dCollection

Ph.D. Dissertation

Design of High-Speed PAM4-Binary Bridge for Memory Testing

메모리 테스트를 위한 고속 PAM4-바이너리 브리지 설계

by

Daeho Yun

August, 2023

Department of Electrical and Computer Engineering
College of Engineering
Seoul National University

Design of High-Speed PAM4-Binary Bridge for Memory Testing

지도 교수 정 덕 균

이 논문을 공학박사 학위논문으로 제출함
2023 년 8 월

서울대학교 대학원
전기·정보공학부
윤 대 호

윤대호의 박사 학위논문을 인준함
2023 년 8 월

위 원 장 _____ (인)

부위원장 _____ (인)

위   원 _____ (인)

위   원 _____ (인)

위   원 _____ (인)

Design of High-Speed PAM4-Binary Bridge for Memory Testing

by

Daeho Yun

A Dissertation Submitted to the Department of
Electrical and Computer Engineering
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
at
SEOUL NATIONAL UNIVERSITY

August, 2023

Committee in Charge:

Professor Suhwan Kim, Chairman

Professor Deog-Kyoon Jeong, Vice-Chairman

Professor Woo-Seok Choi

Professor Yongsam Moon

Professor Jun-Eun Park

Abstract

High-performance computing applications such as machine learning and A/I require high memory bandwidth. Multi-level signaling is being considered to meet the bandwidth demands of DRAM, but it necessitates significant infrastructure changes, particularly for DRAM products produced in mass quantities. In addition, DRAM manufacturers have large-scale facilities for evaluating Non-Return-to-Zero signals, so implementing multi-level signaling support would require costly and time-consuming test facility changes. A bridge chip has been proposed to address this problem by converting input/output data from low-performance test equipment into high-speed PAM4 signals, which are then transmitted to DRAM.

For the first chip, a 32 Gb/s PAM4-Binary bridge for the next-generation memory testing is presented. The bridge incorporates all the required functions to evaluate a high-speed PAM4 memory using a low-speed NRZ tester. The low-speed data transmitted from the NRZ tester to the bridge are converted into high-speed PAM4 data through half-rate clock control, forwarded to the memory, and vice-versa. The ground-terminated PAM4 driver provides the single-ended output by controlling the output current with a 2-tap feed-forward equalizer, achieving a ratio level mismatch (RLM) of 0.95. To minimize the offset at the PAM4 receiver, the offset cancellation circuit with an offset of 2.76 mV consisting of a CTLE and sampling latches is employed, and the horizontal margin of the received PAM4 signal is 50% for BER<10^{-9}. An all-digital PLL integrated in the bridge doubles the 4 GHz WCK used as a forwarded clock for the graphic memory. The count-based PAM4 eye-opening monitor

is also proposed to find the optimal codes for the maximum eye opening using the PRBS7 data sequence. The bridge fabricated in the 40-nm CMOS technology occupies an active area of 1.6 mm^2 and dissipates 132 mW.

The second chip presents a 48 Gbps PAM4 memory interface with a level mismatch adjustment capability for a high-speed PAM4 memory/tester bridge. The bridge incorporates all the required functions to test and validate a high-speed PAM4 memory using a low-speed NRZ tester. The level-adjustable PAM4 TX is designed as a voltage mode CMOS driver and improves the RLM through a calibration circuit. The RX achieves BER less than 10^{-12} through equalizers such as parallel CTLEs and 1-tap DFE. The bridge operates at 48 Gbps per pin and consumes 1.85 pJ/bit and 2.97 pJ/bit for the write and read modes of the PAM4 memory, respectively. The proposed bridge is fabricated in 40 nm CMOS technology, occupying 2.13x1.098 mm^2.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Presently, there is an increase in data traffic, which is projected to persist due to the emergence and growth of the Internet of Things (IoT), Artificial Intelligence (AI), and Machine-To-Machine (M2M) technologies, alongside existing computers, mobile devices, and data centers. Cisco's annual Internet report forecasts that the number of connected devices will reach 29.3 billion by 2023, with M2M technology driving the growth of the Internet of Things on 50 % of all connected devices [1]. The exponential increase in mobile data traffic, including M2M data, is predicted to reach 5 zettabytes per month by 2030, according to the International Telecommunication Union [ITU] (Fig. 1.1). This growth necessitates faster data processing speeds, which may be achieved through improved I/O transmission speed of memory.
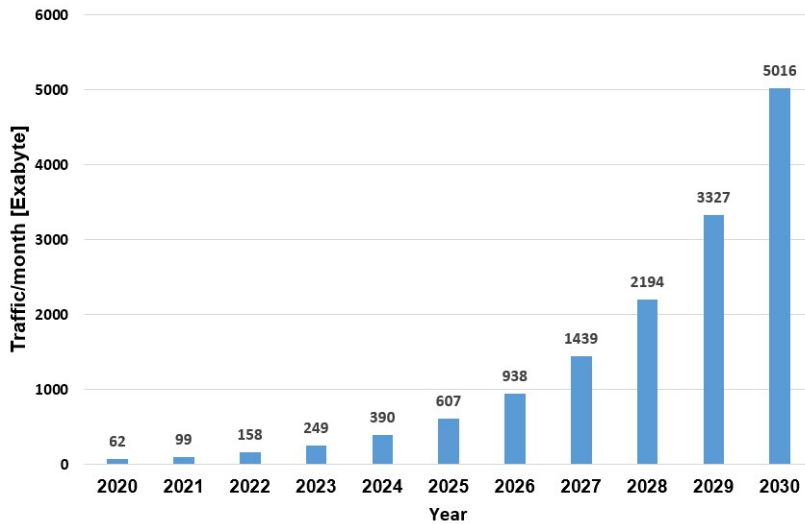
Fig. 1.1 Global mobile data traffic forecast.
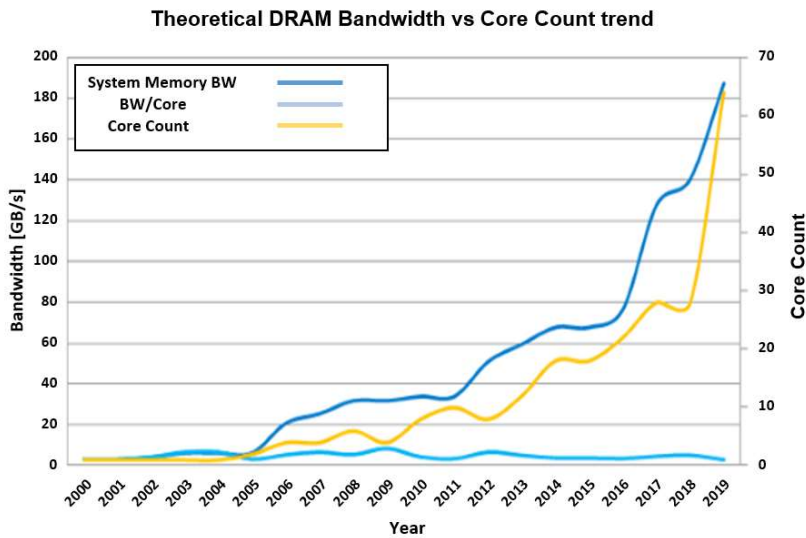
Fig. 1.2 The trend of increasing computing core and memory bandwidth.

The demand for memory bandwidth has increased due to the development of technologies such as AI, GPU accelerators, and cloud services. The "memory wall"

issue was identified in the 1990s, indicating that the rate of improvement in microprocessor performance exceeded that in DRAM memory speed. Furthermore, the recent development of multi-core CPU architecture has led to an unprecedented increase in core counts, significantly increasing the memory subsystem's bandwidth requirements, as shown in Fig. 1.2 [2]. Meanwhile, for the next-generation memory standards, including GDDR7 and Post-DDR5 [3], multi-level signaling such as pulse amplitude modulation (PAM) has been discussed to secure a link margin that has reached its limit with existing binary methods. In accordance with this, a multi-level signaling interface applicable to memory is under development by major memory manufacturers. The pulse amplitude modulation4 (PAM4) signal can transfer twice as much information at the same Nyquist frequency as the non-return to zero (NRZ) signal.

However, existing test solutions using Automatic Test Equipment/System Level Test (ATE/SLT) offer only a low-speed binary mode lacking a multi-level signaling capability. Thus efficient evaluation methods for PAM4 signals must be explored. Furthermore, the DRAM has only used NRZ interfaces so far, and changing the test equipment used in the past is necessary to apply the PAM4 interface to DRAM. For example, as shown in Fig. 1.3, a T5511 (ADVANTEST) tester is popularly used to test and verify the internal and external interface operation in characterizing DRAM interface. Still, only a binary mode is supported, and its maximum clock speed is 4 GHz, thereby limiting the testable data rate to 8 Gb/s. Moreover, new test equipment for PAM4 signaling is not expected to be available soon since the increased interface speed inevitably incurs larger signal attenuation over the lengthy test cable.
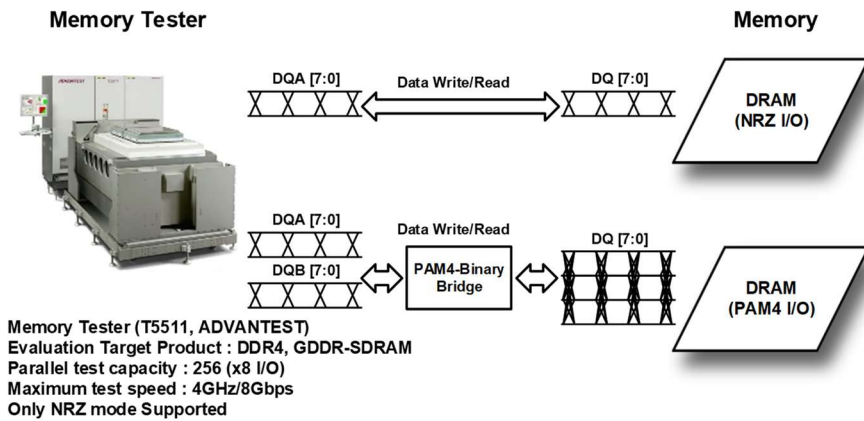
Fig. 1.3 Memory test environment between DRAM and ATE.

A PAM4-Binary Bridge that serves multiple functions such as a translator, an equalizer, and a retimer between a popularly used low-speed binary tester and a high-speed PAM4 memory will play a key role in testing such newly developed memory interfaces. Thus, for future high-speed a PAM4 signaling tests, a bridge chip for memory testing with PAM4 transceiver (TX/RX) of up to 48 Gb/s is required.

This thesis proposes the high-speed PAM4-Binary Bridge architecture, which includes all the necessary functions to test and validate a high-speed PAM4 memory using existing test equipment. Furthermore, the proposed PAM4-Binary Bridge is equipped with the TX circuit capable of adjustable tuning of the RLM, impedance matching, and the RX circuit capable of offset cancellation, therefore achieving high data rates and advanced memory test tasks compared to state-of-the-art circuits.

1.2 Thesis Organization

This thesis is organized as follows. Chapter 2 provides the backgrounds of a high-speed interface and channel equalization. First, the basic concept of multi-level pulse amplitude modulation in a high-speed interface is briefly explained. Then the necessity of the PAM4-Binary Bridge is provided.

Chapter 3 presents a 32 Gb/s PAM4-Binary Bridge for next-generation memory testing. The ground-terminated PAM4 driver provides the single-ended output by controlling the output current with a 2-tap feed-forward equalizer, achieving a RLM of 0.95. To minimize the offset at the PAM4 receiver, the offset cancellation circuit with an offset of 2.76 mV consisting of continuous-time linear equalizers (CTLE) and sampling latches is employed. The horizontal margin of the received PAM4 signal is 50 % for BER<10^{-9}. The count-based PAM4 eye-opening monitor is also proposed to find the optimal codes for the maximum eye opening using the PRBS7 data sequence.

Chapter 4 presents a 48 Gbps PAM4 memory interface with a level mismatch adjustment capability for use in a high-speed PAM4 memory/tester bridge. The level-adjustable PAM4 TX is designed as a voltage mode CMOS driver and improves a RLM through a calibration circuit. The RX achieves BER less than 10^{-12} through equalizers such as parallel CTLEs and 1-tap decision feedback equalizer (DFE).

Chapter 5 details the design for testability and presents the measurement environment. Finally, Chapter 6 summarizes the proposed works and concludes this thesis.

Chapter 2

Background of High-Speed Memory Interface

2.1 Overview

To achieve higher bandwidth, DRAM technology is evolving in two ways as the systems become more advanced. One way is to increase the number of pins for data input and output similar to HBM, while the other is to improve the transmission speed per pin like DDR and GDDR. Table 2.1 outlines the bandwidth, pin speed, and pin count of different DRAM types.

Table 2.1 DRAM bandwidth according to the type of DRAM

Type	HBM2	GDDR5	GDDR6	DDR4	DDR5
# of Pin	1024	16/32	16/32	4/8/16	4/8/16
Data rate	1.7	11.4	21	3.2	6.4
Bandwidth	217.6	45.6	84	6.4	12.8
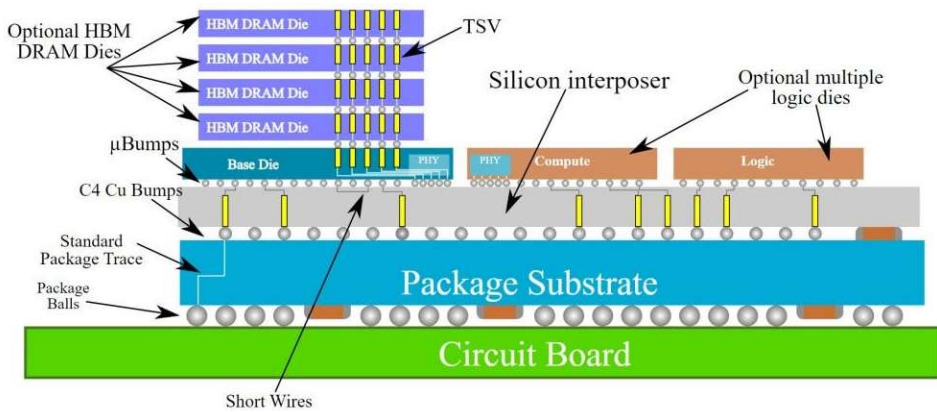
Fig. 2.1 2.5D/3D system architecture with HBM memory.

One way to enhance bandwidth is by increasing the number of I/O pins, which can be accomplished through a structure like HBM, as depicted in Fig. 2.1. This architecture is created by employing the TSV process and silicon interposer, where the DRAM core die stacked with TSV transmits data to the base (bottom) die, and the base die forwards the data to the memory controller via u-bump connected silicon interposer. Utilizing the silicon interposer allows more I/O to be integrated than the conventional wire method. However, a high price is necessary to manufacture HBM because it uses a silicon interposer and TSV. Additionally, the replacement cost is substantial in a single failure.

Researchers continually explore expanding bandwidth to overcome these challenges by increasing the pin speed on existing PCB wires. With each new generation of DDR or GDDR, the pin speed improves. However, as illustrated in Fig. 2.2, increasing the data rate per pin results in a more significant loss and an inability to increase the data rate further using NRZ signaling. Multi-level signaling, particularly PAM4 signaling, is currently being studied in DRAM as an alternative solution to
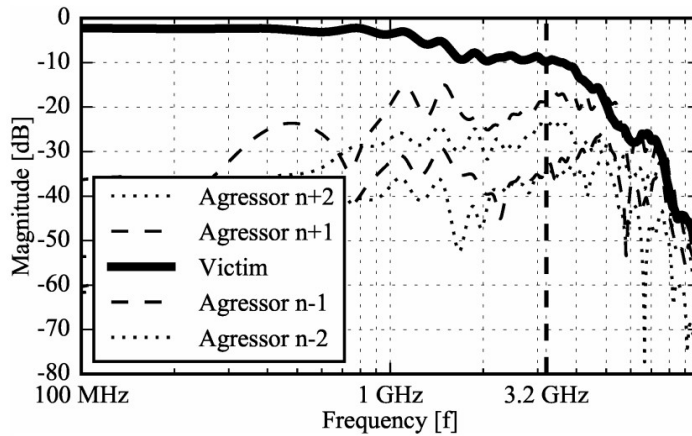
Fig. 2.2 Long channel frequency response for a modern server configuration.

improve bandwidth at the same Nyquist frequency.

Preparations for testing DRAM must be made in advance whenever a new scheme is implemented. For example, HBM replaced the conventional DRAM's metal PAD and PCB wire with u-bump and silicon interposer, respectively. As a result, the DRAM testing process required significant changes, and DRAM companies proposed various solutions to address these challenges [4]. Similarly, the introduction of the multi-signaling to the DRAM interface necessitates significant changes. However, the current DRAM testing equipment is designed for NRZ signaling and cannot test multi-signaling. Even if new equipment that supports PAM4 signaling is obtained, it would require a considerable amount of time and expense to replace existing NRZ equipment. To overcome this challenge, a bridge chip capable of testing PAM4 DRAM with ex-isting NRZ equipment has been proposed [5], [6].

2.2 Basis of DRAM Interface
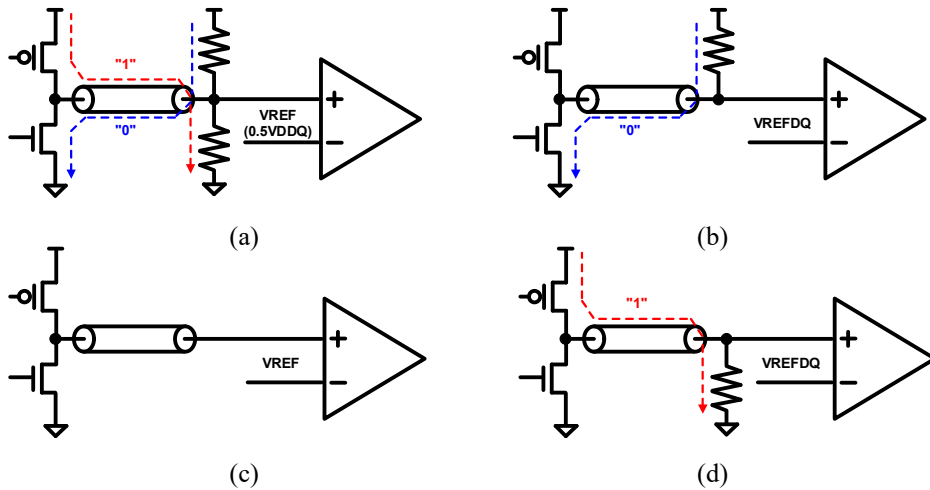
Fig. 2.3 Single-ended interface of memory (a) SSTL, (b) POD, (c) HSUL, (d) LVSTL.

The modern DRAM interface uses asymmetric termination to decrease C_{io}, increasing bandwidth and reducing operating power. Fig. 2.3(a) illustrates the SSTL method used before DDR3, and the center tap method was utilized for termination. The SSTL method involves turning on the N/PMOS and terminating it with 0.5VDDQ. However, it has the drawback of a large C_{io} and DC flowing during termination [7]. To overcome these issues, the POD structure was chosen for DDR4 and GDDR6, as shown in Fig. 2.3(b). The POD structure eliminates the driver's PMOS transistor, which reduces C_{io} and prevents current from flowing during termination. Additionally, the termination method for Mobile DRAM has been modified. Fig. 2.3(c) shows that LPDDR3 employed HSUL, which provided benefits in power consumption and a larger swing at low speed (without termination). However, with increasing required bandwidth, LPDDR4 requires the termination and adopts LVSTL, as shown in Fig.

2.3(d). The reason for this is that LVSTL reduces power consumption, unlike POD. Furthermore, LVSTL terminates with NMOS only, which is advantageous for easy VDD scaling due to low termination. In LPDDR4x, the advantages of LVSTL are further amplified. LPDDR4X uses an N/N driver instead of a CMOS driver to eliminate a relatively large PMOS, thereby reducing the driver's C_{io}. Also, by utilizing NMOS for pull-up, the interface voltage can be substantially reduced. Despite these benefits, LVSTL has the drawback that the receiver must be set up as a PMOS. Therefore, a POD structure for DDR5 utilizes a multi-drop method to connect multiple DRAMs to a relatively long channel. Meanwhile, LPDDR5, which prioritizes reducing power consumption, employs an LVSTL structure that enables VDDQ scaling. For future DRAM developments, it is anticipated that either POD or LVSTL will be chosen based on the channel situation.

The transmitter's final stage includes a driver responsible for driving the transmission line. Two types of drivers can be used: current mode and voltage mode. As illustrated in Fig. 2.4(a), the current mode driver produces an output signal with high impedance and is shunt-connected with a resistor for source termination. However, it consumes more than twice the voltage mode driver's current and is unsuitable for memory applications that require single-ended signaling. On the other hand, the voltage mode driver shown in Fig. 2.4(b) is commonly used in DRAM applications. It produces an output signal with low impedance and has a source termination that is

(a)   (b)
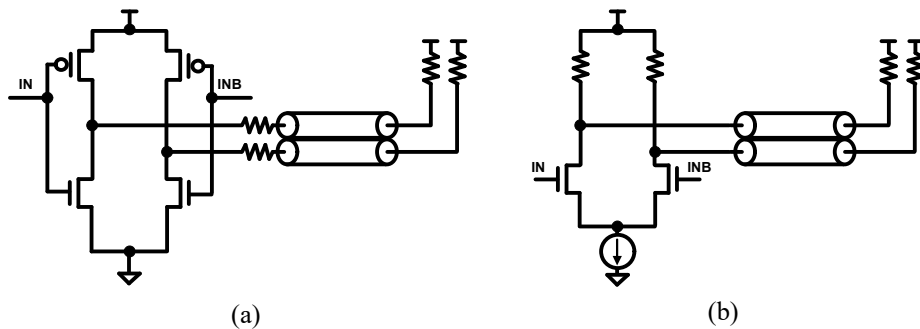
Fig. 2.4 (a) Voltage Mode Driver, (b) Current Mode Driver.

series-connected with a resistor. Voltage mode drivers typically consist of transistors and series resistors. As the transistor operates in linear mode and has a constant R_{on}, the output impedance becomes $R_{on}+R_s$, which requires impedance matching with $Z_0=R_{on}+R_s$. Since R_{on} is sensitive to PVT changes, adjusting R_{on} to a desired value is necessary.

Since DDR3, ZQ calibration is used to adjust the Ron value. The ZQ calibration block diagram for DRAM is illustrated in Fig. 2.5. For high termination, an external reference resistor (Z_0) is connected between the VSS and ZQ pins to match the impedance of the pull-up driver. When ZQ calibration is initiated, the strength of the pull-up driver is adjusted to make the ZQ pin half of VDDQ, and the result is stored in the register (PCODE). After the pull-up driver has been calibrated, the pull-down driver's impedance is calibrated. For this, a reference resistor (Z_0) is connected between the VDD and ZQ pins, and the reference resistor is replaced with an already calibrated pull-up driver. The impedance of the pull-down driver is adjusted to make the node value between the pull-up and pull-down drivers VDDQ/2, which is stored in the register (NCODE). These stored values are applied to all DQs to match the impedance that is sensitive to PVT. To prevent impedance change during operation

due to voltage/temperature changes, the DRAM controller performs calibration periodically by issuing the ZQCS command. (ZQCL is a command used for precise calibration during DRAM initialization, while ZQCS is used for short-term calibration during operation.)

The impedance of DRAM in a PCB board environment can vary due to factors like the permittivity of the board, transmission line structure/width/thickness, and stack. To address this, DRAM is designed with multiple impedance values. For example, as seen in Fig. 2.6, the output driver of conventional DRAM has a pull-up and pull-down driver with 7 legs calibrated to 240 Ω. This configuration enables the DRAM to provide termination impedances of 240, 120, 80, 60, 48, 40, and 34 Ω and output driver impedances of 48, 40, and 34 Ω. Typically, DRAM is matched with a 40 Ω impedance.
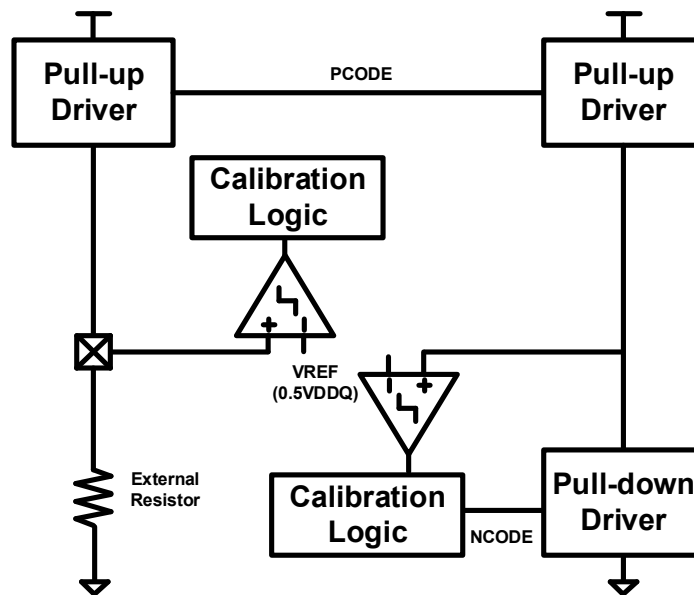
Fig. 2.5 Output driver calibration scheme on DRAM.
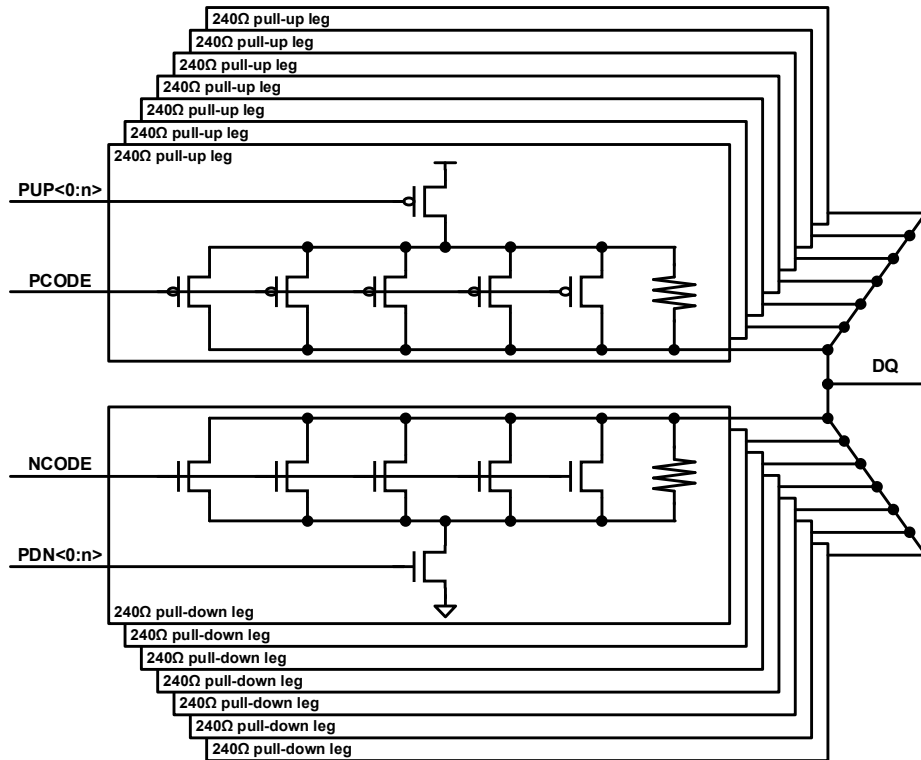
Fig. 2.6 The structure of the output driver of DRAM.

2.3 Architecture in High-speed Interface

2.3.1 Serial Link

Data transmission in a high-speed serial interface involves passing them from a transmitter (TX) through a channel and receiving them at a receiver (RX). A simplified block diagram of a serial link is illustrated in Fig. 2.7. The serialization of data is performed by a serializer (SER), and their deserialization by a deserializer (DES) enables serial communication. The channel used in practice consists of various components, including printed circuit boards (PCBs), cables, connectors, vias, and backplanes. Due to skin effect and dielectric loss, these components exhibit frequency-dependent loss, particularly at high frequencies. Therefore, an equalizer (EQ) is employed to compensate for the channel's low-pass filtering characteristic and expand bandwidth.
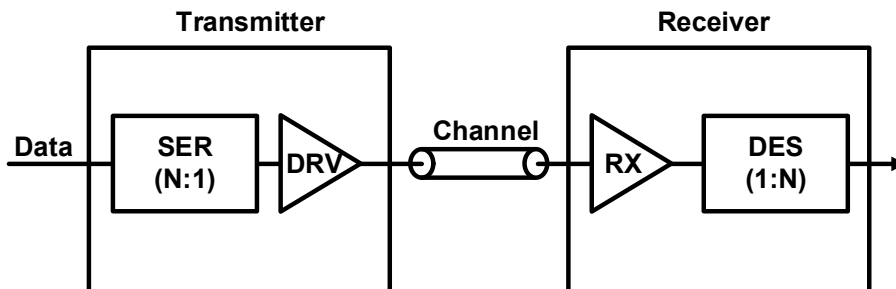
Fig. 2.7 Simplified block diagram of a serial link.

(a)

(b)
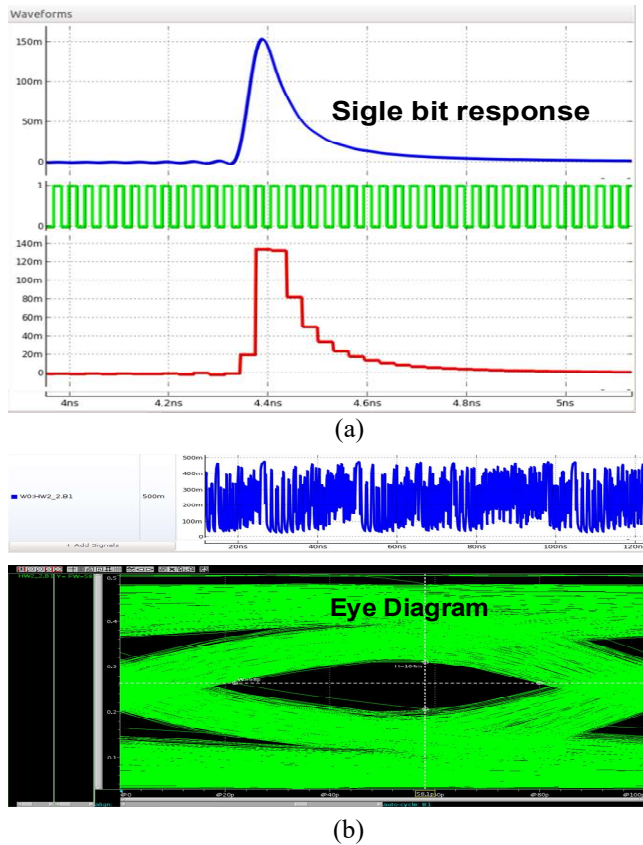
Fig. 2.8 (a) Single bit response; (b) Degraded NRZ eye diagram with ISI.

The frequency response of a channel can be expressed as a time-domain response, and the Single-bit Response (SBR) provides an effective means of visualizing the channel response. The SBR of the channel, denoted as sbr(t), is represented as:

$$sbr(t) = h(t) * \phi(t) \tag{2.1}$$

where h(t) is the impulse response of channel, and $\phi(t)$ is the transmitted single-bit pulse. The continuous-time signal sbr(t) is sampled at the receiver, resulting in a discrete-time representation, such as:

$$sbr[n] = sbr(T_0 + nT_b) \tag{2.2}$$

where T_0 is the sampling time of a main cursor and T_b is a bit period. The value sbr[0], or h_0, is called the main cursor. The values sbr[n] or h_n correspond to the pre-cursor and post-cursor for negative and positive values of n, respectively. Under the assumption that the channel is a Linear Time-Invariant (LTI) system, the received signal at the RX side is expressed as a superposition of SBR separated by a bit-period. The received signal y(t) can be represented as:

$$y(t) = \sum_{k=-\infty}^{\infty} x[n-k] \cdot sbr(t + kT_b) \tag{2.3}$$

where x[n] is the transmitted signal. x[n] can be +1 or -1 in PAM2 or NRZ signaling. As in equation (2.3), sampled received signal y[n] is written as,

$$
\begin{aligned}
y[n] = y(T_0 + nT_b) &= \sum_{k=-\infty}^{\infty} x[n-k] \cdot sbr[k] \\
&= x[n] \cdot sbr[0] + \sum_{k \neq 0} x[n-k] \cdot sbr[k] \\
&= x[n] \cdot h_0 + \sum_{k \neq 0} x[n-k] \cdot h_k
\end{aligned} \tag{2.4}
$$

Equation (2.4) illustrates that y[n] is composed of two terms: the desired signal and the unwanted deterministic dispersion, which is also known as inter-symbol interference (ISI). Fig. 2.8 provides an example of SBR and the corresponding degraded Non-Return-to-Zero (NRZ) eye diagram with ISI. Therefore, equalizers aim to eliminate

the ISI and completely recover the transmitted signal to achieve the target Bit Error Rate (BER), usually less than 10^{-12} for NRZ and 10^{-9} for PAM4. With a reasonable design and in the absence of noise, it is known that the NRZ eye and PAM4 eye start to close at approximately 10 dB and 4.5 dB, respectively.

2.3.2 Multi-level Pulse-amplitude Modulation

The utilization of PAM4 signaling is commonly applied to enhance pin speed. Nonetheless, it has yet to be implemented in DRAM. Since DRAM uses single-end signaling, both PAM3 and PAM4 are being studied. A comparison of the NRZ signal, PAM4 signal, and PAM3 signal is displayed in Fig. 2.9. The order of bandwidth increase at the same Nyquist frequency is PAM4, PAM3, and NRZ. However, the voltage margin decreases. Hence, choosing a suitable method for the channel environment in which DRAM is utilized is crucial. This thesis proposes multilevel signaling as a potential solution to overcome bandwidth limitations due to constraints imposed by channel bandwidth and process integrity.

(a)   (b)   (c)
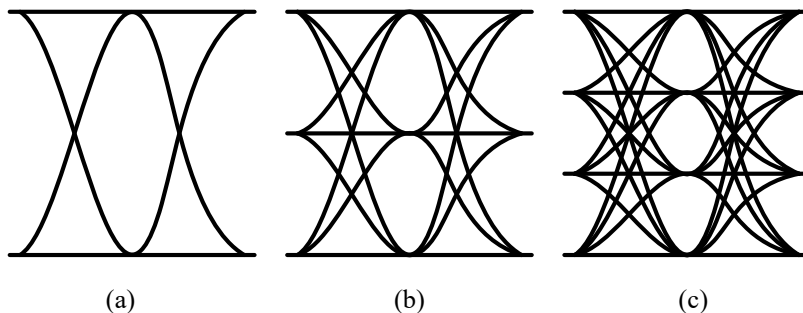
Fig. 2.9 Basic eye diagrams of (a) NRZ, (b) PAM3 and (c) PAM4 signal.
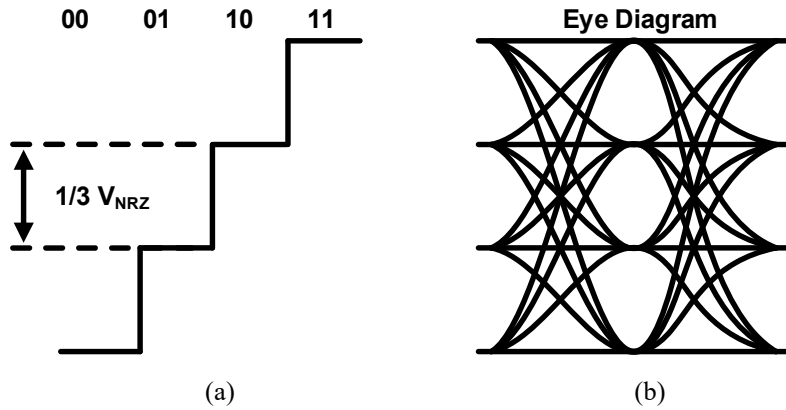
Fig. 2.10 (a) Binary encoded PAM4 signal; (b) PAM4 eye diagram.

PAM4 is an encoding technique that compresses two data bits into one symbol, effectively doubling the number of data bits transmitted within the same time unit interval (UI). In a channel with a Nyquist 12 GHz bandwidth, the PAM4 can achieve a data rate of 48 Gb/s or 24-Gsymbol/s, while NRZ can only transfer data at 24 Gb/s. The eye diagram of PAM4 signaling with binary coding is depicted in Fig. 2.10. Within a single PAM4 symbol, the first bit is referred to as the most significant bit (MSB) and the second bit as the least significant bit (LSB). The LSB in equally-spaced four-level data has a half voltage swing compared to the MSB, equation (2.4) is converted to as,

$$y[n] = x[n] \cdot \frac{h_0}{3} + \sum_{k \neq 0} x[n-k] \cdot \frac{h_k}{3} \tag{2.5}$$

PAM4 uses four amplitude levels, namely, +3, +1, -1, or -3, for signal transmission. These amplitude levels can be represented using different coding techniques such as

binary coding ({00, 01, 10, 11}), thermometer coding ({000, 001, 011, 111}), or integer coding ({0, 1, 2, 3}). Since the vertical eye opening in PAM4 is one-third that of NRZ, the signal-to-noise ratio (SNR) loss of the PAM4 format compared to NRZ format is expressed as,

$$20 \cdot log\left(\frac{1}{3}\right) \sim -9.5 \ dB \tag{2.6}$$

Furthermore, nonlinearity has a greater impact on SNR degradation in PAM4 than NRZ, as previously reported [8]. The linearity of the data eye is a crucial indicator of signal integrity in the PAM4 method, where the smallest eye is an important performance measure. To evaluate the linearity characteristics of PAM4 signals, a criterion known as a RLM is used. The RLM can be obtained from the PAM4 signal by measuring the four output levels (V_A~V_D), determining the eye height through the measured output level, and finding the minimum signal level (S_{min}) among them in equation (2.7). The RLM is defined as shown in equation (2.8). Generally, PAM4 transmitters must satisfy RLM > 0.92 to ensure acceptable signal integrity [9].

$$S_{min} = \frac{1}{2}\min(V_B - V_A, V_C - V_B, V_D - V_C) \tag{2.7}$$

$$R_{LM} = \frac{6S_{min}}{V_D - V_A} \tag{2.8}$$
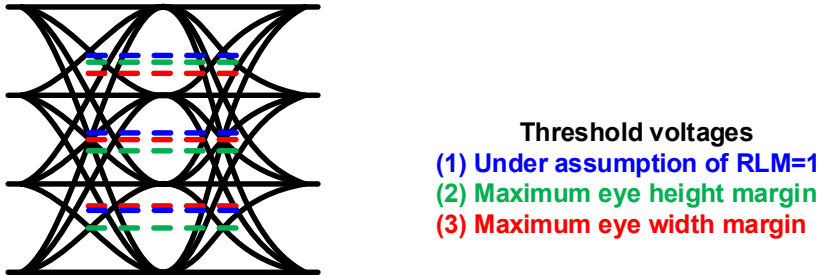
Fig. 2.11 Issues on deciding PAM4 threshold voltages.

The bottleneck of the system margin lies in the worst eye, and a lower RLM directly results in degraded BER. In addition, the threshold voltage is another factor that affects the horizontal eye width and BER. Among the four voltage levels in PAM4 signaling, three threshold voltages bisect the adjacent voltage levels. Deciding on unequally spaced threshold voltages is a concern when the three vertical eyes are not symmetric. Furthermore, the transitions between non-adjacent signal levels take longer time than those between adjacent levels, which narrows the eye further. The issue described above is illustrated in Fig. 2.11.

Various design techniques have been proposed to address the inferior BER characteristic of PAM4 signaling. Gray coding is one technique in which consecutive data levels are encoded to {00, 01, 11, 10} instead of {00, 01, 10, 11} as in binary coding when two bits are mapped to one symbol. This method reduces BER by 33% compared with binary coding, resulting in only a one-bit error per symbol for incorrect decisions. Furthermore, dual mode with NRZ signaling is supported by grounding LSB. Additionally, to improve the RLM performance in TX, several compensation schemes have been proposed so far [10].

2.3.3 PAM4 in DRAM Interface

The use of PAM4 in single-ended signaling for memory applications presents several sources of nonlinearity. One of these sources is the inherent nonlinearity that occurs in the transmitter output, which depends on the architecture of the TX driver. A current-mode driver with a current-mode logic (CML) structure exhibits higher linearity in signaling but consumes twice as much power as necessary. To achieve pin efficiency, memory interface commonly uses a voltage-mode (VM) driver or a source-series terminated (SST) driver. However, the VM driver shows inferior linearity due to impedance variation in the pull-up and pull-down drivers and the mismatched impedance of the channel. In cases where 50 Ω termination is matched through transistor size instead of a passive R, as in recent LPDDR [11], nonlinearity from termination mismatch is increased. Another source of nonlinearity arises from active blocks in the receiver. The data path from the analog front end to samplers consists of equalizers, which will be discussed in the next section. Asymmetric and nonlinear transimpedance or frequency-dependent gain can occur due to mismatch or offset in transistors and segregated paths per clock phase or data level. PAM4 requires a more dynamic voltage range [12], which can cause transistors to experience increased distortion in transimpedance or bandwidth, particularly within single-ended signaling.

The basic operation of the bridge chip transmitter is to convert multiple NRZ signals from the memory tester to PAM4 signals and transmit them to the DRAM. In DRAM testing, both the transmitter and receiver of the DRAM must be evaluated. However, the PAM4 receiver of the DRAM cannot be assessed with an NRZ tester, necessitating the utilization of the bridge chip. The DRAM receiver's characteristics
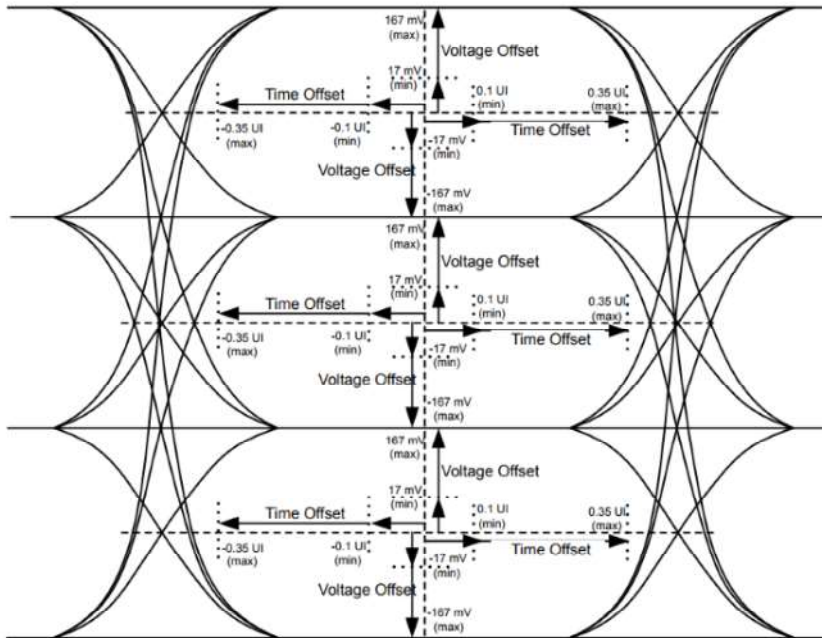
Fig. 2.12 Lane margining in PCIe 6.0

can be determined by varying the timing and voltage margins of the PAM4 signal. As shown in Fig. 2.12, lane margining is an example of this function in PCIe 6.0 [13], which determines the input data margin by measuring several times within the voltage and timing offset at the receiver. To confirm the efficacy of this function, the tester must include the capability to adjust the voltage and timing. The timing margin of the DRAM receiver can be evaluated by adjusting the clock applied to the transmitter of the bridge chip to change the valid data window. In contrast, the voltage margin can be assessed by modifying the output voltage level of the bridge chip transmitter.

2.3.4 Equalizer

Equalizers can be classified based on their location or characteristics, with the possibility of performing equalization at either the transmitter or receiver side. Transmitter-side equalizers commonly include feed-forward equalizer (FFE), while receiver-side options consist of CTLEs and DFE. In addition, these equalizers can be grouped into linear or non-linear types, with DFE as the typical non-linear equalizer and FFE and CTLEs as the linear equalizer representatives.

2.3.3.1 Feed-Forward Equalizer (FFE)

The FFE, a pre-emphasis filter, is a linear equalizer typically located at the transmitter side. Although a receiver-side FFE is available, it increases the complexity of design and power consumption due to the implementation of analog delay. The FFE is a finite impulse response (FIR) filter that includes a shift register, weight multiplier, and summer, as shown in Fig. 2.13. The FFE has advantages such as canceling precursors and having a simple implementation on the transmitter side. However, the
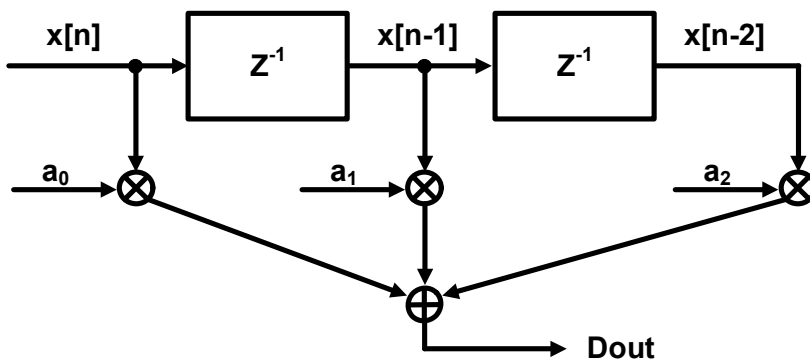
Fig. 2.13 Block diagram of conventional FIR filter.
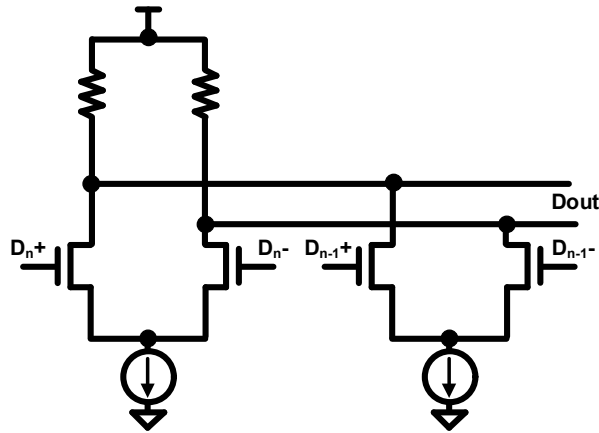
Fig. 2.14 Conventional FFE implemented using CML summer architecture.

FFE attenuates signal amplitude due to a peak swing limit, which is unsatisfactory from the perspective of SNR. Therefore, the FFE de-emphasizes the low-frequency component rather than pre-emphasizing the high-frequency component.

The FFE shown in Fig. 2.14 employs a summer architecture based on current-mode logic (CML) and uses tail currents in each branch to control the tap weights. Input transistors of each branch are controlled by symbols shifted by the unit interval (UI), and the resulting output is summed using current form.

In addition to the CML-type driver, FFE can be utilized with voltage mode or low voltage differential signaling (LVDS) drivers. A voltage-mode source-series-terminated (SST) driver utilizing FFE is illustrated in Fig. 2.15 [14]. The number of slice units establishes the tap weight. Therefore, the de-emphasis level is also determined.

$$Lev_{de-emp}[dB] = 20log_{10}(\frac{m - p}{m + p}) \qquad (2.7)$$

where m and p are the number of slice units driven by main tap and post-cursor
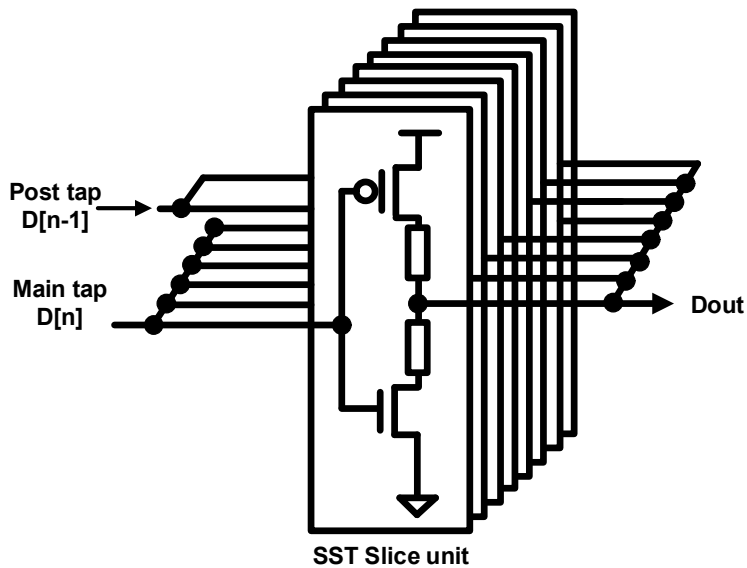
Fig. 2.15 FFE implemented voltage mode source-series-terminated (SST) driver.

tap, respectively.

2.3.3.2 Continuous-Time Linear Equalizer (CTLE)

The CTLE is a type of linear equalizer typically utilized in RX circuits. Its purpose is to compensate for frequency-dependent channel loss using a high-pass filtering transfer function. By ensuring that the transfer function of an equalizer is the inverse of the channel frequency response, multiplied functions result in a flat frequency response. The objective of a practical equalizer is to enhance the signal around the Nyquist frequency, enabling the overall frequency response to be flat up to the Nyquist frequency. The RC-degenerated CTLE is a traditional active linear equalizer, and its fundamental structure and frequency response are depicted in Fig. 2.16. To obtain the
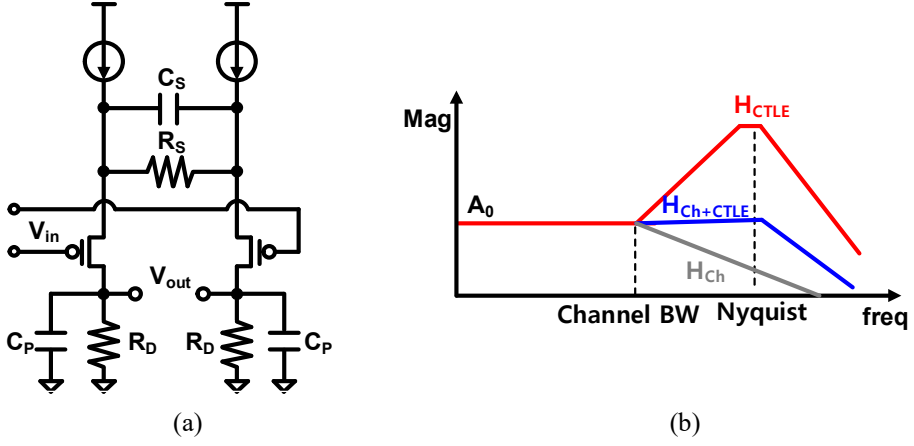
Fig. 2.16 (a) Circuit and (b) frequency response of CTLE.

transfer function, the following steps are followed:

$$H(s) = \frac{g_m R_D}{1 + \frac{g_m R_S}{2}} \frac{(1 + \frac{s}{\omega_z})}{(1 + \frac{s}{\omega_{p1}})(1 + \frac{s}{\omega_{p2}})} \tag{2.8}$$

$$where \ \ \omega_z = \frac{1}{R_S C_S}, \omega_{p1} = \frac{1 + \frac{g_m R_S}{2}}{R_S C_S}, \omega_{p2} = \frac{1}{R_D C_P} \tag{2.9}$$

The transconductance of the input transistor is referred to as g_m, while source degeneration resistance and capacitance are denoted by R_S and C_S, respectively. The circuit produces a maximum boost factor of $1 + \frac{g_m R_S}{2}$ around the Nyquist frequency when ω_z and ω_{p1} are set to $\omega_{channel}$ and $\omega_{Nyquist}$, respectively. Overboosting via the use of the CTLE leads to noise amplification, hence a decrease in dc gain is often applied to adjust the boost factor to various channel losses. Digital control is usually employed to adjust the placement of poles and zero, as well as the DC gain, by manipulating R_S and C_S.

There are two main factors to consider in [15]. Firstly, the boost factor $1 + \frac{g_m R_S}{2}$ trades off with the dc gain $\frac{g_m R_D}{1 + \frac{g_m R_S}{2}}$. To maintain the DC gain at unity and prevent excessive attenuation of received data, the boosting gain must be limited. Secondly, when the output pole ω_{p2} occurs below the degeneration pole ω_{p1} at high speed, the circuit bandwidth is limited. In this scenario, the maximum boost gain cannot be achieved as ω_{p1} and ω_{p2} approach each other in the frequency domain. To increase the boost gain, C_S can be increased, which will increase the distance between poles, but this will result in a lower peaking frequency. Fig. 2.17 shows how adjusting R_S and C_S can modify the DC gain and zero location when $\omega_{p1} > \omega_{p2}$.

Various design techniques have been introduced to overcome the issues above. To
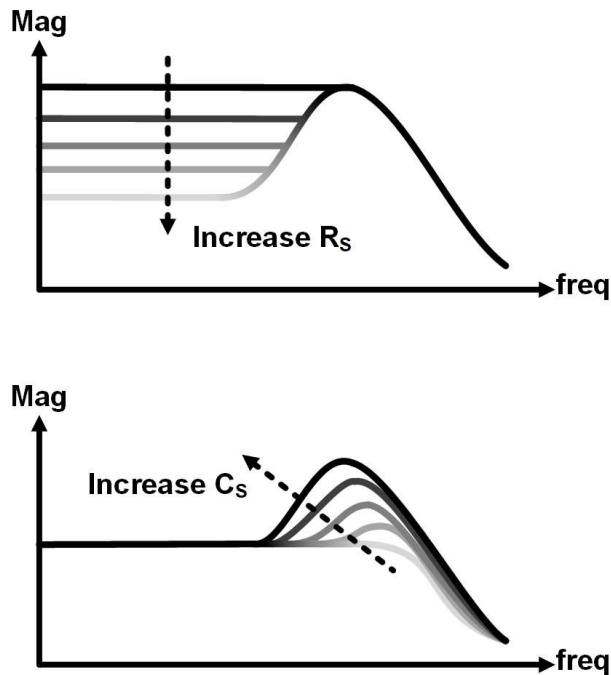
Fig. 2.17 DC gain and zero location adjustments of RC-degenerated CTLE.
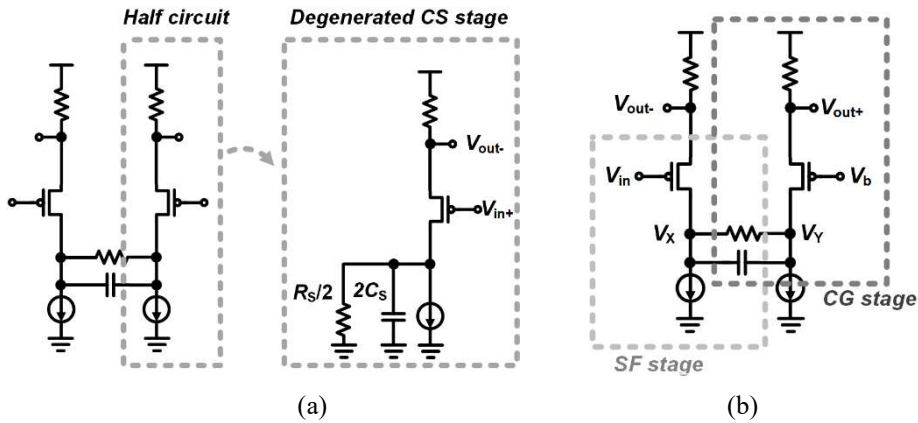
Fig. 2.18 Interpretation of (a) differential input and (b) single-ended input.

improve peaking, multiple CTLE stages can be cascaded, albeit at the expense of decreased bandwidth and higher power consumption. Alternatively, an inductive peaking technique is frequently utilized for greater bandwidth. This method widens the bandwidth without increasing power consumption by incorporating an inductor in series with R_D. However, due to the bulky size of the inductor in the integrated chip, there is a significant area penalty.

A differential CTLE is commonly used in single-ended signaling, such as in memory interfaces, by applying a single-ended signal to one input transistor and a dc common-mode or reference voltage to the other. Fig. 2.18 highlights the difference between sensing differential input and single-ended input in a differential CTLE. Theoretically, the circuit converts a single input to a fully differential output. When a differential input is applied, each half circuit functions as a degenerated common-source stage with an internal virtual ground and $\frac{R_S}{2}, 2C_S$ for each side. However, if a single-ended signal is used (with V_{in} as the input and V_b as the dc voltage), the circuit is no longer interpreted as a half circuit. Instead, V_{in} becomes a degenerated common-

mode stage with parallel R_S, C_S in series with $\frac{1}{g_m}$. To obtain $\frac{V_{out}}{V_{in}}$, the circuit can be broken down into a series of source-follower and common-gate stages. This yields $\frac{V_{out+}}{V_{in}} = \frac{V_b}{V_{in}}$. Under the assumption of an ideal current source and no channel-length modulation, the resulting transfer function $H(s) = \frac{V_{out+} - V_{out-}}{V_{in}}$ is identical to equation (2.11), with $\frac{V_{out+}}{V_{in}}$, $\frac{V_{out-}}{V_{in}}$, and $\frac{V_Y}{V_{in}}$ calculated as follows.

$$\frac{V_{out+}}{V_{in}} = -\frac{V_{out-}}{V_{in}} = \frac{-\frac{g_m R_D}{2}}{1 + \frac{g_m R_S}{2}} \frac{(1 + \frac{s}{\omega_z})}{(1 + \frac{s}{\omega_{p1}})(1 + \frac{s}{\omega_{p2}})} \tag{2.10}$$

$$\frac{V_Y}{V_{in}} \cong \frac{g_m \cdot \frac{1}{g_m}}{1 + g_m \left(\frac{R_S}{1 + s R_S C_S} + \frac{1}{g_m}\right)} \tag{2.11}$$

$$where \;\; \omega_z = \frac{1}{R_S C_S}, \omega_{p1} = \frac{1 + \frac{g_m R_S}{2}}{R_S C_S}, \omega_{p2} = \frac{1}{R_D C_P} \tag{2.12}$$

2.3.3.3 Decision-Feedback Equalizer (DFE)

The DFE, used on the receiver side, is a nonlinear equalizer, as shown in Fig. 2.19. Its main functional blocks include a decision block, feedback filter, and summer. The decision block, a sampler or comparator, determines whether the received data is a logical 0 or 1. This nonlinear behavior is what makes the DFE nonlinear. The sampled data is delayed, multiplied by a tap coefficient, passed through the feedback filter, and subtracted at the summer. In an n-tap DFE, sampled data is delayed for n UI, and multiplied by the nth tap coefficient w_n. Setting w_n to a post-cursor value h_n can
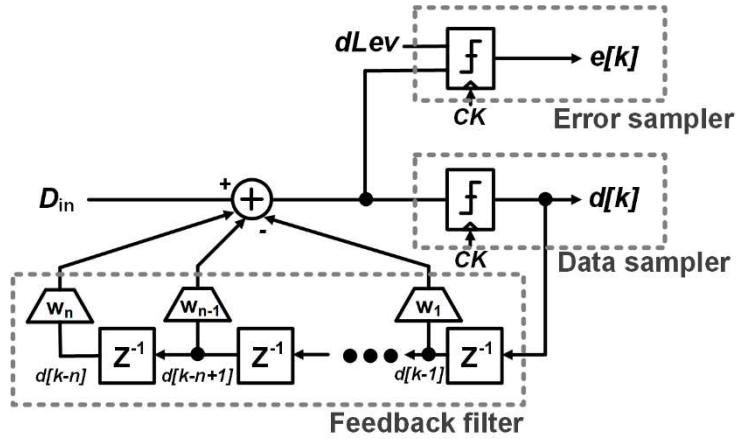
Fig. 2.19 Block diagram of an n-tap DFE.

cancel corresponding post-cursor ISIs from the received signal, as shown in Fig. 2.20. Unlike the CTLE, which enhances a main cursor value and removes long-tail post-cursors, DFE's goal is to eliminate residual post-cursors. Also, DFE does not cause noise amplification through feedback of a hard-decision value. However, it can propagate errors if a previous decision is incorrect.

One critical design challenge of the DFE is the timing constraint of the feedback loop. The sampler must have sufficient setup time, deliver the result after a C-to-Q delay, and settle at the summing node before the next data is sampled. This limitation can be expressed as:

$$t_{c2q} + t_{setup} + t_{settle} < 1 \ UI \tag{2.13}$$

where t_{c2q}, t_{setup} and t_{settle} are C-to-Q delay, a setup time of the sampler, and a settling time at the summing node, respectively. Loop-unrolling DFE, or speculative DFE can be adopted to alleviate this stringent timing constraint.
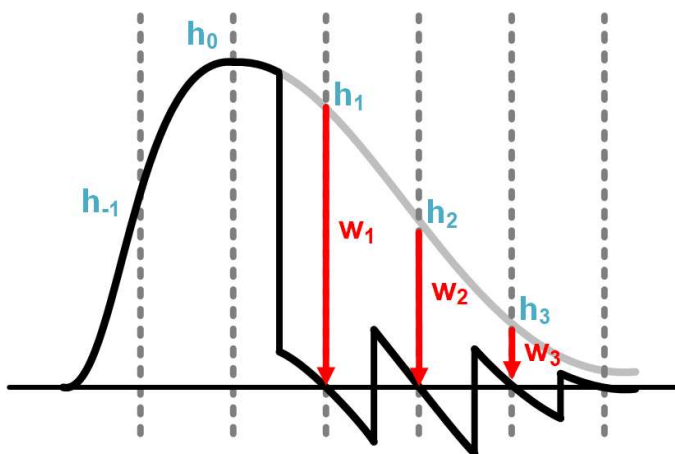
Fig. 2.21 Operation of DFE with single-bit response (SBR).
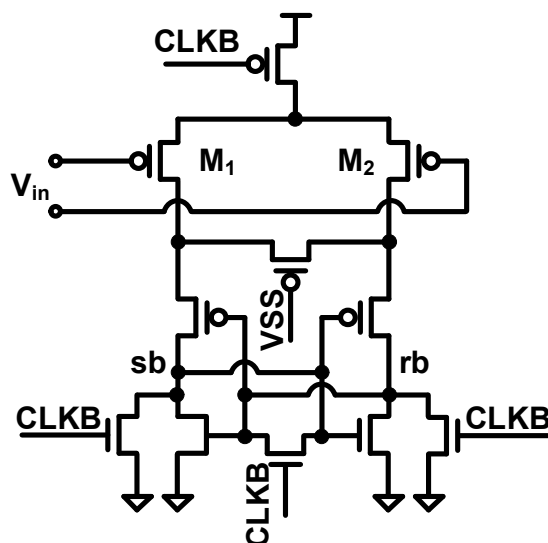
Fig. 2.20 Structure of a StrongArm latch.

The StrongArm latch is the most commonly used sampler topology for three reasons [16]: it consumes no static power, produces rail-to-rail outputs directly, and has an input-referred offset that mainly results from one differential pair, as shown in Fig.

2.21. The hard decision output provides convenience for DFE implementation. Additional input pairs can be utilized to compensate for the input-referred offset or compare the differential input with a differential reference voltage. However, the kickback noise arising from clocked current sources becomes troublesome, even though the differential architecture cancels it out to some extent. Therefore, the bridge design in the next section incorporates the offset removal technique by employing a PMOS input RC-degenerated active linear equalizer (CTLE) with a differential input latch to achieve optimal RX operation.

Chapter 3

Design of 32 Gb/s PAM4-Binary Bridge with Sampler Offset Cancellation for Memory Testing

3.1 Overview

Multi-level signaling interface applicable to memory is under development by major memory manufacturers. However, existing test solutions of using Automatic Test Equipment/System Level Test (ATE/SLT) offer only a low-speed binary mode lacking a multi-level signaling capability, thus efficient evaluation methods for PAM4 signals must be explored. For example, as shown in Fig. 1.3, in characterizing a DRAM interface, a T5511 (ADVANTEST) tester is popularly used to test and verify the internal and external interface operation, but only a binary mode is supported, and

its maximum clock speed is 4 GHz, thereby limiting the testable data rate within 8 Gb/s. Moreover, a new test equipment for PAM4 signaling is not expected to be available soon since the increased interface speed inevitably incurs larger signal attenuation over the lengthy test cable. A PAM4-Binary Bridge that serves multiple functions such as a translator, an equalizer, and a retimer between a popularly used low-speed binary tester and a high-speed PAM4 memory will play a key role in testing such newly developed memory interfaces in a timely manner. Therefore, the optimum structure of the bridge with all the necessary functions required between the tester and the next-generation DRAM under test to ease the speed requirement is proposed. In view of these, this brief describes the design of a bridge for memory testing that is optimized to work with PAM4-based DRAMs with 32 Gb/s per-pin data-rate at 1.25 V interface supply.

3.2 PAM4-Binary Bridge

3.2.1 Architecture

Fig. 3.1 shows the proposed PAM4-Binary Bridge that allows the tester supporting only an NRZ mode to evaluate the memory with PAM4 signaling. In the proposed bridge, a WCK signaling scheme used in the GDDR5/6 interface is employed [17]. The frequency of the WCK used in the bridge is the same as the WCK of the memory side and thus, it is twice as high as the WCK of the NRZ tester. An all-digital PLL (ADPLL) integrated in the bridge with a small area and good process-voltage-temperature (PVT) tolerance doubles the incoming WCK frequency. The internal WCK, the ADPLL output, provides a timing reference for data write/read operations after being phase-adjusted by a phase interpolator (PI) and a duty cycle corrector (DCC). The PI is integrated for bridge self-evaluation as well as for performing data-sensing timing scans in the Eye-Opening Monitoring (EOM) mode. For the write operation with PAM4 signaling, the bridge first receives four NRZ data supporting VSS termination from the tester. The half-rate data are aligned to the internal WCK and the encoder converts the NRZ to a Gray-coded PAM4 [18]. The PAM4 data are then forwarded to the driver with the 2-tap FFE capability. The driver for PAM4 is ground-referenced (VSS-terminated) and provides the single-ended output by controlling the output current with the PMOS switches. For the read operation, the single-ended PAM4 signal from the memory is fed to three CTLEs with three threshold levels selected to bisect the upper, middle, and lower data levels. The four parallel data sampled by the PI-

adjusted clock are Gray-decoded, deserialized, and transmitted to the NRZ tester. The

proposed bridge determines the optimal voltage and sampling time during start-up in

the EOM mode using a PRBS7 sequence and a zero-and-one counter.
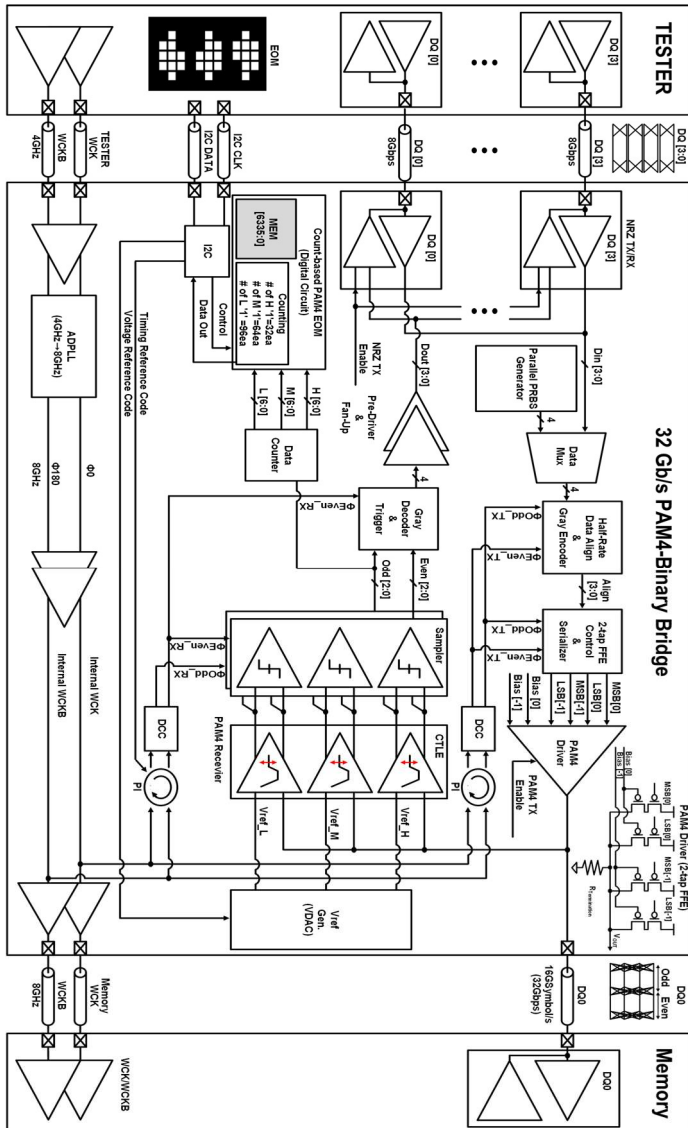
Fig. 3.1 Overall architecture of the proposed PAM4-Binary Bridge.

3.2.2 Training and Normal Operation of PAM4-Binary Bridge
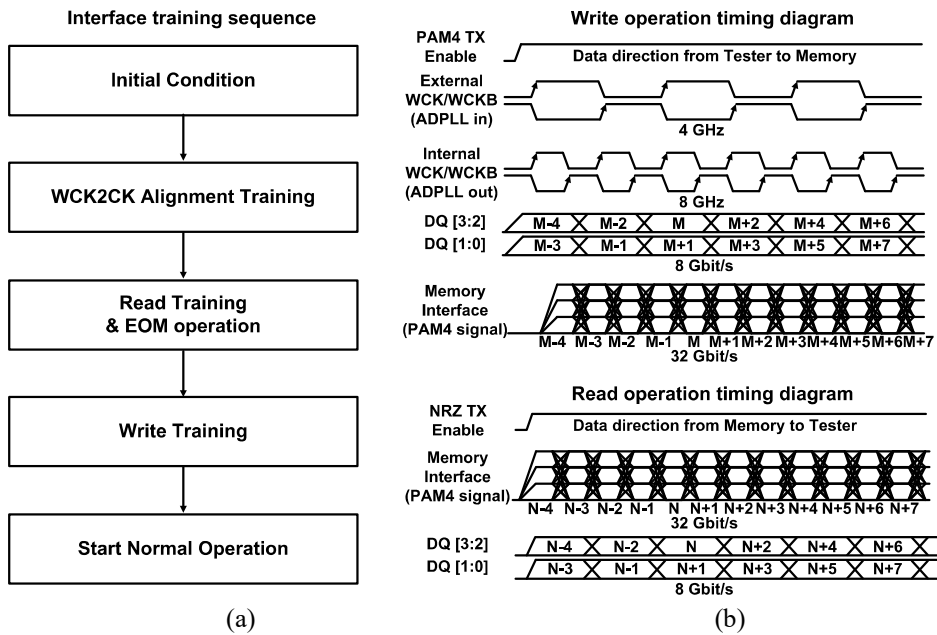
Fig. 3.2 (a) Training sequence and (b) timing diagram of read/write operation.

Fig. 3.2(a) is the bridge's operating flowchart compatible with the training method of the standard GDDR6 [19]. The difference from the GDDR6 WCK2CK training sequence is to compare the WCK between the increased frequency in the PAM4-Binary Bridge and the memory's operating clock. The method of finding the optimal reference voltage levels and timing codes for the VDAC and PI in the reading training mode is added. After the training sequence is finished as shown in Fig. 3.2(b), regular write/read operations are performed on the memory using the PAM4-Binary Bridge according to the timing diagram. The internal WCK, whose frequency is doubled (8

GHz) by the ADPLL, is used for both write/read operations. In the write operation,

four DQ data of 8 Gb/s are combined as the PAM4 signal of 32 Gb/s per pin through

the bridge. When reading memory data, as opposite to the write operation, the single

PAM4 data of 32 Gb/s is converted to four NRZ data of 8 Gb/s each in the bridge.

3.3 Single-ended current mode PAM4 Transmitter

The PAM4 driver is the most important circuit to consider when designing the
PAM4 binary bridge transmitter. The initial step involves deciding whether to design
a voltage mode or current mode driver. Voltage mode drivers are preferable as they
consume less power than current mode drivers. However, they have a disadvantage in

(a)

(b)
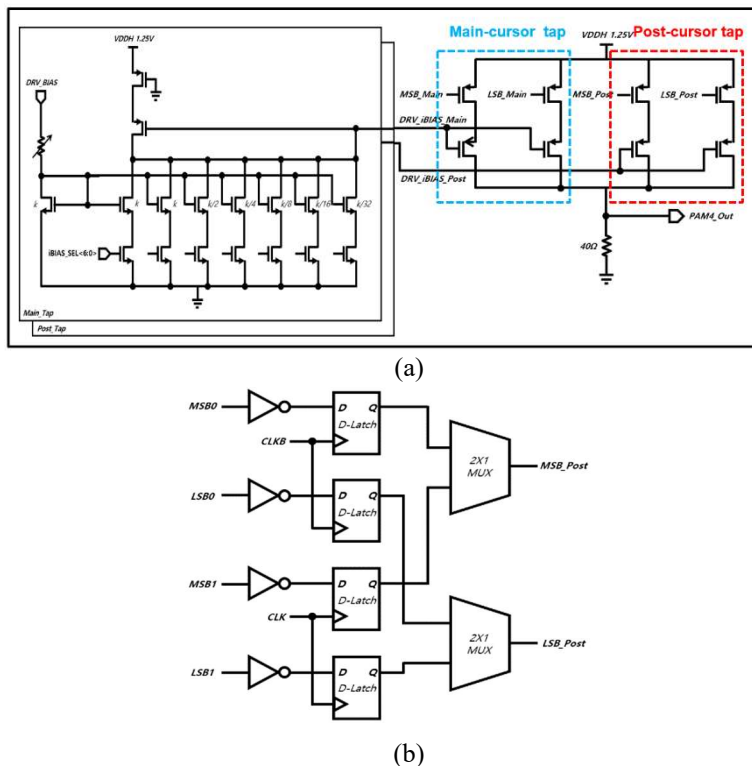
Fig. 3.3 The proposed main driver circuit (a) PAM4 main driver with 2-tap FFE with

Current source circuit, (b) MSB/LSB generator for post-cursor tap.

that matching the output impedance is challenging compared to the current mode
driver. Failure to match the characteristic impedance of the transmission line results
in signal distortion due to the reflected wave. Impedance matching is a critical factor
that directly impacts bandwidth degradation. These points have been discussed in ref-
erences [20]-[24]. The current mode driver has a shorter rise/fall time of the signal
compared to the voltage mode driver, resulting in faster operating speeds. Addition-
ally, impedance matching is relatively easy with parallel termination resistors. Never-
theless, the main disadvantage is that it consumes more power than the voltage mode
driver, as discussed in references [25]-[28].

The current mode driver type was adopted in the transmitter of the first chip. Alt-
hough it consumes more power than the voltage mode driver, it is advantageous in
having a relatively simple impedance-matching structure and a fast operating speed.
Despite the poor SNR characteristics of PAM4 signals compared to NRZ, a pre-em-
phasis circuit for the Feed Forward Equalizer (FFE) function was inserted to compen-
sate for signal distortion. The FFE adds one tab to the main tab for signal compensa-
tion.

It is required to modify the differential pair structure to adapt DRAM to a single-
ended signal method. Furthermore, VSS termination can be achieved by changing to
the PMOS structure instead of the NMOS structure. However, utilizing a multi-slice
structure to regulate the FFE coefficient can generate skew between slices, directly
impacting the main driver's characteristics. To address this issue, a new approach to
adjusting the FFE coefficient was proposed in this thesis, which involves controlling
the driving currents of the main tap and post-cursor tap separately.

The initial step involves inverting the MSB/LSB signal from the prior data, delaying it by 1-UI, and transforming the parallel-aligned signal into a series via a serializer. Fig. 3.3(b) illustrates that the circuit leverages a D-latch and 2x1 MUX. The final driver then receives the main-cursor and post-cursor signals, which have been serialized.

The main driver was intentionally designed in the current mode configuration, which, although it consumes more power, offers the benefits of the faster operation and improved impedance matching. The proposed primary driver circuit is designed with PMOSs, as shown in Fig. 3.3(a), because GDDR DRAM uses VSS termination. It consists of PMOSs that receive driver inputs and resistors. The resistance value was designed as 40 Ω, slightly smaller than the commonly used 50 Ω to take advantage of high-speed characteristics.

Using a multi-slice type FFE tap may result in deteriorated overall I/O characteristics due to skew generation in each slice unit for adjusting the tap coefficient. Therefore, the driver for the post-cursor tap is not designed using this method. Instead, this thesis proposes a design method for the FFE coefficients of the main-cursor and post-cursor taps using a 7-bit thermometer code through I2C to control the current source, as depicted in Fig. 3.3(a).

3.4 Offset cancellation PAM4 Receiver

3.4.1   PAM4 Receiver

Fig. 3.4 shows the PAM4 RX designed for the PAM4-Binary Bridge. The RX consists of three CTLEs and six sampling latches. A DQ input, IN, received from the memory passes through three parallel CTLE stages with three different reference levels. The reference levels, generated in the voltage digital-to-analog converters (VDACs), are selected so that the random offset of the CTLE can be cancelled and the maximum eye height can be obtained. The CTLE output is sampled by two latches
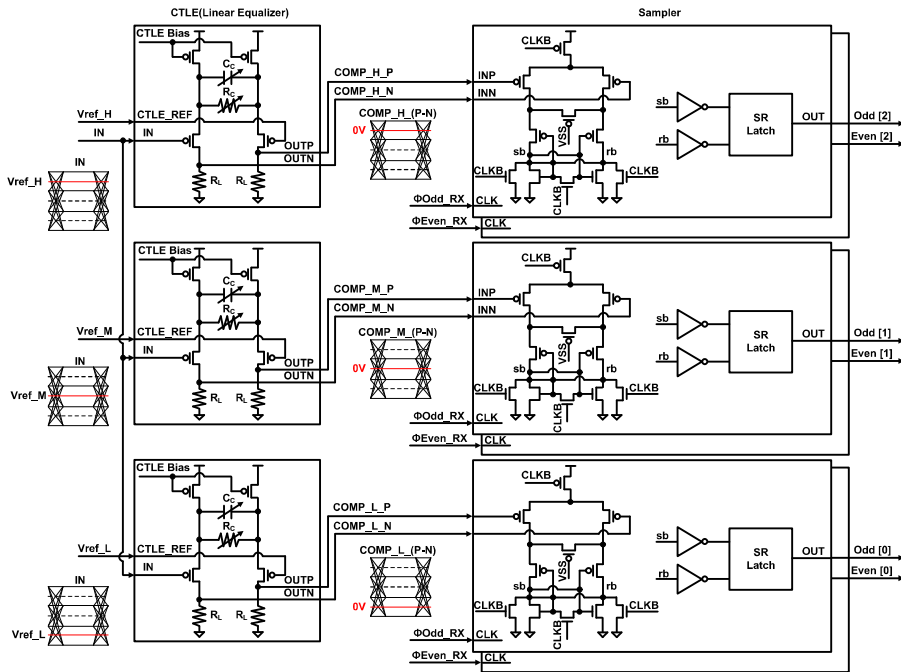
Fig. 3.4 Overall architecture of the proposed PAM4-Binary Bridge.

with half-rate clocks in the RX front-end. The CTLE compares the single-ended

PAM4 input with a reference voltage, amplifies the difference with a linear gain, and

outputs the result in the differential mode (Fig. 3.5). It serves as an offset cancellation

circuit as well as boosting high-frequency signal component for equalization. The ref-

erence levels are adjusted during the training stage based on the EOM measurement

result. The sampling latches operate with even and odd sampling clocks, and they

mark the interface between the analog and digital domains. Since the effective sampler

offset is reduced by the DC gain of the CTLE, the designed samplers do not employ

any offset cancellation circuitry, leading to a simpler RX front-end design.
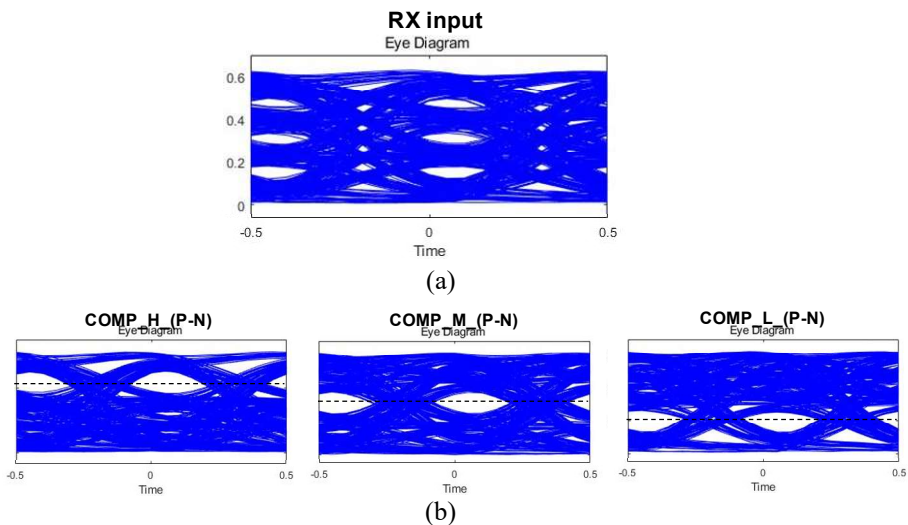
Fig. 3.5 CTLE output signals (Sampler inputs) according to Receiver PAM4 input.

3.4.2 Offset Cancellation Analysis

(a)   (b)
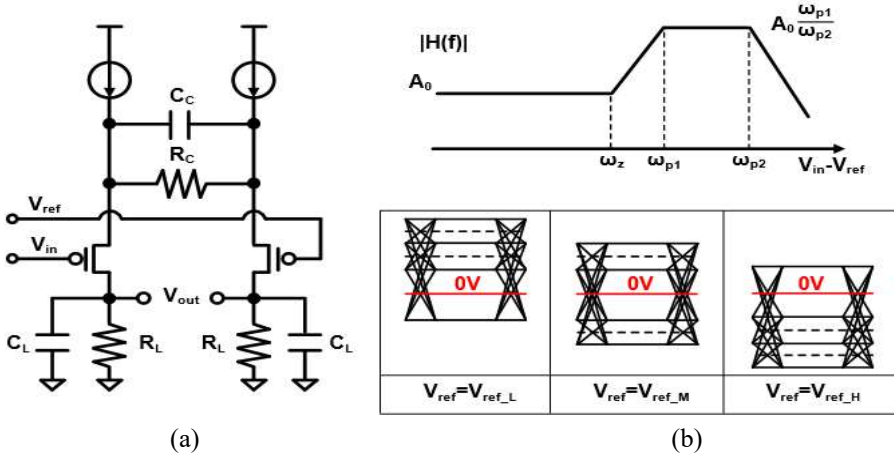
Fig. 3.6 PMOS input RC-degenerated active linear equalizer (a) Circuit imple-
mentation. (b) Frequency response and CTLE differential output(V_{out}) with three ref-
erence input voltages.

A CTLE is a linear equalizer that efficiently boosts high-frequency components
through capacitive degeneration [29]. Fig. 3.6(a) shows the PMOS input RC-degen-
erated active linear equalizer, and its frequency response is depicted in Fig. 3.6(b),
where the DC gain (A_0), zero (ω_z) and pole(ω_{p1}, ω_{p2}) frequencies can be represented
as follow.

$$|\omega_z| = \frac{1}{R_C C_C} \ , |\omega_{p1}| = \frac{1}{R_L C_L}, |\omega_{p2}| = \frac{\left[1 + \frac{(g_m + g_{mb})R_C}{2}\right]}{R_C C_C} \tag{3.1}$$

$$A_0 = \ g_m R_L / [1 + (g_m + g_{mb})R_C/2] \tag{3.2}$$
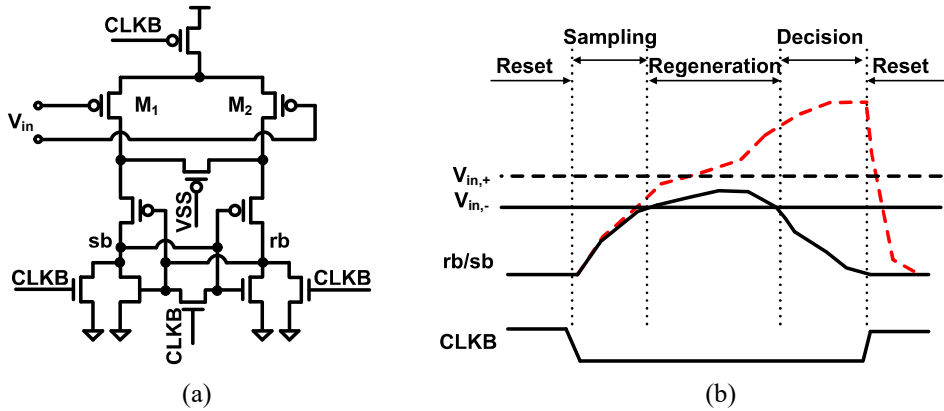
Fig. 3.7 PMOS input sampling latch (a) Circuit implementation (b) Timing diagram. [30].

The designed CTLE achieves 6dB gain at the 8 GHz Nyquist frequency, improving

the incoming 32 Gb/s PAM4 data eye. Fig. 3.6(b) indicates that a CTLE differential

output with three reference input voltages can selectively increase each data eye of

PAM4 by the gain of the CTLE. The sampling latch used for RX is a low-power re-

generative latch that amplifies a static input voltage by converting it into a current,

then integrates that current on a capacitor over a well-defined time window. When

designing the sampling latch, it is necessary to minimize the offset caused by V_{th} and

β mismatch, and layout imbalance appearing in the input pair transistors since the

latch only detects the sign of the (small) input voltage in high-speed operation

(Fig.3.7). The offset is the most critical factor in the sensing phase of Fig. 3.7(b).

Equation (3.3) indicates the offset caused by mismatch of the sampling latch input

[31].
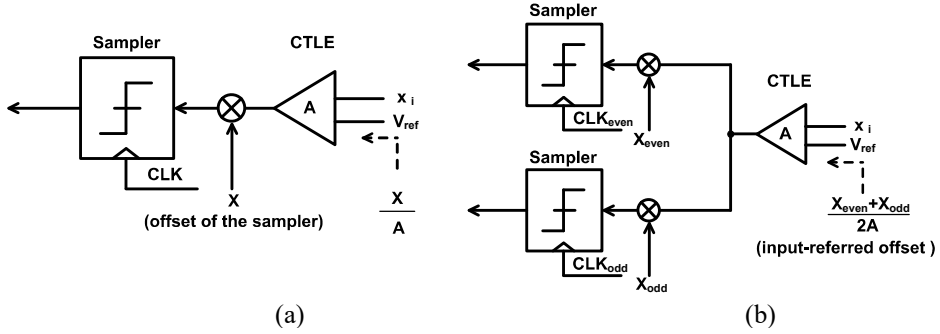
Fig. 3.8 Offset cancellation with (a) one sampler and (b) two samplers with shared CTLE.

$$v_{os}|_{M1,2} = \frac{\Delta I_{M1,2}}{g_{m1,2}} = \Delta V_{t1,2} + \frac{\Delta \beta_{1,2}}{\beta_{1,2}} \cdot \frac{V_{DSAT1,2}}{2} \tag{3.3}$$

To examine the offset of the PMOS input sampling latch of Fig. 3.7, the Monte
Carlo simulation was performed in 40 nm CMOS technology, and the offset evaluated
under the conditions of $V_{DD} = 1.25V$, clock frequency $f_{CLK} = 8$ GHz and the input
common-mode voltage $V_{cm} = 0.2$V was about 9.3 mV$_{rms}$. The conventional method
to reduce the offset of the sampling latch is to add adjustable capacitance to the drain
nodes of the input transistors [32]. However, this method requires additional time for
offset cancellation, and the speed of the sampling latch is reduced due to the increased
capacitance. Moreover, the additional capacitance and its control circuit for calibrat-
ing each sampling latch incur large area overhead.

If the CTLE and sampling latch are connected in series, the CTLE amplifies the
signal around the Nyquist frequency, which is then forwarded to the sampling latch.
Using the training methods that will be presented shortly, the data eye input to the
sampling latch can be maximized by finding an optimal reference voltage applied to

one input of the CTLE. Since the differential input to the sampling latch is amplified
with an enlarged data eye through the CTLE, a sufficiently large differential input
enables high-speed RX operation in the presence of a small sampling latch offset.

Fig. 3.8(a) shows that when one sampler with an offset X and the CTLE with gain
A are connected in series, the input referred offset can be reduced to zero if the refer-
ence level is set to X/A. However, in this design, two samplers with offsets of X_{even}
and X_{odd} operating with even and odd clocks are connected to the common CTLE with
a DC gain of A as shown in Fig. 3.8(b). The reference level of the CTLE is set to the
mid-point of the two sampler offsets during training. Thus, the input-referred offset
in this case is calculated as follows.

$$\sigma\left[\frac{X_{even} - A\frac{X_{even} + X_{odd}}{2A}}{A}\right] = \frac{\sqrt{\sigma_{even}^2 + \sigma_{odd}^2 - 2cov(X_{even}, X_{odd})}}{2A} \tag{3.4}$$

With the simulated offset of the sampler (9.3 mV$_{rms}$), the DC gain of the CTLE (4-
dB), and the estimated covariance (6.2 mV$_{rms}$), the input-referred offset is calculated
as 2.9 mV$_{rms}$. The measured average DC offset of the sampler of the chip was about
2.8 mV$_{rms}$, which corroborates the validity of our assumption.

3.5 Count-Based PAM4 EOM
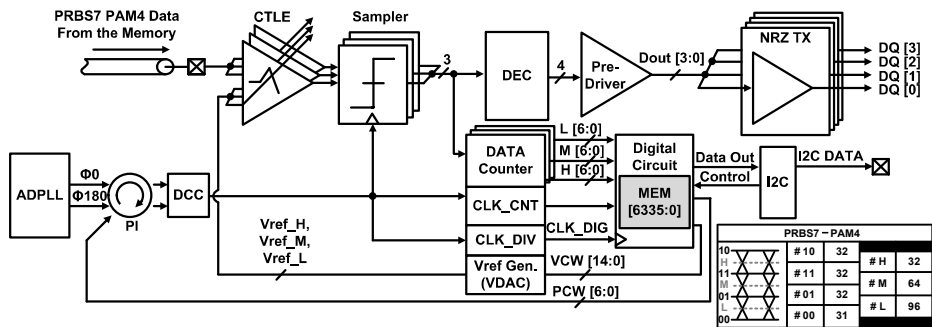
Fig. 3.9 Block diagram and operation of the count-based PAM4 EOM.

This brief proposes the count-based PAM4 EOM as a means to find the optimal voltage reference and timing codes for the VDAC and PI (Fig. 3.9) over all frequencies up to the Nyquist frequency and over all the offsets and worst-case input patterns. The circuit comprises a counter that counts 127 internal clock cycles, a PI that adjusts the WCK timing, a VDAC that provides a reference voltage according to a digital code, and an internal memory set. As shown in Fig. 3.10(a), First, the pre-defined PAM4 PRBS7 data sequence is transmitted from memory to the PAM4-NRZ bridge in a burst mode. Three samplers of the RX divide the eyes of PRBS7 based on three reference levels, REFH, REFM, and REFL, and count/store the number of 1's in the internal memory of the bridge. Note that the number of 1's from the sampler with REFH, REFM, and REFL, becomes 32, 64, and 96, respectively, when the PRBS7 sequence is sampled without any error. When PRBS7 data of one cycle is transmitted starting from VCW=0, the WCK counter of the EOM circuit is reset, and the next

PRBS7 cycle is repeated to complete 64 timing scans (up to PCW=63). After that, if PRBS7 is transmitted 2112 times while sequentially increasing the VCW value to 32 in the same way, the EOM operation is completed. Then the optimal reference and timing code for VDAC and PI can be determined by the eye diagram after reading the stored data from the internal memory to the tester through I2C, as shown in Fig. 3.10(b). Since 2112 data (33 voltages X 64 timings) are required for scanning the entire eye, it takes less than 67us to perform EOM with a 4GHz-speed tester. With the advantage of finding the optimal codes faster than existing EOM methods, the count-based PAM4 EOM method can be applied to the time-limited training operation of the memory.

(a)   (b)
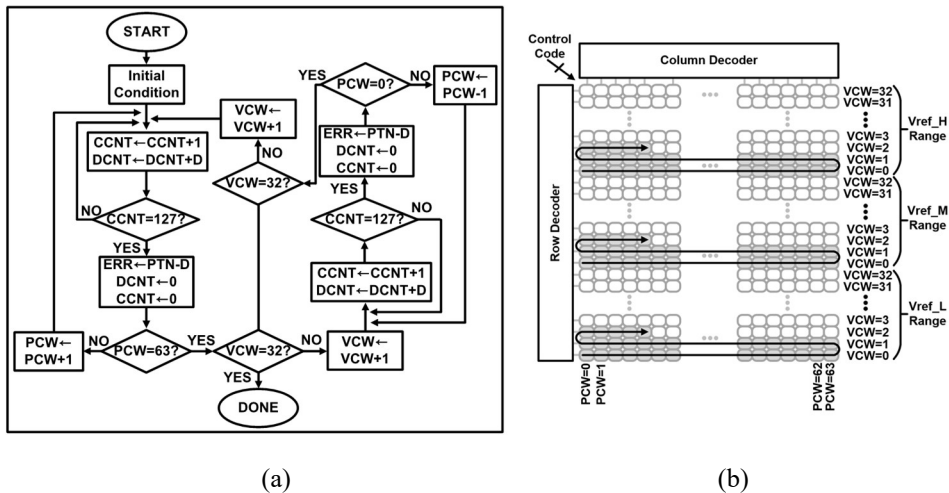
Fig. 3.10 (a) Flow chart and (b) Operation of the count-based PAM4 EOM.

3.6 Measurement
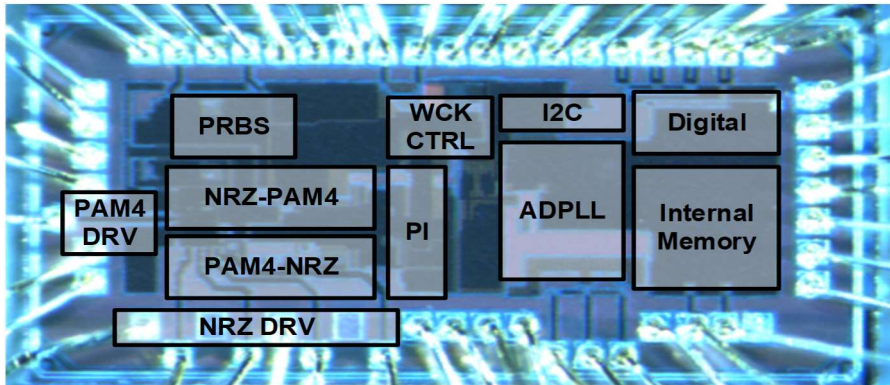
Fig. 3.11 Die photomicrograph.

The PAM4-Binary Bridge is fabricated in a 40 nm CMOS technology. Fig. 3.11 shows the die photomicrograph, along with the layout details of the 1.6 mm^2 PAM4-Binary Bridge.

Fig. 3.12 is the 32 Gb/s PAM4 output waveform captured using the internal PRBS7 generator in the bridge. The 2-tap FFE of the TX can compensate for the signal attenuation caused by the channel in the test environment with 0.95 RLM. In addition, with Gray coding, an NRZ mode is possible up to 16 Gb/s using only MSB bits while LSB bits are fixed as 0s. Fig. 3.13 shows the PAM4 data eye for each operating speed, and Table 3.1 shows the FoM for each operating speed. At the maximum operating speed of 32 Gb/s, FoM achieved 2.83 pJ/bit.

Fig. 3.13 shows the performance of the ADPLL and PAM4 RX characteristics of the bridge. The measured integrated rms jitter of the ADPLL in the bridge is 0.58 ps. With the proposed count-based PAM4 EOM method, the optimal PAM4 voltage and
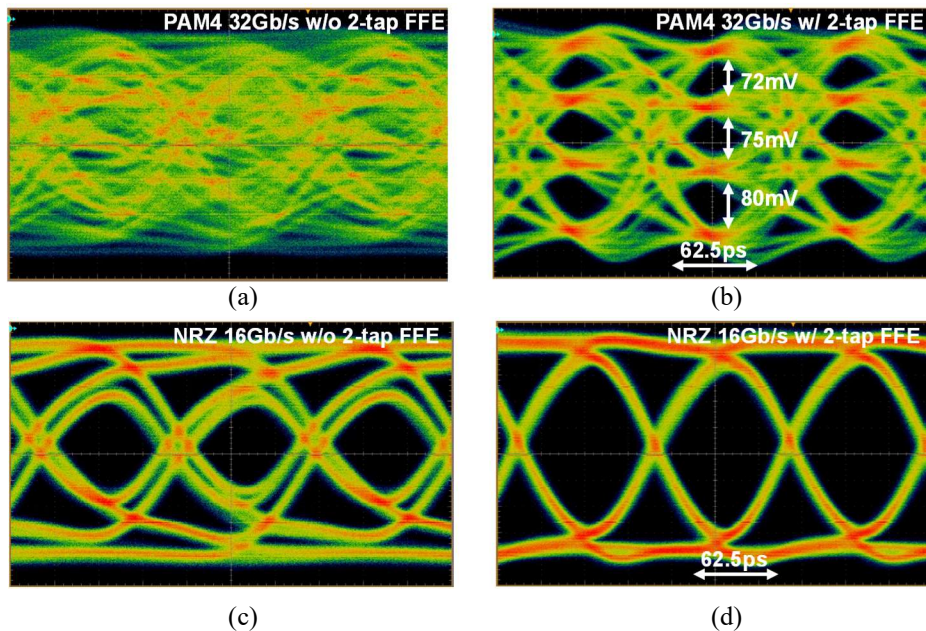
Fig. 3.12 Measured PAM4/NRZ transmitter data eye (a) PAM4 32 Gb/s without 2-tap FFE, (b) PAM4 32 Gb/s with 2-tap FFE, (c) NRZ 16 Gb/s without 2-tap FFE, and (d) NRZ 16 Gb/s with 2-tap FFE.

Table 3.1 Performance summary of Fig. 3.13.

Clock Frequency[GHz]	VDDH[V]	VDD[V]	Power[mW]	Data Rate[Gb/s]	FoM[pJ/bit]
6	1.25	0.9	65.7	24	3.17
7	1.25	0.9	70.2	28	2.95
8	1.25	0.9	90.4	32	2.83

timing codes for VDAC and PI were determined based on the data eye diagram obtained from the internal memory of the bridge. The average measured DC offset is about 2.8 mV$_{rms}$ with the CTLE DC gain of 4 dB and the boost set to 6 dB when an optimal reference voltage at the input of the CTLE is determined during training.

(a)

(b)

(c)

(d)
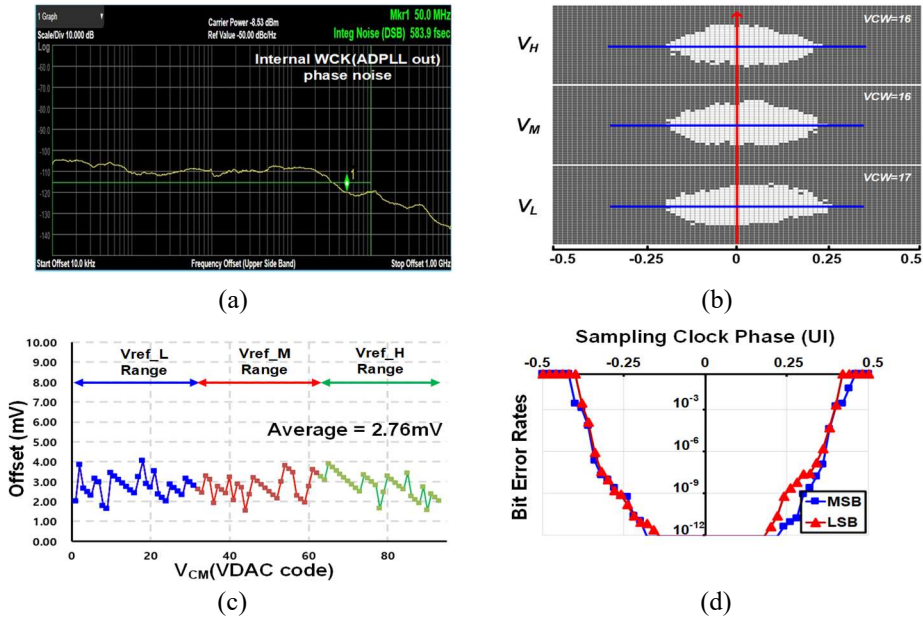
Fig. 3.13 (a) Measured phase noise of ADPLL, (b) count-based EOM, (c) CTLE/sampler offset, and (d) Bathtub curve of PAM4 receiver.

PAM4 eyes are fully open, as demonstrated by the bathtub plot showing 50 % horizontal margin at BER <10^{-9} at the 8 GHz Nyquist frequency.

Table 3.2 shows the performance summary of the proposed PAM4-NRZ bridge and its comparison with a previously introduced work. The proposed PAM4-NRZ bridge achieves a speed of 36 Gb/s per DQ pin of the memory.

Table 3.2 Performance summary and comparison

	ISSCC'21 [3]	ASSCC'21 [6]	ISCAS'19 [10]	JSSC'21 [33]	JSSC'15 [38]	This work
Technology	1Ynm	28nm	65nm	65nm	65nm	40nm
Supply(V)	1.35	1.2/1	1.0	1.0/0.6	1.1	1.25/0.9
Data rate(Gb/s/pin)	22	24	20	28	16.8	32
Tx Driver topology	Voltage-mode SST	Voltage-mode	Voltage-mode SST	Voltage-mode	POD	Current-mode
Tx Equalization	Pulse-based de-emphasis	-	2-tap FFE	2-tap asymmetric FFE	2-tap FFE	2-tap FFE
Rx Slicer topology	CTLE	1tap DFE	-	-	Source follower type CTLE	CTLE
Signaling type	PAM4/NRZ	PAM4/NRZ	PAM4/NRZ	PAM4	NRZ	PAM4/NRZ
Clocking type	No PLL	External	External	External	PLL	ADPLL
EOM function	-	-	-	-	-	O
Rx BER(10⁻⁹)	-	0.36UI	-	0.5UI	0.38UI	0.5UI
Tx Driver RLM	-	0.95@24Gbps	0.98@20Gbps	0.993@28Gbps	-	0.95@32Gbps
Energy efficiency (pJ/bit)	-	-	3.07*	0.64*	5.9**	3.66***

*Tx only   **Includes Rx and PLL   ***Includes Rx, PLL and EOM function

Chapter 4

Design of PAM4 Level Mismatch Adjustment Scheme for 48 Gb/s PAM4 Memory Interface

4.1 Overview

With the ever-increasing bandwidth of memory interfaces, securing link margins is becoming more critical. Starting with the production of memory that employs PAM4 such as GDDR6X, each memory manufacturer aims at designing a robust interface with PAM4 signaling. However, direct test equipment is not available to verify the operation of the memory with the controller. For example, T5511 (ADVANTEST) Tester does not support PAM4 signaling and testable data rates are limited to 8 Gbps. Therefore, development of innovative methods such as the Built-Out Self-Test chip

(BOST) chip [5], [6] is being designed to ease the high-speed testing as well as reducing the increased cost of new test equipment. However, testing the memory requires adjustable tuning of the RLM as well as impedance matching. In this thesis, an optimal bridge architecture with all the necessary functions to test and validate a high-speed PAM4 memory using existing test equipment is proposed. In addition, we describe the method of output level adjustment in the PAM4 driver to improve the RLM.

4.2 PAM4 Memory/Tester bridge
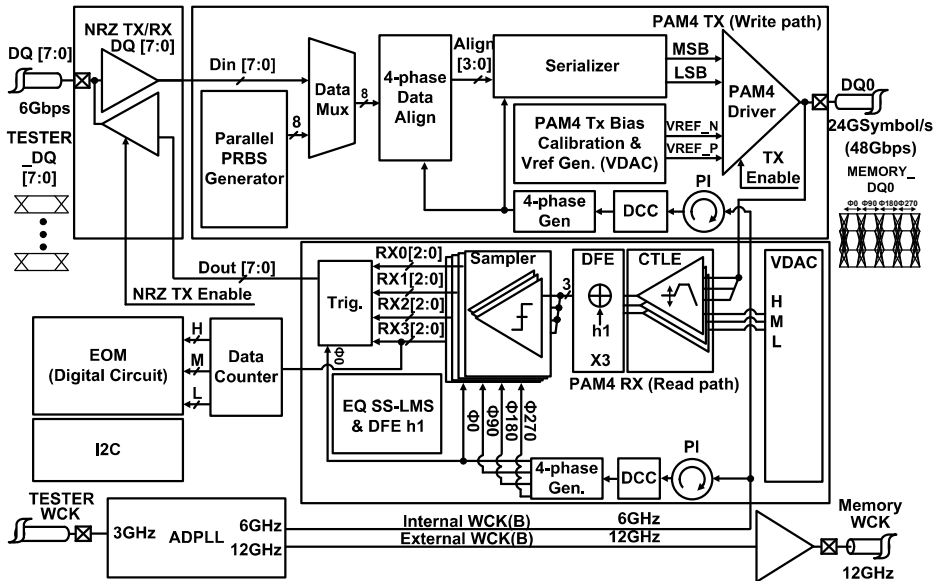
Fig. 4.1 Overall architecture of the proposed PAM4 Memory/Tester bridge.

Fig. 4.1 shows the proposed PAM4 Memory/Tester bridge that can test PAM4 sig-naling memory with a tester that supports only NRZ signaling. The proposed bridge employs a quadruple WCK signaling scheme in the GDDR5/6 interface. To support the write operation of the DRAM, the tester transmits NRZ data on 8 pins to the bridge chip. The input level of the PAM4 driver is adjusted in response to the calibration result to improve the RLM of the PAM4 signal. This 4-level adjustable main driver is employed to reduce the impedance variation of the output driver caused by the changes in the output level [33]. The PAM4 main driver is a single-ended voltage mode CMOS driver with VSS-termination. The receiver incorporates 3 parallel

CTLEs and 1-tap DFE along with digital adaptation algorithm. A direct-feedback DFE eliminates the first post-cursor. The internal WCK, which an ADPLL generates, is phase-adjusted by a PI and a DCC for maximum frequency control within the bridge.

4.3 Level Mismatch Adjustment Transmitter

4.3.1 Overall Architecture
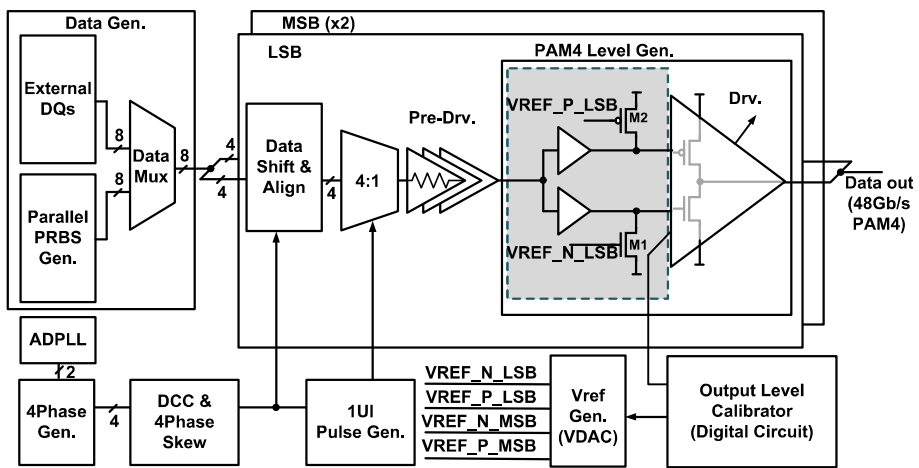
Fig. 4.2 Proposed level adjustable PAM4 TX scheme.

The bridge chip transmitter must be able to generate precise PAM4 signals with a high a RLM value, adjust each output level for DRAM RX testing, and provide various termination and output impedance values as per DRAM interface specifications. To achieve this, a CMOS voltage mode driver was employed to enable impedance control. DRAM typically employs a CMOS driver with 6 legs of 240 Ω each to configure various termination levels and output impedances. Using a voltage mode driver instead of a current mode driver offers an advantage in power consumption. However, there is a need to address the issues of impedance mismatching and the RLM degradation due to transistor nonlinearity in multi-level signaling. To solve this issue, [33]

implemented an N over N voltage mode driver for LPDDR5 using the LVSTL inter-
face. They resolved the problem by adjusting the number of transistors turned on,
similar to the width, based on each output level. The NMOS transistor of the N over
N driver operates in the triode region, and the non-linear change as per output level
and V_{ds} level is controlled through W, as shown in Equation 4.1.

$$I_{ds} = \frac{1}{2}k_n\frac{W}{L}\left[2\left(V_{gs} - V_{th}\right)V_{ds} - V_{ds}^2\right] \tag{4.1}$$

The bridge chip enhanced non-linearity by modifying Vgs rather than rectifying
linearity through width correction. Modifying Vgs enables better precision than ad-
justing the width. In addition, it is employed to enhance the RLM and modify the
output level for evaluating the voltage margin of the DRAM receiver.

Fig. 4.2 shows the level-adjustable PAM4 transmitter architecture of the bridge. 6-
Gbps data generated by the internal PRBS or external 8 DQ pins are used, and internal
data passing through the 4:1 mux is fed to the PAM4 main driver. The data is pro-
cessed at the PAM4 main driver with 24 GS/s through the control of the 4-phase gen-
erator. Since the transistor of the PAM4 driver operates in the linear region, the four
output voltage levels (drain voltage) are changed by the gate-source voltage of the
MOS transistor. Therefore, transistors (M1, M2) are added to adjust the driver's input
node. Additional transistors for the PAM4 driver independently control the input level
of the PAM4 driver to evaluate the voltage margin of the memory and improve the
RLM. The amount of Vgs controlled by the added transistor is determined by the
VREF_P_LSB level obtained through calibration.

4.3.2 Level Adjustment PAM4 Driver
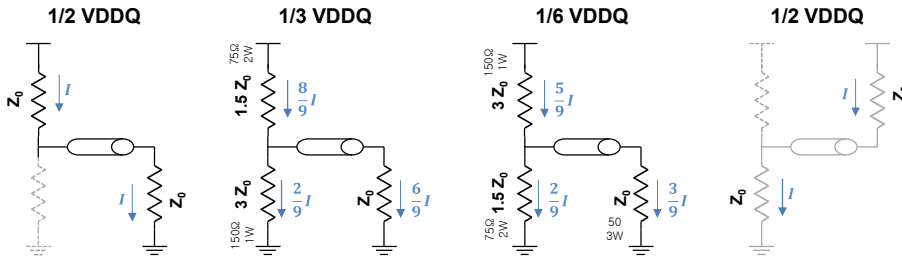
Fig. 4.3 Impedance and current prediction according to output level.

The PAM4 binary bridge transmitter requires precise control over each level of the PAM4 signal and impedance matching capability. Fig. 4.3 illustrates the ideal driver's output impedance for each output level, consisting of parallel PMOS and NMOS transistors and three $3Z_0$ size transistors. The bridge chip defaults to using VSS termination. To produce a 1/2VDDQ output level, all three pull-up transistors are activated while all pull-down transistors are deactivated, creating an output level with an impedance matching that of the termination. For a 1/3VDDQ output level, two pull-up transistors and one pull-down transistor are turned on, resulting in an output impedance of $3Z_0 \| 1.5Z_0 = Z_0$. The pull-up and pull-down transistors' resistances and the termination resistance determine the output level and impedance. The third picture in Fig. 4.3 shows that even when the output level is 1/6VDDQ, it comprises a pull-up driver of $3Z_0$ and a pull-down driver of $1.5Z_0$.

However, the implementation of Fig. 4.3 using real transistors results in non-linear driver operation that does not resemble the resistor in the figure. To mitigate the RLM

Table 4.1 The equation of PAM4 transistor according to output level.

Output Level	PMOS	NMOS
$\frac{1}{2}V_{DDQ}$	$I_O = K_p\frac{3W}{L}[2\{(\gamma - V_{DDQ} - V_{th})\cdot(-\frac{1}{2}V_{DDQ})\} - (-\frac{1}{2}V_{DDQ})^2]$	$I_O = K_n\frac{3W}{L}[2\{(V_{DDQ} - V_{th})\cdot(\frac{1}{2}V_{DDQ})\} - (\frac{1}{2}V_{DDQ})^2]$
$\frac{1}{3}V_{DDQ}$	$\frac{8}{9}I_O = K_p\frac{2W}{L}[2\{(\delta - V_{DDQ} - V_{th})\cdot(-\frac{1}{3}V_{DDQ})\} - (-\frac{1}{3}V_{DDQ})^2]$	$\frac{2}{9}I_O = K_n\frac{W}{L}[2\{(V_{DDQ} - \alpha - V_{th})\cdot(\frac{1}{3}V_{DDQ})\} - (\frac{1}{3}V_{DDQ})^2]$
$\frac{1}{6}V_{DDQ}$	$\frac{5}{9}I_O = K_p\frac{W}{L}[2\{(-V_{DDQ} - V_{th})\cdot(-\frac{1}{6}V_{DDQ})\} - (-\frac{1}{6}V_{DDQ})^2]$	$\frac{2}{9}I_O = K_n\frac{2W}{L}[2\{(V_{DDQ} - \beta - V_{th})\cdot(\frac{1}{6}V_{DDQ})\} - (\frac{1}{6}V_{DDQ})^2]$
Result	$\gamma = \frac{1}{6}V_{DDQ}, \delta = \frac{1}{12}V_{DDQ}$	$\alpha = \frac{1}{12}V_{DDQ}, \beta = \frac{1}{6}V_{DDQ}$

degradation caused by this non-linearity, a voltage mode PAM4 driver is recommended. The CMOS driver operates in the triode region as shown in equation 4.1, which causes non-linearity due to the current changing proportionally with the square of the output voltage, V_{ds}.

To achieve linear operation, corrections can be made by adjusting the width or Vgs according to Vds in equation 4.1. Previous work improved linearity by modifying the width, but in this study, the gate voltage was adjusted for fine resolution to support output level changes for the DRAM RX testing and the RLM improvement. Fig. 4.5 shows that the number of turned-on transistors determines the width, while the output level determines Vd, and the current flowing through the transistor is defined for each output level as in Fig. 4.3. Table 4.1 displays the Fig. 4.3 equation correction values α, β, γ, and δ, which require calibration circuits to fine-tune Vgs values for each output level. However, this calculation result is ideal and may differ from the actual result.

There are two approaches to determine the gate level of the ultimate driver. The first method involves configuring the supply voltage level of the pre-driver based on the preferred gate level, as shown in Fig. 4.4(a). Although it can be conveniently achieved by creating the preferred voltage level using the LDO, implementing this
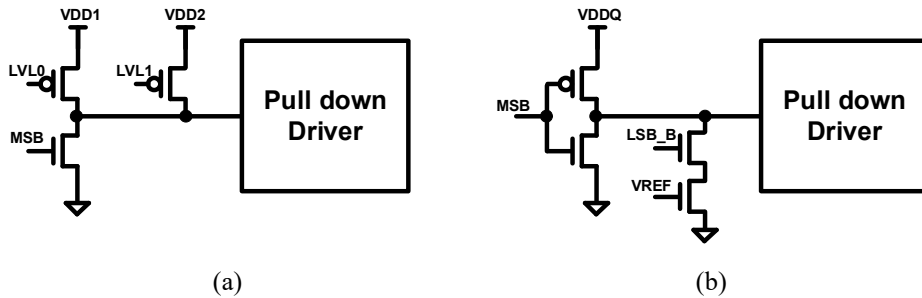
(a)   (b)

Fig. 4.4 Schematic of possible pre-driver.

method requires the placement of four power lines and capacitors to generate the 4

levels (VDDQ-α, VDDQ-β, γ, δ), which causes area overhead and noise control issues.

Additionally, due to the continuous switching between VDD1 and VDD2, a charge

flowing problem from a high voltage to a low voltage is expected. This leads to an

increase in the circuit complexity, particularly when low voltages are high.

In contrast, the second method involves creating a DC path at the output of the

predriver to adjust the output level, as illustrated in Fig. 4.4(b). By regulating the size

of the DC path through reference voltage control, the desired output level of the

predriver is established. Although this approach requires four references, the reference

signals are linked to the gate node of the final driver, which results in less capacitance

and metal line overhead than the first method. While the second method results in

higher current consumption due to the occurrence of a DC path, it was preferred over

the first method since the overhead for four power lines was deemed to be greater in

practical implementation.

The predriver used in this study has a drawback where the slope is lower when

transitioning to VDDQ-α compared to transitioning to VDDQ, causing a decrease in

the main driver's bandwidth. The NODE_X generates VDDQ-α and VDDQ signals

based on the MSB_B signal, as shown in Fig. 4.5(a), and the eye diagram in Fig. 4.5(b)

shows the difference in slope when transitioning between VDDQ-α and VDDQ.

(a)   (b)
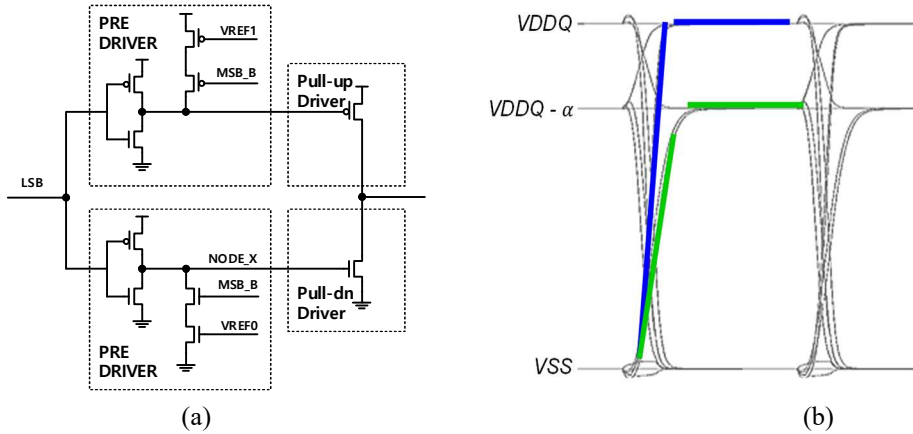
Fig. 4.6 (a) Proposed pre-driver for controlling the output level (b) simulation result of expected problem of the pre-driver.

(a)   (b)
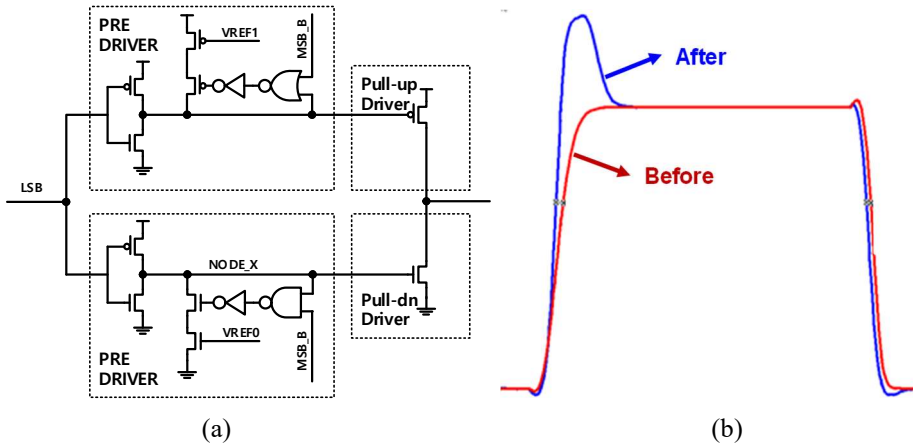
Fig. 4.5 (a) Schematic of overdrive scheme (b) simulation result of NODE_X before and after applying the scheme.

To address this issue, an overdriver scheme was employed to enhance the predriver's slope. Fig. 4.6(a) shows that the predriver's DC path was activated after NODE_X swings enough when outputting VDDQ-α, resulting in the waveform shown in Fig. 4.6(b). Even if there is an overshot at the predriver output, the slope becomes similar to when VDDQ is reached. Fig. 4.7 illustrates the simulation results of the main driver

before and after implementing the overdrive scheme, demonstrating an enhanced output slope that ensures sufficient eye margin. The overall transmitter architecture, combining the predriver and main driver, is presented in Fig. 4.8.

(a)   (b)

Fig. 4.7 The simulated output eye diagram before/after applying the overdrive scheme.
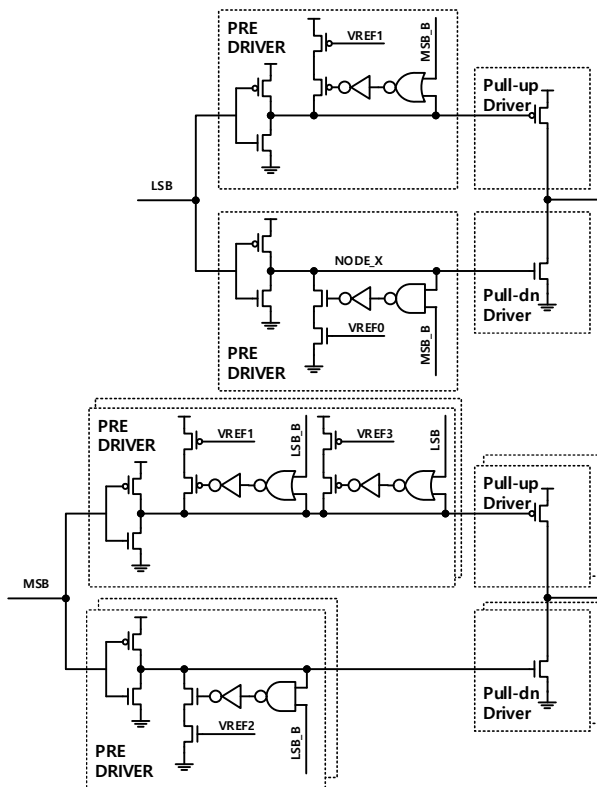
Fig. 4.8 Overall architecture of the driver.

4.3.2   PAM4 Driver input level calibration
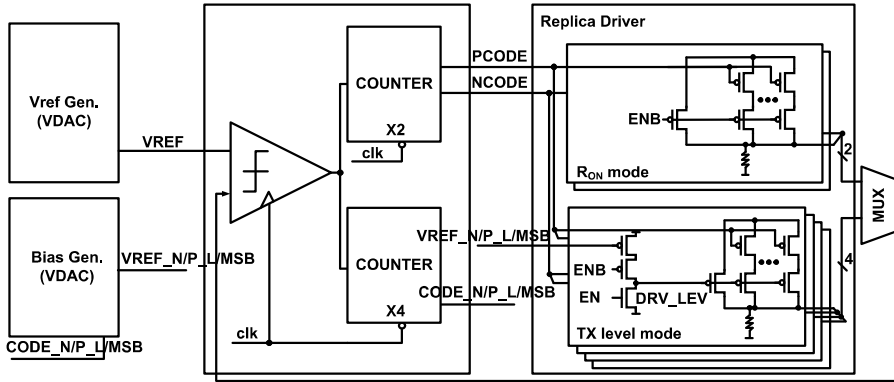
Fig. 4.9 PAM4 Driver input level calibration circuit.

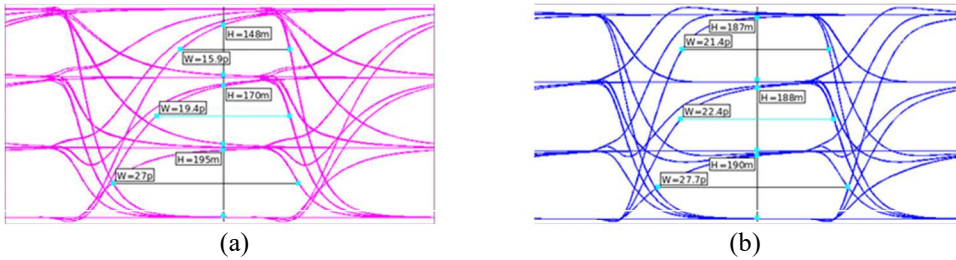To optimize the RLM of the PAM4 output, a suitable reference level needs to be applied to the pre-driver. This thesis suggests a calibration circuit to determine the reference value. Fig. 4.9 depicts the circuit architecture, which includes 6 replica drivers to show the state of the driver for each of the 4 output levels, a comparator to verify the appropriateness of the replica driver's output level, and a digital block implemented with verilog code to determine a proper reference. As illustrated in Fig. 4.10, the reference signal outputs the voltage generated by the resistor divider through the mux. VDAC generates four signals, namely VREF0~VREF3, to specify the output level, and three reference signals (1/2VDDQ, 1/3VDDQ, 1/6VDDQ) for calibration purposes. Fig. 4.11 displays the calibration flowchart, while Fig. 4.12 portrays the concept of a replica driver. The thesis introduces a novel calibration circuit to ascertain the appropriate reference value for maximizing the RLM of the PAM4 output.
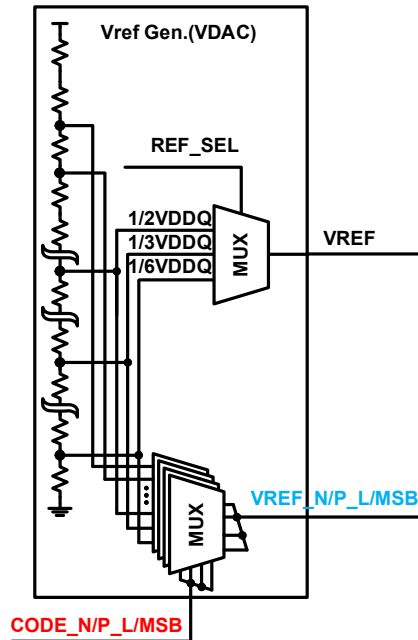
Fig. 4.10 Block diagram of the reference generator.

The calibration process involves six steps. The first and second stages involve determining the widths of the NMOS and PMOS. Due to the PVT variation, the widths of the PMOS and NMOS are determined based on the operating conditions to adjust the output impedance of the NMOS and PMOS to Z0 and 3Z0, respectively. This is done by setting the gate input voltages of NMOS and PMOS to VDD and VSS, respectively, and configuring the replica driver as shown in Fig. 4.12(a). The calibration logic adjusts the widths until the NMOS replica driver outputs 1/2VDDQ level and the PMOS replica driver outputs 1/6 VDDQ level, and the width values obtained through calibration become PCODE and NCODE signals, which are then transmitted to all drivers and replica drivers. Steps 3 to 6 involve finding an appropriate input gate level. Since the suitable width for PVT was selected through
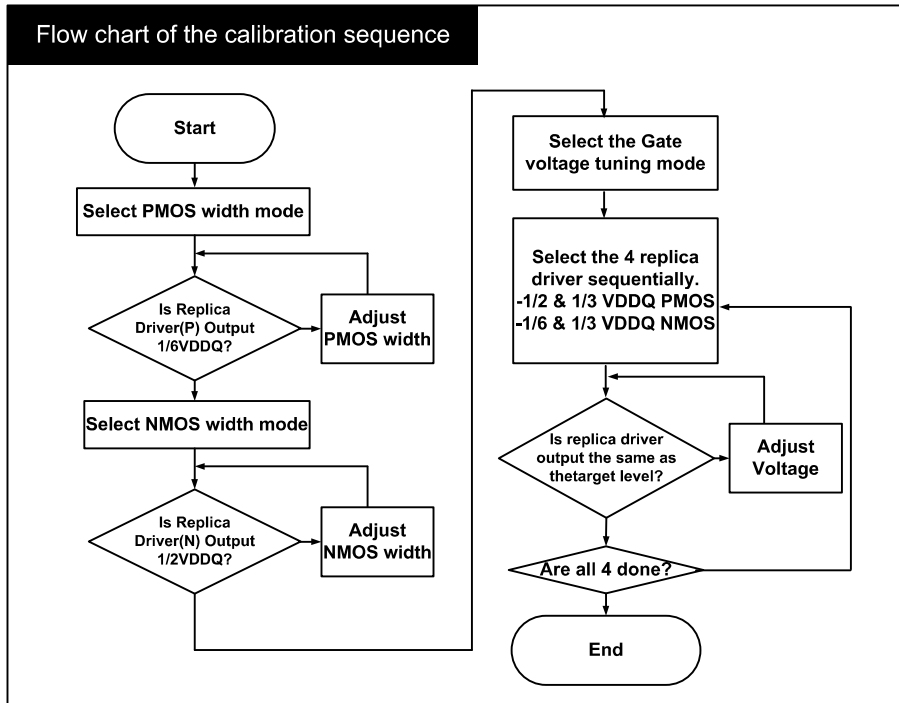
Fig. 4.11 Flow chart of the calibration sequence.

PCODE and NCODE, the next step is to determine a reference voltage that can make the output level linear. Fig. 4.12(b) shows that there are four replica drivers for each output level. For example, if the output level is 1/3VDDQ, the PMOS and NMOS on the right side of Fig. 4.12(b) should be selected and calibrated.

To prevent a DC path caused by the NMOS replica driver, a pull-up resistor with the same function was obtained and simplified. In the NMOS replica driver located at the bottom right of Fig. 4.12(b), the pull-up resistance and the replica termination resistance cause a DC path. To block the DC path, the replica termination resistor was removed, and a pull-up resistor of 6Z0 was used to generate an impedance of 3Z0 and output 1/3VDDQ. The NMOS replica driver operating at 1/6 VDDQ also used the

same method by removing the replica termination resistor and changing the pull-up
to 15/2Z0.

Once the reference level is found for each output level, the calibration process is
complete, and all calibration result values can be read out via I2C. To adjust the output
level for the DRAM RX test, the calibration result code can be read first, changed
accordingly, and then written back to I2C.

The transient simulation results in Fig. 4.13(a) were used to verify the calibration
operation, which shows that the replica driver output (black line) for each mode is
aligned with the reference signal (blue line). Fig. 4.13(b) presents the simulation
outcome of the calibration. Usually, the high-level eyes are relatively large before
calibration due to the nonlinearity of the Ron value according to the V_{ds} level.
However, after calibration, obtaining a uniform eye size for all modes is possible.
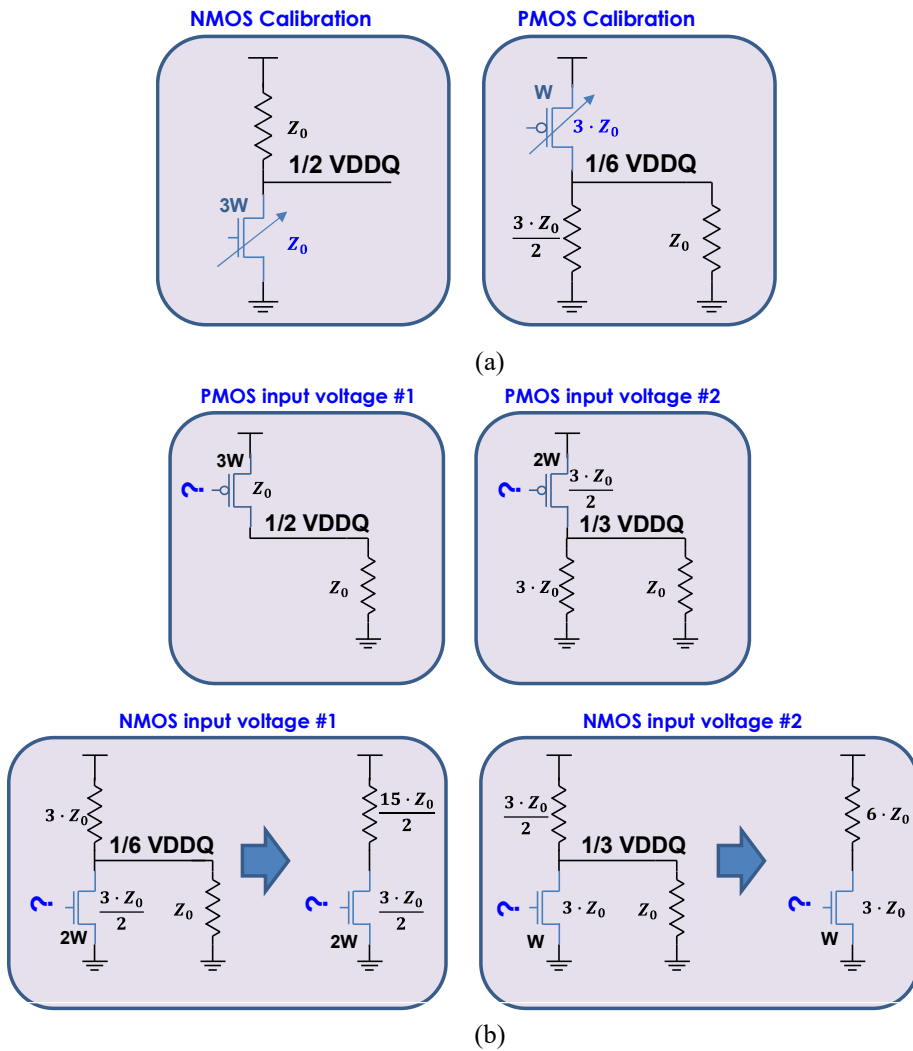
Fig. 4.12 Replica driver of calibration circuit (a) replica drivers that tune the width

(b) replica driver that tune the V_{gs}.
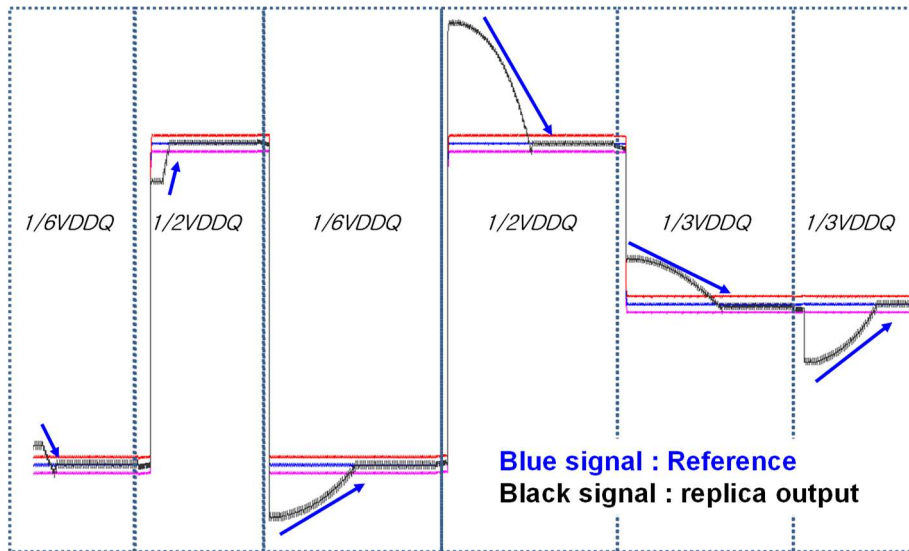
(a)
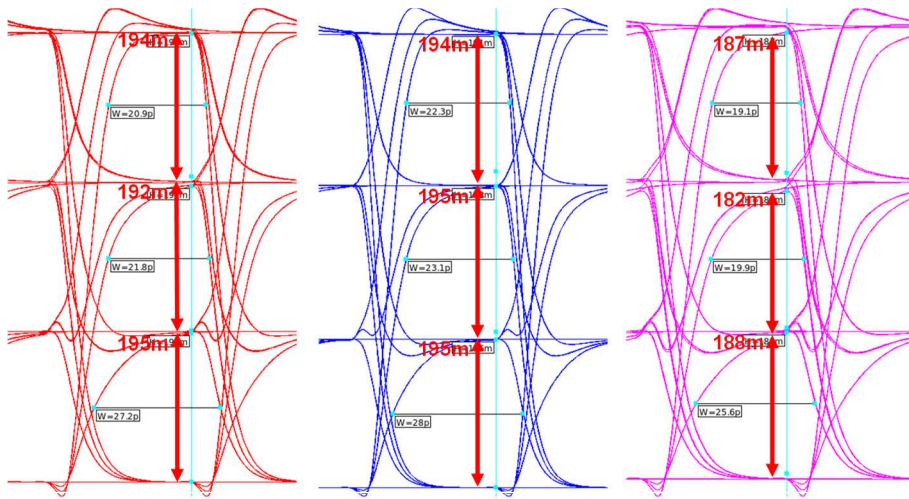
(b)

Fig. 4.13 (a) transient simulation result of calibration circuit and (b) simulated output

eye diagram of transmitter after the calibration.

4.4 PAM4 Receiver with Nonlinearity Compensation
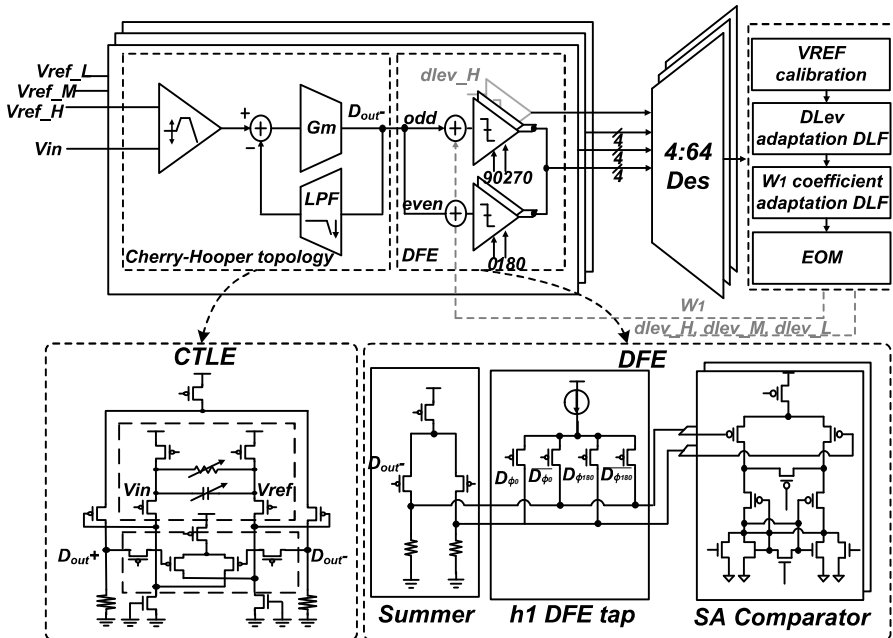
Fig. 4.14 Block diagram of PAM4 RX, including circuit implementation of analog front end.

Fig. 4.14 describes the PAM4 receiver of the bridge. A Cherry-Hooper topology consists of a conventional RC-degeneration CTLE, a transimpedance stage (Gm cell), and an additional peaking stage by the negative feedback of low-pass-filtered signal. PMOS input transistors are utilized from VSS to half VDD to cover the data level sent from DRAM. The Cherry-Hooper CTLE consists of three stages. The first stage, a conventional RC-degeneration stage, provides AC gain peaking around the Nyquist frequency. The second stage, a transconductance or CML stage, provides overall gain
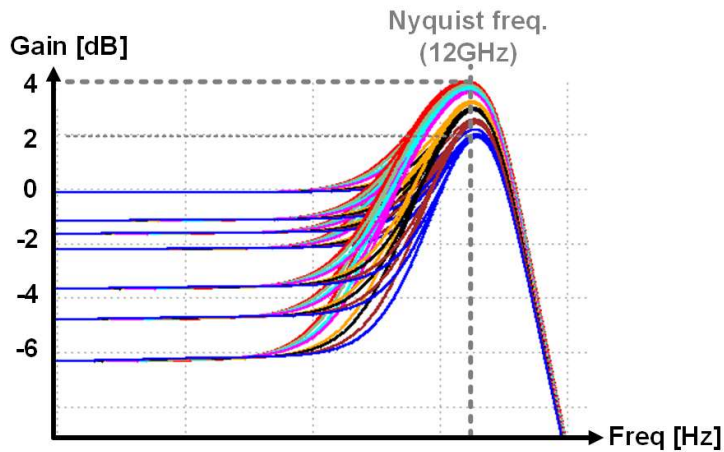
Fig. 4.15 CTLE frequency response post simulation result.

from DC to the Nyquist frequency, similar to a VGA. The third stage is a negative

feedback stage, where the intended resistance in the feedback path creates an

additional pole in the frequency domain that acts as a zero due to negative feedback.

This additional zero helps to increase peaking gain and broaden the bandwidth in the

presence of an output pole. Therefore, the Cherry-Hooper CTLE has several

advantages over the RC-degeneration CTLE, including overall gain, bandwidth, and

peaking gain in the presence of the same output pole. This helps overcome the design

complexity that arises when a single-ended input is used in the presence of channel-

length modulation.
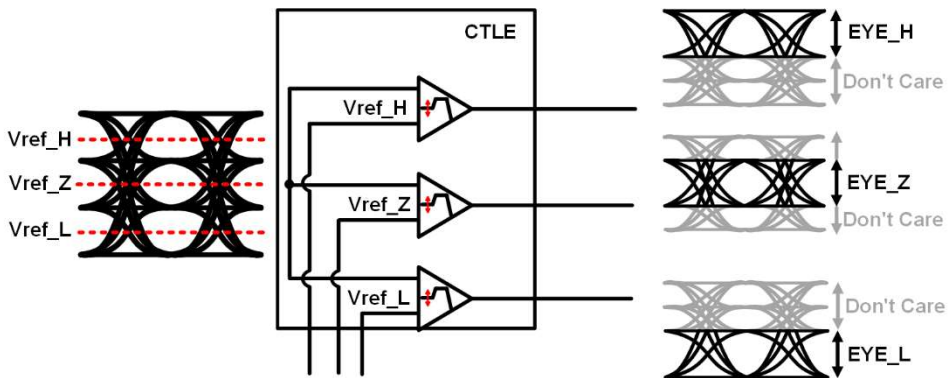
Fig. 4.16 PAM4 data in 3 CTLEs.

The transconductance stage provides an additional benefit to the operation of the subsequent DFE. The summer of the DFE, located immediately after the CTLE, is a CML type. If the input voltage swing is greater than a specific range, linearity is disrupted, and the DC gain decreases. However, the first stage of the CTLE outputs a pseudo-differential swing where one side voltage swing is twice that of the other, and the second CML stage distributes them to nearly equal current since the two input transistors share the same source node and current source. This helps the CTLE output swing fit within the linearity range of the DFE summer.

The source-degenerated resistor and capacitor values are digitally controlled using 3-bit control bits through I2C. The current sources for the three stages are biased from a single bias pad, and their respective 3-bit control bits adjust the biasing current with fine resolution. The resistance in the negative feedback path is digitally controlled by 1 bit. The post-simulation of the CTLE frequency response with various R, C values is depicted in Fig. 4.15, which offers selective 2~8 dB peaking gain around the Nyquist frequency.

For each of the 3 parallel CTLEs input transistors, 3 threshold voltages are applied

on the opposite side of the received PAM4 signal, as illustrated in Fig. 4.16. These
threshold voltages are independently calibrated via the Count-based EOM and
generated from VDAC using a 7-bit control input. Independent calibration of
threshold voltages reduces the impact of random offset caused by transistor mismatch
and data path. Additionally, since each of three CTLEs deals with only 1 data eye, the
design complexity of the PAM4 RX, such as linearity condition and DFE summer
input linearity range, becomes significantly relaxed.

Fig. 4.17 illustrates the circuit implementation of the nonlinearity-compensating
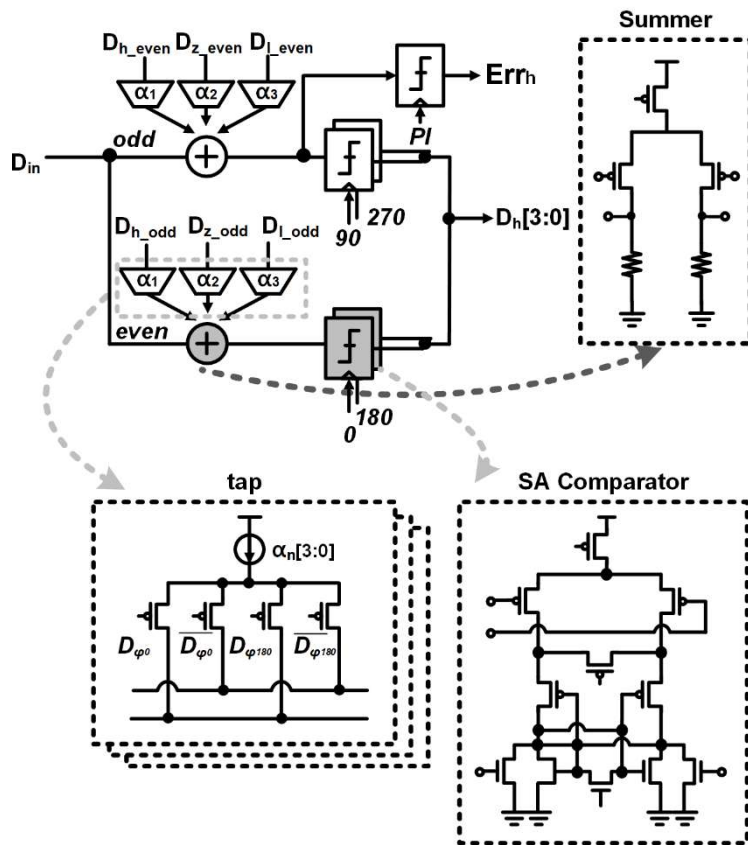
Fig. 4.17 DFE circuit implementation.

1-tap 9-coefficient adaptive DFE proposed. Unlike a speculative DFE, the direct feed-

back DFE has a more stringent timing constraint. To address this, the StrongArm latch

output directly feeds into the CML tap within the buffer without the RS latch, which

converts RZ data to NRZ data, thereby reducing feedback time. In addition, the DFE

employs a shared-summer structure. In a conventional quarter-rate DFE, 4 summers

are required to use the first tap by NRZ data with 4 UI widths. However, using RZ

data, which provides valid data for only 2 UI, 4 summers can be combined into two

paths, even and odd. This results in a reduced summer, which reduces parasitic capac-

itance at the CTLE output node, thus relaxing settling time constraints.

4.5 Measurement
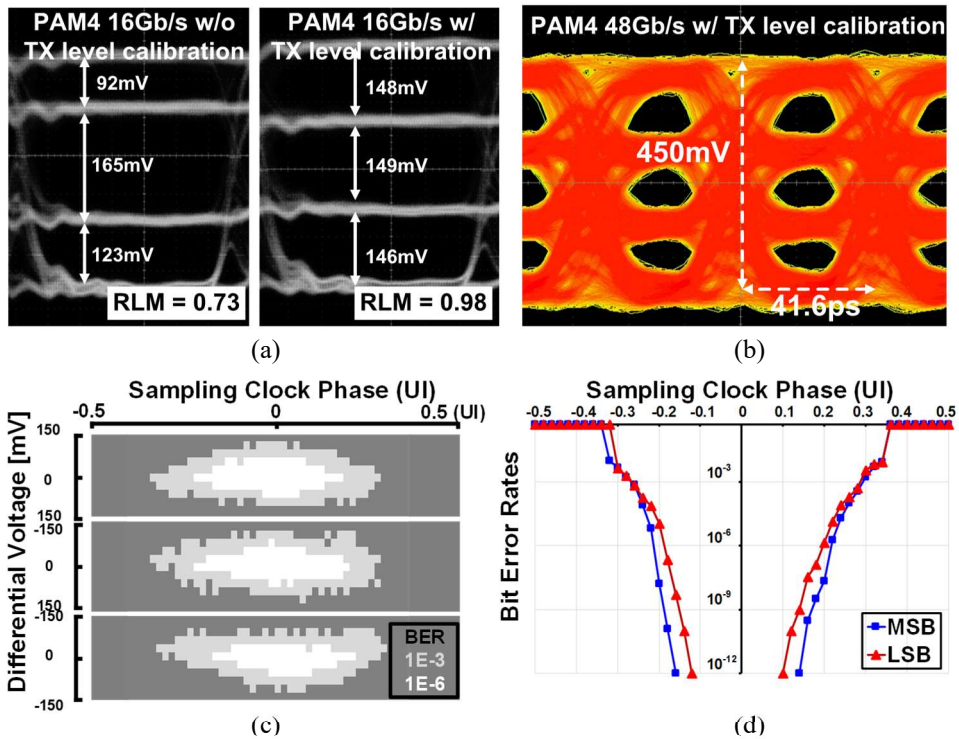
Fig. 4.18 Measured (a) eye-diagrams of PAM4 output before and after calibration, (b) eye-diagram of PAM4 transmitter at 48 Gbps, (c) phase noise of ADPLL, and (d) bathtub curve of PAM4 receiver.

The prototype chip is fabricated in 40 nm CMOS technology and uses a total area of 2.13×1.098 mm^2. Fig. 4.18(a) is the measurement result of the RLM before and after calibration, and the RLM of the PAM4 driver is improved from 0.73 to 0.98 at 16 Gbps. Fig. 4.18(b) shows 48 Gbps PAM4 TX output eye diagram obtained with the same calibration coefficients. Fig. 4.18(c) shows the data eye diagram obtained

from the internal EOM circuit of the bridge. Fig. 4.18(d) shows the performance of

the bridge's RX characteristics. With the use of the PAM4 RX, the BER was measured

at less than 10^{-12}, and an opened eye is obtained at 48 Gbps. Fig. 4.19 illustrates the

chip photomicrograph and the measurement setting. Table 4.2 summarizes the perfor-

mance and compares it to other works. In this work, the speed is improved from the

existing built-out test method, and the PAM4 level can be adjusted, making it possible

to test and characterize the memory interface.
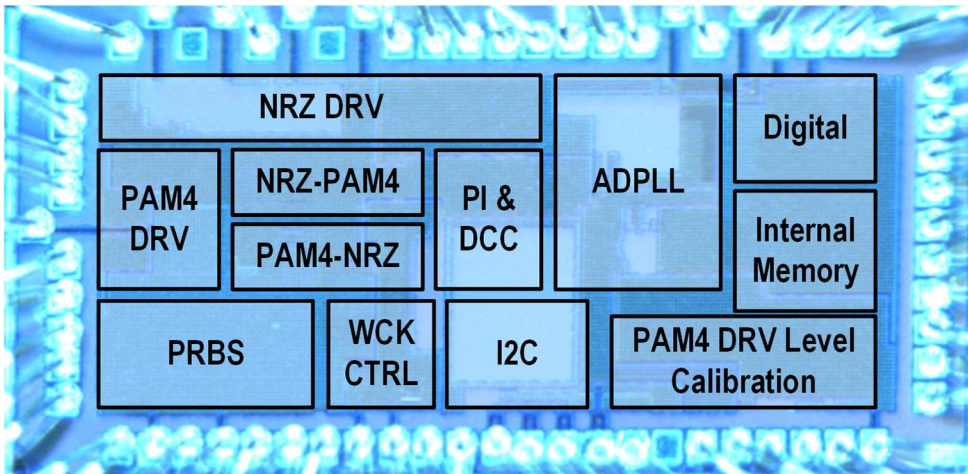
Fig. 4.19 Chip photomicrographs.

Table 4.2 Performance summary and comparison.

	TCAS-II'21 [5]	ASSCC'21 [6]	JSSC'21 [33]	ISSCC'17 [36]	This work
Technology	40nm	28nm	65nm	40nm	40nm
Supply(V)	1.25/0.9	1.2/1	1.00/0.6	1.5/1.0	1.25/0.9
Data rate(Gb/s/pin)	32	24	28	56	48
Tx Driver topology	Current-mode	Voltage-mode	Voltage-mode	Current-mode	Voltage-mode
Tx Equalization	2-tap FFE	-	2-tap FFE	3-tap FFE	2-tap FFE
Rx Slicer topology	CTLE	1 tap DFE	-	CTLE, 3-tap DFE	CTLE, 1-tap DFE
Signaling type	PAM4/NRZ	PAM4/NRZ	PAM4	PAM4/NRZ	PAM4
Clocking type	ADPLL	External	External	PLL	ADPLL
EOM function	O	-	-	O	O
Rx BER(10^{-9})	0.5UI @32Gbps	0.36UI @24Gbps	0.5UI @28Gbps	0.35UI @56Gbps	0.3UI @48Gbps
Tx Driver RLM	0.95 @32Gbps	0.95 @24Gbps	0.95 @28Gbps	-	0.96 @48Gbps
Energy efficiency (pJ/bit)	3.66*	-	0.65**	3.57@Tx * 6.82@Rx	1.85@Tx * 2.97@Rx

*Includes Rx, PLL and EOM function   **Includes Rx and PLL

Chapter 5

Design for testability & measurement setups

5.1 Design for testability

5.1.1 Clock Generator

The tester sends an input clock with a frequency of 4 GHz to the bridge, while an internal clock of 8 GHz is needed for the PAM4 transmitter to function at its maximum capacity of 48 Gb/s. The thesis documents the use of ADPLL to raise the clock speed from 4 GHz to 8 GHz. In addition, the ADPLL was tested and found to have a phase noise level of -99.2 dBc/Hz at a 1 MHz offset.
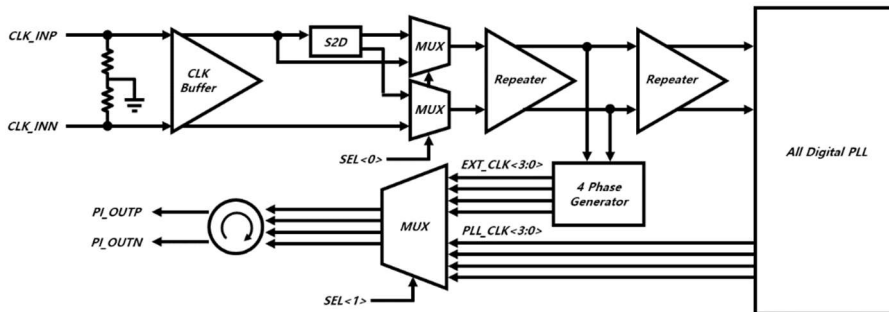
Fig. 5.1 Block diagram of internal clock generator.

Fig. 5.1 shows the inclusion of an external clock path for separate operation verification alongside the ADPLL. When a 4 GHz clock is externally input, the ADPLL outputs a 4-phase 8 GHz clock, which is then transmitted as the input signal of the PI. On the other hand, when an 8 GHz clock is externally input, it is directly transmitted to the PI as the input signal via a 4-phase generator without requiring ADPLL operation. The differential input clock may also be converted to a single-ended clock. The operation modes for each case can be selected through I2C control, with a summary of the operation modes provided in Table 5.1.

Table 5.1 Clock generator mode by SEL code.

Case	Input Source	WCK(B) Input	Single/Differential	SEL<1:0>
1	PLL Clock	4G	Differential	00
2	External Clock	8G	Differential	10
3	External Clock	8G	Single	11

5.1.2 Phase Interpolator

(a)

(b)
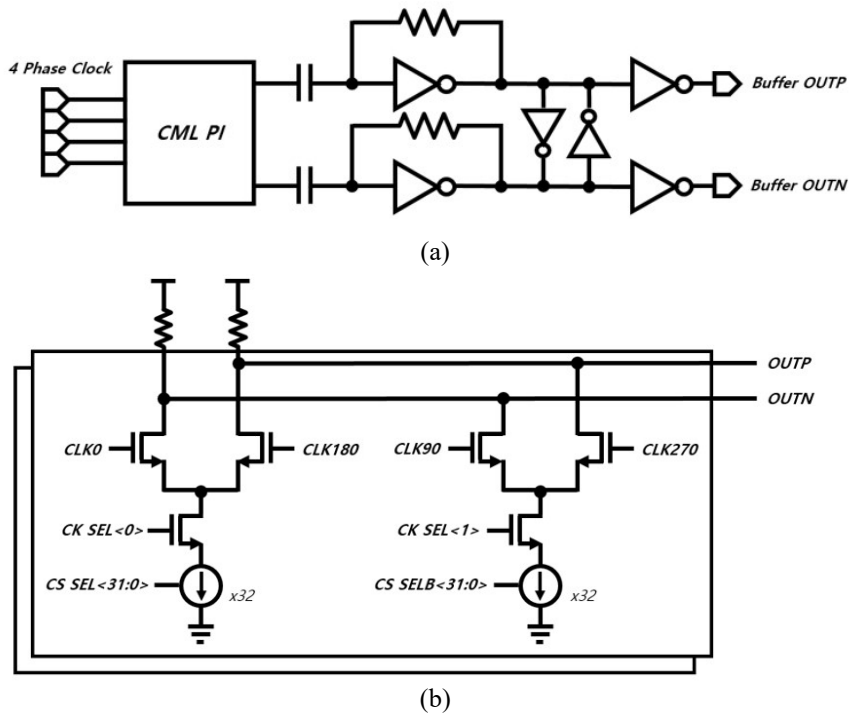
Fig. 5.2 (a) CML PI and AC buffer (b) CML based PI.

The phase interpolator (PI) is responsible for interpolating the 4 phase clock deliv-
ered by the PLL into a differential clock with a specific phase. By adjusting the phase
of the clock used in the serializer, the output phase of WCK can be modified. To
achieve a resolution of 1 ps or less, an 8-bit control PI with a resolution under 1 ps
was designed based on a 4 GHz clock. The structure of the CML-based PI used in the
chip is illustrated in Fig. 5.2. The CML-based PI was preferred due to its supply noise
advantages over CMOS-based PI. It accepts 4 phase clocks, generates 2 phase differ-
ential clocks through CML PI, and sharpens the created clock while correcting duty

Fig. 5.3 Simulation result of PI waveform.

using an AC buffer. The phase is interpolated by adjusting the number of current sources that affect the amount of current flow.

To begin the phase interpolation process, CK_SEL [0:7] signal is utilized to choose two adjacent clocks for interpolation. The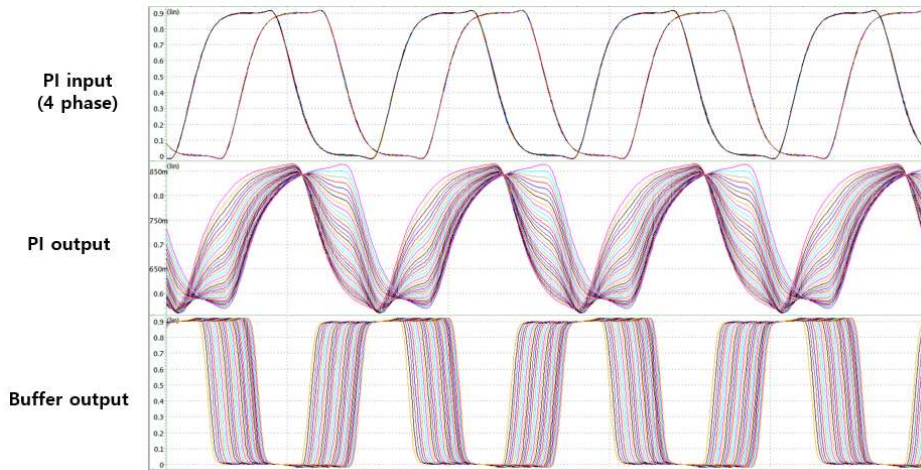 CS_SEL, 32-bit thermometer code, selects the number of current sources operating in each CML, determining the current flow in each CML and, in turn, the interpolation ratio. However, the resulting OUTP and OUTN clocks do not have full swing and have a poor slew. Therefore, an AC buffer and multiple inverters produce a sharp clock with full swing. Fig. 5.3 displays the simulation results of the PI, where the CS_SEL code is swept while feeding in 4-phase clock input. Additionally, Fig. 5.3 illustrates the waveforms when CK0 with a phase of 0° and CK45 with 45° are selected, and CK45 and CK90 are chosen, with the CS SEL code fully adjusted in each case.

The output signal from CML PI, OUTP, has a swing between 650 mV and 850 mV, but this is converted to full swing using the buffer (Buffer OUTP). The resulting
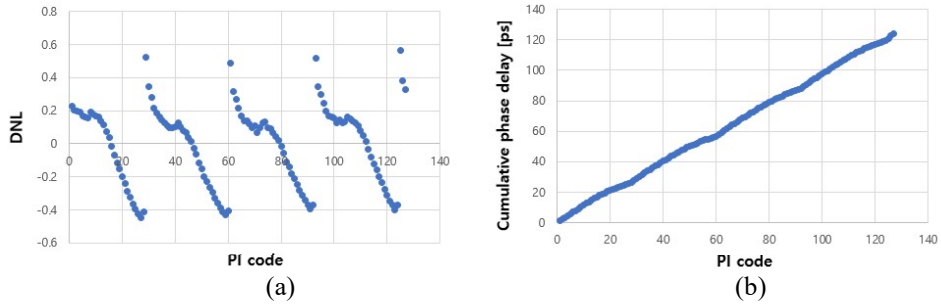
Fig. 5.4 (a) DNL and (b) Cumulative phase delay.

waveform is the final output of the PI, with a relatively constant phase difference for each input phase. Graphs depicting the degree of phase error and phase shift by code are presented in Fig. 5.4(a) and 5.4(b). According to Fig. 5.4(a), the maximum DNL at 8 GHz is 0.59, and increasing the PI code sequentially results in a linear increase in cumulative phase delay. Fig. 5.4(b) illustrates that the phase is linearly shifted by 360° across the whole code.

DNL (Differential Non-Linearity) is an essential measure of linearity, and it can be defined as follows.

$$\text{DNL} = (H_{(i)} - H_{ideal})/H_{ideal} \tag{5.1}$$

The value of H represents the difference in each phase, while the disparity between the Ideal H and the current H is expressed as DNL.

5.1.3 Parallel PRBS Generator
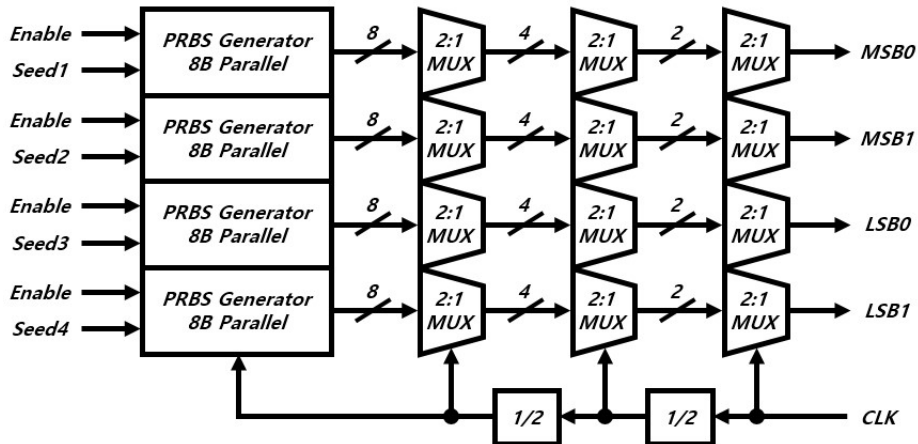
Fig. 5.5 Block diagram of PRBS generator.

The operation of the bridge chip can be verified using up to 8 DQ pads connected to its outside and a tester, while the Pseudo Random Binary Sequence (PRBS) generator circuit generates data input patterns internally without external DQ (Fig. 5.5). This thesis aims to design a transmitter for PAM4 signal output at a maximum speed of 48 Gb/s (24 Gsymbol/s).

Since the maximum frequency of the tester is 4 GHz, and the speed of data input through an external DQ is 8 Gb/s, two MSB and LSB signals must be input in parallel. Therefore, the PRBS generator should output up to 8 parallel PRBS sequence outputs operating at 8 Gb/s. The PRBS core circuit operates at 1 Gb/s, requiring clock dividers and 2:1 serializers to create an input pattern of 8 Gb/s. Each PRBS core circuit outputs 8 bits in parallel, with seed values set differently for various transition patterns based on the PAM4 signal.
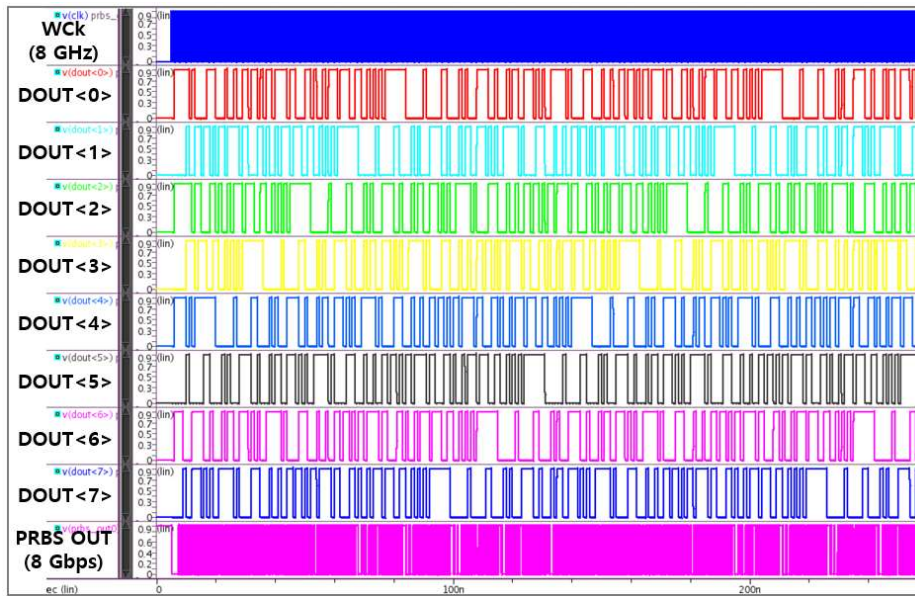
Fig. 5.7 Simulation result of 8-bit parallel PRBS generator.

$$D_K = D_{K-5} {}^\wedge D_{K-6}$$
$$D_{K+1} = D_{K-4} {}^\wedge D_{K-5}$$
$$D_{K+2} = D_{K-3} {}^\wedge D_{K-4}$$
$$D_{K+3} = D_{K-2} {}^\wedge D_{K-3}$$
$$D_{K+4} = D_{K-1} {}^\wedge D_{K-2}$$
$$D_{K+5} = D_K {}^\wedge D_{K-1}$$
$$D_{K+6} = D_{K+1} {}^\wedge D_K = D_{K-6} {}^\wedge D_{K-5} {}^\wedge D_K$$
$$D_{K+7} = D_{K+2} {}^\wedge D_{K+1} = D_{K-6} {}^\wedge D_{K-4}$$

(a) (b)

Fig. 5.6 (a) 8-bit parallel transition matrix. (b) Obtained equations from transition matrix.

The PRBS core circuit used the PRBS-7 sequence, with a length of 2^7-1 and a characteristic polynomial of $x^7+x^6=1$. The digitally designed PRBS core circuit outputs an 8-bit parallel signal with an operating speed of 1 Gb/s from each of the four PRBS Cores with a length of 2^7-1 (Fig. 5.6) [34], [35]. The output signal is converted

into a serial signal with an operating speed of 8 Gb/s through three 2:1 MUXs. Fig. 5.7 shows the simulation result of the 8-bit parallel output of PRBS core and the final serial output through 2:1 MUXs.

5.1.4 Digital-to-Analog Converter

The digital-to-analog converter (DAC) block diagram and its function are presented in Fig. 5.8. To ensure monotonicity, the DAC uses an R-ladder topology. A 7-bit control input VCW is used to select 127 voltage levels that are evenly distributed across the possible output range. Then, the maximum voltage range and corresponding
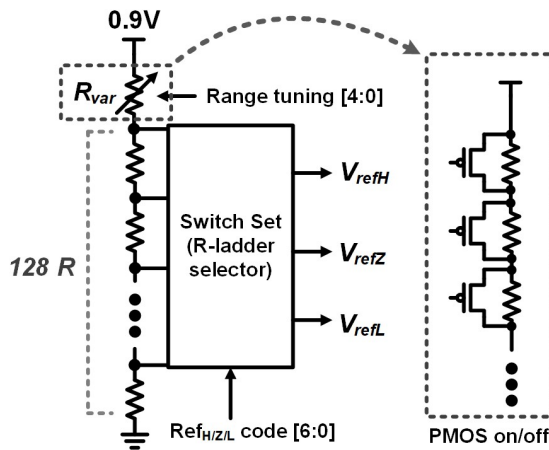
Fig. 5.8 Circuit implementation of R-ladder DAC.

Table 5.2 Post-layout simulation result of DAC voltage range.

	Ref range	1 step	Current	Code
TM All off	0 ~ 0.42V	~ 3.2 mV	130 uA	128 step
TM All on	0 ~ 0.65V	~ 5 mV	165 uA	128 step

voltage step per LSB is adjusted precisely using five pmos transistors in parallel with resistors regulated by a 5-bit digital code. Table 5.2 shows the available output voltage range and voltage step per LSB when all pmos transistors are turned on and off.

5.1.5 Eye-Opening Monitor

The proposed receiver utilizes two types of on-chip eye-opening monitors: the count-based EOM and the XOR-based EOM. The count-based EOM is responsible for identifying the optimal voltage reference and timing codes for the VDAC and PI, as shown in Fig. 5.9 and Fig. 5.10. During EOM training, predefined PRBS-7 data is transmitted from memory to the bridge receiver in a burst mode. The PAM4 signal is bisected into the upper, middle, and lower eyes using three parallel CTLEs, and three

Fig. 5.9 Flow chart of Count-based EOM.

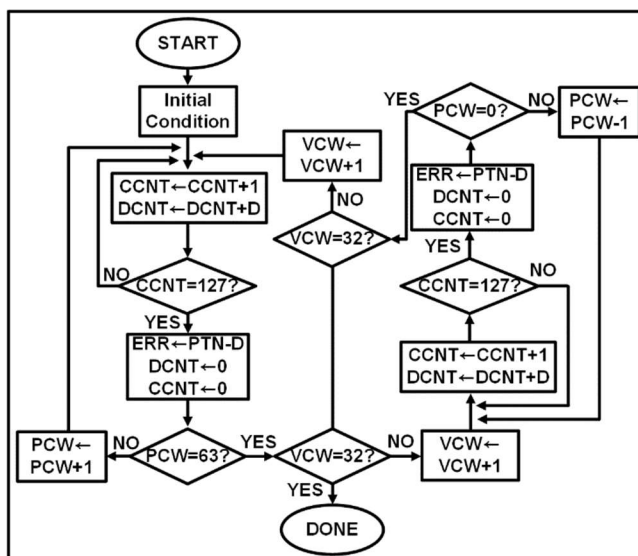data threshold voltages, REFH, REFZ, and REFL, are applied. The subsequent samplers produce a certain number of '1s' out of 127 bits in a single PRBS-7 cycle, which is counted to determine whether a particular voltage and timing code corresponds to a valid sampling point. The reference voltages are generated from VDAC with a 7-bit control input VCW, and the PI generates a clock with a 6-bit control input PCW.

The XOR-based eye opening monitor (EOM) is used to scan 2D data eye and measure bit error rate (BER). Once the DFE is adapted, the 3 error samplers are reused for scanning the 2D eye-opening monitor. These error samplers are XORed with 3 sampled data [36]. If the scanning point or an error sampler is situated inside the data eye, the outcome should be the same as that of the data sampler. By varying the voltage-controlled oscillator (VCW) and phase-controlled oscillator (PCW), 128x64 pixels are generated to illustrate a 2D eye-opening map. The circuit implementation of the XOR EOM is presented in Fig. 5.11. Furthermore, the BER can be measured in 2D using a pseudo-random binary sequence (PRBS) burst pattern.
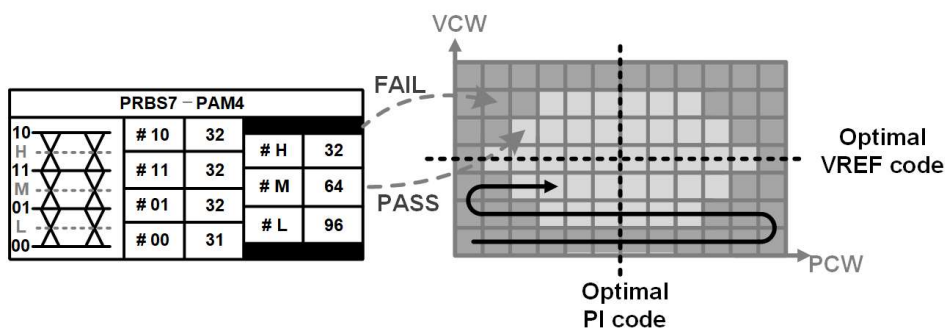
Fig. 5.10 Operation of Count-based EOM.
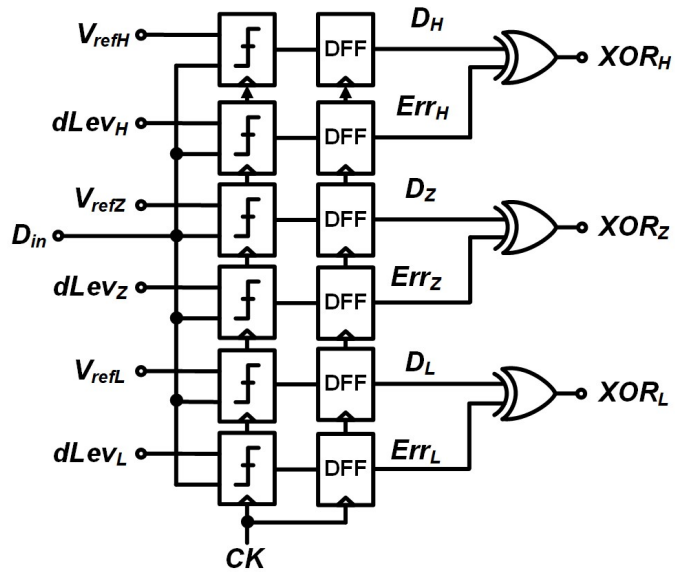
Fig. 5.11 Circuit implementation of XOR-EOM.

5.2 Measurement Setup
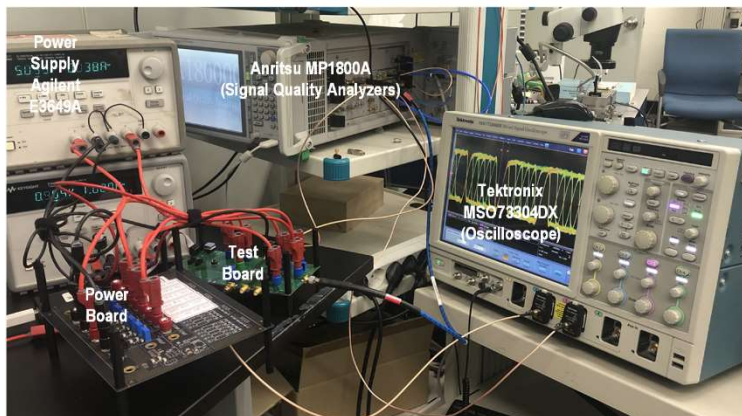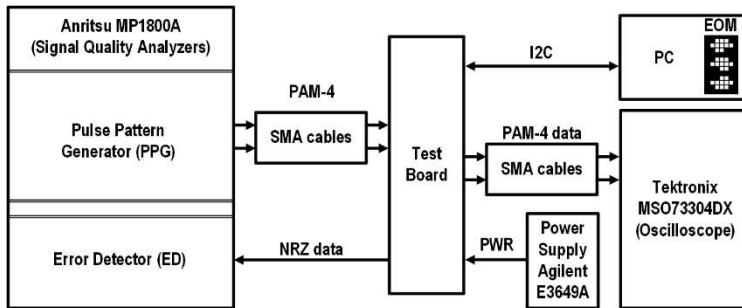
Fig. 5.12 Measurement setup.

The measurement setup utilized to measure the chip discussed in Chapters 3 and 4 of the thesis remains constant. The setup is illustrated in Fig. 5.12 and comprises a signal quality analyzer (Anritsu MU1800) functioning as both a pattern generator and an error detector.

The test environment for measuring the operation of the PAM4-Binary Bridge is similar to that of [5], and its configuration is shown in Fig. 5.12. Python code controls

the I2C, and communication between the PC and I2C occurs via Aardvark. The Anritsu MP1800A equipment is used to input a differential clock, and the Tektronix MSO73304DX oscilloscope is employed to measure the output waveform. However, since the Anritsu MP1800A equipment can only provide 2 NRZ data, it needs to provide more input data for this chip, which requires 8 inputs. Therefore, an internal PRBS generator is employed for measurement.

To measure the operation of the PAM4 receiver, a bit error tester (BERT) produces two binary PRBS data that correspond to MSB and LSB. To make the PAM4 signal, a passive power combiner (HL9404 BALUN) is employed, as demonstrated in Fig. 5.12 [37]. MSB is applied to one port of the passive power combiner, and LSB passed through a 6 dB attenuator, is applied to the other. The channel insertion loss at the Nyquist frequency of 12 GHz in the SMA cables and PCB is measured to be approximately 7dB. The recovered and deserialized NRZ DQ data is fed back to the error detector to measure BER.

The eye diagram of the transmitted PAM4 input signal is measured using the Tektronix MSO73304DX oscilloscope, as shown in Fig. 5.13. The configurations for the DUT are adjusted by an external PC using the I2C protocol. The Agilent E3649A DC power supply is used to input supply power for both the I2C and transmitter.
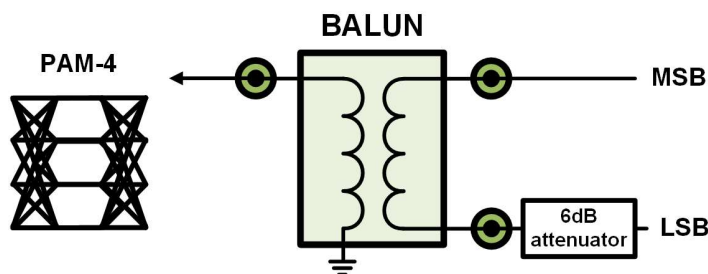
Fig. 5. 1 Production of PAM-4 signal.

Chapter 6

Conclusions

This thesis proposes the PAM4-Binary Bridge that incorporates all the required functions to test a next-generation high-speed PAM4 memory using a popularly used low-speed NRZ tester.

The first PAM4-Binary Bridge chip supports 32-Gb/s operations per DQ pin, which is twice the speed of the current-generation highest-speed memory. It will play a key role in testing newly developed memory interfaces in a timely manner. The low-speed data transmitted from the NRZ tester to the bridge are converted into high-speed PAM4 data through half-rate clock control, forwarded to the memory, and vice-versa. The ground-terminated PAM4 driver provides the single-ended output by controlling the output current with a 2-tap feed-forward equalizer, achieving the RLM of 0.95. To minimize the offset at the PAM4 receiver, the offset cancellation circuit with an offset of 2.76 mV consisting of a CTLE and sampling latches is employed, and the horizontal margin of the received PAM4 signal is 50% for BER<10^{-9}. The bridge fabricated in the 40-nm CMOS technology occupies an active area of 1.6 mm^2 and dissipates 132 mW.

The second chip presents a 48 Gbps PAM4 memory interface with a level mismatch adjustment capability for a high-speed PAM4 memory/tester bridge. The level-

adjustable PAM4 TX is designed as a voltage mode CMOS driver and improves the RLM through a calibration circuit. The RX achieves BER less than 10^{-12} through equalizers such as parallel CTLEs and 1-tap DFE. The bridge operates at 48 Gbps per pin and consumes 1.85 pJ/bit and 2.97 pJ/bit for the write and read modes of the PAM4 memory, respectively. The proposed bridge is fabricated in 40 nm CMOS technology, occupying 2.13×1.098 mm^2. The WCK signaling scheme applied to the GDDR memory was used to operate the two bridge chips. An all-digital PLL integrated into the bridge doubles the up to 6 GHz WCK used as a forwarded clock for the graphic memory. For the measurement of the bridge, a newly proposed count-based PAM4 eye-opening monitor is also proposed to find the optimal codes for the maximum eye opening using the PRBS7 data sequence.

Bibliography

[1] Cisco.com. [Online] [Accessed on 3rd Apr. 2022]. Available: *https://www.cisco.com/c/en/us/solutions/collateral/executive-perspec-tives/annual-internet-report/white-paper-c11-741490.html*.

[2] Micron.com. [Online] [Accessed on 24th Nov. 2022]. *https://www.mi-cron.com/about/blog/2019/june/ddr5-the-next-step-in-system-level-perfor-mance.*

[3] T. M. Hollis *et al.*, "An 8Gb GDDR6X DRAM Achieving 22Gb/s/pin With Single-Ended PAM4 Signaling," *IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers*, 2021, pp. 348-350.

[4] H. Jun, S *et al.*, "High-Bandwidth Memory (HBM) Test Challenges and So-lutions," in *IEEE Design & Test*, vol. 34, no. 1, pp. 16-25, Feb. 2017.

[5] D. Yun *et al.*, "A 32-Gb/s PAM4-Binary Bridge with Sampler Offset Cancel-lation for Memory Testing," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 9, pp. 3749-3753, Sept. 2022.

[6] H. Jin *et al.*, "A 24Gb/s/pin PAM-4 Built Out Tester chip enabling PAM-4 chips test with NRZ interface ATE," 2021 *IEEE Asian Solid-State Circuits Conference (A-SSCC)*, 2021, pp. 1-3.

[7] 최정환. (2012). 고속 DRAM interface. 전자공학회지, 39(7), 20-26.

[8] M. Bassi, F. Radice, M. Bruccoleri, S. Erba and A. Mazzanti, "A High-Swing 45 Gb/s Hybrid Voltage and Current-Mode PAM-4 Transmitter in 28 nm CMOS FDSOI," in *IEEE J. Solid-State Circuits*, vol. 51, no. 11, pp. 2702-2715, Nov. 2016.

[9] R. Stephens, PAM4: Symbol levels and voltage compression meas-urements. EDN Network [Online] [Accessed on 24th Nov. 2022]. Available: *http://www.edn.com*.

[10] C. Hyun *et al.*, "A 20Gb/s Dual-Mode PAM4/NRZ Single-Ended Transmitter with RLM Compensation," in *IEEE Int. Symposium on Circuits and Systems. (ISCAS)*, 2019, pp. 1-4.

[11] J. -S. Heo *et al.*, "A 5Gb/s/pin 16Gb LPDDR4/4X Reconfigurable SDRAM with Voltage-High Keeper and a Prediction-based Fast-tracking ZQ Calibration," *2019 Symposium on VLSI Circuits*, 2019, pp. C114-C115.

[12] N. Dikhaminjia *et al.*, "PAM4 signaling considerations for high-speed serial links," 2016 *IEEE International Symposium on Electromagnetic Compatibility (EMC)*, 2016, pp. 906-910.

[13] PCI Express® Base Specification Revision 6.0 Version 0.7, 2020, [online] Available: *https://pcisig.com/specifications*.

[14] M. Kossel *et al.*, "A T-coil-enhanced 8.5 Gb/s high-swing SST transmitter in 65 nm bulk CMOS with < 16 dB return loss over 10 GHz bandwidth," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2905–2920, 2008.

[15] B. Razavi, "The Design of an Equalizer—Part One [The Analog Mind]," in *IEEE Solid-State Circuits Magazine*, vol. 13, no. 4, pp. 7-160, Fall 2021.

[16] B. Razavi, "The StrongARM Latch [A Circuit for All Seasons]," in *IEEE Solid-State Circuits Magazine*, vol. 7, no. 2, pp. 12-17, Spring 2015.

[17] Y. -J. Kim *et al.*, "A 16Gb 18Gb/S/pin GDDR6 DRAM with per-bit trainable single-ended DFE and PLL-less clocking," in *IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers*, 2018, pp. 204-206.

[18] J. L. Zerbe *et al.*, "Equalization and Clock Recovery for a 2.5-10-Gb/s 2-PAM/4-PAM Backplane Transceiver Cell," *IEEE J. Solid-State Circuits*, vol. 38, pp. 2121–2130, Dec. 2003.

[19] GDDR6 SGRAM Specification (JESD250C), JEDEC Standard, JEDEC solid state technology association, Feb. 2021.

[20] S. Saxena *et al.*, "A 2.8 mW/Gb/s, 14 Gb/s Serial Link Transceiver," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1399–1411, May 2017.

[21] W. Bae *et al.*, "A Supply-Scalable Serializing Transmitter with Controllable Output Swing and Equalization for Next Generation Standards," *IEEE Trans. Industrial Electronics*, vol. 65, no. 7, pp. 5979–5989, 2018.

[22] K. L. Chan *et al.*, "A 32.75-Gb/s Voltage-Mode Transmitter With Three-Tap FFE in 16-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 52, no. 10, pp. 2663–2678, Oct. 2017.

[23] Y. Song and S. Palermo, "A 6-Gbit/s Hybrid Voltage-Mode Transmitter With Current-Mode Equalization in 90-nm CMOS," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 8, pp. 491–495, Aug. 2012.

[24] A. Roshan-Zamir, O. Elhadidy, H. Yang, and S. Palermo, "A Reconfigurable 16/32 Gb/s Dual-Mode NRZ/PAM4 SerDes in 65-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 52, no. 9, pp. 2430–2447, Sep. 2017.

[25] J. F. Bulzacchelli *et al.*, "A 10-Gb/s 5-Tap DFE/4-Tap FFE Transceiver in 90-nm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2885–2900, Dec. 2006.

[26] K. Oh *et al.*, "A 5-Gb/s/pin Transceiver for DDR Memory Interface With a Crosstalk Suppression Scheme," *IEEE J. Solid-State Circuits*, vol. 44, no. 8, pp. 2222–2232, Aug. 2009.

[27] R. Navid *et al.*, "A 40 Gb/s Serial Link Transceiver in 28 nm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 50, no. 4, pp. 814–827, Apr. 2015.

[28] M. Chen *et al.*, "A Fully-Integrated 40-Gb/s Transceiver in 65-nm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 47, no. 3, pp. 627–640, Mar. 2012.

[29] P. Nuzzo *et al.*, "Noise Analysis of Regenerative Comparators for Reconfigurable ADC Architectures," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 6, pp. 1441-1454, July 2008.

[30] T. Kobayashi *et al.*, "A current- controlled latch sense amplifier and a static

power-saving input buffer for low-power architecture," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, pp. 523–527, Apr. 1993.

[31] S. Gondi and B. Razavi, "Equalization and Clock and Data Recovery Techniques for 10-Gb/s CMOS Serial-Link Receivers," *IEEE J. Solid-State Circuits*, vol. 42, pp. 1999-2011, Sept. 2007.

[32] G.Van der Plas, S. Decoutere, and S. Donnay, "A 0.16pJ/Conversion-step 2.5mW 1.25GS/s 4b ADC in a 90nm Digital CMOS Process,'' *IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers*, 2006, pp. 566-567.

[33] Y. -U. Jeong *et al.*, "A 0.64-pJ/Bit 28-Gb/s/Pin High-Linearity Single-Ended PAM-4 Transmitter With an Impedance-Matched Driver and Three-Point ZQ Calibration for Memory Interface," *IEEE J. Solid-State Circuits*, vol. 56, pp. 1278-1287, April 2021.

[34] M. Chen and C. K. Yang, "A low-power highly multiplexed parallel PRBS generator," *Proceedings of the IEEE 2012 Custom Integrated Circuits Conference*, 2012, pp. 1-4.

[35] O'Reilly, "Series-parallel generation of m-sequences," the Radio and Electronic Engineer, Apr. 1975, pp. 171-176.

[36] P. -J. Peng *et al.*, "A 56Gb/s PAM-4/NRZ transceiver in 40nm CMOS," *IEEE International Solid-State Circuits Conference (ISSCC)*, 2017, pp. 110-111.

[37] S. Roh *et al.*, "A 64-Gb/s PAM-4 Receiver With Transition-Weighted Phase Detector," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 9, pp. 3704-3708, Sept. 2022.

[38] H. Lee *et al*., "A 16.8 Gbps/Channel Single-Ended Transceiver in 65 nm CMOS for SiP-Based DRAM Interface on Si-Carrier Channel," in *IEEE J. Solid-State Circuits*, vol. 50, pp. 2613-2624, Nov. 2015.

초 록

　기계 학습 및 A/I 와 같은 고성능 컴퓨팅 애플리케이션에는 높은 메모리 대역폭이 필요하다. 다단계 시그널링은 DRAM 의 대역폭 요구를 충족시키기 위해 고려되고 있지만, 특히 대량 생산되는 DRAM 제품의 경우 상당한 인프라 구조 변경이 필요하다. DRAM 제조업체는 Non-Return-to-Zero 신호를 평가하기 위한 대규모 시설을 갖추고 있으므로 다중 레벨 신호 지원을 구현하려면 비용과 시간이 많이 소요되는 테스트 시설 변경이 필요하다. 이러한 문제를 해결하기 위해 저성능 테스트 장비의 입/출력 데이터를 고속 PAM4 신호로 변환한 다음 DRAM 으로 전송하는 브리지 칩이 제안되었다.

　첫 번째 칩의 경우 차세대 메모리 테스트를 위한 32Gb/s PAM4-바이너리 브리지가 제공된다. 브리지는 저속 NRZ 테스터를 사용하여 고속 PAM4 메모리를 평가하는 데 필요한 모든 기능을 통합한다. NRZ 테스터에서 브리지로 전송되는 저속 데이터는 Half-rate 클록제어를 통해 고속 PAM4 데이터로 변환되어 메모리로 전달되고 그 반대도 마찬가지이다. 접지 종단 PAM4 드라이버는 2-tap feed-forward 이퀄라이저로 출력 전류를 제어하여 단일 종단 출력을 제공하여 비율 레벨 불일치(RLM) 0.95 를 달성한다. PAM4 수신기에서 오프셋을 최소화하기 위해 CTLE 와 샘플링 래치로 구성된 2.76mV 오프셋의 오프셋 제거 회로를 사용하고 수신된 PAM4 신호의 수평 마진은 BER<10^{-9}에 대해 50% 이다. 브리지에 통합된 전체 디지털 PLL 은 그래픽 메모리용 전달 클록으로 사용되는 4GHz WCK 를 두 배로 늘린다. 카운트 기반 PAM4 아이 오프닝 모니터

는 PRBS7 데이터 시퀀스를 사용하여 최대 아이 오프닝에 대한 최적의 코드를 찾기 위해 제안된다. 40nm CMOS 기술로 제작된 브리지는 $1.6mm^2$의 영역을 차지하고 132mW 의 전력을 소모 한다.

두 번째 칩은 고속 PAM4 메모리/테스터 브리지에서 사용하기 위한 레벨 불일치 조정 기능이 있는 48Gbps PAM4 메모리 인터페이스를 제공한다. 브리지는 저속 NRZ 테스터를 사용하여 고속 PAM4 메모리를 테스트하고 검증하는 데 필요한 모든 기능을 통합한다. 레벨 조정이 가능한 PAM4 TX 는 전압 모드 CMOS 드라이버로 설계되었으며 보정 회로를 통해 RLM 을 개선한다. RX 는 병렬 CTLE 및 1 탭 DFE 와 같은 이퀄라이저를 통해 10^{-12} 미만의 BER 을 달성한다. 브리지는 핀당 48Gbps 에서 작동하고 PAM4 메모리의 쓰기 및 읽기 모드에 대해 각각 1.85pJ/bit 및 2.97pJ/bit 를 소비한다. 제안된 브리지는 $2.13x1.098mm^2$ 를 차지하며 40nm CMOS 기술로 제작되었다.