Broadband Linear Modulator Driver Design for High-data-rate Wireline Communications

by

Bolun Cui

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2025

© Bolun Cui 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Sudip Shekhar Professor, Dept. of Electrical and Computer Engineering, The University of British Columbia
Supervisor(s):	John Long Professor, Dept of Electrical and Computer Engineering, University of Waterloo
Internal Member:	Peter Levine Associate Professor, Dept of Electrical and Computer Engineering, University of Waterloo
	David Nairn Associate Professor, Dept of Electrical and Computer Engineering, University of Waterloo
Internal-External Member:	Adrian Lupascu Associate Professor, Dept. of Physics and Astronomy, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis presents the design, analysis, and implementation of a broadband optical modulator driver targeting high-speed wireline communication systems in a 22-nm FD-SOI CMOS technology. The driver is designed to meet stringent specifications for 15.7-dB gain, 90-GHz bandwidth, linear output swing of 4 $V_{pp,d}$, and minimum energy per bit, with a focus on compensating for high-frequency losses in optical modulators through the driver's 9-dB gain peaking.

The core of the design revolves around a two-stage broadband amplifier, incorporating a novel shunt/double-series interstage network to achieve amplitude peaking at 75 GHz and extended bandwidth. The interstage network is optimized to provide a 4-pole, 1-zero transfer function, enabling efficient inductive peaking and minimizing in-band ripple. The preamp stage features a PFET cascode amplifier with a shunt/double-series input matching network, designed to achieve broadband input matching up to 80 GHz. The postamp stage employs a parallel-path NFET differential cascode amplifier with an output combining network, optimized for more than 15.2-dBm output 1-dB compression point up to 75 GHz.

Experimental results demonstrate the driver's performance across a wide range of operating conditions. The prototype achieves a small-signal bandwidth of over 90 GHz, with a peak gain of 20.4 dB at 76 GHz for a 60-Ω differential load. The driver demonstrates a low-frequency gain of 15 dB and maintains a flat group delay of 17 ps up to 60 GHz, peaking at 30 ps at 76 GHz. The driver's linearity is characterized using the 1-dB compression point (OP_{-1dB}) and total harmonic distortion (THD). The OP_{-1dB} remains above 13 dBm, equivalent to a differential output swing of 4 $V_{pp,d}$, up to 75 GHz across a 100-Ω differential load. The THD is measured at 1.6% for a 1-GHz input signal with a 4-$V_{pp,d}$ output swing.

Time-domain measurements demonstrate the driver's ability to transmit 140-GBaud NRZ signals with a 4-$V_{pp,d}$ eye amplitude. For 100-GBaud PAM-4 signals, the driver achieves a 4-$V_{pp,d}$ outer optical-modulation amplitude, but the top and bottom eyes begin to close due to the driver's frequency response and distortion with 100-Ω differential loads. The use of a 4-tap feedforward equalizer (FFE) compensates for frequency response limitations, resulting in improved eye opening. The driver's performance is compared with other state-of-the-art broadband drivers, highlighting its competitive bandwidth, linearity, and energy efficiency. The prototype occupies a core area of 0.17 mm^2 and a total area of 0.7 mm^2, including pads. The results provide valuable insights into the design and optimization of broadband drivers for high-speed communication systems.

Acknowledgements

I would like to express my utmost gratitude to my supervisor Dr. John Long for offering me the opportunity to continue exploring the world of RF IC, especially in a new direction towards high-speed analog circuits, as well as his invaluable support, encouragement and guidance throughout my PhD studies. I would also like to thank Dr. Peter Levine, Dr. David Nairn, Dr. Adrian Lupascu and Dr. Sudip Shekhar for reading my thesis and providing valuable feedback.

I would like to thank all my colleagues and friends in the RFIC group for their company and support. I would like to thank especially Shaeera and Gio for their technical knowledge, experience, and valuable discussions.

Finally, I would like to thank my parents for their deepest love and unconditional support over the years. I also extend my honest appreciation to my grandparents for their selfless support and dedication.

I wish this thesis could inspire other designers.

Dedication

Dedicated to B. C. in 2009, who decided to take the leap

Dedicated to S. L. in 2024, the rejuvenated euphoria sweetens my steeped tea

Dedicated to the gorgeous idealism

Non sunt multiplicanda entia sine necessitate

Table of Contents

Examining Committee ii

Author's Declaration iii

Abstract iv

Acknowledgements v

Dedication vi

List of Figures xi

List of Tables xx

1 Introduction 1

1.1 Overview of DP-QAM transmitters 2

1.2 Design Objectives and Challenges 5

1.2.1 Target Data Rate and Electrical Bandwidth 5

1.2.2 Gain, Peaking amplitude, Input/Output Swing and Linearity 7

1.2.3 Return Loss 9

1.2.4 Other Objectives 10

1.3 Thesis Organization 11

2 Literature Survey of Driver Amplifiers **12**
2.1 Previous Work on Driver Amplifiers . . . 12
2.2 Wide-Swing Linear Driver Amplifiers . . . 13
2.2.1 Cascoding . . . 14
2.2.2 Feedforward breakdown-voltage doubling . . . 16
2.2.3 Feedforward linearization . . . 17
2.2.4 Summary of wide-swing techniques and selected driver examples . . . 18
2.3 Driver-MZM Co-design . . . 20
2.4 Performance Summary . . . 21

3 Interstage Network Design **23**
3.1 A 7-pole 2-zero, shunt/double-series interstage network . . . 23
3.1.1 Analysis of the 4-pole, 1-zero, shunt/double-series interstage network 25
3.1.2 Interstage design example for the driver . . . 41

4 Broadband Driver Design **52**
4.1 Two-stage broadband amplifier with shunt/double-series interstage . . . 52
4.1.1 Input matching network and preamp stage . . . 53
4.1.2 Postamp stage . . . 55
4.1.3 Shunt/double-series interstage . . . 69
4.1.4 Postamp linearization . . . 76

5 Physical Layout and Implementation **80**
5.1 General considerations for MOSFET physical layout in high-power designs 80
5.1.1 Electromigration considerations . . . 81
5.2 Passive layout . . . 82
5.2.1 Input matching network . . . 82
5.2.2 Output combining network . . . 84

5.3 Metal fill and parasitic optimization . . . 85
5.4 Peak gain and peak frequency over process, voltage, temperature variations 86
5.4.1 Peak gain and peak frequency against process, voltage and temperature variations . . . 86
5.4.2 Effects of supply voltage variations (at 27 °C and 'typical' process, MC disabled) . . . 87
5.4.3 Effects of temperature variations (at nominal voltage and 'typical' process, MC disabled) . . . 88
5.4.4 Effects of process variations for various FET corners (at nominal voltage and 27 °C, MC disabled) . . . 89
5.4.5 Effects of global process variations (at nominal voltage and 27 °C, MC enabled) . . . 90
5.4.6 Aggregate effects of process, voltage and temperature variations . . . 91

6 Experimental Results 94
6.1 Calibration and measurement setup . . . 96
6.2 DC bias measurement results . . . 97
6.3 Small-signal measurement results . . . 101
6.4 Large-signal measurement results . . . 108
6.5 Time-domain measurement results . . . 112
6.5.1 Test pattern: PRBS-15 100-/120-/140-GBaud NRZ . . . 116
6.5.2 Test pattern: PRBS-15 50-/100-GBaud PAM-4 . . . 120
6.6 Performance Summary and Comparison . . . 124

7 Summary and Conclusions 128
7.1 Conclusions . . . 128
7.2 Major Contributions . . . 129
7.2.1 Amplitude-peaking filter networks . . . 129
7.2.2 Linear high-swing optical modulator drivers . . . 130
7.2.3 Power-efficient linearizer . . . 131

7.3 Future Work . . . 131
7.3.1 Variable-gain amplifier and multi-stage driver designs . . . 132
7.3.2 Compensation of thermal effects . . . 133
7.3.3 Custom high-Q capacitor design . . . 134

References **136**

Appendices **146**

A The effects of finite Q-factor of series coils in a shunt/double-series network 147
B Analysis of the 4-pole, double-series interstage network . . . 148
B.1 Summary of the 4-pole, double-series interstage design methodology 158
C Target specification for MPWs . . . 159
D Other small-signal measurement results . . . 160
E Measured THD at 10 GHz and 25 GHz . . . 161
F Simulation results of the driver . . . 163

List of Figures

1.1 Simplified block digram of an optical DP-QAM system [1] 2

1.2 Normalized electro-optic frequency response masks for drivers with various symbol rates proposed by the HB-CDM 2.0 standard (NRZ signal assumed; BW_{up} and BW_{lo} represent the upper and lower mask of -3dB bandwidth, respectively) [1] 6

1.3 Differential electrical input reflection coefficient for drivers with various symbol rates proposed by the HB-CDM 2.0 standard ($S_{11,up}$ represent the mask of differential input return loss) [1] 9

1.4 A simplified block diagram of a two-arm, travelling-wave MZM with an differential, electrical driver and far-end termination [2] 10

2.1 Simplified schematics of (a) conventional cascode driver, (b) with a thick-oxide cascode device and (c) with an R-C compensation network at the gate of the cascode [3] 14

2.2 Simplified CG FET with drain-gate feedback resistor to spread voltage stress at low frequencies [4] 15

2.3 Simplified schematic of a breakdown-voltage doubler [5] 16

2.4 Block diagram of a 2-way feedforward linearization network [6] 17

2.5 Simplified schematic of a driver output stage in [4] 19

2.6 Simplified diagram of a multi-electrode MZM (2 segments) with integrated driver [7] 20

3.1 Simplified schematic of (a) the proposed shunt/double-series interstage network and (b) the lumped-element equivalent model 24

3.2 Simplifications of the 7-pole, 2-zero circuit (a) 4-pole, double-series interstage network, (b) 4-pole, 1-zero, shunt/double-series interstage network . . 24

3.3 Pole-zero plot of the 4-pole, 1-zero, shunt/double-series network (zero neglected) with 3 cases of p_1 and p_2 highlighted (Case 1: $\sigma_1 = 2.164$, Case 2: $\sigma_1 = 1.08$, Case 3: $\sigma_1 = 0.26$) 31

3.4 Normalized magnitude response of 4-pole, 1-zero, shunt/double-series networks for various σ_1 (Case 1: $\sigma_1 = 2.079$, Case 2: $\sigma_1 = 1.038$, Case 3: $\sigma_1 = 0.13$) 32

3.5 Normalized bandwidth and peak amplitude against various σ_z for the 4-pole, 1-zero, shunt/double-series network design example (Assumptions: $\omega_{pk} = \omega_3 = 2.6$, $\zeta_3 = 0.05$ and $\sigma_1 = 1.038$. Bandwidth is normalized to 1 rad/s) . 34

3.6 Shunt/double-series network implemented with coupled coils m_1' of 1.77 and m_3' of 0.438 with coupling factor k_{13} of 0.17 for a capacitor ratio $n = C_2/(C_1 + C_2)$ of 0.9 (equivalent to Fig. 3.2b with m_1, m_2 and m_3 of 1.92, -0.15 and 0.588, respectively) 37

3.7 Normalized magnitude response of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$) 37

3.8 Phase response of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$) . 39

3.9 Normalized group delay of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$) 39

3.10 Normalized step response of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$) 40

3.11 Single-ended lumped-element model of the shunt/double-series interstage network 43

3.12 Magnitude responses of the two second-order transfer functions from Eq. 3.33. Parameter $c = 2.02$ is used to set the normalized peak frequency ω_{pk}' while parameter b is 0 ($p_{1,2} = -0.4 \pm j0.6$, $p_{3,4} = \pm j1.42$). The real zero is used to set the DC gain to zero ($z_1 = -0.807$). 45

3.13 Magnitude responses of the two second-order transfer functions from Eq. 3.33. Parameter $c = 2.02$ is used to set the normalized peak frequency ω'_{pk} while parameter b is 2.44×10^{-2} ($p_{1,2} = -0.196 \pm j0.598$, $p_{3,4} = -0.0122 \pm j1.42$). The zero is used to set the DC gain to unity ($z_1 = -0.8$). 47

3.14 Magnitude responses of the two second-order transfer functions from Eq. 3.33. Parameter $c = 2.02$ is used to set the normalized peak frequency ω'_{pk} while parameter b is 2.44×10^{-2} ($p_{1,2} = -0.196 \pm j0.598$, $p_{3,4} = -0.0122 \pm j1.42$). The zero is used to tune A_{pk} ($z_1 = -0.633$). 48

4.1 Simplified block diagram of the broadband driver 52

4.2 Input stage of the two-stage broadband driver with a shunt/double-series input matching network 53

4.3 Simulated reverse isolation against various neutralization capacitances . . . 54

4.4 Simulated differential transconductance and input return loss of the input stage (with input matching network and preamp stage included, output pole excluded) 55

4.5 Postamp stage of the two-stage broadband driver 56

4.6 Simulated differential voltage gain of the postamp stage (with 60-Ω differential load and input pole excluded) 57

4.7 Threshold voltage of NFETs versus body-to-source voltage ($V_{GS} = 0.45$ V, $V_{DS} = 0.8$ V. NFET width of 160 μm with 2 μm/finger, length of 18 nm.) . 58

4.8 Bridging capacitor C_{k3} realizes common-gate compensation in a differential implementation (example of the M_3 common-gate stage from Fig. 4.5) . . . 60

4.9 Time-domain pulse response of a cascode device for various compensation resistance value for R_{gs} (Tested with schematic from Fig. 4.8. Signal source is a 1-$V_{pp,d}$ pulse with rise/fall time of 0.1 ns. Source and load resistances are 50 and 60 Ω, respectively. Output voltage is taken at v_{o+}) 61

4.10 Lumped-element model of the output combining network 63

4.11 A simplified model to show the effect of C_{ds} on R_{out} for a cascode device . 63

4.12 Real part of the output impedance from Fig. 4.11 with a variable C_{ds} in a cascode amplifier ($C_d = C_{ds}$, $r_d = r_{ds} = 55$ Ω, $g_m = 0.33$ S) 65

4.13 A simplified model to show the effect of C_k and C_{dg} on C_{out} for a cascode device 65

4.14 A simplified model to show the effect of C_{dg} and L_d on R_{in} for a common-source amplifier (neglecting C_{ds} and r_{ds}) 67

4.15 Single-ended lumped-element model of the shunt/double-series interstage network, including series ohmic loss of L_1 and L_2 for a Q-factor of 25 (r_{s1} and r_{s2}), series loss of L_3 for a Q-factor of 12, and shunt-equivalent R-C network of the postamp input at 75 GHz (R_2 and C_2) 70

4.16 Equivalent differential lumped-element model of the shunt/double-series interstage network at 75 GHz (magnetic coupling factors: • $k_{11} = 0.35$, ∘ $k_{22} = 0.21$, ★ $k_{33} = 0.28$, ▲■ $k_{23} = 0.26$, △□ $k_{32} = 0.05$). Capacitors C_{1d} to C_{4d} are the differential equivalent capacitors (i.e., half of C_1 to C_4, respectively) 72

4.17 The differential interstage network shown in Fig. 4.16 73

4.18 Frequency responses of the shunt/double-series interstage network loaded by the postamp, with and without the tuning capacitors C_3 and C_4 in Fig. 4.16 75

4.19 Schematic of a linearized postamp utilizing the IM3 cancellation technique 77

4.20 Simulated OP_{-1dB} against various control voltages and centre frequencies (V_C sets the conduction angle of the auxiliary amplifier) 77

4.21 Simulated IM3 against various control voltages and centre frequencies in a two-tone test (input power is -25 dBm and tone spacing is 10 MHz) 78

5.1 Sample FET layout for an 80-μm NFET in Fig. 4.5 (each finger is 2-μm wide and 18-nm long, and 40 fingers are split into upper and lower blocks) 81

5.2 Layout of the input matching network of Fig. 4.2 83

5.3 Physical layout for the output combining network of Fig. 4.10 (outer ground pads adjacent to each RF pad not shown) 84

5.4 RF signal pad with manual metal filling pattern ($M_4 - M_7$ ground plane not shown) 85

5.5 Changes in A_{pk} and f_{pk} against ±5% shifts in supply voltages V_{SSP} and V_{DDN} (Voltage normalizing factor = $V_{DDN}/3 = V_{SSP}/2.97$; Process: TT; Temperature: 27 °C) 87

5.6 Changes in A_{pk} and f_{pk} against various temperatures from -40 °C to +110 °C (Process: TT; Voltage: V_{SSP} of 2.97 V, V_{DDN} of 3 V) 88

5.7 Changes in A_{pk} and f_{pk} against various FET process corners (e.g., SF corner represents slow NFET and fast PFET; Temperature: 27 °C; Voltage: V_{SSP} of 2.97 V, V_{DDN} of 3 V) 89
5.8 Distribution of A_{pk} and f_{pk} resulted from Monte Carlo simulations for process variations (1000 samples in total. Voltage: V_{SSP} of 2.97 V, V_{DDN} of 3 V; Temperature: 27 °C) 90
5.9 Simulations for A_{pk} and f_{pk} against process, voltage and temperature variations (12000 samples in total), minimum occurrence of 100 are shown (Process: randomly distributed, MC-enabled; Voltage: [$|V_{SSP}|$, $|V_{DDN}|$] pairs of [2.82 V, 2.85 V], [2.97 V, 3 V], and [3.12 V, 3.15 V]; Temperature: 0, 27, 60, and 110 °C) 91
5.10 Simulations for A_{pk} and f_{pk} against process, voltage and temperature variations (12000 samples in total), minimum occurrence of 300 are shown (Process: randomly distributed, MC-enabled; Voltage: [$|V_{SSP}|$, $|V_{DDN}|$] pairs of [2.82 V, 2.85 V], [2.97 V, 3 V], and [3.12 V, 3.15 V]; Temperature: 0, 27, 60, and 110 °C) 92
5.11 Distribution of A_{pk} and f_{pk} against process, voltage and temperature variations (12000 samples in total. Process: randomly distributed, MC-enabled; Voltage: [$|V_{SSP}|$, $|V_{DDN}|$] pairs of [2.82 V, 2.85 V], [2.97 V, 3 V], and [3.12 V, 3.15 V]; Temperature: 0, 27, 60, and 110 °C) 93

6.1 Micrograph of the two-stage broadband driver prototype 95
6.2 110-GHz, 4-port frequency-domain measurement setup 95
6.3 Time-domain measurement setup 96
6.4 On-chip probing and back-side temperature control setup 97
6.5 Measured DC bias current for the postamp against various TEC controller temperature settings. Simulated DC bias current is also plotted for the same range of device temperature using Spectre simulations 98
6.6 Measured on-chip termination resistance at the driver input against various TEC controller temperature settings. Calculated resistance is also plotted assuming that the TEC temperature equals circuit temperature 100
6.7 Measured small-signal differential forward gain (S_{21d}) versus various TEC temperature settings for 100-Ω differential source and load impedances. The simulated result for temperature settings of 27°C and 60°C in Spectre are also plotted for comparison 101

6.8 Simulated interstage load resistance and transconductance of a 80-μm NFET at 27°C and 60°C . . . 102

6.9 Measured small-signal differential forward gain (S_{21d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for temperature settings of 27°C and 60°C in Spectre are also plotted for comparison . . . 103

6.10 Measured small-signal differential input reflection coefficient (S_{11d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 27°C in Spectre is also plotted for comparison . . . 105

6.11 Measured small-signal differential stability factor, μ versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison . . . 106

6.12 Measured small-signal differential group delay (GD_{21d}) versus various TEC temperature settings for 100-Ω source and 100-Ω load (differential) impedances. The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison . . . 107

6.13 Measured small-signal differential group delay (GD_{21d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load (differential) impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison . . . 107

6.14 Measured large-signal differential output -1dB compression point (OP_{-1dB}) versus various TEC temperature settings for a 100-Ω differential load . . . 109

6.15 Measured large-signal differential output -1dB compression point (OP_{-1dB}) at 75 GHz for a 100-Ω differential load versus various body-to-source voltages for postamp FETs at a TEC temperature setting of 7°C . . . 110

6.16 Measured gain compression and THD at the driver output, for a 1-GHz sinusoidal input signal and differential 100-Ω load at three TEC temperature settings . . . 111

6.17 Measured large-signal differential output -1dB compression point (OP_{-1dB}) versus frequency for a 100-Ω differential load at a chip backside temperature of 7°C ($V_{BS} = 2$ V, postamp bias current of 151 mA) 112

6.18 Measured single-ended 100-GBaud, 120-GBaud, and 140-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.8, 0.5, and 0.28, respectively. A thru replaces the DUT connected between the AWG (including the input network) and sampling oscilloscope (including the output network) for these measurements 113

6.19 Measured differential 100-GBaud PAM-4 eye pattern with raised-cosine roll-off factors (α) of 1 for a thru interconnection as the DUT connected between the AWG (including the input channel) and sampling oscilloscope (including the output channel) 116

6.20 Measured single-ended 100-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.8 for the driver for three chip backside temperatures . . 117

6.21 Measured single-ended 120-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.5 for the driver at three chip backside temperatures . . 118

6.22 Measured single-ended 140-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.28 for the driver at three chip backside temperatures . . 119

6.23 Measured single-ended and differential 50-GBaud PAM-4 eye patterns with raised-cosine roll-off factors (α) of 1 for the driver at three chip backside temperatures (worst-case eye height and width are labelled) 120

6.24 Measured single-ended and differential 100-GBaud PAM-4 eye patterns for the driver at three chip temperatures, differential output swings, and raised-cosine roll-off factors (worst-case eye height and width are labelled) 122

7.1 Block diagram of a multi-stage driver with EQ and gain control 132

7.2 Cross-sectional views of various capacitor structures (P and N are the positive and negative polarities of capacitors) 135

A.1 Peak amplitude from the dominant second-order transfer function (Eq. 3.50) and roll-off from the non-dominant second-order transfer function (Eq. 3.51) against various Q-factor of the series coils at a normalized frequency of 1.42 148

B.2 Pole-zero plot of the 4-pole, double-series network with 3 cases of p_1 and p_2 highlighted (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$) 152

B.3 Normalized magnitude response of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$) 153
B.4 Phase response of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$) 154
B.5 Normalized group delay of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$) 155
B.6 Normalized step response of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$) 156
D.7 Measured small-signal differential output reflection coefficient (S_{22d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison 160
D.8 Measured small-signal differential reverse isolation (S_{12d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison 161
E.9 Measured gain compression and THD at the driver output, for a 10-GHz sinusoidal input signal and differential 100-Ω load at three TEC temperature settings 162
E.10 Measured gain compression point and total harmonic distortion at the output for a 25-GHz sinusoidal input signal and differential 100-Ω load at various TEC temperature settings 162
F.11 Simulated small-signal differential forward gain ($|S_{DD21}|$) and reverse isolation ($|S_{DD12}|$) for 60-Ω (solid) and 100-Ω (dashed) loads 164
F.12 Simulated small-signal differential input ($|S_{DD11}|$) and output return losses ($|S_{DD22}|$) for 60-Ω (solid) and 100-Ω (dashed) loads 165
F.13 Simulated small-signal common-mode forward gain ($|S_{CC21}|$) and common-mode rejection ratio (CMRR) for 60-Ω (solid) and 100-Ω (dashed) loads . . 165
F.14 Simulated small-signal differential and common-mode μ-stability factor 60-Ω (solid) and 100-Ω (dashed) loads 166
F.15 Simulated differential output and input -1dB compression point 60-Ω (solid) and 100-Ω (dashed) loads 166

F.16 Simulated differential total harmonic distortion 60-Ω (solid) and 100-Ω (dashed) loads 167
F.17 Simulated differential small-signal forward gain and output -1dB compression for 60-Ω load against various postamp back-gate biasing 167
F.18 Simulated differential total harmonic distortion for 60-Ω load against various postamp back-gate biasing 168
F.19 Simulated DC power consumption of the postamp and overall driver against various postamp back-gate biasing 168

List of Tables

2.1 Typical performance of state-of-the-art MZM drivers (recent 10 years) . . . 22

3.1 Design goals for a 4-pole, 1-zero, shunt/double-series network design example 30

6.1 Performance summary and comparison of high-swing, linear modulator drivers 127

B.1 Design goals for a 4-pole, double-series network design example 151

C.3 Design specifications used in various prototype tape-outs 159

F.4 Simulated results of the two-stage broadband amplifier with shunt/double-series interstage 163

Chapter 1

Introduction

The past decade has seen a steady increase in high-speed internet subscriptions and data traffic in Canada due to the emergence of new applications [8]. For example, 75% of the subscriptions are for speeds of >50 and >10 Mbps for download and upload, respectively. Most subscriptions offer download speeds ranging from 100 to 300 Mbps. The average monthly download has increased from 205 GB in the first quarter of 2019 to 395 GB in the third quarter of 2022.

Video streaming services such as YouTube Live® and Twitch® have enabled people to broadcast in real time, making everyone a potential content creator on the internet. Recent 5G network deployment (5G coverage in Canada has reached 88% in 2021 [9]) has allowed wireless streaming on mobile devices with one-tenth of the latency compared to 4G [10]. Short video clips have become the main form of information posted on social networks like TikTok® and Reels®. Cloud based storage, computing, and entertainment services, such as Google Drive®, Amazon Web Services®, and Microsoft XBOX® Game Pass®, have eliminated the need for a powerful local computing device while still providing quality services. Applications powered by generative, pre-trained transformer (GPT) based natural language models (e.g., Writefull® TexGPT, Microsoft Bing® with GPT) are disrupting traditional text writing and web searching.

None of these emerging applications would be possible without data centres as the backbone of data infrastructures. To ensure reliable and high-speed communication, inter- and intra-data centre interconnects are built using fibre-optic links. Currently, there is intensive research and development of wireline systems capable of handling >100-Gb/s throughput, which is the primary application domain for the work presented in this thesis.

1.1 Overview of DP-QAM transmitters

Birefringence is an optical property of a material that causes its refractive index to vary depending on the polarization and direction of the incident light [11]. An electric field can be applied to adjust the birefringence of optical media, a phenomenon known as the Pockels effect [12]. This allows two lightwaves with parallel and perpendicular polarizations (i.e., X and Y polarizations, respectively) to be transmitted together in one medium without interference [11]. As a result, dual-polarization (DP) multiplexing doubles the data rate of a single-polarized modulated signal.

Quadrature-amplitude modulation (QAM) is widely deployed in wireless communication systems, and is being considered as a replacement for non-return-to-zero (NRZ) and pulse-amplitude modulation (PAM) in wireline systems due to its ability to transmit more bits per symbol. Current systems operating at 64-GBaud DP-4QAM [11] offer a maximum theoretical data rate of 256 Gb/s. DP-16QAM and DP-64QAM are potential options for the next generation.

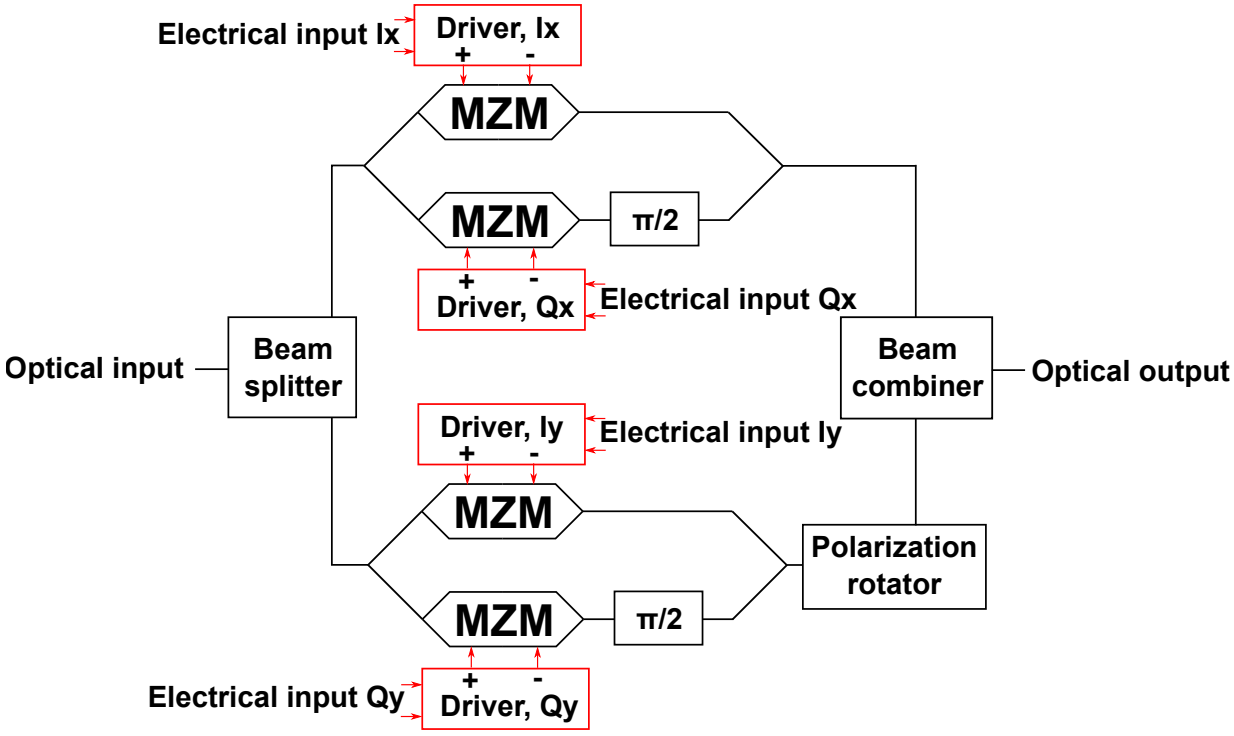


Figure 1.1: Simplified block digram of an optical DP-QAM system [1]

A simplified block diagram of a DP-QAM system is depicted in Fig. 1.1. The optical input from a laser is divided into four branches, each of which is fed to a Mach-Zehnder modulator (MZM). Electrical input signals I_x, Q_x, I_y and Q_y represent differential, quadrature-amplitude modulated (I-Q) data. The I-Q signals are amplified by the modulator driver and used to drive the MZM. The $\pi/2$ block realizes a 90-degree phase shift between the I and Q paths. A polarization rotator inserted into one branch orthogonalizes X and Y polarizations. Finally, the beams are combined and coupled to the optical fibre for transmission. This work focuses on the design of the electrical driver (shown in red). In the systems described above, DP-QAM is implemented in the optical domain. By using modulation schemes such as PAM4 with the electrical driver, the data rate may be doubled again (e.g., a 64-GBaud system with optical DP-QAM driven by electrical drivers with PAM4 signals may reach 512 Gb/s).

Wireless communication utilizes frequency-division multiplexing (FDM) to further boost data transmission rates. This involves dividing the total bandwidth into multiple, non-overlapping channels, each of which is used to transmit a distinct signal. Wavelength-division multiplexing (WDM) is available for optical communication. WDM combines multiple colours of light from various channels into a single beam, which is sent through a fibre-optic waveguide. The light output of a DP-QAM system is transmitted through optical fibre at wavelengths from 1270 nm to 1610 nm with a 20-nm spacing between each band [13]. These wavelengths are chosen because attenuation for single-mode optical fibres is minimal in this range [14], and therefore optical amplifiers are not required. The HB-CDM 2.0 standard specifies the use of the C-band (1530-1565 nm or 191.35-196.2 THz) and the L-band (1565-1625 nm or 186-191.5 THz) for transmission [1]. Cisco® demonstrated an optical module that multiplexes input signals of 60 GBaud and achieves 400 Gb/s with DP-16QAM and WDM [15]. The 800G WDM module offered by Eoptolink® is capable of transmitting at 850 Gb/s across 2 km over single-mode fibre [16].

The optical components of the DP-QAM system depicted in Fig. 1.1 have the following specifications. The peak optical input power specified by HB-CDM 2.0 is 18 dBm, and the maximum insertion loss per polarization (X/Y) is 16 dB [1]. The required optical input and output return losses are 27 dB. The extinction ratio (ER) is the power ratio between an optical '1' and '0'. The typical ER of an individual MZM and the output of X/Y polarization are 25 and 22 dB, respectively. The polarization extinction ratio (PER) defines the ratio of power transmitted by the device when the polarization axes are aligned compared to the condition when the axes are crossed (i.e., cross-talk between X and Y polarization). The typical PER is set to 20 dB for HB-CDM 2.0. The I/Q skew represents the timing difference between the I and Q paths. The maximum I/Q skew for HB-CDM 2.0 is 50 ps with 1.5-ps variation (within X/Y polarization) and a total skew of 100 ps with

4-ps variation at max.

Silicon photonic (SiP) MZMs offer the potential benefit of integrating the optical modulator and electrical driver, which is an advantage over lithium-niobate MZM alternatives [12]. A recent study of SiP MZMs has revealed a -3dB bandwidth of 38 GHz and a half-wave voltage (i.e., V_π, the bias voltage required to attain 180-degree phase shift at the MZM output) of greater than 8 V for a DP-MZM with 3-mm traveling-wave MZM arms [2]. Due to the high V_π, it is common to drive at a fraction of V_π to remain in the linear operating range of SiP MZMs [17], which can lead to a decrease in the ER of PAM-4 levels. For instance, when the driving range is reduced to $0.5 \times V_\pi$, the ER may be reduced by up to 1 dB [2].

In Samani [2], a DP-MZM with a bandwidth of 38 GHz was demonstrated to transmit PAM-4 signals up to 56 GBaud while keeping the bit-error rate (BER) below the hard-decision, forward-error-correction (HD-FEC) threshold of 3.8×10^{-3}. Since the target symbol rate for this work is 130 GBaud, the MZM bandwidth is the limiting factor of the electro-optical frequency response. An objective of this work is to design an electrical driver with a peak in its frequency response to realize a higher electro-optical bandwidth overall by compensating for the modulator response.

The objectives of this project encompass the design and benchmarking of an integrated, differential Mach-Zehnder modulator driver prototype. This prototype is intended for utilization in coherent optical communication systems and is implemented in FD-SOI CMOS technology. The driver's electrical response will be measured for prototype validation. Electro-optical testing is outside the scope of this work.

Design priorities and the chip fabrication technology (22-nm FD-SOI CMOS) were selected after discussion with an industrial partner:

- Bandwidth exceeding 90 GHz;
- Differential peak-to-peak output voltage swing exceeding 4 V for a 60-Ω differential load;
- Output-referred 1-dB compression point surpassing 15.2 dBm (i.e., 4-V linear output swing across a 60-Ω differential load);
- DC power consumption below 650 mW.

The 22-nm fully-depleted silicon-on-insulator (FD-SOI) CMOS used for technology benchmarking is 22FDX® [18] manufactured by GlobalFoundries®. This technology features a maximum transistor transition frequency (f_T) of 270 GHz, making it suitable

for designing wideband circuits operating at RF and millimetre-wave (mm-wave) frequencies [19]. For the back end of line (BEOL), this technology offers four layers of thick metal (3 copper and 1 aluminum, with stack option 19), facilitating the design of low-loss passive components on chip.

Compared to bulk CMOS technologies, FD-SOI exhibits a higher breakdown voltage in the well, and lower junction capacitance due to the ultra-thin buried oxide layer (BOX) beneath the device [18]. This feature allows integration of power amplifiers at mm-wave frequencies by cascoding transistors [20]. The FD-SOI transistors in this technology employ a dual-gate structure with front and back gates (note that the back gate operates at a lower speed). The ability to apply back bias to these transistors allows trade-offs between performance and power consumption in real time, making the overall system more adaptable.

1.2 Design Objectives and Challenges

The design of broadband CMOS modulator drivers come with many challenges. Typical specifications for a driver circuit consist of target data rate, electrical bandwidth, input/output swing, gain, peaking amplitude, linearity (i.e., compression point), return losses, DC power consumption, and system supply voltage. The primary objective of this work is to make a 'best effort' in achieving the highest possible electrical bandwidth for an optical modulator driver, with a minimum bandwidth target of 90 GHz.

1.2.1 Target Data Rate and Electrical Bandwidth

Current commercial systems (e.g., by Cisco® [15] and by Eoptolink® [16]) operate at a symbol rate of 64 GBaud, enabling data rates ranging from 400 to 800 Gb/s. There are three primary methods to achieve such data throughput. The first method involves paralleling multiple 100-Gb/s links in a single package (e.g., aggregating 8x100-Gb/s lanes to achieve 800 Gb/s) [11]. Another method employs modulation schemes beyond on-off keying (OOK), which allows for more bits per symbol (e.g., using 4-level pulse-amplitude modulation, PAM4). Lastly, multiplexing (e.g., WDM) enables multiple signals to be transmitted simultaneously. Consequently, all three approaches can achieve higher data throughput without increasing the electrical bandwidth of the driver, which amplifies the input data signal and controls the optical modulator. To further enhance speed, advanced modulation schemes like 16-QAM can be deployed in a polarization-multiplexed, coherent optical system as depicted in Fig. 1.1 [11].

The prototype modulator driver envisioned in this thesis aims to support a per-channel symbol rate of 130 GBaud, and is intended for deployment in 400 Gb/s systems. Depending on the modulation scheme and pulse-shaping technique employed, this data rate can correspond to various electrical bandwidth requirements. For non-return-to-zero (NRZ), square-pulse signals, the required bandwidth is approximately 90 GHz [21]. With raised-cosine pulse shaping, the bandwidth requirement may be lowered to 75 GHz [11] (with a roll-off factor of 0.15 [21]) for 130-GBaud signals.

According to HB-CDM 2.0 [1] (see Fig. 1.2), a minimum -3dB bandwidth of 35 GHz is recommended for 64-GBaud systems, referred to as Class-40 HB-CDM. For 96-GBaud systems, known as Class-60 HB-CDM, a bandwidth of 55 GHz is required. Finally, for the target symbol rate of 130 GBaud, labeled Class-80 HB-CDM, a specified -3dB bandwidth of 75-90 GHz is established, where 75 GHz and 90 GHz represent the lower and upper masks of the frequency response, respectively.

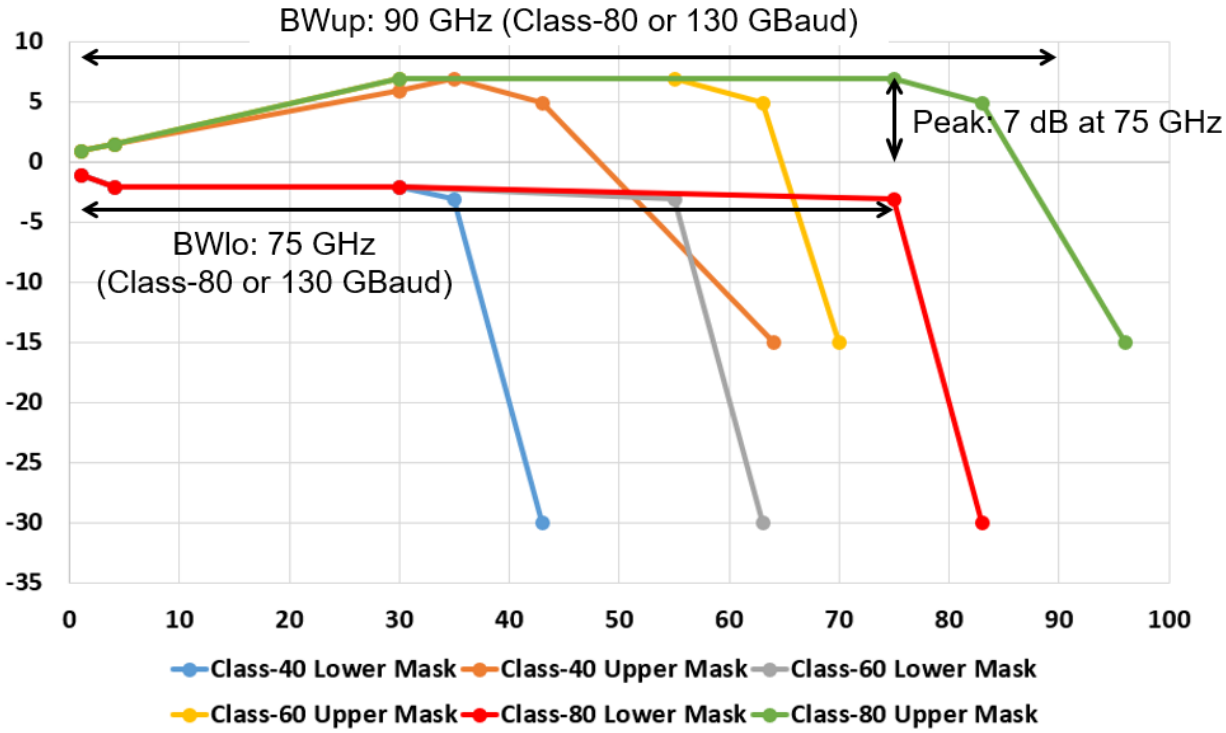


Figure 1.2: Normalized electro-optic frequency response masks for drivers with various symbol rates proposed by the HB-CDM 2.0 standard (NRZ signal assumed; BW_{up} and BW_{lo} represent the upper and lower mask of -3dB bandwidth, respectively) [1]

1.2.2 Gain, Peaking amplitude, Input/Output Swing and Linearity

As mentioned in previous sections, references [2, 17] have indicated that a V_π of greater than 8 V is typical for a 3-mm long silicon-photonic (SiP) MZM. MZMs may be biased at a lower voltage (e.g., $V_\pi/2$ [2]) to reduce the requirement of driver output swing. It is common to use a 6-V peak-to-peak differential ($\mathrm{V_{pp,d}}$) swing when driving a 100-Ω differential load (i.e., 16.5 dBm of RF power, as seen in Zhao [22]), while a 4-$\mathrm{V_{pp,d}}$ swing is sufficient for a 60-Ω differential load (i.e., 15.2 dBm of RF power).

The input range applied to the MZM driver typically falls within the range of 0.3 $\mathrm{V_{pp,d}}$ to 0.6 $\mathrm{V_{pp,d}}$, as indicated in [1]. Consequently, the overall forward transmission coefficient (S_{21}) may be calculated with Eq. 1.1, including the effect of a non-50Ω load (assuming a binary data signal with 50% duty cycle and equal number of 1's and 0's).

$$S_{21} = \frac{V_{out}}{V_{in}}(1+\Gamma_{in})\sqrt{\frac{R_s}{R_L}}, \tag{1.1}$$

where $\mathrm{V_{in}}$ and $\mathrm{V_{out}}$ are the input and output differential, peak-to-peak voltage swing (i.e., $V_{pp,d}$), respectively. The load and source resistances are represented by $\mathrm{R_L}$ and $\mathrm{R_s}$ and Γ_{in} is the input reflection coefficient. Assuming that $\mathrm{V_{out}}$, $\mathrm{V_{in}}$, $\mathrm{R_L}$ and $\mathrm{R_s}$ are 4 $\mathrm{V_{pp,d}}$, 0.3 $\mathrm{V_{pp,d}}$, 30 Ω and 50 Ω, respectively, and the input is matched, the end-to-end S_{DD21} required is 24.7 dB.

To compensate for losses within the package and the frequency response of the MZM, it is necessary to peak the frequency response, especially near the upper cutoff frequency of the driver. As shown in Figure 1.2, a maximum peak of 7 dB is considered acceptable [1]. In the context of a 130-GBaud system, the proposed gain peaking specification is 9 dB at 75 GHz. Gain at lower frequencies can be calculated by subtracting the peaking amplitude (i.e., 9 dB) from S_{DD21} (i.e., 24.7 dB), resulting in a low-frequency gain of 15.7 dB.

In a fully differential amplifier, the common-mode gain (S_{CC21}) is reduced by degenerating the core transconductor through the tail current source in the common mode. However, the use of a tail current source for the output power stage adds to the total DC power consumption. Furthermore, parasitic capacitance of the current source restricts the frequency range for common-mode gain suppression. As common-mode signals in the circuit (e.g., noise from power supplies) may cause oscillation if instability occurs in the common mode. The amplifier should be designed so that it passes a stability test (e.g., μ-stability test [23]) with sufficient margin.

Linearity is another crucial specification, particularly in systems that utilize high-order modulation schemes. In the case of broadband linear drivers, a Class-A amplifier is used at the output. These amplifiers are designed to drive a broadband resistive load while maintaining acceptable linearity. Class-A amplifiers are typically operated below power saturation levels (i.e., in 'backoff'). This backoff operation is required to maintain linearity.

To assess linearity, a widely used metric is the 1-dB compression point ($P_{\text{-1dB}}$) [24]. For a single-stage, Class-A power amplifier (e.g., the output stage), the output 1-dB compression point ($OP_{\text{-1dB}}$, expressed in dBm) can be estimated from [25]

$$OP_{-1dB} \approx 10 \times \log_{10}(\frac{V_{rms}^2}{R_L \times 10^{-3}}) + 2 \text{ dB}, \tag{1.2}$$

where V_{rms} and R_L are the root-mean-square output swing (sinusoidal signal assumed) and load resistance, respectively. An $OP_{\text{-1dB}}$ of 15.2 dBm is required for 4 $V_{\text{pp,d}}$ driving a 60-Ω differential load and 17.2 dBm is preferred (i.e., 2-dB margin added). For broadband operation, the $OP_{\text{-1dB}}$ is expected to be greater than 15.2 dBm up to 75 GHz, where the amplitude response peaks.

Note that eq. 1.2 assumes a single stage. In a cascaded system, the overall $OP_{\text{-1dB}}$ expressed in dBm [24] is given by

$$OP_{-1dB} = 10 \times \log_{10} \frac{1}{\frac{1}{OP_{-1dB,n}} + \frac{1}{OP_{-1dB,n-1}G_n} + ... + \frac{1}{OP_{-1dB,1}G_n...G_2}}, \tag{1.3}$$

where $OP_{-1dB,n}$ and G_n represent the $OP_{\text{-1dB}}$ and the power gain of the n^{th} stage ($G_n = |S_{21n}|^2$), respectively. Equation 1.3 predicts that the $OP_{\text{-1dB}}$ of a cascaded amplifier is mainly determined by the last stage (i.e., stage 'n'). However, pre-amp stages, which usually produce less power than the output stage, are designed to meet overall gain and peaking requirements. They have a much lower $OP_{\text{-1dB}}$ than the output stage, but still have a significant influence on the overall linearity.

For example, an output stage with 17.2-dBm $OP_{\text{-1dB}}$ and 5-dB gain cascaded to a pre-amp with 11.2-dBm $OP_{\text{-1dB}}$ (i.e., scaled down by 4 times) results in an overall $OP_{\text{-1dB}}$ of 13.7 dBm (i.e., 3.5-dB drop in $OP_{\text{-1dB}}$ caused by the pre-amp stage). Eq. 1.3 indicates that the loss in $OP_{\text{-1dB}}$ due to pre-amp stages is suppressed by increasing the gain of the output stage, which favours maximizing the bandwidth of the output stage (i.e., avoiding early gain roll-off is advantageous for broadband linearity).

The ratio between the root sum square of the root-mean-square voltage for each harmonic and the fundamental is known as the total harmonic distortion (THD), which is another specification used to characterize linearity (Eq. 1.4).

$$THD = \frac{\sqrt{v_2^2 + v_3^2 + v_4^2 ...}}{v_1}. \tag{1.4}$$

Typically, THD is measured with a fundamental tone at 1 GHz as seen in [26]. A THD of <6% (or <-24 dB) is targeted for an output of 4-$V_{pp,d}$ with a 60-Ω differential load.

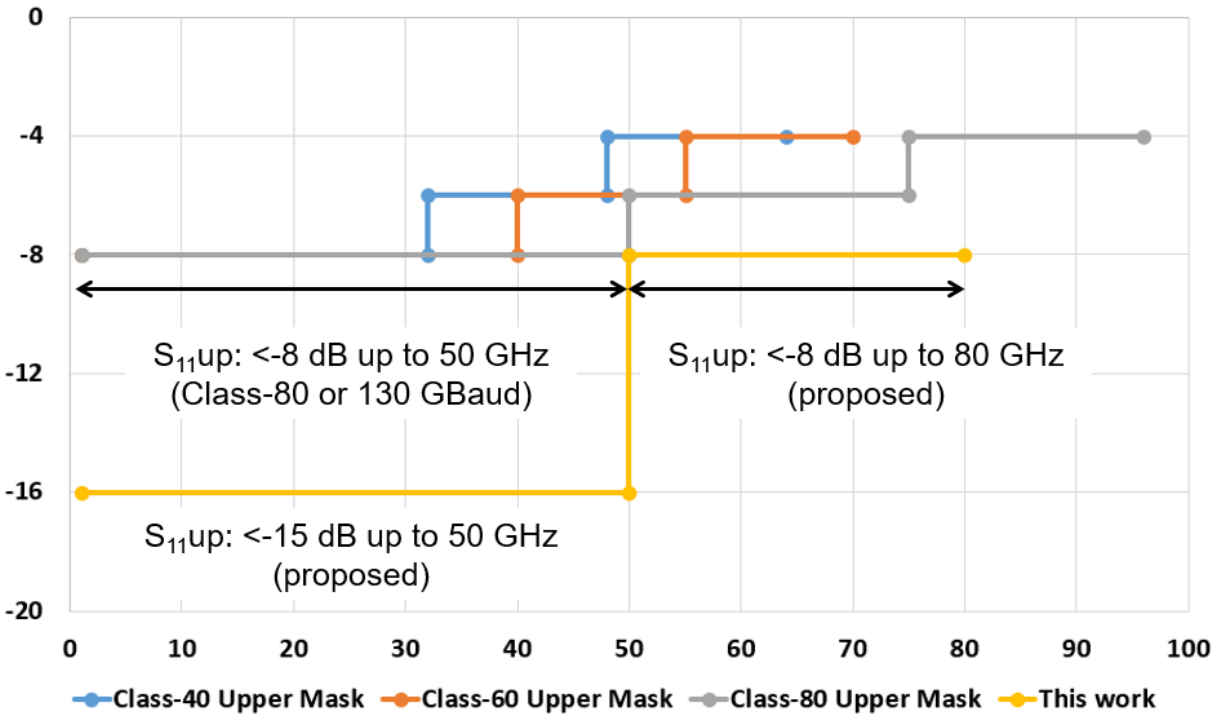


Figure 1.3: Differential electrical input reflection coefficient for drivers with various symbol rates proposed by the HB-CDM 2.0 standard ($S_{11,up}$ represent the mask of differential input return loss) [1]

1.2.3 Return Loss

The input and output return losses are typically evaluated from the differential S-parameters $|S_{DD11}|$ and $|S_{DD22}|$, respectively. In contrast to discrete amplifiers where a

50-Ω interface is required, the driver output is packaged physically close to the MZM in an optical module [12]. Consequently, there is no strict requirement placed on the output return loss. A mask for $|S_{\mathrm{DD11}}|$ in the electrical domain is proposed in HB-CDM 2.0. Input return losses of 8, 6 and 4 dB are stipulated for frequencies up to 50, 75 and 96 GHz, respectively. To impose even stricter requirements, this work aims for input return losses of 15 dB and 8 dB up to 50 GHz and 80 GHz, respectively, as shown in Fig. 1.3.

The common-mode input return loss ($|S_{\mathrm{CC11}}|$) of the driver is set by the lower corner frequency of the on-chip termination and decoupling network at the input. A mask of 2 dB from 10 MHz to 64 GHz is proposed. In this work, the input signal is fed by a vector-network analyzer or an arbitrary-waveform generator, and is expected to be differential. Therefore, $|S_{\mathrm{CC11}}|$ will not affect driver's performance during measurements. In practice (i.e., unclean signal sources), a poor $|S_{\mathrm{CC11}}|$ at the input results in signal reflection in the common mode, which may lead to intersymbol interference if the system is not perfectly balanced.

1.2.4 Other Objectives

One drawback associated with operating in Class A is power efficiency well below 50% [24], leading to high DC power consumption. This is particularly relevant for broadband MZM drivers due to their demanding output voltage swing requirements. For this work, a DC power budget of 650 mW is assumed for the entire driver. According to the HB-CDM 2.0 guidelines, a typical supply voltage for the driver is 3.3 V, with a range of 3.14 V to 3.46 V [1]. From a system perspective, a single 3.3-V supply is preferable, but multiple supplies may be used for prototyping.

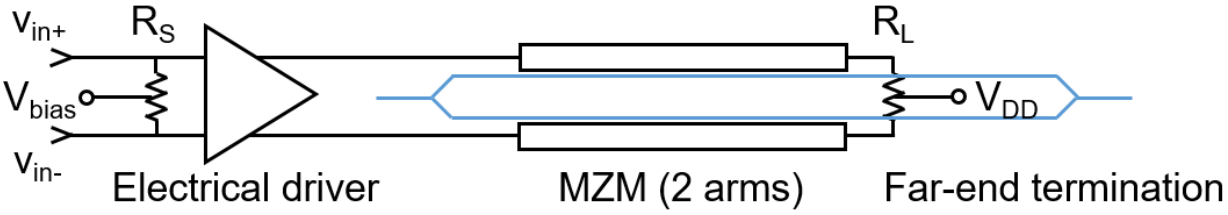


Figure 1.4: A simplified block diagram of a two-arm, travelling-wave MZM with an differential, electrical driver and far-end termination [2]

In a comprehensive optical module, the MZM is supplied by a far-end termination, typically set at 7 V [1], as depicted in Fig. 1.4. As the scope of this work focuses on the

electrical driver only, it is assumed that bias-Ts are available at both the input and output to supply current to the driver.

Electrostatic discharge (ESD) can cause failures during measurement and packaging. A minimum level of 250-V ESD protection is required [1] (0.17-A ESD current for the human-body model). This is achieved by incorporating ESD protection diodes and supply clamps at the relevant pads, as seen in [19].

1.3 Thesis Organization

This thesis consists of six chapters. In chapter two, an extensive literature review of prior research related to the design of broadband, linear, and high-swing drivers is presented. Chapter three focuses on the design of core circuit blocks (e.g., on-chip termination, pre-amp, interstage peaking network, and output power stage), which includes theoretical derivations, a condensed design methodology, and simulation outcomes. The fourth chapter describes the physical layout and implementation of a prototype driver. Chapter five delineates the test strategy and the resulting measurement data, followed by an analysis of the simulated and measured results. The last chapter concludes this thesis with a summary and a proposed plan for future work.

Chapter 2

Literature Survey of Driver Amplifiers

In this chapter, a review of the state-of-the-art in modulator drivers is presented. Designs are divided into three categories: broadband, linear, and wide-swing drivers.

2.1 Previous Work on Driver Amplifiers

On-off keying (OOK) has been the most common modulation technique used in high bit-rate wireline systems. The OOK-modulated signal switches between '0' and '1' levels, and therefore limiting amplifiers (LA) may be used in the driver to provide a high voltage gain, as seen in Galal [27]. On the other hand, Zhao and Vera [22, 28] employed the distributed LA approach. The time delay between LA stages was regulated digitally through latches, and the output waveform could be adjusted in real time to reduce the rise/fall time.

This thesis focuses on the use of broadband, linear, analog driver amplifiers with high output swing applicable to modulation schemes with multiple amplitude levels (e.g., QPSK, PAM-4, etc.). Common broadband circuit techniques include inductive peaking (e.g., shunt peaking by Galal and Hwang [27, 29], series peaking by Choi and Shopov [30, 31], shunt-series peaking by Jyo and Galal [32, 33]), negative shunt capacitance synthesizing by Nakano and Galal [27, 34], f_T doubling by Vera, Kim and Ding [35, 36, 37], distributed amplification by Wakita, Chen and Baker [38, 39, 40] and cascading by Shopov [31].

The breakdown voltage of the device closest to the output restricts the swing. With long-channel FETs, reliable operation is assumed because the breakdown voltage is much

higher than the output swing. However, in deep-submicron CMOS technologies, the breakdown voltage of a thin-oxide transistor is 2-3 V [41], which can cause reliability problems when the required swing is large. A common solution is to use a thick-oxide device with a higher breakdown voltage at the output. Another approach, which is often used in band-limited wireless systems, is to divide the output voltage swing among multiple transistors in series [42]. In this way, stress is shared equally among devices and each transistor is kept below the breakdown limit. Common circuit techniques to implement this include device cascoding by Kim, Li, Shopov, Chen and El-Aassar [3, 4, 31, 39, 43], and actively driving the cascode gate node (also known as breakdown-voltage doubling) by Li, Mandegaran, Ahmed, Agah and Ahmed [4, 5, 26, 44, 45].

Large-signal operation can present difficulties in ensuring linearity, particularly when using high-order modulated signals. To address this, a variety of linearization techniques have been developed, such as feedback, feedforward, harmonic termination, optimal biasing, derivative superposition, and IM3 cancellation [46]. While some of these techniques are only suitable for narrowband systems (e.g., harmonic termination, derivative superposition, and IM3 cancellation), others are sensitive to process and temperature variations (e.g., optimal biasing). Feedback and feedforward, however, have the potential to be applied to broadband designs, since linearization is achieved by suppressing distortion in a closed feedback loop or cancelling distortion using parallel signal paths (i.e., feedforward) [6, 47].

A DP-QAM transmitter is an electro-optic (EO) system, and EO co-design has been proposed to realize adequate bandwidth and swing. In contrast to the traditional travelling-wave MZM, a multi-electrode MZM may reduce the electrical driver's bandwidth or output swing requirements by up to a factor of two [2, 17], although the MZM occupies more area. EO co-design was implemented in Hosseinzadeh [7], and 50 Gb/s transmission was reported.

2.2 Wide-Swing Linear Driver Amplifiers

This section reviews the design of linear, wide-swing drivers. Wide-swing methods include cascoding and feedforward breakdown-voltage doubling. Feedforward linearization methods are also discussed.

2.2.1 Cascoding

A large output voltage swing is required to drive MZMs with a high extinction ratio (e.g., >3 $V_{pp,d}$ for InP and >4 $V_{pp,d}$ for silicon photonics [43]). Therefore, the design of high-swing drivers is mostly confined to SiGe-BiCMOS technologies, where bipolar transistors with >3-V breakdown voltage are offered. Cascoding is used to achieve improved output swings, as the supply voltage and voltage headroom are increased compared to a single device. However, the voltage distribution between transistors in a cascode is non-uniform. That is, the drain-to-source voltage (V_{DS}) across the common-source (CS) transistor M_1 is small compared to V_{DS} across the common-gate device M_2 in a MOS cascode output stage, as depicted in Fig. 2.1a. As a result, more voltage stress is placed across the transistor closest to the output, which may cause breakdown and poor reliability.

To solve this problem, a thick-oxide transistor with a higher breakdown voltage may be utilized to build the last CG cell in CMOS, as shown in Fig. 2.1b. However, the f_T of thick-oxide transistors (e.g., 117 GHz for 1.2-V thick-oxide NFETs from 22FDX®) is typically much lower than that of thin-oxide transistors (e.g., 355 GHz for 0.8-V thin-oxide NFETs from 22FDX®), which decreases GBW of the driver.

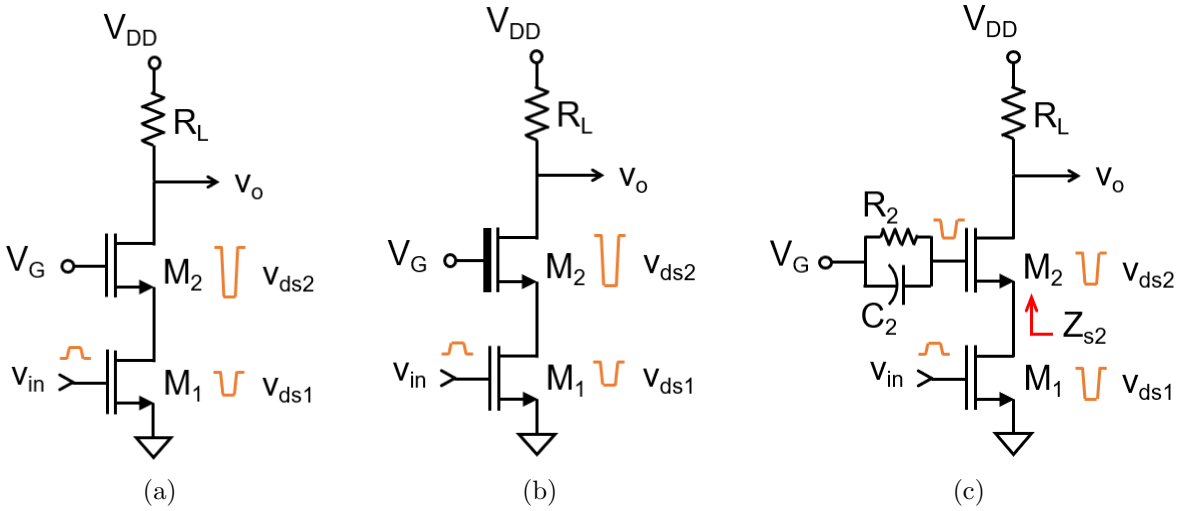


Figure 2.1: Simplified schematics of (a) conventional cascode driver, (b) with a thick-oxide cascode device and (c) with an R-C compensation network at the gate of the cascode [3]

Device cascoding, or 'stacking', employs a gate compensation network, as shown in Fig. 2.1c. Unlike conventional cascode amplifiers where the gate of CG stage is AC grounded, a

parallel RC network is connected from the gate of M_2 to ground. A simplified small-signal model of the CG stage is depicted in Fig. 2.2. Neglecting the drain-to-gate feedback (C_{dg} and R_{dg}), $C_{\rm k}$ (i.e., the k-th device in a cascode amplifier) forms a capacitive voltage divider with the gate-to-source capacitance ($C_{\rm gs}$). As a result, part of the voltage swing at the source of the cascode device appears at the gate of the k-th cascode device. It compensates for the effect of the oxide capacitance of FETs on the frequency response, and lowers the transconductance of CG cells. Assuming that any current fed back via $C_{\rm dg}$ is negligible and that resistor $R_{\rm k}$ used for biasing is large, the input impedance of the k-th cascode stage ($Z_{\rm s,k}$) [3] is approximated by

$$Z_{s,k} \approx \frac{1}{g_{m,k}}(1+\frac{C_{gs,k}}{C_k})(\frac{1}{1+s/\omega_T}), k = 2, 3..., \tag{2.1}$$

where $g_{\rm m,k}$ is the intrinsic transconductance of the k-th CG FET and $\omega_{\rm T}$ is the transition frequency in rad/s. The value of $C_{\rm k}$ is selected such that $Z_{\rm s,k}$ increases as k increases. As a result, the voltage swing at the source of the cascode device increases as k increases. Ideally, each device in a cascoded string experiences the same $V_{\rm DS}$, and $V_{\rm DS}$ is kept below the breakdown voltage. One potential drawback is that a total failure occurs if any one device in the series string fails, which affects reliability when cascoding many devices.

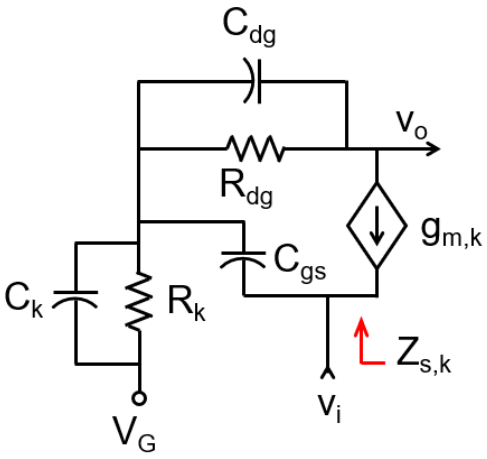


Figure 2.2: Simplified CG FET with drain-gate feedback resistor to spread voltage stress at low frequencies [4]

At frequencies close to DC, $C_{\rm gs}$ of the CG FET becomes an open circuit and the increase in $Z_{\rm s,k}$ is not seen. A drain-gate feedback resistor (R_{dg}) may be used to bias the CG FET

and couple a portion of the output swing to the gate of CG FETs (i.e., relax the V_{DS} stress), as depicted in Fig. 2.2. Another potential solution is to add a resistor in parallel with $C_{\mathrm{gs,k}}$, which compensates for the high-pass response when only using C_{k} at CG-FET gates to extend the desired response to DC.

2.2.2 Feedforward breakdown-voltage doubling

As discussed in previous sections, the total output voltage swing may be spread among multiple FETs by generating a voltage swing at the gate of CG devices. It is accomplished with a passive R-C network in the device-cascoding approach. Alternatively, breakdown-voltage doubling (BVD) uses an active circuit (i.e., an auxiliary path) to apply a signal at the cascode gate, resulting in similar outcomes. A simplified schematic of a BVD is depicted in Fig. 2.3.

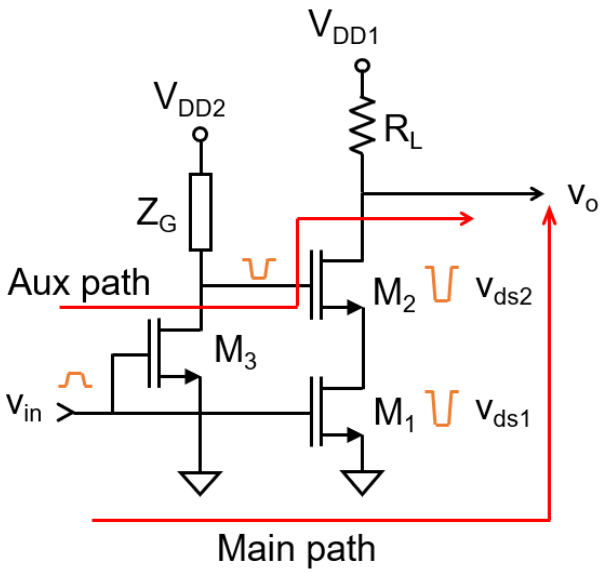


Figure 2.3: Simplified schematic of a breakdown-voltage doubler [5]

The main path is via M_1 and M_2, and the auxiliary path via M_3 applies a signal to the gate of M_2. Note that impedance (Z_G) at the gate of M_2 adjusts the amplitude and phase of the signal in the auxiliary path so that the output swing is divided equally between M_1 and M_2. To avoid excessive DC power consumption, M_3 may be biased at a lower current

than M_1, reducing its transconductance. Increasing the value of Z_G ensures that the swing required at the gate of M_2 remains unchanged.

2.2.3 Feedforward linearization

A feedforward network is typically comprised of a main path and one or more auxiliary paths. An error signal, which contains distortion information, is processed by an error amplifier in the auxiliary path. This signal is then combined with the main path to cancel out its distortion (i.e., linearize signals from the main path). An example of a simplified feedforward linearization network is shown in Figure 2.4.

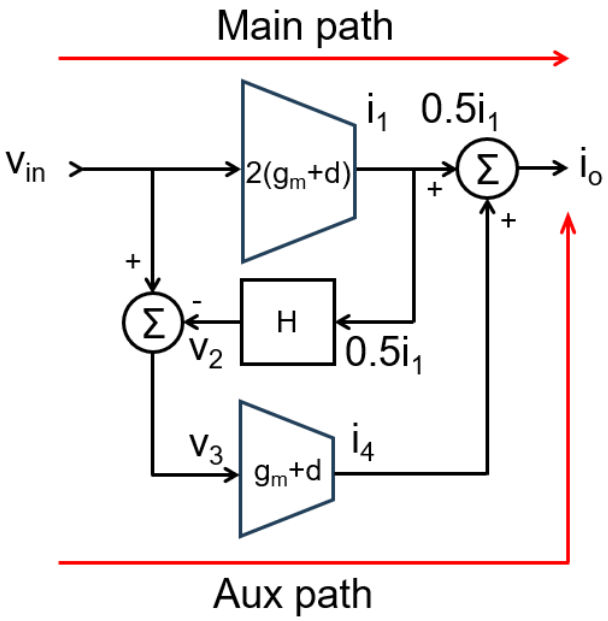


Figure 2.4: Block diagram of a 2-way feedforward linearization network [6]

The main transconductor outputs a current $i_1 = 2v_{in}(g_m + d)$, where g_m and d represent the signal and distortion of the transconductance, respectively. Assuming that half of this current is fed to the H network, the input voltage to the auxiliary path v_3 for H set to $1/g_m$ is given by

$$v_3 = v_{in} - v_2 = -\frac{v_{in}d}{g_m}. \tag{2.2}$$

If the error transconductor in the auxiliary path has transconductance $(g_m + d)$, the output error current $v_3(g_m + d)$ is combined with the main signal, which yields

$$i_o = 0.5i_1 + i_4 = v_{in}(g_m + d) - \frac{v_{in}d}{g_m}(g_m + d) = v_{in}(g_m - \frac{d^2}{g_m}), \tag{2.3}$$

where i_o is the current output. If the magnitude of the distortion is much less than transconductance (i.e., $d \ll g_m$), then the residual distortion in the transconductance (d^2/g_m) is negligible and the system behaviour is linear.

The feedforward linearization technique is attractive for designing a broadband linear driver. Unlike feedback linearization, there are fewer stability concerns. The major disadvantage of feedforward is the DC power consumed by the auxiliary path. Furthermore, half of the current is dissipated in the H network to generate the error signal, again reducing the efficiency.

2.2.4 Summary of wide-swing techniques and selected driver examples

In this section, two wide-swing techniques (cascoding with capacitive gate network and feedforward breakdown-voltage doubling) are discussed. The main limitations of these techniques are the amount of DC power used and the space they occupy. Both techniques introduce signals at the gate of cascode devices to equalize the voltage stress among multiple transistors and keep them away from breakdown.

Adding a capacitive network to the gate of cascode devices is seen in Kim, Li, Shopov, Chen and El-Aassar [3, 4, 31, 39, 43]. It has the advantages of not consuming DC power and occupying little area (i.e., both due to the fact that on-chip, metal-oxide-metal capacitors can be made dense by utilizing low-layer metals). The main drawback is that including a parallel R-C network at the gate of cascode devices results in a high-pass response (shown in Fig. 2.1c), which is not suitable for broadband drivers operating down to DC (i.e., compensation is needed). In the large-signal regime, C_{gs} of cascode devices is nonlinear, which leads to mismatch between C_{gs} and C_k if a constant $Z_{s,k}$ is expected (see Eq. 2.1). Since the required gate swing increases with the number of cascoded devices, the size of the gate capacitor C_k keeps shrinking, which decreases the gain and eventually makes the method not feasible when cascoding many devices.

The feedforward breakdown-voltage doubling (BVD) technique spreads the drain-to-source swing across multiple transistors in series by actively driving the gate of the cascode

devices, as seen in Li, Mandegaran, Ahmed, Agah and Ahmed [4, 5, 26, 44, 45]. The advantage of BVD is that the gate swing may be controlled by adjusting the bias current of the auxiliary path via current mirrors, making the circuit more flexible.

For this thesis, the total DC power target is 650 mW without a specific chip area constraint. Therefore, it is preferred to utilize a passive network at the gate of cascode devices.

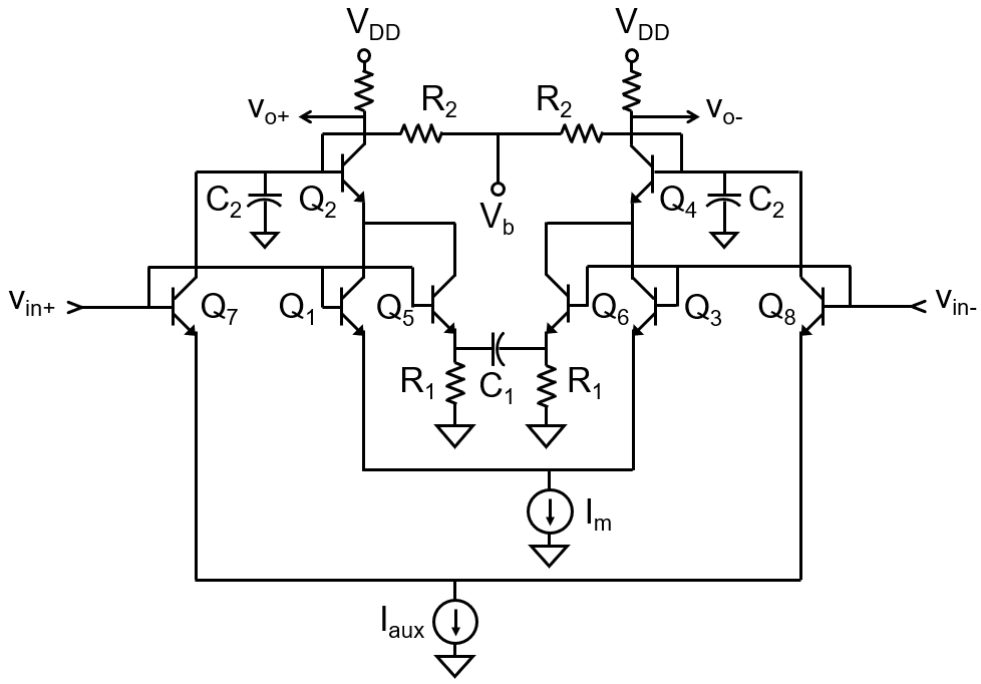


Figure 2.5: Simplified schematic of a driver output stage in [4]

The driver in Li [4] realized in 0.35-um SiGe-BiCMOS technology is comprised of two stages (pre-driver and output stages), and Fig. 2.5 shows the output stage. In this example, both a feedforward gain stage and a passive network are used to generate the desired swing at the base of the cascode devices $Q_{2,4}$. The main path is a differential cascode amplifier (Q_{1-4}) with a bandwidth enhancement circuit (i.e., $Q_{5,6}, R_1$ and C_1 together introduces a zero and peaks the frequency response). The auxiliary path comprises $Q_{7,8}$ and R_2, and is designed to provide a swing at the base of $Q_{2,4}$ with half of the output swing amplitude [4]. Due to the large size of $Q_{2,4}$, the auxiliary load R_2 has to be kept small to ensure a small time constant at the base of $Q_{2,4}$. That is, the auxiliary devices $Q_{7,8}$ have to be scaled up to maintain sufficient gain from the auxiliary path, leading to an increase in DC power

dissipation. However, with the aid of C_2 (i.e., the high-pass passive network at the base of $Q_{2,4}$), a portion of the collector and emitter signals are coupled to the base via C_{cb} and C_{eb} of $Q_{2,4}$ at high frequencies, respectively. This effectively makes the auxiliary path responsible only for signal generation at low frequencies. As a result, R_2 may be sized large and the current consumption of the auxiliary path is reduced. To avoid process and temperature variations, capacitor C_2 is implemented by a BJT with base and emitter shorted together [4].

With a 1-$V_{pp,d}$, PRBS-31 signals at the input, the driver achieved a speed of 10 Gb/s with NRZ signals. The output swing is 9 $V_{pp,d}$ across a 100-Ω load while consuming 1 W from a 6-V DC supply at the output. The measured rise/fall times and jitters are 25/27 ps and 13 ps_{pp}, respectively. The BER measured with a single-ended input amplitude greater than 100 mV_{pp} is less than 10^{-11} [4].

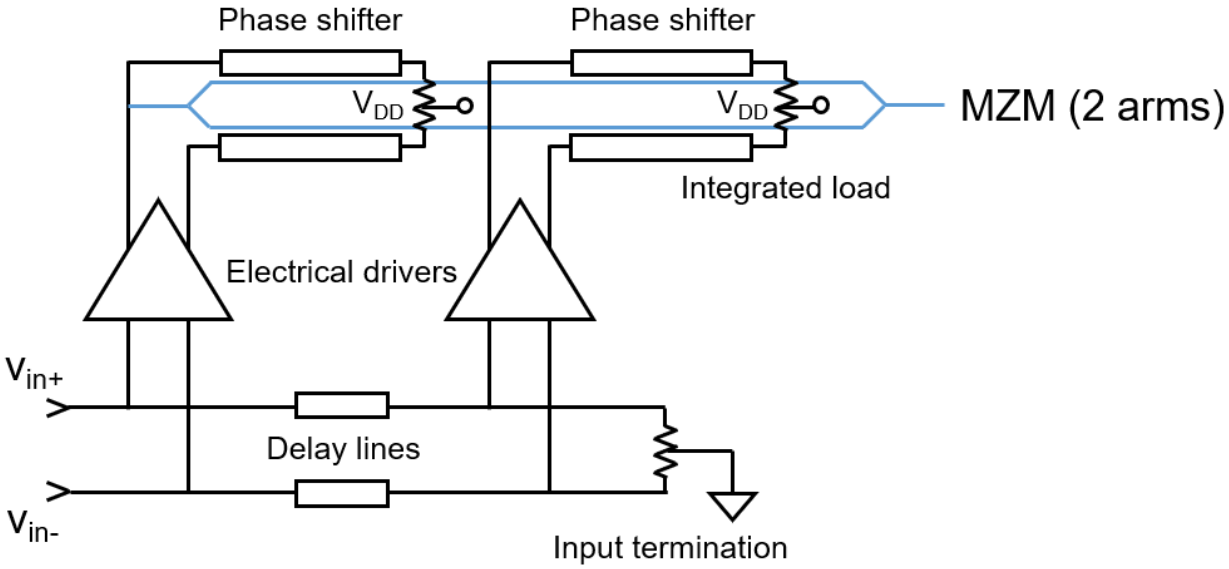


Figure 2.6: Simplified diagram of a multi-electrode MZM (2 segments) with integrated driver [7]

2.3 Driver-MZM Co-design

Integration of the electrical driver with a MZM optimizes the overall EO response. Compared to travelling-wave MZMs (see Fig. 1.4), multi-electrode MZMs have the load resistor

moved to the MZM die and integrated into the driver stage, as depicted in Fig. 2.6. Each segment becomes a short travelling-wave phase shifter with its bandwidth limited by electrode loss [7]. The input signal is fed with a delay between each driver to sync the output (i.e., a distributed input line).

In Hosseinzadeh [7], the electrical driver is wire bonded to connect to the phase shifter. The wirebond inductance leads to series inductive peaking, thus increasing the bandwidth. When transmitting at 30-Gb/s with a pseudorandom binary sequence-31 pattern (PRBS-31), the full link (i.e., driver integrated with the modulator) yields a BER better than 10^{-10} with 2-tap feedforward equalization applied to enhance eye openings. The optical eye is open up to 50 Gb/s.

2.4 Performance Summary

Table 2.1 summarizes the specifications of drivers covered in previous sections. Compared to conventional modulation schemes such as OOK and PAM-4, DP-QAM shows a significant advantage in terms of data throughput (e.g., 552 Gb/s in Ahmed [44] and 640 Gb/s in Jyo [32]). The drawback of DP-QAM systems is the DC power consumption, as four electrical drivers are needed to drive MZMs for both I-Q paths and X-Y polarizations (i.e., 4x DC power consumption compared to a single lane). This may be seen in Ahmed [44] (in total 4 W for the transmitter). The power consumed by the DP-32QAM driver in Jyo [32] is only 900 mW, which is due to the low swing required (1.5 V). Distributed amplification is commonly applied in drivers with more than 50-GHz electrical bandwidth, as seen in Wakita and Baker [38, 40]. For lumped designs, cascading with interstage inductive peaking is widely deployed to attain wide bandwidth, as shown in Nakano, Hwang, Shopov and Jyo [29, 31, 32, 34].

For high-swing drivers, adding a passive R-C network at the cascode device gate to distribute the voltage stress is widely used in MOS designs, such as Hosseinzadeh, Shopov and Dadash [7, 31, 48]. Even swing distribution can also be achieved by actively driving the cascode device gate (i.e., BVD), such as those described in Ahmed and Giuglea [44, 49].

Table 2.1: Typical performance of state-of-the-art MZM drivers (recent 10 years)

Ref.	Technology	Topology	Data Rate (Gb/s)	BER	Elec. Bandwidth (GHz)	Gain (dB)	V_{out} ($V_{pp,d}$)	V_{in} ($V_{pp,d}$)	S_{22} (dB)	P_{DC} (mW)	Energy Efficiency (pJ/b)
JSSC 2017 [28]	130-nm SiGe-BiCMOS	Digitally-controlled, 3-stage distributed, limiting amplifier	40 NRZ	-	-	-	6	0.3	>20, <57 GHz	1920	48
IMS 2020 [7]	45-nm SOI-CMOS & 90-nm SiP	EO co-design 4-segmented MZM distributed driver	50 NRZ	$<10^{-10}$	26	9	4.4[1]	-	N/A Integrated	480	9.6
JSSC 2016 [31]	28-nm FD-SOI CMOS	Series peaking, lumped 6-stage cascade	60 NRZ	10^{-4} at 56 Gb/s	17/35[2]	38/22[2]	4.3[3]	0.4	N/A Integrated	320	5.3
VLSI 2017 [34]	65-nm CMOS	Shunt-series, negative shunt capacitor lumped, 5-stage cascade	80 PAM4	$<10^{-12}$	>32	28	2	0.15	N/A, Integrated	180	2.25
Access 2022 [49]	130-nm SiGe-BiCMOS	Inductorless BVD lumped	90 PAM4	-	61.2	15.5	7.2	0.65	N/A, Integrated	670	7.44
TMTT 2017 [50]	130-nm SiGe:C-BiCMOS	Shunt-series, 6-stage DA	90 PAM4	-	90	12.5	4	1.2	>10, <90 GHz	550	6.1
IMS 2016 [38]	250-nm InP-DHBT	VGA + 6-stage DA	112 PAM4	-	>67	10	1.8	0.7	>10, <67 GHz	840	7.5
BCICTS 2018 [48]	22-nm FDSOI	Lumped cascode (3-stack)	112 PAM4	-	39[2]	16.5[2]	2.4	-	-	140	1.25
BCICTS 2022 [51]	130-nm SiGe-BiCMOS	Shunt-series, lumped 3-stage cascade	128 PAM4	-	>115	18	2	-	>8, <30 GHz	280	2.2
RFIC 2017 [40]	55-nm SiGe-BiCMOS	3-stage cascade, shunt-peaked unit cell in 7-stage DA	168 PAM8	$<10^{-12}$	>70	15	3	-	>10, <67 GHz	1100	6.5
TMTT 2017 [52]	55-nm SiGe-BiCMOS	Shunt peaking lumped, 4-stage cascade	168 PAM8	$<10^{-12}$	57.5	12.8	3.8	0.25	N/A, Integrated	820	4.88
JSSC 2020 [44]	130-nm SiGe-BiCMOS	Inductorless BVD, lumped	552 DP-16QAM	3.2[-2]	>40	20-30 (w/ VGA)	2.4	-	N/A, Integrated	4000	7.2
MWCL 2021 [53]	90-nm SiGe-BiCMOS	series peaking, negative shunt capacitor lumped, 4-stage cascade	632 DP-16QAM	$<2.4\times10^{-6}$	>67	21 (+20 dB @45 GHz)	4.5 @100Ω	-	N/A, Integrated	350	2.2
JSSC 2020 [53]	250-nm InP-DHBT	Analog MUX bandwidth doubling, lumped, 4-stage cascade	1070 PDM-16QAM	-	>110	10.5 (+8 dB @80 GHz)	1.5 @100Ω	-	N/A, Integrated	3960	3.7

[a]Simulated result
[b]Large-signal measured result
[c]Single-ended

Chapter 3

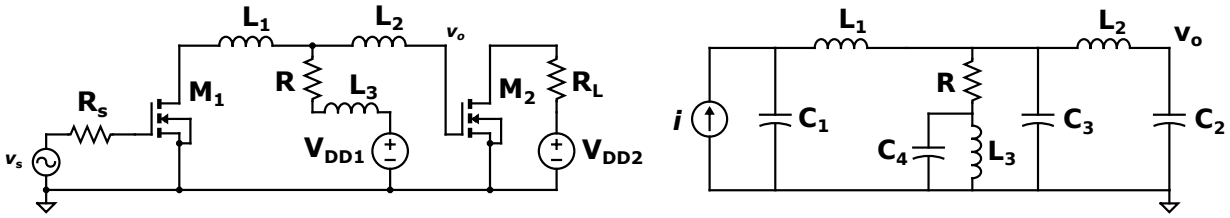
Interstage Network Design

This chapter explores the design of interstage filter networks, focusing on a 7-pole 2-zero shunt/double-series configuration, simplified into two analyzable models: a 4-pole double-series and a 4-pole, 1-zero, shunt/double-series network. The normalized transfer function of the 4-pole, 1-zero network is derived, mapping coefficients to circuit parameters and emphasizing pole-zero optimization for bandwidth extension and peaking frequency. A systematic methodology prioritizes flexibility: capacitor ratio compatibility allows interfacing mismatched stages without iterative tuning. The chapter advances broadband amplifier design by enabling tailored peaking profiles while minimizing in-band ripple and design cycles.

3.1 A 7-pole 2-zero, shunt/double-series interstage network

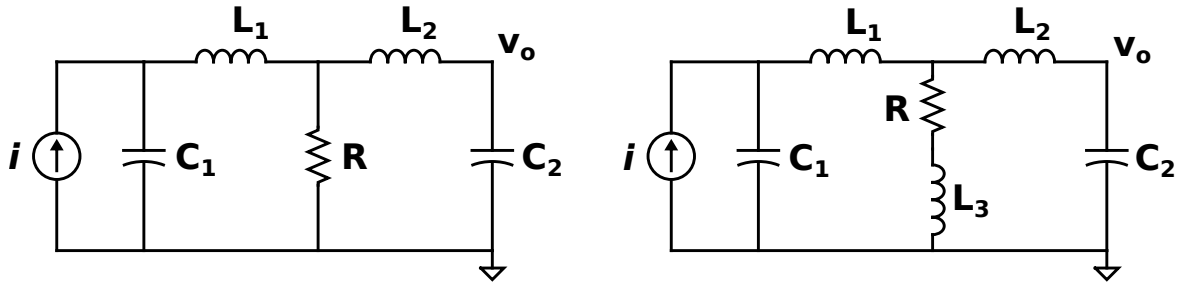
The interstage network coupling two gain stages, and an lumped-element model for the network are shown in Fig. 3.1a and Fig. 3.1b, respectively. Lumped-element models are used instead of distributed-element models, because the physical size of the network is ∼200 μm (detailed in Chapter 5), and the guide wavelength for a frequency of 100 GHz is 1.5 mm. That is, the maximum physical length of passive structures is 0.13x of the guide wavelength, and therefore lumped-element models are valid. The model in Fig. 3.1b includes the output capacitance C_1 looking into the drain of transistor M_1 and the input capacitance C_2 looking into the gate of transistor M_2. The current source i (Fig. 3.1b) represents the small-signal current generated from transistor M_1 in the driving stage (Fig. 3.1a). Series peaking is

realized with L_1, L_2 and R tapped in between. Shunt peaking is achieved by L_3. Capacitors C_1 to C_4 represent the parasitic capacitances associated with the neighbouring components (e.g., dominant parasitics of L_1 are absorbed in C_1 and C_3; parasitics of R are absorbed by C_3 and C_4). The driving and load capacitances of the transistors are included in C_1 and C_2, respectively.



(a) (b)

Figure 3.1: Simplified schematic of (a) the proposed shunt/double-series interstage network and (b) the lumped-element equivalent model



(a) ignoring L_3, C_3 and C_4 (b) ignoring C_3 and C_4

Figure 3.2: Simplifications of the 7-pole, 2-zero circuit (a) 4-pole, double-series interstage network, (b) 4-pole, 1-zero, shunt/double-series interstage network

The transfer function of the complete network of Fig. 3.1b has 7 poles and 2 zeros, which is cumbersome to analyze and design using nodal and mesh analysis. Here, we begin by studying two simplified cases. One case neglects the shunt-peaking network and its parastics (i.e., L_3, C_3 and C_4). The simplified circuit is the 4-pole, double-series network depicted in Fig. 3.2a. The other case (shown in Fig. 3.2b) includes shunt peaking inductor L_3, while parasitic capacitances C_3 and C_4 are neglected (i.e., a 4-pole, 1-zero network). Both simplified cases focus on the components that affect bandwidth (e.g., R, C_1 and C_2

determine a time constant and L_1 to L_3 set the bandwidth extension ratio), while ignoring parasitics that are small (i.e., C_3 and C_4 are typically small compared to the driving and load capacitances of transistors). Therefore, studying these simplified cases provides insight into the operation of the 7-pole, 2-zero network and speeds up the analysis.

3.1.1 Analysis of the 4-pole, 1-zero, shunt/double-series interstage network

The analysis of the 4-pole, double-series network is detailed in Appendix B. Table B.2 in the Appendix B shows that various amplitude peaks may be achieved with different transfer function damping ratio and circuit parameters. However, each design goal usually points to a specific value of n (e.g., to achieve a normalized peak amplitude of 2.078 at frequency of 2.6, n is 0.759 which indicates a ratio of 1:3 between C_1 and C_2. Other inductive peaking techniques, such as series, shunt, and T-coil peaking, could also maximize performance at a specific n [54]). This is cumbersome in circuit realization, because the driving and load capacitances are associated with the transconductance stage design, which may be prioritized for other design specifications (e.g., the output stage has to be sized large to provide sufficient RF power to meet the linearity specification, and the input stage has to be sized properly to realize the desired input-matching bandwidth). To achieve the desired n, it is possible to deliberately introduce extra capacitance. However, this results in a larger time constant at the interstage, which is not advantageous. In addition, improper selection of the non-dominant poles p_1 and p_2 may lead to an in-band local minima (A_{lm}). The 4-pole, 1-zero, shunt/double-series network in Fig. 3.2b is therefore introduced to provide more design freedom and reduce the number of design iterations needed.

The transfer function of the 4-pole, 1-zero, shunt/double-series interstage network is given by

$$F(s) = \frac{v_o}{iR} = \frac{s \cdot \epsilon/(R \cdot \alpha) + 1/\alpha}{s^4 + s^3 \cdot \beta/\alpha + s^2 \cdot \gamma/\alpha + s \cdot \delta/\alpha + 1/\alpha} \quad , \tag{3.1}$$

where the denominator polynomial is written in standard form. Coefficients α, β, γ, δ and ϵ are introduced, which yields

$$\begin{aligned} \alpha &= C_1C_2(L_1L_2 + L_1L_3 + L_2L_3), \\ \beta &= RC_1C_2L_1 + RC_1C_2L_2, \\ \gamma &= C_1(L_1 + L_3) + C_2(L_2 + L_3), \\ \delta &= RC_1 + RC_2, \\ \epsilon &= L_3. \end{aligned} \tag{3.2}$$

To normalize the transfer function, the frequency is normalized to ω_{ref} of $1/RC_T$ rad/s and the amplitude is normalized to iR V. That is, in the sinusoidal steady state ($s = j\omega$), ω is normalized to ω/ω_{ref}. To simplify the expressions, parameters R and C_T are also normalized to 1 Ω and 1 F, respectively. The following parameters are introduced:

$$C_T = C_1 + C_2 \qquad n = \frac{C_2}{C_T} \qquad m_1 = \frac{L_1}{R^2 C_T} \qquad m_2 = \frac{L_2}{R^2 C_T} \qquad m_3 = \frac{L_3}{R^2 C_T}. \tag{3.3}$$

Substituting Eq. 3.3 into Eq. 3.2 yields

$$\begin{aligned} \alpha &= n(1-n)(m_1 m_2 + m_1 m_3 + m_2 m_3), \\ \beta &= n(1-n)(m_1 + m_2), \\ \gamma &= (1-n)(m_1 + m_3) + n(m_2 + m_3), \\ \delta &= 1, \\ \epsilon &= m_3. \end{aligned} \tag{3.4}$$

A generalized 4-pole, 1-zero transfer function with complex poles p_1 to p_4 and a real zero z_1 [54] is given by

$$H(s) = \frac{(-1)^4 p_1 p_2 p_3 p_4}{(s-p_1)(s-p_2)(s-p_3)(s-p_4)} \cdot \frac{(s-z_1)}{-z_1} \quad . \tag{3.5}$$

The numerator of Eq. 3.1 (N_{tf}) may be rewritten as

$$N_{tf} = s \cdot \epsilon/\alpha + 1/\alpha = \epsilon/\alpha(s + 1/\epsilon). \tag{3.6}$$

Compare N_{tf} to the numerator of Eq. 3.5 results in

$$z_1 = -\frac{1}{\epsilon} \qquad\qquad p_1 p_2 p_3 p_4 = \epsilon/\alpha \cdot (-z_1) = \frac{1}{\alpha}. \tag{3.7}$$

The denominator of Eq. 3.1 (D_{tf}) may be expanded as

$$D_{tf} = s^4 + s^3 \cdot a_3 + s^2 \cdot a_2 + s \cdot a_1 + a_0, \tag{3.8}$$

where

$$\begin{aligned} a_3 &= -(p_1 + p_2 + p_3 + p_4), \\ a_2 &= p_1 p_2 + p_1 p_3 + p_1 p_4 + p_2 p_3 + p_2 p_4 + p_3 p_4, \\ a_1 &= -(p_1 p_2 p_3 + p_1 p_2 p_4 + p_1 p_3 p_4 + p_2 p_3 p_4), \\ a_0 &= p_1 p_2 p_3 p_4. \end{aligned} \tag{3.9}$$

By matching the denominator polynomial coefficients of the generalized 4-pole, 1-zero transfer function (Eqs. 3.8 and 3.9) and the normalized transfer function of the shunt/double-series interstage network of Fig. 3.2b (Eqs. 3.1 and 3.4), we have the following equalities:

$$a_3 = \beta/\alpha \qquad a_2 = \gamma/\alpha \qquad a_1 = \delta/\alpha = 1/\alpha \qquad a_0 = 1/\alpha \qquad \Rightarrow \qquad a_1 = a_0. \tag{3.10}$$

Note that the frequency in Eq. 3.1 is normalized to ω/ω_{ref}. For an unnormalized case, the relationship expressed by Eq. 3.10 may be rewritten as

$$a_1 = \frac{1}{\omega_{ref}}\frac{\delta}{\alpha} = \frac{1}{R(C_1+C_2)}\frac{RC_1+RC_2}{\alpha} = \frac{1}{\alpha} \qquad a_0 = \frac{1}{\alpha} \qquad \Rightarrow \qquad a_1 = a_0. \tag{3.11}$$

That is, a_1 also equals a_0 for the unnormalized transfer function. Equating coefficients a_1 and a_0 from Eq. 3.9 gives

$$\begin{gathered} p_1p_2p_3p_4 = -(p_1p_2p_3 + p_1p_2p_4 + p_1p_3p_4 + p_2p_3p_4), \\ \text{or} \qquad 1 + \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4} = 0. \end{gathered} \tag{3.12}$$

Complex poles are utilized to realize efficient inductive peaking as the response becomes underdamped [54]. Assuming that $p_{1,2}$ and $p_{3,4}$ are normalized complex-conjugate pole pairs, and z_1 is the normalized real zero, they may be expressed in Cartesian form as

$$\begin{aligned} p_1 &= -\sigma_1 + j\omega_1, \\ p_2 &= -\sigma_1 - j\omega_1, \\ p_3 &= -\sigma_3 + j\omega_3, \\ p_4 &= -\sigma_3 - j\omega_3, \\ z_1 &= -\sigma_z. \end{aligned} \tag{3.13}$$

Equation 3.12 therefore has four unknowns ($\sigma_1, \sigma_3, \omega_1$ and ω_3).

Recall that in a normalized second-order system with transfer function $F'(s)$, pole locations (p'_1 and p'_2) [55] are

$$\begin{aligned} F'(s) &= \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \\ p'_1 &= -\sigma'_1 + j\omega'_1 = -\zeta\omega_n + j\omega_n\sqrt{1-\zeta^2}, \\ p'_2 &= -\sigma'_1 - j\omega'_1 = -\zeta\omega_n - j\omega_n\sqrt{1-\zeta^2}. \end{aligned} \tag{3.14}$$

The damping factor and undamped natural frequency of the second-order polynomial are ζ and ω_n respectively [55], which are given by

$$\begin{aligned} \omega_n &= \sqrt{\mathrm{Re}(p_1')^2 + \mathrm{Im}(p_1')^2} = |p_1'|\,, \\ \zeta &= \frac{\mathrm{Re}(p_1')}{\sqrt{\mathrm{Re}(p_1')^2 + \mathrm{Im}(p_1')^2}} = \frac{\mathrm{Re}(p_1')}{|p_1'|}. \end{aligned} \tag{3.15}$$

Assuming that the desired system response is underdamped and therefore ζ is small, the peaking frequency ($\omega_{pk,2p}$) may be estimated from the imaginary part of the poles. When ζ approaches 0 [55],

$$\begin{aligned} \omega_{pk,2p} &\approx \omega_n\sqrt{1-2\zeta^2}, \\ \lim_{\zeta\to 0}\omega_{pk,2p} &= \omega_n, \\ \lim_{\zeta\to 0}\mathrm{Im}(p_1') &= \omega_n = \omega_1', \\ \omega_{pk,2p} &\approx \omega_1'. \end{aligned} \tag{3.16}$$

The peak amplitude (A_{pk}) and -3dB bandwidth (w_{-3dB}) in the frequency response of a second-order system may be estimated by [55, 56]

$$\begin{aligned} A_{pk} &= \frac{1}{2\zeta\sqrt{1-\zeta^2}}, \\ \omega_{-3dB} &= \omega_n\sqrt{1-2\zeta^2+\sqrt{(2\zeta^2-1)^2+1}}. \end{aligned} \tag{3.17}$$

The 10%-90% risetime (t_r), settling time (t_s), delay time (t_d) and percentage overshoot (δ_{os}) are important metrics of system step response in the time domain. The 10%-90% risetime is defined as the time required for a pulse to rise from 10% to 90% of its steady-state value. The settling time is the time it takes for the step response to settle within 2% of the final value. The delay time is defined as the time required for the step response to reach 50% of its final value. The percentage overshoot is defined as the ratio between the maximum overshoot and the steady-state value in percentage. For an underdamped second-order system in Eq. 3.14 with approximations made in Eq. 3.16, these time-domain

metrics may be estimated by [56]

$$\begin{aligned} t_r &\approx \frac{1 - 0.4167\zeta + 2.917\zeta^2}{\omega_n}, \\ t_s &\approx \frac{3.2}{\zeta\omega_n}, \\ t_d &\approx \frac{1.1 + 0.125\zeta + 0.469\zeta^2}{\omega_n}, \\ \delta_{os} &= 100e^{-\pi\zeta/\sqrt{1-\zeta^2}}. \end{aligned} \tag{3.18}$$

To design the 4-pole, 1-zero network shown in Fig. 3.2b, p_3 may be chosen as the dominant pole that determine the peaking frequency (but not the amplitude). Therefore, ω_{pk} equals ω_3 when assuming that σ_3 is small (i.e., p_3 and p_4 are close to the $j\omega$ axis in the s-plane), and that σ_3 is much smaller than σ_1 (i.e., poles p_1 and p_2 have minimal effect on the peaking frequency).

The design of the interstage network requires knowledge of the input and output stage parameters C_1, C_2 and R (from Fig. 3.2b). Since the peaking frequency is specified at 75 GHz for this work, the normalized peaking frequency ω_{pk} may be calculated by

$$\omega_{pk} \approx \omega_3 = 75\,\text{GHz} \cdot (2\pi R(C_1 + C_2)), \tag{3.19}$$

with the assumption (made previously) that σ_3 is small and $|\sigma_3| \ll |\sigma_1|$.

The magnitude of the transfer function (with normalized poles) of the 4-pole, 1-zero network (Eq. 3.5) is given by

$$\begin{aligned} |F(\omega)| &= \sqrt{F(j\omega) \cdot F^*(j\omega)} \\ &= \frac{(\sigma_1^2 + \omega_1^2)(\sigma_3^2 + \omega_3^2)\frac{1}{\sigma_z}\sqrt{\sigma_z^2 + (\frac{\omega}{\omega_{ref}})^2}}{\sqrt{[\sigma_1^2 + (\frac{\omega}{\omega_{ref}} + \omega_1)^2][\sigma_1^2 + (\frac{\omega}{\omega_{ref}} - \omega_1)^2][\sigma_3^2 + (\frac{\omega}{\omega_{ref}} + \omega_3)^2][\sigma_3^2 + (\frac{\omega}{\omega_{ref}} - \omega_3)^2]}}. \end{aligned} \tag{3.20}$$

The peak amplitude (A_{pk}) of $|F(\omega)|$ may be expressed as

$$\begin{aligned} A_{pk} &= |F(\omega = \omega_{ref} \cdot \omega_{pk})| \approx |F(\omega = \omega_{ref} \cdot \omega_3)| \\ &= \frac{(\sigma_1^2 + \omega_1^2)(\sigma_3^2 + \omega_3^2)\frac{1}{\sigma_z}\sqrt{\sigma_z^2 + \omega_3^2}}{\sqrt{[\sigma_1^2 + (\omega_3 + \omega_1)^2][\sigma_1^2 + (\omega_3 - \omega_1)^2][\sigma_3^2 + (2\omega_3)^2]\sigma_3^2}}. \end{aligned} \tag{3.21}$$

The normalized -3dB bandwidth ($\omega_{-3dB}/\omega_{ref}$) is defined when the magnitude of the transfer function (Eq. 3.20) equals to $1/\sqrt{2}$ is given by

$$\frac{(\sigma_1^2+\omega_1^2)(\sigma_3^2+\omega_3^2)\frac{1}{\sigma_z}\sqrt{\sigma_z^2+(\frac{\omega_{-3dB}}{\omega_{ref}})^2}}{\sqrt{[\sigma_1^2+(\frac{\omega_{-3dB}}{\omega_{ref}}+\omega_1)^2][\sigma_1^2+(\frac{\omega_{-3dB}}{\omega_{ref}}-\omega_1)^2][\sigma_3^2+(\frac{\omega_{-3dB}}{\omega_{ref}}+\omega_3)^2][\sigma_3^2+(\frac{\omega_{-3dB}}{\omega_{ref}}-\omega_3)^2]}}=\frac{1}{\sqrt{2}}. \tag{3.22}$$

Along with Eqs. 3.12 and 3.21, target A_{pk} and ω_{pk}, a set of $p_{1,2}$ locations may now be determined. As an example, a list of design parameters is shown in Table 3.1. The desired peak amplitude is set to 2 V/V (6 dB), as the total 9-dB peaking could be accumulated from multiple circuit stages (e.g., input and output matching network). Assuming an interstage load of 30 Ω, driving and load capacitances of 55 fF and 130 fF, respectively, the normalized peaking frequency (ω_{pk}/ω_{ref}) calculated via Eq. 3.19 is 2.6. The damping factor is selected at 0.05 (i.e., close to 0. Table B.2 in Appendix B shows that with a damping factor 0.1, A_{pk} = 2.078. Therefore a lower factor is needed to achieve a higher A_{pk}) to ensure that the network has an underdamped response.

The in-band local minima A_{lm} is specified to solve the problem of in-band dipping seen in the 4-pole, double-series networks. The targeted A_{lm} is greater than 1 (i.e., for all frequencies below ω_{pk}, the normalized system magnitude is above the DC gain). The targeted bandwidth extension ratio $\eta_b = \omega_{-3dB}/\omega_{ref}$ is set to 3.

The time-domain response of a second-order system with the same peak amplitude are used as a reference for comparison (i.e., the reference second-order system has a damping factor of 0.259 and a natural frequency of 2.8, which results in the same 2 V/V peak amplitude at 2.6 rad/s). The reference system has a risetime of 0.46 s, settling time of 5.02 s, delay time 0.41 s and percentage overshoot of 43%. Overshoot in the time-domain response is inevitable due to amplitude peaks in the frequency response for an underdamped system. The objective for the 4-pole, 1-zero, shunt/double-series network is to keep the percentage overshoot lower than that of the arbitrary second-order system, while achieving bandwidth extension and improved risetime.

Table 3.1: Design goals for a 4-pole, 1-zero, shunt/double-series network design example

ζ_3	A_{pk} (V/V)	A_{lm} (V/V)	η_b (Hz/Hz)	ω_{pk}/ω_{ref}	δ_{os} (%)
0.05	≥ 2	≥ 1	≥ 3	2.6	≤ 43

By substituting ω_{pk} and ζ_3 into Eqs. 3.15 and 3.19, p_3 and p_4 are found at $-0.13 \pm j2.6$ in

the s-plane. Evaluating Eqs. 3.12, 3.21 and 3.22 with p_3, p_4 and A_{pk} yields

$$\begin{gathered}\sigma_3 = 0.13,\\ \omega_3 = 2.6,\\ 1 + \frac{1}{\sigma_1 + j\omega_1} + \frac{1}{\sigma_1 - j\omega_1} + \frac{1}{\sigma_3 + j\omega_3} + \frac{1}{\sigma_3 - j\omega_3} = 0,\end{gathered} \tag{3.23}$$

$$\begin{gathered}A_{pk} = \frac{(\sigma_1^2+\omega_1^2)(\sigma_3^2+\omega_3^2)\frac{1}{\sigma_z}\sqrt{\sigma_z^2+\omega_3^2}}{\sqrt{[\sigma_1^2+(\omega_3+\omega_1)^2][\sigma_1^2+(\omega_3-\omega_1)^2][\sigma_3^2+(2\omega_3)^2]\sigma_3^2}} \geq 2,\\ \frac{(\sigma_1^2+\omega_1^2)(\sigma_3^2+\omega_3^2)\frac{1}{\sigma_z}\sqrt{\sigma_z^2+\eta_b^2}}{\sqrt{[\sigma_1^2+(\eta_b+\omega_1)^2][\sigma_1^2+(\eta_b-\omega_1)^2][\sigma_3^2+(\eta_b+\omega_3)^2][\sigma_3^2+(\eta_b-\omega_3)^2]}} = \frac{1}{\sqrt{2}}.\end{gathered} \tag{3.24}$$

The corresponding p_1 and p_2 from Eq. 3.23 are solved as

$$\begin{gathered}p_1 = -\sigma_1 + j\omega_1,\\ p_2 = -\sigma_1 - j\omega_1,\\ \omega_1 \approx 0.141\sqrt{-\sigma_1(50.13\sigma_1 + 104.26)},\end{gathered} \tag{3.25}$$

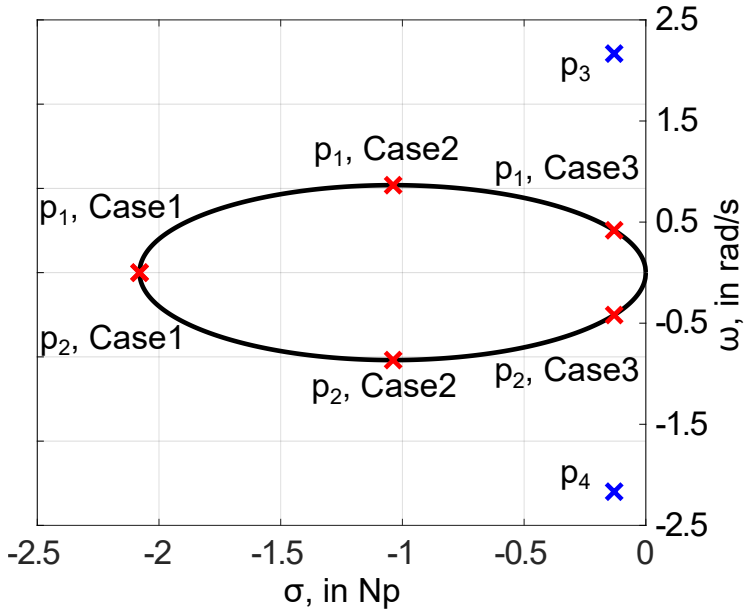


Figure 3.3: Pole-zero plot of the 4-pole, 1-zero, shunt/double-series network (zero neglected) with 3 cases of p_1 and p_2 highlighted (Case 1: $\sigma_1 = 2.164$, Case 2: $\sigma_1 = 1.08$, Case 3: $\sigma_1 = 0.26$)

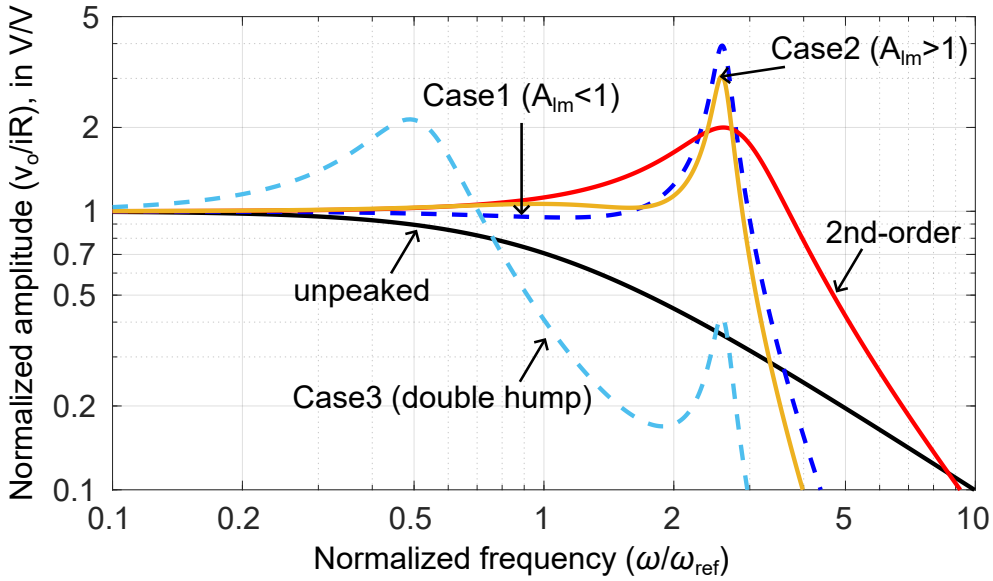


Figure 3.4: Normalized magnitude response of 4-pole, 1-zero, shunt/double-series networks for various σ_1 (Case 1: $\sigma_1 = 2.079$, Case 2: $\sigma_1 = 1.038$, Case 3: $\sigma_1 = 0.13$)

The locations of p_1 to p_4 are depicted in Fig. 3.3. The dominant poles (p_3 and p_4) are shown with blue crosses. All possible solutions of p_1 and p_2 lie on a circle. Three cases of p_1 and p_2 are highlighted with red crosses in Fig. 3.3. Assuming that the real zero z_1 lies at negative infinity, the effects of p_1 and p_2 on the frequency response of the 4-pole, 1-zero transfer function is shown in Fig. 3.4.

The solid black curve shows the unpeaked response with the -3dB bandwidth normalized to 1 rad/s (i.e., $L_1 = L_2 = 0$ and the network becomes an RC circuit). The solid red curve shows the second-order response, which peaks at 2.6 rad/s with an amplitude of 2 V/V and a bandwidth of 4.14 rad/s. Case 1 (blue dashed curve, $\sigma_1 = 2.079$ and $\omega_1 = 0.002$) has the highest peak amplitude of 3.935 V/V with bandwidth of 3.103. The peaking frequency is 2.59, which is close to ω_3 of 2.6 (i.e., Eq. 3.19, the assumption of $\omega_{pk} \approx \omega_3$ is valid). Poles p_1 and p_2 have minimal effect on the overall response, as they are far from the dominant poles. Due to frequency separation between poles, an in-band local minima (A_{lm}) of 0.95 V/V (-0.446 dB) is seen at a frequency of 1.1. Fig. 1.2 shows that 1-dB in-band local minima is allowed, but here we see A_{lm} less than 1 V/V does not meet the specifications. Compared to the second-order reference, Case 1 rolls off much faster as it is a 4-pole system, resulting in a lower bandwidth.

Case 2 (solid orange curve) sets ω_1 at max on the circle (i.e., $\sigma_1 = 1.038$ and

$|\omega_1| = 1.038$). The peak amplitude due to p_3 is reduced to 3.059 V/V at the frequency of 2.6. Since p_1 resides closer to the $j\omega$ axis, it introduces a second peak (near the frequency of 1), which resolves the in-band dipping seen in Case 1 (i.e., for all frequencies below ω_{pk}, the gain is greather than 1 V/V).

Case 3 (cyan dashed curve) has σ_1 equal to σ_3 (i.e., $\sigma_1 = 0.13$ and $|\omega_1| = 0.503$). Since the assumption of $|\sigma_3| \ll |\sigma_1|$ is no longer valid, p_1 instead dominates the peaking profile. The first peak has an amplitude of 2.138 V/V at a frequency of 0.486, which leads to a bandwidth of 0.8.

Due to the zero, z_1 at $-\sigma_z$, Eq. 3.21 and Eq. 3.24 has three unknowns (σ_z, σ_1 and ω_1) and two equations, which does not have an explicit solution (i.e., unlike the 4-pole, double-series network which has two equations and two unknowns). Additional assumptions are needed to simplify Eq. 3.24.

Recall that p_1 may be pushed towards the $j\omega$ axis to eliminate the in-band local minima at the cost of reducing the peak amplitude at ω_{pk}. Therefore, σ_1 may be selected for A_{lm} greater than 1 V/V (i.e., Case 1 and 2 in the design example above). Note that σ_1 may be selected near A_{lm} of 1 V/V to gain more design margin (i.e., avoid trading off too much A_{pk} for optimizing A_{lm}). With the assumed value for σ_1, the relation between z_1 and A_{pk} may be solved.

For simplicity, ω_1 is set at maximum (i.e., Case 2) to suppress the potential in-band ripple caused by pole separation between p_1 and p_3. This results in σ_1 of 1.038 and $|\omega_1|$ of 1.038 from Eq. 3.25. Poles p_1 and p_2 may then the utilized to evaluated A_{pk} and amplitude at the -3dB frequency (i.e., A_{η_b}, where η_b is the -3dB bandwidth) in terms of σ_z using Eq. 3.24 (i.e., the zero z_1 is located at $-\sigma_z$), which leads to

$$A_{pk} \approx 3.054\frac{\sqrt{\sigma_z^2 + 6.76}}{\sigma_z} \text{ for } \sigma_z > 0, \tag{3.26}$$

$$A_{\eta_b} = \frac{14.657\frac{1}{\sigma_z}\sqrt{\sigma_z^2 + \eta_b^2}}{[1.082 + (\eta_b + 1.04)^2][1.082 + (\eta_b - 1.04)^2][0.017 + (\eta_b + 2.6)^2][0.017 + (\eta_b - 2.6)^2]} = \frac{1}{\sqrt{2}}$$
$$\text{for } \sigma_z > 0 \text{ and } \eta_b > 0. \tag{3.27}$$

The positive-real solutions to Eqs. 3.26 and 3.27 are plotted in Fig. 3.5. Both normalized bandwidth (η_b) and peak amplitude (A_{pk}) appear on parabolic curves. The black curve represents η_b, which is greater than 3 for σ_z lower than 9.177. With the zero z_1 located on the negative-real axis in the s-plane, A_{pk} is always greater than 3.054 (shown in the blue

curve). To attain the required A_{pk} and η_b from Table 3.1 simultaneously, z_1 is therefore -9.177 $< z_1 <$ 0.

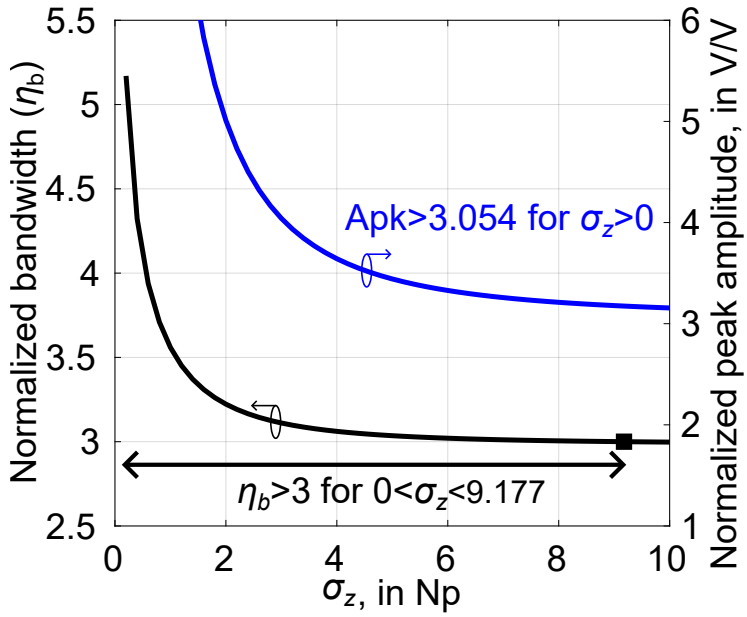


Figure 3.5: Normalized bandwidth and peak amplitude against various σ_z for the 4-pole, 1-zero, shunt/double-series network design example (Assumptions: $\omega_{pk} = \omega_3 = 2.6$, $\zeta_3 = 0.05$ and $\sigma_1 = 1.038$. Bandwidth is normalized to 1 rad/s)

With p_1 to p_4 from the previous assumptions and calculations, given that

$$\begin{aligned} p_1 &= -1.04 + j1.04, \\ p_2 &= -1.04 - j1.04, \\ p_3 &= -0.13 + j2.6, \\ p_4 &= -0.13 - j2.6, \\ -9.177 &< z_1 < 0, \end{aligned} \tag{3.28}$$

p_1 to p_4 are substituted to Eq. 3.9 to solve for coefficients in the 4th-order polynomial. The

coefficients are

$$\begin{aligned} a_3 &= 2(\sigma_1 + \sigma_3) = 2.34 \\ a_2 &= (\sigma_1^2 + \omega_1^2) + (\sigma_3^2 + \omega_3^2) + 4\sigma_1\sigma_3 = 9.481, \\ a_1 &= 2\sigma_3(\sigma_1^2 + \omega_1^2) + 2\sigma_1(\sigma_3^2 + \omega_3^2) = 14.66, \\ a_0 &= (\sigma_1^2 + \omega_1^2)(\sigma_3^2 + \omega_3^2) = 14.66. \end{aligned} \tag{3.29}$$

Equating coefficients a_3 to a_0 with denominator coefficients of the 4-pole, 1-zero, shunt/double-series transfer function Eq. 3.1 and substituting Eq. 3.2 yield

$$\begin{aligned} a_3 &= \frac{\beta}{\alpha} = \frac{m_1 + m_2}{m_1m_2 + m_1m_3 + m_2m_3} = 2.34 \\ a_2 &= \frac{\gamma}{\alpha} = \frac{(1-n)(m_1 + m_3) + n(m_2 + m_3)}{n(1-n)(m_1m_2 + m_1m_3 + m_2m_3)} = 9.481, \\ a_1 &= \frac{\delta}{\alpha} = \frac{1}{\alpha} = \frac{1}{n(1-n)(m_1m_2 + m_1m_3 + m_2m_3)} = 14.66, \\ a_0 &= \frac{1}{\alpha} = \frac{1}{n(1-n)(m_1m_2 + m_1m_3 + m_2m_3)} = 14.66. \end{aligned} \tag{3.30}$$

Recall Eq. 3.7 for the numerator and substituting z_1 yields

$$\begin{aligned} z_1 = -\sigma_z = -\frac{1}{\epsilon} = -\frac{1}{m_3}, \\ m_3 = \frac{1}{\sigma_z}, \\ m_3 > 0.109. \end{aligned} \tag{3.31}$$

There are four unknowns (m_1 to m_3 and n) and three equations (a_1 to a_3) in Eqs. 3.30 and 3.31. Therefore, there will always be mathematical solutions for a given n value, which is known from the capacitances of the source and load stages (i.e., $n = C_2/C_T$). That is, parametric tuning is no longer needed, and only one design iteration is sufficient (if the results are realizable). This is advantageous compared to traditional inductive peaking techniques which involve fitting the transfer function coefficients to known filter coefficients to attain frequency responses such as Butterworth (maximally flat ampltiude), Chebyshev (equal ripple) and Bessel (maximally flat group delay), and often require a specific n value.

Substituting $n = 0.7$ to Eq. 3.30 yields two solutions

$$\begin{array}{llll} n = 0.7 & m_1 = 0.319 & m_2 = 0.441 & m_3 = 0.242, \\ n = 0.7 & m_1 = 0.745 & m_2 = 0.015 & m_3 = 0.413. \end{array}$$

Equation 3.31 may be used to validate these solutions. In this case, both solutions have $m_3 > 0.109$, which means that they would both meet the specifications in Table 3.1. This approach is very robust when interfacing any non-flexible gain stages (i.e., other stages may prioritize other design objectives without worrying about the interstage. The interstage network is designed to interface stages with various values of n, while achieving the desired peak amplitude, bandwidth, and in-band ripple). Note that some solutions may have m_3 less than 0.109, indicating that they would not meet the specifications in Table 3.1. Some solutions may have negative values for one of m_1 or m_2 (i.e., negative inductance), which is nonphysical. To synthesize an equivalent network with one of m_1 and m_2 being negative, coupled inductors may be implemented that utilize mutual inductance to produce the same effect at the terminals. An example is provided for solving Eqs. 3.9 and 3.30 with n of 0.9 and the resulting m_2 is negative:

$$\begin{array}{llll} n = 0.9 & m_1 = 1.27 & m_2 = 0.503 & m_3 = 0.067, \\ n = 0.9 & m_1 = 1.92 & m_2 = -0.15 & m_3 = 0.588. \end{array}$$

To synthesize the second solution with m_2 of -0.15, the equivalent network with parameters m_1' to m_3' may be given by

$$\begin{aligned} m_1' &= m_1 - m_2', \\ m_2' &= -m_2 = k_{13}\sqrt{m_1' m_3'} \\ m_3' &= m_3 - m_2'. \end{aligned} \tag{3.32}$$

The solution for m_1' and m_3' coupled with a factor k_{13} of 0.17 is:

$$n = 0.9 \qquad m_1' = 1.77 \qquad m_2' = 0.15 \qquad m_3' = 0.438,$$

and the equivalent network is shown in Fig. 3.6.

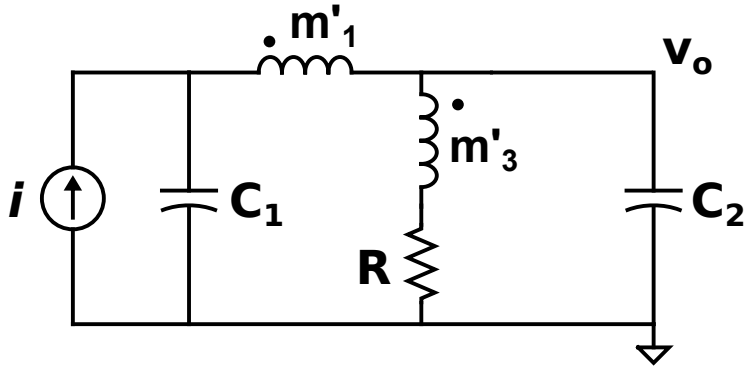


Figure 3.6: Shunt/double-series network implemented with coupled coils m_1' of 1.77 and m_3' of 0.438 with coupling factor k_{13} of 0.17 for a capacitor ratio $n = C_2/(C_1 + C_2)$ of 0.9 (equivalent to Fig. 3.2b with m_1, m_2 and m_3 of 1.92, -0.15 and 0.588, respectively)

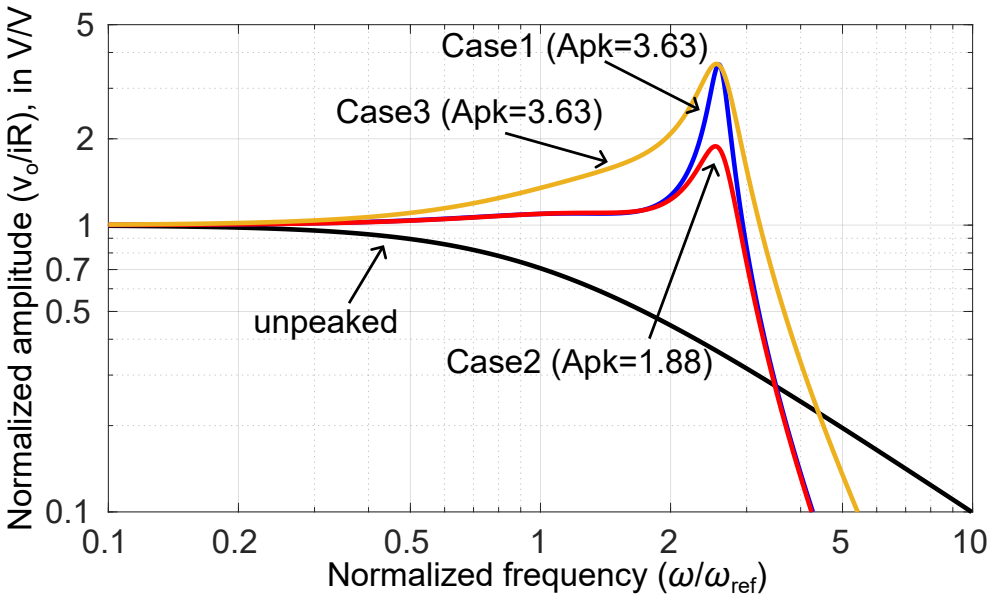


Figure 3.7: Normalized magnitude response of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$)

Normalized magnitude responses are depicted in Fig. 3.7 including three cases. Case 1 (blue curve) is the solution of substituting n of 0.7 to Eq. 3.30 from the previus design example (i.e., $\zeta_3 = 0.05, n = 0.7$ results in $m_1 = 0.319, m_2 = 0.441, m_3 = 0.242$). Since m_3

meets the condition in Eq. 3.31, the design goals from Table 3.1 should be met. Case 2 (red curve) shows the response of increasing ζ_3 to 0.1 while n remains the same (i.e., $\zeta_3 = 0.1, n = 0.7$ yields $m_1 = 0.272, m_2 = 0.529, m_3 = 0.194$). Lastly, Case 3 keeps ζ_3 at 0.1, and adjusts n and m_3 to attain a larger bandwidth while having the peak amplitude remain same as Case 1 (i.e., $\zeta_3 = 0.1, n = 0.93$ and $m_1 = 2.931, m_2 = -0.339, m_3 = 0.756$. Note that this network requires two coils with values of 2.592 and 0.417 and a mutual coupling factor of -0.326). The black curve shows the unpeaked response with the -3dB bandwidth normalized to 1 rad/s (i.e., $L_{1-3} = 0$ and the network becomes an RC circuit).

Case 1 has a peak amplitude A_{pk} of 3.63 V/V with a bandwidth of 3.057. The peaking frequency is 2.59, which is close to ω_3 of 2.6 (i.e., the assumption of $\omega_{pk} \approx \omega_3$ is valid). As pole p_1 is selected to maximize ω_1 from Eq. 3.25 ($p_1 = -1.04 + j1.04$), the in-band dipping effect is minimized. The hump near the frequency of 1.5 may be explained by the peaking effect due to p_1. Overall, Case 1 meets all the specifications proposed in Table 3.1 while the resulting n complies with the assumption made in Eq. 3.19 (i.e., no extra iterations are needed). An optimal location of p_1 may be selected to ensure that A_{lm} is greater than 1 V/V without causing another peak (i.e., pick p_1 with $\sigma_1 > 1.04$ and $0 < |\omega_1| < 1.04$).

Case 2 (red curve) has a -3dB bandwidth of 3.017 with a peak amplitude of 1.88 V/V at a frequency of 2.54. There is no in-band local minima, as this solution is found using the same methodology as Case 1 (i.e., $A_{lm} > 1$ V/V for all frequencies below ω_{pk}). Compared to Case 1, A_{pk} decreases due to the increase in ζ_3 (i.e., more damping leads to less peaking). Note that the bandwidth in Case 2 does not differ from Case 1 by much (3.017 from Case 2 versus 3.057 from Case 1).

Case 3 (orange curve) achieves an A_{pk} of 3.65 V/V which is close to Case 1, and its bandwidth is improved to 3.42. The increase in bandwidth is attributed to two factors. The first factor is the increase in m_3 (i.e., decrease in σ_z since $m_3 = 1/\sigma_z$ from Eq. 3.31). As seen in Fig. 3.5, the normalized bandwidth is inversely proportional to σ_z. Therefore, decreasing σ_z favours wider bandwidth. The second factor is the increase in ζ_3. Recall that for a second-order system, the bandwidth is proportional to the natural undamped frequency ω_n which equals the magnitude of the pole (Eqs. 3.15 and 3.17). Therefore, increasing ζ_3 causes $|p_3|$ to increase and improve the bandwidth. Note that similar bandwidth improvement is not seen in Case 2, as Case 2 has ζ_3 increased and m_3 decreased. The drawback of Case 3 is that it requires an n of 0.93, which disagrees with the assumption, and therefore tuning the driving or load capacitances is required (i.e., may disrupt the design of other amplifier stages).

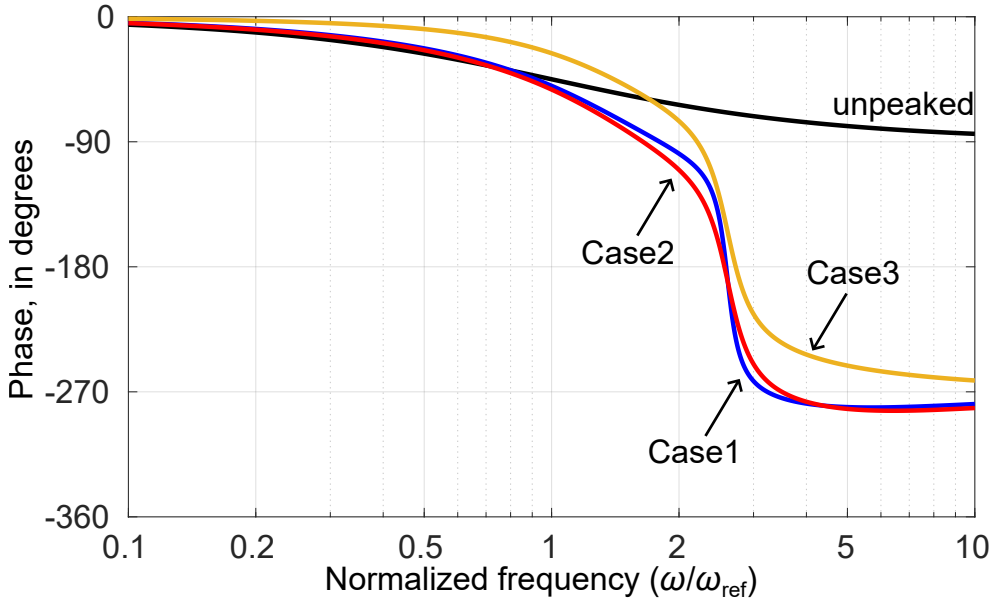


Figure 3.8: Phase response of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$)

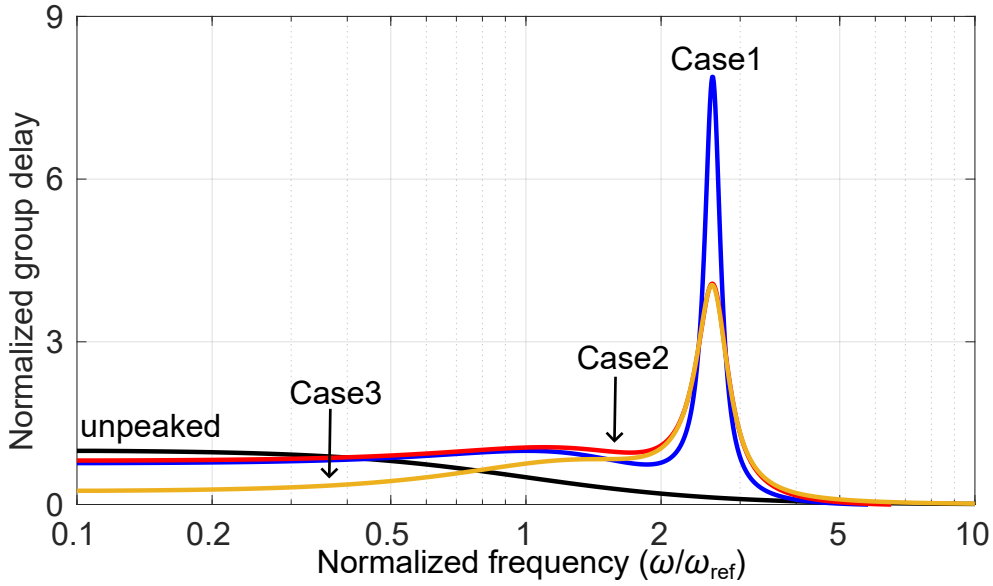


Figure 3.9: Normalized group delay of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$)

The phase and group delay are shown in Fig. 3.8 and Fig. 3.9, respectively. The unpeaked case (black curve) has a slope of -45 degree/decade in phase with a total shift of -90 degrees (i.e., 1-pole low-pass response). Case 1 to Case 3 all have a total phase shift of -270 degrees due to the 4 poles and 1 zero. Due to the dominant pole p_3, a steep phase change is seen near ω_3 of 2.6 in the three peaked cases, leading to a peak in the group delay. Case 1 has the worst peak group delay because it has the lowest ζ_3 of 0.05 (and the highest A_{pk}). Compared to Case 1, the peak group delay for Case 2 is much lower, as ζ_3 for Case 2 increases to 0.1 and A_{pk} is lower than Case 1. Case 3 has the largest m_3 (lowest σ_z) among all three cases. Therefore, the effect of zero shows up earlier in frequency, which leads to a gradual decrease in phase at low frequencies (i.e., lower group delay at low frequencies). At ω_{pk}, Case 3 has the same ζ_3 of 0.1 as Case 2, which leads to a similar peak group delay.

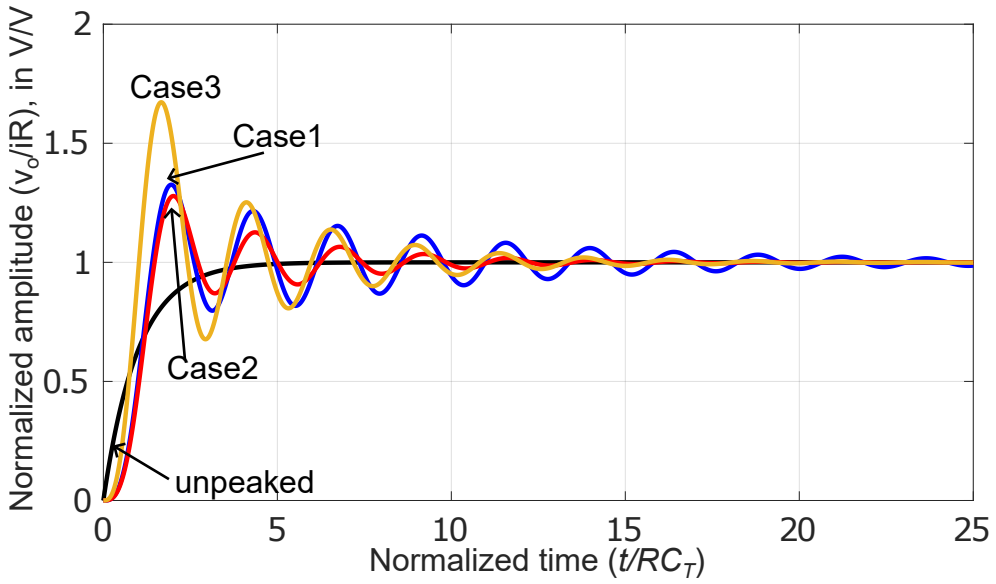


Figure 3.10: Normalized step response of three 4-pole, 1-zero, shunt/double-series networks (Case 1: $\zeta_3 = 0.05, n = 0.7$, Case 2: $\zeta_3 = 0.1, n = 0.7$, Case 3: $\zeta_3 = 0.1, n = 0.93$)

The step responses are shown in Fig. 3.10. In the three peaked circuits, the 10%-90% risetimes are improved from 2.2 to 0.776, 0.817 and 0.577, respectively. Case 3 has the fastest risetime due to its largest bandwidth. The drawback of overshoot has become worse for shunt/double-series networks compared to the double-series network due to the additional zero, as the zero speeds up the circuit and causes extra overshoot [56]. For example, the overshoot for Case 3 is 66.9% which is much worse than the other two cases (32.6% for Case 1 and 27.7% for Case 2) due to its zero being closest to the $j\omega$ axis.

The settling times for Case 1 to Case 3 are 21.44, 10.68 and 13.87, respectively, and the delay times are 0.994, 1.02 and 0.733, respectively. Both specifications are worse than the unpeaked case. Note that the delay time of Case 3 is better than the other two cases, which is due to the negative m_2 in series with the load capacitor causing a reduction in delay.

Comparisons in Figs. 3.7 and 3.10 indicate that the risetime is proportional to m_3 (i.e., the shunt peaking inductance). This is because m_3 determines the location of the zero. As m_3 increases, the zero moves towards the $j\omega$ axis and bandwidth increases. In time domain, m_3 introduces an extra delay between the resistor and the capacitors, forcing the circuit to charge C_1 and C_2 at a faster rate during the start. However, faster risetimes lead to a drawback of large overshoots. In this thesis, the primary goal is to design a broadband driver with amplitude peaking to compensate for the high-frequency losses of the optical modulator. Therefore, a high design priority is given to sufficient peak amplitude and wide bandwidth, which makes large overshoot inevitable. The targeted percentage overshoot is set to less than 43% (i.e., the percentage overshoot of an arbitrary second-order system with 2-V/V normalized peak amplitude at 2.6 rad/s).

3.1.2 Interstage design example for the driver

In this section, design parameters of the interstage filter for the driver are derived. These parameters are used in Chapter 4 to guide the physical design of the interstage. Magnitude responses of the dominant and nondominant second-order transfer functions are plotted to show the effect of parametric tuning to adjust the peak amplitude.

3.1.2.1 Reconstruct the 4-pole, 1-zero transfer function for the shunt/double-series interstage network using two second-order equations

The shunt/double-series interstage network transfer function Eq. 3.1 may be rearranged as a product of dominant and non-dominant second-order transfer functions, given by

$$\begin{aligned} F(s) &= F_{so1}(s) \cdot F_{so2}(s) \\ &= \frac{1}{(s-p_3)(s-p_4)} \cdot \frac{k\cdot(s-z_1)}{(s-p_1)(s-p_2)} = \frac{1}{s^2+b\cdot s+c} \cdot \frac{k\cdot(s-z_1)}{s^2+d\cdot s+e} \\ &= \frac{k\cdot s - k\cdot z_1}{s^4+(b+d)s^3+(c+e+b\cdot d)s^2+(c\cdot d+b\cdot e)s+c\cdot e}, \\ &= \frac{s\cdot\epsilon/(R\cdot\alpha)+1/\alpha}{s^4+s^3\cdot\beta/\alpha+s^2\cdot\gamma/\alpha+s\cdot\delta/\alpha+1/\alpha} \qquad \text{(ref. Eq. 3.1)}, \end{aligned} \tag{3.33}$$

where

$$k = \left| \frac{\prod_1^n p_i}{\prod_1^m z_i} \right|, \tag{3.34}$$

$$p_1 = -\sigma_1 + j\omega_1 \qquad p_2 = -\sigma_1 - j\omega_1 \qquad p_3 = -\sigma_3 + j\omega_3 \qquad p_4 = -\sigma_3 - j\omega_3,$$

$$b = 2\sigma_3 \qquad c = \sigma_3^2 + \omega_3^2 \qquad d = 2\sigma_1 \qquad e = \sigma_1^2 + \omega_1^2. \tag{3.35}$$

Parameter k is a gain compensation factor to normalize the DC gain of the interstage network to 0 dB. Dominant poles p_3 and p_4 deliver the peaking (i.e., dominant second-order transfer function $F_{so1}(s) = 1/(s - p_3)(s - p_4)$).

From Eq. 3.33, the following equalities may be derived:

$$\begin{aligned} b + d &= \frac{\beta}{\alpha}, \\ c + e + b \cdot d &= \frac{\gamma}{\alpha}, \\ c \cdot d + b \cdot e &= \frac{\delta}{\alpha}, \\ c \cdot e &= \frac{1}{\alpha}. \end{aligned} \tag{3.36}$$

Recalling from Eqs. 3.10 and 3.11 that $\frac{\delta}{\alpha} = \frac{1}{\alpha}$, and substituting p_1 to p_4 into Eq. 3.36 yields

$$\begin{aligned} c \cdot d + b \cdot e &= c \cdot e, \\ \frac{b}{c} + \frac{d}{e} &= 1, \\ \frac{2\sigma_3}{\sigma_3^2 + \omega_3^2} + \frac{2\sigma_1}{\sigma_1^2 + \omega_1^2} &= 1, \\ 1 + \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4} &= 0, \end{aligned} \tag{3.37}$$

which matches Eq. 3.12. Therefore, reconstruction of the 4-pole, 1-zero, shunt/double-series transfer function using two second-order equations is verified.

3.1.2.2 Determining the normalized peak frequency ω'_{pk} and the root locus of non-dominant poles p_1 and p_2

In the special case where parameter b in Eq. 3.33 is assumed to be zero, the transfer function simplifies to

$$\begin{aligned} F_{b=0}(s) &= F_{so1,b=0}(s) \cdot F_{so2}(s) \\ &= \frac{1}{s^2 + c} \cdot \frac{k \cdot (s - z_1)}{s^2 + d \cdot s + e} \\ &= \frac{k \cdot s - k \cdot z_1}{s^4 + d \cdot s^3 + (c + e)s^2 + (c \cdot d)s + c \cdot e}, \end{aligned} \tag{3.38}$$

In the dominant second-order polynomial $F_{so1}(s)$, as shown in Eq. 3.35, when parameter b is zero, parameter c equals ω_3^2, which sets ω_3 and the normalized peaking frequnecy ω'_{pk} of the interstage network.

For this example, it is assumed that the input (C_1) and load (C_2) capacitances of the interstage network are 38 fF and 184 fF, respectively (i.e., $n = 184/(184 + 38) = 0.83$), and the interstage load R is 12 Ω, as shown in Fig. 3.11 (justification of these values are presented in Chapter 4). Parasitic capacitance (C_{par}) of 12 fF is also included.

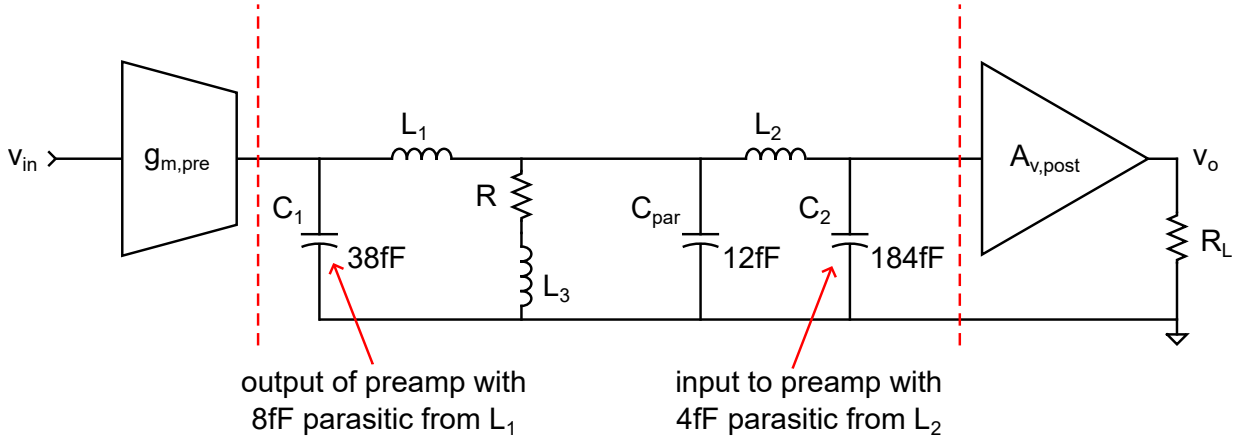


Figure 3.11: Single-ended lumped-element model of the shunt/double-series interstage network

With R, C_1, C_2 and C_{par} known, ω_{ref} and ω'_{pk} are

$$\begin{aligned} \omega_{ref} &= \frac{1}{2\pi\frac{R}{2}C_T} = \frac{1}{2\pi\frac{R}{2}(C_1+C_2+C_{par})} = \frac{1}{2\pi\cdot 12\cdot 234\cdot 10^{-15}} = 56.7 \text{ GHz}, \\ \omega'_{pk} &= \frac{\omega_{pk}}{\omega_{ref}} = \frac{2\pi\cdot 80 \text{ GHz}}{\omega_{ref}} = 1.42, \end{aligned} \tag{3.39}$$

assuming a desired peak at 80 GHz (i.e., adding 5-GHz margin to $\omega_{pk} = 75$ GHz). The interstage is expected to have a peak of at least 12.8 dB (4.36 V/V) at 75 GHz for the driver.

For a desired ω'_{pk} of 1.42 (from Eq. 3.39), the corresponding value of c is 2.02 (i.e., $p_{3,4} = \pm j1.42$ from Eq. 3.35). By substituting p_3 and p_4 to Eq. 3.37, the root locus of the non-dominant poles p_1 and p_2 is defined by the following equations

$$\omega_1 = \sqrt{-\sigma_1^2 - 2\sigma_1}, \tag{3.40}$$

$$\omega_2 = -\sqrt{-\sigma_1^2 - 2\sigma_1}, \tag{3.41}$$

which describe a circle centred at (-1,0) in the s-plane ($\sigma = -1$, $\omega = 0$) with a radius of 1. Poles p_1 and p_2 are selected at $-0.2 \pm j0.6$ to introduce an in-band peak that eliminates any potential in-band dip in the magnitude response caused by the separation of the dominant and non-dominant poles.

Magnitude responses of the two second-order transfer functions and the overall response are plotted in Fig. 3.12. Note that (real) zero z_1 is placed at -0.807, resulting in a gain of 0 dB at DC for simplicity. The blue dashed curve in Fig. 3.12 represents the frequency response of the dominant second-order transfer function. The normalized peak frequency is at 1.42 with a peak amplitude approaching infinity because p_3 and p_4 are assumed to lie on the imaginary axis in the s-plane. The red dashed curve shows the response of the non-dominant, second-order polynomial with an in-band peak at a normalized frequency of 0.6, which corresponds to the peaking introduced by p_1. The overall response is shown in black, and the peaking profile is determined by the dominant second-order equation, as expected.

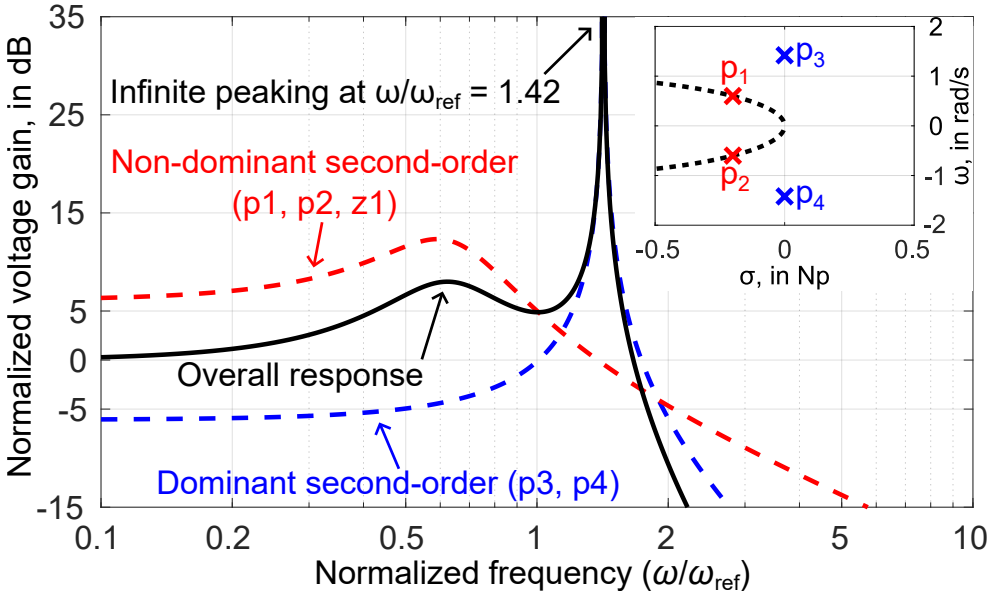


Figure 3.12: Magnitude responses of the two second-order transfer functions from Eq. 3.33. Parameter $c = 2.02$ is used to set the normalized peak frequency ω'_{pk} while parameter b is 0 ($p_{1,2} = -0.4 \pm j0.6$, $p_{3,4} = \pm j1.42$). The real zero is used to set the DC gain to zero ($z_1 = -0.807$).

Parameter c is used to tune ω'_{pk}. When parameter b is neglected (i.e., $b = 0$ assumed), Eq. 3.36 yields

$$c + e = \frac{\gamma}{\alpha}, \tag{3.42}$$

$$c \cdot e = \frac{1}{\alpha}, \qquad \text{or} \qquad e = \frac{1}{\alpha c}, \tag{3.43}$$

and c may be solved from substitution of Eq. 3.43 into Eq. 3.42 as

$$c = \frac{\frac{\gamma}{\alpha} \pm \sqrt{(\frac{\gamma}{\alpha})^2 - \frac{4}{\alpha}}}{2}. \tag{3.44}$$

Note that, ω'_{pk} is heavily dependent on the ratio γ/α. From Eq. 3.4 where coefficients α, β, γ, δ are represented in terms of normalized circuit parameter m_1 to m_3 (i.e., inductance

of L_1 to L_3) and n (capacitance ratio, $0 < n < 1$), γ/α may be expressed as

$$\begin{aligned}\frac{\gamma}{\alpha} &= \frac{(1-n)m_1 + nm_2}{n(1-n)(m_1m_2 + m_1m_3 + m_2m_3)} \\ &\approx \frac{1}{m_2 \cdot n} + \frac{1}{m_1 \cdot (1-n)} \qquad \text{(when } m_3 \text{ is neglected)},\end{aligned} \tag{3.45}$$

Equation 3.45 indicates that normalized inductances m_1 to m_3 and the capacitor ratio n determine γ/α. When n is large (i.e., load capacitor C_2 in Fig. 3.11 dominates), m_1 (i.e., inductor L_1) tunes the normalized peak frequency ω'_{pk}. When n is small (i.e., source capacitor C_1 dominates), m_2 is used to set ω'_{pk}.

3.1.2.3 Determine the peak amplitude A_{pk} for finite parameter b

In a realistic case, parameter b is greater than zero due to the interstage load resistance R of 12 Ω (single-ended) and losses in passive components. Recall that parameters b, c and the damping factor ζ_3 of the dominant second-order transfer function can be expressed in terms of σ_3 and ω_3 (i.e., location of dominant pole p_3). As expressed in Eqs. 3.15 and 3.35, ζ_3 may be written as

$$\zeta_3 = \frac{\sigma_3}{\sqrt{\sigma_3^2 + \omega_3^2}} = \frac{b}{2\sqrt{c}}. \tag{3.46}$$

The peak amplitude A_{pk} for an underdamped second-order transfer function is determined by the damping factor (see Eq. 3.17). Given parameter c of 2.02, when b is set to 2.44×10^{-2}, the corresponding ζ_3 is 8.6×10^{-3} (Eq. 3.46), resulting in an A_{pk} of 35.3 dB (58.1 V/V) at a normalized frequency of 1.42 using Eq. 3.17. The updated magnitude response is shown in the blue dashed curve in Fig. 3.13. Dominant poles p_3 and p_4 are $-0.0122 \pm j1.42$. Compared to Fig. 3.12 where b is zero, A_{pk} becomes finite, while ω'_{pk} remains unchanged. With the updated p_3 and p_4, the new root locus of the non-dominant poles p_1 and p_2 is given by

$$\omega_1 = 0.1417\sqrt{-49.8\sigma_1^2 - 100.83\sigma_1}, \tag{3.47}$$

$$\omega_2 = -0.1417\sqrt{-49.8\sigma_1^2 - 100.83\sigma_1}. \tag{3.48}$$

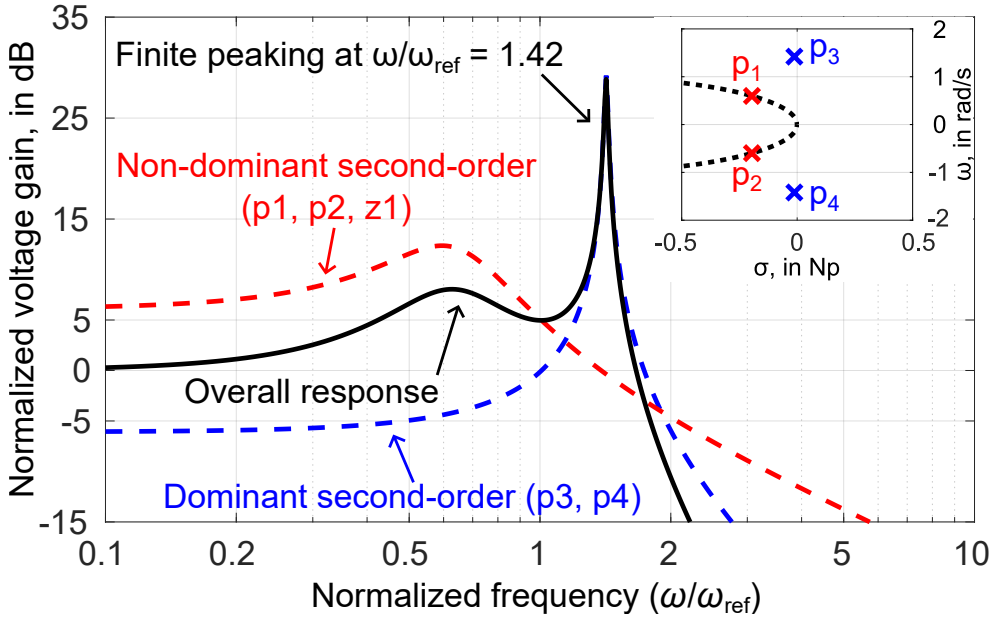


Figure 3.13: Magnitude responses of the two second-order transfer functions from Eq. 3.33. Parameter $c = 2.02$ is used to set the normalized peak frequency ω'_{pk} while parameter b is 2.44×10^{-2} ($p_{1,2} = -0.196 \pm j0.598$, $p_{3,4} = -0.0122 \pm j1.42$). The zero is used to set the DC gain to unity ($z_1 = -0.8$).

Note that assuming non-dominant poles p_1 and p_2 at $-0.196 \pm j0.598$, (i.e., slightly shifted from $-0.2 \pm j0.6$ when $b = 0$), the value of A_{pk} (black solid curve) is 28.8 dB due to 6.5-dB of loss contributed by the non-dominant second-order polynomial of the transfer function (red dashed curve) at ω'_{pk} of 1.42.

3.1.2.4 Determining the zero location and convert the interstage filter transfer function to normalized circuit parameters

Thus far, the zero (z_1) has not yet been utilized in the design. Poles p_1 to p_4 are substituted to Eq. 3.24 to correlate σ_z, A_{pk} and the -3dB bandwidth. When $|\sigma_z|$ is set to 0.633, the predicted A_{pk} is 33 V/V or 30.4 dB (up from 28.8 dB when z_1 is not included), and the predicted -3dB bandwidth of the interstage network is 1.786, or 101 GHz (denormalized). Compared to the required peak amplitude of 12.8 dB (9-dB specification with 3.8-dB roll-off compensation for the preamp and postamp at 75 GHz), the margin for A_{pk} is 17.6 dB The predicted bandwidth meets the 90-GHz requirement. Frequency responses of the non-dominant second-order transfer function and the overall transfer function are shown in

dash red curve and black solid curve in Fig. 3.14, respectively. Compared to the previous case with z_1 at -0.8 in Fig. 3.13, when z_1 moves to -0.633 (toward the $j\omega$ axis in the s-plane), A_{pk} at ω'_{pk} increases due to the increased gain from the non-dominant second-order transfer function. To maintain a DC gain of 0 dB, the gain compensation factor k is set to 1.26 according to Eq. 3.34.

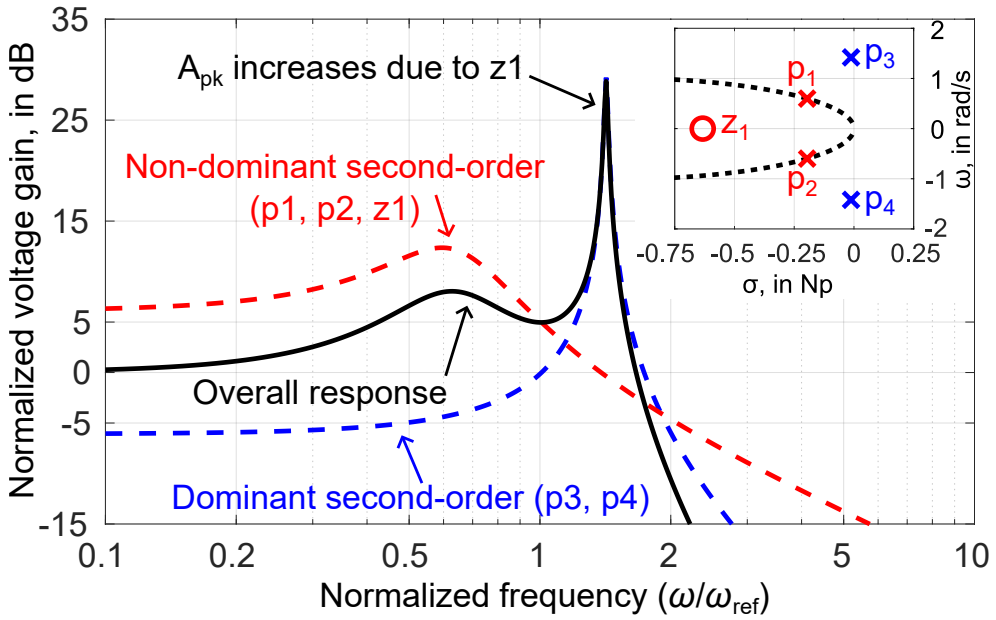


Figure 3.14: Magnitude responses of the two second-order transfer functions from Eq. 3.33. Parameter $c = 2.02$ is used to set the normalized peak frequency ω'_{pk} while parameter b is 2.44×10^{-2} ($p_{1,2} = -0.196 \pm j0.598$, $p_{3,4} = -0.0122 \pm j1.42$). The zero is used to tune A_{pk} ($z_1 = -0.633$).

Substituting n of 0.83 (Eq. 3.3), and p_1 to p_4 to Eqs. 3.9 and 3.30 yields the following normalized design parameters:

Solution-1:	$n = 0.83$	$m_1 = 2.42$	$m_2 = 1.25$	$m_3 = 1.58$,
Solution-2:	$n = 0.83$	$m_1 = 3.7$	$m_2 = -0.01$	$m_3 = 2.41$.

Solution-1 is chosen, as it has a smaller m_1 (less series loss in L_1 leads to better voltage

headroom for the preamp). The transfer functions corresponding to m_1 to m_3 and n are

$$F(s) = \frac{1.269s + 0.803}{s^4 + 0.416s^3 + 2.433s^2 + 0.803s + 0.803}, \tag{3.49}$$

$$F_{so1}(s) = \frac{1}{(s - p_3)(s - p_4)} = \frac{1}{s^2 + 0.0244s + 2.017}, \tag{3.50}$$

$$F_{so2}(s) = \frac{k \cdot (s - z_1)}{(s - p_1)(s - p_2)} = \frac{1.2616s + 0.7986}{s^2 + 0.392s + 0.396}, \tag{3.51}$$

where poles, zero and the gain compensation factor k are given by

$$\begin{aligned} p_1 &= -\alpha_1 + j\omega_1 = -0.196 + j0.598, \\ p_2 &= -\alpha_1 - j\omega_1 = -0.196 - j0.598, \\ p_3 &= -\alpha_3 + j\omega_3 = -0.0122 + j1.42, \\ p_4 &= -\alpha_3 - j\omega_3 = -0.0122 - j1.42, \\ z_1 &= -0.633, \\ k &= \left| \frac{\prod_1^n p_i}{\prod_1^m z_i} \right| = \frac{0.799}{0.633} = 1.2616. \end{aligned} \tag{3.52}$$

These parameters are used in the driver designed in Chapter 4. In summary, the circuit parameters n, m_1, m_2 and m_3 are 0.83, 2.42, 1.25 and 1.58, respectively. The corresponding damping factor for the dominant second-order transfer function ζ_3 is 8.6×10^{-3} and the bandwidth extension ratio is 1.786. The peak gain attained so far is 33 V/V or 30.4 dB (i.e., a 17.6-dB margin for a 12.8-dB specification). The gain margin is required to compensate for the coil losses and the loading effect of the postamp, which are detailed in Chapter 4.

3.1.2.5 Summary of the 4-pole, 1-zero, shunt/double-series interstage design methodology

The design methodology of the 4-pole, 1-zero, shunt/double-series interstage network is summarized as follows:

1. Characterize the driving and load stages to obtain C_{in} and C_{out}, which are used as the initial values for C_1 and C_2. Solve for R required for the desired low-frequency gain. The normalized frequency and gain are $1/R(C_1 + C_2)$ and R, respectively. Circuit parameters are converted to variable n, m_1, m_2 and m_3 using Eq. 3.3. Based on the design specifications given, calculate the normalized peaking frequency (ω_{pk}) using Eq. 3.19.

2. Assuming that the frequency of the dominant pole (p_3) ω_3 is equal to ω_{pk}, select a damping factor (ζ_3). The damping factor should be small to ensure that the system has an underdamped response to achieve peaks in amplitude. Locations of p_3 and p_4 may be calculated using Eq. 3.15.

3. Substitute p_3 and p_4 to Eq. 3.24 to solve for σ_1 and ω_1. Poles p_1 and p_2 should lie on a circle as seen in Eq. 3.25 (p_1 and p_2 should locate in the left-hand side of the s-plane to ensure system stability). The location of p_1 ($p_1 = -\sigma_1 + j\omega_1$) should then be selected to minimize the in-band local minima (A_{lm}). For simplicity, selecting max ω_1 minimizes the distance between p_1 and p_3, resulting in lowest A_{lm}. The location of p_1 may be optimized to attain a smooth gain profile.

4. Substitute p_1 to p_4 to Eq. 3.24 to solve the location of the real zero ($z_1 = \sigma_z$). The zero shows a hyperbolic relationship with the peak amplitude (A_{pk}) and the -3dB bandwidth of the system (i.e., a range of σ_z may be found which meets the specifications of A_{pk} and bandwidth). With p_1 to p_4 known, substitute them into Eq. 3.9 to solve the denominator coefficients of a generalized 4-pole transfer function (i.e., a_0 to a_3).

5. There are four unknowns (n and m_1 to m_3) and three equations in Eq. 3.30. To compute these variables, substitute $C_{out}/(C_{in} + C_{out})$ for n to solve m_1 to m_3 (i.e., C_{in} and C_{out} are given in step 1), and use Eq. 3.31 to validate the result of m_3 (i.e., use the inequality of m_3 shown in Eq. 3.31). Circuit parameters C_1, C_2 and L_1 to L_3 can be calculated with Eq. 3.3.

6. Substitute C_1, C_2, and L_1 to L_3 to Eq. 3.1 to plot the amplitude, phase, group delay and step response. The peak amplitude and bandwith should be known from Eqs. 3.21 and 3.22 already. Check the percentage overshoot from the step response to ensure that it meets the specification (ζ_3 and m_3 should be tuned to adjust the percentage overshoot).

7. Due to losses of coils and parasitics added during implementation (e.g., C_1 and C_2 will increase due to coil parasitics and series losses of L_1 to L_3 will decrease the peak amplitude), steps 1 to 6 should be iterated to converge to the final design. That is, the addition of parasitics reduces ω_{ref}, resulting in an increase of ω_{pk} in Step 1. Including series losses of the inductor coil causes the Q-factor to drop and decreases A_{pk}. To have sufficient design margins, a smaller damping factor ζ_3 should be used in Step 2.

The design methodology highlights the flexibility of the design due to the free choice of parameter n (the ratio between driving and load capacitors). That is, according to Eq. 3.30, there are two solutions for any given n value (but this does not mean that the solutions are always realizable due to physical limitations), which is suitable for interfacing any two stages. This advantage is due to the fact that the addition of the shunt inductor L_3 keeps the denominator of the transfer function in the 4-th order. That is, the order does not increase compared to a double-series network (i.e., without L_3), but the number of variables is increased by 1. Therefore, another variable could be substituted (variable n in this case) to solve for L_3.

Fewer design iterations are required, as the resulting circuit parameters correspond to the n value specified in the first step. Therefore, parametric tuning is not required compared to traditional inductive peaking techniques. The only adjustment is to account for parasitics added during the implementation of passive devices (i.e., briefed in Step 7 of the design methodology). Another key feature is that there is a wide range of achievable amplitude peaks with respect to the location of the zero, as shown in Fig. 3.5. A drawback of the design methodology is that the bandwidth extension ratio is limited by the normalized peaking frequency (ω_{pk}/ω_{ref}) and therefore the approach is best suited for systems demanding large ω_{pk}/ω_{ref}. The reason is that the magnitude is expected to roll off at -60 dB/decade from the peak (due to the 4 poles and 1 zero) and therefore the bandwidth extension ratio is directly proportional to ω_{pk}/ω_{ref}.

Circuit parameters and performance are summarized in Table 3.2. The gain local minima and bandwidth extension ratioare represented by A_{lm} and η_b, respectively. Design goals from Table 3.1 are used as a starting point. Note that unlike the 4-pole, double-series design examples from B.2, the 4-pole, 1-zero, shunt/double-series design examples in Table 3.2 do not have any in-band local minima ($A_{lm} < 0$), as the problem of in-band local minima is resolved during the design.

Table 3.2: Circuit parameters and performance summary for 4-pole, 1-zero, shunt/double-series network design examples

Design highlight	ζ_3	n	m_1	m_2	m_3	A_{pk}	ω_{pk}	A_{lm}	η_b
Baseline	0.1	0.7	0.272	0.529	0.194	1.88	2.54	-	3.02
Improved peak	0.05	0.7	0.319	0.441	0.242	**3.63**	2.59	-	3.06
Improved bandwidth	0.1	0.93	2.931	-0.339	0.756	**3.65**	2.55	-	**3.42**
Driver's interstage	8.6×10^{-3}	0.83	2.42	1.25	1.58	**33**	**1.42**	-	**1.786**

Chapter 4

Broadband Driver Design

This chapter details the design of broadband drivers, with the emphasis on the analysis and realization of the shunt/double-series interstage network. The small-signal and large-signal simulated results are compared to the specifications proposed in Chapter 1. Design nonidealities due to unwanted feedback (e.g., through drain-to-source and drain-to-gate capacitors) are discussed.

4.1 Two-stage broadband amplifier with shunt/double-series interstage

The design of a two-stage broadband amplifier is described in this section. A simplified block diagram of the broadband driver is shown in Fig. 4.1. The input matching network and preamp realize broadband impedance match at the input and provide DC bias for the postamp. The interstage filter is responsible for gain peaking and bandwidth extension. Lastly, the postamp and output combining network are designed to achieve the desired gain and the 4-$V_{pp,d}$ linear output swing for the driver.

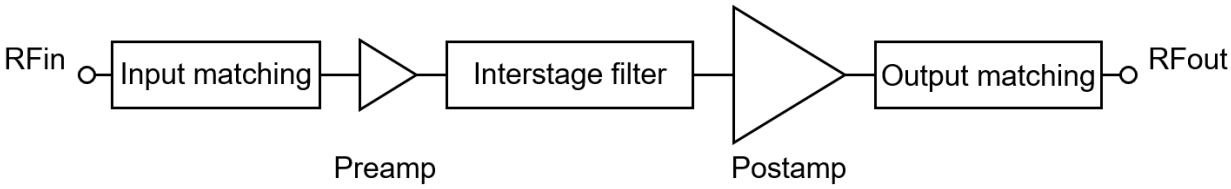


Figure 4.1: Simplified block diagram of the broadband driver

4.1.1 Input matching network and preamp stage

The input stage consists of a PFET cascode amplifier with a shunt/double-series input matching network, as shown in Fig. 4.2. Cascode amplifier M_1 to M_4 is biased by tail current source M_5 and reference diode M_6. Transistors M_1 to M_4 are sized 40-μm wide (W) and 18-nm long (L) with 20×2-μm wide fingers. The common-source gates in Fig. 4.2 are biased at V_G of 1.8 V via off-chip bias-Ts (note that an alternative is to apply V_G at the centre tap of L_3). PFET inputs are utilized to lower the 3-V supply (V_{SSP}) down to a desired common-mode output DC voltage of 0.45 V. The output is DC-coupled to the postamp stage that is implemented with NFETs. Alternating PFET and NFET stages in cascade permits DC coupling if the input and output DC bias voltages are identical. The bias current of each stage is tuned via back gates V_{BG1} and V_{BG2}. The source-to-gate (V_{SG}), source-to-drain (V_{SD}) and source-to-body (V_{SB}) voltages of M_1 to M_4 are 0.45 V, 0.9 V and 2 V, respectively. The differential cascode amplifier is biased at 15 mA per side (I_D), resulting in a current density of 0.021 mA/μm^2 (or I_D/W of 0.375 mA/μm).

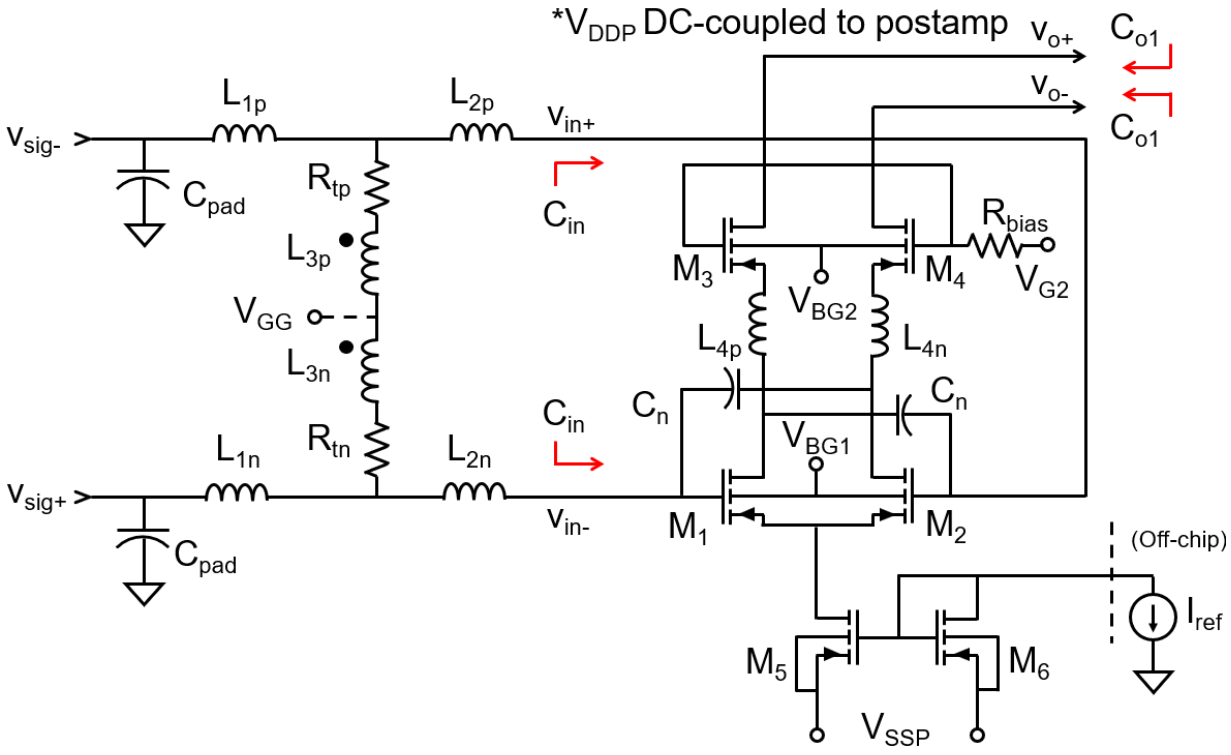


Figure 4.2: Input stage of the two-stage broadband driver with a shunt/double-series input matching network

Current source M_5 is 80-μm wide to supply the DC current and provide common-mode rejection for the cascode amplifier. The back gate of M_5 is connected to its source to reduce the number of backgate bias voltages required. Mirror reference M_6 is biased by an off-chip current source via a DC pad to control the bias of the cascode amplifier (I_{mir} is 6 mA). Cascode devices M_3 and M_4 are biased at V_{G2} of 0.9 V via resistor R_{bias} of 12.5 kΩ. The resistor suppresses the common-mode gain without disrupting the differential mode. A series-peaking inductance $L_{4p,n}$ of 25 pH between the common-source and common-gate devices improves the bandwidth. To improve stage isolation (and thus the driver stability), neutralization capacitors C_n are added to suppress drain-to-gate feedback in M_1 and M_2. The improvement in reverse isolation is shown in Fig. 4.3. To avoid positive feedback due to overcompensation, C_n is selected at 12 fF, resulting in an increase in isolation of 10 dB at 100 GHz (for C_n between 8 fF to 16 fF, the isolation improvement is greater than 7 dB at 100 GHz). The output capacitance of the preamp stage (C_{o1}) is approximately 30 fF, and its effects are addressed during the interstage design in Section 4.1.3.

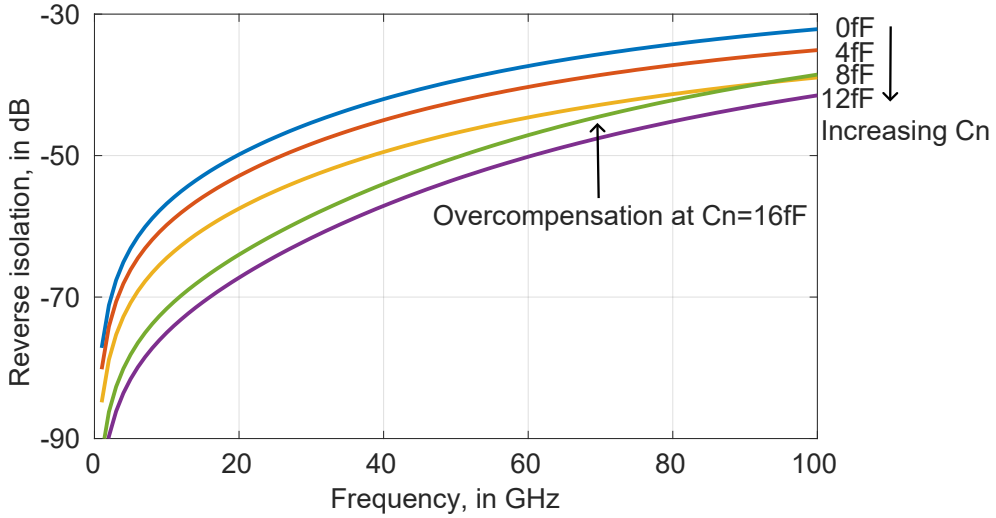


Figure 4.3: Simulated reverse isolation against various neutralization capacitances

The differential shunt/double-series input matching network consists of inductors L_1 to L_3 and 50-Ω on-chip termination resistor $R_{tp,n}$. The inductances of $L_{1p,n}$, $L_{2p,n}$ and L_3 are 115 pH, 11 pH, and 48 pH, respectively. Differential shunt peaking inductors (L_3) are magnetically coupled with a coupling factor of 0.406 (i.e., mutual inductance M of 19.5 pH), increasing the equivalent shunt peaking inductance to 67.5 pH. The source and load capacitors to the shunt/double-series network are input pad capacitance C_{pad} of 24 fF and the equivalent input capacitance of the cascode amplifier C_{in} of 50 fF. The goal of the input

matching network is to meet the specification for return loss (i.e., better than 15 dB up to 50 GHz, and better than 8 dB up to 80 GHz, as shown in Fig. 1.3). The simulated differential transconductance and input return loss are plotted in Fig. 4.4. The transconductance roll-off is 1.1 dB at 75 GHz, which will be peaked by another interstage filter network. The input return loss at 50 GHz and 80 GHz is 24 dB and 13 dB, respectively, giving a margin greater than 5 dB with respect to the design specification.

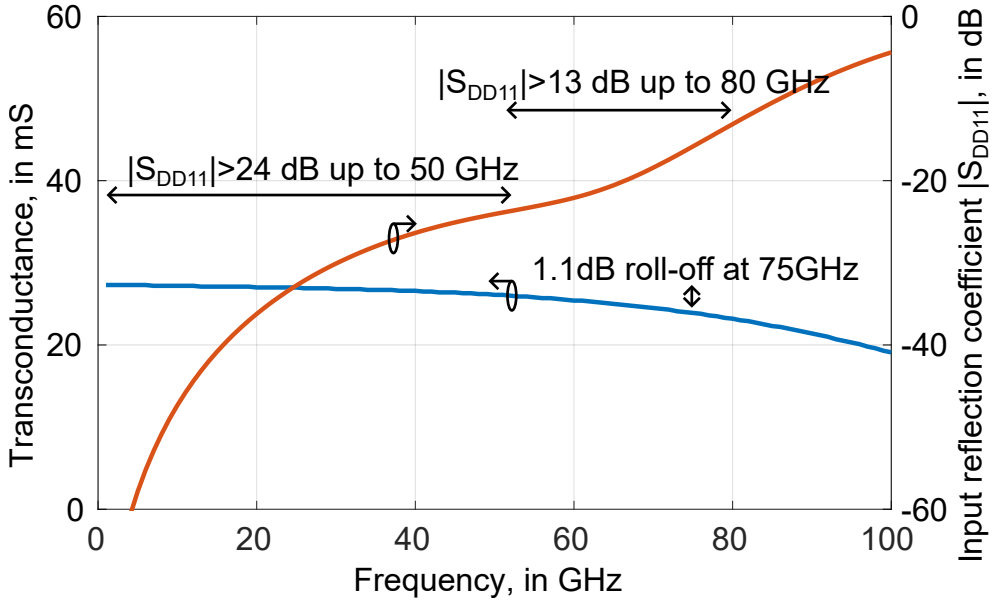


Figure 4.4: Simulated differential transconductance and input return loss of the input stage (with input matching network and preamp stage included, output pole excluded)

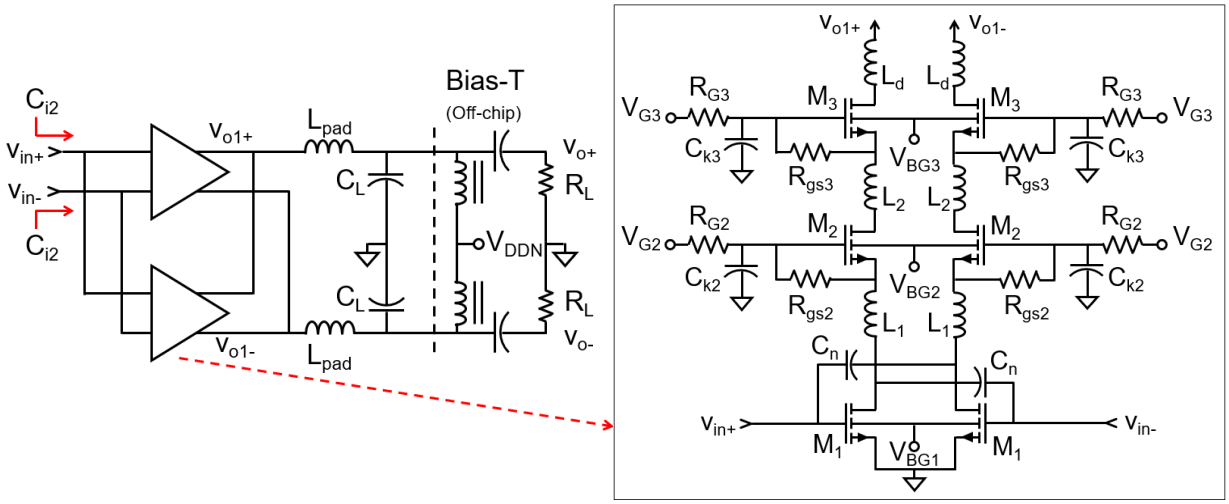
4.1.2 Postamp stage

The parallel-path postamp stage consists of two NFET differential cascode amplifiers combined at the output, as shown in Fig. 4.5. Each cascode amplifier has a common-source FET (M_1) and two common-gate FETs (M_2 and M_3). Compensation capacitors (C_{k2}, C_{k3}) are added in shunt with the gate of M_2 and M_3 to equalize the drain-to-source voltage swings across each FET and avoid breakdown, as explained in Section 2.2.1. Compensation resistors (R_{gs2} and R_{gs3}) are placed between the gate and source of common-gate FETs to reduce the frequency response roll-off at low frequencies. The bias resistor values (R_{G2} and R_{G3}) are selected such that time constants of the gate-to-source network ($R_{gs2,3}$ and $C_{gs2,3}$) and the series gate network ($R_{G2,3}$ and $C_{k2,3}$) are matched. Furthermore, $V_{G2,3}$ is

adjusted to

$$V'_{G2,3} = V_{G2,3} + \frac{V_{GS2,3}}{R_{gs2,3}} \cdot R_{G2,3}, \tag{4.1}$$

which ensures that the operating points of M_2 and M_3 are not disrupted by $R_{gs2,3}$.



(a) circuit block diagram and the output matching network

(b) transconductor unit cell

Figure 4.5: Postamp stage of the two-stage broadband driver

All NFETs are sized at 80-μm width and 18-nm length with 40 × 2-μm wide unit fingers. The gate-to-source (V_{GS}), drain-to-source (V_{DS}) and body-to-source (V_{BS}) voltages of M_1 to M_3 are 0.45 V, 1 V and 2 V, respectively. That is, the total supply voltage (V_{DDN}) is 3 V. The differential cascode amplifier is biased at 40 mA per side, resulting in a current density of 0.027 mA/μm^2 (or I_D/W of 0.5 mA/μm). The drains of M_3 are combined via L_d and L_{pad}, which is connected to an off-chip bias-T to the load. The load resistor R_L is assumed to be 30 Ω for the integrated modulator driver application. Series inductors L_2, L_d and L_{pad} (36 pH, 90 pH and 19 pH) are selected to achieve the desired -1dB compression point at the output (OP_{-1dB}). Inductor L_1 of 18 pH provides series peaking. A load capacitor C_L of 78 fF is required to realize the desired impedance for RF power at 75 GHz. Neutralization capacitors C_n of 18 fF are utilized to suppress the Miller effect and improve isolation at the M_1 inputs. A tail current source is not used for the postamp stage to reduce DC power consumption, and also because the common-mode rejection ratio (CMRR) of the entire

amplifier is dominated by the input stage [57]. The input capacitance of the postamp is approximately 180 fF (C_{i2}), and its effects are addressed later during the interstage design. The simulated voltage gain of the postamp ($A_{v,post}$, i.e., input pole excluded) is plotted in Fig. 4.6. The low-frequency voltage gain is 7.6 V/V and has a roll-off of 2.7 dB at 75 GHz (i.e., will be peaked by the interstage network). The simulated -3dB bandwidth of $A_{v,post}$ is 91 GHz.

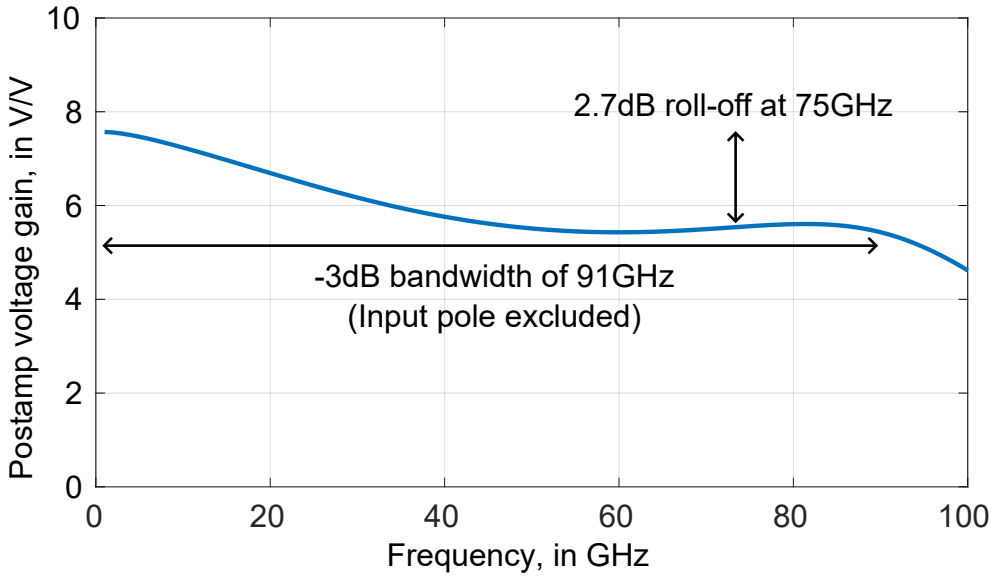


Figure 4.6: Simulated differential voltage gain of the postamp stage (with 60-Ω differential load and input pole excluded)

4.1.2.1 Selection of FET dimension and DC operating point for desired OP_{-1dB} in cascade amplifiers

Recall that the OP_{-1dB} of a Class-A amplifier and a multi-stage amplifier are approximated by Eq. 1.2 and Eq. 1.3, respectively. Given that the required differential linear swing (V_{ppd}) is 4 V across a 60-Ω differential load (i.e., 15.2 dBm of differential RF power for an optimum load of $R_{opt,d} = 60\ \Omega$), the equivalent single-ended root-mean-square (RMS) voltage is 0.707 V_{rms} across a 30-Ω load ($R_{opt,se} = 30\ \Omega$). As shown in Eq. 1.2, the RF output power plus 2 dB is an estimate of the postamp OP_{-1dB} (i.e., single-ended $OP_{-1dB,post}$ of 14.2 dBm), which results in 0.89 V_{rms} and an RMS current of 30 mA_{rms} at the load. For Class-A operation, the current is expected to swing between zero and the

max (I_{max}=$2\sqrt{2}I_{rms}$), and therefore the device is biased at half of I_{max}. The bias current for the desired OP_{-1dB} is set at 42 mA (i.e., $I_{bias} = \sqrt{2}I_{rms}$).

However, Eq. 1.3 predicts that OP_{-1dB} for a multi-stage amplifier is always lower than OP_{-1dB} of the postamp stage alone. This is attributed to the finite gain of the postamp and the significantly lower OP_{-1dB} of the preamp, because the preamp is typically responsible for providing voltage gain and amplitude peaking for the driver (i.e., not a high-power stage). Assuming that the supply voltage and bias current of the preamp are scaled down from those of the postamp by factors of 1.67 and 8.9, respectively, the OP_{-1dB} of the preamp ($OP_{-1dB,pre}$) is expected to be scaled down by a factor of 8.9, resulting in $OP_{-1dB,pre}$ of just 4.73 dBm. To apply Eq. 1.3, gain of the postamp stage is required.

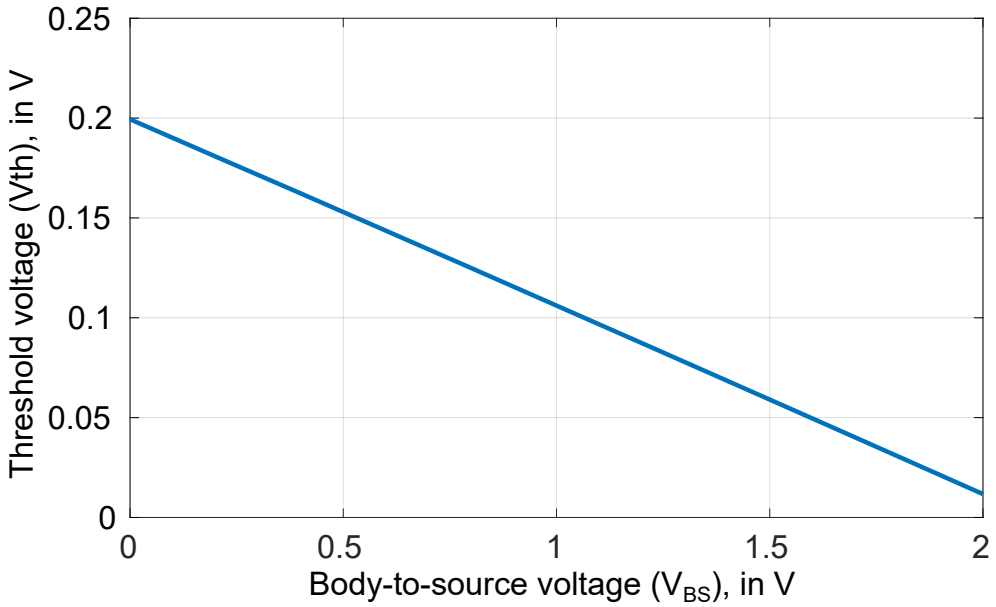


Figure 4.7: Threshold voltage of NFETs versus body-to-source voltage (V_{GS} = 0.45 V, V_{DS} = 0.8 V. NFET width of 160 μm with 2 μm/finger, length of 18 nm.)

The back gate bias (V_{BS}) may be used to tune the threshold voltage (V_{th}) of the NFETs. The effect of V_{BS} on V_{th} is shown in Fig. 4.7. For an increase of V_{BS} from 0 to 2 V, V_{th} decreases from 0.2 V to 0.01 V. Assuming that NFETs of the postamp have V_{BS} of 2 V and parasitic resistances are neglected, the transconductance ($g_{m,post}$) and gain ($A_{v,post}$) of

the postamp are estimated at

$$\begin{aligned} g_{m,post} &= \frac{2I_{bias}}{V_{ov}} = \frac{2 \cdot 0.042}{0.45 - 0.01} = 0.191\ S, \\ |A_{v,post}| &= g_{m,post} R_L = 0.191 \cdot 30 = 5.72\ V/V. \end{aligned} \tag{4.2}$$

The voltage gain of 5.72 V/V is converted to $|S_{21,post}|$ of 7.24 V/V with Eq. 1.1 (assuming that R_G of 12 Ω interfaces with the input capacitance of the postamp, which results in Γ_{in} of 1 near DC). Substituting $OP_{-1dB,post}$, $OP_{-1dB,pre}$ and $S_{21,post}$ into Eq. 1.3 results in an overall OP_{-1dB} of 13.5 dBm, which is 0.7 dB lower than $OP_{-1dB,post}$. Although the overall OP_{-1dB} still meets the 12.2-dBm specification, at high frequencies where the gain rolls off significantly (due to frequency response of the postamp), greater degradation in total OP_{-1dB} is expected.

Assuming that the postamp voltage gain rolls off by 6 dB at 75 GHz, $|S_{21,post}|$ becomes 3.62 V/V. The total OP_{-1dB} at 75 GHz drops to 12 dBm according to Eq. 1.3 (2.2 dB lower than $OP_{-1dB,post}$). To meet the 12.2-dBm specification, a design margin should be included while designing the postamp.

A simple solution is to compensate $OP_{-1dB,post}$ by 2.2 dB (i.e., increase from 14.2 dBm to 16.4 dBm). Recalculating the corresponding voltage and current results in V_{rms} and I_{rms} of 1.153 V and 38 mA, respectively. As a result, the postamp bias current increases from 42 mA to 54 mA. With a knee voltage (V_k) of 0.4 V assumed, the supply voltage should be at least 2 V (i.e., $V_{DD} = V_k + \sqrt{2}V_{rms}$). The updated $g_{m,post}$ and $|S_{21,post}|$ are 0.242 S and 4.58 V/V at 75 GHz, respectively, resulting in an updated cascade OP_{-1dB} of 14.9 dBm (i.e., margin of 2.7 dB). At a current density of 0.027 mA/μm^2 and NFET length of 18 nm, the NFET width is 110 μm, which may be used as a starting point for the postamp design. In reality, $g_{m,post}$ will be lower than the estimate in Eq. 4.2 due to postlayout parasitic resistances and non-ideal current gain of cascode devices at high frequencies. Parasitic capacitances also lead to greater roll-off of $|S_{21,post}|$ at high frequencies. Sufficient margining of the postamp NFET width is therefore required. Since the minimum supply voltage of 2 V is much higher than the nominal voltage of NFETs (i.e., 0.8 V), at least three transistors in series are needed to fit within the supply. The NFET width is finalized at 160 μm (two 80-μm wide units) and the corresponding small-signal transconductance is 0.33 S.

An alternative solution is to compensate the $OP_{-1dB,pre}$. In the previous example, 18 mW is added to $OP_{-1dB,post}$ (i.e., increase from 14.2 dBm to 16.4 dBm). If the preamp is compensated for the same amount, the $OP_{-1dB,pre}$ becomes 13.2 dBm and the cascade OP_{-1dB} results in 13.7 dBm. Compared to the post-compensation 14.9 dBm from the

previous example, compensating the preamp is less effective. This may be explained by Eq. 1.3, where the effect of $OP_{-1dB,pre}$ on the cascade is suppressed by gain of the postamp. Adding margin directly to the postamp improves both $OP_{-1dB,post}$ and $|S_{21,post}|$ simultaneously, and is therefore a more effective approach.

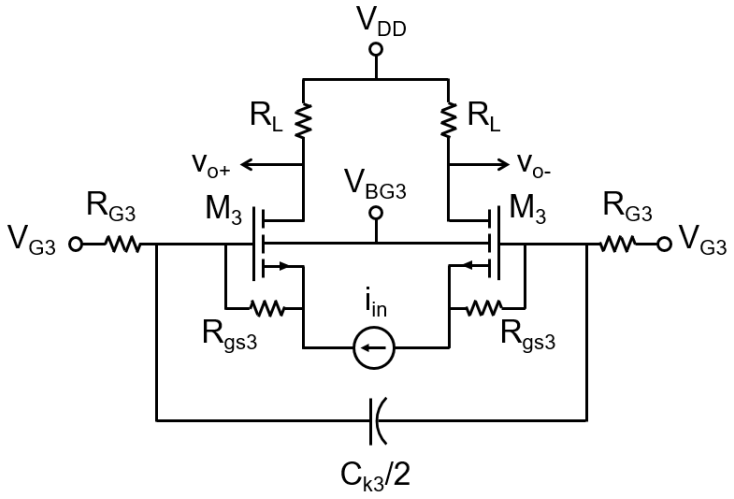


Figure 4.8: Bridging capacitor C_{k3} realizes common-gate compensation in a differential implementation (example of the M_3 common-gate stage from Fig. 4.5)

4.1.2.2 Compensation network design for common-gate FETs

The RC network at the gate of cascode devices increases the impedance $Z_{s,k}$ ($k = 2, 3...$, see Fig. 2.2) looking into the source of the common-gate transistor, as explained in Section 2.2.1. The impedance may be estimated by Eq. 2.1. To evenly spread the total output swing evenly across the drain to source of M_1, M_2 and M_3, C_{k2} and C_{k3} should be selected such that impedances $Z_{s,2}$ and $Z_{s,3}$ are 10 Ω and 20 Ω, respectively, for a 30-Ω load (i.e., $Z_{s,2} = R_{opt,se}/3$ and $Z_{s,3} = 2R_{opt,se}/3$ assuming that the common-gate FETs have an ideal current gain of 1). For C_{gs2} and C_{gs3} of 90 fF, Eq. 2.1 predicts that the required C_{k2} and C_{k3} are 39 fF and 16 fF, respectively (assuming infinite r_{ds} for M_1 to M_3). In a differential realization, a half-size capacitor is utilized between the gates of common-gate FETs instead of two separate capacitors, as shown in Fig. 4.8. The advantage is that, while differential operation remains unchanged, it can be assumed that the gate of M_3 sees high impedance, if large resistance values are selected for R_{G3} and R_{gs3} (e.g., in this design R_{G3} and R_{gs3} are 12.5 $k\Omega$ and 3 $k\Omega$). As a result, the small-signal transconductor is turned off in common

mode, resulting in a voltage division between the output resistance of M_3 (r_{ds3}) and the load resistor R_L. Therefore, the common-mode gain is suppressed to R_L/r_{ds3} (e.g., for r_{ds3} of 55 Ω and R_L of 30 Ω, the common-mode gain of M_3 is 0.55 V/V).

Due to the drain-to-gate feedback via $C_{dg2,3}$ in Fig. 2.2, $Z_{s,k}$ increases to [42]

$$Z_{s2,3} \approx \frac{1}{g_{m2,3}} \frac{C_{gs2,3} + C_{k2,3} + C_{dg2,3}(1 + g_{m2,3}R_{d2,3})}{C_{k2,3} + C_{dg2,3}} \left(\frac{1}{1 + s/\omega_T}\right), \tag{4.3}$$

where $R_{d2,3}$ is the impedance looking out from the drain of second/third FET. For example, substituting C_{gs2} and g_{m2} of 90 fF and 0.33 S, respectively, to Eq. 2.1 results in Z_{s2} of 10 Ω. With additional C_{dg2} and R_{d2} of 41.4 fF and 20 Ω, respectively, Eq. 4.3 yields a Z_{s2} of 16.7 Ω (increased from 10 Ω). That is, C_{k2} and C_{k3} have to increase to lower $Z_{s,2}$ and $Z_{s,3}$ back to 10 Ω and 20 Ω, respectively. The updated values (with finite r_{ds} of 55 Ω for M_1 to M_3) of C_{k2} and C_{k3} are 90 fF and 54 fF from post-layout simulations, respectively (i.e., increased from 39 fF and 16 fF, respectively, estimated from Eq. 2.1).

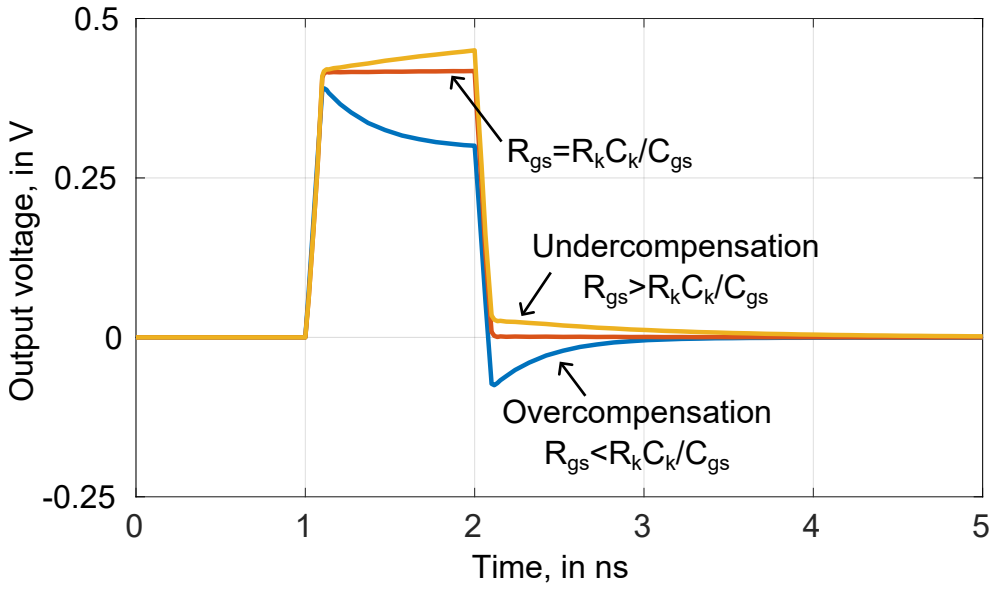


Figure 4.9: Time-domain pulse response of a cascode device for various compensation resistance value for R_{gs} (Tested with schematic from Fig. 4.8. Signal source is a 1-$V_{pp,d}$ pulse with rise/fall time of 0.1 ns. Source and load resistances are 50 and 60 Ω, respectively. Output voltage is taken at v_{o+})

The gate compensation network shown in Fig. 2.2 does not have a frequency response down to DC, as C_{gs}, C_k and R_k form a highpass filter. This leads to slow charg-

ing/discharging for a long input sequence of all 0's or all 1's (i.e., fundamental frequency below $1/R_k(C_k+C_{gs})$ rad/s, which is the corner frequency of the highpass network). Compensation resistor R_{gs} added between the gate and source of the cascode FETs mitigates this problem by creating an explicit low-frequency path via R_{gs} and R_k. The value of R_{gs} is selected to match the time constants of the two parallel RC networks (i.e., $R_{gs} = R_kC_k/C_{gs}$). The drawback is that there will be a DC current of V_{GS}/R_{gs} constantly flowing through R_{gs} and R_k, and the external gate bias V_G has to increase to compensate for the voltage drop in R_k. A large resistance value for R_{gs} is preferred to reduce the DC power consumed by the resistors. The final values of R_{gs2}, R_{k2}, R_{gs3} and R_{k3} are 12.5 $k\Omega$, 12.5 $k\Omega$, 3 $k\Omega$ and 12.5 $k\Omega$, respectively.

The simulated pulse response in Fig. 4.9 shows the effect of adding the compensation resistor. Without R_{gs}, the response takes 4 ns to settle. With R_{gs} set to the critical value (R_kC_k/C_{gs}), the response settles faster. However, if R_{gs} is too small (i.e., overcompensation) unwanted undershoot is observed.

4.1.2.3 Output combining network design

Due to the effect of the drain-to-source capacitance C_{ds}, the optimum resistance $R_{opt,se}$ for maximum output power $P_{opt,se}$ is no longer valid at high frequencies and the optimum impedance $Z_{opt,se}$ should be applied. Load-pull simulations are executed to determine $Z_{opt,se}$ at 75 GHz where the amplitude peaks. Note that due to the frequency response of the postamp, the gain at 75 GHz is 2.7-dB lower than at low frequencies (see Fig. 4.6). Therefore, in load-pull simulations, the input power should be increased accordingly to account for the gain roll-off to attain the most accurate result. The $OP_{-1dB,se}$ with R_L of 30 Ω (i.e., without design an output matching network for $P_{opt,se}$) is 12 dBm. With $Z_{opt,se}$ of $14.7 + j14.7$ Ω, $OP_{-1dB,se}$ increases to 15.5 dBm. Note that $OP_{-1dB,se}$ is lower than 18.3 dBm from the previous calculation, which is due to the non-ideal current gain at high frequencies for the common-gate stages, and losses due to parastics. Since the postamp is divided into two transistor units, $Z_{opt,se}$ of $29.4 + j29.3$ Ω is seen by each transistor drain. This is realized by a combiner at the output, which is modelled as shown in Fig. 4.10.

Drains of transistor M_3 in Fig. 4.5 are represented by D_{1+} to D_{2-}. An inductor L_d of 90 pH connects each drain to the output signal pad. The parasitic capacitance C_d of L_d is 5.5 fF on each side. Since $Z_{opt,se}$ at 75 GHz is sensitive to parasitics, the pad inductance is explicitly modelled at the output, which is 19 pH. Lastly, a load capacitance of 78 fF is required to realize $Z_{opt,se}$. It consists of the pad capacitance (C_{pad}) of 24 fF and a capacitor (C_t) of 54 fF, which is placed between the signal pad and the adjacent ground pad.

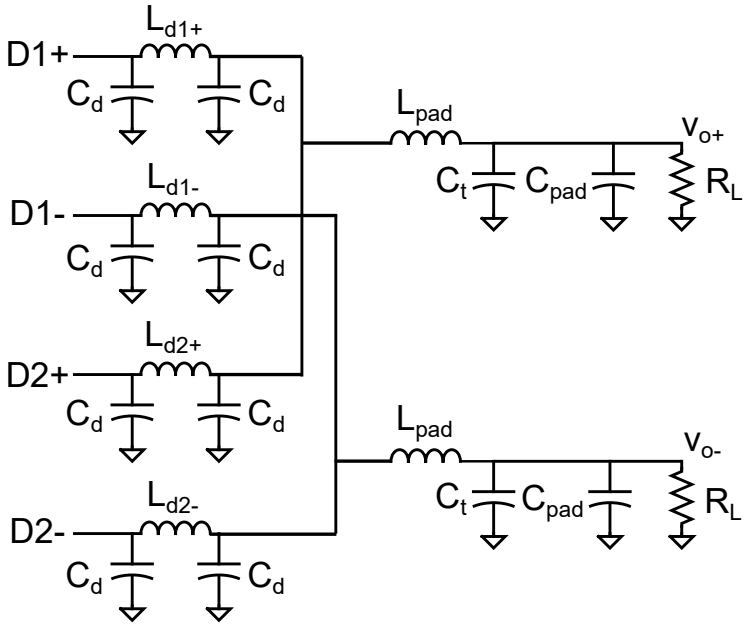


Figure 4.10: Lumped-element model of the output combining network

4.1.2.4 Effects of the drain-to-source feedback and the gate compensation capacitor for cascode devices

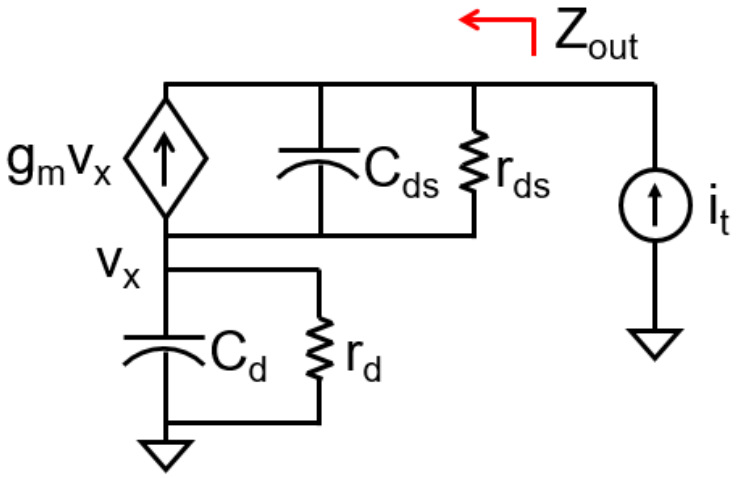


Figure 4.11: A simplified model to show the effect of C_{ds} on R_{out} for a cascode device

The drain-to-source capacitance (C_{ds}) of a cascode device may lead to a negative output resistance (i.e., risk of instability) at high frequencies. A simplified model of a cascode device with C_{ds} is shown in Fig. 4.11. The output impedance Z_{out} is given by

$$Z_{out}(s) = \frac{r_d + r_{ds} + g_m r_{ds} r_d + s \cdot (C_{ds} + C_d) r_{ds} r_d}{1 + s \cdot (C_{ds} r_{ds} + C_d r_d) + s^2 \cdot C_{ds} C_d r_{ds} r_d}, \tag{4.4}$$

where C_{ds} and r_{ds} are the drain-to-source capacitance and resistance of the common-gate transconductor, respectively. The impedance of the preceding stage is modelled by C_d and r_d in parallel (e.g., a common-source or common-gate output impedance). Note that C_{gs} of the cascode device is absorbed in C_d, and v_{gs} is labelled as v_x in Fig. 4.11. For simplicity, assuming that C_d equals C_{ds} and r_d equals r_{ds}, the real part of Z_{out} (R_{out}) is expressed as

$$R_{out} = \mathrm{Re}(Z_{out}(\omega)) = \frac{-g_m r_{ds}^4 C_{ds}^2 \omega^2 + 2 r_{ds}^3 C_{ds}^2 \omega^2 + g_m r_{ds}^2 + 2 r_{ds}}{(1 - r_{ds}^2 C_{ds}^2 \omega^2)^2 + (2 r_{ds} C_{ds} \omega)^2}. \tag{4.5}$$

When ω is large, the condition for a negative value of R_{out} is dominated by the first two terms of the numerator, which are given by

$$\begin{gathered} r_{ds}^4 C_{ds}^2 \omega^2 (-g_m + \frac{2}{r_{ds}}) < 0, \\ g_m r_{ds} > 2. \end{gathered} \tag{4.6}$$

For the postamp, with r_{ds} of 55 Ω and g_m of 0.33 S, R_{out} is plotted against various values of C_{ds} in Fig. 4.12 ($C_d = C_{ds}$ assumed). With no C_{ds}, R_{out} is resistive (around 1.1 $k\Omega$). As C_{ds} increases, R_{out} rolls off as frequency increases. With C_{ds} of 60 fF, R_{out} becomes negative above 55 GHz, which leads to a stability risk. As C_{ds} increases, the corner frequency where R_{out} becomes negative decreases.

In a differential realization, capacitor cross-coupling between the drain and source of the cascode devices may compensate the effect of C_{ds} at the cost of bandwidth. Another approach to minimize the effect of C_{ds} is via layout, as the fringing capacitance between the source and drain backend metals is a major contributor to C_{ds}.

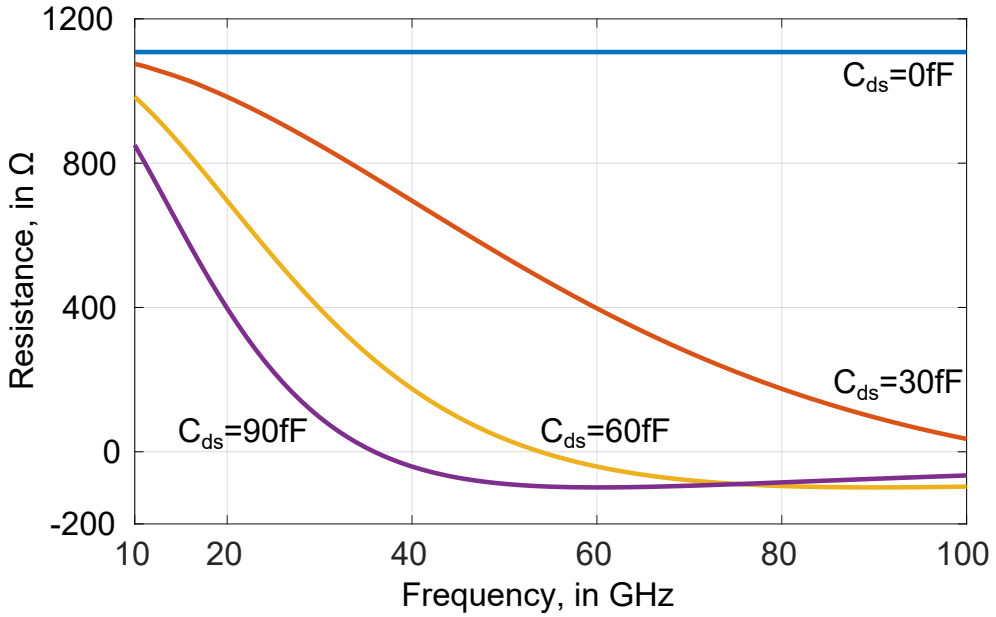


Figure 4.12: Real part of the output impedance from Fig. 4.11 with a variable C_{ds} in a cascode amplifier ($C_d = C_{ds}$, $r_d = r_{ds} = 55\ \Omega$, $g_m = 0.33$ S)

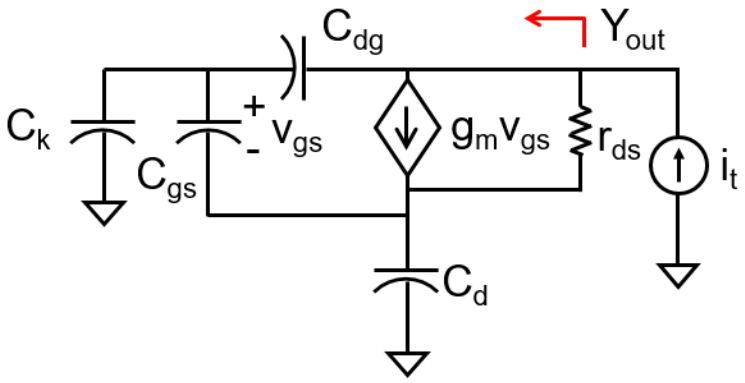


Figure 4.13: A simplified model to show the effect of C_k and C_{dg} on C_{out} for a cascode device

Another undesirable behaviour is the increase in output capacitance C_{out}, caused by the gate compensation capacitor C_k, resulting in a larger time constant at the RF output. A simplified model of a cascode device with C_{dg} and C_k is shown in Fig. 4.13. The output

admittance Y_{out} is given by

$$Y_{out}(s) = \frac{s^2 \cdot \alpha + s \cdot \beta}{s \cdot \gamma + \delta}, \tag{4.7}$$

where

$$\begin{aligned}
\alpha &= C_{dg} r_{ds} (C_{gs} C_d + C_k C_d + C_{gs} C_k), \\
\beta &= g_m r_{ds} C_{dg} (C_k + C_d) + C_d (C_{gs} + C_{dg} + C_k) + C_k (C_{gs} + C_{dg}), \\
\gamma &= r_{ds} (C_d (C_{gs} + C_{dg} + C_k) + C_k (C_{gs} + C_{dg})), \\
\delta &= g_m r_{ds} (C_k + C_{dg}) + (C_{gs} + C_{dg} + C_k).
\end{aligned} \tag{4.8}$$

For a regular cascode (i.e., C_k short circuited), the parallel-equivalent output capacitance C_{out} is given by

$$C_{out} = C_{dg} + \frac{C_{gs} + C_d}{1 + g_m r_{ds}}, \tag{4.9}$$

which is approximately C_{dg} if $g_m r_{ds}$ is large. With C_k in place, C'_{out} may be estimated by

$$C'_{out} = \frac{\mathrm{Im}(Y_{out}(s))}{s} \approx \frac{\beta}{\delta} \approx \frac{C_{dg}(C_k + C_d)}{C_k + C_{dg}}, \tag{4.10}$$

at frequencies where the terms in Eq. 4.7 with $g_m r_{ds}$ dominate. Depending on the values of C_k, C_{dg} and C_d, an increase in C_{out} (and the output time constant) is expected when a gate compensation capacitor is utilized compared to a regular cascode amplifier (e.g., for $C_{k3} = 54$ fF, $C_{gs} = 90$ fF, $C_{gd} = C_d = 40$ fF, $g_m r_{ds} = 18$ V/V, C_{out} increases from 47 fF (cascode) to 61 fF (with gate compensation capacitor), a 30% increase in the output time constant).

4.1.2.5 Effects of drain-to-gate feedback and drain inductance on common-source devices

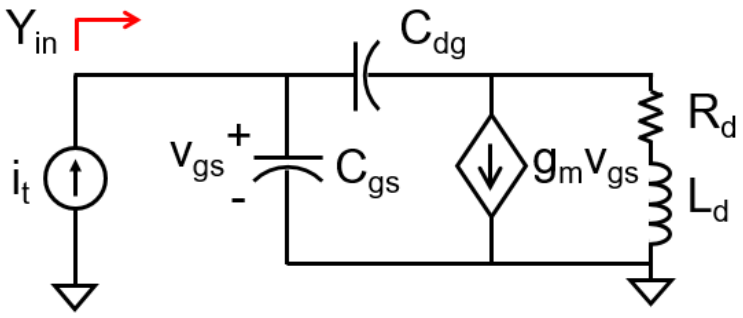


Figure 4.14: A simplified model to show the effect of C_{dg} and L_d on R_{in} for a common-source amplifier (neglecting C_{ds} and r_{ds})

An ideal common-source amplifier should have infinite input resistance at DC if drain-to-gate feedback is neglected, as the input impedance is purely capacitive (i.e., dominated by the gate-to-source capacitor C_{gs}). However, due to the feedback through the drain-to-gate capacitance (C_{dg}), the input resistance becomes finite and decreases at a rate proportional to the square of the frequency (see Eq. 4.13). Moreover, inductance in series with the drain (L_d) of a common-source transistor (commonly utilized for series peaking) may lead to a negative input resistance (i.e., risk of instability) due to feedback via C_{dg}. A simplified model of a common-source device with C_{dg} and L_d is shown in Fig. 4.14 (n.b., C_{ds} and r_{ds} are neglected). The input admittance Y_{in} is given by

$$Y_{in}(s) = \frac{s^3 \cdot \alpha + s^2 \cdot \beta + s \cdot \gamma}{s^2 \cdot \delta + s \cdot \epsilon + 1}, \tag{4.11}$$

where

$$\begin{aligned} \alpha &= C_{dg}C_{gs}L_d, \\ \beta &= C_{dg}C_{gs}R_d + g_m C_{dg}L_d, \\ \gamma &= C_{gs} + C_{dg} + g_m C_{dg}R_d, \\ \delta &= C_{dg}L_d, \\ \epsilon &= C_{dg}R_d. \end{aligned} \tag{4.12}$$

When L_d is neglected, the parallel-equivalent input resistance is given by

$$R_{in} = \frac{1}{\mathrm{Re}(Y_{in}(s))} \approx \frac{R_d}{1 + g_m R_d} + \frac{1}{\omega^2 C_{dg}^2 R_d (1 + g_m R_d)}, \tag{4.13}$$

where the first term represents the input resistance at high frequencies, and the second term shows the frequency dependency of R_{in} (i.e., resistance decreasing as frequency increases, which is undesirable for the interstage design). With L_d included, the parallel-equivalent input resistance R'_{in} is given by

$$\begin{aligned} R'_{in} &= \frac{1}{\mathrm{Re}(Y_{in}(s))} = \frac{\omega^4\delta^2 - \omega^2(2\delta - \epsilon^2) + 1}{\omega^4(\beta\delta - \alpha\epsilon) - \omega^2(\beta - \gamma\epsilon)}, \\ &\approx \frac{\delta^2}{\beta\delta - \alpha\epsilon} = \frac{1}{g_m} \quad \text{(when } \omega^4 \text{ terms dominate)}, \\ &\approx \frac{1}{-\omega^2(\beta - \gamma\epsilon)} + \frac{2\delta - \epsilon^2}{\beta - \gamma\epsilon} \quad \text{(when } \omega^2 \text{ terms dominate)}. \end{aligned} \tag{4.14}$$

At low frequencies, the parallel-equivalent input resistance is infinite (input impedance is determined by C_{in} in Eq. 4.16). When ω^2 terms dominate R'_{in} as shown in Eq. 4.14, there is a transition frequency ω_c, where the sign of R'_{in} changes. That is, R'_{in} is negative when

$$\begin{aligned} &\text{Case 1:} \quad \omega < \omega_c, \text{ for } L_d > \frac{C_{dg}R_d(1 + g_mR_d)}{g_m}, \\ &\text{Case 2:} \quad \omega > \omega_c, \text{ for } L_d < \frac{C_{dg}R_d(1 + g_mR_d)}{g_m}, \\ &\text{where} \quad \omega_c = \sqrt{\frac{1}{2\delta - \epsilon^2}} = \sqrt{\frac{1}{2C_{dg}L_d - C_{dg}^2R_d^2}}, \end{aligned} \tag{4.15}$$

when C_{ds} and r_{ds} of the common-source FET are neglected. In general, Case 2 is preferred, as a potential negative R'_{in} will only occur at frequencies above ω_c. With a proper value of L_d selected, the ω^4 terms dominate in Eq. 4.14 before negative R'_{in} occurs (i.e., R'_{in} approaches $1/g_m$ instead of becoming negative). Therefore, series-inductive peaking applied at the drain of common-source devices must be kept lower than the L_d specified in Eq. 4.15 (Case-2), to avoid potential instability. In this design, L_d is 18 pH with capacitor neutralization applied.

Recall from Eq. 4.11 that the parallel-equivalent input capacitance is given by

$$\begin{aligned} C_{in} = \frac{-\mathrm{Im}(Y_{in}(s))}{\omega} &\approx \gamma = C_{gs} + C_{dg}(1 + g_mR_d) \quad \text{(when linear terms in '}s\text{' dominates)}, \\ &\approx \frac{\alpha}{\delta} = C_{gs} \quad \text{(when } s^3 \text{ terms dominate)}. \end{aligned} \tag{4.16}$$

As shown in Eq. 4.16, C_{in} matches with the Miller's approximation at low frequencies when s terms dominate. At high frequencies, C_{in} decreases to C_{gs} when s^3 terms dominate.

4.1.3 Shunt/double-series interstage

After finalizing the preamp and the postamp, the next objective is to determine the component values for the interstage filter network to attain desired peaked frequency and gain (i.e., 12.8 dB at 75 GHz). Parameters for the interface between the preamp and postamp were used for the design example of Section 3.1.2 and shown on the block diagram of Fig. 3.11 (page 44). The interstage design follows the methodology described in Section 3.1.2.5. The circuit design parameters are attained in Section 3.1.2 and shown in Table 3.2 (i.e., $n = 0.83, m_1 = 2.42, m_2 = 1.25, m_3 = 1.58$). An analysis of coil losses on the peak amplitude covered in this section.

As shown in Fig. 3.11, the required interstage load resistance R which sets the $|S_{DD21}|$ of the overall driver to 15.7 dB (i.e., voltage gain A_v of 4.73 V/V) from the gain specification in Section 1.2.2 may be computed from the DC gain

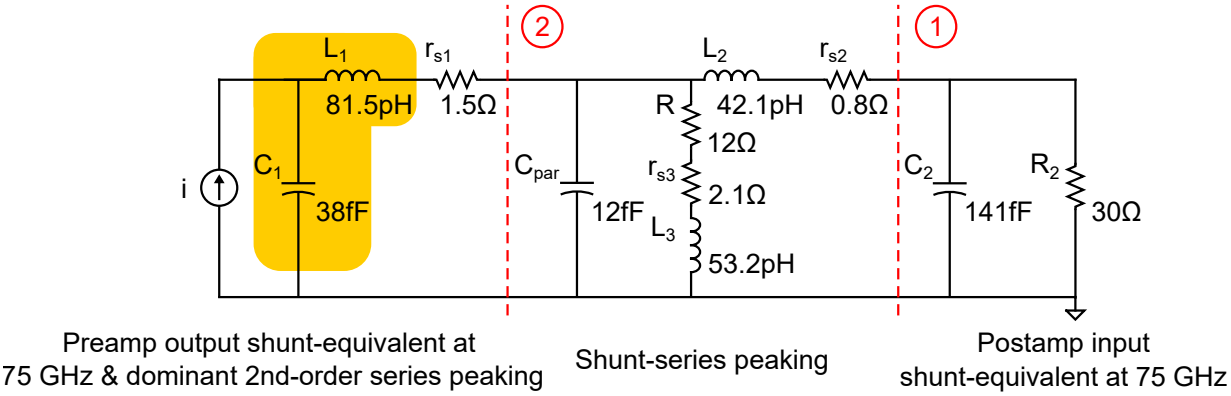
$$\begin{aligned} |A_v| &= A_{v,pre} \cdot A_{v,post} = g_{m,pre} \cdot R \cdot A_{v,post}, \\ \text{and} \qquad R &= \frac{A_v}{g_{m,pre} \cdot A_{v,post}} = \frac{4.73 \text{ V/V}}{26 \text{ mS} \cdot 7.6 \text{ V/V}} = 24 \ \Omega, \end{aligned} \tag{4.17}$$

where $g_{m,pre}$ and $A_{v,post}$ are the preamp transconductance from Fig. 4.4 and postamp voltage gain from Fig. 4.6, respectively. At DC, $g_{m,pre}$ is 27 mS and $A_{v,post}$ is 7.6 V/V, resulting in an R value of 24 Ω (i.e., single-ended equivalent of 12 Ω) to realize the desired driver voltage gain.

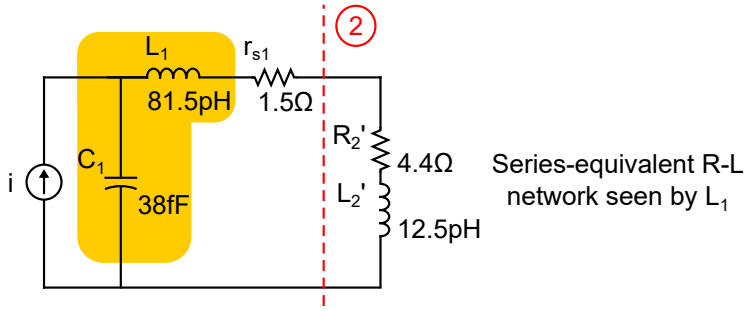
Recall that the preamp stage has a 1.1-dB decrease in gain at 75 GHz (see Fig. 4.4) and that the postamp gain decreases by 2.7 dB at 75 GHz (see Fig. 4.6). To achieve an overall peak of 9 dB at 75 GHz, the interstage is expected to have a peak of at least 12.8 dB (4.36 V/V) at 75 GHz for the driver (including the effects of the output pole of the preamp and the input pole of the postamp). A design margin is also needed for A_{pk}, due to ohmic losses from passive components that dampen the network. Also, the finite input impedance of the postamp loads the interstage network as the frequency increases (as explained in Section 4.1.2.5).

4.1.3.1 Analyzing the effects of coil losses on parameters b and A_{pk}

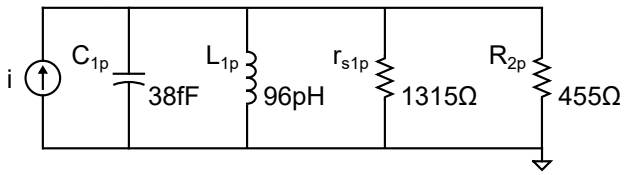
Based on the design parameters in Table 3.2 (i.e., $n = 0.83, m_1 = 2.42, m_2 = 1.25, m_3 = 1.58$), denormalizing m_1 to m_3 with Eq. 3.3 (interstage load $R = 12$ Ω and total capacitance $C_T = 234$ fF) yields circuit parameters L_1 to L_3 of 81.5 pH, 42.1 pH, and 53.2 pH for the single-ended equivalent shunt/double-series network, as shown in Fig. 4.15a.



(a) Dominant 2nd-order network consisting of L_1 and C_1 highlighted (n.b., input impedance of the postamp becomes C_2 of 141 fF in parallel with R_2 of 30 Ω at 75 GHz, instead of simply 184 fF at DC)



(b) Convert the impedance seen at boundary-2 to a series equivalence R_2' and L_2' at 75 GHz



(c) Shunt-equivalent RLC network defining the dominant 2nd-order response at 75 GHz (C_1, L_1, r_{s1} and R_2' in Fig. 4.15b converted to shunt-equivalent)

Figure 4.15: Single-ended lumped-element model of the shunt/double-series interstage network, including series ohmic loss of L_1 and L_2 for a Q-factor of 25 (r_{s1} and r_{s2}), series loss of L_3 for a Q-factor of 12, and shunt-equivalent R-C network of the postamp input at 75 GHz (R_2 and C_2)

Series ohmic losses from physical implementation of the inductor coils are modelled by resistors r_{s1} to r_{s3}. Q-factors of 25, 25, and 12 are applied to inductors L_1 to L_3 at 75 GHz, respectively (details explained in Section 4.1.3.2). Note that the load network

changes from C_2 of 184 fF to an R-C network of C_2 (141 fF) and R_2 (30 Ω) in parallel, which models the shunt-equivalent input impedance of the postamp at 75 GHz. As explained in Section 4.1.2.5, the shunt-equivalent input capacitance and resistance are both frequency dependent, and decrease as frequency increases due to the drain-to-gate feedback. In other words, the postamp input impedance cannot be modelled by a parallel R-C across a wide range of frequencies.

Factors such as the finite Q of the inductors and the input impedance of the postamp at ω'_{pk} damp the $L_1 - C_1$ resonator, resulting in a lower A_{pk}. Q-factors for parallel/series R-L (Q_L) and R-C (Q_C) networks are given by [24]

$$\begin{aligned} Q_L &= \frac{R_p}{\omega L_p} = \frac{\omega L_s}{R_s} \qquad \text{(for parallel/series R-L circuits)}, \\ Q_C &= \omega C_p R_p = \frac{1}{\omega C_s R_s} \qquad \text{(for parallel/series R-C circuits)}, \end{aligned} \tag{4.18}$$

where R_p and R_s are the equivalent resistance for parallel and series inductors L_p and L_s, respectively (same for parallel and series capacitors C_p and C_s). Series-to-parallel (vice versa) conversion exploits the following equations [24]

$$\begin{aligned} L_p &= L_s \frac{1+Q_L^2}{Q_L^2} \text{ and } R_p = R_s(1+Q_L^2) \qquad \text{(for series-to-parallel R-L conversion)}, \\ C_p &= C_s \frac{Q_C^2}{1+Q_C^2} \text{ and } R_p = R_s(1+Q_C^2) \qquad \text{(for series-to-parallel R-C conversion)}. \end{aligned} \tag{4.19}$$

To analyze the change in A_{pk}, parameter b of the dominant second-order polynomial is used, as it models A_{pk}. In Fig. 4.15a, the resonant components corresponding to ω'_{pk} of 1.42 are C_1 and L_1. The schematic in Fig. 4.15a is simplified to a shunt-equivalent RLC network shown in Fig. 4.15c using Eqs. 4.18 and 4.19, where r_{s1p} and R_{2p} are the shunt-equivalent resistance of r_{s1} and the rest of resistances, respectively. Note that the simplification converts the impedance seen at boundary 2 in Fig. 4.15a to R'_2 and L'_2 in series (Fig. 4.15b), resulting in the final parallel RLC network shown in Fig. 4.15c. For a parallel RLC network, the characteristic polynomial is given by [58]

$$\begin{aligned} H_{RLC}(s) &= s^2 + s \cdot 2\alpha + \omega_o^2, \\ \alpha &= \frac{1}{2RC}, \\ \omega_o &= \frac{1}{\sqrt{LC}}, \end{aligned} \tag{4.20}$$

where α and ω_o are the neper and resonant frequencies for the parallel RLC network, respectively. Comparing Eq. 3.33 to Eq. 4.20, it may be concluded that

$$b = 2\alpha = \frac{1}{R_{eq}C_{1p}}/\omega'_{ref} = \frac{R(C_1 + C'_2 + C_{par})}{R_{eq}C_1} = \frac{12}{0.2 \cdot R_{eq}} \approx \frac{60}{R_{eq}}, \tag{4.21}$$

where ω'_{ref} is the updated unpeaked -3dB frequency (i.e., with $C_2 = 141$ fF compared to 184 fF in Fig. 3.11) and parameter b is normalized to $1/\omega'_{ref}$. The total equivalent resistance in Fig. 4.15c is 338 Ω, resulting in a b of 0.177 instead of 0.0244 (the value in Eq. 3.50, which is attained from the driver design example in Section 3.1.2 using lossless lumped elements and neglecting the postamp loading effects at 75 GHz). That is, the interstage response is damped. The peak amplitude of the shunt/double-series interstage with b of 0.177 is 13.3 dB, indicating a 17.1-dB decrease compared to Fig. 3.14 where A_{pk} is 30.4 dB. In detail, the series ohmic loss of L_1 (r_{s1}) and changes in the postamp input impedance at 75 GHz (C_2 in parallel with R_2) cause b to increase by 0.046 and 0.132, respectively (i.e., input impedance of the postamp has a greater impact). Series loss r_{s1}, postamp resistive loading R_2 and the change in C_2 are the major contributors to the reduction in the peak amplitude. Other factors, such as resistive losses in L_2 and L_3 (r_{s2} and r_{s3}), and finite output resistance of the preamp decrease the peak amplitude further.

All of these factors are taken into account during the physical design and electromagnetic simulations of the interstage network.

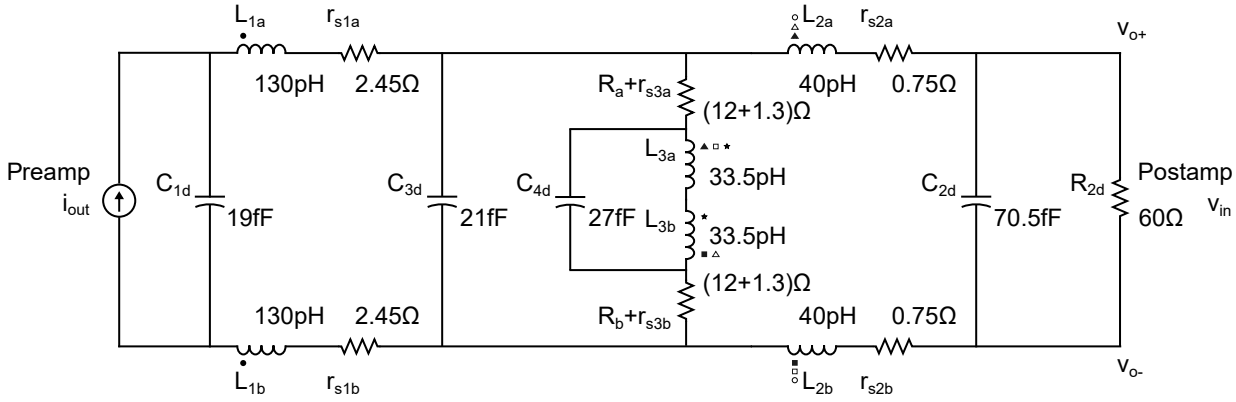
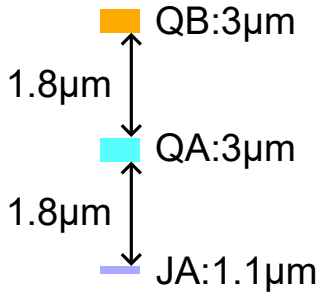


Figure 4.16: Equivalent differential lumped-element model of the shunt/double-series interstage network at 75 GHz (magnetic coupling factors: • $k_{11} = 0.35$, ○ $k_{22} = 0.21$, ★ $k_{33} = 0.28$, ▲■ $k_{23} = 0.26$, △□ $k_{32} = 0.05$). Capacitors C_{1d} to C_{4d} are the differential equivalent capacitors (i.e., half of C_1 to C_4, respectively)

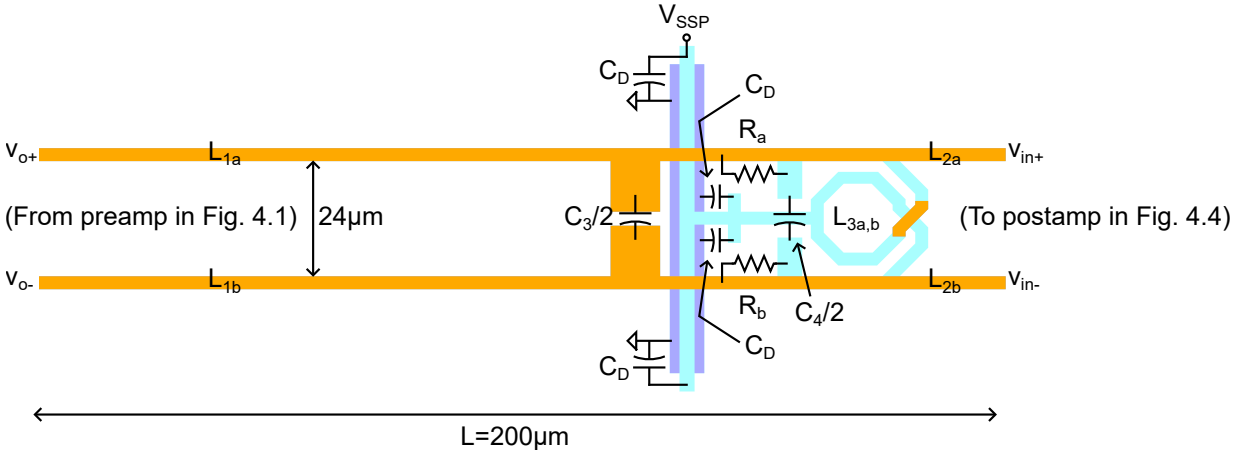
4.1.3.2 Differential shunt/double-series interstage design and physical implementation

$L_{1a,b}$ (pH)	$L_{2a,b}$ (pH)	$L_{3a,b}$ (pH)	$R_{a,b}$ (Ω)
130	40	33.5	12
C_3 (fF)	C_4 (fF)	C_D (fF)	V_{SSP} (V)
42	54	150	0.275

(a) Component values



(b) BEOL metal stack



(c) Physical layout

Figure 4.17: The differential interstage network shown in Fig. 4.16

The differential equivalent model of the shunt/double-series interstage network at 75 GHz is depicted in Fig. 4.16. Capacitors C_{1d} and C_{2d} are the differential equivalent total input and load capacitances, including parasitics from neighbouring elements (i.e., $C_{1d} = C_1/2$ and $C_{2d} = C_2/2$). For example, C_1 of 38 fF consists of C_{o1} of 30 fF and the parasitic capacitance of $L_{1a,b}$, which is 8 fF. Capacitor C_2 of 141 fF consists of C_{i2} of 137 fF and the parasitic capacitance of L_2 which is 4 fF. Resistor R_2 is the shunt-equivalent input resistance of the postamp at 75 GHz. Inductors $L_{1a,b}$ to $L_{3a,b}$ are 130 pH, 40 pH and 33.5 pH, respectively. The Q-factors at 75 GHz associated with $L_{1a,b}$ to $L_{3a,b}$ are 25, 25 and 12, respectively (Q_3 is lower than Q_1 and Q_2 because $L_{3a,b}$ is a spiral inductor using lower metal layers with

underpasses and vias). The expected degradation in peak amplitude due to finite Q for $L_{1a,b}$ and $L_{2a,b}$ is 8 dB (i.e., Fig. A.1 predicts that the peak amplitude decreases from 30 dB to 22 dB). The ohmic loss in $L_{3a,b}$ appears in series with the interstage load resistor, and its effect on A_{pk} is minor compared to the series loss in L_1 and postamp loading.

Magnetic couplings between coils are denoted by symbols in Fig. 4.16 (i.e., ● ○ ★ ▲ ■). For example, k_{11} represents the coupling between $L_{1a,b}$ coils between the differential signal paths (the same applies to k_{22} and k_{33}). Magnetic coupling factors k_{11} to k_{33} are 0.35, 0.21, and 0.28, respectively. The coupling factor k_{23} of 0.26 represents the coupling between L_2 and L_3 on the same side (i.e., L_{2a} to L_{3a} and L_{2b} to L_{3b}). Coupling to the opposite side of the layout is given by k_{32} of 0.05 (i.e., L_{2a} to L_{3b} and L_{2b} to L_{3a}).

The physical layout of the shunt/double series interstage network is shown in Fig. 4.17c, where the series-peaking coils L_1 of 130 pH and L_2 of 40 pH are implemented in 2.4-μm wide QB metal (orange). The shunt-peaking coil L_3 of 33.5 pH is implemented in 2.4-μm wide QA metal (cyan) with a QB overpass. The input and output of the network are connected to the preamp drain and postamp gate, respectively. The gap between the series coils and the outer dimension of the shunt coil are both 24 μm to minimize the physical size of the common-mode current loop (i.e., preamp and postamp differential transistors are 24-μm apart). The proximity of coils where signal current flows in a reduction of net flux caused by opposite directions results in opposing magnetic fields, and lower equivalent coil inductances (i.e., self-inductance is reduced by the mutual inductance). The magnetic coupling between L_{1a} and L_{1b} ($k_{11} = 0.35$) is an example of opposing fields. Therefore, self-inductances of $L_{1a,b}$ have to increase from 81.5 pH to 130 pH to match the single-ended model shown in Fig. 4.15a. On the other hand, differentially excited shunt coils $L_{3a,b}$ in Fig. 4.16) have reinforcing fields ($k_{33} = 0.28$) that increase the self-inductance (i.e., self-inductances for $L_{3a,b}$ decrease from 53.2 pH in the single-ended model to 33.5 pH). Inductor L_{2a} experiences opposing fluxes produced by L_{2b} ($k_{22} = 0.21$) and L_{3b} ($k_{32} = 0.05$), and aiding flux from L_{3a} ($k_{23} = 0.26$), resulting in a self-inductance value of 40 pH, which is slightly decreased compared to the single-ended model of 42.1 pH (vice versa for L_{2b}).

Metal-oxide-metal (MoM) tuning capacitors C_3 of 30 fF and C_4 of 50 fF are added (n.b., including neighbouring parasitics increases C_3 and C_4 to 42 fF and 54 fF, respectively) to peak the response beyond 75 GHz, as depicted in Fig. 4.18 (i.e., the blue curve is peaked by C_3 and C_4 compared to the yellow curve). Capacitor C_3 and C_4 are library MoM cells comprised of metal layers M_3 to M_7 from the 22-nm CMOS design kit. The capacitor C_3 is located near the end of $L_{1a,b}$ and the capacitor C_4 bridges across $L_{3a,b}$.

Decoupling capacitors are included to isolate the interstage network from common-mode supply noise. The centre of $L_{3a,b}$ is tapped and decoupled by a 150-fF capacitor damped

by a 10-Ω resistor in series. The same decoupling network is added to both ends of the V_{SSP} supply line (20 μm away from $L_{1a,b}$) to ensure symmetry, and V_{SSP} is supplied from a DC pad on the left side. A layer of JA ground metal (purple) is placed underneath the QA supply line to reduce the common-mode impedance.

The interstage load resistors $R_{a,b}$ of 12 Ω tap the series coil (at the midpoint of L_{1a} and L_{2a}) and are in series with L_{3a} (vice versa for L_{1b} to L_{3b}). The resistors are 3.5-μm long and 10-μm wide to handle the preamp current of 15 mA per side.

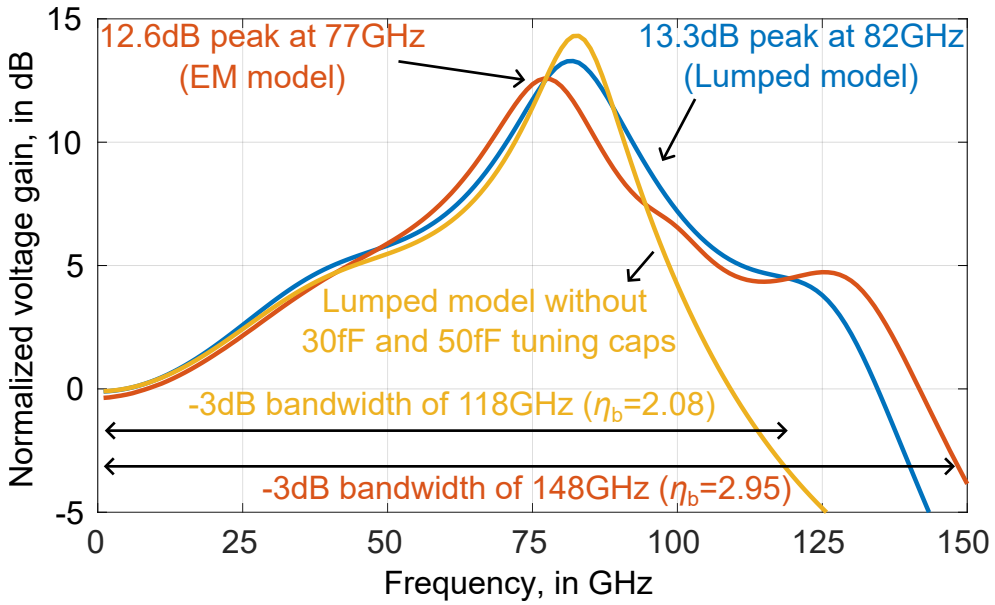


Figure 4.18: Frequency responses of the shunt/double-series interstage network loaded by the postamp, with and without the tuning capacitors C_3 and C_4 in Fig. 4.16

The magnitude response of the differential interstage network with the DC gain normalized to 0 dB is shown in Fig. 4.18. Simulated results using the differential lumped-element model with and without tuning capacitors C_3 and C_4 are shown in blue and yellow in Fig. 4.16, respectively. The blue curve shows a normalized peak amplitude of 13.3 dB at 82 GHz, which matches the mathematical prediction in Section 4.1.3.1 exactly when parameter b increases to 0.177 due to added coil losses and the postamp loading effect. The bandwidth peaking beyond 100 GHz is due to an additional pair of complex-conjugate poles near 125 GHz introduced by tuning capacitors C_3 and C_4 (i.e., the complete network has 7 poles and 2 zeros). The orange curve is the response simulated using an electromagnetic (EM) model of the interstage network (the EM model is generated by

Keysight PathWave® Momentum). The EM model predicts an amplitude peak of 12.6 dB at 77 GHz, and a -3dB bandwidth of 148 GHz. The bandwidth extension ratio (η_b) is 2.95 (n.b., tuning capacitor C_3 appears in shunt with the source/loading capacitors, which adds 30 fF to C_{par} in the low-frequency interstage model in Fig. 3.11, and lowers ω_{ref} from 56.7 GHz to 50.2 GHz). Compared to the yellow curve (without adding C_3 and C_4), the bandwidth extension ratio increases from 2.08 to 2.95. The change in A_{pk} compared to the differential, lumped-element model (blue curve) is within 1 dB, which validates the lumped model. Finally, the in-band amplitude humps near a normalized frequency of 0.6 (34 GHz if denormalized) seen in Figs. 3.12 to 3.14 are relaxed in Fig. 4.18, which is again a result of the frequency-dependent loading effect of the postamp.

In summary, there is a significant decrease in A_{pk} from 30.4 dB (a mathematical prediction using lossless coils and a simple capacitor model for the postamp in Section 3.1.2.3) to 12.6 dB (postlayout simulated result with postamp loading in Section 4.1.3.2). Therefore, the postamp stage should be carefully designed so that its input resistance and capacitance do not drop drastically at the peak frequency. For drivers with more than two stages, the gain peaking requirement may be distributed across multiple interstages, which relaxes the A_{pk} required per from each stage.

4.1.4 Postamp linearization

A Class-A postamp with a width of 160 μm achieves an OP_{-1dB} of 18 dBm at low frequencies. However, due to the frequency response of the postamp, and OP_{-1dB} degradation due to cascading, it is particularly challenging to maintain an OP_{-1dB} of 15.2 dBm at 75 GHz in the design of broadband linear drivers. This thesis describes a postamp utilizing a triple cascode amplifier with passive R-C networks at the gates of common-gate FETs to uniformly distribute the output swing, enhancing driver reliability, as depicted in Fig. 2.1c. However, the effective transconductance in such a cascode topology is anticipated to be lower than that of a conventional cascode amplifier, given that the impedance at each common-gate drain terminal is higher than in a conventional cascode, resulting in less signal current propagating through the cascode to the output load.

To improve linearity at high frequencies, particularly above 50 GHz in this study, one might explore linearization techniques such as feedforward [46]. Cascode compensation (CasComp) is a linearization technique developed for bipolar designs [59]. A similar topology has been applied in a CMOS distributed amplifier [60], achieving linearization through the cancellation of third-order intermodulation distortion (IM3 cancellation). A schematic example is shown in Fig. 4.19.

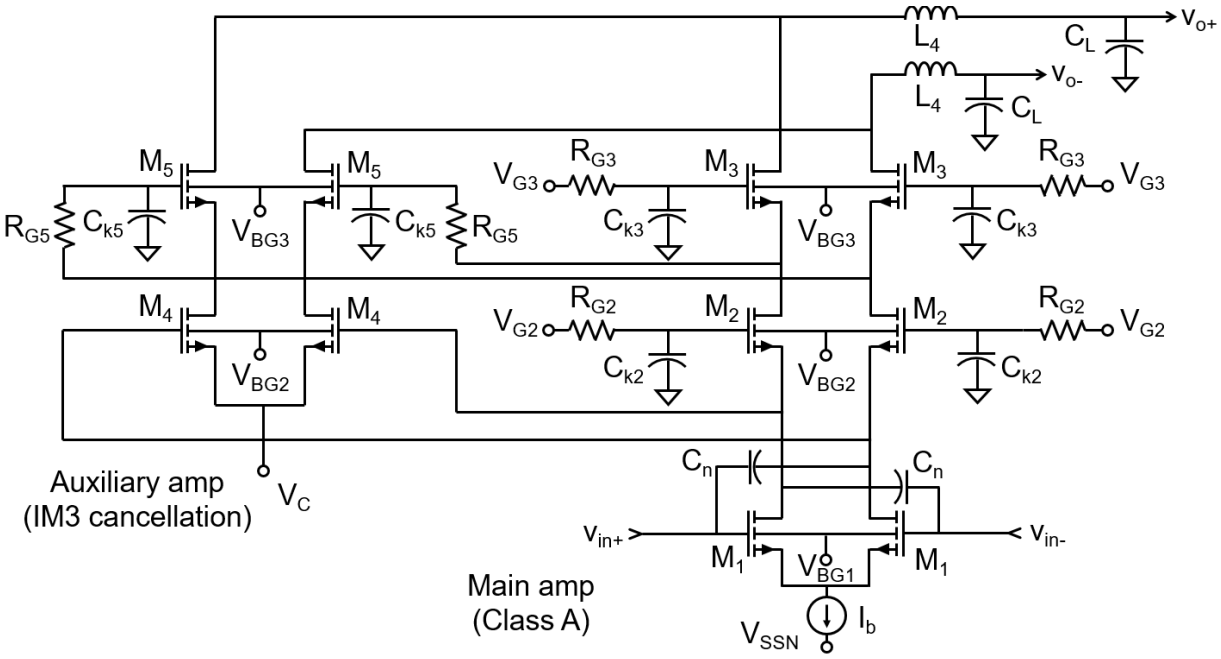


Figure 4.19: Schematic of a linearized postamp utilizing the IM3 cancellation technique

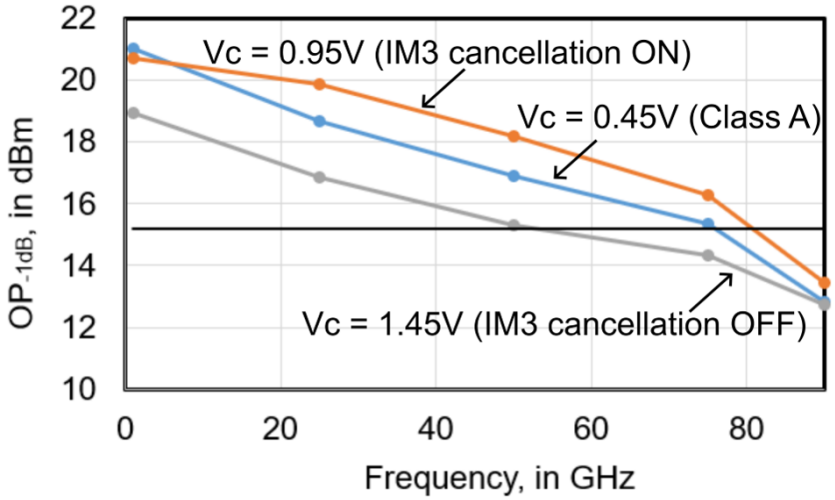


Figure 4.20: Simulated OP_{-1dB} against various control voltages and centre frequencies (V_C sets the conduction angle of the auxiliary amplifier)

The linearized postamp consists of two sub-blocks: the main amplifier, which operates

in Class A, and the auxiliary amplifier, which has a smaller conduction angle to produce the IM3 error signal for IM3 cancellation. The control voltage (V_C) allows dynamically adjusting the conduction angle of the auxiliary amplifier. The improvement in OP_{-1dB} is shown in Fig. 4.20. At a V_C of 1.45 V, the auxiliary amplifier switches off fully, serving as a reference case, and the simulated OP_{-1dB} is 14 dBm at 75 GHz, as depicted by the grey curve, which is below the target of 15.2 dBm. Reducing V_C increases the conduction angle, enabling the auxiliary amplifier to generate IM3. When V_C reaches 0.95 V, IM3 of the main amplifier is cancelled (i.e., main amplifier is linearized, see orange curve in Fig. 4.20). Should V_C continue to rise, the auxiliary amplifier will fully turn on and operate in Class A alongside the main amplifier, indicated by the blue curve. Although there remains a boost in OP_{-1dB} relative to the reference (grey curve), this improvement is less than via IM3 cancellation, and is achieved at a current density of 0.7 mA/μm (at V_C of 0.95 V, the auxiliary amplifier runs at a current density of 0.025 mA/μm). Note that, although the auxiliary amplifier's quiescent current is 1 mA only, as it is biased deep in class AB, the actual DC power consumption is input power dependent. Harmonic balance simulations show that the actual DC power consumption at the input 1-dB compression level at 75 GHz (i.e., input power of -16 dBm) is 5 mA (still very small compared to the differential Class A main amplifier which runs at 140 mA).

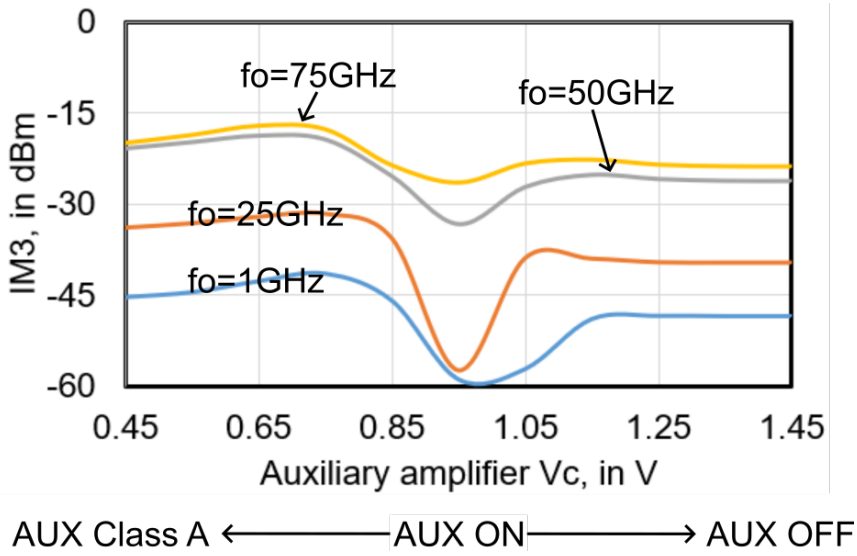


Figure 4.21: Simulated IM3 against various control voltages and centre frequencies in a two-tone test (input power is -25 dBm and tone spacing is 10 MHz)

To examine the relationship between IM3 and V_C, a two-tone test is performed as il-

lustrated in Fig. 4.21. The results reveal a notable reduction in IM3, at a V_C of 0.95 V across centre frequencies ranging from 1 GHz to 75 GHz. The cancellation effect is most pronounced at a centre frequency of 25 GHz and diminishes with increasing frequency. Specifically, at 75 GHz, which is the highest frequency of interest for the linearity specification, OP_{-1dB} improves by 2 dB without incurring additional DC power consumption. Consequently, the IM3 cancellation circuit proves highly beneficial for broadband drivers that must sustain linear operation beyond 50 GHz. However, a limitation of this topology is that the auxiliary amplifier's input signal is drawn from the output of the main amplifier's common-source stage, which increases the R-C time constant at that node, thereby reducing the postamp's bandwidth. Despite this, bandwidth reduction remains acceptable due to the compensated common-gate stages' low input resistance. To recover speed, series-peaking inductors can be employed between the main amplifier stages.

Chapter 5

Physical Layout and Implementation

The 22-nm CMOS-FDSOI technology offers an 11-layer metal stack, with 7 'thin' metal layers (M_1 to M_7) and 4 'thick' metals (M_8 to M_{11}) [61]. Since the postamp stage must deliver substantial RF power at 75 GHz (i.e., 15.2 dBm across a 60-Ω load), the MOSFET layouts are optimized to minimize drain-to-gate and drain-to-source parasitics, and avoid potential instability (see Sections 4.1.2.4 and 4.1.2.5). The implementation of passive networks described in previous chapters (e.g., input matching and output combiner networks) are also presented.

5.1 General considerations for MOSFET physical layout in high-power designs

The transistors are laid out in a grid pattern, with double-gate contacts to reduce the extrinsic gate resistance r_{gg}, as shown in Fig. 5.1. The 80-μm wide transistor has two 40-μm units, where each unit consists of 20×2-μm fingers. The signals access the gate interconnect (green) from the left side and flow vertically to reach the gate grid (gray) in M_1. The gate grid metal (gray) and gate metal (black) are arranged in 20×2-μm unit in each row, as shown in Fig. 5.1. This layout has lower r_{gg} compared to conventional layouts where 40×2-μm fingers are arranged in a single row (e.g., literature reports 66% and 40% reduction in [62], [63], respectively). Simulations show a 8% reduction in r_{gg} for the layout of Fig. 5.1.

The source (red) and drain (blue) interconnects each occupy one-half of the active area to minimize the drain-to-source sidewall capacitance C_{ds}, which is mainly due to fringing

rather than sidewall capacitance. The drain tap (purple) utilizes an M_9 and M_{10} stack to connect to common-gate devices on the top side. The source terminal (yellow) is in a horseshoe shape to relax electromigration constraints by splitting the DC bias current into two paths. The source connects to the second transistor in a differential layout on the right side, using M_9 to minimize the parasitic resistance and common-mode impedance.

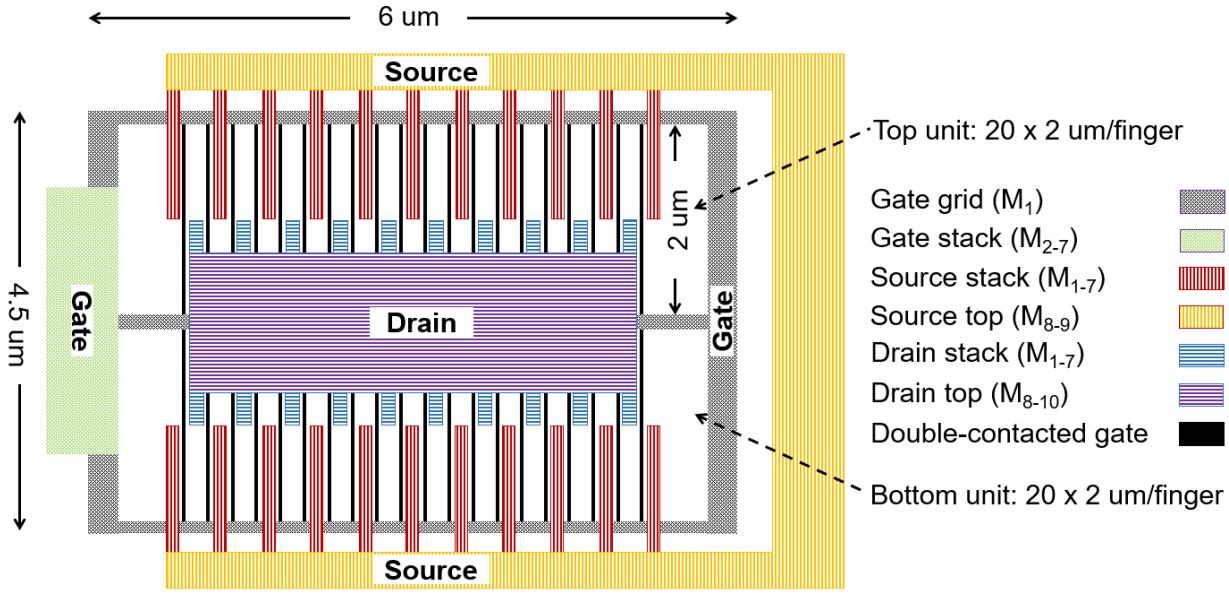


Figure 5.1: Sample FET layout for an 80-μm NFET in Fig. 4.5 (each finger is 2-μm wide and 18-nm long, and 40 fingers are split into upper and lower blocks)

5.1.1 Electromigration considerations

Electromigration is the momentum transfer between conductance electrons and metal ions in the crystal lattice in the direction of DC (or root-mean-square AC) current flow over time. The migration of metal can cause voids and hillocks in the wire, leading to failure of metal connections, and is a well-known reliability concern for IC designers [64].

An 80-μm wide transistor biased at 0.5 mA/μm in the postamp stage requires interconnections that are capable of carrying 40 mA of DC current. This challenges the electromigration limit of BEOL metal layers M_1 to M_7. To satisfy electromigration criteria, the source/drain interconnections utilize a stack of metals (i.e., multiple metal layers in parallel) that rise from the active region to a specific height (M_1 to M_7 in this case). The aggregate current-carrying capability is a multiple of the capacity for a single metal

layer. Current can flow vertically up the stack through vias and distribute itself between layers M_1 to M_7 to reduce the electromigration stress in all of the metals.

The 22-nm CMOS-FDSOI technology provides three different drain-source metal interconnect pitches for the thin-oxide transistor, labeled as 1x, 2x, and 3x FET. Increasing the pitch results in a greater separation between the gate fingers, thereby increasing the overall area occupied by the transistor for a given width. The width of the source/drain metal strips increase as the FET interconnect pitch size increases, so the electromigration capability improves.

The short-length effect reduces electromigration for technologies at the 45-nm node and below [64]. The short-length effect describes the immunity of a wire to electromigration when its length is below the Blech length. At the Blech length, there is a (partial) cancellation between electromigration, due to atoms of metal moving with the current, and stress migration, where metal is moving against the current. Note that aside from the Blech length criteria, the wire width has to be less than three times its minimum width for the short-length effect to occur. At 3x transistor pitch, the source/drain width exceeds three times the minimum width (i.e., short-length exception is no longer valid). Therefore, 2x is selected as the transistor pitch and the gate finger width is kept at 2 μm (i.e., not violating the Blech length criteria) to optimize the electromigration capability. The expected operating lifetime is 100k hours (11 years).

5.2 Passive layout

In this section, the layout of passive networks described in previous chapters, such as the input matching network and the output combiner network, are detailed.

5.2.1 Input matching network

The physical layout of the input matching network of Fig. 4.2 is shown in Fig. 5.2. The red, orange, and cyan colours represent metal layers LB (M_{11}), QB (M_{10}) and QA (M_9), respectively. The 3.6-μm wide LB traces realize the series-peaking inductor L_1 of 115 pH in Fig. 4.2, and connect to the RF input pads at the bottom. The differential lines are separated by 170 μm, which is required by the GSGSG input pad configuration with 100-μm pitch. The differential preamp transistors are connected at the top of the layout, which has a metal-to-metal gap of 24 μm to minimize the differential current loop size. The minimum value for L_1 is determined by the horizontal distance between v_{sig} and v_{in}

(i.e., 70 μm). A 45-degree LB trace is used for L_{1a} and L_{1b}, which is 206-μm long. The second series-peaking inductor L_2 of 11 pH represents the interconnections from the end of L_1 to the gates of the preamp transistors (shown in QB).

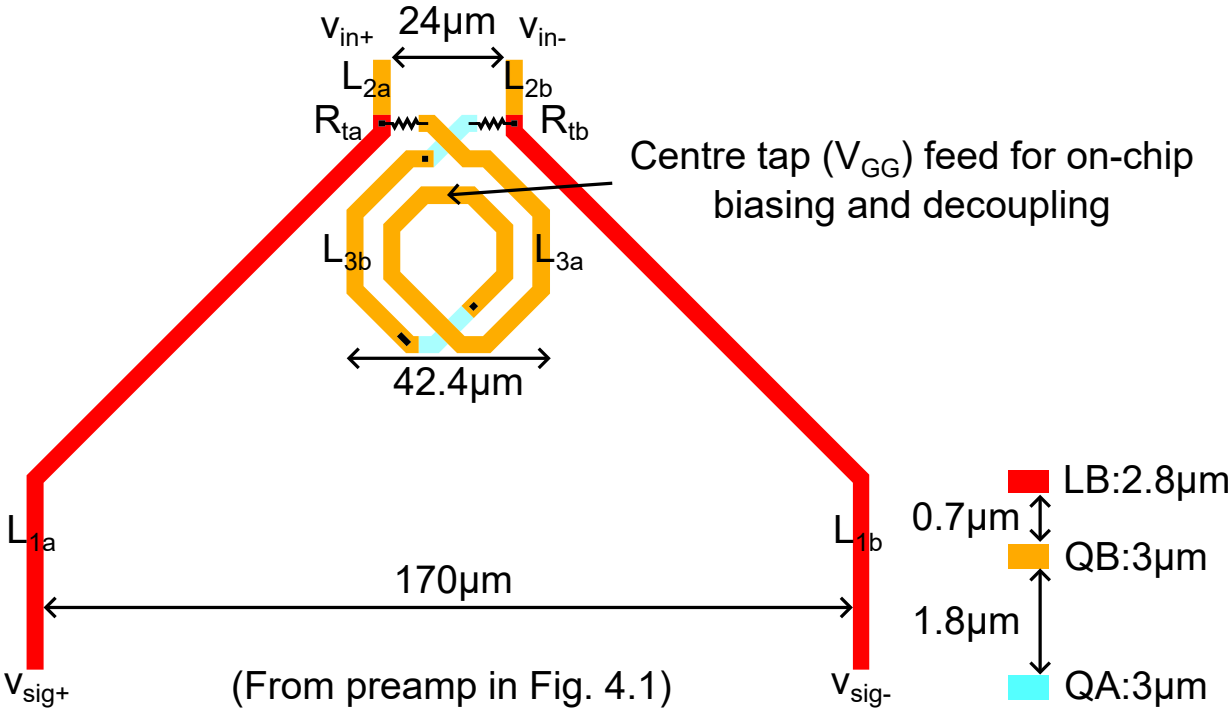
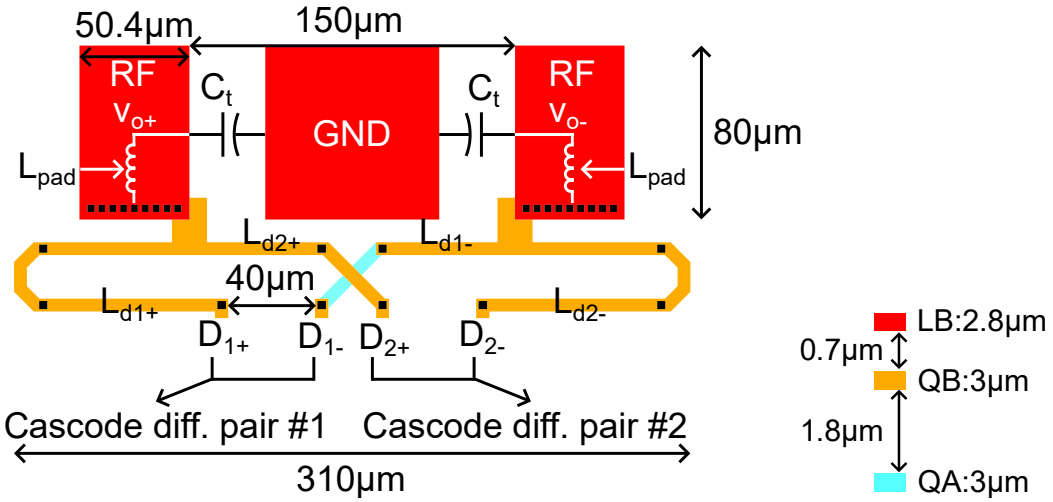


Figure 5.2: Layout of the input matching network of Fig. 4.2

Shunt coil L_3 is a differentially excited spiral inductor realized in the QB layer, and is physically located between L_{1a} and L_{1b}. The outer diameter of L_3 is 42.4 μm and the self inductance is 48 pH per side. The magnetic coupling (recall k of 0.406 from Section 4.1.1) increases the total inductance by the mutual inductance of 19.5 pH (i.e., 67.5 pH overall). The 50-Ω on-chip input termination resistors R_{ta} and R_{tb} connect the ends of L_1 and L_3. They are 4-μm long and 2.5-μm wide. For the on-chip bias option shown in Fig. 4.2, a biasing network may be connected to the centre tap of L_{3a} and L_{3b} in parallel with a decoupling capacitor. This sets the lower cutoff frequency of the driver and determines the common-mode input return loss.

5.2.2 Output combining network

The output combining network transforms a 30-Ω load at the RF output to the optimum single-ended load, $Z_{opt,se}$ of 29.4+j29.3 Ω at the drain of each transistor at 75 GHz. The physical layout of the output combining network is shown in Fig. 5.3, where series coils L_{d1+} to L_{d2-} of 90 pH are realized in 6-μm wide QB metal (orange). The parasitic inductances of 19 pH for the RF signal pads (red) in LB metal are labelled as L_{pad}. The pads are 80-μm long and 50.4-μm wide (aluminum LB top metal) with 100-μm pitch (outer ground pads not shown). The tuning capacitances C_t of 54 fF between the signal and adjacent ground pads together with the parasitic pad capacitance of 24 fF result in 78 fF of total load capacitance at the RF output.



(a) Output combiner (b) Metal stack

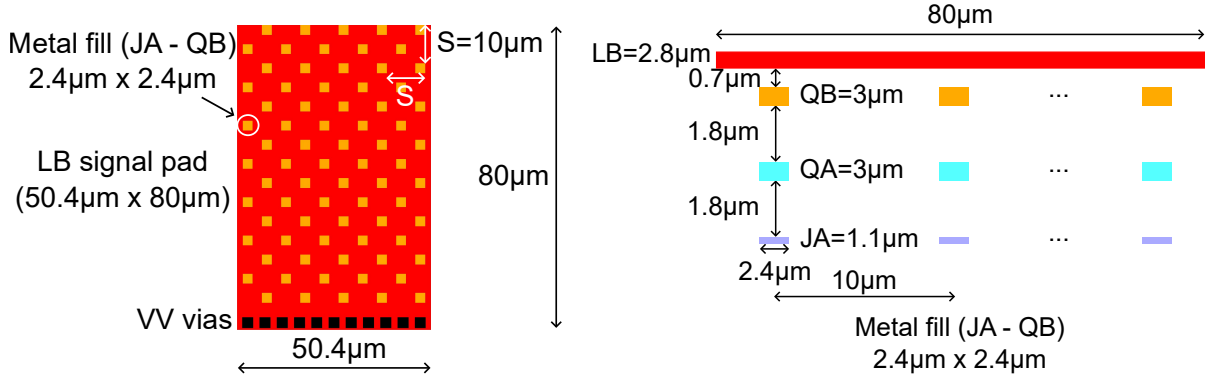
Figure 5.3: Physical layout for the output combining network of Fig. 4.10 (outer ground pads adjacent to each RF pad not shown)

Physical layout of the output combining network is determined by the pitch of the GSGSG output pads. Recall that a bridging capacitor is utilized at the gate of the cascode devices in the differential output stage (see Fig. 4.8). The parasitic inductance of the capacitor in series with the gate must be minimized to ensure stability of the common-gate stages. Therefore, the postamp differential pairs are physically close to each other in the layout (i.e., D_{1+} is close to D_{1-} and D_{2+} is close to D_{2-}). The drains of the cascode device outputs are denoted by D_{1+} to D_{2-}, respectively, in Fig. 5.3a. Nodes D_{1+} and D_{1-} are 40-μm apart (same for D_{2+} and D_{2-}). Since the L_d required is only 90 pH and the signal pads

are 200-μm apart (centre to centre), design priority is given to locating the inner drains (i.e., D_{1-} and D_{2+}) to set L_{d2+} and L_{d2-} at 90 pH. The outer drains (i.e., D_{1+} and D_{2-}) are routed with longer metal traces in a hairpin shape, and mutual inductance between the hairpin branches is utilized to partially cancel the self-inductance of the outer traces. Final values for inductors L_{1+} to L_{2-} are 85 pH, 92 pH, 90 pH and 85 pH, respectively (i.e., close to the desired L_d value of 90 pH).

5.3 Metal fill and parasitic optimization

As mentioned in Section 5.2.2, the impedance Z_{se} seen at each drain node is selected to maximize the linear output power at 75 GHz. The optimum drain impedance ($Z_{opt,se}$) is affected by the pad capacitance, which is substantial due to metal filling underneath the signal pad (metal fill reduces mechanical stress in the BEOL metal stack [65]). For 80×50.4 μm^2 signal pads, the pad capacitance (C_{pad}) is 31 fF when a fill density of 60% is assumed, or 20 fF if there is no metal fill. As a result, Z_{se} changes from $13.5 + j15.2\ \Omega$ for 60% fill (a voltage-standing-wave ratio, VSWR of 2.91) to $15.6 + j14.4\ \Omega$ (VSWR of 2.49) without filling. Since metal fill is time-consuming to simulate accurately due to the complex fill patterns, manual metal filling is applied with the pattern shown in Fig. 5.4.



(a) Top view (b) Side view

Figure 5.4: RF signal pad with manual metal filling pattern ($M_4 - M_7$ ground plane not shown)

The metal fill consists of a stack of JA to QB layers (without vias) underneath the LB pad with a unit size of 2.4×2.4 μm^2. The centre-to-centre spacing for the fill objects is

10 μm. Beneath JA, there is a ground plane comprised of M_4 to M_7 stack that runs across the GSGSG pads (not shown in Fig. 5.4). The pattern satisfies the local minimum metal density requirement for the 22-nm CMOS technology and adds less parasitic capacitance than metal fill applied by automation scripts provided by the design kit. The EM simulated result for C_{pad} with the manual metal fill of Fig. 5.4 is 24 fF. Manual filling also improves simulation accuracy at the output compared to auto-filling, and provides a predictable value for $Z_{opt,se}$ that maximizes the linear output power.

5.4 Peak gain and peak frequency over process, voltage, temperature variations

In the previous context, active and passive components within the driver are assumed to operate at the 'typical' process, 'nominal' supply voltage (i.e., V_{SSP} of 2.97 V as defined in Fig. 4.2, V_{DDN} of 3 V as defined in Fig. 4.5), and an ambient temperature of 27 °C. In reality, the device may stray away from the 'typical' process (e.g., PFETs and NFETs may experience more/less parasitic capacitance and carrier mobility). The supply may vary around the 'nominal' voltage, and environmental temperature may differ from the assumed 27 °C.

This section quantifies changes in the magnitude response (peak gain and peak frequency) due to process, voltage, and temperature (PVT) variations using simulation.

5.4.1 Peak gain and peak frequency against process, voltage and temperature variations

Monte Carlo (MC) simulations are performed to gain insight into distributions of peak gain and frequency of the gain peak (A_{pk} and f_{pk}) against process variations. Voltage and temperature variations are simulated across multiple points. Process, voltage and temperature (PVT) settings are as follows:

1. Process (MC enabled): randomly-distributed (Latin hypercube sampling) process corners
2. Voltage (MC disabled): ±5% of nominal supply, resulting in three pairs of [$|V_{SSP}|$, $|V_{DDN}|$], that is, low-voltage: [2.82 V, 2.85 V], nominal-voltage: [2.97 V, 3 V], high-voltage: [3.12 V, 3.15 V] (V_{SSP} as defined in Fig. 4.2, V_{DDN} as defined in Fig. 4.5).

3. Temperature (MC disabled): 0, 27, 60, and 110 °C (simulation of functionality down to -40 °C for 'typical' process and nominal voltage simulations)

4. Samples: 1000 MC samples over process for each temperature and voltage combinations (i.e., 12000 samples in total)

5.4.2 Effects of supply voltage variations (at 27 °C and 'typical' process, MC disabled)

First, effects on voltage variations are simulated. Figure 5.5 shows changes in A_{pk} and f_{pk} versus various supply voltages at 27 °C and 'typical' process (TT, representing 'typical-speed' NFET and 'typical-speed' PFET). As the voltage setting changes from low-voltage to high-voltage (see voltage settings in Section 5.4.1), A_{pk} is between 8.55 dB and 8.8 dB (blue curve in Fig. 5.5), and f_{pk} remains unchanged at 76 GHz (red curve in Fig. 5.5). These changes are minor, and one can conclude that for nominal temperature and process parameters that the driver is insensitive to supply voltage variations.

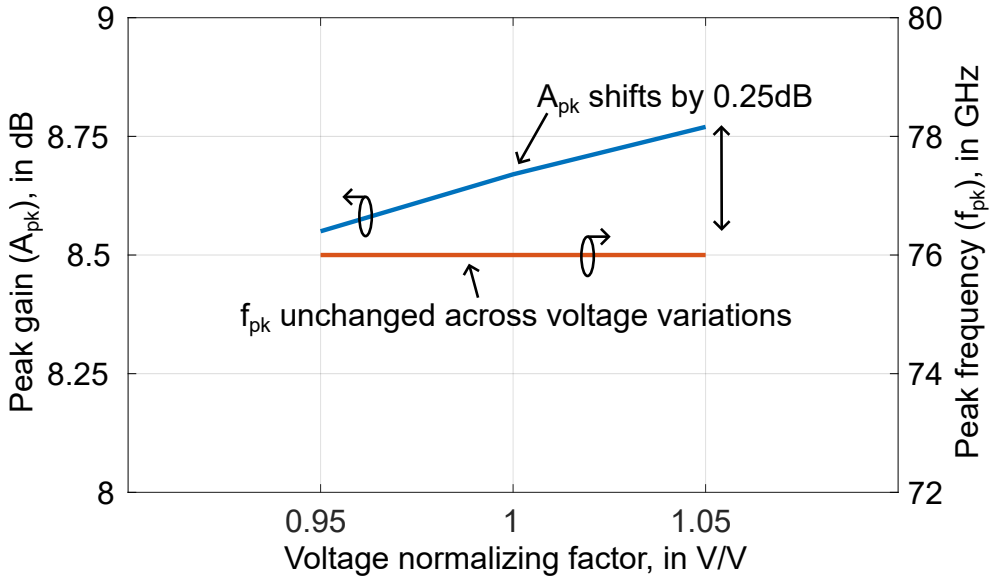


Figure 5.5: Changes in A_{pk} and f_{pk} against ±5% shifts in supply voltages V_{SSP} and V_{DDN} (Voltage normalizing factor = $V_{DDN}/3 = V_{SSP}/2.97$; Process: TT; Temperature: 27 °C)

5.4.3 Effects of temperature variations (at nominal voltage and 'typical' process, MC disabled)

When supply voltages are held at their nominal values (i.e., V_{SSP} and V_{DDN} are 2.97 V and 3 V, respectively), and the temperature is varied from -40 °C to +110 °C at the TT process corner, A_{pk} decreases from 11.8 dB to 4.8 dB. Applying linear regression to the simulation results yields a first-order polynomial relating A_{pk} and temperature (T),

$$\begin{aligned} A_{pk} &= -0.047T + 9.94 \quad (R^2 = 1), \\ f_{pk} &= -0.0273T + 76.89 \quad (R^2 = 0.95). \end{aligned} \tag{5.1}$$

Parameter R^2 is the coefficient of determination representing how close the linear equation is to the data points ($R^2 = 1$ means a perfect fit). Thus, the relationship between A_{pk} in dB and temperature (T in degree Celsius) is linear (blue curve in Fig. 5.6). On the other hand, f_{pk} decreases from 78 GHz to 73 GHz as temperature increases (red curve in Fig. 5.6).

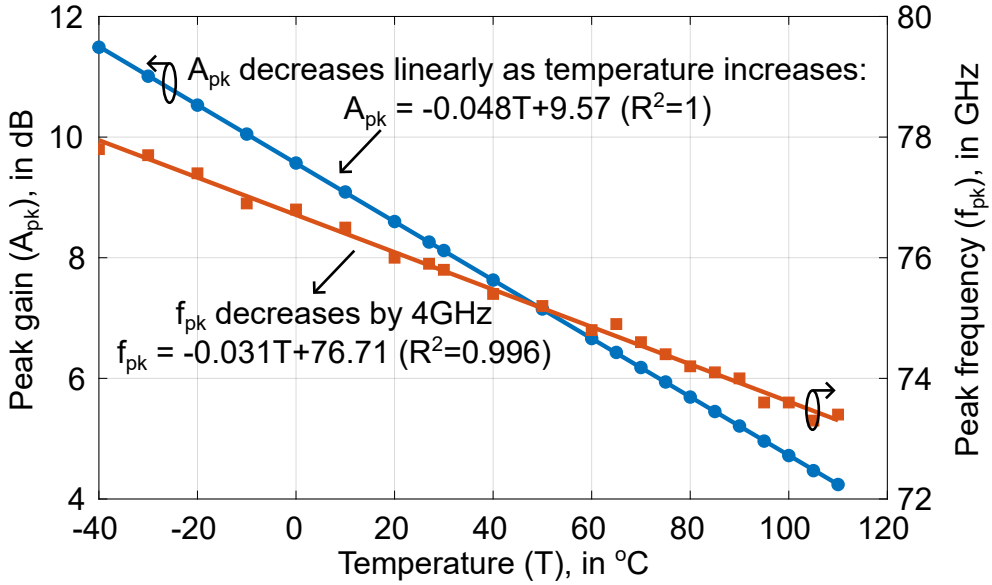


Figure 5.6: Changes in A_{pk} and f_{pk} against various temperatures from -40 °C to +110 °C (Process: TT; Voltage: V_{SSP} of 2.97 V, V_{DDN} of 3 V)

In conclusion, temperature significantly affects A_{pk} and f_{pk}. The gain should be monitored and controlled (e.g., using a feedback loop) to maintain the desired A_{pk} as tempera-

ture changes. The peak frequency may be adjusted by tuning inductor L_1 for the interstage filter network, as explained in Section 3.1.2.2.

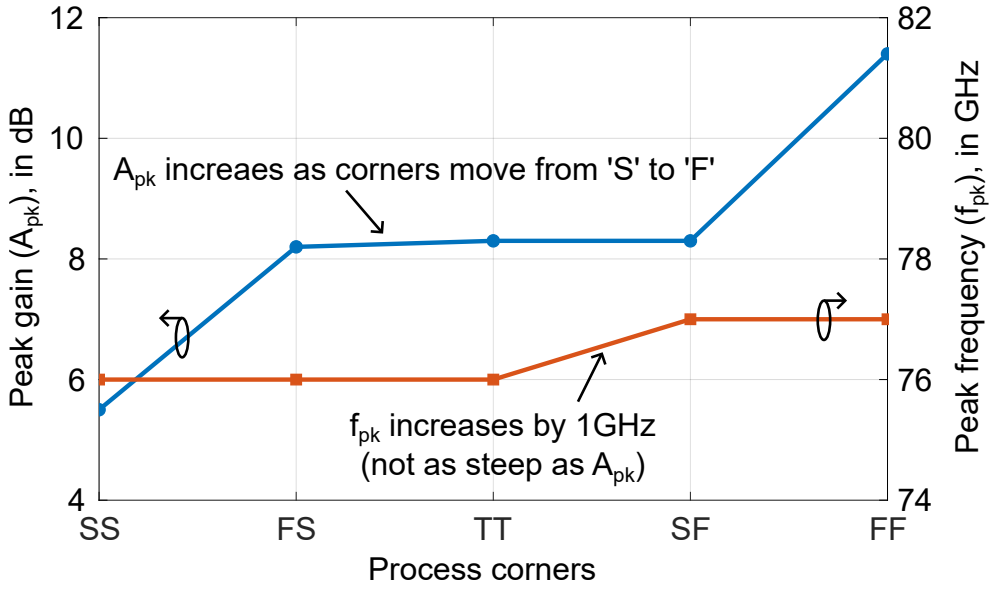
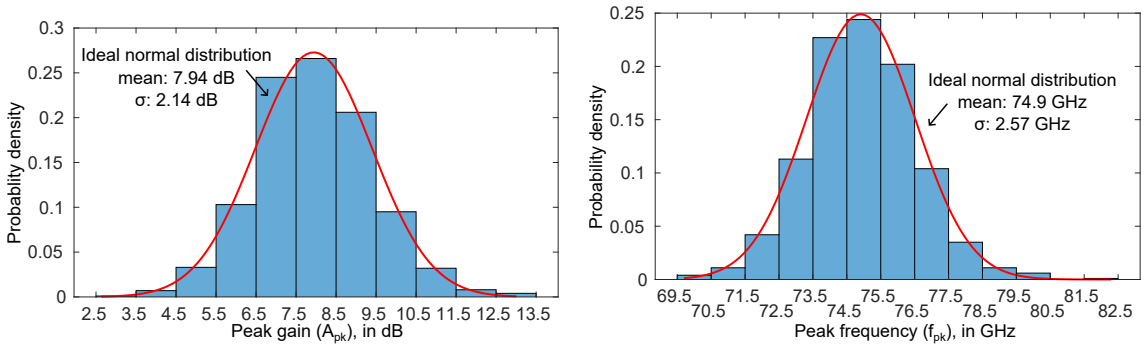


Figure 5.7: Changes in A_{pk} and f_{pk} against various FET process corners (e.g., SF corner represents slow NFET and fast PFET; Temperature: 27 °C; Voltage: V_{SSP} of 2.97 V, V_{DDN} of 3 V)

5.4.4 Effects of process variations for various FET corners (at nominal voltage and 27 °C, MC disabled)

When simulating at 27 °C with nominal supply voltages and varying process for FETs (e.g., from the 'slow' SS process to 'fast' FF process, where SS represents slow mobility and high parasitic capacitance for both NFET and PFET), the resulted A_{pk} and f_{pk} are shown in Fig. 5.7. The peak gain increases from 5.5 dB (SS process) to 11.4 dB (FF process). The other three corners report an A_{pk} of 8.3 dB (3 dB higher than the SS corner and 3 dB lower than the FF corner). That is, A_{pk} increases with increasing carrier mobility and decreasing parasitic capacitance (NFET and PFET mobilities have similar effects on the overall A_{pk}). The peak frequency f_{pk} increases slightly from 76 GHz to 77 GHz from the SS to FF corner.

5.4.5 Effects of global process variations (at nominal voltage and 27 °C, MC enabled)



(a) Distribution of A_{pk}

(b) Distribution of f_{pk}

Figure 5.8: Distribution of A_{pk} and f_{pk} resulted from Monte Carlo simulations for process variations (1000 samples in total. Voltage: V_{SSP} of 2.97 V, V_{DDN} of 3 V; Temperature: 27 °C)

In reality, active and passive devices experience process variations, and Monte Carlo (MC) simulations are performed to capture the effects of process random distribution on A_{pk} and f_{pk}. When simulating at 27 °C with nominal supply voltages and random process (i.e., MC enabled globally), the resulted A_{pk} and f_{pk} are shown in Fig. 5.8.

The mean and variance for A_{pk} (1000 samples) are 7.94 dB and 2.14 dB, respectively (e.g., 68% of the data lie between 6.5 dB and 9.5 dB). The average A_{pk} is 0.5-dB lower than the simulated A_{pk} of 8.5 dB in Table F.4 or 1-dB lower than the 9-dB specification. The distribution is slightly left-skewed (e.g., the 6.5 dB-7.5 dB bin is higher than the 8.5 dB-9.5 dB bin), which could be that the peaking profile depends on the preamp transconductance, input/output capacitances, interstage load resitance and interstage filter network. When process MC is enabled, the passive devices from the interstage filter network which are EM-simulated do not have MC enabled.

For f_{pk}, the mean and variance are 74.9 GHz and 2.57 GHz, respectively (e.g., 68% of the data lie between 72.5 dB and 77.5 dB). The average f_{pk} is close to the desired 75 GHz. Both A_{pk} and f_{pk} pass the normality check (i.e., Kolmogorov-Smirnov test in MATLAB is used, and A_{pk} and f_{pk} could be considered Gaussian distributed).

5.4.6 Aggregate effects of process, voltage and temperature variations

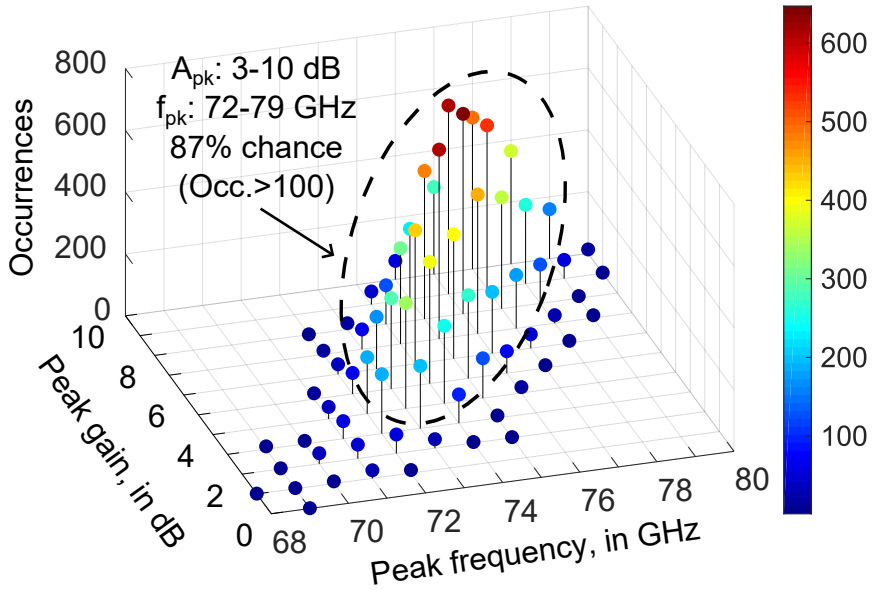


Figure 5.9: Simulations for A_{pk} and f_{pk} against process, voltage and temperature variations (12000 samples in total), minimum occurrence of 100 are shown (Process: randomly distributed, MC-enabled; Voltage: [$|V_{SSP}|$, $|V_{DDN}|$] pairs of [2.82 V, 2.85 V], [2.97 V, 3 V], and [3.12 V, 3.15 V]; Temperature: 0, 27, 60, and 110 °C)

Figure 5.9 shows the aggregate simulated results including process, voltage and temperature variations. The simulation settings comprised of three voltages (low-voltage, nominal-voltage and high-voltage), four temperature (0, 27, 60, and 110 °C), and MC-enabled process (i.e., in total data from 12000 simulation runs). For each coloured dot, the height represents the number of simulation runs (occurrences), which results in a specific combination of peak gain A_{pk} and peak frequency f_{pk}. The bar is comprised of colours from blue to red, where blue and red represent the least and most frequent occurrences, respectively. From the data, the peak gain A_{pk} lies between 3 dB and 10 dB, and the frequency of the gain peak f_{pk} lies between 72 to 79 GHz with a probability of 87%. Figure 5.10 filters out results with less than 300 occurrences, where 54% of samples predict an A_{pk} within 4-10 dB and an f_{pk} within 73-78 GHz. Overall, A_{pk} has a wider variation than f_{pk} (e.g., for a desired peak frequency of 75 GHz, majority of samples lie in the range within ±3 GHz of 75 GHz, whereas for a desired peak gain of 9 dB, A_{pk} varies from 4 dB to 10 dB).

When analyzing A_{pk} and f_{pk} separately, the mean and variance for A_{pk} (12000 samples) are 7.45 dB and 6.32 dB, respectively (e.g., 68% of the data lie between 5 dB and 10 dB). The centre is 1.5 dB lower than the desired 9 dB (i.e., bias control loop or thermal control for the amplifier needed to tune A_{pk}). Compared to MC simulation results for only the process variation (Section 5.4.5), A_{pk} mean decreases by 1 dB and variance increases by 1.7 dB. The distribution of A_{pk} is shown in Fig. 5.11a, which is no longer Gaussian compared to Fig. 5.8a (due to voltage and temperature settings are fixed combinations). For a range within ±2.5 dB of the desired A_{pk} of 9 dB, the aggregate probability is 61.1%. For f_{pk}, the mean and variance (12000 samples) are 75.6 GHz and 4.22 GHz, respectively. The distribution of f_{pk} is shown in Fig. 5.11b (also no more Gaussian compared to Fig. 5.8b). Compared to Section 5.4.5, f_{pk} mean increases by 0.5 GHz and variance decreases by 2.4 GHz. For a range within ±2.5 GHz of the desired f_{pk} of 75 GHz, the aggregate probability is 76.4%.

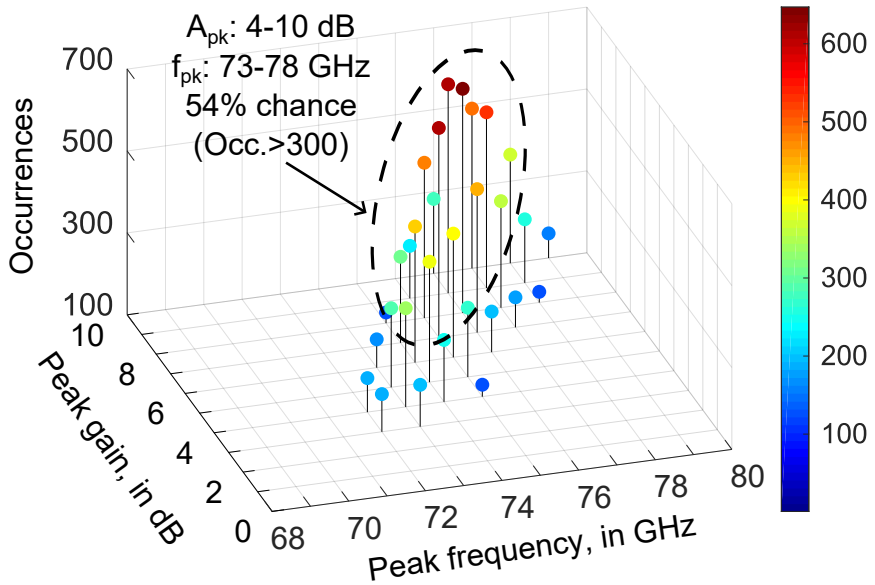
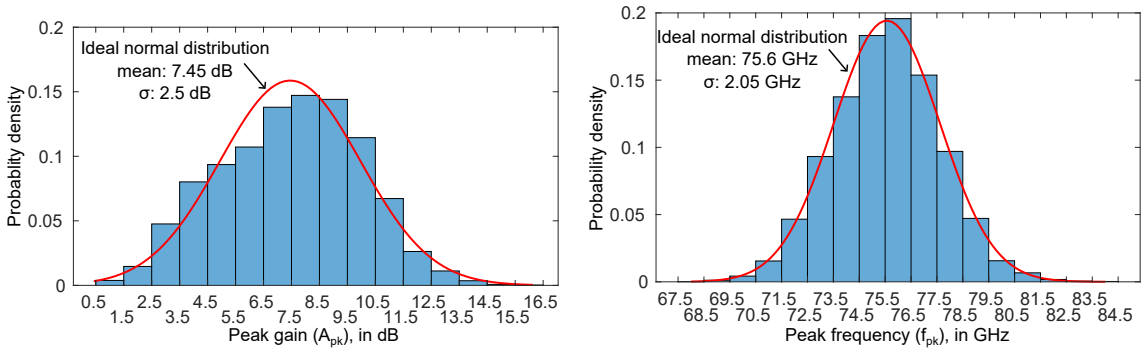


Figure 5.10: Simulations for A_{pk} and f_{pk} against process, voltage and temperature variations (12000 samples in total), minimum occurrence of 300 are shown (Process: randomly distributed, MC-enabled; Voltage: [$|V_{SSP}|$, $|V_{DDN}|$] pairs of [2.82 V, 2.85 V], [2.97 V, 3 V], and [3.12 V, 3.15 V]; Temperature: 0, 27, 60, and 110 °C)



(a) Distribution of A_{pk} (non-Gaussian)

(b) Distribution of f_{pk} (non-Gaussian)

Figure 5.11: Distribution of A_{pk} and f_{pk} against process, voltage and temperature variations (12000 samples in total. Process: randomly distributed, MC-enabled; Voltage: [$|V_{SSP}|$, $|V_{DDN}|$] pairs of [2.82 V, 2.85 V], [2.97 V, 3 V], and [3.12 V, 3.15 V]; Temperature: 0, 27, 60, and 110 °C)

Chapter 6

Experimental Results

This chapter presents the experimental results of a two-stage broadband driver prototype fabricated in a 22-nm CMOS-FDSOI technology. The chapter begins with a detailed description of the calibration and measurement setup, including the instruments and methodologies employed for both frequency-domain and time-domain measurements. The frequency-domain measurements focus on small-signal S-parameters, compression points and total harmonic distortion, while the time-domain measurements evaluate the driver's performance with serial data inputs, including NRZ and PAM-4 modulation schemes. Note that optical measurements are not performed because electro-optical testing is outside the scope of the work.

The driver performance is measured across temperatures, with the chip backside temperature controlled using a thermoelectric heater/cooler. Eye pattern is presented to assess the driver's time-domain performance. The chapter concludes with a comparison with other state-of-the-art broadband drivers, highlighting the prototype's strengths and suggesting areas for improvement. The findings provide valuable insights into the design and optimization of broadband drivers for high-speed communication systems.

The micrograph of the two-stage broadband driver prototype is shown in Fig. 6.1. The total chip area is 0.78×0.88 mm^2, and the driver core area is 0.55×0.31 mm^2.

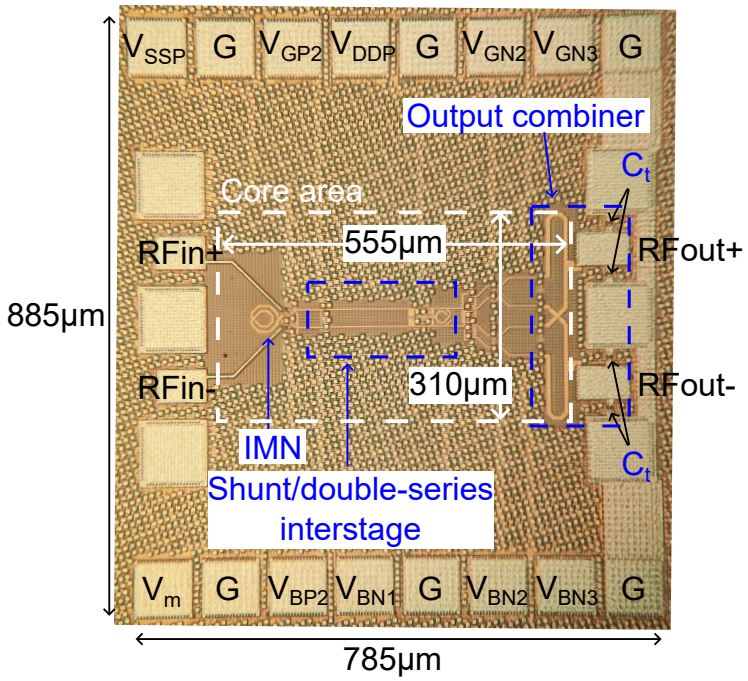


Figure 6.1: Micrograph of the two-stage broadband driver prototype

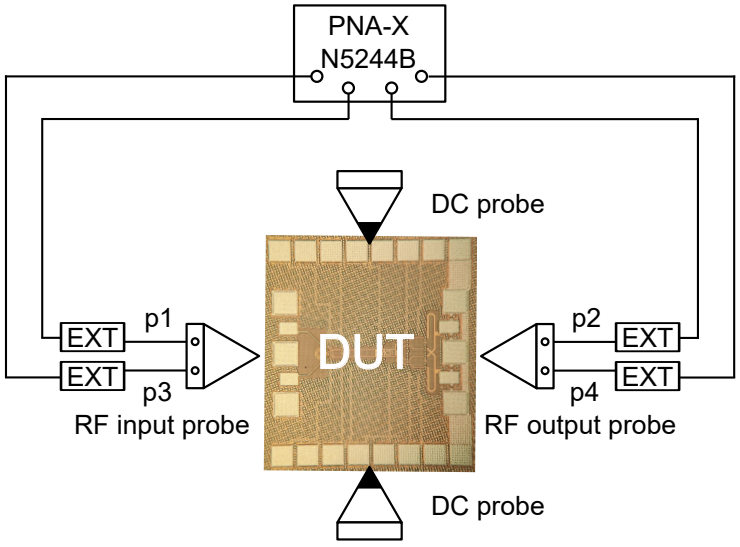


Figure 6.2: 110-GHz, 4-port frequency-domain measurement setup

6.1 Calibration and measurement setup

Four-port small-signal S-parameters (10 MHz - 110 GHz) and the large-signal 1-dB compression point (P_{-1dB}) are measured using a vector network analyzer (Keysight® N5244B) with frequency extender heads (Keysight® N5293AX02 with built-in bias-T, see Fig. 6.2). Calibration is performed with the short-open-load-reciprocal (SOLR) method, and the reference plane for measurement is at the RF on-wafer probe tips. A source power calibration is performed at the RF input ports for (e.g., large-signal compression point) measurements.

The time-domain measurement setup consists of an arbitrary-waveform generator (AWG, Keysight® M8199B, 256 GS/s, 75 GHz analog bandwidth) to create NRZ and PAM-4 test signals that drive the input pads. Channel calibration is performed for all serial data transmission measurements. A sampling scope (Keysight® DCA-X N1000A) with a pair of 100-GHz bandwidth sampling heads (Keysight® N1046A) measure the driver outputs, as shown in Fig. 6.3.

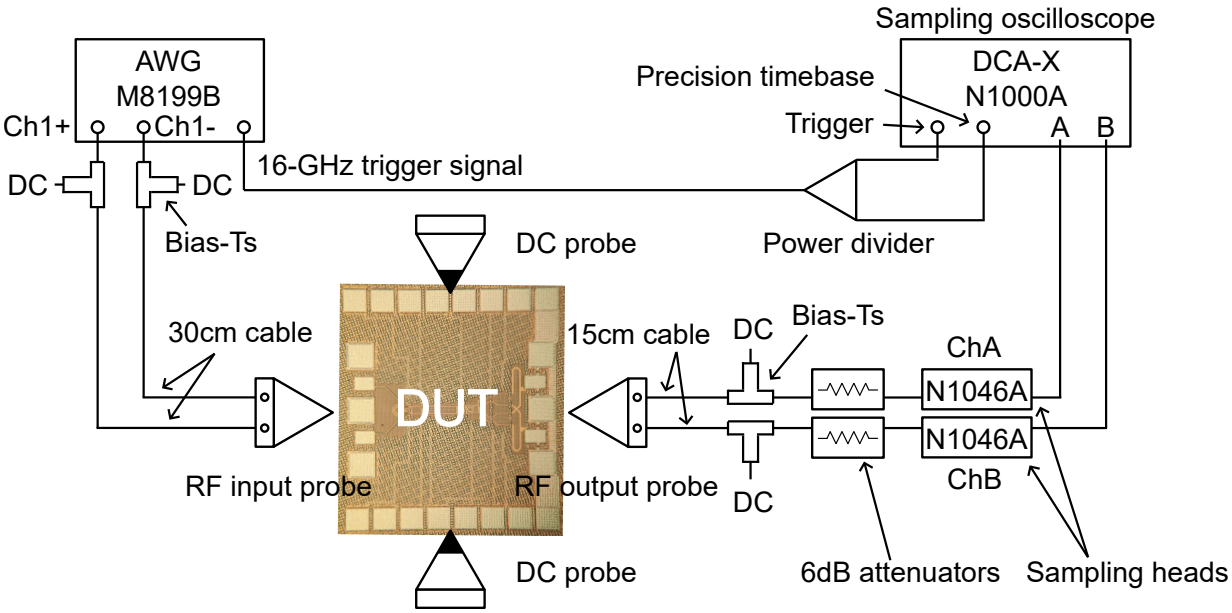


Figure 6.3: Time-domain measurement setup

In Section 5.4, Monte Carlo simulations show that temperature variations have the greatest effect on the peaking amplitude (see Fig. 5.6) compared to voltage and process variations (e.g., a 100°C increase in the simulation temperature leads to 5-dB reduction in

the peaking amplitude), which motivates a temperature-controlled test setup to validate the simulation results.

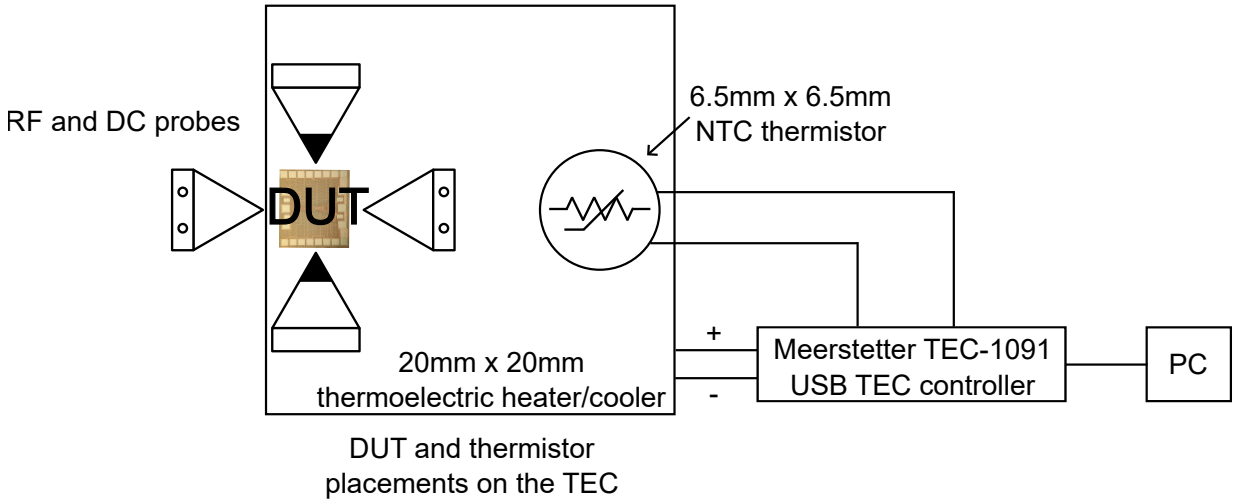


Figure 6.4: On-chip probing and back-side temperature control setup

A thermoelectric heater/cooler (TEC) is used to control the temperature of the backside of the chip. The voltage and current of the TEC are set by a thermoelectric controller (Meerstetter TEC-1091 [66]) with a negative temperature coefficient thermistor sensor in a proportional–integral–derivative (PID) control loop that maintains a constant temperature. The chip sample and thermistor are mounted on the TEC using conductive epoxy adhesive, which ensures good thermal conductivity between the TEC surface, chip backside and thermistor. However, due to space constrains, the thermistor cannot be situated adjacent to the device under test (DUT). Instead, it is placed on the TEC's edge opposite to the DUT (see Fig. 6.4). The thermistor's temperature reading is less than the actual circuit temperature owing to the intervening distance. To achieve a more precise measurement of the circuit temperature, the input termination resistor is employed as an on-chip thermistor to approximate the circuit temperature.

6.2 DC bias measurement results

In this section, results of the DC bias current measurements for the broadband driver are summarized. The driver is biased with a nominal DC current of 106 mA at room temperature for the postamp at a back gate bias of 0.6 V. Note that although room

temperature is 24 °C during measurement, the TEC controller reads 26.5 °C from the NTC thermistor (TEC control off) due to the heat generated by the driver. Moreover, the actual device temperature must be higher than values reported by the TEC controller, as explained in Section 6.1. The DC bias current of the postamp increases as the temperature increases, as shown in Fig. 6.5.

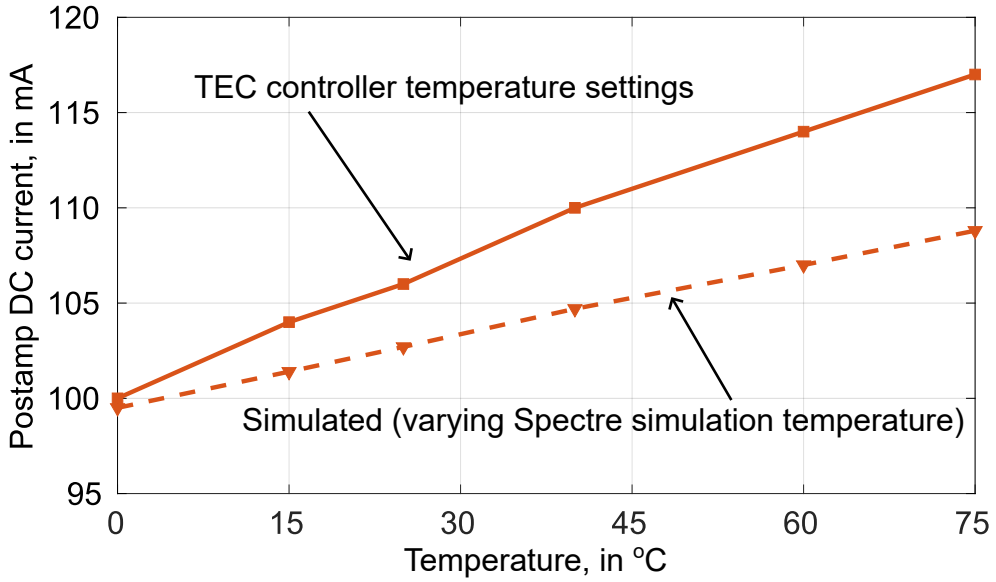


Figure 6.5: Measured DC bias current for the postamp against various TEC controller temperature settings. Simulated DC bias current is also plotted for the same range of device temperature using Spectre simulations

The measured result is shown with the solid line, while the simulated result is depicted by a dashed line. A clear discrepancy is seen between simulated and measured results. The slope of the measured result is steeper than in simulations, indicating that a higher temperature dependency is seen during measurements, which predicts that the temperature read by the off-chip thermistor should not be directly used as the temperature in Spectre simulation. Ideally, if the device temperature during measurement could be set by the TEC (i.e., TEC temperature equals the simulation temperature in Spectre), then the solid and dashed lines would match. However, in reality, the TEC could only set the chip backside temperature (i.e., temperature at the heat sink) and the actual device temperature (i.e., temperature at the heat source) will be higher than TEC due to finite thermal resistance of the substrate [67].

Body-to-source voltages (V_{BS}) for the postamp transistors are set to 0.6 V for small-signal measurements. The measured bias currents of the preamp and postamp are 30 mA and 106 mA at room temperature (TEC off), respectively. The preamp is supplied by 3 V (top rail) and 0.26 V (bottom rail), and the postamp has a single 3-V supply. The DC power consumptions are 82 mW and 318 mW for the preamp and postamp, respectively, at room temperature (i.e., in total 400 mW for the driver).

To detail the temperature delta between the off-chip NEC thermistor reading (i.e., chip backside) and driver in operation, the on-chip input termination is used as temperature sensor (i.e., it is closer to the on-chip heat sources compared to the off-chip NEC thermistor), whose resistance is measured via an ohmmeter (neglecting the input pad and feed line losses at DC). By measuring the termination resistance over multiple TEC controller temperatures and comparing with calculated temperature-dependent resistances, the difference between the NEC thermistor reading and on-chip temperatures are extracted, as shown in Fig. 6.6.

The dashed blue curve shows the calculated value for the input termination resistance using a first-order linear equation provided by the design kit [68], covering a device temperature from 15°C to 105°C. Note that the calculation assumes the device under test temperature is the same as the backside temperature (i.e., controlled by the TEC), which is not the case for SOI devices, as the insulating buried oxide layer has a much lower thermal conductivity compared to a regular bulk CMOS substrate (e.g., the thermal conductivity of bulk CMOS and SOI substrates are approximately 140 W/m·K [69] and 1.4 W/m·K [70] at 300 K, respectively). Therefore, a temperature delta is expected between chip backside (i.e., TEC temperature) and active circuits (i.e., simulation temperature used in Spectre) for SOI substrates. The TEC sets the chip backside temperature to a constant, and therefore only controls the device temperature (i.e., chip topside) indirectly.

The solid red curve shows the measured resistance value when the driver is turned off, while the backside temperature is set by the TEC controller. There is a discrepancy between the slopes of the red and blue curves, which may be related to the metal probes dissipating heat when measuring the input termination resistance, especially at high temperatures. During measurements, a steeper resistance slopes is seen when only the input RF probe is landed on the chip. The red curve is measured with all four probes (2 RF probes and 2 DC probes, as shown in Fig. 6.2) contacted to the chip, which has the best fit to the blue curve (calculated results). When the driver is turned on, the measured input termination resistance is shown in the solid yellow curve, which is 3-Ω higher across TEC temperatures compared to the case when the driver is off (solid red). This is because the same thermal dissipation network (i.e., TEC at the backside and probes at the frontside) now has to dissipate the additional heat generated by the driver (i.e., 400 mW), resulting

in a larger difference compared to the calculated values (dashed blue curve).

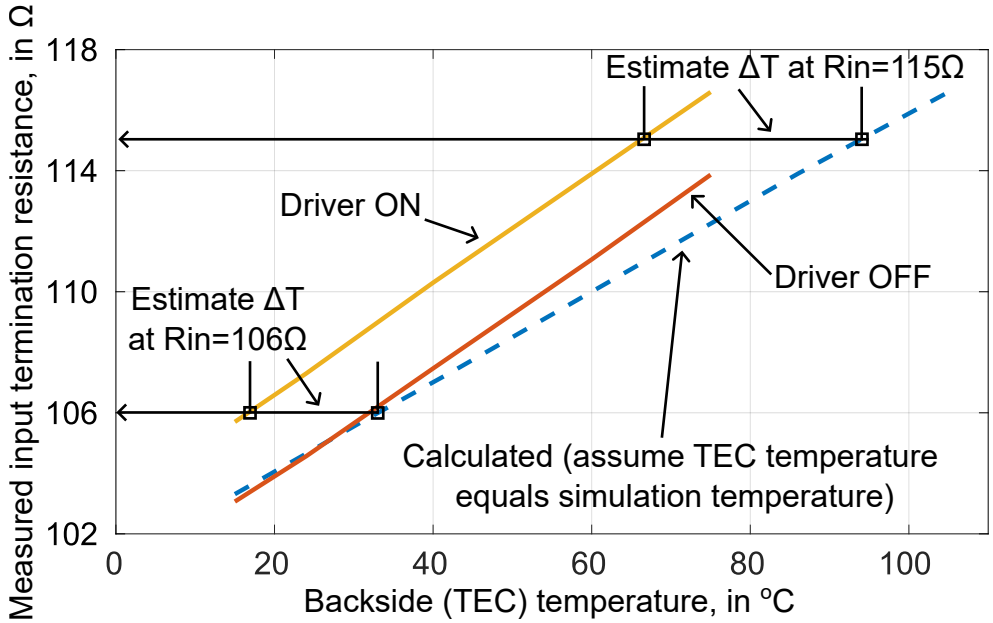


Figure 6.6: Measured on-chip termination resistance at the driver input against various TEC controller temperature settings. Calculated resistance is also plotted assuming that the TEC temperature equals circuit temperature

At a given resistance value, the temperature difference (ΔT) between the measurement (solid yellow) and the calculation using the first-order linear equation provided by the design kit (dashed blue) estimates the delta between the chip backside and circuit temperatures. For example, for an input termination resistance of 105.7 Ω on the y-axis of Fig. 6.6, the corresponding chip backside and circuit temperatures are 15°C and 31°C, respectively, resulting in an estimated ΔT of 16°C. For an input termination resistance of 116.6 Ω, the corresponding chip backside and circuit temperatures are 75°C and 105°C, respectively, resulting in an estimated ΔT of 30°C. Therefore, as the chip backside temperature increases, ΔT increases, indicating a larger delta between the chip backside and circuit temperature. When the three curves are compared together (calculated, measured with driver off and measured with driver on), the measured result with driver off (solid red) has a steeper slope compared to the calculated result (dashed blue), which is due to the on-chip probe setup affecting the heat dissipation. The measured result with driver on (solid yellow) has the same slope as the measured result with driver off (solid red),

indicating that the heat dissipation of the chip is unchanged with the same on-chip probe setup. The shift along the y-axis is due to the additional heat generated by the driver.

Therefore, to attain more accurate simulation results, a ΔT of 20°C is selected because the input termination resistance measured at the backside 7°C matches the simulated result using circuit temperature of 27°C in Cadence Spectre.

6.3 Small-signal measurement results

In this section, measured small-signal results for the broadband driver are summarized. The measurement setup for frequency-domain tests is shown in Fig. 6.2. To illustrate the effects of temperature on small-signal performance, the TEC controller is used to vary the chip backside temperature from 7°C to 75°C. This TEC temperature range is limited by the power specification of the TEC controller.

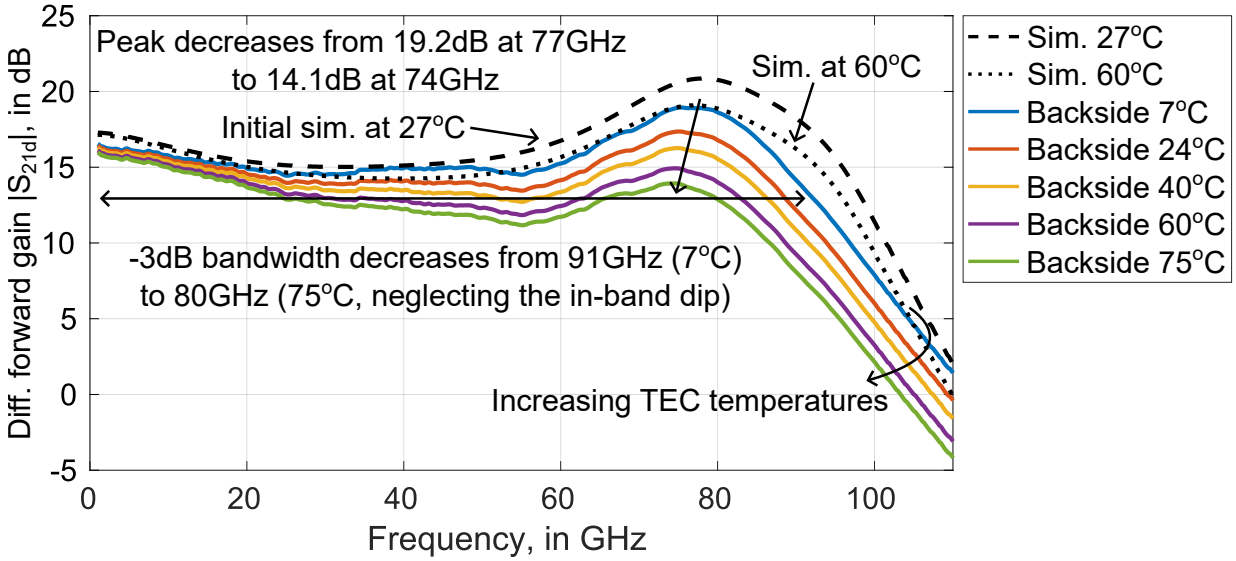


Figure 6.7: Measured small-signal differential forward gain (S_{21d}) versus various TEC temperature settings for 100-Ω differential source and load impedances. The simulated result for temperature settings of 27°C and 60°C in Spectre are also plotted for comparison

The measured differential forward gain ($|S_{21d}|$) for a 100-Ω differential load versus various TEC temperature settings is shown in Fig. 6.7. At 1 GHz, the highest gain is 16.8 dB at backside 7°C and reduces by 0.5 dB at the max TEC temperature setting (75°C). At a

backside temperature of 7°C, the peak gain is 19.2 dB at 77 GHz (i.e., 2.4-dB peaking) and the -3dB bandwidth is 91 GHz. The maximum in-band dip at 7°C is 2 dB from 25 GHz to 55 GHz. As the TEC temperature setting increases from 7°C, the peak gain decreases to 14.1 dB at 74 GHz (i.e., no peaking) for a backside temperature of 75°C and the -3dB bandwidth lowers to 80 GHz. Moreover, the in-band dip increases to 5 dB at 55 GHz.

Note that the gain at 1 GHz fluctuates by 1 dB against changes in chip backside temperature, which is significantly less than the peak gain variation of 5.1 dB. To explain this observation, the interstage resistance and transconductance of an NFET unit cell used in the postamp are simulated at various temperature and frequencies.

The DC resistance of the interstage load resistor is plotted in Fig. 6.8 at 27°C and 60°C. The DC resistance increases as the simulation temperature increases. However, when simulating the short-circuit transconductance of a 80-μm NFET in common-source configuration, the transconductance decreases with increasing temperature across frequencies. These trends lead to different effects at DC and the peaking frequency (75 GHz).

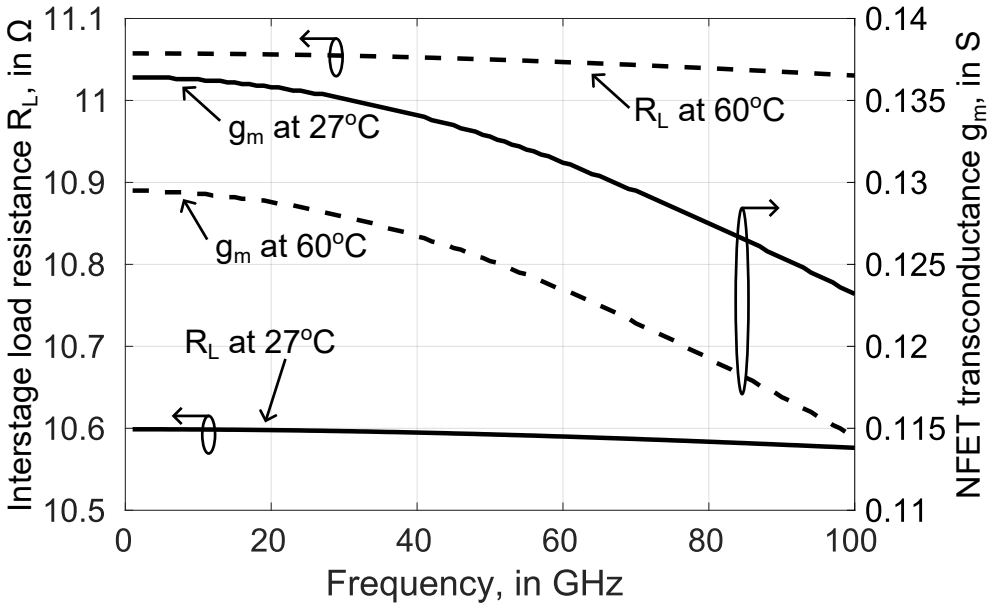


Figure 6.8: Simulated interstage load resistance and transconductance of a 80-μm NFET at 27°C and 60°C

At DC, the decrease in transconductance due to increasing temperature (i.e., driver gain) is compensated by the increase in interstage load resistance. Therefore, the DC

gain does not vary much (see Fig. 6.9) compared to the gain at high frequencies. At the peaking frequency (75 GHz), the increase in resistance due to increasing temperature causes the interstage filter to be damped (see Eq. 4.21), lowering the peak gain at 75 GHz. Moreover, the transconductance decreases across all frequencies as temperature increases, which further lowers the gain at 75 GHz. Therefore, the peak gain reduces significantly compared to the gain at DC with increasing temperature (e.g., measurements show that the peak and DC gains reduce by 6 dB and 1 dB, respectively).

Since the driver is designed for MZMs with a differential characteristic impedance of 60 Ω and the measured differential S-parameters (i.e., with 100-Ω source and load impedances) are impedance-transformed to S-parameters with source and load impedances of 100 Ω and 60 Ω, respectively, as shown in Fig. 6.9. The impedance transformation is detailed in [71], which is cumbersome for direct calculation. Therefore, Keysight ADS is utilized to perform the conversion. The conversion is done by simulating the measured 4-port S-parameter file (in Touchstone format) with a 100-Ω differential source impedance and 60-Ω differential load impedance. The resulting S-parameters are as accurate as the measured results with 100-Ω loads, as the conversion is a mathematical impedance transformation process.

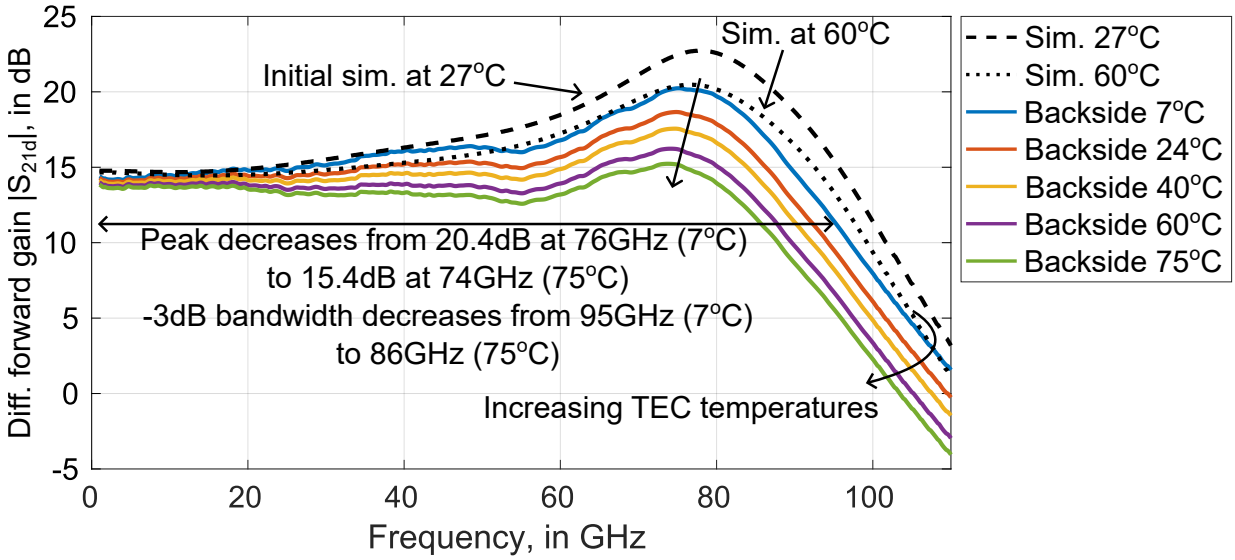


Figure 6.9: Measured small-signal differential forward gain (S_{21d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for temperature settings of 27°C and 60°C in Spectre are also plotted for comparison

For a 60-Ω differential load, the gain after impedance transformation at 1 GHz is 15 dB and there is no in-band dip in gain as seen in the case of 100-Ω load (Fig. 6.7). When the TEC temperature increases from 7°C to 75°C, the peak gain drops from 20.4 dB at 76 GHz (5.4-dB peaking at backside temperature of 7°C) to 15.4 dB at 74 GHz (0.4-dB peaking at backside temperature of 75°C). The -3dB bandwidth decreases from 95 GHz to 86 GHz as the backside temperature increases.

As shown in Figs. 6.7 and 6.9, the blue to green curves show the differential $|S_{21d}|$ for TEC temperature settings of 7°C to 75°C. The dashed and dotted black curves are simulated results at simulated temperatures of 27°C and 60°C, respectively. The simulation temperature of 27°C represents the default room temperature in Spectre. Simulating at 60°C shows a best fit to the measured result at a backside temperature of 7°C (detailed later), which indicates that the output stage is operating at a much higher temperature than predicted using the PDK at 27°C.

Recall that from the previous measurement of input termination resistance (Section 6.2), a ΔT of 20°C accounts for the temperature difference between the temperature of the chip backside and at the driver's input. The gain measured at a TEC temperature of 7°C (i.e., blue curve) is compared with the simulated result at 27°C (i.e., solid blue curve). The largest discrepancy between the simulated and measured results is at the peak gain. The measured peak gain is 20.4 dB at 76 GHz at a TEC temperature of 7°C, which is 2.3 dB lower than the simulated result at 27°C (22.7 dB gain at 78 GHz).

The discrepancy may be due to the 300-μm distance between the on-chip termination at the input and postamp stage, as the postamp consuming 318 mW of DC power is the main on-chip heat source (i.e., the driver's output drain is expected to be hotter than the passive input termination resistor). A ΔT of 20°C only considers the temperature difference between the chip backside and the driver's input. It does not account for the temperature gradient between the driver's input and output. Therefore, a simulation temperature of 27°C (i.e., corresponds to a ΔT of 20°C) does not reflect the performance degradation due to temperature at the drain (driver's output). The simulation temperature is increased to match the blue curve to estimate the realistic chip temperature. At a circuit temperature of 60°C (i.e., dashed black curve), the simulated result matches the measured result with a backside temperature of 7°C.

By comparing simulated results at 27°C and 60°C to the measured result at a backside temperature of 7°C, it can be concluded that the chip temperature impairs the performance, especially the peak gain at high frequencies for this FD-SOI driver prototype. There is a known ΔT of 16°C to 30°C between the chip backside and the driver's input (Section 6.2) and an unknown temperature delta between the driver's input and output (estimated to

be 33°C). Future study is needed to verify the temperature on-chip at the output and compensate for the unwanted thermal effects, which is discussed in a later chapter in detail.

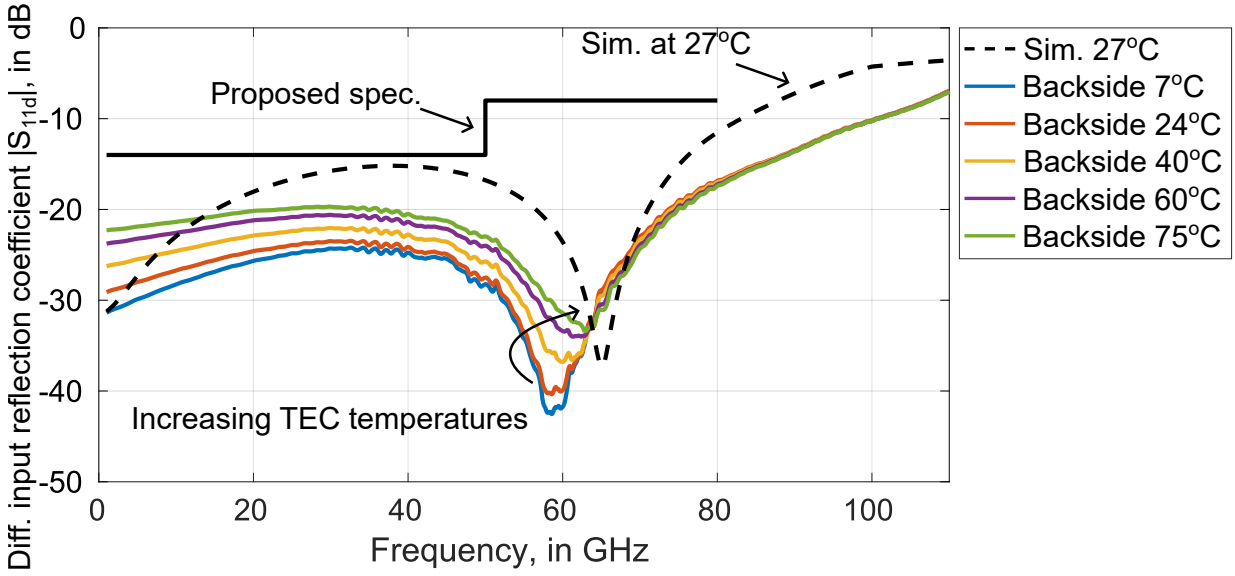


Figure 6.10: Measured small-signal differential input reflection coefficient (S_{11d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 27°C in Spectre is also plotted for comparison

The measured differential input reflection coefficient ($|S_{11d}|$) is depicted in Fig. 6.10. The blue to green curves represent the measured results with backside temperatures of 7°C to 75°C. The proposed specification for input matching is shown in solid black. The dashed black curve is the simulated result at 27°C. Note that 27°C is used because $|S_{11d}|$ is determined by the on-chip termination at the input, which estimates a ΔT of 20°C as explained previously.

The input matching shows a minor temperature dependency at frequencies below 60 GHz, which is dominated by the change in the input termination resistance over temperature. Overall, the measured results are well below the proposed mask for input matching up to 80 GHz. The $|S_{11d}|$ is kept below -10 dB up to 100 GHz (i.e., broadband input matching is achieved). At 1 GHz, the measured result at TEC temperature setting of 7°C and simulated result at 27°C match (i.e., due to a ΔT of 20°C is added to the simulation temperature). The slight difference in frequency of the in-band $|S_{11d}|$ dip (i.e., at 60 GHz

for the measured results and 65 GHz) could be due to parasitics of the input pad and feedline. As the measured results are well below the proposed specification over temperatures, the minor shift is not critical for broadband input match.

The measured stability factor, μ for a differential 100-Ω load is plotted in Fig. 6.11. The stability factor is greater than 1 at all frequencies, indicating that the driver is unconditionally stable. At low frequencies, the open-drain postamp has a high output impedance (e.g., $Z_o > 1k\Omega$ at 1 GHz), giving $|S_{22d}|$ close to 0 dB, which causes the stability factor to approach 1 (i.e., potential instability). Therefore, the real part of the output impedance of the cascode postamp has to be carefully designed to avoid turning negative due to C_{ds} of the common-source FETs, as explained in Section 4.1.2.4. Note that there is a peak at 95 GHz seen in both measured and simulated μ. The peak corresponds to the dip in $|S_{22d}|$ at 95 GHz as shown in Fig. D.7 (i.e., better output matching improves driver's stability). The simulated μ has a higher peak compared to the measured result, which may reflect to the difference in $|S_{12d}|$ (i.e., better isolation improves stability). At the $|S_{21d}|$ peak frequency of 75 GHz, the stability factor degrades slightly as the chip backside temperature increases.

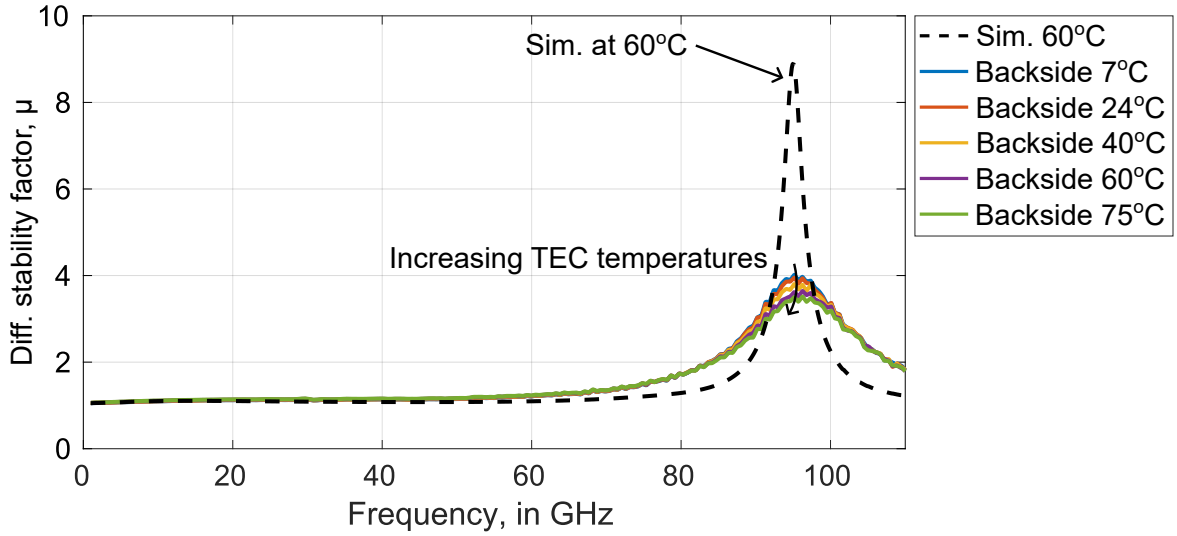


Figure 6.11: Measured small-signal differential stability factor, μ versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison

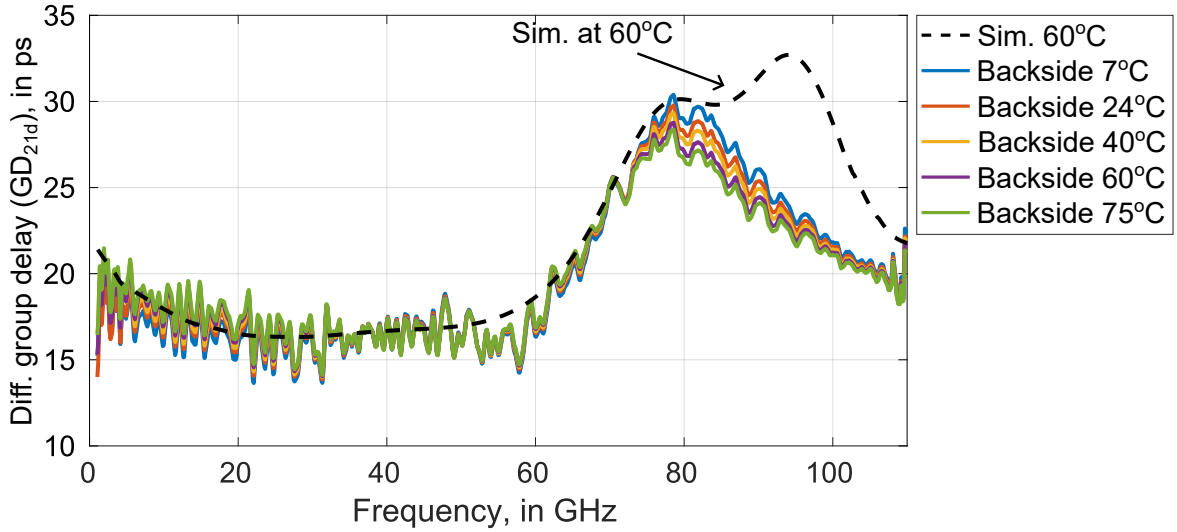


Figure 6.12: Measured small-signal differential group delay (GD_{21d}) versus various TEC temperature settings for 100-Ω source and 100-Ω load (differential) impedances. The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison

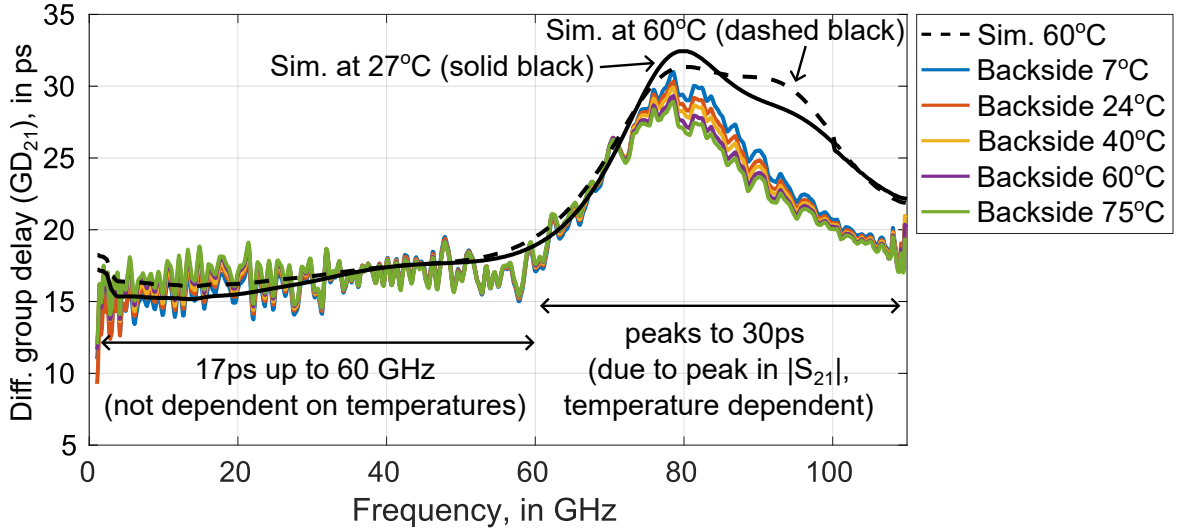


Figure 6.13: Measured small-signal differential group delay (GD_{21d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load (differential) impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison

The measured differential group delay (GD_{21d}) for 100-Ω source and load impedances is plotted in Fig. 6.12. The group delay is 18 ps with a variation of 2 ps up to 60 GHz, and the peak is 30 ps at 77 GHz. The peak in GD_{21d} is expected and corresponds to the peak in $|S_{21d}|$ seen in Fig. 6.7. Compared to the simulated result at a TEC temperature setting of 60°C, the second hump near 90 GHz matches the magnitude response, which is not seen in the measured results. This could be due to parasitic capacitances at the output pads caused by metal filling, resulting in a steeper roll-off in the frequency response. The non-constant group delay is expected to degrade the eye opening when the input data has significant power at frequencies beyond 60 GHz.

The measured group delay for a 100-Ω differential load is used to predict the delay for a 60-Ω differential load, and is depicted in Fig. 6.13. The group delay is 17 ps and it is flat up to 60 GHz. It then peaks at 76 GHz (i.e., also corresponds to the $|S_{21d}|$ peak). As the TEC temperature increases, group delays for 100-Ω and 60-Ω loads decrease, which is consistent with the trend observed in $|S_{21d}|$ against increasing temperatures. The discrepancy beyond 80 GHz matches with the difference in roll-off of the magnitude response between the simulated and measured results. When the driver is loaded by a 60-Ω MZM, the group delay response anticipates that the eye opening is maintained when the input data spectrum is less than 60 GHz, and equalization may be required if the overall electro-optical interface (i.e., driver and MZM in cascade) has non-constant magnitude response or group delay.

6.4 Large-signal measurement results

Linearity of the driver prototype is characterized using two metrics: output -1dB compression point OP_{-1dB} and total harmonic distortion (THD). Note that these two metrics are related as OP_{-1dB} is determined by amplitudes at the fundament and the third harmonic, and THD is determined by the ratio between the sum of harmonics and fundament. The compression point is measured in the frequency domain using a vector network analyzer by sweeping input power levels (the measurement setup in shown in Fig. 6.2). It is also verified in the time domain by sweeping the input amplitude for a single tone from an arbitrary waveform generator (AWG).

The measured differential OP_{-1dB} is plotted in Fig. 6.14. The blue, yellow and green curves are OP_{-1dB} at a TEC temperature setting of 7°C, with TEC turned off (i.e., chip at room temperature of 24°C plus self-heating of the driver), and 75°C, respectively. At 1 GHz, the OP_{-1dB} is 16 dBm for all temperature settings. The linearity shows minor temperature dependency up to 25 GHz, where the OP_{-1dB} gradually decreases to 14.3 dBm.

Beyond 25 GHz, the driver's linearity degrades with increasing temperature, and the degradation becomes more severe as frequency increases. At 75 GHz, there is a 1.7 dB difference in measured OP_{-1dB} between TEC temperatures of 7°C and 75°C. Compared to the simulated result with a differential 100-Ω load (dashed black), the simulated and measured results differ by less than 1 dB across all frequencies when the simulation temperature is set to 60°C. The circuit is designed to drive a differential 60-Ω MZM, so this simulated result is also plotted (dashed black curve with square markers). Note that the test instruments have a differential input impedance of 100 Ω, so the 60-Ω result could only be simulated. For large-signal measurements, the driver's non-linear behaviour (i.e., distortion) changes with input power level and load impedance, therefore the 100-Ω and 60-Ω load results are not related by a linear scale factor.

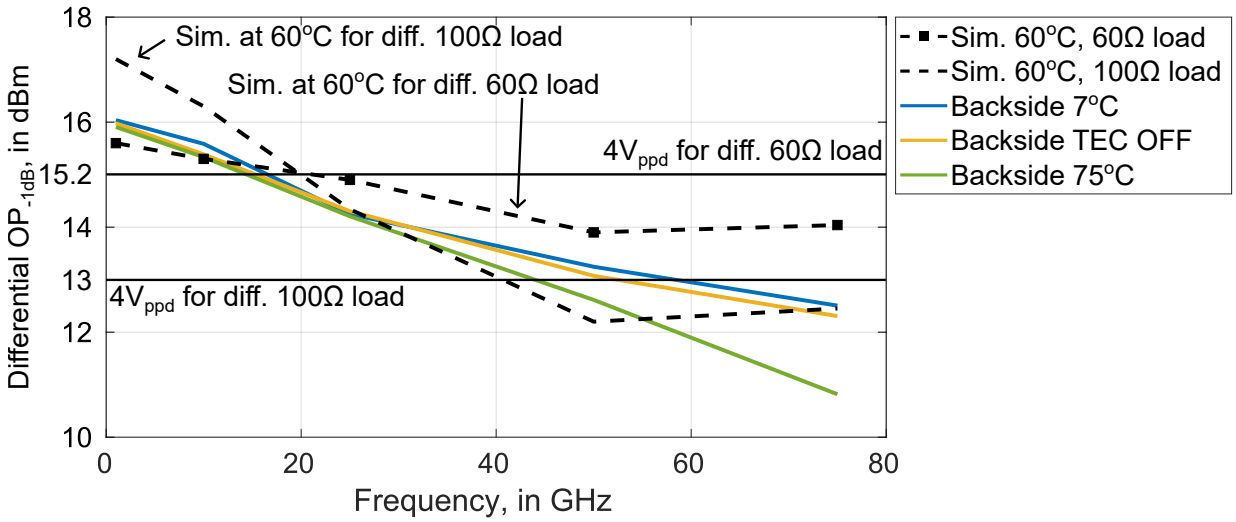


Figure 6.14: Measured large-signal differential output -1dB compression point (OP_{-1dB}) versus various TEC temperature settings for a 100-Ω differential load

The target linear output swing of the driver is 4 $V_{pp,d}$, which may be specified as a OP_{-1dB} of 13 dBm for a differential 100-Ω load, or 15.2 dBm for a differential 60-Ω load. Beyond 50 GHz, the measured OP_{-1dB} falls below 13 dBm when the TEC is off (room temperature of 24°C plus chip self-heating). Therefore, the backgate voltage may be increased (i.e., increase postamp biasing current) to tune the linearity at the cost DC power consumption.

The measured differential OP_{-1dB} at a TEC temperature of 7°C against various body-to-source bias voltages (V_{BS}) for postamp FETs is shown in Fig. 6.15. The baseline case is

taken at a V_{BS} of 0.6 V. The baseline OP_{-1dB} is 12.6 dBm for a postamp DC bias current of 102 mA. When V_{BS} increases to 1 V, OP_{-1dB} is improved by 2 dB to 14.7 dBm at the cost of a 13% increase in the postamp biasing current (i.e., 115 mA). At a V_{BS} of 1.5 V, OP_{-1dB} of 15.3 dBm is achieved while consuming 133 mA DC current for the postamp (i.e., 30% more compared to the baseline case). Increasing the backgate voltage further does not make the driver more linear, as a V_{BS} of 1.5 V (133 mA) biases the postamp near the peak f_T. Further increases in the current density result in a lower f_T, reducing OP_{-1dB} of the postamp. Therefore, for large-signal measurements of THD, a V_{BS} of 1 V is used. The total driver DC power consumption increases from 388 mW to 427 mW at a chip backside temperature of 7°C. The OP_{-1dB} for 60-Ω loads is expected to be better than the measured data with 100-Ω loads due to a better match to the optimum impedance at the output to maximize the RF output power.

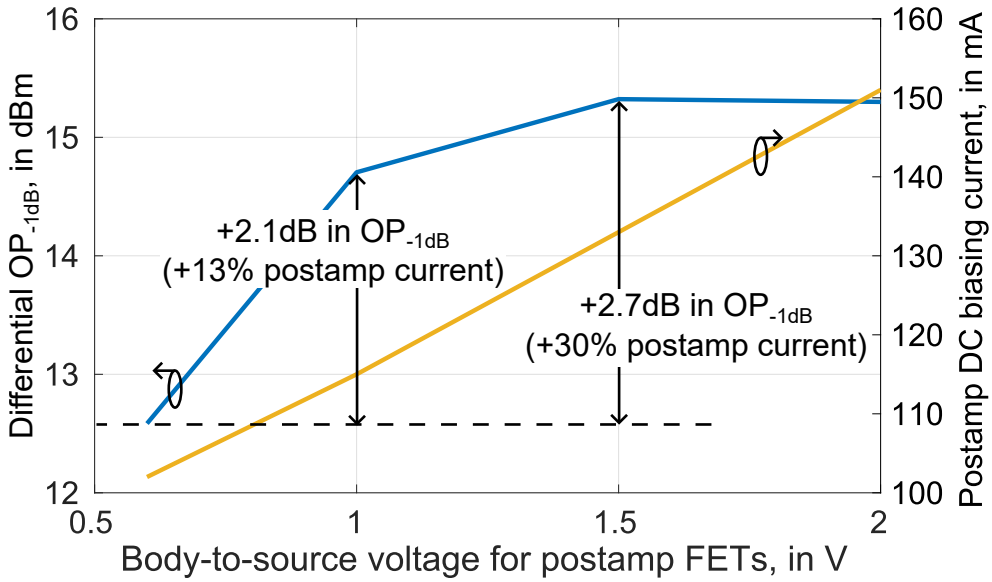


Figure 6.15: Measured large-signal differential output -1dB compression point (OP_{-1dB}) at 75 GHz for a 100-Ω differential load versus various body-to-source voltages for postamp FETs at a TEC temperature setting of 7°C

The differential OP_{-1dB} is verified using a one-tone measurement in the time domain. The same setup is used to measure THD (see Fig. 6.3). Frequencies of the input sinusoidal test tones are 1 GHz, 10 GHz and 25 GHz. The measured results are plotted in Figs. 6.16, E.9 and E.10, respectively. At each frequency, three TEC temperature settings are used: 7°C, TEC turned off (i.e., room temperature of 24°C plus chip self-heating) and 60°C.

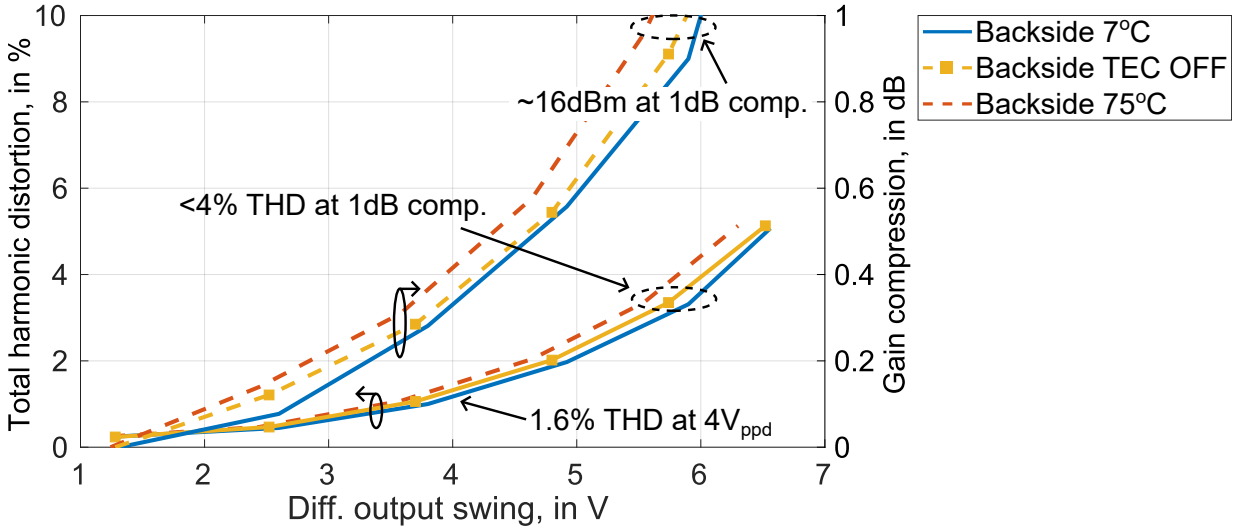


Figure 6.16: Measured gain compression and THD at the driver output, for a 1-GHz sinusoidal input signal and differential 100-Ω load at three TEC temperature settings

For an input signal at 1 GHz and an output swing of 4 $V_{pp,d}$, the gain compression is less than 0.4 dB and the THD is 1.6% for all three temperatures, as seen in Fig. 6.16. The gain is compressed by 1 dB when the output swing is above 5.6 V, matching the measured OP_{-1dB} of 16 dBm from the frequency-domain test (i.e., see Fig. 6.14 at 1 GHz). The THD at OP_{-1dB} is less than 4%. Increasing the chip backside temperature slightly degrades both OP_{-1dB} and THD, matching the trend seen in Fig. 6.14 (i.e., linearity degradation due to temperature is minor at low frequencies).

The measured differential OP_{-1dB} for a 100-Ω differential load from 1 GHz to 75 GHz at a chip backside temperature of 7°C and postamp bias current of 151 mA is plotted in Fig. 6.17. The OP_{-1dB} exceeds 13 dBm (i.e., 4-$V_{pp,d}$ linear output swing) up to 75 GHz. Note that the driver is capable of delivering 4-$V_{pp,d}$ linear output swing up to 50 GHz when the postamp is biased at 102 mA (see Fig. 6.14). Consequently, the backgate voltage of the postamp could be fine-tuned according to the input signal bandwidth for enhanced energy efficiency during serial data transmissions.

In summary, the measured linearity of the driver agrees well with simulations when the simulation temperature is set to 60°C, as previously shown for the small-signal $|S_{21}|$ measurements. Increasing the chip backside temperature lowers the driver's linearity, and the degradation in performance becomes severe at frequencies above 25 GHz. As a result, the driver could not maintain a 4-$V_{pp,d}$ linear output swing beyond 50 GHz when the TEC

is off (i.e., room temperature plus chip self-heating) and a body-to-source voltage (V_{BS}) of 0.6 V for the postamp FETs (i.e., driver DC power consumption of 388 mW). A 1-V V_{BS} attains a 2.1-dB increase in OP_{-1dB} at high frequencies, resulting in a linear output swing above 4 $V_{pp,d}$. However, the cost is 30% more DC power consumption for the postamp. Overall, the total DC power consumption of the driver is 427 mW after the adjustment for V_{BS}.

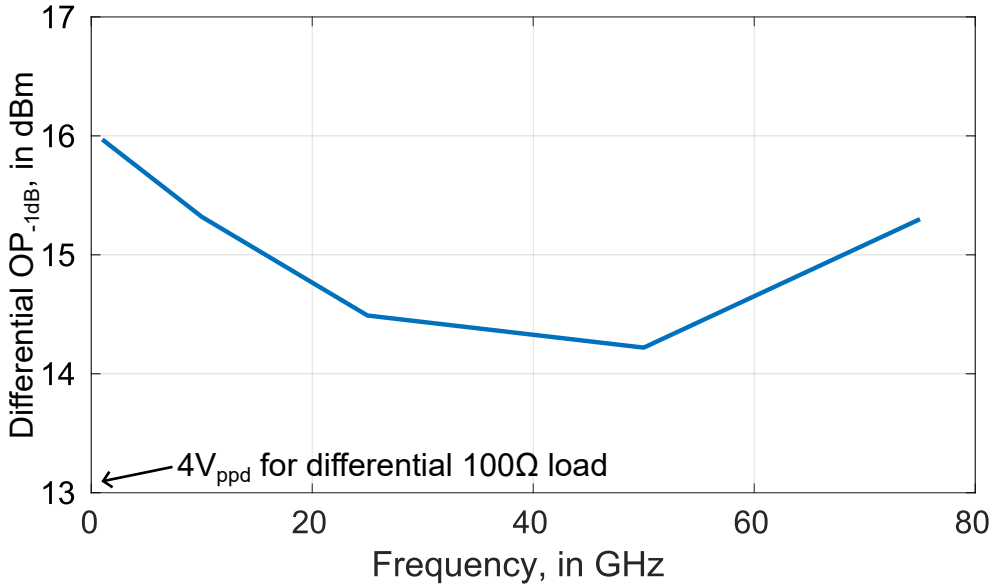
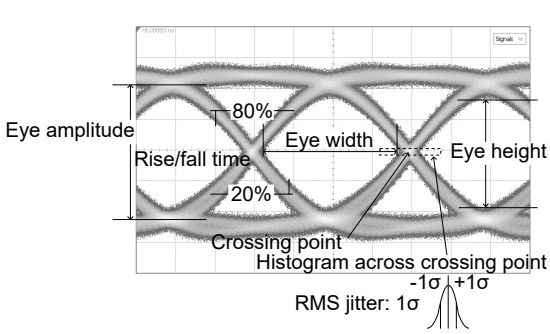


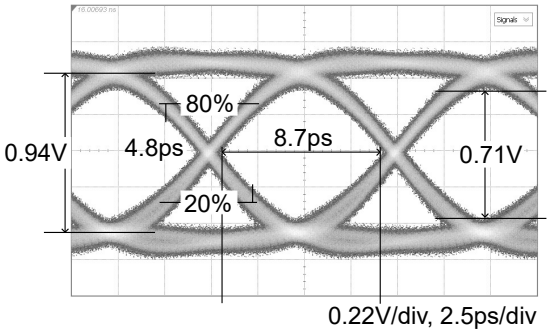
Figure 6.17: Measured large-signal differential output -1dB compression point (OP_{-1dB}) versus frequency for a 100-Ω differential load at a chip backside temperature of 7°C (V_{BS} = 2 V, postamp bias current of 151 mA)

6.5 Time-domain measurement results

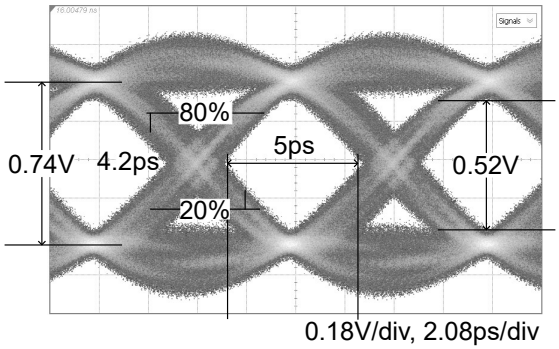
After verifying the small-signal frequency response and large-signal linearity of the driver, time-domain measurements with serial input data are performed. The time-domain setup is shown in Fig. 6.3. A 256-GS/s, 75-GHz analog bandwidth arbitrary-waveform generator (AWG) generates pseudo random binary sequence that is 15 bits in length (PRBS-15) at an input symbol rate above 100 GBaud. The loss and dispersion of the AWG and input network (e.g., the input bias-Ts, adapters, cables and probes) are measured and corrected using Keysight IQTools pre-emphasize the input signal. The resulting signal at the driver's input (i.e., after passing through the input network) is not distorted.



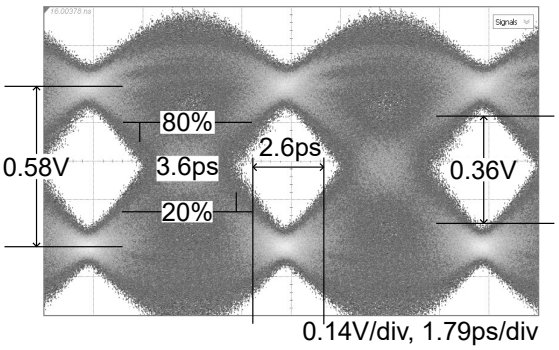
(a) Evaluation metrics for NRZ eye patterns



(b) 100 GBaud (α = 0.8, RMS jitter: 222 fs, SNR: 12.2)



(c) 120 GBaud (α = 0.5, RMS jitter: 560 fs, SNR: 10)



(d) 140 GBaud (α = 0.28, RMS jitter: 766 fs, SNR: 7.7))

Figure 6.18: Measured single-ended 100-GBaud, 120-GBaud, and 140-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.8, 0.5, and 0.28, respectively. A thru replaces the DUT connected between the AWG (including the input network) and sampling oscilloscope (including the output network) for these measurements

At the receive side, 100-GHz-bandwidth sampling heads are used to capture the differential time-domain output waveform and generate eye patterns on a sampling scope. The output channel includes the output probes, cables, bias-Ts and 6-dB attenuators. The frequency response of the output network is corrected at the sampling scope by de-embedding measured S-parameter of the output network to correctly characterize the waveform at the driver's output (i.e., drain of the postamp). Consequently, the output eye quality should reflect the frequency response and linearity of the driver for large-signal inputs (i.e., input and output network effects are corrected by the AWG and sampling scope, respectively). Note that the 100-GHz sampling heads have a maximum differential tracking swing of

1.6 $V_{pp,d}$. With a 6-dB attenuator in each lane as shown in Fig. 6.3, the maximum differential tracking range is 3.2 $V_{pp,d}$. To view a 4-$V_{pp,d}$ output swing, two 6-dB attenuators are cascaded (i.e., 12-dB attenuation in one lane), and the other output is terminated by a 50-Ω load. Therefore, a 2-V_{pp} single-ended eye pattern is measured that is equivalent to 4 $V_{pp,d}$.

Non-return-to-zero (NRZ) and 4-level pulse-amplitude-modulation (PAM-4) signals are measured at various symbol rates. Since the driver prototype has a small-signal bandwidth greater than 90 GHz, a raised-cosine filter is included during signal generation (i.e., part of the AWG test signal generation function) to shape the input signal spectrum to lower than 90 GHz by adjusting the raised-cosine roll-off factor (α) [11]. The raised-cosine shaped bandwidth (where spectrum goes to zero) is given by [21]

$$BW = 0.5 \times BR \times (1 + \alpha), \tag{6.1}$$

where BR and BW are the symbol rate (in Baud) and pulse-shaped spectrum bandwidth (in Hz) of the input signal, respectively. That is, for input signals with symbol rates of 100 GBaud, 120 GBaud and 140 GBaud, the required raised-cosine roll-off factors to set BW to 90 GHz are 0.8, 0.5 and 0.28, respectively. Decreasing α reduces the analog bandwidth required to transmit the same data stream, at the cost of amplitude peaking when the signal transitions between logic levels (i.e., overshoot/undershoot) in the eye pattern. As overshoot/undershoot takes voltage headroom, a higher input level is required to attain the same eye amplitude as α decreases.

A differential thru interconnection [72] (i.e., delay of 2.64 ps) on the calibration standard substrate is used to measure baseline eye patterns. Metrics used to express NRZ eye quality are shown in Fig. 6.18a, including eye amplitude, eye height, eye width, 20%-80% rise/fall time, jitter, and signal-to-noise ratio (SNR) [73]. Note that there is a key difference between eye amplitude and eye height for NRZ data. The eye amplitude is the difference between the one and zero levels, and the calculation uses the mean values of the histograms for level one and level zero [73]. The definition of eye height is derived by computing the difference between the inner 3σ points on the inside of the histograms for level one and level zero [73]. Therefore, the eye height should always be less than, or equal to, the eye amplitude for a given eye pattern. Sufficient eye amplitude is important to drive the optical modulator (MZM) on the transmitter side, and a receiver with sufficient dynamic range will ultimately determine whether the data bit is a '0' or a '1' based on the eye amplitude. SNR is defined as a ratio of the desired signal level to the level of background noise, plus any distortion. Furthermore, any data bits scattered beyond the 3σ points of the level-one and level-zero histograms will result in a possible bit error. Therefore, the bit error rate (BER) depends on the SNR [73].

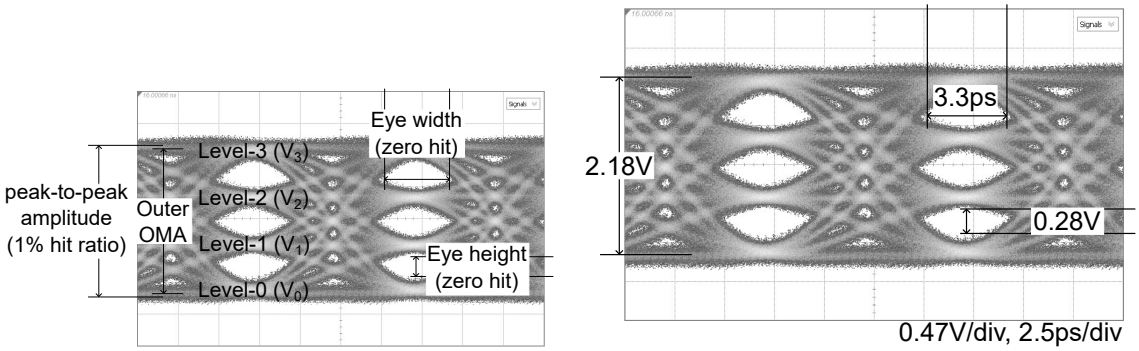
The single-ended eye patterns for 100-GBaud, 120-GBaud, and 140-GBaud PRBS-15 NRZ signals are shown in Figs. 6.18b to 6.18d, respectively. The eye amplitudes of the measured NRZ reference eye patterns (i.e., thru interconnection) are 0.94 V, 0.74 V and 0.58 V for symbol rates of 100 GBaud, 120 GBaud, and 140 GBaud, respectively. The measured eye heights are 0.71 V, 0.52 V and 0.36 V, and the measured eye widths are 8.7 ps, 5 ps and 2.6 ps for 100-GBaud, 120-GBaud, and 140-GBaud NRZ signals, respectively. Increasing the symbol rate decreases the eye height and width, resulting in a decreasing SNR (e.g., SNRs are 12.2, 10 and 7.7 for 100-GBaud, 120-GBaud, and 140-GBaud signals, respectively). Besides amplitude distortion (e.g., SNR), jitter represents random and deterministic time deviations from the ideal timing of a data steam. The time histogram of the rising and falling edges of an eye pattern at the crossing point (see Fig. 6.18b) is analyzed to determine the jitter. The root-mean-square (RMS) jitter is defined as one standard deviation of the histogram. The measured RMS jitter for a thru interconnection is 222 fs, 560 fs and 766 fs for 100 GBaud, 120 GBaud, and 140 GBaud NRZ signals, respectively.

Metrics used to express PAM-4 eye quality are shown in Fig. 6.19a, including outer optical modulation amplitude, peak-to-peak amplitude, eye height, eye width, and ratio level mismatch. For PAM-4 eye patterns, the outer optical modulation amplitude (or outer OMA) is defined similarly to the eye amplitude for NRZ eye patterns. That is, the outer OMA is the difference between the three (V_3) and zero (V_0) levels [74].

In serial data transmission, the peak amplitude of the signal is represented in a histogram [73]. The zero-hit eye opening is defined as the opening area in which there are no database hits [75]. In the IEEE Ethernet standard [74] (IEEE 802.3cu), the peak-to-peak amplitude is taken at a hit ratio of 1% to avoid misleading results due to signal and oscilloscope noise for overshoot and undershoot [74]. The peak-to-peak amplitude is typically greater than, or equal to, the outer OMA, as it includes overshoots and undershoots of the transmitted signal. Lastly, the linearity of a PAM-4 eye is expressed using the ratio level mismatch (RLM) in [74] (IEEE 802.3 Clause 94). It is defined as the ratio of the smallest separation between the adjacent levels and the ideal separation, as shown in Eq. 6.2,

$$RLM = \frac{3 \times \min(V_3 - V_2, V_2 - V_1, V_1 - V_0)}{V_3 - V_0}, \tag{6.2}$$

where V_0 to V_3 are the voltage for level-0 to level-3, respectively, and RLM of 1 represents an ideal linearity for PAM-4 eyes (i.e., separations between all four levels are the same).



(a) Evaluation metrics for PAM-4 eye patterns

(b) 100 GBaud ($\alpha = 1$, RLM: 0.97)

Figure 6.19: Measured differential 100-GBaud PAM-4 eye pattern with raised-cosine roll-off factors (α) of 1 for a thru interconnection as the DUT connected between the AWG (including the input channel) and sampling oscilloscope (including the output channel)

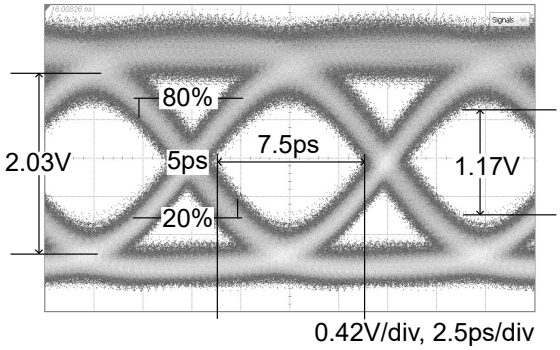
The differential eye pattern for the 100-GBaud PRBS-15 PAM-4 signal is shown in Fig. 6.19b, respectively. The AWG signal amplitude is set to 2.7 $V_{p,d}$ to attain the max available swing and the data length is 32800 to ensure a full PRBS-15 pattern is covered. The raised-cosine roll-off factor is set to 1. The worst-case eye height and eye width are labelled. That is, the minimum eye height and eye width 0.28 V and 3.3 ps for 100-GBaud PAM-4 eyes, respectively. The outer OMA (i.e., similar to the eye amplitude for NRZ eye patterns) is 2.2 V, whereas the peak-to-peak amplitude is 2.24 $V_{pp,d}$. The RLM is 0.97 (lower than 1), which may be due to the nonlinearity of the signal generator and nonlinearity of the sampling heads.

6.5.1 Test pattern: PRBS-15 100-/120-/140-GBaud NRZ

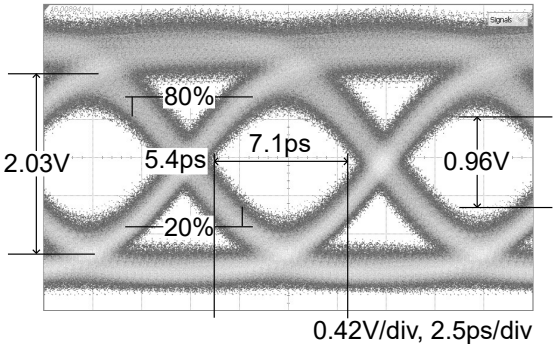
The driver's small-signal frequency response in Fig. 6.7 shows that the peak gain and flatness of the in-band gain depend on the chip temperature. Therefore, three TEC temperature settings: 7°C, TEC turned off, and 60°C, are used to quantify the effects of temperature on the measured eye.

At a symbol rate of 100 GBaud (raised-cosine roll-off factor set to 0.8 for a spectrum bandwidth of 90 GHz), single-ended NRZ eye patterns of the driver at various TEC temperature settings are shown in Fig. 6.20. At all three temperatures, the output swing is set to 2 V_{pp} (i.e., equivalent to 4 $V_{pp,d}$). The eye heights of the measured NRZ signals are 1.17 V_{pp}, 0.96 V_{pp}, and 0.74 V_{pp} for TEC temperatures of 7°C, off (i.e., room temperature

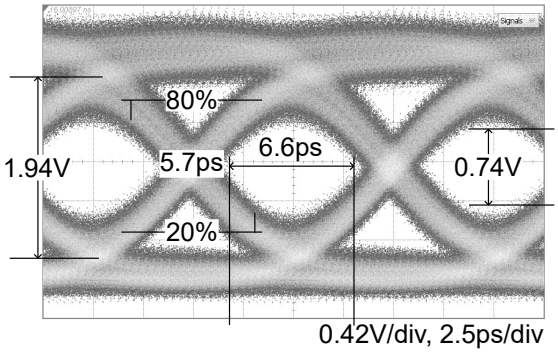
of 24°C plus chip self-heating), and 60°C, respectively. The eye widths have the same trend as the eye heights, which decrease from 7.5 ps to 6.6 ps as temperature increases. The minimum rise/fall time of 5 ps occurs when TEC is set to 7°C (n.b., measured rise/fall time is 4.8 ps using a thru at 100 GBaud NRZ). The RMS jitter increases from 420 fs to 567 fs as temperature increases to 60°C (222 fs for the thru interconnection). Lastly, the SNR degrades from 5.9 at 7°C to 4.32 at 60°C, compared to an SNR of 12.2 for the thru interconnection at 100 GBaud NRZ. Therefore, the quality of the NRZ signal transmission is reduced as the chip backside temperature increases.



(a) TEC 7°C (RMS jitter: 420 fs, SNR: 7.1)



(b) TEC off (RMS jitter: 485 fs, SNR: 5.7)

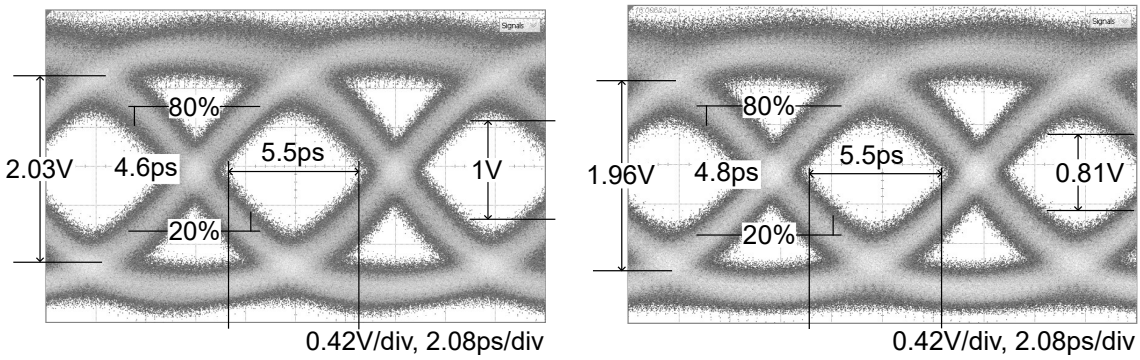


(c) TEC 60°C (RMS jitter: 567 fs, SNR: 4.84)

Figure 6.20: Measured single-ended 100-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.8 for the driver for three chip backside temperatures

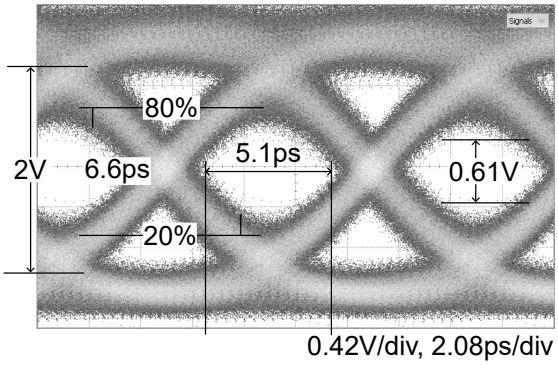
The single-ended NRZ eye patterns at 120 GBaud and 140 GBaud (raised-cosine roll-off factors set to 0.5 and 0.28, respectively, for 90-GHz bandwidth) are depicted in Figs. 6.21 and 6.22, respectively. In both cases, equivalent differential eye amplitudes of 4 $V_{pp,d}$ are

achieved. At the same TEC temperature setting (e.g., at 7°C), the eye height is largest at 100 GBaud (1.17 V_{pp}) compared to 0.83 V_{pp} at 60°C. Although the bandwidth is set to 90 GHz for all three symbol rates (100 GBaud to 140 GBaud), the 140-GBaud NRZ signal with a raised-cosine roll-off factor of 0.28 is more sensitive to gain and linearity at high frequencies, as its spectrum contains more content near the cutoff, compared to 100 GBaud and 120 GBaud. Therefore, the 140-GBaud NRZ eye height (Fig. 6.22a) is lower than the 100-GBaud case (Fig. 6.20a).



(a) TEC 7°C (RMS jitter: 472 fs, SNR: 5.9)

(b) TEC off (RMS jitter: 481 fs, SNR: 5.1)

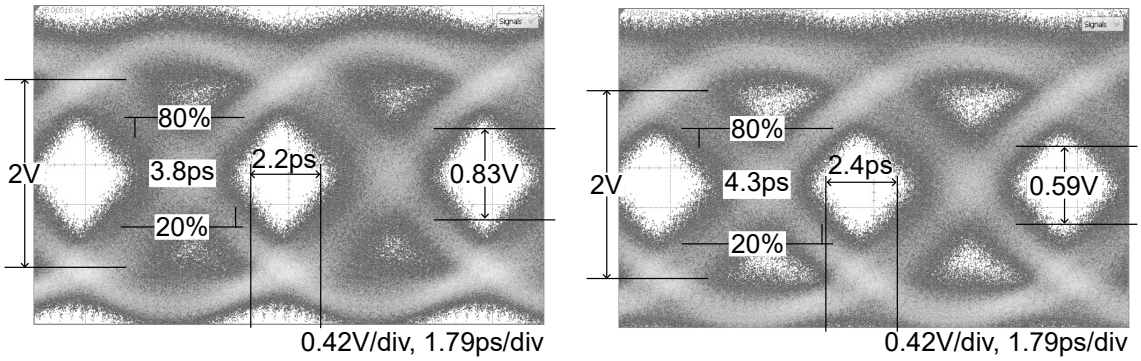


(c) TEC 60°C (RMS jitter: 532 fs, SNR: 4.32)

Figure 6.21: Measured single-ended 120-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.5 for the driver at three chip backside temperatures

The trend of eye quality degradation against increasing temperature seen in the 100-GBaud measurement occurs at 120 GBaud and 140 GBaud as well, expect for the RMS jitter and eye width for 140-GBaud NRZ signals. Interestingly, the eye width increases and

the RMS jitter decreases as the TEC temperature increases, which is opposite to the 100-GBaud and 120-GBaud cases. This may be explained by the change in the group delay and high-frequency peaking of the driver with increasing temperature. At small-signal levels, the peak in the driver's group delay decreases as the temperature increases (see Fig. 6.12). Therefore, for large-signal outputs at 60°C the group delay variation is expected to be lower than at 7°C, resulting in a decrease in the RMS jitter and an increase in the eye width.



(a) TEC 7°C (RMS jitter: 820 fs, SNR: 5.2)

(b) TEC off (RMS jitter: 787 fs, SNR: 4.2)

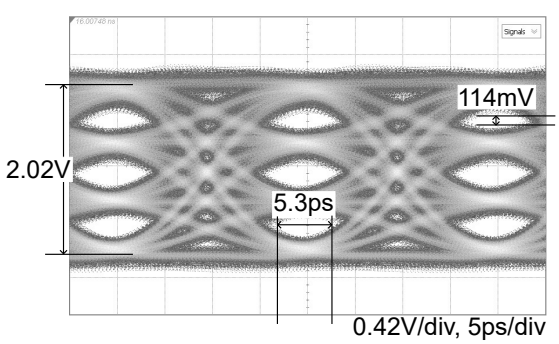
(c) TEC 60°C (RMS jitter: 778 fs, SNR: 3.5)

Figure 6.22: Measured single-ended 140-GBaud NRZ eye patterns with raised-cosine roll-off factors (α) of 0.28 for the driver at three chip backside temperatures

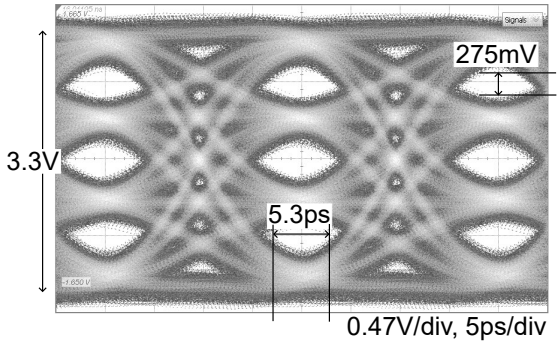
Note that when the output swing is 4 $V_{pp,d}$ for NRZ signals, the DC bias current measured at 7°C is 130 mA (instead of 115 mA at small-signal levels when V_{BS} is 1 V for postamp NFETs, as shown in Fig. 6.15). Furthermore, the off-chip thermistor reads 32°C when the TEC controller is turned off, which is higher than 26.5°C when the driver

is idle (i.e., no serial data transmission). Both observations indicate that excessive heat is generated when the driver outputs serial data at 4 $V_{pp,d}$. Therefore, the driver's frequency response (especially peak gain) is worse than the small-signal response in **??** and a margin must be included during the design phase to compensate for the performance degradation caused by heat generation.

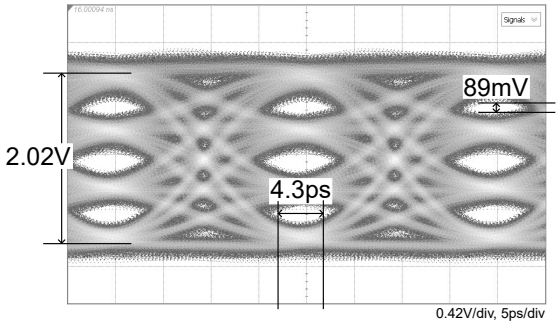
6.5.2 Test pattern: PRBS-15 50-/100-GBaud PAM-4



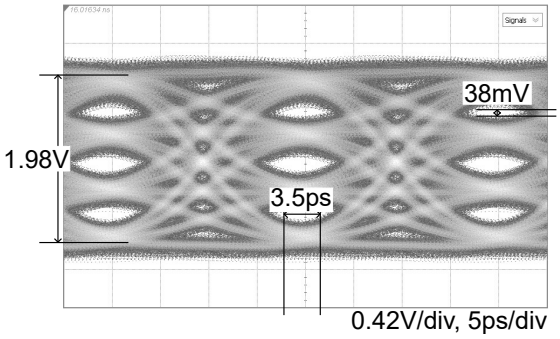
(a) TEC 7°C (single-ended, 1%-hit-ratio: 2.03 V, RLM: 0.93)



(b) TEC 7°C (differential, 1%-hit-ratio: 3.32 V, RLM: 0.97)



(c) TEC off (single-ended, 1%-hit-ratio: 2.07 V, RLM: 0.92)



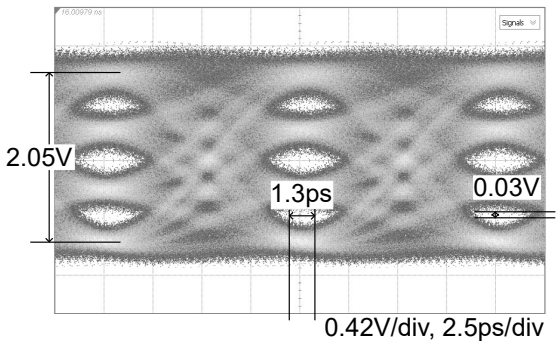
(d) TEC 60°C (single-ended, 1%-hit-ratio: 1.99 V, RLM: 0.92)

Figure 6.23: Measured single-ended and differential 50-GBaud PAM-4 eye patterns with raised-cosine roll-off factors (α) of 1 for the driver at three chip backside temperatures (worst-case eye height and width are labelled)

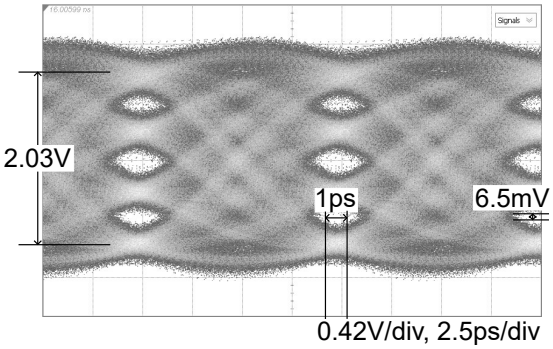
Compared to NRZ, PAM-4 modulation places more stringent requirements on the driver's linearity to achieve robust transmission. At a symbol rate of 50 GBaud (raised-cosine roll-off factor set to 1 for a bandwidth of 50 GHz), single-ended and differential PAM-4 eye patterns of the driver at three TEC temperature settings (i.e., 7°C, off, and 60°C) and two outer optical modulation amplitude (OMA) (i.e., 3.3 $V_{pp,d}$ and 4 $V_{pp,d}$ output swings) are shown in Fig. 6.23. At 50 GBaud and a chip backside temperature of 7°C, the driver's frequency response is not the bottleneck for serial data transmission (i.e., small-signal -3dB bandwidth is well above 50 GHz as seen in Fig. 6.7 and group delay has not yet started to peak, as seen in Fig. 6.12). Therefore, amplitude distortion due to nonlinearity is likely the main factor for PAM-4 eye quality degradation.

Recall that the two 6-dB attenuators from Fig. 6.3 are cascaded in Channel A to permit larger output swing and to view the effects of gain compression on PAM-4 eyes. The single-ended eye patterns (i.e., Figs. 6.23a, 6.23c and 6.23d) are measured at different temperatures, and the outer OMA is set to 2 V_{pp} (i.e., equivalent to 4 $V_{pp,d}$ output swing). The differential eye (Fig. 6.23b) is measured at a chip backside temperature of 7°C and 3.3-$V_{pp,d}$ (i.e., backed off compared to 4 $V_{pp,d}$). Note that the zero-hit eye height and eye width are measured, which show the worst-case eye properties (i.e., the zero-hit eye opening is defined as the opening area in which there are no database hits [75].).

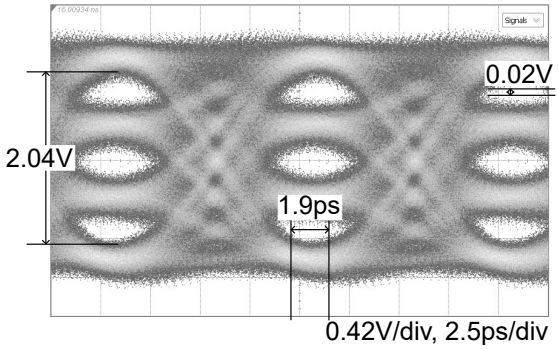
The worst-case eye heights of the measured, singled-ended PAM-4 signals are 114 mV, 89 mV, and 38 mV, respectively, for TEC temperatures of 7°C, off (i.e., room temperature of 24°C plus chip self-heating), and 60°C. A decreasing trend in the eye height is seen for increasing temperature, which is the same trend as seen for the NRZ eye patterns. In all three cases, the top and bottom PAM-4 eyes are compressed and skewed compared to the middle eye. The worst-case eye widths have the same trend as the eye heights, which decrease from 5.3 ps to 3.5 ps as temperature increases. Lastly, the ratio level mismatch (RLM, defined in IEEE 802.3 Clause 94 [74]) is measured to attain the driver's linearity. At 4 $V_{pp,d}$ equivalent output swing, the RLM for TEC temperatures of 7°C is 0.93, and reduces to 0.92 for higher temperatures. When the outer OMA is backed off to 3.3 $V_{pp,d}$ and the TEC temperature remains at 7°C (i.e., Fig. 6.23b), the RLM improves from 0.93 to 0.97 compared to Fig. 6.23a, indicating that less distortion is attained. Less compression and skew are seen in the top and bottom eyes. The worst-case eye height and width are 275 mV (up from 114 mV) and 5.3 ps, respectively (i.e., remains the same). The results show that amplitude distortion of the driver limits the eye opening at 4 $V_{pp,d}$, and phase distortion (i.e., group delay) does not affect the eye quality at 50 GBaud, because the driver's group delay is flat up to 60 GHz.



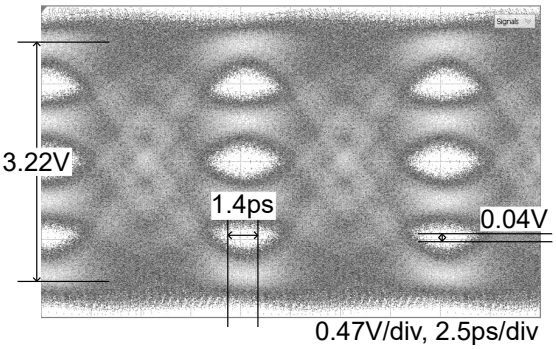
(a) TEC 7°C (single-ended, $\alpha = 1$, 1%-hit-ratio: 2.09 V, RLM: 0.93)



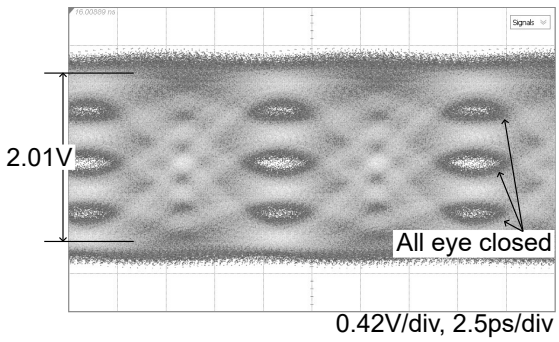
(b) TEC 7°C (single-ended, $\alpha = 0.5$, 1%-hit-ratio: 2.24 V, RLM: 0.93)



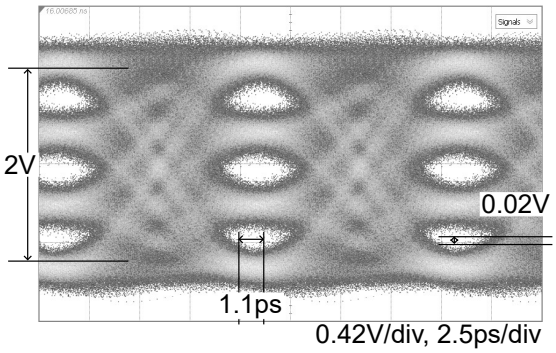
(c) TEC 7°C (single-ended, $\alpha = 1$, 4-tap FFE: on, 1%-hit-ratio: 2.37 V, RLM: 0.93)



(d) TEC 7°C (differential, $\alpha = 1$, 1%-hit-ratio: 3.27 V, RLM: 0.96)



(e) TEC off (single-ended, $\alpha = 1$, 1%-hit-ratio: 2.04 V, RLM: 0.93)



(f) TEC 60°C (single-ended, $\alpha = 1$, 4-tap FFE: on, 1%-hit-ratio: 2.42 V, RLM: 0.92)

Figure 6.24: Measured single-ended and differential 100-GBaud PAM-4 eye patterns for the driver at three chip temperatures, differential output swings, and raised-cosine roll-off factors (worst-case eye height and width are labelled)

The tests conducted at 50 GBaud demonstrate the extent to which the linearity of the driver constrains the integrity of the output PAM-4 eyes at an amplitude of 4 $V_{pp,d}$. Subsequently, the symbol rate is elevated to 100 GBaud to showcase how the driver's frequency response, particularly the non-flat group delay, impacts serial data transmission. Six PAM-4 eyes are compared in Fig. 6.24. The effect of feedforward equalization (FFE) is also shown. FFE is a digital equalization method that mitigates inter-symbol interference [76]. FFE utilizes digital-signal processing (DSP) to realize a linear transfer function which is the inverse of the system end-to-end frequency response (i.e., from AWG to the sampling heads including the driver) to achieve compensation.

The single-ended eye patterns (see Figs. 6.24a to 6.24c, 6.24e and 6.24f) are measured at different temperatures, raised-cosine roll-off factors, FFE settings, and the outer OMA is set to 2 V_{pp} (i.e., equivalent to 4 $V_{pp,d}$ output swing). The differential eye (Fig. 6.24d) is measured for a chip backside temperature of 7°C and 3.2-$V_{pp,d}$ output swing.

At a TEC temperature of 7°C, with a raised-cosine roll-off factor of 1 (spectrum of 100 GHz), and at 4 $V_{pp,d}$ (Fig. 6.24a), the top and bottom PAM-4 eyes experience compression, similar to what is observed in the 50-GBaud scenario. The middle eye remains relatively unaffected. Notably, the top eye closes, suggesting more pronounced compression at 100 GBaud. The bottom eye, being the second most distorted, measures 30 mV in height and 1.3 ps in width. The RLM (i.e., eye linearity) of the eye is 0.93, identical to that of the 50-GBaud case.

Next, the AWG is adjusted to set the raised-cosine roll-off factor (α) to 0.5, resulting in an input bandwidth of 75 GHz (as per Eq. 6.1). The eye measured at the driver output is depicted in Fig. 6.24b. Compared to Fig. 6.24a, where only α differs, there is no enhancement in eye opening, and the RLM remains at 0.93. This finding indicates that even with a reduced input bandwidth of 75 GHz, the driver faces similar eye closure problems, implying that driver bandwidth and linearity above 75 GHz are not limiting the PAM-4 eye opening. Consequently, the linearity and frequency response (especially flatness of the group delay) of the driver below 75 GHz determine the eye quality.

A 4-tap FFE is applied to the measured data from the sampling scope to compensate for the driver's frequency response, and the resulting eye is presented in Fig. 6.24c. It is important to note that the 2-dB dip in the driver's magnitude response observed between 25 GHz and 55 GHz occurs only when a 100-Ω differential load is used (see Fig. 6.7). This anomaly does not appear with a 60-Ω load, such as the intended MZMs, as illustrated in Fig. 6.9 (note that the driver is designed for 60-Ω loads). The primary role of the FFE in this case is to mitigate excessive amplitude peaking and non-flat group delay above 60 GHz. The eye opening is enhanced compared to Fig. 6.24a, which reflects the same conditions

but without FFE. In the worst case, the eye amplitude and width are 20 mV and 1.9 ps, respectively. This demonstrates that the driver's frequency response, particularly below 75 GHz, significantly influences eye quality. The RLM remains at 0.93, as anticipated, since FFE cannot address the driver's nonlinearity (e.g., top and bottom eyes remain skewed). Note that since the driver (especially the output combining network) is designed for 60-Ω MZMs, the frequency response and linearity with 60-Ω loads are expected to be better than measured results with 100-Ω loads. Therefore, less closure in the top and bottom eyes is expected.

The output swing is backed off from 4 $V_{pp,d}$ to 3.2 $V_{pp,d}$ to illustrate the change in linearity of the output eye (see Fig. 6.24d). The RLM increases to 0.96 compared to 0.93 for all the 4 $V_{pp,d}$ cases. Without FFE, the eye opening is greatly improved and the worst-case eye height and width are 40 mV and 1.4 ps, respectively. The skew of the top and bottom eyes is reduced. In conclusion, a similar improvement in output eye linearity is observed as in the 50-GBaud case, when the output swing is backed off from 4 $V_{pp,d}$ to 3.2 $V_{pp,d}$.

The TEC temperature settings are then increased from 7°C (Figs. 6.24a to 6.24d) to room temperature (i.e., TEC turned off) and 60°C, as depicted in Figs. 6.24e and 6.24f, respectively. All three eyes close compared to Fig. 6.24a, with the TEC deactivated and an output maintained at 4 $V_{pp,d}$. This occurs due to a decrease in bandwidth and a greater in-band dip in the driver's frequency response with rising temperature, as demonstrated in Fig. 6.7. When a 4-tap FFE is introduced, the output PAM-4 eyes reopen even with a chip backside temperature of 60°C, and the RLM reduces slightly to 0.92. This may be attributed to a slight decline in the driver's linearity at higher temperatures, which the FFE is unable to rectify.

6.6 Performance Summary and Comparison

The frequency-domain characteristics of the broadband driver prototype are outlined in Table 6.1 and compared to other broadband drivers from the recent literature. The driver discussed in this thesis is implemented in 22-nm FD-SOI CMOS. Other designs using bulk CMOS, bipolar/BiCMOS and InP technologies are listed in the table to compare performance across different processing technologies.

The driver prototype developed in this thsis demonstrates a small-signal low-frequency gain of 16.8 dB and a peak of 19.2 dB at 77 GHz for a 100-Ω differential load. Gain is 15 dB with a peak of 20.4 dB at 76 GHz for a differential load of 60 Ω. The chip backside (TEC)

temperature is 7°C in both cases. The -3dB bandwidths at a TEC temperature of 7°C are 91 GHz for a 100-Ω load, and 95 GHz for a 60-Ω load. Increasing the TEC temperature to 75°C results in a 5 dB decrease in peak gain and a reduction of 10 GHz in -3dB bandwidth for both load conditions. The good impedance match at the input (e.g., >10 dB return loss up to 100 GHz) ensures that the input return loss is well above the proposed mask (i.e., 14 dB up to 50 GHz and 8 dB up to 80 GHz). The group delay is consistently 18 ps below 60 GHz, and peaks at 30 ps at 76 GHz, aligning with the peak frequency of the magnitude response. The driver is unconditionally stable across all frequencies (up to 110 GHz) and temperature settings.

The preamp and postamp operate at DC bias currents of 30 mA and 115 mA, respectively. The driver consumes a total DC power of 427 mW when the TEC is set to 7°C. From measurements of large-signal compression at a DC bias current of 115 mA (i.e., corresponds to a 1-V V_{BS} for FETs in the postamp stage), the OP_{-1dB} remains above 13 dBm, (i.e., a differential output swing of 4 $V_{pp,d}$ across 100-Ω loads) up to a frequency of 75 GHz. Furthermore, adjusting V_{BS} enables a trade-off between DC power consumption and postamp linearity. The design occupies 0.17 mm^2 (core area) and a total area of 0.7 mm^2, including pads. The multiple samples tested showed excellent consistency in small-signal and large-signal performance.

Time-domain measurements report a THD of 1.6% at 1 GHz and 4-$V_{pp,d}$ output swing. In serial data tests, the driver enables transmissions of 140-GBaud NRZ signals with 4-$V_{pp,d}$ eye amplitude at the driver's output. When applying 100-GBaud PAM-4 signals, a 4-$V_{pp,d}$ eye is attained at a TEC temperature of 7°C, but the top and bottom eyes begin to close due to the frequency response and linearity of the driver with 100-Ω loads. With a 4-tap FFE, the driver's frequency response is compensated and eye reopening is observed (eye shrinkage due to compression still remains). The 100-GBaud PAM-4 eye quality is anticipated to improve with 60-Ω loads, as the frequency response and linearity are expected to be better compared to 100-Ω loads.

The 5-stage linear driver discussed in [32] achieves an analog bandwidth of 48 GHz and transmits at 64 GBaud. It offers a low-frequency gain of 13 dB, adjustable to 22.5 dB. With a maximum linear output swing of 1.5 $V_{pp,d}$ at 1 GHz, the total harmonic distortion (THD) is 2.2%. This driver is fabricated using a 65-nm CMOS technology and has a DC power consumption of 225 mW. When compared to the driver detailed in the thesis, [32] reports lower DC power usage, correlated with the restricted output swing (i.e., 225 mW for 1.5 $V_{pp,d}$). Other specifications do not demonstrate advantages over this work.

In [34], a 65-nm CMOS linear driver is detailed, offering 32-GHz bandwidth with a gain adjustable between 20 dB and 28 dB. Linearity is evaluated using THD, measuring under

5% at 1 GHz and less than 6% at 10 GHz. During transmission of 80-Gb/s PAM-4 signals, the driver delivers an output swing of 2 $V_{pp,d}$, with both its speed and output swing falling short compared to this work. The driver is noted for its compact core size of 0.043 mm^2, and a low DC power consumption of 180 mW.

Bipolar technologies in SiGe and InP are prevalent in broadband linear driver designs, as documented in [44, 49, 50, 52, 77, 78]. The driver with an analog multiplexer (i.e., bandwidth doubling by interleaving two PAM-4 signals) in [77] reports the highest bandwidth of >110 GHz (transmits 168-GBaud PAM-4 signals), but requires DSP for digital preprocessing, resulting in a total DC power consumption of 0.99 W for the analog multiplexer and driver. The highest swing of 7.2 $V_{pp,d}$ is achieved in [49] at 45 GBaud for PAM-4 signals and 72 GBaud for NRZ signals. The SiGe BiCMOS driver in [52] transmits 56-GBaud PAM-8 signals at 3.8 $V_{pp,d}$, indicating a potential for transmitting modulation schemes which are more complex than PAM-4. Lastly, the travelling-wave driver in [78] reports the lowest power efficiency of 2.83 pJ/bit among all bipolar designs while transmitting 128-GBaud PAM-4 signal at 2.8 $V_{pp,d}$.

For time-domain measurements with serial-data inputs, the driver prototype reaches a 100-GBaud symbol rate using a PAM-4 modulation scheme with an output swing of 4 $V_{pp,d}$. It is important to note that the driver's postamp consumes 130 mA with large-signal serial data inputs, up from 115 mA in the frequency-domain measurements. This results in a total DC power consumption of 472 mW (recorded in Table 6.1) compared to the previous 427 mW. The resulting energy efficiency is 2.36 pJ/b, making it the lowest in the table, except for [32, 34], which has a lower output swing of 1.5 $V_{pp,d}$ and 2 $V_{pp,d}$, respectively. The drivers in [32, 44, 77] exceed data rates of 500 Gb/s, achieved through polarization-division multiplexing modulation in the optical domain. The linear driver in [52] experiments the transmission of PAM-8 signals and 56 GBaud is achieved (168 Gb/s).

Table 6.1: Performance summary and comparison of high-swing, linear modulator drivers

Parameters	This work	[32] ISSCC 20	[34] VLSI 17	[78] TMTT 24	[49] Access 22	[50] TMTT 17	[77] JSSC 20	[44] JSSC 20	[52] TMTT 17	[79] BCICTS 23
Gain at LF[1](dB)	15[2]/ 16.8[3]	13 - 22.5	20 - 28	11.2 - 21.8	15.5	12.5	10.5	20 - 30	12.8	25
Gain peaking (dB)	5.4[2]/ 2.4[3]	5	N/A	4.3	6	N/A	7.5	N/A	4.2	19
@ Freq.	76 GHz[2,3]	41 GHz	N/A	50 GHz	35 GHz	N/A	85 GHz	N/A	47 GHz	72 GHz
-3 dB BW. (GHz)	95[2]/ 91[3]	48	32	70	61.2	90	>110	>40	62	86
Driver output ($V_{pp,d}$)	4	1.5	2	2.8	7.2	4	1.5	2.4	3.8	3
-10dB S_{DD11} BW. (GHz)	100[2,3]	50	50	60	>67	83	75	40	>67	65
Group delay (flat) (ps)	17[2]/ 18[3]	N/A	N/A	27	11	23	N/A	N/A	N/A	N/A
Group delay (peak) (ps)	30[2,3] @ 76 GHz	N/A	N/A	33	21 @ 38 GHz	70 @ 98 GHz	N/A	N/A	N/A	N/A
$OP_{-1dB,d}$ (dBm)	14.7[3]@	N/A	N/A	15.2 @	N/A	13.5 @	7[5]@	N/A	15 @	13.2 @
@ Freq.	75 GHz	N/A	N/A	1 GHz	N/A	1 GHz	1 GHz	N/A	10 GHz	50 GHz
THD @ Freq. (%)	1.6[3]@ 1 GHz	2.2 @ 1 GHz	5 @ 1 GHz	2 @ 1 GHz	1 @ 1 GHz	<5 @ 1 GHz	3 @ 1 GHz	3.6 @ 1 GHz	10 @ 10 GHz	3[5]@ 1 GHz
@ Output swing (V_{ppd})	4[3]	1.5	2	3.56	6.5	3	1.5	6	5	4
V_{DD} (V)	3	2.6	2.5	4	4.2	5.5	4.5	5.5	6	4.1
Z_{load} (diff. Ω)	100	100	100	100	100	100	100	100	100	100
Modulation	PAM-4	DP-32QAM	PAM-4	PAM-4	PAM-4	PAM-4	DP-16QAM	DP-16QAM	PAM-8	PAM-4
Baud rate (GBaud)	100	64	40	128	45	45	168	69	56	128
Data rate (Gb/s)	200	640[4]	80	256	90	90	1344[4]	552[4]	168	256
P_{DC} (mW/channel)	472	225	180	725	670	550	990	1000	820	540
Energy/bit (pJ/b)	2.36	1.41	2.25	2.83	7.44	5.78	2.95[5]	7.2	4.88	2.11
Core area (mm^2)	0.17	1.75[5,6]	0.04	1.07[5,6]	0.8[6]	1.2[6]	4[6]	0.8[5,6]	0.6[6]	2.1[5,6]
FoM[7](GBaud)	7.06[2]/ 4.24[3]	0.8	1.11	1.73	4.35	1.64	0.48	0.5	1.23	2.67
Process	22-nm FDSOI-CMOS	65-nm CMOS	65-nm CMOS	55-nm SiGe	130-nm SiGe	130-nm SiGe	250-nm InP	130-nm SiGe	55-nm SiGe	55-nm SiGe

[1] a range of low-frequency gain is reported if a VGA presents

[2] for 60-Ω differential load

[3] for 100-Ω differential load

[4] E/O transmitter response

[5] estimated

[6] including pads

[7] FoM $= \frac{\text{Baud rate [GBaud]} \cdot V_{out,ppd}^2[V^2]}{8 \cdot Z_{load,d}[\Omega] \cdot P_{DC}[W/ch]}$

Chapter 7

Summary and Conclusions

This chapter summaries key research findings, contributions, and recommendations for future work. The chapter highlights the development of amplitude-peaking filter networks, linear high-swing optical modulator drivers, and power-efficient linearization techniques. Finally, potential future work is proposed, such as the design of variable-gain stages, multi-stage drivers, and custom high-Q capacitors, to further enhance the driver's performance and adaptability for next-generation communication systems.

7.1 Conclusions

The demand for higher data rate in communication systems drives the development of advanced integrated circuits. This thesis presented the design and implementation of a broadband driver in a 22-nm FD-SOI CMOS technology, targeting over 100-GBaud per wavelength for optical communication. The driver leverages a novel shunt/double-series interstage network design methodology that achieves amplitude peaking and extends bandwidth. The two-stage amplifier architecture, comprising a PFET cascode preamp and an NFET differential cascode postamp, is optimized for broadband input match up to 100 GHz, an output compression point of 15.2 dBm and 4 $V_{pp,d}$ linear output swing across a 60-Ω differential load.

Extensive simulations and experimental results validate the driver's performance, with small-signal measurements showing a peak gain of 20.4 dB at 76 GHz and a -3dB bandwidth of 95 GHz. Large-signal measurements demonstrate the driver's ability to deliver a 4-$V_{pp,d}$ output swing with 1.6% total harmonic distortion, while time-domain measurements reveal

robust eye patterns at symbol rates up to 140 GBaud for NRZ and 100 GBaud for PAM-4 modulation schemes, respectively. The driver's performance is evaluated under varying temperature conditions, with a thermoelectric heater/cooler used to control the chip's backside temperature. Measurements show that the chip temperature significantly affects the performance, and the design flow should be revised to include thermal effects at early stages.

In conclusion, the proposed driver achieves state-of-the-art performance. Future work could focus on optimizing the driver's thermal management, and exploring additional linearization and equalization techniques to enhance performance beyond 75 GHz. The design's flexibility, robustness, and high performance position it as a potential enabler for high-speed communication systems.

7.2 Major Contributions

Contributions to the design of broadband linear drivers including: amplitude-peaking filter networks, linear high-swing optical modulator drivers, and a power-efficient linearizer are summarized in this section.

7.2.1 Amplitude-peaking filter networks

The interstage filter network in the thesis is a 7-pole, 2-zero shunt/double-series network designed to enhance frequency response and control gain peaking at 75 GHz. This network configuration couples two gain stages. The inductance and capacitance values are selected to meet the small-signal bandwidth and peak gain specifications. It has been shown that a simplified 4-pole, 1-zero, lumped-element model predicts the bandwidth extension realized through strategic placement of poles and zeros in the circuit's transfer function. The filter operation is not dependent on a particular source and load capacitor ratio, enhancing the robustness of the proposed design methodology, and minimizing the need for tuning during implementation. Traditional peaking methods, such as series and shunt peaking, generally require a fixed capacitor ratio to obtain a specific bandwidth extension ratio. The approach developed and presented in this thesis is much more adaptable.

The peak-gain profile could be tuned by adjusting the values of the peaking inductors. The analysis indicates that the peaking frequency is primarily determined by a series inductor, which is connected to either the source or load capacitors, whichever has a smaller capacitance value. A damping factor for the dominant poles of 0.05 achieved the

intended peak gain. This optimization extended the network's -3dB bandwidth to 148 GHz, equivalent to a 2.95x bandwidth extension, facilitated by additional shunt tuning capacitors of 30 fF and 50 fF. Furthermore, comprehensive electromagnetic simulations were used to model parasitic components accurately. The detailed lumped-element model was derived to aid in fine tuning the overall system response.

7.2.2 Linear high-swing optical modulator drivers

The postamp stage of the broadband driver is designed to meet RF output power, linearity, bandwidth and gain specifications. This stage consists of two parallel differential cascode amplifiers combined at the output, each with a common-source FET and two common-gate FETs. By combining two 80-μm units in parallel, the BEOL metal on the source and drain terminals (handling both DC and signal currents) doubled, improving the heat dissipation capability, as the postamp consumes 300 mW of DC power. The common-source stage includes neutralization capacitors (18 fF) to mitigate the Miller effect, improving isolation and stability across a wide frequency range. A parallel R-C network is implemented at the gate of common-gate FETs which spreads the output swing evenly across transistors, resulting in better reliability. A bridging resistor is added between the gate and source terminals of the common-gate FETs, creating a low-frequency path to compensate the frequency response (i.e., high-pass) of the gate R-C network.

The output combining network was designed to achieve the desired -1dB compression point at the output, which is essential for maintaining linearity for a differential 60-Ω load. Optimization of the load impedance seen at the drain of the top-most FET was accomplished through output load-pull simulations at 75 GHz, which is the maximum frequency of interest for linearity specifications. Although the optimal load is a narrowband definition, it still helped to improve the OP_{-1dB} beyond 50 GHz (below 50 GHz, the postamp stays linear by selecting a 3-V supply and operating at 0.5 mA/μm current density). The simulated -3dB bandwidth of the postamp stage with respect to its DC gain is 91 GHz (excluding the input pole, as it was addressed during the interstage filter network design), featuring a gain roll-off of 2.7 dB at the target peak-gain frequency of 75 GHz. This confirms that the stage can support high-speed operation (i.e., 91-GHz bandwidth) with minimal signal distortion (linear output power maximized by presenting the optimal impedance to the output drains). The output swing, vital for data centre optical interconnection systems, was designed to fulfill an OP_{-1dB} specification of 15.2 dBm (equivalent to 4-V_{ppd} across a differential 60-Ω load), thus delivering a robust output for passive Mach-Zehnder optical modulators.

7.2.3 Power-efficient linearizer

The linearity of the driver is dominated by the postamp (i.e., output stage), as shown in Eq. 1.3. An auxiliary circuit may be introduced to linearize (i.e., a linearizer) the postamp to optimize the driver's linearity. Such linearization techniques includes feedback, feedforward, and adaptive biasing, and digital predistortion [24].

The postamp linearization in this thesis utilizes the IM3 cancellation technique to enhance linearity at frequencies beyond 50 GHz. This technique is specifically tailored to cancel the third-order intermodulation distortion (IM3), which is directly correlated to OP_{-1dB} of the postamp, achieving a maximum IM3 suppression of 25 dB across the postamp's operational range.

A dual-amplifier configuration is utilized, consisting of a main Class-A amplifier and an auxiliary amplifier set to operate at a reduced conduction angle, specifically optimized for IM3 cancellation. The auxiliary amplifier uses 40-μm wide FETs, which are smaller than those in the main amplifier, generating an error signal that corresponds with the IM3 distortion of the main signal. This design achieves at least 2-dB improvement in OP_{-1dB} while consuming 3 mW, maintaining linearity with minimal bandwidth sacrifice (i.e., the auxiliary inputs are connected at the input of common-gate stages for a cascode main amplifier).

A dynamic control voltage (V_C) enhances postamp efficiency by adjusting the auxiliary amplifier's conduction angle. Ranging from 0.95 V to 1.45 V, this voltage modulates the auxiliary amplifier's operation: at 0.95 V, it achieves optimal IM3 cancellation, maintaining a balance between linearity and power-saving. When V_C reaches 1.45 V, the auxiliary amplifier is deactivated, setting a performance baseline. This dynamic tuning ensures that the IM3 suppression precisely matches the amplifier's requirements at various operating conditions, optimizing both efficiency and linearity. The findings confirm that with IM3 cancellation, the postamp can maintain linear operation above 50 GHz while reducing current consumption (1 mA quiescent current, 5 mA DC current when driving at the input 1-dB compression power level at 75 GHz).

7.3 Future Work

In this section, potential avenues for future research and development are proposed that could improve the performance and functionality of broadband drivers. Key areas of focus include the design of variable-gain amplifiers for improved adaptability, multi-stage driver

architectures to optimize linearity and bandwidth, and the development of high-Q capacitors to minimize losses in interstage networks. Additionally, impacts of postamp source degeneration impedance and the thermal effect on peak gain are considered, and strategies to mitigate these effects are proposed.

7.3.1 Variable-gain amplifier and multi-stage driver designs

The driver gain specification (e.g., 15.7 dB at low frequencies and 24.7 dB at peak) is calculated by Eq. 1.1 with the assumption that the input and output voltage swings are 0.3 $V_{pp,d}$ and 4 $V_{pp,d}$, respectively. According to OIF HB-CDM 2.0 [1], the input sensitivity is specified as a range of 0.3 $V_{pp,d}$ to 0.6 $V_{pp,d}$. Therefore, a variable-gain amplifier (VGA) is needed to adjust the gain. An automatic-gain control loop (AGC) could compensate for any drift in the gain during operation. The VGA should have a minimum range of 6 dB (i.e., for 0.3 $V_{pp,d}$ to 0.6 $V_{pp,d}$ input swing).

Reviewing the published literature (e.g., [26, 32, 34, 53]), the VGA is often the first stage of the driver. To optimize power efficiency, the VGA is usually implemented with narrow-width transistors, leading to limited RF power and linearity at the output. Putting the VGA in the first stage could minimize its effect on the overall linearity of the driver, as seen in Eq. 1.3 (power gain in later stages suppresses the linearity bottleneck of the first stage). Moreover, since the overall driver common-mode rejection ratio (CMRR) is dominated by the first stage, CMRR of the VGA is another critial metric.

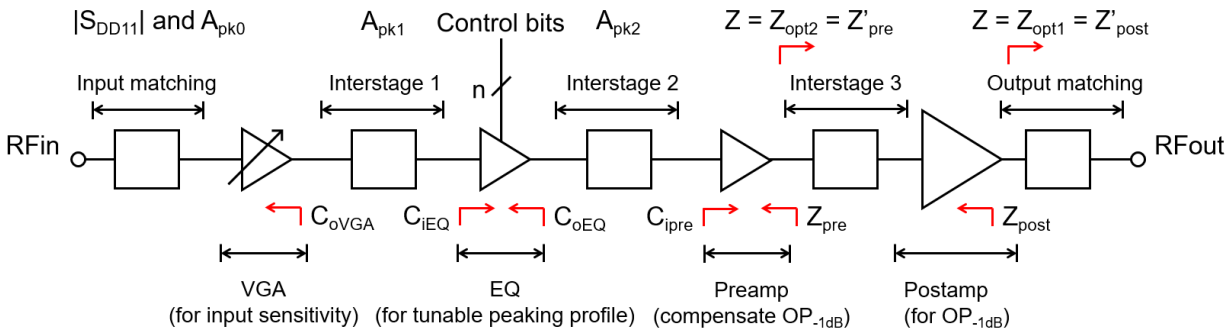


Figure 7.1: Block diagram of a multi-stage driver with EQ and gain control

In a fully developed electrical driver prototype, the driver usually consists of a VGA, multiple equalization stages (EQ), and an output power stage, as seen in [32, 34]. The VGA accommodates the desired input sensitivity, ranging from 300 mV to 600 mV, while

the EQ adjusts both low-frequency and peak gains by tuning the parallel R-C degeneration network to maintain a consistent electro-optical (E/O) response. The output power stage generates the required linear output swing (i.e., 4 V_{ppd}) for the optical modulator (MZM). An additional preamp could be inserted between the EQ and output power stage to mitigate the loss in linearity caused by the EQ and VGA (see Eq. 1.3), as shown in Fig. 7.1.

In a multi-stage driver, the output matching network is used to maximized the linear output power, as implemented in this thesis. The interstage 3 could be reserved for optimizing interstage power (e.g., Eq. 1.3 indicates that the preamp stage could still affect the driver's linearity if the preamp OP_{-1dB} is too low). Therefore, the peaked gain profile has to be implemented by interstages 1 and 2. Capacitances C_{oVGA}, C_{iEQ}, C_{oEQ} and C_{ipre} correspond to the shunt-equivalent source and load capacitances for interstages 1 and 2. These capacitances are required to determine the design values for the shunt/double-series tuning network described in this thesis. By spreading the gain peaking across multiple interstages, the likelihood of instability is diminished, eliminating the need for a highly underdamped network. Moreover, if VGA and EQ are low-power stages, the reduction in peak gain will likely be less significant than if the gain-peaking interstage connected two high-power stages, which would experience local heat accumulation.

Finally, the design of the input matching network must aim to enhance the input matching bandwidth and facilitate cascading peak gain. Note that introducing gain peaking in the input matching network involves a trade-off between $|S_{DD11}|$ bandwidth with peak gain. Therefore, this network should not serve as the primary contributor to the peak gain. The total peak gain is determined by the product $A_{pk0} \cdot A_{pk1} \cdot A_{pk2}$ in Fig. 7.1.

7.3.2 Compensation of thermal effects

Recall that reduction in the peak gain (see Fig. 6.9) is due to the heat generated by self-heating of the driver. For example, the peaking amplitude decreases from the simulated 8.5 dB to measured 6 dB. The reason is explained in Section 6.3. In this section, potential solutions are proposed as future work to measure the temperature at driver's output and compensate for unwanted thermal effects on the driver performance.

A thermal imaging system could be used to map temperatures across the chip [80] and verify the temperature of the postamp (e.g., at the drain of the driver where the maximum current occurs). For open-drain postamps, a differential dummy resistor could also be added between the drains to estimate the temperature at the driver's output. Furthermore, transistor layouts with better heat dissipation could be developed to mitigate the performance degradation due to thermal effects [81].

For the two-stage driver designed in this thesis, the peak amplitude (i.e., delta between the peak and DC gains) may be retained by making the interstage load resistor tunable (e.g., use MOS resistor or a digitally controlled resistor bank). Prioritizing retaining the gain peaking profile, the thermal effect on the interstage load is compensated by lowering the interstage load resistance to the expected value via tuning.

However, the overall gain drop due to reduced postamp transconductance across frequencies still remains. To compensate the temperature-dependent transconductance, a conventional solution is to use a temperature-independent current reference (e.g., a proportional to absolute temperature, or PTAT source) to set the transconductance. However, in this thesis, the postamp omits the tail current source to save DC power consumption. Furthermore, the postamp is biased in Class A near the peak f_T (i.e., 0.36 mA/μm to 0.48 mA/μm) and the output transconductance does not vary with increasing bias current (e.g., Fig. F.17 shows that the gain does not vary much versus increasing postamp current). One future work is to back off the postamp bias to Class AB and increase the bias current to attain higher transconductance to overcome the temperature effect. The tail current source for the PFET preamp has to be adjusted accordingly to ensure DC gain of the driver remains unchanged.

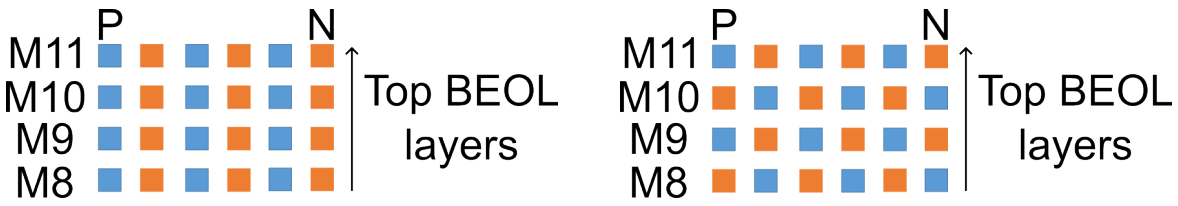
In a multistage driver as shown in Fig. 7.1, to mitigate this problem and recover the peak gain, a variable-gain amplifier (VGA) and a continuous-time linear equalizer (CTLE) are necessary (both covered in Section 7.3.1). VGA and CTLE could be used to adjust the driver's gain and the peak amplitude, respectively. Moreover, the input signal to the driver is typically generated from a digital-signal processing (DSP) unit [77]. The input signal may be pre-emphasized by the DSP to compensate for thermal degradations of the driver.

7.3.3 Custom high-Q capacitor design

The design kit provides alternate-polarity, metal-oxide-metal (APMOM) capacitors and metal-oxide-metal (MOM) capacitors. Cross-sectional views of these capacitors are seen in Fig. 7.2. For MOM capacitors (Fig. 7.2a), the capacitance arises from the sidewall capacitance between the positive and negative finger stacks, with each finger subjected to two sidewall capacitances. Conversely, the APMOM capacitor, shown in Fig. 7.2b, has four sidewall capacitances per finger, increasing the capacitance of the layout for a given area [82].

It is desired to have a custom-designed APMOM capacitor that has the finger width and spacing equivalent to a 10-V MOM capacitor to attain a higher Q-factor compared

to other capacitors from the design kit (i.e., high Q-factor leads to low ohmic loss in the interstage filter network, which preserves the peaking amplitude). The desired APMOM offers the benefit of a higher capacitance density inherent to the APMOM structure, which minimizes the capacitor's footprint and enhances the Q-factor for a fixed capacitance value.



(a) MOM (b) APMOM

Figure 7.2: Cross-sectional views of various capacitor structures (P and N are the positive and negative polarities of capacitors)

References

[1] O. I. Forum, "Implementation agreement for high bandwidth coherent driver modulator (hb-cdm)." White Paper, July 2021.

[2] A. Samani, E. El-Fiky, M. Morsy-Osman, R. Li, D. Patel, T. Hoang, M. Jacques, M. Chagnon, N. Abadía, and D. V. Plant, "Silicon photonic mach–zehnder modulator architectures for on chip pam-4 signal generation," *Journal of Lightwave Technology*, vol. 37, no. 13, pp. 2989–2999, 2019.

[3] J. Kim and J. F. Buckwalter, "A 40-gb/s optical transceiver front-end in 45 nm soi cmos," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 3, pp. 615–626, 2012.

[4] Day-Uei Li and Chia-Ming Tsai, "10-gb/s modulator drivers with local feedback networks," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 5, pp. 1025–1030, 2006.

[5] S. Mandegaran and A. Hajimiri, "A breakdown voltage multiplier for high voltage swing drivers," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 2, pp. 302–312, 2007.

[6] P. Quinn, "A cascode amplifier nonlinearity correction technique," in *1981 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, vol. XXIV, pp. 188–189, 1981.

[7] N. Hosseinzadeh, K. Fang, L. Valenzuela, C. Schow, and J. F. Buckwalter, "A 50-gb/s optical transmitter based on co-design of a 45-nm cmos soi distributed driver and 90-nm silicon photonic mach-zehnder modulator," in *2020 IEEE/MTT-S International Microwave Symposium (IMS)*, pp. 205–208, 2020.

[8] C. Radio-television and T. Commission, "Communications market reports: High-speed broadband." https://crtc.gc.ca/eng/publications/reports/PolicyMonitoring/ban.htm, 2022. (Accessed on 12/10/2023).

[9] C. Radio-television and T. Commission, "Communications market reports: Mobile wireless." https://crtc.gc.ca/eng/publications/reports/PolicyMonitoring/mob.htm. (Accessed on 12/10/2023).

[10] Y. Hao, "Investigation and technological comparison of 4g and 5g networks," *Journal of Computer and Communications*, no. 9, pp. 36–43, 2021.

[11] F. Chang, *Datacenter Connectivity Technologies: Principles and Practice*. River Publishers, 1st ed., 2018.

[12] S. Shekhar, W. Bogaerts, L. Chrostowski, J. E. Bowers, M. Hochberg, R. Soref, and B. J. Shastri, "Silicon photonics – roadmapping the next generation," 2023.

[13] T. S. S. O. ITU, "Spectral grids for wdm applications: Cwdm wavelength grid," tech. rep., INTERNATIONAL TELECOMMUNICATION UNION, 2003.

[14] G. L. B. Collings, F. Helsmann, "Reference guide to fiber optic testing," tech. rep., VIAVI Solutions, 2018.

[15] "Cisco 400g qsfp-dd high-power (bright) optical module," 2023. Cisco Systems, Inc.

[16] "Product details eoptolink's qsfp-dd800 2×fr4 optical transceiver," 2023. Eoptolink Technology Inc., Ltd.

[17] A. Samani, *Silicon Photonic Mach-Zehnder Modulator Architectures for High Order Modulation Formats*. PhD thesis, McGill University, 2018.

[18] "Integrated 5g mmwave mobile fems/trxs using 22fdx™ rf and 22fdx rf+." https://gf.com/wp-content/uploads/2021/10/GF21-5G-mmWave-mobile-FEM-22FDX-RF-0927_0.pdf, 2021. (Accessed on 10/23/2023).

[19] B. Cui and J. R. Long, "A 1.7-db minimum nf, 22–32-ghz low-noise feedback amplifier with multistage noise matching in 22-nm fd-soi cmos," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 5, pp. 1239–1248, 2020.

[20] D. Y. Lie, "Highly efficient wide-band power amplifiers design for millimeter-wave digital arrays (midas)," tech. rep., Texas Tech University, 2022.

[21] Tektronix, "Coherent optical measurements common transmitter and receiver impairments." Poster, Aug 2018.

[22] Y. Zhao, L. Vera, and J. R. Long, "A 10 gb/s, 6 v p-p , digitally controlled, differential distributed amplifier mzm driver," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 9, pp. 2030–2043, 2014.

[23] M. Edwards and J. Sinsky, "A new criterion for linear 2-port stability using a single geometrically derived parameter," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 12, pp. 2303–2311, 1992.

[24] B. Razavi, *RF Microelectronics*. Pearson, 2nd ed., 2011.

[25] D. J. Cassan and J. R. Long, "A 1-v transformer-feedback low-noise amplifier for 5-ghz wireless lan in 0.18-μm cmos," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3, pp. 427–435, 2003.

[26] A. H. Ahmed, D. Lim, A. Elmoznine, Y. Ma, T. Huynh, C. Williams, L. Vera, Y. Liu, R. Shi, M. Streshinsky, A. Novack, R. Ding, R. Younce, R. Sukkar, J. Roman, M. Hochberg, S. Shekhar, and A. Rylyakov, "30.6 a 6v swing 3.6% thd ¿40ghz driver with 4.5× bandwidth extension for a 272gb/s dual-polarization 16-qam silicon photonic transmitter," in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, pp. 484–486, 2019.

[27] S. Galal and B. Razavi, "10-gb/s limiting amplifier and laser/modulator driver in 0.18-μm cmos technology," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2138–2146, 2003.

[28] L. Vera and J. R. Long, "A 40-gb/s sige-bicmos mzm driver with 6-v $_{\text{p-p}}$ output and on-chip digital calibration," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 2, pp. 460–471, 2017.

[29] J. Hwang, G.-S. Jeong, W. Bae, J.-E. Park, C. S. Yoon, J. M. Yoon, J. Joo, G. Kim, and D.-K. Jeong, "A 32 gb/s, 201 mw, mzm/eam cascode push–pull cml driver in 65 nm cmos," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 4, pp. 436–440, 2018.

[30] H.-S. Choi, J. Hwang, G.-S. Jeong, G. Kim, and D.-K. Jeong, "A 35-gb/s 0.65-pj/b asymmetric push-pull inverter-based vcsel driver with series inductive peaking in 65-nm cmos," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 12, pp. 1824–1828, 2018.

[31] S. Shopov and S. P. Voinigescu, "A 3×60 Gb/s transmitter/repeater front-end with 4.3 v_{pp} single-ended output swing in a 28nm utbb fd-soi technology," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 7, pp. 1651–1662, 2016.

[32] T. Jyo, M. Nagatani, J. Ozaki, M. Ishikawa, and H. Nosaka, "12.3 a 48ghz bw 225mw/ch linear driver ic with stacked current-reuse architecture in 65nm cmos for beyond-400gb/s coherent optical transmitters," in *2020 IEEE International Solid-State Circuits Conference - (ISSCC)*, pp. 212–214, 2020.

[33] S. Galal and B. Razavi, "40-gb/s amplifier and esd protection circuit in 0.18-/spl mu/m cmos technology," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2389–2396, 2004.

[34] S. Nakano, M. Nagatani, M. Nogawa, Y. Kawamura, K. Kikuchi, K. Tsuzuki, and H. Nosaka, "A 2.25-mw/gb/s 80-gb/s-pam4 linear driver with a single supply using stacked current-mode architecture in 65-nm cmos," in *2017 Symposium on VLSI Circuits*, pp. C322–C323, 2017.

[35] L. Vera, J. R. Long, and B. J. Gross, "A low-power sige feedback amplifier with over 110ghz bandwidth," in *2014 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*, pp. 1–4, 2014.

[36] J. Kim and J. F. Buckwalter, "Staggered gain for 100+ ghz broadband amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 5, pp. 1123–1136, 2011.

[37] R. Ding, Z. Xuan, P. Yao, T. Baehr-Jones, D. Prather, and M. Hochberg, "Power-efficient low-noise 86 ghz broadband amplifier in 130 nm sige bicmos," *Electronics Letters*, vol. 50, no. 10, pp. 741–743, 2014.

[38] H. Wakita, M. Nagatani, K. Kurishima, M. Ida, and H. Nosaka, "An over-67-ghz-bandwidth 2 vppd linear differential amplifier with gain control in 0.25-µm inp dhbt technology," in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 1–3, 2016.

[39] T. Chen, H. Su, T. Lee, and S. S. H. Hsu, "A 64-gb/s 4.2-vpp modulator driver using stacked-fet distributed amplifier topology in 65-nm cmos," in *2019 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 730–733, 2019.

[40] R. J. A. Baker, J. Hoffman, P. Schvan, and S. P. Voinigescu, "Sige bicmos linear modulator drivers with 4.8-vpp differential output swing for 120-gbaud applications,"

in *2017 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 260–263, 2017.

[41] Q. H. Le, D. K. Huynh, D. Wang, T. Kämpfe, and S. Lehmann, "Dc-110 ghz characterization of 22fdx® fdsoi transistors for 5g transmitter front-end," in *ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC)*, pp. 218–221, 2019.

[42] H.-T. Dabag, B. Hanafi, F. Golcuk, A. Agah, J. F. Buckwalter, and P. M. Asbeck, "Analysis and design of stacked-fet millimeter-wave power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 4, pp. 1543–1556, 2013.

[43] O. El-Aassar and G. M. Rebeiz, "A dc-to-108-ghz cmos soi distributed power amplifier and modulator driver leveraging multi-drive complementary stacked cells," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 12, pp. 3437–3451, 2019.

[44] A. H. Ahmed, A. E. Moznine, D. Lim, Y. Ma, A. Rylyakov, and S. Shekhar, "A dual-polarization silicon-photonic coherent transmitter supporting 552 gb/s/wavelength," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 9, pp. 2597–2608, 2020.

[45] A. Agah, J. A. Jayamon, P. M. Asbeck, L. E. Larson, and J. F. Buckwalter, "Multi-drive stacked-fet power amplifiers at 90 ghz in 45 nm soi cmos," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 5, pp. 1148–1157, 2014.

[46] H. Zhang and E. Sánchez-Sinencio, "Linearization techniques for cmos low noise amplifiers: A tutorial," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 1, pp. 22–36, 2011.

[47] S.-T. Lim and J. Long, "A feedforward compensated high-linearity differential transconductor for rf applications," in *2004 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 1, pp. I–I, 2004.

[48] M. S. Dadash, D. Harame, and S. P. Voinigescu, "Large-swing 22nm si/sige fd-soi stacked cascodes for 56gbaud drivers and 5g pas," in *2018 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, pp. 267–270, 2018.

[49] A. Giuglea, M. M. Khafaji, G. Belfiore, R. Henker, and F. Ellinger, "A low-distortion modulator driver with over 6.5-vpp differential output swing and bandwidth above 60 ghz in a 130-nm sige bicmos technology," *IEEE Access*, vol. 10, pp. 78667–78676, 2022.

[50] P. Rito, I. García López, A. Awny, M. Ko, A. C. Ulusoy, and D. Kissinger, "A dc-90-ghz 4- v_{pp} modulator driver in a 0.13- μm sige:c bicmos process," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 5192–5202, 2017.

[51] F. Iseini, A. Malignaggi, F. Korndörfer, M. Inac, and G. Kahmen, "Lumped ultra-broadband linear driver in 130 nm sige sg13g3 technology," in *2022 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, pp. 136–139, 2022.

[52] A. Zandieh, P. Schvan, and S. P. Voinigescu, "Linear large-swing push–pull sige bicmos drivers for silicon photonics modulators," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 5355–5366, 2017.

[53] T. Jyo, M. Nagatani, Y. Ogiso, S. Yamanaka, and H. Nosaka, "An over 67-ghz bandwidth 21-db gain 4.5-vppd linear modulator driver for 100-gbd coherent optical transmitter," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 6, pp. 705–708, 2021.

[54] E. M. Peter Staric, *Wideband Amplifiers*. Springer New York, 1st ed., 2006.

[55] M. A. Haidekker, *Linear Feedback Controls: The Essentials*. Elsevier Inc., 2nd ed., 2020.

[56] B. C. K. Farid Golnaraghi, *Automatic Control Systems*. Wiley, 9th ed., 2010.

[57] R. Pallas-Areny and J. Webster, "Common mode rejection ratio for cascaded differential amplifier stages," *IEEE Transactions on Instrumentation and Measurement*, vol. 40, no. 4, pp. 677–681, 1991.

[58] J. William H. Hayt, J. E. Kemmerly, and S. M. Durbin, *Engineering Circuit Analysis*. MacGraw-Hill, 8th ed., 2012.

[59] P. Quinn, "A cascode amplifier nonlinearity correction technique," in *1981 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, vol. XXIV, pp. 188–189, 1981.

[60] Z. El-Khatib, L. MacEachern, and S. A. Mahmoud, "Improvement of carrier power to third-order intermodulation distortion power ratio in cmos distributed amplifiers," in *2008 International Conference on Microelectronics*, pp. 98–101, 2008.

[61] S. Ong, S. Lehmann, W. Chow, C. Zhang, C. Schippel, L. Chan, Y. Andee, M. Hauschildt, K. Tan, J. Watts, C. Lim, A. Divay, J. Wong, Z. Zhao, M. Govindarajan, C. Schwan, A. Huschka, A. Bcllaouar, W. LOo, J. Mazurier, C. Grass, R. Taylor, K. Chew, S. Embabi, G. Workman, A. Pakfar, S. Morvan, K. Sundaram, M. T. Lau, B. Rice, and D. Harame, "A 22nm fdsoi technology optimized for rf/mmwave applications," in *2018 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, pp. 72–75, 2018.

[62] H.-S. Kim, K. Park, H. Oh, and E. S. Jung, "Effective gate layout methods for rf performance enhancement in mosfets," *IEEE Electron Device Letters*, vol. 30, no. 10, pp. 1105–1107, 2009.

[63] Z. Zhao, S. Lehmann, L. Lucci, Y. Andee, A. Divay, L. Pirro, T. Herrmann, A. Zaka, R. Sousa, P. J. Artz, K. Hempel, J. Faul, T. Chen, R. Taylor, J. Mazurier, C. Grass, J. Hoentschel, and D. Harame, "22fdx® fmax optimization through parasitics reduction and gm boost," in *ESSDERC 2019 - 49th European Solid-State Device Research Conference (ESSDERC)*, pp. 166–169, 2019.

[64] J. Lienig, "Electromigration and its impact on physical design in future technologies," in *Proceedings of the 2013 ACM International Symposium on Physical Design*, ISPD '13, (New York, NY, USA), p. 33–40, Association for Computing Machinery, 2013.

[65] R. O. Topaloglu, *Characterization, Modeling and Optimization of Fills and Stress in Semiconductor Integrated Circuits*. PhD thesis, UNIVERSITY OF CALIFORNIA, SAN DIEGO, 2008.

[66] M. engineering, "Datasheet - tec controller tec-1091." Datasheet, Oct 2024.

[67] B. Wei, W. Luo, J. Du, Y. Ding, Y. Guo, G. Zhu, Y. Zhu, and B. Li, "Thermal interface materials: From fundamental research to applications," *SusMat*, vol. 4, no. 6, p. e239, 2024.

[68] GlobalFoundries, "22fdx-plus technology design manual." Design Manual, Jun 2023.

[69] C. Jeong, S. Datta, and M. Lundstrom, "Thermal conductivity of bulk and thin-film silicon: A landauer approach," *Journal of Applied Physics*, vol. 111, p. 093708, 05 2012.

[70] M. Asheghi, M. N. Touzelbaev, K. E. Goodson, Y. K. Leung, and S. S. Wong, "Temperature-dependent thermal conductivity of single-crystal silicon layers in soi substrates," *Journal of Heat Transfer*, vol. 120, pp. 30–36, 02 1998.

[71] D. Frickey, "Conversions between s, z, y, h, abcd, and t parameters which are valid for complex source and load impedances," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 2, pp. 205–211, 1994.

[72] M. Wojnowski, V. Issakov, G. Sommer, and R. Weigel, "Multimode trl calibration technique for characterization of differential devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 7, pp. 2220–2247, 2012.

[73] Anritsu, "Understanding eye pattern measurements." Application note, Mar 2010.

[74] L. S. C. of the IEEE Computer Society, "Ieee standard for ethernet," tech. rep., IEEE Computer Society, 2022.

[75] K. Technologies, "Flexdca user's guide for sampling oscilloscope products." https://helpfiles.keysight.com/scopes/FlexDCA-UG/Content/Topics/FlexDCA-UG-Home.htm, 2024. (Accessed on 01/23/2025).

[76] J. G. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill, 5th ed., 2008.

[77] M. Nagatani, H. Wakita, H. Yamazaki, Y. Ogiso, M. Mutoh, M. Ida, F. Hamaoka, M. Nakamura, T. Kobayashi, Y. Miyamoto, and H. Nosaka, "A beyond-1-tb/s coherent optical transmitter front-end based on 110-ghz-bandwidth 2:1 analog multiplexer in 250-nm inp dhbt," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 9, pp. 2301–2315, 2020.

[78] S. Niu, R. Ahmed, J. Lambrecht, J. Declercq, M. Verbeke, M. Verplaetse, B. Moeneclaey, C. Wang, J. V. Kerrebrouck, N. Singh, J. Bauwelinck, and X. Yin, "A linear modulator driver with over 70-ghz bandwidth 21.8-db gain and 3.4-vppd output swing for beyond 120-gbd optical links," *IEEE Transactions on Microwave Theory and Techniques*, vol. 72, no. 7, pp. 4080–4091, 2024.

[79] H. Uemura, T. Misawa, N. Itabashi, M. Kurokawa, Y. Sugimoto, S. Kumagai, M. Takechi, and K. Tanaka, "A 19-db peaking at 72-ghz and 4.1-vppd output swing sige bicmos linear driver with dynamic cascode output buffer," in *2023 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, pp. 159–162, 2023.

[80] F. Duan, K. Chen, S. Wang, L. Wei, Y. Yu, and D. Ban, "Temperature profile and transient response of thermally tunable ridge waveguides with laterally supported suspension," *Applied Physics Letters*, vol. 116, p. 011102, 01 2020.

[81] S. Alluri, J. Jayamon, B. Hanafi, and P. M. Asbeck, "Interconnect effects on thermal resistance of cmos-soi transistors in microwave power integrated circuits," *Solid-State Electronics*, vol. 186, p. 108149, 2021.

[82] J. Shi, A. Sidelnicov, K. W. J. Chew, M. S. Chin, C. Schippel, J. d. Santos, F. Schlaphof, L. Meinshausen, J. R. Long, and D. Harame, "Evolution and optimization of beol mom capacitors across advanced cmos nodes," in *2018 48th European Solid-State Device Research Conference (ESSDERC)*, pp. 190–193, 2018.

[83] C. Radio-television and T. Commission, "Annual highlights of the telecommunications sector 2021," tech. rep., Canadian Radio-television and Telecommunications Commission, 2023.

[84] T. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge Univeristy Press, 2nd ed., 2004.

[85] D. M. Pozar, *Microwave Engineering*. John Wiley Sons, Inc., 4th ed., 2012.

[86] E. Cherry and D. Hooper, *Amplifying Devices and Low-Pass Amplifier Design*. Wiley, 1st ed., 1968.

[87] O. El-Aassar and G. M. Rebeiz, "A cascaded multi-drive stacked-soi distributed power amplifier with 23.5 dbm peak output power and over 4.5-thz gbw," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 7, pp. 3111–3119, 2020.

[88] Y. Baeyens, N. Weimann, P. Roux, A. Leven, V. Houtsma, R. F. Kopf, Yang Yang, J. Frackoviak, A. Tate, J. S. Weiner, P. Paschke, and Young-Kai Chen, "High gain-bandwidth differential distributed inp d-hbt driver amplifiers with large (11.3 v/sub pp/) output swing at 40 gb/s," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 10, pp. 1697–1705, 2004.

[89] Y. Li, P. Chiu, K. Li, D. J. Thomson, G. T. Reed, and S. S. H. Hsu, "A 40-gb/s 4-vpp differential modulator driver in 90-nm cmos," *IEEE Microwave and Wireless Components Letters*, vol. 28, no. 1, pp. 73–75, 2018.

[90] B. Analui and A. Hajimiri, "Bandwidth enhancement for transimpedance amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 8, pp. 1263–1270, 2004.

[91] E. L. Ginzton, W. R. Hewlett, J. H. Jasberg, and J. D. Noe, "Distributed amplification," *Proceedings of the IRE*, vol. 36, no. 8, pp. 956–969, 1948.

[92] S. Galal and B. Razavi, "Broadband esd protection circuits in cmos technology," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2334–2340, 2003.

[93] D. Tang, Z. Li, J. Chen, P. Zhou, Z. Chen, D. Hou, and W. Hong, "An ultra-wideband amplifier with a novel non- distributed butterfly topology achieving 2-250 ghz bandwidth and 4.67 thz gbw in 130nm sige bicmos," in *2023 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1–2, 2023.

[94] J. Kim and J. F. Buckwalter, "Bandwidth enhancement with low group-delay variation for a 40-gb/s transimpedance amplifier," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 8, pp. 1964–1972, 2010.

[95] H. Romain, D. Jean-Yves, K. Agnieszka, J. Filipe, B. Fabrice, R. Muriel, N. Virginie, M. Colin, K. Abed-Elhak, and O. Achour, "Over 70-ghz 4.9-vppdiff inp linear driver for next generation coherent optical communications," in *2019 IEEE BiCMOS and Compound semiconductor Integrated Circuits and Technology Symposium (BCICTS)*, pp. 1–4, 2019.

[96] S. Mohan, M. Hershenson, S. Boyd, and T. Lee, "Bandwidth extension in cmos with optimized on-chip inductors," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 3, pp. 346–355, 2000.

[97] R. F. Stevens, "Polarisation extinction ratio – measurement requirements for optical communication systems," tech. rep., Centre for Electrical and Time Metrology, National Physical Laboratory, Teddington, Middlesex, UK, 2002.

[98] P. Heydari and M. Safiallah, "Evolution of broadband amplifier design: From single-stage to distributed topology," *IEEE Microwave Magazine*, vol. 24, no. 9, pp. 18–29, 2023.

[99] Y. Chen, P.-I. Mak, H. Yu, C. C. Boon, and R. P. Martins, "An area-efficient and tunable bandwidth- extension technique for a wideband cmos amplifier handling 50+ gb/s signaling," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 12, pp. 4960–4975, 2017.

[100] S. Shekhar, J. Walling, and D. Allstot, "Bandwidth extension techniques for cmos amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 11, pp. 2424–2439, 2006.

[101] A. Meaamar, C. C. Boon, K. S. Yeo, and M. A. Do, "A wideband low power low-noise amplifier in cmos technology," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 4, pp. 773–782, 2010.

[102] R. Nagarajan, I. Lyubomirsky, and O. Agazzi, "Low power dsp-based transceivers for data center optical fiber communications (invited tutorial)," *Journal of Lightwave Technology*, vol. 39, no. 16, pp. 5221–5231, 2021.

[103] Q. Fu, M. Tan, D. Xing, S. Shao, Z. Hu, and J. Feng, "A 57.2-gb/s pam4 driver for a segmented silicon-photonics mach-zehnder modulator with extinction ratio ¿9-db in 45-nm rf-soi cmos technology," in *2021 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–5, 2021.

[104] T. Cheung and J. Long, "Shielded passive devices for silicon-based monolithic microwave and millimeter-wave integrated circuits," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 5, pp. 1183–1200, 2006.

Appendices

A The effects of finite Q-factor of series coils in a shunt/double-series network

The analysis for the circuit shown in Fig. 3.11 does not include the effect of ohmic losses from the inductor coils. Series losses of L_1 and L_2 damp the system (i.e., increases $|\sigma_3|$), therefore reducing the peak amplitude (on the other hand, the loss of L_3 appears in series with the interstage load and its effect may be tuned out by adjusting the load). Define Q_{pk} as the Q-factor at the peaking frequency, which is given by

$$Q_{pk} = \frac{\omega_{pk} L_1}{R_1} = \frac{\omega_{pk} L_2}{R_2}, \tag{A.1}$$

where R_1 and R_2 are the intrinsic series resistance of L_1 and L_2, respectively. The effect of Q_{pk} on the peak amplitude is shown in Fig. A.1.

An increase in Q-factor yields a diminishing return for the achievable peak amplitude via the dominant second-order transfer function, but the roll-off due to the non-dominant second-order remains constant at ∼5 dB. A realizable Q-factor for on-chip inductors is less than 50 [104]. Therefore, the interstage amplitude peaking is expected to have a decrease of at least 5 dB due to losses in the series coils (i.e., from the ideal 35 dB to 30 dB).

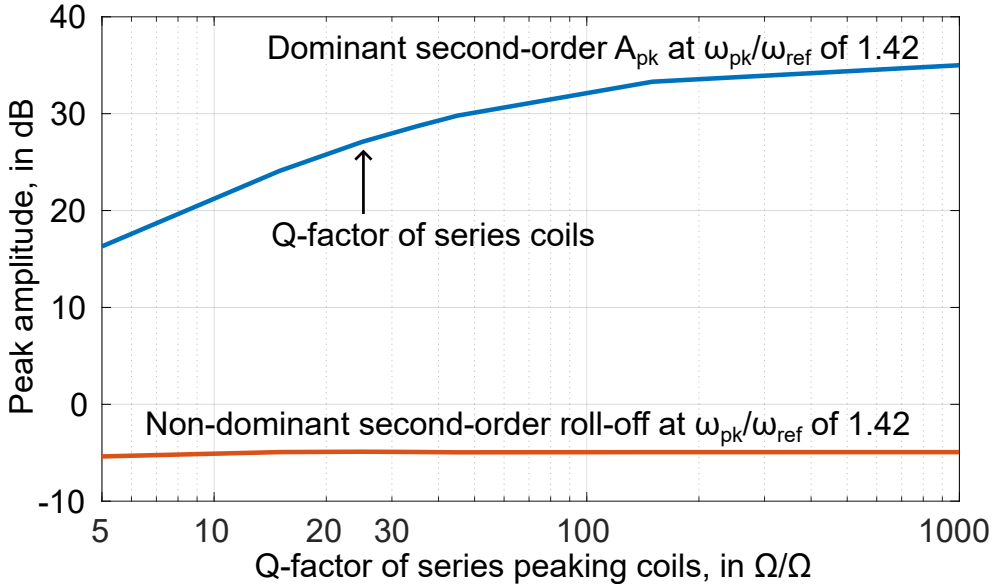


Figure A.1: Peak amplitude from the dominant second-order transfer function (Eq. 3.50) and roll-off from the non-dominant second-order transfer function (Eq. 3.51) against various Q-factor of the series coils at a normalized frequency of 1.42

B Analysis of the 4-pole, double-series interstage network

Recall the 7-pole, 2-zero, shunt/double-series network Fig. 3.1b, the goal is to determine the bandwidth extension ratio and amplitude peaking that may be achieved compared to the unpeaked circuit using a simplified network (i.e., 4-pole, double-series network as shown in Fig. 3.2a) (i.e., at low frequencies, the circuit becomes a parallel R-C network comprised of $R \,||\, (C_1 + C_2)$). The transfer function of the 4-pole, double-series interstage network is given by

$$F(s) = \frac{v_o}{iR} = \frac{1/\alpha}{s^4 + s^3 \cdot \beta/\alpha + s^2 \cdot \gamma/\alpha + s \cdot \delta/\alpha + 1/\alpha} \quad , \tag{B.2}$$

where the denominator polynomial is written in standard form. Coefficients α, β, γ and δ

are

$$\begin{aligned} &\alpha = C_1C_2L_1L_2, \\ &\beta = RC_1C_2L_1 + RC_1C_2L_2, \\ &\gamma = C_1L_1 + C_2L_2, \\ \text{and} \quad &\delta = RC_1 + RC_2. \end{aligned} \tag{B.3}$$

To normalize the transfer function, the frequency is normalized to ω_{ref} of $1/RC_T$ rad/s and the amplitude is normalized to iR V. That is, in the sinusoidal steady state ($s = j\omega$), ω is normalized to ω/ω_{ref}. To simplify the expressions, parameters R and C_T are also normalized to 1 Ω and 1 F, respectively. The following parameters are introduced:

$$C_T = C_1 + C_2 \qquad n = \frac{C_2}{C_T} \qquad m_1 = \frac{L_1}{R^2C_T} \qquad m_2 = \frac{L_2}{R^2C_T}. \tag{B.4}$$

Substituting Eq. B.4 into Eq. B.3 yields

$$\begin{aligned} &\alpha = n(1-n)m_1m_2, \\ &\beta = n(1-n)(m_1+m_2), \\ &\gamma = (1-n)m_1 + nm_2, \\ &\delta = 1. \end{aligned} \tag{B.5}$$

A generalized 4-pole transfer function with complex poles p_1 to p_4 [54] is given by

$$H(s) = \frac{(-1)^4 p_1p_2p_3p_4}{(s-p_1)(s-p_2)(s-p_3)(s-p_4)}, \tag{B.6}$$

and its denominator (D) may be expanded as

$$D = s^4 + s^3 \cdot a_3 + s^2 \cdot a_2 + s \cdot a_1 + a_0, \tag{B.7}$$

where

$$\begin{aligned} a_3 &= -(p_1 + p_2 + p_3 + p_4), \\ a_2 &= p_1p_2 + p_1p_3 + p_1p_4 + p_2p_3 + p_2p_4 + p_3p_4, \\ a_1 &= -(p_1p_2p_3 + p_1p_2p_4 + p_1p_3p_4 + p_2p_3p_4), \\ a_0 &= p_1p_2p_3p_4. \end{aligned} \tag{B.8}$$

By equating the denominator polynomial coefficients of the generalized 4-pole transfer function (Eqs. B.7 and B.8) and the normalized transfer function of the 4-pole, double-series interstage network of Fig. 3.2a (Eqs. B.2 and B.5), we have the following:

$$a_3 = \beta/\alpha \qquad a_2 = \gamma/\alpha \qquad a_1 = \delta/\alpha = 1/\alpha \qquad a_0 = 1/\alpha \qquad \Rightarrow \qquad a_1 = a_0. \tag{B.9}$$

Again, note that the frequency in Eq. B.2 is normalized to ω/ω_{ref}. For an unnormalized case, the relationship expressed in Eq. B.9 may be rewritten as

$$a_1 = \frac{1}{\omega_{ref}}\frac{\delta}{\alpha} = \frac{1}{R(C_1+C_2)}\frac{RC_1+RC_2}{\alpha} = \frac{1}{\alpha} \qquad a_0 = \frac{1}{\alpha} \qquad \Rightarrow \qquad a_1 = a_0. \tag{B.10}$$

That is, a_1 equals a_0 for the unnormalized transfer function. Equating coefficients a_1 and a_0 from Eq. B.8 gives

$$\begin{aligned} -p_1p_2p_3p_4 &= p_1p_2p_3 + p_1p_2p_4 + p_1p_3p_4 + p_2p_3p_4, \\ \text{or} \qquad 1 + \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4} &= 0. \end{aligned} \tag{B.11}$$

Complex poles are utilized to realize efficient inductive peaking as the response becomes underdamped [54]. Assuming that $p_{1,2}$ and $p_{3,4}$ are normalized complex-conjugate pole pairs, they may be expressed in Cartesian form as

$$\begin{aligned} p_1 &= -\sigma_1 + j\omega_1, \\ p_2 &= -\sigma_1 - j\omega_1, \\ p_3 &= -\sigma_3 + j\omega_3, \\ p_4 &= -\sigma_3 - j\omega_3. \end{aligned} \tag{B.12}$$

Equation B.11 therefore has four unknowns: $\sigma_1, \sigma_3, \omega_1$, and ω_3.

To design the 4-pole network shown in Fig. 3.2a, p_3 may be chosen as the dominant pole that determine the peaking frequency (but not the amplitude). Therefore, ω_{pk} equals ω_3 when assuming that σ_3 is small (i.e., p_3 and p_4 are close to the $j\omega$ axis in the s-plane), and σ_3 is much smaller than σ_1 (i.e., poles p_1 and p_2 have minimal effect on the peaking frequency).

The design of the interstage network requires knowledge of the input and output stages parameter (C_1, C_2 and R from Fig. 3.2a). Since the peaking frequency is specified at 75 GHz for this work, the normalized peaking frequency ω_{pk} may be calculated by

$$\omega_{pk} \approx \omega_3 = 75\,\text{GHz} \cdot (2\pi R(C_1+C_2)), \tag{B.13}$$

with the assumptions (made previously) that σ_3 is small and $|\sigma_3| \ll |\sigma_1|$.

The magnitude of the transfer function (with normalized poles) of the generalized 4-pole network (Eq. B.6) is given by

$$\begin{aligned} |F(\omega)| &= \sqrt{F(j\omega)\cdot F^*(j\omega)} \\ &= \frac{(\sigma_1^2+\omega_1^2)(\sigma_3^2+\omega_3^2)}{\sqrt{[\sigma_1^2+(\frac{\omega}{\omega_{ref}}+\omega_1)^2][\sigma_1^2+(\frac{\omega}{\omega_{ref}}-\omega_1)^2][\sigma_3^2+(\frac{\omega}{\omega_{ref}}+\omega_3)^2][\sigma_3^2+(\frac{\omega}{\omega_{ref}}-\omega_3)^2]}}. \end{aligned} \tag{B.14}$$

The magnitude at the normalized peaking frequency ω_{pk} (A_{pk}) may be expressed as

$$\begin{aligned} A_{pk} &= |F(\omega=\omega_{ref}\cdot\omega_{pk})| \approx |F(\omega=\omega_{ref}\cdot\omega_3)| \\ &= \frac{(\sigma_1^2+\omega_1^2)(\sigma_3^2+\omega_3^2)}{\sqrt{[\sigma_1^2+(\omega_3+\omega_1)^2][\sigma_1^2+(\omega_3-\omega_1)^2][\sigma_3^2+(2\omega_3)^2]\sigma_3^2}}. \end{aligned} \tag{B.15}$$

Using Eqs. B.11 and B.15 and the target A_{pk} and ω_{pk}, a set of $p_{1,2}$ locations may be determined.

As an example, a list of design parameters is shown in Table B.1. The desired peak amplitude is set to 2 V/V (6 dB), as the total 9-dB peaking could be accumulated from multiple circuit stages (e.g., input and output matching network). Assuming an interstage load of 30 Ω, driving and load capacitances of 55 fF and 130 fF, respectively, the peaking frequency calculated via Eq. B.13 is 2.6 rad/s. The damping factor is selected at 0.1 (i.e., close to 0) to ensure that the network has an underdamped response. The time-domain metrics of an arbitrary second-order system with the same peak amplitude are used as references (i.e., the reference second-order system has a damping factor of 0.259 and a natural frequency of 2.8, which results in the same 2 V/V peak amplitude at 2.6 rad/s). The reference system has a risetime of 0.46 s, settling time of 5.02 s, delay time 0.41 s and percentage overshoot of 43%. That is, overshoot in the time-domain response is inevitable due to amplitude peaks in the frequency response for an underdamped system. The objective for the 4-pole, double-series network is to keep the percentage overshoot less than the arbitrary second-order system, while achieving bandwidth extension and improvement in risetime.

Table B.1: Design goals for a 4-pole, double-series network design example

A_{pk} (V/V)	ω_{pk}/ω_{ref}	ζ_3	δ_{os} (%)
≥ 2	2.6	0.1	≤ 43

By substituting ω_{pk} and ζ_3 to Eqs. 3.15 and B.13, p_3 and p_4 are found at $-0.26 \pm j2.6$ in the s-plane. Evaluating Eqs. B.11 and B.15 with p_3, p_4 and A_{pk} yields

$$\begin{aligned} \sigma_3 &= 0.26, \\ \omega_3 &= 2.6, \\ 1 + \frac{1}{\sigma_1 + j\omega_1} + \frac{1}{\sigma_1 - j\omega_1} + \frac{1}{\sigma_3 + j\omega_3} + \frac{1}{\sigma_3 - j\omega_3} &= 0, \\ A_{pk} = \frac{(\sigma_1^2 + \omega_1^2)(\sigma_3^2 + \omega_3^2)}{\sqrt{[\sigma_1^2 + (\omega_3 + \omega_1)^2][\sigma_1^2 + (\omega_3 - \omega_1)^2][\sigma_3^2 + (2\omega_3)^2]\sigma_3^2}} &\geq 2. \end{aligned} \tag{B.16}$$

The corresponding locations for poles p_1 and p_2 are solved as

$$\begin{aligned} p_1 &= -\sigma_1 + j\omega_1, \\ p_2 &= -\sigma_1 - j\omega_1, \\ 1.766 \leq \sigma_1 &\leq 2.164, \\ \omega_1 &\approx 0.287\sqrt{\sigma_1(-12.13\sigma_1 + 26.26)}. \end{aligned} \tag{B.17}$$

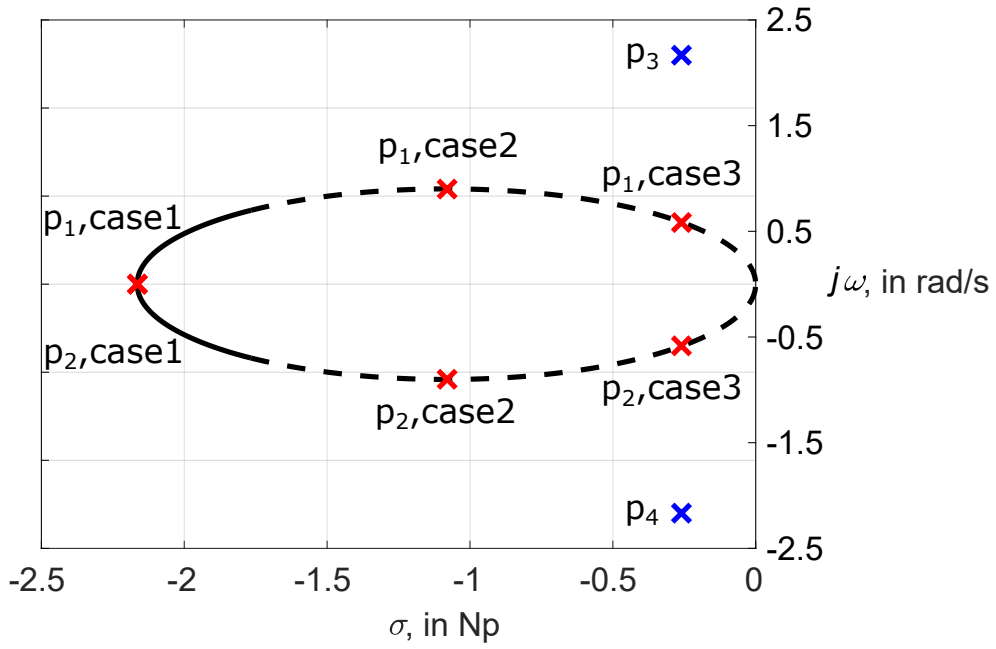


Figure B.2: Pole-zero plot of the 4-pole, double-series network with 3 cases of p_1 and p_2 highlighted (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$)

The locations of p_1 to p_4 are depicted in Fig. B.2. The dominant poles (p_3 and p_4) are shown with blue crosses. All possible solutions of p_1 and p_2 lie on an elliptical trace. The solid line represents ω_1 in the range of $1.766 \leq \sigma_1 \leq 2.164$ (i.e., from Eq. B.17). The dashed line represents the function of ω_1 with σ_1 down to 0 (for the purpose of showing the whole ellipse). Three cases of p_1 and p_2 are highlighted with red crosses in Fig. B.2.

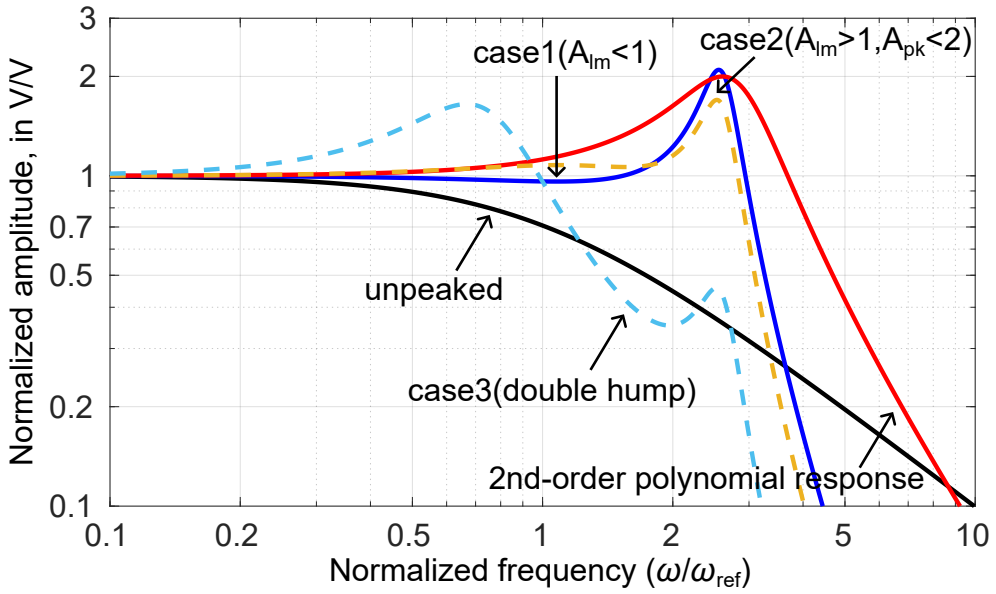


Figure B.3: Normalized magnitude response of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$)

Normalized magnitude responses are depicted in Fig. B.3. The solid black curve shows the unpeaked response with the -3dB bandwidth normalized to 1 rad/s (i.e., $L_1 = L_2 = 0$ and the network becomes an RC circuit). The solid red curve shows a second-order response for reference, which peaks at 2.6 rad/s with an amplitude of 2 V/V and a bandwidth of 4.14 rad/s. Case 1 (solid blue curve, $\sigma_1 = 2.164$ and $\omega_1 = 0.044$) has the highest peak amplitude of 2.078 with bandwidth of 3.091. The peaking frequency is 2.586, which is close to ω_3 of 2.6 (i.e., for Eq. B.13, the assumption of $\omega_{pk} \approx \omega_3$ is valid). Poles p_1 and p_2 have minimal effect on the overall response, as they are far from the dominant poles. Due to frequency separation between poles, an in-band local minima (A_{lm}) of 0.961 (-0.344 dB) is seen at a frequency of 1.094 (i.e., although in Fig. 1.2, the lower mask shows that 1-dB in-band local minima is allowed. Here we see $A_{lm} < 1$ as a drawback). Compared to the second-order response, case 1 rolls off much faster as it is a 4-pole system, resulting in a lower bandwidth.

Case 2 (orange dashed curve) sets ω_1 at max on the ellipse (i.e., $\sigma_1 = 1.08$ and $\omega_1 = 1.08$). The peak amplitude due to p_3 is reduced to 1.686 at the frequency of 2.5 (i.e., no longer meeting the specifications of $A_{pk} \geq 2$). Since p_1 gets closer to the $j\omega$ axis, it introduces another peak, which improves the local minima to 1.063 (i.e., for all frequencies below ω_{pk}, the gain is greather than 1).

Case 3 (cyan dashed curve) has σ_1 equal to σ_3 (i.e., $\sigma_1 = 0.26$ and $\omega_1 = 0.7$). Since the assumption of $|\sigma_3| \ll |\sigma_1|$ is no longer valid, p_1 instead dominates the peaking profile. The first peak has an amplitude of 1.646 at frequency of 0.652, which leads to a bandwidth of 1.153.

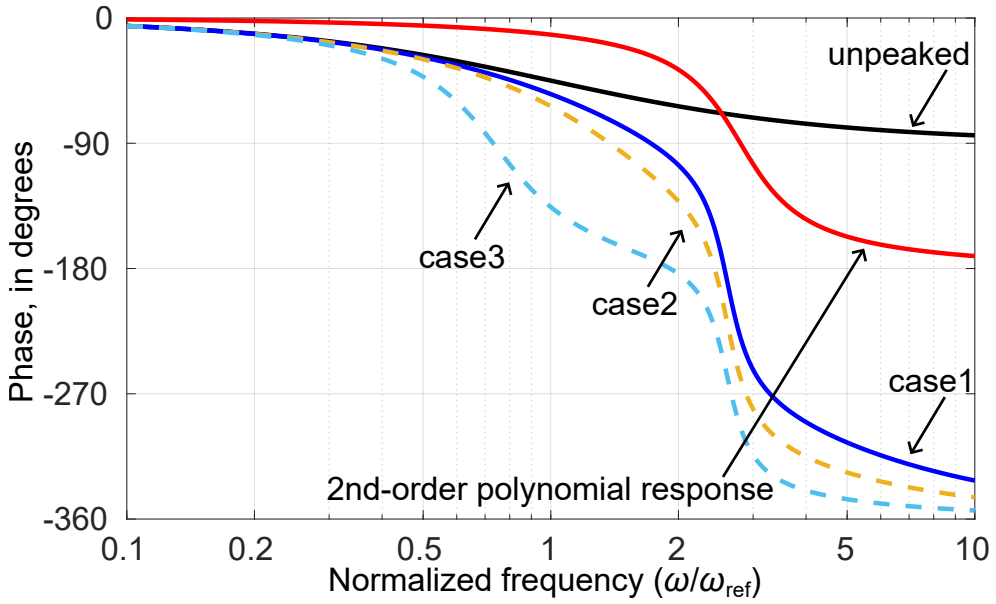


Figure B.4: Phase response of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$)

The phase and group delay of the 4-pole, double-series network are plotted in Fig. B.4 and Fig. B.5, respectively. The unpeaked case (solid black curve) has a slope of -45 degree/decade in phase with a total shift of -90 degrees (i.e., 1-pole low-pass response). In case 1 to case 3, the total phase shift is -360 degrees due to in total 4 poles. In case 1 and case 2, p_3 is the dominant pole for peaking. A steep phase shift is seen near ω_3 of 2.6 in the three cases, leading to a peak in group delay (e.g., group delay greater than 4 seen in Fig. B.5). In case 3, the additional peak introduced by p_1 causes a phase shift near ω_1 of 0.7. Therefore, two peaks in group delay are seen in case 4 (i.e., cyan dashed

curve in Fig. B.5). The second-order system has a total phase shift of -180 degrees for 2 poles. Compared to 4-pole networks, the slope is not as steep, leading to a lower peak in the group delay. Note that the group delay of the arbitrary second-order system is much lower than 1 at low frequency. This is because the system is not derived from a physical circuit with RC time constant of 1 s at low frequency, but rather a mathematical transfer function.

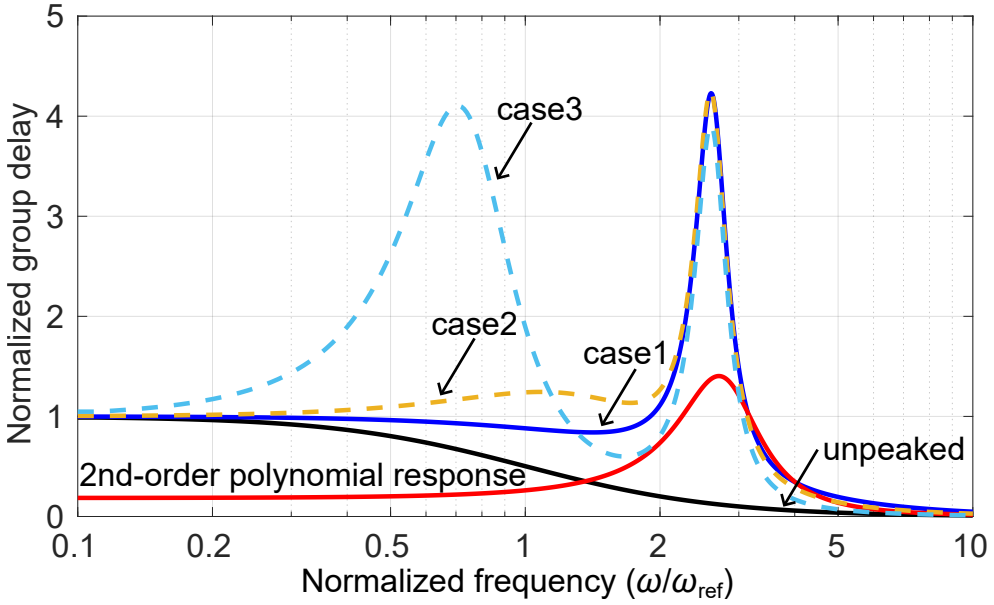


Figure B.5: Normalized group delay of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$)

The step response of the 4-pole, double-series network is shown in Fig. B.6. The unpeaked case (solid black curve) has a normalized 10%-90% risetime, settling time and delay time of 2.2, 3.91 and 0.69, respectively. In case 1 and case 2, the 10%-90% risetimes are improved to 0.805 and 0.872, respectively. That is, an increase in bandwidth leads to a faster risetime in the time domain, as explained in **??**. However, peaks in the amplitude response lead to unwanted overshoot and ringing in the step response. For example, overshoots for case 1 and case 2 are 21.6% and 24.7%, respectively, and settling times for case 1 and case 2 are 11.77 and 10.82, respectively. This implies that p_1 should be separated from p_3 for the larger bandwidth extension, lower risetime and less overshoot. In case 3 where the bandwidth is 1.153, the 10%-90% risetime is 1.421 with an overshoot of 36.3%. That is, moving p_1 close to the $j\omega$ axis degrades the step response compared to case 1 to

case 3 (where p_3 dominates the response at a higher peaking frequency). Compared to 4-pole double series networks, the second-order system shows a much larger overshoot of 43% for the same peak amplitude in the magnitude response (i.e., $A_{pk} = 2$ V/V). Due to the larger normalized bandwidth (4.1 rad/s compared to case 1's 3.1 rad/s), the risetime of the second-order system is faster (0.46). As explained in **??**, series peaking introduces time delay. The delay times of 4-pole, double-series networks are higher than the unpeaked circuit (e.g., case 1 to case 3 have delay times of 1.08, 1.21 and 1.74, respectively).

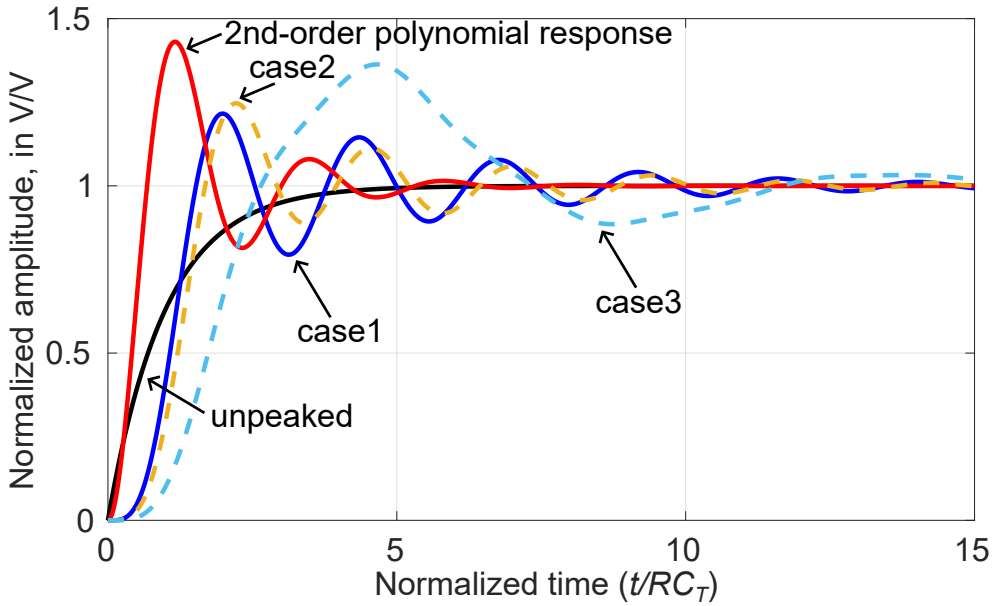


Figure B.6: Normalized step response of 4-pole, double-series networks for various σ_1 (case 1: $\sigma_1 = 2.164$, case 2: $\sigma_1 = 1.08$, case 3: $\sigma_1 = 0.26$)

With the calculated p_1 to p_4 from previous steps, the corresponding circuit parameters can be solved by equating Eq. B.5 and Eq. B.8. Poles for case 1 are used here as an example. Given that

$$\begin{aligned} p_1 &= -2.164 + j0.044, \\ p_2 &= -2.164 - j0.044, \\ p_3 &= -0.26 + j2.6, \\ p_4 &= -0.26 - j2.6, \end{aligned} \tag{B.18}$$

p_1 to p_4 are substituted to Eq. B.8 to solve for coefficients in the 4th-order polynomial.

The coefficients are

$$\begin{aligned} a_3 &= 2(\sigma_1 + \sigma_3) = 4.848 \\ a_2 &= (\sigma_1^2 + \omega_1^2) + (\sigma_3^2 + \omega_3^2) + 4\sigma_1\sigma_3 = 13.763, \\ a_1 &= 2\sigma_3(\sigma_1^2 + \omega_1^2) + 2\sigma_1(\sigma_3^2 + \omega_3^2) = 31.987, \\ a_0 &= (\sigma_1^2 + \omega_1^2)(\sigma_3^2 + \omega_3^2) = 31.987. \end{aligned} \tag{B.19}$$

Matching coefficients a_3 to a_0 with denominator coefficients of the 4-pole, double-series transfer function Eq. B.2 and substituting Eq. B.5 yield

$$\begin{aligned} a_3 &= \frac{\beta}{\alpha} = \frac{m_1 + m_2}{m_1 m_2} = 4.848 \\ a_2 &= \frac{\gamma}{\alpha} = \frac{(1-n)m_1 + n m_2}{n(1-n)m_1 m_2} = 13.763, \\ a_1 &= \frac{\delta}{\alpha} = \frac{1}{\alpha} = \frac{1}{n(1-n)m_1 m_2} = 31.987, \\ a_0 &= \frac{1}{\alpha} = \frac{1}{n(1-n)m_1 m_2} = 31.987. \end{aligned} \tag{B.20}$$

Two solution sets exist. Parameters n, m_1 and m_2 are

$$\begin{array}{lll} n = 0.241 & m_1 = 0.445 & m_2 = 0.384 \\ n = 0.759 & m_1 = 0.384 & m_2 = 0.445. \end{array}$$

The corresponding values of C_1, C_2, L_1 and L_2 can be solved by substituting n, m_1 and m_2 values to Eq. B.4 and scaled up with proper values of R and C_T. Other cases can be solved using the same method.

The drawback of this design approach is that there is a discrepancy between an n of 0.759 and the assumption made in Eq. B.13, where n is 0.7 (i.e., 55 fF and 130 fF are assumed for the driving and load capacitances, respectively). Therefore, parametric tuning is needed for either the interstage network or the driving/load stages to converge to the final design (i.e., multiple iterations are needed).

B.1 Summary of the 4-pole, double-series interstage design methodology

Table B.2: Circuit parameters and performance summary for the 4-pole, double-series network designed for different targets with a peak at normalized frequency of 2.6[1]

Goal	ζ_3	n	m_1	m_2	A_{pk}	ω_{pk}	A_{lm}	η_b	η_r	η_s	η_d	δ_{os} (%)
Maximize A_{pk} @ $\omega_{pk} = 2.6$	0.05	0.8	0.561	0.379	3.888	2.6	0.95	3.11	2.82	0.17	0.64	25.3
Maximize A_{pk} @ $\omega_{pk} = 2.6$	0.1	0.759	0.384	0.445	2.078	2.586	0.961	3.09	2.73	0.33	0.64	21.6
Maximize A_{pk} @ $\omega_{pk} = 2.6$	0.15	0.753	0.316	0.471	1.522	2.511	0.96	3.06	2.68	0.48	0.65	17.7

Circuit parameters and performance are summarized in Table B.2. The gain local minima, bandwidth extension ratio, risetime improvement factor, settling time improvement factor, delay time improvement factor and step-response overshoot are represented by A_{lm}, η_b, η_r, η_s, η_d and δ_{os}, respectively. Various damping ratios of the dominant pole p_3 (ζ_3) are selected to show different achievable A_{pk}.

The design methodology of the 4-pole, double-series interstage network is summarized as follows:

1. Characterize the driving and load stages to obtain C_{in} and C_{out}, which are used as the initial values for C_1 and C_2. Solve for R required for the desired gain. The normalized frequency and gain are $1/R(C_1 + C_2)$ and R, respectively. Circuit parameters are converted to variables n, m_1 and m_2 using Eq. B.4. Based on the design specifications given, calculate the normalized peaking frequency (ω_{pk}) using Eq. B.13.

2. Assuming that frequency of the dominant pole (p_3) ω_3 is equal to ω_{pk}, select a damping factor (ζ_3). The damping factor should be small to ensure that the system has an underdamped response to achieve the amplitude-peak specification A_{pk} (e.g., 2 V/V at ω_{pk} of 2.6 rad/s in the design example). Locations of p_3 and p_4 may be calculated using Eq. 3.15.

[1]Only solutions with $n \geq 0.5$ and positive circuit parameters are listed, since usually the load capacitance is higher than the driving capacitance for a broadband driver, and utilizing negative mutual inductance is not considered

3. Substitute p_3 and p_4 to Eq. B.16 with A_{pk} to solve for σ_1 and ω_1. The values of σ_1 and ω_1 should be related as seen in Eq. B.17 (i.e., p_1 and p_2 are located on an ellipse with a range of σ_1 valid for desired peak amplitude). If no real solutions are found by solving Eq. B.16, repeat step 2 with a smaller ζ_3.

4. Select any p_1 and p_2 from the valid range solved from step 2. With p_1 to p_4 known, substitute them into Eq. B.8 to solve the denominator coefficients of a generalized 4-pole transfer function (i.e., a_0 to a_3). Utilize Eqs. B.5 and B.9 to compute variables n, m_1 and m_2, then translate them back to circuit parameter C_1, C_2, L_1 and L_2.

5. Substitute C_1, C_2, L_1 and L_2 to Eq. B.2 to plot the amplitude, phase, group delay and step response. Check the in-band local minima of the amplitude response to see if it meets the specification. If not, repeat step 4 with p_1 and p_2 brought closer to the $j\omega$ axis along the elliptical curve. This would increase the amplitude of the local minima at the cost of reducing the peak amplitude.

6. Due to the parasitics added during implementation (e.g., C_1 and C_2 will increase due to coil parasitics), steps 1 to 5 should be iterated to converge to the final design.

C Target specification for MPWs

Table C.3 shows changes in target specification of past MPWs. The gain specification received a significant update during the design based on discussions with an industrial partner.

Table C.3: Design specifications used in various prototype tape-outs

Target spec.	MPW2241	MPW2247	MPW2252 (finalized)
$\lvert S_{DD21}\rvert$ (dB)	22	30.7	15.7
Gain peak at 75 GHz (dB)	9	9	9
Bandwidth (GHz)	90	90	90
$\lvert S_{DD11}\rvert$ at 50 GHz (dB)	> 14	> 14	> 14
$\lvert S_{DD11}\rvert$ at 80 GHz (dB)	> 6	> 6	> 6
Diff. OP-1dB across 60-Ω load (dBm)	15.2	15.2	15.2
DC power consumption (mW)	-	-	< 650

D Other small-signal measurement results

The measured differential output reflection coefficient ($|S_{22d}|$) is shown in Fig. D.7. The blue to green curves represent the measured results as backside temperatures varies from 7°C to 75°C. The dashed black curve is the simulated result at 60°C. Note that 60°C is used because $|S_{22d}|$ is taken at the output, which is expected to experience a large rise in temperature due to heat generated by the postamp stage. As the postamp is an open-drain stage (i.e., to be integrated with the MZM), the output matching at low frequencies is expected to be close to 0 dB. The output matching also has a minor temperature dependency. The difference between the simulated and measured results (dip at 93 GHz for measurement versus dip at 95 GHz for simulation) may be due to parasitics of the output pad. For a driver to be integrated with the MZM, such difference in the driver output match is not critical.

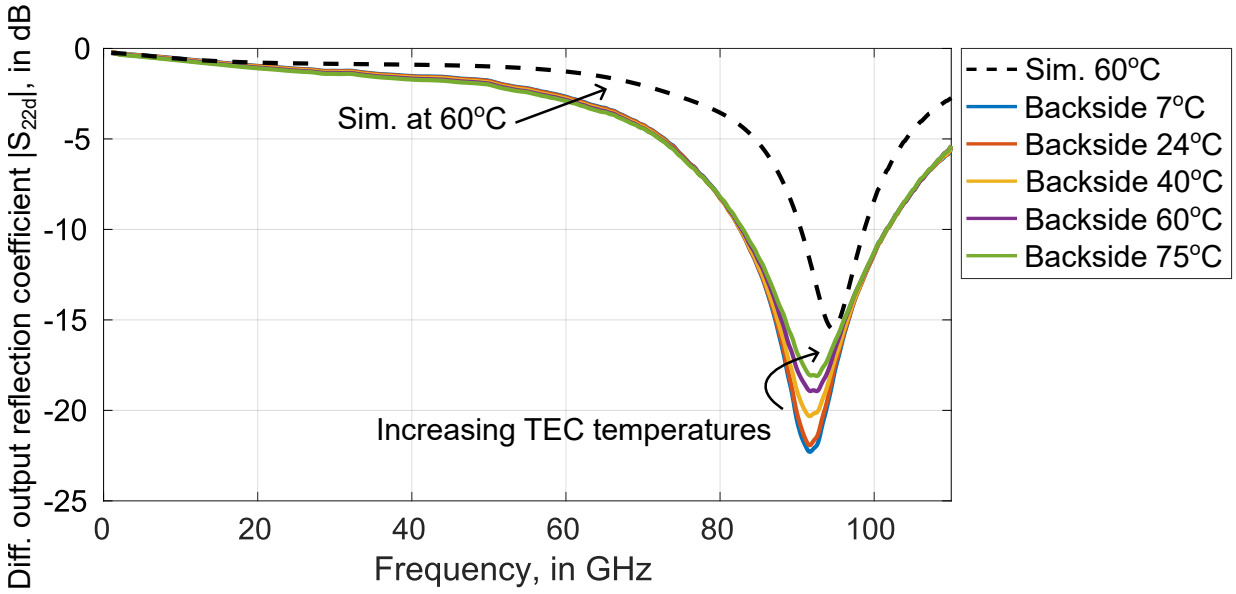


Figure D.7: Measured small-signal differential output reflection coefficient (S_{22d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison

The measured differential reverse isolation ($|S_{12d}|$) is shown in Fig. D.8. Measured results for all TEC temperature settings report an $|S_{12d}|$ of better than -60 dB across all frequencies (i.e., high isolation), and temperature variations do not affect the isolation.

The simulation results are optimistic, as parasitic coupling (e.g., via the external probes and the supply nets) is not included.

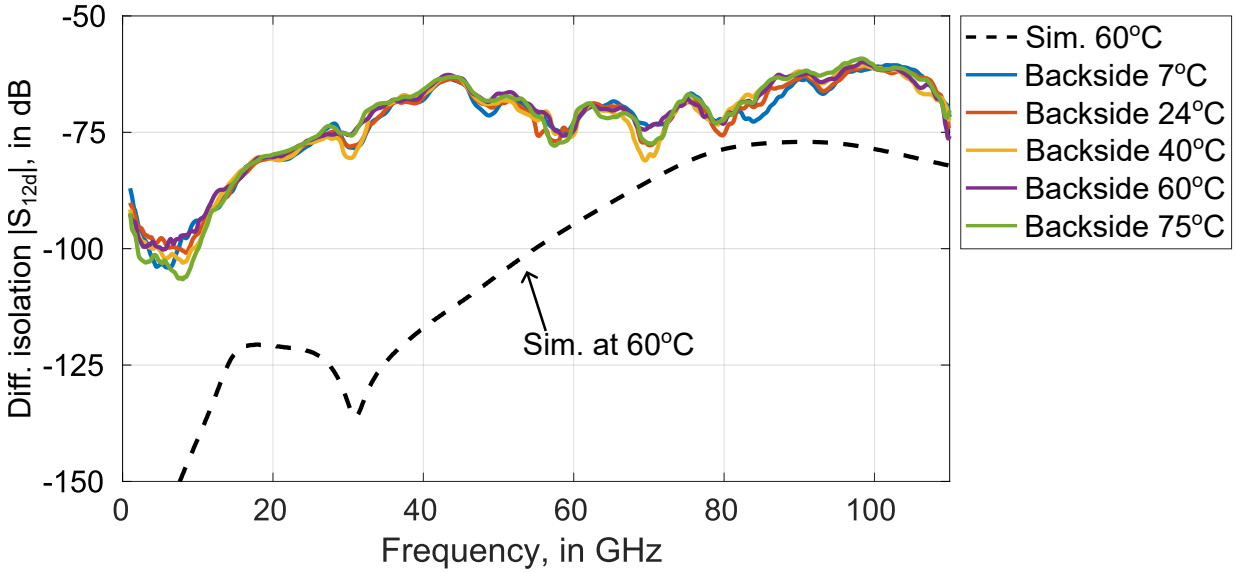


Figure D.8: Measured small-signal differential reverse isolation (S_{12d}) versus various TEC temperature settings for 100-Ω source and 60-Ω load differential impedances (load impedance scaled from the measured results for a 100-Ω load). The simulated result for a temperature setting of 60°C in Spectre is also plotted for comparison

E Measured THD at 10 GHz and 25 GHz

At 10 GHz and 4-$V_{pp,d}$ output swing, the gain compression is 0.4 dB and the THD is 1.6% for all three temperatures, as seen in Fig. E.9. The output 1dB compression point is at 5.5 V, matching the measured OP_{-1dB} of 15.8 dBm from the frequency-domain test (i.e., see Fig. 6.14 at 10 GHz). The THD at OP_{-1dB} is less than 4%.

Linearity of the driver decreases with increasing frequency as predicted from the measured results in the frequency domain. The gain compression increases to 0.7 dB at 4-$V_{pp,d}$ output swing with a 25-GHz input signal, and the THD is 3%, as seen in Fig. E.10. Time-domain measurements reveal an output -1dB compression point of 4.5-$V_{pp,d}$, which is equivalent to 14 dBm for a differential 100-Ω load (i.e., matches the measured result in Fig. 6.14 at 25 GHz). Lower than 3.5% THD is achieved at OP_{-1dB}. All TEC temperature

settings result in similar outcomes for linearity, as the thermal effects become predominant beyond 25 GHz, as seen in Fig. 6.14.

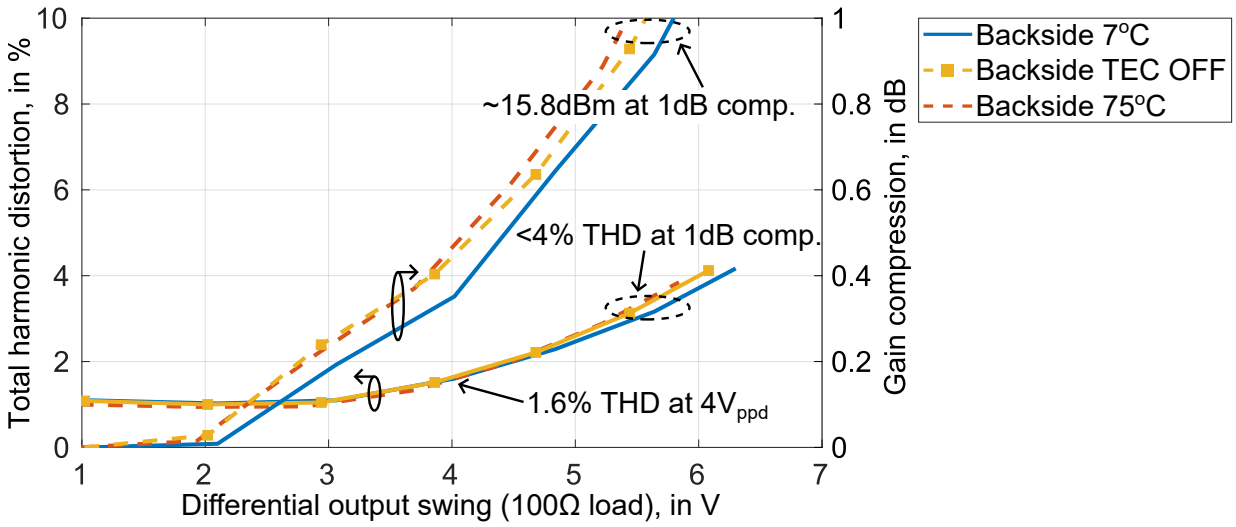


Figure E.9: Measured gain compression and THD at the driver output, for a 10-GHz sinusoidal input signal and differential 100-Ω load at three TEC temperature settings

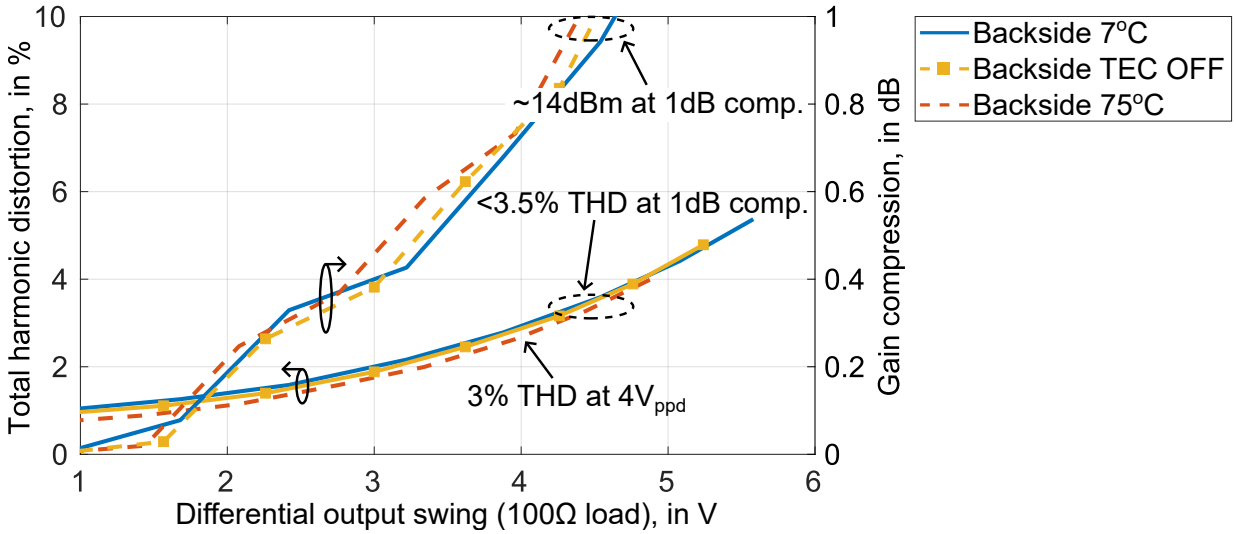


Figure E.10: Measured gain compression point and total harmonic distortion at the output for a 25-GHz sinusoidal input signal and differential 100-Ω load at various TEC temperature settings

F Simulation results of the driver

The postlayout simulation results are summarized in Table F.4. The differential small-signal S-parameters are shown in Figs. F.11 and F.12, which meet the specified gain, peak amplitude, bandwidth, and return losses. The common-mode forward gain ($|S_{CC21}|$) and rejection ratio (CMRR) are depicted in Fig. F.13. The common-mode gain is above 0 dB below 3 GHz, which is due to the pure capacitive impedance in the common mode (i.e., C_{in} in Fig. 4.2). A potential solution is to add decoupling capacitors at the centre tap of input on-chip terminations at the input, which ensures that the input is matched at low frequencies in the common mode. Differential and common-mode μ-stability factors [85] are plotted in Fig. F.14 for 60-Ω and 100-Ω loads. Both factors are greater than 1 up to 100 GHz, which predicts unconditional stability.

Table F.4: Simulated results of the two-stage broadband amplifier with shunt/double-series interstage

Parameter	Target spec.	Sim. results (60-Ω load) (V_{BGN} = 2 V)	Sim. results (100-Ω load) (V_{BGN} = 2 V)	Sim. results (60-Ω load) (V_{BGN} = 0.6 V)	Sim. results (100-Ω load) (V_{BGN} = 0.6 V)
-3dB Bandwidth (GHz)	$\geq$ 90	> 100	99	> 100	> 100
$\lvert S_{DD21}\rvert$ at LF (dB)	15.7	15.4	17.5	15	17.2
$\lvert S_{DD21}\rvert$ peak (dB)	24.7	23.9	22.4	23.9	22.4
$\Delta\lvert S_{DD21}\rvert$ (dB)	9 (60-Ω load)	8.5	4.9	9	5.2
$\lvert S_{DD21}\rvert$ peak frequency (GHz)	75	76	77	77	78
$\lvert S_{DD11}\rvert$ at 50 GHz (dB)	> 14	28	28	28	28
$\lvert S_{DD11}\rvert$ at 80 GHz (dB)	> 8	14	14	14	14
$OP^{1}{}_{-1dB}$ (dBm)	$\geq$ 15.2	15.8 at 75 GHz	14.7 at 75 GHz	14.7[3] at 75 GHz	13.2 at 75 GHz
THD[2] at 1 GHz (%)	$\leq$ 6	3.1 at 4 V_{ppd}	1.3 at 4 V_{ppd}	3.7 at 4 V_{ppd}	1.4 at 4 V_{ppd}
P_{DC} (mW)	$\leq$ 650	590	590	400	400

Harmonic balance simulations characterize large-signal behaviours such as compression points (OP_{-1dB} and IP_{-1dB}) and total harmonic distortion (THD). The OP_{-1dB} remains

[1]The OP_{-1dB} specification is defined across frequencies up to 75 GHz. The required OP_{-1dB} is 15.2 dBm for 4 V_{ppd} across a 60-Ω differential load. During measurement, a 100-Ω load is utilized and a linear 4 V_{ppd} is targeted (i.e., 13 dBm for a differential 100-Ω load).

[2]The simulated THD includes up to the 9th harmonic with a fundamental frequency at 1 GHz. THD is characterized at an output swing of 4 $V_{pp,d}$.

[3]A V_{BGN} of 0.6 V is originally selected to just meet the OP_{-1dB} requirement of 15.2 dBm at 75 GHz, while maximizing peak gain and minimizing DC power consumption. However, due to a change in the post-layout simplification setting, OP_{-1dB} is 0.5 dB below the specification (this could be compensated by adjusting the backgate voltage during measurement).

larger than the required 15.2 dBm up to 75 GHz with a 1-dB margin (for 60-Ω loads) as shown in Fig. F.15. The dip near 75 GHz seen in IP_{-1dB} corresponds to the peak in small-signal gain at the same frequency. The THD depicted in **??** is simulated with a fundamental frequency up to 30 GHz (i.e., the 3rd harmonic is kept within the -3dB bandwidth of the driver). THD at 4-$V_{pp,d}$ and 6-$V_{pp,d}$ outputs are captured for 60-Ω and 100-Ω loads, respectively, and both are lower than 6%. When the output is compressed by 1 dB, THD increases, which is expected.

Performance variations due to back-gate biasing (V_{BGN}) of the postamp are shown in Figs. F.17 and F.18. The small-signal gain and OP_{-1dB} are reduced by 0.5 dB and 1 dB, respectively, and the THD remains below 6% when V_{BGN} decreases from 2 V to 0.5 V while still meeting the specifications. The DC power consumption is 590 mW at V_{BGN} of 2 V. The preamp and postamp consume 85 mW and 505 mW, respectively, as seen in Fig. F.19. At V_{BGN} of 0.5 V, the DC power consumption drops by 33% to 394 mW. The DC power consumed by the preamp remains unchanged at 85 mW as V_{BGN} varies.

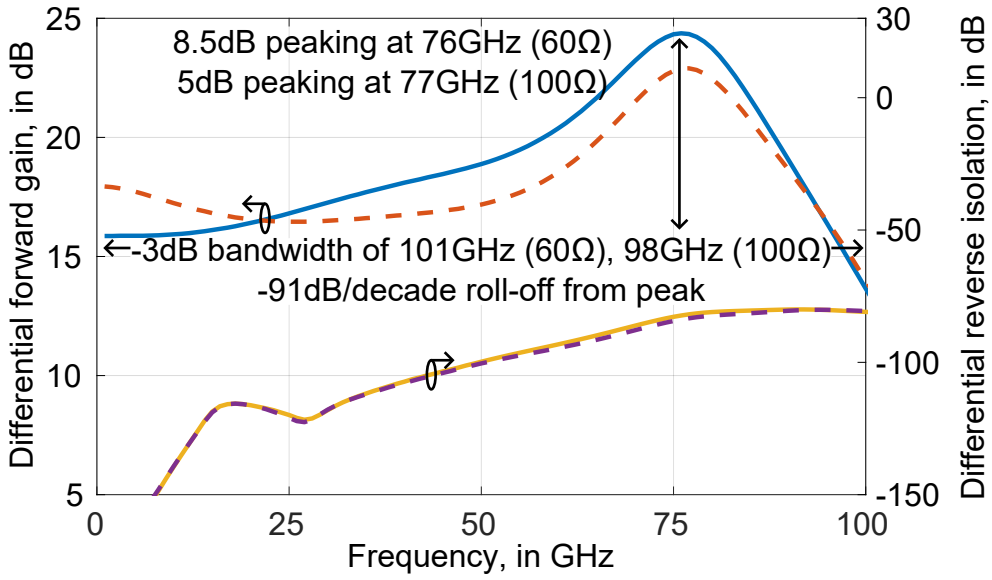


Figure F.11: Simulated small-signal differential forward gain ($|S_{DD21}|$) and reverse isolation ($|S_{DD12}|$) for 60-Ω (solid) and 100-Ω (dashed) loads

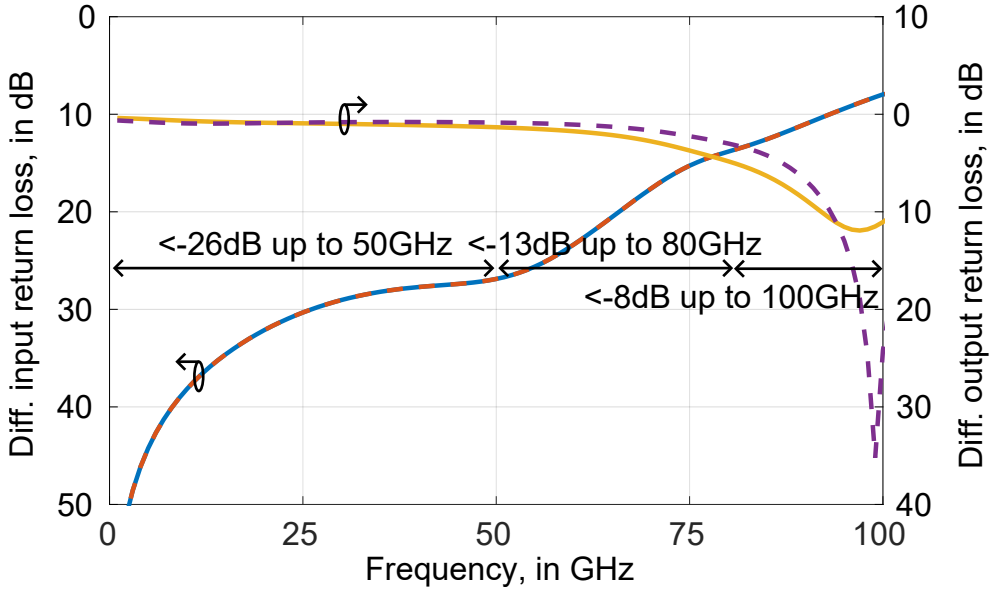


Figure F.12: Simulated small-signal differential input ($|S_{DD11}|$) and output return losses ($|S_{DD22}|$) for 60-Ω (solid) and 100-Ω (dashed) loads

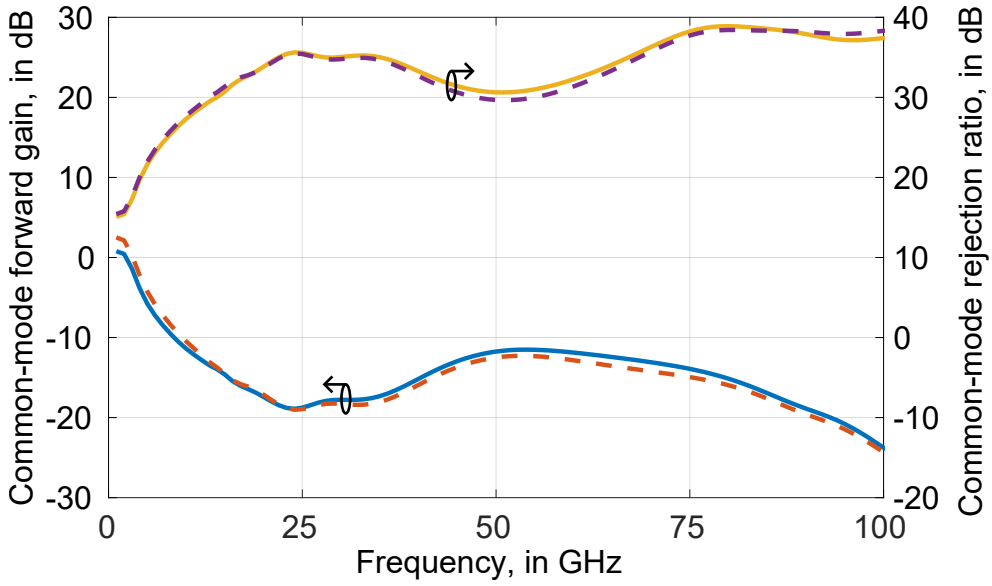


Figure F.13: Simulated small-signal common-mode forward gain ($|S_{CC21}|$) and common-mode rejection ratio (CMRR) for 60-Ω (solid) and 100-Ω (dashed) loads

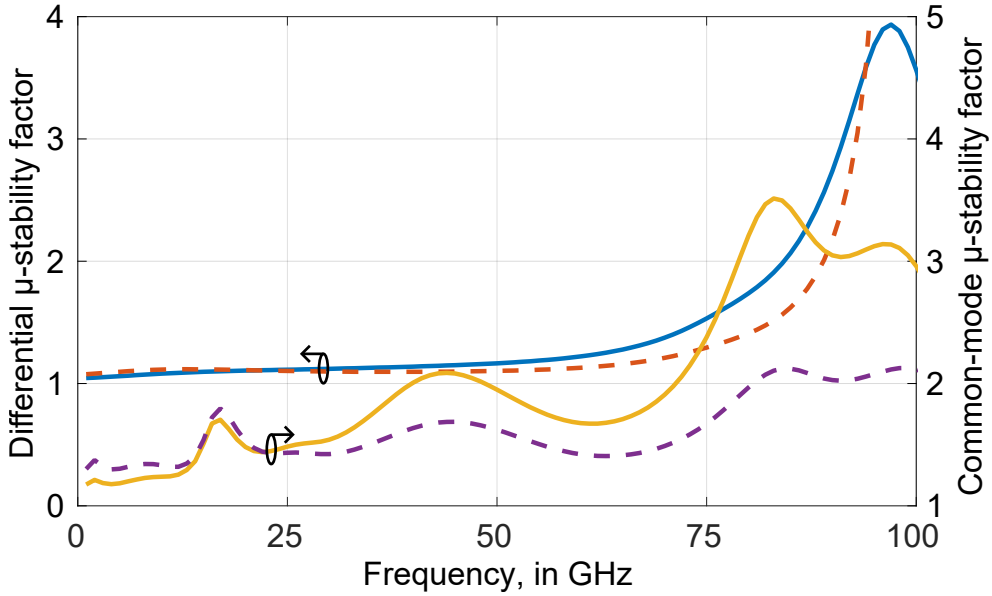


Figure F.14: Simulated small-signal differential and common-mode μ-stability factor 60-Ω (solid) and 100-Ω (dashed) loads

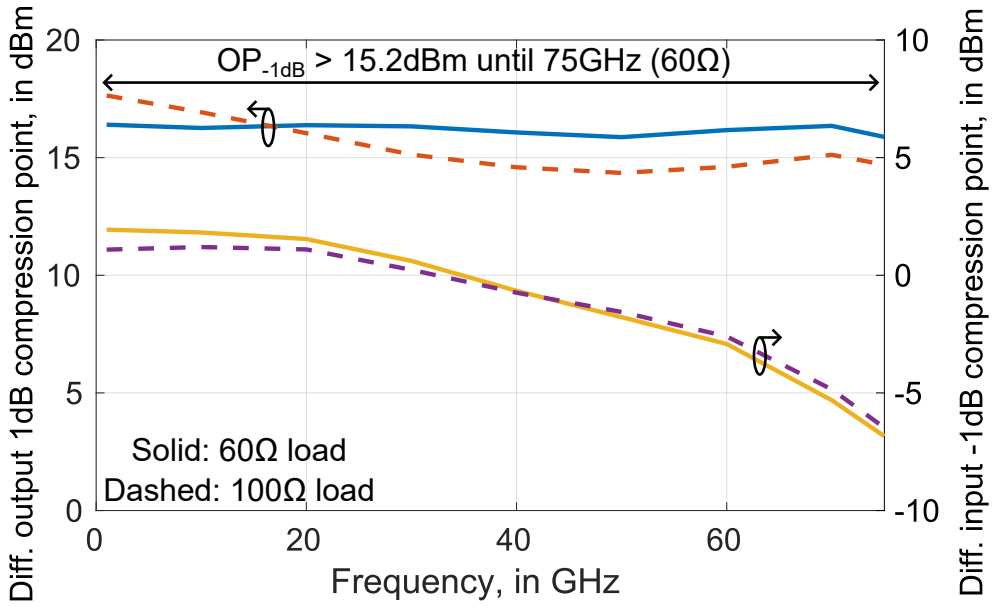


Figure F.15: Simulated differential output and input -1dB compression point 60-Ω (solid) and 100-Ω (dashed) loads

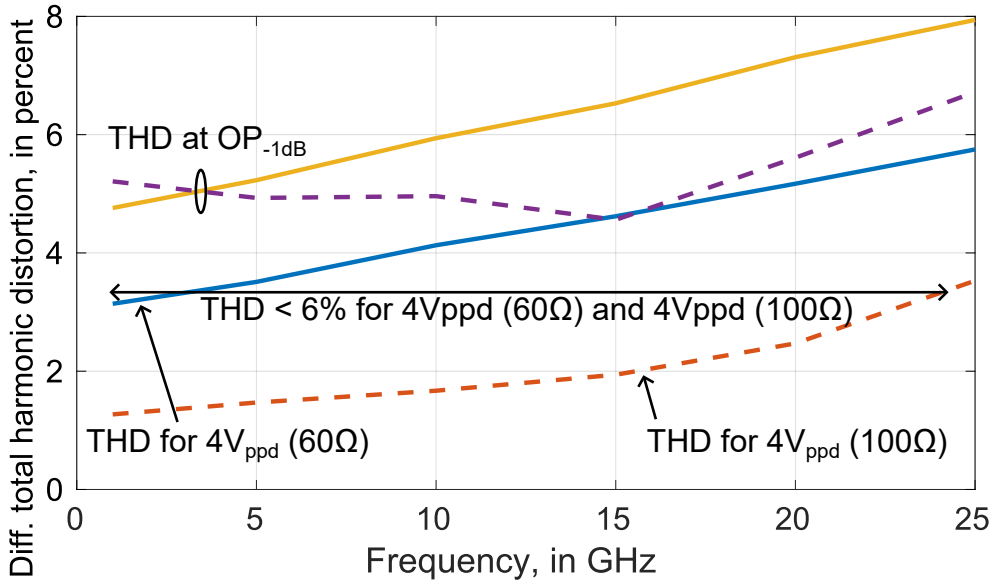


Figure F.16: Simulated differential total harmonic distortion 60-Ω (solid) and 100-Ω (dashed) loads

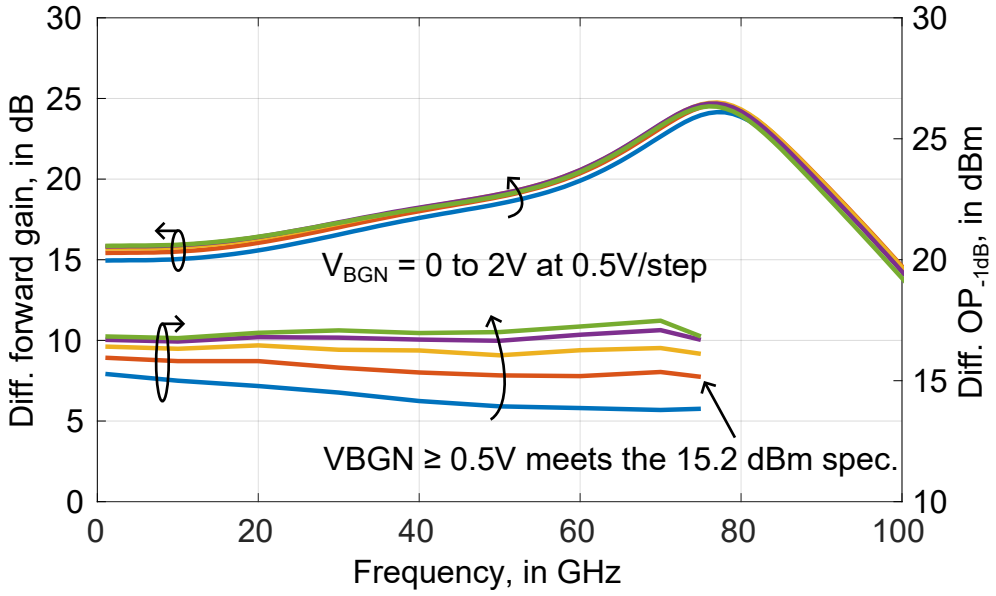


Figure F.17: Simulated differential small-signal forward gain and output -1dB compression for 60-Ω load against various postamp back-gate biasing

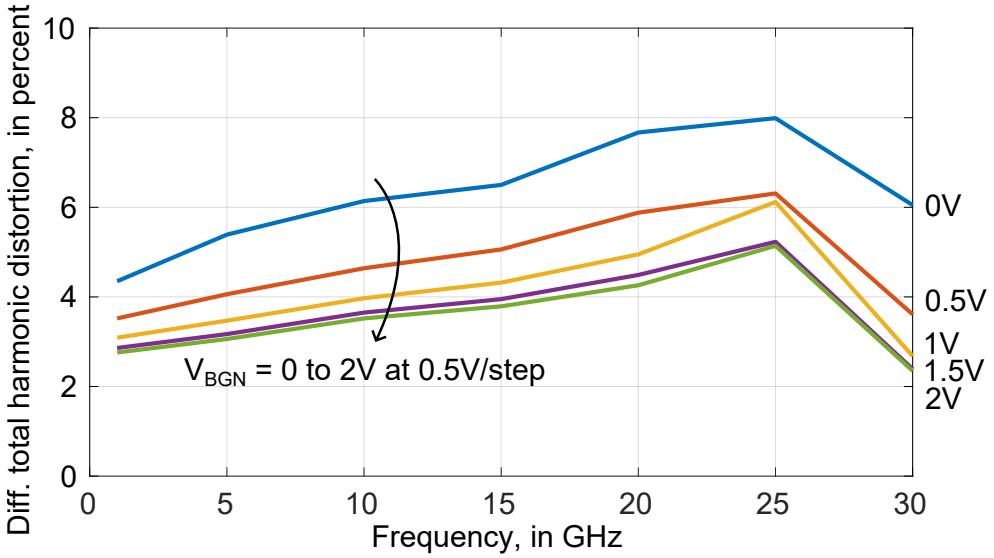


Figure F.18: Simulated differential total harmonic distortion for 60-Ω load against various postamp back-gate biasing

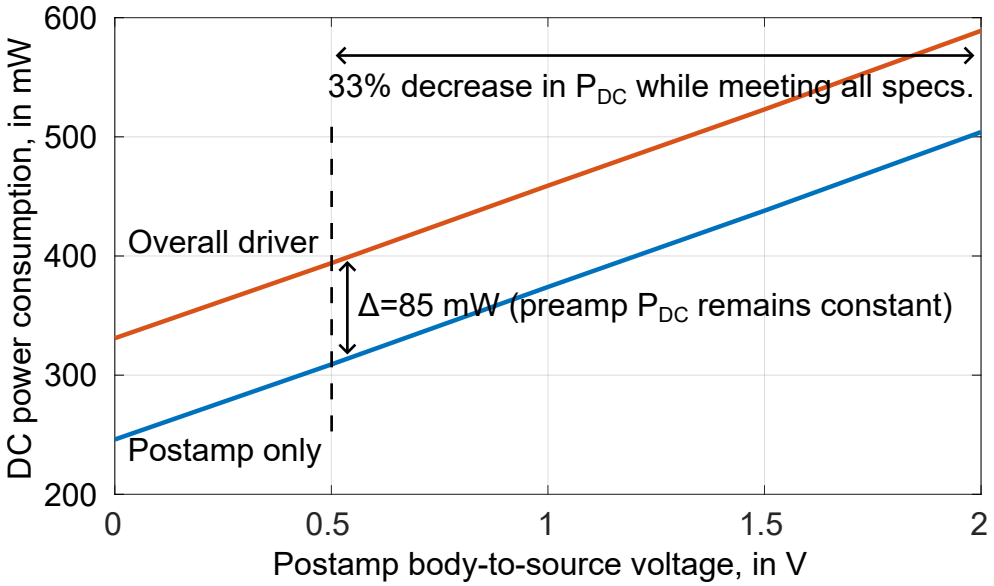


Figure F.19: Simulated DC power consumption of the postamp and overall driver against various postamp back-gate biasing