increased by reducing the resistive load of the CML gate, and adjusting its tail current accordingly. This results in a reduction of the RC time constant at the output of the gate. Increasing the tail current will keep the output voltage amplitude constant, a necessary condition for good operation.

Compared to LC VCO's, ring interpolative oscillators produce higher phase noise. This result is expected since ring oscillators are composed of many active elements whereas LC VCO's are built from a limited number of active components. Reasonably low phase

noise can be achieved by a ring oscillator if careful optimization is done. Because of their compact layout sizes (no bulky passive components needed, i.e. inductors, capacitors) and wide tuning range, they are excellent candidates for high-performance telecommunication applications [Lee02] [Djahanshahi99] [Iravani99].

6.2 Digital frequency locked-loop

The DFLL is better characterized in the time domain. The dynamics of the circuit can be measured by probing the control voltage of the VCO. The control voltage is directly proportional to the digital control word output by the DFLL.

The DFLL can be operating in on of the three following modes:

1. The preset frequency is too low: the VCO frequency needs to be reduced.

2. The preset frequency is too high: the VCO frequency needs to be increased.

3. The preset frequency is outside the valid range: the VCO cannot reach the desired frequency.

In the first two cases, the DFLL will eventually lock to the desired frequency. The locking time is determined by the loop dynamics, i.e. the filter time response. In the last case, the DFLL will remain indefinitely in the acquisition mode. Compared to an analog PLL, where the control voltage will reach the upper or lower limit of the circuit, providing a quasi-fixed frequency, the DFLL will have a constantly changing output frequency. This phenomenon is explained by the overflow of the counter present in the digital frequency detector and the loop filter. If the VCO control voltage was modulated in a particular way, this circuit could be used in spread spectrum clock generation. By digitally controlling the digital sequence applied to the DAC, it would be possible to modify the spectrum, depending on the application and the specifications [Chang03].

The two figures below show the simulated transient response of the DFLL with different frequency settings.

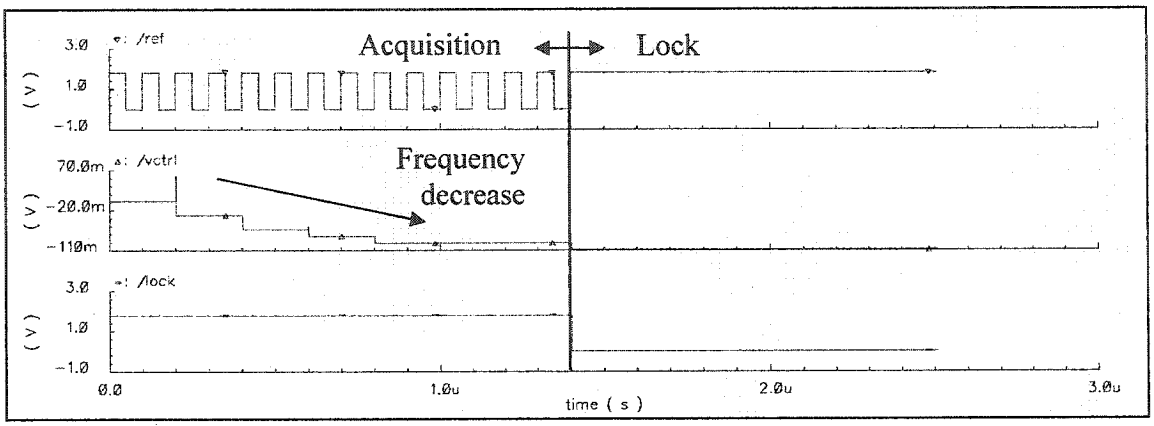


Figure 6.4 Simulation of locking when the VCO frequency is too high.

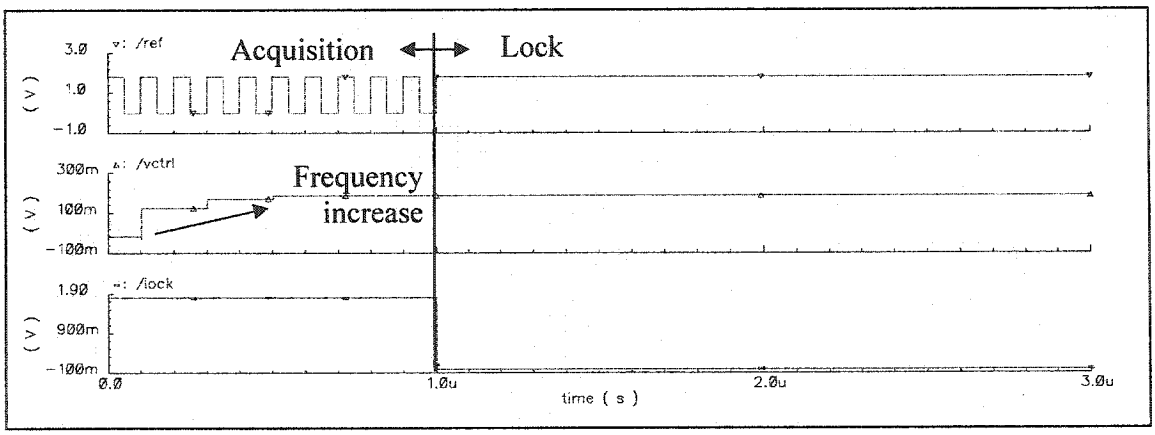


Figure 6.5 Simulation of locking when the VCO frequency is too low.

Three signals are shown on each figure: the *ref* signal is a gated clock, which is provided by an external reference source anded with the lock. By gating the external clock (disabling the digital portion of the circuit), noise coupled to the substrate due to the digital circuitry is greatly reduced when the circuit is locked, enhancing the performance of the analog portion of the chip. The *ctrl* signal is the analog differential control voltage applied to the VCO. And the *lock* signal indicates when the system has reached the desired frequency of operation. It also serves to disable the digital portion of the circuit by gating the input reference clock.

For the two examples above, the VCO is able to achieve the desired frequency within 1.4µsec which is equivalent to 14 reference clock cycles. This might seem excessive compared to the one clock cycle predicted by the theoretical derivation of section 5.3.1. However, this is acceptable since it is difficult to correctly predict the gain of each element in the circuit (e.g. VCO and DAC). Another limitation when implementing the DFLL is the dynamic range of the DAC. The DAC output voltage is bounded between the power supply and a lower bound set by the biasing current and the resistor value used for current-to-voltage conversion. For any reason, if the DAC needs to output a value greater than V_{DD}, the output will be clipped.

The third mode of operation is when the preset frequency lies outside the valid range of operation. Figure 6.6 is a measurement plot of the differential VCO control voltage. In this particular example, the preset frequency is too high compared to the VCO center frequency, producing a positive slope stair case waveform. The digital control wraps from the maximum to the minimum because of the non-saturating property of the counters present in the circuit. This behaviour was discussed earlier.

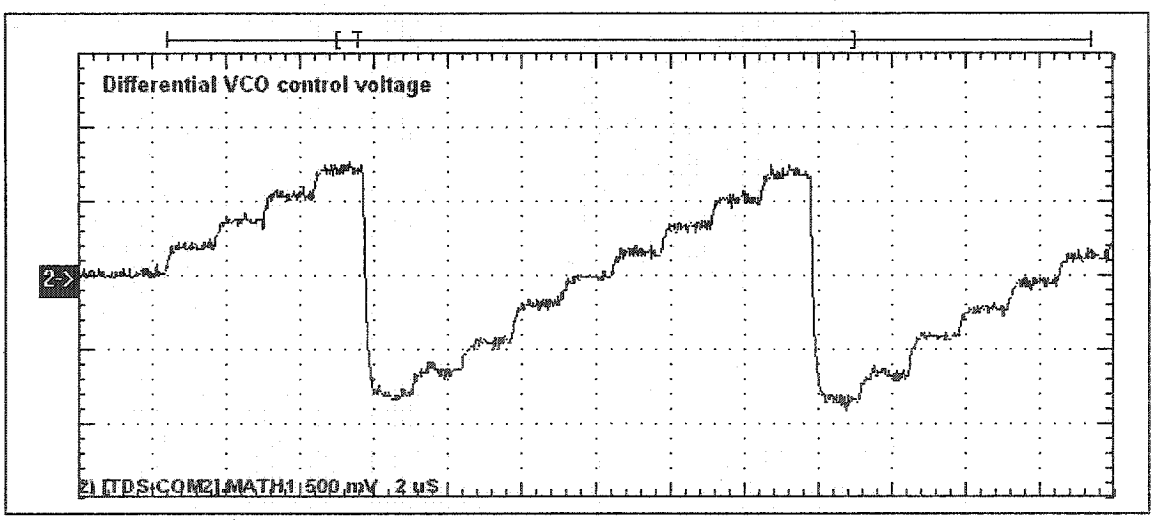


Figure 6.6 Measurement: invalid preset frequency setting.

Figure 6.7 and Figure 6.8 are measurement results of the differential VCO control voltage for different preset value. Two control voltage values are shown in Figure 6.7. The preset value N is set such that the VCO frequency is too low in the case where N=19 and too

high for the case where N=16. As evident, the control voltages increase and decrease respectively to compensate for the difference in VCO cycle counts. Figure 6.8 has the lock signal displayed in addition of the control voltage, showing the locking time of the loop, after a reset is applied. The locking time depends on the preset value. In this example, with N=16, the loop locks itself within 1.3μses. For N=17, the locking time increases to 3.2μses. The locking time is not a linear function of N because of the quantization errors present in the system.

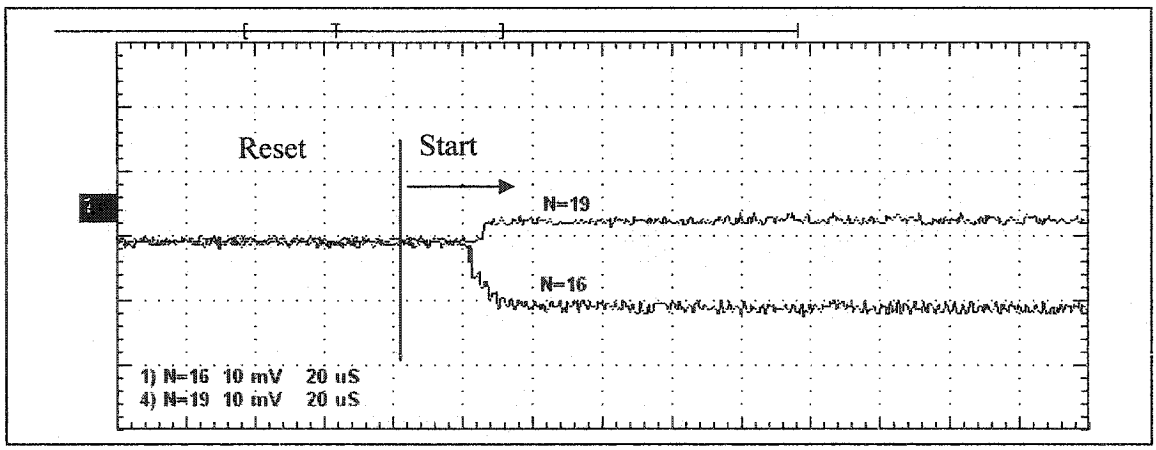


Figure 6.7 Measurement: different control setting.

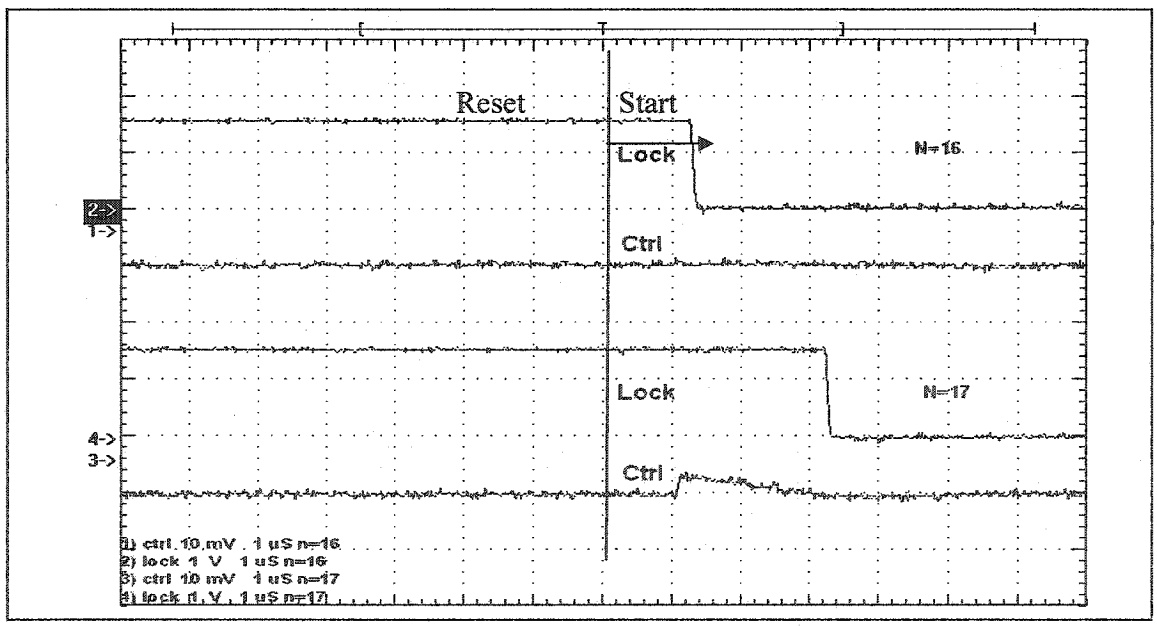


Figure 6.8 Measurement: control and lock signals.

The particular example of Figure 6.8 shows how the quantization error can affect the control voltage and the VCO frequency. Two different runs are plotted on the same graph, all starting at the same moment in time. The first two curves show the *lock* and *ctrl* signals for a setting of N=16. When *lock* is asserted (going to low), the loop has acquired the targeted frequency. This is represented by a differential control voltage of 0V. The second set of curves is for a setting of N=17. The locking time is longer. An explanation is due to the quantization error associated with the discrete nature of the loop. If the reference clock and the on-chip VCO are not in phase, a difference in VCO cycles count can appear (plus or minus one cycle count). This results in an error signal being under or over-estimated. For the example where N=17, the final control voltage value is identical to the control voltage of N=16. As described by equation 5.9, this is a valid value since the circuit has an intrinsic quantization error due to its discrete nature.

The waveform observed in Figure 6.9 is an example of the control voltage observed for an invalid setting. The N preset value was set such that the VCO frequency is too small. This is confirmed by the positive slope of the control voltage when the time is increased. The non-linear waveform behaviour is explained in the next section.

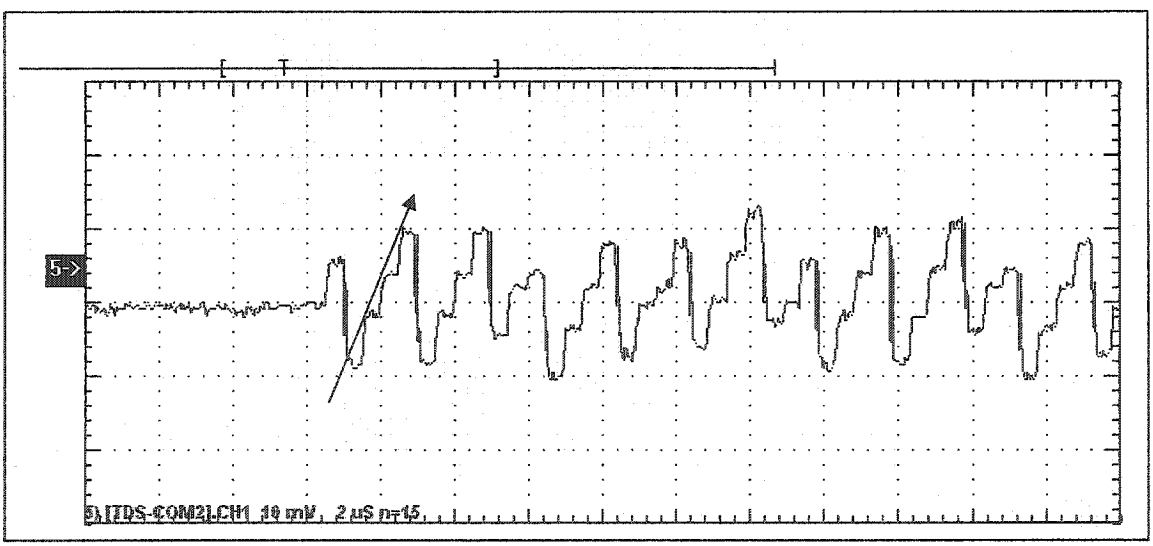


Figure 6.9 Measurement: invalid frequency setting.

6.2.1 Qualitative non-linear analysis of the DFLL

A non-linear model of the loop built in Matlab is required for in depth analysis of the non-linear effects in the circuit (Figure 6.10). Non-linearities are caused by the counter wrap-around feature, the adder overflow, the DAC output voltage clipping, and the VCO frequency tuning range limitation. These are the major contributors setting the limits on the circuit operation.

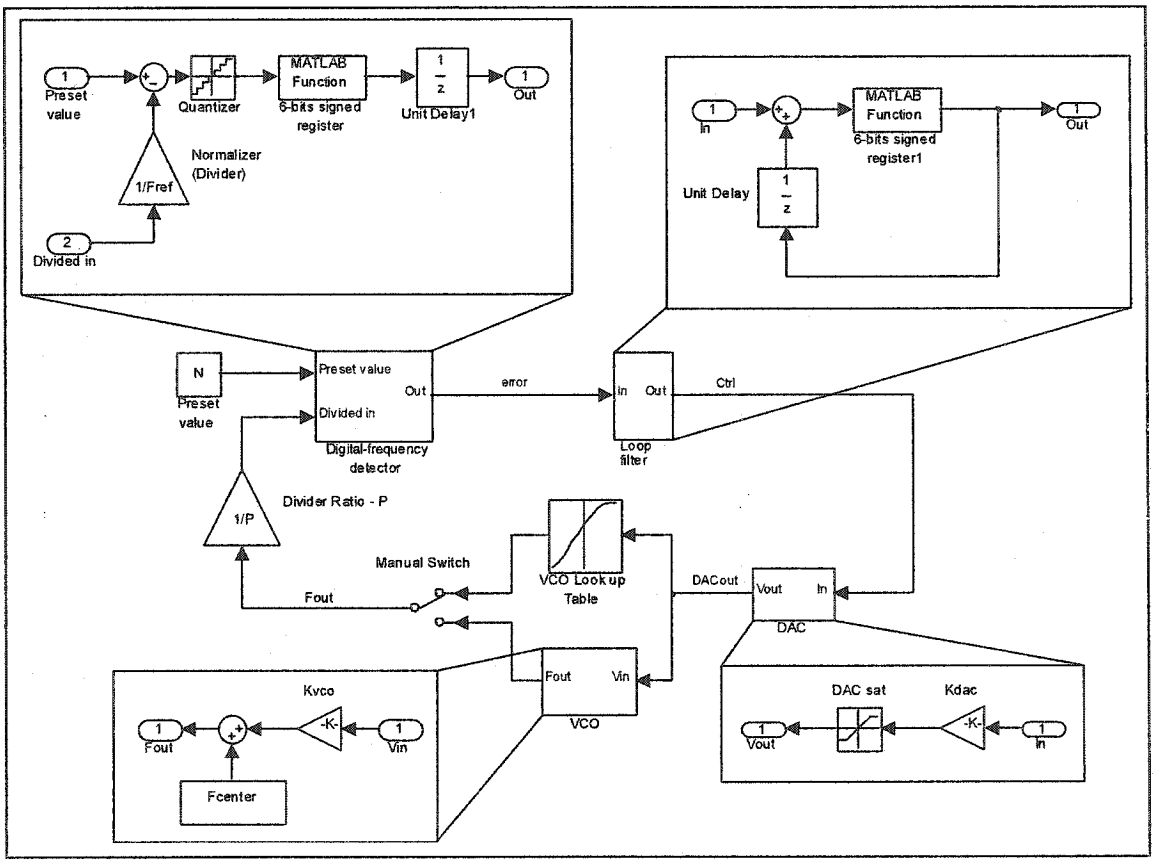


Figure 6.10 Non-linear model used in Matlab.

The digital-frequency detector computes the difference between the preset value N and the actual count of divided VCO cycles. A division by F_{ref} is necessary to normalize the input frequency to the reference. The difference is quantized to reflect the discrete nature of the system. A wrap around function is coded in Matlab to model the wrap-around affect of a digital adder.

```
function out = signed_adder(in, width)

% Usage: Outputs a signed value, depending on the input
%        and the adder width
% April 2003
%
% in    : input value
% width: register width in bits (i.e. 6 for a 6 bit adder)

max = 2^width ;       % maximum unsigned value of the register + 1

val = mod(in,max) ;

if val < (max/2)
    out = val ;
else
    out = val - max ;
end
```

The same function is used in the digital filter, since an adder is also used. The saturation effect of the DAC is also modelled. To simplify the model, limits of ±0.5 are set, which corresponds to the maximum differential voltage output by the DAC when clipped by the power supply. The gain of the DAC can also be changed in order to reflect its effect on the measurements. Two models for the VCO are used. The first model consists of a linear approximation of a real VCO. It consists of a simple multiplier, converting the input value (voltage) into frequency. A better approximation of the VCO behaviour consists of using a lookup table. The table is filled with measurement values. The lookup table has the advantage of reflecting the true behaviour of the VCO. In this case, frequency clipping (i.e. decrease in K_{vco}) is present at the limits of the tuning range.

By replacing the Matlab variables by their measured value, it is possible to observe the same transient operation of the loop. Figure 6.11 is an example of such a non-linear transient behaviour. This control waveform is similar to the one presented in Figure 6.9.

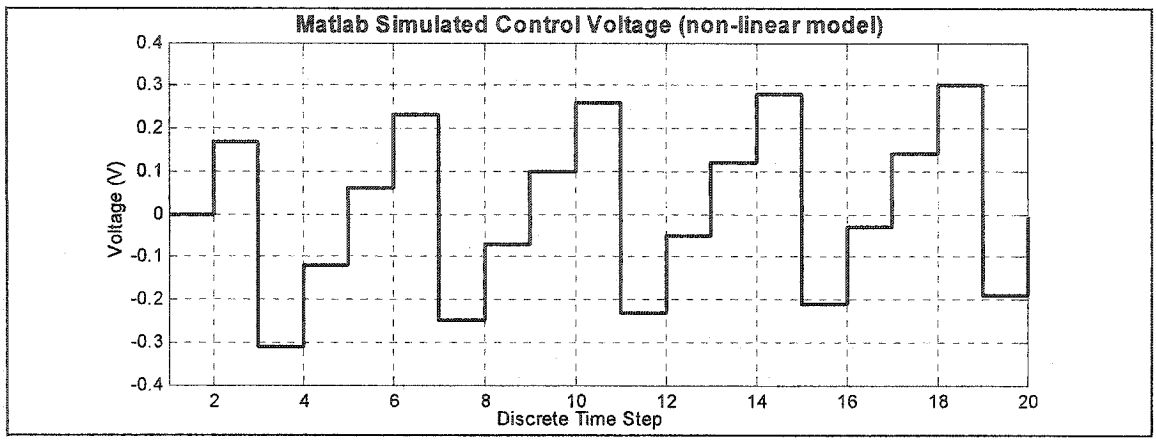


Figure 6.11 Non-linear loop effect.

6.2.2 DFLL VCO Characteristics

The VCO characteristics can be measured indirectly through a divided-by-8 clock signal. The frequency tuning range of the VCO is shown in Figure 6.12. The linear range is between -0.4V and 0.3V. The tuning range is limited to approximately 205MHz or 20.6% of the center frequency of 1.045GHz. This is in agreement with the 2.0GHz stand-alone VCO measurement of section 6.1. The VCO gain (K_{VCO}) is 264MHz/V.

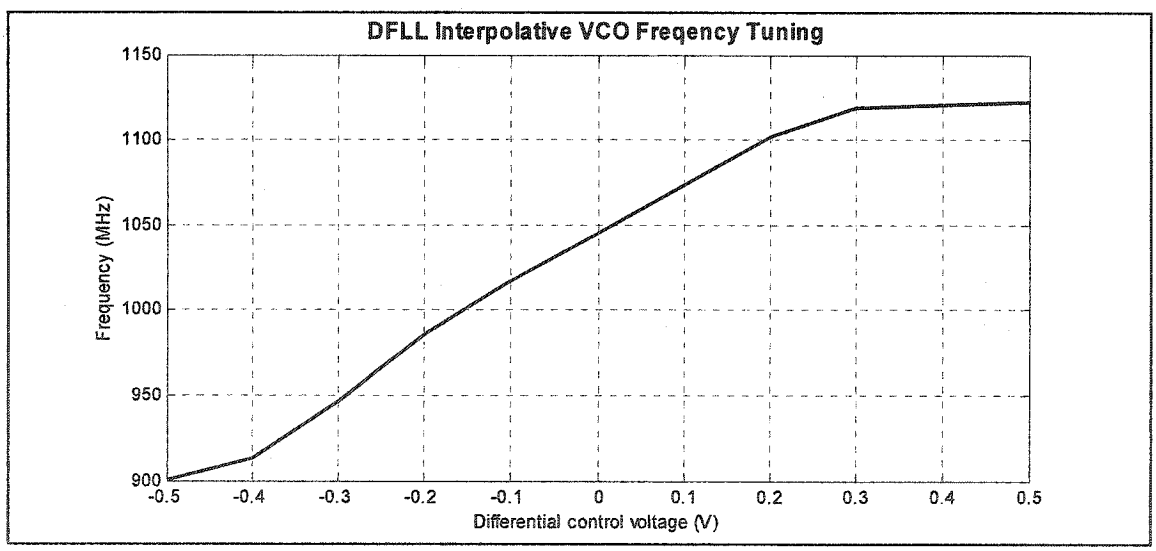


Figure 6.12 DFLL interpolative VCO frequency tuning.

To widen even more the tuning range, one could have added more delay elements in the delay path of the multiplexer (section 5.3.6). By using a lower number of stages (i.e. 2

instead of 4), the center frequency can be increased to compensate for the increase in the delay element.

Having a limited tuning range is problematic if the VCO's center frequency has a major offset from the designed frequency. In this case, the VCO will be operated close to its non-linear region. This will change the behaviour of the loop and can seriously compromise the operation of the system. Therefore, it is important to have a large tuning range to correct for the VCO center frequency offset.

Figure 6.13 is the measured spectrum of the VCO output divided-by-8. The center frequency is 130.2 MHz. This gives a true VCO frequency of 1.042 GHz. Again, the VCO frequency is smaller than the designed/simulated value. Since the VCO structure is based on the same topology of the VCO presented in section 6.1, it is expected to have such a difference. The major difference between the two VCO structures is the number of multiplexers used to create the delay. In the case of the DFLL, 4 delay elements were used where as 2 were used for the stand-alone VCO.

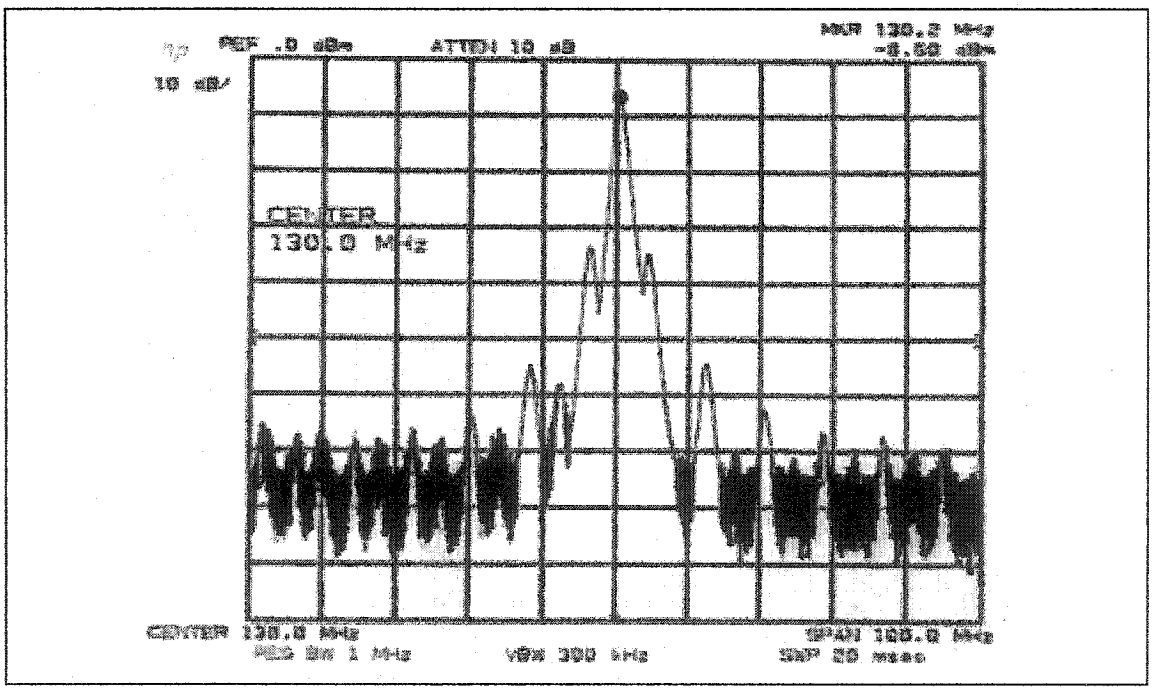


Figure 6.13 Divided-by-8 power spectrum when locked.

Of interest are the side-lobes present on each side of the center tone. These tones are approximately spaced 5MHz apart from the center tone. They cannot come from noise modulated from the VCO, as the interference noise would be around 40MHz (5MHz * Divider ratio of 8 = 40MHz). There is no sub-circuit in the chip operating at this frequency. Another explanation would be a strong interference source present in the lab. This explanation is not valid since the chip was housed in a thick metal box, acting as a Faraday cage (section 6.3). It is believed that this interference comes from coupling between the reference clock signal and the divided output. This is explained by the close proximity of the two signals, on the chip boundary. A solution is to separate sensitive signals, even medium frequency signals by interleaving a ground signal between them. This will help to keep the signal integrity at a higher level.

6.2.3 DC characteristics

The power consumption of the circuit can be measured for each of the two operation modes of the circuit. During acquisition, the power consumption is expected to be higher because of the digital circuitry being in operation. In the hold mode, only the DAC, VCO and dividers are powered on. Table 6.2 shows the simulated and measured power consumptions. All measurements were made for a supply voltage of 1.5V.

Due to the fast behaviour of the circuit, it is difficult to measure precisely the power consumption when the circuit is operating in the acquisition mode. This measurement was not taken. Power measurement in the hold mode is straight forward. Simulated and measured power consumption agree with each other. Because the VCO is built from CML cells drawing a constant current from the power supply, even if the VCO frequency is off. This fact is also confirmed by the stand-alone VCO power supply current measurements.

	Simulation		Measurement	
	Current (mA)	Power (mW)	Current (mA)	Power (mW)
Acquisition	37.7	56.6	-	-
Hold mode	35.9	53.9	35.3	53.0

Table 6.2 DFLL power consumption.

6.3 *Physical test setup*

As described in Chapter 4, the test setup used for characterizing a circuit is as important as the design of the circuit itself. Figure 6.14 shows how the integrated circuit was mounted on a printed circuit board, itself housed in a metal enclosure.

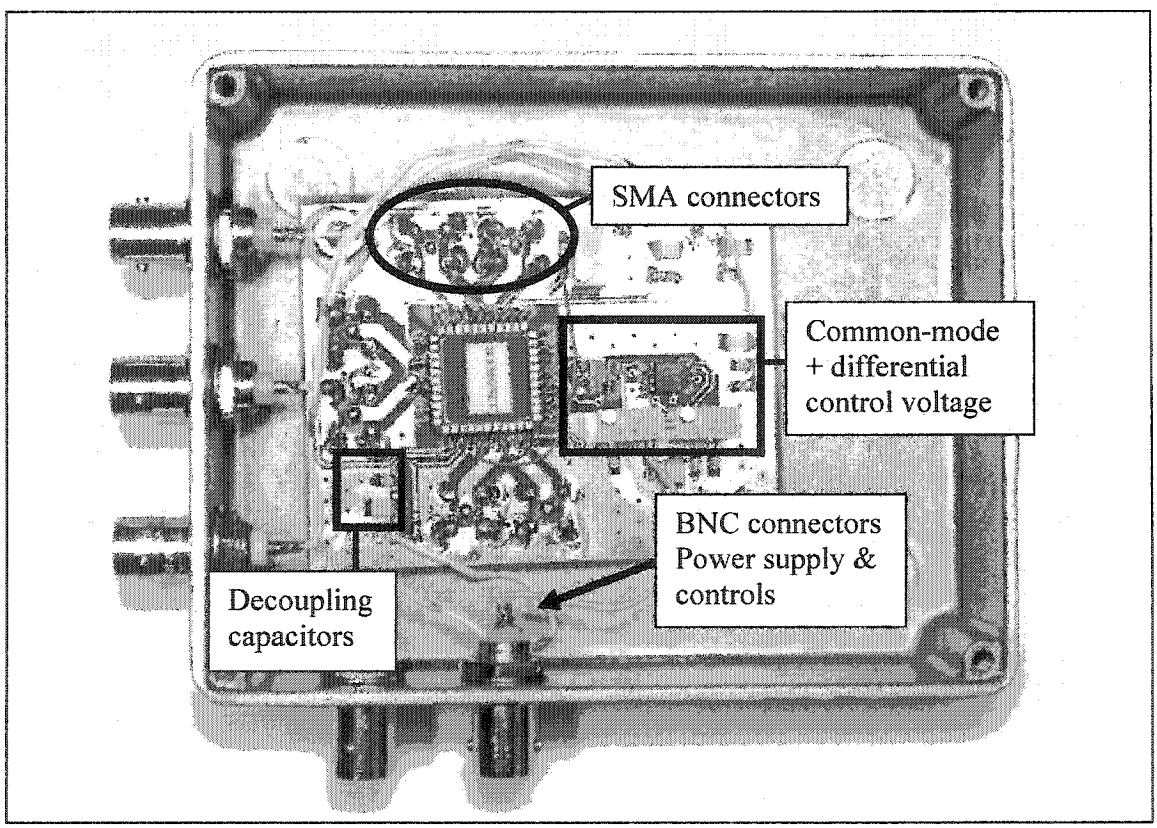


Figure 6.14 The VCO chip test setup.

A custom PCB was built for testing the prototype chip of the interpolative ring oscillator. The custom PCB allows maximum flexibility for signal routing and testing. A standard two-layer FR4 board was used. High-speed signals are routed using transmission lines, to conserve signal integrity. The ground connection is the most important signal in a design. This is why a ground plane, which covers the complete bottom plate of the board, is used to make easy and direct connections to the chip. In addition, the ground plane forms parts of the transmission lines used to route the high-speed signals [Johnson93].

Routing of control, high-speed and power signals was done so that crossing was kept to a minimum. This reduces the crosstalk between different signals to a minimum. To make the ground connections even more accessible for the chip, unused space is flooded with metal, connected through vias to the bottom ground plane. SMA connectors are used for the high-speed signals. Power connections are done using BNC connectors [Beaudoin02].

To reduce the connectors head count, an electronic circuit was added to set the control voltage of the VCO. The circuit, illustrated in Figure 6.15, is a simple voltage follower. By changing the resistor value of P1 and P2, the common-mode voltage can be set, and the differential control voltage is varied.

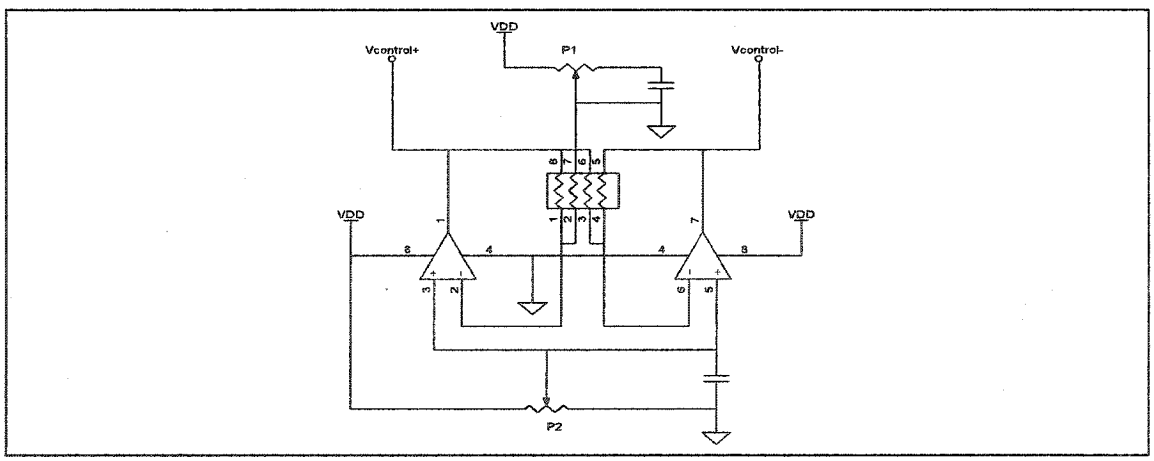


Figure 6.15 Active control voltage setting.

The PCB is housed in a metal enclosure that acts as a Faraday cage. This helps to keep the outside interference to a minimum. It also simplifies the testing. Because the chip is

fully-packaged (i.e. no on-chip probing was possible), and all the signals are accessible through controlled-impedance connectors, the measurements setup can be completed in a matter of minutes. Also, because the circuit is immune to external electro-magnetic noise sources, it is not necessary to have a testbench sitting in a Faraday cage.

The same kind of setup was used successfully to test the prototype chip of the digital-frequency-locked loop (Figure 6.16). To reduce the BNC connectors head count, digital signals such as the preset value N and the internal filter register values are connected to dip-switches and simple through-hole connectors respectively. Analog signals and power supplies are interfaced to the outside world with the help of BNC connectors. Because separate power supplies exist on the chip (i.e. digital, high-speed analog), many power supply pins surround the package. Each power supply can be powered individually, or they can all be connected to the main power supply by a solder-blob jumper. To control the VCO common-mode voltage, the same circuit as explained earlier was used.

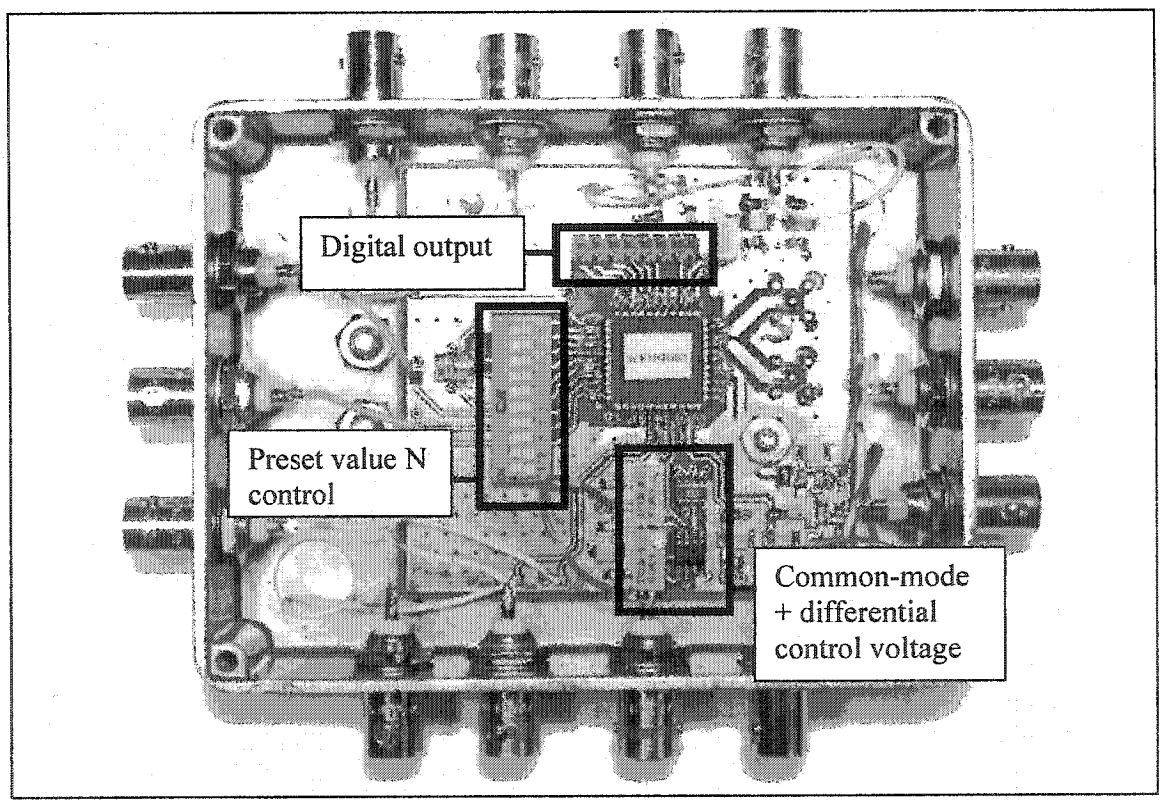


Figure 6.16 The DFLL chip test setup.

6.4 Microphotograph

A microphotograph of the DFLL chip is shown in Figure 6.17. The die size was limited by the number of pads available instead of the size of the circuit itself. The digital portion of the circuit was laied out by hand. Basic cells (NOT, AND, OR, ...) were custom made. The packing density of the digital gates is around 20Kgates/mm^2. This is one fifth of the maximum density that TSMC is claiming (~100Kgates/mm^2) [TSMC] [CMC] for the standard 0.18μm technology. The digital circuitry has a lot of empty space. It should be noted that space and density was not a driving factor for this design.

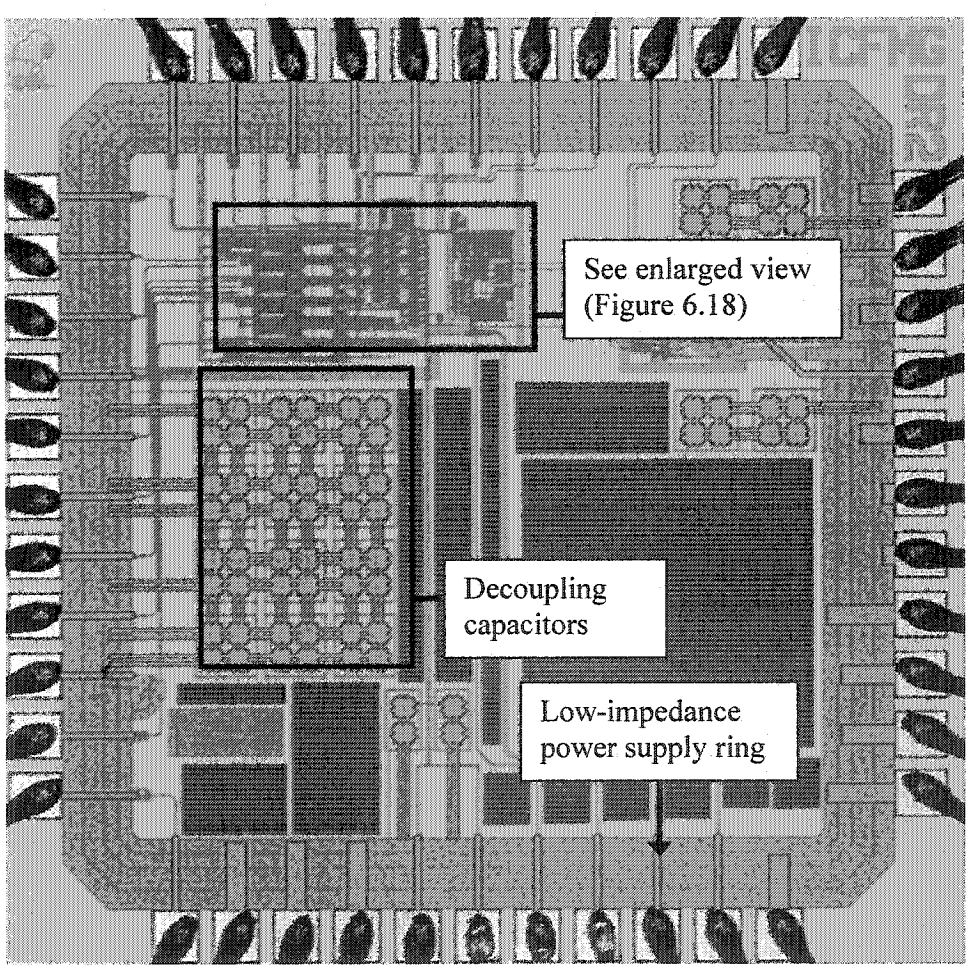


Figure 6.17 Chip microphotograph of DFLL chip.

A power supply ring runs around the chip. It provides very low impedance interconnect to the power supply rails. Keeping the power supply as stable as possible is critical for the analog portion of the circuit. For this design, two rings are used: one for the digital portion and one for the analog portion of the design. Multiple pads are used to connect the power supplies to the off-chip. The die size is approximately 1.4mm x 1.4mm.

Figure 6.18 is an enlarged view of the DFLL circuit itself. To keep the layout effort as low as possible (since it was all done by hand), the digital portion of the circuit was done by repeating bit slices. Each slice was connected to the other at the end. The DAC structure and the VCO are surrounded by guard rings to minimize noise coupling through the substrate. This is of concern, since the digital circuitry is next these to these structures. In addition, the two structures are built from fully differential elements, which rejects some of the common-mode noise. Low impedance ground connections are readily available to the VCO.

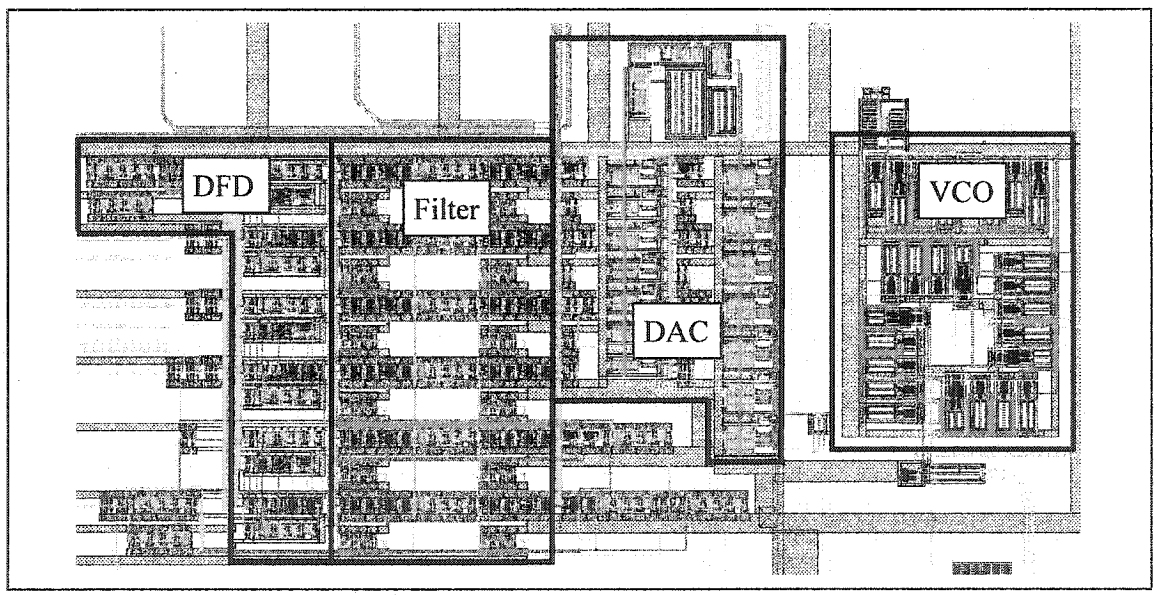


Figure 6.18 DFLL chip zoom on the digital circuitry.

6.5 Conclusion

This chapter presented the simulation and measurement results for a stand-alone ring-interpolative VCO and a digital-frequency-locked loop. Discrepancies between the simulated and measured data can be explained by the added parasitic capacitances present on the internal nodes of the VCO's. These parasitic capacitances were not extracted by the extractor (Diva) because of the tool limitations. A successful prototype with a discussion of the limitations of the DFLL implementation was presented. The circuit has a limited frequency tuning range due to the non-linear effects of the adder use of in the circuit and the VCO tuning range. Finally, a test setup is described to facilitate the testing of the fabricated circuits. By using a metal enclosure and a custom-made PCB, paired with an external circuit built from discrete components, it is possible to do precise measurements, without the use of on-chip probing.

6.6 References

[Mostafa01] A.H. Mostafa, M.N. El-Gamal, R.A. Rafla, "A Sub-1-V 4-GHz CMOS VCO and a 12.5-GHz oscillator for low-voltage and high-frequency applications," IEEE Transactions on Circuits and Systems II, vol. 48, pp. 919-926, Oct. 2001 .

[TSMC] http://www.tsmc.com

[CMC] http://www.cmc.ca

[Greshishchev00] Y.M Greshishchev, P. Schvan, "SiGe clock and data recovery IC with linear-type PLL for 10-Gb/s SONET application," IEEE Journal of Solid-State Circuits, Volume 35, Issue9, pp. 1353-1359, Sep. 2000.

[Lee02] Sang-Hyun Lee, Moon-Sang Hwang, Youngdon Choi, Sungjoon Kim, Yongsam Moon, Bong-Joon Lee, Deog-Kyoon Jeong, Wonchan Kim, Young-June Park, Gijung Ahn, "A 5-Gb/s 0.25-/spl mu/m CMOS jitter-tolerant variable-interval oversampling clock/data recovery circuit," IEEE Journal of Solid-State Circuits, Volume 37, Issue 12, pp. 1822-1830, Dec. 2002.

[Djahanshahi99] H. Djahanshahi, C.A.T. Salama "Differential 0.35 μm CMOS circuits for 622 MHz/933 MHz monolithic clock and data recovery applications," ISCAS '99. Proc. of IEEE International Symposium on Circuits and Systems, vol.2, pp. 93-96, Jul. 1999.

[Iravani99] K. Iravani, F. Saleh, D. Lee, P. Fung, P. Ta, G. Miller, "Clock and data recovery for 1.25 Gb/s Ethernet transceiver in 0.35 μm CMOS," Proc. of CICC, pp. 261-264, 1999.

[Chang03] Hsiang-Hui Chang, Hui Hua, hen-Iuan Liu, "A spread-spectrum clock generator with triangular modulation," IEEE Journal of Solid-State Circuits, Volume 38, Issue 4, pp. 676-676, Apr. 2003.

[Beaudoin02] F. Beaudoin, M. N. El-Gamal, "A 5-Gbit/s CMOS optical receiver frontend," The 2002 45th Midwest Symposium on Circuit and Systems, Volume 3, pp. 168-171, August 2002.

[Johnson93] H. Johnson, M. Graham, "High-speed digital design, a handbook of black magic," Prentice-Hall, New-York, 1993.

CHAPTER 7

Conclusion

This thesis addressed the issues associated with the design, the implementation, and the testing of an optical front-end receiver. Chapter 2 discussed the common metrics used in optical systems. These metrics can give to the designer good insight on the limitations and kind of setups needed for testing. The two following chapters discussed the design and measurement results of a 5.0Gb/s optical front-end.

In addition, a circuit to help broaden the acquisition range of clock-and-data recovery circuits has been presented. This circuit is not limited to CDRs but can be used with any kind of PLL based circuits. Even if the measured results shown a limitation in the tuning range of the circuit, the DFLL of chapter 5 was proven to be a very interesting circuit to be used with future CDR's.

Testing methodologies were presented to simplify the testing of the front-end circuitry and the DFLL chip. Custom-built PCBs are mandatory for good measurement results. If possible, housing the chip in a metal enclosure will reduce the interference from external noise sources.

Several research and development issues that could be pursued by others are presented in the following paragraph.

7.1 Future directions

This section presents, but is not limited to, some interesting topics that would be worth the effort to consider as future research.

1. The design of an offset current compensation circuit to be used at the input of a transimpedance amplifier. The overall circuit will therefore be ready to be used with a real photodiode. The use of a real photodiode for TIA characterization.

2. Porting the TIA and limiting amplifier design to a 0.13μm state-of-the-art CMOS process. With the help of bandwidth extension techniques (passive inductors, active inductors, new TIA topology), greater bandwidth could be achieved.

3. The design and implementation of a complete optical receiver, including a TIA, a limiting amplifier, a clock-and-data recovery circuit, and a demultiplexer.

4. Implementing the DFLL in a specifically targeted CDR application.

5. The design of the optical transmitter: the laser driver. Combined with the optical receiver, this macro could be used for different applications such a Fibre Channel or Gigabit Ethernet.

6. Porting the optical receiver design to a copper-channel medium.

CHAPTER 8

Appendix

8.1 CML gates

Current-mode logic (CML) gates is a family of high performance gates. Speed of several giga-hertz can be attained in standard CMOS technologies. In addition, CML gates provide differential outputs providing better noise immunity.

The principle behind the CML gate is the current steering in a differential pair. By changing the operating point at the input of a differential pair, the current will flow in one branch or the other. Compared to small-signal amplifiers, CML gates employ the large signal behaviour of the differential pair.

8.1.1 Inverters

The basic CML inverter is built as shown in Figure 8.1. When In+ is higher than In-, all the current from the current source I1 will flow through M_2. The voltage created at node Out+ will be R*I less than V_{dd}. Since no current flows in transistor M_1, Out- remains at V_{dd}.

The AC gain of the gate needs to be greater than one in order to regenerate the signal stream through a chain of amplifiers.

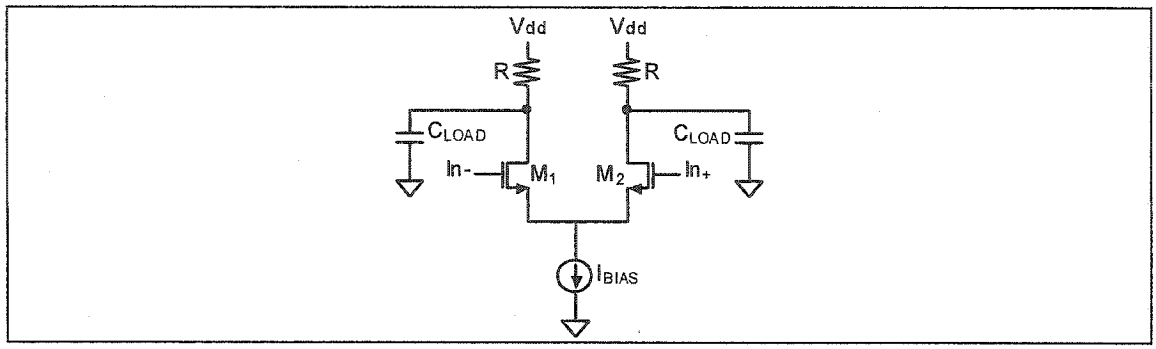


Figure 8.1 CML inverter.

The operating speed of a gate is defined by the product of the pull-up resistance and the capacitance seen at the output node. Since the circuit is differential, the analysis can be done by using the equivalent half circuit.

Optimization of a CML gate is done by changing different component sizing. Power optimization is done by either reducing the biasing current I_{bias} or by reducing the power supply voltage. Reducing the biasing current requires a larger resistance at the drain of the differential pair. This results in an obvious drawback: a larger RC constant is created. Power can also be optimized by reducing the swing at the output of the gate. Hovewer, such a reduction is bounded to some limits. If the output swing is too small, it will impair the switching of the input transistor of the next stage. The minimum power supply for an inverter can be expressed by:

$$V_{dd\_\min} = V_{DSAT}(I_{bias}) + V_{DSAT}(M_1) + V_{swing-SE}. \qquad \textbf{Eq. 8.1}$$

Equation 8.1 can further be expressed in terms of the biasing current, for any kind of gate:

$$V_{dd\_\min} = V_{DSAT}(I_{bias}) + nV_{DSAT} + RI_{bias}, \qquad \textbf{Eq. 8.2}$$

where $V_{DSAT}(I_{bias})$ is the saturation voltage of the current source, n is the number of transistor stacking to produce a logic function, I_{bias} is the biasing current and R is the pull-up resistor.

Because the rise/fall times of CML gates are governed by the RC product, such circuits are very sensitive to parasitic capacitances (i.e. input gate capacitances and interconnect capacitances). If, after layout optimization, the parasitic capacitances cannot be further reduced, a solution would be to add a buffer in series with the critical net. A buffer (which is a simple inverter) is desgined to accommodate the capacitive loading at its output. To build such a circuit, the pull-up resistor needs to be reduced. Therefore, the biasing current needs to be increased resulting, would be in an increase in power consumption.

8.1.2 Other CML gates

All the types of basic logic gates can be built in CML. Figure 8.2 (a) shows a D latch and Figure 8.2 (b) an XOR gate.

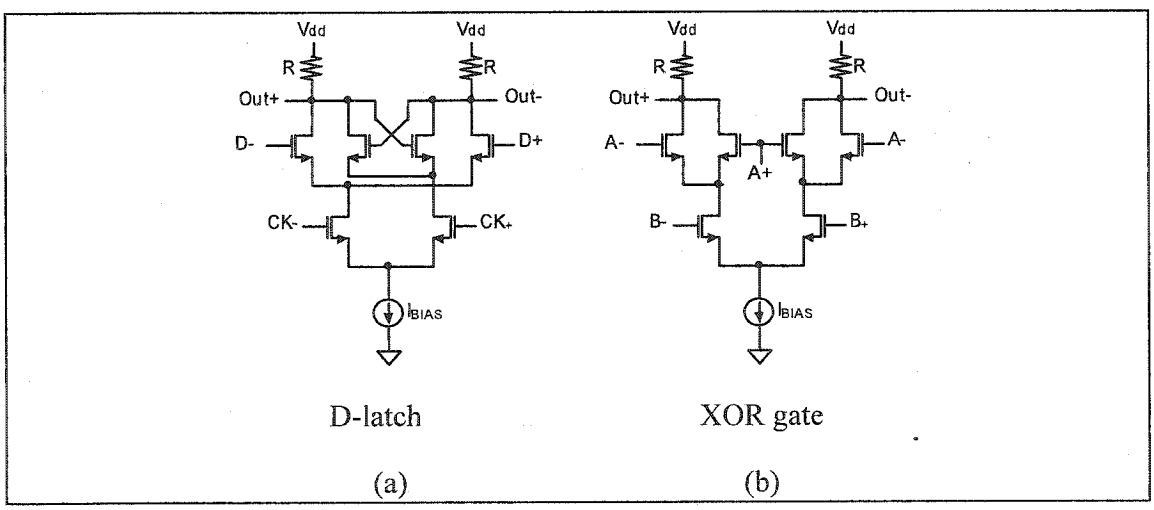


Figure 8.2 Basic CML gates.

Other types of gates can easily be built by changing the interconnections of a basic D-latch. Also, combining gates together will create more complex functions. For example, cross-connecting D-latches can form an edge-triggered D flip-flop. This type of circuit is particularly useful in clock-and-data recovery applications.

8.2 PRBS and eye diagram scripts

Below is the code used to generate a pseudo-random bit sequence usable in Cadence with the VPWLF (voltage piecewise linear read from a file).

```
================================
lfsr.m
================================

srs = 7;

% Size of shift register
r = zeros(1,srs);
r(srs) = 1; % init last bit

%for k = 1: 2^srs
for k = 1: 2^(srs) % limit length
    % Calc next bit
    next = xor(r(7), r(1));

    % Save out bit
    out = r(srs);

    %Shift
    r = [next r(1,1:srs-1)];

    % Save seq
    seq(k) = out;
end

seq = seq' ;
```

An eye diagram, such as the one of Figure 4.7 can be plotted in Matlab with the following script.

```
================================
gen_time.m
================================

function out = gen_time(seq, period, rise, A)


[rows cols] = size(seq);
out = zeros(2*rows, 2);

for k = 1:2*rows

    i = floor((k-1)/2);
```

```
    step = mod((k-1),2);

    if (step == 0)
        out(k,1) = i * period;
        out(k,2) = seq( mod(i - 1 + rows, rows) + 1 );
    else
        out(k,1) = i * period + rise;
        out(k,2) = seq( mod(i + rows, rows) + 1 );
    end
end

out(:,2) = A * (2 * out(:,2) - 1);

save temp.txt out -ASCII
```

```
==================================
eye_plot.m
==================================

function eye_plot(data, window_size, step_size, offset)


[nrows, ncols]= size(data)

figure;
hold on;

for k = 1:step_size:nrows-window_size-offset
   plot( data(1:window_size,1), data(k+offset:k+window_size-
1+offset,2));
end
```