Signal Processing Techniques
for High-Speed Chip-to-Chip Links

by

Mike Bichan

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Electrical and Computer Engineering
University of Toronto

© Copyright by Mike Bichan 2012
Signal Processing Techniques for High-Speed Chip-to-Chip Links

Mike Bichan

Doctor of Philosophy, 2012
Graduate Department of Electrical and Computer Engineering
University of Toronto

Abstract

This thesis tackles the problem of high-speed data communication over wireline channels. Particular attention is paid to backplane channels which have impedance discontinuities and high-frequency loss. These channels require extra equalization effort in order to produce an open eye diagram at the receiver. Three signal processing techniques were investigated in the pursuit of higher data rates over backplane channels: transmit-side FIR filter equalization with variable tap spacing, bidirectional communication using frequency-division multiplexing, and an ADC-based receiver to provide a capability for non-linear equalization. The ADC presented here is a 5-bit flash ADC intended to be time-interleaved to attain a sufficient data rate. This ADC uses redundant comparators to obtain sufficient resolution without an explicit threshold tuning circuit. A resonant clocking line is used to reduce power and increase the maximum clock frequency.
Acknowledgments

Success is the ability to go from one failure to another with no loss of enthusiasm.

__________________________
Winston Churchill

I would first like to thank my supervisor Professor Tony Chan Carusone. His guidance and enthusiasm for electronics made this thesis possible.

I would also like to thank Professors Johns, Sheikholeslami, Anderson, and Ng of the ECE Department as well as Professor van Driel of the Physics Department and Professor Yuan of Ryerson University for serving on my thesis defense committee and providing valuable feedback that helped improve this thesis.

Many thanks go to CMC, CMP, and ST Microelectronics for allowing me access to the 90nm and 65nm CMOS design kits and especially CMC for allocating me die area.

Thank you to my fellow graduate students in BA5158, BA5000, and BA4182 who made my time in school pass more quickly. Especially Kentaro Yamamoto, Shahriar Shahramian, Ricardo Aroca, Ruslana Shulyzki, Farzaneh Shahrokhi, Tina Tahmoureszadeh, Dustin Dunwell, Shayan Shahramian, Hemesh Yasothenan, Alain Rousson, Bert Leesti, Trevor Caldwell, Hamed Jafari, and Karim Abdelhalim. I hope we continue to cross paths professionally and non-professionally for years to come.

My parents made this thesis possible with the foundation of love and support they have provided. This support often goes unrecognized, but I cannot imagine having made it this far without them.

Finally, I would like to thank my wife Danielle. Her consistent belief in my ability to finish what I started has helped me accomplish my goals in graduate school. Her love and support made the easy parts fun and the difficult parts surmountable.
Contents

List of Figures vii

List of Tables xiii

1 Introduction 1
  1.1 Motivation ................................................................. 1
  1.2 Channel Description ..................................................... 2
  1.3 State of the Art Wireline Transceivers ......................... 5
    1.3.1 Transmit Driver — voltage-mode vs. current-mode .... 5
    1.3.2 Equalization — FFE, CTLE, DFE ......................... 8
    1.3.3 Equivalence of DFE and ADC .................................. 11
    1.3.4 Summary — state of the art ................................... 12
  1.4 Thesis Organization ................................................... 13

2 6.5 Gb/s Fractionally-Spaced Feed-Forward Equalizer in 90nm CMOS 14
  2.1 Motivation ................................................................. 14
  2.2 Literature Survey ........................................................ 16
  2.3 Circuit Design ............................................................. 17
  2.4 Measurement Results .................................................. 19
  2.5 Conclusions ............................................................... 27

3 5 Gb/s Downstream, 500 Mb/s Upstream Bidirectional Data Communication 31
  3.1 Motivation ................................................................. 31
  3.2 Literature Survey ........................................................ 34
  3.3 Frequency-Division Bidirectional Communication ............ 36
    3.3.1 AC-Coupled Links ............................................... 36
    3.3.2 AC-DC Bidirectional Link .................................... 37
    3.3.3 Transmitter Configuration .................................... 39
    3.3.4 Receiver ............................................................ 39
    3.3.5 Bidirectional Simulations .................................... 41
  3.4 Measurement Results .................................................. 43
  3.5 Future Work ............................................................. 46
## Contents

3.6 Conclusions .................................................. 46

4  3 GS/s, 17.6 mW Flash ADC in 65nm CMOS 50
   4.1 Motivation .................................................. 50
       4.1.1 Comparison of ADC Topologies ....................... 51
       4.1.2 Proposed ADC ....................................... 54
   4.2 Flash ADC Non-Idealities .................................. 55
       4.2.1 Comparator Offset .................................... 55
       4.2.2 Clock Skew ........................................... 59
   4.3 Redundancy-Based Flash ADC ................................ 62
       4.3.1 SNDR Degradation Due to Offset Variation ........... 64
       4.3.2 Calibration ............................................ 65
   4.4 Resonant Clock Path ......................................... 67
   4.5 Circuit Design ............................................. 68
       4.5.1 Flash ADC Architecture ................................ 68
       4.5.2 Dynamic Comparator .................................... 70
       4.5.3 RS Latch .............................................. 70
       4.5.4 Voltage Regulator ..................................... 73
       4.5.5 Inductor Modeling ..................................... 74
       4.5.6 Clock Path ............................................ 76
       4.5.7 Layout Considerations .................................. 77
       4.5.8 Calibration DAC ....................................... 79
       4.5.9 Biasing Circuits ....................................... 80
   4.6 Selected Simulation Results ................................. 80
       4.6.1 Dynamic Comparator Simulations ....................... 80
       4.6.2 Resonant Clocking Simulations ......................... 87
   4.7 Measurement Results ....................................... 89
       4.7.1 Test Setup ............................................ 89
       4.7.2 ADC Measurements ..................................... 96
       4.7.3 Voltage Regulator Measurements ....................... 104
       4.7.4 Calibration DAC Measurements ......................... 104
   4.8 Conclusions ................................................ 108

5  Conclusion ..................................................... 109

Appendix A  Summary of Contributions .......................... 111

Appendix B  Remedial Simulations ................................ 112
Contents

References  116
List of Figures

1.1 Cisco blade switch ............................................. 1
1.2 HP blade server ................................................ 2
1.3 (a) Backplane channel, (b) frequency response .............. 3
1.4 24" channel: frequency response and impulse response sampled at 5 Gb/s and 8 Gb/s ........................................... 4
1.5 Eye diagrams ..................................................... 4
1.6 Power efficiency vs. data rate for state-of-the-art wireline transceivers. In the bottom figure, marker size is proportional to channel loss equalized ........................................ 7
1.7 Differential signaling: (a) current-mode driver, (b) voltage-mode driver .............................. 8
1.8 Typical implementation of: (a) current-mode driver, (b) voltage-mode driver ........................................ 8
1.9 Linear transceiver with eye diagrams ........................................ 9
1.10 24" backplane channel 5 Gb/s eye. 1-tap decision feedback equalization splits the eye into two based on the previous bit, providing a larger eye to sample ........................................ 10
1.11 24" backplane channel 8 Gb/s eye. A 1-tap DFE is no longer effective. However, splitting the eye into four (2-tap DFE) based on the previous two bits opens the eye ........................................ 10
1.12 (a) Channel frequency responses, (b) maximum data rate possible over a "dirty backplane" channel with different kinds of equalization ........................................ 11
1.13 The flash ADC and lookahead DFE architectures both operate by sampling with several comparators at different thresholds ........................................ 12

2.1 Impulse response with samples taken at three different intervals for a nominal data rate of 5 Gb/s. As the delay between filter taps shrinks, more ISI becomes targetable by the FIR equalizer ........................................ 16
2.2 Transmitter block diagram ........................................ 18
2.3 Schematic diagram of one output driver slice ........................................ 18
2.4 (a) Delay stage. Each delay cell block in Fig. 2.2 contains four such stages, (b) measured tap delay and total power of the delay line ........................................ 19
2.5 Example of two-tap pre-emphasis with the proposed transmitter. Tap weights are set to [+2 -1] ........................................ 20
2.6 Transmitter test board ............................................ 21
2.7 Measurement: 3-bit transmitter amplitude control at 6 Gb/s. Differential eye amplitude as shown. .............................................. 22
2.8 Measurement: Transmitter tap spacing control. Tap spacing shown as control voltage is increased linearly from 100 mV to 800 mV. .............. 22
2.9 Transmit-side DDJ vs. delay per cell of the delay line. The signal goes through the entire delay line before going to the output. ......................... 23
2.10 Measurement: Transmit-side eye diagrams at 5 Gb/s for two-tap equalization as boost is increased. RMS jitter is shown for each eye. ............... 23
2.11 Measurement: Transmit-side eye diagrams at 5 Gb/s for nominally-constant DC eye amplitude and increasing boost. RMS jitter is shown for each eye. 25
2.12 Linearity of the output driver DAC. (a) Output swing for each digital code, (b) DNL and INL ............................................................... 25
2.13 Measured and simulated behavioral model 5 Gb/s eye diagrams for a PRBS7 pattern: (a) Output of unequalized 24" backplane channel, (b) transmitter output for half-baud-spaced jitter-minimizing pulse shape, (c) equalized 24" backplane channel output. ....................................................... 26
2.14 Received 6.5 Gb/s equalized eye diagram for the 24" backplane channel. Tap spacing is 0.53 UI and the optimal tap weights are [+2 -1]. Total transmitter power for this configuration is 42 mW. ......................... 27
2.15 Measured receive-side DDJ for the 24" backplane channel at (a) 6.5 Gb/s, (b) 5.5 Gb/s. Optimal tap weights are recalculated for each tap spacing individually. ................................................................. 28
2.16 The transmitter prototype was implemented in 90 nm CMOS. Die size is 1 mm × 1 mm. ........................................................................ 29
3.1 (a) Power spectra and (b) eye diagrams of the two signals used in the proposed frequency-division bidirectional communication link: 500 Mb/s DC signal and 8 Gb/s AC signal. ............................................ 32
3.2 Three bidirectional communication schemes. ........................................ 35
3.3 Side-view of an AC-coupled link. ...................................................... 37
3.4 (a) Proposed AC/DC bidirectional link, (b) frequency response of different paths in the AC/DC bidirectional link with L =80 nH and C =10 pF. ...... 38
3.5 Five different transmitter pulse-shaping configurations with the power spectra obtained when transmitting a PRBS $2^{11} − 1$ sequence at 5 Gb/s. These pulse shapes produce identical peak voltage swing. ....................... 40
3.6 Hysteresis-based dual-path receiver designed by Masum Hossain. ....... 41
3.7 Simulated eye diagrams showing bidirectional operation. Data rates are 500 Mb/s for the DC channel and 8 Gb/s for the AC channel. Signals (e) & (f) have been recovered by slicing with the hysteresis-based receiver. ...... 42
List of Figures

3.8 Bidirectional test board. .................................................. 44
3.9 Measured 5 Gb/s eye diagrams for the (a) input to the AC receiver and (b) the equalized output. Separate transmitter and receiver chips were flip-chip bonded to a printed circuit board. Hysteresis threshold must be set within the threshold adjustment area shown in (a). .......................... 45
3.10 (a) Received and (b) recovered 5 Gb/s AC-coupled pseudo-random bit sequence (PRBS) sequence of length $2^7 - 1$. The eye diagrams for these signals are shown in Fig. 3.9. ............................................. 45
3.11 Eye diagrams and power spectra of two different 500 Mb/s DC pulse shapes. High-frequency content is reduced for the slew-rate limited signal in (c). . 47
3.12 Power spectra of (a) a 500 Mb/s DC signal and (b) a 5 Gb/s AC signal. . 48
3.13 DC and AC spectra from Fig. 3.12, plotted on a logarithmic frequency scale. . 48
3.14 AC/DC bidirectional link using buried-bump technology. ..................... 49

4.1 Power efficiency vs. sampling rate for recent high-speed CMOS ADCs. In the bottom figure, time interleaved converters are shown in black. The star indicates the work discussed in this chapter. ......................... 52
4.2 (a) Transfer function of a 2-bit A/D converter, (b) quantization noise, (c) squared quantization noise, (d) probability distribution of a sinusoidal input, (e) weighted, squared quantization noise. .............................. 56
4.3 (a) Transfer function of a 2-bit A/D converter with one threshold shifted from its ideal location, (b) quantization noise, (c) squared quantization noise with shaded area showing additional squared noise due to threshold offset. 57
4.4 Effect on SNDR of a single threshold shift in a 5-bit quantizer. ................... 58
4.5 Calculated PDFs of (a) offset for a single comparator, (b) $P_{\text{noise,added}}$ for a 5-bit quantizer, (c) SNDR for a 5-bit quantizer when $\sigma = \frac{1}{4} \text{LSB}$. ............................ 59
4.6 Results of behavioural simulation for a 5-bit ADC (10,000 trials): (a) SNDR histograms with threshold standard deviations as marked, (b) average SNDR vs. $\sigma_{\text{offset}}$ for behavioural simulation and analysis, (c) standard deviation of SNDR. .................................................... 60
4.7 Effect of ADC clock skew on sampling threshold for a sinusoidal input. ... 61
4.8 Change in threshold voltage (a) vs. threshold voltage for a given clock skew, (b) vs. clock skew for a given threshold voltage. There are two solutions in each case. ................................................................. 62
4.9 $\Delta v$ vs. $T_{\text{skew}}$ for each of the 31 threshold levels in a 5-bit ADC. ......... 62
4.10 (a) traditional 2-bit flash ADC, (b) example of a redundancy-based flash ADC with redundancy of $k = 4$. ................................. 63
4.11 In a redundancy-based flash ADC, threshold variance decreases as redundancy increases. .......................................................... 64
4.12 SNDR of the best 95% of devices vs. threshold standard deviation in a traditional 5-bit flash ADC. SNDR improves with redundancy factor $k$ for a given unit-size comparator offset variance $\sigma_0$. ............................. 65
4.13 Result of simple calibration illustrated with randomly generated offset values. There are 31 blocks and eight comparators per block. The chosen comparators are circled. .................................................. 65
4.14 Result of calibration with $v_{\text{ladder,diff}} = 50 \text{ mV}$ illustrated with randomly generated offset values. .................................................. 66
4.15 Interconnect modeled as a transmission line periodically loaded by clock buffers. .......................................................... 67
4.16 Clock skew vs. position for (a) the interconnect shown in Fig. 4.15, (b) the same interconnect terminated with a 5.6 nH inductor. ............. 68
4.17 Maximum SNDR achievable for two different clocking schemes, considering only the SNDR penalty due to clock skew. ....................... 69
4.18 Simplified block diagram of the ADC test chip. .......................... 69
4.19 Detailed block diagram of the ADC test chip. ............................ 71
4.20 Dynamic comparator. Kickback cancellation MOS capacitors are shown on the left-hand side. ....................................................... 72
4.21 RS latch. ............................................................................. 72
4.22 Voltage regulator. ................................................................. 73
4.23 Voltage regulator load regulation circuit. .................................... 73
4.24 Separating the load regulation and line regulation functions leads to less output ripple. .............................................................. 74
4.25 Inductor modeling: (a) $2\pi$ inductor model, (b) comparison between parameters computed from ASITIC s-parameters and parameters computed from frequency response of the above $2\pi$ model. Lines with square markers are values derived from ASITIC, while lines with no markers are the result of a spice simulation of the optimized $2\pi$ model. ...................... 75
4.26 Diagram of the clock path. ..................................................... 76
4.27 Simulated clock path differential input impedance with distributed transmission line model and two different inductor models. .......... 77
4.28 Diagram of the clock gating circuit. ......................................... 78
4.29 (a) In-order comparators, single resistor string lead to routing congestion, (b) routing is simplified with out-of-order comparators, doubled-back resistor string. ..................................................... 79
4.30 (a) Example of a segmented 4-bit DAC. This is a hybrid 2-bit resistor string, 2-bit current-mode DAC. (b) DC nonlinearity of the DAC. .............. 81
4.31 (a) DAC modified to reduce nonlinearity, (b) DC nonlinearity. ........ 81
4.32 Schematic of the opamp bias circuit. ....................................... 82
4.33 Schematic of the opamp circuit. ................................................. 82
4.34 Testbench used for dynamic comparator simulations. ................. 83
4.35 Veriloga-aided simulation converges on the comparator offset. The ideal comparator (no parasitics) has no offset, whereas the RCc extracted comparator has an offset of -95 mV. The table at right shows the largest parasitic capacitances connected to the output nodes of the comparator. Comparator nodes names are shown in Fig. 4.20. ................................................. 83
4.36 (a) Simulated comparator offset dependence on output capacitance imbalance with no other parasitics. (b) Comparator offset dependence on input common-mode voltage. (c) monte carlo simulation of ideal comparator (30 trials). ................................................. 84
4.37 (a) Hysteresis of the RCc extracted comparator. (b) Monte carlo simulation of hysteresis with $f_{cl}=3$ GHz, no extracted parasitics (30 trials). ............. 85
4.38 (a) Average threshold for all ADC codes for different values of $V_{ref,diff} = V_{ref,p} - V_{ref,n}$. (b) Standard deviation of threshold for all ADC codes. ............. 86
4.39 Decreasing the total reference ladder impedance by ten times (from 18 kOhm to 1.8 kOhm) reduces the average nonlinearity of the ADC but increases threshold variation. (a) average threshold, (b) average DNL, (c) sigma (30 trials). ................................................. 86
4.40 (a) Comparator offset variation decreases with a larger ladder resistance due to kickback onto the reference input. (b) Directly decreasing the common mode of the reference voltage decreases offset variation. ......................... 87
4.41 (a) True and complementary transient waveforms on the resonant clock line with an additional 2 fF at each of the 256 comparator inputs. (b) Corresponding waveforms after buffering. ................................................. 88
4.42 Clock duty cycle with respect to frequency. Capacitor values represent additional loading at each node of the clock line. (a) true clock, (b) complementary clock. ................................................. 89
4.43 Altera Stratix IV Signal Integrity Evaluation Board. ..................... 91
4.44 ADC test board. ........................................................................... 92
4.45 Simplified diagram of the test setup. ............................................. 93
4.46 Expanded diagram of the test setup. ALTGX is Altera’s name for their high-speed transceiver IP block. NIOS II is a synthesizable CPU that can be programmed in C. ................................................. 94
4.47 ADC configuration to facilitate deskew. Only one comparator per output has been activated, the rest are powered down with output low. ................................. 96
4.48 Measurement: comparator offset voltage histogram. ................... 97
4.49 Measurement: offset of the 256 comparators in the ADC. The comparators chosen by a simple calibration scheme are highlighted. ................................. 98
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.50</td>
<td>Measurement: histograms of comparator offset before and after calibration. Test conditions are the same as those shown in Fig. 4.49.</td>
</tr>
<tr>
<td>4.51</td>
<td>Measurement: comparator hysteresis vs. $f_{\text{clk}}$ for $v_{\text{in,cm}}, v_{\text{ladder,p}}, v_{\text{ladder,n}}$ all equal to 0.7 V. This measurement is for the 128 comparators driven by the positive edge of the clock.</td>
</tr>
<tr>
<td>4.52</td>
<td>Measurement: 4096-point ADC output spectrum with simple calibration. Computed SNDR is 21.6 dB.</td>
</tr>
<tr>
<td>4.54</td>
<td>Measurement: offset of the 256 comparators in the ADC with reference voltage applied. (a) after simple calibration and application of $v_{\text{ladder,diff}} = 250 \text{ mV}$, and (b) after recalibration.</td>
</tr>
<tr>
<td>4.55</td>
<td>Measurement: 4096-point ADC output spectrum after recalibration. Computed SNDR is 26.0 dB.</td>
</tr>
<tr>
<td>4.56</td>
<td>Measurement: 4096-point ADC output spectrum with $f_{\text{clk}} = 3 \text{ GHz}$. Computed SNDR is 23.5 dB.</td>
</tr>
<tr>
<td>4.57</td>
<td>Measurement: (a) SNDR vs. input power, (b) SNDR vs. sampling frequency, (c) SNDR vs. input frequency. The measurements in (b) and (b) were all taken at a constant input power.</td>
</tr>
<tr>
<td>4.58</td>
<td>Measurement: ADC power (a) with $f_{\text{clk}} = 3 \text{ GHz}$ (0.38 mW per comparator), (b) with 32 comparators turned on.</td>
</tr>
<tr>
<td>4.59</td>
<td>Die photo of the test chip with ADC and inductor shown. Die area is 1 mm$^2$. ADC and inductor combined area is $0.07 \text{ mm}^2 + 0.01 \text{ mm}^2 = 0.08 \text{ mm}^2$.</td>
</tr>
<tr>
<td>4.60</td>
<td>Measurement: DC voltage regulator output.</td>
</tr>
<tr>
<td>4.61</td>
<td>Measurement: DC voltage regulator output impedance.</td>
</tr>
<tr>
<td>4.62</td>
<td>Measurement: linearity of the on-chip calibration DAC: (a) before calibration, (b) after calibration with two 10b off-chip DACs.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1</td>
<td>ADC nodes of interest.</td>
</tr>
<tr>
<td>B.2</td>
<td>Extracted simulations with $f_{\text{clk}} = 3 \text{ GHz}$: (a) input bandwidth, (b) SNDR vs. $f_{\text{in}}$ at two different points in the circuit.</td>
</tr>
<tr>
<td>B.3</td>
<td>Floorplan of the ADC.</td>
</tr>
<tr>
<td>B.4</td>
<td>Comparison of simulated (RCc extracted) and measured SNDR. The 5 dB discrepancy is due to device mismatch and noise which are not included in this simulation.</td>
</tr>
</tbody>
</table>
List of Tables

1.1 State of the art wireline transceivers, ordered by data rate. 6
2.1 State of the art wireline transmitters. 15
3.1 State of the art bidirectional wireline transceivers. 33
A.1 Summary of contributions. 111
List of Acronyms

2-PAM  two-level pulse amplitude modulation
4-PAM  four-level pulse amplitude modulation
ADC    analog-to-digital converter
APS    Advanced Parallel Simulator
BER    bit error rate
CDF    cumulative distribution function
CDR    clock and data recovery
CML    current-mode logic
CMU    clock multiplier unit
CTLE   continuous-time linear equalizer
DAC    digital-to-analog converter
DDJ    data-dependent jitter
DFE    decision-feedback equalizer
DNL    differential nonlinearity
DUT    device under test
FFE    feed-forward equalizer
FIR    finite impulse response
IC     integrated circuit
INL    integral nonlinearity
I/O    input/output
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISI</td>
<td>inter-symbol interference</td>
</tr>
<tr>
<td>NEXT</td>
<td>near-end crosstalk</td>
</tr>
<tr>
<td>NRZ</td>
<td>nonreturn-to-zero</td>
</tr>
<tr>
<td>PCB</td>
<td>printed circuit board</td>
</tr>
<tr>
<td>PD</td>
<td>phase detector</td>
</tr>
<tr>
<td>PDF</td>
<td>probability distribution function</td>
</tr>
<tr>
<td>PLL</td>
<td>phase-locked loop</td>
</tr>
<tr>
<td>PRBS</td>
<td>pseudo-random bit sequence</td>
</tr>
<tr>
<td>PSD</td>
<td>power spectral density</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>SAR</td>
<td>successive-approximation register</td>
</tr>
<tr>
<td>SNDR</td>
<td>signal-to-noise-and-distortion ratio</td>
</tr>
<tr>
<td>THA</td>
<td>track-and-hold amplifier</td>
</tr>
<tr>
<td>TI</td>
<td>time-interleaved</td>
</tr>
<tr>
<td>UI</td>
<td>unit interval</td>
</tr>
<tr>
<td>VCO</td>
<td>voltage-controlled oscillator</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Motivation

POWER efficiency and data rate are the two main measures of any high-performance chip-to-chip data communication system. Implicit in this statement is the requirement that bit error rate (BER) be maintained at a sufficiently low level. In the case of a CPU-memory channel this could be as low as $10^{-16}$ while for a computer-to-computer Ethernet channel it might be as high as $10^{-6}$. Achieving low BER at high data rate requires compensating for non-ideal characteristics of the channel such as limited bandwidth, crosstalk, and random noise. To this end, equalization of high-frequency signals is needed.

At the same time, power efficiency is a primary concern. Electric power consumed by information technology equipment in data centres accounts for 0.5% of total worldwide electricity use, while another 0.5% is needed for cooling and power distribution [1]. Energy used per year is thus approximately equivalent to the potential energy stored in 88 million barrels of oil, or about 18 times the amount of oil released in the 2010 Gulf of Mexico oil spill. Notwithstanding the effect on electricity usage, increasing power efficiency allows more functionality to be implemented in a system, for example more high-speed transceivers in an FPGA.

In recent years there has been increasing commercial interest in blade servers as a replacement for traditional rack-mount servers. Blade servers feature a smaller form factor so that up to 128 servers can be

Figure 1.1: Cisco blade switch.
located in a single chassis. Each server is essentially a large line card that plugs into a common motherboard, or backplane. An example of a blade-style ethernet switch is shown in Fig. 1.1. The worldwide server industry has total revenues of approximately $10 billion\(^1\). High-speed transceivers are also heavily used in FPGAs ($2.75 billion worldwide) where the ability to adapt to different channel conditions is necessary.

The backplane channel is extremely inhospitable to high-speed communication. As shown in Fig. 1.2 a single chassis can be fairly large, implying channel lengths up to 1 m. In addition, the dielectric material used in older systems is often FR-4 which has significant loss at high frequency. These backplanes can comprise up to 16 layers and contain many vias. All of these factors mean that precise control of channel impedance is difficult and reflections due to impedance discontinuities are common.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig12.png}
\caption{HP blade server.}
\end{figure}

\section*{1.2 Channel Description}

Wireline systems consist of two chips communicating over a physical channel. While copper wires have traditionally dominated this application, there is constant debate about the potential advantages of optical media because of their low loss and relaxed equalization requirements \cite{2}. However, turn of the century speculation that “optoelectronic interconnects may revolutionize system implementation in the next decade” has not been borne out \cite{3}. Poor performance of integrated silicon photodetectors limits data rates to 5 Gb/s \cite{4}. Multi-chip solutions that take advantage of discrete photodetectors have trouble competing with single-chip CMOS solutions on the basis of cost. In this work we do not consider optical channels further, but instead focus on copper channels and electrical signaling.

\footnote{http://www.idc.com/getdoc.jsp?containerId=prUS22360110}
Copper channels normally consist of some combination of bondwires, package traces, printed circuit board (PCB) traces, connectors, and cables. These components introduce frequency-dependent attenuation as a result of skin effect and dielectric losses. Transceiver equalization performance is often quoted in terms of the amount of loss (at half the data rate) that can be equalized. While low-frequency communication is possible on unterminated channels, maximizing the data rate requires termination to the channel’s characteristic impedance. Termination, often to 50 Ω, ideally removes all reflections so that the next symbol can be transmitted without waiting for the line to “ring up”. In practice, impedance discontinuities from packages, connectors, and the chips themselves cause significant reflections that impede communication. Via stubs can also be a source of reflections [5]. Stubs can be counterbored to reduce parasitic inductance and capacitance, but at some cost. The presence of reflections can be seen in the channel frequency response, where they show up as regularly-spaced bumps.

While relatively accurate analytical models can be constructed for simple channels like coaxial cables, for backplane channels this is not trivial. However, it is fairly easy to simply measure the frequency response of the channel of interest and compute the impulse response for use in system-level simulations. A backplane channel is pictured in Fig. 1.3(a). The frequency response of this channel is shown in Fig. 1.3(b) for two different motherboard lengths. Note that reflections are present in the 4” channel. The 24” channel has similar impedance discontinuities, but the reflections have been attenuated over the longer length of that channel so they are not as evident in the frequency response. Intentional roughness

Figure 1.3: (a) Backplane channel, (b) frequency response.
in the copper surface also contributes to ripples in the channel frequency response [6].

The 24" channel is shown with an expanded frequency axis in Fig. 1.4. The computed impulse response is also shown, sampled at 5 Gb/s and at 8 Gb/s. At 5 Gb/s the inter-symbol interference (ISI) is concentrated in one post-cursor sample, whereas at 8 Gb/s one pre-cursor and two post-cursor samples have significant ISI.

Eye diagrams corresponding to the 24' backplane are shown in Fig. 1.5. Even though the impulse response sampled at 5 Gb/s contained only one significant ISI sample, the eye diagram is almost closed. At 5 Gb/s and above, equalization is required for this channel.
1.3 State of the Art Wireline Transceivers

A comparison of recent state-of-the-art wireline transceivers is shown in Table 1.1. The table is ordered by data rate, but also includes information about power dissipation, equalization strategy, and fabrication technology. This information is plotted graphically in Fig. 1.6. These wireline transceivers can be categorized in several ways:

1.3.1 Transmit Driver — voltage-mode vs. current-mode

The fastest transmitters use current-mode signaling, while the most power-efficient use voltage-mode signaling. Fig. 1.7 illustrates the two types of signaling in a differential system with receiver termination. With current-mode signaling only one quarter of the supply current is delivered to the load. The other three quarters is dissipated in the transmitter. By contrast, with voltage-mode signaling all of the supply current is delivered to the load. Voltage-mode drivers are also capable of higher swing, with $V_{Rx} = \frac{1}{2}V_{DD}$ in the case of transmitter and receiver both terminated to the same impedance. In a current-mode driver the switches are typically implemented as NMOS transistors (see Fig. 1.8), and such a large swing would most likely put them in the triode region.

Since the transistors in a current-mode driver operate exclusively in the saturation region, they tend to be faster than voltage-mode drivers whose transistors cycle between saturation, triode, and cutoff. Current-mode drivers generate little power supply noise because the same current is simply switched between two neighbouring resistors. Voltage-mode drivers often require power supply regulation, both to reduce the effect of power supply noise and to provide low-impedance nodes at voltages other than the main supply voltage [13,23]. Another advantage of current-mode drivers is that it is easier to perform equalization while maintaining an impedance match. Voltage-mode drivers require a high degree of segmentation to enable equalization in the first place, and maintaining a constant impedance at the same time is a challenge. For these reasons, the highest speed transceivers use current-mode drivers.
### Chapter 1 Introduction

<table>
<thead>
<tr>
<th>Reference</th>
<th>Data Rate (Gb/s)</th>
<th>Power Efficiency (mW/Gb/s)</th>
<th>Channel (attenuation)</th>
<th>Tx FFE taps</th>
<th>Tx driver (V/I)</th>
<th>CTLE</th>
<th>Rx DFE taps</th>
<th>ADC</th>
<th>Tech. (nm)</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>40</td>
<td>90.0</td>
<td>1 m cable (4.6 dB)</td>
<td>-</td>
<td>I</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>130</td>
<td>2009</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>4.1</td>
<td>40 cm FR4 (21 dB)</td>
<td>3</td>
<td>I</td>
<td>✓</td>
<td>1</td>
<td>-</td>
<td>65</td>
<td>2009</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>11.0</td>
<td>2&quot; backplane (20 dB)</td>
<td>4</td>
<td>I</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>90</td>
<td>2006</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>6.3</td>
<td>25 cm FR4</td>
<td>3</td>
<td>I</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>90</td>
<td>2008</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>13.0</td>
<td>3&quot; FR4 (15 dB)</td>
<td>5</td>
<td>I</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>65</td>
<td>2009</td>
</tr>
<tr>
<td>12</td>
<td>15</td>
<td>6.5</td>
<td>18&quot; backplane (21 dB)</td>
<td>3</td>
<td>I</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>65</td>
<td>2008</td>
</tr>
<tr>
<td>13</td>
<td>12.5</td>
<td>1.0</td>
<td>27 cm FR4 (12 dB)</td>
<td>2</td>
<td>V</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>65</td>
<td>2010</td>
</tr>
<tr>
<td>14</td>
<td>12.5</td>
<td>26.4</td>
<td>30 cm low-loss PCB</td>
<td>4</td>
<td>I</td>
<td>-</td>
<td>5</td>
<td>-</td>
<td>4.5b flash</td>
<td>2x TI</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>3.15</td>
<td>6&quot; FR4 (10 dB)</td>
<td>2</td>
<td>I</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>32</td>
<td>2009</td>
</tr>
<tr>
<td>16</td>
<td>11.8</td>
<td>6.6</td>
<td>24&quot; PCB (25 dB)</td>
<td>3</td>
<td>I</td>
<td>✓</td>
<td>4</td>
<td>-</td>
<td>32</td>
<td>2010</td>
</tr>
<tr>
<td>17</td>
<td>11.3</td>
<td>8.4</td>
<td>24&quot; backplane (13 dB)</td>
<td>4</td>
<td>I</td>
<td>✓</td>
<td>5</td>
<td>-</td>
<td>40</td>
<td>2011</td>
</tr>
<tr>
<td>18</td>
<td>10.3</td>
<td>25.2</td>
<td>75 cm backplane (25 dB)</td>
<td>3</td>
<td>I</td>
<td>✓</td>
<td>1</td>
<td>-</td>
<td>90</td>
<td>2009</td>
</tr>
<tr>
<td>19</td>
<td>10</td>
<td>50.0</td>
<td>backplane (26 dB)</td>
<td>3</td>
<td>I</td>
<td>adj.</td>
<td>6b flash</td>
<td>-</td>
<td>65</td>
<td>2010</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>1.8</td>
<td>17&quot; backplane (18 dB)</td>
<td>2</td>
<td>V/I</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>45</td>
<td>2010</td>
</tr>
<tr>
<td>21</td>
<td>8.9</td>
<td>1.9</td>
<td>16&quot; backplane (27 dB)</td>
<td>-</td>
<td>V</td>
<td>-</td>
<td>IIR</td>
<td>-</td>
<td>65</td>
<td>2009</td>
</tr>
<tr>
<td>22</td>
<td>8</td>
<td>29.0</td>
<td>160 cm backplane (37 dB)</td>
<td>5</td>
<td>V</td>
<td>✓</td>
<td>2</td>
<td>-</td>
<td>90</td>
<td>2008</td>
</tr>
<tr>
<td>23</td>
<td>6.25</td>
<td>2.2</td>
<td>80 cm FR4 (15 dB)</td>
<td>-</td>
<td>V</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>90</td>
<td>2007</td>
</tr>
<tr>
<td>24</td>
<td>6.25</td>
<td>69.0</td>
<td>36&quot; backplane (20 dB)</td>
<td>4</td>
<td>I</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>130</td>
<td>2005</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>56.0</td>
<td>cable (15 dB)</td>
<td>2</td>
<td>I</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>5b flash</td>
<td>4x TI</td>
</tr>
<tr>
<td>26</td>
<td>4.3</td>
<td>3.3</td>
<td>3&quot; FR4 (4 dB)</td>
<td>-</td>
<td>V</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>40</td>
<td>2010</td>
</tr>
<tr>
<td>27</td>
<td>3.6</td>
<td>7.5</td>
<td>8 cm FR4 (12 dB)</td>
<td>2</td>
<td>V</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>180</td>
<td>2004</td>
</tr>
</tbody>
</table>

1. tx driver: voltage-mode (V) or current-mode (I)
2. CTLE: continuous-time linear equalizer
3. TI: time-interleaved
4. IIR: infinite impulse response

**Table 1.1**: State of the art wireline transceivers, ordered by data rate.
Figure 1.6: Power efficiency vs. data rate for state-of-the-art wireline transceivers. In the bottom figure, marker size is proportional to channel loss equalized.
Chapter 1 Introduction

Figure 1.7: Differential signaling: (a) current-mode driver, (b) voltage-mode driver.

Figure 1.8: Typical implementation of: (a) current-mode driver, (b) voltage-mode driver.

1.3.2 Equalization — FFE, CTLE, DFE

Most of the transceivers in Table 1.1 use some kind of equalization. The lone exception is [26] which transmits at a relatively low rate over a short channel. The remaining transceivers use a combination of transmit-side feed-forward equalizer (FFE), continuous-time linear equalizer (CTLE), inductive termination, receive-side decision-feedback equalizer (DFE), and full ADC. Simple channels with smooth frequency responses can be equalized with only a two-tap FFE and a CTLE as shown in Fig. 1.9. The gentle roll-off in gain can be compensated for by reducing the DC swing with pre-emphasis and providing high-frequency boost at the receiver.

A channel frequency response that contains a notch cannot be so easily equalized because
no amount of boost is sufficient for power to be received at the notch frequency. Instead, a non-linear circuit such as a DFE is needed. In a DFE, knowledge of previously received symbols is used to cancel post cursor ISI. This cancellation can be done using a single comparator with a threshold that is dependent on previous bits [28]. However, with only 1 UI between decisions (125 ps at 8 Gb/s) it is difficult to choose the correct threshold in time to sample the next bit. Instead, the DFE is often implemented in a lookahead fashion, where the received eye is sampled multiple times and the correct version subsequently chosen with a mux [29].

Fig. 1.10 shows the benefit of a DFE. In the 24" backplane channel, the eye is almost closed at 5 Gb/s. If we use a 1-tap DFE and choose a different sampling threshold depending on whether the previous bit was 1 or 0, we effectively split the eye into two. These two eyes have larger vertical and horizontal openings and can be sampled without error.

Fig. 1.11 shows the results obtained with the 24" channel at 8 Gb/s. With a 1-tap DFE the two eyes are still closed, however a 2-tap DFE opens the eye. Continuing to increase the number of taps in this way, all post-cursor ISI can be cancelled. However, in the lookahead architecture the number of comparators increases exponentially with the number of taps, making long lookahead DFEs impractical.

Equalizers with more taps are able to equalize lossier channels. However, after a certain point there are diminishing returns to increasing equalizer complexity and the only result is higher power. Fig. 1.12 illustrates this effect for a backplane channel. With no CTLE, increasing Tx FIR taps from one to two produces a sizeable increase in maximum data

Figure 1.9: Linear transceiver with eye diagrams.
Chapter 1 Introduction

Figure 1.10: 24” backplane channel 5 Gb/s eye. 1-tap decision feedback equalization splits the eye into two based on the previous bit, providing a larger eye to sample.

Figure 1.11: 24” backplane channel 8 Gb/s eye. A 1-tap DFE is no longer effective. However, splitting the eye into four (2-tap DFE) based on the previous two bits opens the eye.
rate. However, further increases are only marginally beneficial. Increasing the number of DFE taps appears to always help, however the geometric increase in number of filter taps shown in the figure hides the diminishing effectiveness of each additional tap. Finally, if a DFE skips the first post-cursor tap its effectiveness is greatly reduced.

1.3.3 Equivalence of DFE and ADC

While many-tap DFEs are impractical due to tight timing constraints in the feedback path, the same post-cursor equalization ability can be achieved using an analog-to-digital converter (ADC). As shown in Fig. 1.13, both lookahead DFEs and flash ADCs depend on an array of comparators sampling the input. In a DFE the correct sampling threshold is quickly determined and the information from the remaining comparators is discarded. By contrast, an ADC keeps the information obtained from all comparators which allows processing in the digital domain. While the resulting digital block may be fairly complex, the area and power of such a block is subject to scaling. Each new process node allows more digital functionality for the same power and area.

ADCs and DFEs differ in their choice of comparator thresholds. While DFE thresholds are often chosen adaptively based on BER or eye opening criteria, ADC thresholds are

\footnote{According to Intel system-level simulations}
Chapter 1 Introduction

Figure 1.13: The flash ADC and lookahead DFE architectures both operate by sampling with several comparators at different thresholds.

typically uniformly-spaced in order to minimize quantization noise. For an ADC in a wireline receiver, compressing the full scale range can improve BER for a given ADC resolution [31]. This BER improvement is due to the fact that comparator thresholds far from the zero crossing of the input signal provide little additional information. Also, non-uniform threshold spacing may reduce the number of comparators needed although more investigation is required [32]. However, while the number of comparators is reduced, the required precision of the remaining comparators is still high.

1.3.4 Summary — state of the art

Looking again at Table 1.1, we see that the most power efficient transceiver [13] uses a voltage-mode driver with two-tap pre-emphasis and a receive-side CTLE. This configuration delivers the most “bang for the buck” in terms of channel loss equalized for a given power efficiency.

The winner in terms of most channel loss equalized is [22] which uses a 5-tap FFE, CTLE, and a 2-tap DFE. This equalization performance is achieved at the expense of
higher power and a lower data rate.

The highest data rates are achieved by sacrificing all other attributes, as with [7]. This transceiver uses only transmit- and receive-side CTLEs with several transformers to provide 4.6 dB of equalization at 40 Gb/s. Much of the boost is used to increase the limited bandwidth of the circuit itself as opposed to the channel bandwidth.

The three ADC-based transceivers shown are not competitive in terms of power efficiency and are middle-of-the-pack in terms of channel loss equalized. Their advantage lies in the fact that complex routines such as MIMO equalization or crosstalk cancellation can be implemented in low-cost digital circuits. That the reported transceivers have not leveraged that advantage into increased equalization effectiveness suggests that more bits of resolution may be needed.

1.4 Thesis Organization

With the backplane channel of Fig. 1.3 as a given, this thesis looks at ways to improve chip-to-chip communication performance through equalization. Performance is measured by data rate, power efficiency, and type and length of channel equalized.

Chapter 2 describes the design, implementation, and measurement of a 6-tap fractionally-spaced FFE for equalizing long backplane channels at 6.5 Gb/s. Chapter 3 describes a frequency-division bidirectional transceiver assembled using the FFE from Chapter 2 along with a dual-path hysteresis-based receiver designed by Masum Hossain, a colleague. Chapter 4 describes the design, implementation, and measurement of a redundancy-based 5-bit flash ADC with resonant clocking for use in an ADC-based receiver. Some analysis of the redundancy-based ADC concept and the resonant clocking strategy is also presented.
Chapter 2

6.5 Gb/s Fractionally-Spaced Feed-Forward Equalizer in 90nm CMOS

This chapter describes a 6.5 Gb/s transmitter for use in backplane links. This transmitter, originally presented in [33], incorporates a finite impulse response (FIR) filter with programmable tap spacing in the output driver to compensate for ISI. Using jitter-minimizing tap weights computed with a behavioural model of the transmitter, it is shown that at 6.5 Gb/s jitter is reduced by over 50% by using a tap spacing of 0.53 unit interval (UI) instead of the usual 1 UI.

2.1 Motivation

Chapter 1 mentioned the diminishing returns obtained when the number of FFE taps is increased beyond a certain point. The channel impulse response only contains a certain number of non-zero ISI terms that can be neutralized. Making the filter response longer than the channel response results in little benefit. Typically, the delay between filter taps in an FFE is one UI. However, if we allow fractional UI spacing between taps, it is possible to shape the impulse response more precisely in order to reduce received jitter and increase the received eye opening. Fig. 2.1 shows the same impulse response sampled with different time spacings between filter taps. Note how shrinking the tap spacing has increased the number of ISI-contributing taps.
## Table 2.1: State of the art wireline transmitters that are not already included in Table 1.1.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Data Rate (Gb/s)</th>
<th>Power Efficiency (mW/Gb/s)</th>
<th>FFE taps</th>
<th>Tap Spacing (UI)</th>
<th>FFE bits</th>
<th>Driver Mode</th>
<th>Output Swing</th>
<th>VDD (V)</th>
<th>Area (um²)</th>
<th>Tech. (nm)</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>[34]</td>
<td>50</td>
<td>10.6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>I</td>
<td>1200</td>
<td>2.0</td>
<td>2.1</td>
<td>65</td>
<td>2009</td>
</tr>
<tr>
<td>[35]</td>
<td>20</td>
<td>5.3</td>
<td>2</td>
<td>½</td>
<td>I</td>
<td>-</td>
<td>1.8</td>
<td>0.7</td>
<td>90</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>[36]</td>
<td>20</td>
<td>2.9</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>I</td>
<td>300</td>
<td>1.2</td>
<td>0.89</td>
<td>65</td>
<td>2010</td>
</tr>
<tr>
<td>[37]</td>
<td>16</td>
<td>3.6</td>
<td>4</td>
<td>1</td>
<td>V</td>
<td>-</td>
<td>500</td>
<td>1.0</td>
<td>0.013</td>
<td>65</td>
<td>2007</td>
</tr>
<tr>
<td>[38]</td>
<td>12</td>
<td>11.1</td>
<td>10</td>
<td>½</td>
<td>I</td>
<td>-</td>
<td>1.0</td>
<td>0.18</td>
<td>90</td>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>[39]</td>
<td>10</td>
<td>17.4</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>I</td>
<td>450</td>
<td>1.65</td>
<td>-</td>
<td>90</td>
<td>2005</td>
</tr>
<tr>
<td>[40]</td>
<td>9.6</td>
<td>6.3</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>I</td>
<td>500</td>
<td>2.5</td>
<td>0.56</td>
<td>130</td>
<td>2005</td>
</tr>
<tr>
<td>[41]</td>
<td>8.5</td>
<td>11.3</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>V</td>
<td>500</td>
<td>1.5</td>
<td>0.065</td>
<td>65</td>
<td>2008</td>
</tr>
<tr>
<td>[42]</td>
<td>8</td>
<td>20.5</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>I</td>
<td>700</td>
<td>1.2</td>
<td>0.68</td>
<td>130</td>
<td>2007</td>
</tr>
<tr>
<td>[43]</td>
<td>7.4</td>
<td>4.3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>V</td>
<td>400</td>
<td>1.1</td>
<td>-</td>
<td>45</td>
<td>2010</td>
</tr>
<tr>
<td>[44]</td>
<td>6</td>
<td>4.0</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>I</td>
<td>300</td>
<td>1.0</td>
<td>0.036</td>
<td>90</td>
<td>2006</td>
</tr>
<tr>
<td>[45]</td>
<td>5</td>
<td>10.4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>I</td>
<td>400</td>
<td>1.8</td>
<td>0.39</td>
<td>180</td>
<td>2010</td>
</tr>
<tr>
<td>This work</td>
<td>6.5</td>
<td>6.5</td>
<td>6</td>
<td>var.</td>
<td>4</td>
<td>I</td>
<td>300</td>
<td>1.2</td>
<td>0.1</td>
<td>90</td>
<td>2008</td>
</tr>
</tbody>
</table>

1driver mode: voltage-mode (V) or current-mode (I)
2swing: single-ended
Figure 2.1: Impulse response with samples taken at three different intervals for a nominal data rate of 5 Gb/s. As the delay between filter taps shrinks, more ISI becomes targetable by the FIR equalizer.

2.2 Literature Survey

Table 2.1 shows the recent standalone wireline transmitters, which can be split into voltage-mode and current-mode drivers. These drivers all provide pulse shape control, with either analog or digital control of amplitude. While output driver linearity limits the equalization performance of the transmitter, it is difficult to provide linearity while maintaining adequate voltage swing. The values of $V_{\text{DD}}$, $V_{\text{th}}$, and $V_{\text{swing}}$ determine linearity. The tail current source can be designed with a lower $V_{\text{eff}}$ alleviating the headroom problem somewhat at the expense of larger devices. Large devices have more parasitic capacitance which reduces their high-frequency output impedance. Transmitters in the recent literature that report a $\frac{V_{\text{swing}}}{V_{\text{DD}}}$ ratio higher than 30% are either voltage-mode transmitters [37, 41, 43] or use inductors to boost the output voltage higher than the supply [34, 42]. Larger swing is often obtained by simply using a higher supply voltage and thick-oxide devices [40].

Tap spacing is usually fixed at 1 UI. This spacing is easy to generate with a series of flip flops all clocked from the same source [36], although usually a half-rate [39] or
quarter-rate [44] architecture is used to reduce clocking power. The series of flip flops is attractive because jitter is the same for all taps. Some designs use a tap spacing of \( \frac{1}{2} \) UI to facilitate duobinary signaling [38] or to enable communication at 20 Gb/s [35]. The \( \frac{1}{2} \) UI spacings are also generated with cascaded flip flops. Finally, RF-style transmitters use transformer coupling to sum the contributions of several amplifiers with different frequency responses [34, 42]. These designs cannot be said to have a certain tap spacing, but they produce a similarly pre-emphasized output. They also produce a large output swing, though at the expense of increased area used by the transformers.

Current- and voltage-mode drivers each have their advantages. Current-mode drivers often use a polysilicon resistor to provide termination, which is relatively constant up to the frequency of the output pole. Summing multiple signals is easily accomplished by shorting the corresponding differential pairs and connecting them to the termination resistor. Complex pulse shapes can be created in this way, although bandwidth and power efficiency decline as complexity increases. Most transmitters make a tradeoff between flexibility and efficiency. For example in [46] the maximum currents of the four taps are scaled by the ratio 1:4:2:1. While this configuration limits the amplitude of the pre-cursor tap to at most one quarter that of the main tap, it has half the parasitic capacitance of the case where the taps are scaled by the ratio 4:4:4:4. There are many variations on this technique.

### 2.3 Circuit Design

The proposed FIR equalizer, shown in Fig. 2.2, consists of a six-tap delay line with a variable gain stage for each tap. The delay line is tuneable to allow the use of different tap spacings and different bit rates. Each tap has adjustable gain with 4-bit resolution. The variable tap spacing can be used as an additional knob to tune the impulse response of the filter. Another benefit of variable spacing could be to produce a roaming filter tap used to cancel reflections that occur at unpredictable time offsets from the main pulse. Another idea is to make a multi-tap filter where the delay of each tap is independently tunable. Though interesting, these extensions of the variable tap delay concept are not explored further here.
The variable gain stage is broken into six slices, one for each tap of the delay line. The currents from the six slices are summed in 50 Ω load resistors to produce the output voltage. As shown in Fig. 2.3, each slice of the gain stage consists of a sign selection switch followed by a preamplifier and a 3-bit adjustable output driver, for a total of 4 bits. When a given filter tap is not in use, the corresponding gain stage slice is automatically shut down in order to save the power that would have been burned in the preamplifier. Unfortunately, digital-to-analog converter (DAC) linearity limits performance. However, as long as we can characterize the nonlinearity we can take it into account when optimizing the pulse shape for equalization of a given channel.

The delay line generates six phase-shifted versions of the input data using five delay cells. Each delay cell consists of four stages like the one in Fig. 2.4(a) with a differential pair driving symmetric loads. This type of load acts as a resistance with an adjustable
value. As the resistance increases, the delay of the stage increases while the voltage swing is kept constant by a replica bias feedback loop that reduces the tail current. Four stages are required in order for each cell to generate a significant delay while still maintaining the required bandwidth. With fewer stages the delay per cell would be too small.

The delay cell is tuneable from 62–216 ps, as shown in Fig. 2.4(b). Increasing the delay reduces the power dissipation from 40 down to 12 mW, providing power scaling for slower data rates. At a delay of 120 ps the bandwidth of the delay line is sufficient for data rates only up to 5 Gb/s. To achieve higher delay while maintaining bandwidth, multiple delay cells can be used as one by simply turning off the intervening taps. For example, to generate a 125 ps tap spacing, we can use either one delay cell with a delay of 125 ps or two cascaded delay cells each with a delay of 62.5 ps. The latter option will result in higher bandwidth. Sample waveforms for this transmitter are shown in Fig. 2.5.

2.4 Measurement Results

The transmitter test board is shown in Fig. 2.6. The main feature of the transmitter is tuneability of output amplitude and tap spacing. Measurements of 3-bit transmitter
amplitude control are shown in Fig. 2.7. Measurements of delay line control are shown in Fig. 2.8. Both the amplitude and delay vary non-linearly. In these figures only one filter tap is active at a time.

As the delay of the line is increased, the bandwidth of each delay cell goes down resulting in deterministic jitter at the transmitter output. Fig. 2.9 shows transmit side data-dependent jitter (DDJ) vs. tap spacing for a PRBS7 pattern at two different bit rates. When the delay reaches 0.7 UI the delay line can no longer operate as the jitter has increased too much. This limitation forces the use of multiple delay cells in cascade when implementing longer delays.

With more than one tap active, equalization can be performed. Fig. 2.10 shows two-tap equalization with seven different values for the second tap. The limited bandwidth of the output node can be seen in this series of eye diagrams. Although all eyes except for (a) are programmed for equalization, only eyes (e)-(g) show any overshoot. In eye (d) tap weights [+7 -3] are just enough to equalize the transmitter’s own high-frequency loss.

Fig. 2.11 shows two-tap equalization for a nominally-constant DC eye amplitude. For a perfectly linear transmitter, the DC eye amplitude should be constant as the tap weights are increased from [+2 -1] to [+7 -6]. Clearly nonlinearity is a factor, as the transmit-side
Figure 2.6: Transmitter test board.
Figure 2.7: Measurement: 3-bit transmitter amplitude control at 6 Gb/s. Differential eye amplitude as shown.

Figure 2.8: Measurement: Transmitter tap spacing control. Tap spacing shown as control voltage is increased linearly from 100 mV to 800 mV.
Figure 2.9: Transmit-side DDJ vs. delay per cell of the delay line. The signal goes through the entire delay line before going to the output.

Figure 2.10: Measurement: Transmit-side eye diagrams at 5 Gb/s for two-tap equalization as boost is increased. RMS jitter is shown for each eye.
eye is closed with tap weights of [+7 -6].

For a 6-tap filter with 4-bit resolution for each tap weight, there are $(2^4)^6 \approx 1.6$ million possible filter configurations for a given delay setting. To help with tap weight selection a behavioural model of the transmitter and channel was created. To start, we measure the nonlinearity of the output driver DAC. The nonlinearity was characterized by measuring the output swing for all tap weight settings with only one tap operational. The linearity of the output driver DAC is shown in Fig. 2.12. Since the linearity is far from ideal, we take it into account when choosing the transmitted pulse shape.

A slew rate limitation is then added, with the value that is observed in measurement. With six output driver slices whose currents are summed together, there is significant parasitic capacitance at the output node that limits the slew rate.

This slew-rate-limited signal is then put through an RLC circuit that models the parasitics of the QFN package. Finally a frequency response computed from the measured s-parameters of the backplane channel gives us the resulting receive-side signal.

Using this model, behavioral simulations were performed to evaluate all possible tap weight settings. For each configuration (channel length, number of taps, and tap spacing) the tap weights resulting in the lowest simulated jitter were selected.

Fig. 2.13 shows the match between model and measurement for three sample eye diagrams at 5 Gb/s. Fig. 2.13(a) shows the unequalized signal received over a 24" backplane, while (b) and (c) show the transmit- and receive-side signals for the optimal transmitted pulse shape. The tap weights for the optimal pulse in this case are [+1 +5 -3] with a tap spacing of 100 ps.

An eye diagram of the received signal over 24" backplane for optimal tap weights and tap spacing at 6.5 Gb/s is shown in Fig. 2.14. In this case, the optimal tap spacing was 0.53 UI, much less than the typical tap spacing of 1 UI.

The tap weights were chosen only to minimize DDJ, so we use the jitter decomposition feature of the oscilloscope to examine only the part of the jitter that is due to ISI. Random jitter remains roughly constant across all transmitter configurations around 1 ps rms.

Fig. 2.15(a) shows the improvement in DDJ as the tap spacing is varied for a data rate of 6.5 Gb/s. Since DDJ is bounded, unlike random jitter, peak-to-peak DDJ is plotted. At 6.5 Gb/s the equalizer benefits significantly from using a smaller tap spacing; DDJ can
Chapter 2 6.5 Gb/s Fractionally-Spaced Feed-Forward Equalizer in 90nm CMOS

Figure 2.11: Measurement: Transmit-side eye diagrams at 5 Gb/s for nominally-constant DC eye amplitude and increasing boost. RMS jitter is shown for each eye.

Figure 2.12: Linearity of the output driver DAC. (a) Output swing for each digital code, (b) DNL and INL.
Figure 2.13: Measured and simulated behavioral model 5 Gb/s eye diagrams for a PRBS7 pattern: (a) Output of unequalized 24" backplane channel, (b) transmitter output for half-baud-spaced jitter-minimizing pulse shape, (c) equalized 24" backplane channel output.
Figure 2.14: Received 6.5 Gb/s equalized eye diagram for the 24” backplane channel. Tap spacing is 0.53 UI and the optimal tap weights are [+2 -1]. Total transmitter power for this configuration is 42 mW.

be halved compared with conventional baud-rate tap spacing. As shown in Fig. 2.15(b), even at 5.5 Gb/s the jitter varies by almost 10 ps as the tap spacing is changed, underlining the importance of choosing the optimal tap spacing. The total power of the delay line is also plotted for both of these graphs. Note that the jump in the middle of these graphs is caused by the limited bandwidth of the delay line. Once the desired tap spacing is greater than twice the minimum delay of the line (62.5 ps), two delay cells are combined with their individual delays halved.

A die photo of the transmitter, implemented in 90 nm CMOS, is shown in Fig. 2.16. The power varies from 40–80 mW depending on the tap spacing and output swing. At 5 Gb/s with half-baud tap spacing and a 500 mV output swing, 63 mW is consumed. Of this, 23 mW is consumed in the delay line and 40 mW is consumed in the output drivers.

2.5 Conclusions

The design of this transmitter aims for flexibility in the way the transmitted pulse is shaped by providing six taps with variable spacing and variable amplitude. In circuit design, adding flexibility or reconfigurability means trading off other characteristics such as bandwidth. In this case the trade off was made so that a wider range of transmitter configurations could be explored with a single circuit. However, transmitter non-idealities
Figure 2.15: Measured receive-side DDJ for the 24" backplane channel at (a) 6.5 Gb/s, (b) 5.5 Gb/s. Optimal tap weights are recalculated for each tap spacing individually.
such as output driver nonlinearity and delay line jitter turn out to limit performance such that no optimal pulse shapes use more than three taps. The superfluous taps serve only to decrease the bandwidth at the output node and increase the power dissipated in the delay line.

In the proposed transmitter, the size of each output driver slice is the same. This choice leads to maximum flexibility, as any of the six taps can be the main tap that provides the bulk of the power. However, that flexibility is not needed in the context of equalizing a single backplane link (it may help perform transmit-side deskew). A more efficient solution would be to fix one tap as the main tap and shrink the device sizes in the remaining taps to keep the output pole frequency as high as possible. Another solution would be to segment the output driver into $N$ equal slices, each of which can act as any of the six taps. However, this solution requires extra digital processing of signals before the output driver and starts to approach the complexity of a current-mode DAC.

The measurement results presented here show the benefit of fractional tap spacing for equalizing a backplane link. However, this tap spacing was implemented with a variable delay line dissipating 23 mW at 5 Gb/s. A $\frac{1}{2}$ UI-spaced delay line implemented with a series of flip flops may be more power efficient. In addition, output driver non-linearity could be reduced in future with calibration of tail current sources. This change would
likely improve equalization performance.
Chapter 3

5 Gb/s Downstream, 500 Mb/s Upstream
Bidirectional Data Communication

This chapter describes a bidirectional transceiver combining the transmitter from Chapter 2 with a receiver presented in [47] which was designed by my colleague Masum Hossain. This transceiver, presented in [48], uses frequency-division multiplexing to allow bidirectional communication over a single differential channel. In this application, the flexibility of the 6-tap FFE allows it to be reconfigured to transmit either a low data rate signal with little high frequency content, or a high data rate signal with no DC content. Simultaneous bidirectional communication is shown in simulation, while only sequential bidirectional communication is shown in measurement because the prototype transceiver was not fully integrated. Impedance discontinuities in the prototype board result in a closed eye when both transmitters are operating.

3.1 Motivation

The goal of a wireline communication system is to maximize the throughput between two chips. This can be accomplished by increasing the data rate over a single channel, up to the point where high-frequency attenuation limits BER. Multiple pins can be dedicated to I/O to increase aggregate data rate further. However, as gate lengths scale down the number of pins on a chip has not kept pace with the functionality that can be integrated. The result is that increasing the number of pins dedicated to I/O takes pins away from other functions such as providing low-impedance power supply rails. A
solution to this problem is to enable bidirectional communication over a single channel. If communication can take place at least half as quickly as in the unidirectional case, then bidirectional communication can provide some benefit.

Aside from the goal of maximizing throughput, enabling data transmission from the receiver to the transmitter has its own benefits. This backchannel allows adaptation of transmit-side equalization based on the received signal and the decisions made in the receiver. Because only a small amount of data is needed for adaptation, this kind of bidirectional link can have highly asymmetrical data rates. The asymmetrical link can be useful as long as the low-speed backchannel does not significantly impact the high-speed forward channel.
Chapter 3 5Gb/s Downstream, 500Mb/s Upstream Bidirectional Data Communication

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[49]</td>
<td>10</td>
<td>1:1</td>
<td>26</td>
<td>3m AWG28</td>
<td>I</td>
<td>D</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.2</td>
<td>1.02</td>
<td>110</td>
<td>2007</td>
</tr>
<tr>
<td>[50]</td>
<td>8</td>
<td>1:1</td>
<td>14</td>
<td>4.6' FR4</td>
<td>V</td>
<td>S</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3.3</td>
<td>0.11</td>
<td>350</td>
<td>2004</td>
</tr>
<tr>
<td>[11]</td>
<td>8</td>
<td>1:1</td>
<td>13</td>
<td>3&quot; PCB</td>
<td>I</td>
<td>D</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>1.2</td>
<td>0.17</td>
<td>65</td>
<td>2009</td>
</tr>
<tr>
<td>[51]</td>
<td>5.008</td>
<td>625:1</td>
<td>40</td>
<td>20' FR4</td>
<td>I</td>
<td>D</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>1.3</td>
<td>0.3</td>
<td>130</td>
<td>2005</td>
</tr>
<tr>
<td>[52]</td>
<td>5</td>
<td>1:1</td>
<td>24</td>
<td>5m AWG28</td>
<td>V</td>
<td>D</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.8</td>
<td>0.9</td>
<td>180</td>
<td>2001</td>
</tr>
<tr>
<td>[53]</td>
<td>4</td>
<td>1:1</td>
<td>40</td>
<td>4cm PCB</td>
<td>I</td>
<td>D</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.8</td>
<td>0.13</td>
<td>180</td>
<td>2003</td>
</tr>
<tr>
<td>[54]</td>
<td>4</td>
<td>1:1</td>
<td>18</td>
<td>10cm FR4</td>
<td>I</td>
<td>S</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>1.8</td>
<td>0.6</td>
<td>180</td>
<td>2005</td>
</tr>
<tr>
<td>[55]</td>
<td>4</td>
<td>1:1</td>
<td>7</td>
<td>5cm PCB</td>
<td>V</td>
<td>S</td>
<td>✓</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.8</td>
<td>0.022</td>
<td>100</td>
<td>2005</td>
</tr>
<tr>
<td>This work</td>
<td>2.75</td>
<td>10:1</td>
<td>26</td>
<td>12&quot; cable</td>
<td>I</td>
<td>D</td>
<td>-</td>
<td>-</td>
<td>✓</td>
<td>-</td>
<td>1.2</td>
<td>0.145</td>
<td>90</td>
<td>2009</td>
</tr>
</tbody>
</table>

1 driver mode: voltage-mode (V) or current-mode (I)

Table 3.1: State of the art bidirectional wireline transceivers.
Chapter 3  5Gb/s Downstream, 500Mb/s Upstream Bidirectional Data Communication

3.2 Literature Survey

Table 3.1 shows the recent bidirectional wireline transceivers. The concept of pins as a scarce resource forces us to evaluate transceivers in terms of their total throughput per pin. That is, the sum of the data rates in two directions divided by the number of pins required. In this light, a differential transceiver must operate at twice the rate of a single-ended transceiver in order to compete. So while the fastest bidirectional transceiver operates differentially [49], we find that single-ended transceivers are indeed competitive despite the additional impairments they encounter [50,54]. The most power-efficient transceiver is also single-ended, in addition to making use of multi-level signaling [55].

Fig. 3.2 shows block diagrams for the two most popular bidirectional implementations. Fig. 3.2(a) is representative of most bidirectional communication schemes that use echo cancellation to separate the two signal directions. See for example [49,50,52,53,55]. In echo cancellation systems it can be difficult to achieve good matching between the transmit path and the replica path, especially for a range of channels and packages. Fig. 3.2(b) is a frequency-division system some variation of which is seen in many RF transceivers. In [54] this type of system is used in a wireline transceiver. However, the complexity of a full RF transceiver is not needed to achieve frequency-separation between signal paths. Fig. 3.2(c) shows a simplified frequency-division system that uses first-order filters. Shifting some of the signal-processing burden from the filters to the transmitter and receiver enables this simplification.

The ability of bidirectional transceivers to equalize difficult channels is limited. While several unidirectional transceivers from Table 1.1 can provide at least 20 dB of boost, most of the bidirectional transceivers in Table 3.1 operate over benign channels. The challenge of removing the unwanted signal is difficult enough without adding a harsh backplane environment. This discrepancy calls into question the basis for using bidirectional communication in a wireline system, as the channel must be of sufficient length for this effort to be worthwhile. Nevertheless, some applications may benefit from the use of bidirectional techniques.
Chapter 3 5Gb/s Downstream, 500Mb/s Upstream Bidirectional Data Communication

Figure 3.2: Three bidirectional communication schemes.

(a) echo cancellation

(b) frequency-division multiplexing

(c) frequency-division multiplexing with simple filters


text content
transitions at the transmitter. The receiver must be able to recover the original NRZ data from these AC pulses. Since the DC content is not received in any case, it can be eliminated from the transmitted signal as well.

AC coupled receivers typically include a memory element that stores the previously received bit until a transition is detected. When the received signal is corrupted by noise and a bit error occurs, the receiver continues to make errors until the next transition occurs. The impact of these error runs can be minimized by coding the data stream to contain frequent transitions [61].

### 3.3.2 AC-DC Bidirectional Link

To make use of the high frequency channel bandwidth we propose an AC channel using the same physical wire as the baseband channel. The AC and DC channels are frequency-separated by a simple filter consisting of passive elements. The two channels can then be used as a bidirectional link by sending the frequency-separated signals along the wire in opposite directions. A schematic of the bidirectional link is shown in Fig. 3.4(a).

The data rates of the AC and DC channels must be chosen so as to maximize the aggregate data rate while still allowing the two receivers to recover independent data simultaneously. As seen in Fig. 3.1, both signals contain negligible energy at frequencies equal to their respective data rates. So intuitively it is attractive to set the data rate of the AC signal to twice that of the DC signal to place the frequency of maximum AC power at
Chapter 3  5Gb/s Downstream, 500Mb/s Upstream Bidirectional Data Communication

Figure 3.4: (a) Proposed AC/DC bidirectional link, (b) frequency response of different paths in the AC/DC bidirectional link with $L = 80$ nH and $C = 10$ pF.

A null in the DC power. In practice, the data rate of the AC signal must be several octaves higher than that of the DC signal in order to allow filtering of the two frequency bands.

Frequency selection is performed by simple filters, as mentioned earlier in Fig. 3.2. In this case the filters are implemented using inductors and capacitors. These elements were chosen because the transmitter and receiver test chips were in our possession and simple passive filters could easily be implemented on-board. In a monolithic implementation, the filters could also be integrated on-chip. These simple filters do not strongly separate the two frequency bands. The frequency responses of the various paths through the system are shown in Fig. 3.4(b). There remains significant near-end crosstalk (NEXT) between the AC (DC) transmitter and DC (AC) receiver. The -3 dB frequency of the filters should be chosen to limit this interference and permit signal detection at both ends of the link.

In addition, to produce a maximally-flat frequency response for both the AC and DC channels, the impedance seen by the transmitter should be equal to 50 Ω across all frequencies. Unfortunately, the -3 dB frequencies of the AC and DC channels cannot be more than one octave apart without affecting the characteristic impedance of the channel. The -3 dB frequency of the AC channel should be chosen first to allow recovery of the high-frequency AC data. Then, the -3 dB frequency of the DC channel should be chosen to satisfy the requirement of 50 Ω impedance across all frequencies. The remaining low-frequency bandwidth is then used for DC data. The relationship between the data rates...
and -3 dB frequencies is discussed further with the simulation results in this section.

### 3.3.3 Transmitter Configuration

The transmit-side equalizer is used to shape the signal before it is launched into the channel. This pulse-shaping eases the requirements on the signal separation filter and allows it to be implemented with a single inductor and capacitor. While the same transmitter is used for both AC and DC channels, it is configured differently in these two cases to maximize signal power in the desired frequency range. Five possible transmitter configurations are shown in Fig. 3.5. The power spectra of the signals are plotted for a PRBS $2^{11} - 2$ pattern at 5 Gb/s. The transmit-side filter allows the signal power to be controlled by emphasizing different parts of the spectrum. The typical pre-emphasis pulse shape in Fig. 3.5(b) boosts power by 6 dB at half the data rate. The AC pulse shape in Fig. 3.5(c) produces zero DC power and a 20 dB/dec rolloff below half the data rate. The pulse shapes with fractionally-spaced taps allow attenuation of the high-frequency content. The slew-rate limited pulse in Fig. 3.5(d) is close to optimal from a standpoint of maximizing the ratio of power below the data rate to power above the data rate.

Additional shaping of the transmitted spectrum can be achieved by coding the data. Codes such as 8B10B, [61], are often used to guarantee a certain transition density for the purposes of clock and data recovery. A similar code could be used to reduce the signal power at low frequency.

### 3.3.4 Receiver

The receiver passes the incoming signal through a pre-amplifier before equalizing it using the weighted sum of two paths. The first path is a hysteresis circuit that can recover a full-swing NRZ signal from low-swing AC pulses. The second path is a linear amplifier that amplifies the data transitions to compensate for the reduced bandwidth of the hysteresis circuit. A schematic of the receiver is shown in Fig. 3.6.

This receiver was first presented in [47] as an AC-coupled receiver. Because of the adjustable weighted summer, when no equalization is required the hysteresis path can be turned off which increases the bandwidth and reduces jitter. Measurement results of the
Figure 3.5: Five different transmitter pulse-shaping configurations with the power spectra obtained when transmitting a PRBS $2^{11} - 1$ sequence at 5 Gb/s. These pulse shapes produce identical peak voltage swing.
receiver recovering data from 14 Gb/s AC pulses are shown in [47].

3.3.5 Bidirectional Simulations

Unidirectional simulations of the AC and DC channels are shown in Fig. 3.7(a,b). These are the outputs of the two channel directions when one transmitter is using the AC pulse (Fig. 3.5(c)) and the other transmitter is using the slew-rate limited pulse (Fig. 3.5(d)). The limited bandwidth of the DC channel causes additional ISI at the output of the channel. The transmitted AC pulse has been decreased in amplitude by the attenuation of the channel.

Bidirectional simulation results for the AC and DC channels are shown in Fig. 3.7(c–f). The maximum achievable data rates are 500 Mb/s in the DC band and 8 Gb/s in the AC band, limited by crosstalk between the two links. In both cases, the receiver is set up so that most of the signal comes through the hysteresis path, with only a small amount coming through the linear path. In Fig. 3.7, comparing (a)/(c) and (b)/(d) shows the interference introduced by bidirectional operation.

The link was simulated using a 10 cm microstrip line and filter values $L = 80 \, nH$ and $C = 10 \, pF$. These values set the cutoff frequency separating the two bands to 180 MHz. This frequency cannot be increased to permit a higher data rate in the DC link because the AC signal still has significant energy down to a few hundred megahertz. When this energy is filtered out, the received AC signal becomes unrecoverable. In contrast, the AC data
Figure 3.7: Simulated eye diagrams showing bidirectional operation. Data rates are 500 Mb/s for the DC channel and 8 Gb/s for the AC channel. Signals (e) & (f) have been recovered by slicing with the hysteresis-based receiver.
rate can be reduced arbitrarily as long as the pulses sent contain sufficient high frequency content.

To allow a lower ratio of forward- to back-channel data rates, the cutoff frequency would have to be moved higher towards the AC data rate. Recovery of AC pulses with significant low-frequency energy removed requires a topology more sophisticated than that of the dual-path receiver shown here.

### 3.4 Measurement Results

The bidirectional test board is shown in Fig. 3.8. To show the feasibility of the proposed bidirectional communication channel, we present measurement results of the receiver recovering data sent by the transmitter over a short SMA cable channel. In these results the passive filter is imitated by changing the pulse shape used by the programmable transmitter.

To demonstrate the feasibility of bidirectional operation, we show:

- The receiver can recover high-speed AC data.
- The receiver can recover low-speed DC data.
- The power spectra of the DC and AC signals are sufficiently separated to allow filtering by simple passive elements.

The eye diagrams in Fig. 3.9 show a 5 Gb/s AC pulse input to the receiver and the open eye at the output of the receiver. To recover the data with no errors, the adjustable hysteresis threshold of the receiver must be set precisely in the threshold adjustment area shown in Fig. 3.9(a). This hysteresis threshold allows the receiver to distinguish between genuine pulses that should be amplified and noise or reflections that should be attenuated.

For a linear equalizer, an open eye at the output of the receiver is a good indication that no errors are being made. However, because our receiver contains a nonlinear hysteresis block, it is possible to obtain an open eye even when bit errors are present. To verify correct operation, a complete PRBS sequence of length $2^7 - 1$ is shown in Fig. 3.10 for both the input to the AC receiver and the recovered data.
Chapter 3  5Gb/s Downstream, 500Mb/s Upstream Bidirectional Data Communication

Figure 3.8: Bidirectional test board.
Chapter 3 5Gb/s Downstream, 500Mb/s Upstream Bidirectional Data Communication

**Figure 3.9:** Measured 5Gb/s eye diagrams for the (a) input to the AC receiver and (b) the equalized output. Separate transmitter and receiver chips were flip-chip bonded to a printed circuit board. Hysteresis threshold must be set within the threshold adjustment area shown in (a).

**Figure 3.10:** (a) Received and (b) recovered 5Gb/s AC-coupled PRBS sequence of length $2^{27} - 1$. The eye diagrams for these signals are shown in Fig. 3.9.
Recovery of low-speed data can be easily accomplished by the DC receiver. In addition, the spectrum of the transmitted DC signal can be shaped by adjusting the pulse shape produced by the transmitter. Figs. 3.11(a) and (b) show the eye diagram of a normal DC pulse along with its power spectrum. By shaping the transmitted pulse to produce longer rise and fall times as shown in Fig. 3.11(c), the high-frequency content of the DC pulse can be suppressed by 12 dB as shown in Fig. 3.11(d).

By comparing the power spectrum of this 500 Mb/s slew-rate limited DC signal with the power spectrum of a 5 Gb/s AC signal we can see that the majority of the power in each signal occupies a frequency band that is underused by the other signal. The relevant spectra are shown in Fig. 3.12. The spectrum of the 5 Gb/s AC signal consists of a number of peaks rising out of the noise floor because the signal is a PRBS sequence that repeats every $2^7 - 1$ bits. The peaks occur at frequencies that are integer multiples of $\frac{5 \text{GHz}}{2^7 - 1} = 39.4 \text{MHz}$. If we plot only the peaks of the two power spectra against a logarithmic frequency scale, we obtain the graph shown in Fig. 3.13 where we can see that the AC and DC signals occupy distinct frequency bands.

### 3.5 Future Work

In [62], a technique was presented that allows both DC and AC connections to the same chip. This technique, shown in Fig. 3.14, uses capacitive plates to send the AC signal from one chip to another. The DC contact is made by a solder bump. Alignment of the capacitive plates is made possible because the solder bump is buried in the substrate, allowing the capacitive plates to be sufficiently close together. The buried-bump technique fits this bidirectional application because the AC coupled interconnect provides the required capacitance for free while the DC contact through the buried solder bump leaves us free to set the series inductance on-chip.

### 3.6 Conclusions

This chapter showed how a 6-tap FIR transmitter with variable tap spacing could be used to facilitate bidirectional communication over a single wireline channel. This transmitter
Figure 3.11: Eye diagrams and power spectra of two different 500 Mb/s DC pulse shapes. High-frequency content is reduced for the slew-rate limited signal in (c).
Figure 3.12: Power spectra of (a) a 500 Mb/s DC signal and (b) a 5 Gb/s AC signal.

Figure 3.13: DC and AC spectra from Fig. 3.12, plotted on a logarithmic frequency scale.
allows the power spectrum of a transmitted signal to be shaped so that high-frequency and low-frequency bands can more easily be separated. A drawback of this simple frequency-division bidirectional technique is that the upstream and downstream signals must be widely separated in frequency.

While simulation results show simultaneous bidirectional operation, this was not possible in measurement due to impedance discontinuities on the test board. The separation between dies on the test board led to reflections that closed the received eye. Future work on this topic would involve integrating the transmitter and receiver on the same die with either on-chip or in-module passive elements to provide filtering. With all components integrated on-chip, reflections would be minimized.

Figure 3.14: AC/DC bidirectional link using buried-bump technology.
Chapter 4

3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

This chapter describes the design, implementation, and measurement of a 5-bit flash ADC in 65 nm CMOS. Novelties include resonant clock path and use of redundant comparators. Resonant clocking overcomes the problem of clock skew in an ADC without a track-and-hold amplifier (THA) and reduces clocking power. Redundant comparators eliminate the need for separate threshold tuning circuits for each comparator. ADC nonidealities are investigated and an analytical relationship between offset variation and SNDR is derived. The ADC test chip also includes a voltage regulator, calibration DAC, 256-bit shift register, and five 50 Ω output buffers. Testing is accomplished with the help of a Stratix IV FPGA used to capture 5-bit-wide ADC output data at the full rate. Although the ADC achieves an SNDR greater than 24 dB at sampling frequencies between 2 and 3 GHz, input bandwidth is a disappointing 20 MHz due to excessive parasitic capacitance in the Gray encoder circuit. The ADC consumes 17.6 mW at a sampling frequency of 3 GHz.

4.1 Motivation

ANALOG techniques such as FIR filtering, decision-feedback equalization, and the use of peaking amplifiers have dominated high-performance wireline transceivers in the recent past. But the endpoint of wireline communication is a fully ADC-based transceiver that would move complex equalization and clock recovery blocks into the digital domain. Advancements in fabrication technology can then be exploited to reduce power and area. Results have shown that feed-forward digital clock and data recovery is possible up to
12.5 Gb/s [14]. An ADC-based receiver also allows complex clock and data recovery (CDR) circuits to be implemented entirely with synthesized logic, eliminating challenging analog blocks [63].

The ADC resolution required for effective equalization varies with the channel under consideration. Two ADC-based receivers intended for backplane or multimode fiber channels used a resolution of 6 bits [19,64]. By allowing non-uniform quantization even 2 bits can be enough to allow an open eye at 10 Gb/s over a 3' backplane channel [32]. Since a 1- or 2-tap DFE is often enough to equalize a wireline channel, this observation bolsters the idea of DFE and ADC equivalence mentioned in Sec. 1.3.3. However, in cases where uniform quantization is used several results indicate that at least 4.5 bits are necessary [14,65]. One exploration of ADC-based receivers for backplane channels showed that 4.5 to 5.5 bits are needed depending on the length of the channel [66]. Higher resolution enables higher data rates just as more DFE taps enable higher data rates (shown in Fig. 1.12).

Fig. 4.1 gives the power efficiency of recently-reported high-speed CMOS ADCs. Also plotted is a line equal to 2 mW/Gb/s which is an approximate limit above which an ADC-based receiver is no longer competitive with a traditional analog receiver. The number 2 mW/Gb/s is, if anything, conservative given the number of full transceivers in Fig. 1.6 that achieve a lower power dissipation using traditional analog techniques. Fig. 4.1 makes clear that at sampling rates greater than 2 GS/s ADC power dissipation increases to the point where a single-ADC-based solution has so far not been viable. To achieve higher data rates the circuit designer must either resort to time-interleaving multiple sub-ADCs or find another way to reduce power.

### 4.1.1 Comparison of ADC Topologies

Reported gigasample-per-second ADCs have been implemented using flash, pipeline, and successive-approximation register (SAR) topologies. Typically pipelined converters are used in applications that require a combination of moderate resolution (~10 bits) and moderate sampling rate (hundreds of MS/s). See [91] for example. In this application space, flash converters cannot compete due to the large power that would be required for so many comparators, while SAR converters would find it difficult to resolve that many bits in such a short sampling period. At higher sampling rates, pipelined converters do not fare
Figure 4.1: Power efficiency vs. sampling rate for recent high-speed CMOS ADCs. In the bottom figure, time interleaved converters are shown in black. The star indicates the work discussed in this chapter.
as well, although some time-interleaved pipelined converters have been reported \cite{70,80}.

Similarly, SAR cannot compete with flash in terms of the speed of a single ADC; the fastest reported CMOS single SAR ADC has a sampling rate of 2.5 GS/s \cite{78} while several flash converters have attained that speed and the fastest can reach 5 GS/s \cite{83}. The advantage of the SAR topology lies in the fact that the number of comparisons required increases linearly with the number of bits, while the number of comparisons required in the flash topology increases exponentially. In practice, SAR converters are more energy efficient at six bit resolution and higher, while flash converters are more efficient at four bit resolution and lower \cite{92}. At the crossover point of 5 bit resolution, the preferred topology depends on circuit implementation.

Based purely on power efficiency, an argument can be made that SAR should be the topology of choice in an ADC-based wireline receiver. This is especially true for channels that require more than 5 bits to equalize. However, another factor is the latency of the converter. Recent DSP-based wireline receivers have used Müller-Mueller timing recovery \cite{14,19,64}, one advantage of which is fast convergence to the optimal sampling phase \cite{93}. Significant ADC latency nullifies this advantage somewhat. In addition, ADC latency can decrease high-frequency jitter tolerance and lead to instability of the CDR loop \cite{94}.

**Time-Interleaved**

Time-interleaving is an attractive technique to increase the sampling rate and/or the power efficiency of an ADC. In fact, the only ADC in Fig. 4.1 that achieves sub-10-mW/GS/s efficiency at a sampling rate greater than 2.2 GS/s employs a time-interleaved flash architecture \cite{68}. That work used a reduced-rate clock which is time-shifted and distributed to the various sub-ADCs. The power benefit comes from the fact that individual sub-ADCs typically have higher power efficiency at lower sampling rates and therefore the aggregate power is lower than that of an equivalent single ADC. However, special attention must be paid to timing skew and other mismatches to avoid undue SNDR degradation. Several authors have investigated these effects and proposed techniques to deal with them \cite{95–99}.

Other designers use time-interleaving to push the sampling rate as high as possible, though at the cost of increased die area and power dissipation. For example, in \cite{78} an
array of 16 SAR ADCs was used to achieve an aggregate sampling rate of 40 GS/s. This work performed calibration using a sine-wave input, synthesized on-chip with the help of a narrowband amplifier. In [100], 160 time-interleaved pipeline ADCs were used to achieve 8 bit resolution at 20 GS/s at the expense of a 9 W power dissipation and a die area of 196 mm².

Time-interleaved ADCs in general have more complicated calibration procedures because of the need to independently tune the gain, offset, linearity, and timing skew for each sub-ADC. Calibration becomes more difficult as the number of sub-ADCs increases. Because of the higher sampling rate possible with the flash topology, a lower interleaving factor can be used which eases the calibration requirements.

**Summary**

While time-interleaving is necessary to achieve the highest sampling rate or the best power efficiency, in this work we do not consider it further. Instead we limit ourselves to discussing techniques for reducing power dissipation in a single ADC, keeping in mind that several of these ADCs might be time-interleaved in a practical wireline communication system. In particular we consider the flash topology for the advantages detailed above: its power efficiency at low resolution, low latency, and higher sampling rate.

### 4.1.2 Proposed ADC

A typical high-speed ADC consists of a THA, a reference generation circuit, a clock buffer, an array of comparators, and some decoding logic. From a power standpoint, we can neglect the contribution of the reference circuit which can be as low as 50 µW [69]. We propose a three-part solution to reduce the power dissipated by the remaining blocks:

- Use redundant comparators to address comparator offset variation. This technique allows the use of smaller, lower-power comparators that would normally have a large variance in offset. Redundancy allows us to choose the devices with the correct threshold and turn off the rest so that they dissipate no power.

- Remove the THA and analyze the effect of timing skew on signal-to-noise-and-distortion ratio (SNDR). Previous work has analyzed the SNDR penalty associated
with clock skew in time-interleaved ADCs [98] but not the effect of clock skew in a single ADC lacking a THA.

- Add an inductor to terminate the clock network, reducing the required clock buffer power and practically eliminating clock skew at a particular frequency. This technique has been used in a wireline transceiver [23] but is also useful in the clock path of an ADC.

Sections 4.2.1 and 4.2.2 analyze the effect of threshold variation and clock skew on the SNDR of the ADC. Section 4.3 analyzes the expected threshold variation in the redundant comparator scheme. Section 4.4 analyzes the clock skew that can be expected in the proposed resonant clocking scheme. Section 4.7 shows the results of simulation and measurement of the proposed 5-bit flash ADC.

4.2 Flash ADC Non-Idealities

4.2.1 Comparator Offset

We would like to find the SNDR of a quantization-noise-limited N-bit ADC:

$$SNDR = \frac{P_{signal}}{P_{noise}}$$

(4.1)

The quantizer is defined by a set of threshold levels $m_i$ and output levels $n_i$ located at:

$$m_i = \frac{1}{2} \text{LSB}, \frac{3}{2} \text{LSB}, ..., \left(\frac{2^N - 1}{2}\right) \cdot \text{LSB}$$

(4.2)

$$n_i = 0, \text{LSB}, ..., \left(\frac{2^N - 1}{2}\right) \cdot \text{LSB}$$

(4.3)

where:

$$\text{LSB} = \frac{V_{ref}}{2^N}$$

(4.4)

The ADC transfer function is shown in Fig. 4.2(a) for the case of a 2-bit converter. Fig. 4.2(b) shows the quantization noise for a given input. Fig. 4.2(c) shows the squared quantization noise. For an input signal with a uniform probability distribution, the av-
Figure 4.2: (a) Transfer function of a 2-bit A/D converter, (b) quantization noise, (c) squared quantization noise, (d) probability distribution of a sinusoidal input, (e) weighted, squared quantization noise.

Average squared quantization noise gives us the quantization noise power. For a sinusoidal input, the quantization noise must be weighted by the input distribution, as shown in Fig. 4.2(d,e). The threshold levels $m_i$ are chosen to maximize SNDR for a sinusoidal input rather than to minimize maximum quantization noise. If the thresholds shift from these locations, additional quantization noise results. For a sinusoidal input the quantization noise added depends on which threshold is shifted. However, in the following analysis we assume a uniform input distribution because it introduces negligible inaccuracy for converters with more than 4 bits. We can describe the SNDR as:

$$SNDR = \frac{P_{signal}}{P_{noise,ideal} + P_{noise,added}}$$  \hspace{1cm} (4.5)
Figure 4.3: (a) Transfer function of a 2-bit A/D converter with one threshold shifted from its ideal location, (b) quantization noise, (c) squared quantization noise with shaded area showing additional squared noise due to threshold offset.
where $P_{\text{noise,ideal}}$ is the quantization noise power when the decision thresholds are at their optimal locations. The offset of a single threshold causes additional quantization noise power equal to the shaded area in Fig. 4.3(c) divided by the full scale range:

$$P_{\text{noise,added,single}} = \frac{1}{V_{\text{FSR}}} \int_{m_i}^{m_i+v_{\text{offset}}} \Delta V_Q^2 \cdot dv_{\text{in}}$$  \hspace{1cm} (4.6)$$

$$= \left( \frac{1}{(2^N - 1) \text{ LSB}} \right) \int_{m_i}^{m_i+v_{\text{offset}}} \left[ (v_{\text{in}} - n_i)^2 - (v_{\text{in}} - n_{i+1})^2 \right] dv_{\text{in}}$$  \hspace{1cm} (4.7)$$

$$= \left( \frac{1}{(2^N - 1) \text{ LSB}} \right) (1 \text{ LSB}) v_{\text{offset}}^2$$  \hspace{1cm} (4.8)$$

$$= \left( \frac{v_{\text{offset}}^2}{2^N - 1} \right)$$  \hspace{1cm} (4.9)$$

The effect of a single threshold offset on SNDR is shown in Fig. 4.4. For a single threshold shift of $\frac{1}{2}$ LSB, the SNDR penalty is 0.4 dB. Since there are $2^N - 1$ thresholds, the total noise added due to threshold offsets on all comparators is:

$$P_{\text{noise,added}} = \sum_{i=1}^{2^N - 1} \left( \frac{v_{\text{i,offset}}^2}{2^N - 1} \right) \text{(uniform input distribution)}$$  \hspace{1cm} (4.10)$$

If each $v_{\text{i,offset}}$ is a Gaussian random variable with zero mean and variance $\sigma^2$, then $P_{\text{noise,added}}$ is a chi-squared random variable with $2^N - 2$ degrees of freedom [101]. This random variable has the following properties:

$$\text{mean} (P_{\text{noise,added}}) \approx \sigma^2$$  \hspace{1cm} (4.11)$$

$$\text{stdev} (P_{\text{noise,added}}) \approx \frac{1}{2} \sigma^2$$  \hspace{1cm} (4.12)$$
The PDFs of $v_{\text{offset}}$, $P_{\text{noise, added}}$, and SNDR are plotted in Fig. 4.5 for the case of a 5-bit quantizer.

The pdf of SNDR can be used to find the yield of an ADC given the expected threshold variance. Fig. 4.5(c) shows us that for the case of a 5-bit quantizer with $\sigma = \frac{1}{4}$ LSB, 95% of the ADCs produced will have SNDR greater than 28.6 dB. To evaluate the inaccuracy added by the uniform input distribution assumption, we compare these calculations with the result of 10,000 behavioural simulations. The results are shown in Fig. 4.6. The behavioural simulations use a sinusoidal input and all thresholds in the ADC are varied randomly about their nominal positions. The analytical formulation is slightly pessimistic as $\sigma_{v_{\text{offset}}}$ approaches $\frac{1}{2}$ LSB. If mismatch exceeds that amount, the circuit would best be redesigned as a 4-bit ADC.

### 4.2.2 Clock Skew

Skew in clock arrival time to different points in the comparator array degrades the performance of a flash ADC. This skew effectively introduces a threshold offset as shown in Fig. 4.7. If the input is a sinusoid, it can be represented as:

$$v = \sin (2\pi f_{\text{in}} t),$$

(4.13)
Figure 4.6: Results of behavioural simulation for a 5-bit ADC (10,000 trials): (a) SNDR histograms with threshold standard deviations as marked, (b) average SNDR vs. $\sigma_{\text{offset}}$ for behavioural simulation and analysis, (c) standard deviation of SNDR.
and we can calculate the effective threshold offset as follows:

\[
v + \Delta v = \sin[2\pi f_{in}(t + \Delta t)]
\]
(4.14)

\[
\Delta v = \sin[2\pi f_{in}(t + \Delta t)] - v
\]
(4.15)

\[
\Delta v = \sin\left[2\pi f_{in}\left(\frac{\sin^{-1}(v)}{2\pi f_{in}} + \Delta t\right)\right] - v
\]
(4.16)

\[
\Delta v = \sin\left[\sin^{-1}(v) + 2\pi T_{skew}\right] - v,
\]
(4.17)

which has two solutions \(\Delta v_1\) and \(\Delta v_2\) shown in Fig. 4.8. The two solutions correspond to sampling the rising and falling edges of the input. These solutions depend only on the threshold voltage, \(v\), and the sampling skew as a fraction of the input period, \(T_{skew} = f_{in}\Delta t\).

We can substitute Eq. (4.17) into Eq. (4.10) to treat skew as if it were threshold offset:

\[
P_{\text{noise,added}} = \sum_{i=1}^{2^{N}-1} \left(\frac{1}{2^{N} - 1}\right)\Delta v_i^2
\]
(4.18)

where the \(\Delta v_i\) are the skew-induced threshold shifts for each comparator based on a sinusoidal input. Fig. 4.9 plots \(\Delta v\) vs. \(T_{skew}\) for each of the 31 threshold levels in a 5-bit ADC. Given a set of clock skews \(\{T_{skew,i}, i = 1 \ldots 2^N - 1\}\), we can calculate a set of threshold shifts \(\{T_{skew,i}, i = 1 \ldots 2^N - 1\}\). In Section 4.4 we use the simulated clock skew from two different clocking schemes and find their effect on SNDR.
Figure 4.8: Change in threshold voltage (a) vs. threshold voltage for a given clock skew, (b) vs. clock skew for a given threshold voltage. There are two solutions in each case.

Figure 4.9: $\Delta v$ vs. $T_{skew}$ for each of the 31 threshold levels in a 5-bit ADC.

### 4.3 Redundancy-Based Flash ADC

Redundancy has been used extensively in pipelined converters to eliminate missing decision levels and allow digital calibration to improve SNDR [102]. Two-times redundancy was used in the flash converter in [103] to reduce the range of offset-correction that had to be accommodated by an output-connected current DAC. The second comparator was only used if the first comparator needed maximum calibration. Another flash ADC used $4 \times$ redundancy along with a fault-tolerant thermometer-to-binary converter to provide an 8-bit output with 6-bit linearity [104].

Other previously-reported redundancy schemes have suffered from the problem that the comparator thresholds are out of order. If not addressed, out-of-order thresholds create bubble errors which can decrease the SNDR. Solutions have involved reordering the thresholds through comparator reassignment [105] and using an adder after the comparator array so that order of thresholds is unimportant [90, 106]. These solutions require more complicated decoding logic which is less suitable for high-speed ADCs. Another solution to
this problem is to individually tune the threshold of each comparator. The tuning voltage in this case typically comes from a multiple-output resistor-string DAC [68,69,81].

In the proposed redundancy scheme each comparator in a traditional flash ADC is replaced by an array of \( k \) comparators, exactly one of which is active at any given time. Fig. 4.10 illustrates the difference between a traditional flash ADC and the proposed approach. In this example, each comparator has been replaced with an array of four comparators and a 4-to-1 multiplexor. If we are allowed to choose the single lowest-offset comparator out of a group of \( k \) comparators, the offset becomes

\[
\nu_{\text{offset, group}} = \text{sign}(\nu_{\text{offset, i}}) \cdot \min_{i=1...k} \{|\nu_{\text{offset, i}}|\}.
\] (4.19)

From [107], the cumulative distribution function (CDF) of \( \nu_{\text{offset, group}} \) is

\[
F_{\nu_{\text{offset, group}}}(x < a) = \frac{1}{2} + \text{sign}(a) \cdot \frac{1}{2} \left( 1 - \left[ 1 - \int_{-|a|}^{|a|} f_{\nu_{\text{offset}}}(x) dx \right]^k \right),
\] (4.20)

and the probability distribution function (PDF) can be obtained by differentiation. Fig. 4.11 illustrates the reduction in threshold variance that can be achieved using redundancy, assuming the threshold offset of a single comparator is a Gaussian random variable with zero mean. Assuming the same area is divided into smaller comparators, the \( \sigma^2_{\nu_{\text{offset}}} \) of individual comparators goes up in inverse proportion to the area of the individual comparators but
4.3.1 SNDR Degradation Due to Offset Variation

Now we can repeat the exercise of Sec. 4.2.1 and find the achievable SNDR for different redundancy factors. The result is shown in Fig. 4.12. We use $\sigma_0$ as the baseline standard deviation of offset for a unit-size comparator, so that a half-size comparator has $\sigma_{v_{\text{offset}}} = \sqrt{2} \cdot \sigma_0$. We are making this comparison on a constant-area basis, so that the line $k = 2$ in Fig. 4.12 represents two comparators with the same total area as the single comparator represented by the line $k = 1$. This graph suggests that any desired SNDR can be achieved without resorting to an explicit threshold tuning mechanism. Although as the redundancy factor increases there is more overhead in terms of power and area that is not captured here.
Chapter 4 3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

Figure 4.12: SNDR of the best 95% of devices vs. threshold standard deviation in a traditional 5-bit flash ADC. SNDR improves with redundancy factor $k$ for a given unit-size comparator offset variance $\sigma_0$.

Figure 4.13: Result of simple calibration illustrated with randomly generated offset values. There are 31 blocks and eight comparators per block. The chosen comparators are circled.

### 4.3.2 Calibration

Calibration consists of choosing the comparators with the lowest offset. The simplest calibration procedure involves setting the reference input to zero and sweeping the ADC input until a given comparator experiences a transition. An example is shown in Fig. 4.13. For each block of $k$ comparators, the comparator with offset closest to zero will be turned on and all others in that block will be turned off. Subsequently, the reference voltage is set and the ADC can be used normally. A single on-chip calibration DAC can be used to perform this calibration.

One problem with this approach is that the relationship between threshold and reference
voltage varies between comparators due to mismatch. So the comparators which are optimal when reference voltage equals zero may no longer be optimal when reference voltage equals $+200 \text{ mV}$. In this case calibration needs to be performed with the reference voltage set to its nominal value, and comparators would be chosen based on their threshold voltages’ proximity to the ideal threshold for that comparator. Fig. 4.14 illustrates the result of a calibration performed in this way. In general, comparators chosen in each block will be different than those chosen in Fig. 4.13.

The two calibration procedures described above use a DC input to the ADC. More precise calibration could be achieved with a sinusoidal input at a frequency of interest. However, this type of calibration may only be useful in the lab due to the high cost of generating a sine wave which is more precise than the ADC itself. Future work could examine ways of performing background calibration while the ADC is operating. The redundant comparators could be chosen by comparison with a single reference comparator designed for low offset. This reference comparator might be used solely for calibration and would be shut down thereafter to save power. One-time calibration can be used to compensate for process variations, but multiple-calibration is needed to counteract the effects of temperature variation on comparator offset.
4.4 Resonant Clock Path

The resonant clocking scheme considered here involves using an inductor to resonate out the interconnect parasitic capacitance as well as the input capacitance of the subsequent clock buffer stages. A similar technique was used in [23] to distribute the clock in a wireline transceiver system. Here the same technique is used to distribute the clock to a comparator array in a flash ADC. The main difference is that in a multi-channel wireline transceiver clock skew between lanes is removed by per-channel clock recovery. In a flash ADC with no THA, clock skew between comparators directly degrades SNDR and must therefore be analyzed. Resonant clocking avoids the high power of conventional, active clock distribution networks such as the current-mode logic (CML) clock tree used in [108].

Using the relationship between clock skew and threshold shift derived in Sec. 4.2.2, we can evaluate the performance of this resonant clocking scheme and compare it to the performance achieved with simple interconnect. The interconnect model is shown in Fig. 4.15.

From simulation, Fig. 4.16 shows clock skew along the length of an RLC delay line for three different clock rates. This simulation assumes that the comparators in the ADC are arranged in order of threshold voltage along a delay line from which they receive the sampling clock. Another comparator ordering would show a different dependence of SNDR on input frequency. In a traditional clocking scheme skew measured in UI increases as the sampling frequency increases. In contrast, resonant clocking can produce approximately zero skew at any single desired frequency. This tuning results in negligible SNDR penalty due to clock skew at that clock frequency. We can translate the skew information taken from simulation into an effective threshold shift using the relationships given in Fig. 4.9. These
threshold shifts can then be related to SNDR by the techniques of Sec. 4.2.1. Fig. 4.17 shows the maximum SNDR achievable in the two clocking schemes. At the clock distribution’s resonant frequency (5 GHz in his example) the SNDR penalty due to clock skew is negligible.

### 4.5 Circuit Design

#### 4.5.1 Flash ADC Architecture

A simplified diagram of the flash ADC test chip is shown in Fig. 4.18. The ADC consists of an array of 256 comparators arranged in 32 groups of 8. Each comparator is followed by an RS latch. Each group of 8 comparators is followed by an 8-to-1 switch. The 31 thermometer-coded outputs are then decoded into a 5-bit gray code. The 32nd comparator is present for balance. All of these blocks are powered by an on-chip regulator. There is also an on-chip calibration block that includes a differential 8-bit DAC as well as memory to store the locations of the chosen comparators. A switch at the input of the ADC allows either normal operation or calibration mode, but does not act as a track and hold.

A more detailed diagram is shown in Fig. 4.19. The clock is taken single-endedly from off-chip before a complimentary clock is generated by using a single inverter. Similarly, the
Chapter 4 3GS/s, 17.6mW Flash ADC in 65nm CMOS

Figure 4.17: Maximum SNDR achievable for two different clocking schemes, considering only the SNDR penalty due to clock skew.

Figure 4.18: Simplified block diagram of the ADC test chip.
output of the gray code logic is single-ended, but a pseudo-differential signal is generated to drive the 50 Ω CML output buffers. A 256-bit shift register stores the configuration of the ADC and peripheral test circuits. A 32-input mux allows low-speed readout of comparator output values to aid in computer-driven testing. Three bias-current inputs are not shown.

### 4.5.2 Dynamic Comparator

A schematic of the dynamic comparator is shown in Fig. 4.20. This is a standard topology that has been frequently used in the literature, see [109] for example. Its great advantage is that positive feedback from the cross-coupled inverters increases the sensitivity so that small differential input voltages can be resolved to full CMOS levels in a short time. Since the comparator consumes no static power, it can be disabled by gating the clock.

Intentional offset can easily be added in several ways: by varying W and L of the input devices [110], by adding additional capacitance to one side of the comparator [69], by attaching a differential current DAC to the output of the comparator [103], or by adding devices in parallel with the input pair [68]. In the proposed ADC we use the latter method, with the gates of the parallel devices attached to a reference ladder to generate 31 different offsets.

In our implementation the transistor sizes are made fairly small because redundancy is being used to correct for offset. A width of 2 μm was used for the input devices. While this width could have been decreased further, parasitics become more significant as finger widths get smaller and smaller.

### 4.5.3 RS Latch

A schematic of the RS latch is shown in Fig. 4.21. It consists of two cross-coupled NAND gates with devices sized so that the input transistors can easily override the retention transistors. Here, the retention transistors are given the minimum width that still allows three contacts to each source and drain. This circuit is also commonly used, see [111] for example.
Figure 4.19: Detailed block diagram of the ADC test chip.
Figure 4.20: Dynamic comparator. Kickback cancellation MOS capacitors are shown on the left-hand side.

Figure 4.21: RS latch.
Chapter 4 3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

4.5.4 Voltage Regulator

A block diagram of the voltage regulator is shown in Fig. 4.22. This design has been adapted from [112]. The main feature of the design is that the functions of load regulation and line regulation have been separated. A low-frequency feedback loop uses a scaled-down copy of the regulator to set the input voltage $v_{loop}$ of the main regulator. The feedback loop ensures that the main regulator output $v_{dd}$ tracks $v_{ref}$ under the condition of zero load current. The load regulation circuit, shown in Fig. 4.23, then adjusts the bias of the main output transistor in response to changing current draw at the output.

The advantage of this topology is that it leads to smaller output ripple when the load

![Voltage regulator](image)

**Figure 4.22:** Voltage regulator.

![Voltage regulator load regulation circuit](image)

**Figure 4.23:** Voltage regulator load regulation circuit.

<table>
<thead>
<tr>
<th>Device</th>
<th>$W$ ($\mu$m)</th>
<th>$L$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>62.4</td>
<td>1000</td>
</tr>
<tr>
<td>$M_2$</td>
<td>0.52</td>
<td>60</td>
</tr>
<tr>
<td>$M_3$</td>
<td>11.44</td>
<td>60</td>
</tr>
<tr>
<td>$M_4$</td>
<td>0.52</td>
<td>500</td>
</tr>
<tr>
<td>$M_5$</td>
<td>31.2</td>
<td>60</td>
</tr>
</tbody>
</table>
current consists of pulses of current. Fig. 4.24, taken from [112], illustrates this effect. Because the low-frequency feedback loop does not respond to changes in load current, the output $v_{dd}$ does not settle back to its nominal voltage before the next current pulse arrives. In the circuit used here, the opamp used in the feedback loop is powered from the scaled-down load regulation circuit. The opamp draws $30 \mu A$, which does not substantially affect the DC value of $v_{loop}$. There is no start-up problem with this circuit, because if $v_{loop} = 0$ and $v_{dd} = 0$, all the current from $M_1$ in Fig. 4.23 will be used to pull down the gate of $M_5$, causing $v_{dd}$ to rise and the opamp to turn on.

### 4.5.5 Inductor Modeling

The clock path requires an inductor with a value of $5.6 \text{nH}$ to tune out its parasitic capacitance. During simulation the inductor must be replaced by a model that includes such non-idealities as interwinding capacitance, series resistance, parasitic capacitance to ground, and substrate resistance. We use the $2-\pi$ model shown in Fig. 4.25(a). To ensure that this model is an accurate representation, we use an electromagnetic simulator to compute the $s$-parameters of a given inductor. From the $s$-parameters we can obtain series resistance, series inductance, parallel capacitance, and capacitance from each terminal to ground. These parameters derived from electromagnetic simulation can then be compared to parameters similarly derived from the $2-\pi$ model. The parameters in the $2-\pi$ model are then adjusted until there is good agreement between the frequency-dependent $R$, $L$, and $C$ values. The final comparison of these values is shown in Fig. 4.25(b).
Chapter 4 3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

(a)

(b)

Figure 4.25: Inductor modeling: (a) 2-π inductor model, (b) comparison between parameters computed from ASITIC s-parameters and parameters computed from frequency response of the above 2-π model. Lines with square markers are values derived from ASITIC, while lines with no markers are the result of a spice simulation of the optimized 2π model.
4.5.6 Clock Path

Several considerations are important in the design of the clock path, shown in Fig. 4.26. The comparators in the ADC are clocked by a single-ended single-phase clock, so it would be natural to use a single-ended clock throughout the clock path. However, because the clock line is terminated by an inductor it makes sense to distribute the clock differentially. It is easier to model a differential inductor than a single-ended one. A single-ended inductor would have one end connected to ground, but in the context of multi-GHz signals it is not obvious exactly how ground should be modeled. Improper inductor modeling can cause the predicted resonant frequency and/or the Q of the clock line to be incorrect. With a differential inductor the ground modeling problem disappears because the centre of the inductor is AC ground as long as the signals seen by the two ends of the inductor are differential.

The differential clock line could be provided from off-chip through two pins, possibly with differential buffers in between. However, providing a differential clock signal from test equipment is not trivial, since most high-quality signal sources have only a single-ended output. Therefore, we provide a single-ended to differential conversion on-chip by simply passing the clock through an inverter to generate a complementary clock. The input
resistance seen from off-chip is provided by a 50 Ω resistor and a self-connected inverter ensures that the input inverter is biased at its threshold voltage.

As seen in Fig. 4.26 there is some feedthrough of the sinusoidal clock to the ground pin due to wirebond inductance. The inductance is high enough that the impedance of the wirebond begins to approach 50 Ω at the clock frequency of interest and so acts as a voltage divider. This feedthrough does not affect the operation of the clock path, but it would affect other circuits on-chip. As a result, this ground is kept separate from other grounds on the test chip.

Using the 2-π inductor model from Sec. 4.5.5 and the distributed model of the clock line from Fig. 4.15 we can simulate to find the bandwidth of the clock line. The differential impedance of the clock line is plotted vs. frequency in Fig. 4.27.

To save power, unused comparators are shut off by gating the clock. The clock gating subcircuit is shown in Fig. 4.28.

### 4.5.7 Layout Considerations

Layout of the comparator array has a significant impact on ADC performance. The impact may be especially large for the case of a redundancy-based flash ADC because of the larger number of comparators, although each comparator is individually small. In the test chip,
Chapter 4 3GS/s, 17.6mW Flash ADC in 65nm CMOS

Figure 4.28: Diagram of the clock gating circuit.

each comparator takes an area of $1.8 \mu m \times 43 \mu m$ while the entire array occupies $460.8 \mu m \times 43 \mu m$. The ADC also consists of a resistive reference ladder, clock distribution line, RS latches, switches, and decoding logic. Ideally all of these subcircuits would be close together so that the connections between them would be short. However, because of the 2-dimensional nature of the die not everything can be adjacent to everything else. Some connections must, by necessity, be longer.

If the comparators were layed out in-order (from high threshold to low threshold) next to a single-string resistive ladder, the routing between comparators and ladder would be extremely congested. See Fig. 4.29(a) for example. In Fig. 4.29(b) the ladder is now doubled-back on itself and the order of comparators has been rearranged. This change has greatly simplified some of the routing, now none of the connections between ladder and comparators cross each other.

From the point of view of clock skew, the order of the comparators can also be important. As discussed in Sec. 4.2.2, a conventional clock distribution line implies different skew at different points along the line. So comparator order would clearly have an effect on SNDR in that case. In the case of resonant clocking, however, there is negligible skew at the resonant frequency and so in this case comparator order is not a consideration.

Finally, connections between 8-to-1 switches and decoding logic would also ideally be
short although it may not be possible to satisfy these multiple constraints simultaneously. In the case where these lines are forced to be longer, additional buffers can be added to trade off increased delay for higher bandwidth.

### 4.5.8 Calibration DAC

The test chip uses a dual 8-bit DAC for calibration of the comparators. A simplified 4-bit version of the DAC is shown in Fig. 4.30. This is a hybrid DAC in which half of the bits are obtained from a traditional resistor string. Switches are used to select two adjacent nodes on the resistor string. The remaining bits are determined by a row of PMOS source-followers in which adjacent devices are connected by a switch. Opening exactly one of the switches will produce a weighted average of the two resistor-string voltages. This weighted average is shifted up by $V_{gs}$ due to the source follower. The current source provides the bias current for the source followers.

As depicted in Fig. 4.30(a), this DAC has the advantage of strict monotonicity shared by a traditional resistor string DAC, while occupying less area due to the smaller number of resistors. However, since the current source is in practice not an ideal one, finite current source output impedance causes the step size to shrink as the output voltage gets closer to $V_{dd}$. This effect is shown in Fig. 4.30(b).
To reduce the integral nonlinearity (INL), we add a mechanism to increase the PMOS tail current as the output voltage becomes too high. The modified DAC is shown in Fig. 4.31. As the output voltage rises, additional current sources are switched on to maintain a constant current. With the coarse 2-bit control of tail current shown in Fig. 4.31(a), INL is reduced from 10 LSBs to 4 LSBs though at the cost of surrendering monotonicity.

### 4.5.9 Biasing Circuits

Bias currents were necessary for opamps used in the voltage regulator, calibration DAC, and at the node setting the input common-mode voltage level. A schematic of the opamp bias circuit is shown in Fig. 4.32. Where device lengths are non-minimum, this has been done to minimize transistor variation. A schematic of the biasing opamp is shown in Fig. 4.33.

### 4.6 Selected Simulation Results

This section presents simulations results for more in-depth exploration of two critical features of the proposed ADC: the dynamic comparator itself and the resonant clocking scheme.

#### 4.6.1 Dynamic Comparator Simulations

The testbench used for these dynamic comparator simulations is shown in Fig. 4.34. A veriloga block is used to generate the inputs to the comparator in response to its decisions. This block performs a binary search to find the offset of the comparator. Once the offset is found, the input amplitude is swept back and forth to find the hysteresis. Another veriloga block connects the reference input of the comparator to a resistor ladder. This testbench allows the offset and hysteresis of the comparator to be determined with a single run, which facilitates monte carlo and sweep-based simulations.

Fig. 4.35 shows the result of one simulation with this testbench. While the comparator without parasitics has no offset, the RCc extracted comparator shows an offset of -95 mV. The table at right in the figure shows the parasitic capacitances in the comparator. Node
Figure 4.30: (a) Example of a segmented 4-bit DAC. This is a hybrid 2-bit resistor string, 2-bit current-mode DAC. (b) DC nonlinearity of the DAC.

Figure 4.31: (a) DAC modified to reduce nonlinearity, (b) DC nonlinearity.
Figure 4.32: Schematic of the opamp bias circuit.

Figure 4.33: Schematic of the opamp circuit.
Chapter 4 3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

Figure 4.34: Testbench used for dynamic comparator simulations.

Figure 4.35: Veriloga-aided simulation converges on the comparator offset. The ideal comparator (no parasitics) has no offset, whereas the $RC_c$ extracted comparator has an offset of $-95\, \text{mV}$. The table at right shows the largest parasitic capacitances connected to the output nodes of the comparator. Comparator nodes names are shown in Fig. 4.20.

<table>
<thead>
<tr>
<th>node1</th>
<th>voutp</th>
<th>voutn</th>
</tr>
</thead>
<tbody>
<tr>
<td>gnd</td>
<td>1.1 fF</td>
<td>0.9 fF</td>
</tr>
<tr>
<td>vdd</td>
<td>0.8 fF</td>
<td>0.4 fF</td>
</tr>
<tr>
<td>vout[n/p]</td>
<td>0.7 fF</td>
<td>0.7 fF</td>
</tr>
<tr>
<td>X1</td>
<td>0.2 fF</td>
<td>0.8 fF</td>
</tr>
<tr>
<td>X2</td>
<td>0.3 fF</td>
<td>0.2 fF</td>
</tr>
</tbody>
</table>
voutp has 0.6 fF more capacitance to supply, while node voutn has 0.5 fF more capacitance to nodes X1 and X2. Since the supplies are constant, whereas X1 and X2 drop quickly at the sampling instant this capacitance imbalance results in a systematic negative offset. Care must be taken to ensure a balanced layout, although this goal is in conflict with the desire for a small footprint.

To confirm the sensitivity of the comparator to small differences in parasitics connected to the output node, Fig. 4.36 shows the change in offset of an ideal comparator as an unbalanced load capacitor (to ground) is increased from zero to 5 fF. Such a capacitance imbalance is exacerbated by increasing the common-mode input voltage. We can see this both in Fig. 4.36(b) where the existing imbalance in the parasitics is made worse at higher common-mode input voltages, and in Fig. 4.36(c) where the perfectly balanced ideal comparator exhibits greater variation in offset at higher input common mode.

While a higher common mode worsens mismatch, it increases the maximum sampling rate of the comparator. Fig. 4.37 shows that hysteresis increases for high sampling rates when the common-mode input voltage is decreased. Hysteresis is an indication of insufficient comparator speed, or equivalently, insufficient time for the precharge or decision phases. Input common-mode voltage must be chosen with an eye to optimizing the trade-off between speed and mismatch. In my measurements, a common-mode level of 700 mV is used for both reference and signal inputs.

Another trade-off exists in the choice of differential reference voltage, $V_{\text{ref,diff}}$. As $V_{\text{ref,diff}}$ is
increased, greater offset variation is tolerable because the full-scale range is being increased. Fig. 4.38 shows the behaviour of the ADC threshold voltages as $V_{\text{ref,diff}}$ is increased. Although the average threshold voltage follows the expected linear relationship, the threshold voltage variation increases quickly once $V_{\text{ref,diff}}$ is greater than 250 mV.

While the signal input to the ADC receives significant attention, the reference input is equally important. The reference is essentially another signal input to the ADC, albeit a DC input. In the proposed ADC an 18 kOhm resistive ladder provides reference voltages to the 256 comparators. Choosing a higher resistance results in less power dissipation but more kickback (despite the kickback cancellation capacitors). Fig. 4.39 shows the effect of decreasing the reference ladder resistance by ten times. The increased kickback from a lower reference ladder resistance temporarily reduces the common-mode level on the reference input. This reduction has the effect of decreasing the threshold voltage variation of the comparator. Fig. 4.40 shows how decreasing the common-mode level of the reference affects the offset.
Figure 4.38: (a) Average threshold for all ADC codes for different values of $V_{\text{ref},\text{diff}} = V_{\text{ref},p} - V_{\text{ref},n}$. (b) Standard deviation of threshold for all ADC codes.

Figure 4.39: Decreasing the total reference ladder impedance by ten times (from 18 kOhm to 1.8 kOhm) reduces the average nonlinearity of the ADC but increases threshold variation. (a) average threshold, (b) average DNL, (c) sigma (30 trials).
Figure 4.40: (a) Comparator offset variation decreases with a larger ladder resistance due to kickback onto the reference input. (b) Directly decreasing the common mode of the reference voltage decreases offset variation.

4.6.2 Resonant Clocking Simulations

Simulations of the resonant clock path were performed with the distributed-element model shown in Fig. 4.15. There are two pseudo-complementary clock lines each with 128 elements, terminated in a differential inductor. The clock line is loaded by 256 dynamic comparators each with their own clock buffers. The measures of performance here are the maximum clock rate, the skew between clocks, and the duty cycle distortion.

In the proposed clocking scheme, one half of the comparators are driven by a complementary clock while the other half are driven by the true clock. As a result, half of the clock buffers include an extra inversion. In addition, the complementary clock input to the resonant line is simply generated by inverting the true clock. The pseudo-complementary nature of these clocks increases duty cycle distortion. Fig. 4.41 shows the waveforms on the clock line for several frequencies. Also shown are the corresponding waveforms after buffering inside the comparators.

Fig. 4.42(a) shows the duty cycle of the clock at the buffer output for various frequencies. While ideally the clock line operates at 5 GHz, with additional parasitics of only 2 fF at each clock node the optimal frequency drops to 4 GHz. The duty cycle of the complementary clock is shown in Fig. 4.42(b). The duty cycle distortion shown here could be corrected
Figure 4.41: (a) True and complementary transient waveforms on the resonant clock line with an additional 2 fF at each of the 256 comparator inputs. (b) Corresponding waveforms after buffering.
by using an inverter biased at its threshold voltage as the input to each comparator clock buffer, though at the cost of additional power. Alternatively, a truly complementary clock could be used at the input to the resonant clock line to ensure equal clock waveforms at the comparators.

4.7 Measurement Results

4.7.1 Test Setup

The primary challenge in high-speed ADC testing is to capture the ADC output. One solution to this problem is the use of an on-chip DAC that takes the digital output of the ADC and converts it back into an analog signal that can be easily output over one pin. The signal can then be examined using a spectrum analyzer in order to measure the parameters of interest such as SNDR. To make sure that the non-idealities of the DAC do not overshadow the performance of the ADC, the DAC should be at least two bits more linear than the device under test [113]. For low speed and/or low resolution data converters this requirement is not a problem. However at high data rates designing a DAC with that linearity is a significant challenge.
A second problem with cascaded ADC-DAC testing is the fact that the signal is shaped by the transfer function of the DAC. Even an ideal DAC has a frequency response that follows a $\sin(x)/x$ function and therefore changes the observed linearity of the ADC. This degradation can be compensated for, but not without increased uncertainty about the accuracy of the results. Typically, this type of testing is used in cases where the ADC in question is pushing the bounds of the technology and any other type of testing is impossible [108, 114].

Another solution to the data capture problem is to deserialize the ADC output bits on-chip and preserve them using some type of on-chip storage [85]. The captured data can later be transferred off the chip at a lower rate into a computer. One drawback of this approach is the large amount of storage needed. For a 5-bit ADC a 4096-point FFT is often used to construct the output spectrum. This measurement results in $5 \times 4096 = 20480$ bits. While a 20 Kb SRAM block would not take up significant area, it would be a nontrivial task to successfully add this block to a test chip layout.

A final frequently-used data capture method is to take the ADC outputs off-chip at high speed. In this scenario, the data is captured by high-speed test equipment such as a ParBERT [115] or a logic analyzer [67]. Sometimes the output bits cannot all be output at one time, and are instead multiplexed [69, 116]. The output bits are then realigned in software. This drawback with this method is in the high cost of test equipment. For companies with access to enough resources, an on-chip 10 Gigabit Ethernet PHY can transfer data at the required rate [19].

Our objective here is to achieve the benefits of the high-speed output method without incurring the high costs. To this end, we acquired the Altera Stratix IV Signal Integrity Evaluation Board shown in Fig. 4.43. This FPGA is equipped with high-speed transceivers that can be used to receive serial data from an ADC. The ADC test board is shown in Fig. 4.44.

A simplified diagram of the test setup is shown in Fig. 4.45. Aside from the FPGA board, only two signal sources and a computer are required. An expanded diagram of the test setup is shown in Fig. 4.46. The ADC test chip contains five 50 $\Omega$ output drivers that can simultaneously send 5 bits off chip. These bits are sent along a 10 cm microstrip line on board before going through an SMA connector and a short SMA cable to the FPGA board.
Figure 4.43: Altera Stratix IV Signal Integrity Evaluation Board.
Figure 4.44: ADC test board.
The bits are then received by the high-speed transceivers on the FPGA, deserialized to 20 parallel lines each, aligned to the same clock by a FIFO block, and finally stored in RAM on the FPGA. Once the data has been stored in RAM, it can be read off at a lower rate by a NIOS II CPU that has been synthesized in the FPGA. This CPU can perform some basic calculations on the data or simply transfer it to a PC via a USB interface. The CPU is also used to send configuration data to a 24-bit shift register on the test board and a 256-bit shift register on the test chip.

In a test setup where multiple bits are sent off-chip at high speed it is a challenge to maintain alignment of the bits so that they can be correctly interpreted. In fact, skew between bits is added at many stages between the ADC and the FPGA so that there is no chance that the bits will be properly aligned once they are captured in RAM. The data must be deskewed in software after being captured to restore the original signal sent by the ADC. There are two ways to deskew the data, based on the two modes of operation of the FPGA high-speed transceivers.

**Lock-to-Data Mode**

In lock-to-data mode, each receive channel in the high-speed transceiver quad independently recovers a sampling clock from the received data using a built-in CDR circuit. The CDR ensures that sampling will occur in the middle of the eye for each channel. As long as the input data contains enough transitions the CDR will remain locked and will always correctly sample the incoming data. In our measurements, an input frequency greater than 10 MHz was required to produce the required transition density. Once all five receive
Figure 4.46: Expanded diagram of the test setup. ALTGX is Altera's name for their high-speed transceiver IP block. NIOS II is a synthesizable CPU that can be programmed in C.
CDRs are locked, the relative skew between channels does not change. However, every time the system is powered up each CDR can have a different latency which means that the relative skew between channels is random (within a certain range, e.g. zero to twenty unit intervals). In this case, the five channels can be deskewed in the following manner:

- Capture 4096 samples in RAM.
- Consider only bit<4> (MSB) and bit<3>. Combine them to form a 2-bit number. Find the FFT of the sequence of 4096 2-bit numbers. Compute the SNDR.
- Shift bit<3> with respect to bit<4> and recompute the SNDR.
- Repeat this operation for all shifts of bit<3> between -20 and 20 bits.
- Find the shift that produces the highest SNDR and call that the skew between bit<4> and bit<3>.
- Repeat this operation, adding the next lowest bit each time.

**Lock-to-Reference Mode**

In lock-to-reference mode, the transceivers accept a low frequency reference clock which is then multiplied up to the bit rate by a clock multiplier unit (CMU). The comparator in each receiver is then triggered by this clock, even though this may not result in sampling in the middle of the received eye. This mode is intended to be used as a “stepping stone” to get the frequency of the voltage-controlled oscillator (VCO) in the CDR close enough to the data rate that the CDR can achieve lock within a reasonable time. However, it is also possible to operate the transceiver in lock-to-reference mode to receive data. The advantage of this mode is that the skew between channels is fixed and can be measured using the following process:

- Set the ADC reference input to zero so that all comparators have a nominal threshold voltage of zero.
- Turn on only one comparator for each output bit, turn off all other comparators. This configuration is shown in Fig. 4.47.
Figure 4.47: ADC configuration to facilitate deskew. Only one comparator per output has been activated, the rest are powered down with output low.

- Send a large-swing sinusoid as the input to the ADC.
- Capture these signals in RAM and observe the square-wave signal at each output of the ADC, noting the time delay between rising edges of the five signals.
- Shift the signals by an appropriate amount according to the observed skew.
- Set the ADC reference back to normal, turn on all 31 comparators, and proceed with measurements.

4.7.2 ADC Measurements

Comparator Offset Voltage

A measurement of comparator offset voltage is shown in Fig. 4.48, with the condition of the device under test (DUT) during the measurement shown on the left-hand side. In this case the differential ladder voltage is zero so the expected comparator threshold offset is
zero. From this measurement of 256 comparators, we can observe that there is a systematic offset of $-44.7\, mV$ in addition to the expected Gaussian spread of offset values.

This systematic offset is a result of unbalanced parasitic capacitance at the output of the comparator array. Pre-tapeout simulations were performed with lumped C extractions of the ADC sub-blocks only and did not show this imbalance. As mentioned in Sec. 4.6.1 post-tapeout simulations that included a Cc extraction of the complete ADC showed that certain coupling capacitances differing by 0.5 fF caused this imbalance. This was the case even though the total capacitance at the output nodes was roughly equal.

**Calibration**

In order to perform standard ADC measurements, calibration must first take place. Fig. 4.49 shows the comparator offset voltages organized by block (from 1 to 31). A simple calibration procedure would observe the offsets available in each block and choose the one that is closest to zero. Because of the systematic offset here, we instead choose the comparator in each block that is closest to the mean offset, $-44.7\, mV$. The chosen comparators are circled in Fig. 4.49.
Chapter 4 3GS/s, 17.6mW Flash ADC in 65nm CMOS

Figure 4.49: Measurement: offset of the 256 comparators in the ADC. The comparators chosen by a simple calibration scheme are highlighted.

Figure 4.50: Measurement: histograms of comparator offset before and after calibration. Test conditions are the same as those shown in Fig. 4.49.
Chapter 4 3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

Max
Avg

<table>
<thead>
<tr>
<th>Hysteresis (mV)</th>
<th>Comparator Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>96</td>
</tr>
<tr>
<td>5</td>
<td>128</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>9</td>
<td>96</td>
</tr>
<tr>
<td>10</td>
<td>128</td>
</tr>
</tbody>
</table>

$\text{f}_{\text{clk}} = 2.75\ \text{GHz}$

$\text{f}_{\text{clk}} = 3.25\ \text{GHz}$

$\text{f}_{\text{clk}} = 3.75\ \text{GHz}$

$\text{f}_{\text{clk}} = 4.25\ \text{GHz}$

Figure 4.51: Measurement: comparator hysteresis vs. $\text{f}_{\text{clk}}$ for $\text{v}_{\text{in,cm}}, \text{v}_{\text{ladder,p}}, \text{v}_{\text{ladder,n}}$ all equal to 0.7 V. This measurement is for the 128 comparators driven by the positive edge of the clock.

Clock Path Bandwidth

Next we determine the bandwidth of the clock path. We can do this by measuring the hysteresis of the comparators. Hysteresis exists when the threshold depends on the previous decision; it is equal to the difference between the two observed thresholds. Ideally a comparator would have zero hysteresis, but it can occur when there is insufficient time to precharge the comparator or insufficient time to make a decision. We cannot observe the on-chip clock waveform directly, but the presence or absence of hysteresis provides some information about it. Fig. 4.51 shows comparator hysteresis for half of the 256 comparators as clock frequency is changed.

ADC Output Spectra

The plot of hysteresis vs. clock frequency shows us that the centre frequency of the resonant clock path is approximately 2.5 GHz with acceptably low hysteresis between 2 and 3 GHz. Now we can use the results of the simple calibration to take an ADC measurement at the centre frequency of 2.5 GHz. The results of the 4096-point measurement are shown in
Chapter 4 3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

Regulator

\[ v_{\text{ladder}, p} = 0.825 \text{ V} \]
\[ v_{\text{ladder}, n} = 0.575 \text{ V} \]
\[ v_{\text{ref}} = 1.205 \text{ V} \]
\[ f_{\text{in}} = 3.0518 \text{ MHz} \]
\[ v_{\text{in}, \text{cm}} = 0.71 \text{ V} \]
\[ f_{\text{clk}} = 2.5 \text{ GHz} \]
\[ v_{\text{reg}, \text{in}} = 1.3 \text{ V} \]
\[ v_{\text{dd}} = 1.175 \text{ V} \]

**Figure 4.52:** Measurement: 4096-point ADC output spectrum with simple calibration. Computed SNDR is 21.6 dB.

Fig. 4.52. The computed SNDR is 21.6 dB, or 3.3 bits.

From the transient waveform, we can see that there are some problems with the codes at the extreme bottom and top. The reason for this is that we have chosen the comparators based on their thresholds when the reference voltage is zero. The effect of applying a reference voltage is shown in Fig. 4.53. Ideally, applying a reference voltage would shift each comparator’s threshold by an amount proportional to its place on the reference ladder. In Fig. 4.53, we can see that this relationship holds for small values of \( v_{\text{ladder}, \text{diff}} \), but with \( v_{\text{ladder}, \text{diff}} = 150 \text{ mV} \) the linear relationship breaks down for the comparators which experience the biggest reference voltage. However, the input full scale range of the ADC is approximately equal to the differential reference voltage, so we are forced to push the comparators into nonlinearity in order to covert a signal of appreciable amplitude.

To compensate for this effect, we recalibrate the ADC based on the threshold obtained with \( v_{\text{ladder}, \text{diff}} = 250 \text{ mV} \). A comparison of the offset voltages used in the original calibration and the offset voltages used after recalibration are shown in Fig. 4.54. The resulting measurement, shown in Fig. 4.55, gives an SNDR of 26.0 dB or 4.0 effective bits. In
addition to choosing slightly different comparators, a higher SNDR is obtained with an input signal that is smaller than the full scale range. This was not the case after the simple calibration, but it indicates that the thresholds at the extremes are not situated properly.

Setting the clock frequency to the upper end of the resonant clock path bandwidth, we take an ADC measurement with $f_{\text{clk}} = 3\,\text{GHz}$. The result is shown in Fig. 4.56. SNDR is 23.5 dB. SNDR measurements for a sweep of input power, input frequency, and clock frequency are shown in Fig. 4.57. While SNDR is maintained above 24 dB for sampling frequencies from 2–3 GHz, unfortunately input bandwidth is only 20 MHz. In order to find the cause of this problem, simulation is required.

A note about the common-mode level at the input to the ADC. A common-mode level of 0.71 V was used even though it would appear that a higher level would minimize the on-resistance of the PMOS input switches. Higher input levels were found to exacerbate the existing systematic offset of the comparators. A compromise was made between input switch resistance and added systematic offset.

ADC power dissipation is shown in Fig. 4.58. Static power is dissipated in the clock input match circuit and the voltage regulator. A photo of the test chip die is shown in Fig. 4.59.
Figure 4.54: Measurement: offset of the 256 comparators in the ADC with reference voltage applied. (a) after simple calibration and application of $v_{\text{ladder, diff}} = 250 \text{ mV}$, and (b) after recalibration.

Figure 4.55: Measurement: 4096-point ADC output spectrum after recalibration. Computed SNDR is 26.0 dB.
Figure 4.56: Measurement: 4096-point ADC output spectrum with $f_{\text{clk}} = 3$ GHz. Computed SNDR is 23.5 dB.

Figure 4.57: Measurement: (a) SNDR vs. input power, (b) SNDR vs. sampling frequency, (c) SNDR vs. input frequency. The measurements in (b) and (b) were all taken at a constant input power.
4.7.3 Voltage Regulator Measurements

DC voltage regulator measurements are shown in Fig. 4.60. Measurements of the voltage regulator output impedance are shown in Fig. 4.61.

4.7.4 Calibration DAC Measurements

Calibration DAC measurements are shown in Fig. 4.62. Linearity is worse than predicted by simulation. Power supply resistance may have reduced the voltage headroom available to the current mirror transistor. The test board provides inputs to the DAC from a 10-bit tunable resistor and a 10-bit DAC. Using these inputs to calibrate the on-chip 8-bit DAC reduced the INL from 30 LSBs to 3 LSBs.

Before using the on-chip DAC to calibrate the ADC, it was determined which combination of off-chip resistance (10 bits) and off-chip voltage (10 bits) would produce the desired DAC output for each combination of 8 DAC input bits. This set of 28 bits was stored in the FPGA-based CPU for each 8-bit DAC word. The CPU would then translate the 8 input bits into 28 bits which would be used to control the on-chip DAC as well as the on-board DAC and tunable resistor. While not elegant, this solution allowed automatic control of the calibration voltage provided to the ADC.
Figure 4.59: Die photo of the test chip with ADC and inductor shown. Die area is 1 mm$^2$. ADC and inductor combined area is 0.07 mm$^2 + 0.01 $mm$^2 = 0.08 $mm$^2$. 
Chapter 4  3 GS/s, 17.6 mW Flash ADC in 65nm CMOS

Figure 4.60: Measurement: DC voltage regulator output.

Figure 4.61: Measurement: DC voltage regulator output impedance.
Figure 4.62: Measurement: linearity of the on-chip calibration DAC: (a) before calibration, (b) after calibration with two 10b off-chip DACs.
4.8 Conclusions

The flash ADC proposed here used two novel techniques. First, comparator redundancy for total replacement of calibration circuitry. Second, resonant clocking for sampling bandwidth extension and clock power reduction.

Calibration through comparator redundancy allowed us to eliminate explicit digital calibration circuits and their associated loading and area. The absence of calibration circuits meant that we could use a second input differential pair to add offset with voltages generated from a resistor ladder. This method of threshold-setting is more predictable than methods that rely on unbalanced devices and also allows for an adjustable reference using the terminal nodes of the ladder. With no calibration circuits attached to the output of each comparator, output bandwidth of these devices is increased.

One drawback of redundancy is the decreased input bandwidth. This problem was evident in the measurement results shown here, although with some changes to the layout the situation could be ameliorated. A similar decrease in sampling bandwidth was compensated for by the adoption of resonant clocking, which allowed an increase in the clock rate and a reduction in the drive strength of clock buffers. The drawbacks of resonant clocking are the additional area taken by the inductor and the fact that low-speed testing is made more difficult or impossible. However, comparator selection at high speed is desirable because dynamic behaviour of the comparators is accounted for.
Chapter 5

Conclusion

This thesis has dealt with three aspects of the path from transmitter to receiver in a wireline communication channel: transmit-side equalization, bidirectional communication, and ADC-based receive-side equalization.

The transmit-side equalizer proposed here uses fractional tap spacings that are adjustable with an analog voltage. Measurement results show that less receive-side jitter is obtained when using a tap spacing of $\frac{1}{2}$UI. In addition, tap spacing tuneability provides an additional knob that could be used in equalizer adaptation. Future work could examine adaptation schemes using this feature, perhaps with independent control over individual spacings.

A technique for bidirectional communication is proposed which relies on the transmit-side equalizer discussed above. This transmitter is able to shape the spectrum of the transmitted signal so that two signals can be simultaneously received across the same channel. In this case the link has asymmetric data rates in the up- and down-stream directions which could be used to send data in one direction and equalizer adaptation information in the other direction. Although simulation results show simultaneous bidirectional operation, measurements were not able to show this due to impedance discontinuities in the test board. Future work could use the buried bump technique along with integration of the transmitter and receiver on a single die in order to eliminate this problem.

Some attempts have been made recently to design an ADC-based receiver for wireline communication. The main obstacle preventing acceptance of this technique is the high power consumed by the ADC. Therefore, Chapter 4 looks at an ADC design driven by the requirements of minimum power and 5-bit resolution. Non-idealities introduced by leaving out the track and hold circuit are analyzed. Explicit calibration circuits are ruled out in
favour of a redundancy-based calibration scheme. This scheme allows the ADC to achieve
the desired resolution without increasing power, although additional area is required and
the input capacitance is increased. A resonant clocking scheme is used to reduce clocking
power and increase the maximum clock frequency. While the target input bandwidth
was not achieved in the test circuit due to parasitic capacitance, it is expected that this
problem could be rectified with some additional effort. However, this design failure raises
the point that verification of a large high-speed circuit is in itself a major challenge. Future
efforts should focus on verification techniques that are time-efficient and tractable for the
simulator.
# Appendix A

## Summary of Contributions

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Contribution</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>• backplane transmitter with variable tap spacing</td>
<td>[33]</td>
</tr>
<tr>
<td>3</td>
<td>• bidirectional communication using pulse-shaping transmitter</td>
<td>[48]</td>
</tr>
<tr>
<td>4</td>
<td>• derivation of SNDR-offset relationship</td>
<td>[107]</td>
</tr>
<tr>
<td></td>
<td>• flash ADC (redundancy-based, resonant clocking)</td>
<td></td>
</tr>
</tbody>
</table>

*Table A.1: Summary of contributions.*
Appendix B

Remedial Simulations

To find the circuit problems that limit the input bandwidth to 20 MHz, we perform an extracted simulation of the entire ADC. This had not been possible with spectre (or hspice) due to memory constraints, but became possible post-tapeout using the newly-available Advanced Parallel Simulator (APS) from Cadence. The extracted RCc netlist is 1.6 million lines and occupies 90 MB.

In simulation, we can find the SNDR at various points in the circuit. Fig. B.1 shows the points of interest. Fig. B.2(a) shows the input bandwidth of the ADC, which is limited by the on-resistance of the PMOS switches and the parasitic resistance and capacitance. From the difference between the RCc and Cc simulations, we can see that parasitic resistance is a significant factor. Fig. B.2(b) shows the SNDR at two different points in the circuit. SNDR at the switch output is limited by the input bandwidth. However, SNDR at the output of the gray decoder is even more severely limited. To investigate this further, we look at the parasitics present in the thermometer-to-gray decoder.

Fig. B.3 shows the floorplan of the ADC with dimensions of all blocks. Large parasitic capacitance at the input to the Gray encoder is noted. The large capacitance causes long rise times which in some cases exceed $1\text{UI} = \frac{1}{3\text{GHz}} = 333\text{ps}$. This capacitance is the most likely cause of limited input bandwidth in the ADC, however it could be mitigated by inserting buffers between the 8:1 switch and the Gray encoder at the expense of some added latency.

A comparison of measured and simulated SNDR at the output is shown in Fig. B.4. The 5 dB difference between the curves is most likely due to device mismatch which was not included in this simulation.
Appendix B Remedial Simulations

Figure B.1: ADC nodes of interest.

Figure B.2: Extracted simulations with $f_{clk} = 3$ GHz: (a) input bandwidth, (b) SNDR vs. $f_{in}$ at two different points in the circuit.
Appendix B Remedial Simulations

Figure B.3: Floorplan of the ADC.
Appendix B Remedial Simulations

Figure B.4: Comparison of simulated (RCc extracted) and measured SNDR. The 5 dB discrepancy is due to device mismatch and noise which are not included in this simulation.
References


References


References


References


