Decision Feedback Equalizer Design for 60GHz Mobile Transceivers

Chintan Thakkar

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2012-190
http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-190.html

August 20, 2012

Decision Feedback Equalizer Design for 60GHz Mobile Transceivers

by Chintan S. Thakkar

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

Committee:

Professor Elad Alon
Research Advisor

Date

* * * * * *

Professor Ali M. Niknejad
Second Reader

Date

Contents

List of Figures

Acknowledgments

The authors would like to thank sponsors of the Berkeley Wireless Research Center (BWRC), NSF Infrastructure Grant No. 0403427, wafer fabrication donation by STMicroelectronics, the UC-Micro program and Berkeley Design Automation.

Chapter 1

Introduction

The 7GHz of bandwidth available in the 60GHz band [1] offers an opportunity to achieve multi-Gb/s wireless links. For this reason, ultra-high data-rate wireless communication over the 60GHz band has gained increasing interest from both academic [3] [4] and industrial groups [5] [6]. Commercial transceiver solutions for applications such as home HDTV streaming are now also available [7]. Since designs such as [7] are targeted at wall-powered set-top boxes, communication techniques with relatively high levels of circuit and signal processing complexity such as OFDM are often applied, which necessitates the use of data-converters and digital signal processors (DSPs) operating at multi-GS/s rates.

Fig. 1.1 shows the block diagram of a typical radio receiver. The RF front-end performs low-noise amplification followed by downconversion and (optionally) filtering, while the baseband comprises of the data-converter and DSP. The DSP performs the requisite signal-conditioning to counter the non-idealities of wireless communication, such as mitigating inter-symbol interference, performing carrier phase/frequency recovery, clock/data recovery, error-control coding/decoding etc.

Many previous works have focused on the data-converter(s) as one of the key challenges in such 60GHz links. However, recent publications [12] [13] [14] have demonstrated energy-efficient designs with figure-of-merit (FOM)[1] of 50-500 fJ/conversion step at sampling rates of 2-10GS/s for

[1]FOM = $P/(2^{ENOB} \cdot f_s)$, where P – power dissipation, $ENOB$ – effective number of bits, f_s – sample-rate.

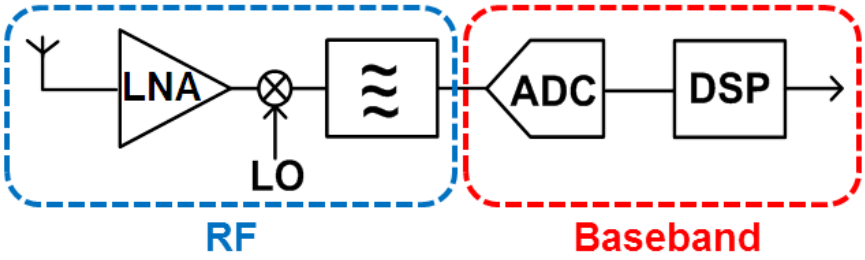


Figure 1.1: Typical implementation of a wireless receiver

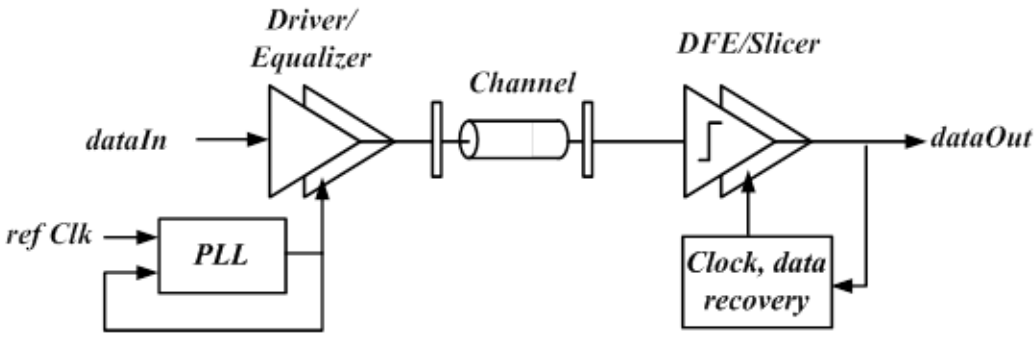


Figure 1.2: Typical implementation of a high-speed wireline transceiver

4-6 bits of dynamic range. For sampling rates near the bandwidth limits of 60GHz communication (e.g. 5GS/s) and moderate resolution (4 bits), the power consumption of these state-of-the-art ADCs would be under 50mW. In a GS/s baseband, the ADC will most likely not be the power bottleneck.

On the contrary, and especially at GS/s data-rates, the power consumption of the baseband is dominated by the various DSP blocks. This can be illustrated by the power dissipation of state-of-the-art digital designs with similar data-rates and functionality as the constituent signal processing blocks required in a 60GHz transceiver. For example, a 12-tap FIR filter in 90nm CMOS demonstrated in [15] consumes 320mW at 12GS/s. Similarly, a 10GBASE-T ethernet low-density parity-check (LDPC) decoder in 65nm CMOS in [16] running at a 6.67Gb/s throughput consumes 144mW from a scaled (0.7V) supply. The combined power of even these highly optimized exemplary DSP blocks can therefore easily be on the order of $\sim$1W [7]. As a result, DSP-based multi-GS/s designs used by OFDM-based transceivers are only suitable for wall-powered applica-

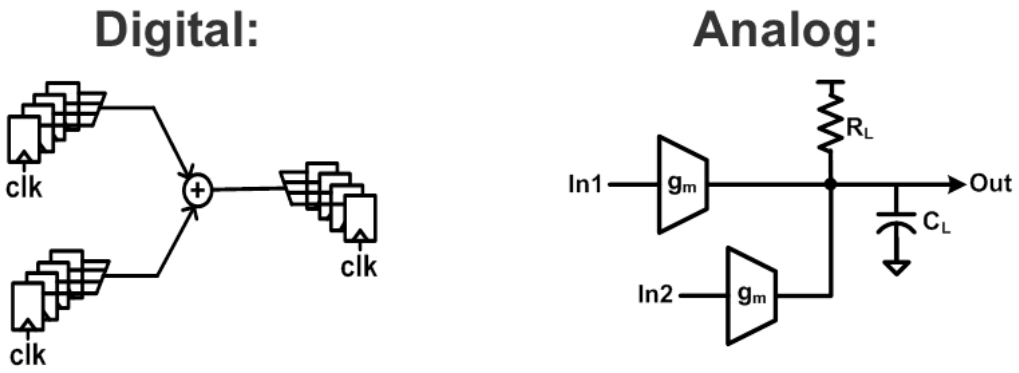


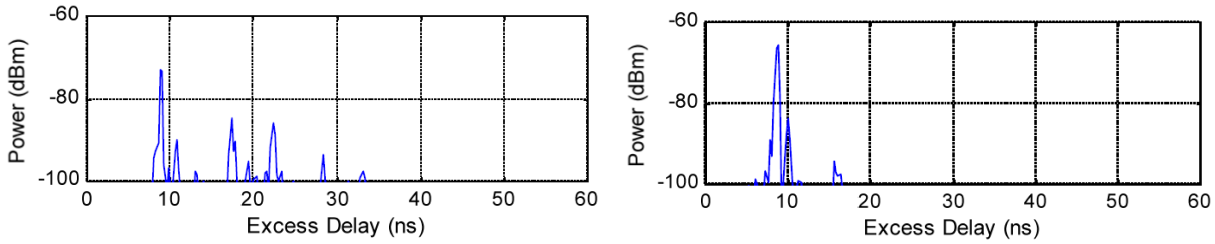
Figure 1.3: Comparison between digital and analog processing: Digital processing typically uses more capacitance per operation and hence consumes more power.

tions but will likely be infeasible for mobile/hand-held devices.

High-speed chip-to-chip serial links (Fig. 1.2) offer a stark contrast to these high dynamic range mostly-digital wireless transceiver basebands. These designs have shown that for high bandwidths and relatively low dynamic range (implying simple modulation such as 2-PAM), analog processing and a minimal number of comparators is significantly more efficient than multi-bit ADC/DSP-based solutions. The energy-efficiency achieved by current state-of-the-art serial link designs using mostly analog processing is $<$2mW/Gb/s [17] [18], which is orders of magnitude lower than the aforementioned DSP-based solutions. While these links do use digital processing, their scope is mostly limited to low-speed digital calibration/control circuitry which does not accrue a power penalty.

Fig. 1.3 intuitively illustrates the difference between digital and analog processing. For any signal processing operation, the energy per computation is set by the effective capacitance of the signal processing circuit. A digital processor typically employs multiple blocks per operation, set by the requirements on resolution. Due to minimum device size constraints on each of these constituent blocks, a digital processor will almost invariably have larger effective capacitance (dominated by timing elements such as flip-flops). On the contrary, an analog processor usually has a smaller total device size per operation (as long as it is not constrained by noise or linearity[2]) and can hence achieve lower power.

[2]In modern sub-micron technologies, up to about 5 bits of linearity can be achieved without incurring a penalty on device size.



(a) Using omnidirectional receiver (Rx) antenna

(b) Using 90^o HPBW Rx antenna

Figure 1.4: 60GHz indoor channel measurements [2]

Inspired by such low-power high-speed serial links using analog processing, in this project we aim the design of an energy-efficient mobile 60GHz baseband, with specific emphasis on time-domain equalization of line-of-sight (LOS) multi-path channels. A mixed-signal decision feedback equalizer (DFE) that operates by summing current steering DACs [3] [8] is an excellent candidate for efficiently removing post-cursor inter-symbol interference (ISI). One of the challenges of using such a DFE for the targeted symbol rates as high as 5GS/s is that a 60GHz channel may exhibit 30-50 complex[3] taps of post-cursor ISI, even with a directional front-end [2] (Fig. 1.4(b)). Previous solutions employing mixed-signal DFEs for 60GHz channels have either been low-speed ([8] operates at 500MS/s) or implement only a few taps of equalization ([3] incorporates only a 5-tap complex DFE).

These limitations can be understood by observing the structure in Fig. 1.5 which shows a conventional mixed-signal DFE. The circuit cancels post-cursor ISI by subtracting currents representing the ISI taps at a resistive load. The ISI cancelation currents are implemented by using current steering DACs, whose magnitude and sign represent the ISI magnitude and direction respectively. As will be described in detail later, the summing node of this structure is loaded by the parasitic capacitance of the current-steering switches. Therefore, in typical DFE implementations, only a relatively limited number of taps can be implemented before this self-loading makes it infeasible (at any power) to achieve the bandwidth required for multi-GS/s operation.

[3]Complex refers to the presence of both in-phase and quadrature-phase components. Therefore, in the context of an I/Q baseband, each complex tap consists of (a) a direct tap from I-to-I or Q-to-Q channel and (b) a cross-tap from I-to-Q or Q-to-I channels.

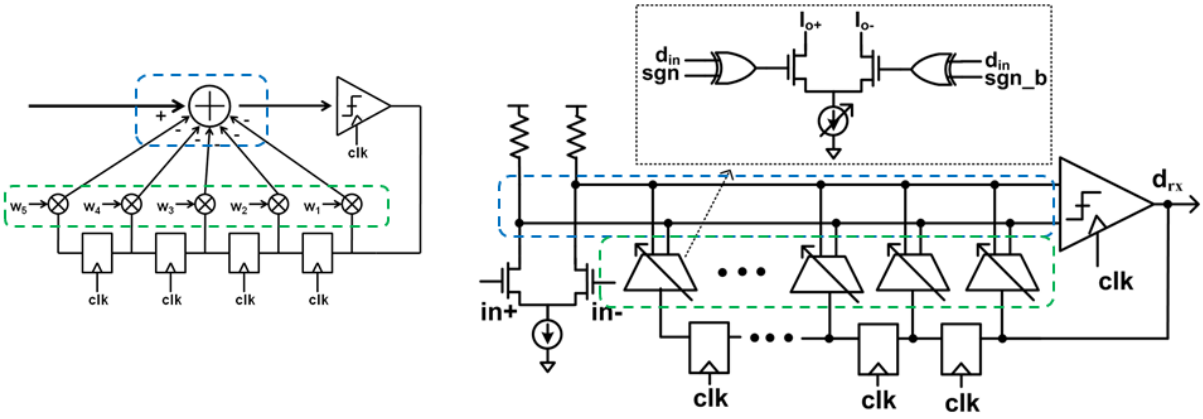


Figure 1.5: Block Diagram of a DFE (left); Conventional mixed-signal implementation (right)

This work proposes a cascode-summation structure that significantly increases the number of ISI taps that can be efficiently canceled by the DFE, while enabling 10Gb/s quadrature phase shift keying (QPSK) communication. The topology leverages the fact that in any channel, the total multi-path amplitude (and energy) is bounded. To demonstrate this approach, a 65nm CMOS test-chip was designed that included a mixed-signal DFE capable of handling 20 complex ISI taps at 10Gb/s while consuming only 14mW of power.

This report first presents the design methodology for a mixed-signal DFE. The technique can be used to explore the design space and compute the power dissipation of a DFE as a function of data-rate and the number of taps (i.e. the channel characteristics). The methodology is also used to highlight the shortcomings of the conventional DFE summing structure, which motivates the proposed cascode current-summing structure to increase the number of feasible taps. The design framework is then extended to exhibit these improvements, followed by the key design challenges of implementing one such prototype 20-complex cascode-summation structure. Finally, we conclude with a discussion of the test-chip measurement results.

Chapter 2

Mixed-Signal Decision Feedback Equalizers

Prior to designing an energy-efficient DFE suitable for a 60GHz wireless channel, we will first illustrate the design methodology of a conventional mixed-signal DFE. In particular, it is important to understand the dependence of power dissipation on the data-rate, the ISI profile/number of taps in the DFE, as well as technology-related parameters. This dependence illustrates the limitations of the conventional current summation, and motivates the improvements brought about by the proposed cascode current summation.

Fig. 2.1 shows a conventional mixed-signal DFE summing amplifier structure, and its small-signal model. To simplify the analysis, the output resistance of all transistors has been ignored. The input cursor amplitude is V_{in}[1], while the summing amplifier has a DC gain of G_{dfe} and operates at a data-rate of f_s symbols per second. The DFE has N_{taps} taps, each of which can cancel ISI up to a maximum amplitude of k times the cursor amplitude[2]. Therefore, the maximum tap current, I_{tap} is

$$I_{tap} = k \cdot g_{m,cursor} \cdot V_{in} \tag{2.1}$$

In order to properly cancel the ISI, the tap-DAC current (I_{tap}) must be steered by the differential data signal (shown as d and $\overline{db}$) from feedback shift register and satisfactorily settled at the

[1]Cursor amplitude V_{in} does not include the ISI.

[2]Due to multi-path interference, a wireless channel can in general have ISI of up to a certain amplitude for all tap positions.

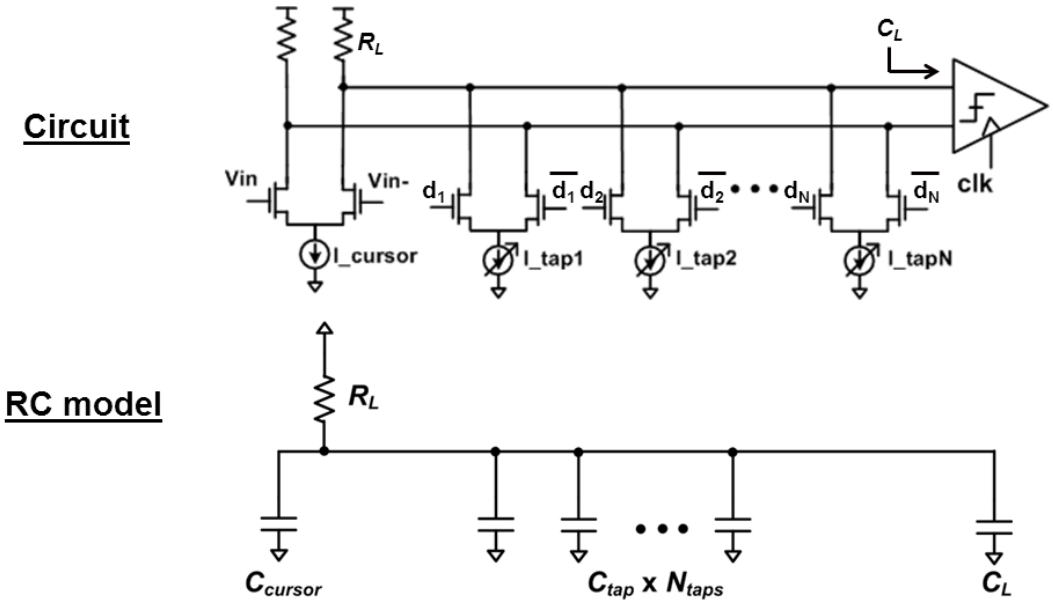


Figure 2.1: Gain/bandwidth analysis of a conventional mixed-signal DFE

summation node before the next data bit is resolved by the comparator. A full-rate[3] DFE therefore has a 1UI timing constraint for the settling of each tap. This timing constraint can be partitioned as $(1-\alpha)$UI for the digital delay of the flip-flop and the XOR gate (for choosing the sign of the tap), and αUI for the analog settling of the tap current at the summation node ($\alpha < 1$). The time-constant τ of this settling is given by the RC product

$$\tau = R_L \cdot (C_L + C_{cursor} + N_{taps} \cdot C_{tap}) \tag{2.2}$$

where R_L is the summation load resistance, C_{cursor} is the drain capacitance of the input transistor, C_{tap} is the drain capacitance of the current steering switch at each tap, and C_L is the loading from the next stage, which is typically the comparator (often with a preamp input-stage to mitigate kickback). Since the input pair is a gm-stage, $g_{m,cursor} = I_{cursor}/V^*$, where I_{cursor} is the DC bias current of the input pair[4]. The DC gain is therefore given by

$$G_{dfe} = g_{m,cursor} \cdot R_L = \frac{I_{cursor}}{V^*} \cdot R_L \tag{2.3}$$

Of the three capacitors at the summation node (2.2), C_{cursor} and C_{tap} are attributed to the internal self-loading of the structure and are functions of the summing amplifier currents, while

[3]Full-rate implies that the sampling CLK frequency is equal to the data-rate.

[4]V^* is defined as $V^* = 2I_{bias}/g_m$.

C_L is the fixed external capacitive loading. The internal capacitors can be expressed in terms of technology parameters, C_dI_{cursor} and C_dI_{tap}, where C_dI denotes transistor drain capacitance per unit drain current.

$$C_{cursor} = C_dI_{cursor} \cdot \frac{I_{cursor}}{2} \tag{2.4}$$

$$C_{tap} = C_dI_{tap} \cdot I_{tap} = C_dI_{tap} \cdot (k \cdot g_{m,cursor} \cdot V_{in}) = C_dI_{tap} \cdot \left(k \cdot \frac{I_{cursor}}{V^*} \cdot V_{in}\right) \tag{2.5}$$

Substituting for R_L, C_{cursor}, and C_{tap} in terms of I_{cursor} (from 2.3, 2.4 and 2.5 respectively), the time constant in (2.2) can be expressed as

$$\tau = \frac{G_{dfe} \cdot V^*}{I_{cursor}} \cdot \left(C_L + C_dI_{cursor} \cdot I_{cursor} + N_{taps} \cdot C_dI_{tap} \cdot k \cdot \frac{I_{cursor}}{V^*} \cdot V_{in}\right) \tag{2.6}$$

If 1UI is $T = 1/f_s$, then the analog settling constraint implies that

$$n_\tau \cdot \tau = \alpha \cdot T \tag{2.7}$$

where n_τ is the required number of time constants of settling. Combining (2.6) and (2.7) gives a complete expression for the cursor current:

$$I_{cursor} = \frac{C_L \cdot \left(\frac{n_\tau f_s}{\alpha}\right) \cdot G_{dfe} \cdot V^*}{1 - \left(\frac{n_\tau f_s}{\alpha}\right) \cdot G_{dfe} \cdot V^* \cdot \frac{C_dI_{cursor}}{2} \cdot \left(1 + N_{taps} \cdot k \cdot \frac{V_{in}}{V^*} \cdot \frac{2C_dI_{taps}}{C_dI_{cursor}}\right)} \tag{2.8}$$

Since the tap currents (I_{tap}) are proportional to the cursor current (I_{cursor}), the total power dissipation of the summing amplifier is also proportional to I_{cursor}. Therefore, power dissipation of a conventional DFE, P_{conv} is of the form

$$P_{conv} \propto \frac{I_{nom}}{1 - \frac{GBW}{\omega_T} \cdot \gamma \cdot \left(1 + N_{taps} \cdot k \cdot \frac{V_{in}}{V^*} \cdot \frac{2C_dI_{taps}}{C_dI_{cursor}}\right)} \tag{2.9}$$

where γ is the ratio of drain to gate capacitance, $GBW = G_{dfe} \cdot \left(\frac{n_\tau f_s}{\alpha}\right)$ is the gain-bandwidth product, and $I_{nom} = C_L \cdot GBW \cdot V^*$ is the nominal current consumption of a class-A amplifier without self-loading.

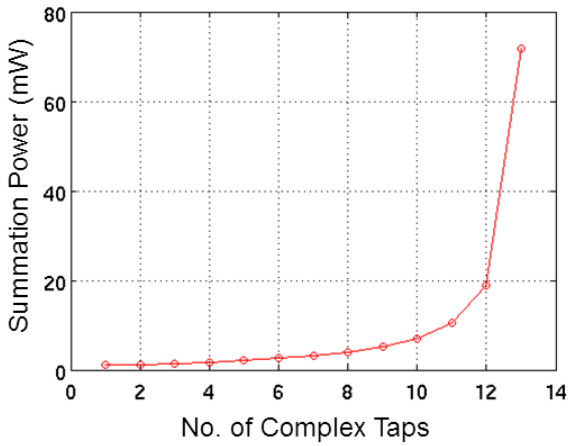


Figure 2.2: Conventional current-summing DFE: Power vs. no. of complex taps for 10Gb/s QPSK. Each tap can cancel ISI upto half the cursor amplitude

The form of equation (2.9) illustrates how a conventional mixed-signal DFE can only support a limited number of taps. When N_{taps} is small, to handle the extra capacitance of every additional tap, the load resistance can be moderately decreased and the current increased in order to maintain a constant gain and bandwidth. However, once the product of GBW and N_{taps} becomes comparable to the ω_T of the technology, the DFE becomes self-loaded to the point that it cannot handle more taps for any increase in power, as seen in Fig. 2.2. At the desired data-rate of 5GS/s in a 65nm CMOS technology, a conventional DFE structure can implement only about 10 complex taps efficiently. Clearly, such a structure is incapable of being directly used in a 60GHz transceiver, which typically needs almost 30-50 taps of equalization (Fig. 1).

Since the channel to be equalized is typically unknown ahead of time, the DFE needs to incorporate a certain amount of reconfigurability into each tap. Implementing such flexibility invariably involves an overdesign of the taps in terms of their current-handling capability, which exacerbates the self-loading of a conventional summing structure. As will be shown in the next chapter, cascode current-summation alleviates the penalty associated with this flexibility and is able to significantly extend the number of feasible taps.

Chapter 3

Cascode Current-Summing Decision Feedback Equalizer

The previous chapter highlighted the shortcomings of a conventional DFE. The DFE structure is primarily constrained by self-loading of its taps. In addition, since the channel to be equalized is not fixed, the DFE requires a certain degree of flexibility, which limits the number of taps that can be implemented. In this chapter, we will show that by making key observations about the channel, a cascode current-summing structure is able to incorporate the requisite flexibility while notably improving the number of feasible taps.

3.1 Concept

Since a wireless channel is time-varying by nature, each tap needs to be designed to cancel a certain maximum magnitude of ISI. From a design standpoint, this sets the size of the current steering switch of each tap to handle this maximum ISI current. If the capacitive loading of each tap handling the maximum ISI current is C_{par}, the total loading from N taps is $N \cdot C_{par}$. However, since the received signal has a limited sum of ISI magnitude (due to finite transmit power), not all

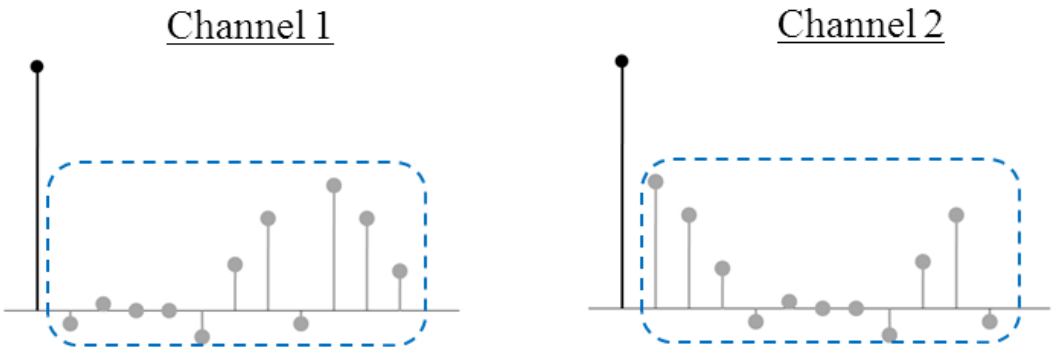


Figure 3.1: Wireless channel response: While each tap can have a variable weight, not all taps will be at their maximum weight altogether. However, the sum of all tap magnitudes is bounded due to finite transmit power.

taps need to be set to their maximum magnitude at the same time (Fig. 3.1). In other words,

$$\|ISI\|_1 < \sum_{i=1}^{N_{taps}} |ISI_{i,max}| \tag{3.1}$$

which in turn means that if the maximum current in each tap is I_{max} and the maximum possible sum of currents in all taps is $I_{ISI,max}$, then $I_{ISI,max} < N_{taps} \cdot I_{max}$. Therefore, loading the summation node with a capacitance of $N \cdot C_{par}$ is inefficient.

Conceptually, the ideal design would be one in which the taps only load the summation node with capacitance corresponding to the maximum possible sum of ISI. One way to realize this would be by using a fully digital FIR filter in the feedback path to sum all taps and cancel the ISI using a DAC, as shown in Fig. 3.2. Since the DAC current would be bounded, the DAC can be designed with bounded capacitive loading at the summation node, thus reducing the power of the summation amplifier. However, since the latency of this FIR/DAC needs to be $<$ 1UI, at GS/s rates the power consumed by the filter in summing such a large number of digital tap values would be unacceptably high.

The proposed cascode current summation structure realizes limited capacitive loading by summing all current through a cascode transistor, as shown in Fig. 3.3. The cascode transistor can be sized to handle the bounded ISI current. The cascode transistor width is therefore much smaller than the sum of current steering switch widths, thus reducing the loading at the output of the summing amplifier. Furthermore, the large capacitance of these switches is moved to the low-

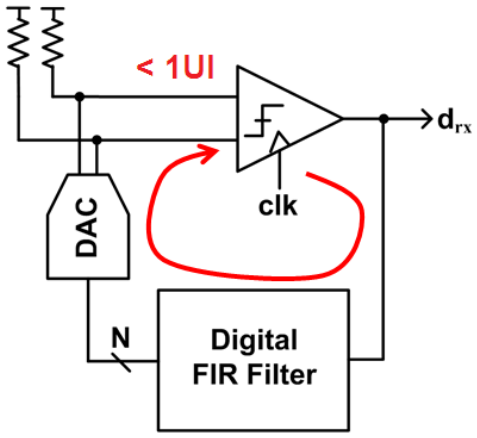


Figure 3.2: A fully digital FIR/DAC implementation of the DFE limits the self-loading at the summing node. However, the 1UI latency constraint on the first tap feedback makes the FIR adder unacceptably expensive in power

impedance source node of the cascode. To maintain a high bandwidth at the cascode source, the g_m of the cascode transistor is increased by applying additional common mode current. As will be shown in the analysis that follows, this structure significantly extends the number of taps that can be implemented by the DFE.

3.2 Analysis

Fig. 3.4 shows the small-signal equivalent of the cascode current-summing structure. As with the analysis of the conventional DFE, the input cursor amplitude is V_{in}, the DC-gain is G_{dfe}, the data-rate is f_s symbols per second, and the number of taps is N_{taps}. The total cursor DC biasing current is I_{cursor}, while the common-mode current added to each side is $I_{bleeder}$. Each of the taps can cancel ISI up to a maximum amplitude of k times the cursor, and a total ISI of amplitude of ISI_{max} times the cursor amplitude. Therefore,

$$I_{tap} = k \cdot g_{m,cursor} \cdot V_{in} = k \cdot \frac{I_{cursor}}{V^*} \cdot V_{in} \tag{3.2}$$

$$I_{taps,total} = ISI_{max} \cdot \frac{I_{cursor}}{V^*} \cdot V_{in} \tag{3.3}$$

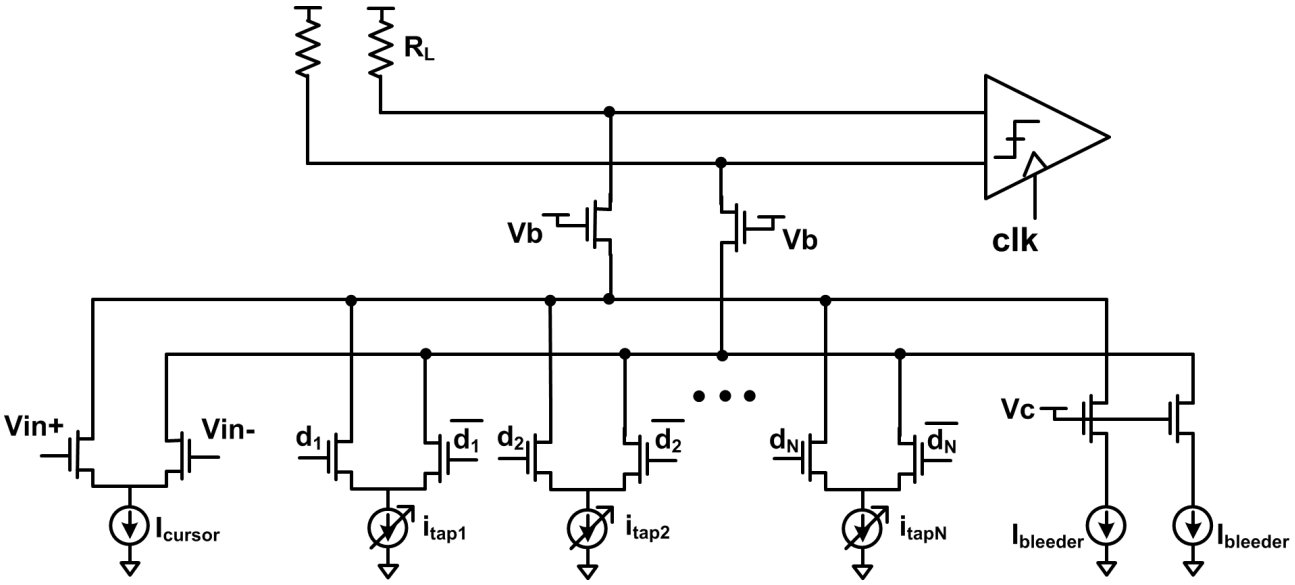


Figure 3.3: Cascode current summing structure: Tap switches moved to low-impedance cascode, and capacitive load at output node reduced

The cascode-based structure has two poles – one at the cascode source where all taps are summed ($\omega_{p,1}$) and another at the output node ($\omega_{p,2}$). It is desirable to place both the poles at approximately the same frequency, since making one pole larger than the other has diminishing returns for effective bandwidth of the summing amplifier in terms of power dissipation. However, the presence of two poles implies that in order to get the same bandwidth as that of the conventional summing structure, both the poles should be $\sqrt{2}$ times larger. Therefore, the αUI analog settling constraint leads to

$$\sqrt{2} \cdot n_\tau \cdot \tau = \alpha \cdot T \tag{3.4}$$

The RC time-constants of the two poles are approximately:

$$\tau_1 = R_1 \cdot C_1 = \frac{1}{g_{m,casc}} \cdot (N_{taps} \cdot C_{tap} + C_{casc,src} + C_{cursor} + C_{bleeder}) \tag{3.5}$$

$$\tau_2 = R_2 \cdot C_2 = R_L \cdot (C_L + C_{casc,drain}) \tag{3.6}$$

where $g_{m,casc}$ is the g_m of the cascode transistor, C_{tap} is the current steering switch drain capacitance, $C_{casc,src}$ is the cascode source capacitance, C_{cursor} is the input transistor drain capacitance, $C_{bleeder}$ is the bleeder transistor drain capacitance, R_L is the summation load resistance, C_L is the

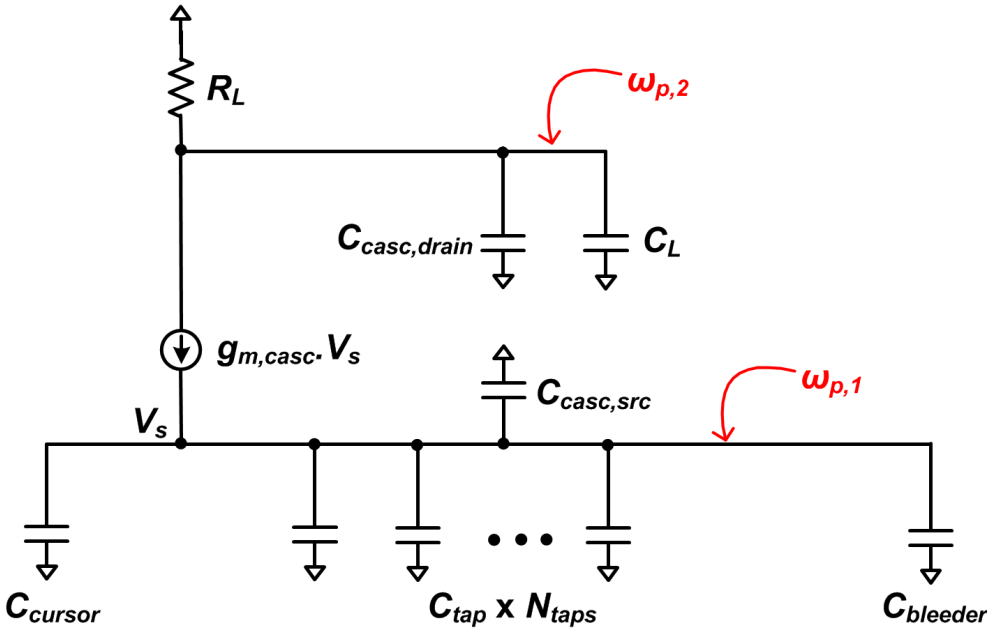


Figure 3.4: Small-signal model of a cascode current-summing DFE

loading from the next stage (i.e. the preamp/comparator), and $C_{casc,drain}$ is the cascode transistor drain capacitance. R_L is set by the DC gain requirements of the DFE, and given by

$$R_L = G_{dfe} \cdot \frac{V^*}{I_{cursor}} \tag{3.7}$$

The g_m of the cascode can be calculated as

$$g_{m,casc} = \frac{2 \cdot \left(\frac{I_{cursor}}{2} + I_{bleeder} - \frac{V_{in}}{2} \cdot \frac{I_{cursor}}{V^*} \right)}{V^*} \tag{3.8}$$

The negative term in the numerator accounts for the cursor small-signal current fully steered one way, which gives the worst-case g_m of the cascode.

The capacitors internal to the structure can be expressed in terms of technology parameters, C_dI_{cursor}, C_dI_{tap}, $C_dI_{bleeder}$, C_dI_{casc}, C_sI_{casc} where C_dI and C_sI respectively denote drain and source capacitance per unit current, and the subscripts refer to cursor, tap switch, bleeder and cascode transistors respectively.

$$C_{cursor} = C_dI_{cursor} \cdot \frac{I_{cursor}}{2} \tag{3.9}$$

$$C_{tap} = C_dI_{tap} \cdot \left(k \cdot \frac{I_{cursor}}{V^*} \cdot V_{in} \right) \tag{3.10}$$

$$C_{bleeder} = C_d I_{bleeder} \cdot I_{bleeder} \tag{3.11}$$

$$C_{casc,drain} = C_d I_{casc} \cdot I_{casc} = C_d I_{casc} \cdot \left(\frac{I_{cursor}}{2} + I_{bleeder} + ISI_{max} \cdot \frac{I_{cursor}}{V^*} \cdot V_{in} \right) \tag{3.12}$$

$$C_{casc,src} = C_s I_{casc} \cdot I_{casc} = C_s I_{casc} \cdot \left(\frac{I_{cursor}}{2} + I_{bleeder} + ISI_{max} \cdot \frac{I_{cursor}}{V^*} \cdot V_{in} \right) \tag{3.13}$$

Since all currents scale with proportionally with the cursor bias current, the bleeder current can be computed as $I_{bleeder} = p \cdot I_{cursor}$, where

$$p = \frac{k \cdot V_{in} \cdot C_d I_{taps} \cdot N_{taps} + \frac{V^*}{2} \cdot (C_s I_{casc} + C_d I_{cursor}) + ISI_{max} \cdot V_{in} \cdot C_s I_{casc} - \frac{\alpha \cdot (1 - V_{in}/V^*)}{2\sqrt{2} \cdot n_\tau \cdot f_s}}{\frac{\alpha}{\sqrt{2} \cdot n_\tau \cdot f_s} - V^* \cdot (C_d I_{bleeder} + C_s I_{casc})} \tag{3.14}$$

Simplifying the above equations gives a complete expression for the cursor current:

$$I_{cursor} = \frac{C_L \left(\frac{\sqrt{2} n_\tau f_s}{\alpha} \right) G_{dfe} V^*}{1 - \left(\frac{\sqrt{2} n_\tau f_s}{\alpha} \right) G_{dfe} V^* \frac{C_d I_{cursor}}{2} \left(U + N_{taps} k \frac{V_{in}}{V^*} \frac{2 C_d I_{taps}}{C_d I_{cursor}} \frac{C_d I_{casc}}{\frac{\alpha}{\sqrt{2} f_s n_\tau V^*} - (C_s I_{casc} + C_d I_{bleeder})} \right)} \tag{3.15}$$

where

$$U = 1 + 2 \cdot ISI_{max} \frac{V_{in}}{V^*} + \frac{V^* (C_s I_{casc} + C_d I_{cursor}) + 2 \cdot ISI_{max} V_{in} C_s I_{casc} - \frac{2\alpha (1 - V_{in}/V^*)}{2\sqrt{2} f_s n_\tau}}{\frac{\alpha}{\sqrt{2} f_s n_\tau} - V^* (C_d I_{bleeder} + C_s I_{casc})} \tag{3.16}$$

Adding the total tap current from (3.3) and $I_{bleeder} = p \cdot I_{cursor}$, the total power consumption, P_{casc} can be expressed (after some approximation) as:

$$P_{casc} \propto \frac{I_{nom}}{1 - \frac{GBW}{\omega_T} \cdot \gamma \cdot \left\{ U + N_{taps} \cdot k \cdot \frac{V_{in}}{V^*} \cdot \frac{2 C_d I_{taps}}{C_d I_{cursor}} \left(\frac{2\gamma}{G_{dfe}} \cdot \frac{GBW}{\omega_T} \right) \right\}} \cdot \left(1 + 2p + ISI_{max} \frac{V_{in}}{V^*} \right) \tag{3.17}$$

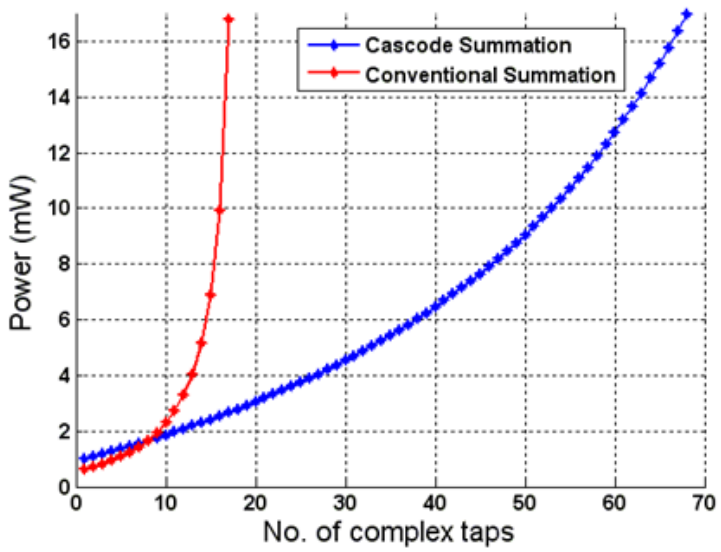


Figure 3.5: Summing amplifier power vs. no. of complex taps for conventional and cascode-summing structures (10Gb/s QPSK)

Similar to the power consumption of the conventional DFE in (2.9), $GBW = G_{dfe} \cdot \left(\frac{\sqrt{2}n_\tau f_s}{\alpha}\right)$ is the gain-bandwidth product, $I_{nom} = C_L \cdot GBW \cdot V^*$ is the nominal current consumption of a class-A amplifier without self-loading, and γ is the ratio of drain to gate capacitance. Compared to the power consumption of a conventional current summing structure, the self-loading term for cascode current summing increases $\frac{G_{dfe}}{2\gamma} \cdot \frac{\omega_T}{GBW}$ times slower with N_{taps}. Intuitively, this benefit is proportional to the ratio $\frac{\omega_T}{\omega_{p,1}}$ ($\omega_{p,1}$ being the cascode bandwidth) since the cascode source bandwidth without external loading from the taps is nominally ω_T.

Due to the decreased rate of increase in self-loading with number of taps, the cascode summing structure will hit the self-loading limit much later than the conventional summing structure, as illustrated by Figure 3.5. It may be noticed that for a small number of taps, the cascode structure actually consumes more power than the conventional structure due to the presence of two higher frequency poles and the higher (1.2V) supply required to accommodate the cascode. Beyond 10 complex taps, however, it is clearly beneficial to use the cascode structure.

3.3 Key Design Issues

While the cascode current summing structure can significantly extend the number of post-cursor ISI cancelation taps, key design issues must addressed in order to implement data-rates of 5GS/s at a BER of $< 10^{-12}$.

3.3.1 Tap-DAC Resolution

The BER of the received data-bits is set by the ratio of the received signal power to the noise of the constituent receiver blocks (input-referred to the comparator). The upper bound on this noise power sets requisite low-noise constraints on both RF and baseband gain blocks in front of the comparator. Since thermal noise is typically the primary constraint, it is necessary that the contribution of all other noise components be much smaller. Implementing as many as 20 complex (i.e. a total of 40) taps in the DFE can potentially accrue sizeable quantization noise. To ensure that the quantization noise of 40 taps is less than the thermal noise, each tap DAC requires 7 bits of resolution[1]. Due to matching limitations, such a high resolution necessitates a large DAC size of more than 50μm x 50μm.

Furthermore, a compact layout of the DFE core to enable minimum loading on the high-speed timing paths requires that these large DACs be physically located 100s of μms of distance away[2] (as shown later in Fig. 3.15). As a result, there exists a large capacitance at the drain of each DAC (which is also the tail-node of the tap switching pair), thus creating a relatively low frequency pole (typically at a hundred MHz). As will be discussed later in the circuit design section, this low frequency pole necessitates the use of low-swing drivers for the current steering switches of the taps.

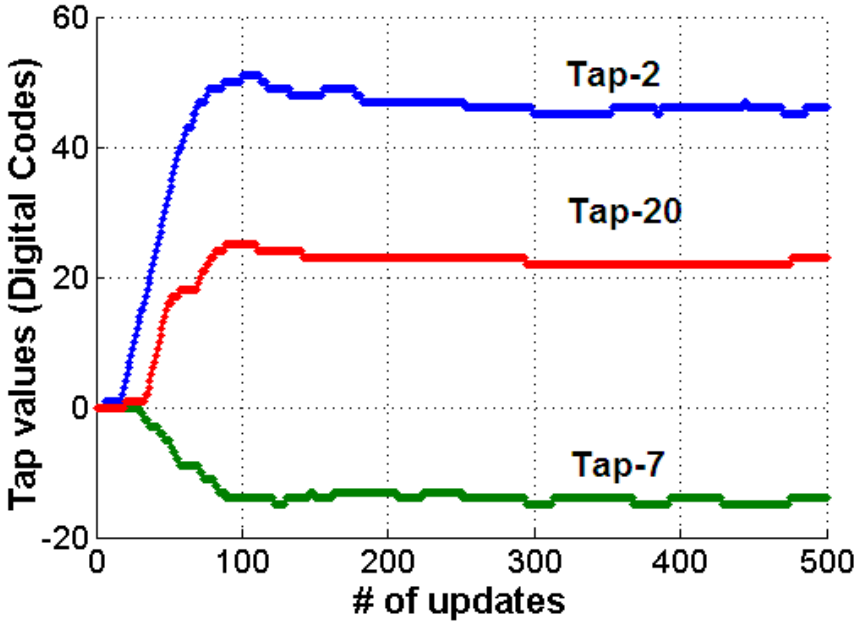


Figure 3.6: DFE tap value adaptation and steady-state dithering for I-to-I taps 2, 7, and 20

3.3.2 Effect of Dithering on DAC Resolution

Since the interference profile of a wireless channel is time-varying by nature, the DFE taps need to be continuously adapted to be able to track these variations. Once a tap is 'locked', the digital code of its DAC invariably dithers between at least two adjacent values (as shown in Fig. 3.6). If the LSB of each DAC is Δ, then the quantization noise power associated with the closest digital code is $\frac{\Delta^2}{12}$, and with the second closest digital code is $\frac{7 \cdot \Delta^2}{12}$. Assuming that the dithering is uniformly distributed between the two digital values, the average quantization noise is $\frac{4 \cdot \Delta^2}{12}$, which is twice as large (in voltage) as compared to always selecting the closest digital code. It must be noted that even if the adaptation is frozen at some particular setting, the expected value of quantization noise is still $\frac{4 \cdot \Delta^2}{12}$. This loss in resolution was taken into account while determining the 7-bit resolution requirement on the tap DACs.

[1]Computed for a signal amplitude of 120mV (differential, peak-to-peak) at the comparator input, and BER $< 10^{-12}$. The signal amplitude was set by a typical 60GHz link budget for 3-5m distances with a moderately directional front-end.

[2]It should be noted that the DACs themselves are outside the high-speed feedback path.

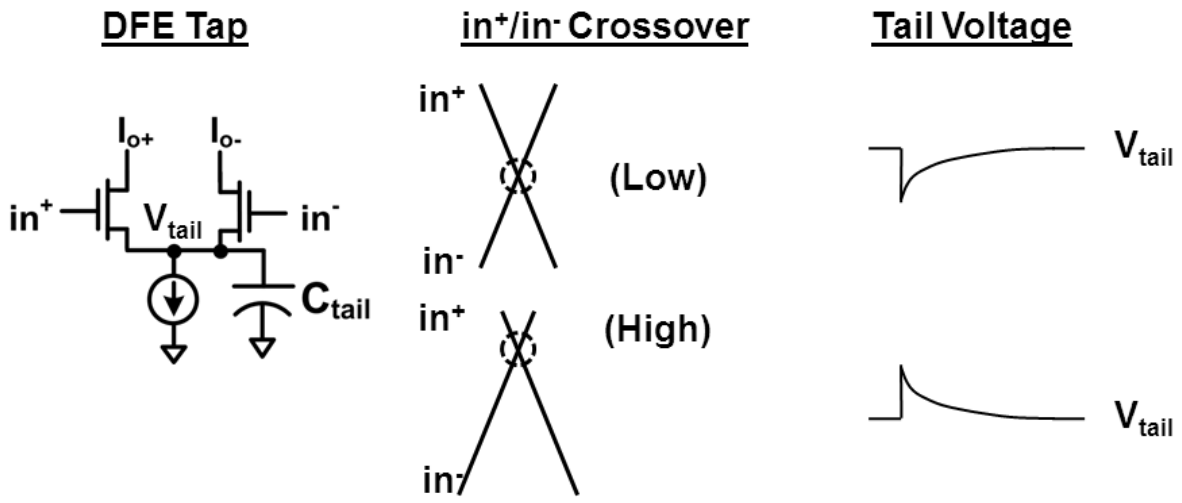


Figure 3.7: Infinite impulse response (IIR) effects due to tap switching. The crossover points of in+/in- are exaggerated to highlight the imbalance in voltage.

3.3.3 IIR Effects

The current-steering tap switches are driven differentially by the low-swing driver. However, there invariably exists an imbalance in the differential driving of the gates of these switches for steering the current to one side as compared to the other, which causes glitches at the tail node. As seen in Fig. 3.7, if the crossover point of the differential signals is too high (low), the the tail node glitches high (low). The glitches eventually settle to the equilibrium value which is effectively the average tail voltage. As discussed earlier, since the tail node (by account of its large capacitance) is relatively slowly settling as compared to the symbol period, the glitch settles over multiple symbol periods as an infinite impulse response (IIR). The IIR effect is more pronounced over long runs of the same data-bit when the tap current drifts and reduces the effective height of the data-signal.

Since the tail node glitches are equal in magnitude about the equilibrium position but opposite in sign for the two directions of current steering, the IIR phenomenon can be mathematically expressed as a convolution of the taps with $\epsilon \cdot (1 - z^{-1})$ on the feedback path, where ϵ is the relative magnitude of the glitching error current with respect to the steady-state tap current. If the DFE taps were matched exactly to the ISI profile of the channel (referred to as the "True channel" on the left of Fig. 3.8), this convolution would cause a spurious component in the DFE feedback (referred to as the "Bad" DFE in the same figure). Fortunately, if the taps are continuously adapted,

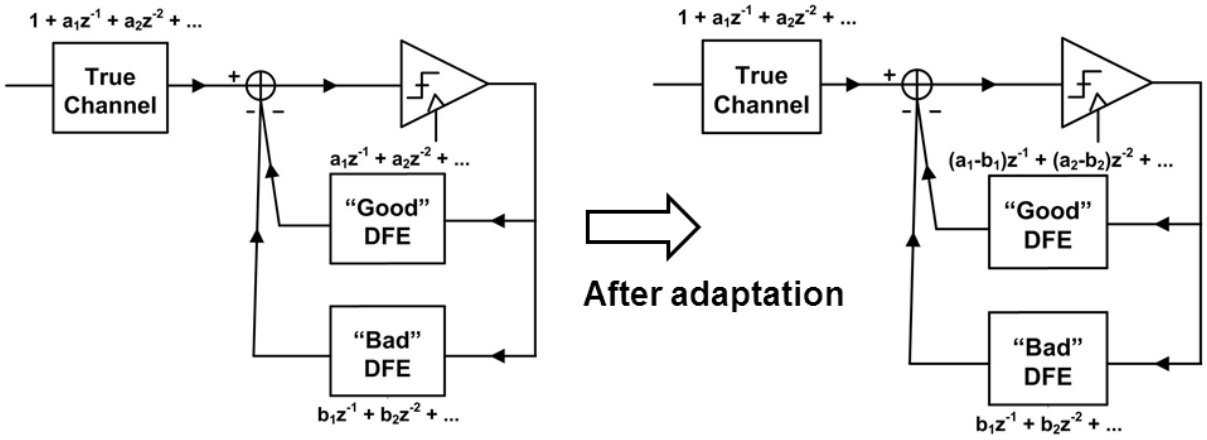


Figure 3.8: Modeling IIR effects: (left) before and (right) after correction by adaptation

these undesired components are absorbed into the taps (as shown on the right of Fig. 3.8). The corrections themselves recursively produce additional IIR effects which eventually decay below the LSB of the tap DACs. The recursion does however mean that a DFE requires a few more taps than the channel to simply correct for its own IIR profile.

To mitigate the impact of IIR effects, the tail node must either settle quickly with respect to a UI, or stay constant over a long run of the same data-bit. It is therefore desirable to make the tail node bandwidth either (a) very high, so that the tail settles within a symbol period, thus avoiding the error altogether, or (b) very low, so that the tail node stays remains unchanged after glitching. In this case, the glitch shows up as an offset when input referred to the comparator, and can be absorbed by the offset-cancellation circuitry.

An additional effect of the switching pattern of the taps is that it effectively causes a non-linear distortion in the output impedance of the summing amplifier. The implication of this time-varying impedance is that currents from both the cursor and the taps should always be summed at the same node so that they see an identical impedance (which makes the changes in this impedance irrelevant).

3.4 Critical Circuit Blocks

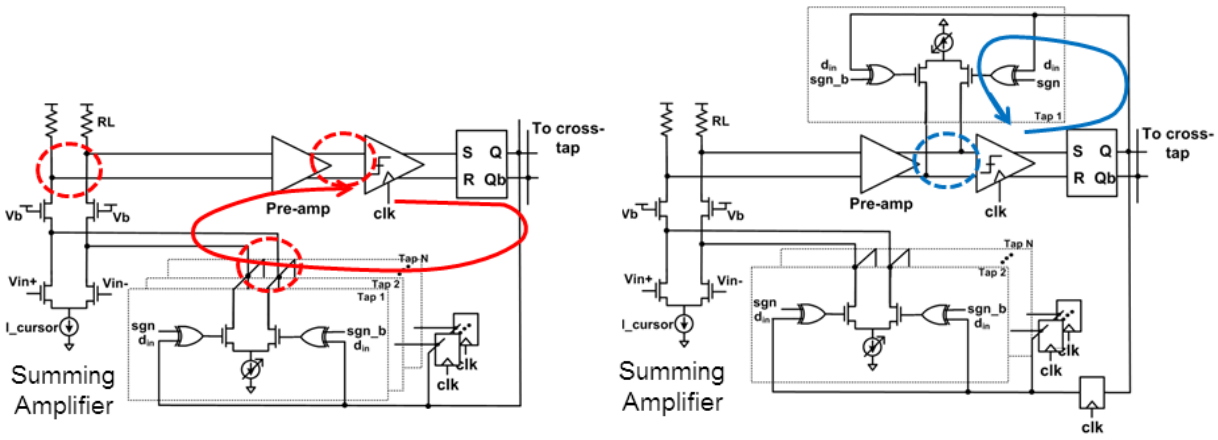
Following the design issues addressed in the previous section, this section discusses the critical circuit design components, namely (a) the high-speed timing paths, and (b) the low-swing drivers of the current-steering tap switches.

3.4.1 High-Speed Timing Paths

Since the comparator must sample the input, resolve its value, and then subtract a signal proportional to that value from the input - all within in one symbol time (200ps at 5GS/s) - the first post-cursor tap of the DFE's feedback filter is typically the most difficult to implement. Loop unrolling [19] has been shown to relax this tight timing constraint by making multiple decisions each cycle, and reducing the critical path to a digital multiplexer (MUX) delay. However, the disadvantage of loop unrolling is that is it exponentially increases the number of comparators as a function of the number of taps unrolled. For a complex DFE, unrolling one complex tap necessitates the use of four comparators [3], which increases both sampler power dissipation and loading at the preceding summing amplifier. Furthermore, loop unrolling increases the complexity of clock and data recovery (CDR) due to the need for filtering edge updates [10] (which also reduces the CDR bandwidth).

Fig. 3.9(a) shows the critical timing path for the first tap in a cascode current summing DFE, which involves settling through three poles of the summing amplifier and the preamp in addition to the comparator resolution delay. Without the use of unrolling, satisfying the timing constraint for the first tap would necessitate a significant increase in the preamp and summing amplifier bandwidth, increasing power dissipation sharply (as predicted by equation 2.8). Therefore, to efficiently relax the timing constraint, the first tap can be directly summed at the preamp, bypassing summing amplifier altogether (similar to [9]). Using this technique, the analog portion of the settling delay for tap-1 involves only a single pole (as shown in Fig. 3.9(b)) of the preamp (which typically has much higher bandwidth than the summing amplifier[3]).

[3]The preamp's primary function is to isolate the summing amplifier from the comparator kickback. Since the preamp should be able to settle this kickback within a very small fraction of the UI, the preamp should have a much higher bandwidth than the summing amplifier.



(a) All taps summed together - slow settling

(b) Tap-1 summed at comparator input - fast settling

Figure 3.9: Tap-1 feedback

Summing the first direct feedback and cross taps at the comparator input adds to the self-loading of the preamp structure. However, the power overhead is small since the self-loading at the preamp is primarily dominated by the offset cancelation current switches. In addition to local summing at the preamp, the timing overhead of the first tap is further reduced by implementing the sign selection (XOR) in domino logic.

3.4.2 Low-Swing Drivers

As explained in the previous section, the tail node of the current-steering pair of each tap tends to have a lower bandwidth as compared the data-rate of the DFE. All of these slow-moving tail nodes are only isolated from the amplifier output voltage variations by the output impedance of the cascode transistor and the tap switches. The low intrinsic gain and output impedance of transistors in sub-micron technologies therefore requires that both the cascode and the tap switches be in saturation to ensure sufficient isolation. These headroom requirements necessitate the use of low-swing XOR-drivers for the tap switches.

Fig. 3.10 shows the design for these drivers which operates from a 0.6V supply. The driver

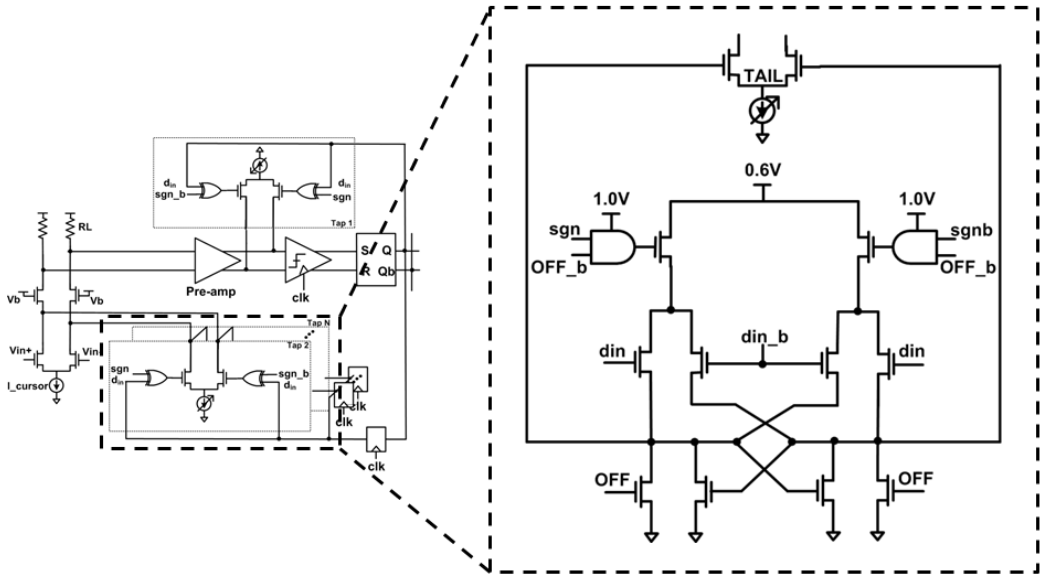


Figure 3.10: Low-swing drivers with embedded XOR for current steering switches

inputs are full-swing (1V) differential digital signals (din and din_b) from the feedback shift registers. In addition to implementing XOR functionality for sign selection (using sgn and $sgnb$), the drivers also completely turn off the tap when required (for OFF = '1', OFF_b = '0'). Since the top-most NMOS transistors of the driver are fed by static signals ($sgn \cdot OFF\_b$ and $sgnb \cdot OFF\_b$)[4], they are typically sized larger than the other transistors to reduce the driver delay without incurring a power penalty.

3.5 Simulations, Test-Chip and Measurements

Fig. 3.11 shows the schematics of the prototype complex cascode current-summing DFE. To adaptively perform tap adaptation using sign-sign LMS and an edge-based CDR [10], additional 'ADAPTIVE' and 'EDGE' samplers are used. While 20 complex taps would nominally require 40 flip-flops in the feedback shift register chain, the floorplan for cascode current summation necessitates dedicated shift registers for direct and cross feedback to both I and Q channels. This

[4] sgn and OFF change when a tap changes its sign or is turned off, which is only when the channel response changes. Since the channel varies more than 4 orders of magnitude slower than the UI, the signals can be considered to be static for all practical purposes.

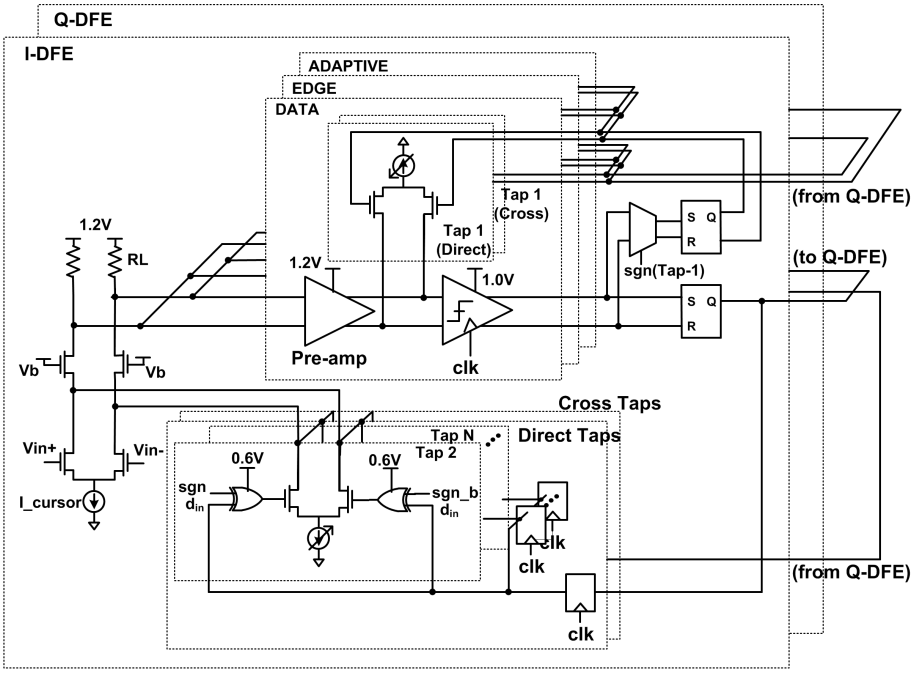


Figure 3.11: 20-complex-tap cascode current-summing DFE prototype

requirement doubles the requisite number of flip-flops to 80. At 10Gb/s, the DFE has a total power consumption of 14mW. From this total, 3mW is consumed by the 2 summing amplifiers, 4mW by the 6 preamps and comparators, and the other half of the power (7mW) is consumed by the feedback shift register chain of 80 flip-flops. The peculiar nature of this power breakdown (Fig. 3.14) once again illustrates the power penalty of digital gates at GS/s rates.

Fig. 3.12 shows post-layout simulations of the DFE using a PRBS-7 input convolved with a representative 60GHz channel. The output of the summing amplifier is free of all but the first tap of ISI, which is canceled separately at the input of the comparator. The comparator input eye highlights the difference in settling behavior of the first ISI tap cancelation current which transitions after the latter taps.

In order to validate cascode current summing, a 65nm CMOS baseband test-chip was fabricated with a 20-complex tap DFE supporting 10Gb/s (5GS/s) QPSK. Besides the DFE, the chip also comprised of CLK generation and data recovery (CDR) circuits, a variable gain amplifier (VGA), and a phase rotator (similar to [3]) to perform carrier phase/frequency recovery. A 500MS/s digital engine was also implemented for on-chip DFE tap adaptation and carrier recovery. Fig. 3.13 shows

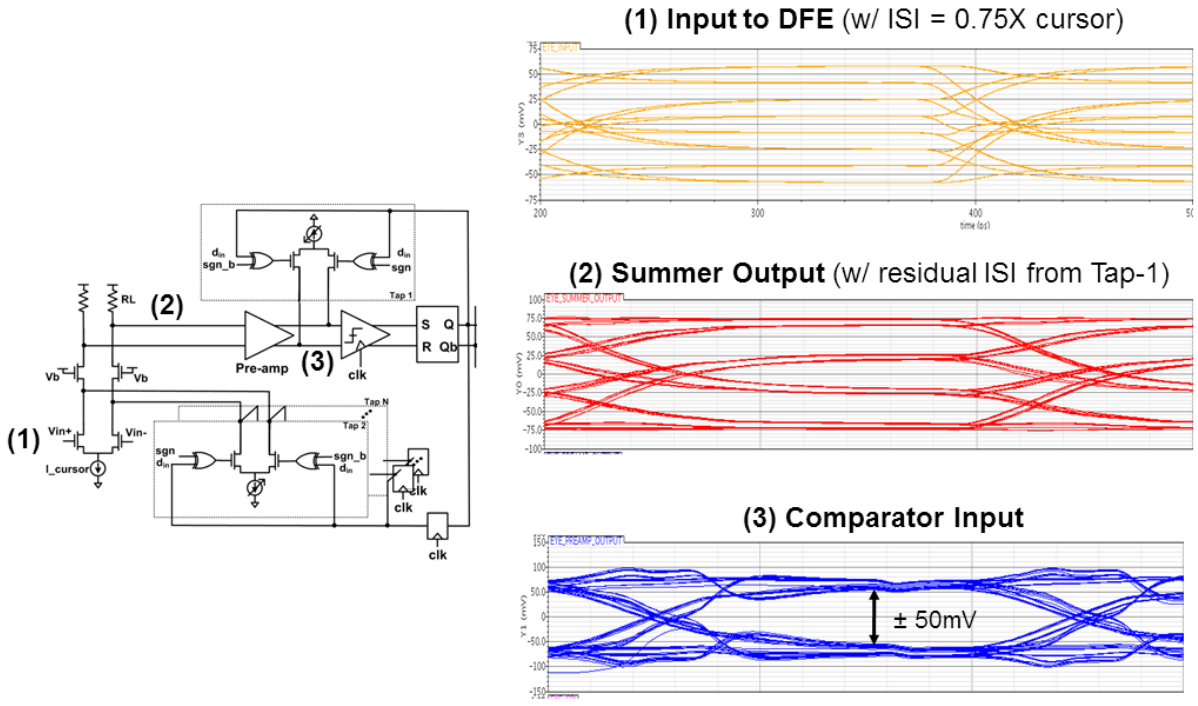


Figure 3.12: Eye diagrams with 5GS/s PRBS-7 input (post-layout simulation)

the block diagram of the entire baseband.

The chip measured 1.7mm x 1.1mm (Fig. 3.15), and was tested using a 5GS/s 2-channel Arbitrary Waveform Generator (AWG). The AWG was programmed to mimic a multi-path channel with ISI magnitude up to 2.5 times the cursor amplitude while generating $2^7 - 1$ and $2^9 - 1$ PRBS data on the I and Q channels respectively. The ISI was distributed randomly from taps 1-20 (both direct and cross feedback) over different measurements to test the operation of all taps. In order to match the received signal amplitude to a realistic 60GHz channel, the AWG amplitude and VGA gain were adjusted to receive a 120mV (diff, p-p) signal at the comparator.

When on-chip DFE adaptation was applied (along with the other on-chip adaptations for CDR and carrier recovery), the test-chip was able to receive 10Gb/s QPSK data with BER $< 10^{-12}$ measured by using an on-chip PRBS checker. Fig. 3.16 plots measured BER vs. hard-coded timing offset while receiving 10Gb/s data, before and after turning on the DFE. It can be seen that BER is $< 10^{-12}$ over 0.2UI of timing offset, thus validating 10Gb/s QPSK operation.

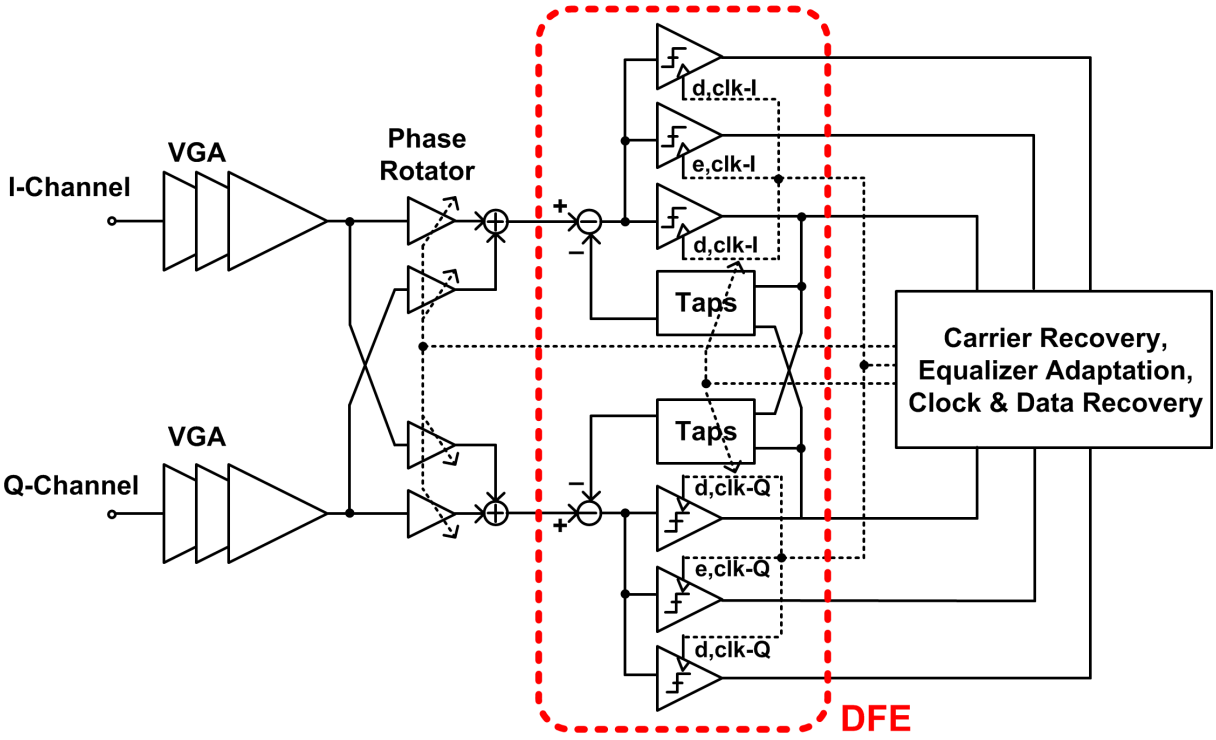


Figure 3.13: 60GHz baseband test-chip block diagram

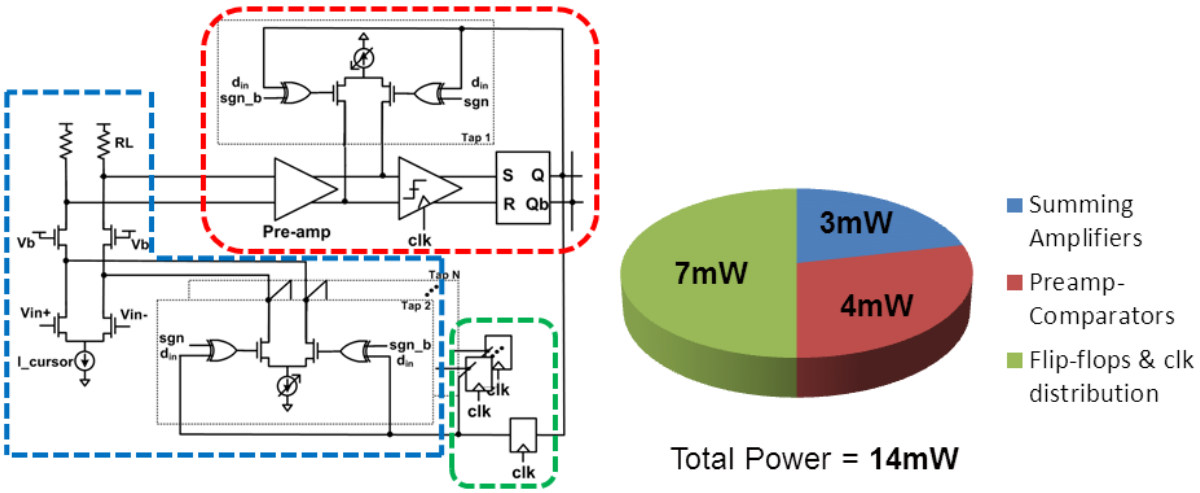


Figure 3.14: DFE power breakdown

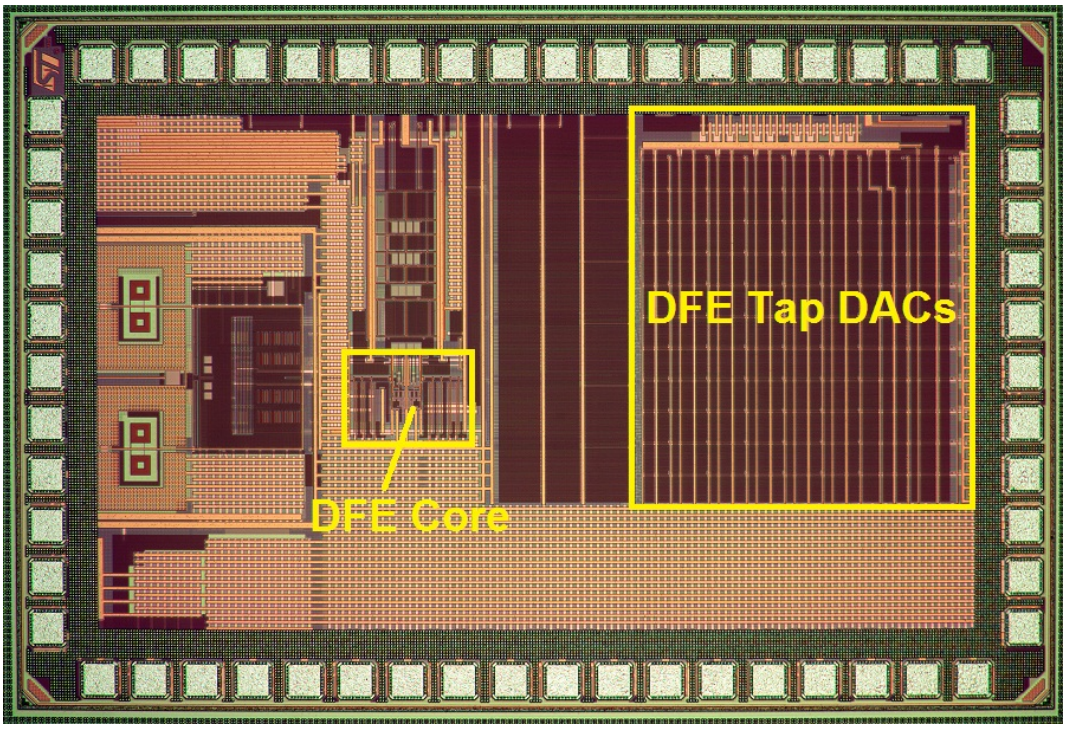


Figure 3.15: 60GHz baseband die micrograph: 1.7mm x 1.1mm

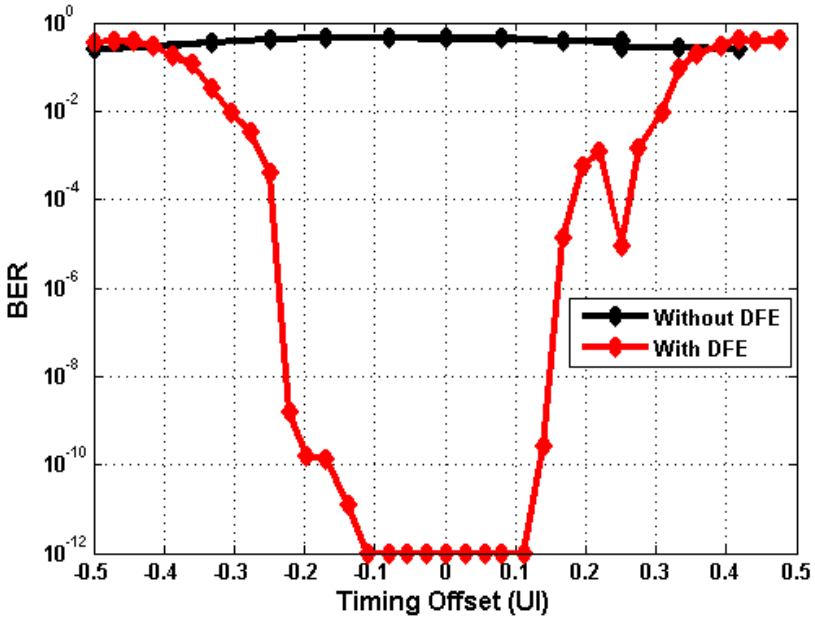


Figure 3.16: BER vs. timing offset (UI) with and without the DFE turned on.

Chapter 4

Conclusion

This report presents a mixed-signal approach to the design of a multi-Gb/s 60GHz transceiver baseband. Inspired by high-speed chip-to-chip serial links using analog/mixed-signal processing and simple modulation schemes like QPSK, this work offers a compelling lower power alternative to multi-bit OFDM-based wireless baseband solutions that tend to dissipate multiple Watts of power at GS/s rates. The techniques discussed in this work are an integral part of the effort to ease the power bottleneck for incorporating 60GHz transceivers into mobile hand-held devices.

A decision feedback equalizer (DFE), which is one of the key constituent blocks of the baseband, is presented as a representative design using mixed-signal processing. A design methodology was first developed to achieve the power-optimal DFE design for a given data-rate and expected interference profile. Using this design framework, we also derived the fundamental limits on a conventional current-summing DFE structure due to self-loading. The constraints due to self-loading are found to significantly limit the time-span of post-cursor ISI that can be canceled by such a structure, making the topology unsuitable for a 60GHz channel. A cascode current-summing structure then was proposed to relax these self-loading constraints. By making key observations about the channel and summing the ISI cancelation currents through a cascode transistor, this proposed structure can equalize a significantly longer ISI profile that is typical of a 60GHz channel response.

An important conclusion of this work is that at GS/s rates, it is much more efficient to use analog processing techniques with moderate resolution (5-6 bits) and simple modulation schemes,

as compared to multi-bit digital processing and modulation schemes with high complexity. This conclusion is validated by a prototype cascode current-summing DFE in 65nm CMOS with 20 complex post-cursor ISI taps that was shown to operate up to data-rates of 10Gb/s for BER less than 10^{-12} while consuming only 14mW of power. The energy efficiency of this prototype compares very favorably with OFDM-based solutions [7] which consume $\sim$1W of power at lower data-rates.

Bibliography

[1] Amendment of Part 2 of the Commissions Rules to Allocate Additional Spectrum to the Inter-Satellite, Fixed, and Mobile Services and to Permit Unlicensed Devices to Use Certain Segments in the 50.250.4 GHz and 51.471.0 GHz Bands. *Federal Commun. Comm.*, FCC ET Docket No. 99-261, Dec. 2000.

[2] M. R. Williamson, et al.,"Investigating the effects of antenna directivity on wireless indoor communication at 60GHz," *Proceedings of 1997 PIMRC Conference*, pp. 635-639.

[3] C. Marcu, et al., "A 90 nm CMOS Low-Power 60 GHz Transceiver With Integrated Baseband Circuitry," *IEEE Journal of Solid State Circuits*, pp. 3434-3447, vol. 44, no. 12, Dec. 2009.

[4] A. Tomkins, et al., "A Zero-IF 60 GHz 65 nm CMOS Transceiver With Direct BPSK Modulation Demonstrating up to 6 Gb/s Data Rates Over a 2 m Wireless Link," *IEEE Journal of Solid State Circuits*, pp. 2085-2099, vol. 44, no. 8, Aug. 2009.

[5] A. Valdes-Garcia, et al., "A SiGe BiCMOS 16-element phased-array transmitter for 60GHz communications" *IEEE ISSCC Digest of Technical Papers, 2010*, pp. 218 - 219.

[6] S. Emami, et al., "A 60GHz CMOS Phased-Array Transceiver Pair for Multi-Gb/s Wireless Communications" *IEEE ISSCC 2011*, to be presented.

[7] http://www.sibeam.com

[8] D. A. Sobel, R. W. Broderson, "A 1 Gb/s Mixed-Signal Baseband Analog Front-End for a 60 GHz Wireless Receiver," *IEEE Journal of Solid State Circuits*, pp. 1281-1289, vol. 44, no. 4, Apr. 2009.

[9] Leibowitz, et al., "A 7.5Gb/s 10-Tap DFE Receiver with First Tap Partial Response, Spectrally Gated Adaptation, and 2nd-Order Data-Filtered CDR," *IEEE ISSCC Digest of Technical Papers, 2007*, pp. 228 - 229.

[10] V. Stojanovic, et al., "Autonomous Dual-Mode (PAM2/4) Serial Link Transceiver With Adaptive Equalization and Data Recovery," *IEEE Journal of Solid State Circuits*, pp. 1012-1026, vol. 40, no. 4, Apr. 2005.

[11] M. Chen, M.-C.F. Chang, "A 2.2Gb/s DQPSK Baseband Receiver in 90-nm CMOS for 60GHz Wireless Links," *2007 IEEE Symposium on VLSI Circuits*, pp. 56-57.

[12] H. Chung et al., "A 7.5-GS/s 3.8-ENOB 52-mW flash ADC with clock duty cycle control in 65nm CMOS," *2009 IEEE Symposium on VLSI Circuits*, pp. 268-269.

[13] B. Verbruggen et al., "A 2.6mW 6b 2.2GS/s 4-times interleaved fully dynamic pipelined ADC in 40nm digital CMOS," *IEEE ISSCC Digest of Technical Papers, 2010*, pp. 296-297.

[14] M. El-Chammas, B. Murmann, "A 12-GS/s 81-mW 5-Bit Time-Interleaved Flash ADC with Background Timing Skew Calibration," *2009 IEEE Symposium on VLSI Circuits*, pp. 157-158.

[15] A. Amirkhany et al., "A 24Gb/s Software Programmable Multi-Channel Transmitter," *2007 IEEE Symposium on VLSI Circuits*, pp. 38-39.

[16] Z. Zhang et al., "An Efficient 10GBASE-T Ethernet LDPC Decoder Design With Low Error Floors," *IEEE Journal of Solid State Circuits*, pp. 843-855, vol. 45, no. 4, Apr. 2010.

[17] J. Poulton et al., "A 14-mW 6.25-Gb/s Transceiver in 90-nm CMOS," *IEEE Journal of Solid State Circuits*, pp. 843-855, vol. 42, no. 12, Dec. 2007.

[18] F. O'Mahony et al., "A 47x10Gb/s 1.4mW/(Gb/s) parallel interface in 45nm CMOS," *IEEE ISSCC Digest of Technical Papers, 2010*, pp. 156-157.

[19] K. K. Parhi, "High-Speed architectures for algorithms with quantizer loops," *Proc. IEEE Int. Symp. Circuits and Systems, 2010*, vol. 3, pp. 23572360.