Title
Design of Multi-Gb/s Multi-Coefficient Mixed-Signal Equalizers

Permalink
https://escholarship.org/uc/item/7s60c53n

Author
Thakkar, Chintan S.

Publication Date
2012

Peer reviewed|Thesis/dissertation

Design of Multi-Gb/s Multi-Coefficient Mixed-Signal Equalizers

by

Chintan S. Thakkar

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Elad Alon, Chair
Professor Ali M. Niknejad
Professor Andrew Packard

Fall 2012

Design of Multi-Gb/s Multi-Coefficient Mixed-Signal Equalizers

Abstract

Design of Multi-Gb/s Multi-Coefficient Mixed-Signal Equalizers

by

Chintan S. Thakkar

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Elad Alon, Chair

The explosion of personal devices that need ubiquitous connectivity is making both wireless and wireline communication experience increasingly rapid growth in data-rates. Wireless channels have been 'fortunate' to see new channels/standards being made available over the past decade to meet up to multi-Gb/s demands. One such medium is the wideband 60GHz channel. Wireless mediums, by definition however, are thwarted by multi-path reflection-based inter-symbol interference (ISI) – a problem which becomes only worse at higher speeds. For decades, equalizers have been used efficiently to mitigate such interference. However, wireless equalizers in commercial CMOS products are typically implemented in DSP along multi-level modulation schemes like OFDM, which when scaled to Gb/s speeds dissipate substantial power. This is particularly detrimental for handheld/mobile devices with limited battery capacity.

To ease the power bottleneck for equalization, this work instead proposes using mixed-signal techniques. As opposed to classic multi-level ADC/DSP design, such techniques are inspired by high-speed chip-to-chip wired communication that advocates the use of simple modulation schemes (such as QPSK) with few comparators. Since wireless channels suffer ISI with longer delay spreads than their wired counterparts, previously developed wireline equalizers cannot be directly ported. This work therefore enables energy-efficient equalizers to cancel extremely long ISI delay spreads. Our first prototype demonstrated a 40-coefficient complex (I/Q) decision feedback equalizer (DFE) in 65nm CMOS to enable 10Gb/s rates over line-of-sight (LOS) 60GHz channels, while consuming only 14mW of power. The second prototype in 65nm low-power (LP) CMOS enables non-line-of-sight (NLOS) channel equalization as well, by using a 32-coefficient receiver feedforward equalizer (FFE) and a longer 100-coefficient DFE, achieving 3.5-8Gb/s rates while consuming 20-67mW.

While the equalizer prototypes in this dissertation have been targeted towards 60GHz channels, the techniques enable energy-efficient equalization for long ISI delay spreads for any high-speed wireless or wireline communication link.

To my Parents and Grandparents

For making me worthy of writing this.

Contents

List of Figures

List of Tables

Acknowledgments

Over the last five and half years of doctoral research, I've lost count of the number of times I must have written this page in my head. For being able to reach here, I wish to wholeheartedly thank ...

... my research adviser, Prof. Elad Alon for his pool of infinite energy and enthusiasm, relentless attack of integrated circuit problems (and beyond), meticulous hands-on tutorship of circuit design, presentation and writing skills, enabling the exposure to industry, the 3AM phone calls – in a nutshell, for being the most super-human mentor possible.

... my qualifying examination and dissertation commmittee members – Prof. Ali Niknejad for his creative feedback and letters of recommendation, Prof. Jan Rabaey for his solid encouragement, and Prof. Andrew Packard, among other things, for the red-eye flight to Chicago (to be with his ailing father) and back a day before making it to my qualifying examination.

... the Berkeley Wireless Research Center (BWRC) as an outstanding environment for infrastructure, collaboration and creativity.

... BWRC sponsors and member companies, especially UC Discovery, Panasonic, Nokia, and Qualcomm for funding our research, and NSF Infrastructure Grant No. 0403427 for creating BWRC in brick and mortar.

... Brian Richards for wearing numerous hats in technical support, and being the go-to person for anything CAD under the sun.

... Susan Mellers and Fred Burghardt for the unenviable task of managing an enviable arsenal of equipment in the BWRC lab.

... Leslie Nishiyama and Olivia Nolan for imparting a charming spunk to BWRC.

... Tom Boot for his exceptional warmth.

... all other BWRC executives and staff, especially Gary Kelson, Ken Tang, Kevin Zimmerman, Pierce Chua, and Ubirata Chaves Coelho.

... Intel Corporation for their generous Ph.D. Fellowship.

... my managers during internships in industry – Ken Chang (Rambus Inc.), Bryan Casper and Christopher D. Hull (Intel Corp.) – for giving me opportunities to work in the 'real' world of semiconductors.

... my mentors and colleagues in industry – Hae-Chang Lee, Ting Wu, Brian Leibowitz (Rambus Inc.), and Randy Mooney, Farhana Sheikh, Frank O'Mahony, James Jaussi, Sudip Shekhar, Glenn Murata, Jay Wang, and Jerome Carmon (Intel Corp.).

... Vinayak Nagpal and Bhavna Kapoor for their warmth, hospitality and recreational inspiration, and Animesh Kumar for his invaluable friendship.

... Debopriyo Chowdhury and Cristian Marcu for their leadership and patiently answering all my questions during the formative years.

... Lingkai Kong for all his pace-setting help throughout the course of this dissertation (especially with the first 60GHz transceiver and baseband projects), and for sharing 5+ years of technical humor across the cubicle.

Chapter 1

Introduction

The explosion of personal devices that need ubiquitous connectivity is making both wireless and wireline communication experience increasingly rapid growth in data-rates. For wireless communication, new channels and standards have been made available over the past decade to meet up to multi-Gb/s demands (e.g. [1]). Communication over wireless channels is prone to multi-path reflection which causes inter-symbol interference (ISI). As data-rates increase, this interference effectively spreads over a larger number of symbol periods, thus making it more challenging to counter.

Wireline channels on the other hand offer a more controlled environment and typically have shorter ISI delay spreads. However, ever-increasing data-rate requirements have outpaced the bandwidth of these channels. Due to cost limitations, legacy wireline mediums such as PCB traces and backplanes continue to be used. To keep up with networking demands, I/O data-rate requirements for backplanes supporting routers and servers are being pushed by upcoming standards to as high as 25-28Gb/s [2]. When these bandwidth-limited legacy channels are used at such high throughputs, the ISI spreads over a larger number of symbol periods, similar to wireless channels.

A classic technique to counter ISI is by using an equalizer which filters out this ISI in either the time or the frequency domain. Wireline transceivers operating at GS/s rates employ a combination of both frequency and time-domain equalizers. To equalize a wide range of channel responses, time-domain equalizers offer flexibility in shaping the transfer function to equalize the channel response. Such flexibility can also be leveraged for a wireless channel in equalizing multi-path reflections at different time-delays. While this flexibility can also be achieved by linear equalization in the frequency-domain, it requires conversion between time and frequency domains by using FFT/IFFT blocks which tend to be power-intensive. Time-domain equalizers on the other hand have been typically implemented as mixed-signal circuits (most extensively in wireline transceivers) with excellent energy efficiency. The equalization capability required of time-domain equalizers in terms of number of symbol-rate taps is directly proportional to the channel ISI delay spread and the data-rate. As a result, both wireless and wireline communication with high delay spread and/or high data-rates respectively require increased equalization capabilities. The focus of this dissertation

is the design of such time-domain equalizers requiring many taps of equalization.

While wireline transceivers use energy-efficient mixed-signal equalization, wireless basebands have classically been relegated to using DSP. Wireless equalization is also typically carried out in the frequency domain using multi-level schemes such as orthogonal frequency division multiplexing (OFDM) and single-carrier frequency-domain equalization (SC-FDE). As will be shown in detail shortly, when this approach is scaled to GS/s rates, the power consumption of the DSP is substantial to be used in mobile/handheld devices. So far, the more efficient mixed-signal wireline techniques have not been applied for wireless most likely because wireless equalization requirements in terms of number of coefficients are significantly higher.

The primary focus of this dissertation is therefore the implementation of multi-coefficient time-domain equalizers using energy-efficient analog/mixed-signal circuitry. Aided by these equalizer designs, we will also demonstrate that using mixed-signal circuits in wireless basebands not only dramatically reduces power consumption but also enables significantly higher data-rates. To demonstrate the efficacy of such equalization, in this dissertation we shall use the wireless 60GHz band as a demonstration platform. By means of its wide bandwidth, the 60GHz channel enables multi-Gb/s communication, as will be shown shortly. To make such high data-rates an attractive proposition for mobile/hand-held devices, the baseband power consumption must be kept low. Since equalization consumes a significant portion of the total baseband power, we will demonstrate low-power equalizer design prototypes in this work. Despite the focus of these multi-coefficient designs for wireless channel equalization, the techniques and design frameworks developed in this dissertation will continue to be relevant for wireline equalization, especially in the light of increasing wireline data-rates.

We will first give a quick introduction to 60GHz communication and describe the current state-of-the-art of its mostly digital baseband design that is typical to wireless transceivers. To enable more energy-efficient equalization, we shall then introduce the two main categories of time-domain mixed-signal equalizers – namely decision feedback and feedforward equalizers. Finally, we shall examine the 60GHz channel characteristics across different environments to determine the requirements for these equalizers.

1.1 60GHz Wireless: A Demonstration Platform for Multi-Coefficient Equalization

The 60GHz wireless band with its 7GHz of bandwidth [1] has enabled ultra-high data-rate wireless communication. The attractiveness of multi-Gb/s wireless has drawn increasing interest from both academic [3][4][5][6] and industrial groups [7][8][9][10][11]. Commercial transceiver solutions for applications such as home HDTV streaming using the WirelessHD standard [12] are now also available[13].

Despite operating at much higher carrier frequencies and bandwidths, today's 60 GHz radios often bear significant resemblance to lower frequency designs. For example, similar to
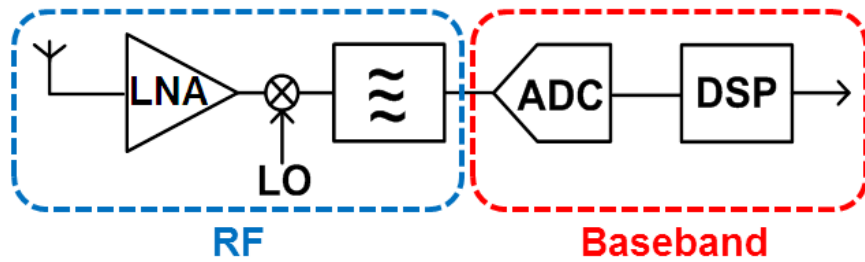
Figure 1.1: Typical implementation of a wireless receiver

today's 802.11a/b/g/n wireless LAN radios, current 60 GHz radios often rely on Orthogonal Frequency DivisionMultiplexing (OFDM). This choice of modulation scheme requires relatively high levels of circuit and signal processing complexity, and hence typical 60 GHz implementations utilize the traditional system partitioning shown in Fig. 1.1. The RF front-end provides directionality (typically using a phased array [8]), and performs low-noise amplification followed by downconversion and (optionally) filtering, while the baseband comprises of the data-converters and digital signal processing (DSP). The DSP performs the requisite signal-conditioning to counter the non-idealities of wireless communication, such as OFDM de-modulation (or compensating inter-symbol interference in single-carrier link), performing carrier phase/frequency recovery, clock/data recovery, and error-control coding/decoding.

Following this traditional radio implementation strategy, todays commercial 60 GHz basebands dissipate roughly 1W of power [13] which is likely to limit the use of these designs in mobile/hand-held devices. Although at first glance one may assume that the multi-GS/s data converters are responsible for the majority of this power, recent advances in such data-converter designs [14][15][16][17][18][19][20][6] 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 4-8 bits of dynamic range. This indicates that even a 5 GS/s ADC with 5-bit resolution would dissipate well under 25 mW. In a GS/s baseband therefore, the ADC will most likely not be the power bottleneck.

A closer examination thus reveals that the power of a typical 60 GHz baseband is dominated by the various DSP blocks. This can be illustrated by the power dissipation of digital 60GHz basebands presented recently by Okada *et al.* in [6] and Hsiao *et al.* in [21]. The digital baseband in [6] which incorporates a relatively short 8-tap time-domain FIR filter-based equalizer and a power-optimized LDPC decoder, besides carrier recovery and automatic gain control consumes 206mW and 224mW for 3.5Gb/s and 7Gb/s throughputs using QPSK and 16-QAM respectively. The digital solution in [21] with more comprehensive equalization – specifically 512-pt FFT based orthogonal frequency division multiplexing (OFDM) and single-carrier frequency domain equalizer (SC-FDE) based digital equalizers – consumes 150mW and 208mW respectively for the two equalization modes at 7Gb/s using

---

[1]FOM $= P/(2^{ENOB} \cdot f_s)$, where P – power dissipation, $ENOB$ – effective number of bits, f_s – sample-rate.
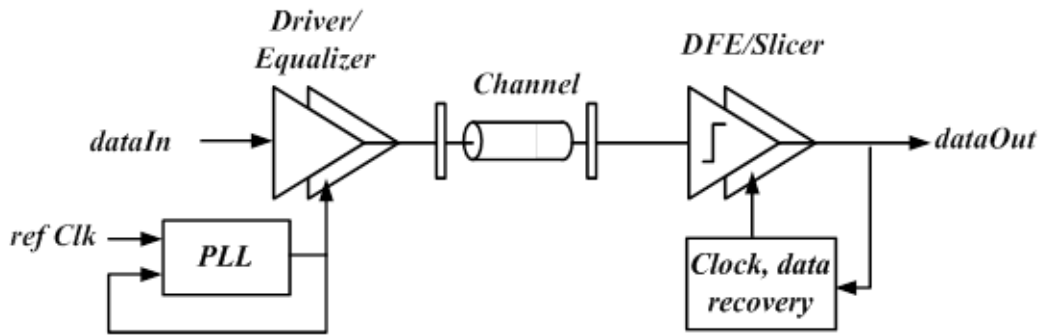
Figure 1.2: Typical implementation of a high-speed wireline transceiver

16-QAM.

Furthermore, the above-mentioned designs have only focused on current 60GHz standards [22][23] agreed upon by consortiums/alliances [12][24] supporting a maximum sample-rate of 1.76GS/s. While requirements on throughputs are increasing at breakneck speed, the power efficiency of digital designs – which are already operating close to the limit of feasible speeds in their respecive CMOS technology nodes – scales poorly for data-rates higher than 1.76GS/s. These DSP-based multi-GS/s designs which consume hundreds of mW of power (and potentially up to ∼1W for higher throughputs) will therefore likely be infeasible for mobile/hand-held devices. It is worthwhile examining alternative communication systems and implementations to seek inspiration for lower-power solutions.

High-speed chip-to-chip electrical 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. Specifically, the energy-efficiency achieved by current state-of-the-art serial link designs using mostly analog processing is 1mW/Gb/s [25][26] while operating at up to 12 Gb/s. For applications with even higher channel losses (25-35 dB) such as backplane-based transceivers requiring high transmit swings (0.6-1.2 V) and powerful equalizers (with as many as 14 taps of equalization), recent serial link designs [27], [28] have been able to achieve data-rates of 12-16 Gb/s with energy efficiencies of 8-15 mW/Gb/s.

Fig. 1.3 illustrates the key benefit offered by analog processing over digital 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
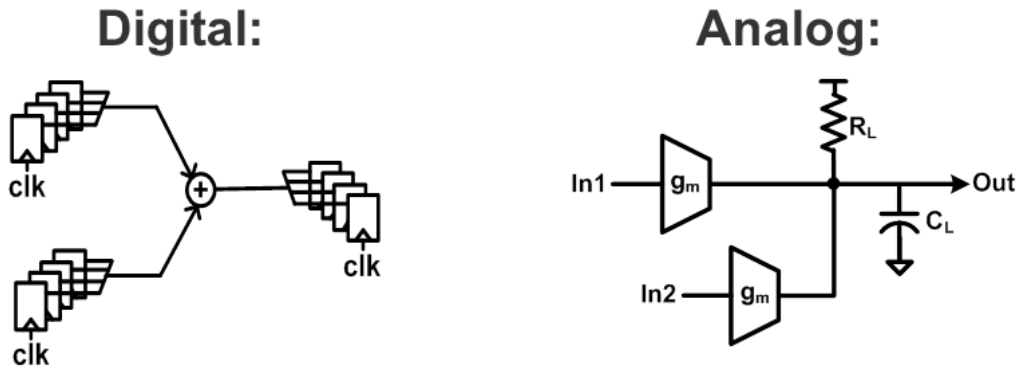
Figure 1.3: Comparison between digital and analog processing: Digital processing typically uses more transistor width (and hence more capacitance) per operation and consumes more power.
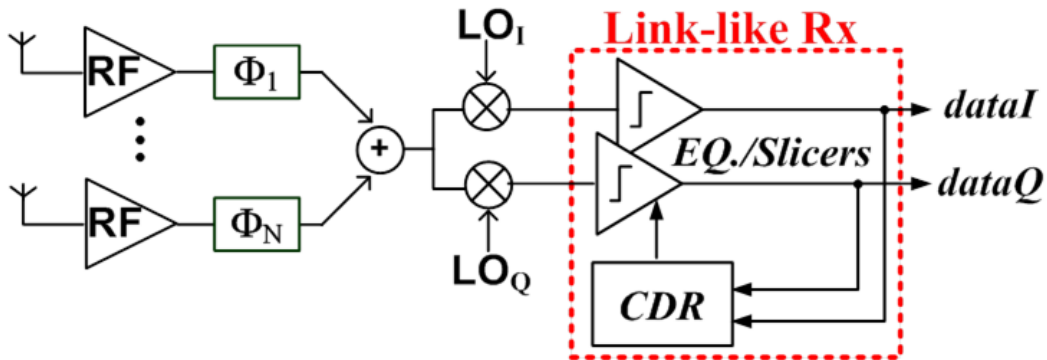
Figure 1.4: 60GHz receiver with mixed-signal baseband

as it is not constrained by noise or linearity[2]) and can hence achieve lower power. Note that serial links do use digital processing for certain high precision computations; however, since this processing is mostly limited to low-speed calibration/control circuitry, it does not incur a significant power penalty.

Inspired by such low-power high-speed serial links using analog processing, we aim for the design of an energy-efficient mobile mixed-signal 60GHz baseband (Fig. 1.4). Due to multi-path reflections, however, the wireless 60GHz channel (as will be shown in detail in section 1.3) has a much longer delay spread than a typical wireline channel. Mitigating the inter-symbol-interference (ISI) over such a long spread would require multi-coefficient time-domain equalizers that lie beyond the design space of serial links. The emphasis of this dissertation is therefore on designing feedback and feedforward time-domain equalizers for

---

[2]In modern sub-micron technologies, up to ~5 bits of linearity can be achieved without incurring a penalty on device size.
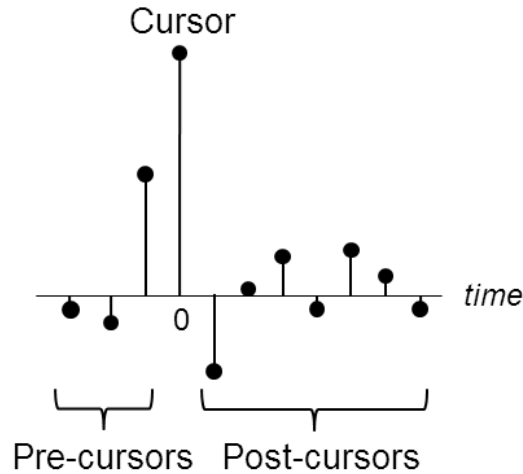
Figure 1.5: Channel impulse response
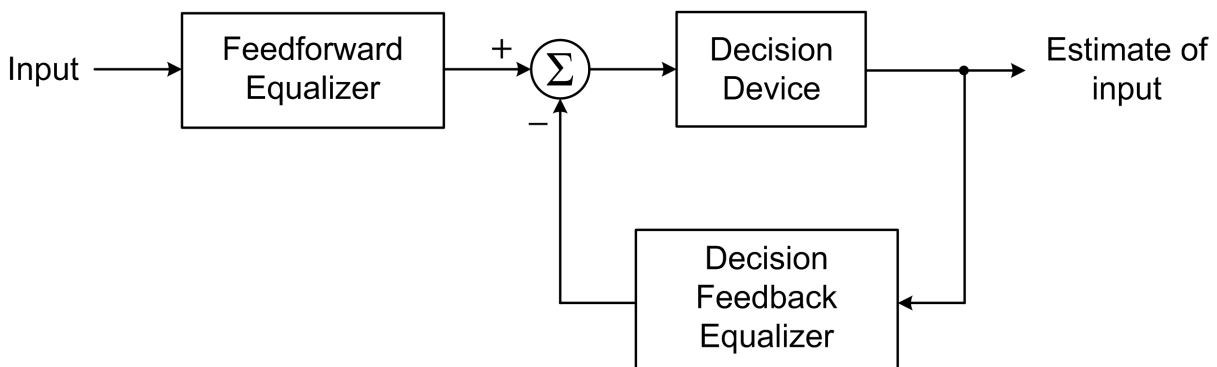
Figure 1.6: Block diagram of feedforward and decision feedback equalization [29].

line-of-sight (LOS) and non-line-of-sight (NLOS) multi-path 60GHz channels.

In the next two sections, we shall first introduce mixed-signal decision feedback and feed-forward equalizers, and then examine channel models for 60GHz wireless links to determine the requirements for these equalizers.

1.2   Mixed-Signal Time-Domain Equalizers

Fig. 1.5 shows a generic baseband channel response with ISI. Typically, the strongest response tap is assigned as the cursor. All taps occutring before the cursor in time are labelled as pre-cursors, while all taps occurring after the cursor in time are labelled as post-cursors. Since the ISI tends to reduce the effective received signal-to-noise (SNR) ratio, the presence of ISI is detrimental to the reception of transmitted bits. An equalizer improves reception
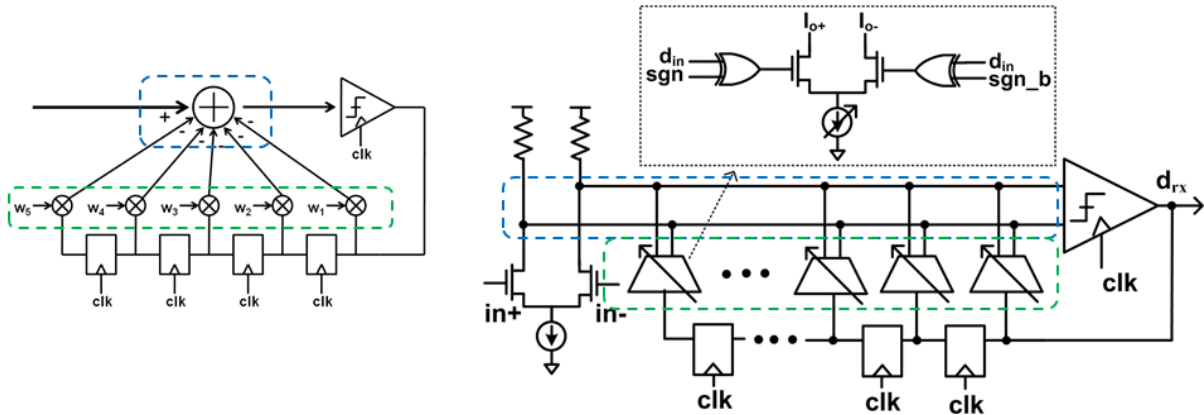
Figure 1.7: Block diagram of a decision feedback equalizer (DFE) (left); Conventional mixed-signal implementation using resistively-loaded summing of current-steering DACs representing ISI weights (right)

by removing the ISI. A time-domain equalizer filter in general can be divided into two categories – feedback (usually referred to as 'decision feedback') and feedforward, as shown in Fig. 1.6. The following subsections introduce the functionality and implementation of decision feedback and feedforward equalizers.

1.2.1   Decision Feedback Equalizer

A decision feedback equalizer (DFE) (Fig. 1.7 (left)) is essentially a non-linear filter which makes use of previous decisions in attempting to estimate the current symbol. Any trailing (post-cursor) ISI caused by previous symbols is reconstructed and then subtracted. Since a DFE generates the ISI canceling signal from the quantized estimate of the received signal, its primary advantage is that it does not amplify the received noise. While receiving signals with very low SNR (and therefore relatively high BER), the DFE is prone to making incorrect decisions, further aggravated by error propagation. For this work, however, which targets BER$< 10^{-3}$, the DFE is not impaired by error propagation, thus making it an excellent candidate to cancel the post-cursor portion of ISI.

DFE's can be implemented efficiently as a mixed-signal circuit, as shown in Fig. 1.7 (right). This conventional mixed-signal DFE structure cancels post-cursor ISI by subtracting currents representing the ISI coefficients at a resistive load. The cancelation currents are implemented by using current-steering DACs, whose values are adapted to the ISI. Within every UI period, every ISI coefficient needs to be settled starting from the respective feedback shift register flip-flop and ending with current settling at the RC-load. For the 1st coefficient, since this involves the slicer conversion of a small-signal to CMOS-levels, meeting the feedback latency constraint is particularly challenging. In mixed-signal DFEs, therefore, resolving this constraint has been the single-most explored design challenge.
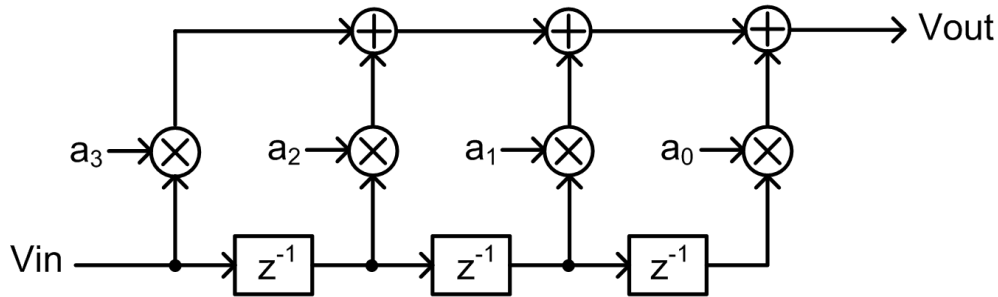
Figure 1.8: FFE Block Diagram

However, in the case of high-throughput wireless communication such as the wideband 60GHz channel, the more significant challenge is related to the delay spread of the ISI. As will be detailed in section 1.3, for targeted symbol rates as high as 5GS/s, a 60GHz channel may exhibit 30-50 complex[3] taps of post-cursor ISI, even with a directional front-end. Adding such a large number of weights in a mixed-signal structure is an unprecedented challenge at GS/s speeds. Previous solutions employing mixed-signal DFEs for 60GHz channels have either been relatively low-speed ([30] operates at 500MS/s, [31] operates at 1.76GS/s) or have limited equalization capabilities ([3] incorporates only a 5-tap complex DFE).

As will be described in detail in Chapter 2, the summing node of the structure is loaded by the parasitic capacitance of the current-steering switches. Therefore, in typical DFE implementations, only a relatively limited number of coefficients can be implemented before this self-loading makes it infeasible (at any power) to achieve the bandwidth required for multi-GS/s operation. In this work, we therefore propose 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. For equalizing channels with even longer delay spreads, the cascode summation approach can be combined with other power-efficient current summation topologies (such as current integration, as described in Chapter 3) to further extend the number of implementable taps.

1.2.2 Feedforward Equalizer

A feedforward or linear equalizer is essentially a finite-impulse-response (FIR) filter (Fig. 1.8) whose function is to collect the energy from a set of ISI taps onto the cursor. Given enough

---

[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.

taps, a feedforward equalizer (FFE) may be programmed to mitigate ISI from all taps – pre-cursor and post-cursor. However, since FFE filtering can potentially amplify thermal noise and given that a DFE is efficient at canceling post-cursor ISI without such amplification, the functionality of an FFE is typically limited to handling only pre-cursor taps.

For a wireless channel, in order to ease tap adaptation, an FFE is more convenient to implement at the receiver (RX). However, this creates a requirement for analog delay elements. As shall be seen in detail in Chapter 4, implementing analog delay leads to a tradeoff between linearity, area, and power. The tradeoffs are further aggravated with the requirement of up to 16 taps of feedforward equalization for the wireless 60GHz channel (as detailed in the section 1.3). In this work, we will propose an architecture to implement analog delay efficiently by using a switching matrix. This matrix is effectively a 1-to-N deserializer followed by N-parallel N-to-1 serializers to achieve an N-element delay line. The proposed architecture enables efficient implementation of the delay line, thus easing the power requirements for the FFE. To demonstrate this approach, a 65nm LP CMOS test-chip was designed that included a complex 16-tap (64 I/Q coefficient) RX-FFE capable of operation up to 8Gb/s, with the switching matrix consuming only 26mW of power. In the following section, we will determine the requirements on both feedback and feedforward equalizers by examining the baseband equivalent channel models for 60GHz wireless links.

1.3 Channel Models for 60GHz Wireless

The 60GHz channel, with its wide bandwidth, makes it amenable to achieve GS/s wireless data-rates. The high throughputs however also imply that even moderate delay spreads in the channel lead to a significant number of ISI taps. In order to quantify the ISI profiles in terms of delay spread and relative magnitude, empirical channel models have been developed in literature, most notably in [33], [32] and [34]. These works have conducted extensive experiments using 60GHz links in both office and residential spaces. The methodology and channel models in [32] and [34] have been used by the IEEE 802.11ad committee for standardization of 60GHz channel models. This section summarizes the findings of these channel models, and computes the requirements for time-domain equalization (TDE) using mixed-signal techniques.

According to the well-known Friis transmission equation[35], the small wavelength in the mm-wave band inherently causes high free-space path loss of as much as 68dB for a 1m TX-RX distance. Experiments with 60GHz radiation conclude that penetration (reflection) through (from) humans, standard building materials and surfaces such as walls and furniture also causes significant attenuation to 60GHz signals. Therefore in contrast with typical lower frequency wireless channels (such as WiFi and cellular frequencies), the 60GHz channel is more deterministic, and can actually be well understood by ray tracing techniques. Due to the finite smoothness of every surface, reflection rays each appear as a group of consecutive ISI taps, and can be modeled by the clustering approach [36]. Human blockers tend to diffract signals, thus raising the rms delay spread [34]. However, the high attenuation of
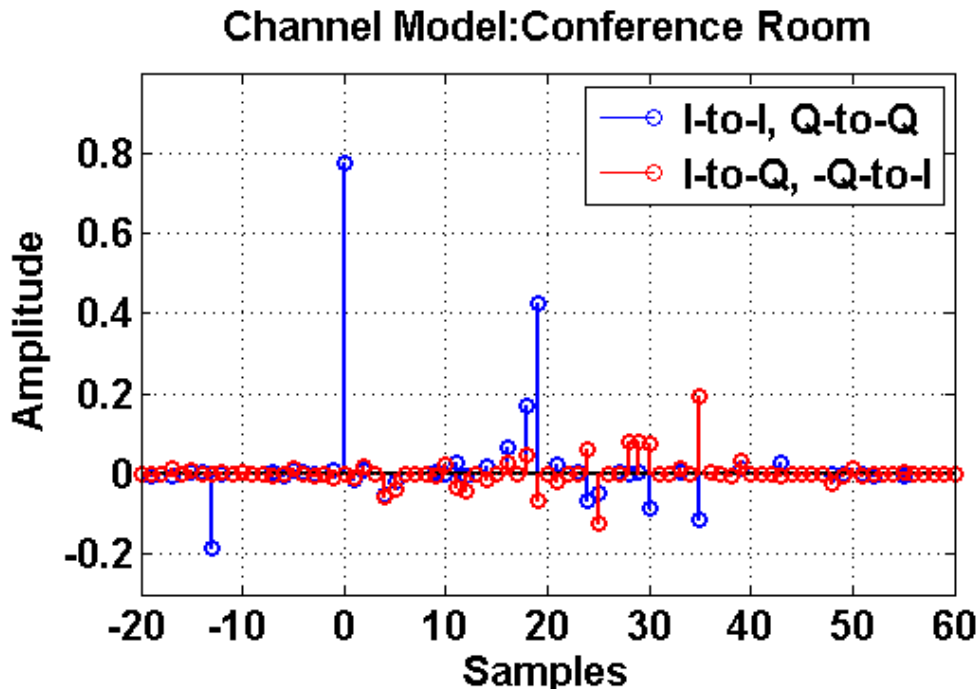
Figure 1.9: 60GHz NLOS channel model of conference room using model from [32]. Datarate: 4GS/s, TX-RX distance: 3m, TX, RX antenna half-power beamwidth (HPBW): 70°.

materials causes most reflections above second order to disappear below the thermal noise floor [33][32]. Since the high loss at mm-wave enforces the need for directional communication [37] (typically using phased arrays [38]) to satisfy the link budget, this further limits the number of reflections.

Based on the guidelines by A. Maltsev *et al.* [32] for the 802.11ad standard for 'WiGig' specifications [24], a statistical model was created. These instantiations were ranked as per their total ISI magnitude relative to the cursor. Fig. 1.9 shows an instantiation of channel models for a conference room around the 95th percentile by total ISI magnitude. In order to quantify the requirements on the equalizer in terms of number of taps, a coverage model was then developed. In this context, coverage is defined as the percentage of statistical channel instantiations that can be equalized by the time-domain equalizer canceling pre-cursor and post-cursor ISI (using an FFE and DFE respectively) to below a target uncoded bit-error-rate (BER). The coverage is also compared to that of a digital SC-FDE using 512-point FFT. Figs. 1.10 and 1.11 show the coverage for an NLOS living room model with high directionality (30° TX, RX antenna half-power beamwidth (HPBW)) and conference room model with lower directionality (70° TX, RX antenna half-power beamwidth (HPBW)). For both these plots, the analysis was done for the WiGig-standard-based 1.76GS/s for an uncoded symbol-rate of BER< 10^{-3} versus number of equalizer taps, before and after equalization. As seen in the figures, at this BER, the SC-FDE achieves 95% coverage. Based on the coverage analysis, it can be concluded that to realize a post-equalization BER of 10^{-3},
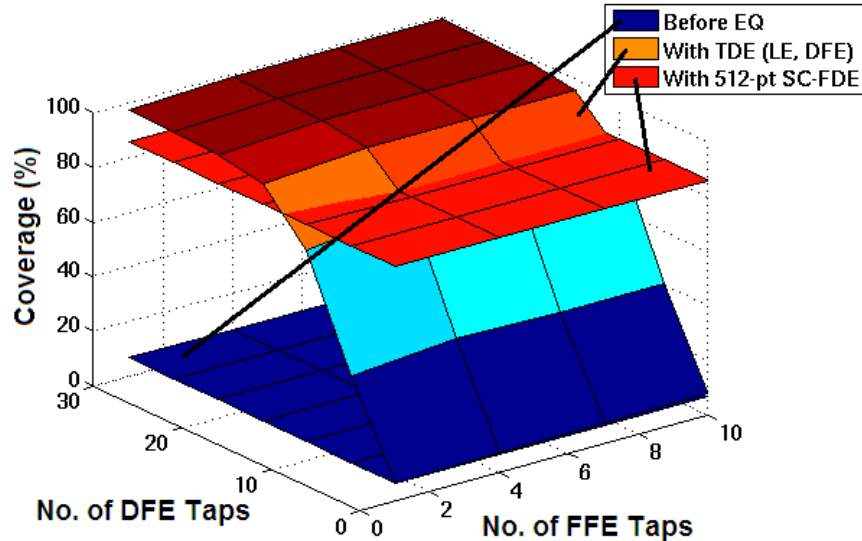
Figure 1.10: Time-domain equalizer coverage (percentage of channels equalized to BER< 10^{-3}) vs. 512-pt FFT-based SC-FDE for LOS living-room channel model[34]). Data-rate: 1.76GS/s, TX-RX distance: 3m, TX, RX antenna HPBW: 30°.

the TDE needs 16 pre-cursor taps and 20 post-cursor taps at 1.76GS/s.

Based on these channel models, the focus of the equalizer design in this dissertation will be two-pronged. The first design prototype strives to enable communication across relatively short 1-1.5m LOS channels with reasonable directionality and a moderate number (20) of post-cursor taps. The equalization of post-cursors is done by using a mixed-signal DFE. While a DFE is an efficient way to cancel post-cursor ISI, the total number of coefficients for this channel is significantly higher than any prior DFE art. Consequently, to implement a power-efficient QPSK design in 65nm CMOS working up to 10Gb/s is challenging.

Building upon the techniques developed in the first prototype, the second design targets the more exhaustive NLOS channels for longer (up to 3m) distances. The targeted data-rates vary from WiGig-specified 3.5Gb/s to a maximum of 8Gb/s in a 65nm LP CMOS process.[4] To handle pre-cursor ISI, this design has an additional 16-tap (32-coefficient) I/Q FFE at the receiver. While a 16-tap FFE can comprehensively equalize moderately directional NLOS channels only up to 1.76GS/s, higher data-rates will require increased directionality to satisfy the link budget. The increased directionality limits the precursor delay spread and hence the FFE tap requirements. To handle the increased delay spread of post-cursor ISI, a 50-tap (100-coefficient) I/Q decision feedback equalizer is designed.

---

[4]Since this is a commercial low-power technology without a provision for low-threshold transistors, the fanout-of-4 (FO4) delay is ∼2X slower than the corresponding GP CMOS process. The peak data-rate targeted is therefore 20% lower than the first prototype.
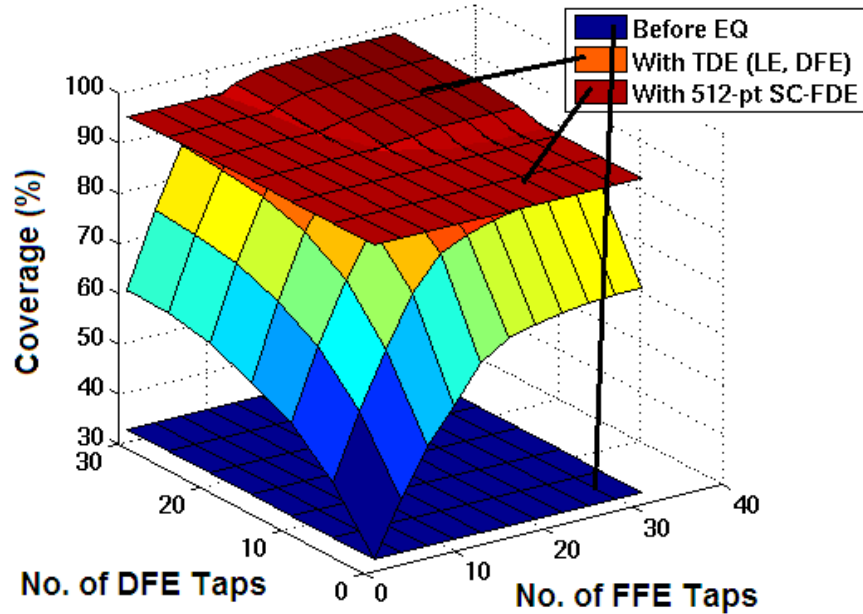
Figure 1.11: Time-domain equalizer coverage (percentage of channels equalized to BER< 10^{-3}) vs. 512-pt FFT-based SC-FDE for NLOS conference-room[32]). Data-rate: 1.76GS/s, TX-RX distance: 3m, TX, RX antenna HPBW: 70°.

1.4 Organization of Dissertation

The dissertation first presents the design methodology for a mixed-signal DFE in Chapter 2. The methodology is 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 highlight these improvements, followed by the implementation of one such prototype 20-complex tap cascode-summation structure in 65nm CMOS operating at 10Gb/s. To enable the implementation of an even higher number of coefficients, Chapter 3 analyzes other coefficient summation techniques developed recently in literature, namely current integration based current summation and switched capacitor based voltage summation. The analysis identifies the architecture most suited for implementing many coefficients.

As a next step towards dealing with the more generic NLOS channels, RX-FFEs are described in Chapter 4. This chapter examines prior RX-FFE art using rotating coefficients [39] and interleaved sampling [40], and explains the limitations in power efficiency of both these architectures for implementing multiple coefficients. The limitations motivate the proposed switching-matrix based architecture that provides an efficient soluton for multi-coefficient RX-FFE implementation. Finally, Chapter 5 describes the design and measurement results of a prototype 8Gb/s 32-coefficient RX-FFE, 100-coefficient DFE in 65nm LP CMOS.

As mentioned earlier, while the mixed-signal equalizer prototypes demonstrated for this work were targeted for multi-Gb/s 60GHz wireless communication, the design techniques that have been developed can be used for any high-speed wireless and wireline communication link requiring energy-efficient multi-coefficient equalization.

Chapter 2

DFE Design using Cascode Current-Summing

As a first step towards designing equalizers for high throughput channels with long multi-path delay spreads – such as the wideband 60GHz channel – this chapter addresses the mitigation of post-cursor ISI. As briefly introduced in the previous chapter, a mixed-signal decision feedback equalizer (DFE) is an excellent candidate for efficiently canceling such post-cursor ISI. Such mixed-signal DFEs have been widely used for wireline channel equalization with up to ∼15 taps of equalization [2]. As shown in the previous chapter, however, multi-GS/s communication over wireless channels such as 60GHz could require mitigation of up to 40-100 post-cursor ISI coefficients. This increase in the number of DFE coefficients, which is almost an order of magnitude higher than prior art, calls for a deeper understanding of the design limitations of conventional mixed-signal DFE architectures.

The first part of this chapter analyzes the power consumed by a conventional current summation structure as a function of the number of coefficients, and illustrates the limitations of this architecture. These limitations and their dependence on the nature of channel motivate the improvements that can be brought about by a proposed cascode summer design described in the second part of the chapter. An analysis of the cascode-summing architecture is then performed, showing how it significantly eases the dependence of power on the number of coefficients. Finally, to illustrate the power-efficiency of the proposed architecture, a 65nm CMOS prototype of a 40-coefficient cascode-summing DFE for 60GHz channels is described, which operates up to 10Gb/s while consuming only 14mW of power.

2.1 Conventional Current Summing DFE Architecture

Representing ISI coefficients as currents and adding them at a resistive load has been the most widely implemented style of DFE summing. The technique has been most popular in wireline receivers to mitigate channel bandwidth limitations, reflections due to impedance mis-
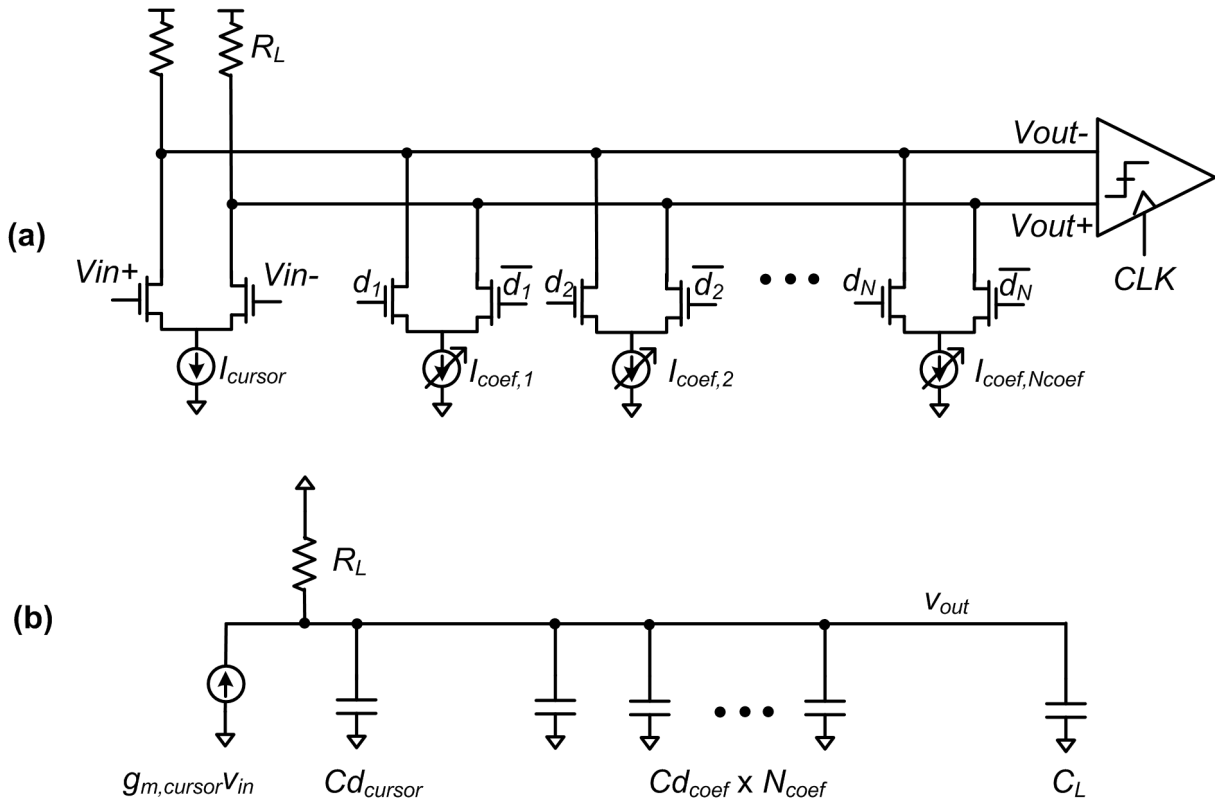
Figure 2.1: Gain/bandwidth analysis of a conventional resistively-loaded current-summing DFE. (a) Schematics (feedback shift register not shown) and (b) Single-ended small-signal model.

matches, or some combination thereof. Initial implementations [41][42] used triode-PMOS loads; however to decrease intrinsic capacitive loading, PMOS loads were eventually replaced by physical resistors [43]. While designing a resistively-loaded DFE, the primary specifications are similar to a Class-A amplifier – voltage gain (to achieve reasonable output swing for the following slicer) and bandwidth (to settle the per-UI switching ISI currents at the RC-load with sufficient accuracy). Given these constraints, it is informative to understand the dependence of power dissipation on the data-rate, the ISI profile/number of coefficients in the DFE, as well as technology-related parameters. As will be shown through the analysis, this dependence highlights the limitations of conventional 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} (excluding the ISI), while the summing amplifier has a DC gain of G and operates at a data-rate of f_s symbols per second. The DFE has N_{coef} number of coefficients that can be adapted to work across multiple channels. To enable this flexibility, each DFE coefficient can cancel ISI up to a maximum amplitude of k times the cursor amplitude. While this analysis assumes, for simplicity, the same maximum across all

coefficients, if certain coefficients magnitudes are bounded differently than the others, it is easy to introduce a variable maximum magnitude as a function of coefficient position (say, k_i for the i-th coefficient). In this analysis, the maximum coefficient current, I_{coef} would be

$$I_{coef} = k \cdot g_{m,cursor} \cdot v_{in} \tag{2.1}$$

In order to properly cancel the ISI, the coefficient-DAC current (I_{coef}) must be steered by the differential data signal (shown as d and $\bar{d}$) from the feedback shift register and satisfactorily settled at the summation node before the next data bit is resolved by the comparator. A full-rate[1] DFE therefore has a 1UI timing constraint for the settling of each coefficient. 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 coefficient), and αUI for the analog settling of the coefficient current at the summation node ($\alpha < 1$). The time-constant τ of this settling is given by the RC product

$$\tau = R_L \cdot (C_L + Cd_{cursor} + N_{coef} \cdot Cd_{coef}) \tag{2.2}$$

where R_L is the summation load resistance, Cd_{cursor} is the drain capacitance of the input transistor, Cd_{coef} is the drain capacitance of the current steering switch at each coefficient, 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 differential g_m-stage, $g_{m,cursor} = \dfrac{I_{cursor}}{V^*}$, where I_{cursor} is the DC bias current of the input pair[2]. The DC gain is therefore given by

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

Of the three capacitors at the summation node in (2.2), Cd_{cursor} and Cd_{coef} are attributed to the internal self-loading of the structure and are functions of the summing amplifier currents, while C_L is a fixed external loading. The internal capacitors can be expressed in terms of technology parameters, $C_{dI,cursor}$ and $C_{dI,coef}$, where C_{dI} denotes transistor drain capacitance per unit drain current.

$$Cd_{cursor} = C_{dI,cursor} \cdot \frac{I_{cursor}}{2} \tag{2.4}$$

$$\begin{aligned}
Cd_{coef} &= C_{dI,coef} \cdot I_{coef} = C_{dI,coef} \cdot (k \cdot g_{m,cursor} \cdot v_{in}) \\
&= C_{dI,coef} \cdot \left(k \cdot \frac{v_{in}}{V^*_{cursor}}\right) \cdot I_{cursor}
\end{aligned} \tag{2.5}$$

---

[1]Full-rate implies that the sampling CLK frequency is equal to the data-rate.
[2]V^* is defined as $V^* = 2I_{bias}/g_m$.

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

$$\tau = \frac{G \cdot V^*}{I_{cursor}} \cdot \left(C_L + C_{dI,cursor} \cdot I_{cursor} + N_{coef} \cdot C_{dI,coef} \cdot k \cdot \frac{I_{cursor}}{V^*_{cursor}} \cdot v_{in} \right) \tag{2.6}$$

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:

$$I_{cursor} = \frac{C_L \cdot \left(\frac{n_\tau f_s}{\alpha} \right) \cdot G \cdot V^*}{1 - \left(\frac{n_\tau f_s}{\alpha} \right) \cdot G \cdot V^*_{cursor} \cdot \frac{C_{dI,cursor}}{2} \cdot \left(1 + N_{coef} \cdot k \cdot \frac{v_{in}}{V^*_{cursor}} \cdot \frac{2\, C_{dI,coef}}{C_{dI,cursor}} \right)} \tag{2.8}$$

Now, if unity gain frequency, ω_T is defined as

$$\omega_T = \frac{g_m}{Cg} = \frac{2 \cdot I_{bias}}{V^*} \frac{\gamma}{Cd} \tag{2.9}$$

in which γ is the ratio of drain to gate capacitance, then C_{dI} in (2.8) can be re-written as

$$\frac{Cd}{I_{bias}} = C_{dI} = \frac{2\gamma}{V^* \cdot \omega_T} \tag{2.10}$$

Now, in (2.8), let us define $GBW = G \cdot \dfrac{n_\tau \cdot f_s}{\alpha}$ as the gain-bandwidth product. Let us also define $I_{nom} = C_L \cdot GBW \cdot V^*$ as the nominal current consumption for a class-A amplifier without self-loading. I_{cursor} may then be written in a simplified form as:

$$I_{cursor} = \frac{I_{nom}}{1 - \left(\gamma \cdot \frac{GBW}{\omega_{T,cursor}} \right) \left(1 + N_{coef} \cdot k \cdot \frac{v_{in}}{V^*_{coef}} \cdot \frac{2\, \omega_{T,cursor}}{\omega_{T,coef}} \right)} \tag{2.11}$$

Since the coefficient currents (I_{coef}) 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 resistively current-summed DFE, $P(I, res)$, as a function of N_{coef} also takes the form of (2.11), and can be expressed as:

$$P(I, res) \propto \frac{\left(\frac{n_\tau}{\alpha} \right) \cdot I_{nom} \cdot V_{dd}}{1 - \left(\gamma \cdot \frac{GBW}{\omega_{T,cursor}} \right) \left(1 + N_{coef} \cdot k \cdot \frac{v_{in}}{V^*_{coef}} \cdot \frac{2\, \omega_{T,cursor}}{\omega_{T,coef}} \right)} \tag{2.12}$$
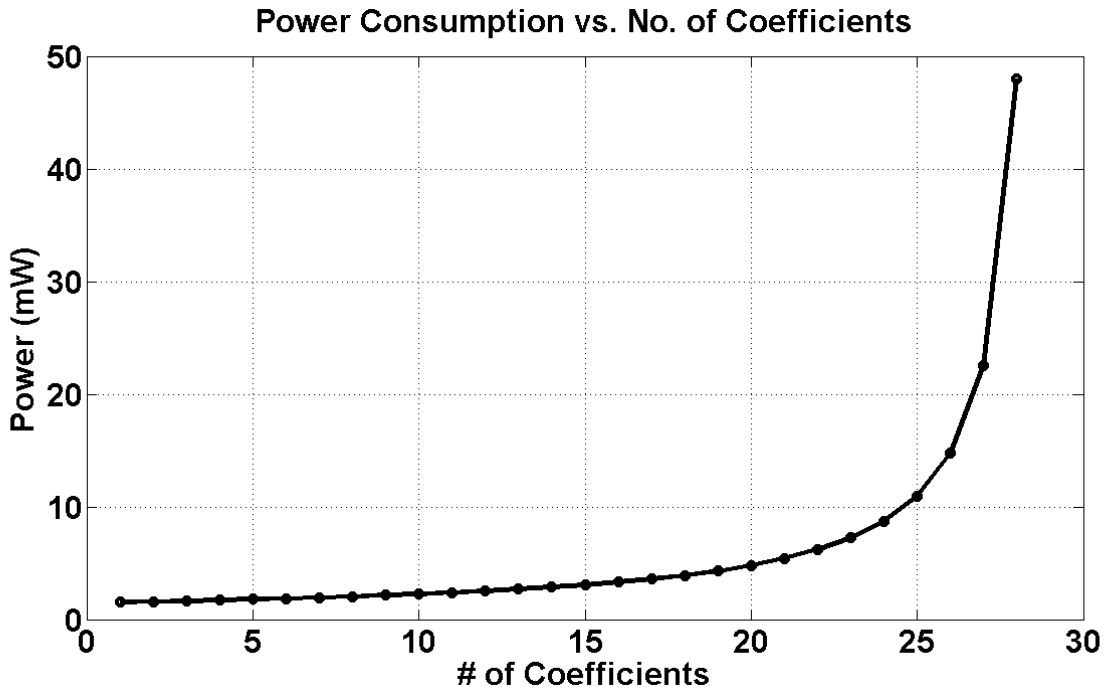
Figure 2.2: Conventional current-summing DFE power vs. no. of coefficients at 5GS/s, with $k_{max} = 0.5$ in 65nm CMOS.

The form of equations (2.11), (2.12) illustrates how a conventional mixed-signal DFE can only support a limited number of coefficients. When N_{coef} is small, to handle the extra capacitance of every additional coefficient, 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_{coef} becomes comparable to the ω_T of the technology, the DFE becomes self-loaded to the point that it cannot handle more coefficients for any increase in power, as seen in Fig. 2.2. At the desired data-rate of 5GS/s in a 65nm GP CMOS technology, a conventional DFE structure can implement only $\sim$10 complex taps (i.e. 20 I/Q coefficients) efficiently. Clearly, such a structure is incapable of being directly used in a 60GHz transceiver, which, as seen in the channel modeling analysis, typically needs up to 50 taps (i.e. up to 100 I/Q coefficients) of equalization.

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 section, the proposed cascode current-summation technique alleviates the penalty associated with this flexibility and is able to significantly extend the number of feasible taps.
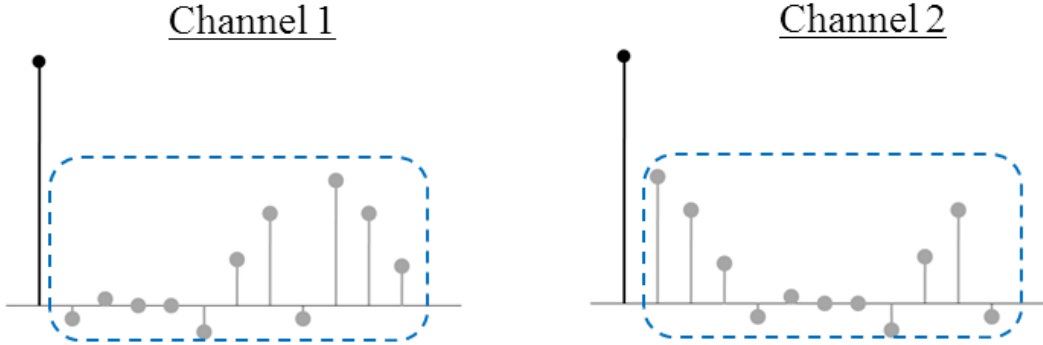
Figure 2.3: Wireless channel response: While each tap can have a variable weight, not all taps will be at their maximum weight altogether. However, the sum of all tap magnitudes is bounded due to finite transmit power.

2.2   Proposed Cascode Current-Summing DFE

The previous section highlighted the shortcomings of a conventional DFE. The DFE structure is primarily constrained by the 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 section, we will show that by making key observations about the impact of channel variability on the DFE design, a cascode current-summing structure is able to incorporate the requisite flexibility while notably improving the number of feasible taps.

2.2.1   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 taps need to be set to their maximum magnitude at the same time (Fig. 2.3). In other words,

$$\|ISI\|_1 < \sum_{i=1}^{N_{coef}} |ISI_{i,max}| \tag{2.13}$$

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 $I_{ISI,max} < N_{coef} \cdot I_{max}$. 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
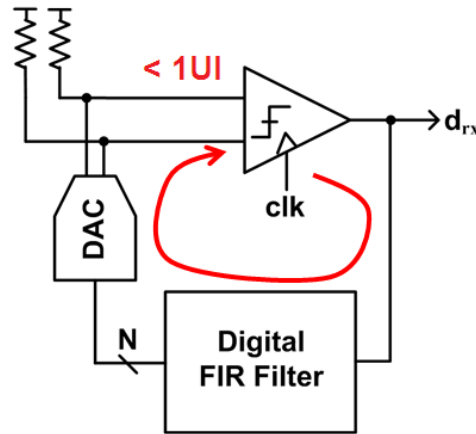
Figure 2.4: A fully digital FIR/DAC implementation of the DFE limits the self-loading at the summing node. However, the 1UI latency constraint on the first tap feedback makes the FIR adder unacceptably expensive in power

this would be to use a fully digital FIR filter in the feedback path to sum all taps and cancel the ISI using a DAC, as shown in Fig. 2.4. 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. 2.5. 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-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.

2.2.2 Analysis

Fig. 2.6 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, the data-rate is f_s symbols per second, and the number of coefficients is N_{coef}. The total cursor DC biasing current is I_{cursor}, while the common-mode current added to each side is I_{boost}. Each of the coefficients can cancel ISI up to a maximum amplitude of k times the
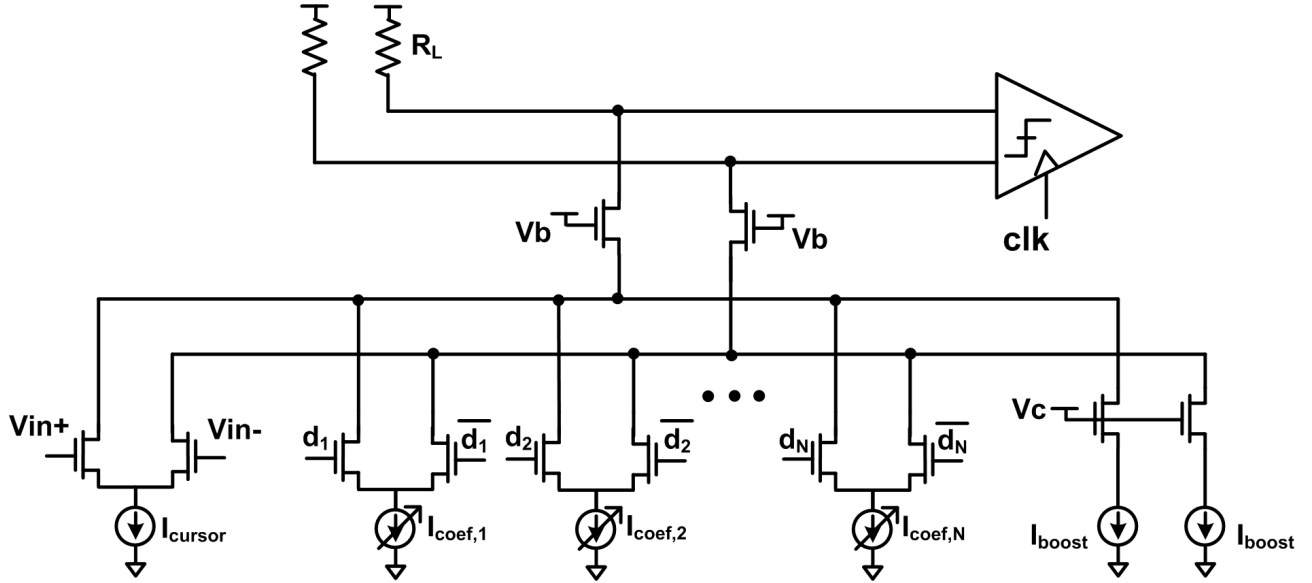
Figure 2.5: Cascode current summing structure: Tap switches moved to low-impedance cascode, and capacitive load at output node reduced

cursor, and a total ISI of amplitude of $k_{ISI,max}$ times the cursor amplitude. Therefore,

$$I_{coef} = k \cdot g_{m,cursor} \cdot v_{in} = k \cdot \frac{I_{cursor}}{V^*} \cdot v_{in} \tag{2.14}$$

$$I_{coefs,total} = k_{ISI,max} \cdot \frac{I_{cursor}}{V^*} \cdot v_{in} \tag{2.15}$$

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

$$\sqrt{2} \cdot n_\tau \cdot \tau = \alpha \cdot T \tag{2.16}$$

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

$$\tau_1 = R_1 \cdot C_1 = \frac{1}{g_{m,casc}} \cdot (N_{coef} \cdot Cd_{coef} + Cs_{casc} + Cd_{cursor} + Cd_{boost}) \tag{2.17}$$

$$\tau_2 = R_2 \cdot C_2 = R_L \cdot (C_L + Cd_{casc}) \tag{2.18}$$
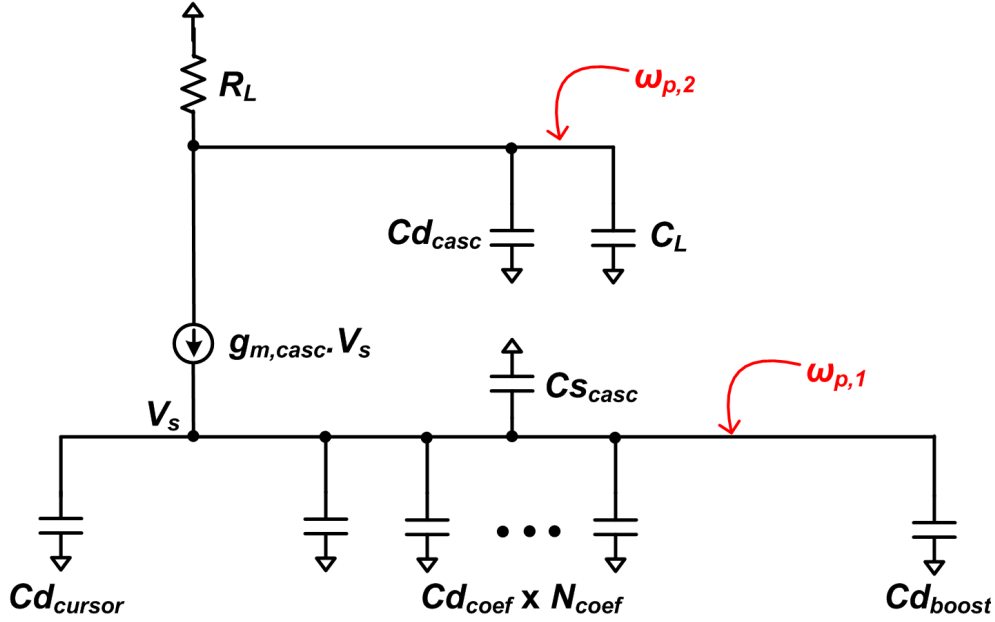
Figure 2.6: Small-signal model of a cascode current-summing DFE

where $g_{m,casc}$ is the g_m of the cascode transistor, Cd_{coef} is the current steering switch drain capacitance, Cs_{casc} is the cascode source capacitance, Cd_{cursor} is the input transistor drain capacitance, Cd_{boost} is the common-mode boost transistor drain capacitance, R_L is the summation load resistance, C_L is the loading from the next stage (i.e. the preamp/comparator), and Cd_{casc} is the cascode transistor drain capacitance. R_L is set by the DC gain requirements of the DFE, and given by

$$R_L = G \cdot \frac{V^*}{I_{cursor}} \tag{2.19}$$

The g_m of the cascode can be calculated as

$$g_{m,casc} = \frac{2 \cdot \left(\frac{I_{cursor}}{2} + I_{boost} - \frac{v_{in}}{2} \cdot \frac{I_{cursor}}{V^*} \right)}{V^*} \tag{2.20}$$

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. This equivalently specifies the linear region of the cursor swing, v_{in}.

The capacitors internal to the structure can be expressed in terms of technology parameters, $C_{dI,cursor}$, $C_{dI,coef}$, $C_{dI,boost}$, $C_{dI,casc}$, $C_{sI,casc}$ where C_{dI} and C_{sI} respectively denote drain and source capacitance per unit current, and the subscripts respectively refer to cursor, coefficient switch, boost and cascode transistors.

$$Cd_{cursor} = C_{dI,cursor} \cdot \frac{I_{cursor}}{2} \tag{2.21}$$

$$Cd_{coef} = C_{dI,coef} \cdot \left(k \cdot \frac{I_{cursor}}{V^*} \cdot v_{in} \right) \tag{2.22}$$

$$Cd_{boost} = C_{dI,boost} \cdot I_{boost} \tag{2.23}$$

$$Cd_{casc} = C_{dI,casc} \cdot I_{casc} = C_{dI,casc} \cdot \left(\frac{I_{cursor}}{2} + I_{boost} + k_{ISI,max} \cdot \frac{I_{cursor}}{V^*} \cdot v_{in} \right) \tag{2.24}$$

$$Cs_{casc} = C_{sI,casc} \cdot I_{casc} = C_{sI,casc} \cdot \left(\frac{I_{cursor}}{2} + I_{boost} + k_{ISI,max} \cdot \frac{I_{cursor}}{V^*} \cdot v_{in} \right) \tag{2.25}$$

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

$$p = \frac{k \cdot v_{in} \cdot C_{dI,coef} \cdot N_{coef} + \dfrac{V^*}{2} \cdot (C_{sI,casc} + C_{dI,cursor}) + k_{ISI,max} \cdot v_{in} \cdot C_{sI,casc} - \dfrac{\alpha \cdot (1 - v_{in}/V^*)}{2\sqrt{2} \cdot n_\tau \cdot f_s}}{\dfrac{\alpha}{\sqrt{2} \cdot n_\tau \cdot f_s} - V^* \cdot (C_{dI,boost} + C_{sI,casc})} \tag{2.26}$$

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

$$I_{cursor} = \frac{C_L \left(\dfrac{\sqrt{2} n_\tau f_s}{\alpha} \right) G V^*}{1 - \left(\dfrac{\sqrt{2} n_\tau f_s}{\alpha} \right) G V^* \dfrac{C_{dI,cursor}}{2} \left(1 + U + N_{coef} k \dfrac{v_{in}}{V^*} \dfrac{2 C_{dI,coef}}{C_{dI,cursor}} \dfrac{C_{dI,casc}}{\dfrac{\alpha}{\sqrt{2} f_s n_\tau V^*} - (C_{sI,casc} + C_{dI,boost})} \right)} \tag{2.27}$$

where

$$U = 2 \cdot k_{ISI,max} \frac{v_{in}}{V^*} + \frac{V^* (C_{sI,casc} + C_{dI,cursor}) + 2 \cdot k_{ISI,max} V_{in} C_{sI,casc} - \dfrac{2\alpha (1 - v_{in}/V^*)}{2\sqrt{2} f_s n_\tau}}{\dfrac{\alpha}{\sqrt{2} f_s n_\tau} - V^* (C_{dI,boost} + C_{sI,casc})} \tag{2.28}$$

Since both I_{boost} and the tap currents (equation 2.15) are proportional to I_{cursor}, the total power consumption, $P_{I,res,casc}$ can be expressed as:

$$P(I, res, casc) \propto \frac{I_{nom,casc} \cdot V_{dd}}{1 - \dfrac{GBW}{\omega_T} \cdot \gamma \cdot \left\{ 1 + U + N_{coef} \cdot k \cdot \dfrac{v_{in}}{V^*} \cdot \dfrac{2 C_{dI,coef}}{C_{dI,cursor}} \left(\dfrac{2\gamma}{G} \cdot \dfrac{GBW}{\omega_T} \right) \right\}} \tag{2.29}$$
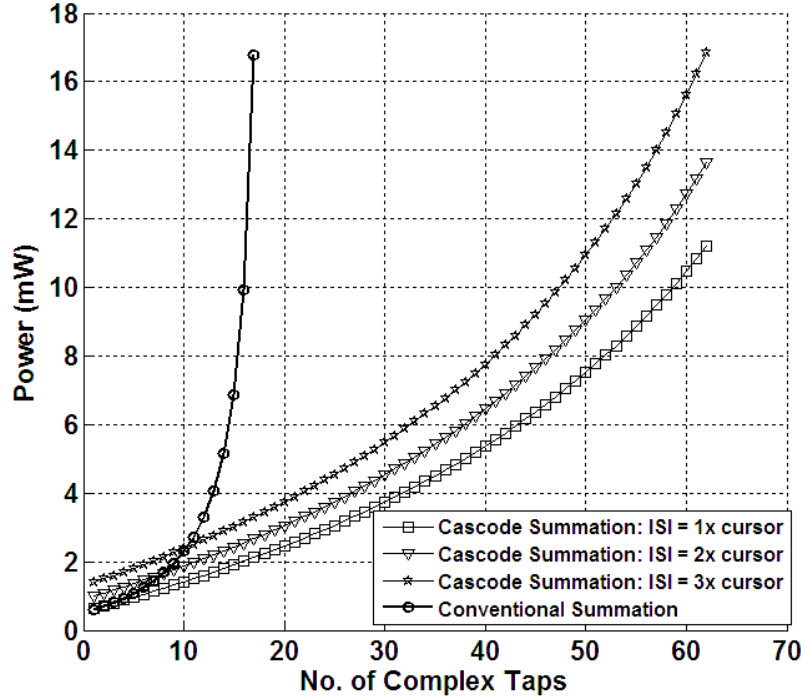
Figure 2.7: Summing amplifier power vs. no. of complex taps for conventional and cascode-summing structures (10Gb/s QPSK).

Similar to the power consumption of the conventional DFE in (2.12), $GBW = G \cdot \left(\frac{\sqrt{2} n_\tau f_s}{\alpha} \right)$ 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 $\frac{G}{2\gamma} \cdot \frac{\omega_T}{GBW}$ times slower with N_{coef}. 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 coefficient switches 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 2.7. While (2.29) contains an additional fixed self-loading term proportional to $k_{ISI,max}$ in U, for a large number of coefficients, the self-loading is dominated by the term proportional to $k \cdot N_{coef}$. From the 60 GHz channel models shown in Chapter 1, it can be concluded that for typical TX-RX separations (3-5 m) and moderately directional RF frontends, k and $k_{ISI,max}$ are 0.5 and 2 respectively. Under these conditions, the cascode summer can support 3-5 times more taps than conventional summation at equal power and data-rate. Importantly, as long as $k_{ISI,max}$ is substantially smaller than $k \cdot N_{coef}$,

the cascode summer power is only weakly sensitive to $k_{ISI,max}$, as shown in Fig. 2.7.

2.3 40-coefficient, 10Gb/s Cascode-Summing DFE Prototype

To prove the efficacy of the proposed cascode-summing DFE architecture, this section describes a 65nm CMOS prototype for 60GHz LOS communication with data-rates up to 10Gb/s using QPSK modulation. Since this test-chip was targeted towards relatively directional LOS channels, the DFE was implemented with 20 complex taps (i.e. 40 I/Q coefficients). To preclude the need for relatively high power consuming error-control-coding mechanisms such as LDPC [44], the design targeted an uncoded bit-error-rate (BER) of $< 10^{-12}$. The discussion in this section includes key design issues, the critical circuit blocks and finally measured results achieved by this prototype.

2.3.1 Key Design Issues

While the cascode current summing structure can significantly extend the number of post-cursor ISI cancelation taps, key design issues must be addressed in order to implement data-rates of 5GS/s at a BER of $< 10^{-12}$. Firstly, the ISI coefficient DACs must be provided with enough resolution to ensure that the resultant quantization noise is less than the thermal noise. The resolution must also take into account the effect of dithering from adapting these coefficients. Finally, the imperfect switching of the current steering DACs creates undesired self-inflicted ISI, which can fortunately be corrected by adapting the latter taps.

2.3.1.1 Coefficient-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 requirements 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) DACs in the DFE can potentially accrue sizeable quantization noise. To ensure that the quantization noise of 40 coefficient DACs is less than the thermal noise, each tap DAC requires 7 bits of resolution[3]. Due to matching limitations, such a high resolution necessitates a large DAC size of $> 50\mu$m 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

---

[3]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.
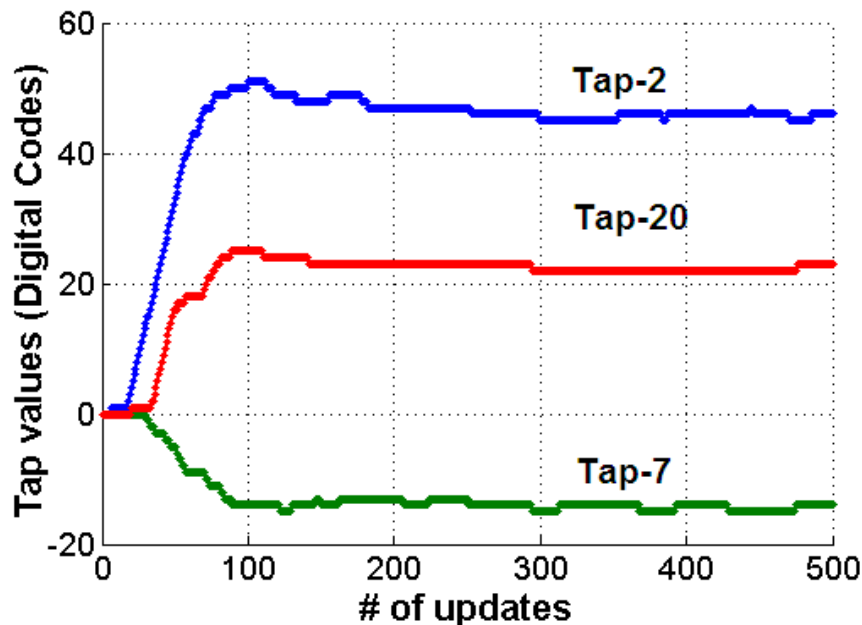
Figure 2.8: DFE tap value adaptation and steady-state dithering for I-to-I taps 2, 7, and 20

distance away[4] (as shown later in Fig. 2.17). 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 $\sim$100MHz). As will be discussed later in the circuit design subsection, this low frequency pole necessitates the use of low-swing drivers for the current steering switches of the taps.

2.3.1.2   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. 2.8). 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.

---

[4]It should be noted that the DACs themselves are outside the high-speed feedback path.
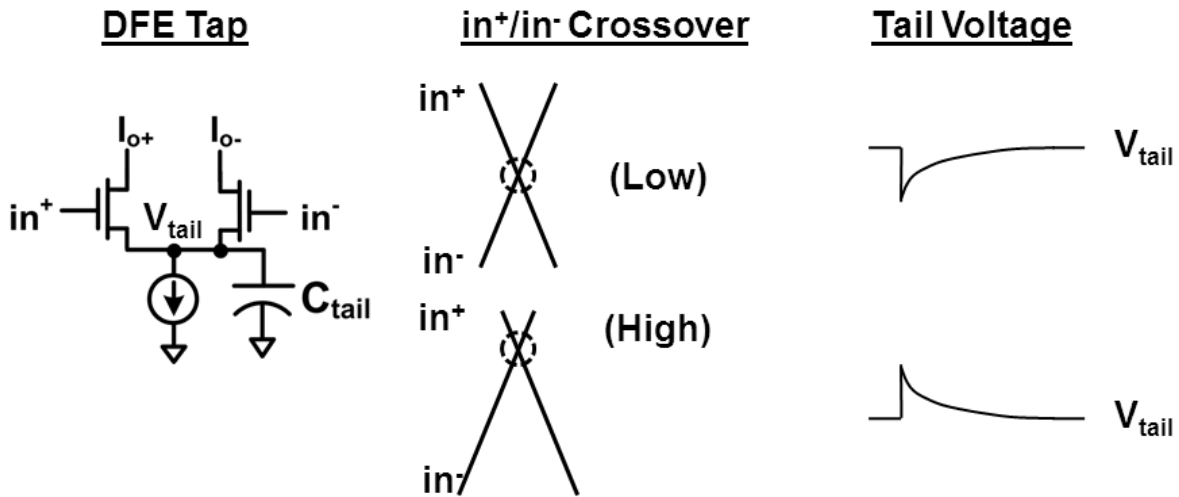
Figure 2.9: Infinite impulse response (IIR) effects due to tap switching. The crossover points of in+/in- are exaggerated to highlight the imbalance in voltage.

2.3.1.3 IIR Effects

While the current-steering tap switches are driven differentially by the low-swing driver, imbalances are inevitable whenever whenever the gates of these switches are driven to steer current from one side to the other. As seen in Fig. 2.9, if the crossover point of the differential gate-drive signals is too high (low), 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) filter. 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 roughly 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. 2.10), 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, these undesired components are absorbed into the taps (as shown on the right of Fig. 2.10). 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
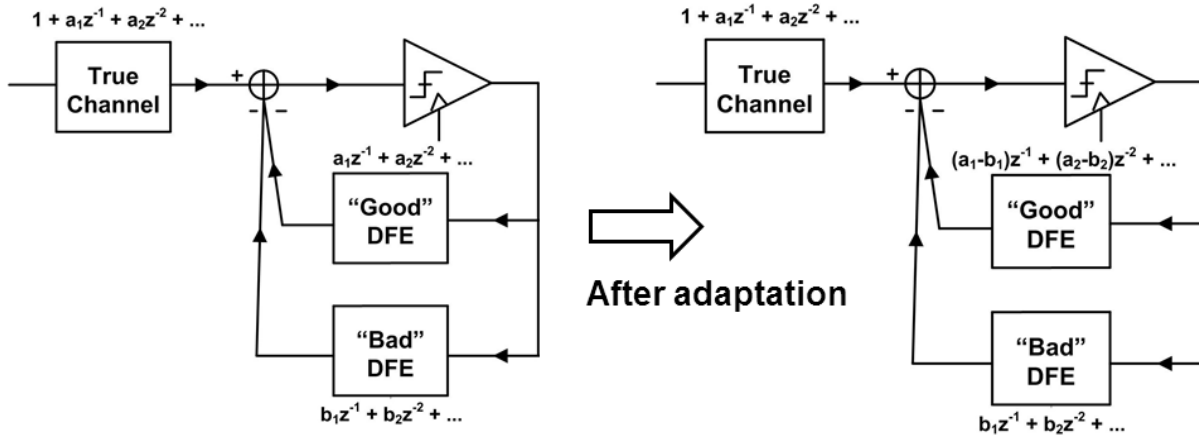
Figure 2.10: Modeling IIR effects: (left) before and (right) after correction by adaptation

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.

2.3.2 Key Circuit Blocks

Following the design issues addressed in the previous sub-section, this sub-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.

2.3.2.1 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 [45][46] 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. Importantly, loop unrolling increases the feedback delay for the latter taps (especially the 2nd tap). Therefore, as opposed to eliminating the feedback delay constraint, this often results in merely shifting the burden of the critical path. Finally, loop unrolling increases the complexity of clock and data recovery (CDR) due to the need for filtering edge updates [47] (which also reduces the CDR bandwidth).
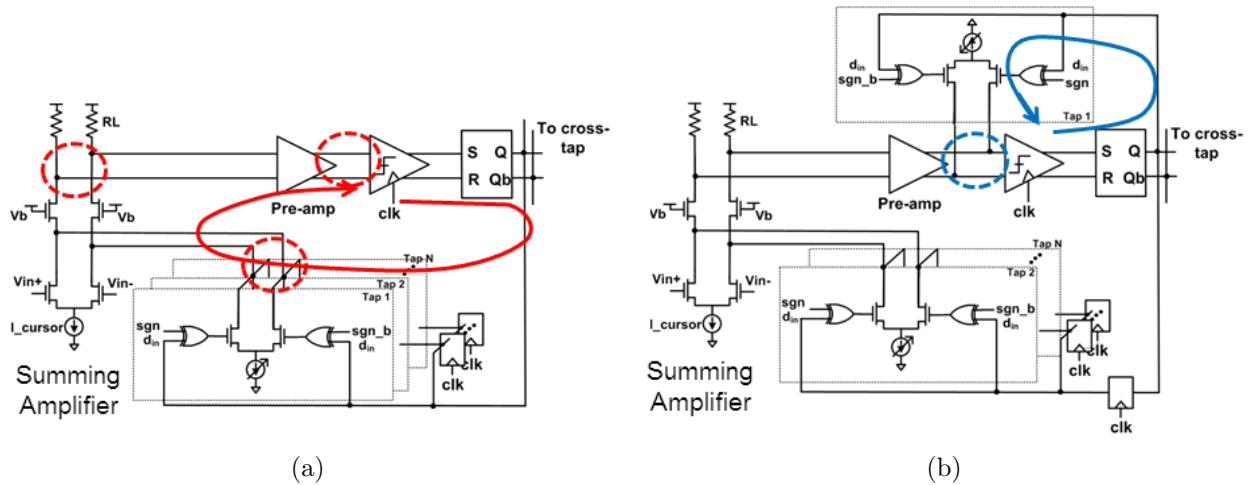
Figure 2.11: Tap-1 feedback (a) All taps summed together: slow settling; (b) Tap-1 summed at comparator input: fast settling.

Fig. 2.11a 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 the summing amplifier altogether (similar to [48]). Using this technique, the analog portion of the settling delay for tap-1 involves only a single pole (as shown in Fig. 2.11b) of the preamp (which typically has much higher bandwidth than the summing amplifier[5]).

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.

2.3.2.2 Low-Swing Drivers

As explained in the previous sub-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

---

[5]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.
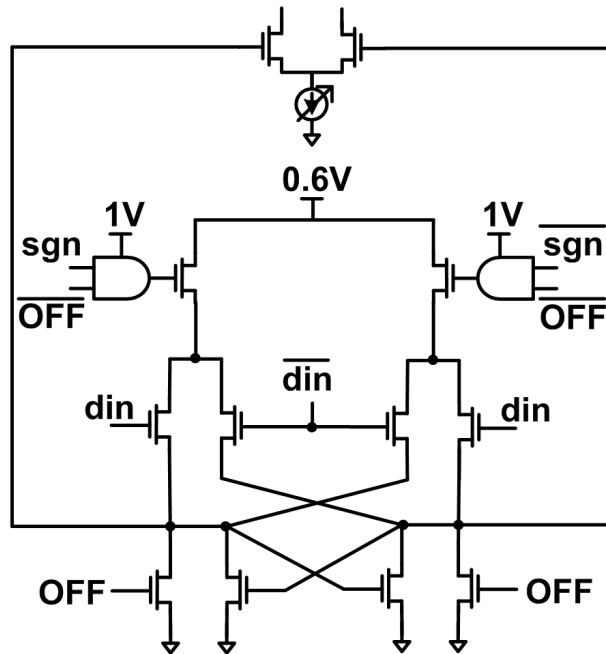
Figure 2.12: Low-swing drivers with embedded XOR for current steering switches

cascade 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. 2.12 shows the design for these drivers, which operate from a 0.6V supply. The driver inputs are full-swing (1V) differential digital signals (din and $\overline{din}$) from the feedback shift registers. In addition to implementing XOR functionality for sign selection (using sgn and $\overline{sgn}$), the drivers also completely turn off the tap when required (for $OFF = 1$, $\overline{OFF} = 0$). Since the top-most NMOS transistors of the driver are fed by static signals ($sgn \cdot \overline{OFF}$ and $\overline{sgn} \cdot \overline{OFF}$)[6], they are typically sized larger than the other transistors to reduce the driver delay without incurring a power penalty.

2.3.3   Simulations, Test-Chip and Measurements

Fig. 2.13 shows the schematics of the prototype complex cascode current-summing DFE. To enable tap adaptation using sign-sign LMS and an edge-based CDR [47], 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 in this test-chip necessitated dedicated shift registers for direct and cross feedback to both I and Q channels. This requirement doubled 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

---

[6]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.
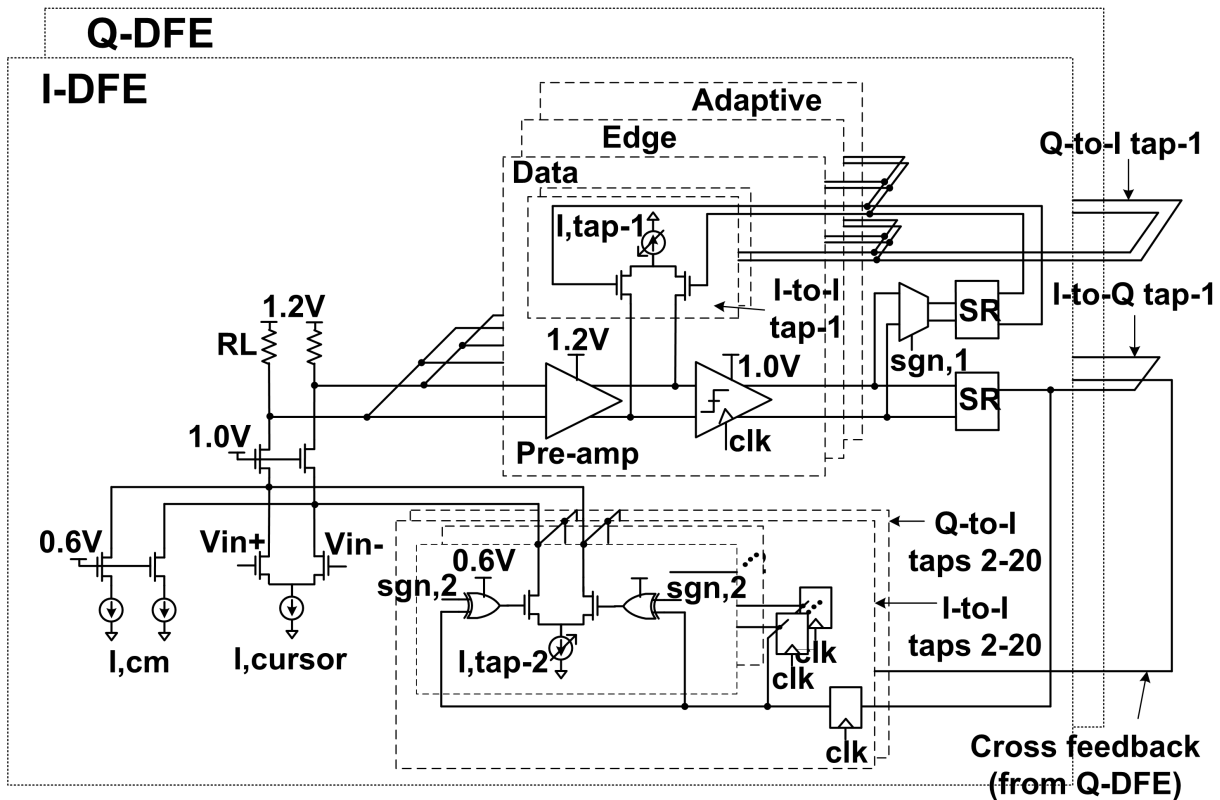
Figure 2.13: 20-complex-tap cascode current-summing DFE prototype

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. This power breakdown (Fig. 2.14) once again illustrates the power penalty of digital gates at GS/s rates.

Fig. 2.15 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 this 20-complex tap prototype DFE supporting 10Gb/s (5GS/s) QPSK [49]. 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. 2.16 shows the block diagram of the entire baseband.

The chip measured 1.7mm x 1.1mm (Fig. 2.17), 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
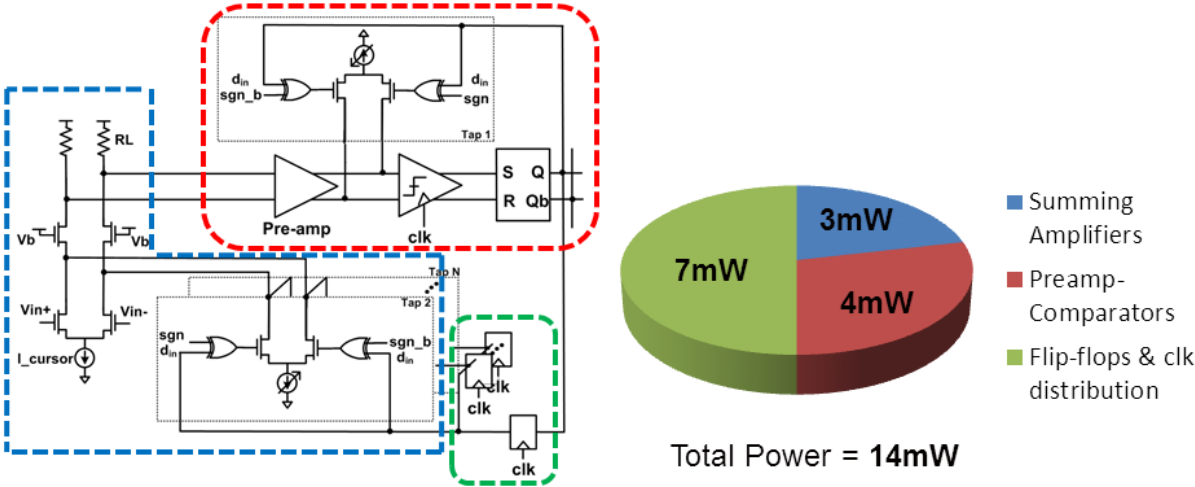
Figure 2.14: Cascode-summation DFE power breakdown at 10Gb/s operation
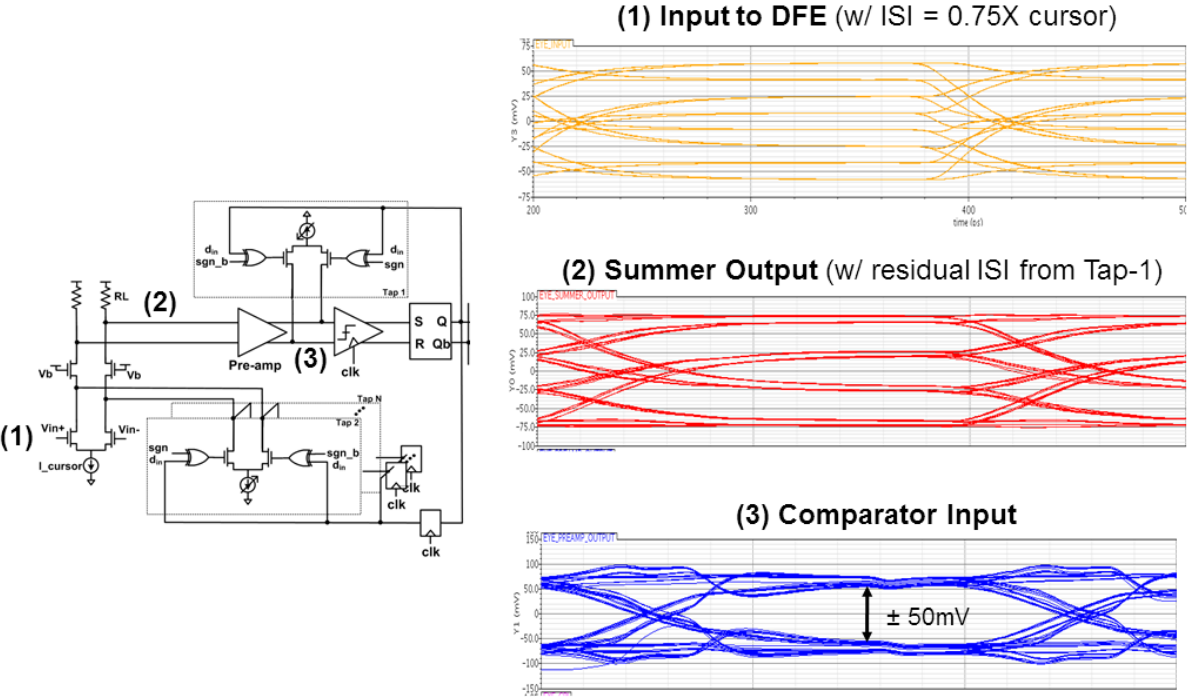
Figure 2.15: Post-layout simulations of eye diagrams with 5GS/s PRBS-7 input
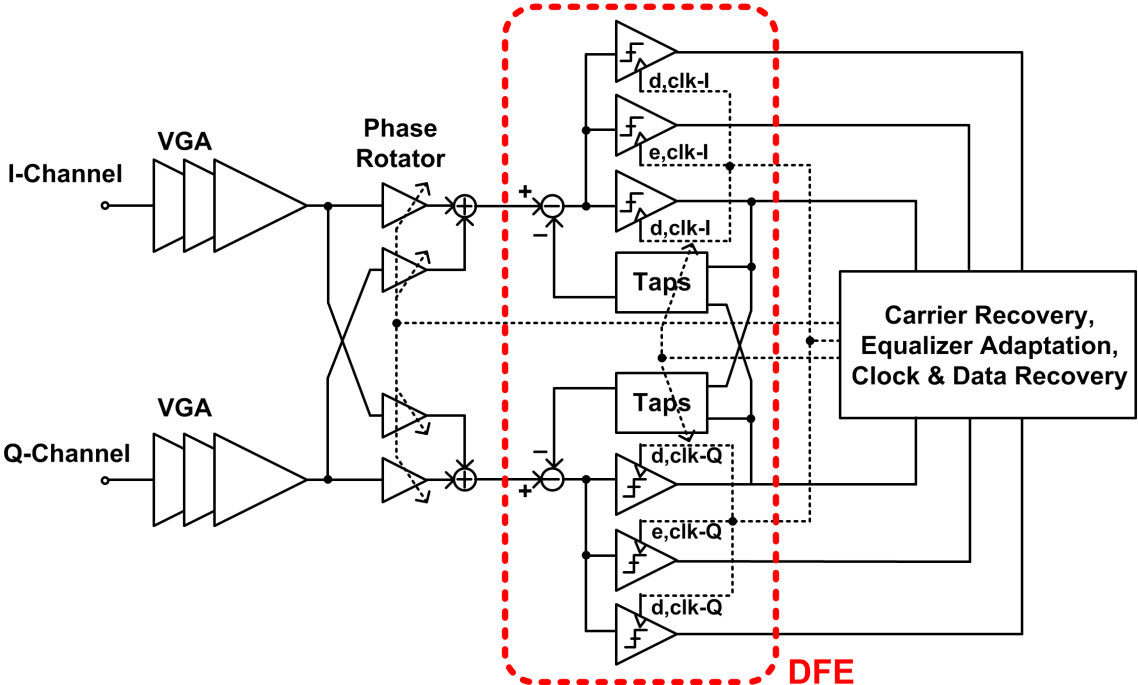
Figure 2.16: Block diagram of 60GHz baseband test-chip with prototype I/Q DFE
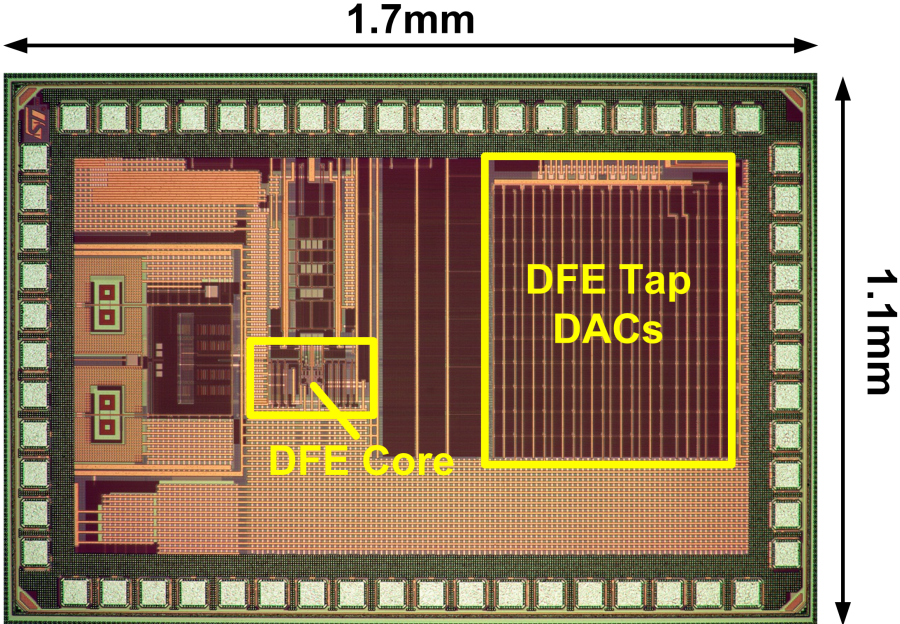
Figure 2.17: 60GHz baseband die micrograph of 65nm CMOS prototype, with DFE overlaid
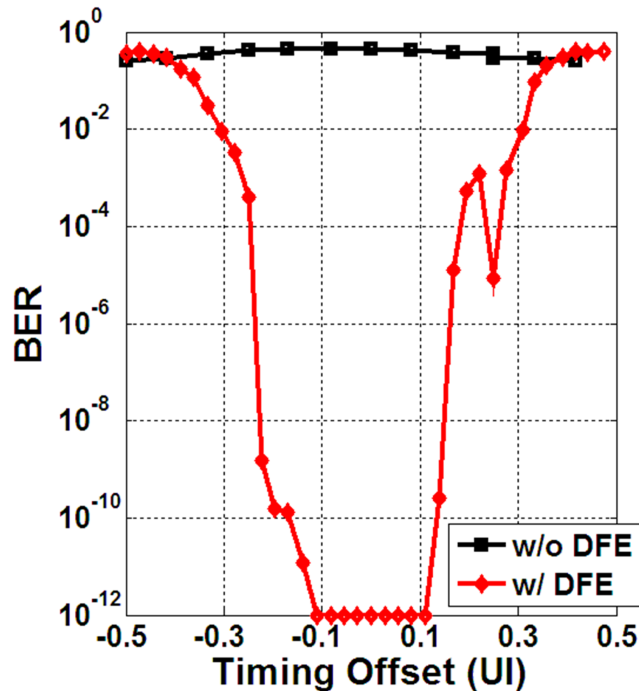
Figure 2.18: BER vs. timing offset (UI) with and without the DFE turned on.

$2^9 - 1$ PRBS data on the I and Q channels respectively. To test the operation of all taps, the ISI was initially distributed randomly from taps 1-20 (both direct and cross feedback) over different measurements. 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 an on-chip PRBS checker. Fig. 2.18 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.

2.4 Conclusion

In this chapter, a design methodology was 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 communication over a channel with long post-cursor delay profile, such as the multi-GS/s 60GHz channel.

A cascode current-summing structure was then 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. The proposed design is validated by a prototype cascode current-summing DFE in 65nm CMOS with 20 complex post-cursor ISI taps. The prototype was shown to operate up to data-rates of 10Gb/s for BER less than 10^{-12} while consuming only 14mW of power. The power efficiency of the proposed DFE with many coefficients is the first major step towards easing the power bottleneck of equalizing LOS channels with wide delay spread.

Chapter 3

Alternate DFE Summation Architectures

In the previous chapter, we discussed a mixed-signal DFE design using resistively loaded cascode-based current summation. While this technique enabled the implementation of many coefficients suitable for a directional LOS wireless channel, the equalization of an NLOS wireless channel requires an even larger number of DFE coefficients and FFE coefficients. For achieve this, in this chapter, we shall examine several alternative summing architectures developed in the past few years that offer improvements over classic resistively loaded summing. These architectures have proposed the use of techniques such as current integration [50], switched capacitor based voltage summation [51], or a combination of both [52].

All proposed techniques claim to offer improvements over resistively loaded current summation. However, since the designs have each been implemented for different technology nodes, data-rates and number of feedback coefficients, a fair way to compare their efficacy requires an analytical framework. In order to determine which technique is most optimal, this chapter analyzes all architectures to compute the power dissipation of a DFE as a function of data-rate and the number of coefficients (i.e. the channel characteristics).

This chapter concludes by determining which technique can be scaled most efficiently to implement a large number of coefficients. For the latter part of this dissertation, we shall leverage the most optimal summing technique determined in this chapter to enable NLOS equalization.

3.1   Architectures

3.1.1   Current Integration on a Capacitive Load

As shown in the previous chapter, resistively loaded current summation is primarily constrained by self-loading of the coefficients at the fast-settling summation node. To relax this
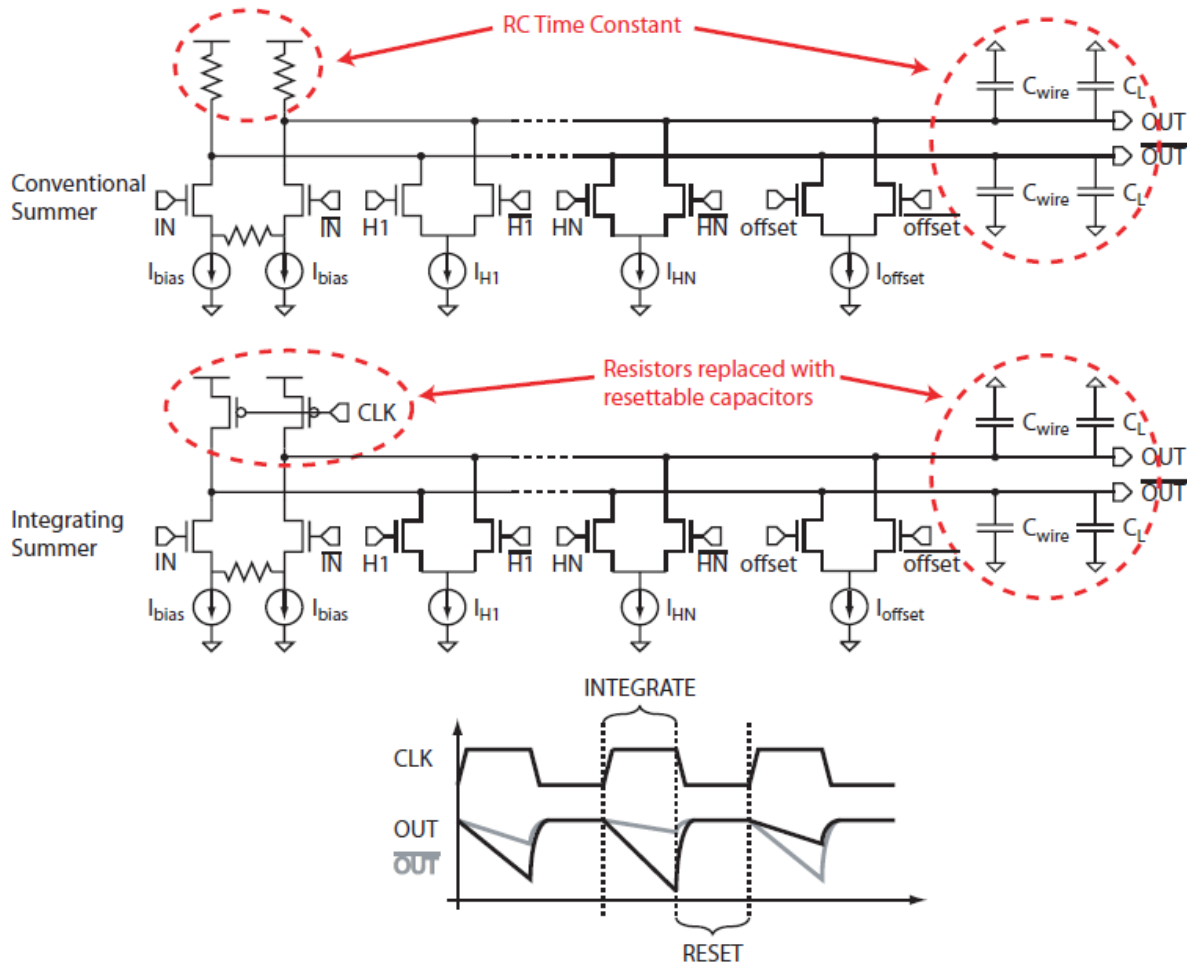
Figure 3.1: A 2-tap current-integrating DFE and its comparison with a conventional resistively loaded DFE [50].

settling time constraint, an alternative was provided by current integration-based summing – a technique first used by M. Park *et al.* in [50] for a 2-tap DFE and C. Ezekwe, B. Boser in [53] for a switched capacitor integrator-based accumulator. Fig. 3.1 illustrates the operation of the 2-tap current integrating DFE in [50], and compares it with a conventional resistively loaded DFE. The resistors are replaced with reset PMOS switches, thus making the load mostly capacitive when the switches are off. During the 'reset' phase ($CLK=0$), both ends of the differential output are set close to V_{dd}. During the 'integrate' phase ($CLK=1$), the cursor and DFE taps integrate charge off the outputs, thus creating a differential voltage, which then is sampled by the following slicer (not shown explicitly) at the end of the phase.

As will be shown in detail in the analysis section, an integrator offers higher low-frequency gain than a resistively loaded amplifier for equal power consumption. Therefore, replacing an RC-load by a C-only load makes the summing more power-efficient. If the input of the cursor or DFE taps were to change during the integration window, it would create a systematic
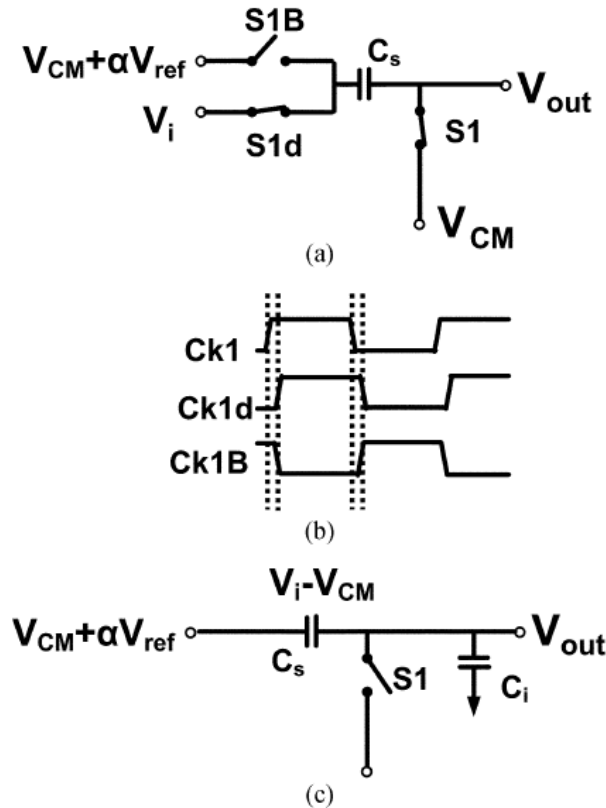
Figure 3.2: Switched-capacitor summation [51]: (a) Front-end design. (b) Clocking scheme. (c) Equivalent circuit in the equalization phase.

frequency-dependent loss. Since an integrator offers highest gain at DC, this loss can be prevented by using a S/H at the inputs $IN, \overline{IN}$ [54]. Since the taps are driven by digital shift registers whose outputs are typically settled before integration starts, the tap inputs are 'DC' by default.

3.1.2 Switched Capacitor-based Voltage Summing

The DFE summation techniques discussed so far have all been in the current domain. A. Emami-Neyestanak *et al.* instead proposed a voltage-mode summing technique by using switched capacitors in [51]. Fig. 3.2 shows the schematics and clocking waveforms of this single-tap[1] switched-capacitor DFE implementation. The tap summation is achieved by subtracting tap feedback voltage αV_{ref} from input voltage V_i. The voltage difference $(V_i - \alpha \cdot V_{ref})$ is realized on the tap capacitor C_s by the clocking waveforms $Ck1$, $Ck1d$ and $Ck1b$ as shown in Fig. 3.2 (b). The tap weight is set by adjusting the weighting factor α on the reference voltage V_{ref}. As compared to resistive current-summation and current integra-

---

[1]The DFE in [51] has two taps, but only tap-2 is implemented by a switched capacitor. Tap-1 is unrolled.

tion, the advantage of capacitive summation is that its settling time is only of the order of the rise/fall time of the switch drivers. The benefit of reduced settling time eases the speed (and therefore power[2]) of the feedback shift register flip-flops.

3.1.3 Combination of Current Integration and Switched Capacitor Summing

T. Toifl *et al.* introduced a hybrid technique using both current integration and switched capacitor summation in an 8-tap DFE [52]. The cursor gain is provided by current integration, while the tap summation is implemented with voltage-mode capacitive feedback. In contrast to the switched capacitor DFE in [51] which uses a voltage DAC for tap weighting, however, this architecture uses a capacitor DAC for each tap. Fig. 3.3 shows the schematics and equalizing waveforms. While this architecture can be scaled to many coefficients, the resolution of the capacitor DAC directly impacts capacitive loading on the high-speed summer output.

3.2 Analysis

This section provides an analytical framework that predicts the power as a function of the number of taps for the current integrating [50] (sub-section 3.1.1) and hybrid switched capacitor – current integrating DFE [52] (sub-section 3.1.3) architectures, similar to the one for conventional and cascode-based current summing DFEs in Chapter 2. While the voltage-mode only switched capacitor DFE [51] technique (sub-section 3.1.2) is not explicitly discussed, it will be analyzed as a special case of the hybrid switched capacitor current integrating DFE.

3.2.1 Current Integration on a Capacitive Load

Before beginning to mathematically analyze the power consumption of a capacitively-loaded current integrating DFE-summer, it is worthwhile to intuitively understand how an integrator-based amplifier is intuitively different than a resistively-loaded amplifier. Consider two common-source amplifiers with identical bias currents and gm-transistors. One of the amplifiers is loaded with a resistor R_L and the other loaded with a precharge PMOS transistor, as shown in Fig. 3.4. Assume that R_L is less than the output resistance of the common-source transistor, r_o. The output capacitance to be driven, C_L is large enough that it dominates the intrinsic capacitance of the amplifiers at the output node.

The gain-bandwidth plots of the two circuits are conceptually shown in Fig. 3.5. To address the variation in r_o across different technology nodes, the figure also shows the vari-

---

[2]This benefit is assuming that the feedback shift register flip-flops and sign-select drivers were originally bigger than minimum size in order to meet the feedback delay constraints for resistive summation or current integration.
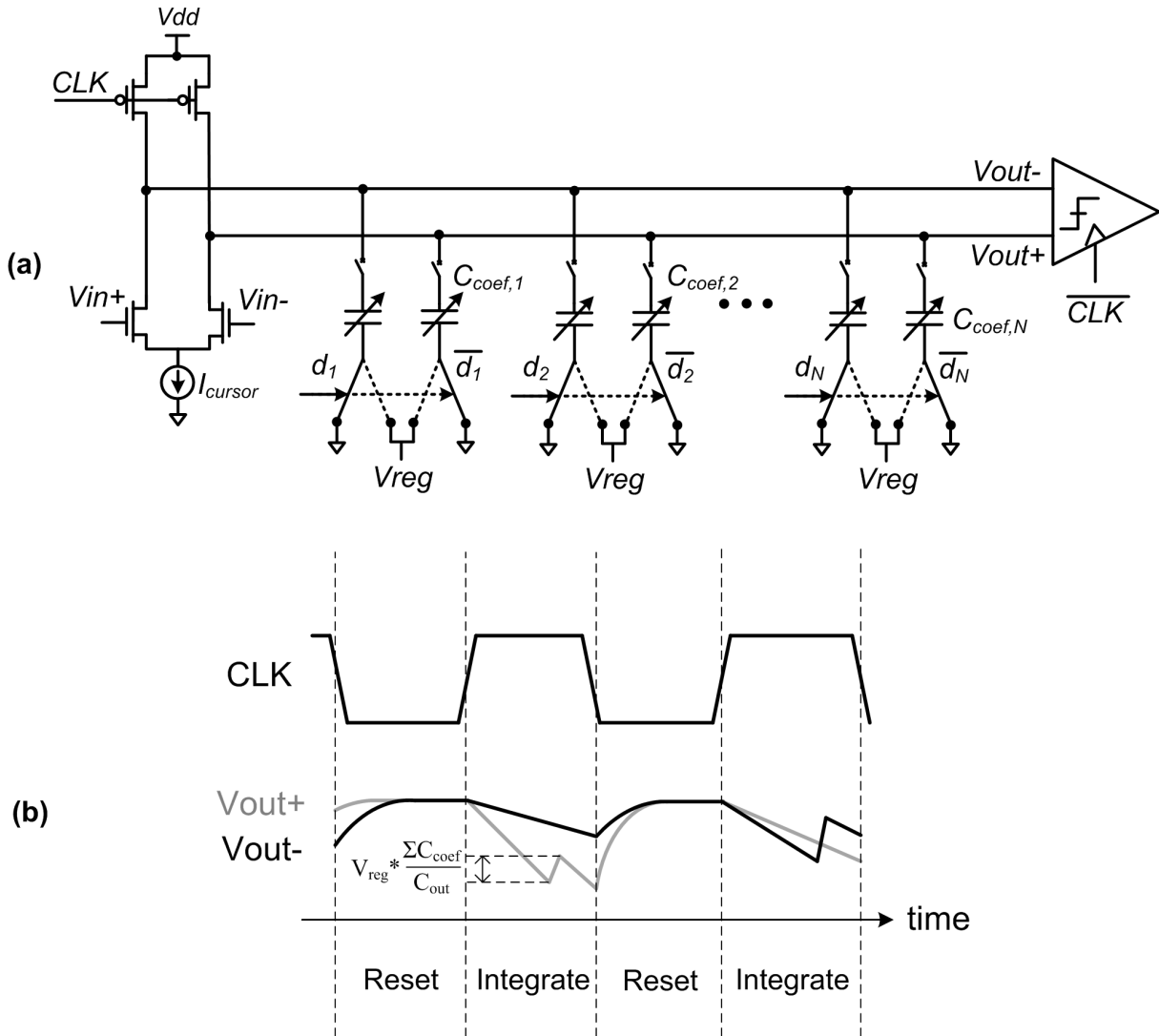
Figure 3.3: Combination of current integration and switched capacitor summation [52]: (a) Schematics. (b) Clocking and waveforms.

ation in the gain-bandwidth curves with r_o. For equal current consumption, the integrating structure achieves a gain higher than or equal to that of the resistively loaded amplifier across all frequencies. If the input signal is a sampled DC value, the integrator will always achieve higher gain. Since the current consumption of a class-A amplifier is proportional to gain, conversely, to achieve equal DC gain as its resistively loaded counterpart, the integrator would consume less current.

Along with its higher DC gain, a standalone integrator with low bandwidth would clearly be prone to self-inflicted ISI. To avoid this problem, the integrator is provided with a fast reset between two integration periods by using precharge PMOS transistors. Thus, for sampled
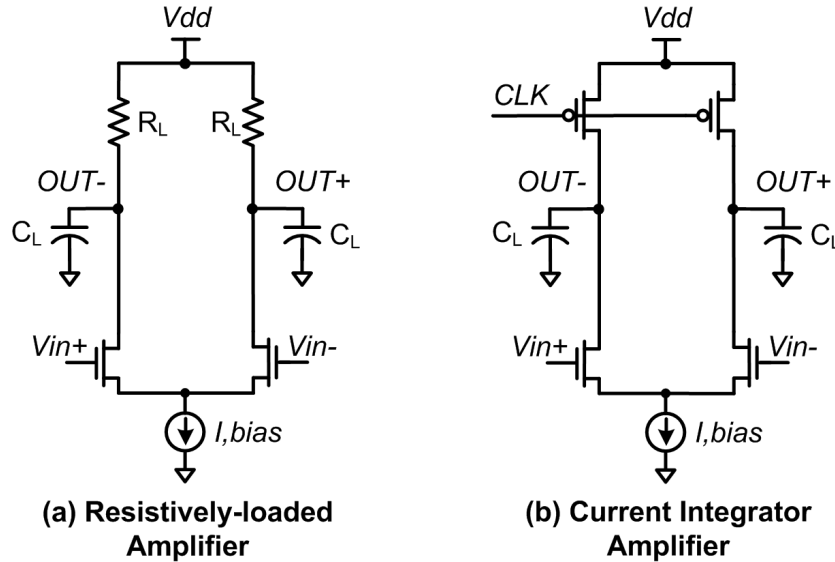
Figure 3.4: Resistively-loaded and current-integrating amplifiers

inputs the integrator-based summer with reset capability is more power-efficient than the resistively loaded summer.

In addition to improved cursor gain, it is also necessary to understand through a complete mathematical analysis the impact of current integration on coefficient self-loading and hence power as a function of the number of coefficients. Fig. 3.6 shows a capacitively loaded current-integrating mixed-signal DFE summing amplifier structure, and its small-signal model. Sim-
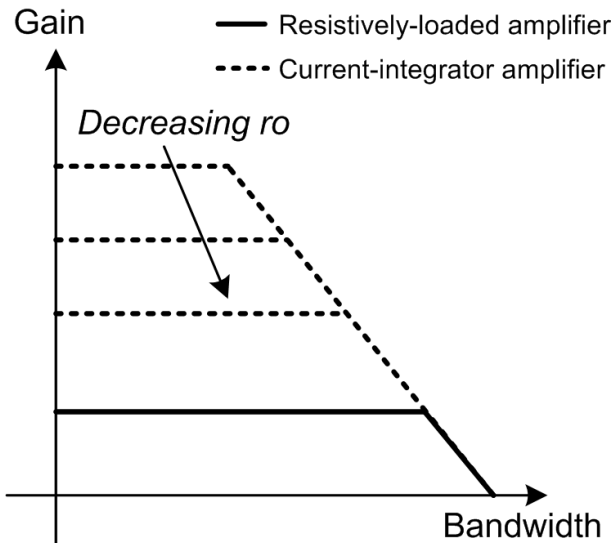
Figure 3.5: Resistively-loaded and current-integrating amplifiers: Gain vs. Bandwidth. For equal current consumption, current integration provides higher DC gain.
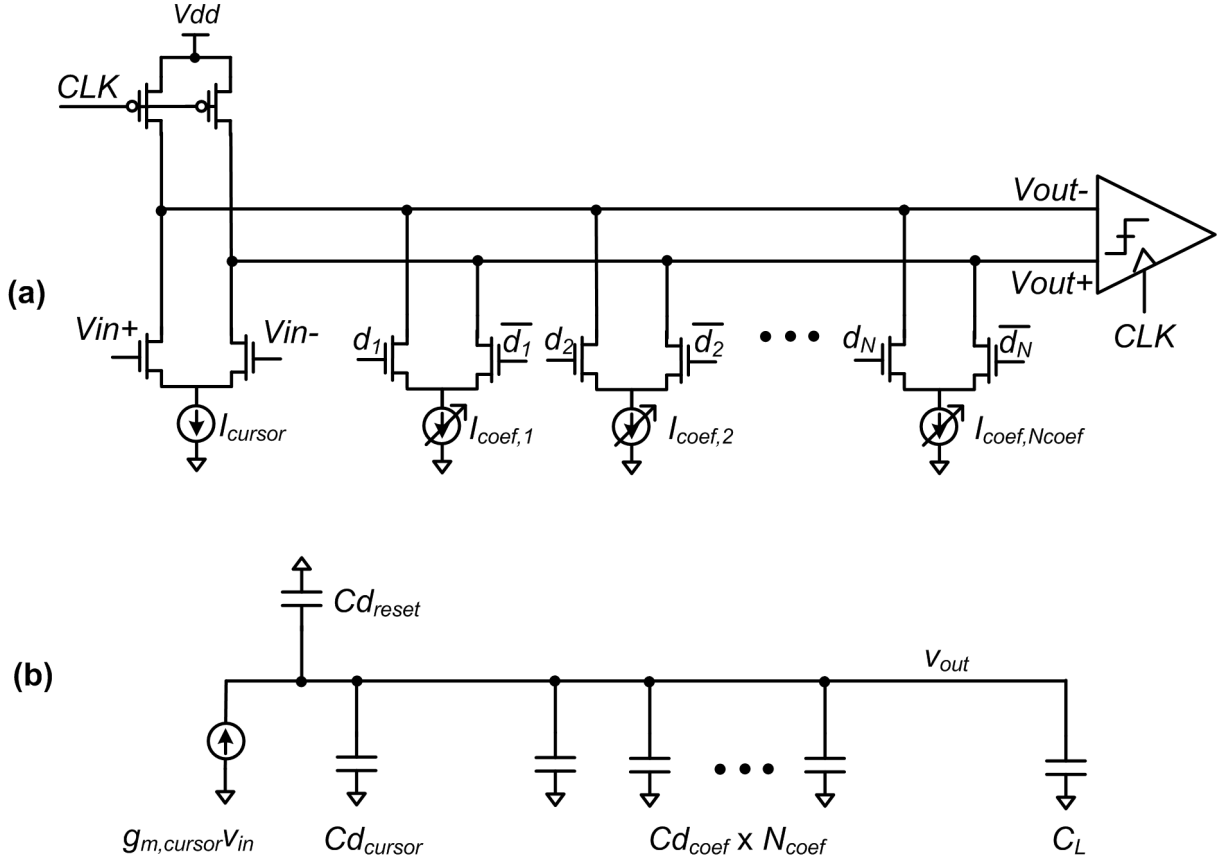
Figure 3.6: Gain/bandwidth analysis of a current-integrating DFE [50]. (a) Circuit (top) and (b) Single-ended small-signal model (bottom)

ilar to the resistively loaded current summing DFE, the notations v_{in}, G, f_s, N_{coef}, and k respectively stand for the cursor input amplitude (excluding ISI), the DC gain, the data-rate (symbols/sec), the number of DFE coefficients and the maximum per-coefficient magnitude (relative to cursor-only amplitude). Similar to resistively loaded current-summing, coefficient current is given by:

$$I_{coef} = k \cdot g_{m,cursor} \cdot v_{in}$$
$$= \left(k \cdot \frac{v_{in}}{V^*_{cursor}} \right) \cdot I_{cursor} \tag{3.1}$$

The total capacitance, $C_{T,out}$, at the output node is

$$C_{T,out} = C_L + Cd_{cursor} + N_{coef} \cdot Cd_{coef} + Cd_{reset} \tag{3.2}$$

where Cd denotes drain capacitance and suffixes *cursor* and *DFE* denote the cursor, DFE switch and reset transistors respectively. During integration, since the reset transistor is off,

the capacitance at its drain terminal comprises of only diffusion capacitance. Assuming a 50% clock duty cycle, the reset transistor should be sized to be able to precharge the output well within a half clock cycle. If the digital delay of the precharge is a fraction α_{reset} of this half cycle (such that $\alpha_{reset} < 1$), the delay can be expressed in terms of the equivalent switch resistance of the reset transistor R_{reset} as follows

$$\alpha_{reset} \cdot \frac{T}{2} = R_{reset} \cdot C_{T,out}$$

$$i.e., \quad \frac{\alpha_{reset}}{2 \cdot f_s} = R_{reset} \cdot Cd_{reset} \cdot \left(1 + \frac{C_L + Cd_{cursor} + N_{coef} \cdot Cd_{coef}}{Cd_{reset}} \right) \tag{3.3}$$

To facilitate a comparison of the current integrating DFE with the digitally driven switched capacitor-based architecture (in the following sub-section), we may compute precharge delay by using fanout-of-4 digital delay T_{FO4} of the technology as an indicator of digital gate delay. By then assuming that equal PMOS and NMOS resistance requires a 2:1 transistor width ratio, the self-loading delay of the reset transistor will be

$$R_{reset} \cdot Cd_{reset} = \frac{2}{3} \cdot \frac{\gamma \cdot T_{FO4}}{\gamma + 4} \tag{3.4}$$

where γ is the drain-to-gate capacitance ratio. Using this value of $R_{reset} \cdot C_{reset}$ in (3.3) gives:

$$Cd_{reset} = \frac{1}{\dfrac{3}{4} \cdot \dfrac{\alpha_{reset} \cdot (\gamma + 4)}{f_s \cdot \gamma \cdot T_{FO4}} - 1} \cdot (C_L + Cd_{cursor} + N_{coef} \cdot Cd_{coef}) \tag{3.5}$$

Let us express the reset transistor drain capacitance Cd_{reset} in terms of the sum of all other capacitors at the summing node in terms of a factor, k_{reset}.

$$Cd_{reset} = k_{reset} \cdot (C_L + Cd_{cursor} + N_{coef} \cdot Cd_{coef}) \tag{3.6}$$

Intuitively, k_{reset} is the factor by which the capacitance of any node needs to be increased in order to include PMOS reset capability. k_{reset} may therefore be termed as the 'reset factor'. Comparing (3.5) and (3.6), it is easy to see that

$$k_{reset} = \frac{1}{\dfrac{3}{4} \cdot \dfrac{\alpha_{reset} \cdot (\gamma + 4)}{n_\tau \cdot f_s \cdot \gamma \cdot T_{FO4}} - 1} \tag{3.7}$$

By replacing Cd_{reset} in terms of the other capacitances by k_{reset}, the total summing node capacitance $C_{T,out}$ from equation (3.2) can be expressed as

$$C_{T,out} = (1 + k_{reset}) \cdot (C_L + Cd_{cursor} + N_{coef} \cdot Cd_{coef}) \tag{3.8}$$

Similar to (2.5), the cursor and DFE coefficient capacitances can be expressed in terms of their respective currents and capacitance per unit current (C_{dI}) as

$$Cd_{cursor} = C_{dI,cursor} \cdot \frac{I_{cursor}}{2}$$

$$Cd_{coef} = C_{dI,coef} \cdot \frac{k \cdot v_{in}}{V_{cursor}^*} \cdot I_{cursor} \tag{3.9}$$

By plugging this relation in (3.8), $C_{T,out}$ is expressed in terms of I_{cursor} as

$$C_{T,out} = (1 + k_{reset}) \cdot \left\{ C_L + \left(\frac{1}{2} \cdot C_{dI,cursor} + \frac{k \cdot v_{in}}{V_{cursor}^*} \cdot C_{dI,coef} \cdot N_{coef} \right) \cdot I_{cursor} \right\} \tag{3.10}$$

Now, for the current integration operation, the small-signal output voltage can be computed as

$$v_{out} = \frac{i_{cursor} \cdot T_{int}}{C_{T,out}} \tag{3.11}$$

Since the integration is for a half-cycle duration ($T_{int} = T/2 = 1/2f_s$),

$$v_{out} = \frac{g_{m,cursor} \cdot v_{in}}{2 \cdot f_s \cdot C_{T,out}} \tag{3.12}$$

Therefore, the DFE integrator gain can be computed as

$$\frac{v_{out}}{v_{in}} = G = \frac{I_{cursor}}{2 \cdot f_s \cdot C_{T,out} \cdot V_{cursor}^*} \tag{3.13}$$

Substituting for $C_{T,out}$ from (3.10) in (3.13) and simplifying to obtain I_{cursor} gives

$$I_{cursor} = \frac{2 \cdot G \cdot f_s \cdot V_{cursor}^* \cdot C_L \cdot (1 + k_{reset})}{1 - G \cdot f_s \cdot V_{cursor}^* \cdot C_{dI,cursor} \cdot (1 + k_{reset}) \cdot \left(1 + 2 \cdot N_{coef} \cdot k \cdot \frac{v_{in}}{V_{cursor}^*} \cdot \frac{C_{dI,coef}}{C_{dI,cursor}} \right)} \tag{3.14}$$

To make the relation more intuitive and similar to the form of the resistively loaded current summing DFE in (2.11), C_{dI} may be expressed terms of ω_T from (2.10). To be able to compare the current consumption to conventional resistive current summation, let us also define the gain-data-rate product GDR and an I_{nom} as

$$GDR = G \cdot f_s$$

$$I_{nom} = C_L \cdot G \cdot f_s \cdot V^* \tag{3.15}$$

$$= C_L \cdot GDR \cdot V^*$$

In terms of GDR, I_{nom} and ω_T, I_{cursor} can be expressed as:

$$I_{cursor} = \frac{2 \left(1 + k_{reset} \right) \cdot I_{nom}}{1 - 2 \left(1 + k_{reset} \right) \left(\gamma \cdot \frac{GDR}{\omega_{T,cursor}} \right) \left(1 + N_{coef} \cdot k \cdot \frac{v_{in}}{V_{coef}^*} \cdot \frac{2 \, \omega_{T,cursor}}{\omega_{T,coef}} \right)} \tag{3.16}$$

Figure 3.7: Current integrating vs. resistively loaded current-summing DFE summer power vs. no. of complex taps at 5GS/s for $k_{max} = 0.5$ in a 65 nm CMOS technology. (1 complex tap = 2 I/Q coefficients).

In comparison, the cursor current for conventional current summation from equation (2.11) is:

$$I_{cursor}\left(conv\right) = \frac{\left(\dfrac{n_\tau}{\alpha}\right)I_{nom}}{1 - \left(\dfrac{n_\tau}{\alpha}\right)\cdot\left(\gamma\cdot\dfrac{GDR}{\omega_{T,cursor}}\right)\left(1 + N_{coef}\cdot k\cdot\dfrac{v_{in}}{V_{coef}^*}\cdot\dfrac{2\ \omega_{T,cursor}}{\omega_{T,coef}}\right)} \qquad (3.17)$$

In the above equation, the gain-bandwidth product (GBW) for conventional current summation from (2.11) is replaced by the gain-data-rate product (GDR) as:

$$\begin{aligned}
GBW &= G\cdot BW = G\cdot\left(\frac{n_\tau}{\alpha}\cdot f_s\right) \\
GBW &= \left(\frac{n_\tau}{\alpha}\right)\cdot(G\cdot f_s) \\
GBW &= \left(\frac{n_\tau}{\alpha}\right)\cdot GDR
\end{aligned} \qquad (3.18)$$

In comparison to I_{cursor} for the resistively loaded current summing DFE in (2.11), the self-loading of the current integrating DFE is lower by a factor of $\dfrac{2\left(1 + k_{reset}\right)}{(n_\tau/\alpha)}$. In 65nm CMOS technology, with 95% settling for the resistively loaded summer $(n_\tau = 3)$ with reasonably
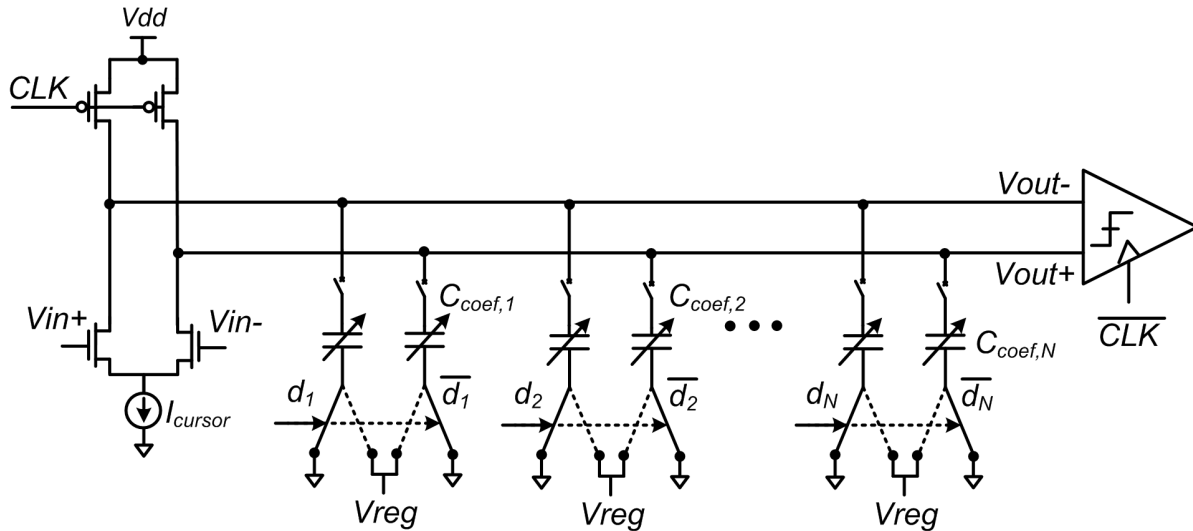
Figure 3.8: Schematics of switched-capacitor feedback-based current-integrating DFE [52]

fast shift register flops and DFE drivers ($\alpha = 0.5$), and ($k_{reset} = 0.1$) for the integrator, the integrator self-loading is reduced by a factor of 3. Since the multiplying factor to I_{nom} is also improved by the same factor of 3, the current integrator exhibits a significant improvement over the resistive summer. The improvements can be visualized by comparing the trends for cursor current consumption versus number of DFE complex taps (1 complex tap = 2 I/Q coefficients) for both in Fig. 3.7.

3.2.2 Combination of Current Integration and Switched Capacitor Summing

Now that a current integrating DFE has been analyzed, we will now consider its hybrid architecture with switched capacitor based coefficient summation [52] from sub-section 3.1.3.

Fig. 3.8 shows the summer schematics. The capacitor DAC used to implement ISI cancelation using capacitive feedback is shown in Fig. 3.9. Each capacitor DAC is driven by voltage V_{reg}. To isolate the unused capacitance of each of the capacitor DAC coefficients from the high-speed summing node, a fraction of each capacitor DAC coefficient is gated by using a PMOS switch per leg (as shown in Fig. 3.9 which has the least 3 significant bits ungated). Ideally, the entire capacitor could have such gating capability. However, for very small LSB capacitor sizes, the gating switch capacitance becomes comparable to DAC leg capacitance, for which, the gating switch becomes an overhead and should be removed. For 5GS/s operation in a 65nm CMOS technology, given the smallest realizable LSB capacitor and switch sizes of $50aF$ (as realized in [18] in 65nm CMOS by using standard MOM capacitors) and $0.12\mu m$ respectively, it is best to leave the smallest 2 LSB segments ungated.

The summer output is always loaded with the ungated proportion p of the switch capaci-
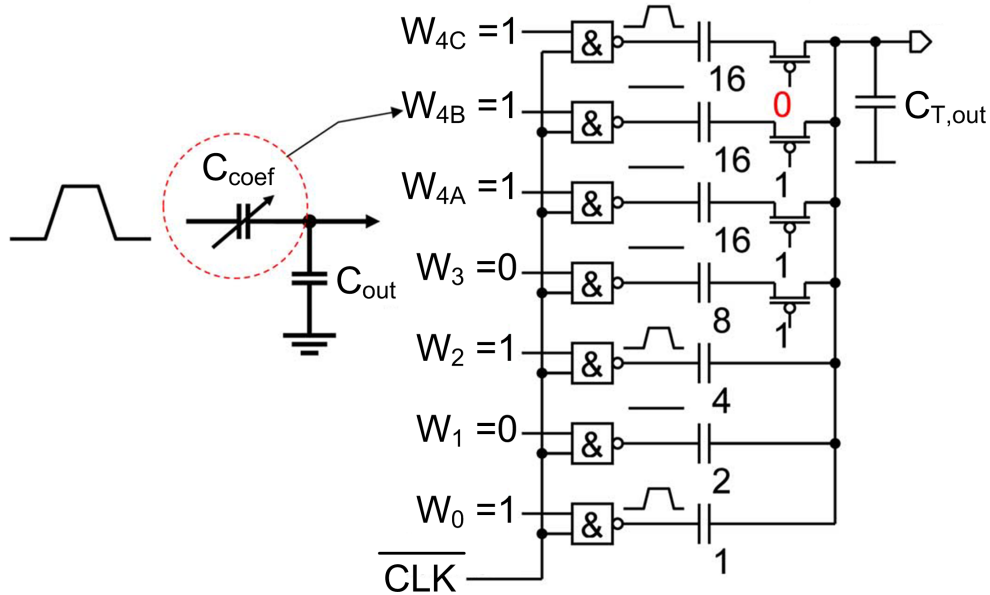
Figure 3.9: Schematics of switched capacitor DAC [52]

tance and the isolation switch drain capacitances ($C_{d,sw}$ per coefficient). Sizing the capacitor driver and isolation switch for delay requires the total isolation switch drain capacitance $C_{d,sw}$ to be approximately proportional to the total coefficient capacitance C_{coef}:

$$C_{d,sw} = \beta \cdot C_{coef} \tag{3.19}$$

As shown in the detailed analysis of the switched capacitor architecture in Appendix A using a digital switch model, for optimal driver sizing to meet a data-rate of f_s, β is determined to be:

$$\beta = \frac{\left(\dfrac{4\gamma}{3}\right) \cdot (1-p)}{\dfrac{\gamma+4}{T_{FO4}} \cdot \dfrac{\alpha}{f_s} - (4\gamma+1)} \tag{3.20}$$

where T_{FO4} is the inverter fanout-of-4 delay and γ is the drain to gate capacitance.

The required coefficient capacitance C_{coef} can be determined as a function of the maximum ISI cancelation voltage $max\,(V_{ISI})$, the coefficient driver voltage V_{reg}, and by the capacitor divider with the total output capacitance $C_{T,out}$.

$$max\,(V_{ISI,i}) = v_{in} \cdot k_{max} = V_{reg} \cdot \frac{C_{coef}}{max\,(C_{T,out})} \tag{3.21}$$

where k_{max} is the maximum magnitude of per-coefficient ISI relative to the cursor magnitude.

Similarly, for the sum of all DFE coefficients to cancel up to a certain total maximum ISI magnitude $k_{ISI,max}$ times the cursor,

$$v_{in} \cdot k_{ISI,max} = V_{reg} \cdot \frac{max \left(\sum_{i=1}^{N_{coef}} C_i \right)}{max \left(C_{T,out} \right)} \tag{3.22}$$

Finally, the current integrator gain G (from equation (3.13)) in terms of cursor current I_{cursor}, gain G and data-rate f_s is:

$$G = \frac{I_{cursor}}{2 \cdot f_s \cdot V_{cursor}^* \cdot max \left(C_{T,out} \right)} \tag{3.23}$$

Using all of the above constraints (equations (3.19)-(3.23)) and following the simplification in Appendix A, I_{cursor} (equation A.24) is:

$$I_{cursor} = \frac{2 \cdot (1 + k_{reset}) \cdot I_{nom}}{1 - 2 \left(1 + k_{reset} \right) \left(\gamma \cdot \frac{GDR}{\omega_{T,cursor}} \right) \left[1 + \frac{k_{max} \, v_{in} \, \omega_{T,cursor}}{\gamma \, V_{reg} \, G \, f_s} \left\{ \frac{k_{ISI,max}}{k_{max}} + (\beta + p) \, N_{coef} \right\} \right]} \tag{3.24}$$

In this equation, as with all previous DFE analyses, $I_{nom} = G \cdot f_s \cdot V_{cursor}^* \cdot C_L$ is the current consumption of a class-A amplifier without self-loading. Additionally, $GDR = G \cdot f_s$ is the gain-data-rate product, k_{reset} is the relative capacitance overhead due to the reset transistor (3.7) and $C_{dI,cursor}$ is the cursor transistor drain capacitance per unit current.

The form of the expression in (3.24) for switched capacitor based current integration looks similar to that for vanilla current integration in (3.16), with a difference in the multiplying factor to the number of coefficients. Since the switched capacitor voltage divider has coefficient driver voltage V_{reg} close to the supply voltage (1V), this helps in relaxing the coefficient capacitor values with respect to the total output loading. This benefit effectively reduced the self-loading. However, minimum-size constraints on the coefficient capacitors and its driver (isolation) gates (switches) limit this benefit. For instance, with the given data-rate and technology, and driving a capacitive load of $10fF$, the total capacitor and switch size per DAC are of the order of $\sim$2.5-5fF and $\sim$1-2μm respectively (depending on the number of coefficients). To achieve quantization noise below the noise floor requires $\sim$5-7 bits of segmentation of these devices (depending on the target BER). Due to a lower limit on practically achievable capacitor and driver/isolation switch values, this segmentation runs into the minimum-size barrier. To achieve the requisite segmentation, therefore, the entire summing circuitry needs to be sized up, thus raising both static and dynamic power consumption.

Fig. 3.10 shows the variation in I_{cursor} with number of complex taps (1 complex tap = 2 I/Q coefficients) for the switched capacitor architecture, and compares it with resistive current summation and current integration. The step-like variation of I_{cursor} is due to the
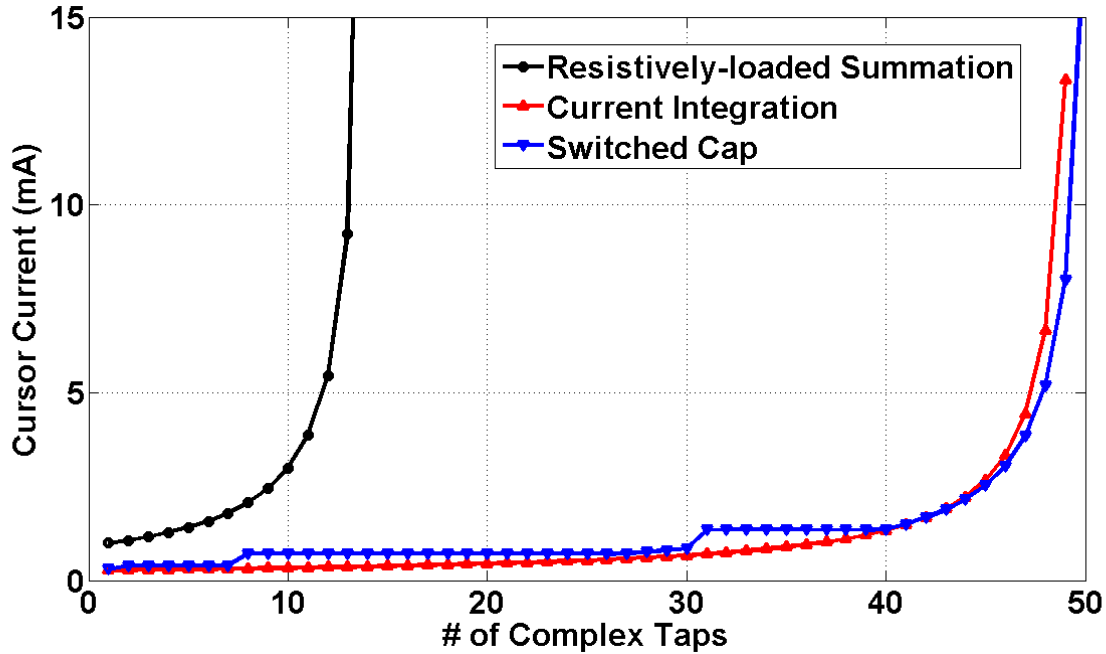
Figure 3.10: Switched capacitor, current-integrating, and resistively loaded current-summing DFE: Summer cursor current vs. no. of complex taps at 5GS/s for $k_{max} = 0.5$, using a 65nm CMOS technology. (1 complex tap = 2 I/Q coefficients).

increase in the number of bits of capacitor DAC-segmentation. As seen in the figure, the static current consumption for the switched capacitor architecture is similar or higher than that for current integration. It is observed that the potential benefit expected by reduced self-loading is overshadowed by the minimum-size constraint of $0.12\mu m$ on the isolation switch (despite not having an explicit switch on the smallest 2 LSB segments).

Apart from the higher static curent dissipation, the switched capacitor DFE consumes significant digital clocking power in the NAND drivers for the capacitors (Fig. 3.9 earlier showed the drivers with in-built sign-select). After accounting for this dynamic power consumption, the total power consumption becomes significantly worse, as shown in Fig. 3.11. Intuitively, this is because of a relatively high NAND input capacitance per coefficient capacitance being driven with rail-to-rail signals every clock cycle. Since the NAND gate is sized proportionally to the coefficient capacitance, the dynamic power consumption is also sensitive to the smallest realizable LSB capacitor size and the smallest available isolation switch size. As compared to current integration, for the given minimum switch and capacitance sizes, the switched capacitor architecture consumes up to $\sim$2.5X higher total static and dynamic power.
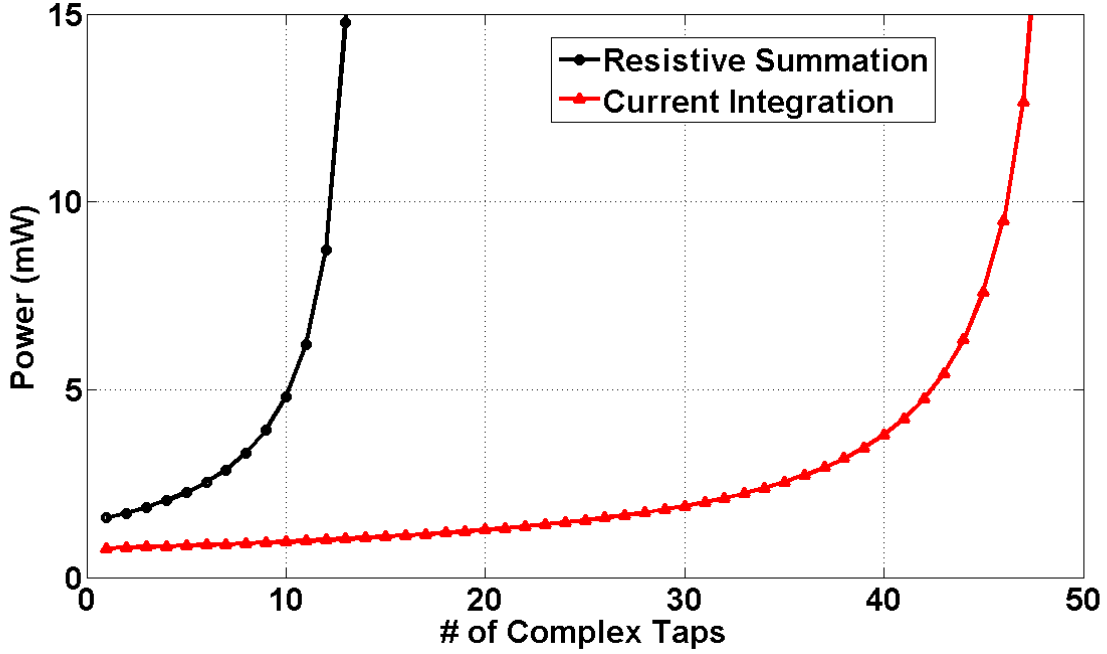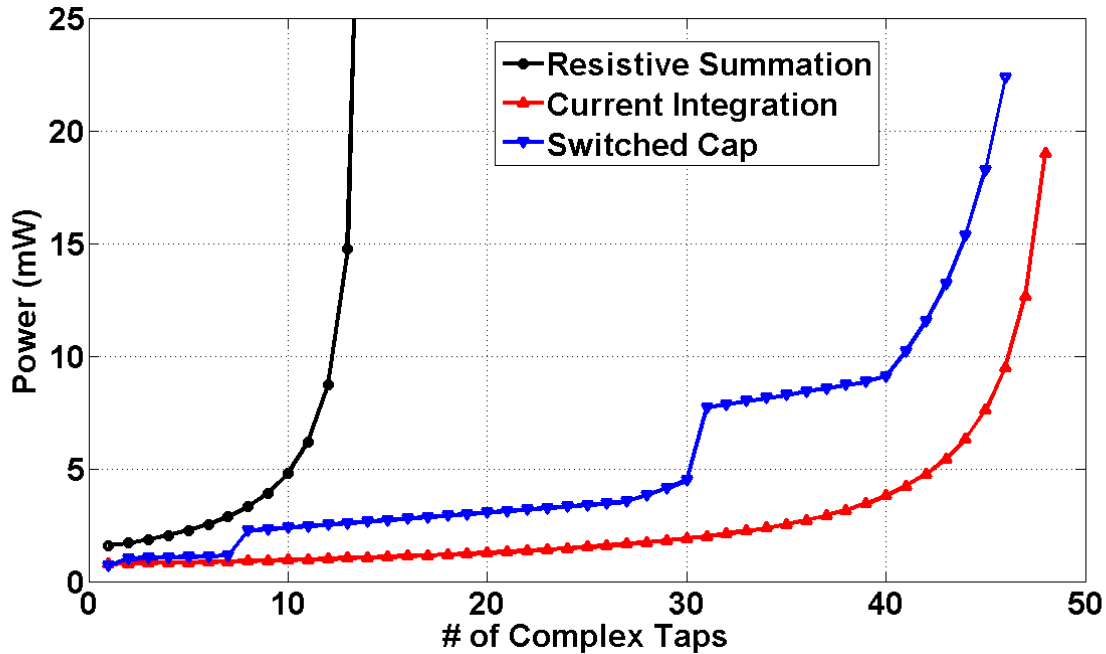
Figure 3.11: Switched capacitor, current-integrating, and resistively loaded current-summing DFE: Summer power vs. no. of complex taps at 5GS/s for $k_{max} = 0.5$, using a 65nm CMOS technology. (1 complex tap = 2 I/Q coefficients).

3.2.3 Switched Capacitor-Based Voltage Summing

In comparison to the hybrid switched capacitor current integrating DFE [52] which uses a fixed voltage and capacitor DAC per coefficient, the voltage-mode only switched capacitor summing in [51] (as discussed in sub-section 3.1.2) instead uses a fixed capacitor and voltage DAC per coefficient. This technique of coefficient implementation can be analyzed as a special case of the capacitor DAC where the capacitor is neither segmented nor needs isolation switches. In this scenario, the loading seen from each coefficient is its full-scale capacitance, irrespective of the ISI value. Therefore, the self-loading and power consumption of such a summer structure would be significantly worse than using a capacitor DAC.

3.3 Conclusion

In this chapter, we discussed and analyzed primarily two alterate DFE summing styles – one using current integration and another using voltage feedback using switched capacitors. Both these techniques promise lower power summer designs than conventional resisively loaded current summation.

The current integration technique [50] enables the use of a low-bandwidth summing amplifier by using sampled (DC) inputs. The output is differentially reset every symbol cycle

to mitigate any effect of the previous data value. In comparison to conventional current summing at a resistive load, current integration significantly reduces the impact of self-loading from the DFE coefficients, and enables the implementation of ∼3-5X more coefficients at equal power.

The switched capacitor based voltage feedback technique uses capacitive coupling to implement ISI cancelation. The DFE coefficients may be implemented either as a capacitor DAC driven by a fixed supply voltage or a fixed capacitor driven by a voltage DAC. With the capacitor DAC-based technique [52], the unused DAC segments for a particular ISI configuration may be shielded from the sensitive summing node by gating switches. The use of a high supply voltage driver with the capacitor DAC potentially reduces the capacitor size and hence the relative self-loading of the coefficients at the summing node. However, the segmentation of the capacitors and their gating switches to achieve the requisite resolution is limited by the minimum implementable device size. This requires an upsizing of the entire circuitry and raises the total power consumption to ∼2.5X higher than current integration. Effectively, the direct interaction of coefficient resolution with the shared summing node is detrimental to the architecture. Using a fixed capacitor with a voltage DAC [51] leads to a worst-case capacitor loading for every coefficient irrespective of the coefficient setting. This results in significantly worse self-loading than the use of a capacitor DAC with segment-wise gating. Since both switched capacitor DFE architectures suffer from a certain fixed self-loading irrespective of the coefficient setting[3], the resultant power penalties can be mitigated by using cascode-based summation proposed earlier for resistively loaded current summation in Chapter 2.

The benefit offered by capacitive voltage feedback however is that of reduced analog settling time for the coefficients. This settling time is of the order of the rise/fall time of digital gates. In comparison, current integration typically requires about a half clock cycle for power-efficient integration. The reduced analog settling delay eases the timing constraints and (and therefore power consumption[4]) of the feedback shift register flip-flop and sign-selection drivers. Since capacitive summing techniques nevertheless consume significant digital driving power, they are useful when the benefit of reduced analog bandwidth on static power consumption outweighs the digital power penalty.

Of all the DFE architectures discussed so far in this dissertation – namely current summation on a resistive load, current integration on a capacitive load, and switched capacitor based voltage summing – current integration is shown to be the most power-efficient summing architecture. This architecture reduces self-loading from the coefficients and inherently makes power consumption independent of coefficient resolution. A combination of current integration and cascode summation (which was introduced in Chapter 2) can be combined to efficiently implement even more coefficients than the first 40-coefficient DFE prototype. As will be shown in detail in the next chapter, this structure can also incorporate FFE

---

[3]The capacitor-DAC based architecture suffers from a fixed self-loading from the isolation switches

[4]Power savings are achieved if the flip-flops and drivers were originally sized bigger than minimum transistor size for the requisite driving strength.

coefficients to mitigate pre-cursor ISI. Using this cascoded current integrating summer enables the implementation of an efficient merged FFE-DFE summing amplifier which eases the bottleneck for equalization across NLOS channels.

Chapter 4

Receive-side Feedforward Equalizer (RX-FFE) Design

The previous chapters highlighted the design of decision feedback equalizers (DFEs) that are efficient at canceling post-cursor inter-symbol-interference (ISI). The fundamental limitation of a DFE, however, is that it cannot mitigate pre-cursor ISI. In certain wireline and most wireless channels, which tend to be non-line-of-sight (NLOS), the presence of pre-cursor ISI makes it mandatory to have a separate feed-forward equalizer (FFE).

While a transmit-side (TX) mixed-signal FFE is implemented conveniently by using digital delay elements and current-summing DACs [55], its adaptation requires receiver feedback. TX-adaptation data is typically handled in wireline transceivers by using a back-channel communication path [47] from the receiver. Wireless transceivers usually do not have a reliable back-channel to begin with, thus making such adaptation techniques more difficult to implement. This leaves receive-side (RX) FFEs as the only viable alternative to enable communication across NLOS wireless channels with pre-cursor ISI.

Implementing an RX-FFE in the analog domain requires analog delay elements as well as analog weighting circuitry. As will be shown in this chapter, implementing an RX-FFE with multiple coefficients invariably involves several tradeoffs between linearity, resolution, tuning range, and power. It is therefore not surprising that as compared to mixed-signal TX-FFEs, analog RX-FFEs have been very scarcely implemented. In this chapter, we will first discuss prior-art in RX-FFE architectures and their problems with scaling to a multi-coefficient implementation. We will then describe the proposed switching-matrix architecture for a multi-coefficient RX-FFE, its design methodology and core circuitry. This architecture addresses many of the above-mentioned tradeoffs. While the power of the proposed architecture still scales non-linearly with the number of coefficients (as do all other architectures), through the analyses presented in this chapter, we shall show that the total power consumption is lower in magnitude. The advantages of the proposed techniques are illustrated with the proof-of-concept design of a 65nm CMOS prototype I/Q 32-coefficient FFE to equalize a 60GHz non-line-of-sight (NLOS) channel and achieve 8Gb/s QPSK.
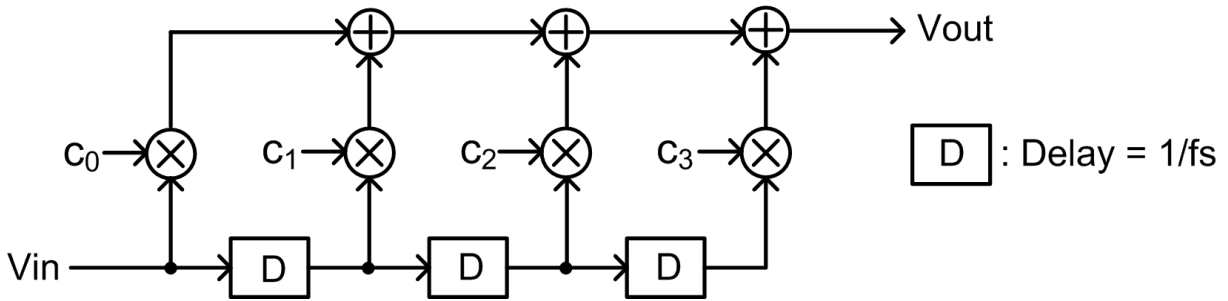
Figure 4.1: FFE Block Diagram
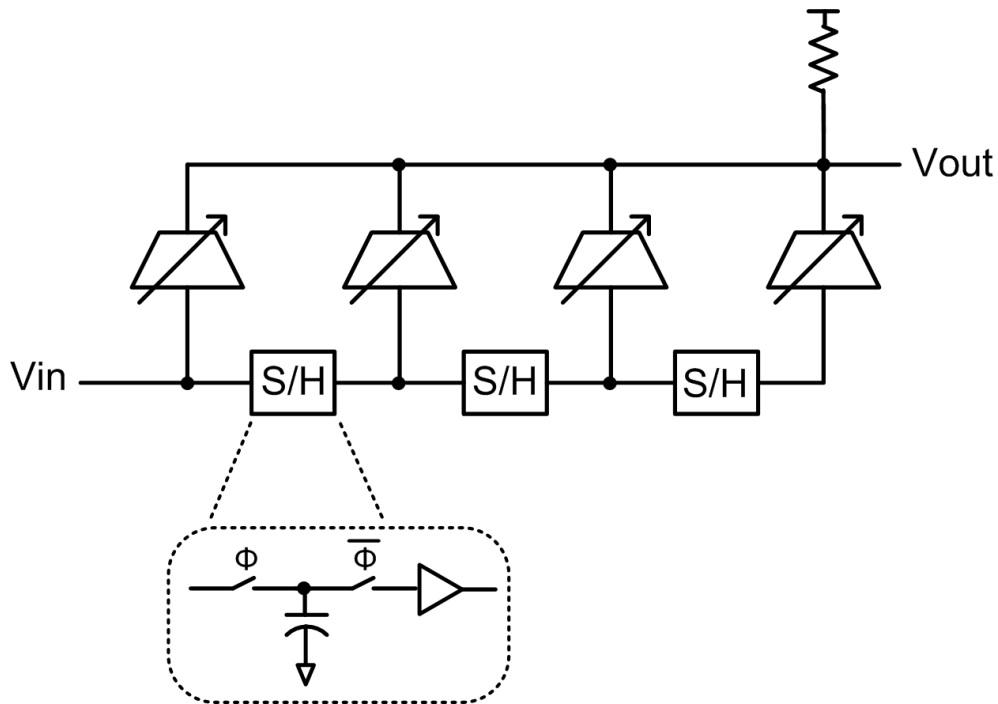
Figure 4.2: Rx-FFE implementation (4-tap example) using cascaded S/H and gain compensation.

4.1 Prior Art

An FFE is essentially an FIR filter, as shown in Fig. 4.1. An RX-FFE requires the delay (D) blocks to be analog delay elements – either in continuous time or discrete time. In order to be able to operate across a variety of channels, wireless receivers typically need to be amenable to variable data-rate operation. Among previously demonstrated designs with continuous-time delay, the 7-tap FFEs in [56] and [57] used LC-based delay lines with active buffers. This passive-intensive design technique suffers from a large area penalty from
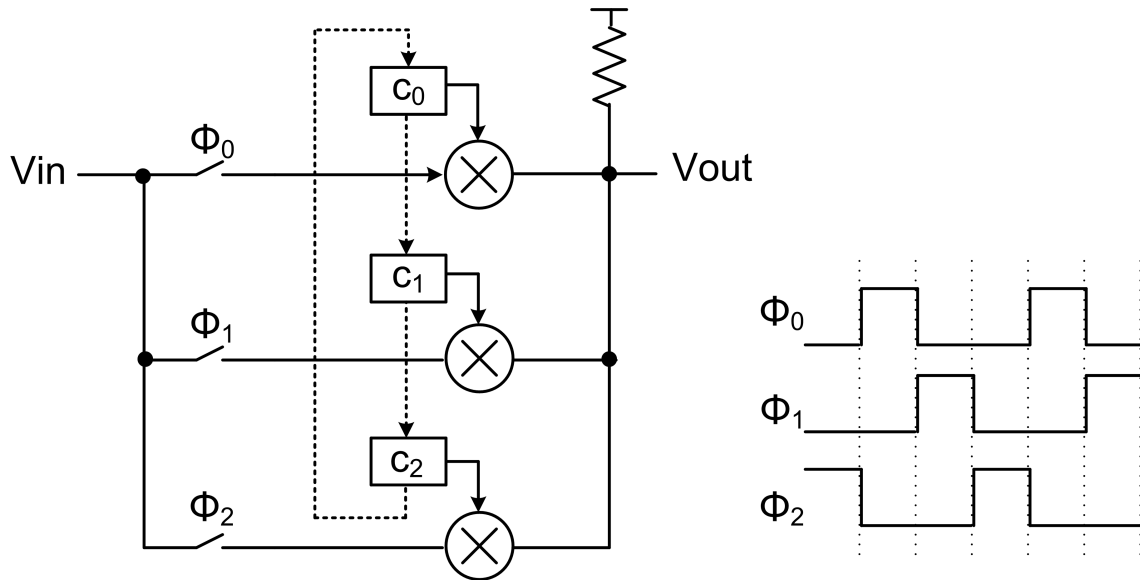
Figure 4.3: Rx-FFE implementation (3-tap example) using rotating coefficients [39].

the inductors. LC-based techniques also have limited tuning range for achieving variable data-rate operation. The 4-tap FFE in [58] uses relatively compact active-inductor based delay-cells, but also has a limited adjustment range (2.5-3.5Gb/s).

As opposed to continuous-time delay elements, the use of discrete-time delay blocks makes it convenient to implement a wide range of data-rates. The most intuitive way to implement discrete-time delay is by using capacitor-based analog sample/hold (S/H) blocks, as shown conceptually in Fig. 4.2. To compensate for per-stage loss in voltage swing due to charge sharing, a delay line with a cascade of S/H blocks would need intermediate buffers. For a reasonably long FFE, such a cascade would suffer from an accumulation of kT/C noise, gain mismatch and DC offset. kT/C noise may be reduced by upsizing each cell, while gain-mismatch and DC offset can be mitigated by calibration and/or feedback. Effectively, while all of the effects may be individually addressed, such mitigation often leads to increased power consumption and/or the overhead of inconvenient block-wise and mostly-offline calibration. Most recent implementations of RX-FFEs in scaled CMOS technologies therefore obviate the use of a S/H-cascade, and replace it with interleaved S/Hs using multi-phase clocks, as will be described next.

4.1.1 Rotating-Coefficients FFE

The RX-FFE proposed by T.-C. Lee and B. Razavi in [39] is implemented by using an interleaved S/H front-end followed by a bank of DACs, as shown in Fig. 4.3. The weight of each DAC is designed to shuffle through the FFE coefficients. The operation is shown conceptually in Fig. 4.4 where all DAC coefficients are changed every data-cycle. This necessitates the equivalent of a symbol-rate switching DAC for every equalizer coefficient. When scaled to a
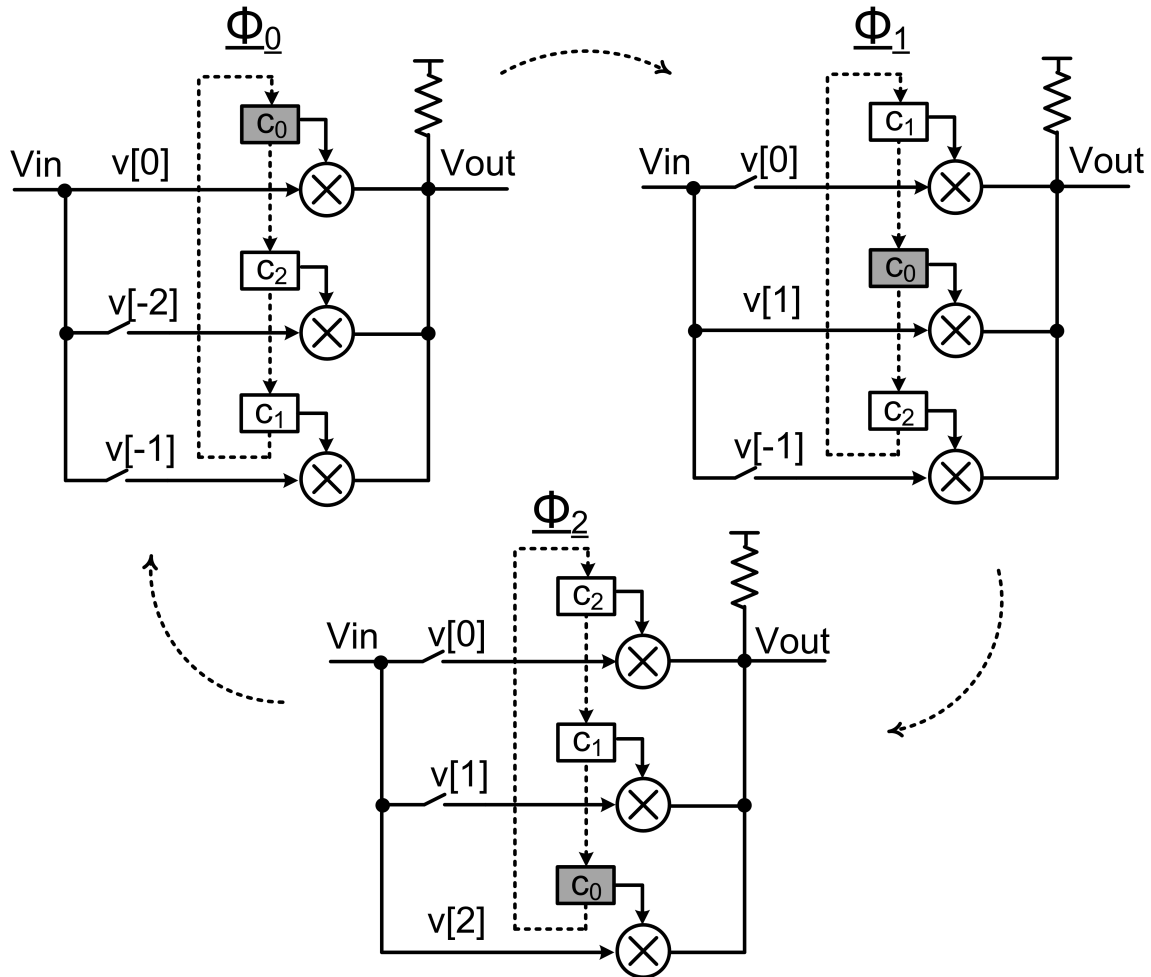
Figure 4.4: Time evolution of rotating-coefficients FFE [39].

larger number of FFE coefficients, in order to maintain equal total quantization noise, each DAC requires higher resolution. This coefficient resolution scales as $O\left(log\ N\right)$. Therefore, the power of the structure scales as $O\left(N \cdot log\ N\right)$.

Since each coefficient DAC requires digital flip-flop drivers on a per-bit basis that switch every clock period, this architecture consumes significant total digital driving power. For a large number of FFE coefficients, the power of such a DAC array becomes prohibitively large.

4.1.2   Time-interleaved FFE

The architecture by J. Jaussi *et al.* [40] implements interleaved S/H followed by with parallelized analog scaling/summation and data slicing blocks as shown in Fig. 4.5. Unlike coefficient-rotation, this architecture allows the use of low-speed DAC-based coefficients.
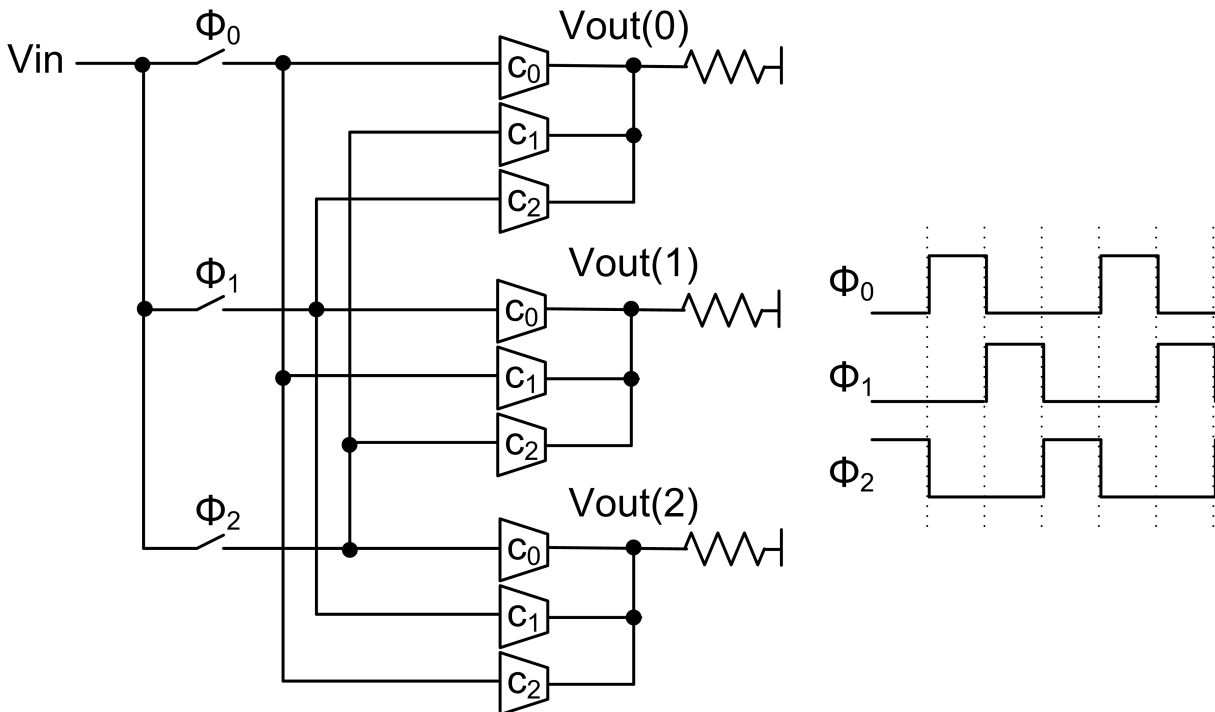
Figure 4.5: Rx-FFE implementation (3-tap example) using interleaving [40].

The interleaving also reduces speed constraints on all constituent blocks, potentially reducing total power consumption. The 4-tap FFE in [40] in 0.13μm CMOS is promising and implements a 4-tap FFE working at 8Gb/s. However, when this architecture is scaled to a higher number coefficients, it incurs repeated analog weighting circuitry, thus suffering from significant area overhead, wire length and consequently power overhead.

Since the FFE is typically followed by a DFE, the impact on the DFE circuitry must also be understood. Parallelized data-slicing when used along with loop-unrolling[45][46], helps with the closure of the timing loop of the 1st DFE tap. Typically, the degree of unrolling and parallelizing are kept equal. Parallelization with a few segments eases the slicer feedforward timing constraints, thus enabling a reduction of per-slicer size and power. However, once the slicer is pared down to the smallest possible size, the only other way that parallelizing can be used to save slicer power is by dropping the supply voltage. However, any additional degree of parallelization with fixed size per slicer proportionately scales up the total output loading from these slicers to the preceding summing amplifiers, thus raising total summing power. Apart from increased total summer power (which dominates reduced slicer power), a high degree of parallelization also causes repeated DFE analog circuitry, which leads to an area, wiring length, and power overhead despite the relaxed timing constraints from parallelization. To mitigate the DFE power overheads arising from excessive parallelization, the parallel FFE outputs may be re-serialized before slicing (with additional digital power penalty).
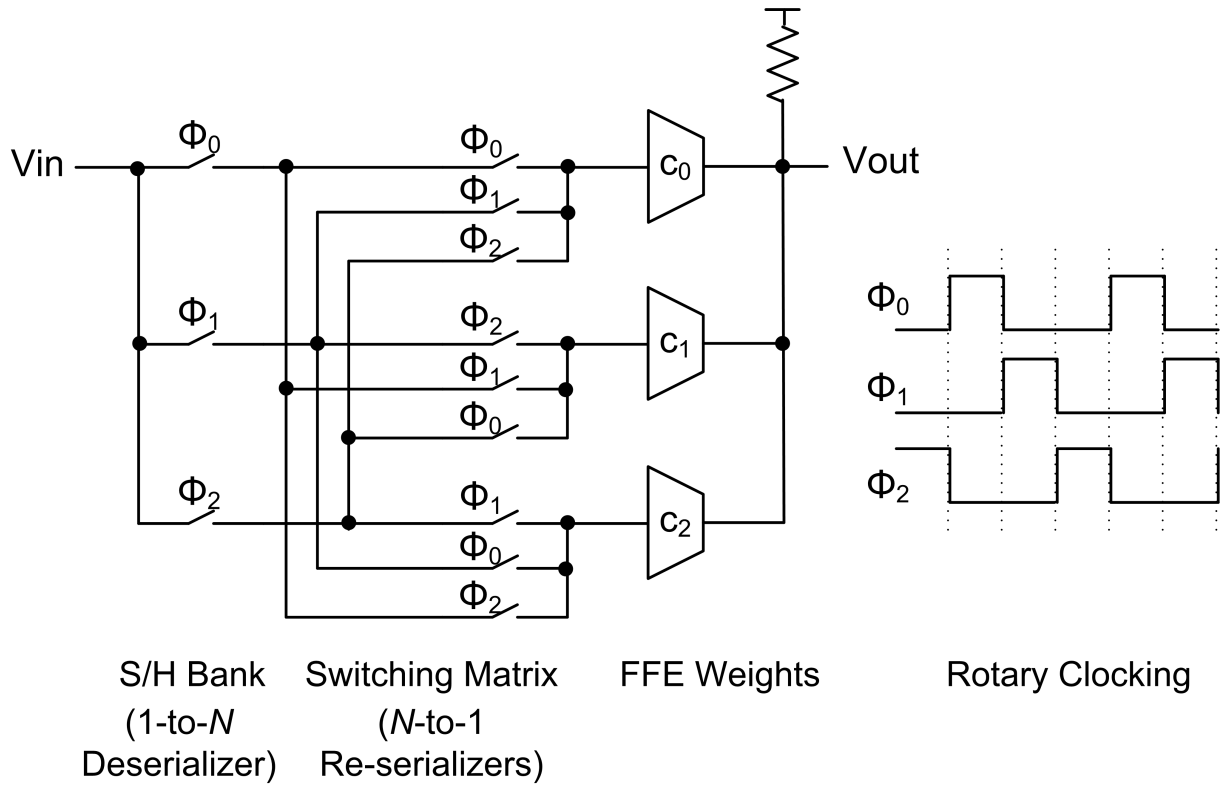
Figure 4.6: Proposed Rx-FFE implementation (3-tap example) using a switching matrix.

4.2 Proposed Switching-Matrix-Based FFE

The proposed RX-FFE architecture obviates both high-speed DACs and the area penalty of interleaved slicing by re-serializing the outputs of the S/H bank used in both of the previous architectures. This is achieved by a switching matrix, as shown conceptually for a 3-tap example in Fig. 4.6. Fig. 4.7 shows the phase-wise working of the switching matrix architecture. Generalizing to N taps, the structure is effectively a 1-to-N deserializer followed by N parallel N-to-1 serializers, such that driving it with a rotary N-phase clock creates an analog S/H delay line. By mitigating the area, wiring and power penalty of interleaved slicing (a problem described in sub-section 4.1.2), the re-serialization also makes it efficient to combine the FFE with the following DFE.

Before quantifying the benefits of the proposed architecture and comparing against prior art, it is instructive to first walk through the operation of the architecture. The switching matrix will then be quantified in terms of power dissipation versus number of analog delay elements. The framework is used to design a 16-tap I/Q prototype FFE prototype in 65nm LP CMOS for a 60GHz NLOS channel. Finally, the framework and circuit details will be used as a baseline to compare of the proposed architecture with prior art using rotating coefficients [39] and interleaved slicing [40] in terms of power consumption versus number of taps.
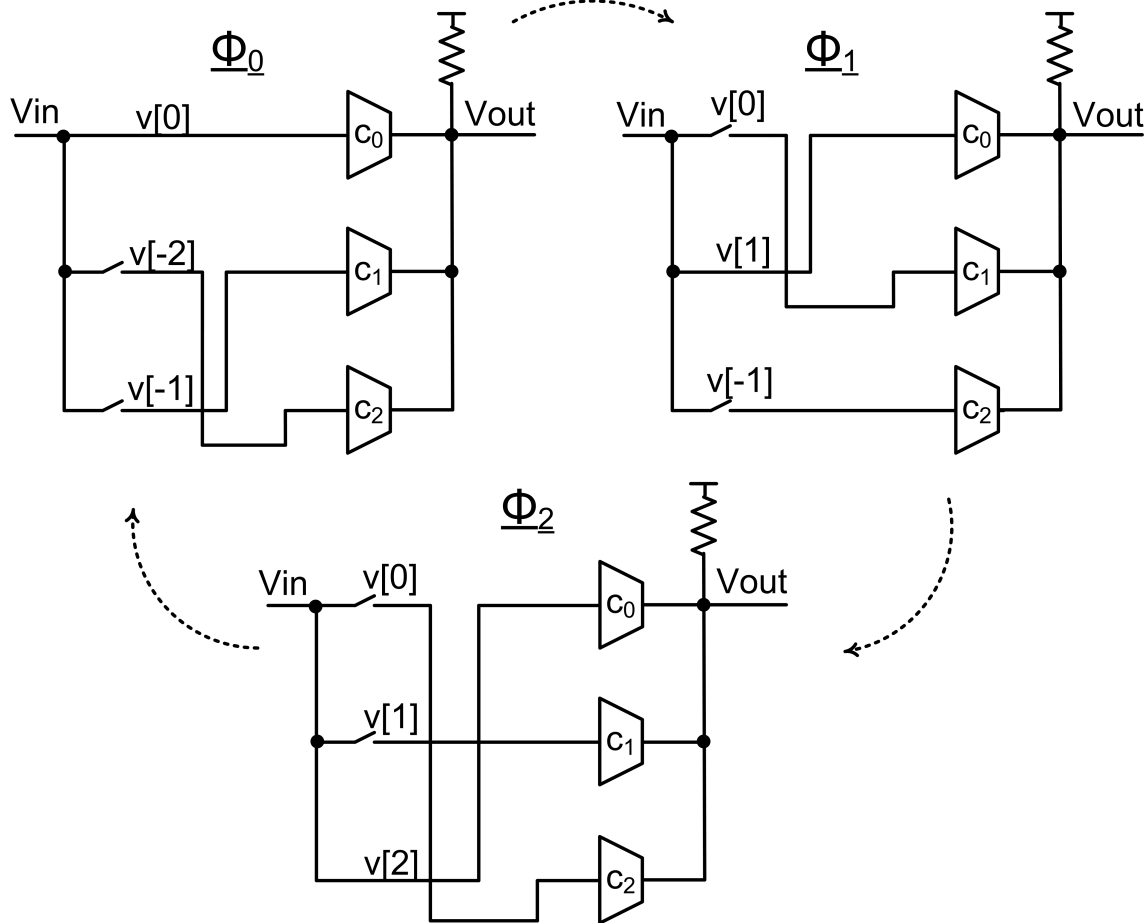
Figure 4.7: Phase-wise working of the switchiing matrix (3-tap example).

4.2.1   Swiching Matrix Core

An efficient way to implement the switching matrix is by using current integration [50]. This technique was earlier discussed in Chapter 3 for the efficient implementation of a DFE with many coefficients. Fig. 4.8 shows the schematics of such a 16-tap prototype switching matrix. As shown in the schematic, switching matrix connectors act as cascodes in each current integrator leg. The integrators are set up such that the current from every g_m cell is passed from one integrator output to the next through these cascode switches. Any mismatch in the gm-cells would cause each matrix output to suffer from rotating offsets that change every UI. Since these rotating offsets cannot be statically canceled after FFE-weight multiplication, the matrix input is provided with per-row offset cancelation capability.

To perform a reset operation at every clock cycle, current integrators typically [50] use clocked PMOS devices to precharge the output nodes to Vdd. Having high common mode current, however, could lead to a significant drop in the output common mode during integration, consequently reducing the integrator gain and linearity. In the 8-tap DFE current
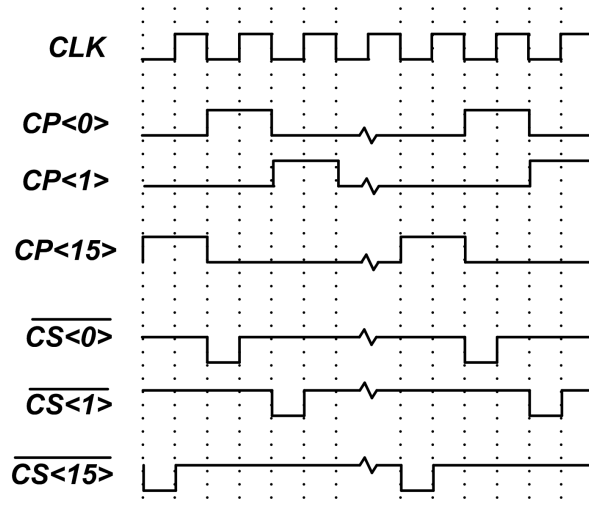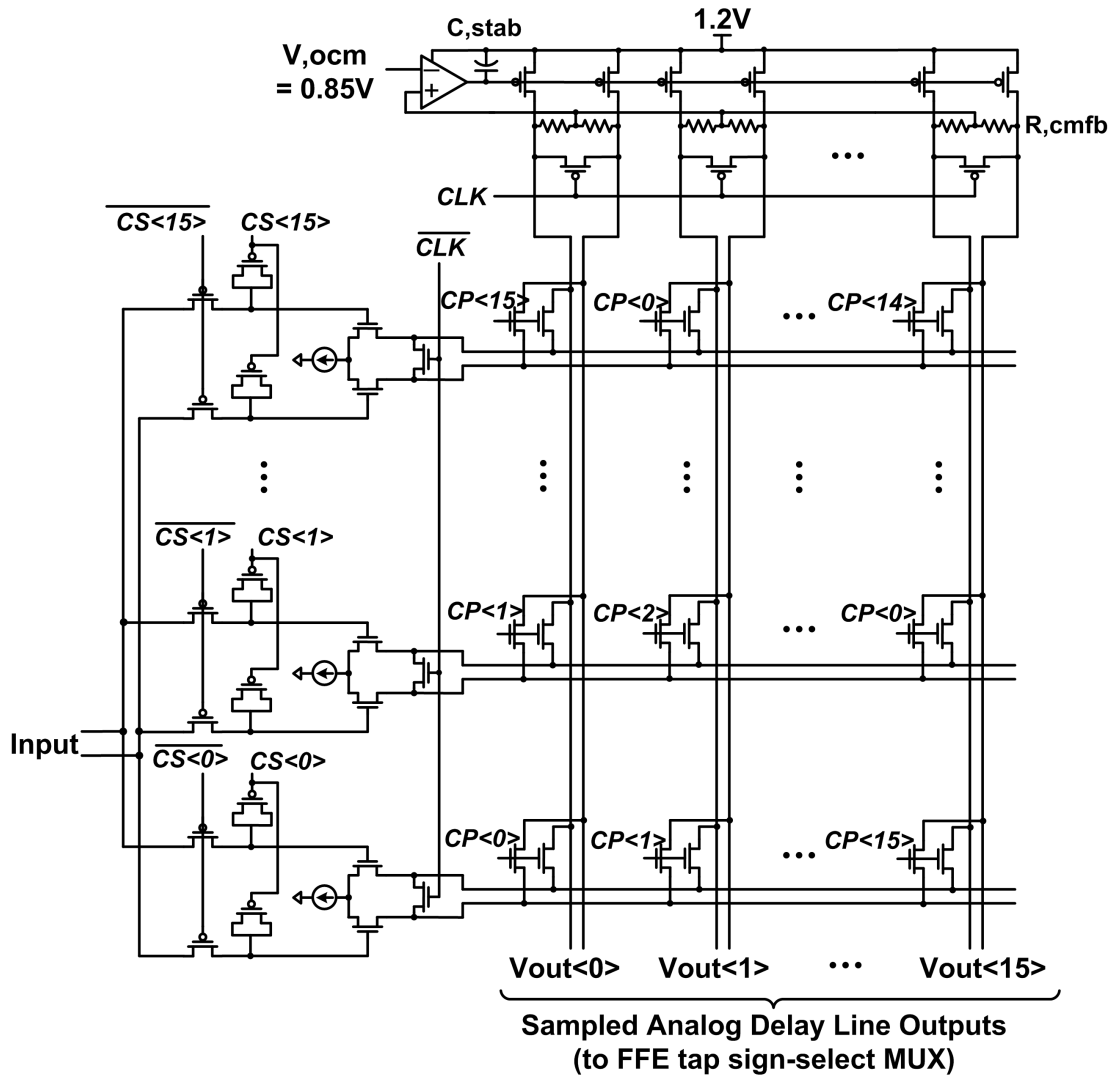
Figure 4.8: Schematics of prototype 16-element switching matrix and associated clocking waveforms.
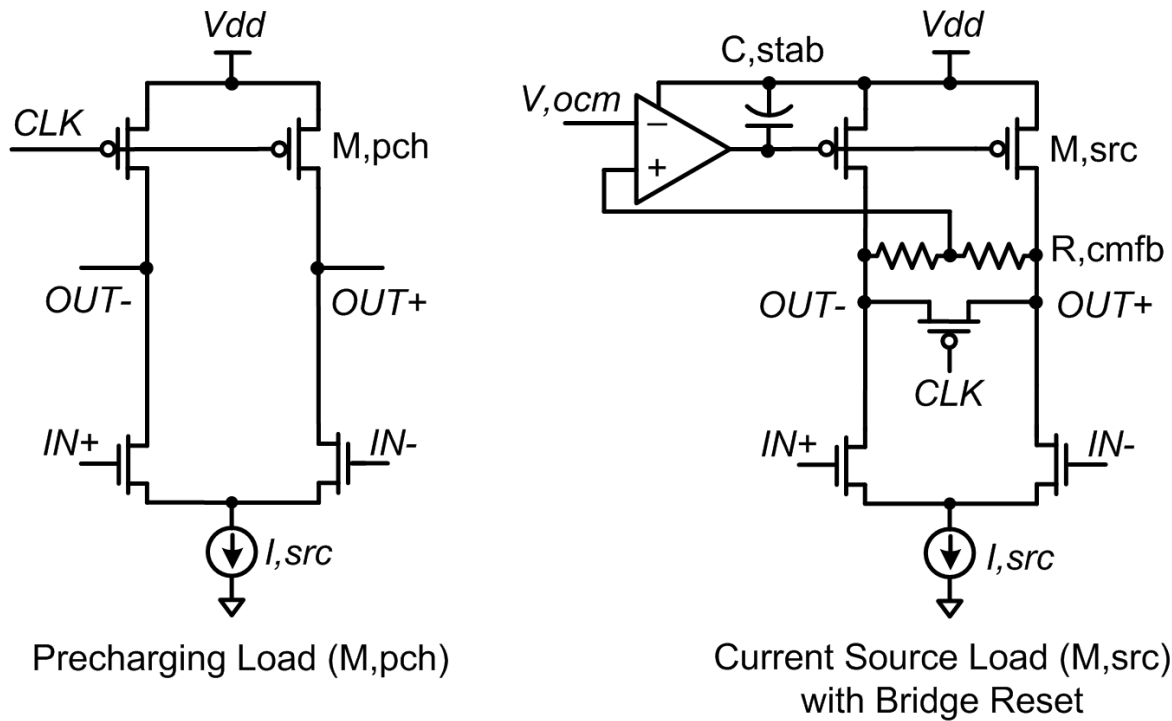
Figure 4.9: Typical current integration with a precharging PMOS load [50] (left); current source load with CMFB, bridge reset (right).
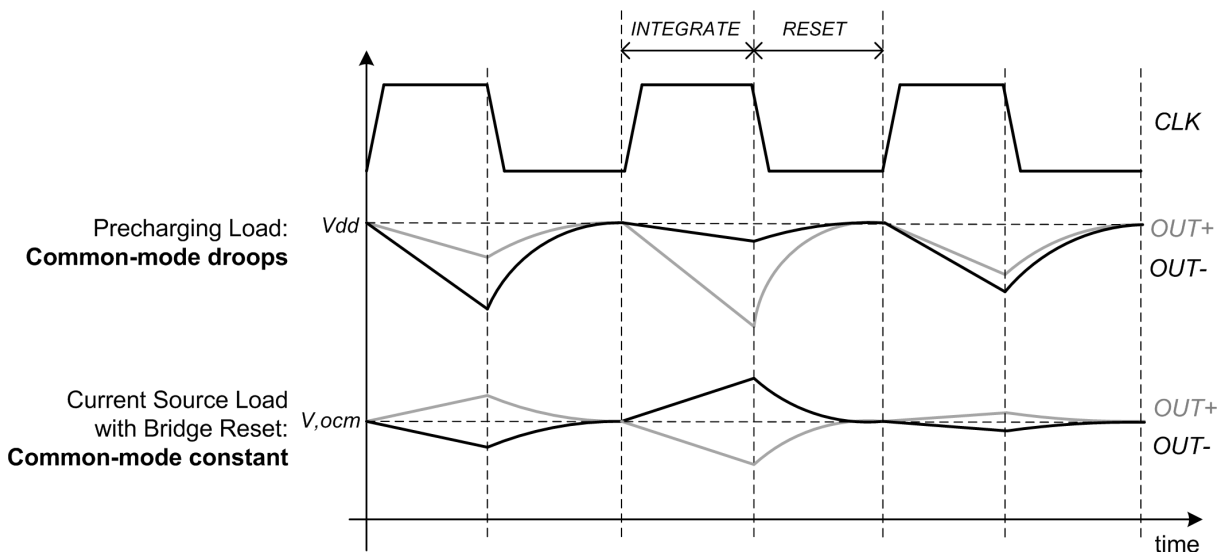
Figure 4.10: Current integrating waveforms with precharging PMOS load [50], and current source load with bridge reset (proposed).

integrator by T. Toifl *et al.* in [52], this problem is handled by a separate common-mode boosting circuit using capacitive coupling. The current integrator by A. Agarwal *et al.* in [59] for a 4-tap RX-FFE mitigates the droop by adding a common-mode current during the integration phase.

In the proposed implementation, the common mode droop is entirely avoided by designing the load as a long-channel PMOS current source with common-mode feedback (CMFB), as shown in Fig. 4.9 (for simplicity, the input cascode devices are not shown). The length of this current source is chosen to tradeoff between high DC output impedance and low capacitive loading overhead. To keep the switching matrix core compact and reduce the CMFB wiring overhead, the CMFB amplifier is shared across all outputs. The CMFB resistance (R_{cmfb}) is implemented with a minimum-width long-channel PMOS in triode operation. The reset operation is performed by a bridge PMOS switch [60]. Fig. 4.10 shows the integrate/reset waveforms for both precharging and current source loads. To ensure high output impedance, the PMOS current source is maintained sufficiently in saturation by using CMFB to place the output common-mode (V_{ocm}) at 850mV with a supply voltage of 1.2V. In this commercial LP process devoid of low-V_{th} transistors, since transistor thresholds are almost half the rail-to-rail voltage, it is extremely critical to ensure sufficient headroom to maintain transistors in saturation. Since each cascode gate is driven by digital gates running off the nominal 1.2V supply, in order to provide reasonable headroom for the gm-stage and its current source, the bulk voltage needs to be raised. All cascode devices are therefore laid out in a shared triple well, with the p-well set to a fixed 400mV.

During the integration phase (CLK=‘1’), the cascode switches steer and integrate the gm-cell current at the high impedance output, which is then multiplied with the corresponding FFE weights. Integration is followed by reset (CLK=‘0’) through the bridge PMOS switch. At any phase, since only one of the N cascodes connected to a gm-cell is on, the cascode source node has low bandwidth and also needs to be reset by a bridge NMOS switch.

During the reset phase, the FFE tap input is isolated from the resetting switching matrix output by using a clocked FFE sign-select MUX (with turn-off capability[1]), as shown in Fig. 4.11. The operation of the MUX is illustrated in Fig. 4.12. Just before the start of the integration phase while the MUX is still off, the input and output of the MUX hold different values. The tap side of the MUX holds the previous data while the switching matrix side – which is being reset to initialize integration for the next cycle – has zero differential voltage. Therefore at the onset of integration when the MUX turns on, charge sharing from the tap input to the switching matrix output disturbs the initial zero differential voltage. Since this charge sharing is from the previous data on to the current data that is starting to be integrated, the effect is equivalent to ISI from the first post-tap relative to the current position on the delay line. Fortunately, since all FFE/DFE taps are continually adapted,[2] this ISI will be adaptively corrected.

---

[1]When a tap is to be turned off completely, to avoid any stray capacitive coupling from the tap input to output, the MUX is provided with OFF functionality by setting $sign$=‘0’, $\overline{sign}$=‘0’.

[2]Technically, the adaptation runs only during a packet header, but the effective channel stays unchanged during the packet payload.
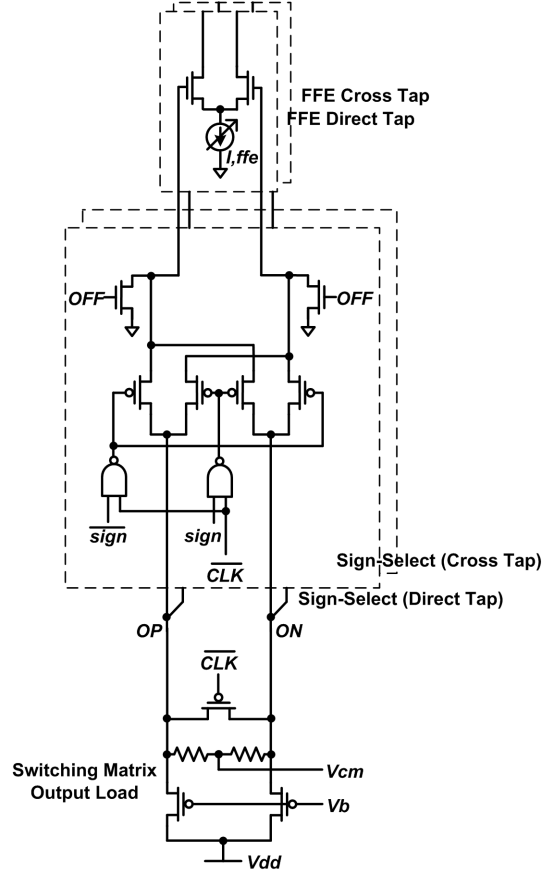
Figure 4.11: Sample/hold from switching matrix output to FFE tap, with built-in sign-select and turn-off capability.

Now that the switching matrix core operation is understood, the next step is to compute the power required by the switching matrix to support a certain number of FFE delay elements, N_{seg}. The power is computed by evaluating the per-segment bias current I_{seg} required by each g_m-stage to support a certain gain G, data-rate f_s and external capacitive load C_L. As derived in the detailed analysis shown in Appendix B, I_{seg} is:

$$I_{seg} = \frac{G \cdot V_{in}^* \cdot f_s \cdot C_L}{(1 + k_{reset})^2 \, (4\pi \cdot f_s) \left(N_{seg} + \gamma \cdot \dfrac{\omega_{T,casc}}{\omega_{T,in}} \cdot \dfrac{V_{casc}^*}{V_{in}^*}\right)} - \dfrac{\gamma \cdot G \cdot f_s}{\omega_{T,casc}} \left(N_{seg} \cdot \dfrac{V_{in}^*}{V_{casc}^*} + \dfrac{\omega_{T,casc}}{\omega_{T,load}} \cdot \dfrac{V_{in}^*}{V_{load}^*}\right)}$$

$$(4.1)$$

In the above equation, subscripts in, $casc$, and $load$ correspond to input g_m-stage, cascode, and PMOS load transistors respectively, while k_{reset} is the reset-factor (first introduced in equation 3.7) to support the capacitance of the bridge reset transistor. Intuitively, the negative term in the denominator – which is proportional to $G \cdot f_s \cdot N_{seg}$ – is an indicator of output self-loading, similar to a standard current integrating summer (equation 3.14).

(a) Hold-mode   **(b) Sample-mode**
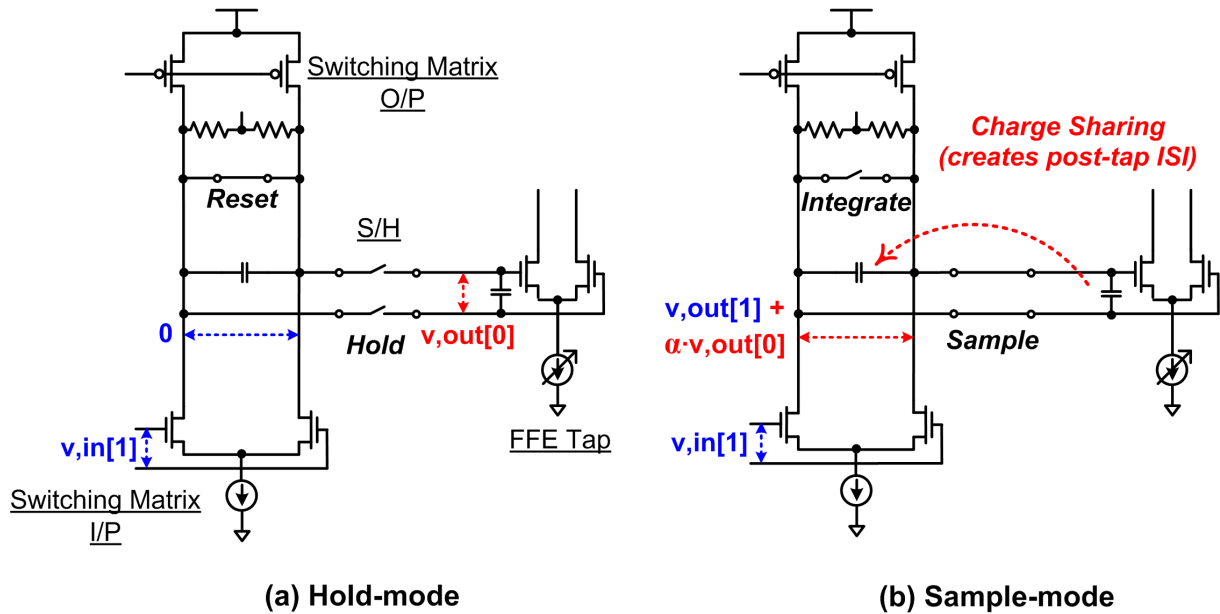
Figure 4.12: Working of S/H (switching matrix and sign-selection details not shown): (a) Hold-mode: Switching matrix resets, tap input holds; (b) Sample-mode: Switching matrix integrates, tap input samples.

Additionally, the first term in the denominator – which is approximately proportional to $\omega_{T,casc}/N_{seg}$ – is an indicator of the self-loading at the cascode source by the N_{seg} other switches connected here. This self-loading degrades the cascode bandwidth $\omega_{T,casc}$ by N_{seg}.

From the above equation, I_{seg} can be simplified to a form:

$$I_{seg} \propto \frac{I_0}{\dfrac{A}{(N_{seg}+1)} - B \cdot (N_{seg}+c)} \tag{4.2}$$

where I_0, A and B are constants dependent on gain, data-rate, external loading, and technology. For a PMOS load with $L = 0.18\mu m$, $c \sim 3$. This form of I_{seg} implies that the switching matrix structure can support a certain number of segments efficiently before it reaches its self-loading limit. More importantly, the form suggests that if the structure is reasonably away from its self-loading limit, $I_{seg} \propto N_{seg}$. This dependence would imply that the total power of the matrix $P_{matrix} \propto N_{seg}^2$. While this dependence on the number of taps is similar as that of a time-interleaved FFE, the absolute values of I_{seg} and P_{matrix} are lower (as will be shown in detail in section 4.2.2).

To illustrate this, a MATLAB simulation was done using the 65nm LP CMOS technology parameters used to design the prototype[3] with $G = 0.6$, $f_s = 5GS/s$ and $C_L = 8fF$. All

---

[3]Since this prototype was implemented in a slow 65nm LP CMOS technology, the gain was intentionally kept low to compensate for the cascode source self-loading of the 5GS/s 16-element delay line. If the GP equivalent of the technology were to be used, it would be able to achieve $G=1$.
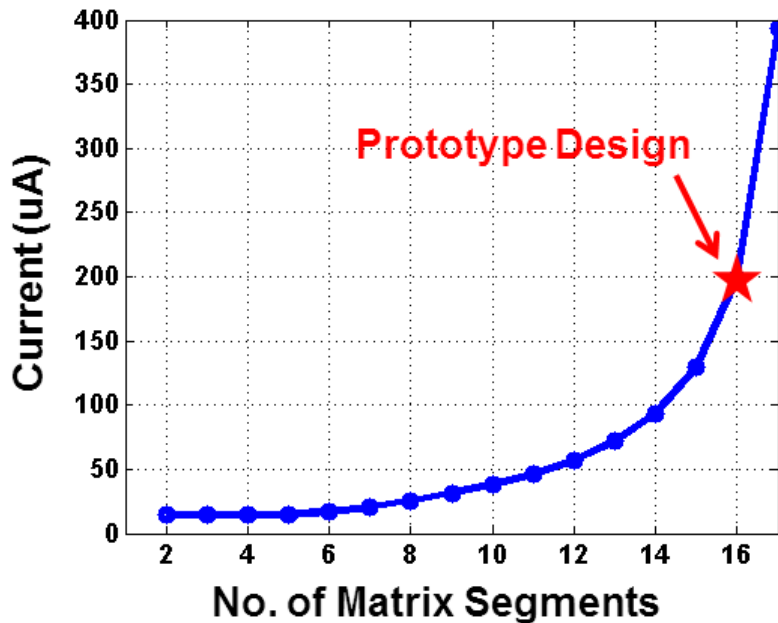
Figure 4.13: Switching matrix current per segment (I_{seg}) vs. No. of FFE Taps, for $f_s = $ 5GS/s, $G = 0.6$, $V^* = 200mV$, $C_L = 8fF$, PMOS load L = $0.18\mu m$.

.

transistors in this design have $V^* = 200mV$, and the PMOS load has $L = 0.18\mu m$. Fig. 4.13 shows the trend expected in equation (4.1) for I_{seg} vs. number of taps. To support the prototype 16-tap FFE for 5GS/s operation, each segment requires $200\mu A$ bias current. While this region of the design is close to the self-loading limit, as will be shown in the comparison analysis that follows, the total power consumption nevertheless compares favorably with prior FFE art using rotating coefficients and time interleaving.

4.2.2 Comparison with Prior Art

Now that we have a framework for analyzing the proposed switching-matrix based FFE architecture, we shall compare it with previously developed FFE architectures using rotating coefficients [39] (sub-section 4.1.1) and time interleaving [40] (sub-section 4.1.2). The comparison is done in terms of power consumption versus number of FFE coefficients. For fairness, the analysis assumes that all designs are implemented in the same 65nm LP CMOS technology node as the proposed 16-tap prototype. The comparison is made for the FFE supporting only real coefficients on only one channel (i.e. I or Q only). To account for unequal input loading across different architectures, the analysis also includes the power required to drive the entire interleaved S/H bank with a class-A amplifier with unity gain. This driving amplifier is designed with enough bandwidth to ensure >95% settling during the UI/2-duration sampling operation. The analysis also includes the power of the summing

Architecture	Rotating Coeffs.	Time Interleaved	Switching Matrix
# of Summers	1	N	1
# of Unrolled Taps	0	1	0
# of Slicers	1	$2N$	1
# of Interleaving	1	N	1
Per-Slicer Power	P_{slicer}	$\dfrac{P_{slicer}}{2 \cdot N}$	P_{slicer}
Tot. Slicer Power	P_{slicer}	$\left(\dfrac{P_{slicer}}{2N}\right) \cdot (2N)$	P_{slicer}
Per-Summer Power	P_{summer}	P_{summer}	P_{summer}
Tot. Summer Power	P_{summer}	$N \cdot P_{summer}$	P_{summer}
Bits/FFE Coef.	$n_{fxd} + log\,(N)$	$n_{fxd} + log\,(N)$	$n_{fxd} + log\,(N)$
Dig. Power/FFE Coef.	$P_{flop} \cdot [n_{fxd} + log\,(N)]$	0	0
Tot. Coef. Dig. Power	$N \cdot P_{flop} \cdot [n_{fxd} + log\,(N)]$	0	0
Sw.mat. # of Legs	0	0	N
Bias/Leg (B.26)	0	0	$I_{seg} \cdot N$
Tot. Sw.mat. Bias	0	0	$I_{seg} \cdot N^2$
CLK Power/Switch	0	0	$P_{switch} \cdot N$
Duty-Cycle/Switch	0	0	$\dfrac{1}{N}$
Sw.mat.CLK Power	0	0	$\left(\dfrac{P_{switch} \cdot N}{N}\right) \cdot N^2$

Table 4.1: Comparison of FFE Architectures.

amplifier(s) and preamp/slicer pair(s) following the analog delay line. Since a DFE is also typically included in most equalizers, the samplers are sized to meet the latency constraints of the initial DFE taps. However, DFE tap power is excluded from the total.

Table 4.1 compares the three architectures by breaking down the power dissipation as a function of their respective components. The number of elements and amount of interleaving listed in this table are better understood by re-visiting Fig. 4.3, Fig. 4.5 and Fig. 4.6 respectively. While these figures each illustrate a 3-tap example, this table generalizes the results to N taps. The notations in the figures are: N - No. of FFE taps, P_{slicer} - slicer power to enable direct DFE tap-1 feedback, P_{summer} - summer analog power to support a unit slicer load (corresponding to P_{slicer}), n_{fxd} - FFE tap resolution required for a single tap, P_{flop} - power of a unit flip-flop, I_{seg} - switching matrix segment bias current for 1x1 matrix, P_{switch} - switch driving power for 1x1 matrix.

For DFE loop-unrolling with the time-interleaved architecture, a high degree of unrolling would reduce the feedback constraint to the sum of few MUX delays. Simultaneously, the slicer feedforward constraint is significantly relaxed which could ideally lead to a proportional easing in slicer resolution speed and power. The savings in power are realized by down-sizing
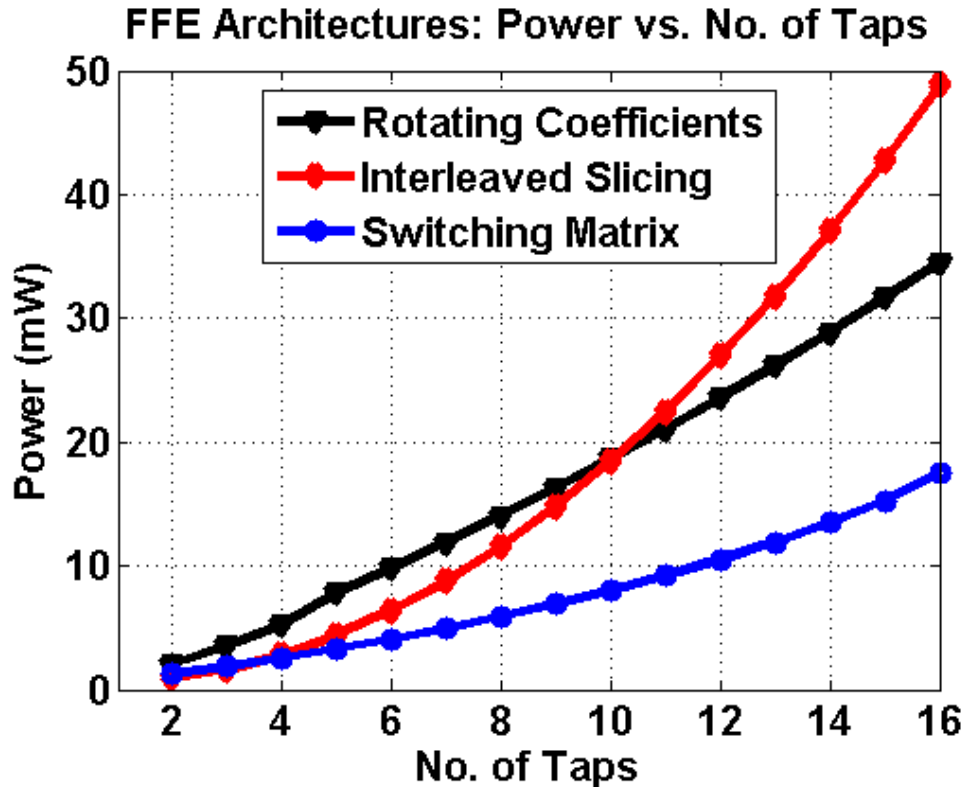
Figure 4.14: Power vs. No. of FFE Taps at 5GS/s for (a) rotating coefficients [39], (b) interleaved slicing [40], (c) switching matrix (this work). Note: Power consumption includes the analog buffer driver, FFE analog delay implementation, FFE-DFE summer/slicers (DFE tap power excluded) and clocking. Power consumption is for one channel only (i.e. I or Q).

the slicer. However, minimum device size constraints restrict slicer down-sizing[4] at best to ~2X. In the analysis, therefore, the loop unrolling is limited to the first tap.

Given the component-wise breakdown in Table 4.1, Fig. 4.14 shows a simulation of power scaling across different architectures. The rotating coefficients architecture is dominated by the power of the sequential delay elements needed to drive the high-speed DACs switching at UI-rate. The jumps in the power plot for this architecture come from the need for increased FFE-tap resolution, which exacerbates the total power consumption. For 16 taps, the power dissipation is ~2X higher than that of the switching matrix.

For the interleaved sampling architecture, with a small number of taps, the data slicers can take advantage of the relaxed latency and may hence be scaled down in size (and power). As a result, the FFE taps and in-turn the S/H circuitry and its driving amplifier may also be sized down to save power. However, once the slicers have been sized down to minimum transistor size, they cannot be further sized down. This causes the FFE tap, S/H circuitry

---

[4]The down-sizing factor of 2X is computed by running simulations for (StrongArm) slicer resolution speed and (dynamic, fast) MUX delay in this 65nm LP CMOS technology.

and the driver amplifier power consumption to increase proportional to the number of taps. The figure shows that at 16 taps, the power dissipation is ∼3X higher than that of the switching matrix. It must be noted that the analysis does not include the power overhead incurred by the increased wiring complexity. Therefore, the computed plot is only a lower bound on the power consumption of this architecture. Qualitatively, it is easy to see for a large number of taps how the interconnections from the S/H bank to the FFE taps and finally the summers expand the area of the FFE-DFE summer core leading to an even higher power penalty. For such a high number of taps (16) as required for the wireless 60GHz NLOS channels, the area and power overheads of time interleaving make it more efficient to use the proposed swiching matrix based architecture.

Now that the advantages of using a switching matrix are well understood, we shall next describe the other key circuit blocks for the FFE, namely the FFE weight design, the interleaved S/H, and the clocking circuitry around them to realize RX-FFE functionality.

4.3 Key Circuit Blocks

4.3.1 FFE Weight Design

After the switching matrix design, the next important design aspect of the FFE is the implementation of coefficient multiplication. As compared to a DFE in which coefficients are multiplied by a single-bit input, an FFE involves multiplication with analog inputs of multi-bit precision.

Prior art on FFE design uses primarily two types of techniques to implement analog weighting – translating analog values to current/gm-magnitudes, or into time pulses. [40] uses binary weighted g_m-cells [40] to implement adjustable coefficient weighting. The primary issue with this technique is that the division of total transistor width into binary segments quickly runs into minimum device size constraints. For a multi-coefficient FFE, this would incur a substantial power penalty. Among the time-based techniques, [61] translates the analog data voltage to a proportionally wide time pulse, for whose duration a capacitor is integrated by a current proportional to coefficient weight. [59] instead translates the coefficient weight into a time-pulse, for whose duration the analog data's g_m cell-based current is integrated. Since both these techniques move the resolution into the time domain requiring rail-to-rail swing clock pulses, they incur a power penalty. [59] suffers from an additional penalty from the per-tap phase interpolation requirement.

To avoid the power-resolution tradeoff, this design used an FFE coefficient based on a current-DAC-based g_m-cell with a variable tail current, as shown earlier in Fig. 4.11. While the g_m-cell does introduce non-linearity at lower gain settings (as shown by simulation in Fig. 4.15) when it operates in sub-threshold, the effect of non-linearity is balanced out by the lower magnitude of gain. By simulating across representative channel instantiations, it was ensured that the combined effect of this non-linearity of all FFE taps is below the targeted noise floor. It is important to note that while the design has two coupled sources

Figure 4.15: Simulated FFE tap distortion vs. input amplitude, across different tap gain settings (max. gain = 1). At max. gain, V* = 200mV.

Figure 4.16: FFE: Measured coefficient weight (full-scale normalized to 1) vs. digital code setting.

Figure 4.17: S/H circuitry with feedthrough cancelation (intrinsic inter-finger capacitors are shown with dotted lines) and dummy switches for charge injection cancelation.

of non-linearity – i.e. from the signal swing and from the coefficient weighting for small weights – it nevertheless exploits the relatively high targeted BER of 1E-3 to trade linearity with power. If a lower BER were to be required, it would be favorable to use one of the previously described techniques that effectively decouple the non-linearity from the input signal and the coefficient weighting. Due to the non-linear nature of the gain vs. tail current characteristics (Fig. 4.16), for an effective 6-bit g_m resolution (which ensures sufficiently low quantization noise for BER $< 10^{-3}$), the FFE DACs need to be designed with 8-bit current resolution.

4.3.2 Input S/H

Fig. 4.17 shows a single slice of the S/H bank at the input of the switching matrix. For charge injection cancelation at every clock transition, the PMOS sampling switch is provided with a dummy switch. In this commercial LP process without low-V_{th} transistors, since the PMOS transistor threshold is equal or higher than $Vdd/2$, the clock drivers to both sets of PMOS switches are skewed to favor the falling transition. In the 'hold' phase when the switches are off, the inter-finger S-D metal capacitance of the sampling switches (shown as dotted capacitors in Fig. 4.17) causes feedthrough from the input. This creates an input pattern-dependent coupling on the sampled outputs. To mitigate the effect of this coupling, identical cross-coupled metal capacitors are added.

Figure 4.18: Ring Counter: (a) Schematics and (b) CLK Waveforms.

Since the input of the S/H bank is the farthest point in the receiver cascade where a continuous time signal exists, it is also used for clock and data recovery (CDR). For bang-bang phase detection, data and edge samples ($d, uneq$ and $e, uneq$ respectively) are detected by using two additional S/H legs followed by preamp/slicer pairs. These S/H legs and slicers are driven by full-rate opposite-phased clocks to obtain 2X oversampling. While the edge sample contains ISI, since the targeted SNR (cursor to rms-noise ratio) of 9dB is not very high (i.e. the received signal is fairly noisy), an additional edge-equalizer it is not worth the power consumption that it would incur. The unequalized edge updates may be somewhat compensated by using a relatively low bandwidth for the CDR loop. It was seen by simulation and verified eventually by measuring the prototype that using unequalized edge samples does not significantly alter the baseband clock locking phase.

The sliced full-rate unequalized data samples $d, uneq$ are used to provide early/late information for CDR and, information for sign-sign LMS-based FFE coefficient adaptation. Updates for FFE coefficients are computed by correlating the sign of these data samples with the sign of error samples off the data-level (dLev) slicer at the equalizer output[62].

4.3.3 CLK Design

Since the switching matrix and its input S/H slices need multi-phase clocks with different duty-cycles, the next key block is clock generation. At the heart of the clock generation is the 16-bit ring counter (Fig. 4.18) that generates the UI-width 16-phase clocks $CP \langle 0 : 15 \rangle$ for the switching matrix. The half-UI-width signals for driving the input S/H legs ($\overline{CS \langle 0 : 15 \rangle}$) and dummy switches ($CS \langle 0 : 15 \rangle$) are generated by AND-ing the full-UI-width CLK phases with the full-rate CLK signal.

To ensure a smooth handover during the CP clock transitions and ensure that each gm-cell is always connected to at least one output, the cascode drivers are skewed to favor the

Figure 4.19: Monte-Carlo simulation of post-layout extracted $CP\langle2\rangle$, $CP\langle3\rangle$ at 5GS/s.

rising transition. Since the prototype was implemented in a relatively slow technology node, the generation of UI/2 (100ps) pulses at 5GS/s required the clock distribution to be implemented with conservative fanout-of-2 (FO2) chains. Since this leads to a relatively high insertion delay, the CP clock drivers at the cascodes are prone to mismatch and therefore delay variation through the distribution network. The post-extraction Monte-Carlo simulation (100 runs, device mismatch only) of two such clocks ($CP\langle2\rangle$ and $CP\langle3\rangle$) in Fig. 4.19 highlights a 15ps variation (1σ) in the arrival of the clocks. If the following clock – in this case, $CP\langle3\rangle$ – arrives before the current data has completed integrating at the output, the resultant overlap from the next data causes precursor ISI. While this ISI may be adaptively canceled by the FFE itself, the resulting noise enhancement degrades the BER.

For the S/H network, to ensure minimal residual charge injection, both $CS\langle0:15\rangle$ and $\overline{CS\langle0:15\rangle}$ driving the relatively high-V_{th} PMOS sampling and dummy switches (respectively) are skewed to favor falling transitions. Similarly, the clocks to the bridge reset PMOS and sign-select MUX are also skewed to favor falling transitions.

4.4 Conclusion

This chapter describes the design of an efficient receive-side feedforward equalizer (RX-FFE) supporting many coefficients. Unlike TX-FFEs which may be implemented conveniently using digital delays (i.e. flip-flops), RX-FFEs require analog delay elements. While analog delay may be implemented with sample/hold (S/H) blocks using sampling capacitors and switches, a S/H cascade suffers from an accumulation of gain mismatch and thermal noise. Prior RX-FFE art [39][40] dealt with these issues by using a time-interleaved S/H bank to avoid the cascade. However, these techniques further necessitated the use of either high-speed FFE DACs [39] or interleaved FFE summing/slicing [40] with significant repitition of analog circuitry – both of which incur a high power overhead while implementing multiple FFE coefficients.

The proposed architecture obviates high power requirements by using a switch matrix to create an analog delay line. The functionality is achieved by effectively implementing an N-parallel N-to-1 re-serializerization of the interleaved S/H bank. Through an analytical framework, the proposed technique is shown to be 2-3X more efficient in power than prior RX-FFE art while supporting 16 FFE taps.

Having designed an efficient architecture to implement an analog delay line, the final design challenge in implementing an NLOS equalizer is to enable summing many FFE and DFE coefficients together. The next chapter describes the implementation of this multi-coefficient summer and a proof-of-concept 32-coefficient FFE and 100-coefficient DFE in 65nm LP CMOS.

Chapter 5

32-Coefficient FFE, 100-Coefficient DFE Prototype

So far in this dissertation, we have talked about two prime aspects of RX equalizer design – firstly, mixed-signal DFE architectures for energy efficient cancelation of post-cursor ISI (Chapter 2, 3), and secondly, analog delay line architectures to enable FFE design for efficient cancelation of pre-cursor ISI (Chapter 4). The final aspect towards completing the design is that of an efficient summer to add many FFE, DFE coefficients together. In this chapter, we will realize this summer by combining the two most efficient summing techniques discussed in this dissertation – cascode current summation (Chapter 2) and current integration [50] (Chapter 3). Enabled by this summer and a switching matrix-based analog delay line (Chapter 4), we demonstrate an NLOS 60GHz equalizer supporting 32 FFE and 100 DFE I/Q coefficients. We shall also discuss the key design aspects of this 65nm LP CMOS prototype that enable equalizer operation up to 8Gb/s QPSK.

5.1   FFE-DFE Summer Circuit Design

This section describes the circuit design of the 32-coefficient FFE and 100-coefficient DFE summer. The discussion first focuses on designing the summer to add all 132 coefficients as currents at a single node. The second part of the discussion describes techniques to meet the stringent timing margins at sampling rates as high as 5GS/s in the relatively slow 65nm LP CMOS process.

5.1.1   Summer Design

To achieve the summing of 32 FFE and 100 DFE coefficients in a power-efficient manner, the equalizer utilizes a combination of cascode current summation (as first proposed in Chapter 2) and current integration [50] (analyzed in detail in Chapter 3). Like the switching matrix, a long-channel PMOS current source with Miller-compensated CMFB prevents a
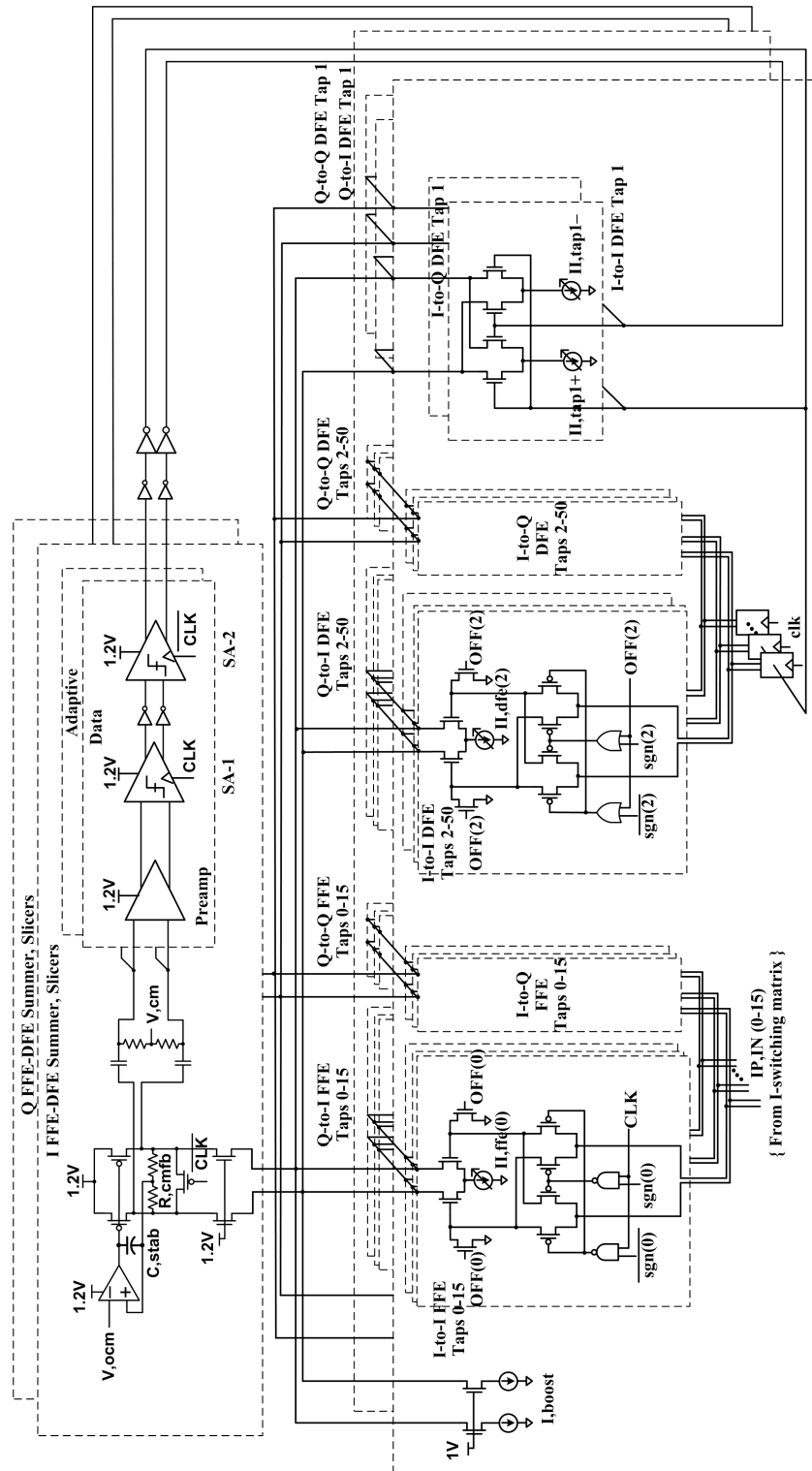
Figure 5.1: Detailed schematics of the I/Q 32-coefficient FFE, 100-coefficient DFE summer-integrator.
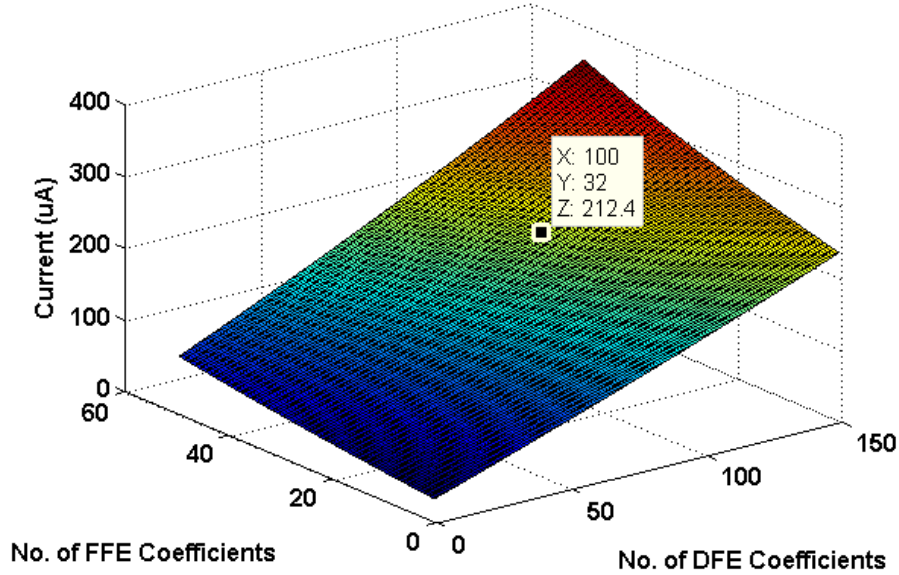
Figure 5.2: FFE-DFE Summer: FFE cursor current ($I_{FFE,0}$) vs. No. of FFE, DFE Coefficients, for $f_s = 5$GS/s, $G = 1$, $V^* = 200mV$, $C_L = 10fF$, $C_{f,FFE} = C_{f,DFE} = 1fF$, PMOS load L = $0.18\mu m$.

large drop in the output common mode voltage during integration. As with the cascode current summation technique, to boost the bandwidth of the heavily loaded cascode source node, common mode current (I_{boost}) is added as shown in the summer schematics in Fig. 5.1. The cascode source therefore does not need a bridge reset switch.

To compute the summer power consumption in terms of number of FFE and DFE coefficients (N_{FFE} and N_{DFE} respectively), an analysis is carried out in Appendix C. The analysis shows (equation (C.12)) that to support a data-rate f_s, gain G, and load capacitance C_L, the required FFE cursor current $I_{FFE,0}$ is:

$$I_{FFE,0} = \frac{(2GV^*_{FFE}f_sC_L)\left(\beta_{0,N} + \beta_{1,FFE}N_{FFE} + \beta_{1,DFE}N_{DFE}\right)}{1 - \left(\frac{2\ GDR}{\omega_{T,FFE}}\right)\left(\beta_{0,D} + \beta_{2,FFE}N_{FFE} + \beta_{2,DFE}N_{DFE}\right)} \tag{5.1}$$

In this equation, the βs are ratios expressed in equations C.13).

The variation of $I_{FFE,0}$ as a function of the number of FFE and DFE coefficients as shown in equation (5.1) can be visualized in Fig. 5.2, for $G = 1$, $f_s = 5GS/s$, and $C_L = 10fF$. For the designed summer with 32 FFE and 100 DFE coefficients, it can be seen that the variation is fairly linear with the number of taps, showing that the structure is away from its self-loading limit.

Fig. 5.3 plots the effective cascode bandwidth (simulated) and shows that it is >7.5GHz. This confirms that the summer maintains high cascode bandwidth in spite of the capacitive loading from the multiple coefficients. The maximum total current consumption for the two I/Q current summer-integrators with 32 FFE, 100 DFE coefficients each while canceling
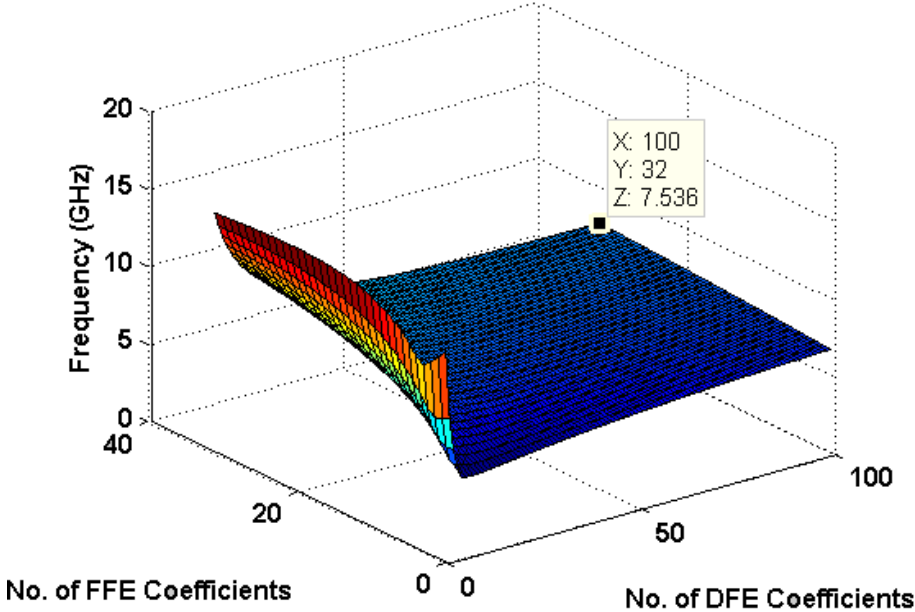
Figure 5.3: FFE-DFE Summer: Obtained integrator pole frequency vs. No. of FFE, DFE Coefficients.
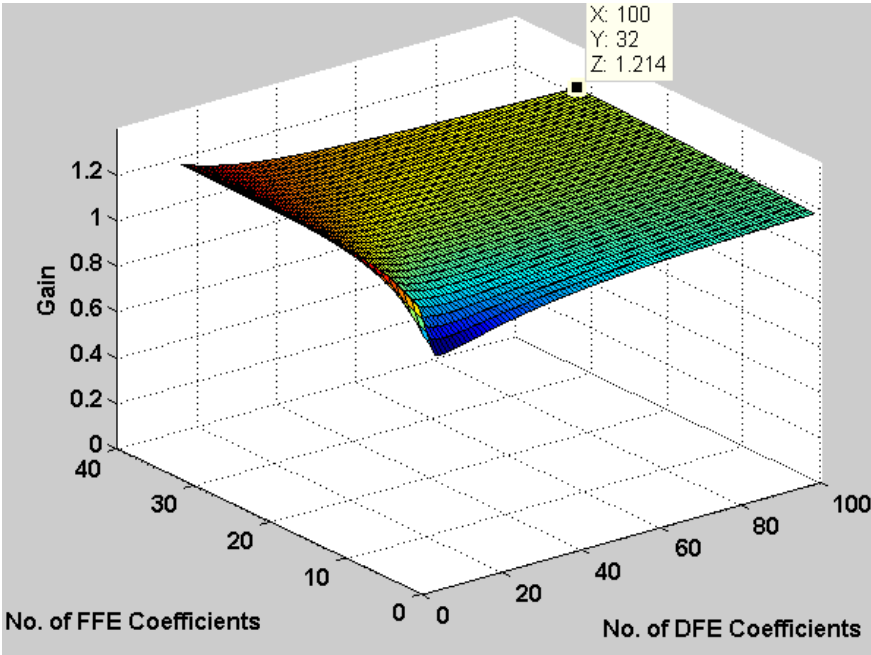
Figure 5.4: FFE-DFE Summer: Obtained gain vs. No. of FFE, DFE Coefficients. Targeted gain=1.

at most 2X precursor ISI, 2X post-cursor ISI and totally 3X ISI (all in terms of the FFE cursor magnitude) is only $4mA$. This demonstrates the extremely high energy-efficiency of cascoded current integration.

5.1.2   Tap-1 Feedback

While the summer-integrator design makes it extremely power-efficient to add multiple FFE and DFE coefficients, there are additional issues that need to be solved to enable reception at data-rates as high as 4GS/s. The key design challenge is to satisfy the DFE coefficient feedback latency of 250ps in a 65nm LP CMOS process that does not include any low-V_t devices and is hence 2X slower than the general purpose (GP) process in the same technology node. This section details the design of a fast slicer feedback configuration for the 1st DFE tap.

The dynamically regenerated outputs of the Strong-Arm typically need to be latched during its precharging phase. The latching stage typically adds a significant delay overhead to the feedback path. In this design, however, the integrate-reset nature of the summer is leveraged to get rid of the slow latch and replace it with another Strong-Arm, as seen earlier in Fig. 5.1. This second Strong-Arm (SA2) is clocked oppositely to the first slicer (SA1) and shares its precharging clock phase (CLK=1) with the reset phase of the summer. The total power consumption of all I/Q preamps and comparators (consisting of the unequalized data, edge comparators, and the post-summation data, adaptive comparators) is 3mW at 4GS/s. Finally, the timing overhead from sign selection for the 1st DFE taps is eliminated by embedding it into the tap as a Gilbert structure. The cost of this tap embedding is only a very marginal increase in the capacitance of the cascode source.

While SA2 is faster than SA1 by account of its CMOS-level inputs, it needs to be buffered to drive the differential legs of both direct and cross taps and the long wire feeding back to the taps. As a result, the total SA2 delay eats into the integration time window of tap-1. To achieve the same maximum ISI cancelation capability as the other taps, in order to integrate equal voltage in the half clock cycle of integration, the full-scale tap-1 currents need to be made larger than that of taps 2-50. If the buffered SA2 delay is T_{SA2}, the size-up factor is $\dfrac{1}{1 - \dfrac{T_{SA2}}{(UI/2)}}$. To achieve a peak data-rate of 4GS/s, the tap-1's are sized up by 2.5X. To ensure that quantization noise remains nearly unchanged, the tap-1 current DACs are provided with one additional bit of resolution.

Fig. 5.5 summarizes the clocking operation of the RX-cascade starting with the FFE (interleaved S/H, and the integrate/reset-based switching matrix) and ending at summing amplifier (integrate/reset-based summer, Strong-Arm latch-based feedback to DFE tap-1). As with the cascode-summation DFE prototype demonstrated in Chapter 2, the power dissipation of clocking the feedback shift register chain dominates total power consumption. At a data-rate of 4GS/s (8Gb/s QPSK), the entire clocking power including DFE flip-flops, tap drivers, and clock distribution is 33.3 mW at 8 Gb/s, i.e. an efficiency of 4.2 mW/Gb/s.
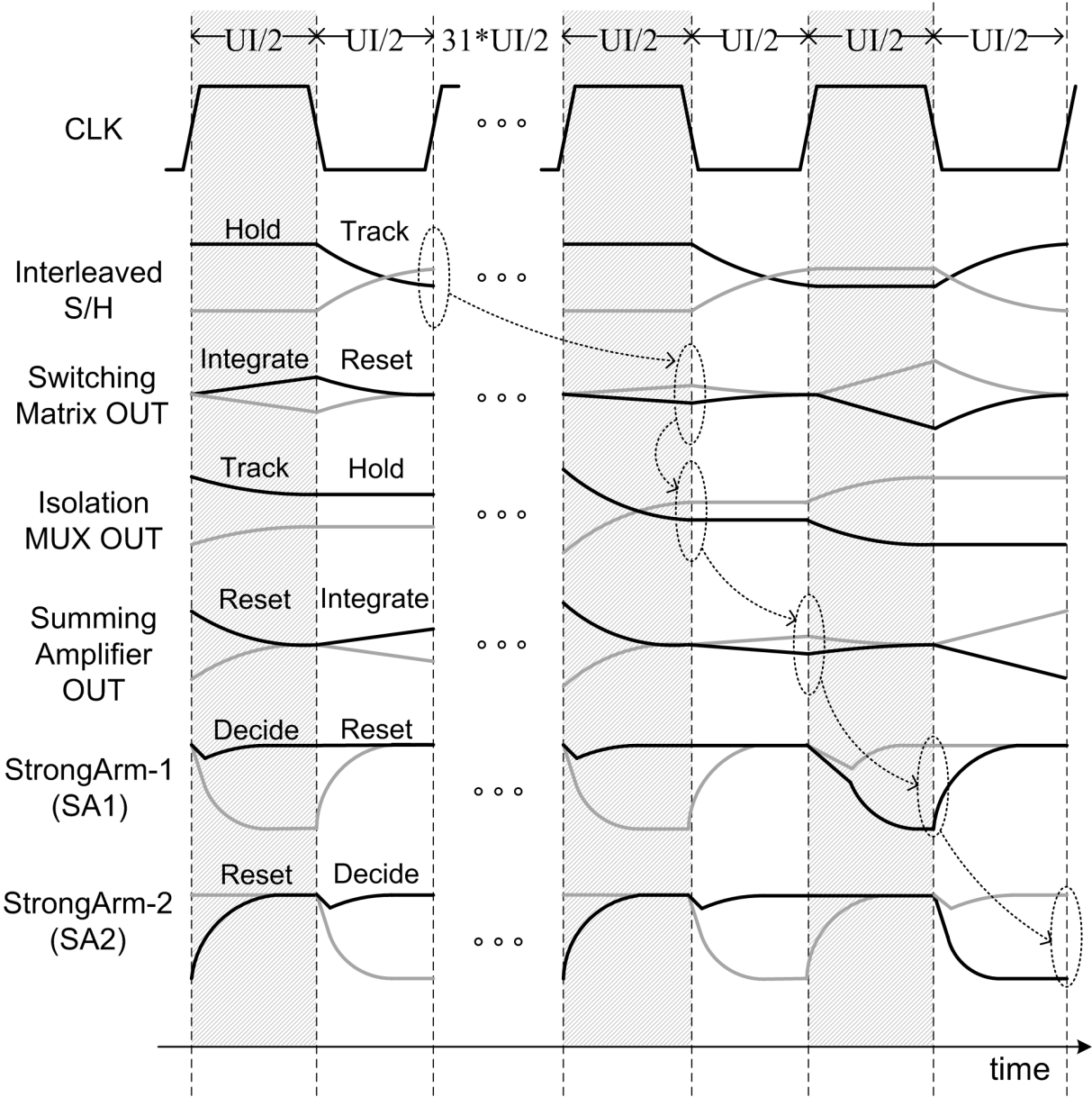
Figure 5.5: Step-wise details of clocking mechanism through the RX FFE/DFE cascade.
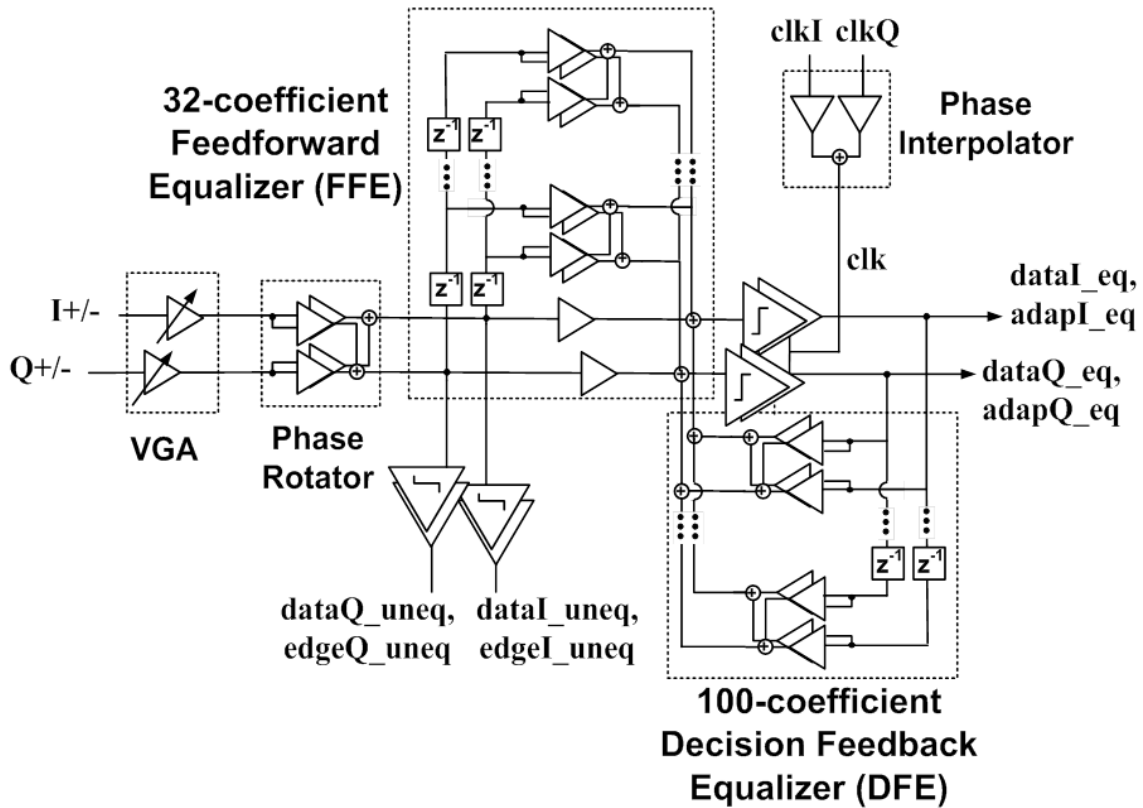
Figure 5.6: Block diagram of prototype NLOS 60GHz baseband receiver with 32-coefficient complex FFE and 100-coefficient DFE. The baseband also includes variable gain amplifiers, phase rotator, and phase interpolation circuitry.

To improve power efficiency of clocking at lower data-rates, the clock distribution supply voltage is lowered.

Now that the key design aspects have been discussed, the next section describes the prototype baseband test-chip (Fig. 5.6) that was fabricated to validate the proposed designs.

5.2   Prototype Measurement Results

The proposed 32-coefficient FFE and 100-coefficient DFE were implemented as a part of a 60GHz NLOS baseband test-chip which was fabricated in a TSMC 65nm LP CMOS process. As shown in Fig. 5.6, the chip also includes a variable-gain amplifier (VGA), analog phase rotator, and baseband phase interpolator circuitry. Fig. 5.7 shows a die-microphotograph [63]. The chip was tested by using a 4GS/s 2-channel Arbitrary Waveform Generator (AWG). All analog blocks were designed to enable QPSK operation up to 10Gb/s. However, due to the relatively slow LP process which hindered the slicer resolution speed and first DFE tap
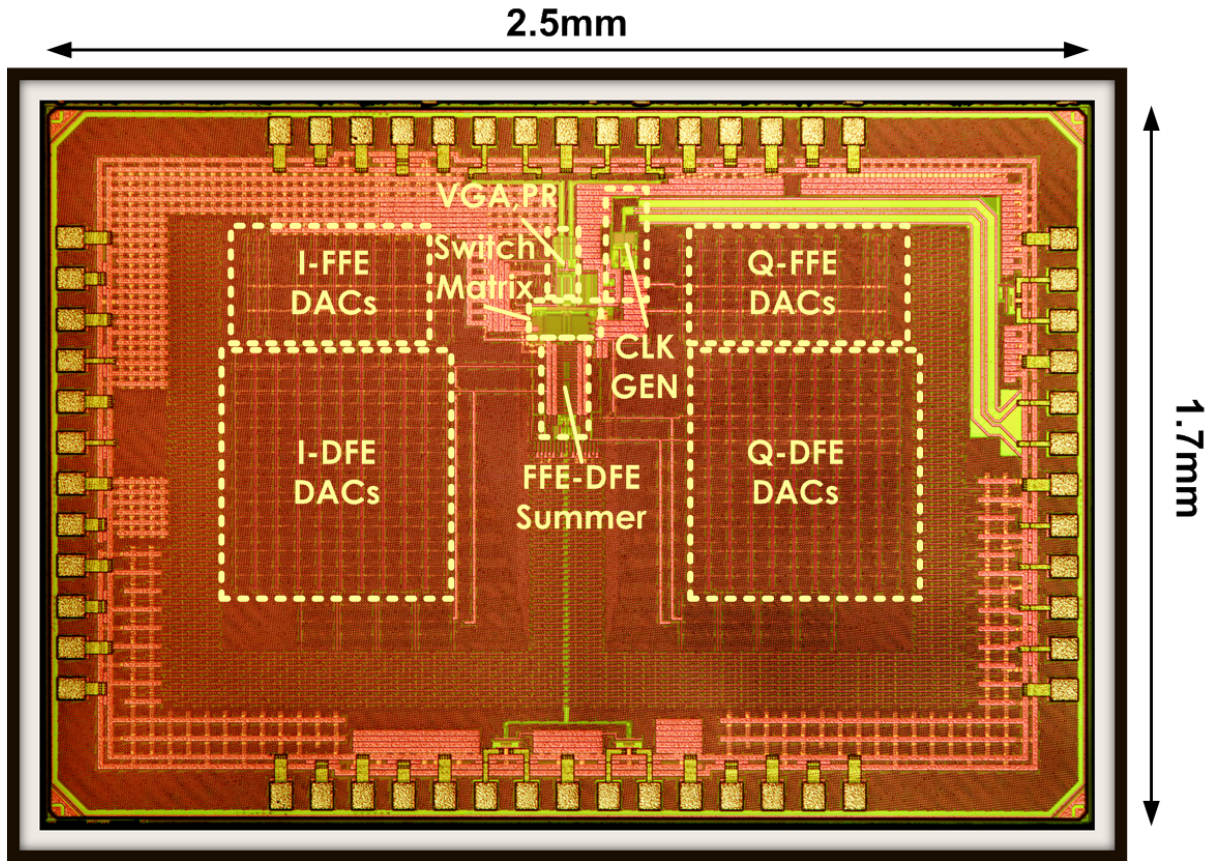
Figure 5.7: Die microphotograph overlaid with key design blocks.

feedback latency, the maximum achievable speed for the DFE after post-layout simulation was 8.7Gb/s.

Firstly, the FFE coefficient weights were characterized as a function of their respective digital code setting, by sweeping the adaptive comparator offset to determine signal swing. Fig. 5.8 shows the measured weight of the first FFE pre-cursor coefficient at 8Gb/s operation, normalized to maximum coefficient weight. Since FFE weight is changed by linearly stepping through its g_m cell bias current, this curve is nonlinear as expected with larger gain steps at lower code settings when the FFE g_m cell operates in subthreshold. In comparison, the DFE weight is measured to change linearly with code setting.

Next, the weight of the first DFE coefficient was measured as a function of data-rate, as seen in Fig. 5.9. For this coefficient, since the weight changes directly according to feedback delay from the slicer's resolution of a small-signal to CMOS levels, this measurement is particularly important. Since a higher data-rate amounts to larger feedback delay and a smaller integration time (both, in terms of cycle time), the first coefficient weight (in terms of cursor magnitude) falls at higher data-rates. Beyond the designed maximum data-rate of 8.7Gb/s, the coefficient weight falls rapidly to zero.
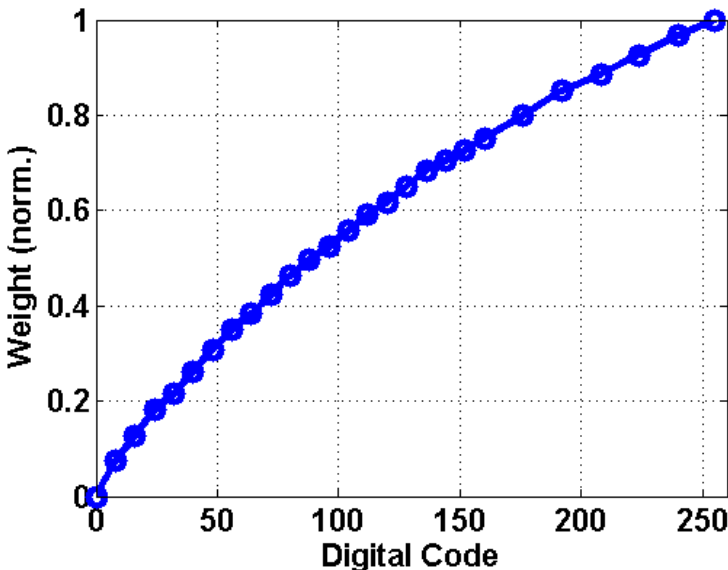
Figure 5.8: FFE: Measured coefficient weight (full-scale normalized to 1) vs. digital code setting.
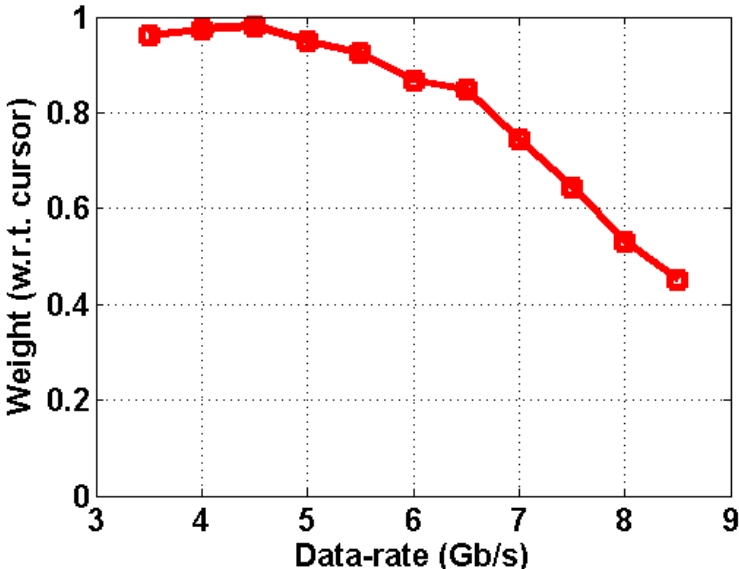
Figure 5.9: DFE: Measured weight of 1st coefficient vs. data-rate.
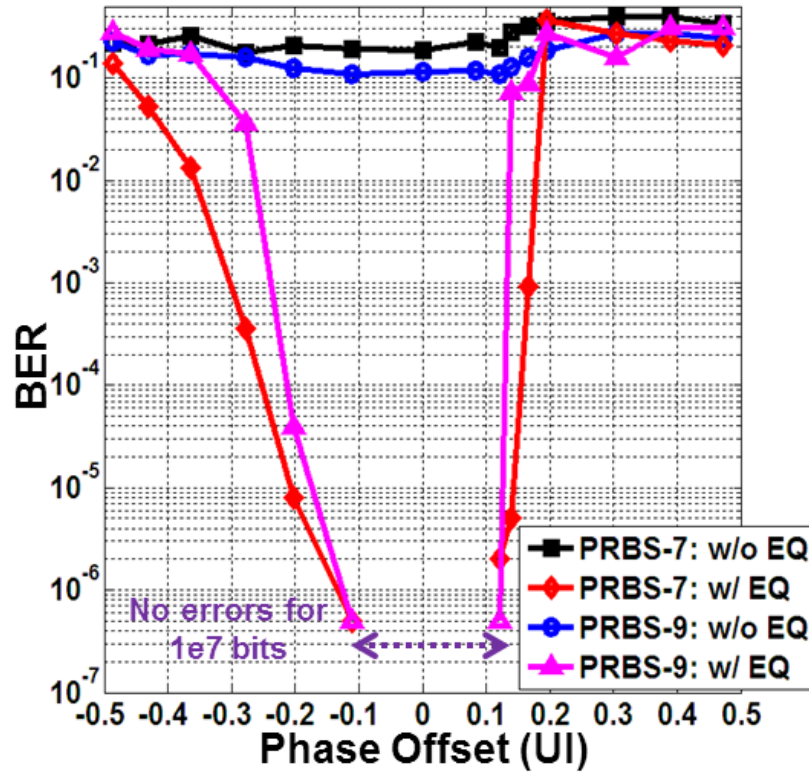
Figure 5.10: Bathtub curve: BER vs. CLK phase (UI) offset, with 8Gb/s effective through-put from PRBS-7 and PRBS-9 on I and Q channels, canceling total ISI of 2X cursor strength.

The AWG was then programmed to mimic a conference room NLOS 60GHz channel with 12ns delay spread and total ISI magnitude of 2-3X the cursor as well as white noise while generating $2^7 - 1$, $2^9 - 1$ PRBS data on the I, Q channels. The ISI in the AWG was programmed as per the WiGig channel models [24] (as discussed in Chapter 1). Following equalization, as shown by the bathtub plot in Fig. 5.10, the stand-alone baseband achieves error-free operation over 10^7 bits.

With white noise added to the input of the baseband, Fig. 5.11 shows the variation in measured BER vs. cursor-to-thermal-noise ratio for a LOS channel with ISI magnitude of 2X the cursor. Here, the cursor strength refers to the signal amplitude without ISI. For comparison, the ideal curve for an AWGN channel and that for an ideal MMSE equalizer are also included. The measured curve differs from the ideal MMSE equalizer due to uncanceled data-dependent feedthrough from the switching matrix S/H circuitry. Since there is no correlation between the interleaving factor and the PRBS sequence length, the feedthrough appears like random noise with magnitude proportional to the signal swing, which degrades the effective SNR and hence the BER. Additionally, the FFE and DFE weighting coefficient quantization noise shows up at higher SNR, further degrading the BER.

Fig. 5.12 shows the scaling in power consumption and power efficiency with data-rate.
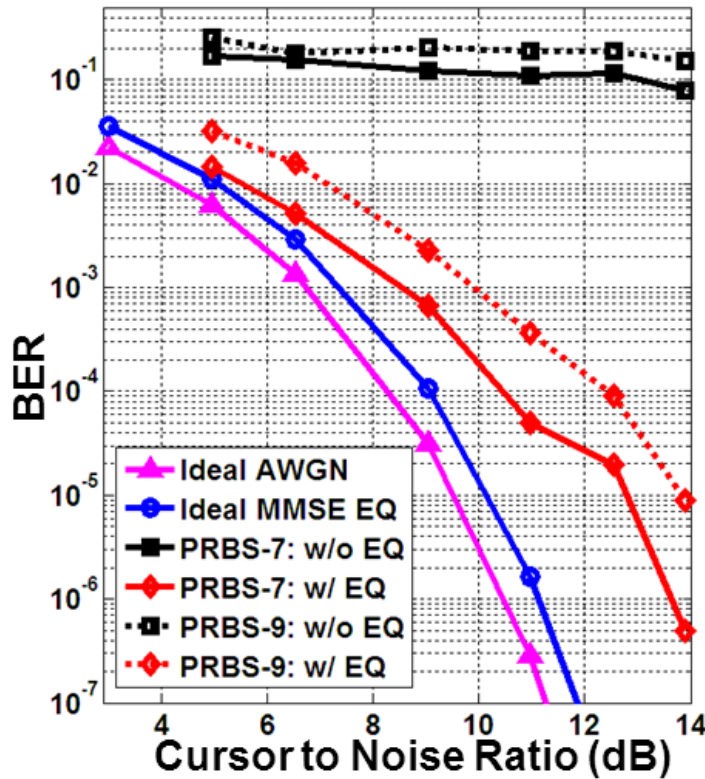
Figure 5.11: Measured BER vs. SNR (cursor to thermal noise ratio), and comparison with an ideal AWGN RX and an ideal MMSE equalizer.

At higher data-rates, the power consumption is dominated by digital power due to the clock distribution. To achieve improved digital power efficiency at lower data-rates, the digital supply for clock distribution is lowered. At lower data-rates, the bias current of all analog circuitry is also lowered; however, due to linearity constraints, this current cannot be scaled proportionally. As a result, the analog bias current consumes a greater share of the total power at lower data-rates. At the WiGig-specified QPSK data-rate of 3.5Gb/s, the overall efficiency is 7pJ/bit, which is 20% better than at peak data-rate of 8Gb/s. Since the delay spread of the channel is fixed, a smaller number of equalizer coefficients would actually be required at lower rates. If a portion of the DFE digital shift register was accordingly power-gated, the total power and efficiency would scale with as shown for the dotted curves in Fig. 5.12. If such gating was available, the improvement in efficiency at 3.5Gb/s would be 30%.

Fig. 5.13 and Table. 5.1 compare this design to prior RX-equalizer art for 60GHz channels. This prototype achieves 2.3X higher throughput than [31] and up to 3.4X better power efficiency than [21]. These results should be viewed in light of the LP flavor of process technology for this test-chip, which has almost 2X lower speed than the GP process on the same node. As noted in Fig. 5.13, mixed-signal techniques not only enable low-power
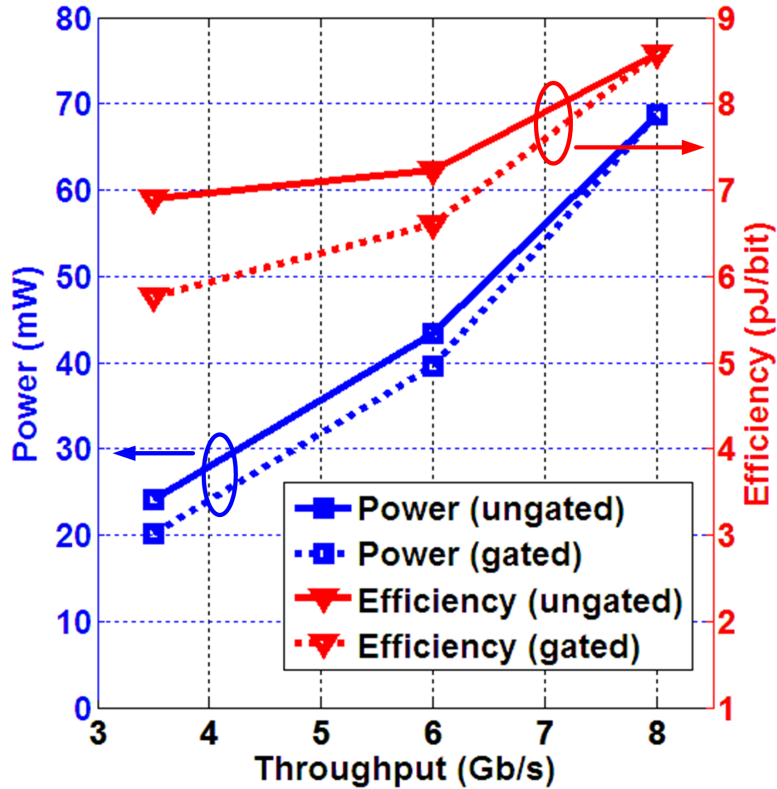
Figure 5.12: Total measured power and efficiency vs. throughput. 'Gated' refers to power gating the latter DFE flip-flops at lower data-rates for equal ISI delay-spread.

equalizer designs for standard specified data-rates of 3.5Gb/s, but also maintain this power efficiency at higher throughputs.

5.3 Conclusion

In this chapter, we examined the design of mixed-signal summers supporting many equalizer coefficients. The summer design was developed by building upon two of the most effective techniques proposed/discussed earlier in this dissertation – namely cascode current summation (Chapter 2) and current integration [50]. By using a combination of the above-mentioned techniques, the proposed cascode current-summing I/Q integrators can support 100 DFE and 32 FFE coefficients at sampling rates up to 4GS/s while consuming only $4mA$ of total static current.

In spite of being able to implement such a low-power summer, the total power consumption of the equalizer ends up being dominated by the digital power, mainly due to clock distribution for the DFE flip-flops and ring-counter-based clock generation for the FFE. Nevertheless, the analog and mixed-signal techniques developed in Chapter 4 and
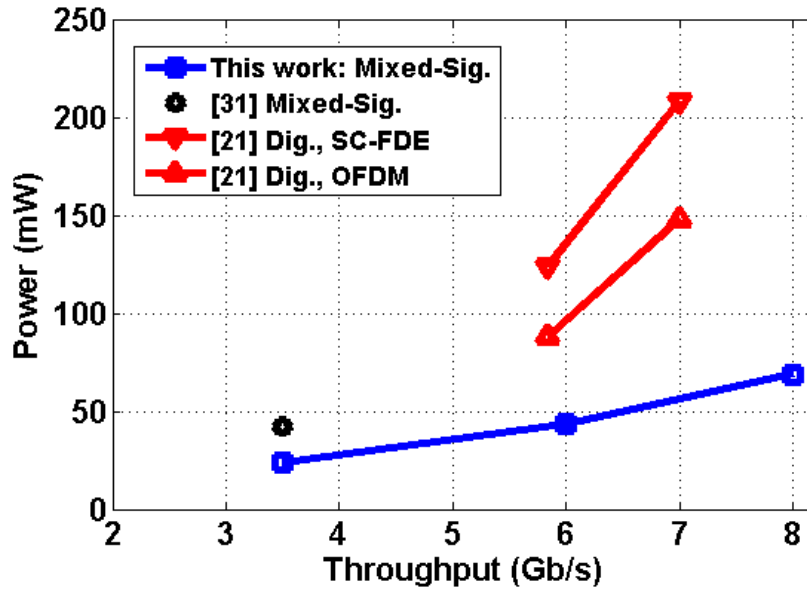
Figure 5.13: Measured power vs. throughput, and comparison with prior art

	[31]	[21]	**This Work**
Technology	65nm GP CMOS	65nm GP CMOS	**65nm LP CMOS**
FO4 Delay	1X	1X	**2X**
Equalizer	FFE, DFE	OFDM/SC-FDE	**FFE, DFE**
Time/Frequency Domain	Time-domain	Frequency-domain	**Time-domain**
Digital/Mixed-Signal	Mixed-Signal	Digital	**Mixed-Signal**
Total I/Q Coefficients			
FFE	12	NA	**64**
DFE	70	NA	**200**
FFT	NA	512-pt	**NA**
Modulation	QPSK	16-QAM	**QPSK**
Sampling Rate	1.76 GS/s	1.76 GS/s	**4.0 GS/s**
Highest Throughput	3.5 Gb/s	7.0 Gb/s	**8.0 Gb/s**
Power	42 mW	148/208 mW	**66 mW**
Efficiency	11.9 pJ/bit	21.0/29.5 pJ/bit	**8.3 pJ/bit**
Energy/Coefficient	145.5 fJ/bit/coef.	NA	**31.3 fJ/bit/coef.**

Table 5.1: Comparison with prior 60GHz NLOS equalizer art.

this chapter make it feasible to achieve time-domain equalization capabilities as powerful as digital frequency-domain FFT-based equalizers at only a fraction of total power consumption. Equally importantly, these mixed-signal techniques enable scaling to significantly higher sampling rates than digital solutions, while simultaneously maintaining good energy efficiency.

Chapter 6

Conclusion

The need for an individual to be ubiquitously connected is sharply increasing data-rate requirements on both user and infrastructure sides of all communication links. The wireless medium that is most typically used by handheld/portable devices is now capable of supporting multi-Gb/s links courtesy of newer standards that have been made available over the past decade. Wireless communication is typically limited by multi-path reflection-based inter-symbol interference (ISI) over relatively large time delay spreads. While ISI may be mitigated by using equalizers, the need for multi-Gb/s communication creates stringent requirements for these circuits primarily in terms of the sheer number of equalizer coefficients. The infrastructure side of the link, primarily consisting of network routers and servers, uses wired links which have shorter ISI delay spreads. However, since such backbone systems need to support data-rates of almost an order of magnitude higher, this leads to substantial channel equalization requirements. This dissertation therefore focuses on the energy-efficient implementation of all such multi-coefficient high-speed equalizers. To demonstrate the efficacy of the proposed equalizer designs, the platform used by this dissertation is the wireless 60GHz channel which is capable of multi-Gb/s communication. However, the design techniques and frameworks developed are equally relevant for any communication link/standard requiring multi-coefficient feedback and feedforward equalization.

Commercial 60GHz radios employ conventional multi-bit OFDM-based wireless baseband solutions using digital signal processing (DSP) classically used for low data-rate wireless. Since these techniques are power intensive at Gb/s rates, we present a mixed-signal approach to the design. 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 low power alternative. 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 mixed-signal decision feedback equalizer (DFE) has been shown in literature to be an excellent candidate to cancel the post-cursor portion of the ISI. 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 channels with a long delay spread, such as that of a 60GHz channel. A cascode current-summing structure then was proposed to relax these self-loading constraints [49] by making a key observation about wireless channels that not all ISI taps concurrently have maximum magnitude. Therefore, by summing the ISI cancelation currents through a cascode transistor, this proposed structure can equalize a long ISI profile that is typical of a line-of-sight (LOS) 60GHz channel response.

Wireless channels that are non-line-of-sight (NLOS) or have only moderately directional RF front-ends have even longer delay spreads that cannot be combated by just using cascode current-summation. Therefore, as a next step we reviewed relatively recently introduced summation techniques in literature using current integration [50] and switched capacitor-based voltage feedback [52]. The analysis of these techniques showed that while the switched capacitor technique is promising, its efficiency is severely limited by the dependence of power dissipation on the power dissipation of the digital driver circuitry. Current integration from a capacitive load with per-symbol reset allows for the use of low-bandwidth summation, thus reducing current consumption and significantly easing the self-loading from equalizer coefficients. As a result, current integration is the most optimal to implement coefficient summation. A combination of current integration and the above-mentioned cascode-based summation is therefore the most promising summation technique.

In addition to longer post-cursor delay spreads, reception over NLOS channels is affected by pre-cursor ISI, which cannot be mitigated by a DFE. To counter pre-cursor ISI, an analog receive-side feedforward equalizer (RX-FFE) is presented. The primary challenge with designing RX-FFEs is the implementation of an efficient analog delay line. Most prior designs avoid the problems of analog delay cell cascades – namely noise and gain mismatch accumulation – by using interleaved sampling with either reconfigurable FFE weighting or time-interleaved summing. Reconfiguration using per-UI switching DACs [39] leads to high digital driver power dissipation, while interleaved weighting and summation [40] causes area and wiring overhead, eventually leading to a power penalty. For a large number of FFE taps – as required for equalizing an NLOS 60GHz channel – these techniques are inefficient. The need for high-speed DACs or heavy interleaving is obviated by re-serializing the interleaved samples using a proposed switching matrix architecture, which enables reserialization of the interleaved samples before summation. Using this architecture enables low-power feedforward equalization with 2-3X better efficiency than prior RX-FFE art.

The proposed techniques for energy-efficient DFE and FFE implementation are validated by two prototypes in 65nm CMOS. The first prototype is a cascode current-summing DFE in 65nm 1V 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 [49]. The second prototype is a 16 complex-tap RX-FFE and 50 complex-tap DFE in 65nm LP 1.2V CMOS with only standard-V_t transistors that was shown to operate with data-rates of 3.5-8Gb/s for BER less than 10^{-6} while consuming 25-67mW of power [63]. If the second prototype was implemented in a 65nm GP process with low-V_t transistors (the same process

as the first prototype), it would dissipate ∼50% lower total power.

At the architecture level, an important conclusion of this work is that 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. The energy efficiency of the equalizer prototypes showcased in this dissertation compares extremely favorably with OFDM and SC-FDE based solutions [21] which consume 150-200mW of power at lower sampling rates and are inefficient to scale to higher sampling rates. Both scaling and summing implementations using analog processing are extremely energy-efficient. While analog summers have fundamental limits in implementation of the number of ISI taps, these limits can be pushed to a large enough number of taps to realize energy efficient summing even for channels with very wide multi-path delay spread. However, a majority of the total equalizer power dissipation is associated with implementing delay. While digital delay implementation using flip-flops is expensive in power (as seen with both DFE prototypes), analog delay implementation using S/H – which is in turn driven by digital circuitry – is even more expensive (as demonstrated by the FFE prototype). Implementation of delay is therefore the limiting factor of the energy-efficiency of these equalizers.

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." In: *Federal Commun. Comm.* FCC ET Docket No. 99-261 (Dec. 2000).

[2] J. F. Bulzacchelli et al. "A 28-Gb/s 4-Tap FFE/15-Tap DFE Serial Link Transceiver in 32-nm SOI CMOS Technology". In: *Solid-State Circuits, IEEE Journal of* 47.12 (Dec. 2012).

[3] C. Marcu et al. "A 90 nm CMOS Low-Power 60 GHz Transceiver With Integrated Baseband Circuitry". In: *Solid-State Circuits, IEEE Journal of* 44.12 (Dec. 2009), pp. 3434–3447.

[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". In: *Solid-State Circuits, IEEE Journal of* 44.8 (Aug. 2009), pp. 2085–2099.

[5] Huaide Wang et al. "A 60-GHz FSK transceiver with automatically-calibrated demodulator in 90-nm CMOS". In: *VLSI Circuits (VLSIC), 2010 IEEE Symposium on.* June 2010, pp. 95–96.

[6] K. Okada et al. "Full Four-Channel 6.3-Gb/s 60-GHz CMOS Transceiver With Low-Power Analog and Digital Baseband Circuitry". In: *Solid-State Circuits, IEEE Journal of* 48.1 (Jan. 2013).

[7] A. Valdes-Garcia et al. "A SiGe BiCMOS 16-element phased-array transmitter for 60GHz communications". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International.* Feb. 2010, pp. 218–219.

[8] S. Emami et al. "A 60GHz CMOS phased-array transceiver pair for multi-Gb/s wireless communications". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International.* Feb. 2011, pp. 164–166.

[9] A. Siligaris et al. "A 65-nm CMOS Fully Integrated Transceiver Module for 60-GHz Wireless HD Applications". In: *Solid-State Circuits, IEEE Journal of* 46.12 (Dec. 2011), pp. 3005–3017.

[10] T. Mitomo et al. "A 2Gb/s-throughput CMOS transceiver chipset with in-package antenna for 60GHz short-range wireless communication". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*. Feb. 2012, pp. 266–268.

[11] V. Vidojkovic et al. "A low-power 57-to-66GHz transceiver in 40nm LP CMOS with -17dB EVM at 7Gb/s". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*. Feb. 2012, pp. 268–270.

[12] *WirelessHD Consortium*. URL: http://www.wirelesshd.org/pdfs/WirelessHD-Specification-Overview-v1.1May2010.pdf.

[13] *Silicon Image*. URL: http://www.siliconimage.com/.

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

[15] B. Verbruggen et al. "A 2.6 mW 6 bit 2.2 GS/s Fully Dynamic Pipeline ADC in 40 nm Digital CMOS". In: *Solid-State Circuits, IEEE Journal of* 45.10 (Oct. 2010), pp. 2080–2090.

[16] M. El-Chammas and B. Murmann. "A 12-GS/s 81-mW 5-bit Time-Interleaved Flash ADC With Background Timing Skew Calibration". In: *Solid-State Circuits, IEEE Journal of* 46.4 (Apr. 2011), pp. 838–847.

[17] Wei-Hsiang Ma, J.C. Kao, and M. Papaefthymiou. "A 5.5GS/s 28mW 5-bit flash ADC with resonant clock distribution". In: *ESSCIRC (ESSCIRC), 2011 Proceedings of the*. Sept. 2011, pp. 155–158.

[18] D. Stepanovic and B. Nikolic. "A 2.8GS/s 44.6mW time-interleaved ADC achieving 50.9dB SNDR and 3dB effective resolution bandwidth of 1.5GHz in 65nm CMOS". In: *VLSI Circuits (VLSIC), 2012 Symposium on*. June 2012, pp. 84–85.

[19] Yun-Shiang Shu. "A 6b 3GS/s 11mW fully dynamic flash ADC in 40nm CMOS with reduced number of comparators". In: *VLSI Circuits (VLSIC), 2012 Symposium on*. June 2012, pp. 26–27.

[20] I-Ning Ku et al. "A 40-mW 7-bit 2.2-GS/s Time-Interleaved Subranging CMOS ADC for Low-Power Gigabit Wireless Communications". In: *Solid-State Circuits, IEEE Journal of* 47.8 (Aug. 2012), pp. 1854–1865.

[21] F. Hsiao et al. "A 7Gb/s SC-FDE/OFDM MMSE equalizer for 60GHz wireless communications". In: *Solid State Circuits Conference (A-SSCC), 2011 IEEE Asian*. Nov. 2011, pp. 293–296.

[22] *IEEE Std., 802.15.3c-2009, Oct. 2009*. URL: http://standards.ieee.org/getieee802/download/802.15.3c-2009.pdf.

[23] *IEEE Std., IEEE802.11ad*. URL: http://standards.ieee.org/develop/project/802.11ad.html.

[24]   *Wireless Gigabit Alliance (WiGig)*. URL: http://www.wirelessgigabitalliance.org/specifications/.

[25]   K. Fukuda et al. "A 12.3mW 12.5Gb/s complete transceiver in 65nm CMOS". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*. Feb. 2010, pp. 368–369.

[26]   F. O'Mahony et al. "A 47 10 Gb/s 1.4 mW/Gb/s Parallel Interface in 45 nm CMOS". In: *Solid-State Circuits, IEEE Journal of* 45.12 (Dec. 2010), pp. 2828–2837.

[27]   M. Ramezani et al. "An 8.4mW/Gb/s 4-lane 48Gb/s multi-standard-compliant transceiver in 40nm digital CMOS technology". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*. Feb. 2011, pp. 352–354.

[28]   A.K. Joy et al. "Analog-DFE-based 16Gb/s SerDes in 40nm CMOS that operates across 34dB loss channels at Nyquist with a baud rate CDR and 1.2Vpp voltage-mode driver". In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*. Feb. 2011, pp. 350–351.

[29]   S.S. Haykin. *Digital communications*. Wiley, 1988. ISBN: 9780471629474.

[30]   D.A. Sobel and R.W. Brodersen. "A 1 Gb/s Mixed-Signal Baseband Analog Front-End for a 60 GHz Wireless Receiver". In: *Solid-State Circuits, IEEE Journal of* 44.4 (Apr. 2009), pp. 1281–1289.

[31]   Ji-Hoon Park and Borivoje Nikolic. "Power-efficient Design of Multi-Gpbs Wireless Baseband". PhD thesis. EECS Department, University of California, Berkeley, Dec. 2011. URL: http://www.eecs.berkeley.edu/Pubs/TechRpts/2011/EECS-2011-135.html.

[32]   A. Maltsev et al. "Experimental investigations of 60 GHz WLAN systems in office environment". In: *Selected Areas in Communications, IEEE Journal on* 27.8 (Oct. 2009), pp. 1488–1499.

[33]   Hao Xu, V. Kukshya, and T.S. Rappaport. "Spatial and temporal characteristics of 60-GHz indoor channels". In: *Selected Areas in Communications, IEEE Journal on* 20.3 (Apr. 2002), pp. 620–630.

[34]   M. Jacob et al. "A ray tracing based stochastic human blockage model for the IEEE 802.11ad 60 GHz channel model". In: *Antennas and Propagation (EUCAP), Proceedings of the 5th European Conference on*. Apr. 2011, pp. 3084–3088.

[35]   H.T. Friis. "A Note on a Simple Transmission Formula". In: *Proceedings of the IRE* 34.5 (May 1946), pp. 254–256.

[36]   A.A.M. Saleh and R. Valenzuela. "A Statistical Model for Indoor Multipath Propagation". In: *Selected Areas in Communications, IEEE Journal on* 5.2 (Feb. 1987), pp. 128–137.

[37] M.R. Williamson, G.E. Athanasiadou, and A.R. Nix. "Investigating the effects of antenna directivity on wireless indoor communication at 6O GHz". In: *Personal, Indoor and Mobile Radio Communications, 1997. 'Waves of the Year 2000'. PIMRC '97., The 8th IEEE International Symposium on.* Vol. 2. Sept. 1997, 635–639 vol.2.

[38] M. Tabesh et al. "A 65 nm CMOS 4-Element Sub-34 mW/Element 60 GHz Phased-Array Transceiver". In: *Solid-State Circuits, IEEE Journal of* 46.12 (Dec. 2011), pp. 3018–3032.

[39] Tai-Cheng Lee and B. Razavi. "A 125-MHz CMOS mixed-signal equalizer for Gigabit Ethernet on copper wire". In: *Custom Integrated Circuits, 2001, IEEE Conference on.* 2001, pp. 131–134.

[40] J.E. Jaussi et al. "8-Gb/s source-synchronous I/O link with adaptive receiver equalization, offset cancellation, and clock de-skew". In: *Solid-State Circuits, IEEE Journal of* 40.1 (Jan. 2005), pp. 80–88.

[41] Bang-Sup Song and D.C. Soo. "NRZ timing recovery technique for band-limited channels". In: *Solid-State Circuits, IEEE Journal of* 32.4 (Apr. 1997), pp. 514–520.

[42] Young-Soo Sohn, Seung-Jun Bae, and Hong-June Park. "A 1.35Gbps decision feedback equalizing receiver for the SSTL SDRAM interface with 2X oversampling phase detector for skew compensation between clock and data". In: *Solid-State Circuits Conference, 2002. ESSCIRC 2002. Proceedings of the 28th European.* Sept. 2002, pp. 787–790.

[43] R. Payne et al. "A 6.25-Gb/s binary transceiver in 0.13- mu;m CMOS for serial data transmission across high loss legacy backplane channels". In: *Solid-State Circuits, IEEE Journal of* 40.12 (Dec. 2005), pp. 2646–2657.

[44] Zhengya Zhang et al. "An Efficient 10GBASE-T Ethernet LDPC Decoder Design With Low Error Floors". In: *Solid-State Circuits, IEEE Journal of* 45.4 (Apr. 2010), pp. 843–855.

[45] K.K. Parhi. "High-speed architectures for algorithms with quantizer loops". In: *Circuits and Systems, 1990., IEEE International Symposium on.* May 1990, 2357–2360 vol.3.

[46] S. Kasturia and J.H. Winters. "Techniques for high-speed implementation of nonlinear cancellation". In: *Selected Areas in Communications, IEEE Journal on* 9.5 (June 1991), pp. 711–717.

[47] V. Stojanovic et al. "Autonomous dual-mode (PAM2/4) serial link transceiver with adaptive equalization and data recovery". In: *Solid-State Circuits, IEEE Journal of* 40.4 (Apr. 2005), pp. 1012–1026.

[48] B.S. 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". In: *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International.* Feb. 2007, pp. 228–599.

[49]   C. Thakkar et al. "A 10 Gb/s 45 mW Adaptive 60 GHz Baseband in 65 nm CMOS". In: *Solid-State Circuits, IEEE Journal of* 47.4 (Apr. 2012), pp. 952–968.

[50]   M. Park et al. "A 7Gb/s 9.3mW 2-Tap Current-Integrating DFE Receiver". In: *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*. Feb. 2007, pp. 230–599.

[51]   A. Emami-Neyestanak et al. "A 6.0-mW 10.0-Gb/s Receiver With Switched-Capacitor Summation DFE". In: *Solid-State Circuits, IEEE Journal of* 42.4 (Apr. 2007), pp. 889–896.

[52]   T. Toifl et al. "A 2.6 mW/Gbps 12.5 Gbps RX With 8-Tap Switched-Capacitor DFE in 32 nm CMOS". In: *Solid-State Circuits, IEEE Journal of* 47.4 (Apr. 2012), pp. 897–910.

[53]   C.D. Ezekwe and B.E. Boser. "A Mode-Matching $\delta\sigma$ Closed-Loop Vibratory-Gyroscope Readout Interface with a $0.004A^o/\mathrm{s}/\sqrt{Hz}$ Noise Floor over a 50Hz Band". In: *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International*. Feb. 2008, pp. 580–637.

[54]   T.O. Dickson, J.F. Bulzacchelli, and D.J. Friedman. "A 12-Gb/s 11-mW half-rate sampled 5-tap decision feedback equalizer with current-integrating summers in 45-nm SOI CMOS technology". In: *VLSI Circuits, 2008 IEEE Symposium on*. June 2008, pp. 58–59.

[55]   J.L. Zerbe et al. "Equalization and clock recovery for a 2.5-10-Gb/s 2-PAM/4-PAM backplane transceiver cell". In: *Solid-State Circuits, IEEE Journal of* 38.12 (Dec. 2003), pp. 2121–2130.

[56]   S. Reynolds et al. "A 7-tap transverse analog-FIR filter in 0.12 mu;m CMOS for equalization of 10Gb/s fiber-optic data systems". In: *Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International*. Feb. 2005, 330–601 Vol. 1.

[57]   A. Momtaz and M.M. Green. "An 80 mW 40 Gb/s 7-Tap T/2-Spaced Feed-Forward Equalizer in 65 nm CMOS". In: *Solid-State Circuits, IEEE Journal of* 45.3 (Mar. 2010), pp. 629–639.

[58]   Xiaofeng Lin et al. "A 2.5- to 3.5-Gb/s Adaptive FIR Equalizer With Continuous-Time Wide-Bandwidth Delay Line in 0.25- CMOS". In: *Solid-State Circuits, IEEE Journal of* 41.8 (Aug. 2006), pp. 1908–1918.

[59]   A. Agrawal et al. "A 19-Gb/s Serial Link Receiver With Both 4-Tap FFE and 5-Tap DFE Functions in 45-nm SOI CMOS". In: *Solid-State Circuits, IEEE Journal of* 47.12 (Dec. 2012).

[60]   G.R. Gangasani et al. "A 16-Gb/s Backplane Transceiver With 12-Tap Current Integrating DFE and Dynamic Adaptation of Voltage Offset and Timing Drifts in 45-nm SOI CMOS Technology". In: *Solid-State Circuits, IEEE Journal of* 47.8 (Aug. 2012), pp. 1828–1841.

[61] Sanquan Song and Stojanovic V. "A 6.25 Gb/s Voltage-Time Conversion Based Fractionally Spaced Linear Receive Equalizer for Mesochronous High-Speed Links". In: *Solid-State Circuits, IEEE Journal of* 46.5 (May 2011), pp. 1183–1197.

[62] B. Widrow et al. "Stationary and nonstationary learning characteristics of the LMS adaptive filter". In: *Proceedings of the IEEE* 64.8 (Aug. 1976), pp. 1151–1162.

[63] C. Thakkar et al. "A Mixed-Signal 32-Coefficient RX-FFE, 100-Coefficient DFE for an 8Gb/s 60GHz Receiver in 65nm LP CMOS". In: *Solid-State Circuits Conference, 2013. ISSCC 2013. Digest of Technical Papers. IEEE International*. Feb. 2013.

Appendix A

Analysis of Switched Capacitor-Based DFE

This appendix gives a detailed analysis of the hybrid current integrating switched capacitor-based DFE architecture [52] described in Chapter 3. The anaysis is used to determine the power consumption of this DFE as a function of the number of coefficients.

Fig. A.1 shows the schematics and small-signal model (single ended) of the architecture. v_{in}, G, f_s, N_{coef}, and k_{max} respectively stand for the cursor input amplitude (excluding ISI), the DC gain, the data-rate (symbols/sec), the number of DFE coefficients and the maximum per-coefficient magnitude (relative to cursor-only amplitude). Each capacitor DAC is driven by voltage V_{reg}. To isolate the unused capacitance of each the capacitor DAC coefficient from the high-speed summing node, a fraction of each capacitor DAC coefficient is gated by using a PMOS switch per leg (as shown in Fig. A.2(a) which has the least 3 significant bits ungated). Ideally, the entire capacitor could have such gating capability. However, for very small LSB capacitor sizes, the gating switch capacitance becomes comparable to DAC leg capacitance, for which, the gating switch becomes an overhead and should be removed.

The computation of current consumption (cursor current, I_{cursor}) can be done in three steps. Firstly, the full-scale capacitance per DFE coefficient, C_{coef} is calculated as a function of C_L, v_{in}, V_{reg}, k_{max}, $k_{ISI,max}$, N_{coef} and T_{FO4}. The computed value of C_{coef} is then used to calculate the total output capacitance $C_{T,out}$. Finally, by using the relation between current integrating gain G, cursor current I_{cursor}, and $C_{T,out}$, the current I_{cursor} is finally computed as a function of N_{coef}.

Firstly, the total output capacitance, $C_{T,out}$ can be written as a sum of external loading and internal transistor capacitances as:

$$C_{T,out} = C_L + Cd_{cursor} + \sum_{i=1}^{N_{coef}} C_i + N_{coef} \cdot C_{off} + Cd_{reset} \tag{A.1}$$

In this equation, C_i is the capacitance of the ith coefficient for a certain ISI cancelation configuration. C_{off} is the capacitance per coefficient independent of the on/off state of
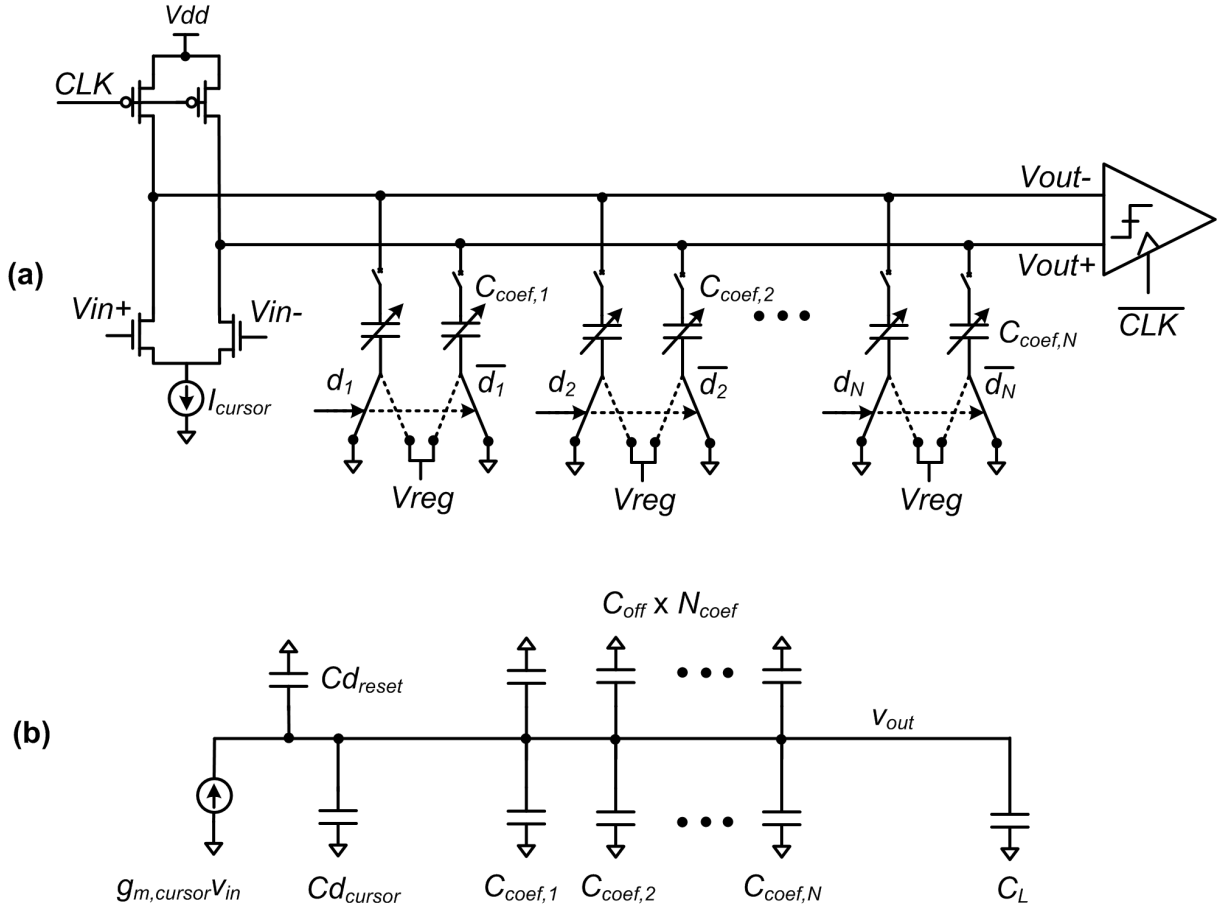
Figure A.1: Gain/bandwidth analysis of a switched-capacitor feedback-based current-integrating DFE [52]. (a) Circuit (top) and (b) Single-ended small-signal model (bottom)

its isolating switches. C_{off} comprises of the ungated capacitance for each coefficient DAC ($p \cdot C_{coef}$, with $p < 1$) and the isolation switches' drain capacitances ($C_{d,sw}$). Therefore,

$$C_{off} = p \cdot C_{coef} + C_{d,sw} \tag{A.2}$$

The proportion p of the capacitance that is ungated can be expressed in terms of the ungated DAC bits $nBits, wo$ out of the total of number of DAC bits, $nBits, tot$.

$$p = \frac{2^{nBits,wo} - 1}{2^{nBits,tot} - 1} \tag{A.3}$$

The bottom-plate capacitance of these ungated capacitor segments may also be added to p.

As will be shown shortly, the capacitor driver and isolation switch sizing for delay can be used to express $C_{d,sw}$ as a function of C_{coef}:

$$C_{d,sw} = \beta \cdot C_{coef} \tag{A.4}$$
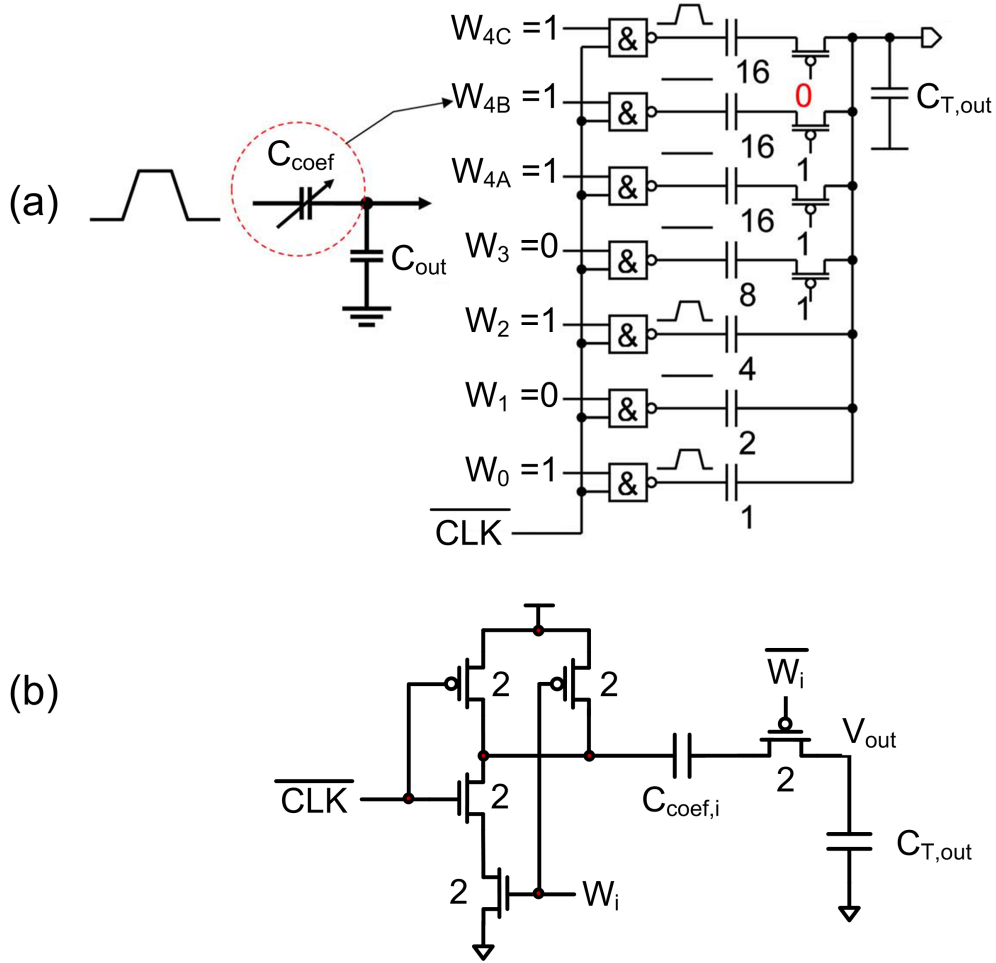
Figure A.2: Switched capacitor DAC: (a) Schematics [52], (b) Sizing for optimal delay

Therefore,

$$C_{off} = (\beta + p) \cdot C_{coef} \tag{A.5}$$

Now, the required coefficient capacitance C_i can be determined as a function of the maximum ISI cancelation voltage $k_{max} \cdot v_{in}$, the coefficient driver voltage V_{reg} and the capacitor divider with total output capacitance $C_{T,out}$. This ISI cancelation voltage $V_{ISI,i}$ of coefficient-i is given by the capacitor divider ratio as:

$$V_{ISI,i} = V_{reg} \cdot \frac{C_i}{C_{T,out}} \tag{A.6}$$

For every DFE coefficient to be able to cancel up to a maximum ISI magnitude of k_{max} times

the cursor,

$$V_{ISI,i} = v_{in} \cdot k_{max} = V_{reg} \cdot \frac{max\,(C_i)}{max\,(C_{T,out})}$$

$$= V_{reg} \cdot \frac{C_{coef}}{C_L + Cd_{cursor} + max\left(\sum_{i=1}^{N_{coef}} C_i\right) + N_{coef} \cdot C_{off} + Cd_{reset}} \quad (A.7)$$

Just like for current-integration based summing, the reset transistor capacitance Cd_{reset} can be expressed in terms of the sum of the other capacitances at the summing node in terms of a reset factor k_{reset} as first introduced in (3.6).

$$Cd_{reset} = k_{reset} \cdot \left[C_L + Cd_{cursor} + max\left(\sum_{i=1}^{N_{coef}} C_i\right) + N_{coef} \cdot C_{off}\right] \quad (A.8)$$

Therefore, equation (A.7) above becomes:

$$k_{max} \cdot \frac{v_{in}}{V_{reg}} \cdot (1 + k_{reset}) = \frac{C_{coef}}{\frac{C_{dI,cursor}}{2} \cdot I_{cursor} + C_L + \left\{\frac{k_{ISI,max}}{k_{max}} + N_{coef} \cdot (\beta + p)\right\} \cdot C_{coef}} \quad (A.9)$$

Similar to the expression for maximum ISI magnitude per coefficient in (A.7), for the sum of all DFE coefficients to cancel up to a certain total maximum ISI magnitude, $k_{ISI,max}$ times the cursor,

$$v_{in} \cdot k_{ISI,max} = V_{reg} \cdot \frac{max\left(\sum_{i=1}^{N_{coef}} C_i\right)}{max\,(C_{T,out})}$$

$$= V_{reg} \cdot \frac{max\left(\sum_{i=1}^{N_{coef}} C_i\right)}{C_L + Cd_{cursor} + max\left(\sum_{i=1}^{N_{coef}} C_i\right) + N_{coef} \cdot C_{off} + Cd_{reset}} \quad (A.10)$$

By taking a ratio of equations (A.10) and (A.7), it is easy to see that

$$max\left(\sum_{i=1}^{N_{coef}} C_i\right) = C_{coef} \cdot \frac{k_{ISI,max}}{k_{max}} \quad (A.11)$$

The only other capacitance contributor to $C_{T,out}$ left is from the per-coefficient off capacitance, which is in turn a function of the isolation switch size. This switch size (and therefore the switch capacitance) is computed in terms of C_{coef} by sizing the capacitor driver circuitry for delay. This dependence, which is described in the following part of the analysis, helps to simplify $C_{T,out}$ and express it as a function of C_{coef}.

Similar to the resistively-loaded current summer, after subtracting the shift-register flip-flop delay, the driver (with embedded sign selection NAND logic) has αUI portion of the clock cycle to settle. Therefore,

$$t_{delay} = \frac{\alpha}{f_s} \tag{A.12}$$

Fig. A.2(b) shows the coefficient NAND driver to switch each leg of the DFE coefficient DAC capacitance and the isolation PMOS switch. If the total input capacitance at the NAND driver is C_{drive}, the time delay (t_{delay}) for switching the DFE coefficient capacitance, according to the sizing shown in Fig. A.2(b), can be computed by using the Elmore delay model as:

$$t_{delay} = t_{inv} \cdot \left\{ \frac{4}{3} \cdot \left(2\gamma + \frac{1}{4} \right) + \frac{2}{3} \cdot (2\gamma + 1) + \frac{8}{3} \cdot \frac{1}{C_{drive}} \cdot \frac{(1-p) \cdot C_{coef} \cdot max\,(C_{T,out})}{(1-p) \cdot C_{coef} + max\,(C_{T,out})} \right\} \tag{A.13}$$

where $t_{inv} = \dfrac{T_{FO4}}{\gamma + 4}$, and $(1-p) \cdot C_{coef}$ is the part of the coefficient capacitance that is gated. In the above equation (A.13), the first term is from the self-loading of the NAND gate, the second term is from driving the PMOS isolation/pass-gate while the third term is from driving the series combination of $(1-p) \cdot C_{coef}$ and $max\,(C_{T,out})$. From (A.7), we already know C_{coef} as a function of the total output capacitance, $C_{T,out}$:

$$C_{coef} = \left(k_{max} \cdot \frac{v_{in}}{V_{reg}} \right) \cdot max\,(C_{T,out}) \tag{A.14}$$

The delay equation (A.13) above may be further simplified by exploiting the magnitude of C_{coef} as compared to $C_{T,out}$. In equation (A.14), the ISI magnitude k_{max} is less than 1. The input voltage v_{in} is typically a small-signal at best of the order of 50-100mV, while the capacitor driver voltage V_{reg} is close to the supply ($\sim$1V). Therefore, C_{coef} will be much smaller than $C_{T,out}$. It may then be approximated that the series combination of capacitors $(1-p) \cdot C_{coef}$ and $max\,(C_{T,out})$ is equivalent to the former (smaller) capacitor. In other words,

$$\frac{(1-p) \cdot C_{coef} \cdot max\,(C_{T,out})}{(1-p) \cdot C_{coef} + max\,(C_{T,out})} \approx (1-p) \cdot C_{coef} \tag{A.15}$$

Using this approximation, (A.13) can then be simplified to

$$t_{delay} = \frac{t_{inv}}{3} \cdot \left\{ 12\gamma + 3 + 8\,(1-p) \cdot \frac{C_{coef}}{C_{drive}} \right\} \tag{A.16}$$

Combining the delay equations in (A.16) and (A.12) and expressing t_{inv} in terms of T_{FO4}, C_{drive} can be obtained in terms of C_{coef}:

$$C_{drive} = \left\{ \frac{\left(\frac{8}{3}\right) \cdot (1-p)}{\frac{\gamma+4}{T_{FO4}} \cdot \frac{\alpha}{f_s} - (4\gamma+1)} \right\} \cdot C_{coef} \tag{A.17}$$

From the sizing shown in Fig. A.2(b), the the isolation switch capacitance at the summing node $C_{d,sw}$ can be directly expressed in terms of NAND input drive capacitance C_{drive} as:

$$C_{d,sw} = \frac{\gamma}{2} \cdot C_{drive} \tag{A.18}$$

$C_{d,sw}$ can therefore finally be expressed in terms of C_{coef} (as expected by (A.4)) as:

$$C_{d,sw} = \left\{ \frac{\left(\frac{4\gamma}{3}\right) \cdot (1-p)}{\frac{\gamma+4}{T_{FO4}} \cdot \frac{\alpha}{f_s} - (4\gamma+1)} \right\} \cdot C_{coef} \tag{A.19}$$

$$= \beta \cdot C_{coef}$$

such that

$$\beta = \frac{\left(\frac{4\gamma}{3}\right) \cdot (1-p)}{\frac{\gamma+4}{T_{FO4}} \cdot \frac{\alpha}{f_s} - (4\gamma+1)} \tag{A.20}$$

Finally, the cursor current can be obtained from the current integrator gain expression from (3.13). The expression for this gain is:

$$G = \frac{I_{cursor}}{2 \cdot f_s \cdot V^*_{cursor} \cdot max\left(C_{T,out}\right)} \tag{A.21}$$

Rewriting for $C_{T,out}$,

$$max\left(C_{T,out}\right) = \frac{I_{cursor}}{2 \cdot G \cdot f_s \cdot V^*_{cursor}} \tag{A.22}$$

Substituting for $max\left(C_{T,out}\right)$ from (A.7) leads to:

$$C_{coef} = k_{max} \cdot \frac{v_{in}}{V_{reg}} \cdot \frac{I_{cursor}}{2 \cdot G \cdot f_s \cdot V^*_{cursor}} \tag{A.23}$$

Substituting for C_{coef} from (A.23) in (A.9), after simplification I_{cursor} is:

$$I_{cursor} = \frac{2 \cdot (1 + k_{reset}) \cdot I_{nom}}{1 - 2 (1 + k_{reset}) \left(\gamma \cdot \dfrac{GDR}{\omega_{T,cursor}} \right) \left[1 + \dfrac{k_{max} \, v_{in} \, \omega_{T,cursor}}{\gamma \, V_{reg} \, G \, f_s} \left\{ \dfrac{k_{ISI,max}}{k_{max}} + (\beta + p) \, N_{coef} \right\} \right]}$$

$$\beta = \frac{\left(\dfrac{4\gamma}{3} \right) \cdot (1 - p)}{\left(\dfrac{\gamma + 4}{T_{FO4}} \right) \cdot \left(\dfrac{\alpha}{f_s} \right) - (4\gamma + 1)}$$

$$(A.24)$$

Since β is a function of p, the proportion of coefficient capacitance that should be gated can be optimized for a given technology and data-rate. For 5GS/s operation in a 65nm CMOS technology with $T_{FO4} = 30ps$, the optimal DAC gating is when the least significant 2 bits are left ungated.

Appendix B

Analysis of FFE Switching Matrix

This appendix analyzes the power required by the switching matrix (described in Chapter 4) to support a certain number of FFE delay elements, N_{seg}. Since the switching matrix uses current integration, its analysis is similar to the analysis of the current integrating summer earlier in Chapter 3 (sub-section 3.2.1). The difference between the two analyses, however, is the presence of a cascode device between the g_m stage and the integrating load.

The power consumption of the switching matrix is computed by evaluating the per-segment bias current I_{seg} required by each g_m-stage to support a certain gain G, data-rate f_s and external capacitive load C_L. Fig. B.1 shows the equivalent circuit and small-signal model for each leg of the matrix, comprising of the input gm-stage, the matrix cascode switches, and the PMOS current load with CMFB (the details of which have been excluded for simplicity). Since the sizing of R_{cmfb} is largely independent of the other transistors,[1] its capacitance can be absorbed into C_L. To simplify the model, The offset cancelation circuitry is not explicitly included in the model and may be absorbed into the g_m cell. The total capacitance at the cascode source and drain is initially condensed into $C_{T,casc}$ and $C_{T,out}$ respectively as shown.

Firstly, the gain may be computed by finding the equivalent integrating capacitance C_{eq} seen by input current $i_{in} = g_{mi} \cdot v_{in}$. The integrated output voltage v_{out} will be

$$v_{out} = \frac{i_{in}}{s \cdot C_{eq}} \tag{B.1}$$

By appplying KCL at the cascode source and drain nodes, we get:

$$i_{in} = V_s \cdot s \cdot C_{T,casc} + g_{mc} \cdot V_s + \frac{V_s - V_o}{r_{oc}} \tag{B.2}$$

$$g_{mc} \cdot V_s = \frac{v_o - v_s}{r_{oc}} + \frac{v_o}{r_L} + v_o \cdot s \cdot C_{T,out} \tag{B.3}$$

---

[1]The CMFB resistor is designed to be significantly larger than the PMOS load output resistance. For the simplicity of the model, therefore, it is convenient to assume the two resistances to be independent.
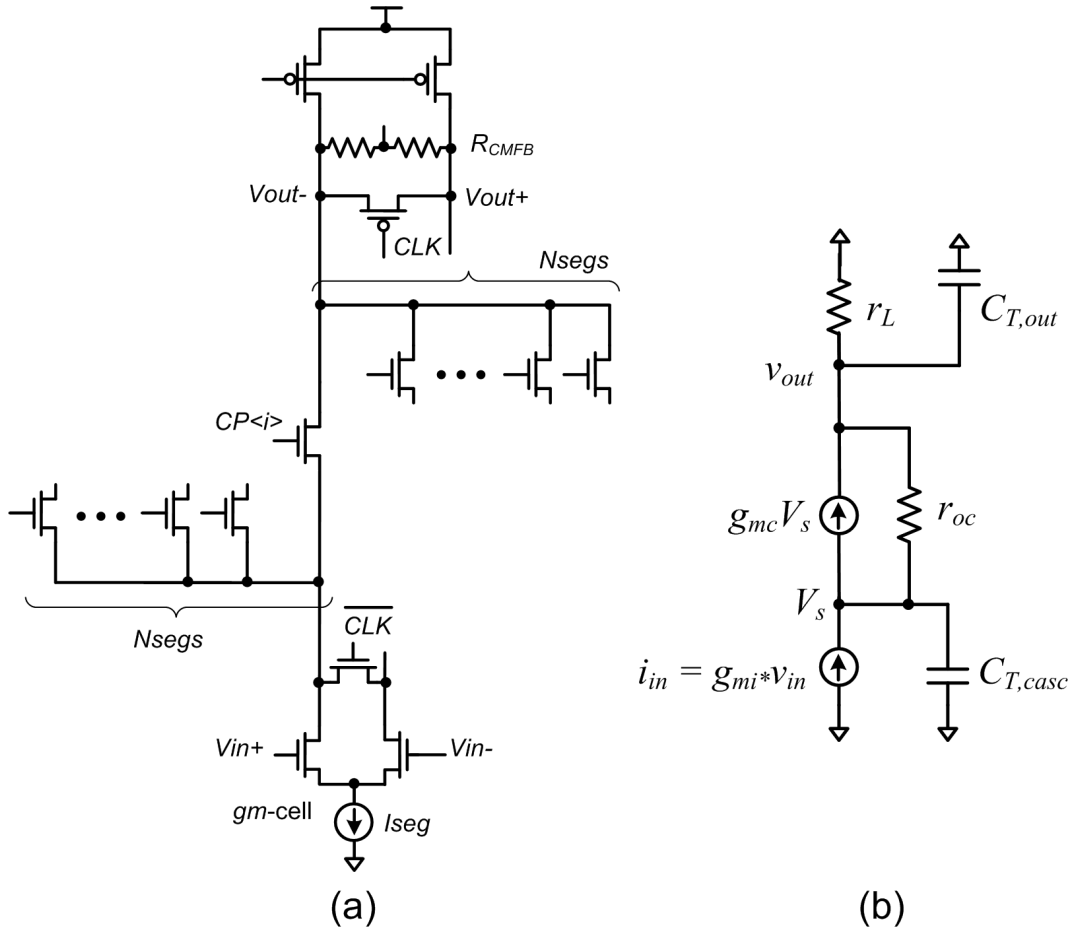
Figure B.1: (a) Equivalent circuit of each switching-matrix segment, and (b) its small signal model. For simplicity, the cascode switches are drawn as single-ended, and CMFB details have been excluded.

Solving (B.1) through (B.3) for v_{out} gives:

$$v_{out} = \frac{i_{in} \cdot r_L}{s^2 \left\{ \dfrac{r_{oc} \cdot r_L \cdot C_{T,out} \cdot C_{T,casc}}{(1 + g_{mc}r_{oc})} \right\} + s \left\{ \dfrac{(r_L + r_{oc}) \cdot C_{T,casc} + (1 + g_{mc}r_{oc}) \cdot r_L \cdot C_{T,out}}{(1 + g_{mc}r_{oc})} \right\} + 1}$$

(B.4)

To simplify the equation, we can assume[2] that $(1 + g_{mc}r_{oc}) \approx g_{mc}r_{oc}$. Since the integrator is low-bandwidth relative to the operating data-rate, the magnitude of the first two denomi-

---

[2]Since the prototype uses the LP flavor of a 65nm CMOS process, the intrinsic transistor gain is reasonably high to make the assumption valid.

nator terms $\gg 1$. Therefore, the equation is now

$$v_{out} = \frac{i_{in}}{s^2 \left(\dfrac{C_{T,out} \cdot C_{T,casc}}{g_{mc}} \right) + s \left\{ \left(\dfrac{1}{g_{mc}r_{ocL}} + \dfrac{1}{g_{mc}r_L} \right) \cdot C_{T,casc} + C_{T,out} \right\}} \tag{B.5}$$

r_{ocL} and r_L may be combined as

$$\frac{1}{r_{ocL}} = \frac{1}{r_{ocL}} + \frac{1}{r_L} \tag{B.6}$$

which simplies v_{out} to

$$v_{out} = \frac{i_{in}}{s \cdot C_{T,out} \left\{ s \cdot \dfrac{C_{T,casc}}{g_{mc}} + \left(1 + \dfrac{C_{T,casc}}{C_{T,out} \cdot g_{mc}r_{ocL}} \right) \right\}} \tag{B.7}$$

Since both cascode and PMOS current load transistors carry equal bias currents, the quantity $g_{mc}r_{ocL}$ is a technology-based constant. Also, since the PMOS load is a long-channel device while the cascode is short-channel, r_{ocL} will be only marginally lower than r_{oc}. Comparing (B.1) and (B.7) gives an equivalent integrating capacitance

$$C_{eq} = C_{T,out} \left(1 + \frac{C_{T,casc}}{C_{T,out} \cdot g_{mc}r_{ocL}} \right) \cdot \left\{ 1 + \frac{s}{\left(\dfrac{g_{mc}}{C_{T,casc}} \right) \left(1 + \dfrac{C_{T,casc}}{C_{T,out} \cdot g_{mc}r_{ocL}} \right)} \right\} \tag{B.8}$$

This form of C_{eq} makes cascode-based current integration look like conventional current integration with an additional pole $\omega_{p,Ceq.}$ at

$$\omega_{p,Ceq.} = \left(\frac{g_{mc}}{C_{T,casc}} \right) \left(1 + \frac{C_{T,casc}}{C_{T,out} \cdot g_{mc}r_{ocL}} \right) \tag{B.9}$$

Now that C_{eq} for cascode-based integration is well understood as a function of cascode transistor parameters, $C_{T,casc}$ and $C_{T,out}$ may be expanded into their respective individual transistor capacitance contributions:

$$C_{T,out} = C_L + N_{seg} \cdot Cd_{casc} + Cd_{load} + C_{reset,out} \tag{B.10}$$

$$C_{T,casc} = N_{seg} \cdot Cg_{casc} + Cd_{in} + C_{reset,casc} \tag{B.11}$$

In the above equations, Cd_{casc} and Cs_{casc} are respectively the per-switch drain and source capacitance for each cascode transistor, Cd_{load} is the PMOS current source drain capacitance, Cd_{in} is the g_m-cell drain capacitance, and $C_{reset,out}$ and $C_{reset,casc}$ are load capacitances due to the reset/bridge switches at the output and cascode nodes respectively. Since the reset

switches will be sized to resettle their respective differential nodes within a given period (within $UI/2$), the reset switch capacitance at each node will be equal to the rest of the capacitance at the node scaled by a certain technology-dependent factor k_{reset} (similar to the current integrator in Chapter 3). In other words,

$$C_{reset,out} = k_{reset} \cdot (C_L + N_{seg} \cdot Cd_{casc} + Cd_{load}) \tag{B.12}$$

$$C_{reset,casc} = k_{reset} \cdot (N_{seg} \cdot Cs_{casc} + Cd_{in}) \tag{B.13}$$

Now, for equal bias current I_{seg}, the long-channel PMOS load has larger drain capacitance than the g_m-stage drain capacitance. Also, the cascode gate and drain caps are approximately equal. Both conditions may be summarized as:

$$\begin{aligned} Cd_{out} &> Cd_{in} \\ Cd_{casc} &\approx Cg_{casc} \\ N_{seg} \cdot Cd_{casc} &\approx N_{seg} \cdot Cg_{casc} \end{aligned} \tag{B.14}$$

Equations (B.10)-(B.14) therefore suggest that

$$C_{T,out} > C_{T,casc} \tag{B.15}$$

For reasonably large intrinsic gain, $g_{mc}r_{ocL} \gg 1$, therefore,

$$\frac{C_{T,casc}}{C_{T,out} \cdot g_{mc}r_{ocL}} \ll 1 \tag{B.16}$$

For the switching matrix, therefore, the value of C_{eq} from (B.8) may then be simplified to

$$C_{eq} \approx C_{T,out} \cdot \left\{ 1 + \left(\frac{C_{T,casc}}{g_{mc}} \right) \cdot s \right\} \tag{B.17}$$

The form of $C_{T,casc}$ in (B.11) suggests that

$$\frac{g_{mc}}{C_{T,casc}} = \frac{\omega_{T,casc}}{(1 + k_{reset}) \left(\gamma \cdot \frac{Cg_{casc}}{Cg_{in}} + N_{seg} \right)} \tag{B.18}$$

where $w_{T,casc}$ is the cascode unity current-gain frequency, γ is the drain to gate capacitance ratio, and Cg_{in} is the gate capacitance of the g_m-cell. At data-rates of 5GS/s, using a 65nm CMOS technology and for a reasonable number of delay-line segments (N_{seg}), it can be computed that the circuit operates beyond the cascode 3dB cut-off frequency. In this region of operation, C_{eq} from (B.17) can be further simplified as

$$\begin{aligned} C_{eq} &\approx C_{T,out} \cdot \left(\frac{C_{T,casc}}{g_{mc}} \cdot s \right) \\ |C_{eq}| &= C_{T,out} \cdot \left(\frac{C_{T,casc}}{g_{mc}} \right) \cdot (2\pi f_s) \end{aligned} \tag{B.19}$$

The transistor capacitances may themselves be expressed in terms of technology-based parameters, $C_{dI,casc}$, $C_{sI,casc}$, $C_{dI,load}$, $C_{dI,in}$, where C_{dI} and C_{sI} respectively denote drain and source capacitance per unit current, and the subscripts *casc*, *load*, and *in* refer to cascode, PMOS-load and input g_m-stage transistors respectively.

$$
\begin{aligned}
Cd_{casc} &= C_{dI,casc} \cdot \frac{I_{seg}}{2} \\
Cs_{casc} &= C_{sI,casc} \cdot \frac{I_{seg}}{2} \\
Cd_{load} &= C_{dI,load} \cdot \frac{I_{seg}}{2} \\
Cd_{in} &= C_{dI,in} \cdot \frac{I_{seg}}{2}
\end{aligned}
\tag{B.20}
$$

Combining (B.10) – (B.13) with (B.20) gives

$$
\begin{aligned}
C_{T,out} &= (1 + k_{reset}) \cdot (C_L + N_{seg} \cdot Cd_{casc} + Cd_{load}) \\
&= (1 + k_{reset}) \cdot \left\{ C_L + \frac{I_{seg}}{2} \cdot (N_{seg} \cdot C_{dI,casc} + C_{dI,load}) \right\}
\end{aligned}
\tag{B.21}
$$

$$
\begin{aligned}
C_{T,casc} &= (1 + k_{reset}) \cdot (N_{seg} \cdot Cs_{casc} + Cd_{in}) \\
&= (1 + k_{reset}) \cdot \left\{ \frac{I_{seg}}{2} \cdot (N_{seg} \cdot C_{sI,casc} + C_{dI,in}) \right\}
\end{aligned}
\tag{B.22}
$$

For an integration time T_{int} of a half-clock period, i.e. $1/(2f_s)$, equation (B.1) can be re-written as

$$
v_{out} = \frac{i_{in} \cdot T_{int}}{C_{eq}} = \frac{1}{C_{eq}} \cdot \left(\frac{g_{mi} v_{in}}{2 f_s} \right)
\tag{B.23}
$$

Replacing $g_{mi} = I_{seg}/V_{in}^*$, $g_{mc} = I_{seg}/V_{casc}^*$ and expressing C_{eq} from (B.19) in terms of $C_{T,out}$ and $C_{T,casc}$ from (B.21) and (B.22) respectively leads to:

$$
v_{out} = v_{in} \cdot \frac{\left(\dfrac{I_{seg}}{V_{in}^*} \right)}{(1 + k_{reset})^2 \cdot \left(2\pi V_{casc}^* f_s^2 \right) \cdot \left\{ C_L + \dfrac{I_{seg}}{2} \cdot (N_{seg} \cdot C_{dI,casc} + C_{dI,load}) \right\} \cdot (N_{seg} \cdot C_{sI,casc} + C_{dI,in})}
\tag{B.24}
$$

Substituting $G = v_{out}/v_{in}$ and simplifying for I_{seg} gives:

$$
I_{seg} = \frac{G \cdot V_{in}^* \cdot f_s \cdot C_L}{\dfrac{1}{(1 + k_{reset})^2 (2\pi \cdot f_s \cdot V_{casc}^*)(N_{seg} \cdot C_{sI,casc} + C_{dI,in})} - \dfrac{G \cdot V_{in}^* \cdot f_s}{2}(N_{seg} \cdot C_{dI,casc} + C_{dI,load})}
\tag{B.25}
$$

This expression may be written in terms of the w_T of each transistor as

$$I_{seg} = \frac{G \cdot V_{in}^* \cdot f_s \cdot C_L}{(1 + k_{reset})^2 (4\pi \cdot f_s) \left(N_{seg} + \gamma \cdot \dfrac{\omega_{T,casc}}{\omega_{T,in}} \cdot \dfrac{V_{casc}^*}{V_{in}^*} \right)} - \frac{\gamma \cdot G \cdot f_s}{\omega_{T,casc}} \left(N_{seg} \cdot \frac{V_{in}^*}{V_{casc}^*} + \frac{\omega_{T,casc}}{\omega_{T,load}} \cdot \frac{V_{in}^*}{V_{load}^*} \right)$$

$$\text{(B.26)}$$

where the subscripts *in, casc* and *load* correspond to input g_m-stage, cascode and PMOS load transistors respectively.

Appendix C

Analysis of Combined FFE-DFE Summer

This appendix analyzes the cascode-based FFE-DFE summer-integrator to compute its total current consumption as a function of the total number of FFE and DFE coefficients. The analysis builds upon the framework presented for the cascode-based integrator for the switching matrix in Appendix B.

Fig. C.1 shows the schematics of the summer-integrator and its simplified small-signal model to compute the gain of the FFE cursor coefficient (FFE,0). The DC bias current of the cursor $I_{FFE,0}$ is determined as a function of the number of FFE coefficients (N_{FFE}) and DFE coefficients (N_{DFE}). Throughout the analysis, I and i denote DC and small signal currents respectively.

The two primary constraints in this analysis are (a) maintaining a certain integrator bandwidth proportional to data-rate f_s, and (b) achieving a certain integrator gain, G. The first constraint helps compute the cascode bandwidth-boosting common-mode current I_{boost} in terms of $I_{FFE,0}$. Using this relation, the second constraint then helps compute $I_{FFE,0}$ itself in terms of N_{FFE} and N_{DFE}. The total power consumption is proportional to the sum of $I_{FFE,0}$ and $2 \cdot I_{boost}$.

The total capacitance at the cascode node ($C_{T,casc}$) can be expressed in terms of individual transistor capacitances as:

$$C_{T,casc} = Cd_{cursor} + N_{FFE} \cdot Cd_{FFE} + N_{DFE} \cdot Cd_{DFE} + Cd_{boost} + Cs_{casc} \qquad (C.1)$$

In this equation, Cd_{cursor}, Cd_{FFE}, Cd_{DFE}, and Cd_{boost} are respectively the FFE cursor, FFE non-cursor per-coefficient, DFE per-coefficient and booster drain capacitances respectively, Cs_{casc} is the cascode source capacitance, and N_{FFE} and N_{DFE} are respectively the number of FFE and DFE coefficients. Similar to the switching matrix analysis, the capacitors can
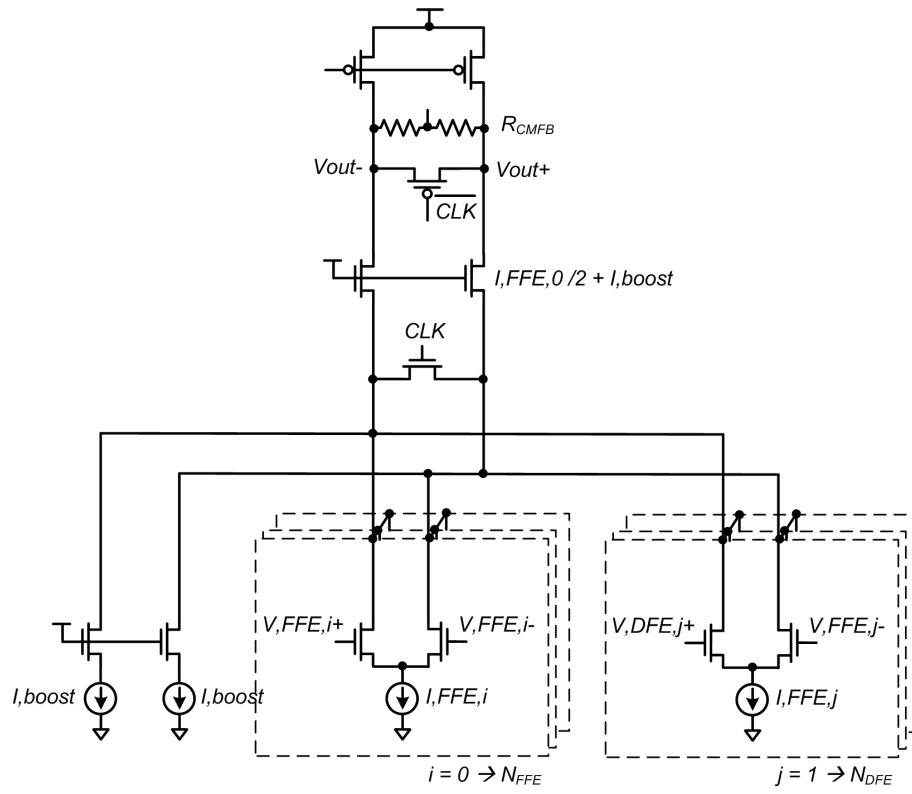
(a)
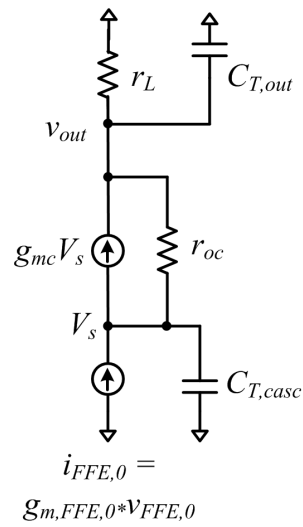
(b)

Figure C.1: (a) Equivalent circuit of the FFE-DFE summer-integrator (CMFB details excluded), and (b) its simplified small signal model.

.

be expressed in terms of their respective currents and capacitance per unit current as

$$
\begin{aligned}
Cd_{cursor} &= C_{dI,FFE} \cdot \frac{I_{FFE,0}}{2} \\
Cd_{FFE} &= C_{dI,FFE} \cdot \frac{I_{FFE,\sim 0}}{2} + C_{f,FFE} \\
Cd_{DFE} &= C_{dI,DFE} \cdot I_{DFE} + C_{f,DFE} \\
Cd_{boost} &= C_{dI,boost} \cdot I_{boost} \\
Cs_{casc} &= C_{sI,casc} \cdot \left(\frac{I_{cursor}}{2} + I_{boost} \right)
\end{aligned}
\tag{C.2}
$$

In these equations, C_{dI} and C_{sI} respectively denote drain and source capacitance per unit current, while the subscripts FFE, DFE, $boost$ and $casc$ denote FFE coefficient, DFE coefficient, booster and cascode transistors respectively. $C_{f,DFE}$ and $C_{f,FFE}$ are the fixed FFE and DFE capacitance per coefficient, mainly from the wiring at the cascode summing node which tends to be physically long. For the FFE coefficients, since the 0th coefficient tends to have higher current than the rest, the subscripts further denote 0 as the cursor and ~ 0 for the other coefficients. The cascode transistor is sized to carry a bias current of only $\left(\frac{I_{cursor}}{2} + I_{boost} \right)$. When the FFE and/or DFE coefficients are active, the total cascode bias current is kept constant by adaptively reducing I_{boost} by the sum of average FFE and DFE currents of all coefficients. In other words,

$$
I_{boost}(i,j) = max(I_{boost}) - \sum_{i}^{N_{FFE}} I_{FFE}(i) - \sum_{j}^{N_{DFE}} \frac{I_{DFE}(j)}{2}
\tag{C.3}
$$

where indices i and j indicate the i-th FFE and j-th DFE coefficient[1] respectively. Now, if the non-cursor FFE and DFE coefficients are at most k_{FFE} and k_{DFE} times larger in magnitude than the FFE cursor coefficient, they may be simplified as:

$$
\begin{aligned}
I_{FFE,\sim 0} &= k_{FFE} \cdot I_{FFE,0} \\
I_{DFE} &= k_{DFE} \cdot i_{FFE,0} \\
&= k_{DFE} \cdot (g_{m,FFE,0} \cdot V_{in}) \\
&= \left(k_{DFE} \cdot \frac{V_{in}}{V_{FFE}^*} \right) \cdot I_{FFE,0}
\end{aligned}
\tag{C.4}
$$

Substituting (C.2) and (C.4) into (C.1), $C_{T,casc}$ can be expressed in terms of $I_{FFE,0}$ and I_{boost} as:

$$
\begin{aligned}
C_{T,casc} = {}& (N_{FFE} \cdot C_{f,FFE} + N_{DFE} \cdot C_{f,DFE}) \cdot (1 + k_{reset}) + I_{boost} \cdot (1 + k_{reset}) \cdot (C_{dI,bleed} + C_{sI,casc}) \\
& + I_{FFE,0} \cdot (1 + k_{reset}) \cdot \left\{ \left(\frac{1 + N_{FFE} \cdot k_{FFE}}{2} \right) \cdot C_{dI,FFE} + N_{DFE} \cdot k_{DFE} \cdot \frac{V_{in}}{V_{FFE}^*} \cdot C_{dI,DFE} + \frac{C_{sI,casc}}{2} \right\}
\end{aligned}
\tag{C.5}
$$

---

[1]Since the DFE coefficient current is hard steered, its average will equal half the DC value.

Similarly, $C_{T,out}$ can be expressed in terms of $I_{FFE,0}$ and I_{boost} as:

$$
\begin{aligned}
C_{T,out} = \; & C_L \\
& + I_{boost} \cdot (1 + k_{reset}) \left(C_{d,casc} + C_{d,load} \right) \\
& + I_{FFE,0} \cdot (1 + k_{reset}) \left(\frac{C_{d,casc} + C_{d,load}}{2} \right)
\end{aligned}
\tag{C.6}
$$

The effective cascode bandwidth is the pole frequency ω_c from equation (B.8):

$$
\omega_c = \left(\frac{g_{mc}}{C_{T,casc}} \right) \left(1 + \frac{C_{T,casc}}{C_{T,out} \cdot g_{mc} r_{ocL}} \right)
\tag{C.7}
$$

In order to obtain I_{boost} as a function of $I_{FFE,0}$, (C.7) can easily be expressed in terms of I_{boost} and $I_{FFE,0}$ by replacing $C_{T,casc}$ and $C_{T,out}$ respectively from (C.5) and (C.6). However, this makes the relationship between I_{bleed} and $I_{FFE,0}$ too complicated to retain an intuitive understanding of the analysis. To simplify the analysis, the factor of $\dfrac{C_{T,casc}}{C_{T,out} \cdot g_{mc} r_{ocL}}$ may be ignored. To compensate for this approximation, a lower value for ω_c is targeted. The factor by which this targeted ω_c is made lower can be computed from a first-cut design obtained by this approximate analysis. Continuing this analysis,

$$
\omega_c = \frac{g_{mc}}{C_{T,casc}} = \frac{I_{FFE,0} + 2I_{boost}}{V^*_{casc} \cdot C_{T,casc}}
\tag{C.8}
$$

Substituting for $C_{T,casc}$ from (C.5) after simplification gives I_{boost}:

$$
I_{boost} = I_0 + \beta \cdot I_{FFE,0}
\tag{C.9}
$$

in which

$$
I_0 = \frac{V^*_{casc} \cdot \omega_c \cdot (1 + k_{reset}) \cdot (C_{f,FFE} \cdot N_{FFE} + C_{f,DFE} \cdot N_{DFE})}{2 - V^*_{casc} \cdot \omega_c \cdot (1 + k_{reset}) \cdot (C_{dI,boost} + C_{sI,casc})}
$$

$$
\beta = \frac{\dfrac{V^*_{casc}\omega_c (1 + k_{reset})}{2} \left\{ (1 + N_{FFE} k_{FFE}) C_{dI,FFE} + 2 N_{DFE} k_{DFE} \dfrac{V_{in}}{V^*_{FFE}} C_{dI,DFE} + C_{gI,casc} \right\} - 1}{2 - V^*_{casc} \cdot \omega_c \cdot (1 + k_{reset}) \cdot (C_{dI,boost} + C_{sI,casc})}
\tag{C.10}
$$

Finally, the gain of a cascode current integrator from equation (B.23) is:

$$
G = \frac{g_{m,FFE0}}{2 \cdot f_s \cdot C_{T,out} \left(1 + \dfrac{C_{T,casc}}{g_{mc} r_{ocL} \cdot C_{T,out}} \right)}
\tag{C.11}
$$

In (C.11), substituting for $C_{T,out}$ and $C_{T,casc}$ from (C.6) and (C.5) respectively, and writing I_{boost} in terms of $I_{FFE,0}$ from (C.9) after simplification gives:

$$
I_{FFE,0} = \frac{(2GV^*_{FFE} f_s C_L) (\beta_{0,N} + \beta_{1,FFE} N_{FFE} + \beta_{1,DFE} N_{DFE})}{1 - (2G) (\beta_{0,D} + \beta_{2,FFE} N_{FFE} + \beta_{2,DFE} N_{DFE})}
\tag{C.12}
$$

where all βs are defined as follows:

$$\beta_{0,N} = 1 + \frac{I_0}{C_L} \left(1 + k_{reset}\right) \left(2\gamma\right) \left\{ \frac{1}{g_{mc}r_{ocL}} \left(\frac{1}{V^*_{boost}\omega_{T,boost}} + \frac{1}{\gamma V^*_{casc}\omega_{T,casc}} \right) + \frac{1}{V^*_{load}\omega_{T,load}} + \frac{1}{V^*_{casc}\omega_{T,casc}} \right\}$$

$$\beta_{1,FFE} = \frac{\left(1 + k_{reset}\right) \cdot C_{f,FFE}}{g_{mc}r_{ocL}C_L}$$

$$\beta_{1,DFE} = \frac{\left(1 + k_{reset}\right) \cdot C_{f,DFE}}{g_{mc}r_{ocL}C_L}$$

$$\beta_{0,D} = V^*_{FFE}f_s \left(1 + k_{reset}\right) \gamma \left[\frac{1}{V^*_{casc}\omega_{T,casc}} + \frac{1}{V^*_{load}\omega_{T,load}} + \frac{1}{2g_{mc}r_{ocL}} \left(\frac{1}{V^*_{FFE}\omega_{T,FFE}} + \frac{1}{\gamma V^*_{casc}\omega_{T,casc}} \right) \right] +$$

$$V^*_{FFE}f_s \left(1 + k_{reset}\right) \left(2\gamma\right) \beta \left[\frac{1}{V^*_{casc}\omega_{T,casc}} + \frac{1}{V^*_{load}\omega_{T,load}} + \frac{1}{g_{mc}r_{ocL}} \left(\frac{1}{V^*_{boost}\omega_{T,boost}} + \frac{1}{\gamma V^*_{casc}\omega_{T,casc}} \right) \right]$$

$$\beta_{2,FFE} = \left(1 + k_{reset}\right) \cdot \frac{\gamma \cdot f_s \cdot k_{FFE}}{\omega_{T,FFE} \cdot g_{mc} \cdot r_{ocL}}$$

$$\beta_{2,DFE} = \left(1 + k_{reset}\right) \cdot \frac{2 \cdot v_{in} \cdot f_s \cdot k_{DFE}}{V^*_{DFE} \cdot \omega_{T,DFE} \cdot g_{mc} \cdot r_{ocL}}$$

$$(C.13)$$