Ultra-High-Speed Digital-to-Analog Converter for Optical Communications

Von der Fakultät Informatik, Elektrotechnik und Informationstechnik
der Universität Stuttgart zur Erlangung der Würde
eines Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von

Hao Huang

aus Liuzhou, China

Hauptberichter:	Prof. Dr.-Ing. Manfred Berroth
Mitberichter:	Prof. Dr.-Ing. Friedel Gerfers

Tag der mündlichen Prüfung: 23. September 2019

Institut für Elektrische und Optische Nachrichtentechnik

der Universität Stuttgart

2019

Table of Contents

Symbols and Abbreviations

Abbreviations

AC	Alternating current
ADC	Analog-to-digital converter
BER	Bit error rate
BiCMOS	Bipolar-CMOS
BJT	Bipolar junction transistor
CLK	Clock
CML	Current mode logic
CMOS	Complementary metal oxide semiconductor
CO-DWDM	Coherent dense wavelength division multiplexing
CS-DAC	Current-steering digital-to-analog converter
D/A	Digital-to-analog
DAC	Digital-to-analog converter
DP	Dual polarization
dB	Decibel
dBc	Decibels relative to the carrier
DC	Direct current
DNL	Differential non-linearity
DSP	Digital signal processing
EM	Electromagnetic
ENOB	Effective number of bits
ESD	Electrostatic discharge
FEC	Forward error correction
FET	Field effect transistor
FS	Full-scale
FPGA	Field programmable gate array
GCPW	Grounded coplanar wave-guide
GPIB	General purpose interface bus
GPIO	General-purpose input/output
GS	Giga-samples
HD	Harmonic distortion
INL	Integral non-linearity
InP	Indium Phosphide

IM-DD	Intensity-modulation direct-detection
IP	Internet protocol
ISI	Inter-symbol interference
I/Q	In-phase and quadrature-phase
LSB	Least significant bit
LVDS	Low voltage differential signaling
NMOS	N-channel metal oxide semiconductor
NRZ	Non-return-to-zero
MOSFET	Metal oxide semiconductor field effect transistor
MSB	Most significant bit
MUX	Multiplexer
OFDM	Orthogonal frequency division multiplexing
OSR	Oversampling ratio
PAM	Pulse-amplitude modulation
PBS	Polarization beam splitter
PC	Personal Computer
PCB	Printed Circuit Board
PMOS	P-channel metal oxide semiconductor
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
RF	Radio frequency
RX	Receiver
RZ	Return-to-zero
SAR	Successive approximation register
SC	Switched capacitor
SFDR	Spurious-free dynamic range
SiGe	Silicon Germanium
SMP	Sub-miniature push-on
SNDR	Signal-to-noise-and-distortion ratio
SNR	Signal-to-noise ratio
SRAM	Static random access memory
THD	Total harmonic distortion
TIA	Transimpedance amplifiers
TI-DAC	Time-interleaved DAC
TX	Transmitter
WDM	Wavelength division modulation

Symbols

A_{out}	Analog output signal (current, voltage or charge)
A_{FS}	Full-scale analog output signal (current, voltage or charge)
$A_{V_{\text{T0}}}$	Technology-dependent constant of the threshold voltage V_{T0}
A_v	Voltage gain
A_β	Technology-dependent constant of the current amplification factor β
$b_0, b_1, b_2, \cdots b_{N-1}$	The bits of the input digital code of a DAC
C	Capacitance
C'	Capacitance per unit length of a transmission line
C_{ox}	Oxide capacitance
C_{sub}	Substrate capacitance
D_{FS}	Full-scale digital input value
f	Frequency
f_{c}	Cut-off frequency
f_{s}	Sampling frequency
f_{sig}	Signal frequency
g_m	Small-signal transconductance
g_{DS}	Drain-source conductance
G'	Conductance per unit length of a transmission line
$h(t)$	Impulse response
$H(f)$	Transfer function
I	Current
I_{out}	Output current
I_D	Drain current
I_{LSB}	LSB current
$k_{1,2,3\ldots.}$	Mutual coupling factor of the Inductors
L	Length or inductance
L'	Inductance per unit length of a transmission line
$p(\varepsilon_Q)$	Probability of the quantization error ε_Q
P_{sig}	Power of the signal
P_{sin}	Power of the sine wave
$P_{\text{sp\_max}}$	Power of the highest spur
P_k	Power of the k$^{\text{th}}$ harmonic distortion
P_{Q}	Power of the quantization noise
P_{n}	Power of the noise

Q	Quality factor
R	Resistance
R_{L}	Load resistance
R'	Resistance per unit length of a transmission line
r_{o}	Small signal output resistance
V	Voltage
V_{DD}	Positive supply voltage
V_{GS}	Gate-Source voltage
V_{in}	Input voltage
V_{out}	Output voltage
V_{ref}	Reference voltage
V_{swing}	Output swing
V_{th}	Threshold voltage
V_{th0}	Threshold voltage at 0 V substrate voltage
W	Width
$x(t)$	Continuous-time signal
$x_S(t)$	Sampled signal
$x_{SH}(t)$	Sampled and hold signal
$X(f)$	Fourier transformation of a continuous-time signal $x(t)$
$X_S(f)$	Fourier transformation of a sampled signal $x_S(t)$
$X_{SH}(f)$	Fourier transformation of a sampled and hold signal $x_{SH}(t)$
Z_0	Characteristic impedance or unit current cell impedance
β	Current amplification factor of a transistor
Δ_{Q}	Quantization interval
Δ_{ref}	The reference quantity (voltage, current, etc.) of a DAC
α_{s}	Attenuation in micro strip line due to skin effect
ε_Q	Quantization error
ε_{r}	Relative permittivity
v_{p}	Wave velocity
σ	Standard deviation
τ_{D}	Propagation delay
$\phi(f)$	Phase angle
ω	Angular frequency

Abstract

The optical communication network is continuously developing to match the demand of the growing global data traffic. For today's transmission systems, transceivers based on high performance digital signal processors (DSPs) integrated with high-speed data converters are the key to realizing coherent networks, which are essential to ultra-high data-rate communications in long-haul networks (e.g., city-to-city, transcontinental fiber networks), metro networks and data centers (fiber networks within and between data centers). For next generation 100 Gb/s and 400 Gb/s fiber optical transceivers, DACs with low or medium resolutions (around 6 to 8 bit) and ultra-high sampling rates over 50 GS/s are required to enable pre-equalizations and (multi-level) modulation formats like dual-polarization (DP) quadrature phase-shift keying (QPSK), quadrature amplitude modulation (QAM) and orthogonal frequency-division multiplexing (OFDM).

The objective of this work is to explore the design techniques for ultra-high sampling-rate DACs and to design a highest speed (up to 100 GS/s) DAC using advanced CMOS technology for the next generation optical communications. To achieve such a high sampling rate, two different time-interleaving techniques are employed: analog-multiplexing technique and parallel-path technique. The analog-multiplexing technique combines two sub-DACs' current outputs with an analog multiplexer (MUX). The MUX operates at double the sub-DAC sampling rate, hence doubling the overall sampling rate and halving the pulse width. The halved pulse width doubles the bandwidth of the sinc-function envelope, increasing the limit of the analog bandwidth. By directly summing up the two sub-DAC outputs using the parallel-path time-interleaving technique, the overall sampling rate is doubled but the frequency response does not change. Even though there is no gain of the bandwidth, the double sampling rate moves the images originally located in the second Nyquist frequency band of the sub-DACs to much higher frequencies far away from the signal, hence greatly reducing the requirement of the output reconstruction filter. The output signal frequency range can be extended much closer to the Nyquist frequency without interfering adjacent frequency bands, thus increasing the useful range of the output spectrum. One of the drawbacks is the large output capacitance introduced by the parallel structure, which reduces the output bandwidth. Therefore, the DAC output stage is reconstructed in a distributed style to increase the bandwidth by distributing the large output capacitances over artificial transmission lines. One of the major benefits of using advanced CMOS technologies is the ability to integrate large number of transistors due to their small size. Therefore, a 1 kByte full-custom designed memory with minimum area requirement is integrated on the same chip to generate data pattern to the DAC, which greatly reduces the complexity of measurements and the overall cost. Finally, an 8 bit 100 GS/s DAC chip with a 1 kByte memory is implemented in 28 nm CMOS technology to evaluate the concept.

The static and dynamic performances of the DAC are characterized by measurement. The measured integrated nonlinearity (INL) and the differential nonlinearity (DNL) are 1.4 and 2.0 LSB (least significant bits), respectively. The 3 dB bandwidth is over 13 GHz at 100 GS/s.

The effective number of bits (ENOB) ranges from 5.3 bit to 3.2 bit from low frequencies up to Nyquist frequency (24.9 GHz) at 100 GS/s, respectively. A 50 Gbaud pulse amplitude modulation with 4 levels (PAM4) and a 40 Gbaud PAM8 eye-diagram show an electrical transmission capability up to 120 Gbit/s. In an optical transmission experiment, error-free transmission rates up to 45 Gbit/s are demonstrated using the DAC. Higher transmission rate is limited by the optical measurement setup. To the best of the author's knowledge, this is the world's first DAC that achieves a sampling rate of 100 GS/s.

Kurzfassung

Das rasante Wachstum des globalen Datenverkehrs führt zu einer starken Belastung der heutigen Kommunikationsnetze und erfordert ständige Kapazitätserweiterungen. Um mit der wachsenden Bandbreitenanforderung Schritt zu halten, müssen sich die optische Übertragungsnetze immer weiterentwickeln. Die optischen Übertragungssysteme, insbesondere die kohärenten optischen Übertragungssysteme, haben eine zentrale Funktion in Langstrecken-Netzwerken (z.B. bei der Verbindungen zwischen Regionen und zwischen Kontinenten), regionale Netzwerken und Rechenzentren-Netzwerken. Die Sender und Empfänger, die auf Hochleistungs-Digitaler-Signalprozessoren (DSP) mit integrierten schnellen Analog-Digital- und Digital-Analog-Wandlern (engl. analog-digital and digital-analog converter, ADC and DAC) basieren, sind eine Schlüsselkomponente für die Realisierung des kohärenten optischen Übertragungssystems. Für Sender und Empfänger der nächsten Generation mit Übertragungsraten von 100 Gbit/s und 400 Gbit/s, sind schnelle DACs mit einer niedrigen bis mittleren Auflösung (6 bis 8 bit) und einer sehr hohen Datenrate über 50 GS/s notwendig, um komplexe Vorentzerrungen und Modulationsverfahren, wie z.B. Quadraturphasenumtastung (engl. quadrature phase-shift keying, QPSK), Quadratur-Amplituden-Modulationen (engl. quadrature amplitude modulation, QAM) und orthogonale Frequenzmultiplexverfahren (engl. orthogonal frequency-division multiplexing, OFDM) zu realisieren.

In der vorliegenden Dissertation wird die Schaltungstechnik für schnelle DACs untersucht und ein DAC mit einer Umsetzungsrate bis zu 100 GS/s und 8 bit nomineller Auflösung in 28 nm CMOS Technologie entworfen. Solche hohen Umsetzungsraten werden realisiert durch die Verwendung von Zeitverschachtelungstechniken. Dabei geht es um zwei unterschiedliche Verfahren: Das Analogmultiplex- und das Parallelpfadverfahren. Bei den Analogmultiplexverfahren sind die Stromausgänge von zwei Teil-DACs durch einen analogen Multiplexer kombiniert. Dadurch ist die Umsetzungsrate des kombinierten DACs verdoppelt und führt dadurch zu einem flacheren sinc-förmigen Amplitudengang, was eine höhere Ausgangsbandbreite bedeutet. Im Vergleich dazu sind bei dem Parallelpfadverfahren die Stromausgänge von zwei Teil-DACs direkt verbunden. Zwar ist die Umsetzungsrate auch verdoppelt, der Amplitudengang ist jedoch, genau wie bei den einzelnen Teil-DAC, gleich geblieben. Obwohl es mit dieser Konstruktion keine höhere analoge Bandbreite gewonnen werden kann, ist es so möglich, die Spiegelfrequenzen von den nutzbaren Signalfrequenzen weit weg zu schieben. Im Vergleich zu den einzelnen Teil-DAC sind die Anforderungen an den Ausgangsfilter stark reduziert. Die nutzbaren Signalfrequenzen können erhöht werden, ohne das benachbarte Frequenzband zu stören. Die mehrfache Parallelisierung der Teil-DACs erzeugt große parasitäre Kapazitäten am DAC-Ausgang und führt damit zu einer geringeren Bandbreite. Um die Ausgangsbandbreite zu erhöhen, ist die Ausgangsstufe mit einer verteilten Struktur konstruiert. Dabei sind das Stromsummationsnetzwerk und die Taktverteilung an der DAC-Ausgangsstufe mittels künstlich konstruierten Leitungen realisiert, um die parasitären

Kapazitäten auf die künstliche Leitung zu verteilen. Für die Charakterisierung des DACs ist ein 1 kByte Speicher integrierte, der zyklisch ausgelesen werden kann, um die Eingangsdatenströme für den DAC zu erzeugen. Dank der fortgeschrittenen CMOS Technologie kann der Speicher mit minimaler Chipfläche realisiert werden, womit der Messaufbau stark vereinfacht und die Kosten deutlich reduziert werden, im Vergleich zu der Lösung mit externem Signalgenerator.

Die statischen und dynamischen Eigenschaften des DACs sind durch Messungen charakterisiert. Die differenzielle Nichtlinearität (DNL) und die integrale Nichtlinearität (INL) betragen jeweils 2,0 und 1,4 LSB (engl. least significant bits). Die maximale Bandbreite beträgt 13 GHz bei einer Abtastrate von 100 GS/s. Die effektive Anzahl von Bits (engl. effective number of bits, ENOB) beträgt 5,3 bit bei niedrigen Ausgangsfrequenzen und reduziert sich auf 3,2 bit bei 24,9 GHz mit einer Abtastraten von 100 GS/s. Das 40 Gbaud Augendiagramm der Datenmuster, welches von achtstufiger Pulsamplitudenmodulation (PAM8) moduliert ist, zeigt die Realisierbarkeit einer Übertragungsrate von bis zu 120 Gbit/s. Der DAC wurde auch in einem optischen Übertragungsexperiment eingesetzt und hat dabei eine Übertragungsrate von bis zu 45 Gbit/s erreicht. Nach sorgfältiger Prüfung und bestem Wissen des Autors ist dieser DAC der weltweit erste DAC, der eine Abtastrate von 100 GS/s erreicht hat.

1 Introduction

1.1 Motivation

Smartphones, cloud services, online videos and games are continually driving the growth of the global data traffic. Global internet protocol (IP) traffic is expected to increase nearly threefold from 2016 to 2021, at a rate of 24% per year, according to Cisco Visual Network Index [1]. This results in a high demand on the development of transmission technologies for additional capacity, higher spectral efficiency and lowering the cost per bit in the communication networks. One of the key points to start with is the optical communication network, which is the foundation of the long-distance data transmission, due to its many advantages, e.g., huge bandwidth, extremely low signal attenuation, low signal distortion, low power, low material usage, small space requirement and low cost.

The optical bandwidth of a single-mode fiber (tens of THz) is much larger than the bandwidth of electrical data transceivers (tens of GHz). To fully utilize the optical bandwidth with existing electrical transceivers, wavelength-division modulation (WDM) is widely used [2]. The idea is to subdivide the optical transmission spectrum into a number of non-overlapping wavelength bands, with each wavelength driven by an electrical transceiver at its maximum data rate and multiplexed to one optical fiber. With this technique, the data rate per fiber can be improved by increasing the number of channels and increasing the data rate per channel. Over the past years, the optical transmission rate is moved from single channel system of 1 Gb/s with intensity-modulated direct-detection (IM-DD) to today's 100 Gb/s or 200 Gb/s per wavelength and 100 wavelengths per fiber with coherent dense wavelength division multiplexing (CO-DWDM) [3]. The next-generation 400-G systems are also in development and will be available very soon.

The block diagram of a single wavelength coherent optical transceiver is shown in Fig. 1. On the transmitter (TX) side, the TX DSP receives the data from the client side. The data is encoded with forward-error-correction (FEC) and digitally modulated to two complex signals (in-phase (I) and quadrature-phase (Q)) on both the X- and Y-polarizations. The data is digitally filtered before being sent to the four DACs. The output of the DACs are linearly amplified to drive the optical modulators. On the receiver (RX) side, the optical signal is converted to four electrical signals through a polarization beam splitter (PBS), optical hybrids and balanced photo detectors. The electrical signals are then amplified by the transimpedance amplifiers (TIAs) and the linear amplifiers, and are converted to digital signals by the analog-to-digital converters (ADCs). After the digital filtering, the clock, carrier and polarization recoveries are performed. Finally, the signals are demodulated before FEC decoding and are sent to the client side.

The CMOS transceiver chip integrated with powerful CMOS DSPs and high speed ADCs and DACs is the key component for most of the optical transceivers [4]. While keeping the optical channel spacing, e.g., 25 GHz or 50 GHz for DWDM according to TU G.694.1 [5], multilevel modulation, e.g., QAM, as well as frequency multiplexing technique, e.g., OFDM, at the transmitter is mandatory for targeting high bit rate per wavelength [6]. At the same time, the

lowest cost of transmission is usually achievable with each electro-optic conversion operating at its maximum capacity. This requires a maximum DAC sampling rate and bandwidth. Therefore, low-power and high-speed DACs with a high bandwidth and a moderate resolution are required.

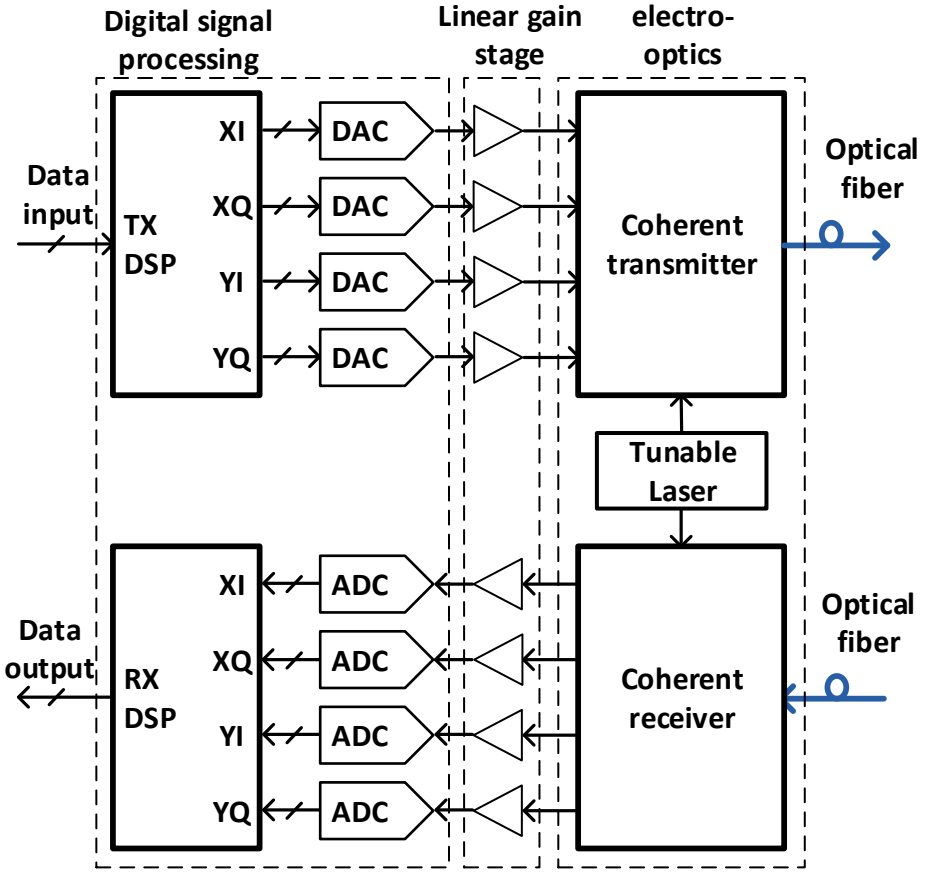


Fig. 1: An optical coherent transceiver [4]

1.2 State-of-the-Art DACs

Driven by the development of the next-generation optical transceivers, several works on high-speed DACs have demonstrated success in achieving high sampling rates in the past few years. In [6–9], DACs using Silicon Germanium (SiGe) Bipolar-CMOS (BiCMOS) technologies with a resolution of 6 bit and with sampling rates from 30 GS/s to 72 GS/s are reported. In [8], the DAC achieves a 45 GS/s sampling rate and a 45 GHz bandwidth. With 6 V differential peak-to-peak output swing, it is capable to directly drive optical modulators without extra buffering, hence saving a lot of power. Despite the excellent analog performances of the SiGe bipolar transistors compared to that of CMOS technologies, e.g., larger transcondutance, higher voltage gain, higher bandwidth, higher breakdown voltage, better matching and lower noise [10], the integration with DSPs is the biggest issue. Most DSPs with high speed and high complexity are implemented in the latest possible CMOS nodes (e.g., 28 nm to 7 nm etc.,) for reducing digital power and chip area. These CMOS nodes are not available in the latest BiCMOS technology. A multi-chip solution can provide flexibilities to choose the best technology for each digital

and analog circuit, but the high power consumptions for the interface between chips, low yield and the high cost make it unsuitable for mass production [11].

Besides SiGe technologies, alternative technologies like Indium Phosphide (InP) also show superior performances due to higher electron mobility, semi-insulated substrate and higher breakdown voltage. 6 bit DACs implemented in InP technologies have achieved a sampling rate up to 80 GS/s [12, 13], which can potentially directly drive the optical modulator and can be integrated with other optical components. However, the yield in InP technology is much lower compared to SiGe and CMOS [14].

Thanks to aggressive progress and continuous scaling on CMOS technologies, the speed of today's MOSFET can also be competitive with bipolar transistors. The peak transit frequency of a 28 nm MOSFET can achieve more than 300 GHz [15], which is comparable to or even higher than the SiGe and InP technologies mentioned above. This makes the design of ultra-high-speed DACs feasible. Recently, CMOS DACs with sampling rates over 50 GS/s have been reported in [16–19]. In 2011, the first CMOS DACs over 50 GS/s sampling rate was reported in [16], which had a 6 bit resolution and was implemented in 65 nm CMOS. An 8 bit DAC in 40 nm CMOS with a sampling rate up to 65 GS/s and a 13 GHz bandwidth was commercially available in 2012 [17]. In 2014, an 8 bit DAC with a sampling rate of 100 GS/s in 28 nm low-power CMOS was published in [18]. In 2017, a transmitter for 100 Gb/s coherent network with 4 fold 64 GS/s 8 bit DACs in 20 nm was reported in [19]. The latest one achieves the lowest power consumption of 620 mW per DAC.

Table 1.1 lists the state-of-the-art high-speed DACs with sampling rates over 50 GS/s. It can be seen that SiGe and InP DACs can achieve higher bandwidth compared to CMOS DACs. However, their power consumption is much higher even without considering the power for interfacing between DSPs and DACs.

Table 1.1: State-of-the-art high-speed DACs over 50 GS/s

	[18] this work	[16] Ciena	[17] Fujitsu	[19] Broadcom	[8] Uni Toronto	[9] Micram	[20] NTT	[13] Anritsu
Resolution [bit]	**8**	**6**	**8**	**8**	**6**	**6**	**6**	**6**
Sampling rate [GS/s]	**100**	**56**	**65**	**64**	**56**	**72**	**60**	**80**
Differential peak-to-peak output swing [V]	**1**	**0.6**	**1.2**	**0.7**	**6**	**1.3**	**1**	**0.8**
Bandwidth 3 dB [GHz]	**13**	**n.a**	**13**	**20**	**45**	**13**	**n.a**	**40**
Power [W]	**2.5**	**0.75**	**0.75**	**0.65**	**5.2**	**>10**	**1.8**	**5.3**
Technology	**28 nm CMOS**	**65 nm CMOS**	**40 nm CMOS**	**20 nm CMOS**	**130 nm SiGe/ BiCMOS**	**SiGe HBT/Bi CMOS**	**0.5 µm InP HBT**	**0.7 µm InP DHBT**
Year	**2014**	**2011**	**2012**	**2017**	**2016**	**2014**	**2011**	**2016**

1.3 Outline of This Work

This work employs time-interleaving techniques and distributed architecture to increase the sampling rate and the bandwidth of the high-speed CMOS DACs. To show the feasibility of this concept, an 8 bit 100 GS/s DAC with an on-chip memory is implemented in a 28 nm CMOS technology. The following chapters focus on the design techniques of high speed DACs, the DAC chip implementation and the measurement of the DAC.

Chapter 2 introduces the basic fundamentals of the DACs. It first introduces the behavior of an ideal DAC, e.g., the output characteristics, the quantization error, the zero-order-hold output and the output spectrum. Then, the static and dynamic performances are discussed. Afterwards, the common implementations (charge mode, voltage mode and current mode) for high-speed DACs and the common architectures (binary-coded, thermometer-coded and segmented) of the DACs are presented.

Chapter 3 focuses on the circuit design techniques of the current-steering DACs. After a brief introduction to the basic structure of the current-steering DAC, error sources that potentially limit the DAC performances and the possible design solutions to these errors are reviewed and analyzed. The time-interleaving techniques for increasing the sampling rate and the bandwidth extension techniques for the clock buffers and the DAC output stages are presented.

Chapter 4 presents the implementation of the 8 bit 100 GS/s distributed DAC. After introducing the chip architecture, the design of the building blocks, e.g., the DAC output stage, the clock circuits, the data MUXs and the memory are discussed. The chip layout is shown at the end of the chapter.

Chapter 5 deals with the measurement of the DAC. First, it introduces the test board design and the measurement setup. Then, the measurement results of the static and dynamic performances are presented and analyzed. The time-domain performances are shown by the eye-diagrams of the multi-level modulated output signals afterwards. Finally, a single carrier optical transmission experiment is presented to show the capability of high-speed transmission using the DAC.

Chapter 6 gives a conclusion of this work. It summarizes the key performances of the DAC chip and an outlook of future work for further possible improvements is discussed.

2 Digital-to-Analog Converter Basics

The digital-to-analog converter is an essential component in data processing systems that provides a link from the digital to the analog world. The analog output of a DAC can be a voltage or a current or a charge, which is proportional to the digital inputs. The DACs are employed in a wide spectrum of applications, e.g., audio/video system, motor control, software-defined radio system and communication system etc. Each application has its own particular requirements on the DAC's performances, which can be categorized into static and dynamic performances. For data communications, the most important performances are dynamic performances, especially frequency-domain performances.

Different circuit implementations and architectures of the DACs have their own advantages and limitations regarding linearity, speed and power. To achieve the overall system requirements, an appropriate architecture and implementation must be carefully chosen and the trade-off between the performances must be balanced.

This chapter introduces the basic fundamentals of the DACs. It begins with an introduction to the ideal DAC, e.g., the output characteristics, the time-domain and the frequency-domain behavior. Next, the static and dynamic performance metrics are explained. Finally, the common implementations (e.g., charge mode, voltage mode and current mode) and the common architectures (e.g., binary-coded, thermometer-coded and segmented architectures) of the DACs are presented.

2.1 Ideal DAC

A DAC converts a digital input (D_{in}) to an analog output (A_{out}). The input code usually uses an *N* bit binary format, i.e., $D_{\text{in}} = (b_{N-1} \ldots b_1 b_0)$, and its decimal value can be written as

$$D_{\text{in}} = 2^{N-1} b_{N-1} + \cdots + 2^1 b_1 + 2^0 b_0 \,, \tag{2.1}$$

where b_{N-1} is referred to the most significant bit (MSB) and b_0 is referred to the least significant bit (LSB). The analog output value is proportional to the digital input code. The relationship between the digital inputs and the analog output of an *N* bit DAC can be expressed as

$$A_{\text{out}} = \Delta_{\text{ref}} D_{\text{in}}, \tag{2.2}$$

where Δ_{ref} is the reference quantity e.g., voltage, current, charge, and is referred to 1 LSB of the output value.

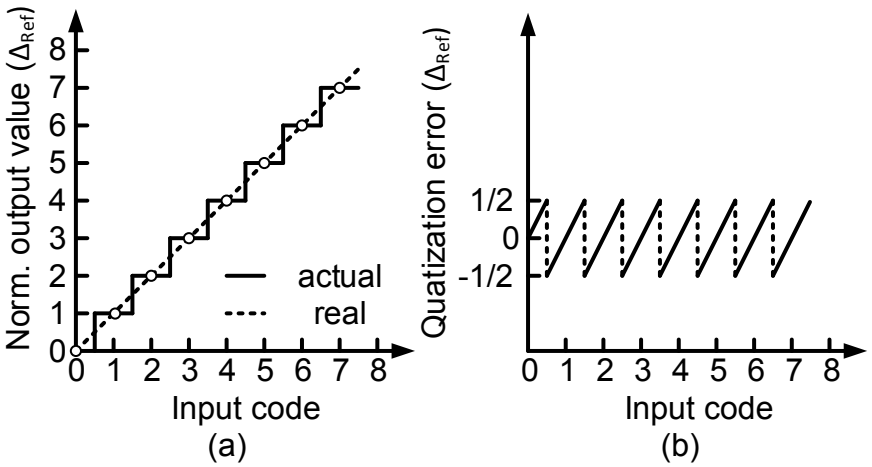


Fig. 2: (a) Output transfer characteristics (b) Quantization error

The transfer characteristics of an ideal DAC is shown in Fig. 2 (a). The solid line illustrates the actual output and the dashed line is the output with an infinite resolution. With a finite resolution, the value between two levels cannot be generated, since any value between two levels is quantized to one (the same) output level. The difference between the desired value and the actual quantized value is defined as the quantization error. It is important to notice that the actual quantization process is done in the digital domain before the DAC input and hence the actual digital-to-analog conversion has no quantization error. Nevertheless, the definition of the quantization error can be still used to present the difference between the DAC outputs with an infinite resolution and the DAC outputs with a finite resolution. For ideal digital-to-analog conversion, the quantization error is always smaller than 1 LSB, e.g., -½ LSB – ½ LSB. The transfer function of the quantization error is illustrated in Fig. 2 (b).

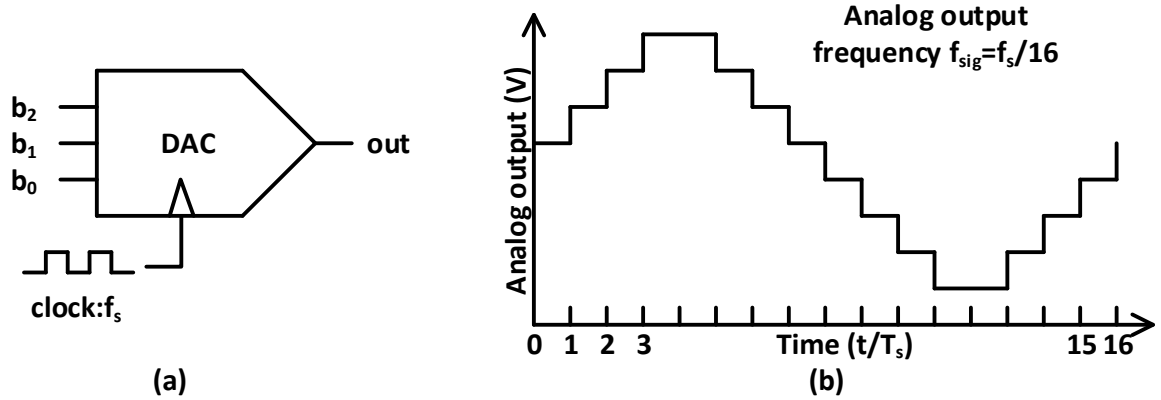


Fig. 3: (a) Block diagram of a 3 bit DAC and (b) its output in time domain

In the time domain, the conversion is normally triggered by a clock signal with a frequency of f_S. At each clock rising (or falling) edge, the DAC samples the input code and converts it to an analog value, thus having a sampling frequency of f_S equal to the clock frequency. The analog value is held for a whole clock period T_S. This behavior is illustrated in Fig. 3 with an example of a 3 bit DAC and its output waveform. It can be seen that the output waveform consists of a sequence of rectangle pulses with a width of $T_S = 1/f_S$. The hold behavior of the DAC output

is known as zero-order hold (ZOH). The impulse response of a ZOH function is shown in Fig. 4 (a) and can be expressed with a rectangular function, i.e.,

$$h(t) = \mathrm{rect}\left(t - \frac{T_\mathrm{s}}{2}\right) = \begin{cases} 1/T_\mathrm{s} & \text{for } 0 \le t < T_\mathrm{s} \\ 0 & \text{otherwise} \end{cases}, \tag{2.3}$$

where T_s is the sampling period; $\mathrm{rect}(t)$ is the rectangular function.

The frequency response can be calculated by applying a continues Fourier transformation to the impulse response, i.e.,

$$H(f) = e^{-\mathrm{j}\pi f T_\mathrm{s}} \left(\frac{\sin(\pi f T_\mathrm{s})}{\pi f T_\mathrm{s}}\right) = e^{-\mathrm{j}\pi f T_\mathrm{s}} \mathrm{sinc}(f T_\mathrm{s}) , \tag{2.4}$$

where $\mathrm{sinc}(x)$ is the normalized sinc function, i.e., $\mathrm{sinc}(x) = \sin(\pi x)/\pi x$.

The magnitude and the phase of $H(f)$ can be expressed as

$$|H(f)| = |\mathrm{sinc}(f T_\mathrm{s})| \tag{2.5}$$

and

$$\phi(f) = \angle \mathrm{sinc}(f T_\mathrm{s}) - \pi f T_\mathrm{s} . \tag{2.6}$$

Note that $\angle \mathrm{sinc}(f T_\mathrm{s})$ shows a phase jump of -π at integer multiple of $1/T_s$.

The magnitude $|H(f)|$ and wrapped phase $\phi(f)$ is shown in Fig. 4 (b) and (c), respectively.

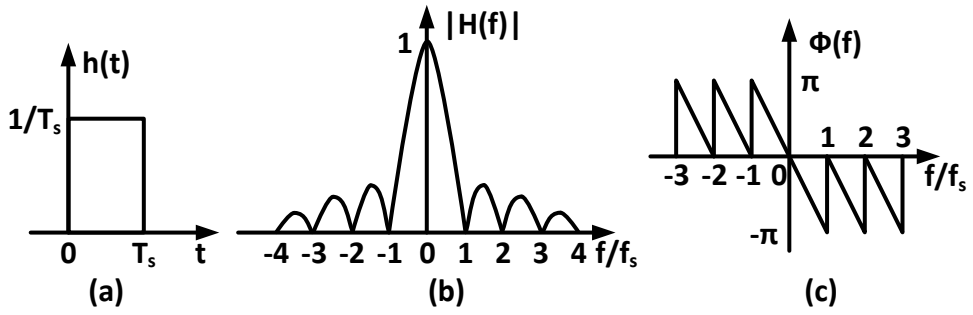


Fig. 4: Zero-order hold: (a) impulse response, (b) magnitude and (c) phase of the frequency response

Assuming that the DAC input is the sampled signal from an analog signal $x(t)$, and ignoring the quantization error, i.e., infinite resolution, the sampled signal $x_\mathrm{S}(t)$ can be defined as

$$x_\mathrm{S}(t) = x(t) \sum_{n=-\infty}^{+\infty} \delta(t - n T_\mathrm{s}) , \tag{2.7}$$

where $x(t)$ is the original signal; $x_\mathrm{S}(t)$ is the sampled signal; $\sum_{n=-\infty}^{+\infty} \delta(t - n T_\mathrm{s})$ is the dirac comb.

Using convolution theorem, the spectrum of $x_\mathrm{S}(t)$ can be expressed as

$$X_S(f) = X(f) * \frac{1}{T_s} \sum_{n=-\infty}^{+\infty} \delta\left(f - \frac{n}{T_s}\right). \quad \textbf{(2.8)}$$

Using the definition of periodic summation, $X_S(f)$ can be written as

$$X_S(f) = \frac{1}{T_s} \sum_{n=-\infty}^{+\infty} X\left(f - \frac{n}{T_s}\right). \quad \textbf{(2.9)}$$

Thus, the spectrum of the sampled signal $X_S(f)$ is a periodic replication of the input signal $X(f)$ with a period of $1/T_s$.

The sampled and hold DAC output $x_{SH}(t)$ can be expressed as the convolution of the sampled signal $x_S(t)$ and the ZOH function $h(t)$, i.e.,

$$x_{SH}(t) = x_S(t) * h(t) = x_S(t) * \text{rect}\left(t - \frac{T_s}{2}\right), \quad \textbf{(2.10)}$$

In the frequency domain, the Fourier transformation of $x_{SH}(t)$ can be written as

$$X_{SH}(f) = X_S(f) \cdot H(f) = \frac{1}{T_s} \sum_{n=-\infty}^{+\infty} X\left(f - \frac{n}{T_s}\right)\left(\frac{\sin(\pi f T_s)}{\pi f T_s}\right) e^{-j\pi f T_s} \quad \textbf{(2.11)}$$

The spectrum of a sampled signal with ZOH is thus the spectrum of an ideal sampled signal attenuated by a sinc function. The spectrum is repeated at multiples of the sampling frequency.

For a sine wave with a signal frequency f_{sig} generated by a DAC with a sampling rate of f_s, the signal frequency is replicated at $f_s \pm f_{sig}$, $2f_s \pm f_{sig}$, $3f_s \pm f_{sig}$ and so on, as shown in Fig. 5. Note that the negative frequency -f_{sig} is not shown in the spectrum. Since the signals are repeated at multiples of the sampling frequency, for signal frequencies higher than the Nyquist frequency ($f_{Nyquist} = f_s/2$), a replication of the signal appears at a frequency below $f_s/2$, causing aliasing. The aliasing can be avoided by keeping the signal frequency below the Nyquist frequency. The frequency bands $0 \sim f_{Nyquist}$, $f_{Nyquist} \sim 2f_{Nyquist}$, … are called 1st, 2nd , … Nyquist band or Nyquist zone, respectively.

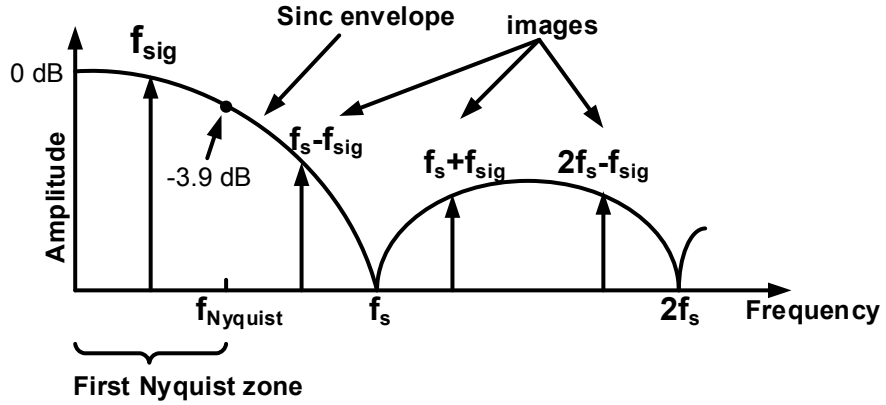


Fig. 5: Ideal DAC output spectrum for sinusoidal signal

2.2 Static Performance

Due to imperfections of real devices, there are differences between the ideal transfer characteristics and the real transfer characteristics. These differences can be characterized with offset error, gain error, differential non-linearity (DNL) and integrated non-linearity (INL).

Offset and gain errors

The offset and gain errors of the normalized and interpolated output transfer characteristic are shown in Fig. 6 (a) and (b), respectively.

The real output transfer curve is shifted by the offset error from the ideal transfer curve. The offset error can be measured by the difference between the real output and the ideal output with input code equal to zero.

The gain is defined as the slope of the line that results from the interpolation of the output transfer characteristics. Assuming that the gain is constant over all input code, it can be written as

$$gain = \frac{A_{\mathrm{FS}}}{D_{\mathrm{FS}}}, \tag{2.12}$$

where D_{FS} is the full-scale digital input and A_{FS} is the full-scale analog output. The gain error ($\varepsilon_{\mathrm{gain}}$) can be defined by the deviation of the real gain ($gain$) from the ideal gain ($gain_{\mathrm{id}}$), or the real full-scale value (A_{FS}) to the ideal full-scale value (A_{FSid}) [21], i.e.,

$$\varepsilon_{\mathrm{gain}} = \frac{gain}{gain_{\mathrm{id}}} - 1 = \frac{A_{\mathrm{FS}}}{A_{\mathrm{FSid}}} - 1. \tag{2.13}$$

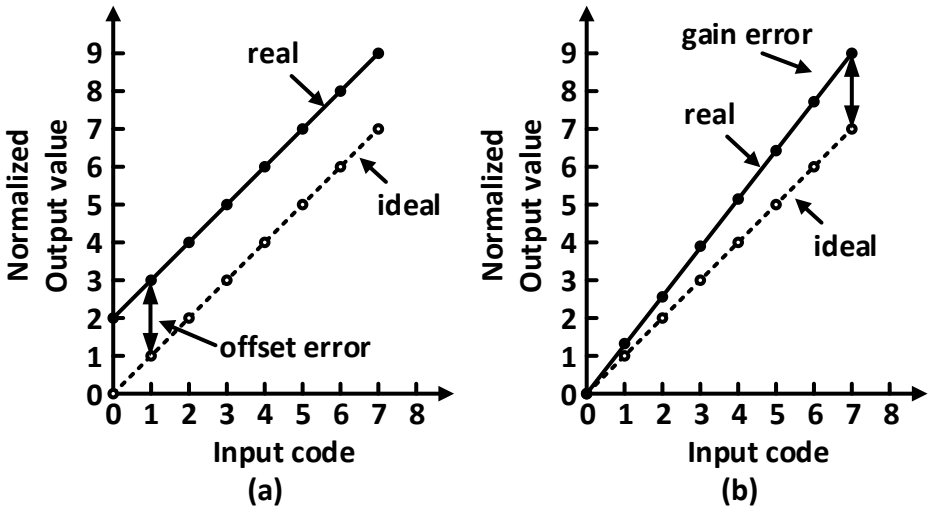


Fig. 6: (a) Offset error and (b) gain error

DNL and INL

The static non-linearity can be characterized with DNL and INL. Before measuring the INL and DNL, the offset and gain error has to be removed to fit the actual output line to the ideal output line. The INL and DNL of a DAC output are shown in Fig. 7 (a).

DNL is the difference between actual output step size and the ideal output step size of each code. The ideal step size is referred to one LSB. The DNL of a given input code i can be written as

$$DNL_i[\text{LSB}] = A_{i+1} - A_i - 1, \tag{2.14}$$

where A_i is the output with an input code i.

INL is the difference between the actual output (solid line) and the reference output (dashed line). The INL of the input code i is defined as

$$INL_i[\text{LSB}] = A_i - A_{\text{id},i}\,, \tag{2.15}$$

where $A_{\text{id},i}$ is the ideal output with an input code i, which is equal to i LSBs

The INL can be also derived from DNL, i.e.,

$$INL_i[\text{LSB}] = \sum_{k=0}^{i-1} DNL_k\,. \tag{2.16}$$

The INL and DNL over the input code extracted from Fig. 7 (a) are shown in Fig. 7 (b) and (c), respectively.

To summarize the DNL and INL performance of a DAC, the maximum absolute values are commonly reported.

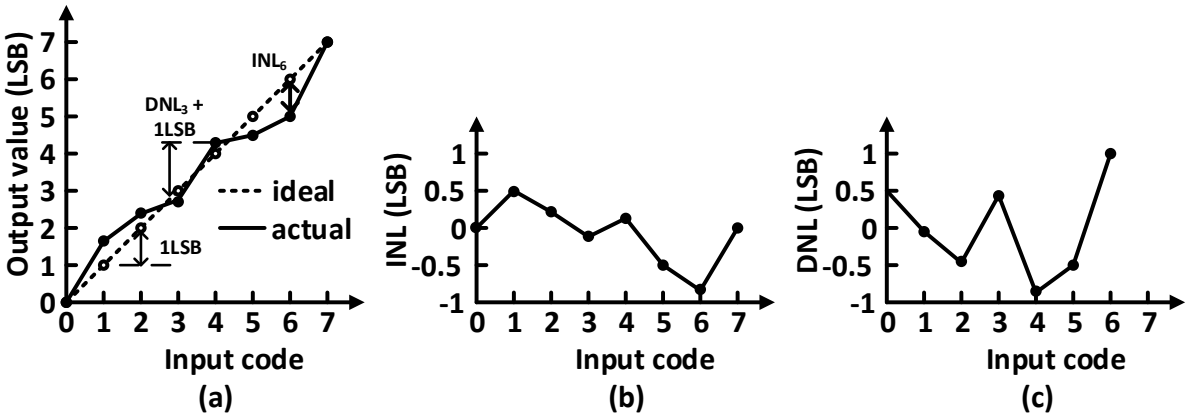


Fig. 7: (a) INL and DNL in an output transfer characteristics and extracted (b) INL and (c) DNL

2.3 Dynamic Performance

The dynamic performances include noise, distortion, linearity, bandwidth etc., which are frequency and amplitude dependent. The characterization can be done by analyzing the DAC output spectrum as shown in Fig. 8. The next sections introduce the most important performance metrics and the fundamental limitations, e.g., quantization noise and sampling jitter.

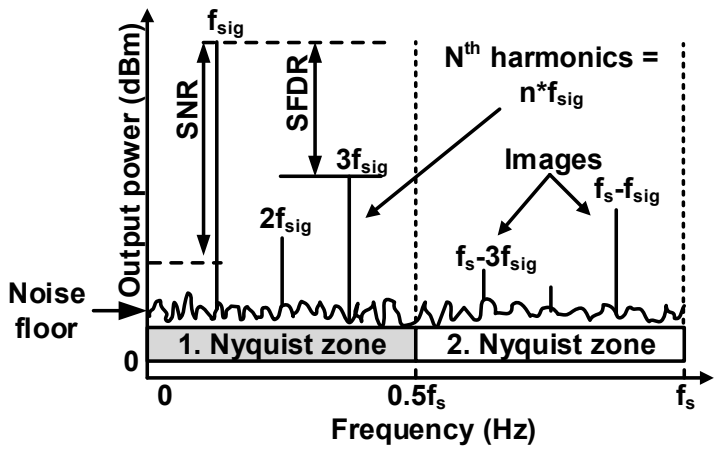


Fig. 8: DAC output spectrum for sinusoidal signal

2.3.1 Performance Metrics

Signal to noise ratio (SNR)

Signal to noise ratio (SNR) is the ratio between the signal power (P_{sig}) and the noise power (P_{n}) within the frequency range of interest, i.e.,

$$SNR\ [\text{dB}] = 10\log\frac{P_{\text{sig}}}{P_{\text{n}}}. \tag{2.17}$$

Total harmonic distortion (THD)

When generating a fundamental frequency f_{sig} with a DAC, its non-linearity causes harmonic distortions (HDs). The k$^{\text{th}}$ harmonic distortion (HD_k) is the ratio between the power of k$^{\text{th}}$ harmonic distortion (P_{k}) and the power of the fundamental signal (P_{sig}), i.e.,

$$HD_k\ [\text{dB}] = 10\log\frac{P_{\text{k}}}{P_{\text{sig}}}. \tag{2.18}$$

Total harmonic distortion (THD) is defined by the ratio between total harmonic power and the signal power within the frequency range of interest, i.e.,

$$THD\ [\text{dB}] = 10\log\frac{\sum_2^N P_{\text{k}}}{P_{\text{sig}}}. \tag{2.19}$$

where N is the number of harmonics that is taken into account.

Signal to noise and distortion ratio (SNDR)

Signal to noise and distortion ratio (SNDR) is defined by the ratio between the signal power and the total noise and distortion power within the frequency range of interest, i.e.,

$$SNDR\ [\text{dB}] = 10\log\frac{P_{\text{sig}}}{P_{\text{n}} + \sum_2^N P_k}. \tag{2.20}$$

Effective number of bits (ENOB)

Effective number of bits (ENOB) is the effective resolution of a converter considering noise and distortion. It can be expressed as [21]

$$ENOB = \frac{SNDR[\text{dB}] - 1.76}{6.02}. \tag{2.21}$$

Spurious free dynamic range (SFDR)

Spurious free dynamic range (SFDR) is the ratio between the signal power and the power of the highest spur ($P_{\text{sp\_max}}$) within the frequency range of interest, i.e.,

$$SFDR\ [\text{dB}] = 10\log\frac{P_{\text{sig}}}{P_{\text{sp\_max}}}. \tag{2.22}$$

2.3.2 Quantization Noise

One of the fundamental limitations of the dynamic performances is the quantization noise, which is caused by the quantization error. The quantization error is an error due to finite resolution. For sufficiently high resolution, the noise floor of an ideal DAC without other noise sources is determined by quantization noise. For low resolution, the quantization error generates more distortion-like components in the spectrum [21].

The quantization error is smaller than one quantization interval Δ_{Q}. Assuming that it is equally distributed within one quantization interval, the probability of the quantization error ε_{Q} is

$$p(\varepsilon_{\text{Q}}) = \begin{cases} 1/\Delta_{\text{Q}} & \text{for } -\Delta_{\text{Q}}/2 \le \varepsilon_{\text{Q}} < \Delta_{\text{Q}}/2 \\ 0 & \text{otherwise} \end{cases}. \tag{2.23}$$

The power of the quantization noise can be determined by

$$P_{\text{Q}} = \int_{-\Delta_{\text{Q}}/2}^{\Delta_{\text{Q}}/2} {\varepsilon_{\text{Q}}}^2\, p(\varepsilon_{\text{Q}})\mathrm{d}\varepsilon_{\text{Q}} = \frac{{\Delta_{\text{Q}}}^2}{12}. \tag{2.24}$$

For a sine wave generated by an ideal DAC with finite resolution, the SNR can be determined by the ratio between the signal power and the quantization noise power. The full-scale output A_{FS} of an N bit DAC is $2^N\Delta_{\text{Q}}$. Thus, the signal power of the sine wave is

$$P_{\text{sin}} = \frac{1}{T_{\text{sin}}}\int_0^{T_{\text{sin}}} \left(\frac{A_{\text{FS}}}{2}\right)^2 \sin^2\left(2\pi\frac{t}{T_{\text{sin}}}\right)\mathrm{d}t = \frac{A_{\text{FS}}^2}{8} = \frac{\left(2^N\Delta_{\text{Q}}\right)^2}{8}, \tag{2.25}$$

where P_{sin} is the time-averaged power of the sine wave; T_{sin} is the period of the sine wave.

With sufficient resolution and using (2.17) (2.24) and (2.26), the SNR of the sine wave is given by

$$SNR_{\text{sin}}\ [\text{dB}] = 10\log\frac{\frac{\left(2^N\Delta_{\text{Q}}\right)^2}{8}}{\frac{{\Delta_{\text{Q}}}^2}{12}} = 10\log\left(\frac{3}{2}\cdot 2^{2N}\right) = 6.02N + 1.76. \tag{2.26}$$

It can be seen that every additional bit can improve the SNR by 6.02 dB, i.e., the quantization noise power reduces by a factor of four for every bit improvement.

The quantization noise power is spread out over the whole Nyquist band. If the frequency range of interest is only a portion of the Nyquist band, e.g., with oversampling, the noise power is reduced after filtering out the noise outside the frequency band of interest. In the case of oversampling, the sampling frequency is much higher than the signal bandwidth, i.e., the frequency range of interest. The ratio between the Nyquist frequency $f_s/2$ and the signal bandwidth $f_{\text{sig\_bw}}$ is defined by the oversampling ratio (OSR), i.e.,

$$OSR = \frac{f_s}{2f_{\text{sig\_bw}}}. \tag{2.27}$$

The quantization noise power of an oversampling system is given by

$$P_{\text{Q}} = \frac{{\Delta_{\text{Q}}}^2}{12 \cdot OSR}. \tag{2.28}$$

Therefore, the signal to noise ratio can be written as

$$SNR_{OSR}\ [\text{dB}] = 6.02N + 1.76 + 10\log(OSR)\ . \tag{2.29}$$

It shows that increasing the OSR by a factor of four improves the SNR by 6 dB, which is equal to 1 bit resolution.

2.3.3 Sampling Clock Jitter

The dynamic performance is also affected by the sampling clock uncertainty, which is referred to as sampling clock jitter. The jitter in the sampling clock causes variations of the actual sampling period at the DAC output, i.e., the period of each sample differs from its ideal value and varies from sample to sample.

Assuming that the timing error of i^{th} sample is $\delta_{\text{jitter}}(i)$, the output error $x_{\text{j\_err}}(i)$ caused by the timing error is the product of the output difference $\Delta x_{\text{S}}(i)$ between i^{th} and $(i-1)^{th}$ sample (i.e., $\Delta x_{\text{S}}(i) = x_{\text{S}}(i) - x_{\text{S}}(i-1)$) and the timing error $\delta_{\text{jitter}}(i)$, i.e.,

$$x_{\text{j\_err}}(i) = \frac{\Delta x_{\text{S}}(i)\delta_{\text{jitter}}(i)}{T_{\text{s}}}, \tag{2.30}$$

where T_{s} is the ideal sampling period. In other words, the output error is the product of the deviation of the ideal output and the relative timing error ($\delta_{\text{jitter}}(i)/T_{\text{s}}$).

For a sine wave input with an amplitude of A and a frequency of f_{sig} (i.e., $x_{\text{in}}(t) = A\sin(2\pi f_{\text{sig}}t)$), the error of the sampled output is given by

$$\Delta x_{\text{S}}(i) = A2\pi f_{\text{sig}}T_{\text{s}}\cos(i \cdot 2\pi f_{\text{sig}}T_{\text{s}})\ . \tag{2.31}$$

According to (2.30), the power of the output jitter error ($P_{\text{j\_err}}$) is determined by the product of the mean power of the error amplitude ($2A^2\pi^2 f_{sig}^2$) and the power of the root-mean-square (RMS) jitter (σ_{jitter}^2).

$$P_{\text{j\_err}} = 2A^2\pi^2 f_{sig}^2\ \sigma_{jitter}^2 \tag{2.32}$$

Since the power of the sine wave is $A^2/2$, the signal to noise ratio with respect to the jitter (SNR_{jitter}) is

$$SNR_{jitter}\ [\text{dB}] = -20\log\left(2\pi f_{\text{sig}}\sigma_{jitter}\right). \tag{2.33}$$

It can be seen that, for a desired SNR, the jitter has to decrease with increased signal frequency. Fig. 9 plots the SNR and the ENOB versus frequencies for different RMS jitter. It shows that to achieve an ENOB of 6 bit at 20 GHz the RMS jitter should be less than 100 fs.

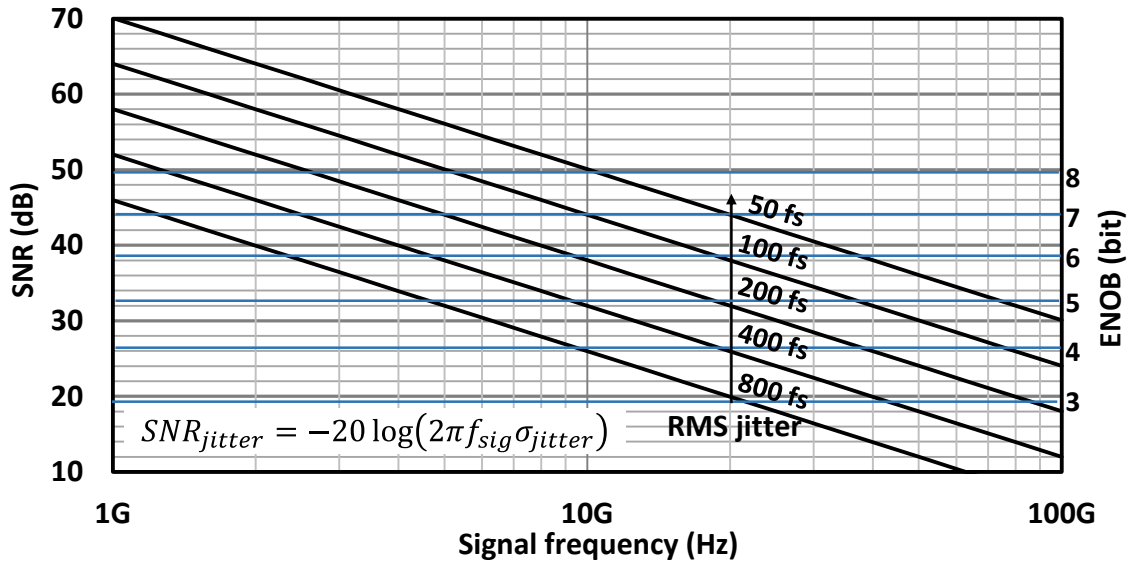


Fig. 9: SNR and ENOB versus frequencies for different RMS jitter

2.4 DAC Implementations

There are different ways to implement the digital-to-analog conversion. Since the focus is on high-speed communications, only high-speed DAC implementations are discussed. There are generally three types of DAC implementations [22]: charge-redistribution mode, voltage mode and current mode.

2.4.1 Charge-Redistribution DAC

The charge-redistribution DAC can be seen as a switched-capacitor (SC) DAC. It is widely used in successive-approximation-register (SAR) ADCs [23] .

Fig. 10 shows a 3 bit charge-redistribution DAC operating in reset phase. It consists of an array of binary-weighted capacitors and switches. The bottom plate of the capacitors are connected to a reference voltage via the switches. The output is connect to the top plate of the capacitor array.

During the reset phase, all switches are connected to ground to reset the voltages across the capacitors to zero. The top and bottom plate can also be connected to different voltages to reset the capacitor voltages to a non-zero value. After that, the *reset* switch is disconnected from ground and the bottom plate can be switched to the reference voltage V_{ref} according to the input code. This process causes a charge redistribution and changes the output voltage. The output voltage can be calculated with the digital input code and the reference voltage, i.e.,

$$V_{\text{out}} = V_{\text{ref}}\left(\frac{b_2}{2^1} + \frac{b_1}{2^2} + \frac{b_0}{2^3}\right) \tag{2.34}$$

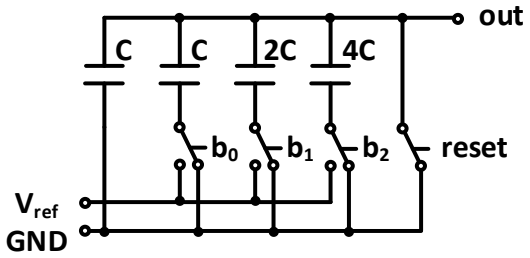


Fig. 10: Charge-Redistribution DAC

For a high impedance load, e.g., comparator input in SARADC, the DAC output is capable of directly driving the load without buffering. For applications using a lower impedance load, the output has to be buffered with an operational amplifier, which limits the settling time and hence reducing the conversion rate.

2.4.2 R-String DAC

An R-string DAC is a voltage mode DAC, which consists of a string of resistors connected to two reference voltages at their two ends. The output voltage is generated by connecting one of the switches to the output. Fig. 11 shows an example of a 3 bit R-string DAC [24].

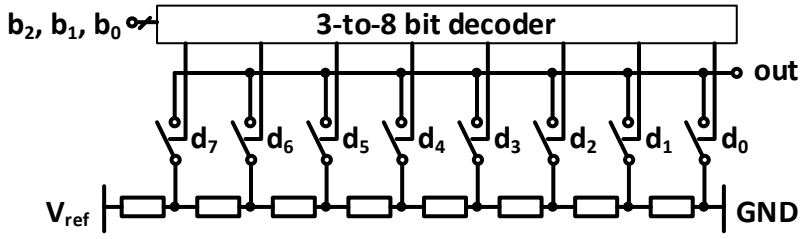


Fig. 11: 3 bit R-String DAC

The advantage of this type of DAC is that the structure is simple and it is inherently monotonic, i.e., the output voltage always either remains constant or changes in the same direction as the input code. The drawback is that the number of components increases exponentially with the

number of input binary bits. Since the output impedance is code dependent and relatively high, a buffer at the output is normally required to improve linearity and increase the output bandwidth. Even with an output buffer, the settling time is still dependent on the RC constant, which becomes slower for higher resistance [25].

2.4.3 R-2R Ladder DAC

The limitation of an R-string DAC is that the number of resistors increases exponentially with the number of input binary bits. This problem can be solved by using an R-2R ladder structure, which consists of a number of resistors with the values R and $2R$. The number of the resistors in an R-2R ladder DAC increases linearly with the number of input bits. An R-2R ladder DAC can be implemented in voltage mode or current mode.

Fig. 12 shows a 4 bit voltage-mode R-2R ladder DAC. The output impedance of the R-2R ladder is R and is independent of the input code. The output voltage can be written as [24]

$$V_{\text{out}} = V_{\text{ref}}\left(\frac{b_3}{2^1} + \frac{b_2}{2^2} + \frac{b_1}{2^3} + \frac{b_0}{2^4}\right) \tag{2.35}$$

The voltage-mode R-2R ladder DAC output requires buffering to drive the output load. The reference voltage generator needs to have sufficient low output impedance, since the load on the V_{ref} node is code dependent and hence generates distortion at DAC output.

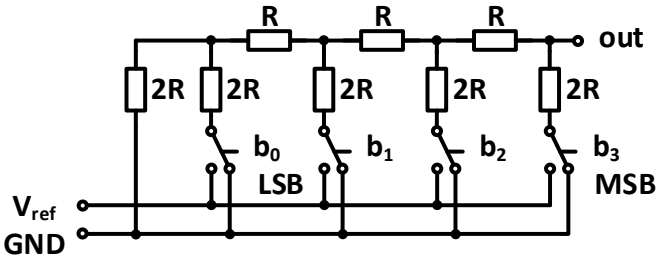


Fig. 12: Voltage-mode R-2R ladder DAC

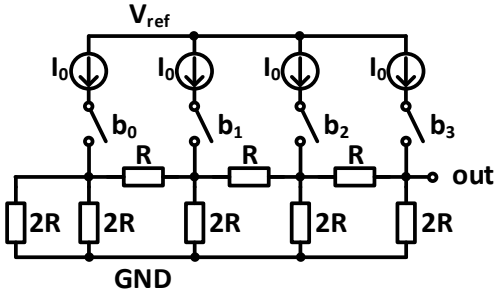


Fig. 13: Current mode R-2R ladder DAC

The current mode R-2R ladder DAC is shown in Fig. 13. The output voltage is given by

$$V_{\text{out}} = R \cdot I_0 \cdot \left(\frac{b_3}{2^0} + \frac{b_2}{2^1} + \frac{b_1}{2^2} + \frac{b_0}{2^3}\right) \tag{2.36}$$

With the output resistance $R = 50\ \Omega$, the output can be directly used in a 50 Ω system without extra buffering.

One of the biggest advantages of R-2R ladder DACs is the use of identical structures that consist of a current source and two resistors with fixed values of R and $2R$. The use of identical current sources is especially beneficial for the bipolar technology, since the bipolar junction transistors (BJTs) require much higher minimum current compared to MOSFETs for proper operation and thus the LSB currents cannot be directly generated from the BJT current sources in a high-resolution DAC. To keep the proper operating range of the current sources, instead of scaling down the current sources, the R-2R ladders divide down the currents from the same current sources.

One of the biggest disadvantages is the higher power consumption compared to normal current-steering DAC, since most of the currents are wasted in the R-2R ladder current divider. For low-impedance resistors, the parasitic capacitance is relatively high due to the large area, which is constrained by its maximum current density. The delay on the R-2R ladder is code dependent, which generates code dependent glitches and thus increases distortion.

2.4.4 Current-Steering DAC

The current-steering (CS) DAC consists of an array of current sources and switches. The output current is the sum of the currents that are switched to the output, controlled by the input code. The output voltage is determined by the output current I_{out} and the output load R, i.e.,

$$V_{\text{out}} = R \cdot I_{\text{out}} = R \cdot I_0 \cdot (2^3 b_3 + 2^2 b_2 + 2^2 b_2 + 2^0 b_0) \tag{2.37}$$

where I_0 is the LSB unit current.

Fig. 14 shows a 4 bit CS-DAC consisted of binary weighted current sources.

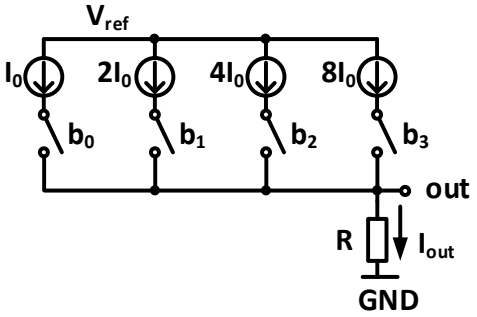


Fig. 14: Current-Steering DAC

CS-DACs are widely used in high-speed DAC design. It offers higher switching speed and higher output bandwidth compared to the voltage-mode and the charge-redistribution-mode DACs, because the transistor switching is much faster in current mode than in voltage mode due to lower input capacitance and lower output swing. The voltage-mode DACs require amplifiers with feedback technique at the output as a buffer for better linearity, which limits the bandwidth, whereas CS-DACs can drive the output resistor directly without buffering.

2.5 DAC Architectures

For CS-DACs, the current sources controlled by the input codes can be binary weighted or unary weighted. According to the input code style and the weight of the current sources, the CS-DAC can be classified as binary-coded, thermometer-coded or segmented DACs.

2.5.1 Binary-Coded DAC

Binary-coded DACs consist of binary-weighted current sources as shown in Fig. 14. The current sources are directly controlled by the binary input code and therefore it is simple to implement. The major drawbacks of the binary-weighted architecture are relatively poor DNL and large glitches [21], since the number of switches to be switched is not proportional to the change of the input code and the output value. For example, to go across the middle scale point from 011…11 to 100…00 by one LSB, all switches have to be switched, which therefore generates the largest glitch error. The worst-case DNL is also at the middle code transition due to the mismatch between MSB and LSB current sources.

2.5.2 Thermometer-Coded DAC

Thermometer-coded DACs consist of unary weighted current sources as shown in Fig. 15 (a). The binary input code ($b_3 - b_0$) is decoded to thermometer-code ($t_{15} - t_0$) to select the unary weighted current sources. The number of switches to be switched is proportional to the change of the input code and thus the output glitch energy is proportional to the change in the output value, reducing the glitch error. In contrast to the binary-coded DAC, the DNL errors are equal for all the input codes. The major drawbacks of thermometer-coded DACs are the need of binary-to-thermometer decoders and the larger number of current sources and switches. The complexity of the decoder logic and the layout increases exponentially with the number of the binary bits, which results in large layout area.

2.5.3 Segmented DAC

In order to reduce the decoder and layout complexity, segmented architectures are commonly used for high-resolution DACs. A segmented DAC is a combination of binary-coded and thermometer-coded DACs. A 4 bit segmented CS-DAC is shown in Fig. 15 (b). The two LSB inputs (b_1, b_0) are binary coded and the two MSB inputs (b_3, b_2) are decoded to 3 bit thermometer code (t_2, t_1, t_0). Compared to the 4 bit binary-coded CS-DAC, the segmented DAC has only one additional bit, but the worst-case DNL and the worst-case glitch error are reduced to half of those of the binary-coded CS-DAC. The complexity is also much lower compared to the pure thermometer-coded DAC. With proper ratio of segmentation, an optimum balance between the linearity, the complexity, the area, the speed and the power is achievable.

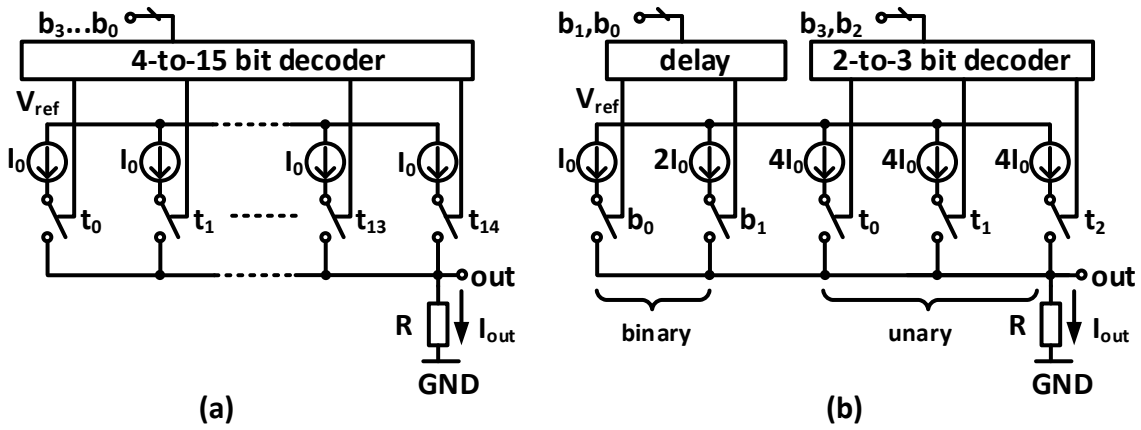


Fig. 15: (a) 4 bit thermometer-coded and (b) 4 bit segmented CS-DAC

3 Current-Steering DAC

This chapter first introduces the basic structure of a current-steering DAC. A number of common error sources that potentially limit the DAC performances and the possible design solutions are reviewed and analyzed. The time-interleaving techniques, which are used to increase the sampling rate, are presented along with the analysis of their spectrums and their error mechanisms. The bandwidth extension techniques – including inductive peaking and distributed amplifier techniques – for increasing the bandwidth of the clock buffer and the DAC output stage are discussed.

3.1 Basic Structure

As discussed in the previous chapter, current-steering (CS) DACs provide the highest conversion speed, which are best suited for high-speed communication applications. The DAC consists of an array of current sources and current switches weighted in binary or unary or segmented, as depicted in Fig. 16. The current switches direct the current either through the positive or through the negative output node, depending on the input data. Since the current switches only redirect a fixed amount of current generated by corresponding current source, it can operate very fast.

Depending on the DAC architecture, the thermometer bits are generated by a thermometer decoder and the binary bits are matched to the decoder delay. The data signals applied to the current switches are synchronized by the clock and amplified by the switch drivers.

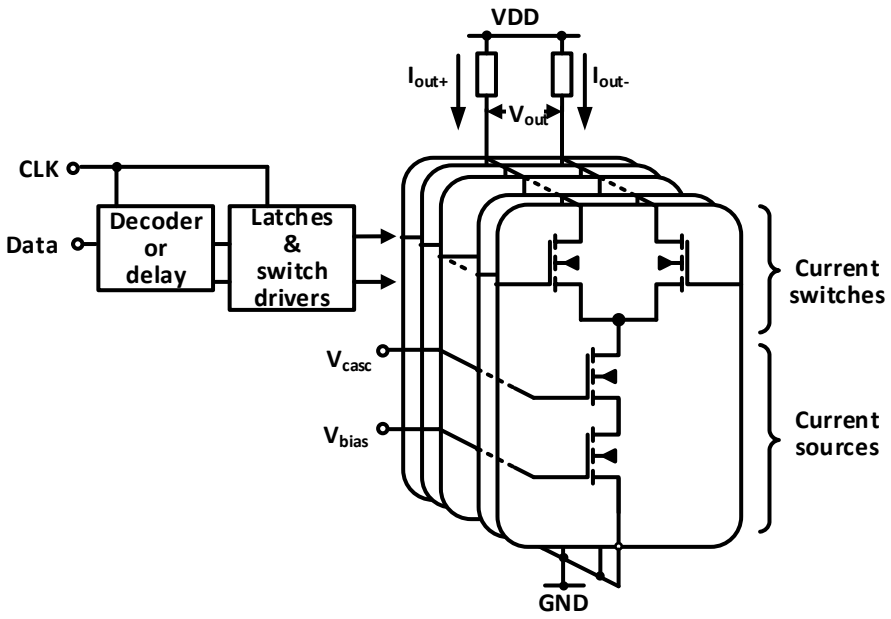


Fig. 16: Basic current-steering DAC structure

3.2 Design Challenges

Non-ideal behaviors of the real devices cause errors in the digital-to-analog conversion and limit the performances of the CS-DACs. The errors include static errors, e.g., mismatches in the current sources, the finite output impedance, etc., and include the dynamic errors, e.g., timing errors, glitches, etc. Low frequency performances are mainly limited by the static errors, while high frequency performances are affected by both the static and the dynamic errors. Most of the known error mechanisms have been well studied, and the methods to mitigate the errors are also proposed in a range of research works [26–28]. This section reviews the important error sources and their common solutions.

3.2.1 Mismatch of Current Sources

The DAC linearity is directly related to the mismatch of the current sources. The matching properties in the CMOS process have been analyzed in [29] and the standard deviation of the output current of a MOSFET is given by

$$\frac{\sigma^2(I_\mathrm{D})}{I_\mathrm{D}^2} = \frac{4\sigma^2(V_\mathrm{th0})}{(V_\mathrm{GS}-V_\mathrm{th0})^2} + \frac{A_\beta^2}{WL}\ , \tag{3.1}$$

where I_D is the drain current; V_th0 is the threshold voltage at 0 V substrate voltage; W and L are the width and length of the MOS transistor, respectively. A_β is a technology-dependent constant of the current amplification factor β and it is related to the length, the width, the oxide thickness and the mobility variations of the MOS transistor.

The variation of the threshold voltage V_th contributes to the major part of the current source mismatch. According to Pelgrom's law [29], the variance of V_th is inversely proportional to the area, i.e.,

$$\sigma^2(V_\mathrm{th0}) = \frac{A^2{}_{V_\mathrm{th0}}}{WL}\ , \tag{3.2}$$

where A_{V_th0} is a technology-dependent constant related to the threshold voltage.

From (3.1) and (3.2), it can be seen that increasing the overdrive voltage ($V_\mathrm{ov} = V_\mathrm{GS} - V_\mathrm{th}$) and the area ($WL$) improves the matching. However, these two parameters affect also other performances. The transistor has to be operated in the saturation region for high output resistance (r_o). Therefore, the minimum saturation drain-source voltage ($V_\mathrm{DS,sat} = V_\mathrm{ov}$) has to be increased with the overdrive voltage, increasing the power. Larger area increases the parasitic capacitances, decreasing the output impedance at high frequencies. For this reason, the size of the transistors has to be carefully chosen to meet the static matching requirement whilst not degrading the high frequency performance.

The INL is commonly used as an indicator of the static performance. It varies across chips due to the mismatch of the current sources. To describe the INL variation across chips, a parameter called INL_Yield is commonly used. The INL_Yield is defined by the percentage of the DACs

with INL smaller than 0.5 LSB [30]. Assuming that there is no correlation between the output currents, the INL_Yield of an N bit DAC can be obtained by multiplying the probabilities that each output value has an error smaller than 0.5 LSB [30], i.e.,

$$INL\_Yield = \prod_{i=2}^{2^N-1} \mathrm{erf}\left(\frac{Q_i}{\sqrt{2}}\right), \tag{3.3}$$

where $\mathrm{erf}(x)$ is the Gauss error function,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt\,, \tag{3.4}$$

and

$$Q_i = 2^{-(N+1)} \left[\frac{\overline{A}_{\mathrm{out},i}(1-\overline{A}_{\mathrm{out},i})}{2^N-1}\right]^{-\frac{1}{2}} \left[\frac{\sigma(I_{\mathrm{LSB}})}{I_{\mathrm{LSB}}}\right]^{-1}, \tag{3.5}$$

where N is the number of bits; $\overline{A}_{\mathrm{out},i}$ is the normalized mean output at code i, i.e., $\overline{A}_{\mathrm{out},i} = i/(2^N-1)$; $\sigma(I_{\mathrm{LSB}})/I_{\mathrm{LSB}}$ is the relative standard deviation of the LSB current.

However, the assumption of uncorrelated output currents results in an overly pessimistic estimation. Therefore, an adjustment of INL_Yield formula is proposed in [31],

$$INL\_Yield = \prod_{2^{N-1}-1}^{2^{N-1}} \mathrm{erf}\left(\frac{Q_i}{\sqrt{2}}\right), \tag{3.6}$$

where only the MSB switching is taken into account, which results in a more realistic but overly optimistic estimation.

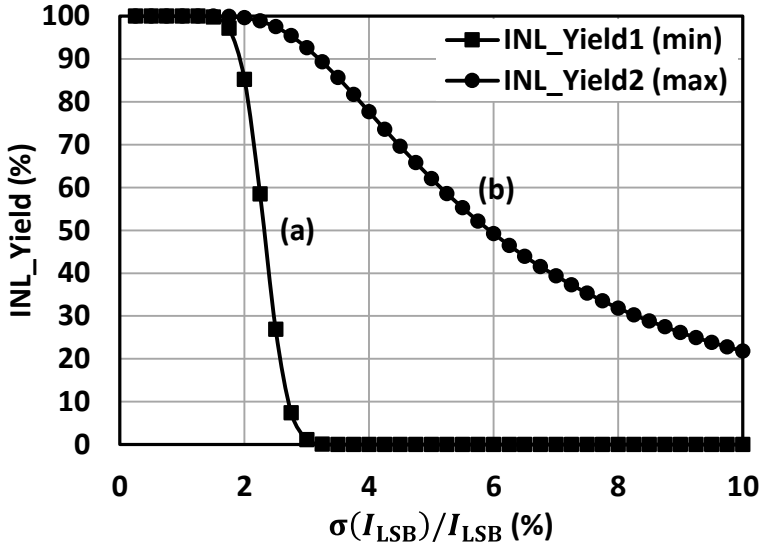


Fig. 17: INL_Yield as a function of current-source mismatch (a) using equation (3.3) and (b) using equation (3.6) for an 8 bit DAC

Therefore, the correct yield estimation is located somewhere between the two results of (3.3) and (3.6). The INL_Yields resulted from the two equations for an 8 bit DAC are depicted in Fig. 17. For example, a 2% standard deviation of the LSB current results in an INL_Yield between 85% – 99%. Although a more accurate model of INL_Yield has been proposed in [32], the most accurate and reliable INL_Yield is still estimated by Monte-Carlo-simulations [33], which can be very time consuming for high resolution DACs.

To reduce systematic errors, layout techniques such as common-centroid and inter-digit placement can be employed. With proper statistical analysis and layout techniques, 14 bit resolution can be achieved [34]. The drawback of the statistical approaches is the requirement of large area. This limits the DAC speed due to large parasitic capacitances and large parasitic resistances caused by large transistors and long routing distance. Furthermore, the yield relies strongly on the mismatch model provided by individual technology process, which cannot be modified.

To further increase the yield and to reduce the area, calibration techniques are often employed for high speed DACs. The calibration techniques are usually categorized in foreground and background calibration. The background calibration can be performed during normal DAC operation, while the DAC operation has to be turned off during the foreground calibration.

Segmented architectures are often employed to reduce the glitch error generated by the switching of the binary weighted current sources. In segmented DACs, the MSBs are thermometer coded and the LSBs are binary coded. The LSB binary current sources normally have lower impact on the DAC linearity. Therefore, an acceptable INL_Yield regarding the binary LSBs is achievable with a reasonable area [35]. Consequently, the calibration is usually applied to each individual unary current cell and to the summed current of the entire binary current cells [19, 35–39].

One common approach of the current source calibration is illustrated in Fig. 18 (a). During calibration, the current source is connected to an extra resistor through an additional switch. The binary current sources are summed up to one current output I_{binary}. To match the unary current to the summed binary currents, an additional dummy LSB current is added to the summed binary current output to obtain the following relation [38, 39] :

$$I_{\mathrm{binary}} = I_{\mathrm{unary}} = I_{\mathrm{ref}} = 2^n I_{\mathrm{LSB}} . \tag{3.7}$$

I_{ref} is first compared to I_{binary} and its value is trimmed until it is equal to I_{binary}. After that, each unary current cell is connected to the calibration node and trimmed against I_{ref}.

The error between the I_{binary} and I_{unary} should be smaller than half an LSB current (I_{LSB}) to meet the DNL constraint, i.e.,

$$\left| I_{\mathrm{unary}} - I_{\mathrm{binary}} \right| < \frac{1}{2} I_{\mathrm{LSB}}. \tag{3.8}$$

The trimming of the current cells can be implemented with a calibration current DAC (CALDAC) as shown in Fig. 18 (b). The total current of the CALDAC has to cover the unary

current mismatch spread, e.g., $\pm 3\sigma(I_{\text{unary}})$. Furthermore, the linearity of the CALDAC has to be sufficiently high to fulfil the matching requirement between the I_{binary} and I_{unary} shown in (3.8). Therefore, the calibration step size, e.g., the LSB current of the CALDAC $I_{\text{cal\_LSB}}$, must be sufficiently small, e.g., $I_{\text{cal\_LSB}} \leq I_{\text{LSB}}$. To keep the layout compact for high-speed operation, the current sources of the DAC and the CALDAC can be placed in a separate location [16].

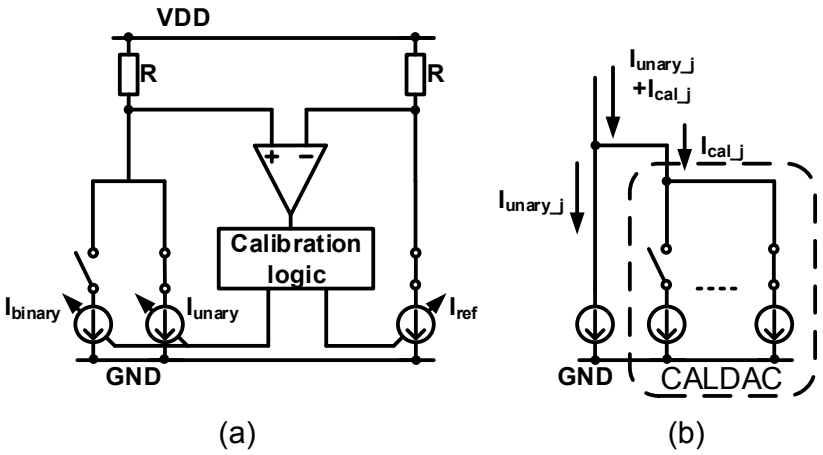


Fig. 18: (a) Current source calibration (b) calibration current DAC

3.2.2 Finite Output Impedance

Due to the finite output impedance of the current sources, the DAC output impedance varies with the input code and leads to distortions. At low frequencies, the impedance at each positive or negative output node of a unit current cell is either equal to Z_0, when the current switch is turned on, or is approximately infinite, when the current switch is turned off [40]. Fig. 19 shows a DAC model with finite output impedance.

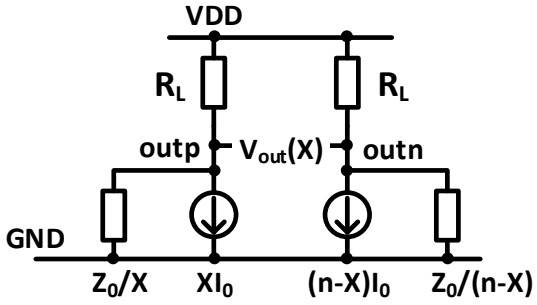


Fig. 19: The effect of the finite output impedance

The number of current sources connected to the output varies with the input code and thus the output impedance, in parallel with the output load (R_{L}), is code dependent. For an N bit DAC with a total number of $n = 2^N - 1$ unit current sources, the impedance at the DAC output nodes are determined by the load resistance R_L and the total output impedance of current cells at each node, i.e., $Z_{\text{outp}}(X)$ and $Z_{\text{outn}}(X)$, where X is the input code ($0 \leq X \leq$ n).

Assuming the output impedance of the unit current source is equal to Z_0, the impedance of the positive and negative outputs with an input code X are

$$Z_{\mathrm{outp}}(X) = \frac{Z_0}{X} \; and \; Z_{\mathrm{outn}}(X) = \frac{Z_0}{n - X}, \tag{3.9}$$

respectively.

Assuming the unit current is equal to I_0, the output currents at the output nodes are

$$I_{\mathrm{outp}}(X) = XI_0 \; and \; I_{\mathrm{outn}}(X) = (n - X)I_0 \, , \tag{3.10}$$

respectively.

The differential output voltage can be calculated by

$$V_{\mathrm{out}}(X) = \frac{XI_0}{\frac{1}{R_{\mathrm{L}}} + \frac{X}{Z_0}} - \frac{(n - X)I_0}{\frac{1}{R_{\mathrm{L}}} + \frac{(n - X)}{Z_0}}. \tag{3.11}$$

Applying the Taylor series to (3.11) and ignoring the terms after the third order, the differential output voltage can be expressed as

$$V_{\mathrm{out}}(X) \approx 2R_{\mathrm{L}}nI_0 \left(\frac{\left(\frac{X - \frac{n}{2}}{n}\right)}{\left(1 + \frac{R_{\mathrm{L}}}{Z_0}\frac{n}{2}\right)^2} + \frac{\left(\frac{X - \frac{n}{2}}{n}\right)^3}{\left(1 + \frac{R_{\mathrm{L}}}{Z_0}\frac{n}{2}\right)^4} \left(\frac{R_{\mathrm{L}}n}{Z_0}\right)^2 \right). \tag{3.12}$$

The even order harmonic distortions are cancelled out due to differential operation, thus the third-order harmonic distortion becomes the biggest distortion.

Setting $X = n$ and ignoring the sinc frequency response, the maximum amplitude of the third order harmonic distortion $(\mathrm{HD3})$ normalized with respect to the signal amplitude can be calculated by

$$\mathrm{HD3\ [dB]} = 20\log\left(\frac{\frac{R_{\mathrm{L}}n}{2|Z_0|}}{1 + \frac{R_{\mathrm{L}}}{Z_0}\frac{n}{2}} \right)^2 - 20\log\left[\left(\frac{\frac{R_{\mathrm{L}}n}{2|Z_0|}}{1 + \frac{R_{\mathrm{L}}}{Z_0}\frac{n}{2}} \right)^2 + 1 \right]. \tag{3.13}$$

Assuming $R_{\mathrm{L}}n \ll 2Z_0$, the HD3 can be simplified to

$$\mathrm{HD3\ [dB]} = 20\log\left(\frac{R_{\mathrm{L}}n}{2|Z_0|} \right)^2. \tag{3.14}$$

The HD3 with respect to $R_{\mathrm{L}}/|Z_0|$ for 6 bit, 8 bit and 10 bit DACs are shown in Fig. 20. It can be observed that: 1. with one additional bit, the output impedance of a unit current cell has to be doubled to obtain the same HD3; 2. increasing the output impedance Z_0 by a factor of two improves the HD3 by 12 dB.

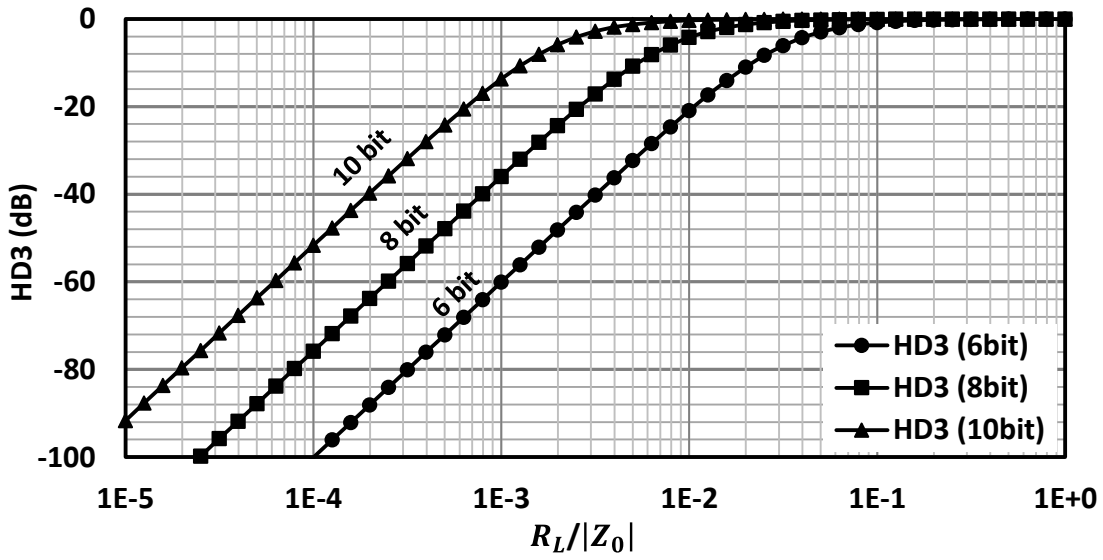


Fig. 20: HD3 versus $R_L/|Z_0|$ for 6 bit, 8 bit and 10 bit DACs

At low frequencies, a reasonable HD3 is achievable with single or double cascode current sources. At high frequencies, the parasitic capacitance of the transistor must be taken into account. Thus, the output impedance of a unit current cell can be modelled as a parallel connection of the resistor R_0 and the capacitor C_0, i.e.,

$$Z_0 = R_0 || \left(\frac{1}{j\omega C_0}\right) = \frac{R_0}{1 + j\omega C_0 R_0}. \tag{3.15}$$

Due to the effect of the parasitic capacitance, low harmonic distortion is quite difficult to achieve at higher frequencies. For example, to achieve a $HD3$ better than -36 dB for a 6 bit thermometer-coded DAC with 64 unit current cells and with 50 Ω load resistance, the output impedance Z_0 must be larger than 12.7 kΩ. This value can be easily achieved with cascode current sources at low frequencies. At higher frequencies ($\omega > 1/R_0C_0$), the parasitic capacitance C_0 starts to dominate the output impedance, i.e., $Z_0 \approx 1/\omega C_0$. At frequencies above 5 GHz, C_0 must be smaller than 3 fF, which is very challenging to achieve.

However, the analysis above assumes an infinite output impedance at the output node where the switch is turned off, which is not valid at high frequencies. As pointed out in [27], the effective output impedance due to the impedance difference between the ON- and OFF-state of an output node of a unit current cell leads to harmonic distortion. The effective output impedance of a unit current cell (Z_{0_eff}) can be calculated by

$$\frac{1}{Z_{0\_\mathrm{eff}}} = \frac{1}{Z_{0\_\mathrm{ON}}} - \frac{1}{Z_{0\_\mathrm{OFF}}}, \tag{3.16}$$

where Z_{0_ON} and Z_{0_OFF} are the output impedances in ON- and OFF-state, respectively. Ideally, the output impedance is Z_0 in ON-state and is infinite in OFF-state. Thus, the effective output impedance is Z_0 at lower frequencies. Due to the presence of parasitic capacitances, the output impedance falls with increased frequency in both states. According to (3.13), the distortions due to the finite output impedance can be minimized by keeping the effective output impedance

as large as possible. This can be achieved by minimizing the difference between Z_{0_ON} and Z_{0_OFF}. If the output impedance of the current cell does not change between the ON- and OFF-state, the output impedance of the DAC output is independent of the input code and does not cause distortion at the DAC output. Therefore, the output impedance (dominated by the output capacitance) at high frequencies has to be as equal as possible. This effectively relaxes the requirement on the output capacitance.

A unit current cell with parasitic capacitances is depicted in Fig. 21. The positive output is switched on and the negative output is switched off. Z_{0_ON} of the positive output is affected by the switch transistor capacitance ($C_{\mathrm{sw\_ON}}$), the common source node capacitance (C_{CSC}) and the bias node capacitance (C_{bias}). Z_{0_OFF} of the negative output is dominated by the drain capacitance $C_{\mathrm{sw\_OFF}}$ of the switch transistor.

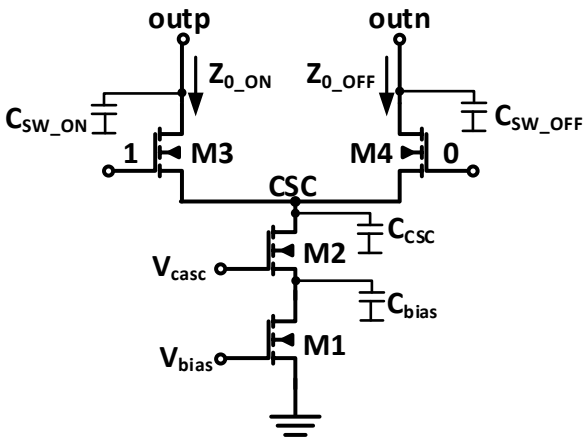


Fig. 21: Parasitic capacitances in a current cell

The output impedances over frequencies is shown in Fig. 22. While Z_{0_OFF} is much higher than Z_{0_ON} at low frequencies, they become much closer at the high frequencies above 10 MHz. The capacitance $C_{\mathrm{sw\_ON}}$ varies with the drain voltage of *M3*, which depends on the input code. The lower drain-source voltage causes larger $C_{\mathrm{sw\_ON}}$ and reduces the output impedance. Therefore, the effective impedance Z_{0_eff} is close to Z_{0_ON} as shown in Fig. 22.

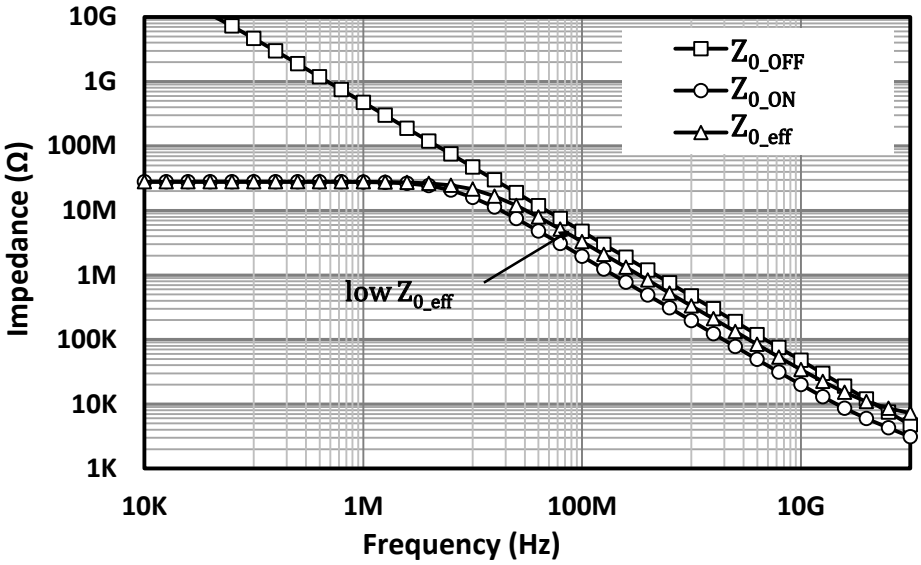


Fig. 22: Output impedances Z_{0_ON}, Z_{0_OFF} and Z_{0_eff} versus frequency for a simple current cell without output cascode transistors

To minimize the output impedance's dependency on the drain-source voltage, additional cascode transistors (*M5, M6*) and additional small trickle currents ($I_{trickle}$) can be added to the differential pair output as shown in Fig. 23 [27]. The small trickle currents keep the cascode transistors turned on, regardless whether the switching transistors (*M3, M4*) are on or off. Since both cascode transistors are always on, the variation of the gain and the drain capacitance with the input code is much smaller. This effectively increases Z_{0_eff} as shown in Fig. 24. Although Z_{0_eff} can be further improved by increasing the trickle current, the size of the cascode transistor and the power consumption also increase, reducing the output bandwidth.

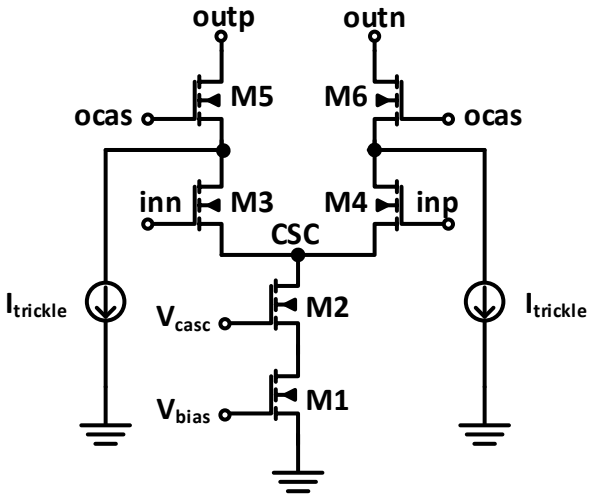


Fig. 23: Cascode transistors and trickle currents added to the current switch outputs

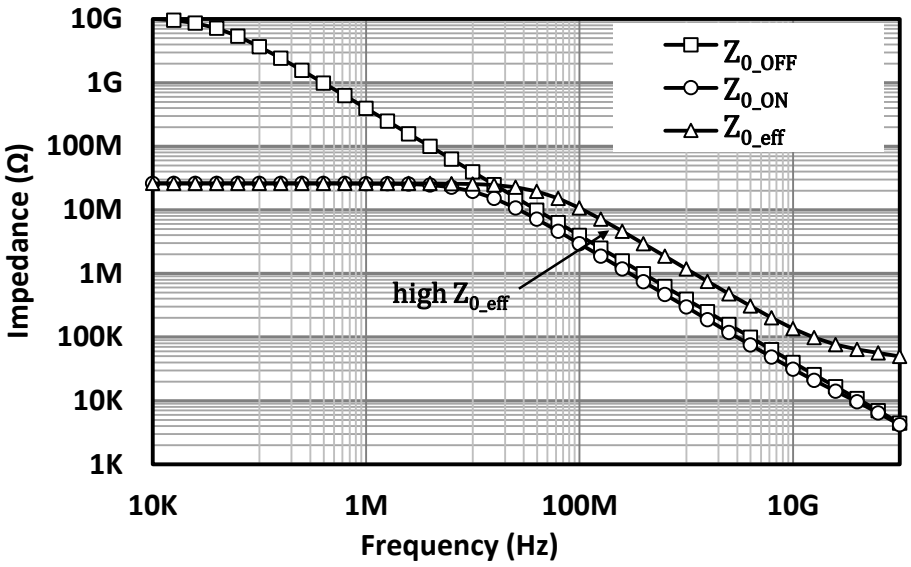


Fig. 24: Output impedances Z_{0_ON}, Z_{0_OFF} and Z_{0_eff} versus frequency for a current cell with cascode transistors and trickle currents added to the current switch output

3.2.3 Noise in Current Sources

The DAC performance is also limited by the circuit noise appearing at the DAC output. To achieve the target resolution, the circuit noise power has to be well below the quantization noise power within the frequency band of interest. One of the major noise contributors is the noise in the current sources, which include its intrinsic noise and the noise from the reference bias that is amplified by the current mirror, as shown in Fig. 25.

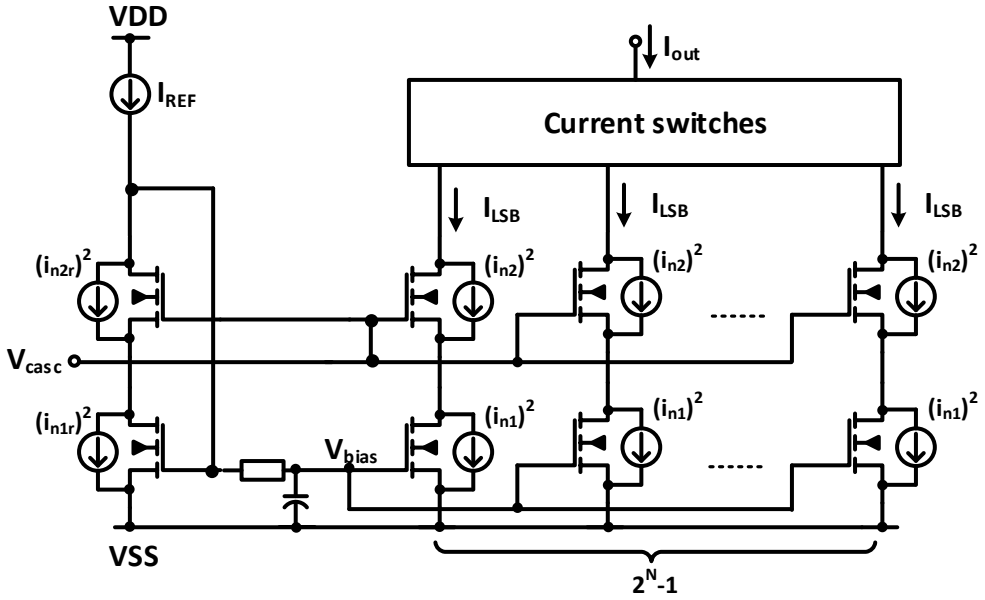


Fig. 25: Noise in the current sources of a DAC

The noise of a MOS transistor is dominated by the thermal noise of the channel resistance. The noise spectral density of the output current can be determined by [41],

$$\overline{i_n^2} = 4k_B T\gamma g_m \ , \tag{3.17}$$

where i_n is the noise current; $k_B = 1.38 \cdot 10^{-23}$ J/K is the Bolzmann constant; T is the temperature; g_m is the transconductance; γ is a process-defined parameter and is equal to 2/3 for long-channel MOS transistors. For a cascode current source, as long as the transconductance (g_{m2}) in the cascode transistor is much higher than the output admittance of the bias transistor ($1/r_{o1}$) the output current noise is dominated by the bias transistor [41]. This can be explained with the schematic shown in Fig. 26. The output noise current (i_{n2o}) is equal to the noise current to the drain node of *M1* (i_{nD1}), i.e., $i_{n2o} = i_{nD1}$, according to Kirchhoff´s current law. If $r_{o1} \gg 1/g_{m2}$, almost no current is flowing to *M1*, and there is no noise current contributing to the output current.

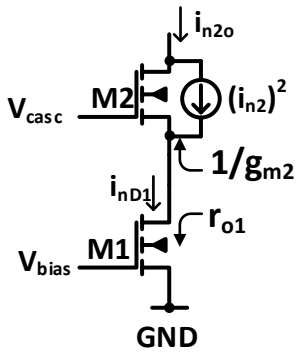


Fig. 26: Noise of the cascode transistor in the cascode current source

At low frequencies, the current noise contributed from the cascode transistor thermal noise can be calculated with

$$\overline{i_{n2o}^2} = \frac{\overline{i_{n2}^2}}{(1 + g_{m2} r_{o1})^2} \ , \tag{3.18}$$

where i_{n2} is the noise current of the cascode transistor and i_{n2o} is the noise current at the output. At higher frequencies, the capacitance at the drain node of the bias transistor reduces the output impedance, hence increasing the output noise.

Similarly, the noise contribution of the current switches is also negligible. This is because that the current switch transistors are either switched off or in their linear region, both has very small current gain. Therefore, their current noise become negligible small according to (3.18). Comparing to noise, the offset caused by random mismatch has much higher impact on the output noise and distortion.

For a current mirror, the noise current in the reference bias is mirrored by the current source output with a current mirror ratio of M: 1, i.e., the ratio between the reference current and the output current. Considering the reference bias noise and the intrinsic current source noise, the total DAC output current noise can be determined by

$$i_{n\_\mathrm{DAC}}^2 = 4k_B T\gamma g_{m\mathrm{LSB}}(2^N - 1)\left(f_{B\_\mathrm{cs}} + \frac{(2^N - 1)}{M} f_{B\_\mathrm{ref}}\right), \quad \textbf{(3.19)}$$

where $g_{m\mathrm{LSB}}$ is the transconductance of the bias transistor in the LSB unit current source; f_{B_cs} and f_{B_ref} are the bandwidth of the current source noise and the reference current noise, respectively; M is the ratio between the reference current and the LSB unit current, i.e., $M = I_{\mathrm{REF}}/I_{\mathrm{LSB}}$. For $(2^N - 1)/M \gg 1$, the output noise is dominated by the reference current noise, if f_{B_cs} and f_{B_ref} are equal. Therefore, the reference current noise has to be filtered out by a low pass filter to limit its bandwidth.

Assuming the full-scale signal power is $[(2^N - 1)I_{\mathrm{LSB}}]^2/2$ and substituting $g_{m\mathrm{LSB}} = 2I_{\mathrm{LSB}}/(V_{GS} - V_{th})$, and the reference current source noise is small because of the low pass filter, the signal-to-noise ratio (SNR) considering only thermal noise is given by

$$SNR\ [\mathrm{dB}] \approx 10\log\frac{I_{\mathrm{LSB}}(V_{GS} - V_{th})}{16k_B T\gamma f_{B\_\mathrm{cs}}}, \quad \textbf{(3.20)}$$

It can be seen that higher SNR can be achieved by increasing the LSB current I_{LSB} or increasing the overdrive voltage $(V_{GS} - V_{th})$ of the current source transistor.

For example, for an 8 bit DAC with a full-scale output of 500 mV and a 50 Ω output resistance, the LSB current is $I_{\mathrm{LSB}} \approx$ 0.08 mA; assuming that $V_{GS} - V_{th} =$ 400 mV and $f_{B\_\mathrm{cs}} =$ 20 GHz, the SNR is around 69 dB, which is much better than the required 50 dB SNR.

3.2.4 Switching Errors

3.2.4.1 Input signal crossing point

To reduce the glitch energy at the output, it is important to keep the current flowing through the switching pair constant during switching. The output current glitch of a current cell is mainly affected by the crossing point of the input signals, which needs to be optimized [26].

The output current of the current source is directly related to the common source node (*CSC*) voltage. The behavior of the common source node voltage (V_{CSC}) during switching at different crossing points is illustrated in Fig. 27. The voltage changes of *CSC* results in extra current flowing to the parasitic capacitor (C_{CSC}). To keep the current constant, V_{CSC} at the crossing point has to be equal to the voltage where one of the switches is completely turned on. This minimizes the disturbance of the voltage at *CSC* during the switching, and the disturbance on *CSC* is symmetric to its average value, as shown in the middle of the waveform in Fig. 27.

If the crossing point is too low, V_{CSC} has to be reduced to keep the current of switch transistors constant. In the worst case, the current source can be completely switched off and V_{CSC} is forced to ground. When the current source is switched on again, the large voltage difference between the outputs (*outp, outn*) and *CSC* causes large glitches. Similarly, a too-high crossing point causes positive voltage glitches at *CSC*.

The simulation results shown in Fig. 28 depict a sweep of different crossing points from 0 V to 1 V in 5 steps. It is observed that the disturbance on *CSC* can be minimized by selecting the

optimum crossing point. In [42], the calibration of the crossing point is done by measuring the glitch energy on the common source node and adjusting the clock circuitry by skewing the differential input signals accordingly.

The mismatch between *M3* and *M4* generates offset at the switch input, causing variation of the switching point and thus variation of the switching time between the current cells. To reduce this offset, one can reduce the random mismatch by increasing the area of the switch transistors while keeping the W/L ratio. Larger current gain can also reduce the offset, therefore one can also increase the W/L ratio. However, larger area increases the input capacitance, which lowers the slope of the input signal and increases jitter. To reduce jitter, a larger input driver and therefore more power is needed to maintain the steepness of the input signal slope.

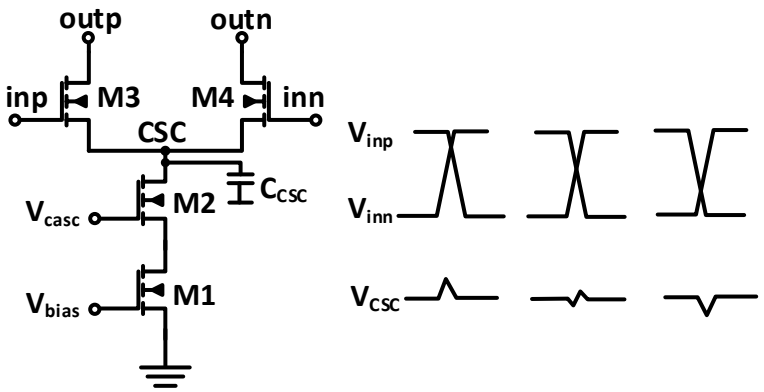


Fig. 27: Switching behavior of the common source node at different input crossing points

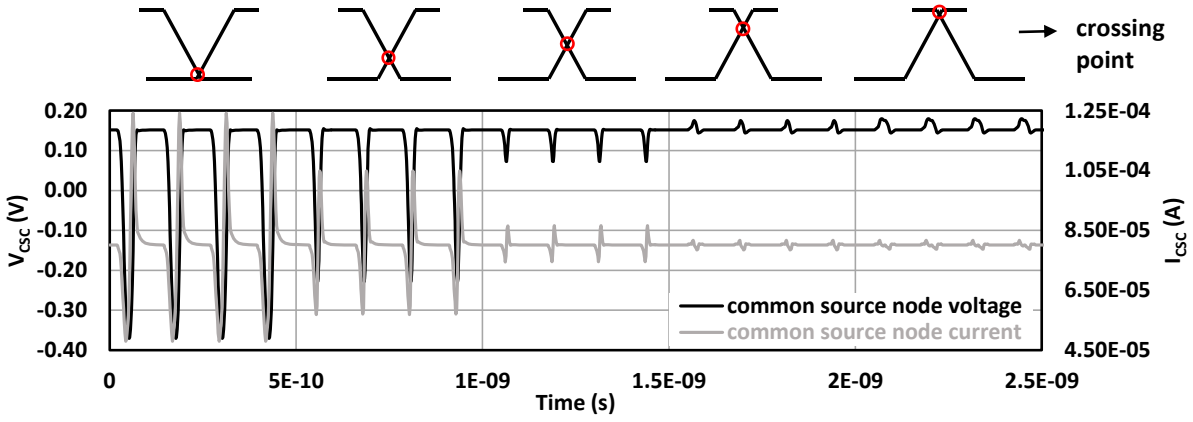


Fig. 28: Simulation of the voltage and current of the common source node by sweeping the crossing point

3.2.4.2 Signal dependent current switching

Due to the finite voltage gain ($g_m r_o$) of the switch transistors, a small amount of the output signal can feed through the switch transistors to the common source node [26]. The ratio between the output signal and the common source node voltage is defined by the gain of the switch transistors, i.e.,

$$\Delta V_{\mathrm{outp,outn}} = \Delta V_{\mathrm{CSC}}\, g_{\mathrm{m}} r_{\mathrm{o}}\,, \tag{3.21}$$

where g_{m} is the transconductance of the switch transistor transconductance; r_{o} is the output resistance of the switch transistor; V_{CSC} is the common source node voltage; $V_{\mathrm{outp,outn}}$ is the voltage of the node outp or outn, depending on which switch transistor (*M3* or *M4*) is turned on. Moreover, the output signals can also be coupled to the common source node at high frequencies through parasitic capacitance of the switch transistors.

The common source node dependency on the output signals is illustrated in Fig. 29. The differential output signals ($V_{\mathrm{outp}}, V_{\mathrm{outn}}$) are shown in the top graph and the voltage on *CSC* is shown in the bottom graph. According to (3.21), V_{CSC} is an attenuated version of the output signal. If *M3* is on (*M4* is off), *CSC* is connected to *outp* through *M3*, and V_{CSC} follows V_{outp}. If *M3* is off (*M4* is on), V_{CSC} follows V_{outn}.

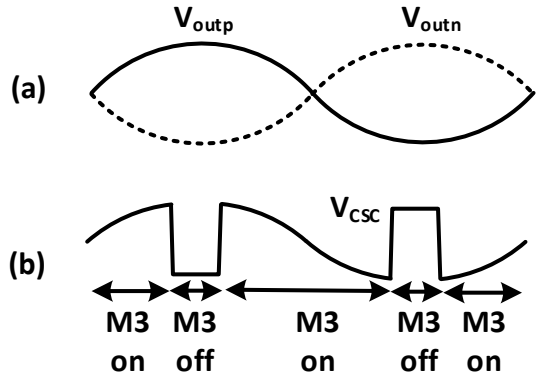


Fig. 29: Common source node dependency on the output signals

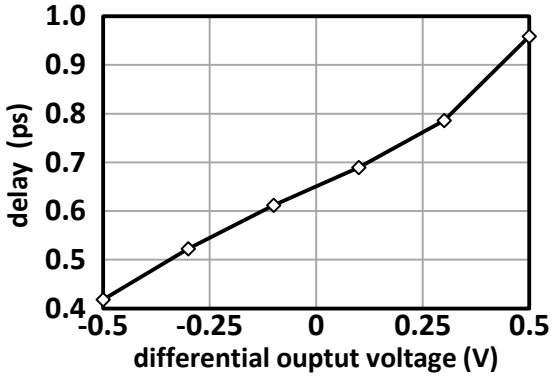


Fig. 30: Output delay versus the differential output voltage

These voltage variations on *CSC* result in variations of the gate-source voltage, and hence modulating the switching point of the switch pair. The actual voltage variation on *CSC* is determined by the differential output voltage where the switching occurs. Therefore, the switching point of the switch transistors is also a function of the output voltages. This causes a signal dependent switching at the output, i.e., signal dependent delay and slew rate of the output pulse, and effectively adds jitter to the output. The amplitude of the jitter depends on the differential output swing and the voltage gain of the switch transistors. Fig. 30 shows the output

delay of a LSB unit current cell versus the differential output voltage for one switching direction (i.e., from the positive node to the negative node). It can be seen that the variation of the output voltage results in a variation of the delay of 0.5 ps.

To reduce the signal dependency, one has to reduce the output swing on switch transistors. Since a cascode transistor can attenuate the voltage from its drain node to its source node, adding cascode transistors between the output and the drain node of the switching transistors, as shown in Fig. 23, can greatly attenuate the signal from the output to the switching pairs, minimizing the voltage variations on *CSC*.

3.2.4.3 Load mismatch between switch drivers

The load seen by the switch drivers determines the rise and fall time of the input signals of the current switches. The mismatches between the current switches cause timing errors between the current cells and thus increases the DAC output noise. The systematic mismatch is mainly caused by the scaling of the switch size in binary-weighted current cells.

Segmented DACs employ both unary-weighted and binary-weighted current cells. The unary current cells have identical current sources and current switches. For the binary-weighted current cells, the transistor width of the current switches are scaled with the current sources to keep the same current density. By doing so, the switching speeds between binary current cells are matched. Because of the scaling, the LSBs of the binary-weighted current cell exhibit lower input load compared to MSBs of the binary-weighted current cells.

To compensate for the differences in the input load between the binary-weighted current cells, dummy transistors can be added to the switch transistor inputs. As shown in Fig. 31, the MSB binary current is half of the unary current, i.e., $I_{\mathrm{b}} = I_{\mathrm{u}}/2$. The width of the switch transistor is scaled with the current accordingly, i.e., $W_{\mathrm{b}} = W_{\mathrm{u}}/2$. To match the input loads, dummy transistors with width of $W_{\mathrm{dummy}} = W_{\mathrm{u}}/2$ are added to both binary switch inputs.

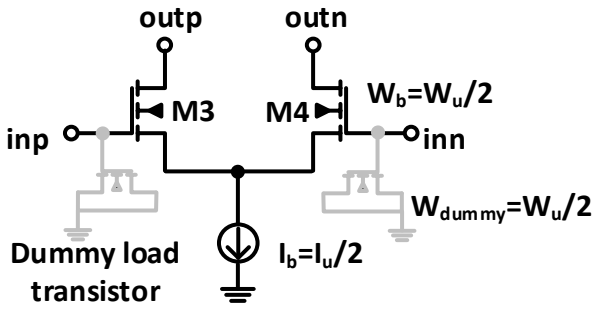


Fig. 31: Matching the input load with dummy transistors

3.2.4.4 Data-dependent switching noise

The digital data stream fed to the current switch input suffers from inter-symbol interference (ISI), i.e., the pulse width and the transition time of one symbol is interfered with by previous symbols. Due to ISI, the glitch energy at the current switch output is code dependent. The disturbance at the common source node of the switch pair is also data dependent, which affects

the current switching speed and generates more data-dependent jitter. Furthermore, for each clock cycle, the number of transitions depends on the input data pattern, which causes data-dependent errors. All of these data-dependent effects result in noise and distortion within the signal band of interest.

One of the possible solutions for reducing the data-dependent switching noise is the return-to-zero (RZ) switching method where the signal switches to zero during half the period of each clock cycle. The pulse width can be kept constant since each data pulse starts from one defined value. Ideally, the distortions caused by an uneven pulse width due to ISI can be eliminated. However, at higher frequencies, if the pulse does not return to its zero state, the memory effect of the input data steam cannot be eliminated. Thus, data-dependent errors are reduced but still remain. Furthermore, with RZ switching, the bandwidth and the clock frequency are doubled for the same sampling rate, which increases the power consumption and reduces the spectrum efficiency.

Based on RZ switching, the data dependent error can be further reduced by keeping constant switching activity. With constant switching activity, the distortion or noise is concentrated as a tone at the sampling frequency outside of the signal band. To realize constant switching activity, quad-switching can be used. This technique was first proposed in [43] and later employed in several designs [42, 44, 45].

The schematic of a quad-switching current cell is shown in Fig. 32 (a). It employs two, instead of one, differential pairs to switch one current source. The transistors *M1*, *M2*, *M3* and *M4* share the same common source node *CSC*. To show the operation of the quad-switching, a signal waveform with a data pattern of “011001” is shown in Fig. 32 (b). In each DAC clock cycle, one of the four switch transistors turns on and one turns off while the other two transistors remain off. If the input data stays unchanged, the current is re-routed through one other transistor to the same output node. If the input data is changed, the current is re-routed through one other transistor to another output node. This kind of switching can effectively be seen as time-interleaving of two RZ switches. Because of the constant switching activity regardless of the input data pattern, the disturbance to the common source node, the output node and the adjacent current cells are data-independent. The switching noise is thus moved to the sampling frequency, which is far away from the frequencies of interest, hence improving the high-frequency performance. One of the drawbacks of the quad-switching is an increase in design complexity, since it requires two return-to-zero signals to be generated at the input and the timing of the two input signals has to be well aligned to avoid additional errors. The dynamic power is also increased due to higher switching activities.

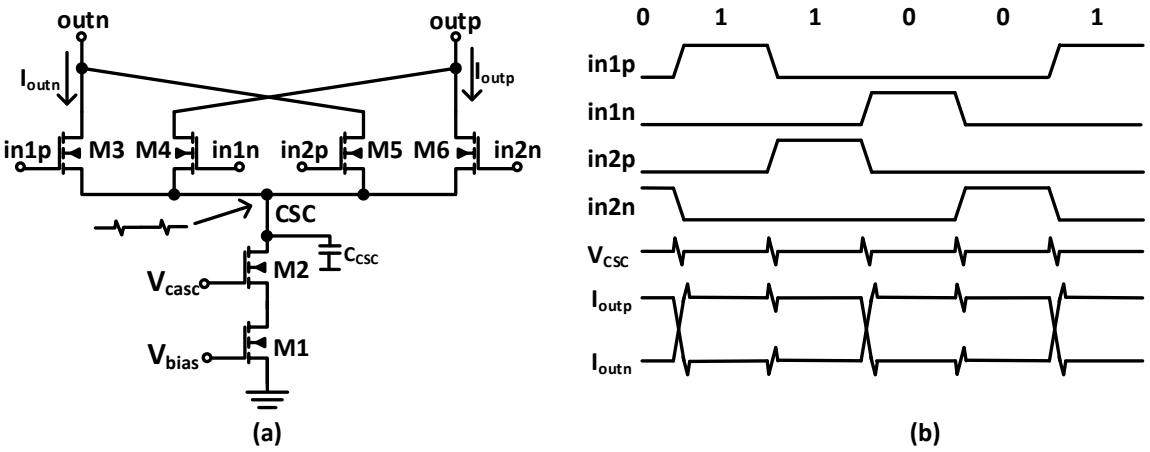


Fig. 32: (a) Quad-switching current cell, (b) input and output waveform

Another important data-dependent error source is the varying load seen by the clock driver of the last retiming latch or last MUX stage as reported in [46]. The load presented at the latch clock input depends on whether the latch state is changed or not, and thus affects the clock rising and falling edge. Assuming that all the retiming latches are driven by the same clock buffer, the clock-to-output delay depends on the activity of the latches. To eliminate this code transition dependent delay, a similar approach with constant activity can be applied to the latches. In [46], a retiming latch with constant load using the principle of constant activity is presented.

3.2.5 Output Swing

The full-scale output swing $V_{\text{out\_FS}}$ of an N bit DAC is set by the total current and the given load resistance R_L, i.e.,

$$V_{\text{out\_FS}} = 2^N \cdot I_{\text{LSB}} \cdot R_{\text{L}} \,, \tag{3.22}$$

where I_{LSB} is the LSB unit current.

The minimum LSB current is limited by the amplitude of the noise current and the matching requirement. The LSB unit current, together with the resolution, determine the total output current. The lower limit of the full-scale output swing is thus determined by the minimum LSB unit current.

A large output swing is desirable for many applications to mitigate noise limitations. For optical communications, a larger output swing can relax the requirement of the modulator driver. However, the maximum full-scale output swing also exhibits limitations.

First, the output voltage should not exceed the breakdown voltage of the output transistors. To achieve a high output bandwidth and a high switching speed, transistors with low threshold voltage and thin gate oxide are commonly used. The operation voltage of such types of transistors is normally below 1 V. The use of thick gate oxide transistors with high breakdown voltage increases the maximum output swing but decreases the bandwidth and the sampling

frequency. Higher supply voltage is necessary because of the larger output swing, which increases the power consumption.

Second, smaller output swing leads to less variation in the transistor operation point, which is beneficial for high switching speed and high linearity.

Third, large output swing increases the impedance difference between two output nodes, hence reducing the effective output impedance and increasing the distortion.

3.2.6 Glitch Error

The glitch errors in CS-DACs are mainly caused by the timing skew between the current switches. For example, if the MSB switches are slower than the other switches at the code transition from 100…00 to 011…11, the code 111…11 appears at the output for a short time period, causing a large glitch error. The power of the glitch error depends on the change in the input code and the switching activities. The glitch errors affect the output settling time and cause harmonic distortions inside the frequency band of interest.

Although the glitches cannot be eliminated, the glitch power can be reduced with segmented structure. For thermometer-coded DACs, the number of switched current cells is proportional to the change in the input code. However, the complexity and the area of the binary-to-thermometer decoder and the output stage limits the maximal number of bits of the thermometer code. The optimum ratio of segmentation is analyzed in [47], which trades the area and complexity for the desired linearity. For high-speed DACs, the sampling rate and the power consumption further limit the segmentation ratio. Less segmentation, i.e. using less unary cells, is preferred to reduce the number of current cells, hence reducing area and power as well as increasing the switching speed.

To further reduce the glitch error, the timing skew between the current cells has to be matched. Therefore, the current cells should be switched at the same time, and the output signal of each current cell should arrive at the output node at the same time. In order to minimize the skew, the most common approach is to use a binary-tree structure to distribute the clock and sum the output currents as shown in Fig. 33 (a). The binary-tree clock distribution exhibits an equal length from the clock driver to each current cell; the output summing network also has an equal length from each current cell to the output node. This kind of topology can achieve a very good timing matching, and the clock delay does not have to be matched to the output delay.

For a larger segmentation ratio, i.e., using more unary current cells, the area increases and the clock needs to be routed through a longer track to the current switches. The output summing network becomes larger as well. The increased area leads to larger parasitic resistances and capacitances in the clock distribution and output summing network, and thus decreases the maximum clock frequency and the output bandwidth. A trade-off between the timing matching and the parasitic reduction must be carefully chosen.

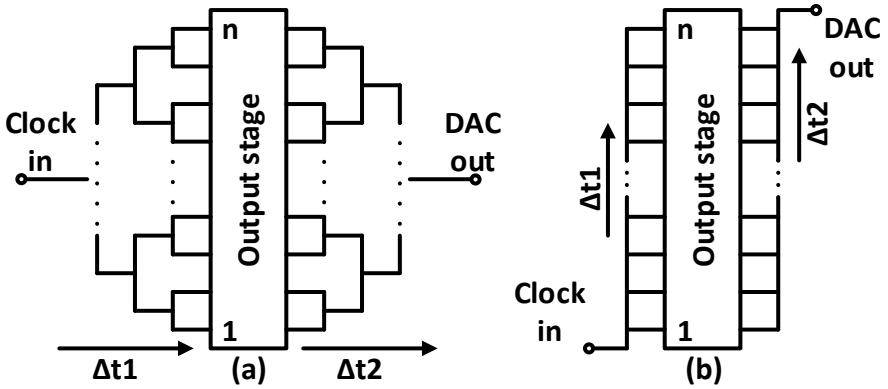


Fig. 33: (a) Binary-tree distribution (b) Propagation-delay matching distribution

Another approach for reducing the glitch error is the propagation-delay matching distribution as shown in Fig. 33 (b). In this structure, a straight clock line and a straight output line connect all current cell clock inputs and current outputs, respectively. Assuming that the delay along the clock distribution line from current-cell-1 to current-cell-*n* is $\Delta t1$ and the delay along the output line is $\Delta t2$, the delay on the clock line and the output line should be well matched, i.e., $\Delta t1 = \Delta t2$, to reduce glitch. Since the signal level degrades along the clock line and the output line, the resistance of the clock line and the output line must be well controlled to avoid too much performance degradation. The major benefits of this structure are the simple layout and the low output parasitics.

3.3 Time-Interleaved DACs

For conventional non-time-interleaved CS-DACs, the parallel inputs' data rate at the output stage has to be equal to the sampling rate of the converter. At very high frequencies, the design of the retiming latch and the switch driver becomes very challenging. The limited bandwidth of the switch drivers causes data-dependent jitter, i.e., ISI, with an increased frequency, hence increasing distortion at the current switch output and degrading the dynamic performances. The dynamic errors at the DAC output caused by the timing skews between the parallel inputs also increase with frequency. To overcome these problems for achieving higher sampling rates, time-interleaving techniques are commonly used.

3.3.1 Architectures

A time-interleaved (TI) DAC consists of *N* sub-DACs, each clocked at one of the *N* phases of the sampling clock. The input of each sub-DAC is generated by the same signal sampled at one of the *N* phases of the sampling clock. The original input signal can be reconstructed at the output by properly combining the outputs of *N* successively operated sub-DACs. The overall DAC sampling rate can be increased to *N* times the single sub-DAC sampling rate. There are two approaches to combining the sub-DAC outputs: analog multiplexing and parallel path.

The most common approach is to combine *N* sub-DACs' current outputs with an analog current multiplexer (MUX) operating at *N* times the sub-DAC sampling frequency. Fig. 34 (a)

illustrates a two-fold analog-multiplexed (AMUX) TI-DAC consisting of two sub-DACs. The two sub-DACs operate at the same clock frequency with opposite clock phases (180° out of phase). During one half clock period, one sub-DAC current output is connected to the output load whilst the other sub-DAC output is connected to the dummy output, representing zero at the output. Therefore, the output of the AMUX-TI-DAC can be seen as the sum of two opposite-phase return-to-zero (RZ) sub-DAC outputs. The waveforms of the current outputs are shown in Fig. 34 (b). The upper two curves show the current outputs of the two sub-DACs. During the output transition of each sub-DAC, their outputs are connected to the MUX dummy output. The transition errors thus do not affect the actual MUX output. Once the sub-DAC output currents are settled, they are routed to the actual MUX output. Therefore, the multiplexed output current (I_{out}) only consists of the settled points of the sub-DAC outputs and the output glitches are greatly reduced. The major drawback of the interleaved architecture is the higher area requirement and the higher power consumption, since two instead of one sub-DAC output stages are employed and half of the static power is wasted at the dummy output. This interleaved architecture has been employed in many high-speed DAC designs [48–52].

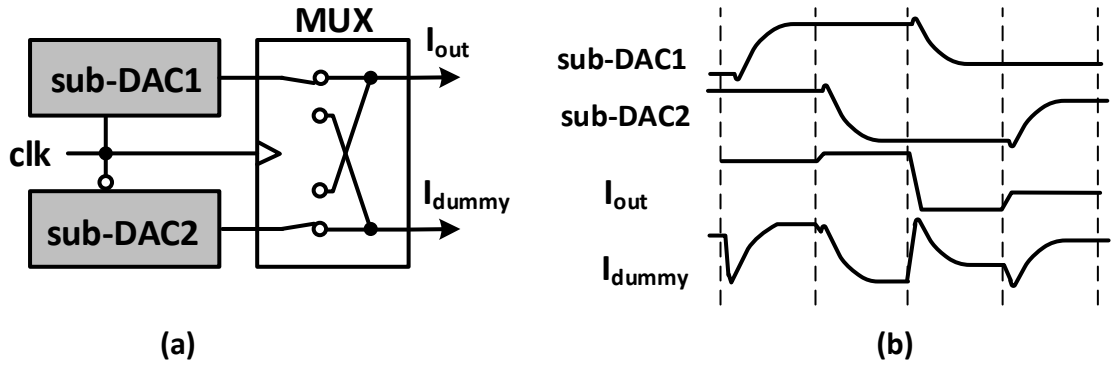


Fig. 34: (a) AMUX-TI-DAC, (b) output waveform

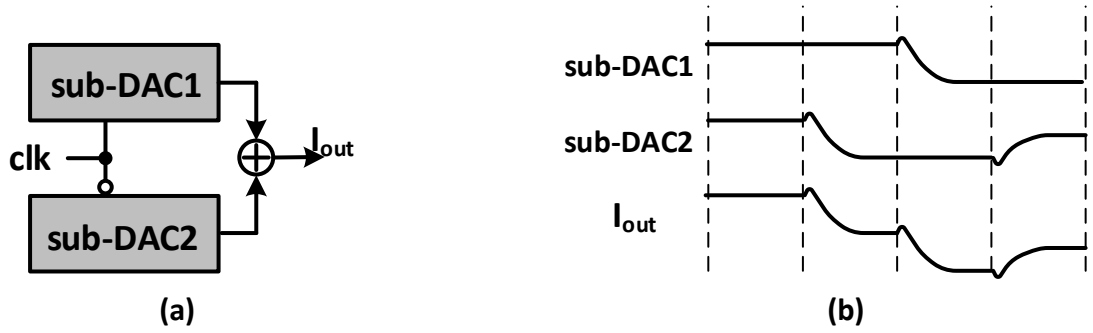


Fig. 35: (a) Parallel-path-TI-DAC, (b) output waveform

The design of analog MUXs can become much more challenging at very high frequencies due to the extremely short pulse width and very high timing requirement. To overcome this problem, another time-interleaving approach based on the parallel-path method proposed in [53] can be used. A two-fold parallel-path (PP) TI-DAC is shown in Fig. 35 (a), where both sub-DAC outputs are simply summed up. Each sub-DAC operates on the opposite phase of the clock and each output value is held for a whole clock period. Therefore, the overall DAC output always

consists of two consecutive sub-DAC output values. The waveform of both sub-DAC outputs and the summed output are illustrated in Fig. 35 (b). Since the sampling rate of the PP-TI-DAC is doubled, the image frequencies of the same signal are shifted with the sampling frequency to much further locations in the spectrum. This can also be understood as the images of the two sub-DAC in their 2nd and 3rd Nyquist band being cancelled out. The images have to be filtered out to avoid interferences. Since the images are far away from the signal frequency, the requirement on the output reconstruction filter is greatly reduced. The output signal bandwidth can be extended much closer to the Nyquist frequency of each sub-DAC without being interfered with by the image frequencies, hence increasing the utility of the output spectrum. However, since the hold time of the individual sub-DAC does not change, the sinc-function frequency response of the PP-TI-DAC output are still the same as that of a single sub-DAC, limiting the analog bandwidth.

3.3.2 Spectrum of TI-DACs

As mentioned in the previous section, combining the two sub-DAC outputs with an analog MUX can be seen as summing up two opposite-phase RZ-sub-DAC outputs. The return-to-zero operation halves the pulse width of each individual sample to $T_s/2$, where T_s is the sampling period of each single sub-DAC. According to (2.5), the sinc frequency response of the RZ-sub-DAC output is $|½\mathrm{sinc}(f/(2f_s))|$ with f_s as the sampling frequency of the sub-DACs. Therefore, the zeros in the spectrum are located at multiples of $2f_s$, i.e., $i \cdot 2f_s$ with i as a non-zero integer. Although the (RMS) amplitude of a RZ-DAC is only half that of a normal non-return-to-zero (NRZ)-DAC, the frequency response is much flatter.

In the PP-TI-DAC, two sub-DAC output are directly summed up without a return-to-zero operation. The hold time of each sample does not change and thus the sinc frequency response of the PP-TI-DAC output is still the same as that of a single sub-DAC with double the amplitude. Fig. 36 illustrates the frequency responses of the NRZ-sub-DAC, the RZ-sub-DAC, the AMUX-TI-DAC and the PP-TI-DAC.

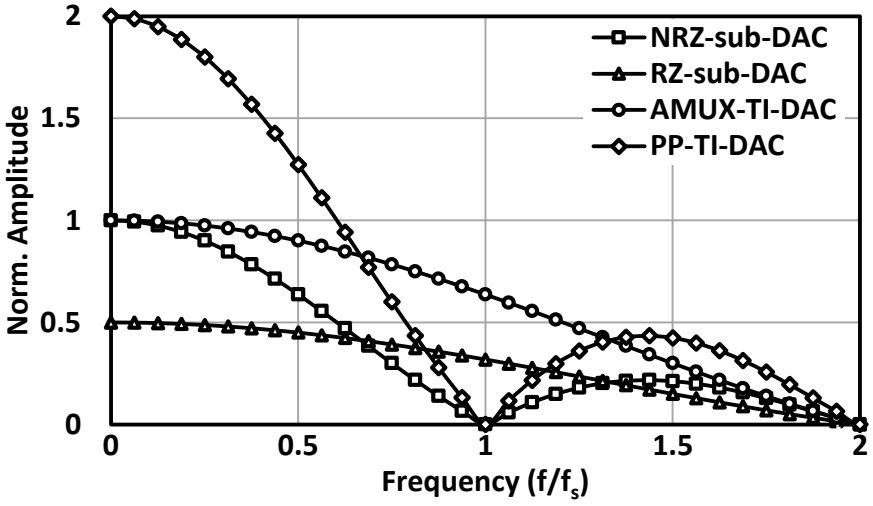


Fig. 36: Output frequency responses of the NRZ-sub-DAC, the RZ-sub-DAC, the AMUX-TI-DAC and the PP-TI-DAC

To gain better insight, the mathematical analysis of the TI-DAC spectrum is also given here.

The input data is generated by sampling the same input data with a skew of half the sampling period. Fig. 37 shows an example of the input data sampled from a sine wave. The inputs of sub-DAC1 are sampled at $i \cdot T_\mathrm{s}$, while the inputs of sub-DAC2 are sampled at $(i+\frac{1}{2})T_s$ with i as a positive integer.

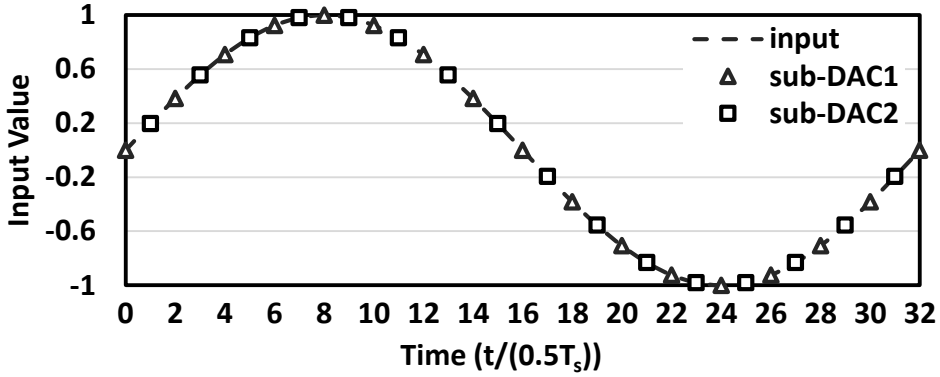


Fig. 37: Input data of the sub-DACs in TI-DACs

With (2.11) and the shift property of Fourier Transformation, the spectrum of a single sub-DAC with delay ΔT can be written as

$$X_{\text{SH-sub}}(f) = \frac{1}{T_\mathrm{s}} \sum_{n=-\infty}^{+\infty} \left(X\left(f - \frac{n}{T_\mathrm{s}}\right) e^{-\mathrm{j}2\pi n \Delta T/T_\mathrm{s}} \right) e^{-\mathrm{j}\pi f T_\mathrm{s}} \left(\frac{\sin(\pi f T_\mathrm{s})}{\pi f T_\mathrm{s}} \right). \qquad \textbf{(3.23)}$$

In an ideal two-fold TI-DAC, the ΔT is set to 0 and $T_\mathrm{s}/2$ for sub-DAC1 and sub-DAC2, respectively.

For the PP-TI-DAC, the frequency response can be determined by summing up the spectrum of both sub-DACs.

$$X_{\text{SH-PP}}(f) = \frac{1}{T_\mathrm{s}} \sum_{n=-\infty}^{+\infty} \left(X\left(f - \frac{n}{T_s}\right)\left(1 + e^{-\mathrm{j}n\pi}\right) \right) e^{-\mathrm{j}\pi f T_\mathrm{s}} \left(\frac{\sin(\pi f T_\mathrm{s})}{\pi f T_\mathrm{s}} \right). \qquad \textbf{(3.24)}$$

Note that the terms of $X_{\text{SH-PP}}(f)$ is zero if n is odd and only the terms with n as an even number remain. It implies that the images of the two sub-DACs around $\pm 1f_\mathrm{s}, \pm 3f_\mathrm{s}, \pm 5f_\mathrm{s}, \ldots$ have opposite phases and are cancelled out in the output spectrum. An example of the spectrums of the sub-DACs and of the PP-TI-DAC is shown in Fig. 38 (a).

Substituting n with $2k$, where k is a integer number, (3.24) can be written as

$$X_{\text{SH-PP}}(f) = \frac{2}{T_\mathrm{s}} \sum_{k=-\infty}^{+\infty} \left(X\left(f - \frac{k}{\frac{T_\mathrm{s}}{2}}\right) \right) e^{-\mathrm{j}\pi f T_\mathrm{s}} \left(\frac{\sin(\pi f T_\mathrm{s})}{\pi f T_\mathrm{s}} \right). \qquad \textbf{(3.25)}$$

This shows that the sampling rate of the PP-TI-DAC is double that of the individual sub-DAC whilst the sinc frequency response remains the same as that of the sub-DACs.

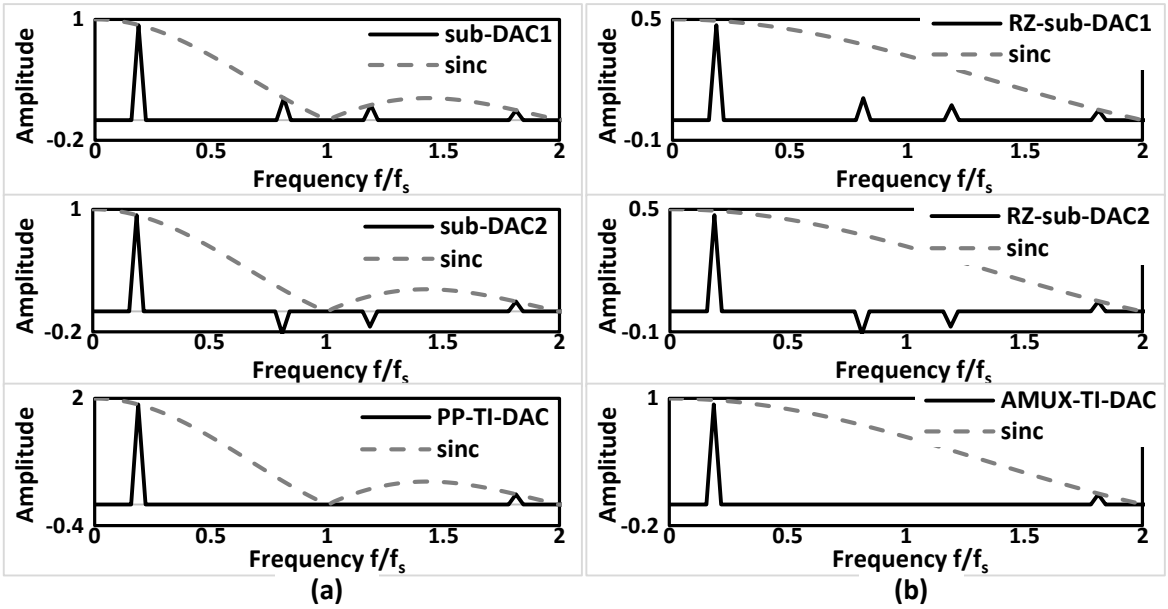


Fig. 38: Frequency response and spectrum (a) of sub-DAC1, sub-DAC2 and the PP-TI-DAC and (b) of RZ-sub-DAC1, RZ-sub-DAC2 and the AMUX-TI-DAC

A similar approach can be applied to analyze the output spectrum of the AMUX-TI-DAC. The AMUX-TI-DAC output can be seen as the sum of the outputs of two RZ-sub-DACs due to the analog-MUX at the output. The spectrum of both RZ-sub-DAC outputs can be written as

$$X_{\text{SH-RZ-sub1}}(f) = \frac{1}{2T_\text{s}} \sum_{n=-\infty}^{+\infty} X\left(f - \frac{n}{T_\text{s}}\right) e^{-\text{j}\pi f\frac{T_\text{s}}{2}} \left(\frac{\sin\left(\pi f \frac{T_\text{s}}{2}\right)}{\pi f \frac{T_\text{s}}{2}}\right) \tag{3.26}$$

and

$$X_{\text{SH-RZ-sub2}}(f) = \frac{1}{2T_\text{s}} \sum_{n=-\infty}^{+\infty} X\left(f - \frac{n}{T_\text{s}}\right) e^{-\text{j}n\pi}\, e^{-\text{j}\pi f\frac{T_\text{s}}{2}} \left(\frac{\sin\left(\pi f \frac{T_\text{s}}{2}\right)}{\pi f \frac{T_\text{s}}{2}}\right), \tag{3.27}$$

respectively.

The spectrum of the AMUX-TI-DAC is thus the sum of the two spectrums, i.e.,

$$X_{\text{SH-AMUX}}(f) = \frac{1}{2T_\text{s}} \sum_{n=-\infty}^{+\infty} X\left(f - \frac{n}{T_\text{s}}\right)\left(1 + e^{-\text{j}n\pi}\right) e^{-\text{j}\pi f\frac{T_\text{s}}{2}} \left(\frac{\sin\left(\pi f \frac{T_\text{s}}{2}\right)}{\pi f \frac{T_\text{s}}{2}}\right). \tag{3.28}$$

Similarly, only the terms with n as an even number remain and the images of the two sub-DACs around $\pm 1f_\text{s}, \pm 3f_\text{s}, \pm 5f_\text{s}, \ldots$ are cancelled out in the output spectrum. An example of the spectrums of the RZ-sub-DACs and of the AMUX-TI-DAC are illustrated in Fig. 38 (b).

Substituting n with $2k$, where k is a integer number, $X_{\text{SH-AMUX}}(f)$ can be written as

$$X_{\text{SH-AMUX}}(f) = \frac{1}{T_\text{s}} \sum_{k=-\infty}^{+\infty} X\left(f - \frac{k}{\frac{T_\text{s}}{2}}\right) e^{-\text{j}\pi f\frac{T_\text{s}}{2}} \left(\frac{sin\left(\pi f \frac{T_\text{s}}{2}\right)}{\pi f \frac{T_\text{s}}{2}}\right). \tag{3.29}$$

The spectrum of the AMUX-TI-DAC is thus the same as that of the single sub-DAC operating at double the sampling rate.

3.3.3 Errors in TI-DACs

Although the interleaved structure, especially the AMUX-TI-DACs, can reduce many dynamic errors in a single sub-DAC, it also introduces some new dynamic errors, which may limit the performance and has to be dealt with carefully. The major errors are the offset error, the gain error and the duty-cycle error.

3.3.3.1 Offset and gain error

To analyze the impact of the offset and gain error in an AMUX-TI-DAC, the sub-DAC outputs can be presented by the sum of the ideal output and the error part. By setting one sub-DAC output (sub-DAC1) as a reference, the errors are only applied to sub-DAC2. Therefore, the real output signal of sub-DAC2 at i sample $x_{\mathrm{sub2}}(i)$ can be written as the sum of the ideal sub-DAC2 output $x_{\mathrm{sub2\_id}}(i)$ and the error term $x_{\mathrm{err}}(i)$, i.e., [49]

$$x_{\mathrm{sub2}}(i) = x_{\mathrm{sub2\_id}}(i) + x_{\mathrm{err}}(i)\,. \tag{3.30}$$

Fig. 39 (a) shows a sine wave signal generated by the DAC. Assuming the offset error is ε_{os}, the output of sub-DAC2 is the sum of its ideal output and the error term $x_{\mathrm{err}}(i) = \varepsilon_{os}$. The error signal extracted from the ideal output signal is shown in the lower graph. It can be seen that the error signal generates a squared waveform with a frequency of half the sampling rate of the TI-DAC, i.e., $f_{\text{s-TI}}/2$. In the output spectrum, this causes a spur at DC and at $f_{\text{s-TI}}/2$.

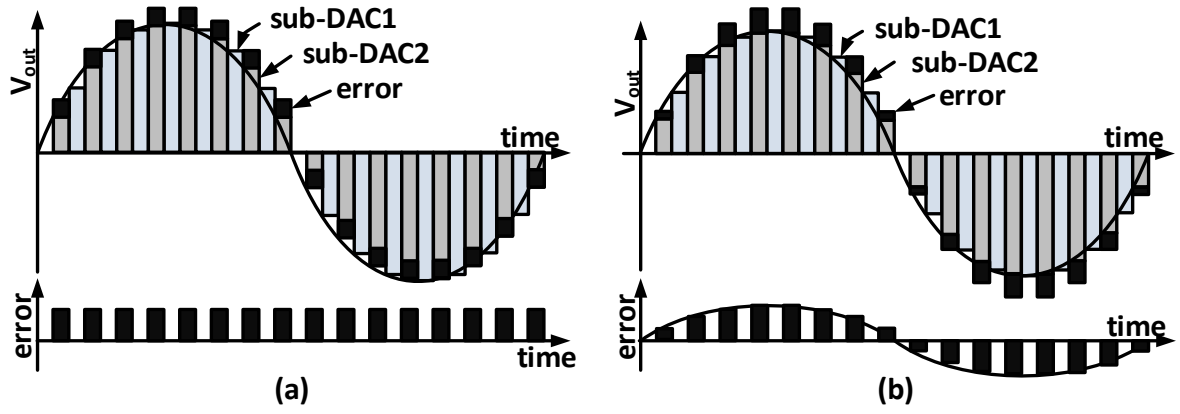


Fig. 39: (a) offset error and (b) gain error of AMUX-TI-DACs

A gain error has no impact on the output spectrum for a single non-interleaved DAC, since it only changes the signal amplitude. In an AMUX-TI-DAC, a gain error ($\varepsilon_{\mathrm{gain}}$) of sub-DAC2 with respect to the gain of sub-DAC1 generates an error signal proportional to the sub-DAC2 output,

$$x_{\mathrm{err}}(i) = \varepsilon_{\mathrm{gain}} x_{\mathrm{sub2\_id}}(i). \tag{3.31}$$

In the top graph of Fig. 39 (b), the sine wave outputs of both sub-DACs are shown. The output of sub-DAC2 consists of an ideal output and a gain error term. In the lower graph, the gain error signal is shown separately. The error signal generates an RZ-sine-wave proportional to the ideal sub-DAC2 output. The error signal has the same frequency as the DAC output signal, thus it adds to the signal of the output spectrum. The image frequencies of the error signal are around half the sampling frequency and generate a spur within the signal band, i.e.,

$$f_{\mathrm{sp}} = \frac{f_{\text{s-TI}}}{2} - f_{\mathrm{sig}}\,. \tag{3.32}$$

where f_{sig} is the signal frequency and f_{sp} is the spur frequency.

The normalized spur power with respect to the signal (P_{sp}) in decibel due to the gain error can be determined by

$$P_{\mathrm{sp}}[\mathrm{dB}] = 20\log\varepsilon_{\mathrm{gain}} - 6 - X_{\mathrm{sinc}}\,, \tag{3.33}$$

where X_{sinc} is the ratio between the signal and the image due to the sinc frequency response of the DAC output. The 6 dB is coming from the return-to-zero behavior, i.e, the RMS amplitude is reduced to half. Due to the sinc frequency response, both the signal and the image of the error signal reduce with the frequency. X_{sinc} can be written as

$$X_{\mathrm{sinc}}[\mathrm{dB}] = 20\log\frac{\mathrm{sinc}\left(\frac{f_{\mathrm{sig}}}{f_{\text{s-TI}}}\right)}{\mathrm{sinc}\left(\frac{\frac{f_{\text{s-TI}}}{2} - f_{\mathrm{sig}}}{f_{\text{s-TI}}}\right)} \tag{3.34}$$

Note that the ratio X_{sinc} is frequency dependent. At low frequencies near DC and at the Nyquist frequency, X_{sinc} is approximately equal to 4 dB and - 4 dB, respectively. Therefore, the normalized spur power at both frequencies can be written as

$$P_{\mathrm{sp,DC}}[\mathrm{dB}] = 20\log\left(\varepsilon_{\mathrm{gain}}\right) - 10 \tag{3.35}$$

and

$$P_{\mathrm{sp,Nyquist}}[\mathrm{dB}] = 20\log\left(\varepsilon_{\mathrm{gain}}\right) - 2, \tag{3.36}$$

respectively.

3.3.3.2 Duty-cycle Error

In the AMUX-TI-DAC, the final multiplexing is sensitive to both edges of the $f_{\text{s-TI}}/2$ clock as new data is presented to the output on both edges. For clocks with a 50% duty cycle, the output of each channel has the same sampling period of $1/f_{\text{s-TI}}$ as the desired one. If the duty cycle differs from 50%, then the sampling period varies from its ideal value, causing spurs in the output spectrum. The frequency of the timing error is half the sampling frequency, i.e., $f_{\text{s-TI}}/2$.

The output error is then the intermodulation product of the timing error frequency and the signal frequency, and generates a spur inside the signal band at a frequency of $f_{\text{s-TI}}/2 - f_{\text{sig}}$.

The duty-cycle error of a TI-DAC output is depicted in Fig. 40 [49]. The top graph shows the sub-DAC outputs with duty-cycle error while the lower graph shows only the error signal. The duty-cycle errors generate a pulse train at half the sampling frequency. Assuming that the error pulse width is ΔT, the duty-cycle error ε_{dc} is determined by the ratio between the timing error ΔT and half the sampling clock period $f_{\text{s-TI}}/2$, i.e.,

$$\varepsilon_{\text{dc}} = \frac{\Delta T f_{\text{s-TI}}}{2}. \tag{3.37}$$

The error pulse height is the difference between consecutive samples. Thus, the error pulses can be approximated with the derivation of the output signal. Ignoring the sinc frequency response of the signal frequency, the error can be determined by

$$x_{\text{err}}(i) = \frac{2\pi f_{\text{sig}}}{f_{\text{s-TI}}} \cos\left(\frac{2\pi f_{\text{sig}}}{f_{\text{s-TI}}} i\right). \tag{3.38}$$

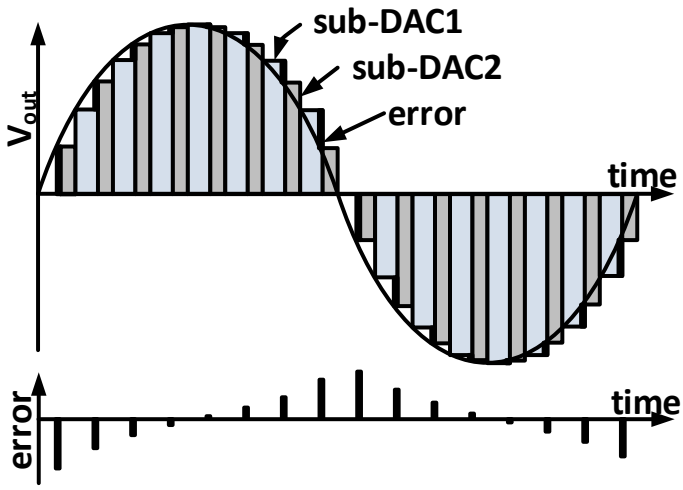


Fig. 40: Duty-cycle error of AMUX-TI-DACs

For small duty-cycle error, the error pulse width is small compared to the sampling period; therefore, the sinc amplitude roll-off is negligible in the first Nyquist band. The amplitude of the spur in the output spectrum is determined by the product of the duty-cycle error ε_{dc} and the error amplitude in (3.38). Since the error and the signal have the same sinc frequency response, the normalized spur power can be calculated by

$$P_{\text{sp}}[\text{dB}] = 20 \log\left(\frac{2\pi f_{\text{sig}} \varepsilon_{\text{dc}}}{f_{\text{s-TI}}}\right). \tag{3.39}$$

Duty-cycle error can be corrected at the clock path and at the output stage. A correction circuit for the clock duty cycle is shown in Fig. 41. The negative feedback resistor automatically sets the common mode voltage of the inverter, which results in an 50% duty cycle. Additional tunable capacitor at the output can be used to tune the skew between the differential clocks.

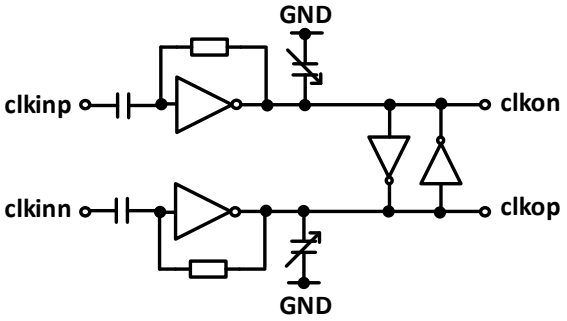


Fig. 41: Duty-cycle and skew correction of the pseudo differential input clock

At the output stage, duty-cycle error is caused by the switch offset. The offset error can be corrected by adjusting the common mode voltage of the input clocks, i.e., Vcalp and Vcaln, as shown in Fig. 42 (a). If a silicon-on-insulator (SOI) technology is available, the backgate voltage can be used to adjust the threshold voltage of the switch transistors to compensate the offset, as shown in Fig. 42 (b).

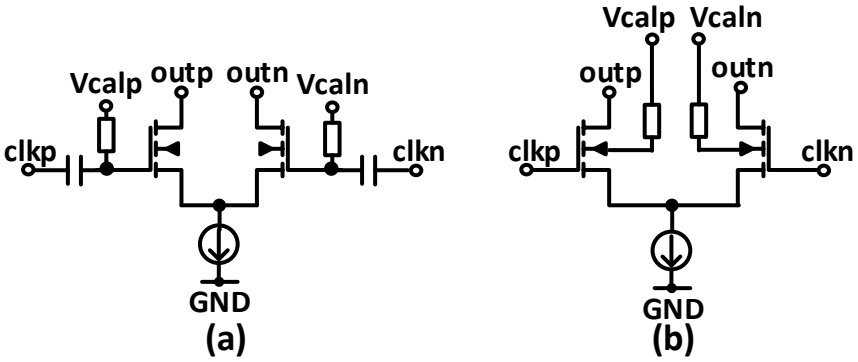


Fig. 42: Duty-cycle error correction for output switches: (a) adjust the input common mode voltage, (b) adjust the backgate voltage in a SOI technology

The duty-cycle error has to be measured to allow error correction. One idea is to output rectangle pulse at the DAC output and measure the average output voltage as shown in Fig. 43. First, sub-DAC1 output is set to maximum and sub-DAC2 is set to minimum. The average output voltage (V_{avg1}) is measured with a low-speed high-resolution ADC. Afterwards, both outputs are swapped to measure V_{avg2}. The minimum duty-cycle error is achieved at which the difference between V_{avg1} and V_{avg2} becomes minimum.

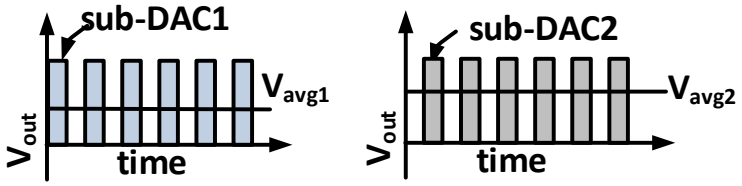


Fig. 43: Duty-cycle error measurement

3.3.4 Circuit Implementation

A widely used circuit implementation for multiplexing the output currents is shown in Fig. 44 [48, 50, 52, 54]. Additional clock switches are added to the output of the data switches of each normal current cell. The clock switches route the current to the outputs (*outp*, *outn*) at the clock high level and reroute the current to the dummy outputs (*dummyp*, *dummyn*) at the clock low level in each current cell. Therefore, the currents of both current cells are routed to the outputs alternatively, realizing the multiplex operation. The data switches control whether the currents are flowing through positive output nodes or negative output nodes.

Fig. 44 (b) illustrates the timing diagram for outputting a data pattern of "011010". The signals *in1, in2, CLK* are the differential inputs and the differential clock signal, respectively. The values of *in1* and *in2* are converted to the output at the rising edge and at the falling edge of *CLK*, respectively. Input *in2* is skewed by half the data period against input *in1*. This allows more settling time for the data switch. Since all output currents are resampled by the clock switches, the timing alignment of the output signal between the current cells is improved.

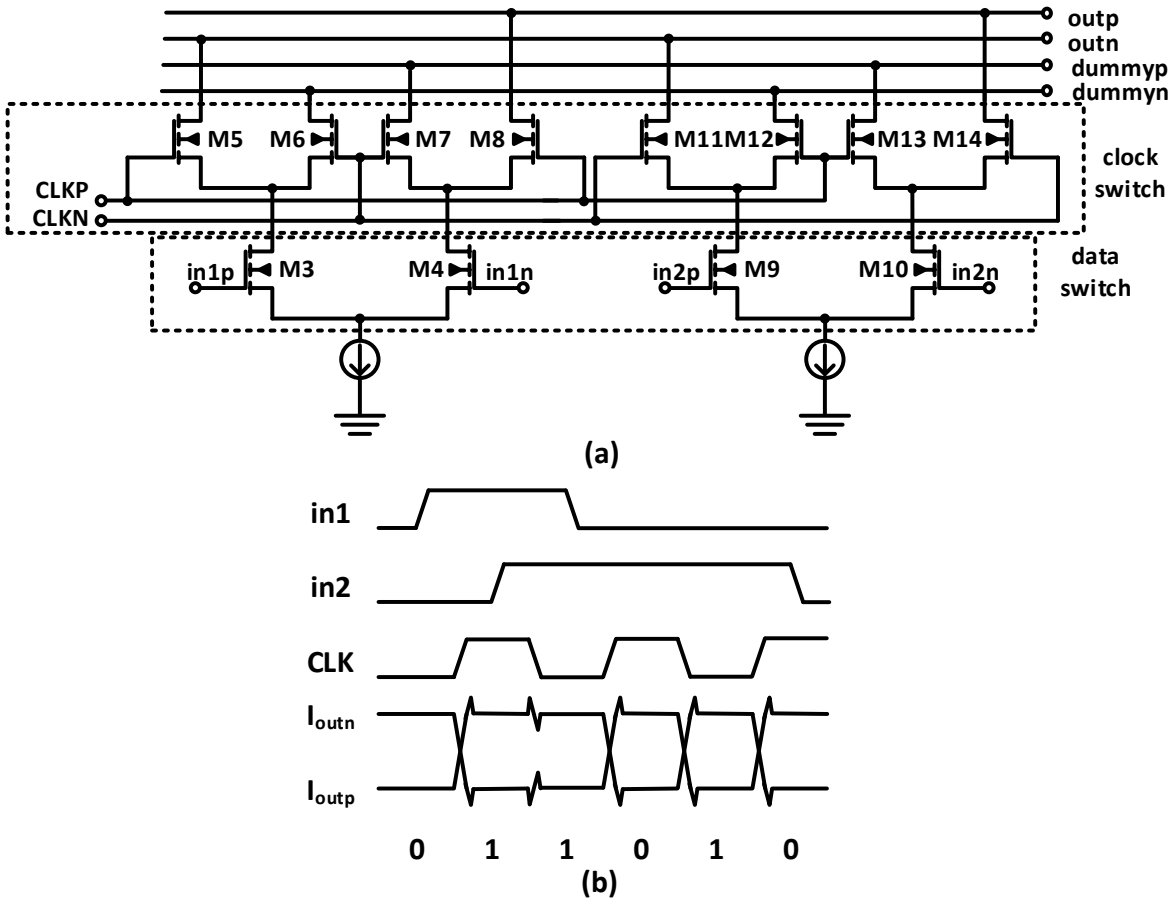


Fig. 44: (a) Current switch with multiplexer and its (b) timing diagram

To further reduce the dynamic error due to the data dependent switching effect, one can combine the quad-switching and the multiplexing technique as reported in [49]. In this case, the number of data switch transistors is increased from two to four. This additional complexity

could introduce some new errors and parasitics, which might result in more negative than positive impact on the circuit performance.

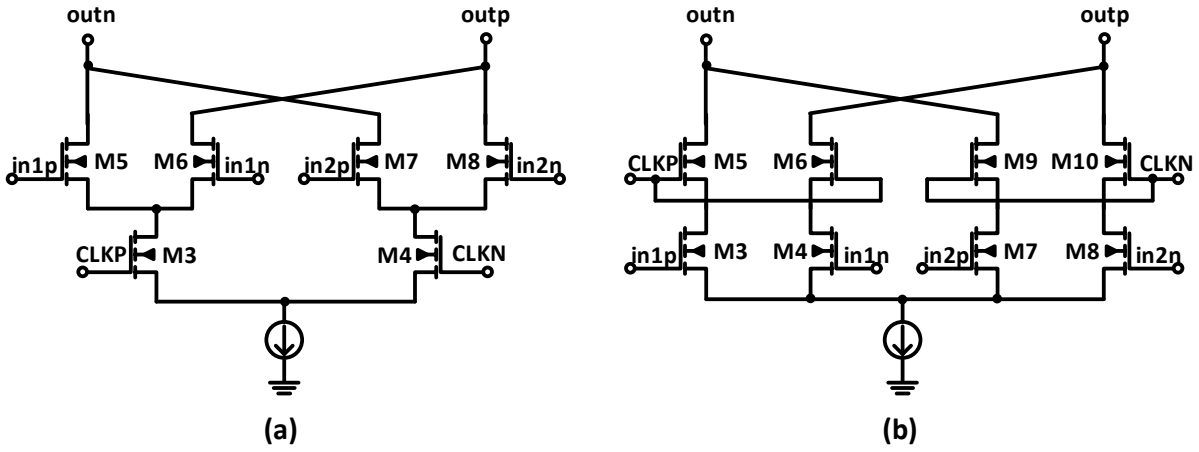


Fig. 45: Low-power current cells: (a) typical CML multiplexer, (b) CML multiplexer with the clock switches at the top of data switches

One of the drawbacks of the current cell shown in Fig. 44 is the high power consumption due to the use of two current sources. The two current sources also introduce extra mismatches. To minimize the power consumption and the mismatches, alternative current cell structures with only one current source, as shown in Fig. 45, can be used. Fig. 45. (a) shows a typical CML multiplexer in which the clock switches are below the data switches, connecting the current source output and the common-source node of the data switches. The current flows through one of the clock switches to the corresponding data switches, which controls the polarity of the current output. However, the typical CML multiplexer has also its drawbacks compared to the current multiplexer shown in Fig. 44. First, since the data switches are directly connected to the output, the code dependent switching glitches appear directly at the current output without masking by the clock switches. Second, the data-switch drivers are normally CMOS buffers with rail-to-rail output swing, which is around 1 V. Therefore, the clock input common-mode level has to be shifted down to keep the clock switch transistors in saturation region, which can cause that the drain-source voltage of the clock switches exceeds the maximum allowed voltage. The shifting-down of the voltage level can also lead to a higher power consumption for the clock driver. Third, the data switches act as a common-gate stage between the output and clock switches. However, they are biased at the low-gain region during the clock switching, thus reducing the switching speed of the clock switches.

To alleviate some of the issues of the typical CML multiplexer, the clock switches can be moved to the top of the data switches, as shown in Fig. 45 (b). This structure eliminates the need of level-shifting of the clock signal. During the data switching, the corresponding clock switches are turned off, isolating the data-dependent switching glitches from the output. Since the clock switches are still connected to the output, small glitch power can coupled to the output via parasitic capacitance. During the off state of one of the clock switch pairs (e.g., M5 and M6), one source node of the two clock transistors is floating, which affects the switching speed of

the clock transistor and causes distortion at the output. One of the possible solution to this issue is to add a trickle current to each source node of the clock transistors, similar to the one shown in Fig. 23, to keep the transistors turned on, at the cost of increasing the power consumption.

3.4 Bandwidth Extension Techniques

An amplifier in current-mode logic (CML) provides higher bandwidth compared to a CMOS buffer. However, the bandwidth of the buffer is still limited by the output RC low pass filter and drops with increased load, e.g., large clock distribution. To further extend the bandwidth of an amplifier, there are two common techniques: inductive peaking and distributed structure.

3.4.1 Inductive Peaking

The bandwidth of a CML buffer is limited by the switching speed of the transistor and the output impedance. A large capacitive output load reduces the output impedance at high frequencies and thus limits the output bandwidth. The idea of the inductive peaking is to increase the output impedance at high frequencies by adding inductors to the output load.

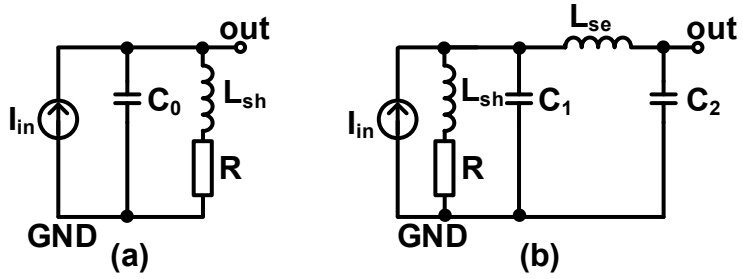


Fig. 46: Equivalent circuit of (a) shunt peaking and (b) shunt-and-series peaking

Shunt peaking

The equivalent circuit of the shunt inductive peaking is shown in Fig. 46 (a). The peaking inductor L_{sh} in series with the load resistor R blocks the current flow at high frequencies, so that most of the current flows into the load capacitor C_0 and the output voltage can make a sharper transition. The output impedance Z_{out} can be calculated by

$$Z_{\mathrm{out}} = \frac{V_{\mathrm{out}}}{I_{\mathrm{in}}} = (\mathrm{j}\omega L_{\mathrm{sh}} + R) \| \frac{1}{\mathrm{j}\omega C_0} = \frac{\mathrm{j}\omega L_{\mathrm{sh}} + R}{(\mathrm{j}\omega)^2 L_{\mathrm{sh}} C_0 + \mathrm{j}\omega R C_0 + 1}. \tag{3.40}$$

Compared to a simple RC network that has only one pole, the peaking inductor adds an additional zero and pole to the output impedance. The zero term (numerator) increases with frequency and thus increases the bandwidth.

It can be seen that the value of R, C_0 and L_{sh} determine the behavior of the impedance Z_{out} over frequencies. Substituting $m = R^2 C_0 / L_{\mathrm{sh}}$ and $\tau = L_{\mathrm{sh}} / R$, Z_{out} can be written as

$$Z_{\mathrm{out}} = \frac{R(\mathrm{j}\omega\tau + 1)}{(\mathrm{j}\omega)^2 \tau^2 m + \mathrm{j}\omega\tau m + 1}. \tag{3.41}$$

Normalizing the impedance with respect to its DC value R, the equation can be written as

$$\frac{|Z_{\text{out}}|}{R} = \sqrt{\frac{(\tau\omega)^2 + 1}{(1 - \omega^2\tau^2 m)^2 + (\omega\tau m)^2}}. \tag{3.42}$$

According to [55], the -3 dB angular frequency ω_0 of the normalized impedance can be written as a function of m, i.e.,

$$\frac{\omega_0}{\omega_1} = \sqrt{(-\frac{m^2}{2} + m + 1) + \sqrt{\left(-\frac{m^2}{2} + m + 1\right)^2 + m^2}}, \tag{3.43}$$

where $\omega_1 = 1/(RC_0)$ is the -3 dB angular frequency without the inductor.

Since different applications have different requirements on the bandwidth, the frequency response and the group delay, the parameter *m* has to be chosen carefully. The bandwidth gain and the peak frequency response regarding *m* for different conditions are summarized in Table 3.1 [55].

Table 3.1 summary of shunt peaking performance regarding m [55]

Conditions	m = R²C/L	Bandwidth	Peak frequency response
Maximum bandwidth	1.41	1.85	1.19
|Z|=R @ ω=1/RC	2	1.8	1.03
Maximum flat frequency response	2.41	1.72	1
Maximum flat group delay	3.1	1.6	1
Without peaking	∞	1	1

In digital data transmission, a fast rising edge, a smaller settling time and a flat group delay are needed to reduce ISI effect. That means a value between the maximally flat frequency response and the best group delay has to be chosen. For clock signal, there is no requirement on ISI, but a longer settling time with ringing is not desirable. Therefore, an optimum between the maximum bandwidth and the maximum flat frequency response has to be chosen.

Shunt-and-Series Peaking

To further increase the bandwidth, a combination of shunt peaking and series peaking as shown in Fig. 46 (b) can be exploited. This peaking technique adds an inductor between the drain capacitor C_1 of the driver and the gate capacitor C_2 of the output amplifier, resulting in a π-network. Similarly, the series inductor L_{se} delays the current flow so that the current first discharges the drain capacitor C_1 and then the gate capacitor C_2. Depending on the ratio between C_1 and C_2, the ratio between the shunt inductance L_{sh} and the series inductance L_{se} can be varied to achieve a maximum bandwidth extension. In [56], with C_1=C_2 and $L_{\text{sh}} = 0.5L_{\text{se}}$, 1.8-dB peak frequency response and 3.5-times bandwidth gain are achieved.

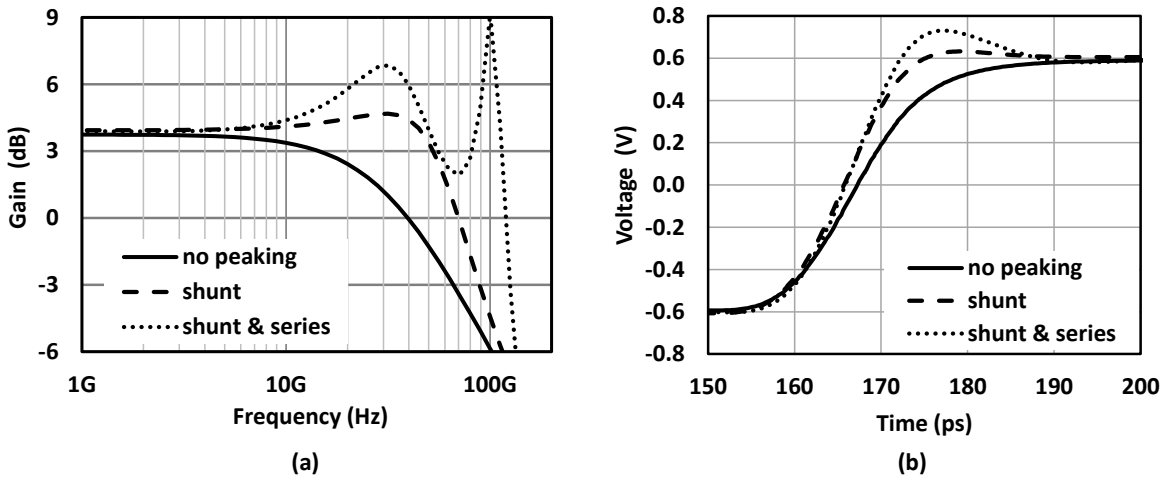


Fig. 47: Comparing CML buffer with shunt peaking, with shunt-and-series peaking and without peaking: (a) frequency response, (b) step response.

A comparison of the frequency response and the step response of a CML buffer with the two peaking techniques is shown in Fig. 47. The schematic simulations are done with a buffer fan-out of one. The small-signal bandwidth gains are 1.8 and 3.3 with shunt peaking and with shunt-and-series peaking, respectively.

3.4.2 Distributed Structure

Distributed structure is another technique to extend the bandwidth of an amplifier, especially if the amplifier has to be large to drive a large output load. The input load of the amplifier also limits the bandwidth of the input signal. The basic idea is to decompose a large amplifier into many small equal units and distribute them along a transmission line. By doing this, the input and output capacitance is absorbed into the transmission line.

3.4.2.1 Artificial transmission line

The key component for a distributed structure is the artificial transmission line, which approximates the behavior of a normal transmission line. The transmission line model is shown in Fig. 48.

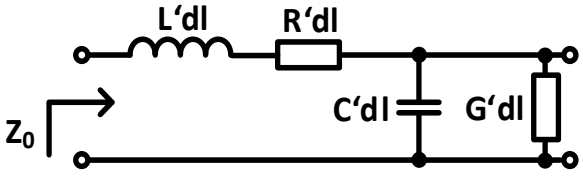


Fig. 48: Lumped circuit model of a transmission line

According to [55], the characteristic impedance of the transmission line can be written as

$$Z_0 = \sqrt{\frac{R' + j\omega L'}{G' + j\omega C'}}. \tag{3.44}$$

The R', L', G' and C' are resistance, inductance, conductance and capacitance per unit length, respectively. If the resistance R' and the conductance G' are negligibly small, i.e., $j\omega L' \gg R'$ and $j\omega C' \gg G'$, the characteristic impedance is a function of only $L'$and C', i.e.,

$$Z_0 = \sqrt{\frac{L'}{C'}}. \tag{3.45}$$

According to [57], the wave velocity on an ideal transmission line is determined by

$$v_{\mathrm{p}} = \frac{1}{\sqrt{L'C'}}. \tag{3.46}$$

Thus, the propagation delay of a transmission line with a length l can be calculated by

$$\tau_{\mathrm{D}} = \frac{l}{v_{\mathrm{p}}} = l\sqrt{L'C'} = \sqrt{LC}\,. \tag{3.47}$$

The artificial transmission line can achieve a similar behavior as the ideal transmission line by approaching the distributed inductance and capacitance with lumped inductors and capacitors. The capacitor of an artificial transmission line includes the parasitic capacitor of the inductor itself and the input or the output capacitor of the amplifiers.

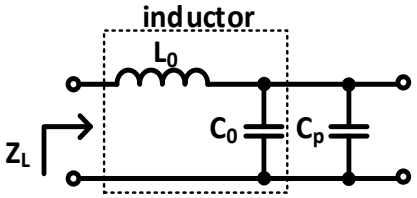


Fig. 49: Circuit model of an artificial transmission line

The circuit model of an artificial transmission line segment is shown in Fig. 49. The characteristic impedance of the artificial transmission line segment with a load (Z_{L}) can be expressed as

$$Z_{\mathrm{L}} = \sqrt{\frac{L_0}{C_0 + C_{\mathrm{p}}}}, \tag{3.48}$$

where L_0 and C_0 are the inductance and the parasitic capacitance of the lumped inductor, respectively; and C_{P} is the input or output capacitance of the amplifier. Similarly, the propagation delay τ_{DL} of the line segment can be determined by

$$\tau_{\mathrm{DL}} = \sqrt{L_0(C_0 + C_{\mathrm{p}})}\,. \tag{3.49}$$

While an ideal transmission line has a constant characteristic impedance over an infinite bandwidth, the characteristic impedance of an artificial transmission line with lumped element is approximately constant only over a limited frequency range. The cut-off frequency f_{c} of an

artificial transmission line, where the impedance is reduced by a factor of $1/\sqrt{2}$, can be determined by [55]

$$f_{\mathrm{c}} = \frac{1}{\pi\sqrt{L_0(C_0 + C_{\mathrm{p}})}}. \tag{3.50}$$

It shows that the impedance of an artificial transmission line degrades at a much lower frequency compared to that of an ideal transmission line. Therefore, to maximize the artificial transmission line performance, the cut-off frequency has to be much higher than the frequency of interest, e.g., the desired amplifier bandwidth. According to (3.48), the parameter of the lumped inductor (L_0 and C_0) can be determined by a given load capacitance C_{p} and a given impedance Z_{L}.

Inductors can be implemented by a finite-length transmission line. In this case, the characteristic impedance Z_0 of the transmission line has to be much higher than the required impedance Z_{L} to compensate the load capacitance. However, an on-chip transmission line exhibits relatively high parasitic capacitance due to the small distance between the top and bottom metal layers. Since the inductance per unit length L' is relatively small, the length has to be very large for a given inductance, hence increasing the resistance and consequently increasing the loss significantly.

Another approach to implementing a lumped inductor is using a spiral inductor. The benefit is that the spiral inductor provides much higher inductance than a transmission line for the same length of metal line. Therefore, the loss on the spiral inductor can be much smaller for the same inductance compared to the finite-length transmission line.

3.4.2.2 Distributed amplifier

The basic structure of a distributed amplifier is shown in Fig. 50. It consists of several equal unit amplifiers. The inputs and outputs of the amplifiers are tapped to the input and output transmission line, respectively. Each transmission line segment has an characteristic impedance of Z_0. The input and output capacitance of each transistor decreases the characteristic impedance of the loaded lines. Assuming that the input and the output of each transistor exhibit the same parasitic capacitance, the input and output transmission lines would have the same characteristic impedance Z_{0L}.

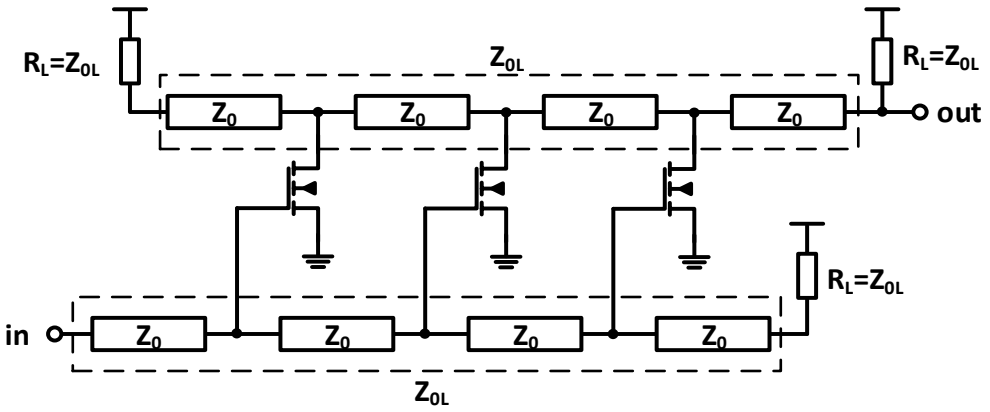


Fig. 50: Basis structure of distributed amplifier

The time-domain behavior of the distributed amplifier is analyzed with a step response. A voltage step propagates along with the input line and appears at each input gate sequentially. The output current I_{out} of each transistor along the transmission line is equal to the product of the input voltage V_{in} and the transconductance g_{m} of the amplifier transistors, i.e.,

$$I_{\mathrm{out}} = g_{\mathrm{m}} V_{\mathrm{in}} \,. \tag{3.51}$$

Since the impedances seen at the output of each transistor are equal for both directions, the current is split in half and creates a voltage step to both direction. The voltage gain of each amplifier is the product of the current gain g_{m} and the output impedance, i.e.,

$$A_{\mathrm{v}} = g_{\mathrm{m}} Z_{0\mathrm{L}}/2 \,. \tag{3.52}$$

The delays of the input and the output line have to be equal to prevent the output signal distortion and to achieve a maximum power at the output node, i.e.,

$$\tau_{\mathrm{D\_in}} = \tau_{\mathrm{D\_out}} \,. \tag{3.53}$$

The output transmission line is terminated with its characteristic impedance to avoid reflections and signal distortions.

3.4.2.3 Performance Limitations of Distributed Amplifier

As pointed out in [58], non-ideal effects in real devices limit the gain and bandwidth of the distributed amplifier.

Resistive loss of the input transmission line attenuates the signal traveling along the input line. The signal amplitude at the far end of the input line will become much smaller if the input line is long or if the resistivity is high. According to (3.51), the output current of each transistor decreases along the output line. For a very large loss, the amplitude of the input signal could be too small to turn on the transistor. Resistive loss in the output line also attenuates the output signal and reduces the whole gain. The finite output resistance of the common source amplifier adds additional loss in the output line and lowers the impedance of the output line as shown in Fig. 51.

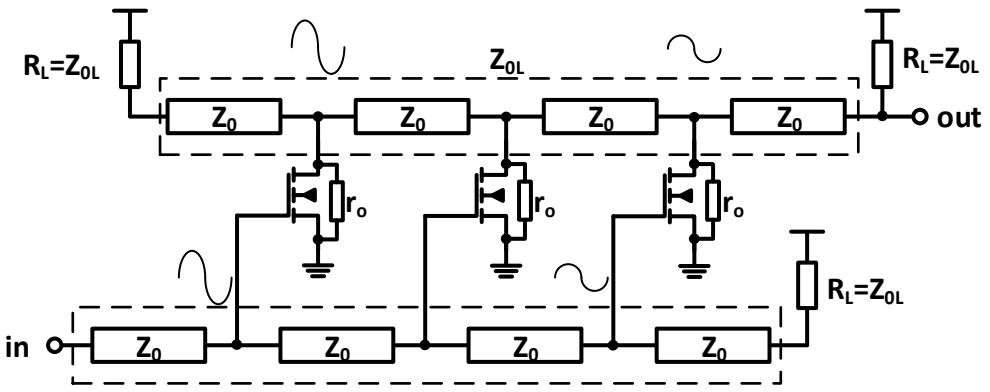


Fig. 51: Finite output resistance in a distributed amplifier

Due to the Miller effect, the input capacitance of each amplifier increases along the input line as the gain increases, as shown in Fig. 52. This effect lowers the characteristic impedance of the input line and increases the delay. Due to different delays on the input and output line, the output signals do not arrive at the output node at exact the same time, further degrading the output power.

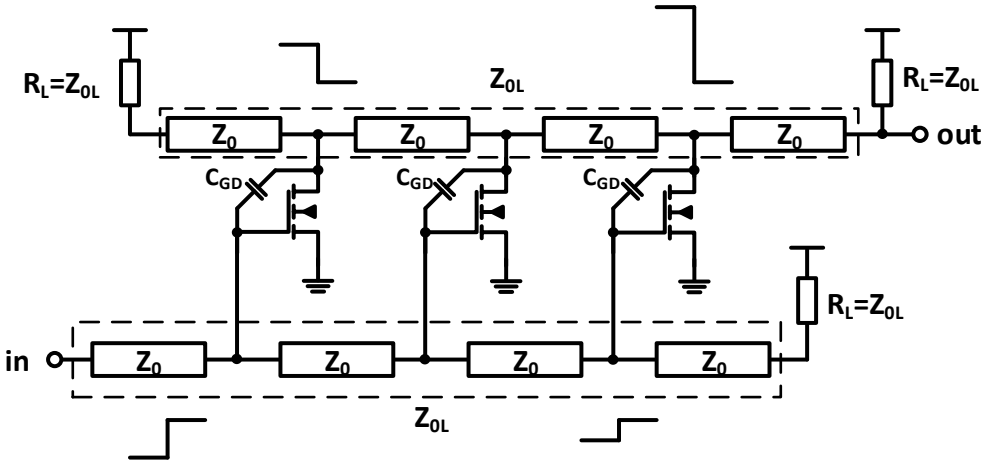


Fig. 52: Non-uniform capacitance due to the Miller effect

3.4.3 Passive Elements Design

3.4.3.1 Microstrip Line

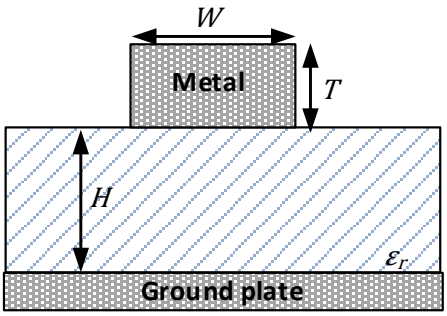


Fig. 53: Cross view of a microstrip line

The Microstrip line shown in Fig. 53 is one of the most widely used on-chip transmission line types. The characteristic impedance and the loss are two key performance metrics of the microstrip line, which mainly depend on its physical structure. For high-speed circuit design, a high characteristic impedance and a low loss are desirable. However, on-chip microstrip lines have some restrictions that limits their performances.

A high characteristic impedance requires a low capacitance of the strip line according to Eq. (3.45), which requires a larger distance (H) between the strip line and the ground plate and a smaller width (W). However, a smaller width increases the metal resistance and leads to higher loss.

Due to the skin effect, the resistance and the loss of the metal line increases with frequency. To reduce the skin effect, thick top metal layers are normally used for the strip line. However, the metal thickness is technology dependent and cannot be modified, which may not able to reduce the skin effect sufficiently.

Because of the lossy substrate in CMOS technology, the ground plate has to be one of the lower metal layers to reduce the substrate loss, which means the distance between the strip line and the ground plate is small, hence exhibiting large capacitance and lowering the characteristic impedance.

The exact value of characteristic impedance can be only determined through electromagnetic (EM) simulations. But an approximate value can be obtained by the following equation according to [57], using the dimensions of the microstrip line, i.e.,

$$Z_0 = \begin{cases} \frac{60}{\sqrt{\varepsilon_r}} \ln\left(\frac{8H}{W} + \frac{W}{4H}\right) & \text{for } W/H \leq 1 \\ \frac{120\pi}{\sqrt{\varepsilon_r}\left[\frac{W}{H} + 1.393 + 0.667\ln\left(\frac{W}{H} + 1.444\right)\right]} & \text{for } W/H \geq 1 \ . \end{cases} \quad \textbf{(3.54)}$$

where ε_r is the relative permittivity. The attenuation in the microstrip line due to the skin effect can be written as

$$\alpha_s = \frac{R_s}{Z_0 W} \ \text{Np/m} = 8.685 \frac{R_s}{Z_0 W} \ \text{dB/m} \ , \quad \textbf{(3.55)}$$

where $R_s = \sqrt{\omega\mu_0/2\sigma}$ is the surface resistivity of the metal. Here, σ is the conductivity, μ_0 is the permeability constant and ω is the angular frequency.

3.4.3.2 On-chip Inductors

Two types of on-chip spiral inductors are used in this work for the artificial transmission line inductor and for the inductive peaking inductor: 2-port single-ended inductor and 3-port symmetric differential inductor with center tap. The layout of these two types of inductors are shown in Fig. 54.

Although inductors are available from the design kit of each technology, they are normally too large and are designed for lower frequencies. Therefore, the models are inaccurate at higher frequencies and are not optimized for high-speed designs. Therefore, custom designed inductors are required. The inductors are implemented with the top two copper metal layers because of its smaller layout size compared to the top aluminum metal layer. The S-parameter characteristics of the inductors are extracted by using EM simulation tools such as ADS momentum. A lumped circuit model is generated for time-domain circuit simulation based on its frequency-domain characteristics. A 2-π circuit model shown in Fig. 55 (a) -similar to the one in [59]- is used to model the inductor. Compared to the single-π circuit model in [60], it is more accurate for modeling the skin effect, the proximity effect and the distributed characteristics at higher frequencies.

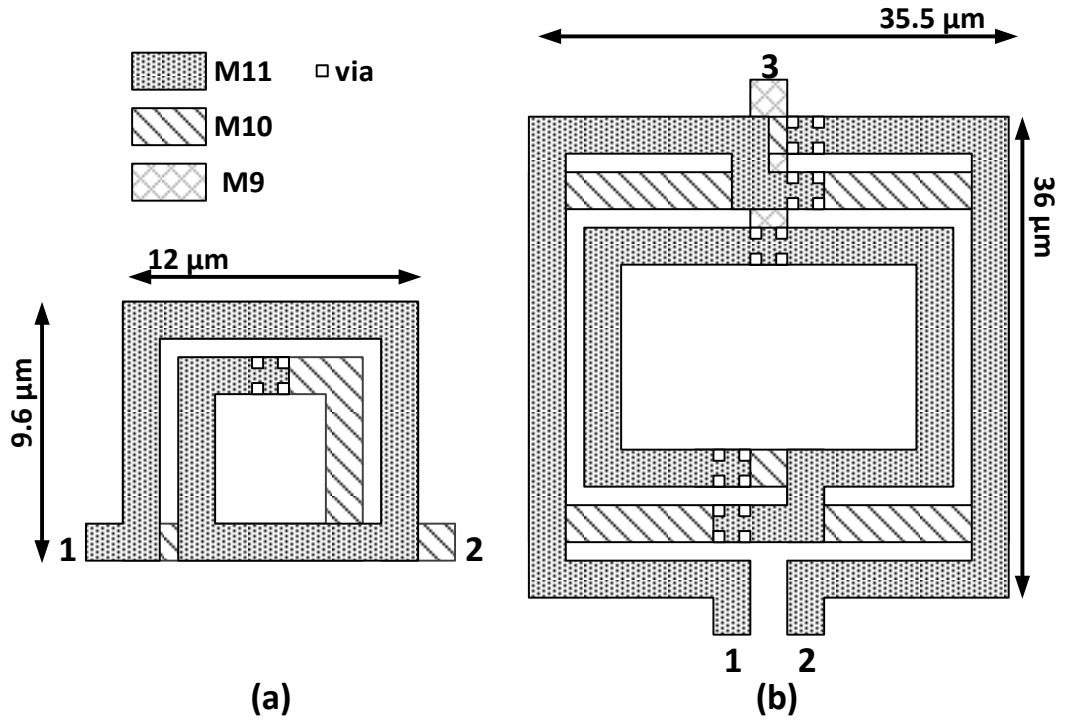


Fig. 54: Layout of on-chip planar spiral inductors: (a) 2-port single-ended inductor, (b) 3-port symmetrical differential inductor with center tap.

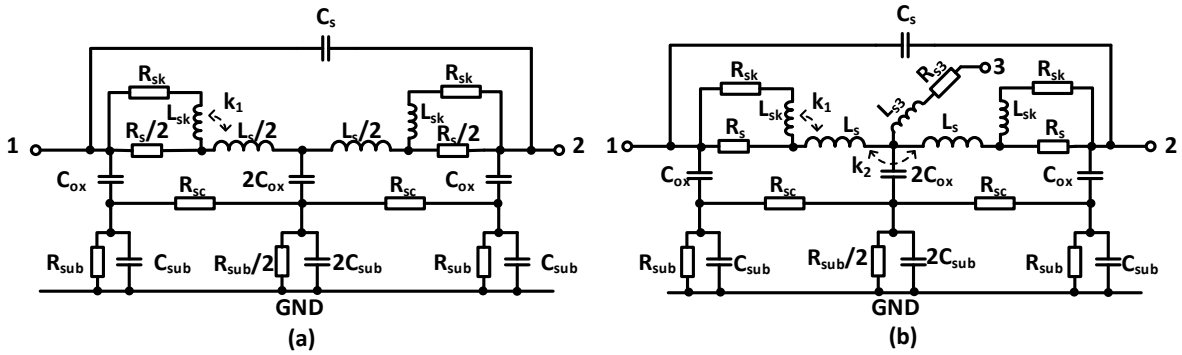


Fig. 55: Broadband lumped model of on-chip planar spiral inductors: (a) the 2-port inductor, (b) the 3-port symmetrical differential inductor with center tap.

The main electrical characteristics of the spiral inductor are modelled with the inductance L_s, the series resistance R_{s}, and the line-to-line coupling capacitance C_{s}. The skin effects and proximity effects are modeled with the series circuit of L_{sk} and R_{sk} and with the mutual coupling k_1 between L_{s} and L_{sk}. The skin effect describes the effect that the current is pushed toward the surface of the conductor at higher frequencies. To describe this effect, an additional R_{sk} and L_{sk} ladder structure is introduced in parallel to R_{s}. The skin effect can be modeled up to very high frequencies with more accuracy by adding more ladder structures according to [61]. To simplify the circuit model, only one stage of the ladder structure is employed in the model. The proximity effect describes the effect that the magnetic field from adjacent metal lines changes the current distribution and causes a higher current density at the edges of the

metal lines. This also increases the resistance at higher frequencies and might have a bigger impact than the skin effect. High-frequency leakage currents through the silicon substrate are modeled by the oxide capacitances C_{ox}, the substrate resistances R_{sub}, and the substrate capacitances C_{sub}. R_{sc} describes the coupling between the lines through the dielectric materials. The 3-port symmetric differential inductor has a similar model to the 2-port inductor and is shown in Fig. 55 (b). The center tap is described with the series circuit of L_{s3} and R_{s3}.

The model parameters of the model are difficult to determine with numeric methods. Therefore, a similar numeric method as in [62] and manual optimization are combined to extract the model parameters, based on the S-parameter results from the EM simulations. Table 3.2 shows the extracted model parameters for the circuit models shown in Fig. 55.

Table 3.2 Extracted model parameter for the 2-port and the 3-port inductor

2-port Inductor								
L_s [pH]	R_s [Ω]	L_{sk} [pH]	R_{sk} [Ω]	C_{ox} [fF]	R_{sub} [Ω]	C_{sub} [fF]	R_{sc} [Ω]	k_1
35	0.82	29	3	15	62k	0.6	160	0.3
3-port differenctial Inductor with centre tap								
L_s [pH]	R_s [Ω]	L_{sk} [pH]	R_{sk} [Ω]	C_{ox} [fF]	R_{sub} [Ω]	C_{sub} [fF]	R_{sc} [Ω]	k_1
122	1.7	700	120	2	1.25k	1.7	400	0.3
L_{s3} [pH]	R_{s3} [Ω]	k_2	C_s [fF]					
122	1.7	0.5	7.6					

To verify the accuracy of the circuit model, the equivalent resistance R, the equivalent inductance L and the quality factor Q of the circuit model and the EM simulation are compared. The equivalent L and R can be extracted from the Y-parameter Y_{11}, i.e.,

$$L = \frac{\mathrm{Im}\left(\frac{1}{Y_{11}}\right)}{\omega}, \tag{3.56}$$

$$R = \mathrm{Re}\left(\frac{1}{Y_{11}}\right). \tag{3.57}$$

The quality factor Q is defined by the ratio of the stored energy to the dissipated total energy, i.e.,

$$Q = -\frac{\mathrm{Im}(Y_{11})}{\mathrm{Re}(Y_{11})}. \tag{3.58}$$

For symmetrical differential inductors, L_{df}, R_{df}, Q_{df} can be determined by [63]

$$L_{\mathrm{df}} = \frac{\mathrm{Im}\left(\frac{1}{Y_{11}-Y_{12}}\right)}{\omega}, \tag{3.59}$$

$$R_{\mathrm{df}} = \mathrm{Re}\left(\frac{1}{Y_{11}-Y_{12}}\right), \tag{3.60}$$

$$Q_{\mathrm{df}} = -\frac{\mathrm{Im}(Y_{11}-Y_{12})}{\mathrm{Re}(Y_{11}-Y_{12})}. \tag{3.61}$$

After manual optimization of the model parameters, the results of the circuit model and the EM simulation are matched very well up to 100 GHz, as shown in Fig. 56.

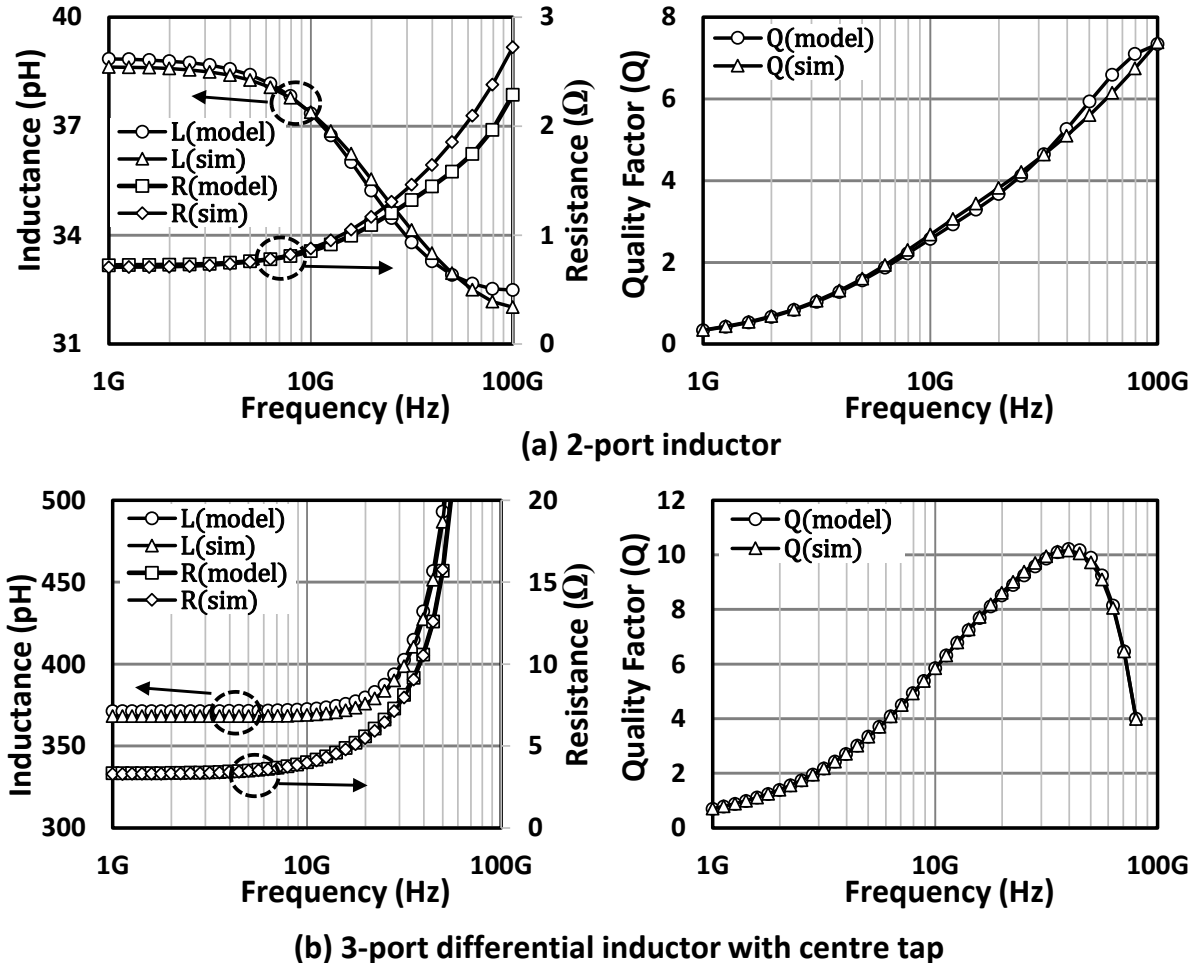


Fig. 56: Comparing the broadband lumped model and the EM simulation of on-chip planar spiral inductors: Inductance, resistance and quality factor of (a) the 2-port inductor and (b) the 3-port symmetrical differential inductor with center tap.

4 Implementation of an 8 bit 100 GS/s Distributed DAC

This chapter presents the circuit implementations of an 8 bit 100 GS/s DAC [64]. The DAC exploits a time-interleaved architecture for an ultra-high sampling rate and a distributed architecture to enable a high output bandwidth. The first section introduces the architecture of the DAC and briefly describes the functionality of the important circuit blocks. The design of the individual circuit blocks – including the distributed DAC output stage, the clock divider and distribution, the phase alignment, the data MUX and the memory block – are demonstrated and analyzed in the following sections. Finally, the layout of the DAC is presented.

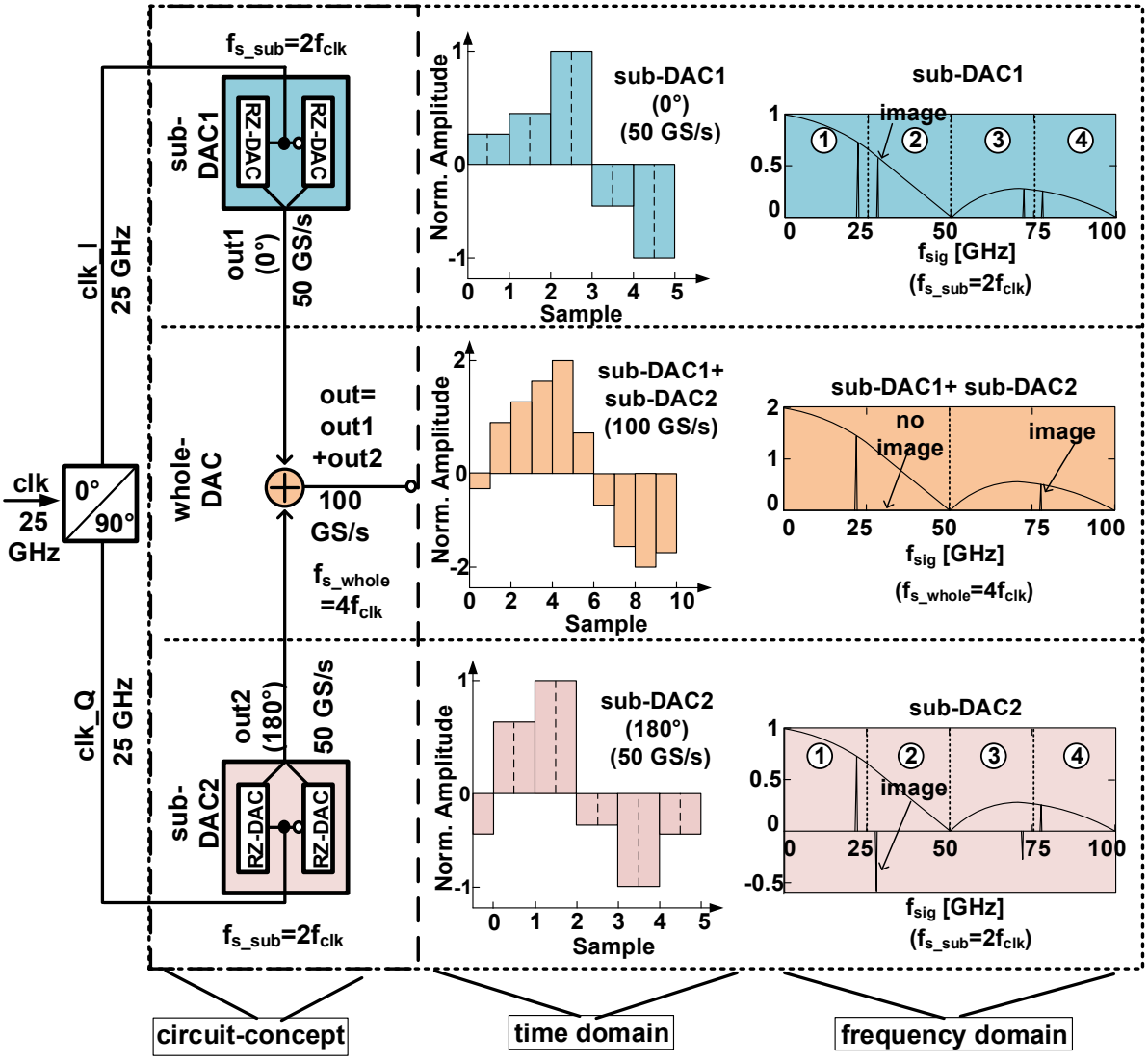


Fig. 57: Concept of the 100 GS/s DAC

4.1 DAC Architecture

Fig. 57 shows the basic circuit concept, the time-domain and the frequency-domain characteristics of the 100 GS/s DAC. The time-interleaving techniques introduced in section 3.3 are employed to implement the time-interleaved output stage.

The DAC chip employs negative supplies. The positive supply voltage VDD is connected to GND. The negative supply voltage of CML circuits (VSSCML) and of (VSSCMOS) are connected to -1.8 V and -1 V, respectively.

The whole DAC can be decomposed into two stages of TI-DACs. In the first stage, two 50 GS/s sub-DACs (sub-DAC1 and sub-DAC2), clocked by 90° phase-shifted 25 GHz clocks, combine their outputs by summing up their output current directly, i.e., parallel-path time interleaving. The analog output waveforms of the two sub-DACs and their summed output waveform are illustrated in the center of the figure. In the frequency domain, the signal frequency component of the two sub-DACs are in-phase in the first Nyquist frequency band, while the images in the second and the third Nyquist band are 180° out-of-phase. By summing the two outputs, the images in those Nyquist bands are canceled out.

In the second stage, each sub-DAC consists of two time-interleaved RZ-DACs clocked with two 25 GHz clocks with 180° phase shift, i.e., analog-multiplexing time interleaving. The sub-DACs convert on the falling and rising clock edge because of the time-interleaving of the two RZ-DACs, hence the sampling frequency is equal to two times the clock frequency, i.e., $f_{\mathrm{s\_sub}} = 2f_{\mathrm{clk}}$.

The 90° phase shift of the two sub-DAC input clocks introduces a skew of half of the sampling period, i.e., $1/(2f_{\mathrm{s\_sub}})$. Therefore, the sum of both outputs generate an overall sampling rate of $f_{\mathrm{s\_whole}} = 2f_{\mathrm{s\_sub}} = 4f_{\mathrm{clk}}$.

Based on the circuit concept in Fig. 57, the DAC chip architecture is illustrated in Fig. 58. The whole DAC consists of two identical sub-DAC blocks. Each sub-DAC block consists of an output stage, a two-stage 2-to-1 MUX array, a pseudo-decoder array, a clock-generation and phase-alignment block and a 512 Byte SRAM memory.

The DAC core, which is the distributed DAC output stage, consists of two sub-DAC output stages. The output currents of the current cells in the output stage are summed at the output artificial transmission line, which has a characteristic impedance of 50 Ω. The left end of the artificial transmission line is terminated with 50 Ω resistors directly on chip to GND; the other end is connected to the output pads and is terminated with external 50 Ω resistors of the following measurement or transmitter device. The external 50 Ω resistors should be terminated to GND for DC measurement. The single-ended and the differential full-scale output swings are 500 mV and 1 V, respectively.

The 8 bit DAC employs a pseudo-segmented structure similar to the one in [52], which uses binary code to directly control the unary-weighted current cells. To be more specific, the 4 LSB binary bits control 4 binary-weighted current cells and the 4 MSB binary bits control 15 unary-weighted current cells. As a result, each RZ-DAC in one sub-DAC consists of 19 current

cells. A pseudo-decoder block is used to perform the 8 bit to 19 bit conversion, which is mainly a rerouting of the 8 bit data input to the 19 bit DAC input.

The 4x19 bit 6.25 Gb/s data stream is serialized to a 19 bit 25 Gb/s data stream by a two stage 2-to-1 MUX arrays. This reduces the operation frequency of the pseudo decoder and the memory. The first stage of the MUX array (8:4 MUX) and the second stage of MUX array (4:2 MUX) are implemented with CMOS logic and CML logic, respectively. The two input data streams fed to the sub-DAC output stage are skewed by half of the sampling period by the second-stage MUX array.

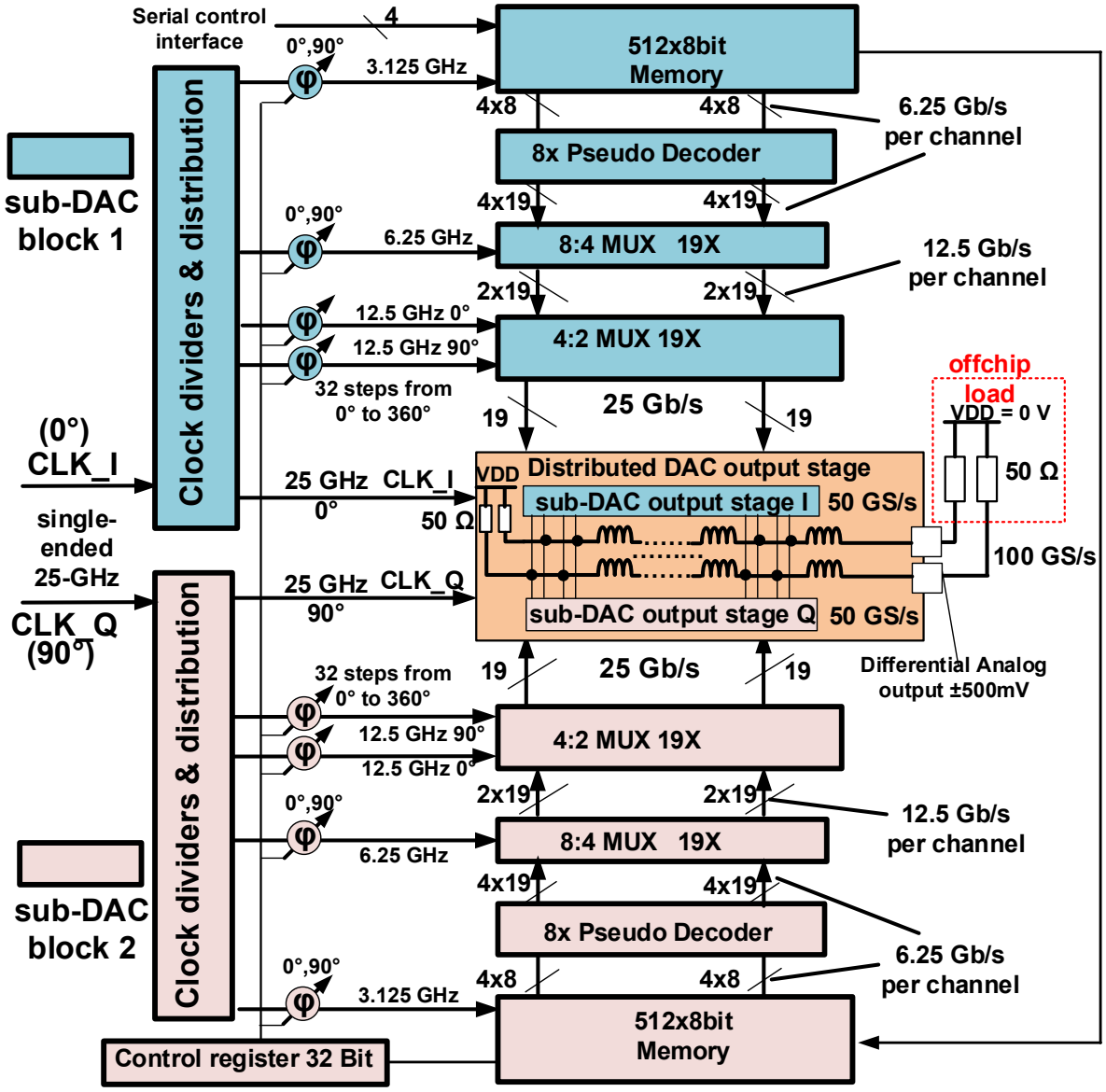


Fig. 58: Chip architecture

The input data for the DAC can be stored in the on-chip memory. This drastically reduces the test complexity compared to other high-speed DACs with a real-time data interface. The two memory blocks in both the sub-DAC blocks can store 1 kByte of data in total, i.e., 1 k symbols for an 8 bit DAC. The memory can be switched between *write* mode and *read* mode by an

external control signal. In the *write* mode, the memory can be written via a serial interface at a low data rate. After the memory is fully written, it can be switched to the *read* mode, in which the data is cyclically read out with a clock derived from the DAC clock and be fed to the DAC output stage. Different signal waveforms, e.g., sine waves, pulse-amplitude-modulated signals, OFDM signals etc., can be easily generated and written into the memory to test the DAC performance up to 100 GS/s.

The multi-phase and -frequency clocks for the individual circuit blocks are derived from two single-ended 25 GHz input clocks. The phase between the data and clock at the inputs of the DAC output stage can be aligned by a phase calibration circuit. The data phase can be shifted with a 5 bit phase rotator. The correct phase skew between data and clock is measured with a phase detector. The phase calibration circuits in both sub-DAC output stages can be tuned individually to ensure optimum phase alignment, with a resolution of 2.5 ps for 25 GHz input clock. The phase calibration circuits with a two-phase selection (0° and 90°) are also implemented for the 8:4 MUX clock and the memory clock to ensure correct timing in the data path. The control of the phase rotators is done by programming the internal control register. With an external evaluation program, an automatic phase-calibration process is established.

4.2 Circuit Blocks

4.2.1 Distributed DAC Output Stage

4.2.1.1 Floorplan

Fig. 59 shows a simple (left) and a detailed block diagram (right) of the DAC output stage. Each RZ-DAC consists of 15 unary-weighted current cells (D18 – D4) and 4 binary-weighted current cells (D3 – D0). The table in Fig. 68 shows the mapping of the current cells with their connections to each binary bit. Bit0 to Bit3 are connected to the binary-weighted current cells D0 to D3. The remaining bits (Bit4, Bit5, Bit6 and Bit7) are connected to 1 (D4), 2 (D5–D6), 4 (D7–D10) and 8 (D1–D18) unary current cells, respectively. Since the decoding is done before the data multiplexer stage as shown in Fig. 58, all current cells have identical input drivers. Therefore, the loads of the input drivers are all matched.

The current cells of each sub-DAC are placed horizontally in one line. The two sub-DAC output stages are identical and are placed symmetrically to the x-axis. The 4 LSB current cells are placed in the middle and the unary current cells controlled by the same binary bit are placed on both sides symmetrically to the middle. The two RZ-DACs in one sub-DAC output stage are placed in an interdigitated way, i.e., the current cells from each RZ-DAC with the same weight are placed next to each other to improve the matching. Each sub-DAC consists of 38 current cells and its total length is about 360 µm, which is dominated by the artificial transmission line

The output currents are summed on the output artificial transmission lines, which are matched to 50 Ω. The outputs of 4 unary-weighted current cells, each from one of the 4 RZ-DACs, are tapped together to one transmission line segment. The outputs of all binary-weighted current cells are tied together and tapped to one transmission line segment. The left end of the output

lines are terminated with on-chip 50 Ω resistors. The right end of the output line is the actual output and is terminated off-chip with 50 Ω resistors.

The clock driver drives the current cell inputs through the artificial clock transmission line. The clock lines are double terminated at both ends. The delay on the clock line and the output line are matched for correct signal reconstruction at the analog output pads.

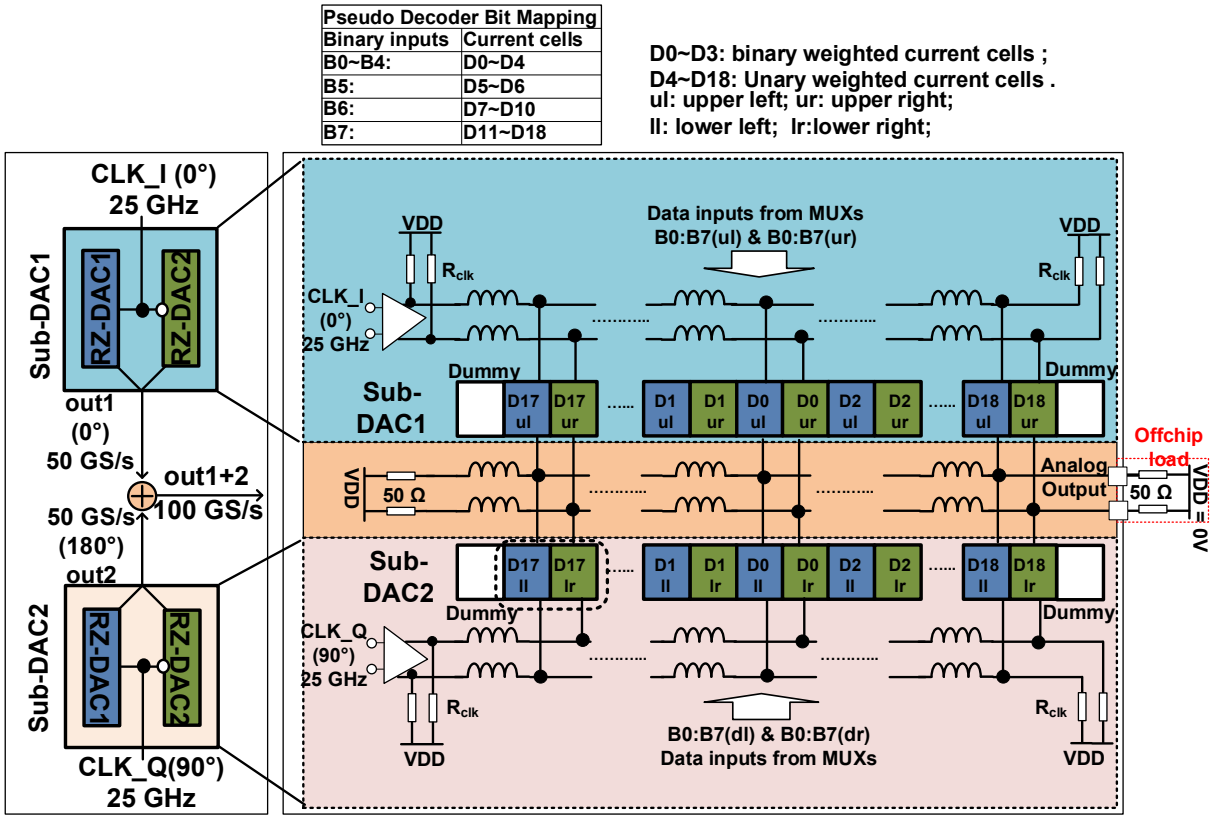


Fig. 59: Block diagram of the distributed DAC output stage

4.2.1.2 Current cell

The current cells employ the same structure as the one described in Section 0. They consist of a cascode current source, a data switch and a clock switch as shown in Fig. 60.

The unary-weighted and binary-weighted currents are generated by cascode current sources. The data switch routes the current to either the positive output branch (*outp, dummyp*) or to the negative output branch (*outn, dummyn*), depending on the input data applied. The clock switch routes the output current from the data switch to the analog outputs (*outp, outn*) when the clock is on a high level and redirects the current to the dummy outputs (*dummyp, dummyn*) when the clock is on a low level. During half of the clock period, where the clock is on a low level, the current from the current cell to the actual output is zero.

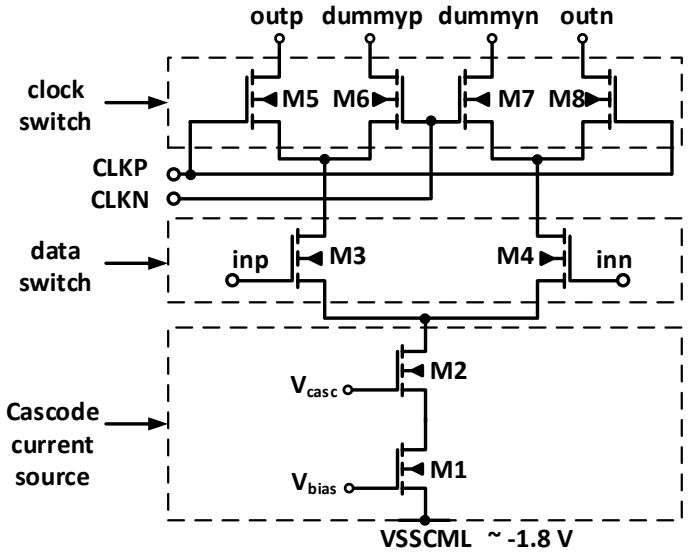


Fig. 60: The current cell

4.2.1.3 Current source

For the current source design in the DAC output stage, the following limitations have to be taken into account:

First, the output current is defined by the maximum output voltage and the output load. For a 500 mV single-ended full-scale output and a 25 Ω load, the total current of the output stage is 20 mA. For an 8 bit DAC, the LSB output current is 78.125 µA. Considering the output to be the sum of two sub-DACs, the LSB current of the sub-DAC is halved, i.e., 39 µA.

Second, the output resistance has to be sufficiently high to ensure high linearity. For a single-transistor current source, as shown in Fig. 61 (a), higher output resistance can be achieved by lowering the transistor W/L ratio. According to [41], the output resistance r_{out} of a transistor is given by

$$r_{\mathrm{out}} = \frac{1}{I_{\mathrm{D}}}\left(\frac{1}{\lambda} + V_{\mathrm{DS}}\right), \tag{4.1}$$

where I_{D} is the drain current; V_{DS} is the drain-source voltage; and λ is the channel-length modulation parameter and is proportional to the channel length. Stacking the transistors can effectively increase the channel length and thus the output resistance. For the cascode current source shown in Fig. 61 (b), the output resistance is given by [41]

$$r_{\mathrm{out}} = g_{\mathrm{m2}} r_{\mathrm{o2}} r_{\mathrm{o1}}\,, \tag{4.2}$$

where g_{m2} and r_{o2} are the transconductance and the output resistance of the cascode transistor *M2*, respectively; r_{o1} is the output resistance of the bias transistor *M1*. The output resistance of the bias transistor is thus increased by a factor of $g_{\mathrm{m2}} r_{\mathrm{o2}}$ in a cascade configuration. For the same output resistance, r_{o1} can be much smaller in a cascade structure, thus reduce the channel

length and the size of the bias transistor *M1*. This reduces parasitic capacitances and improves the output impedance at high frequencies.

The output resistance of a cascode current source can be further increased by the gain-boost technique. By adding an amplifier with an open-loop gain A_{OL}, as shown in Fig. 61 (c), the output resistance of the cascode current source can be increased by a factor of A_{OL} , i.e.,

$$r_{\mathrm{out}} = A_{\mathrm{OL}}(g_{\mathrm{m2}} r_{\mathrm{o2}} r_{\mathrm{o1}}) \tag{4.3}$$

Thus, a further reduction of the transistor size and the parasitic capacitance can be achieved.

The simulated output resistances of the three current sources are shown in Fig. 62. ALL transistors have the same size, i.e., W/L = 2.4 µm / 280 nm, and a maximum operation voltage of 1.8 V. The output current I_{out} is biased to 78.125 µA. The output resistance is the low frequency output impedance resulted from AC simulation.

To achieve a *HD*3 lower than -50 dB for 8 bit resolution, the output resistance has to be larger than 57 kΩ according to (3.14). Therefore, the cascode current source is chosen.

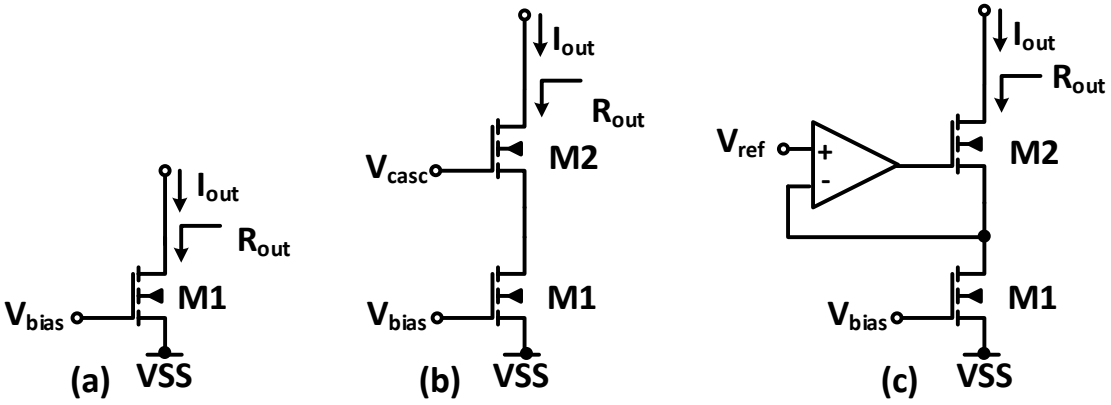


Fig. 61: (a) Single-transistor current source (b) cascode current source (c) cascode current source with gain boost

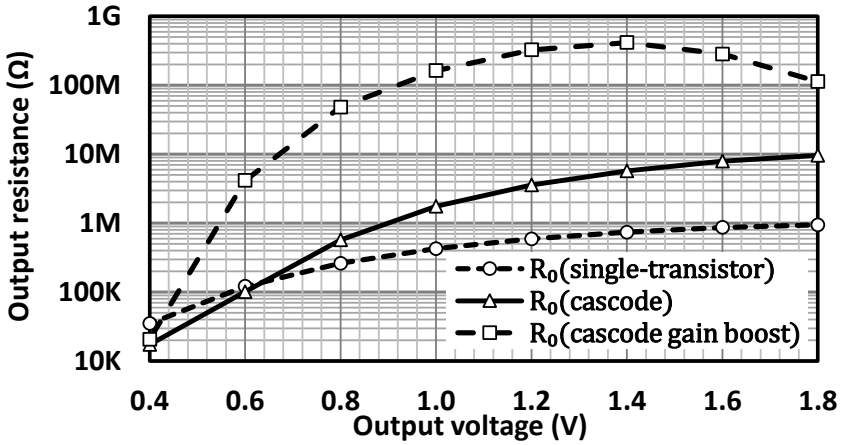


Fig. 62: Comparison of the output resistances of the three current sources

Third, the mismatch between the current sources has to be sufficiently low. Increasing the transistor size reduces the mismatch but also increases parasitics and reduces the high-frequency impedance. The use of calibration techniques reduces the transistor size while maintaining sufficient matching performance, but also adds more design complexity. In this work, intrinsic matching, together with advanced layout techniques, are applied to meet the matching requirement. The matching performance is verified by Monte-Carlo-(MC) simulations. The overdrive voltage of the bias transistor (V_{GS} - V_{th}) is around 300 mV, and V_{th} of the 1.8 V transistor is about 500 mV. The standard deviation of the LSB output current $\sigma(I_{\mathrm{LSB}})$ ($\frac{\sigma(I_{\mathrm{LSB}})}{I_{\mathrm{LSB}}}$) is about 1.9 µA at $I_{\mathrm{LSB}} = 78.125$ µA in a 100 run MC-simulation at the typical corner. The unary current I_{unary} is 16 times the LSB current (i.e., $I_{\mathrm{unary}} = 16 * I_{\mathrm{LSB}}$), and the standard deviation of the unary output current is determined by $\sigma(I_{\mathrm{unary}}) = \sqrt{16} * \sigma(I_{\mathrm{LSB}})$, which is equal to 7.6 µA. The DNL in a segmented DAC is determined by $\sigma(I_{\mathrm{unary}})$. The SNR introduced by the DNL can be determined by

$$SNR_{DNL} = 20\log\frac{(2^n - 1)I_{\mathrm{LSB}}}{3\sigma(I_{\mathrm{unary}})} \tag{4.4}$$

where n is the number of bits of the DAC and 3 sigma variation is presumed.

With the simulation results above, SNR_{DNL} is equal to 59 dB, which is sufficient for 8 bit DAC.

4.2.1.4 Bias voltage generation

The bias reference voltage of a single-transistor current source can be generated by a reference current source and a transistor in a diode connection as shown in Fig. 63 (a). The bias voltages of a cascode current source can be generated by the circuit shown in Fig. 63 (b). This configuration lowers the required headroom up to a threshold voltage (V_{th1}) of the bias transistor *M1*, compared to that of a regular cascode bias generator by cascoding two diode-connected transistors. To keep both bias transistor *M1* and cascode transistor *M2* in the saturation region, i.e., $V_{\mathrm{DS}} > V_{\mathrm{GS}} - V_{\mathrm{th}}$, *M1* and *M2* have to fulfil the following conditions:

$$V_{\mathrm{DS2}} < V_{\mathrm{th1}} \text{ and } V_{\mathrm{GS2}} < V_{\mathrm{th1}} + V_{\mathrm{th2}}\,. \tag{4.5}$$

The actual circuit to generate the reference current is shown in Fig. 63 (c). The goal is to provide a current for a 500 mV single-ended full-scale voltage at the DAC output stage. For a given reference voltage, the reference current is adjusted automatically according to the process variation of the resistor. The - 500 mV reference voltage V_{ref} is generated with an on-chip voltage divider and can be tuned by an external reference voltage. A CML differential amplifier with an active current mirror load (*M5, M6*) acts as an error amplifier. The output of the error amplifier is connected to the gates of *M7* and *M8*. They generate the reference current for the cascode bias circuit. The bias voltages are connected to a replication circuit of the DAC output current cell. The output current of *M8* is mirrored to the cascode current source. Together, they build the second stage amplifier with a voltage gain of $g_{\mathrm{m8}}R_2$, assuming that *M11* and *M12* have the same size. The voltage over R_2 provides the feedback V_{FB} for the error amplifier. V_{FB} tracks

V_{ref} via the control loop. The compensation capacitor C_{comp} is used for the loop stability. The replication of the cascode reference generator provides the bias voltages to the DAC current sources. It decouples the load from the control loop, thus increasing the phase margin and reducing the noise injection back to the control loop.

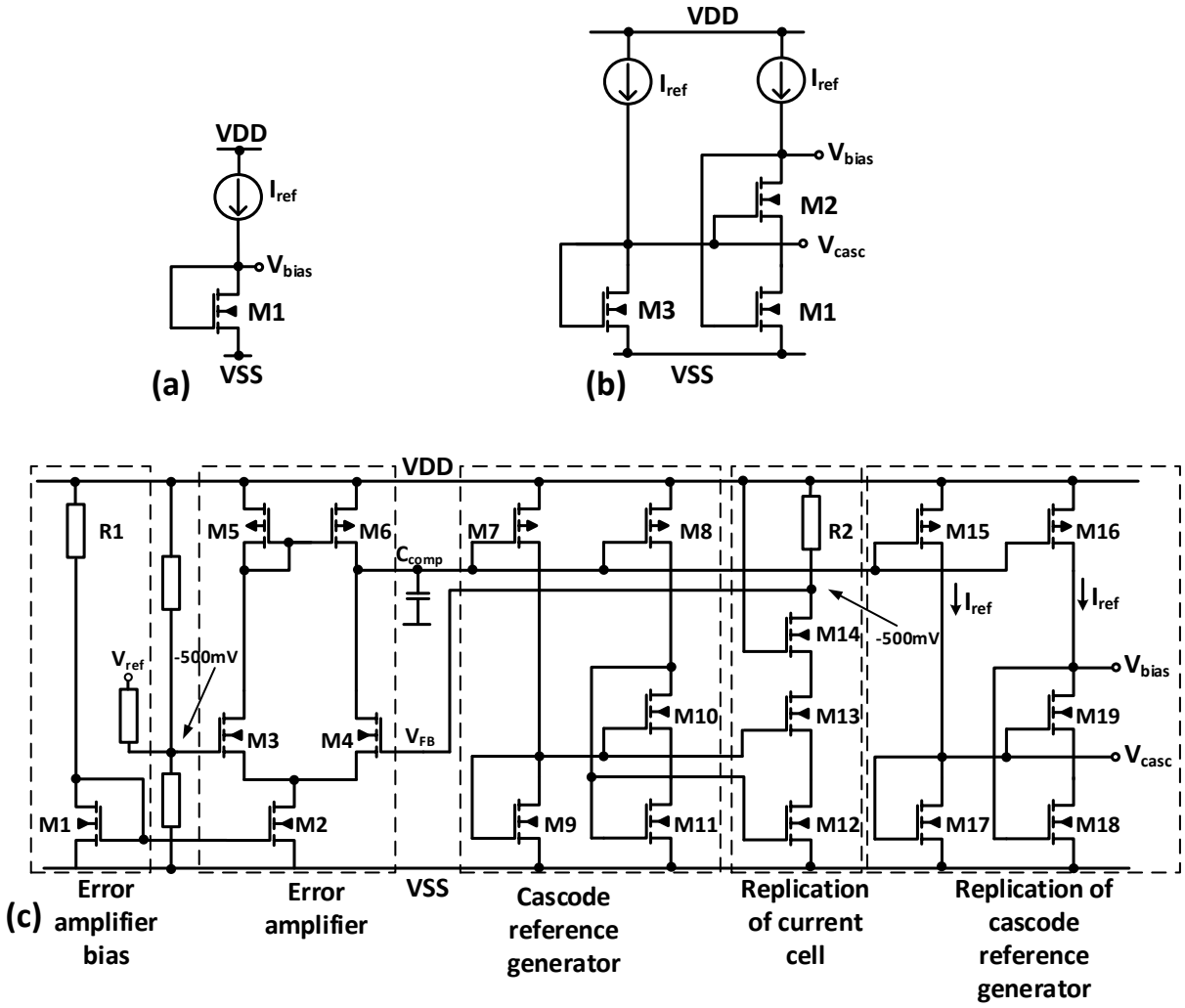


Fig. 63: (a) Bias voltage generator of the single-transistor current source (b) bias voltage generator of the cascode current source (c) bias voltage generator of the cascode current source with external adjustable reference voltage

4.2.1.5 Current switches

Similar to the current source, the dimension of the switch transistors has to be carefully designed with regard to the full-scale output swing, the frequency response and the mismatch performance.

First, the maximum output swing of 500 mV and the 25 Ω load results in a total current of 20 mA at the DAC output stage. Considering the maximum operating voltage of 1.1 V for the super-low-threshold-voltage (slvt) transistors, the drain-source voltage V_{DS} and the gate-source voltage V_{GS} should be lower than 1.1 V. For a given input swing, the W/L ratio of the switch

transistors has to be sufficiently large to provide enough current gain, to be able to switch the tail current completely.

Second, the mismatch between the differential pair generates an offset at the input, which leads to duty-cycle errors, generating distortion at the DAC output. To minimize the mismatch, one can increase the transistor size with the same W/L ratio. This increase the parasitic capacitance at both input and output. Larger input driver is therefore needed to maintain the steepness of the input signal slope. The output bandwidth also decreases due to increased output capacitance.

Third, it is desirable to increase the bandwidth for high-speed design. The bandwidth of a differential pair is limited by the output resistance and the parasitic capacitance. Therefore, using the smallest possible transistor is beneficial for the bandwidth.

The voltage swing of the current switch inputs is around 600 mV by using CML buffers as input driver. To keep the switch transistor in the saturation region at the switching point, the voltage level of the data switch inputs are shifted down by 300 mV, as shown in Fig. 64. The common source node voltage waveform of the clock switches with and without level-shifting the data input is shown in Fig. 65. The real output and the dummy output are set to a constant voltage of -500 mV. It can be seen that the rising edge of the common source node exhibits a similar rise time in both cases while the falling edge exhibits a much shorter fall time with level-shifting. It shows that the level shift helps to improve the off-switching of the data transistor. The lower two curves show that the common source node of the data switches are shifted with the input voltage by 100 mV. The disturbance on the common source node of the data switches is much larger with level-shifting than without level-shifting, since the crossing point is effectively shifted down by 200 mV for the data switches.

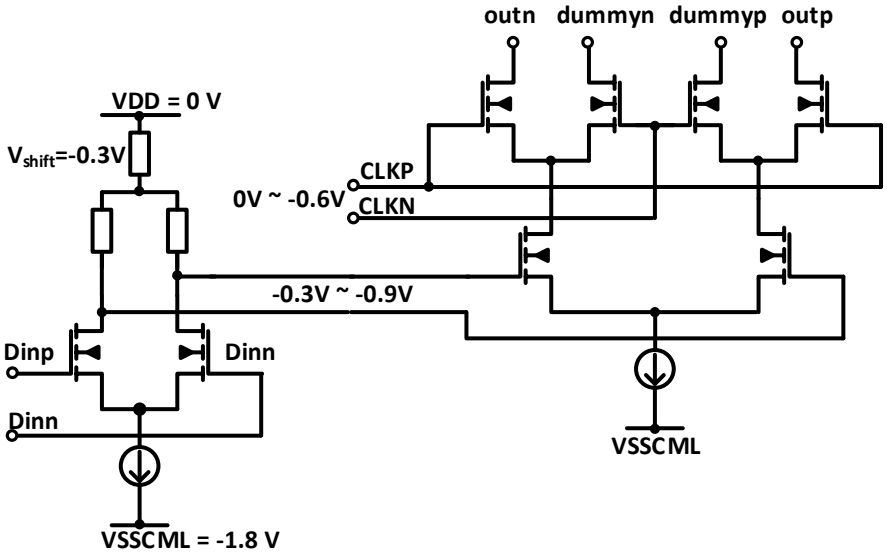


Fig. 64: Current cell with level-shifted input driver

Although shifting down the data inputs increases the switching speed, it also increases the headroom and, thus, the power consumption of both the input buffer and the DAC output stage. The trade-off between speed and power has to be carefully balanced. To lower the power consumption by lowering the supply voltages, a sufficiently large overdrive voltage has to be

applied to reduce the drain-source voltage of the data switch transistors. For the same current gain g_m, larger overdrive voltages reduce the W/L ratio, and hence increasing the bandwidth. In this case, a CMOS input driver with full rail-to-rail CMOS voltage swing might be beneficial.

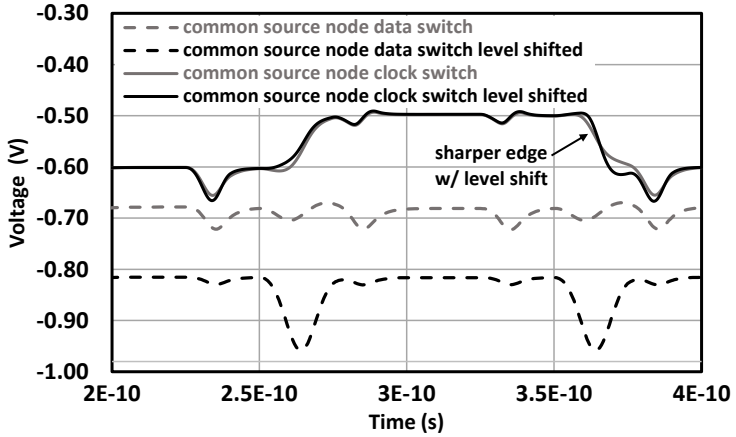


Fig. 65: Common source node voltage of the data switches and the clock switches with and without level shifting the data input

4.2.1.6 Artificial transmission line

The DAC output bandwidth is limited by the sinc frequency response, due to ZOH output. For NRZ DACs, the sinc amplitude roll-off is about 3.9 dB at Nyquist frequency. For low sampling rates, the sinc frequency response is normally the major limitation of the bandwidth. As the sampling rate increases, the output parasitics become the major limitation. In this proposed DAC with parallel-path time-interleaving architecture, the theoretical maximum bandwidth is around 23 GHz at 100 GS/s, limited by the sinc-function envelope, because the pulse width of the sub-DACs is unchanged at the output and therefore there is no bandwidth gain as shown in Fig. 57. The 4-fold time-interleaved structure (4x RZ-DACs) introduces large parasitics at the DAC output. The estimated total output capacitance is close to 300 fF, including the capacitance from the output transistors, the metal connection and the output pad. This limits the analog bandwidth to below 23 GHz with 25 Ω output resistance. Therefore, the DAC output stage is constructed in a distributed style to distribute the output capacitance and increase the bandwidth.

The principle of the artificial transmission line has been introduced in Section 3.4.2.2. Fig. 66 (a) shows the simplified schematic of one segment of the distributed output stage in a single-ended form. Two clock lines are connected to the gate of the clock switches at the top and bottom. The parasitic capacitance at the current cell output C_{pout} and at the clock switch gate C_{pclk} determine the inductance of the output line and the inductance of the clock line, respectively. The output capacitance of the individual current cells varies with the state of the switch transistor and the output voltage. Therefore, the averaged value of the capacitance is used for the inductance calculation. The equivalent circuit of the output line is shown in Fig. 66 (b).

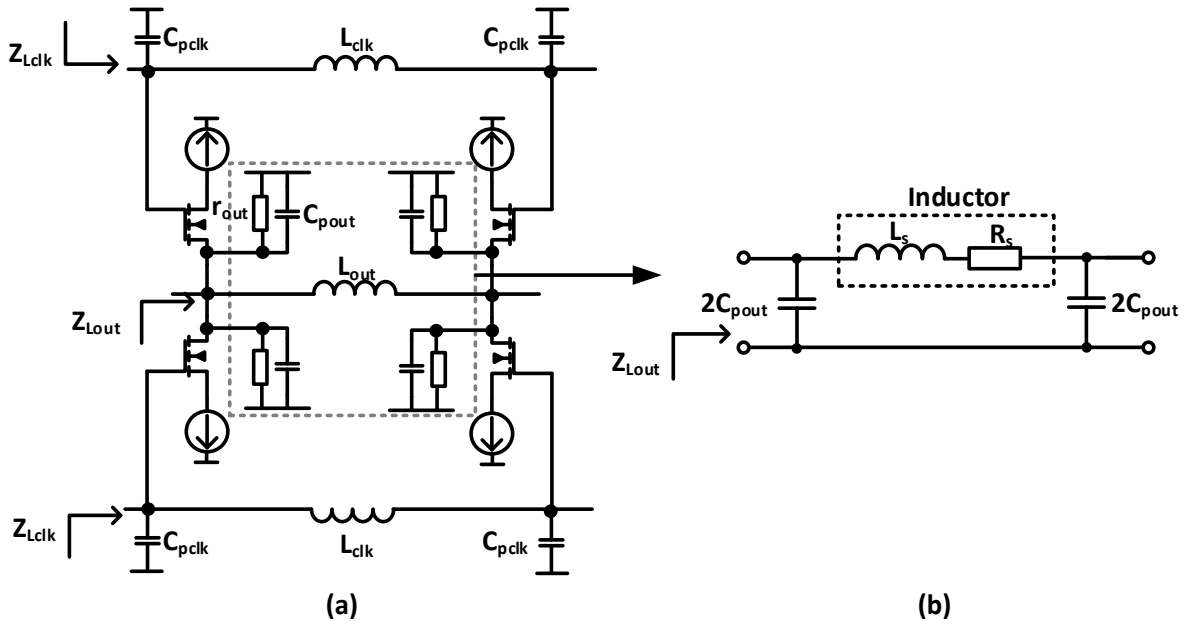


Fig. 66: (a) Simplified schematic of one output stage segment in single-ended form, (b) equivalent circuit of the output transmission line segment

The transmission line segments are implemented with on-chip spiral inductors. Compared to a transmission line, the inductance of a spiral inductor is much higher for the same wire length. This reduces the resistance and the loss on the line segment for the same inductance.

The impedance of the output line has to be matched to 50 Ω. The delay time on the output line has to be matched to the delay on the clock line. With (3.48) and (3.49), the inductance of the output line segment L_{out} and the clock line segment L_{clk} can be determined.

The circuit model of the inductor in the output line has already been described in Section 3.4.3.2 and shown in Fig. 55. A spiral inductor as shown in Fig. 54 (a) is used to approximate the pre-calculated inductance. The S-parameter is simulated with the ADS Momentum EM simulator and the model parameters are extracted from the EM simulation results for the time-domain simulations.

For both the output line and the clock line, the parasitic capacitance, the calculated inductances, the simulated inductances and the resistance are shown in Table 4.1. It can be seen that the variations of the inductance and the resistance over the frequencies are not negligible. This leads to a strong frequency dependency of the characteristic impedance and the delay matching between the output and the clock line.

As discussed in Section 3.4.2.3, the series resistance causes signal attenuation along the artificial transmission line towards the output direction. This effect causes gain degradation along the transmission line toward the output. The Miller-effect on the gate-drain capacitor also increases the capacitance of the clock line toward the output, which has a negative impact on the delay matching between output line and the clock line. These effects worsen the bandwidth and the ENOB of the DAC.

Table 4.1: Parameters of the equivalent circuit model of the output and the clock transmission lines

Output line		Clock line	
C_{pout_total}	11.8 fF	C_{pclk_total}	13.0 fF
$L_{s,out,calc}$	29.5 pH	$L_{s,clk,calc}$	26.7 pH
$L_{s,out,sim\ @\ DC}$	38 pH	$L_{s,clk,sim\ @\ DC}$	35 pH
$R_{s,out,sim\ @\ DC}$	0.8 Ω	$R_{s,clk,sim\ @\ DC}$	0.75 Ω
$L_{s,out,sim\ @\ 25\ GHz}$	35 pH	$L_{s,clk,sim\ @\ 25\ GHz}$	32 pH
$R_{s,out,sim\ @\ 25\ GHz}$	1.24 Ω	$R_{s,clk,sim\ @\ 25\ GHz}$	1.20 Ω

With proper calibration technique, the current source can be tuned to compensate the gain loss of the transmission line. However, the code dependent output frequency causes a code dependent delay on the output line. Because of the skin effect, the series resistance is also frequency dependent. All these effects distort the output signal and cannot be calibrated.

Because of the series resistance on the output line, the current flowing to the output increases along the transmission line towards the output. This effect is illustrated in Fig. 67.

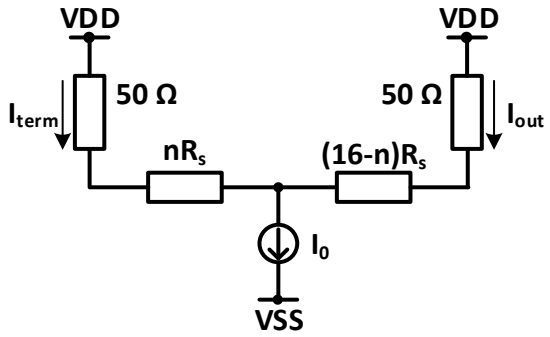


Fig. 67: Effect of the resistance on the output line

The total current connected to one line segment is the unary current or the summed binary currents. Assuming one unary current cell with a current I_0 is connected to the nth segment of the total 16 segments of the output line, the correspondent current at the DAC output $I_{\text{out\_n}}$ can be calculated with

$$I_{\text{out\_n}} = I_0 \frac{50 + nR_s}{100 + 16R_s}. \tag{4.6}$$

where R_s is the parasitic resistance of one output line segment. The ideal output current without resistance should be $I_0/2$. Therefore, the current error $I_{\text{err\_n}}$ is

$$I_{\text{err\_n}} \approx \frac{I_0}{2} \frac{2(n-8)R_s}{100 + 16R_s}. \tag{4.7}$$

The maximum error occurs if the current cell is located at the left or right end, i.e., n = 0 or 16. With $R_s = 1.24\ \Omega$, the maximum current error $I_{\mathrm{err\_max}}$ is equal to $0.164\frac{I_0}{2}$.

Since the unary current is 16 times the LSB unit current, i.e., $I_0 = 16I_{\mathrm{LSB}}$, the maximum error is around 2.6 LSB, which causes large DNL and INL if the unary-weighted current cell are controlled by thermometer code.

One way to reduce such error is to distribute the unary-weighted current cells symmetric to the center and to control them directly with binary code as shown in Fig. 68. For bit 7 to bit 5, the same number of unary current cells are placed symmetrically to the center position. Bit 4 and the 4 LSB binary-weighted current cells are placed in the center.

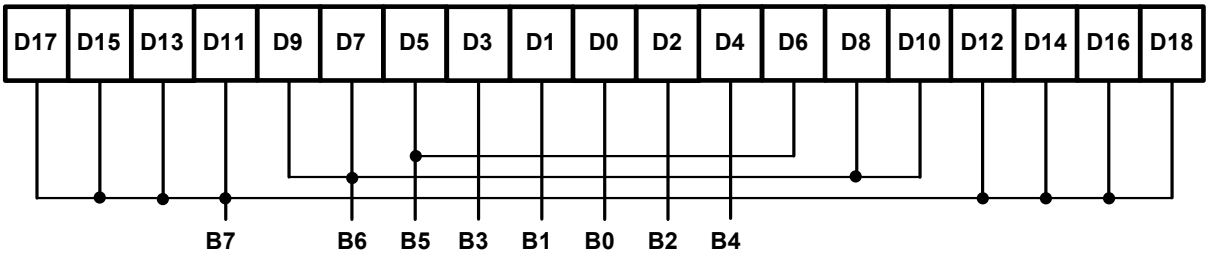


Fig. 68: Symmetric binary-coded structure for reducing error

Although the binary-coded structure naturally introduces more DNL errors than the segmented structure due to current source mismatch, it can drastically reduce the current errors introduced by the output line resistance. The maximum current error $I_{\mathrm{err\_max}}$ of the symmetric binary-coded structure is around $0.01\frac{I_0}{2}$, i.e., 0.16 LSB. The simulated DNL and INL of both structures with ideal current sources are shown in Fig. 69 (a) and (b), respectively. It shows that the DNL and INL can be improved from 2.5 LSB to below 0.2 LSB with the symmetrical binary-coded structure.

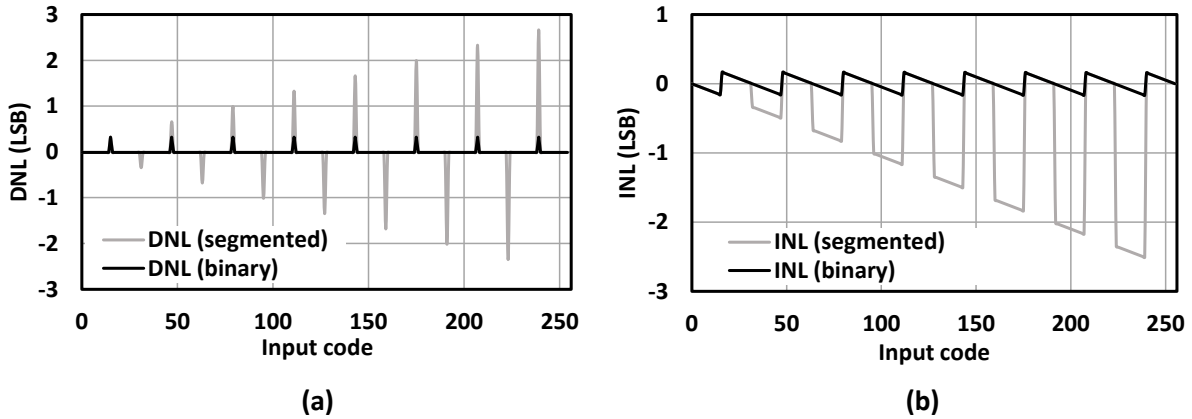


Fig. 69: Comparison of (a) DNL and (b) INL between segmented and binary-coded structure.

4.2.1.7 Bandwidth and resolution

In Fig. 70, the simulated frequency responses of the distributed and the non-distributed DAC are compared. The bandwidth is increased from 15 GHz to 23 GHz with a distributed structure.

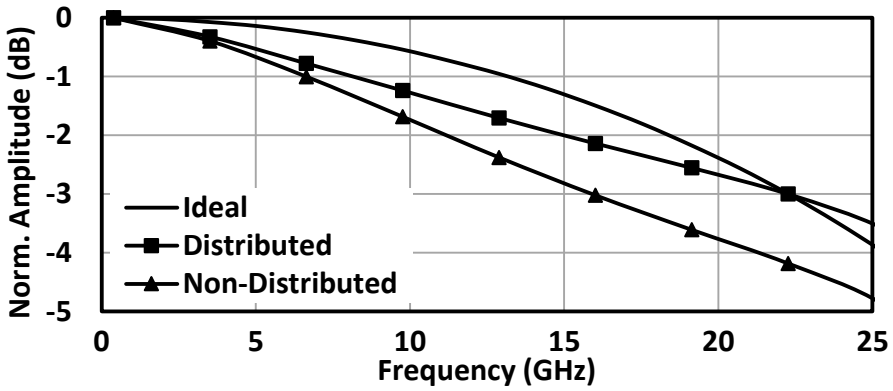


Fig. 70: Frequency response of the distributed and the non-distributed DACs [65]

However, these simulation results are too optimistic due to underestimation of the parasitic capacitance tapped to the transmission lines. The artificial transmission line is designed and optimized based on the pre-estimated parasitics of the current cells. Some parasitics (e.g., parasitic capacitance of the metal connections) are not included. The additional parasitics cause several problems and lower the output bandwidth significantly.

First, the impedance of the artificial transmission line drops much faster with increased frequency due to the additional parasitic capacitance. As shown in Fig. 71, with increased load capacitance, the transmission (S21) is reduced and the reflection (S22) is increased, hence reducing the signal amplitude. The corresponding impulse responses of the output transmission line with different load capacitances are illustrated in Fig. 72. The reduction of the signal amplitude on the clock transmission lines further degrades the output amplitude. Fig. 73 shows the frequency responses of the DAC with increased load capacitance on the output transmission line. It can be seen that the output amplitude is reduced by 1.4 dB at about 19 GHz, by increasing the load capacitance from 5 fF to 15 fF. The bandwidth is reduced from 19 GHz to below 15 GHz, correspondingly.

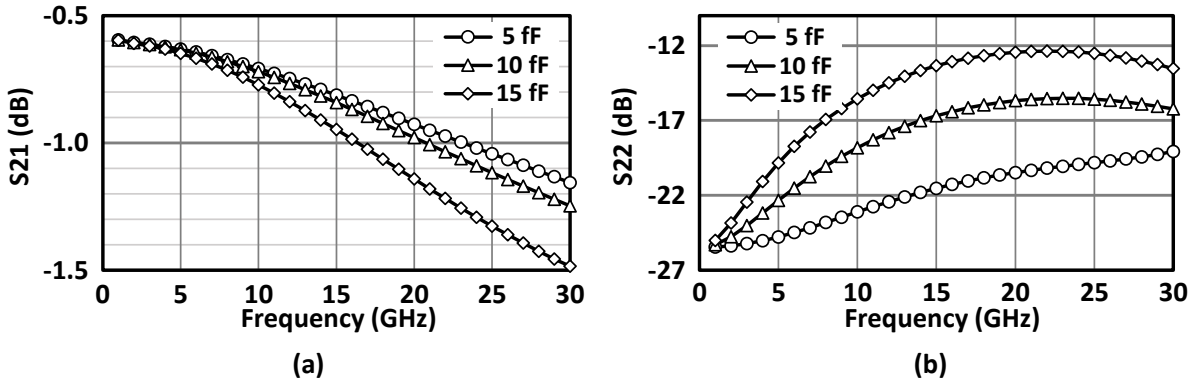


Fig. 71: (a) Transmission (S21) and (b) reflection (S22) of the output transmission line (from center tap to the output node) with different output load capacitances

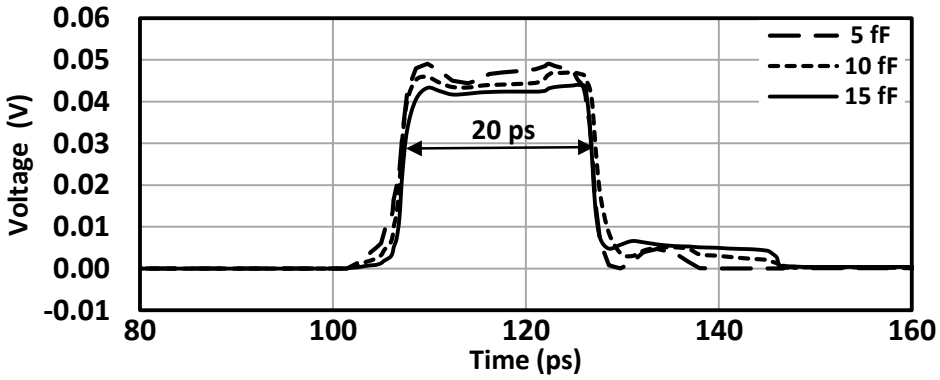


Fig. 72: Impulse response of the output transmission line from center tap to the output node with different output load capacitances

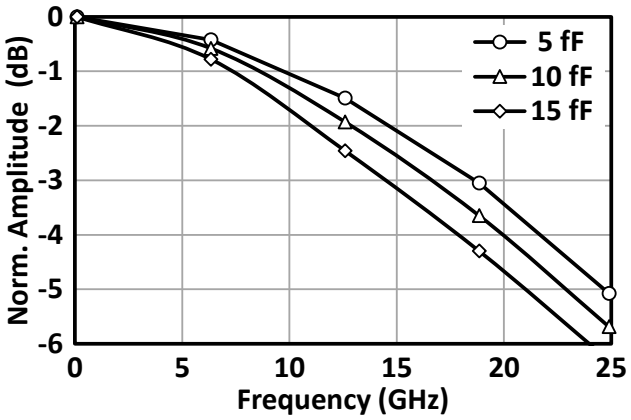


Fig. 73: Output frequency response with different load capacitances on the output transmission line

Second, they worsen the delay matching between the clock and output line, hence causing distortion and reducing the output amplitude. Fig. 74 shows the delay of the output transmission line versus load capacitance. The delay is increased from 12 ps to 15 ps by increasing the load capacitance from 5 fF to 15 fF. The amplitude dependency on the delay mismatch between the clock and the output transmission line can be observed by varying the delay on the clock transmission line. As shown in Fig. 75, the amplitude of a 24.9 GHz sine wave varies in a range of 0.5 dB with a large delay range from 9 ps to 20 ps. It shows that the timing mismatch has relatively small effect on the output bandwidth.

Fig. 76 shows the output frequency responses of a schematic simulation and of a fully extracted layout simulation. The bandwidth is reduced from 20 GHz to 15 GHz. To analyze the individual effect of parasitic resistance and capacitance, simulations with only parasitic capacitance are also performed. The result shows that the major bandwidth degradation comes from the additional parasitic capacitance.

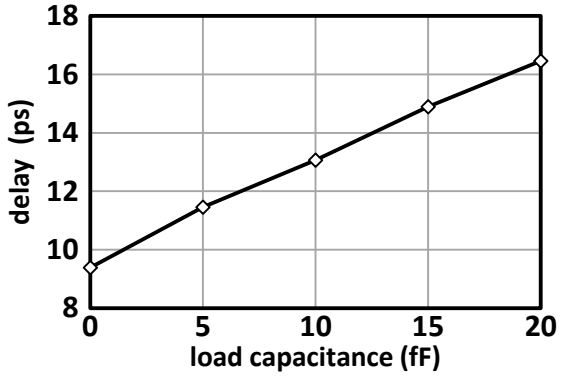


Fig. 74: Delay of the output transmission line versus parasitic load capacitance

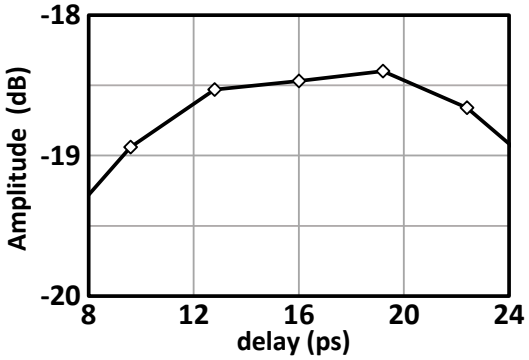


Fig. 75: Output amplitude versus delay of the clock transmission line for 24.9 GHz sine wave at 50 GS/s using one sub-DAC

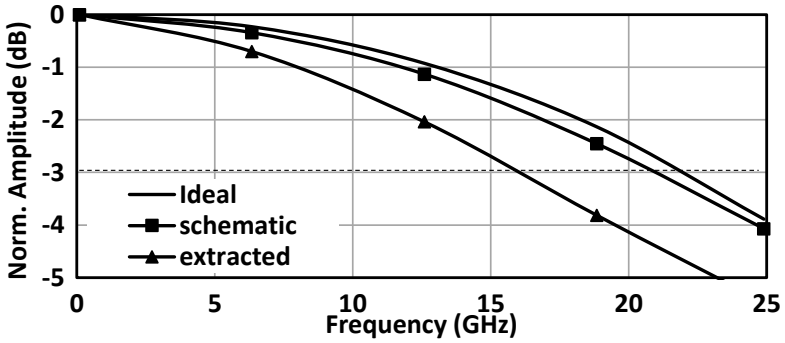


Fig. 76: Parasitic effect on the output frequency response

The ENOB of the DAC over output frequencies is shown in Fig. 77. The results are calculated from transient simulations without noise and mismatch. In the schematic simulation, the ENOB achieves nearly 7 bit at low frequencies and reduces to 5.6 bit at Nyquist frequency. In the extracted simulation, the ENOB achieves 6.5 bit at low frequencies and 5.6 bit at Nyquist frequency. One of the major error sources is the frequency dependent delay mismatch between the clock and output line. At low frequencies, the timing mismatch has a relatively low impact on the ENOB, because the signal varies slowly. As the frequency increases, the timing mismatch becomes larger and degrades the ENOB. However, as the output frequency increases to near the clock frequency, the signal delay between the output and the clock is smaller. The

ENOB degradation due to the delay mismatch at the output frequency near the clock frequency is reduced. The output frequency of each single sub-DAC becomes also very low at the highest frequency. This further improves the timing between the data and the clock. However, since the signal amplitude is lower, the ENOB is lower than that at low frequencies for the same noise floor.

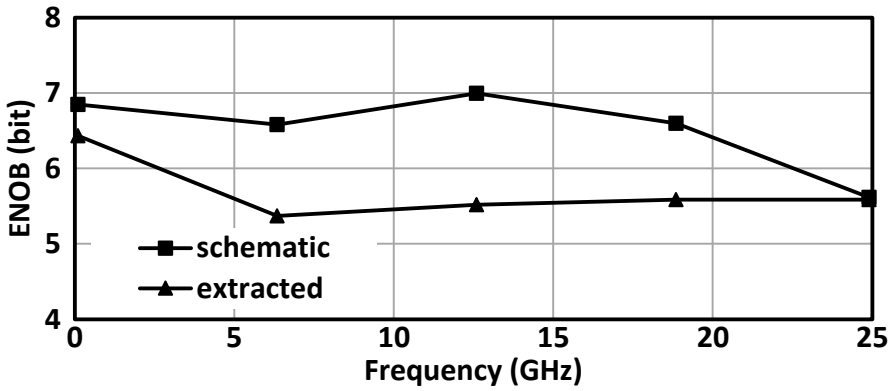


Fig. 77: ENOB versus frequencies with and without metal connection parasitics

4.2.2 Clock Generation

The multi-frequency and multi-phase high-frequency clocks are derived from two single-ended clock inputs with 90° phase shift. The block diagram of the clock dividers and its distribution is shown Fig. 78.

The single-ended clock input is amplified by several stages of the CML buffer to drive the DAC output stage. The 25 GHz clock is divided down to 12.5 GHz 4-phase clocks by a frequency divider. Two 5 bit phase rotators adjust the clock phase with a step size of 11.25° for the phase alignment. Another two frequency dividers generate the clocks for the data MUXs and the on-chip memory. The phase of these clocks can also be adjusted in 90° steps for phase alignment.

The shunt-and-series inductive peaking structure is employed to enhance the bandwidth of the high-speed clock buffers. For the lower speed clock, CMOS buffers are used to drive the CMOS data mux.

The termination voltage of the input 50 Ω termination resistor is generated by resistor voltage divider, providing a proper input common mode voltage (-300 mV) for the CML buffer to operate in a high gain region.

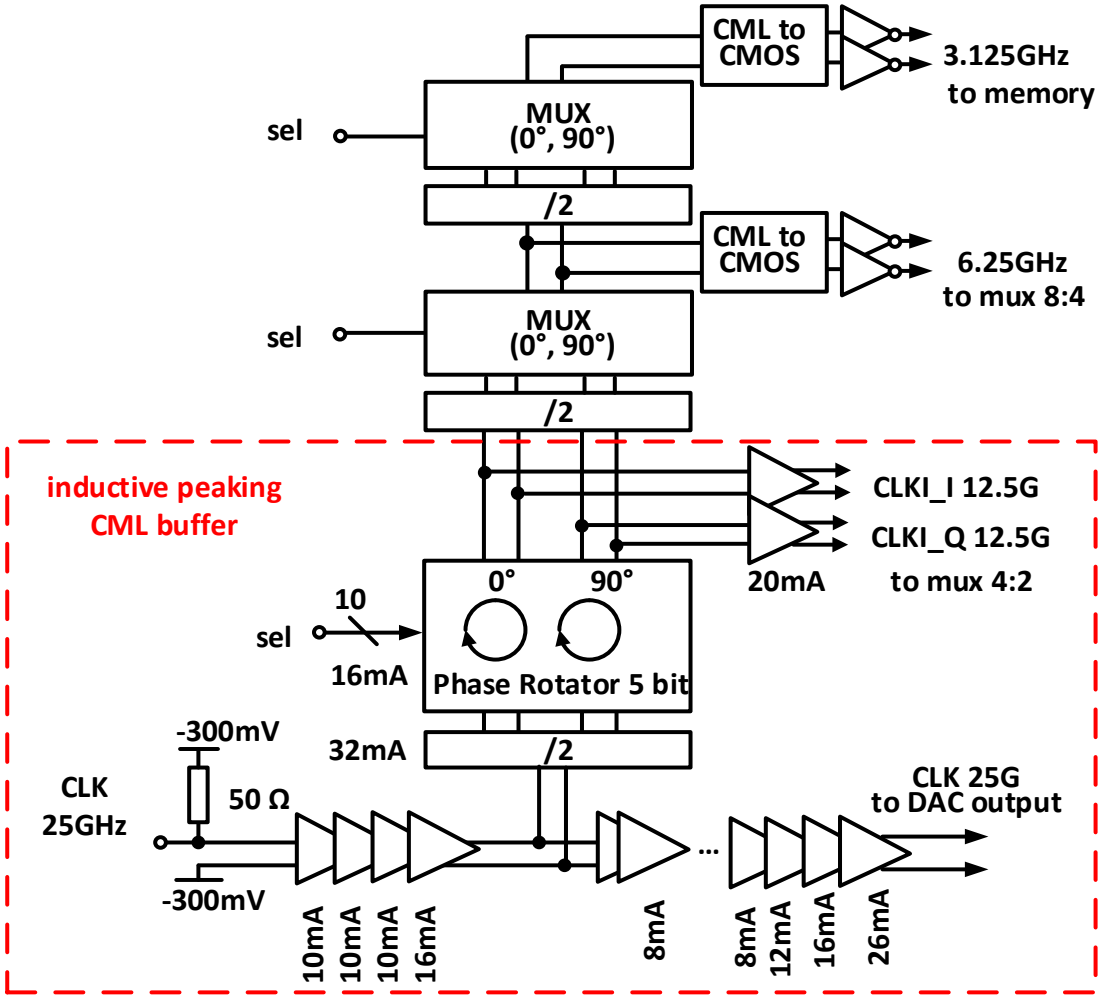


Fig. 78: Block diagram of the multi-phase clock divider and distribution

4.2.2.1 CML to CMOS conversion

The CML-to-CMOS conversion is shown in Fig. 79 (a). It consists of two stages. The first stage is a CML buffer with PMOS diode loads (*M5, M6*) and a cross-coupled pair (*M3, M4*). The diode loads provide sufficiently high output impedance and are necessary to generate a large output amplitude. The cross-coupled pair provides negative impedance, which increases the gain of the CML buffer. The small signal gain of the CML buffer A_{CML} at low frequencies can be calculated by

$$A_{\mathrm{CML}} = \frac{g_{\mathrm{m1}}}{g_{\mathrm{m5}} - g_{\mathrm{m3}}}. \tag{4.8}$$

where g_{m1}, g_{m3}, g_{m5} are the transconductance of the transistors *M1*, *M3* and *M5*, respectively. The transistors (*M3, M4, M5* and *M6*) are identical to achieve the highest DC gain at the common mode voltage, i.e., $g_{\mathrm{m5}} = g_{\mathrm{m3}}$ and $g_{\mathrm{m4}} = g_{\mathrm{m6}}$ at the common mode voltage. The second stage is a PMOS amplifier with an NMOS current mirror load to generate rail-to-rail CMOS output swing. The rising and falling edges of the CMOS outputs (*outn, outp*) are not balanced because the input of the transistors *M8* and *M10* are not equal, which results in a slower falling edge and distorts the duty cycle. The duty cycle error at lower clock frequency

(up to 6 GHz) is relatively small in the simulation over corners. Therefore, only skew correction shown in Fig. 79 (b) is used to correct the skew between the pseudo differential clock signals. The cross-coupled inverter pair generates negative feedback gain, reducing the skew or slew rate differences between the two output nodes. The correction factor is determined by the fixed ratio (6:1) between the CMOS buffer and the cross-coupled inverter pair. However, the negative gain should be relatively small to keep a sufficient overall gain.

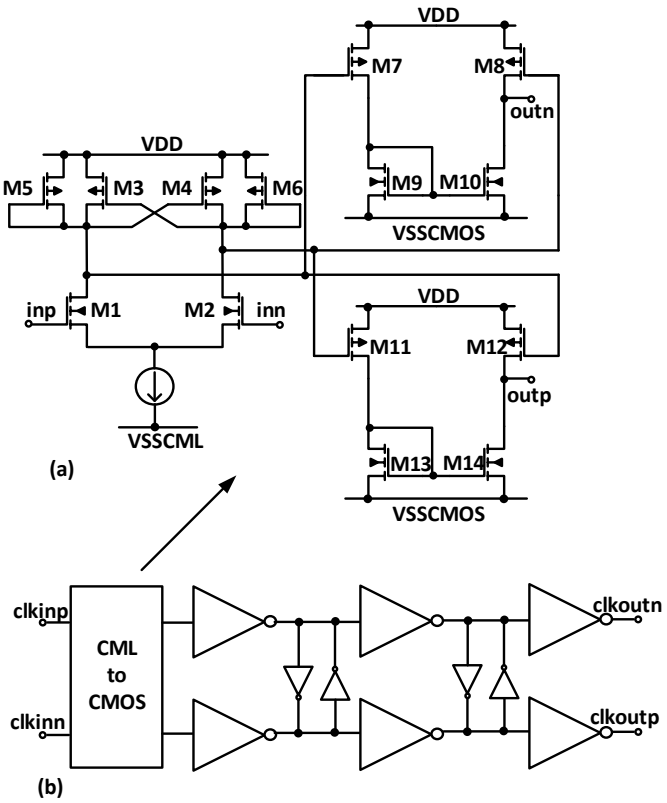


Fig. 79: (a) CML-to-CMOS conversion, (b) skew correction

4.2.2.2 Phase alignment

The timing between the data and the clock at the DAC output stage is critical, especially at very high sampling rates. The clock pulse has to be located in the center of the input data eye to avoid dynamic errors. The intrinsic delay of the clock path and data path cannot guarantee the timing requirement over PVT (process, voltage and temperature). Therefore, the phase alignment shown in Fig. 80 (a) similar to [50] is used to solve the issue. The clock phase for the last stage MUX can be shifted with a 5 bit phase rotator. The phase step size is around 10° of the clock period, i.e., 2.5 ps for 12.5 GHz clock. The LSB of the data input is used for phase detection because the LSB current cell is located in the middle of the DAC output stage and has the lowest input load. The phase detector is implemented with a flip-flop as shown in Fig. 80 (b). The 25 GHz clock is connected to the data input and the LSB data is connected to the clock input of the flip-flop. The state of the phase detector output PD_{out} is determined by the phase position of the data as shown in Fig. 80 (c). If the switching point of the data input D_{in} is at the positive half of the clock period, PD_{out} is high (H). If it is at the negative half of the clock period, PD_{out} is low (L). Sweeping the phase of the input data across the clock edges, PD_{out}

changes its polarity. The phase of the zero crossing can be identified. Afterwards, a fixed phase offset can be added to the data input to center the clock to the data eye.

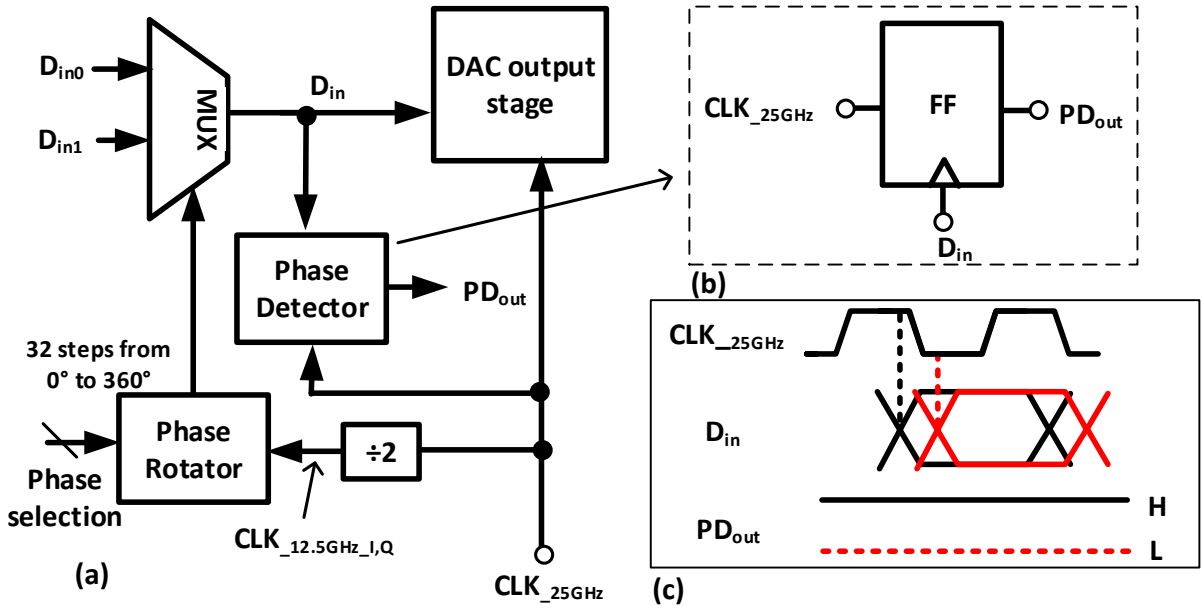


Fig. 80: (a) Phase alignment at the DAC output stage, (b) Implementation of the phase detector, (c) Timing diagram of the phase detector

The phase rotator consists of two differential pairs with inputs connected to the quadrature 12.5 GHz clocks as shown in Fig. 81. The output currents are summed at the output load. The phase can be adjusted by steering the current between the differential pairs. The current sources for each differential pair are binary weighted. The quadrature phase can be selected with the 2 MSBs via the MUX. The sweep of the output clock waveform between 0° to 90° is shown in Fig. 82. To improve the linearity of the phase rotator, the unary-weighted current sources and the thermometer code can be used. The slew rate of the clock input and the gain variation of the differential pair also affect the linearity of the phase rotation.

The phase calibration is also implemented to the 6.25 GHz clock and the 3.125 GHz clock, to allow independent phase calibration and to prevent setup and hold violations. For these two clocks, the phase rotator can only shift the phase in 90° steps. The phase detector measures the phase position between the data and the clock. The outputs of the phase detectors can be selected individually to be connected to an output pad.

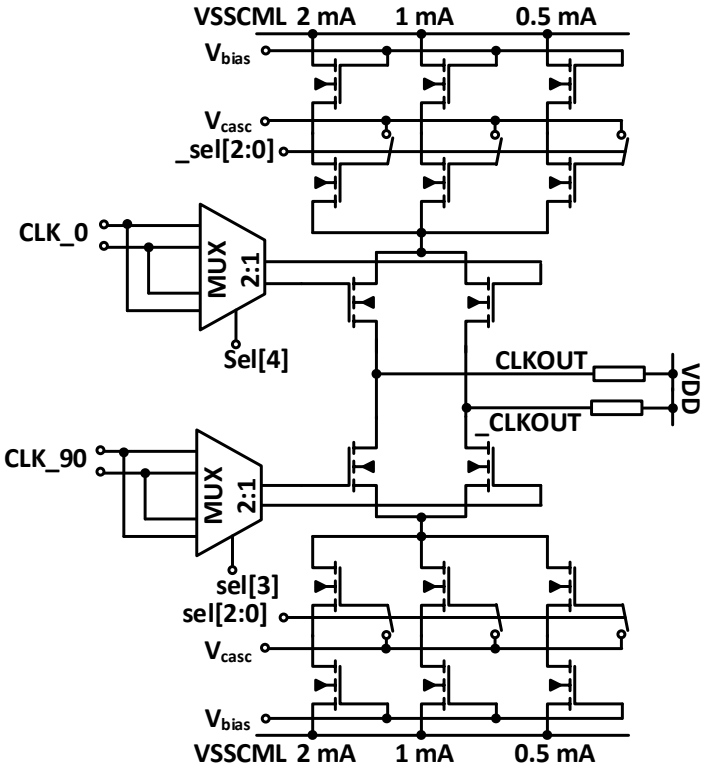


Fig. 81: Phase rotator

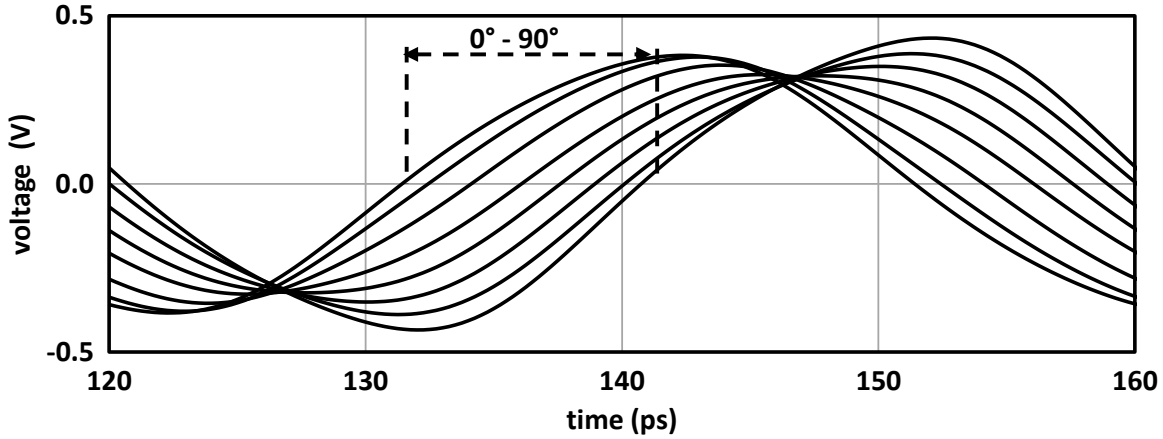


Fig. 82: Sweeping the phase rotator output waveforms from 0° to 90°

4.2.3 Data Multiplexers

The last stage data multiplexing from 12.5 Gb/s to 25 Gb/s is performed by a 4:2 multiplexer driven by 12.5 GHz clocks. A half-bit skew between the two MUX outputs is provided by the quadrature-phase 12.5 GHz clocks. This skew is necessary for the current multiplexing at the DAC output stage. The active phase of the 25 GHz clock has to be placed in the center of the data bit. This prevents data transition errors, i.e., ISI and glitches from becoming visible at the DAC output.

The block diagram of the 4:2 MUX is shown in Fig. 83, similar to the one in [58]. One MUX is driven by the 12.5 GHz in-phase clock (CLK12G5I) and the other by the quadrature phase clock (CLK12G5Q), introducing a half-bit skew. The 4 input data are resampled and shifted by latches and flip-flops to a proper phase position for the correct final multiplexing.

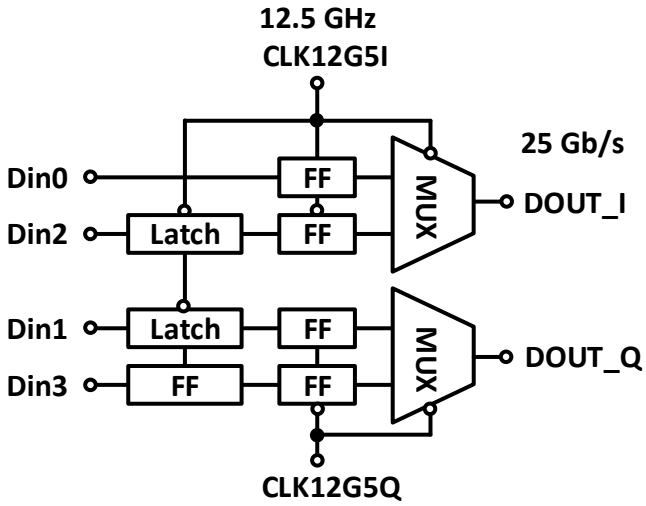


Fig. 83: 4-to-2 MUX

The artificial clock transmission line introduces a delay to each current cell segment along the line. If all input data arrive at the same time, the active clock phase cannot be centered to the data bit for all the current cells as shown in Fig. 84. Therefore, the MUX output signals must also be shifted to compensate the delay on the clock line. Additional delays for the input data signals are generated with the delay buffers. As illustrated in Fig. 84, without delay buffers, the sampling phase of the clock moves to the transition area of the data eye, causing errors at the output. With the delay buffers, the input data signals are moved to compensate for the clock delay. Therefore, the sampling phase of clock can still stay in the valid area of the input data eye.

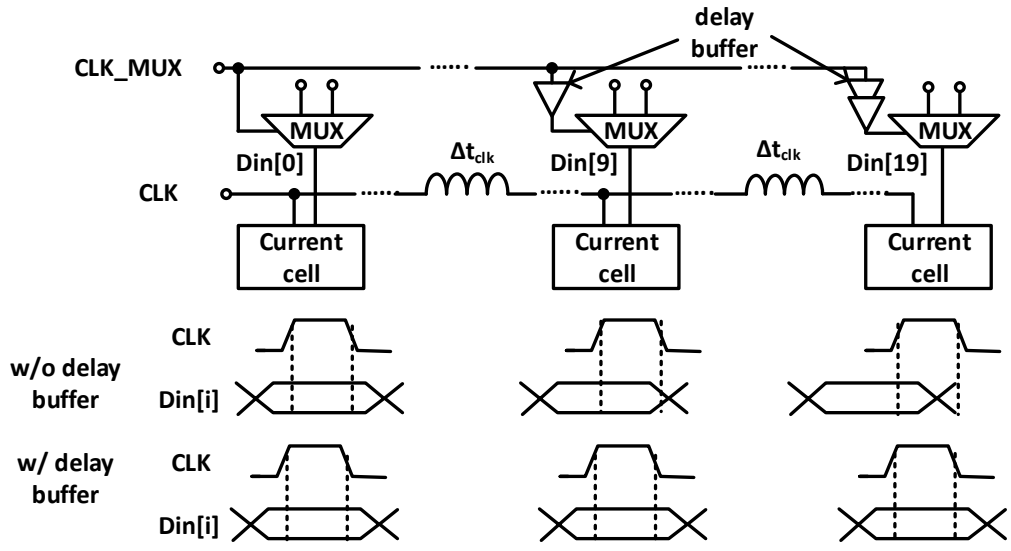


Fig. 84: Timing alignment between the data inputs and the DAC clock along the clock transmission line

4.2.4 On-Chip Memory

4.2.4.1 Architecture

The test method should be considered at an early stage of the design. To test a DAC, a digital data pattern generator must be connected to the digital inputs of the DAC.

For DACs with real-time data interfaces, an FPGA is often used to generate the input data pattern. For sampling rates less than 1.25 GS/s, the LVDS interface on the Xilinx Virtex-4 FPGA can be used to provide a synchronous parallel data stream to the DAC [66]. For sampling rates higher than 2 GS/s, it becomes difficult to generate the synchronous parallel data stream from the FPGA to the DAC data interface. The complexity of the board design and the measurement setup increase drastically. Additionally, the real time interface also consumes a large chip area because of the large number of differential input data pads.

To ease the test for the DAC and to save the chip area, a 1 kByte memory is embedded on the chip to store the digital data pattern for the DAC. The architecture of the on-chip memory is shown in Fig. 85, similar to the one in [67]. Five control inputs (*ckin, Din, Dout, En, rw*) are required to control and program the memory.

The memory can be switched between two operating modes (*read*, *write*) via the pin *rw*. In *write* mode, the parallel outputs of the shift registers are connected to the SRAM bit lines (BL). The address bits (*addr <3:0>*) are generated by a 4 bit counter. The counter is triggered by *En* in *write* mode. For each address, the row decoder selects the correspondent row by generating a pulse on the correspondent word line (WL). The contents of the shift registers are then written to the SRAM cells in the selected row. The timing diagram of the *write* mode is illustrated in Fig. 86.

In *read* mode, the BLs are disconnected from the shift registers. The counter is triggered by the clock CK1G5, which is derived from the DAC clock. In the first half period of the clock, the BLs are precharged to half the supply voltage $-(V_{DD} - V_{\mathrm{SSCMOS}})/2$; in the second half, a correspondent WL is set to high, allowing the contents of the SRAM cells in this row to overwrite the BLs. The parallel data stream on the BLs (*BL<31:0>*) is serialized to an 8 bit data stream at the output via the MUX. The timing diagram of the *read* mode is shown in Fig. 86.

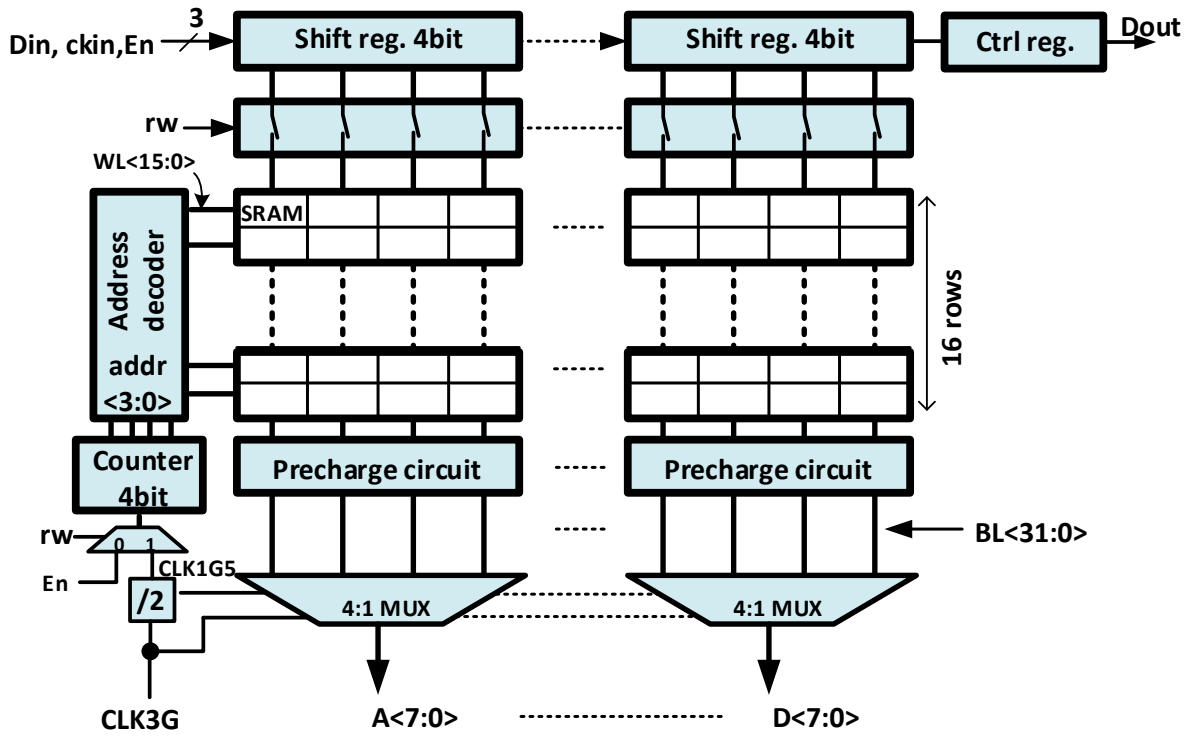


Fig. 85: Memory architecture (1 of 4 memory blocks)

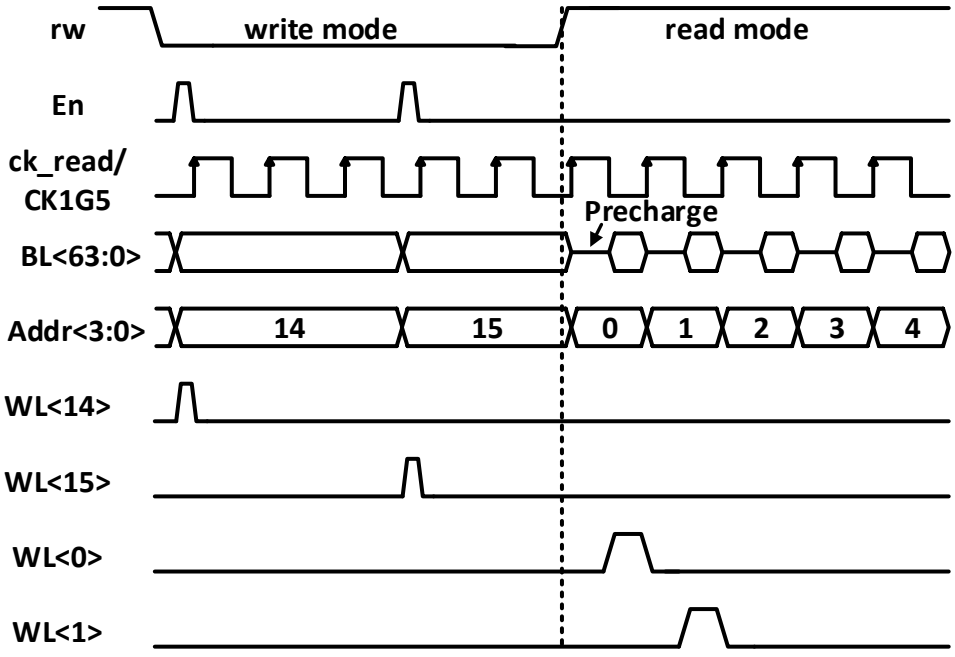


Fig. 86: Timing diagram of the memory

4.2.4.2 Circuit design

Fig. 87 shows an SRAM cell consisting of 6 transistors [68]. The cross-coupled inverter latch (*M1, M2, M3, M4*) can be accessed by two access-transistors (*M5, M6*) via the word line (WL) and the bit lines (BL, BLB).

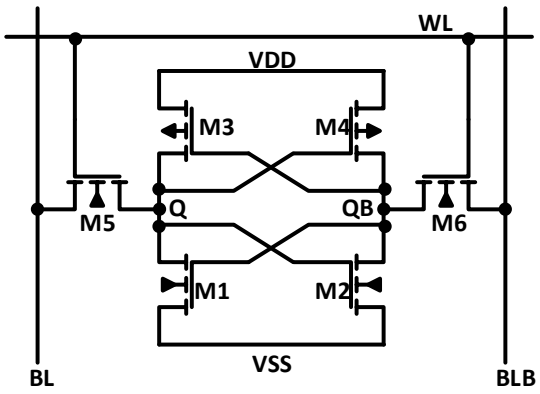


Fig. 87: SRAM cell

To write the SRAM cell, the bit lines are driven by a large driver with complementary values. By setting the WL to high, the access-transistors are turned on and the state of the latch can be overwritten by the value on the BL and BLB. By setting the WL to low, on the other hand, the value stored in the latch remains unchanged.

For reading, the bit line driver has to be disconnected from the BLs. A precharge driver charges both BLs to half the supply voltage $-(V_{DD} - V_{\mathrm{SSCMOS}})/2$ and is disconnected from the BLs afterwards. By setting the WL to high, the latch changes the BL to its stored value. The precharging prevents the previous state of the BLs changing the state of the latch due to large parasitic capacitance on the long BL track. The precharging also increases the reading speed because of the reduced output swing. The bit lines are directly connected to the output inverters without any sensing amplifier. Since the load on the BLs is relatively low, the latches are able to drive the inverter directly with reasonable speed.

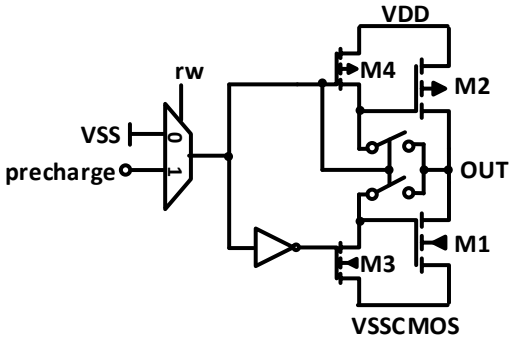


Fig. 88: Precharge driver

The precharge driver is shown in Fig. 88. The precharge voltage is generated by the inverter (*M1, M2*) with its input and output shorted via a transmission gate. The precharge voltage is around $V_{DD}/2$, depending on the ratio between *M1* and *M2*. In *write* mode, the input is connected to VSSCMOS. The gate of *M2* and *M1* are connected to VDD and VSSCMOS, respectively. The output is thus floated, allowing the BLs to be changed by the BL drivers. In *read* mode, the precharge output is controlled by the precharge signal (*precharge*). If *precharge* is set to high, the BLs are charged to (VDD – VSSCMOS)/2. If *precharge* is set to low, the output is floated, allowing the BLs to be overwritten by the SRAM cell.

The address counter is made up of a four-stage cascade frequency divider. The output of each stage corresponds to one address bit, as shown in Fig. 89. The counter is triggered by the input clock *CK*. In *write* mode, the clock input is driven by the *En* signal. In *read* mode, the clock input is connected to the divided DAC clock.

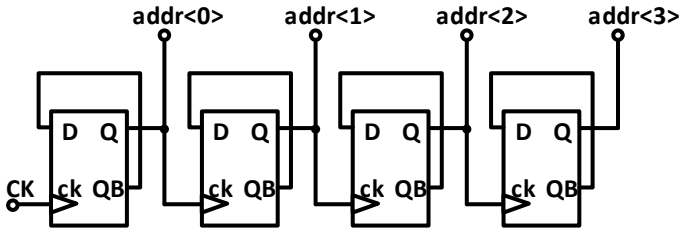
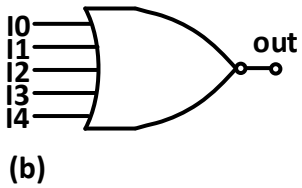


Fig. 89: Counter

According to the address, a corresponding WL is selected via the row decoder. The truth table of the row decoder is shown in Fig. 90 (a). The row decoder can be implemented with 16 fold 5 bit NOR gates according to the truth table. The symbol of the NOR gate and its circuit implementation are shown in Fig. 90 (b) and (c), respectively. The NOR gate is implemented in pseudo NMOS style to reduce the input load. The counter clock *CK* is connected to one of the NOR input in *read* mode, generating a pulse for the precharging. The waveform of the counter and the row decoder is illustrated in Fig. 91.

Addr <3:0>	WL<15:0>
0000	000000000000001
0001	000000000000010
0010	000000000000100
.....	.....
1110	010000000000000
1111	100000000000000



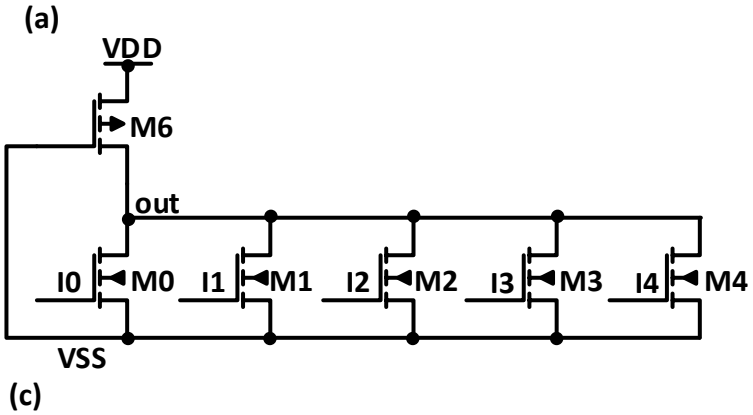


Fig. 90: (a) Row decoder true table (b) Symbol of a 5 bit NOR gate (c) Schematic of a 5 bit pseudo NMOS NOR gate.

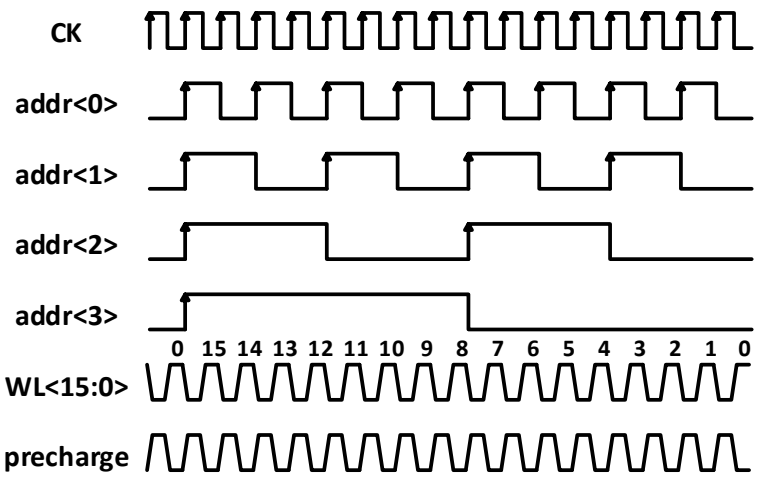


Fig. 91: Waveform of the counter and the row decoder

4.2.4.3 Synchronization

There are two memory blocks located at the top and the bottom of the DAC chip. Each memory block has its own clock (e.g., CKI16, CKQ16) derived from the two input clock via several stage of frequency dividers, as shown in Fig. 92 (a). During programming (*write* mode), the memory is toggled by the *En* signal via external pin. After programming, the memory has to be switched back to the *read* mode by switching the clock input from *En* to CKI16 or CKQ16. The switching of the clock input is triggered by the signal *rw*, which is synchronized to each memory clocks (e.g., CKI16, CKQ16). The phase between the two memory clocks has to be aligned to reproduce the samples at the DAC output in the correct order. However, due to the random start of the frequency dividers, the phase skew between the two memory clocks can be any one of the 16 phases in one clock cycle for each restart, as shown in Fig. 92 (b). Since there is no possibilities to detect the phase skew between the memory clocks on the chip, the phase skew has to be measured at the DAC output, by outputting a pulse at each sub-DAC output and measuring the delay of the two pulses using an oscilloscope. The phase skew can be tuned by shifting one clock using the phase rotator until the correct phase skew is obtained, or reprogramming the memory with a pre-skewed sample order that the incorrect output phase skew is compensated. However, both methods require an oscilloscope to measure the DAC output, which is impractical to perform the phase calibration. To improve the process of the clock phase alignment, a phase detector similar to the one shown in Fig. 80 can be used. One of the two memory clocks can be used to sample the state of the other clock. By shifting the clock phase using the phase rotator, the position of the clock edge can be detected and recorded. With this information, the clock can be shifted back to the correct phase position using the phase rotator. The calibration can be done automatically with existing programming interface, without the need of using an oscilloscope, thus simplifying the calibration process.

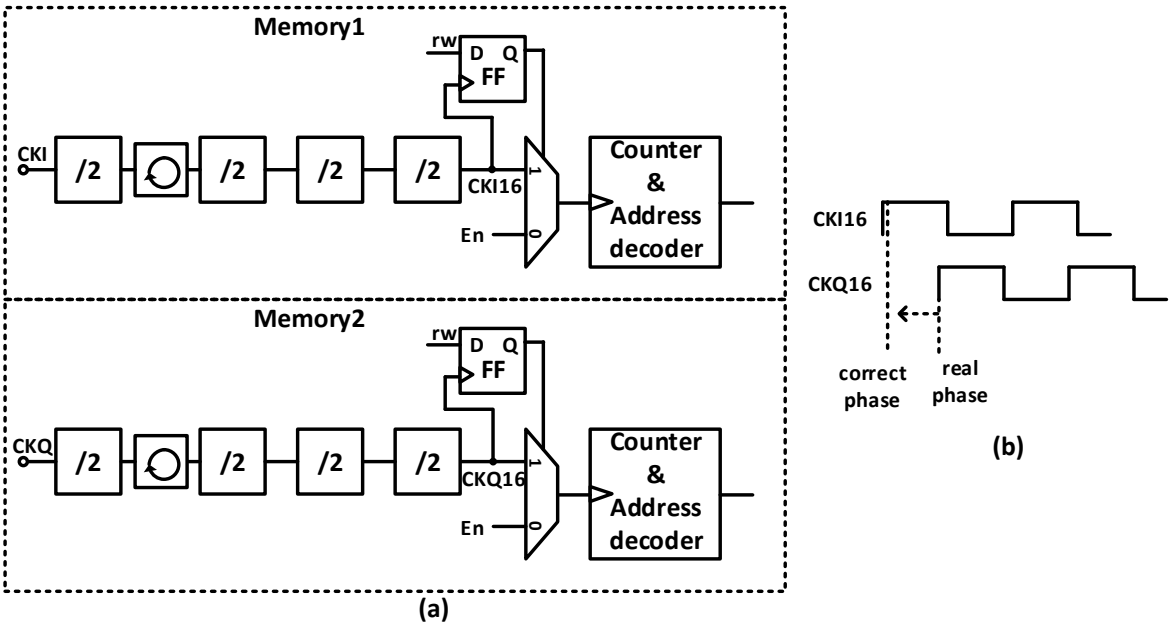


Fig. 92: (a) Clock generation and selection circuit for the memory, (b) phase skew between the two memory clocks

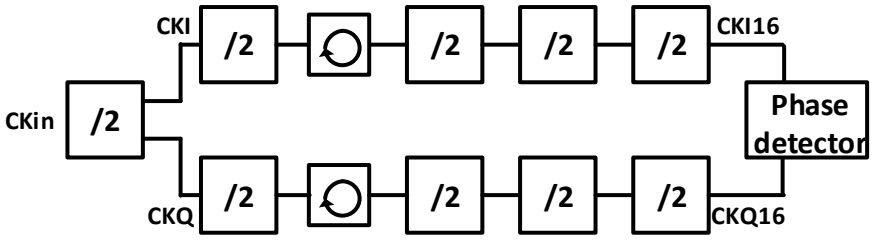


Fig. 93: A possible solution for the clock alignment

4.3 Chip Layout

The chip micrograph is shown in Fig. 94 (a). The chip area is 1.44 mm^2. The two interleaved sub-DACs including the data MUXs, the clock dividers and distribution and the two memory blocks are horizontally symmetric. The two high frequency clock inputs are on the left side and the DAC outputs are on the right side.

A zoom-in to the DAC output stage is shown in Fig. 94 (b). The inductors of the clock lines and the output line are implemented with the top copper metal. The wide top aluminium metal is used for the power supply routing to minimize the IR drop. To reduce the parasitic capacitance, no metal is routed under the clock and output transmission lines.

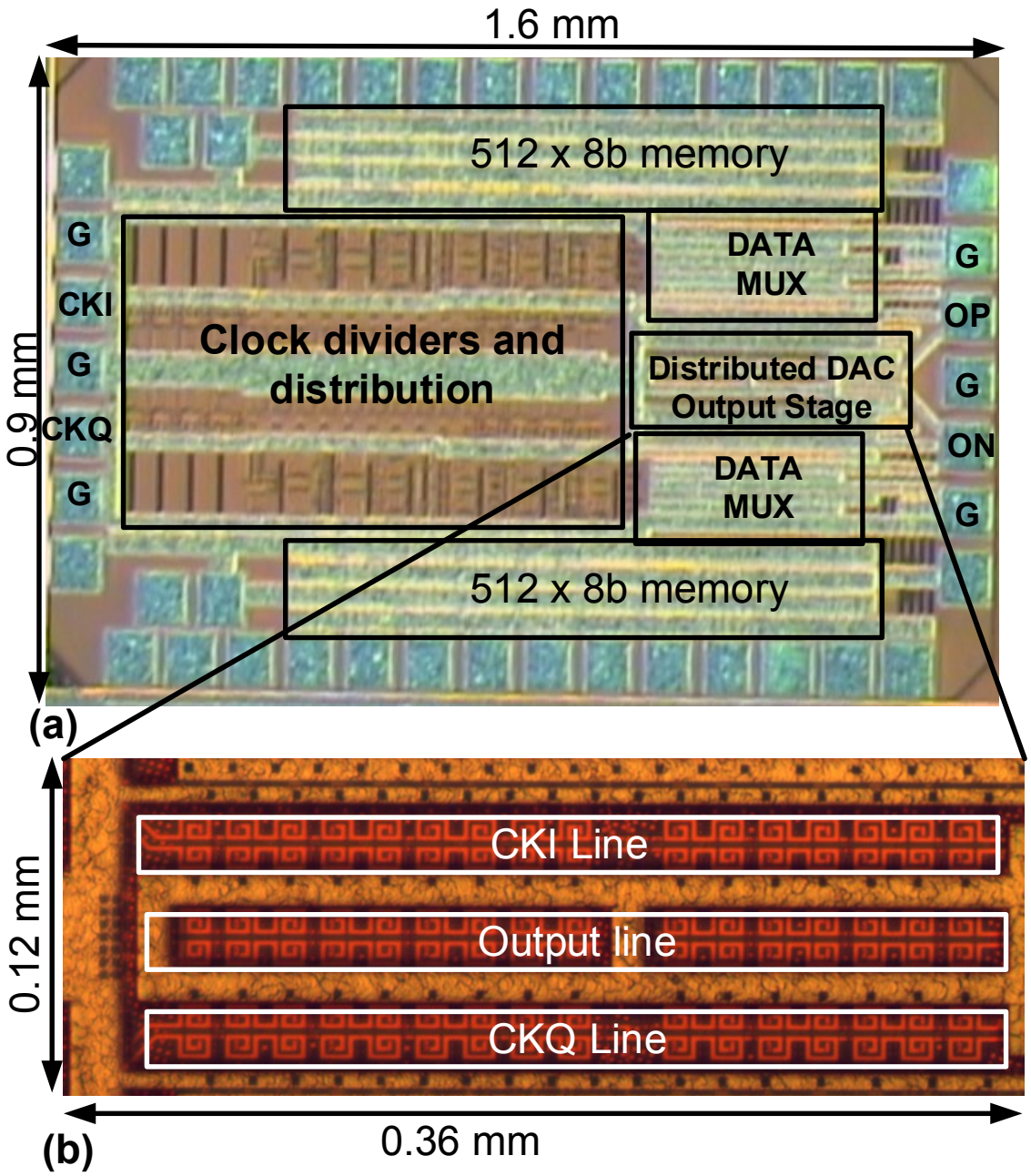


Fig. 94: (a) Chip micrograph (b) Micrograph of the distributed DAC output stage

5 Measurement

This chapter presents the measurement results of the DAC chip. It first introduces the measurement setup and the test board design. Then, the static performances – including the output characteristics, the INL, the DNL and the channel mismatches – are presented, followed by the frequency-domain performances, e.g., the output spectrum, the ENOB and the bandwidth. After that, the eye-diagrams of the multi-level modulated outputs demonstrate the time-domain dynamic performance. Finally, the DAC is used in an optical transmission experiment to show its capability of high-speed transmission.

5.1 Measurement Setup

The measurement setup for characterizing the DAC is illustrated in Fig. 95. The 25 GHz clock comes from a signal generator. The clock is split by a power splitter whose outputs are connected to the I and Q clock inputs of the DAC. The 90° phase skew is achieved by a mechanical phase shifter. The DAC outputs are measured with a subsampling oscilloscope with a 65 GHz bandwidth. The signals are then stored in a PC via a general-purpose interface bus (GPIB) and are post processed with MATLAB. The on-chip memory is programmed by an external program device based on a single-board PC via general-purpose input/output (GPIO) [69]. A level-down shifter shifts the logic levels from the GPIO level (0 V – 3.3 V) to the CMOS logic level on the DAC chip (-1 V – 0 V).

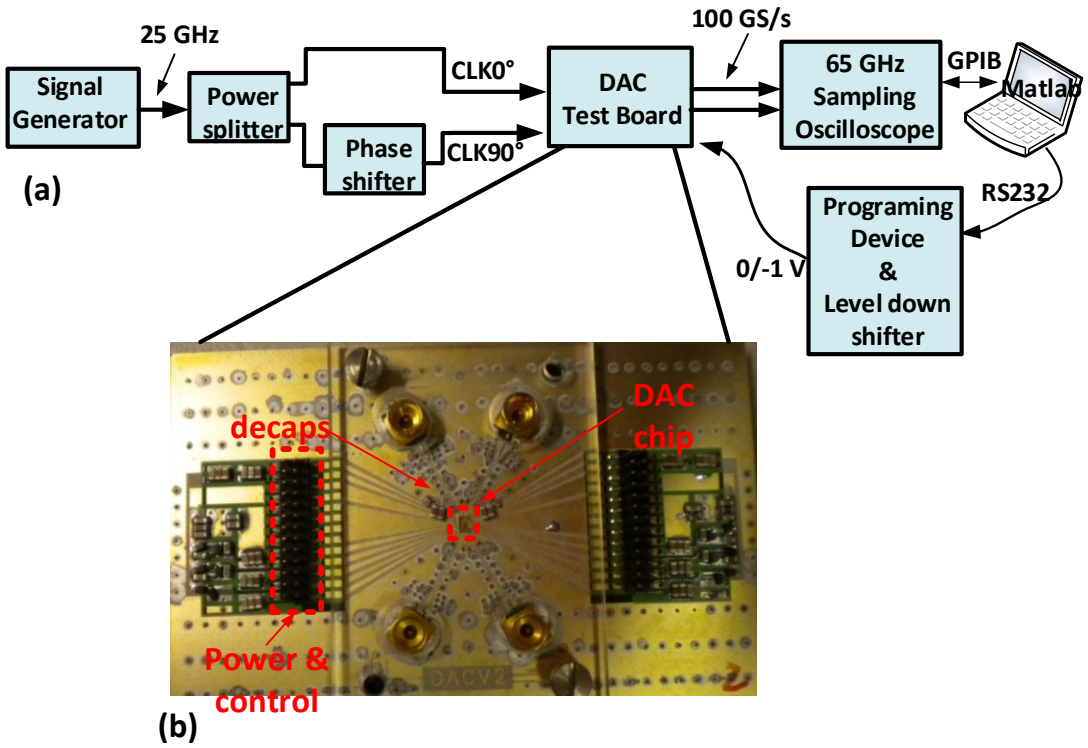


Fig. 95: (a) Measurement setup (b) DAC test board.

The DAC test chip is directly bonded and mounted on an RF test board. The analog outputs can be directly probed with 40 GHz bandwidth on-wafer probes or bonded onto the RF-PCB. The RF-PCB is made from Taconic RF-60A, which provides high RF performances. The PCB is directly mounted on a brass sheet for better mechanical stability. The chip is placed into a PCB cavity and is attached to the brass directly. This lowers the chip surface to the same height as the PCB surface and shortens the bond wire distance between the pads on the chip and the trace on the PCB. The sub-miniature push-on (SMP) connectors are used for the RF connections. A 50 Ω grounded coplanar wave-guide (GCPW) is designed for RF routing on the PCB. The decoupling capacitors are placed on both sides of the PCB. A few of them are placed very close to the chip to reduce the equivalent series resistance (ESR) and inductance (ESL) to improve the high frequency filtering. The sockets on both sides are connectors for the power supply, the bias references and the digital control signals. The electrostatic discharge (ESD) diodes protect the CMOS inputs.

5.2 Static Performance

Based on the output characteristics of the four individual RZ DAC cores, the static performances, including the channel mismatches, the INL and the DNL, can be determined. To measure the output characteristics, a ramp function is stored in the memory of one sub-DAC while the other sub-DAC is programmed to a constant value. The ramp function outputs two samples for each code and is swept from -128 to 127. The outputs of the two RZ-DACs in each sub-DAC can thus be measured from the same sweep.

The DAC output is averaged over 64 times to avoid random errors. However, the ramp function period is sampling rate dependent. Due to the settling time, the measurement results are affected by the sampling rate. Ideally, the static measurement should be done with a low sampling rate with a sufficient settling time. Because of the measurement setup limitation, the sampling rate is limited to several GS/s, which could have a negative impact on the results.

The output characteristics of the four RZ-DACs are shown in the top left graph in Fig. 96 (a). Without calibrating the output full-scale voltage, the maximum differential full-scale voltage is 0.44 V and the LSB voltage is 1.72 mV. The channel mismatches between the two RZ-DACs in one sub-DAC and across the two sub-DACs are shown in the top right graph. The maximum mismatch is 2.2 LSB. The maximum DNL and INL are 2.0 LSB and 1.4 LSB, respectively.

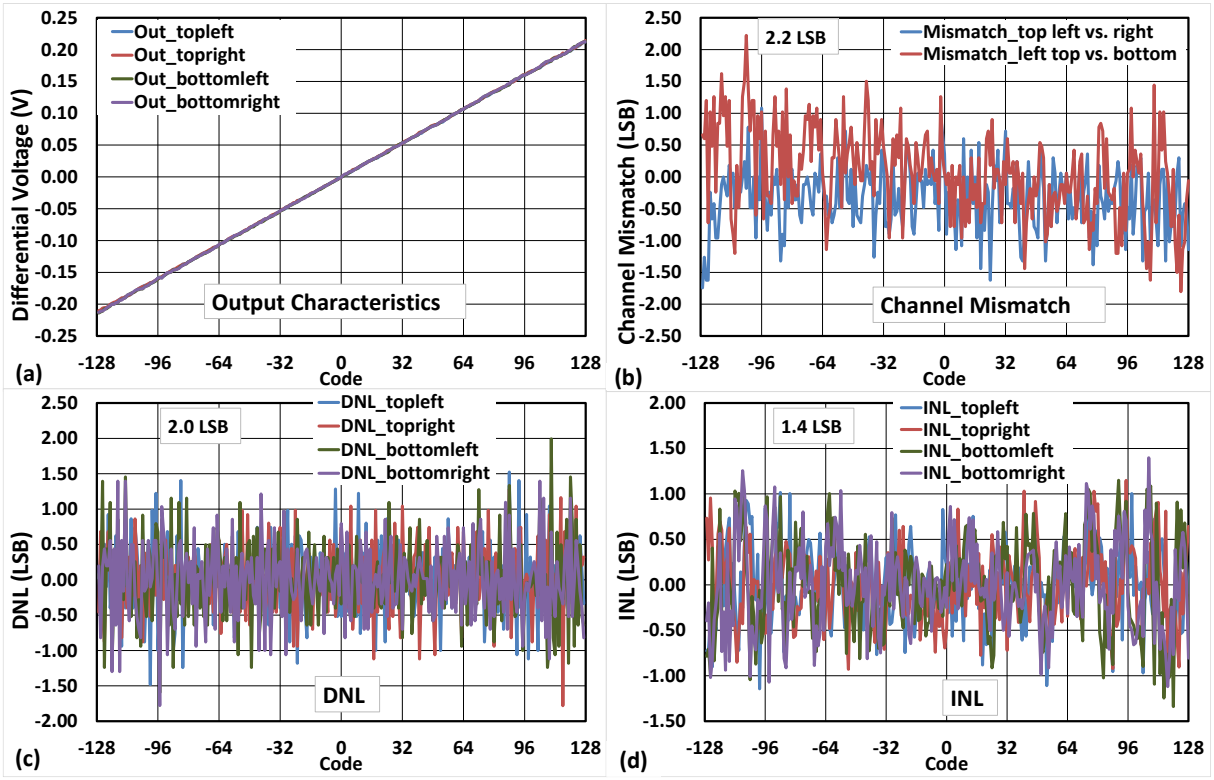


Fig. 96: Measurement results extracted from the ramp-function outputs: (a) output characteristics, (b) channel mismatch, (c) DNL and (d) INL

5.3 Frequency-Domain Performance

The frequency-domain performance characterization focuses on the ENOB, the SFDR and the bandwidth. The input of the DAC is a quantized sine wave with various frequencies stored in the on-chip memory. The DAC output is probed with on-wafer probes and measured with a subsampling oscilloscope. The measured waveform data is stored in a PC and post-processed with fast Fourier transform (FFT) to analyze the spectrum. Alternatively, the output spectrum can be obtained directly with a spectrum analyzer. However, the bandwidth of the available spectrum analyzer in the lab is much less than the Nyquist frequency (50 GHz) at the highest sampling rate (100 GS/s). Therefore, the sampling oscilloscope with a 65 GHz analog bandwidth, a 14 bit vertical resolution and a 1 fs horizontal resolution [70] is best suited for measuring the 100 GS/s 8 bit DAC.

To calculate the spectrum of the sine wave, the FFT window length must be an integer number multiple of the sine wave period, i.e., coherent sampling. With coherent sampling, the signal frequency and its harmonics fall on the frequency bins of the FFT spectrum. If not, it will lead to discontinuity between the last and the first sample of the successive FFT period, thus causing many unwanted harmonics. The power of the signal frequency component is spread out over the frequencies in the spectrum, i.e., spectral leakage. Using the windowing method can reduce the discontinuity and, hence, the spectral leakage, but the leakage cannot be eliminated [21].

To ensure the coherent sampling, the relation between the signal frequency, the sampling rate and the FFT sampling frequency has to be [71]

$$f_{\mathrm{sig}} = \frac{M_{\mathrm{sig}}}{N_{\mathrm{sample}}} f_{\mathrm{s}} = \frac{M_{\mathrm{FFTsig}}}{N_{\mathrm{FFT}}} f_{\mathrm{FFT}} \tag{5.1}$$

where f_{sig}, f_{s} and f_{FFT} are the signal frequency, the DAC sampling frequency and the FFT sampling frequency, respectively; M_{sig}, N_{sample}, M_{FFTsig} and N_{FFT} are the number of cycles of the output signal, the number of samples of the DAC output, the number of cycles of the sampled output signal and the number of samples of the FFT, respectively.

It is not required, but is a common choice, that the number of FFT points is the power of two because this allows the FFT calculation to be most efficient. M_{sig} and N_{sample} as well as M_{FFTsig} and N_{FFT} have to be relatively prime [71], i.e., they have no common factors. The common factors lead to different harmonics located on the same frequency bin in the FFT after aliasing, effectively reducing the number of the frequency bins and the measurement accuracy. Therefore, M_{sig} and M_{FFTsig} have to be odd or a prime number.

Since the on-chip memory can store 1 k 8 bit samples, the maximum period of the output sine wave at 100 GS/s is 10.24 ns. The captured waveform of the sine wave with single-ended peak-to-peak voltage of 500 mV is shown in Fig. 97.

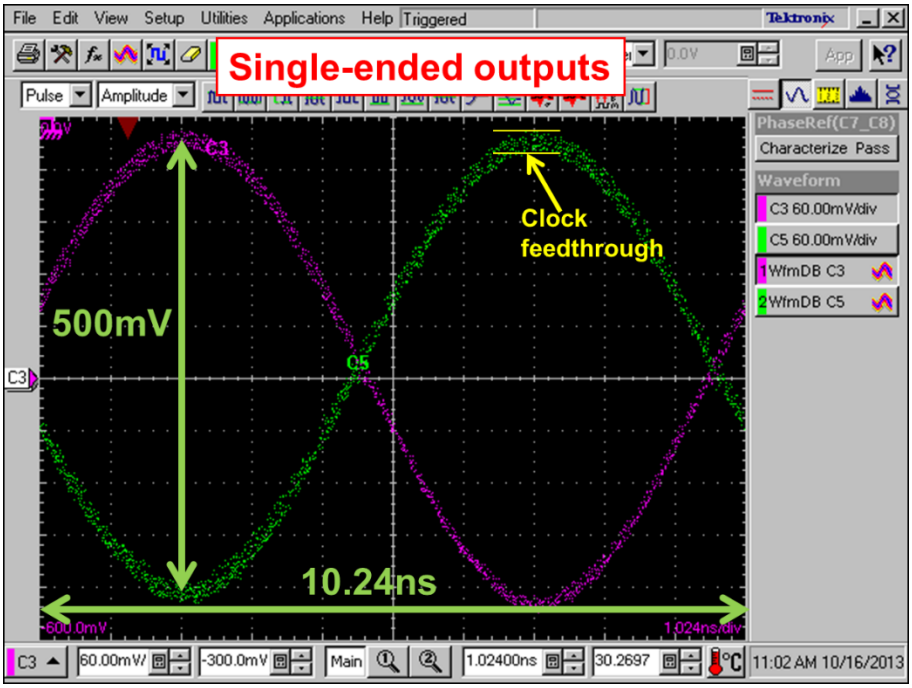


Fig. 97: 97.65 MHz sine wave of the DAC output

The differential output voltage can be determined directly by the mathematical functions in the oscilloscope. Before calculating the differential output, the timing skew between the two single-ended signals has to be removed with the de-skewing function of the oscilloscope. The skew is measured at the maximum DAC output frequency. The signal of one channel in the oscilloscope is shifted until the amplitude of the differential output reaches its maximum.

To increase the measurement accuracy, the vertical and the horizontal resolution of the oscilloscope have to be set to their maximum. Therefore, the time span of the horizontal axis is

set to the maximum period of the DAC output sine wave, e.g., 10.24 ns for 100 GS/s. For the oscilloscope, a maximum of 4000 data points can be saved. Each data point is averaged by 16 times to reduce the impact of random events. The averaging also reduces the noise at the DAC output. The data points are linearly interpolated and resampled with 4096 points to allow coherent sampling. The linear interpolation is employed because the sampling frequency is much higher than the signal bandwidth, hence providing enough accuracy. To reduce the aliasing of the high frequency harmonics and the noise folded back to the first Nyquist frequency band, a 4096 point FFT is applied, which corresponds to 4 times oversampling of the DAC output samples at 100 GS/s.

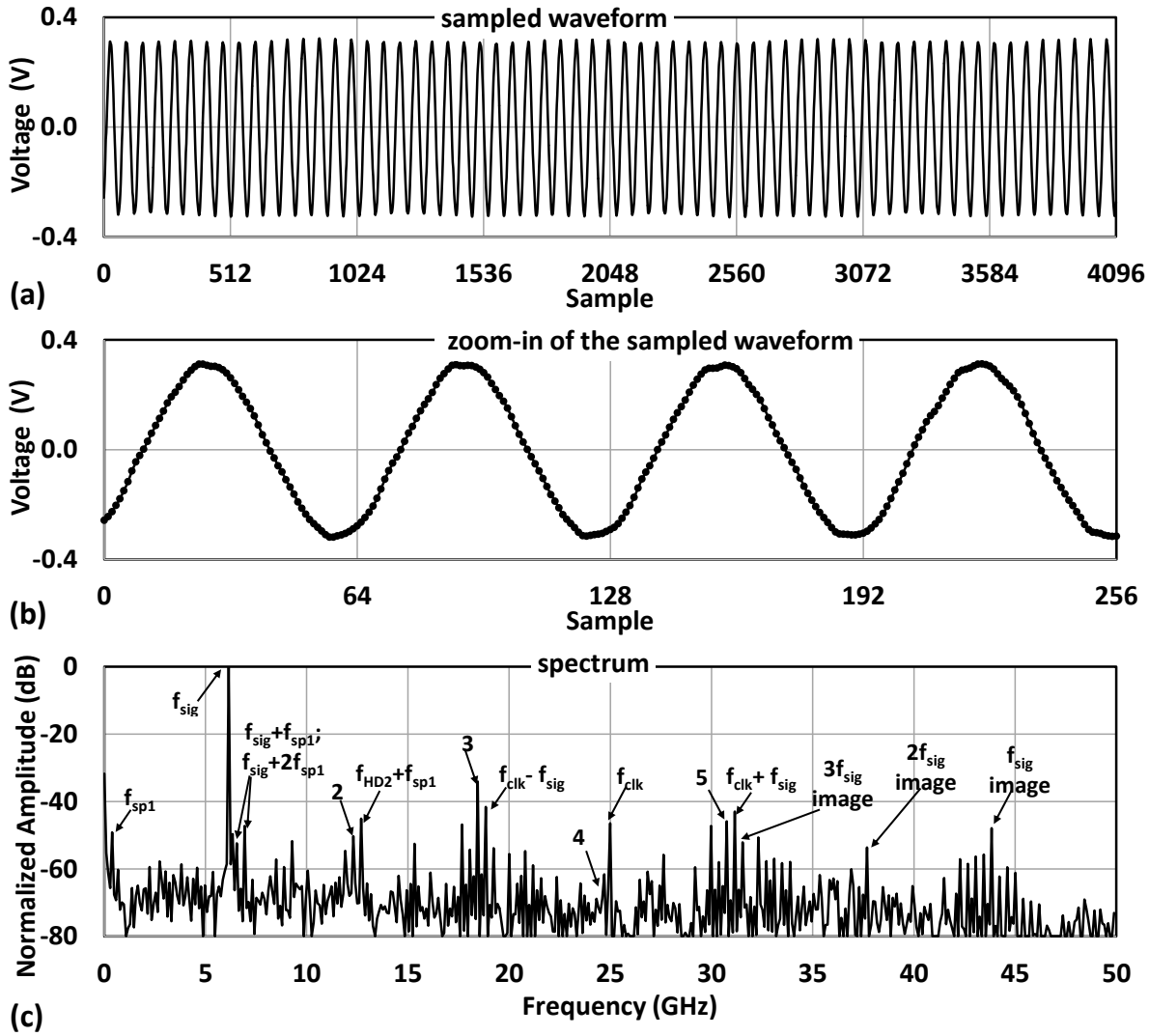


Fig. 98: Measurement results of a 6.152 GHz sine wave output at 100 GS/s with M_{sig}=63 and N_{sample}=1024: (a) resampled sine wave with 4096 points, (b) zoom-in to the sampled sine wave, (c) spectrum of the measured sine wave

An example of the measurement results is shown in Fig. 98. To achieve the coherent sampling shown in (5.1), the signal frequency of the DAC output is set to 63/1024·100GS/s ≈ 6.152 GHz. The measured output data is resampled with 4096 points, as shown in Fig. 98 (a). A zoom-in view of the sampled waveform is shown in Fig. 98 (b). The spectrum shown in Fig. 98 (c) is

generated by applying a 4096 point FFT to the sampled sine wave. The amplitude of the spectrum is normalized to the signal amplitude. The frequency range of the spectrum is 8 times the Nyquist frequency, which is far away from the frequency range of interest. Therefore, only the first 512 frequency bins are shown to analyze the spectrum of the first and the second Nyquist frequency band.

The most important frequency components are marked on the spectrum to analyze the errors and the performance of the DAC. It can be seen that the second harmonic distortion is still visible, which indicates an imbalance of the phase and the amplitude between the differential DAC outputs. The imbalance can come from the DAC output stage due to an imbalanced layout and an imbalanced output load. The imbalance of the differential signals from the output pads to the sampling oscilloscope, and the imperfect de-skewing of the two outputs, are also the major contributors to the second-order distortion. The third-order harmonic distortion is directly related to the linearity of the DAC, which is affected by the mismatch of the current source, the finite output impedance of the current cells and the resistance of the output transmission line, etc. Since the third-order harmonic distortion is the highest spur in the spectrum, it directly limits the SFDR performance. The spurs at $f_s/4 \pm f_{sig}$ are caused by gain mismatch between the sub-DACs and by the duty-cycle error. The images of the fundamental frequency and its harmonics are also marked in the spectrum. The amplitude of the signal frequency image is around -50 dB. Considering the sinc amplitude roll-off (ca. -17 dB at this frequency), the signal image is suppressed by 33 dB by the parallel-path time-interleaving. The spurs around the fundamental frequency and its harmonic can be the intermodulation product of the low frequency spur f_{sp1} (ca. -50 dB) and the frequencies $f_{\mathrm{sig}}, 2f_{\mathrm{sig}}, 3f_{\mathrm{sig}}$, etc. The low frequency spur can be caused by the coupling from the data pattern frequency to the DAC output.

To show the image frequency suppression of the parallel-path time-interleaving, the DAC outputs and their spectrum, with only one single sub-DAC operating at 50 GS/s (AMUX time-interleaving) and with two sub-DACs operating at 100 GS/s (parallel-path time-interleaving), are compared in Fig. 99. The output frequencies in both cases are chosen to be the same (21.78 GHz). The output waveforms and their zoom-in view are shown in Fig. 99 (a) and (b) as well as (c) and (d), respectively. The spectrum of the output waveforms are shown in Fig. 99 (e) and (f). It can be seen that the amplitude of the fundamental image, i.e., $f_{\mathrm{s}} - f_{\mathrm{sig}}$ at 50 GS/s and $f_{\mathrm{s}}/2 - f_{\mathrm{sig}}$ at 100 GS/s, is reduced by more than 30 dB. However, some other spurs, e.g., the third order harmonics, the gain and duty circle error at ($f_{\mathrm{s}}/4 \pm f_{\mathrm{sig}}$), are increased in parallel-path time-interleaving mode at 100 GS/s. One reason for the increase of the errors is that two sub-DACs are operating instead of one, and the errors of both sub-DACs at the same frequencies are summed up.

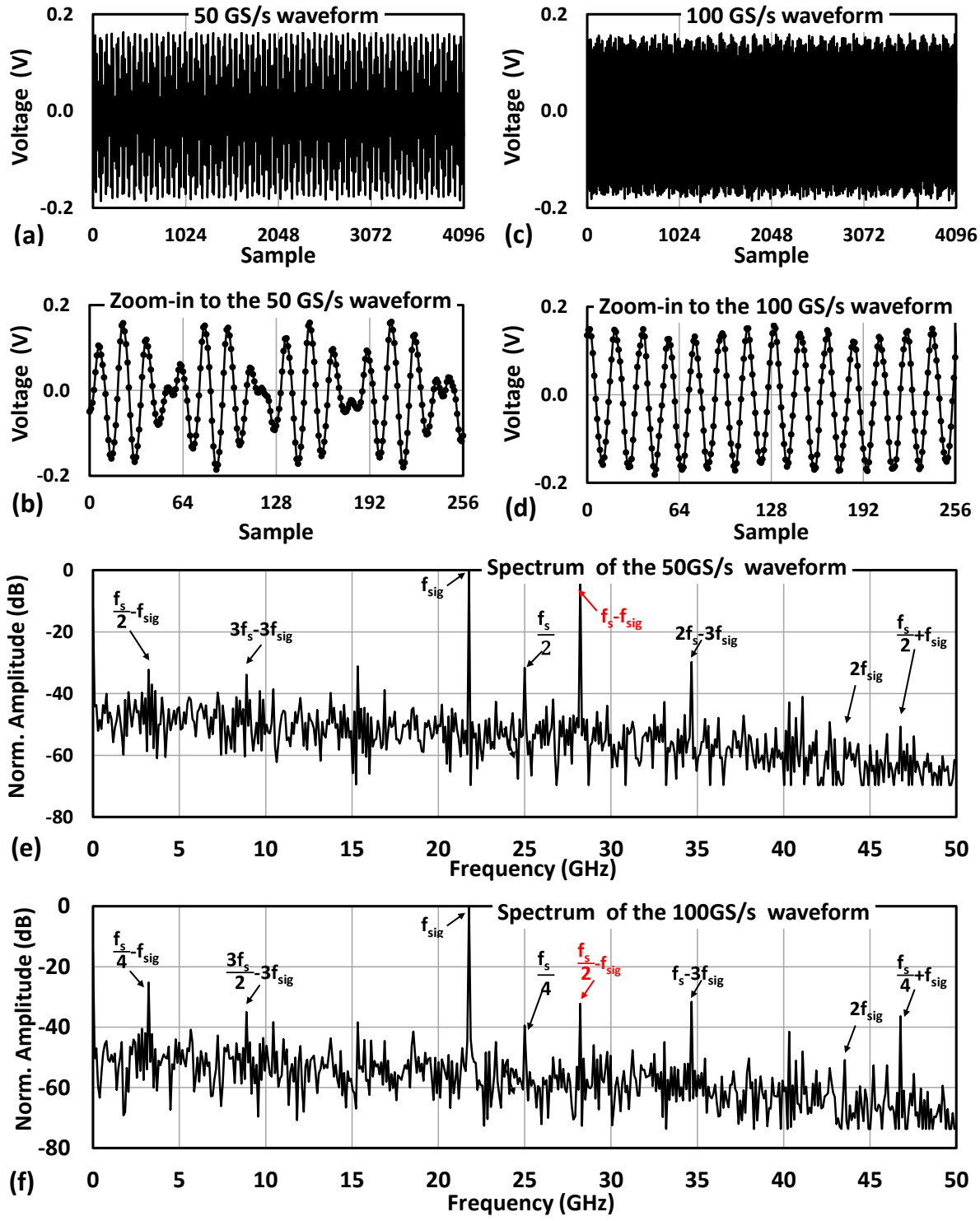


Fig. 99: Comparing the DAC outputs and their spectrum with only one single sub-DAC operating at 50 GS/s (AMUX time-interleaving) and with two sub-DACs operating at 100 GS/s (parallel-path time-interleaving). The signal frequency is 21.78 GHz for both modes, and the image amplitude is reduced by more than 30 dB in later case. (a) 21.78 GHz sine wave generated by a single sub-DAC operating at 50 GS/s. (b) Zoom-in to the 50 GS/s waveform. (c) 21.78 GHz sine wave generated by the two sub-DACs operating at 100 GS/s. (d) Zoom-in to the 100 GS/s waveform. (e) Spectrum of the 50 GS/s waveform. (f) Spectrum of the 100 GS/s waveform.

The ENOB and SFDR at 100 GS/s and 80 GS/s are shown in Fig. 100 (a). All the frequency components up to half the sampling frequency (50 GHz and 40 GHz, respectively) are included in the SFDR/ENOB calculation. At both conversion rates, the ENOB is 5.3 bit at low frequencies and reduces to 3.2 bit at the maximum signal frequency (24.9 GHz and 19.9 GHz, respectively). The SFDR at both conversion rates is 41 dB and 37 dB at low frequencies near DC, respectively, and reduces to 27 dB at the maximum frequency. To measure a single sub-DAC, one of the two sub-DACs is "switched off" by outputting a constant value. The measured ENOB and SFDR at 50 GS/s and 40 GS/s are shown in Fig. 100 (b). At 40 GS/s, the ENOB and SFDR achieve 5.9 bit and 45 dB at the lowest frequency and reduce to 3.8 bit and 33 dB at the Nyquist frequency (20 GHz), respectively. Both values become lower at 50 GS/s.

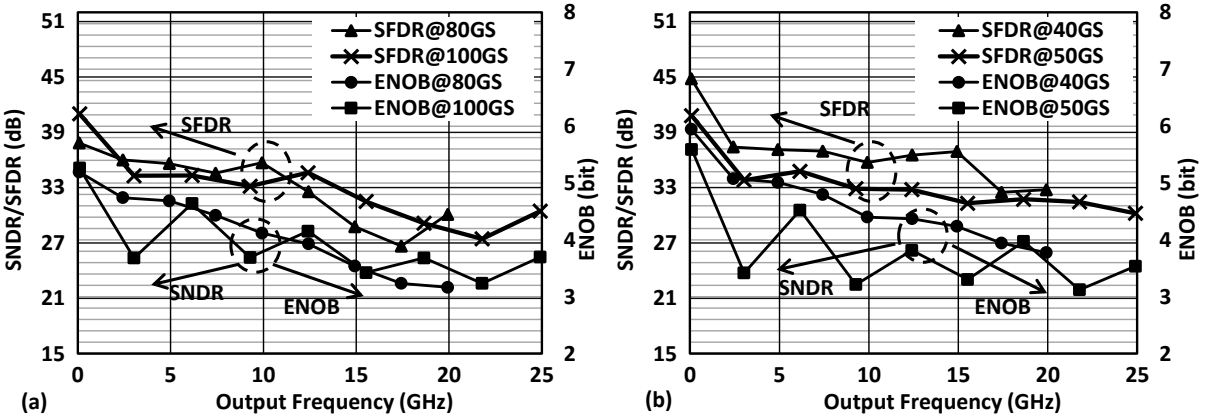


Fig. 100: ENOB and SFDR of (a) whole DAC operating at 80 GS/s and 100 GS/s, and (b) a sub-DAC operating at 40 GS/s and 50 GS/s

As shown in Fig. 101, the 3 dB analog output bandwidth including the sinc amplitude roll-off is 10 GHz and 9.2 GHz at 100 GS/s and 80 GS/s, respectively. Removing the loss of the on-wafer probes and the RF-cable, the bandwidth of the DAC at the output pad is about 13 GHz at 100 GS/s, which is very close to the simulation results with full extracted parasitic elements. The relatively large difference between the schematic and the extracted simulation is due to an underestimation of the parasitic wiring capacitance when designing the artificial transmission line.

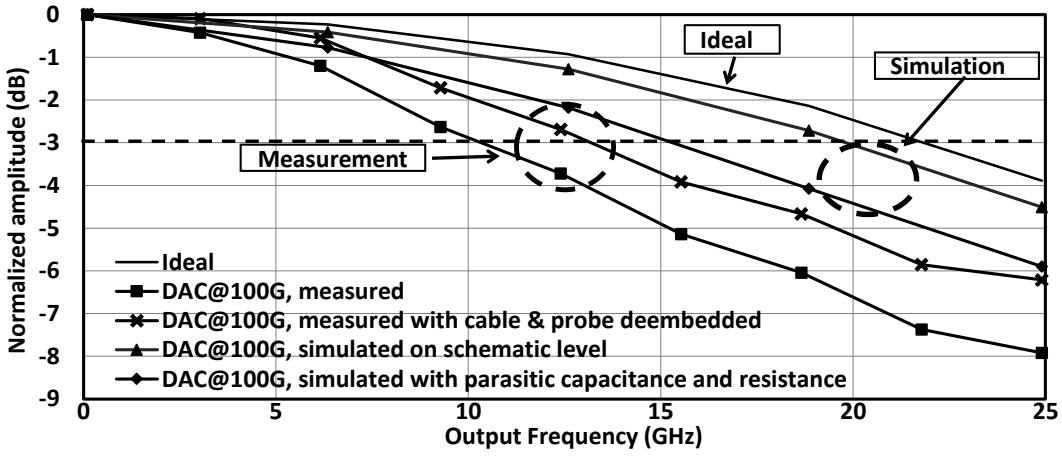


Fig. 101: DAC output frequency response

5.4 Time-Domain Performance

The time-domain characteristics are examined with pulse amplitude modulated signals using only one sub-DAC. Fig. 102 (a) to (c) show the 512-random-symbol pulse-amplitude-modulation with 2 levels (PAM2), PAM4 and PAM8 eyes measured with on-wafer probes without de-emphasis (left) and with 10% digital de-emphasis using one post tap (right) at 40 GS/s. In the case with de-emphasis, the eyes are well opened for all pulse amplitude modulation schemes. At 50 GS/s, a very clearly opened PAM4 eye can be observed in Fig. 102 (d) even with the DAC outputs bonded on the PCB. These measurements show that a transmission rate of up to 120 Gb/s (PAM8 at 40 GS/s) with the DAC test chip is possible.

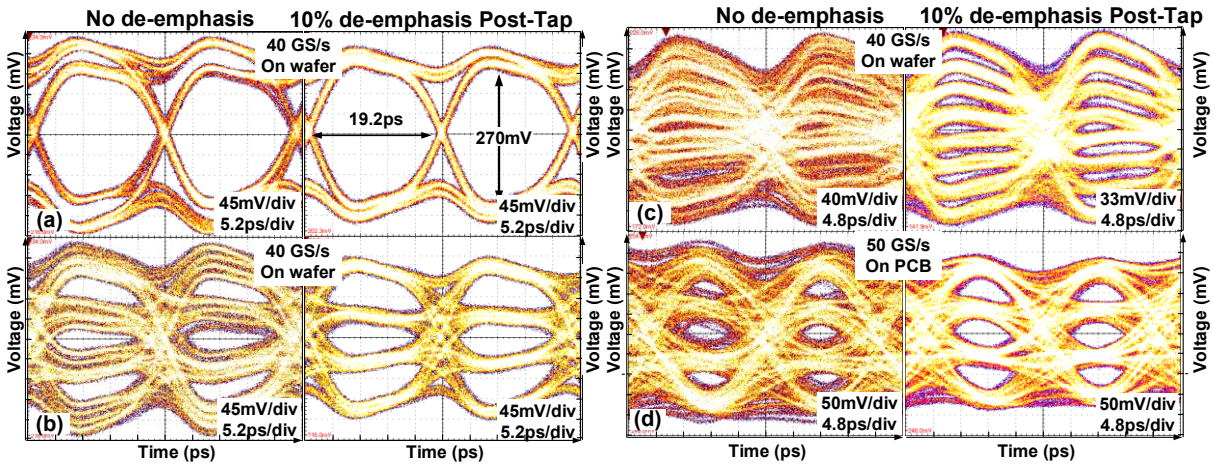


Fig. 102: Eye diagrams of 512-random-symbol PAM2 (a), PAM4 (b) and PAM8 (c) at 40 GS/s and PAM4 (d) at 50°GS/s without de-emphasis (left) and with 10% digital de-emphasis (right) of one sub-DAC (sub-DAC2 outputs zero)

5.5 Optical Transmission Experiment

The DAC is also used in a single-polarization single-carrier optical experiment. The single-ended DAC output is connected to a linear amplifier to drive the optical modulator as shown in Fig. 103 (a). The optical signal is amplified with an optical amplifier and is converted to an electrical signal via a photodiode. The output of the photodiode can be measured either with a 50 GS/s real-time oscilloscope (to collect the transmission data) or with an 80 GHz bandwidth subsampling oscilloscope (to examine the eye diagram of the optical signal received).

In this experiment, only one sub-DAC is used while the other sub-DAC is “switched off”. The conversion rate is limited to 30 GS/s because of the limited sampling frequency of the real-time oscilloscope. The captured eye-diagrams of PAM2, PAM3 and PAM4 with digital de-emphasis at 30 Gbaud using the sampling oscilloscope are shown in Fig. 103 (b). The transmitted data modulated with PAM2, PAM3 and PAM4 without de-emphasis are captured with the real time oscilloscope; the received data is post-processed with a digital filter using Matlab and the bit error rates (BERs) are evaluated. With PAM2 and PAM3, the transmitted data is received correctly, but with PAM4 the received data contains errors, i.e., a transmission rate up to 45 Gb/s (PAM3 at 30 Gbaud) is achieved [64].

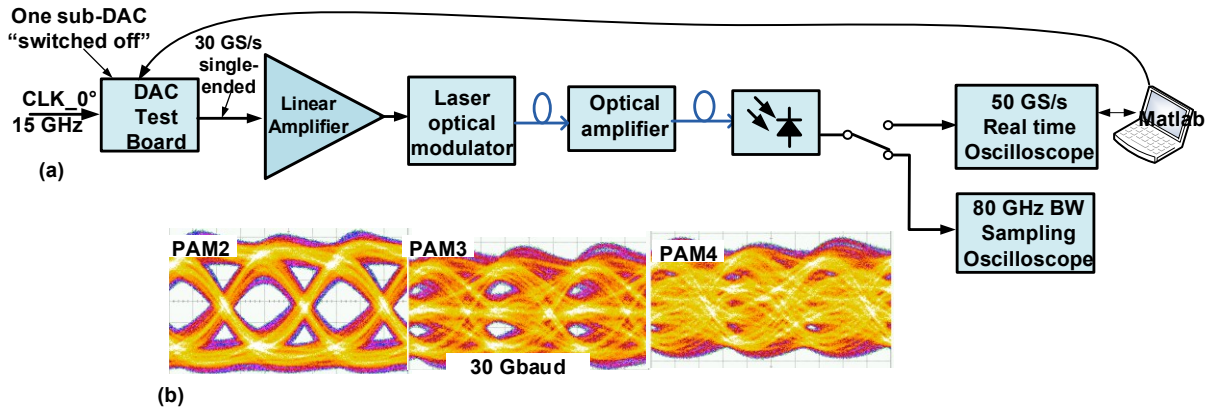


Fig. 103: (a) Measurement setup of electro-optical transmission experiment (b) Optical eye diagrams after receiver photodiode with PAM2, PAM3 and PAM4 at 30 Gbaud

6 Conclusions

6.1 Summary

The rapid growth of the 100/400 Gb fiber optic network has created a high demand for ultra-high-speed transmitters with integrated DACs. These integrated DACs are the key components to enable communications with high-degree modulation, to increase the transmission rate with existing bandwidth. Although recently published DACs have achieved a conversion rate above 60 GS/s, a higher conversation rate and a higher bandwidth are still very much in demand.

At high conversion rates, time-interleaving techniques that reduce the data and clock frequency become attractive. However, the output capacitances are potentially increased by connecting several sub-DACs in parallel, hence reducing the overall output bandwidth. To alleviate this effect, aggressive segmentation and lumped shunt peaking at the DAC output are normally employed [19]. While the shunt peaking is relatively easy to implement, the bandwidth is still limited by the total output capacitance. In this work, the DAC output stage is designed in a distributed style, similar to the distributed amplifiers. The large capacitance at the output node is decomposed into many small unit capacitances and distributed over the output artificial transmission line. As a result, the overall output bandwidth is mainly limited by the small unit capacitance, hence improving the output bandwidth significantly. Nevertheless, the distributed structure exhibits also several disadvantages, which has to be carefully dealt with. For example, the nonlinearity caused by the series resistance of the artificial transmission line is significantly reduced by using a symmetrical layout together with pseudo segmentation, instead of using the traditional segmentation with the thermometer code; the frequency- and code-dependent delay mismatch between the clock and the output transmission lines can be reduced by reducing the code dependency of the output impedance, which can be realized by adding a cascode stage and trickle current to the current cell output.

A combination of the two time-interleaving techniques (analog multiplexing and parallel path) is employed in the proposed DAC, which reduces the maximum data rate at the DAC output stage to 25 Gbit/s for a 100 GS/s conversion rate. While the analog-multiplexing time interleaving increases both the sampling rate and the sinc filter bandwidth, the parallel-path time interleaving only increases the sampling rate without changing the sinc frequency response. Nevertheless, increasing the sampling rate shifts the image to a higher frequency far away from the signal, relaxing the output filter design. The signal band can be extended to a higher frequency without interfering with the neighboring frequency band, hence increasing the utilized bandwidth. The phase alignment between the clock and the data at the DAC output stage is critical and cannot be achieved over PVT with an intrinsic delay matching, even at a reduced data rate due to the time interleaving. Therefore, a phase calibration circuit composed of phase rotators and phase detectors is used to realize the phase alignment. The calibration can be automated with the help of an external digital control. One of the most important advantages of the advanced CMOS technologies is the capability to integrate a large number of transistors.

Therefore, a 1 kByte full-custom-designed memory is integrated on-chip with a minimum requirement of chip area, allowing the DAC to be tested at its maximum speed without external data sources. The memory composed of two memory blocks, which can be synchronized by shifting the memory clock or shifting the data sequence in the memory. The proposed 8 bit 100 GS/s DAC chip is designed and manufactured in a 28 nm low-power CMOS process.

The static and dynamic performances of the DAC chip are characterized by measurements. The INL and the DNL are 2.0 LSB and 1.4 LSB, respectively. The 3 dB bandwidth exceeds 13 GHz at 100 GS/s after removing the losses of the measurement equipment. The ENOB, for frequencies from low frequency near DC up to 24.9 GHz at 100 GS/s, is from 5.3 bit to 3.2 bit, respectively. 100 Gbaud PAM4 and 80 Gbaud PAM8 eye-diagrams are shown. Transmission rates of 45 Gb/s are obtained in an optical transmission experiment. To the best of the author's knowledge, this is the world's first published electric DAC at 100 GS/s.

6.2 Future Works

Although this work has achieved good results regarding high sampling-rate and high bandwidth using time-interleaving techniques and distributed structure, additional work can be done to further improve the bandwidth, linearity, power consumption, etc.

Due to underestimation of the output parasitic capacitances, the delay and the impedance of the artificial transmission line exhibit relatively large deviations to its optimum value, causing a lower output bandwidth than expected. Therefore, a careful extraction of all possible parasitics is required for the future simulations. Further investigation of parasitics reduction is necessary to improve the performance of the artificial transmission line. A smaller capacitance can reduce the required inductance and the series resistance on the transmission line, which not only reduce the attenuation but also reduce the delay mismatch between the input and output line. There are a few possibilities to lower the output capacitance, e.g., adding a cascode stage at the current cell output as mentioned in Section 3.2.2. and using a more compact output stage with fewer segmentations. The inductor design can be improved by using a differential layout style and using a low resistive substrate option. Calibration methods to compensate the gain attenuation over the transmission line are also worth investigating. The current source calibration mentioned in Section 3.2.1 not only improves the linearity, but also reduces the parasitics, since smaller current source transistors can be used. Additionally, calibration methods to compensate the duty-cycle error also have to be considered to reduce the related spurs at the DAC output.

For low cost and high complexity design, the power consumption is one of the most critical parameter, which, however, is not optimized in this work. The most power-consuming blocks are the clock distribution and the data path using CML logics, which also require a large area, introducing more parasitics. Therefore, low-power and high-speed CMOS logics and buffers are needed to replace the CML logics. Because of the 4-fold time-interleaved structure, the DAC output stage consumes 4 times the necessary current. Alternative current multiplexing topologies, such as the CML multiplexers mentioned in Section 0, that require only one current source can be considered to reduce the power consumption.

Appendices

Transmitter Equalization

The transmitter and the transmission media are bandwidth-limited. When the signal travels from the transmitter to the receiver, the high frequency components are attenuated more than the low frequency components, causing distortion. When transmitting a sequence of broadband signals, the distortion causes ISI that degrades noise margins and leads to errors on the receiver side. To deal with ISI, equalization circuits – specifically feed-forward equalizers (FFE) – are widely adopted both at the transmitter and receiver to restore the transmitted bits on the receiver side to reduce the bit-error-rate (BER). With an FFE equalizer, the loss on the channel can be compensated for, either by boosting the high frequency component (pre-emphasis) or attenuating the low frequency component (de-emphasis). Due to the limited amplitude, de-emphasis is normally employed for the transmitters.
FFE can be realized with a finite impulse response (FIR) filter, whose basic structure is shown in Fig. 104. A sequence of input signals x[n] propagates along a delay line composed of delay elements. The delay can be 1 unit-interval (UI) or a half UI or other fractions of a UI [72]. The delayed signals are then scaled by proper coefficients and combined. The tap in the center is referred to as the main-tap; the tap that follows the main-tap is named the post-tap, and the tap that precedes the main tap is called the pre-tap. The weight of each tap is controlled by its corresponding coefficient (C_{-1}, C_0 or C_1). The scaling coefficients can be determined by different techniques, e.g., zero-forcing (ZF) or least-mean-square (LMS) [73]. With the FFE, a higher voltage swing is applied to the signal when a transition occurs; and the signal swing decreases to a lower value after the transition. The high frequency gain is effectively boosted compared to the low frequency gain, resulting in a flatter frequency response at the receiver side due to the loss on the channel. Fig. 105 shows a comparison of the transmitter output waveforms with post-tap on and post-tap off, by outputting a repeated pattern 1111100000

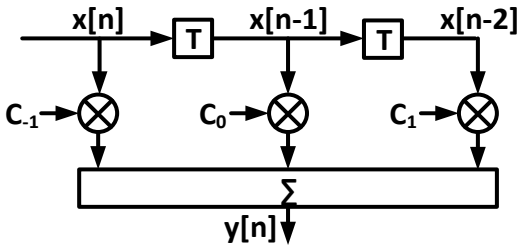


Fig. 104: De-emphasis digital filter using 3-tap FIR filter: the pre-tap, main tap and post-tap

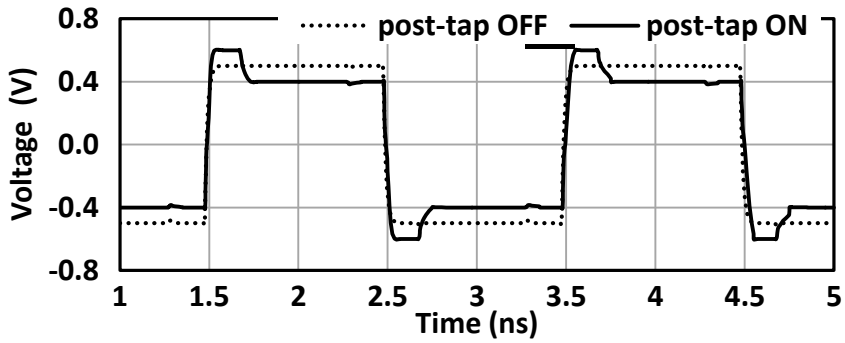


Fig. 105: Waveform of the transmitter output with post-tap on and post-tap off.

Acknowledgments

I will always consider my experience at the Institute of Electrical and Optical Communications Engineering (INT) in the past few years as one of the most wonderful and enjoyable parts of my life.

First of all, I would like to express my gratitude to Prof. Berroth for the opportunity to work in his institute. He has encouraged and guided me through all my research work.

I would also like to thank Prof. Friedel Gerfers for being the co-examiner and for his valuable advices.

I own my thanks to Dr. Grözing, team leader of the IC group at INT, who has introduce me to the world of IC design and measurement. He has always supported and advised me whenever I encountered technical problems. It would not be possible to accomplish the dissertation without his constant encouragement.

I want to say thanks to all my colleagues of the INT for their professional help during my time at the institute.

Thanks all the students whom I supervised for their part of contributions to this work during their thesis work at the institute.

Last, but certainly not the least, I would like to thank my family. The accomplishment of this work would not be possible without them. My parents are the great source of support and encouragement to my research and life. My wife and my children, thank you for your great support.

Bibliography

[1] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2016–2021," 2017. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf.

[2] B. Mukherjee, "WDM optical communication networks: progress and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1810–1824, 2000.

[3] A. H. Gnauck, R. W. Tkach, A. R. Chraplyvy, T. Li, "High-Capacity Optical Transmission Systems," *Journal of Lightwave Technology*, vol. 26, no. 9, pp. 1032–1045, 2008.

[4] C. Laperle, M. O'Sullivan, "Advances in High-Speed DACs, ADCs, and DSP for Optical Coherent Transceivers," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 629–643, 2014.

[5] ITU, "Spectral grids for WDM applications: DWDM frequency grid," 2012. [Online]. Available: https://www.itu.int/rec/T-REC-G.694.1-201202-I/en.

[6] T. Ellermeyer, R. Schmid, A. Bielik, J. Rupeter, M. Möller, "DA and AD converters in SiGe technology: Speed and resolution for ultra high data rate applications," *2010 36th European Conference and Exhibition on Optical Communication. I E E E*, pp. 1–6, 2010.

[7] A. Balteanu, P. Schvan, S. P. Voinigescu, "A 6 bit segmented RZ DAC architecture with up to 50 GHz sampling clock and 4 V_{pp} differential swing," *2012 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pp. 1–3, 2012.

[8] A. Balteanu, P. Schvan, S. P. Voinigescu, "A 6 bit Segmented DAC Architecture With up to 56 GHz Sampling Clock and 6 Vpp Differential Swing," *IEEE Transactions on Microwave Theory and Techniques*, pp. 1–11, 2016.

[9] S. Randel, S. Corteselli, P. J. Winzer, A. Adamiecki, A. Gnauck, S. Chandrasekhar, A. Bielik, L. Altenhain, T. Ellermeyer, U. Dümler, H. Langenhage, R. Schmid, "Generation of a Digitally Shaped 55 GBd 64-QAM Single-Carrier Signal Using Novel High-Speed DACs," *2014 Optical Fiber Communication Conference and Exposition (OFC)*, pp. 1–3, 2014.

[10] A. H. Pawlikiewicz, S. E. Elrai, "RF CMOS or SiGE BiCMOS in RF and Mixed Signal Circuit Design," *2007 14th International Conference on Mixed design of integrated circuits and systems (MIXDES 2007). Politechnika Łódzka. Wydział Elektrotechniki i Elektroniki. Katedra Mikroelektroniki i Technik Informatycznych*, pp. 333–338, op. 2007.

[11] Dedic Ian, "56 Gs/s ADC : Enabling 100 GbE," *Conference on Optical Fiber Communication, collocated National Fiber Optic Engineers Conference (OFC/NFOEC 2010). IEEE*, pp. 1–3, 2010.

[12] M. Nagatani, H. Nosaka, S. Yamanaka, K. Sano, K. Murata, "Ultrahigh-Speed Low-Power DACs Using InP HBTs for Beyond-100-Gb/s/ch Optical Transmission Systems," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 10, pp. 2215–2225, 2011.

[13] Y. Arayashiki, M. Ikeda, Y. Amano, "80 GBd 6 bit DAC in InP DHBT for arbitrary waveform generator," *Electronics Letters*, vol. 52, no. 23, pp. 1937–1938, 2016.

[14] R. Driad, J. Rosenzweig, R. E. Makon, R. Losch, V. Hurm, H. Walcher, M. Schlechtweg, "InP DHBT-Based IC Technology for 100 Gb/s Ethernet," *IEEE Transactions on Electron Devices*, vol. 58, no. 8, pp. 2604–2609, 2011.

[15] F. Ellinger, M. Claus, M. Schroter, C. Carta, "Review of advanced and Beyond CMOS FET technologies for radio frequency circuit design," *2011 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC 2011). IEEE*, pp. 347–351, 2011.

[16] Y. M. Greshishchev, D. Pollex, C. W. Shing, M. Besson, P. Flemeke, S. Szilagyi, J. Aguirre, C. Falt, N. Ben Hamida, R. Gibbins, P. Schvan, "A 56 GS/S 6 b DAC in 65nm CMOS with 256×6 b memory," *2011 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, pp. 194–196, 2011.

[17] Fujitsu, "Digital to Analog Converter- Factsheet LEIA 55-65 GSa/s 8 bit DAC," 2011. [Online]. Available: http://www.fujitsu.com/downloads/MICRO/fme/documentation/c60.pdf.

[18] H. Huang, J. Heilmeyer, M. Grözing, M. Berroth, "An 8 bit 100 GS/s distributed DAC in 28 nm CMOS," *2014 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 65–68, 2014.

[19] J. Cao, D. Cui, A. Nazemi, T. He, G. Li, B. Catli, M. Khanpour, K. Hu, T. Ali, H. Zhang, H. Yu, B. Rhew, S. Sheng, Y. Shim, B. Zhang, A. Momtaz, "A transmitter and receiver for 100 Gb/s coherent networks with integrated 4×64 GS/s 8 b ADCs and DACs in 20 nm CMOS," *2017 IEEE International Solid- State Circuits Conference - (ISSCC). IEEE*, pp. 484–485, 2017.

[20] M. Nagatani, H. Nosaka, K. Sano, K. Murata, K. Kurishima, M. Ida, "A 60 GS/s 6 Bit DAC in 0.5 μm InP HBT Technology for Optical Communications Systems," *IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS), 2011: technical digest. IEEE*, pp. 1–4, 2011.

[21] F. Maloberti, *Data converters.* Springer, Dordrecht, Netherlands 2007.

[22] J. J. Wikner, "Studies on CMOS digital-to-analog converters," Dissertation, Department of Electrical Engineering, Linköpings Universitet, 2001.

[23] J. L. McCreary, P. R. Gray, "All-MOS charge redistribution analog-to-digital conversion techniques. I," *IEEE Journal of Solid-State Circuits*, vol. 10, no. 6, pp. 371–379, 1975.

[24] W. Kester, *Data conversion handbook.* Analog Devices series. Elsevier; Newnes, Amsterdam, London 2005.

[25] M. Pelgrom, "A 50 MHz 10 bit CMOS digital-to-analog converter with 75 Ω buffer," *1990 37th IEEE International Conference on Solid-State Circuits. IEEE*, pp. 200–201, 1990.

[26] D. Mercer, "A study of error sources in current steering digital-to-analog converters," *Proceedings of the IEEE 2004 Custom Integrated Circuits Conference*, pp. 185–190, 2004.

[27] H. L. Chi, van der Goes, Frank M. L., J. R. Westra, J. Mulder, Y. Lin, E. Arslan, E. Ayranci, X. Liu, K. Bult, "A 12 bit 2.9 GS/s DAC with IM3 < -60 dBc Beyond 1 GHz in 65 nm CMOS," *IEEE Journal of Solid State Circuits*, vol. 44, no. 12, pp. 3285–3293, 2009.

[28] A. van den Bosch, M. Steyaert, W. Sansen, "SFDR-bandwidth limitations for high speed high resolution current steering CMOS D/A converters," *ICECS'99. 6th IEEE International Conference on Electronics, Circuits and Systems*, pp. 1193–1196, 1999.

[29] M. J. M. Pelgrom, A. C. J. Duinmaijer, A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid State Circuits*, vol. 24, no. 5, pp. 1433–1439, 1989.

[30] K. R. Lakshmikumar, R. A. Hadaway, M. A. Copeland, "Characterisation and modeling of mismatch in MOS transistors for precision analog design," *IEEE Journal of Solid-State Circuits*, vol. 21, no. 6, pp. 1057–1066, 1986.

[31] C.S.G. Conroy, W. A. Lane, M. A. Moran, K. R. Lakshmikumar, M. A. Copeland, R. A. Hadaway, "Comments, with reply, on 'Characterization and modeling of mismatch in MOS transistors for precision analog design'," *IEEE Journal of Solid-State Circuits*, vol. 23, no. 1, pp. 294–296, 1988.

[32] A. van den Bosch, M. Steyaert, W. Sansen, "An accurate statistical yield model for CMOS current-steering D/A converters," *ISCAS 2000 Geneva. 2000 IEEE International Symposium on Circuits and Systems. Emerging Technologies for the 21st Century. Proceedings*, pp. 105–108, 2000.

[33] J. Bastos, A. M. Marques, M.S.J. Steyaert, W. Sansen, "A 12 bit intrinsic accuracy high-speed CMOS DAC," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 1959–1969, 1998.

[34] J. Vandenbussche, G. van der Plas, A. van den Bosch, W. Daems, G. Gielen, M. Steyaert, W. Sansen, "A 14 b 150 Msample/s update rate Q² random walk CMOS DAC," *1999 IEEE International Solid-State Circuits Conference. Digest of Technical Papers. ISSCC. First Edition*, pp. 146–147, 1999.

[35] Y. Cong, R. L. Geiger, "A 1.5 V 14 bit 100 MS/s self-calibrated DAC," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2051–2060, 2003.

[36] W. Schofield, D. Mercer, L. S. Onge, "A 16 b 400 MS/s DAC with <-80 dBc IMD to 300 MHz and <-160 dBm/Hz noise power spectral density," *IEEE International Solid-State Circuits Conference*, pp. 126–482, 2003.

[37] A. R. Bugeja, B.-S. Song, "A self-trimming 14 b 100 MS/s CMOS DAC," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 12, pp. 1841–1852, 2000.

[38] A. Nazemi, K. Hu, B. Catli, D. Cui, U. Singh, T. He, Z. Huang, B. Zhang, A. Momtaz, J. Cao, "A 36 Gb/s PAM4 transmitter using an 8 b 18 GS/S DAC in 28 nm CMOS," *2015 IEEE International Solid- State Circuits Conference - (ISSCC)*, pp. 1–3, 2015.

[39] H.-H. Chen, J. Lee, J. Weiner, Y.-K. Chen, J.-T. Chen, "A 14 b 150 MS/s CMOS DAC with Digital Background Calibration," *2006 Symposium on VLSI Circuits, 2006. Digest of Technical Papers*, pp. 51–52, 2006.

[40] J. J. Wikner, N. Tan, "Modeling of CMOS digital-to-analog converters for telecommunication," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 5, pp. 489–499, 1999.

[41] B. Razavi, *Design of analog CMOS integrated circuits,* Tata McGraw-Hill ed. Tata McGraw-Hill Pub. Co. Ltd, New Delhi, New York 2010.

[42] B. Schafferer, R. Adams, "A 3 V CMOS 400 mW 14 b 1.4 GS/s DAC for multi-carrier applications," *Solid-State Circuits Conference, 2004. Digest of Technical Papers. ISSCC. 2004 IEEE International*, pp. 360–532, 2004.

[43] S. Park, G. Kim, S. C. Park, W. Kim, "A digital-to-analog converter based on differential-quad switching," *IEEE Journal of Solid State Circuits*, vol. 37, no. 10, pp. 1335–1338, 2002.

[44] E. Gil, K. Shawn, R. Steve, "A 14 b 3/6 GHz current-steering RF DAC in 0.18 µm CMOS with 66 dB ACLR at 2.9 GHz," *2012 IEEE International Solid-State Circuits Conference*, pp. 458–460, 2012.

[45] E. Olieman, A.-J. Annema, B. Nauta, "A 110 mW, 0.04 mm^2, 11 GS/s 9 bit interleaved DAC in 28 nm FDSOI with >50 dB SFDR across Nyquist," *2014 Symposium on VLSI Circuits Digest of Technical Papers*, pp. 1–2, 2014.

[46] D. A. Mercer, "Low-Power Approaches to High-Speed Current-Steering Digital-to-Analog Converters in 0.18 µm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 8, pp. 1688–1698, 2007.

[47] H. L. Chi, K. Bult, "A 10 b, 500 MSample/s CMOS DAC in 0.6 mm^2," *IEEE Journal of Solid State Circuits*, vol. 33, no. 12, pp. 1948–1958, 1998.

[48] E. Olieman, A.-J. Annema, B. Nauta, "An Interleaved Full Nyquist High-Speed DAC Technique," *IEEE Journal of Solid State Circuits*, vol. 50, no. 3, pp. 704–713, 2015.

[49] J. Savoj, A. Abbasfar, A. Amirkhany, M. Jeeradit, B. W. Garlepp, "A 12 GS/s Phase-Calibrated CMOS Digital-to-Analog Converter for Backplane Communications," *IEEE Journal of Solid State Circuits*, vol. 43, no. 5, pp. 1207–1216, 2008.

[50] B. Brandt, D. McMahill, M. Wu, P. Kalthoff, A. Kuckreja, G. Ostrem, "A 14 b 4.6 GS/s RF DAC in 0.18 µm CMOS for cable head-end systems," *2014 IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 390–391, 2014.

[51] T. Alpert, F. Lang, D. Ferenci, M. Grözing, M. Berroth, "A 28 GS/s 6 b pseudo segmented current steering DAC in 90 nm CMOS," *2011 IEEE MTT-S International Microwave Symposium Digest (MTT)*, pp. 1–4, 2011.

[52] J. Deveugele, P. Palmers, M.S.J. Steyaert, "Parallel-path digital-to-analog converters for Nyquist signal generation," *IEEE Journal of Solid State Circuits,*, vol. 39, no. 7, pp. 1073–1082, 2004.

[53] S. van de Frank, L. Nico, D. Filip, H. Zeger, A. Alvin, B. Peter, A. William, S. P. Jeffery, M. Brian, R. Michael, W. Mike, J. Bob, L. Jacky, W. Jake, P. Ken, "A 7.2 GSa/s, 14 Bit or 12 GSa/s, 12 Bit Signal Generator on a Chip in a 165 GHz f_T BiCMOS Process," *IEEE Journal of Solid State Circuits*, vol. 47, no. 4, pp. 1003–1012, 2012.

[54] B. Jewett, J. Liu, K. Poulton, "A 1.2 GS/s 15 b DAC for precision signal generation," *ISSCC. 2005 IEEE International Digest of Technical Papers. Solid-State Circuits Conference, 2005*, pp. 110–587, 2005.

[55] T. H. Lee, *The design of CMOS radio frequency integrated circuits.* Cambridge Univ. Press, Cambridge 1998.

[56] S. Galal, B. Razavi, "40 Gb/s amplifier and ESD protection circuit in 0.18 µm CMOS technology," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2389–2396, 2004.

[57] D. M. Pozar, *Microwave Engineering,* 4th ed. Wiley, Hoboken 2012.

[58] B. Razavi, *Design of integrated circuits for optical communications,* 2nd ed (Online-ausg.). Wiley, Hoboken, N.J 2012.

[59] Y. Cao, R. A. Groves, X. Huang, N. D. Zamdmer, J.-O. Plouchart, R. A. Wachnik, T.-J. King, C. Hu, "Frequency-independent equivalent-circuit model for on-chip spiral inductors," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3, pp. 419–426, 2003.

[60] C. P. Yue, S. S. Wong, "Physical modeling of spiral inductors on silicon," *IEEE Transactions on Electron Devices*, vol. 47, no. 3, pp. 560–568, 2000.

[61] S. Kim, D. P. Neikirk, "Compact equivalent circuit model for the skin effect," *1996 IEEE MTT-S International Microwave Symposium Digest*, pp. 1815–1818, 1996.

[62] K. Myounggon, J. Gil, S. Hyungcheol, "A simple parameter extraction method of spiral on-chip inductors," *IEEE Transactions on Electron Devices*, vol. 52, no. 9, pp. 1976–1981, 2005.

[63] K. Okada, K. Masu, *Modeling of Spiral Inductors.* INTECH Open Access Publisher 2010.

[64] H. Huang, J. Heilmeyer, M. Grozing, M. Berroth, J. Leibrich, W. Rosenkranz, "An 8 bit 100 GS/s Distributed DAC in 28 nm CMOS for Optical Communications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 4, pp. 1211–1218, 2015.

[65] Johannes Heilmeyer, "Distributed D/A converter," Diplomarbeit, University of Stuttgart, 2013.

[66] T. Alpert, M. Werz, F. Lang, D. Ferenci, M. Masini, M. Groezing, M. Berroth, "Arbitrary Waveform Generator Based on FPGA and High-Speed DAC with Real-Time Interface," *2012 8th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME),*, pp. 1–4, 2012.

[67] Yuka Saeki Cardona, "Fast Memory in 65 nm CMOS-Technology," Master Thesis, University of Stuttgart, 2008.

[68] R. J. Backer, *CMOS: Circuit Design, Layout, and Simulation,* 3. ed. IEEE Press series on microelectronic systems. IEEE Press, Piscataway, NJ 2010.

[69] Jonathan Falk, "Entwurf eines automatischen Programmiergerätes für einen 100 GS/s DAC mit On-chip Speicher," Bachelor Thesis, University of Stuttgart, 2014.

[70] Tektronix, "DSA8200 Digital Serial Analyzer Specifications and Performance Verification Technical Reference," 2010. [Online]. Available: https://www.tek.com/oscilloscope/80e06-manual/dsa8200.

[71] IEEE Standards Association, "IEEE Standard for Terminology and Test Methods of Digital-to-Analog Converter Devices," 2011. [Online]. Available: https://ieeexplore.ieee.org/document/6152113.

[72] A. Momtaz, M. M. Green, "An 80 mW 40 Gb/s 7-Tap T /2-Spaced Feed-Forward Equalizer in 65 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 3, pp. 629–639, 2010.

[73] TI, "Equalization Concepts: A Tutorial," 1994. [Online]. Available: http://www.ti.com/lit/an/spra140/spra140.pdf.

List of Publications

[1]. H. Huang, M. Grözing, T. Alpert, M. Berroth "12.5 Gbit/s Configurable Threefold 2:1 MUX and 1:2 DEMUX Chips in 130 nm CMOS Technology," *2010 European Solid-State Circuits Conference (ESSCIRC) Fringe Session*, 2010.

[2]. H. Huang, M. Grözing, J. Digel, D. Ferenci, F. Lang, M. Berroth, "A 6 GS/s 6 bit time interleaved SAR-ADC," *2013 IEEE European Microwave Integrated Circuit Conference (EuMic)*, pp. 37–40, 2013.

[3]. H. Huang, J. Heilmeyer, M. Grözing, M. Berroth, "An 8 bit 100 GS/s distributed DAC in 28 nm CMOS," *2014 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 65–68, 2014.

[4]. H. Huang, J. Heilmeyer, M. Grozing, M. Berroth, J. Leibrich, W. Rosenkranz, "An 8 bit 100 GS/s Distributed DAC in 28 nm CMOS for Optical Communications," *IEEE Transactions on Microwave Theory and Techniques (TMTT)*, vol. 63, no. 4, pp. 1211–1218, 2015.