Application of Active Inductors in High-Speed I/O Circuits

by

Yen-Sung Michael Lee

B.Eng., Carleton University, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Applied Science

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2008

© Yen-Sung Michael Lee, 2008
ABSTRACT

This thesis explores the use of active inductors as a compact alternative to the bulky passive spiral structures in high-speed I/O circuits. A newly proposed PMOS-based topology is introduced and used in active-inductor terminations. The 1st prototype design fabricated in a 90-nm CMOS process consists of an output driver using active-inductor terminations to provide channel equalization and output impedance matching. From measurement results, the use of active inductors in the termination, as compared to when the active inductor is disabled, increases the vertical eye opening in the receiver side by a factor of two and reduces the $jitter_{pp}$ by 30% of the transmitted 10 Gb/s ($2^{31}-1$) pseudo-random binary sequence pattern, over a 6-inch FR4 channel. An output impedance matching with $S_{22}$ less than -10 dB over a bandwidth of 20 GHz is achieved. The pair of active-inductor terminations occupies $17\times25\mu m^2$ and has a low overhead power consumption of 0.8 mW. In the 2nd prototype design, a 4-stage output buffer with active-inductor loads is designed and implemented in a 65-nm CMOS process. Simulation results verify that when operating at 31.25 Gb/s, the output eye of the active-inductor load buffer compares favorably with that of the passive-inductor load buffer. For a similar eye-height and 78% less timing jitter the active-inductor load design’s speed (31.25 Gb/s) is 25% faster than the passive-resistor load design (25 Gb/s). The active-inductor load output buffer achieves comparable performance in terms of speed, power, and output swing with other reported designs using passive inductors. Its total area is $135\times30 \mu m^2$ (including three differential active inductors) which is comparable to the size of a single passive spiral inductor having a 0.5~1 nH inductance.
# TABLE OF CONTENTS

Abstract.............................................................................................................................. ii
Table of Contents............................................................................................................... iii
List of Tables ...................................................................................................................... v
List of Figures................................................................................................................... vi
List of Abbreviations ......................................................................................................... viii
Acknowledgment............................................................................................................... ix
1 Introduction ..................................................................................................................... 1
   1.1 Motivation .............................................................................................................. 1
   1.2 Research Objectives ............................................................................................. 5
   1.3 Thesis Organization .............................................................................................. 5
2 Background .................................................................................................................... 6
   2.1 Inductive Peaking in High-Speed I/O .................................................................. 6
   2.2 Inductor Implementation ...................................................................................... 8
   2.3 Basics of Active-Inductor Shunt Peaking ............................................................ 9
   2.4 Design Challenges .............................................................................................. 12
      2.4.1 Voltage Headroom .................................................................................. 12
      2.4.2 Nonlinearity ............................................................................................ 14
      2.4.3 Noise ...................................................................................................... 15
   2.5 Eye Diagram ....................................................................................................... 16
3 Prototype Design I ........................................................................................................ 18
   3.1 Inductive Peaking Termination .......................................................................... 18
   3.2 PMOS-based Active Inductor ............................................................................ 19
      3.2.1 Circuit Topology ...................................................................................... 19
      3.2.2 Small-Signal Analysis ........................................................................... 21
      3.2.3 Large-Signal Operation ......................................................................... 21
      3.2.4 Linearity Enhancement ......................................................................... 22
   3.3 Output Driver with Active Inductor Load ........................................................... 23
      3.3.1 Full Schematic ......................................................................................... 23
      3.3.2 Biasing Voltage ....................................................................................... 24
      3.3.3 Testing Consideration ............................................................................. 25
   3.4 Experimental Results ......................................................................................... 27
      3.4.1 Driver Frequency Response .................................................................... 28
      3.4.2 Channel Frequency Response ................................................................ 29
      3.4.3 Eye-Diagram Measurement ...................................................................... 31
      3.4.4 Output Matching Measurement .............................................................. 34
4 Prototype Design II ...................................................................................................... 37
   4.1 Output Buffer ...................................................................................................... 37
      4.1.1 Introduction .............................................................................................. 37
      4.1.2 Design Consideration .............................................................................. 38
      4.1.3 Architecture ............................................................................................ 39
LIST OF TABLES

Table 4.1. Summary of the eye-diagram simulation....................................................... 57
Table 4.2. Performance summary of the output buffer designed with active-inductor ....... 59
LIST OF FIGURES

Fig. 1.1. Trend in microprocessor I/O bandwidth................................................................. 1
Fig. 1.2. Frequency responses of backplane channels............................................................ 3
Fig. 2.1. Shunt peaking in a CS amplifier........................................................................... 6
Fig. 2.2. Effect of inductive peaking on magnitude response............................................... 7
Fig. 2.3. Inductive (RL) terminations used for equalization................................................. 8
Fig. 2.4. Conventional active inductor: (a) schematic and (b) small-signal mode................. 9
Fig. 2.5. Active-inductor load: (a) more detailed model and (b) its passive RLC equivalent. ...............................................................................................................................................11
Fig. 2.6 (a). Conventional active inductor with $V_{HIGH}$ (b) folded active inductor .............. 12
Fig. 2.7. Construction of an eye diagram (a) a data stream (b) overlapping each bit-time segment of the data stream to form a data eye. ......................................................................................... 16
Fig. 2.8. Illustration of trailing edges and the effect of ISI.................................................. 17
Fig. 3.1. Doubly terminated I/O blocks with inductive peaking terminations..................... 18
Fig. 3.2. Schematic of the proposed active-inductor termination circuit............................. 19
Fig. 3.3. Simulated $R_s$ and the corresponding $S_{22}$ vs. $I_D$. ........................................... 22
Fig. 3.4. Full schematic of the CML output driver with PMOS-based active-inductor loads. .............................................................................................................................................. 23
Fig. 3.5. On-chip reference voltage circuits ......................................................................... 24
Fig. 3.6. Ideal design with both transmit-side and receive-side terminations...................... 26
Fig. 3.7. Actual fabricated chip with only transmit-side termination................................... 26
Fig. 3.8. Die Micrograph.................................................................................................... 27
Fig. 3.9. Simulated frequency response of the output driver with active inductance turned off and turned on while sweeping $V_{G2}$ ........................................................................................................... 28
Fig. 3.10. Test setup for the measurement of channel frequency response......................... 29
Fig. 3.11. Channels used for measuring the eye-diagrams at the receiver side..................... 30
Fig. 3.12. Measured responses of a 6-inch FR4 trace and a 4-m RG-58/U cable.................. 30
Fig. 3.13. Test setup for eye-diagram measurements ........................................................... 31
Fig. 3.14. Received eye-diagrams for a 5 Gb/s $2^{31}-1$ PRBS pattern through a 4-m RG-58/U cable (a) active inductance off (b) active inductance on ...................................................... 32
Fig. 3.15. Received eye-diagrams for a 10 Gb/s $2^{31}-1$ PRBS pattern through a 6-inch FR4 trace (a) active inductance off (b) active inductance on.......................................................... 33
Fig. 3.16. Test setup for $S_{22}$ measurements ..................................................................... 34
Fig. 3.17. Measured $S_{22}$ of the output driver with different DC current in the active-inductor load ............................................................................................................................................... 36
Fig. 4.1. Output buffer architecture...................................................................................... 40
Fig. 4.2. Topology of the NMOS-based active-inductor load.............................................. 40
Fig. 4.3. $g_{m1}$ vs. $V_{GS1}$ ($W_1/L_1$)=(12um/60nm)............................................................... 41
Fig. 4.4. Output buffer with transistor sizes, tail currents, $C_{PAD}$ and $R_s$ and $L$ values of active-inductor loads labeled........................................................................................................... 45
Fig. 4.5. Three-stage pre-driver with active-inductor loads and transistor sizes, tail currents and $R_{deg}$ values labeled........................................................................................................... 46
Fig. 4.6. Layout of the output buffer with active-inductor loads........................................46
Fig. 4.7. Input-output characteristic (pre-driver with active-inductor loads).........................48
Fig. 4.8. $I_D$ vs. $V_{in\_diff}$ (Pre-driver with active-inductor loads).................................................................49
Fig. 4.9. Differential $G_m$ vs. $V_{in\_diff}$ (Pre-driver with active-inductor loads)......................50
Fig. 4.10. Frequency responses of the entire pre-drivers and 1st-stages designed using different loads...................................................................................................................51
Fig. 4.11. Step response (rising) of the 1st-stage pre-driver......................................................52
Fig. 4.12. Step response (falling) of the 1st-stage pre-driver.....................................................53
Fig. 4.13. Step response (rising) of the entire pre-driver..........................................................54
Fig. 4.14. Step response (falling) of the entire pre-driver..........................................................54
Fig. 4.15. Eye-diagram of the output buffer using active-inductor loads at 31.25Gb/s.............55
Fig. 4.16. Eye-diagram of the output buffer using passive-inductor loads at 31.25Gb/s...........55
Fig. 4.17. Eye-diagram of the output buffer using passive-resistor loads at 31.25Gb/s...........56
Fig. 4.18. Eye-diagram of the output buffer using passive-resistor loads at 25Gb/s..............58
LIST OF ABBREVIATIONS

BER: Bit-Error Rate
CML: Current-Mode Logic
CS: Common-Source
DEMUX: Demultiplexer
DSM: Deep Submicron
FO: Fanout
I/O: Input/Output
ISI: Inter-Symbol Interference
MUX: Multiplexer
PRBS: Pseudo-Random Binary Sequence
PVT: Process, Voltage, and Temperature
UI: Unit Interval
ACKNOWLEDGMENT

I would like to express my deepest gratitude to my advisor, Prof. Shahriar Mirabbasi, without whom I would not have the opportunity to come to UBC and this work* would not have been made possible. I thank him for his support, guidance, and immense amount of patience and understanding.

I am also hugely in debt to my senior lab colleague, Samad Sheikhaei, who has played an invaluable role in my work. Samad often spends a lot more time than he has available, to work together with and help me through the many obstacles that I have encountered in this work. Without his inputs, it would have been much more difficult, if possible at all, to obtain my chip measurement and submitting a conference paper that we worked on together before the deadline, and to submit this thesis to the committee in time.

I would like to thank our test lab manager, Dr. Roberto Rosales, for teaching me how to use the testing equipments and being always so generous to offer helps whenever I need them.

I want to thank Prof. David Michelson and Prof. Alireza Nojeh for accepting to serve on my defense committee and providing me with their valuable comments.

I was lucky to overlap my Master program with Daryl Van Vorst, who is also in Shahriar’s group. Daryl has broad and in-depth knowledge in RF designs and electrical engineering

---

This work is supported in part by Intel Corporation and the Natural Sciences and Engineering Research Council of Canada (NSERC). CAD tool support is provided by Canadian Microelectronics Corporation (CMC Microsystems).
in general. I usually go to him when I need someone with industrial background to discuss with. I want to thank him for reading my thesis and the many useful comments he’s given to me on the thesis. He was also the one who helped me build the FR4 board for channel measurements.

I would also like to thank Li Chen (Nanoelectronics Lab) for explaining to me the device questions that I have come across to in this work and Ben He (Communications Lab) for helping me with the coding of the Matlab scripts that were used in the eye-diagram simulations.

I also appreciate the loyal friendships and supports that my bio-MEMS roommate, Michael Chen, and race-car engineering friend, Gary Chou (who both understand the pronunciation of sesame in the Min-Nan dialect of Taiwan very well) have been giving to me. The people who made my time in the SoC lab memorable are Lyon Lee, Danny Yoo, Arash Zargaran, Cindy Mark, Andrew Lam, Paul Teehan, Rod Froist, and Cathy Yuan.

Beyond any words in my command, I am extremely grateful for having had the loves and supports of Pinky Li and Lydia Lee during different periods of time in the past two years of my graduate study.

Lastly, I want to dedicate this thesis to my most beloved grandma and parents for their unconditional love and sacrifices.
To

Ah Po

(grandma in the Hakka dialect of Taiwan).
1 INTRODUCTION

1.1 Motivation

Device scaling and architecture design advances have fueled the rapid growth in on-chip processing power for the past decades. This continuing growth in computing capacity demands that the off-chip bandwidth also scale in step. Following the exponential trend of growth in the demand for higher microprocessor I/O bandwidth as depicted in Fig. 1 [1], computing platforms that use multiple cores ([2], [3], and [4]) and high-performance graphics will require an off-chip bandwidth as high as several terabits-per-second in the near future to fully realize their computing power.

![Fig. 1.1. Trend in microprocessor I/O bandwidth.](image)

Another driver for higher I/O bandwidth comes from network communication systems such as Synchronous Optical NETwork (SONET) and Ethernet which are constantly evolving in
response to the ever increasing demand for higher network service capacity. Take the Internet for example, apart from the drastic increase in the number of internet users in developing nations such as China and India (India’s internet users grew by 33% between the years 2006 and 2007 [5]), emerging online entertainments like real-time video streaming and multiplayer games have been further increasing the global network traffic. Last year, the popular video site YouTube launched only three years ago, alone consumed as much bandwidth as the entire Internet did in 2000 [5]. This rapidly increasing usage in internet network services as well as the need for higher data transfer rates are afforded by high aggregate-bandwidth communication technologies such as 10 Gb/s SONET for backbone or metro area networks (that have been developed), and other Gb/s technologies including 10Gb/s Ethernet, the Infiniband system, and 40 Gb/s SONET (which are undergoing development) [6]. The aggregate bandwidth in the key components (e.g., routers, framers, and switches) in these systems will soon reach 800 Gb/s and beyond [7].

To deliver the large off-chip bandwidth required in high-performance computer and network communication systems, many high data-rate/pin I/Os (e.g., 10 Gb/s/pin and 40 Gb/s/pin) have to be integrated on the same chip to curb the overall cost. There are a lot of challenges in realizing power-efficient I/O links operating at such high data rates. One of the main challenges is to maintain good signal integrity while transmitting data through physically impaired and band-limited channels. Every physical electrical channel exhibits different degrees of non-ideal characteristics such as DC loss, skin effect, dielectric loss, and impedance discontinuity. In the example of typical backplane channel responses shown in Fig. 1.2 [8], the signal loss could be as large as 30 dB at 3 GHz.
Another challenge in pushing the speed envelope of I/O links is to increase the maximum data rate at which on-chip I/O circuits can operate for a given process technology, and at the same time keep the overall I/O power consumption low. Some of the most speed-critical building blocks in an I/O transceiver are multiplexer, output buffer, decision circuit, voltage-controlled oscillator, and clock buffer since they usually have to drive large capacitive loads while operating at the highest clock rate.

A very power-efficient way to boost the bandwidth of I/O circuits is to employ passive filtering (e.g., shunt and series peaking) [9][10], which uses inductors to trade off bandwidth versus peaking in the magnitude response. It is also of great advantage, in the view of minimizing power consumption, to use on-chip passive filters to provide equalization [11][12]. The main problem with the deployment of passive filters is the large area needed to implement the on-chip passive inductors (also known as spiral inductors). The amount of area (on the low side) needed to realize a pair of differential spiral inductors with an inductance around 1-1.5 nH is on the order of 100 μm by 100 μm [9]. The area
occupied by spiral inductors usually determines the overall area of an I/O building block. Unlike active devices, spiral inductors do not scale with technologies. As more and more I/Os are needed to be integrated on the same chip to achieve a large aggregate off-chip bandwidth, the area consumed by extensive use of spiral inductors in I/O building blocks will be prohibitively large.

A much more compact alternative to realize on-chip inductors for bandwidth enhancement as well as equalization is to use active inductors. Both the area and speed of active inductors scale with technology and they are easy to implement in standard digital CMOS processes since they can be designed using only active devices. These favorable features make their usages compatible with the trend of very high level of integration, particularly in systems that are enabled by CMOS technology (e.g., System-on-Chip).

The state-of-the-art microprocessors [13] [14] and communication SoC chips [15] [16] continue to migrate to deeper submicron CMOS process technologies (such as 65-nm and 45-nm nodes) due to the performance and economic incentives brought by scaling. To evaluate the suitability of using active inductors in the I/Os of current and future high-performance digital chips, in this thesis we are motivated to design and implement prototype active inductor circuits in the most advanced CMOS technologies that we have access to.
1.2 Research Objectives

The research objectives of this work are as follows:

- Investigate the potentials and challenges of using active inductors to provide area-efficient and low-power bandwidth enhancement in I/O circuits and compensation for channel losses.

- Develop techniques and design methods to enhance the performance of active inductors when applied to I/O circuits.

- Design and implement prototype circuits in advanced deep submicron CMOS technologies and verify the performance improvement, if at all, brought by the use of active inductor.

1.3 Thesis Organization

The remaining of the thesis is organized as follows. Chapter 2 provides the background on the basics of inductive peaking, inductor implementations, principle and analysis of the operation of active-inductor shunt peaking, and the design challenges associated with active inductors. Chapter 3 presents the design and silicon validation of an output driver with active-inductor termination to provide output impedance matching and channel loss compensation at 5 and 10 Gb/s data rates. Chapter 4 presents the design of a 30 Gb/s output buffer with active-inductor loads for bandwidth enhancement. The bandwidth enhancement effect of active-inductor shunt peaking is compared with passive-inductor shunt peaking and no shunt peaking. Chapter 5 concludes the thesis and discusses future work.
2 BACKGROUND

2.1 Inductive Peaking in High-Speed I/O

As the operating speed of I/O circuitry reaches well into the multi-Gb/s range, inductors are used more and more extensively in high-speed I/O building blocks such as multiplexers/demultiplexers (MUX/DEMUX), buffers, and output drivers. For example, around 400 inductors are used in the 40 Gb/s/pin serial-link transmitter presented in [17]. In such applications, inductors are primarily used to provide bandwidth enhancement in the heavily capacitive-loaded nodes by means of shunt or serial peaking.

Fig. 2.1 illustrates the basic idea of how shunt peaking works in a simple common-source (CS) amplifier. By inserting an inductor $L$ in series with the resistor load $R_D$ to resonate with the load capacitor $C_L$, a zero is created in the frequency response of the amplifier.
which moves the band-limiting pole ($\omega_p = 1/(R_D C_L)$) to a higher frequency. In the time domain this bandwidth extension can be intuitively explained by applying an input step pulse to the amplifier: The inductor impedes the instantaneous change in the output current $I_D$ and acts as an open circuit, allowing all the current to flow through the capacitor rather than through the resistor. As a result, the output voltage changes faster and thereby enables the amplifier to operate at a higher speed.

![Fig. 2.2. Effect of inductive peaking on magnitude response.](image)

Inductive peaking is also useful in implementing passive equalization to allow higher off-chip signaling rates while keeping the power consumption at reasonable levels [1], [11], and [12]. Despite the word peaking, when used for on-chip bandwidth enhancement the magnitude of peaking is to be kept to a minimum (or ideally no overshoot) to allow a well-behaved response to random data (minimum settling time). As shown in Fig. 2.2, when $L$ is equal to $L_{opt}$, a bandwidth extension of up to 70% can be achieved without peaking. $L_{opt}$ is the value needed to result in a maximally flat response in a 2nd-order RLC network.
But when used to assist channel equalization, \( L \) should be made larger than \( L_{\text{opt}} \) to create the peaking that is needed to compensate the high-frequency signal loss in the channel. In [1] it was shown that by choosing proper values of inductance and resistance in the inductive terminations of the transmitter and receiver, up to 6 dB of gain in the overall channel magnitude response can be achieved.

![Inductive terminations used for equalization.](image)

Fig. 2.3. Inductive (RL) terminations used for equalization.

### 2.2 Inductor Implementation

On-chip inductors are typically realized by passive spiral inductors. Advantages of using passive spiral inductors as opposed to active inductors include lower power consumption, lower voltage-headroom requirement, lower noise, and better linearity. However, the implementation of spiral inductors typically occupies a large fraction of the silicon area and they do not scale with process technology. In addition, the ever increasing demand for higher aggregate bandwidth and the need to support different communication standards in SoC design require more I/O blocks to be integrated on the same chip. In such circuits, the overhead space needed to prevent mutual coupling (cross-talk) between inductors of different I/O channel circuits on the same chip would make the implementation of the spiral inductors even more costly in terms of area. A more area-efficient alternative to realize on-chip inductors is to use active inductors [19] [20]. Active inductors are particularly
suitable for implementing shunt peaking in current-mode logic (CML) circuits due to both the low-Q requirement and low-swing nature of such circuits. One of the most attractive features of active inductors is that, unlike their passive counterparts, both their area and resonant frequency scale with technology. Also, they can be implemented in a standard digital process, since they can be designed using only active devices. In addition, the tunable nature of active inductors allows for gain tuning over the frequency range of interest. This property can be taken advantage of in channel loss compensation and equalization. Furthermore, active inductors can be tuned to compensate for the effects of process, voltage, and temperature (PVT) variations in the circuit.

2.3 Basics of Active-Inductor Shunt Peaking

Fig. 2.4. Conventional active inductor: (a) schematic and (b) small-signal mode.
Although there exist different topologies that a shunt-peaking active inductor can assume, their characteristics and principles of operation are very similar. To illustrate how such circuits work, we start by looking at the conventional structure of an active inductor used in the load of a CS amplifier as shown in Fig. 2.4 (a) [21]. The active inductor shown inside the dashed box consists of an NMOS transistor $M_1$ and a resistor $R_G$ connected between $V_{DD}$ and the gate of $M_1$. The impedance looking into the active-inductor load is denoted by $Z_L$. Fig. 2.4(b) shows the simplified small-signal model of the active-inductor load. If a test signal $V_x$ is applied to the source of $M_1$, the voltage across $C_{gs1}$ is given by

\[
V_{gs1} = V_x \frac{sC_{gs1}}{R_G + \frac{1}{sC_{gs1}}} \]

\[
= V_x \frac{1}{sR_G C_{gs1} + 1}. \tag{2.1}
\]

The corresponding current generated that flows into the drain of $M_1$ will be

\[
I_{d1} = g_{m1} V_{gs1} \]

\[
= g_{m1} V_x \frac{1}{sR_G C_{gs1} + 1}. \tag{2.2}
\]

The impedance $Z_2$ looking into the drain, from the source, of $M_1$ can then be obtained by

\[
Z_2 = \frac{V_x}{I_{d1}} \]

\[
= \frac{sR_G C_{gs1} + 1}{g_{m1}} \]

\[
= \frac{sR_G C_{gs1} + 1}{g_{m1}}. \tag{2.3}
\]

The first term of $Z_2$, i.e., $(sR_G C_{gs1})/g_{m1}$, is linearly proportional with frequency and is what gives rise to the inductive property of the active-inductor load. Now with a more precise
small-signal model of the active-inductor load shown in Fig. 2.5 (a), which includes the drain-to-source conductance $g_{ds1}$ and drain-to-source capacitance $C_{ds1}$ of $M_1$, the impedance $Z_L$ is given by

$$Z_L = \frac{1 + sC_{gs1}R_G}{s^2R_GC_{ds1}C_{gs1} + s(C_{gs1} + C_{ds1} + R_GC_{gs1}g_{ds1}) + (g_{ds1} + g_{m1})}. \quad (2.8)$$

Fig. 2.5. Active-inductor load: (a) more detailed model and (b) its passive RLC equivalent.

For frequencies well below the resonance, the impedance looking into the active-inductor load can be approximated by $R_s$, $L$, and $C_p$ of an RLC network as shown in Fig. 2.5 (b).

$$R_s = \frac{1}{g_{ds1} + g_{m1}} \quad (2.9)$$

$$L = \frac{R_GC_{gs1}}{g_{ds1} + g_{m1}} \quad (2.10)$$

$$C_p = C_{ds1} \quad (2.11)$$

$$\omega_0 = \sqrt{\frac{g_{ds1} + g_{m1}}{R_GC_{gs1}C_{ds1}}} \quad (2.12)$$
Note that from (2.9) the value of series resistance $R_s$ of the active inductor depends on both $g_{m1}$ and $g_{ds1}$ which are associated with the drain current and DC bias point of $M_1$ transistor. Also it can be seen from (2.10) the inductance of the active inductor $L$ can be tuned independently by changing the value of $R_G$ while keeping other parameters constant. The resonant frequency ($\omega_0$) of the active inductor is approximately proportional to the square root of the $f_T$ of transistor $M_1$, which scales with technology. $L$ is proportional to $R_G$ as shown in (2.10), which implies that for a larger inductance the operating speed of the active inductor would be lower (since $\omega_0 \propto \sqrt{1/R_G}$). Such characteristic is also common to spiral inductors.

### 2.4 Design Challenges

#### 2.4.1 Voltage Headroom

![Fig. 2.6 (a). Conventional active inductor with $V_{HIGH}$ (b) folded active inductor.](image)
The conventional active inductor topology shown in Fig. 2.6 (a) requires a large DC voltage drop to ensure that $V_{GS1}$ is higher than the threshold voltage of $M_1$ transistor ($V_{th1}$) so that $M_1$ does not turn off when the output of the amplifiers experiences a large signal swing. This voltage headroom requirement is made worse by the body effect of the NMOS transistor $M_1$. This is because the threshold voltage $V_{th}$ of a MOS transistor is an increasing function of its source-to-body voltage $V_{SB}$. Since $V_B$ of all NMOS transistors have to be at the lowest potential of the entire chip (ground in this case) for CMOS processes that use p-type substrates (as are the 90-nm and 65-nm processes used in this thesis). The NMOS transistors that do not have their source nodes $V_S$ connected to the same potential as $V_B$ would have higher $V_{th}$ and thus require higher voltage headroom (i.e., $V_{GS}$ and $V_{DS}$) to keep them properly biased. As can be seen from (2.12), the resonant frequency $\omega_0$ of the active inductor improves with transconductance ($g_{m1}$) given the same device dimensions in $M_1$. To find the expression for $g_{m1}$, we consider the I-V characteristic of a short channel MOSFET,

$$I_D = W_{sat} C_{ox} \left( \frac{(V_{GS} - V_{th})^2}{(V_{GS} - V_{th}) + E_{crit} L} \right) (1 + \lambda V_{DS}),$$  \hspace{1cm} (2.13)$$

where $v_{sat}$ is the saturation velocity ($v_{sat} \approx 10^7$ cm/s) and $E_{crit}$ is the critical field ($E_{crit} \approx 6 \times 10^4$ V/cm for electrons). $g_m$ can be found by taking the derivative of $I_D$ with respect to $V_{GS}$:

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = W_{sat} C_{ox} \frac{(V_{GS} - V_{th})^2 + (V_{GS} - V_{th})E_{crit} L}{(V_{GS} - V_{th})^2 + (V_{GS} - V_{th})E_{crit} L + (E_{crit} L)^2}.$$  \hspace{1cm} (2.14)$$

Since $E_{crit} L$ is a constant term, it can be shown that from (2.14), $g_m$ increases with $V_{GS}$. Thus it is of advantage to have a larger $V_{GS1}$ as it allows a higher resonant frequency in the active inductor for a given $M_1$ size. Moreover, as will be discussed in the next section, having a larger $V_{GS1}$ also improves the linearity of the active-inductor load.
Most current-mode logic circuits employ tail current transistors and mandate a sufficient amount of drain-source voltage ($V_{DS}$) to ensure their core transistors (i.e., differential-pair driver) remain in saturation at all times. The limited voltage headroom that is left available for a conventional active-inductor load would largely compromise its usefulness, if working at all. The voltage headroom problem is more pronounced if the circuit is to operate from a low supply voltage, i.e., 1.2 V or 1 V, which is common in advanced deep submicron (DSM) CMOS technologies such as the 90-nm and 65-nm nodes. Two solutions have been previously proposed to mitigate such headroom constraints in conventional active inductors. The first solution [19], as shown in Fig. 2.6 (a), uses a voltage-boosting technique to provide a gate bias voltage ($V_{HIGH}$) that is higher than the supply voltage ($V_{DD}$). The cost of this method is increased design complexity and area associated with the addition of voltage-boosting circuitry. Another solution shown in Fig. 2.6 (b) [20] adopts a folded topology which allows the gate-source voltage of the NMOS transistor to be biased at much higher than its $V_{th}$ level without additional circuitry and at the same time eliminates the body effect. However, this approach consumes extra power and has a narrower bandwidth since its load is composed of the output resistance of $M_2$ transistor in parallel with the folded active inductor.

### 2.4.2 Nonlinearity

As compared to a CML circuit that uses passive-inductor loads, the circuit with active-inductor loads does not have a very well-defined DC output level, owing to the nonlinear I-V characteristics of active devices. If the DC output level variations in circuits that use active-inductor loads are too large, two adverse effects might occur. One is the reduced output voltage swing if the DC output level is higher than the original design. Another is
when the DC output level (which is also the input bias voltage to the next stage) drops lower than what is needed for the following stage circuits to remain in saturation, the gain and swing of those stages could be severely degraded.

The active inductors used in I/O circuits are intended to operate with a relatively large signal swing (i.e., up to a few hundred micro-volts). In CML-based I/O circuits, the load impedance and the resonant frequency of the active inductor are bias dependent as shown in (2.9), (2.10), and (2.12) and could vary quite significantly as the output level changes with time. Such nonlinear effects could result in different rise and fall times and introduce more jitter in the output waveform. Hence the effect of the distortion caused by the impedance variations must be investigated and kept to an acceptable level for active-inductor loads to be a viable alternative to passive-resistor or passive-inductor loads.

2.4.3 Noise

Active inductor loads have higher noise levels as compared to their passive counterparts. The two main noise contributions come from the thermal noise associated with the channel resistance and the flicker noise due to charge trapping in transistor $M_1$. The thermal noise also includes a high-frequency component, which is caused by the gate current and is relevant in I/O circuits operating at multi-Gb/s. Since MOS transistors conduct current near the surface of silicon where surface acts as traps that capture and release current carriers, their flicker noise components can be large. These traps capture and release carriers in random fashion and the trapping times are distributed in a way that lead to a $1/f$ spectrum. Thus, the flicker noise power is mainly concentrated on the lower frequencies. However, the magnitude of the noise introduced by active inductors is typically negligible as
compared to the voltage swing level in most I/O circuits, except for the front end of highly sensitive receivers [22].

2.5 Eye Diagram

![Eye Diagram](image)

Fig. 2.7. Construction of an eye diagram (a) a data stream (b) overlapping each bit-time segment of the data stream to form a data eye.

The eye diagram is an intuitive graphical representation of (electrical and optical) data communication signals. A data eye can be obtained by overlapping many bit-long signal traces from a random data stream onto a single bit-time interval (e.g., for 10 Gb/s, bit-time = 100ps) as shown in Fig. 2.7 [23]. The vertical eye opening (or eye height) and horizontal eye opening (or eye width) are important characteristics of the eye diagram that aid in quantifying the signal quality. The vertical eye opening is measured at the sampling instant and is strongly related to the amount of inter-symbol interference (ISI) presented in the signal. The horizontal eye opening is usually expressed as the percentage of one unit bit width, also known as the unit interval (UI), and is largely influenced by ISI as well. In a band-limited system (channel or electronics or both), each transition in the input logic level would produce an exponential output response; the narrower bandwidth of the system, the longer the exponential tails. The longer the output response tails, the greater the effect it has.
on corrupting the amplitude of the subsequent bits. This increases the probability of
detecting the wrong logic level in the receiver. Such effect is known as ISI and is illustrated
in Fig. 2.8 [8]. The two rectangular step pulses represent the input bits (101) to a
band-limited channel and the three curves represent the responses (100, 001, and 101)
appeared at the output of the channel after a finite delay. Note that one of the output
responses (101) is the superposition of the other two responses (100 and 001). As can be
seen from Fig. 2.8, when the input pattern 101 is being applied to the band-limited channel,
the ISI introduced by the long trailing edges between the transitions of the 1st to 2nd output
bits (in 100) and the 2nd to 3rd output bits (in 001) can cause the 2nd output bit (in 101) to
take a wrong value (a logic 1 instead of 0).

![Fig. 2.8. Illustration of trailing edges and the effect of ISI.](image)

By inspecting the amount of ISI presented in the eye diagram at different operating speed, a
quick estimate on the bandwidth or maximum achievable data rate for a given bit-error rate
(BER) can usually be obtained.
3 Prototype Design I

This chapter presents the prototype design of a novel PMOS-based active inductor (first proposed by the author of this thesis [22]) which provides a compact alternative for implementing inductive termination and offers channel loss compensation through tunable peaking. The prototype chip was fabricated by STMicroelectronics (STM) in a 90-nm CMOS technology and measurement results are presented. Note that the 90-nm technology was the most advanced CMOS process node that was available at the time to the author.

3.1 Inductive Peaking Termination

Fig. 3.1. Doubly terminated I/O blocks with inductive peaking terminations.

Fig. 3.1 shows a doubly terminated I/O circuit using two inductive terminations with impedance $Z_T$. In this figure, $C_{PAD}$ represents the lumped parasitic capacitance seen at the I/O terminals that is typically dominated by the parasitic capacitance of the electrostatic discharge (ESD) protection circuit and the bonding pad. $Z_0$ is the characteristic impedance
of the channel. The termination impedance $Z_T$ should ideally exhibit a real part ($R_s$) that is equal to $Z_0$ and constant over the bandwidth of interest. $Z_T$ have an imaginary part ($j\omega L$) which can be used to create a peak at the target transmission frequency so as to partially compensate the high-frequency signal loss in the channel. If the data rate of the transmitted pulse is higher than the frequency of the pole caused by the equivalent RC low-pass filter seen at the termination node, the inductor value can also be chosen to resonate with the parasitic capacitance seen at the node to improve the I/O bandwidth.

### 3.2 PMOS-based Active Inductor

As discussed in the previous chapter, it is of interest to realize inductive peaking circuits using active inductors to save area and provide tuning capability. A PMOS-based active inductor that is capable of implementing the inductive termination as shown in Fig. 3.1 is presented in the following sub-sections.

#### 3.2.1 Circuit Topology

![Schematic of the proposed active-inductor termination circuit.](image-url)

Fig. 3.2. Schematic of the proposed active-inductor termination circuit.
The proposed active inductor topology is shown inside the dashed box in Fig. 3.2. It has an active resistor \( (M_2, \text{which operates in deep-triode region}) \) through which the output node is coupled to the gate of the PMOS transistor \( M_1 \) via a source follower \( M_3 \). A level shifter, consisting of a source follower \( M_3 \) and a current source \( M_4 \), is inserted between \( M_1 \) and \( M_2 \) to allow a lower gate bias voltage for \( M_1 \). This topology allows the active inductor to operate with low voltage headroom without the need to use a separate voltage higher than \( V_{DD} \) for biasing the gate of \( M_1 \) as it does in NMOS-based active inductor [19]. By tuning \( V_{G2} \) to change the resistance of \( M_2 \) (i.e., \( R_G \)), the inductance \( L \) of the active inductor can be independently controlled (2.10). The proposed active inductor also consumes less power than [20] while exhibiting a low resistance (i.e., 50 \( \Omega \)) for termination. Note that [20] needs to sink an additional, relatively large amount of bias current to its folded active inductor for it to exhibit a \( R_s \) of 50 \( \Omega \) (\( R_s \propto 1/g_m \)) on top of the tail current that is already needed in the core driver transistors. The proposed topology reuses the tail current for biasing and consumes no extra current. The topology is based on a PMOS transistor \( (M_t) \) which does not suffer from body effect and hence has a lower and relatively constant \( V_{th1} \) (since \( V_{S1} \) is connected to \( V_{DD} \) so \( V_{SB1} \) shall always remain constant) as compared to when a NMOS transistor is used. Having a lower \( V_{th} \) level that does not increase with the output level almost always serves to the advantage of the circuit operation in analog design. This is especially true when the circuit is implemented in DSM CMOS technologies where supply voltages are expected to continue to drop. In the case of the proposed active inductor the lower \( V_{th} \) level in \( M_1 \) serves to help ease the voltage headroom requirements of the active inductor (i.e., \( V_{GS1} \) and \( V_{DS1} \)) and can also be translated to a higher output swing tolerance given the same voltage headroom since \( V_{D,sat} \) is made higher (\( V_{D,sat}=V_{GS1}-V_{th1} \)). It is worth noting that although \( g_{m3}, g_{ds3}, \) and \( g_{mb3} \) of the source follower transistor \( M_3 \) vary with the output level
due to body effect, such variations have a limited impact on the gain of the level shifter which is given by $1/(1+(g_{ds4}+g_{ds3}+g_{mb3})/g_{m3})$. The drawbacks of the proposed topology are the additional overhead power consumed by the level shifter and the fact that PMOS transistors are slower than NMOS transistors due to the lower carrier mobility. Also, the non-idealities of the level shifter, such as the parasitic capacitances and less-than-unity gain, could adversely affect the speed of the proposed topology.

### 3.2.2 Small-Signal Analysis

Assuming the level shifter has an unity gain over the entire bandwidth of interest, the small-signal analysis of the proposed active inductor would be the same as that derived for the conventional active inductor in Section 2.2. The termination impedance of the active inductor is thus (2.8) and the expressions for its passive RLC equivalent model are (2.9), (2.10), (2.11), and (2.12).

### 3.2.3 Large-Signal Operation

The presented active inductor is intended to operate under a relatively large signal swing of the output and its termination impedance and resonant frequency are bias dependent as discussed in Section 2.4.2. To achieve an acceptable impedance matching, the variation of the termination impedance, in particular the real part $R_s$ has to be kept low. Figure 3.3 shows the change of $R_s$ in the active-inductor load when the current through it ($I_D$) varies from 0 to 10 mA and the corresponding effect this variation has on the output matching $S_{22}$. The active inductor is designed to operate with an $I_D$ of 4 mA when the differential driver is balanced (i.e., each branch has the same current). Note that the $S_{22}$ shown in Fig. 3.3 assumes an
infinite output resistance $r_0$ in the driver transistors. Low frequency operation is also assumed in this simulation since at high frequencies the termination impedance will have to account for the effects of its imaginary component ($j\omega L$) as well as the capacitive load seen at the pad ($C_{PAD}$). From this plot, it can be seen that by confining $I_D$ between 2 mA and 6 mA, a relative constant $R_s$ can be obtained.

![Plot showing $R_s$ and $S_{22}$ vs. $I_D$.](image)

**Fig. 3.3.** Simulated $R_s$ and the corresponding $S_{22}$ vs. $I_D$.

### 3.2.4 Linearity Enhancement

A method to improve the linearity of the termination impedance while fully switching the tail current is to add a common-mode degeneration resistor ($R_{deg}$) between the differential outputs of the driver [24]. Fig. 3.4 shows the full schematic of the differential version of the output driver with a common-mode degeneration resistor $R_{deg}$ and the proposed active-inductor loads. With the addition of $R_{deg}$ the active-inductor load should exhibit a $R_s$ value such that the effective resistive load seen by the driver is $R_{s,eff}=R_s \parallel 0.5R_{deg}=50$ Ω (for termination) and part of the switching current would flow through $R_{deg}$ depending on the
ratio of $R_s$ and $R_{deg}$. This way a portion of the current is always preserved in one branch of the active-inductor load so as to reduce the variation of the termination impedance while allowing complete current switching.

3.3 Output Driver with Active Inductor Load

3.3.1 Full Schematic

As shown in Fig. 3.4, the load to the driver is a pair of inductive terminations implemented by the PMOS-based active inductors. The differential driver transistors $M_{diff}$, upon the application of the input voltage $V_{in}$, steers the tail current $I_{tail}$ between the loads accordingly.
to create the output logic pulses needed for data transmission. The on-chip bandwidth of the driver is limited by the RC time constant seen at the output nodes. The $R$ term (of the RC) is 25 $\Omega$, resulting from the parallel of the effective termination resistance $R_{s,\text{eff}}$ and the characteristic impedance $Z_T$ of the transmission line (both are 50 $\Omega$). The $C$ term (of the RC) is determined by $C_{PAD}$ which is dominated by the parasitic capacitances of the ESD protect device and the bond pad. Since no ESD protection circuits were used in the design, two MIM capacitors each with a capacitance of 700 fF were placed at the output nodes of the driver to emulate the parasitic capacitance of the ESD protection devices. Unfortunately, the fab mistakenly excluded the MIM option and thus these two capacitors were not fabricated on the chip. As a result, the $C_{PAD}$ seen by the driver is mainly due to the pad which is on the order of 100 fF. The off-chip bandwidth limits are set by the physical impairments of the channel such as skin effect, dielectric loss, and impedance discontinuity. If the bandwidth of the on-chip circuitry is higher than the bandwidth of channel, the peaking created by the active inductor can be used to compensate the part of the high-frequency signal loss in the channel.

3.3.2 Biasing Voltage

![Fig. 3.5. On-chip reference voltage circuits](image-url)
Other than the supply voltage $V_{DD}$, which is provided directly from an off-chip power supply, all three biasing voltages $V_{TAIL}, V_{G2}$, and $V_{G4}$ are generated on-chip using the circuits shown in Fig. 3.5. The voltages (i.e., $V_{DD,TAIL}, V_{DD,G2}$, and $V_{DD,G4}$) used to power the reference voltage circuits are supplied from external power supplies to make the testing easier. Note that however in a commercial chip these DC bias and tuning voltages (i.e., $V_{DD,TAIL}, V_{DD,G2}$, and $V_{DD,G4}$) are typically provided on-chip by bandgap reference generators [25] and digital self-tuning circuitries to reduce the number of pins and the complexity in board design. With the use of reference voltage circuits shown in Fig. 3.5 it is easier to tune the bias voltages with a finer scale since the voltage drop across the diode-connect transistor changes more gradual than the external power-supply voltage. This also makes the bias voltages less sensitive to the variation in external power-supply voltage. Moreover, since the reference-voltage circuit is also a current mirror, it is easy to estimate $I_{TAIL}$ and $I_{D4}$ (current through $M_4$ transistor) by considering the size ratios between the $M_{TAIL}$ and $M_{V,TAIL}$, and $M_4$ and $M_{V,G4}$.

### 3.3.3 Testing Consideration

Ideally, the prototype driver should be designed to have two active inductor terminations, one at the transmit-side and one at the receive side (Fig. 3.6), to provide a higher equalization gain. The measurement setup for evaluating the equalization effect of both the transmit-side and receive-side inductive terminations on the channel response would look like that illustrated in Fig. 3.6. The data pulses transmitted by the output I/O driver (which is equipped with a pair of transmit-side terminations) should travel via the off-chip channel typically (a PCB trace, coax cable, or copper wire) to the receive-side terminations that are also implemented on the same chip as the transmitter driver. To obtain measurements on the
received signals, a high input impedance (e.g., 1000 Ω) oscilloscope has to be used (so as to not load the received signal) by connecting it to the receive-side termination nodes. However, since on-chip probes with 4 high-speed pins were not available, probing four high-speed input and four high-speed output pads on the same chip would not have been made possible. Oscilloscopes with high input impedances that could measure signals at the GHz range were not available either. Thus, in the fabricated chip, only the transmit-side terminations are implemented as shown in Fig. 3.7.

Fig. 3.6. Ideal design with both transmit-side and receive-side terminations.

Fig. 3.7. Actual fabricated chip with only transmit-side termination.
3.4 Experimental Results

The prototype active inductor circuit as the load of a CML output driver is fabricated in STM’s 7-metal 90-nm CMOS process. Fig. 3.8 shows the micrograph of the die under probing. The area of the active inductor circuits measures 17 µm × 25 µm. The circuit operates from a 1 V supply. Note that the circuit only draws current from the 1 V supply while the other off-chip DC voltages (e.g., $V_{DD,TAIL}$, $V_{DD,G2}$, and $V_{DD,G4}$) are used for bias only. The simulation and measurement results are presented in the following sub-sections.
3.4.1 Driver Frequency Response

Due to the limitation of equipment, only the simulated frequency response of the circuit is presented. The simulation was done using Cadence Spectre in the 90-nm design kit with the transistor models provided by the foundry. Since the active-inductor termination is primarily designed for compensating high-frequency signal loss in the channel, the use of active inductor for bandwidth enhancement (i.e., ideally wants to achieve maximum on-chip bandwidth extension without peaking) will not herein be discussed.

![Simulated frequency response of the output driver with active inductance turned off and turned on while sweeping $V_{G2}$](image)

Fig. 3.9. Simulated frequency response of the output driver with active inductance turned off and turned on while sweeping $V_{G2}$

By changing the inductance of the active-inductor load, through adjusting $V_{G2}$, the peaking of the driver’s transfer function can be controlled. Fig. 3.9 shows the simulated $S_{21}$, both with the active inductance on (and while sweeping $V_{G2}$) as well as active inductance off. The active inductance can be turned on and off by adjusting the bias of the active inductor circuit. When the active inductance is off ($L=0$) the peaking is removed and it mimics a
typical $RC$ response when a passive resistor is used in the load. Although both are without peaking, passive resistors are more linear and hence are expected to have somewhat less ISI than in the case when active inductance is turned off. The bandwidth of the driver, when the active inductance is turned on, is observed to be at around 20 GHz in the simulation and is sufficient for the intended data rates in this work (5 Gb/s and 10 Gb/s). Note that this simulation is carried out with an output capacitance $C_{PA D}$ of 100 fF. A smaller bandwidth is expected if a larger $C_{PA D}$ is present. As shown in Fig. 3.9, by tuning $V_{G2}$, the channel can be compensated by up to 3 dB and it allows the peaking frequency to be varied between 2 GHz and 10 GHz.

3.4.2 Channel Frequency Response

![Test setup for the measurement of channel frequency response.](image)

Fig. 3.10. Test setup for the measurement of channel frequency response.
Fig. 3.11. Channels used for measuring the eye-diagrams at the receiver side (a) A 6-inch FR4 board trace (b) A 4-m RG-58/U cable with BNC connectors.

The test setup for measuring the channel frequency responses is depicted in Fig. 3.10. The two ports of the Agilent 8510C vector network analyzer (VNA) are connected to the two ends of the FR4 board and RG-58/U coaxial cable shown in Fig. 3.11, respectively.

Fig. 3.12. Measured responses of a 6-inch FR4 trace and a 4-m RG-58/U cable

Fig. 3.12 shows the measured frequency responses of the 6-inch FR4 trace and the 4-m long coaxial cable with BNC connectors. These two channels are used to demonstrate the tunable peaking that the active inductor circuit can provide to help compensating the channel loss.
For the FR4 trace and RG-58/U cable the losses are about 3.1 dB at 5 GHz and 6.5 dB at 2.5 GHz, respectively.

3.4.3 Eye-Diagram Measurement

![Test setup for eye-diagram measurements](image)

Fig. 3.13. Test setup for eye-diagram measurements

The test setup for the eye-diagram measurement is shown in Fig. 3.13. The differential pseudo-random binary sequence (PRBS) with a magnitude of 300 mV_{p-p} and a DC offset of 700 mV is generated by the pulse pattern generator (Anritsu MP1763B) and fed to the inputs of the driver via SMA cables. The four DC voltage supplies ($V_{DD}$, $V_{DD\_TAIL}$, $V_{DD\_G2}$, and $V_{DD\_G4}$) are provided by four external power supplies. The outputs of the driver are connected to the oscilloscope input ports via the channels. The signals are terminated by the internal 50-Ω resistors of the oscilloscope.
Fig. 3.14. Received eye-diagrams for a 5 Gb/s $2^{31}-1$ PRBS pattern through a 4-m RG-58/U cable (a) active inductance off (b) active inductance on
Fig. 3.15. Received eye-diagrams for a 10 Gb/s $2^{31-1}$ PRBS pattern through a 6-inch FR4 trace (a) active inductance off (b) active inductance on.
Figs. 3.14 and 3.15 show the received eye-diagrams when a 5 Gb/s and a 10 Gb/s 2\(^{31}\)-1 PRBS data patterns are transmitted over the FR4 and coaxial-cable channels, respectively, by the output driver. In Figs. 3.14(a) and 3.15(a) the active inductances are turned off. In Figs. 3.14(b) and 3.15(b) the active inductances are turned on and tuned to give maximum gains at the data-rate frequencies. In Figs. 3.14(a) and 3.14(b) the vertical eye openings are 64 mV and 127 mV with a differential swing slightly less than 300 mV\(_{p-p}\). Peak-to-peak jitters are about 0.41 UI and 0.285 UI, respectively in Fig. 3.14(a) and (b). The vertical eye openings in Figs. 3.15(a) and (b) are 67 mV and 133 mV with a differential swing of 300 mV\(_{p-p}\) and the peak-to-peak jitters are 0.51 UI and 0.35 UI, respectively. The pair of active-inductor loads consumes 0.8 mW overhead powers compared to passive loads due to the use of level shifter. This overhead power is independent of the amount of tail current being used by the driver.

3.4.4 Output Matching Measurement

![Fig. 3.16. Test setup for S\(_{22}\) measurements](image)

Agilent 8510C
VNA

Port 1

Port 2

Design Under Test

V\(_{IN}^+\)
V\(_{IN}^-\)
V\(_{DD}\)
V\(_{DD\_TAI}\)
V\(_{DD\_G2}\)
V\(_{DD\_G4}\)

V\(_{OUT}^+\)
V\(_{OUT}^-\)
The test setup for the output impedance-matching $S_{22}$ measurements is shown in Fig. 3.16. The differential inputs of the driver are connected to the same external DC supply for biasing. The other four DC bias voltages are connected in the same way as described in Section 3.4.3. One of the driver’s outputs is open circuit while another connects directly to a VNA port through the DC-block capacitor. Note that the calibration kit available at the time of measurement, for this VNA, can only calibrate up to the 3.5-mm connector of the VNA cable. The remaining electrical connections through which the signal has to travel to reach the integrated circuit (IC), including the SMA-to-3.5mm connectors and SMA cables, were not calibrated.

Since the impedance of the active-inductor load varies with the bias current as previously shown in Fig. 3.3, it is important to take $S_{22}$ measurements with different DC currents flowing through the active inductor. When the inputs of the driver are biased by the same common-mode DC voltage, the differential pair is said to be balanced and the current through each active-inductor load is just half of the tail current $I_{\text{TAIL}}$. $I_{\text{TAIL}}$ can be adjusted by tuning $V_{D_{\text{TAIL}}}$. 
Fig. 3.17. Measured $S_{22}$ of the output driver with different DC current in the active-inductor load

Figure 3.17 shows the measured $S_{22}$ of the output driver when varying the currents in the active-inductor load. The driver has an $I_{TAIL}$ of 8 mA and $R_{deg}$ is designed to have 2 mA flowing through it when $I_{TAIL}$ is being completely switched by the differential drivers. This means that the current flowing through each branch of the active-inductor loads will vary between 2 to 6 mA (since the tail current is 8 mA) while generating maximum output swing. As shown in Fig. 3.17, $S_{22}$ of -10 dB or lower can be achieved when the current is being confined in this range. $S_{22}$ of -10 dB signifies a 10% signal power reflection (or loss) at the output node of the circuit; in other words, 90% of the incident signal power is absorbed by the circuit. A -10 dB of impedance matching is a typical minimum requirement for most communication systems.
4 Prototype Design II

The focus of this chapter is to demonstrate the bandwidth improvements and saving of area (compared to the use of passive spiral inductors) brought by the use of active-inductor loads in an output buffer designed for high-speed I/O transmitters. The 4-stage cascaded output buffer with active-inductor loads was designed in STM’s 65-nm CMOS process (which was made available to the author some time after the submission of the 1st prototype chip) with a supply voltage of 1.2 V and has been taped out for fabrication. The chip is expected to return by November. Design and simulation results are presented in the following.

4.1 Output Buffer

4.1.1 Introduction

An output buffer enables internal logic circuits to drive the off-chip 50-Ω loads and the large capacitances seen at the outputs of an I/O transmitter. When incorporated into the design of high-speed multiplexers (MUX) in serial-link transmitters, the use of output buffers can cut down the area and power consumption needed to implement the MUX. This is because the MUX can be designed using smaller circuits since the core MUX circuit does not have to drive the off-chip 50 Ω load directly.

An output buffer usually consists of a cascade of multiple driver stages that are tapered in
device dimensions and bias current from the first stage to the last so as to maintain its
bandwidth while delivering high output currents to it loads. At a speed higher than 10 Gb/s
CML drivers are usually called for in the design of an output buffer (since they can achieve
the highest speed of operation among all families of logic circuits, due to the
current-steering and low-swing natures). CML drivers are particularly suitable for
implementing active-inductor shunt peaking since the core driver only has two stacked
transistors, one for the differential pair and one for the current source, which allows for
allocation of voltage headroom for the active-inductor loads. Output buffers, particularly
the ones used in high-speed parallel links, are excellent examples of where the area-saving
benefits of using active inductors (as opposed to spiral inductors) can prove to be very
significant as a large number of such buffers (e.g., in 32-bit or 64-bit processor I/Os), each
employing multiple stages of cascaded CML drivers, are typically needed.

4.1.2 Design Consideration

The power consumption and data rate are the two most important measures of performance
as well as tradeoffs in the design of an output buffer. To achieve a reasonable balance
between the two, a data rate of 30 Gb/s is targeted in this design. Output swing is also
another important performance metric since it directly trades with power consumption. The
magnitude of the output swing is proportional to the vertical eye opening of the received
signal in the receiver side of I/O links, thus a large maximum output swing is usually
desired. Moreover, the output swing should ideally be controlled independently to facilitate
the implementations of transmit pre-emphasis and scalable I/O’s [1]. The fanout (FO) of
each stage in the output buffer is dependent on the bandwidth requirement in each stage.
The number of stages needed in a tapered buffer is dependent on the total fanout ratio \( \text{FO}_{\text{tot}} \).
(ratio between the output capacitance seen by the last stage and the input capacitance of the first stage). A larger \( FO_{tot} \) allows the internal logic or MUX circuits preceding the output buffer to be made smaller so as to save power and area for a given output capacitor load \( C_{PAD} \).

### 4.1.3 Architecture

The architecture used in the output buffer design is illustrated in Fig. 4.1. To allow a controllable output swing, the main driver of the output buffer uses only passive resistor loads and has a separate bias voltage \( V_{TAIL2} \) for its tail current transistor. The pre-driver consists of three cascaded drivers, each employing active-inductor loads that are represented by a pair of resistors \( (R_s) \) and inductors \( (L) \) enclosed by dashed boxes. The three current sources in the pre-driver are biased by the same voltage \( V_{TAIL1} \). Note that the \( RC \) time constant seen at the output node of the main driver is relative small \( (R=25 \, \Omega \, \text{due to the parallel of a transmit-side and a receive-side 50-}\Omega \, \text{terminations and a small } C_{PAD} \, \text{of 250 pF}), \) thus it has a sufficiently large bandwidth even without the use of inductive peaking.
4.2 Active-Inductor Load

4.2.1 Topology

Fig. 4.2. Topology of the NMOS-based active-inductor load.
Fig. 4.2 depicts the NMOS-based active-inductor load with $V_{\text{HIGH}}$ that was discussed in Section 2.4.1, except the replacement of the passive resistor $R_G$ by a PMOS active resistor $M_2$. This active inductor topology is chosen as the loads for the three stages of the pre-driver to implement active shunt peaking. The advantages of this topology are higher speed due to the higher charge-carrier mobility of NMOS transistors as compared to the PMOS transistors and zero power overhead. The only power overhead that is needed is to power the voltage boosting circuitry that might be needed for generating $V_{\text{HIGH}}$. It is also possible to provide $V_{\text{HIGH}}$ through off-chip at the cost of an additional pin. In SoC designs very often there are multiple $V_{\text{DD}}$’s, and the ones used in the I/O’s are often higher than the ones used in the logic cores. To comply with the signaling levels required for different I/O standards, voltage supplies as high as 1.8 V and 3.3 V are usually available even in designs realized in more advanced CMOS processes such as the 90-nm and 65-nm nodes. In such cases the use of voltage-boosting circuitry would not be necessary and the additional power cost for using NMOS-based active inductors is further reduced.

4.2.2 Biasing Voltage

![Fig. 4.3. $g_m$ vs. $V_{GS1} (W_1/L_1)=(12\text{um}/60\text{nm})$.](image)

Fig. 4.3. $g_m$ vs. $V_{GS1} (W_1/L_1)=(12\text{um}/60\text{nm})$.
By inspecting (2.10) and (2.12), it can be found that the resonant frequency $\omega_0$ of the shunt-peaking active inductor is proportional to $g_{m1}$ for a given $M_1$ transistor size and a given inductance value $L$ (in the active inductor). The choice of the gate biasing voltage $V_{HIGH}$ for $M_1$ transistor should be made based on optimizing $g_{m1}$. This should be done while ensuring that the gate-source voltage $V_{GS}$ of the $M_1$ transistor (which is $V_{HIGH}-V_{out}$) does not exceed 1.2V (the supply voltage $V_{DD}$) at all time of operation, so as not to damage the thin gate oxide in the transistor. A single NMOS transistor $M_1$ is simulated in the 65-nm design kit using Cadence Spectre with models provided by the foundry. Fig. 4.3 plots $g_{m1}$ versus $V_{GS1}$ with different $V_{DS1}$ for a $M_1$ transistor having a drawn gate width and length of 12 µm and 60 nm, respectively. It can be seen that for all three plots, $g_{m1}$ peaks at around $V_{GS1}=0.8V$ and remains relatively constant (except for $V_{DS}=0.6V$, which drops a little) up to $V_{GS1}=1.2V$. Note that the drop in the $V_{DS1}=0.6V$ plot is attributed to the fact that $M_1$ transistor starts to enter into triode region of operation as its overdrive voltage ($V_{GS1}-V_{th1}$) exceeds its $V_{DS1,sat}$. This figure suggests that the optimal $V_{GS1}$ value is around 1 V for an output swing of ± 200 mV ($\Delta V_{out}$) in the driver (source of $M_1$ transistor is the same node as $V_{out}$ of the driver). The DC output level and output swing in the pre-driver should be set to around 0.8 V and ± 200 mV (which corresponds to 0.2 V $\leq V_{DS1} \leq 0.6$ V), respectively. This implies a bias voltage $V_{HIGH}$ of 1.8 V is required to result in optimal $V_{GS1}$.

4.3 Design Procedure

4.3.1 Main Driver

The loads of the main driver are designed with two 50-Ω poly resistors to provide output impedance matching. At a multi-Gb/s link the receiver almost always has to have 50-Ω
on-die terminations to minimize reflection and signal power loss. Therefore the main driver at the transmitter effectively sees a 25-Ω load at its outputs. The tail current of the main driver is determined by the division between the desired output swing and the 25-Ω load (i.e., \( I_{TAIL} = (2 \times \Delta V_{out})/25\Omega \)). The output swing of the main driver is designed to be maximized (i.e., larger than that of the pre-driver) to increase the overall signal integrity (±350 mV). The maximum output swing it can achieve is largely limited by the modest input drive voltage of ±200 mV provided by the pre-driver. The lower swing level in the pre-driver helps lower the overall power consumption at the cost of a reduced noise margin. To deliver such a large current to its loads the differential pair transistors \( M_{diff} \) in the main driver have to be sized sufficiently large. The large input capacitances associated with the main driver could limit the bandwidth of its preceding stages.

### 4.3.2 Pre-Driver

To drive the large input capacitances of the main driver, the pre-driver is tapered into three cascaded stages. When \( N \) stages are cascaded as an amplifier (e.g., limiting amplifiers), the bandwidth of each stage has to be made larger than the overall bandwidth of the amplifier. Assuming each stage has the same -3-dB bandwidth \( \omega_0 \) and the overall circuit has a -3-dB bandwidth of \( \omega_{3dB} \), the relationship between the two is given by [26]:

\[
\omega_{3dB} = \omega_0 \sqrt{\frac{N}{2}} - 1
\]  

(4.1)

Thus, for the 3-stage pre-driver,

\[
\omega_{3dB} \approx 0.51\omega_0
\]  

(4.2)

This implies that for the pre-driver to operate at the maximum data rate with reasonable ISI, each stage has to have a -3-dB bandwidth of around twice as large as that of the entire
pre-driver circuit. To increase the overall bandwidth of the circuit, active-inductor shunt peaking is employed in all three stages. The resistance $R_s$ in the active-inductor load of each stage is chosen based on the load capacitance $C_L$ and the bandwidth requirement of each stage. From the $R_s$ and the voltage swing ($\pm 200$ mV), the tail current of each stage can be determined. The driver transistors of each stage are sized to guarantee the desired output swing can be achieved and propagate down the driver chain. Since the gain does not need to be particularly high (slightly higher than unity in this case), the degree to which the output common-mode levels in the 3-stage pre-driver could fluctuate in the presence of PVT variations is also limited. Thus, common-mode feedback is not required. The inductance $L$ of the active-inductor load is designed to give a maximally flat response ($L=0.4 \times R_s^2 \times C_L$) [18] in each stage.

4.3.3 Tail Current

Due to the use of active-inductor loads in the pre-driver, special attention has to be given to the output common-mode level as it is more sensitive to PVT variations mainly because the non-linear I-V characteristics of the active-inductor loads. As such, the output impedance of the tail current transistors in each stage of the output buffer has to be made high (by using longer channels and allocating sufficient voltage headroom to the tail current transistors,) so as to minimize $I_{TAIL}$ variation due to the change in the common-mode output level of previous stage. The two biasing voltages for the tail currents transistors $V_{TAIL1}$ and $V_{TAIL2}$ are provided by the same current-mirror circuits as described in Section 3.3.2. The schematic of the pre-driver and main driver is shown in Fig. 4.3. Transistor sizes, $R_s$ and $L$ values in each active-inductor load, $C_{PAD}$ value, and tail current in each driver are labeled in the schematic as well.
4.3.4 Active-Inductor Load

As shown in Fig. 4.6 output degeneration resistors $R_{\text{deg}}$ are used in all three stages of the pre-driver to improve the linearity of the active-inductor loads. The width of $M_1$ transistor in each stage’s active inductor is sized to exhibit an effective resistance $R_{s,\text{eff}}$ that will result in the desired $R_s$ as shown in Fig. 4.5 (i.e., $R_s = R_{s,\text{eff}} // 0.5R_{\text{deg}}$). Minimum length is used in the $M_1$ transistors of all three stages. The inductance $L$ of the active-inductor load can be independently tuned through $V_{G2}$, which controls the resistance of the active resistor implemented by $M_2$ transistor. To reduce the number of pads needed for biasing voltages (since the chip will be tested using on-chip probing and the number of pins on the probe is limited), the tuning voltage $V_{G2}$ is shared by all three active resistors and is provided directly from off-chip. Thus, the $M_2$ transistors in all three stages have to be sized in proportion to their desired $L$ values for a given $V_{G2}$ voltage. Fig. 4.6 shows the three-stage pre-driver with active-inductor loads, each with transistor sizes, $R_{\text{deg}}$, and $V_{\text{HIGH}}$ labeled.
Fig. 4.5. Three-stage pre-driver with active-inductor loads and transistor sizes, tail currents and $R_{\text{deg}}$ values labeled.

![Three-stage pre-driver](image)

Fig. 4.6. Layout of the output buffer with active-inductor loads.

The final layout of the entire output buffer with active-inductor loads in the pre-driver stages is shown in Fig. 4.6. It occupies an area of 130$\mu$m × 30$\mu$m, which is about the size of a typical spiral inductor having an inductance of 0.5 nH ~ 1 nH.
4.4 Simulation Results

To evaluate the effectiveness of the bandwidth enhancement gained from the use of active-inductor shunt peaking, as opposed to when passive inductors or no shunt peaking are used, pre-drivers using passive-resistor loads and passive-inductor loads are also designed and simulated. The values of load resistors and inductors used in the passive-load implementations are the same as that shown in Fig. 4.5. In an attempt to make the comparisons fair, all three implementations have the same levels of tail currents, output swings, and same device dimensions in the differential pairs and current source transistors, in each stage. The only difference is that with a swing of ±200 mV, the passive implementations have an output common-mode level of 1 V while the active implementation has an output common-mode level of about 0.8 V. The additional 200 mV voltage headroom is used to allow a more optimal biasing for the active-inductor loads so that they can achieve a faster speed and operate more linearly. The tradeoff is the reduced voltage headroom for the tail current transistors $M_{TAIL}$ in the pre-driver using active-inductor loads; the bias voltage for these $M_{TAIL}$ transistors thus is adjusted to match the tail current levels of that in the two pre-drivers using passive loads. All simulations are done using Cadence Spectre at the schematic level and presented in the following.

4.4.1 DC Transfer Characteristics

To gain insight into the difference that might have resulted in the DC transfer characteristics of the pre-driver due to the use of active-inductor loads as opposed to passive loads, the DC input voltage is swept against the output voltages and currents of the circuit.
The input/output characteristic of the three-stage pre-driver with active-inductor loads is shown in Fig. 4.6. This graph validates that the pre-driver has a large-signal gain of about unity. With a differential input voltage of 400 mV applied to the pre-driver, the differential swing generated at its outputs is also about 400 mV. The lowest level that the output signal can reach is limited to above 600 mV because the common-mode level is designed to be slightly higher than 800 mV. The intent for making the CM level slightly higher is to reduce the chances for the voltage between the gate and source ($V_{GS}$) of $M_1$ transistor to exceed 1.2 V under the presence of PVT variations (Note that $V_{HIGH}$ is 1.8 V so $V_{out}$ cannot go below 600 mV).
Fig. 4.7 shows the changes in the drain currents of the $M_1$ transistors in the active-inductor loads of the 3$^{rd}$-stage pre-driver vs. the differential input swept on the 1$^{st}$-stage pre-driver. The tail current in the 3$^{rd}$ stage is designed to be around 7.8 mA when the differential pair is balanced. That is, the point when differential input is zero. As can be seen from Fig. 4.7, the tail current is not completely switched between the two active-inductor loads, even at a high level of signal input. Nonetheless the tail current is completely switched between the driver’s differential pair transistors $M_{diff}$. This phenomenon is mainly due to the presence of the degeneration resistors $R_{deg}$, through which a portion of the switching current detours from the higher potential node to the lower potential node of the differential outputs when they are unbalanced. This way the current flowing into one of the $M_{diff}$ transistors is always the sum of the current from the active-inductor load in its branch and the current diverted from the opposite branch of active-inductor load via $R_{deg}$. As described in Section 3.2.3, the purpose of the output degeneration resistor is to moderate the impedance variation in the active-inductor loads.

Fig. 4.8. $I_D$ vs. $V_{in\_diff}$ (Pre-driver with active-inductor loads).
Fig. 4.9. Differential $G_m$ vs. $V_{in\_diff}$ (Pre-driver with active-inductor loads).

Shown in Fig. 4.8 is the total transconductance $G_m$ (i.e., $\partial \Delta I_{D4} / \partial \Delta V_{in}$) of the pre-driver versus the differential input voltage to the pre-driver. Similar to a typical differential pair with resistive loads, $G_m$ peaks when the circuit is balanced and decays to zero as the current switches from one branch to another.
4.4.2 Frequency Response

As shown in Fig. 4.9, the small-signal bandwidth of the pre-driver (when every differential driver stage is balanced) gives a quick estimate and comparison on the maximum speed that each of the three output buffers, two with bandwidth enhancement (through the use of active-inductor loads and passive-inductor loads) and one without bandwidth enhancement (using only passive-resistor loads), can achieve. The -3-dB bandwidth of the pre-driver using active-inductor loads is about 11 GHz, slightly better than the 10 GHz in the case when passive-inductor loads are used. The pre-driver designed with passive-resistor loads has a -3-dB bandwidth of 7.5 GHz. Note that the pre-driver using active-inductor load has a low-frequency gain 1-dB lower than that of the two pre-drivers using passive loads. This is because, in large signal operations the small-signal magnitude responses of all three cases change with their bias points constantly and thus cannot represent the actual (large-signal) transient voltage gain of the pre-drivers. For example, at other bias points (other than...
balanced mode) the pre-driver with active-inductor loads might have higher small-signal gains than the pre-drivers with passive loads such that the overall gain of each of the three pre-drivers will end up the same in the transient responses. The transient responses of the three buffer circuits will be presented in the following sub-section. The frequency responses of the 1st-stages of the pre-drivers designed are also plotted in Fig 4.9. The -3-dB bandwidths in the cases when active-inductor, passive-inductor, and passive-resistor loads are used are about 20 GHz, 23 GHz, and 17 GHz, respectively. Note that when operated under large signal, the bandwidth of the drivers using active-inductor loads varies more than the drivers using passive loads. This is because both the resistance $R_s$ of the active inductor and the output resistance $r_0$ of the driver transistor change with the output level. In the case when passive loads are used only $r_0$ changes. Therefore, to more accurately predict the speed of the circuits using active-inductor loads under large-signal operation, analysis on the transient responses is helpful and is presented in the next sub-section.

### 4.4.3 Step Response

![Step Response Graph](image)

Fig. 4.11. Step response (rising) of the 1st-stage pre-driver.
To examine the differences in the large-signal transient responses between the pre-drivers designed using active-inductor, passive-resistor, and passive-inductor loads, two input step voltages of 400 mV (peak-to-peak swing), one with a rise time of 5 ps and another with a fall time of 5 ps, are applied to the inputs of the three pre-drivers. Fig. 4.10 shows the transient responses of the 1st-stage pre-drivers in all three cases when the rising input step is applied. The 10% - 90% rise times for the output signals, in the cases where active-inductor and passive-resistor loads are used, are about the same (~26 ps). The transient responses when a falling input step is applied are plotted in Fig. 4.11. The 90% - 10% fall time for the output response in the case of active-inductor load is about the same as that of the passive-inductor load (~13 ps). The response speed of the 1st-stage pre-driver using active-inductor loads, upon the application of the falling input step, is about twice faster than that when the rising input step is applied. This can be attributed to the reduced output RC time constant as $R_{s,eff}$ of the active-inductor load decreases when the output level swings down (i.e., $r_0$ of driver transistors $M_{diff}$ decreases).
Fig. 4.12 and 4.13 illustrate the rise-time and fall-time step responses of the three pre-drivers. Both the rise-time and fall-time performances of the active-inductor load implementation are about in the midway between the passive-inductor and passive-resistor implementations.
4.4.4 Eye-Diagram

Fig. 4.15. Eye-diagram of the output buffer using active-inductor loads at 31.25Gb/s.

Fig. 4.16. Eye-diagram of the output buffer using passive-inductor loads at 31.25Gb/s.
Fig. 4.17. Eye-diagram of the output buffer using passive-resistor loads at 31.25Gb/s.

The eye-diagram simulations are performed by applying a 31.25 Gb/s ($2^{31}-1$) PRBS with ±200 mV magnitude generated by a Matlab script written by the author, to the inputs of the three output buffers. The eye-diagrams of the output buffers using active-inductor, passive-inductor, and passive-resistor loads in their pre-drivers are shown in Fig. 4.14, 4.15, and 4.16, respectively. Table 4.1 summaries the eye-diagram simulation results of the three cases as well as the relative differences between the active inductor and passive inductor cases and the active inductor and passive resistor cases, respectively.
Table 4.1. Summary of the eye-diagram simulation.

<table>
<thead>
<tr>
<th></th>
<th>Vertical Eye Opening</th>
<th>Jitter_{p-p}</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Active-Inductor Load</td>
<td>710 mV</td>
<td>~1.4 ps</td>
</tr>
<tr>
<td>II. Passive-Inductor Load</td>
<td>760 mV</td>
<td>~1.9 ps</td>
</tr>
<tr>
<td>III. Passive-Resistor Load</td>
<td>515 mV</td>
<td>~5.8 ps</td>
</tr>
</tbody>
</table>

I. relative to II.  
- 6.5%  
+36%

I. relative to III.  
+37.8%  
+340%

To illustrate the speed improvement in the case when active-inductor loads are used, as compared to when passive-resistor loads are used, the output buffer using only passive-resistor loads are simulated with the same PRBS input pattern at lower data rates. It has been found that the buffer using passive-resistor loads operating at 25 Gb/s has a comparable output eye as that of the buffers using active-inductor or passive-inductor loads operating at 31.25 Gb/s. The output eye-diagram of the buffer with passive-resistor loads operating at 25 Gb/s is shown in Fig. 4.17. The vertical eye opening is about 740 mV and the peak-to-peak jitter is approximately 2.5 ps. The speed improvement of the active-inductor over the passive-resistor loads is about 25% while the jitter of the buffer with active-inductor loads (~1.4 ps) is about 78% less than the buffer with passive-resistor loads (~2.5 ps).
Table 4.2 gives the summary on the performance of the output buffer designed using active-inductor loads and compares the results with other output buffers reported in literature that are used in high-speed data transmitters.
Table 4.2. Performance summary of the output buffer designed with active-inductor loads and comparisons with other reported output buffers.

<table>
<thead>
<tr>
<th></th>
<th>This work&lt;sup&gt;a&lt;/sup&gt;</th>
<th>[27]&lt;sup&gt;b&lt;/sup&gt;</th>
<th>[28]&lt;sup&gt;b&lt;/sup&gt;</th>
<th>[29]&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS Technology</td>
<td>65-nm</td>
<td>80-nm</td>
<td>90-nm</td>
<td>0.15-µm</td>
</tr>
<tr>
<td>$V_{DD}$ (V)</td>
<td>1.2</td>
<td>1</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Data Rate (Gb/s)</td>
<td>31.25</td>
<td>40</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>Power (mW)</td>
<td>41.2</td>
<td>24</td>
<td>50.4</td>
<td>45</td>
</tr>
<tr>
<td>Differential $V_{out}$ (mV)</td>
<td>710</td>
<td>660</td>
<td>400</td>
<td>900</td>
</tr>
<tr>
<td>Fanout ratio&lt;sup&gt;c&lt;/sup&gt;</td>
<td>41.6</td>
<td>23.75</td>
<td>8&lt;sup&gt;d&lt;/sup&gt;</td>
<td>*</td>
</tr>
<tr>
<td># of spiral inductors</td>
<td>0</td>
<td>10</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Area</td>
<td>130µm × 30µm</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

<sup>a</sup> Simulated.  
<sup>b</sup> Measured.  
<sup>c</sup> Estimated.  
<sup>d</sup> Not specified.  
<sup>e</sup> Division between $C_L$ and $C_{in}$ (input capacitance of the 1<sup>st</sup>-stage pre-driver) of the output buffer (i.e., 250fF/6fF).  


5 CONCLUSIONS

5.1 Conclusion

This thesis explores the use of active inductors as a compact alternative to passive spiral inductors for implementing shunt-peaking to extend on-chip bandwidth and providing equalization for channel loss through tunable peaking.

A novel PMOS-based active inductor topology that can operate with low voltage headroom and requires no voltage boosting has been proposed. The first prototype design is an output driver circuit using the PMOS-based active inductor as its termination load and is implemented in a 90-nm CMOS process. The peaking frequency and its corresponding magnitude of the active inductor circuit can be adjusted to facilitate channel loss compensation. Operating at 10 Gb/s over a 6-inch FR4 channel, as compared to the case when the active inductor structure is disabled, the use of active inductor in the transmit-side termination increases the vertical eye opening at the receiver side by a factor of two and reduces the received peak-to-peak jitter by 30%. Appropriate output impedance matching of $S_{22}$ less than -10 dB is achieved by the active-inductor termination. The pair of active inductor circuits occupies $17 \times 25 \, \mu m^2$ and has a low overhead power consumption of 0.8 mW.

The second prototype design is a 4-stage output buffer using NMOS-based active-inductor shunt peaking for bandwidth extension and has been taped out for fabrication in a
cutting-edge 65-nm CMOS process. Through simulations it was verified that the output buffers with active-inductor shunt peaking compared favorably with passive-inductor shunt peaking. The peak-to-peak jitter when the active-inductor shunt peaking was used was more than three times lower and the vertical eye opening is more than 30 % larger than that when no shunt peaking was used. When operating at 25% higher data rate than the buffer designed using passive-resistor loads (i.e., 31.25 Gb/s vs. 25 Gb/s), the use of active-inductor loads achieves the similar vertical eye opening and 78% better jitter performance. The output buffer with active-inductor shunt peaking achieved a data rate of 31.25 Gb/s, a peak-to-peak output swing of 710 mV, and a total fanout ratio of 41.6 while dissipating 41.2 mW and occupies an area of $135 \times 30 \mu m^2$.

5.2 Contributions

The main contributions of this thesis are summarized as following:

- Introduction of a novel active inductor topology
  
  - Low voltage headroom (200 ~ 300 mV)
  
  - Needs no voltage boosting and consumes little power overhead (0.8mW)
  
  - Published in *IEEE International Symposium on Circuits & Systems (ISCAS) 2008*

- First to propose using active inductors for high-speed I/O terminations
  
  - Increases received voltage margin by a factor of two over a 6-inch FR4 channel
  
  - Achieves $S_{22}$ of -10 dB and better
  
  - Accepted by *IEEE Asian Solid-Solid Circuits Conference (A-SSCC) 2008*

- Design of three-stage pre-driver using active-inductor loads in a 30 Gb/s output buffer
  
  - Fastest shunt-peaking active inductors that have been reported up to date
(i.e., 40 Gb/s in each stage in simulation)

- Speed improves by 25% while having 78% less jitter compared to the same output buffer designed using passive-resistor loads
- Consumes no extra power comparing to when passive loads are used

5.3 Future Work

For active inductors to be reliably deployed in high-speed I/O circuits, the robustness of such circuits in the presence of PVT variations has to be more thoroughly investigated, especially when implemented in very DSM CMOS technologies (i.e., 65 nm and below) in which process variations are getting much worse. It is thus one of the most important follow-up works to be considered. Another future work will be testing the 2\textsuperscript{nd} prototype design that has been presented (the 4-stage output buffer using active-inductor loads) when the chip returns in November. Also of interest is to apply the proposed active-inductor shunt peaking technique to other speed critical I/O building blocks such as MUX and DEMUX. Most MUX and DEMUX operating at 10 Gb/s or higher employ the tree architecture [26] in which CML latches and selectors are used extensively as their constituent sub-circuits. Both CML latch and selector need three stacking transistors excluding the loads (CML driver needs only two). It is thus expected to be more challenging to use active-inductor loads in these structures due to the reduction in available voltage headroom.
REFERENCES


