

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

LO Generation and Distribution for 60GHz Phased Array Transceivers

Permalink

<https://escholarship.org/uc/item/767642t2>

Author

Marcu, Cristian

Publication Date

2011

Peer reviewed|Thesis/dissertation

LO Generation and Distribution for 60GHz Phased Array Transceivers

by

Cristian Marcu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ali M. Niknejad, Chair

Professor Elad Alon

Professor Paul K. Wright

Fall 2011

LO Generation and Distribution for 60GHz Phased Array Transceivers

Copyright 2011

by

Cristian Marcu

Abstract

LO Generation and Distribution for 60GHz Phased Array Transceivers

by

Cristian Marcu

Doctor of Philosophy in Electrical Engineering

University of California, Berkeley

Professor Ali M. Niknejad, Chair

Increased memory capacity and processing power in mobile devices has created a need for radios that can transmit data at multi-Gb/s rates over a short range. However, battery capacity has not kept pace with these advances so power consumption must be kept to a minimum to maintain long battery life. Furthermore, consumer devices require low cost components due to the strong market pressures continuously driving down Average Selling Prices (ASP) leading to diminishing margins. This means a fully integrated solution including RF and baseband components is more attractive than a modular solution.

The allocation of 7GHz of unlicensed bandwidth in the 60GHz band and the increasing speed of CMOS technology provides an excellent opportunity for low cost, high data rate, fully integrated radios to fulfill the unique requirements of modern mobile devices. Phased array transceivers using simple modulation schemes should be used due to their high energy efficiency. Phased arrays use spatial power combining to help overcome the high path loss at 60GHz and also provide beam-steering capabilities which can help to overcome fading issues and create a secure means of communication.

Significant progress has been made recently in the design of mm-wave CMOS building blocks and transceivers, including some phased array transceivers. However, very little attention has been paid to systematic optimization and design of the LO generation and distribution subsystem. In this thesis we use the baseband phase shifting architecture as a vehicle for optimizing LO generation and distribution in phased array transceivers. We propose strategies for optimal low power design with a focus on holistic optimization from architectural choices down to block level design resulting in an optimal and scalable LO distribution methodology. Finally, we present sample designs of building blocks such as oscillators and phase locked loops as well as a full LO generation and distribution subsystem for a 4-element baseband phased-array transceiver in a standard digital 65nm CMOS process.

To my wife Alex, my mom and dad, and my sister Gabi.

I couldn't have done it without your love and support.

Contents

List of Figures	v
List of Tables	x
1 Introduction	1
1.1 The 60GHz Band	1
1.2 CMOS for 60GHz	4
1.3 60GHz Transceivers	5
1.3.1 Link Budget Analysis	7
1.3.2 Phased Arrays	9
1.4 Related Work	12
1.5 Thesis Outline	13
1.5.1 Design Methodology	13
2 Passive Design	14
2.1 Lumped Resonant Tanks	14
2.2 Distributed Resonant Tanks	17
2.3 Tapered Transmission Line Resonators	22
2.4 MEMS Resonators	25
2.5 Passive Components	28
2.5.1 Inductors	28
2.5.2 Capacitors	31
2.5.3 Varactors	34
2.5.4 Transmission Lines	39

2.A	Derivation of Lumped Resonant Tank Bandwidth	44
2.B	Derivation of Distributed Resonant Tank Bandwidth	45
2.C	Series-to-Parallel Transformation	47
3	Voltage Controlled Oscillator	49
3.1	A Short Introduction to Oscillators	49
3.2	Design of a Cross-Coupled Oscillator	50
3.2.1	Startup Conditions	51
3.2.2	Tuning the Tank	56
3.2.3	Phase Noise	59
3.2.4	Design Optimization	64
3.3	Other Fundamental Mode Oscillator Topologies	67
3.3.1	Colpitts	67
3.3.2	Common-Drain Colpitts	73
3.3.3	Differential Versions	77
3.4	Cross-Over Frequency	77
3.5	The Push-Push Oscillator	82
3.6	Design Case Studies	90
3.6.1	Push-push Oscillator Prototype	90
3.6.2	Fundamental Oscillator Prototype	92
3.6.3	Performance Summary and Comparison	96
4	Low Power Phase Locked Loop Design	99
4.1	Phase Locked Loop Dynamics	101
4.1.1	The Linear Phase Domain Model	101
4.1.2	First Order PLL	104
4.1.3	Second Order PLL	106
4.1.4	The Charge Pump and Phase Frequency Detector	107
4.1.5	The Charge Pump PLL	111
4.2	Noise in Charge Pump Phase Locked Loops	114
4.2.1	Noise Contributors	115

4.2.2	Design Optimization	117
4.3	Frequency Dividers	119
4.3.1	Flip-Flop Dividers	119
4.3.2	Injection Locked Dividers	124
4.3.3	Regenerative Dividers	126
4.3.4	Prescalers	127
4.4	Sample Design	131
4.A	Spectral Purity Metrics	138
5	LO Distribution	140
5.1	Mixer LO Requirements	140
5.2	LO Generation Strategy	144
5.3	Mixer LO Buffer Design Methodology	148
5.3.1	Scalable Amplifier Model	149
5.3.2	Scalable Transformer Model	150
5.3.3	Equation Based Buffer Design	151
5.3.4	Optimization Based Buffer Design	156
5.3.5	Comparision Between Buffer Design Methods	158
5.3.6	Injection Locked Oscillator As an LO Buffer	160
5.4	LO Distribution Strategy	163
5.5	Design Case Study	169
6	Conclusion	173
	Bibliography	174

List of Figures

1.1	Attenuation due to molecular resonances in the atmosphere (sea-level, $25^{\circ}C$, $7.5g/m^3$ water vapor density).	2
1.2	Constellations of simple modulation schemes.	2
1.3	Evolution of WLAN data rates.	3
1.4	ITRS Roadmap for RF CMOS Technology.	4
1.5	Direct conversion transceiver block diagram.	6
1.6	QPSK constellation with noisy carrier.	7
1.7	BER as a function of SNR for different modulation schemes.	8
1.8	Uniform linear 8-element phased array transceiver block diagram.	9
1.9	Phased array architectures.	10
2.1	Lumped resonant tanks.	16
2.2	Transmission line with arbitrary load.	17
2.3	<i>RLGC</i> ladder representation of transmission line.	18
2.4	Ideal transmission line input impedance.	19
2.5	Lossy transmission line input impedance (plotted for $Q=10$).	21
2.6	Current and voltage standing waves for a quarter-wavelength transmission line.	22
2.7	A tapered quarter wave transmission line utilizes wide width and large gap spacing when the current is high (voltage is low) and narrow width and small gap when the voltage is high (current is low).	23
2.8	The layout of the optimized quarter wave line. The characteristic impedance, Z_o , is non-constant. Slotting is introduced to satisfy design rules.	24
2.9	The optimum characteristic impedance profile.	25
2.10	MEMS resonator model.	26

2.11	MEMS resonator impedance.	27
2.12	Single turn ring inductor.	28
2.13	Wideband lumped element inductor model.	29
2.14	Simplified inductor model valid over a narrow frequency range.	31
2.15	On-chip capacitor structures.	32
2.16	Wideband lumped element capacitor model.	33
2.17	Simplified capacitor model valid over a narrow frequency range.	33
2.18	Switched capacitor.	34
2.19	Diode varactor.	35
2.20	MOS varactor.	37
2.21	Distributed channel impedance model for MOS varactor.	37
2.22	MOS varactor layout.	39
2.23	On-chip transmission lines.	40
2.24	CPW design space.	41
2.25	CPW design space.	42
2.26	E-fields for the two modes present in the CPW structure.	43
2.27	Series and parallel representations of a complex impedance.	47
3.1	Mechanisms of oscillation.	50
3.2	Cross-coupled differential pair VCO with tuning.	51
3.3	Cross-coupled differential pair input impedance.	52
3.4	Large signal G_m	54
3.5	Current limited vs. voltage limited operation. ($Z_o = 50, Q_T = 10$)	55
3.6	Variable capacitor architectures.	56
3.7	Oscillator output spectrum.	59
3.8	Oscillator LTI noise model.	60
3.9	Phase noise: Leeson's model.	62
3.10	Output waveform and ISF of an ideal sinusoidal oscillator.	64
3.11	Phase noise optimization.	67
3.12	Colpitts oscillator schematic.	68
3.13	Capacitive divider as ideal transformer.	68

3.14 Colpitts oscillator effective model (biasing omitted).	69
3.15 Colpitts startup constraint.	70
3.16 Colpitts oscillator waveforms.	72
3.17 Common-Drain Colpitts Oscillator.	74
3.18 Common-drain Colpitts oscillator provides buffering for free.	75
3.19 Differential Colpitts Oscillators	76
3.20 Transistor small-signal model.	77
3.21 Oscillator small signal models including R_g and C_{gs} .	79
3.22 Performance comparisons of Colpitts versus cross-coupled core. All simulations are performed in a 65nm process node.	81
3.23 Frequency multipliers.	82
3.24 Push-push principle	84
3.25 Fundamental vs. push-push.	85
3.26 Tank quality factor.	88
3.27 60GHz push-push oscillator design space.	89
3.28 Push-push oscillator prototype.	91
3.29 Push-push oscillator measured tuning range.	92
3.30 Push-push oscillator measured phase noise.	93
3.31 Cross-coupled oscillator schematic.	93
3.32 Cross-coupled oscillator die photo. (Area shown: $490\mu m \times 380\mu m$)	94
3.33 Cross-coupled oscillator measured tuning range.	96
4.1 A general phase locked loop.	99
4.2 Phase domain model.	101
4.3 XOR phase detector.	102
4.4 First order loop response.	105
4.5 A first-order loop filter.	106
4.6 Second order loop response (dotted: $\zeta = 1/2$, $PM = 52^\circ$; dashed: $\zeta = 1/\sqrt{2}$, $PM = 65^\circ$; solid: $\zeta = 2$, $PM = 86^\circ$).	108
4.7 Adding an integrator to the loop filter.	109
4.8 A flip-flop based Phase Frequency Detector (PFD).	110

4.9	Charge pump based PLL.	111
4.10	Second order loop filter for charge pump PLL.	112
4.11	Third order Type-II loop response (dotted: $f_p/f_z = 4$, $PM = 37^\circ$; dashed: $f_p/f_z = 16$, $PM = 62^\circ$; solid: $f_p/f_z = 64$, $PM = 76^\circ$).	113
4.12	PLL noise sources.	114
4.13	Selection of PLL bandwidth based on phase noise.	118
4.14	Master-slave flip-flop divide by 2.	119
4.15	Common edge-triggered flip-flop topologies.	121
4.16	CML pulsed-latch divider.	123
4.17	Injection locked dividers.	125
4.18	Regenerative (Miller) divider.	126
4.19	2/3 prescaler.	128
4.20	Program/Swallow Counter.	128
4.21	Vaucher modular prescaler.	130
4.22	Integer-N PLL block diagram.	131
4.23	Schematic of injection locked divider.	132
4.24	Simplified charge pump schematic.	133
4.25	Measured VCO and injection locked divider tuning range. (Measurement of divider tuning range limited by VCO.)	134
4.26	Spectrum of locked PLL at 61GHz downconverted with external mixer to allow measurement with Agilent E4440A Spectrum Analyzer. Reference spurs are less than -40dBc	135
4.27	Measured phase noise at 61GHz for two loop bandwidth settings, measured using Agilent E4440A Spectrum Analyzer	136
4.28	Comparison between measured PLL phase noise and system simulation (in- cluding individual contributors).	137
5.1	Comparison of mixer gain vesus LO amplitude.	141
5.2	Common mixer topologies.	142
5.3	Mixer input admittance.	142
5.4	LO buffer topologies.	143
5.5	LO generation strategies.	145
5.6	LO buffer schematic.	147

5.7	Width-scalable transistor model.	149
5.8	Transformer models.	150
5.9	Model of transformer with effective mixer load.	152
5.10	Matched buffer driven by source with impedance Z_o	155
5.11	Simulated 1 : 2 transformer parameters as a function of L_p	157
5.12	Optimum design of Mixer LO buffer versus mixer switch size.	161
5.13	Current required for an ILO as a function of mixer switch size and loop gain (optimized buffer result added for comparison).	163
5.14	90° Hybrids	164
5.15	Tree distribution networks.	166
5.16	Wilkinson power splitter.	166
5.17	Loss of LO distribution network for a sample 16 element linear array with 250 μm pitch.	168
5.18	Active splitter.	168
5.19	4-element 60GHz phased array transceiver die photo.	169
5.20	Phased array transceiver LO block diagram.	170
5.21	Compact 2-way Wilkinson power divider layout.	171
5.22	Schematic of VCO, LO buffers, and LO distribution chain including Wilkinson power splitters and transformer coupled lumped quadrature hybrid.	172

List of Tables

1.1	Phased array architecture comparison	12
2.1	Frequency-Q product of MEMS resonators.	28
3.1	Sample values of n_{opt} for Colpitts oscillator.	73
3.2	Push-push versus fundamental oscillator selection.	89
3.3	Push-Push oscillator performance summary and comparison	97
3.4	Fundamental oscillator performance summary and comparison	97
4.1	PLL performance summary and comparison	138

Acknowledgments

This work is the culmination of collaborations with many wonderful and talented people and would not have been possible without their contributions. I would first like to thank my advisor Prof. Ali M. Niknejad for his help and support. From teaching, to research, to the job search, he was always available with great advice. I would like to also thank Prof. Elad Alon for invaluable guidance and feedback over many years and numerous projects. I would also like to thank Prof. Paul K. Wright for being a part of my thesis committee and providing valuable feedback.

Graduate school is a significant undertaking with many challenges both personal and professional. I would like to thank two people in particular for taking on these challenges with me and making my time at Berkeley such a memorable experience, Jesse Richmond and Michael Mark. I thank you both deeply for your friendship and I can say without a doubt that meeting and getting to know you was the best part of grad school. Thank you also to my friends from BWRC that always kept life fun and interesting: Mervin John, Louis Alarcon, Debopriyo Chowdhury, Amin Arbabian, and Maryam Tabesh. Thank you also for all of our technical discussions and late night coffee runs to keep us working into the wee morning hours. Last, but definitely not least, I would like to thank all the people that I've had the pleasure to work with over the years and who have provided so much insight and technical advice: Mounir Bohsali, Babak Heydari, Wei-Hung Chen, Zhiming Deng, Patrick Reynaert, Ehsan Adabi, Bagher Afshar, Lingkai Kong, Jiashu Chen, Jungdong Park, and Chintan Thakkar.

Finally, I would like to deeply thank my family for all of their love and support. My mom, Domnica, and my dad, Mihai, have sacrificed so much so I could be where I am today. You have always pushed me to be the best I can be and have supported me no matter what I chose to do. To my sister, Gabi, thank you for always being there for me. We've come a long way from where we started and been all over the world. I hope I can be as helpful and supportive to you as you've been for me. Special thanks have to go to my loving wife, Alex. Your love and support have kept me going in good times and bad and I couldn't have asked for a better woman to share my life with. I thank you from the bottom of my heart.

Chapter 1

Introduction

Increased memory capacity and processing power in mobile devices has led to demands of ever increasing data rate communication in order to enable fast synchronization capabilities for mobile devices as well as media sharing and access. However, battery life is still at a premium in mobile devices, and power consumption must be kept low despite the increased data rate. The demands on these systems results in a need for radios that can transmit data at multi-Gb/s rates over a distance of less than 10m while consuming power on the order of hundreds of mW or less.

1.1 The 60GHz Band

The 60GHz band offers an attractive opportunity to build radios to meet these requirements. The FCC has allocated a license free band of approximately 9GHz from 57GHz-66GHz with similar allocations available around the world. Attenuation due to oxygen in the atmosphere is on the order of 10-20dB/km in this band (Fig. 1.1), while rain can introduce as much as 40-50dB/km of additional attenuation [Liebe81, Smith82]. Due to this high loss, the 60GHz band is not useful for long range communication. At medium range, however, the attenuation provides security and the possibility of frequency reuse since the signal of one transmitter will not interfere with another placed only a few km away. Medium range uses in the past have included point-to-point links over distances of 1-2km with highly directional antennas for applications such as fiber extension or cellular backhaul.

At very short range, on the order of 10s of meters or less, we do not have to contend with atmospheric absorption. However, common building materials are very lossy. Transmitting a signal through a wall, for example, can attenuate it by 40-50dB while a reflection results in 10-20dB of attenuation. This implies that 60GHz links are most efficient in line-of-sight arrangements and will mostly be limited to operation within the same room or floor due to power constraints.

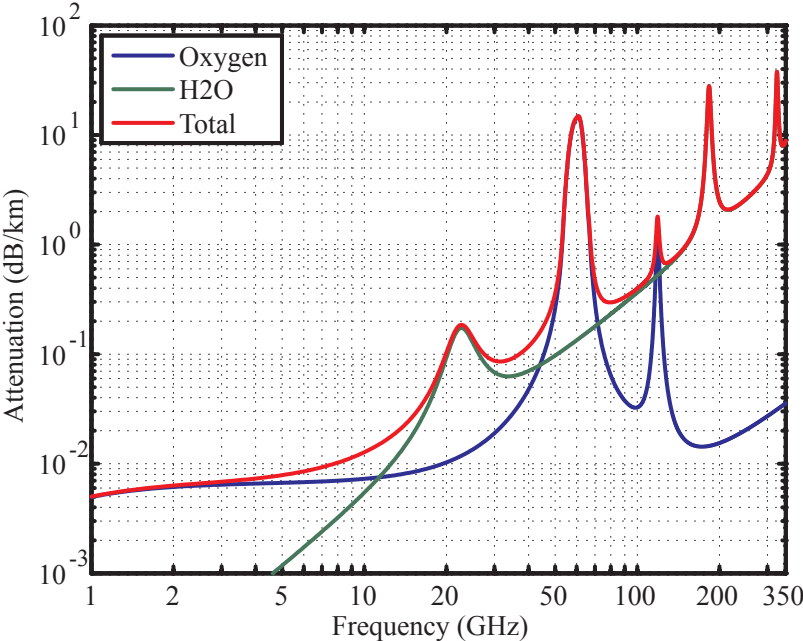


Figure 1.1: Attenuation due to molecular resonances in the atmosphere (sea-level, 25°C, 7.5g/m³ water vapor density).

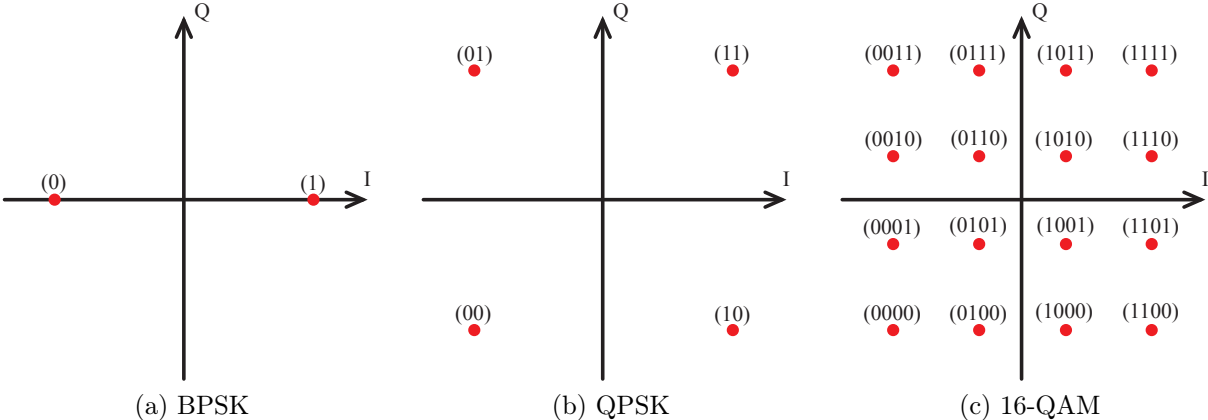


Figure 1.2: Constellations of simple modulation schemes.

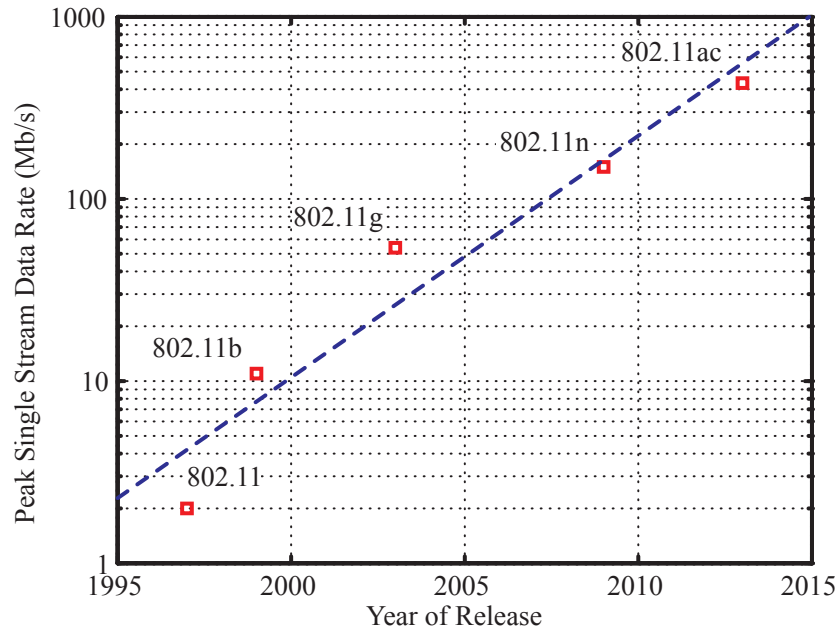


Figure 1.3: Evolution of WLAN data rates.

The large bandwidth allocated around 60GHz means that high data rate communication can be achieved with very simple modulation schemes (Fig. 1.2). These types of simple modulation schemes are not very spectrally efficient but are power efficient since they do not require complicated baseband processing, and the requirements on front-end linearity and carrier spectral purity are relaxed when compared to Orthogonal Frequency Division Multiplexing (OFDM) schemes. Binary Phase Shift Keying (BPSK), and Quadrature Phase Shift Keying (QPSK) are constant envelope modulation schemes. Non-constant envelope modulation such as Quadrature Amplitude Modulation (QAM) can be used to increase data rates further but this type of modulation does require moderate linearity in the front-end and a high purity carrier for low Bit Error Rate (BER) communication. The modulation depth is thus practically limited to 16-QAM for 60GHz transceivers due to these requirements and the need for low power consumption.

Fig. 1.3 shows the evolution of data rates for Wireless Local Area Network (WLAN) radios based on the IEEE 802.11 standard first introduced in [IEE97] for operation in the 2.4GHz ISM band and later ammended to include the much larger 5GHz ISM band as well as data rate enhancements in both bands.¹ Existing WLAN solutions can provide service throughout a house with multiple floors at data rates as high as 150Mb/s. The upcoming 802.11ac ammendment will extend the maximum single stream data rate up to 433Mb/s between transceivers of the type suitable for mobile applications. The 60GHz band is thus well situated to provide short-range, high data rate communication within a line-of-sight environment, augmenting the lower data rates but higher range of legacy WLAN solutions.

¹http://grouper.ieee.org/groups/802/11/Reports/802.11_Timelines.htm

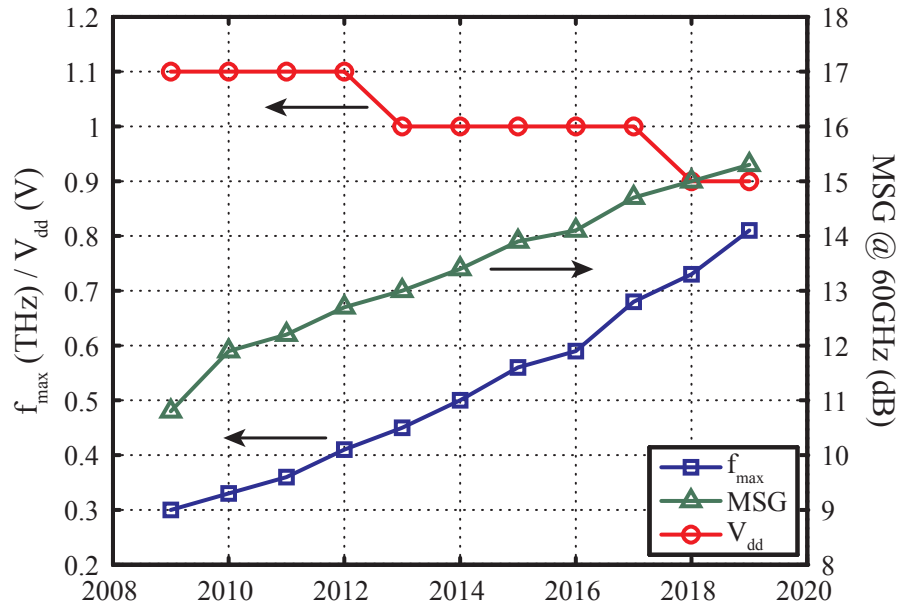


Figure 1.4: ITRS Roadmap for RF CMOS Technology.

Two competing standards are currently making their way through the IEEE standards-making bodies to address this opportunity. One is an amendment to the existing 802.11 standard, dubbed 802.11ad and has a focus on backward compatibility and interoperability with legacy 802.11 devices. The other standard is 802.15.3c [IEE09] and comes from the Wireless Personal Area Network (WPAN) community normally associated with low power and low data rate Wireless Sensor Networks (WSN). Both standards are actually very similar and market forces will likely drive them closer together and toward interoperability. The challenge then from the circuit side is to build a compact 60GHz transceiver to meet the high data rate requirements of these standards with low power consumption and cost since these are the key drivers in consumer mobile applications.

1.2 CMOS for 60GHz

Traditionally, mm-wave design has been limited to expensive III-V compound technologies due to their higher speed, and thus higher gain, compared to Silicon technologies. However, the microprocessor and memory industry has continued to push advances in Silicon, increasing the speed of Complementary Metal Oxide Semiconductor (CMOS) processes year by year (Fig. 1.4).² Today, CMOS technologies have more than enough speed for mm-wave design and the gain available at 60GHz continues to increase. This increase in speed, however, has to come at the expense of a continued reduction in supply voltage. The dynamic power

²<http://www.itrs.net/reports.html>

consumption of a digital gate is given by

$$P_{dyn} = \alpha CV_{dd}^2 f \quad (1.1)$$

where α is the activity factor, C is the total load capacitance, V_{dd} is the supply voltage, and f is the clock frequency. Therefore, for digital signal processing, lowering the supply voltage helps to significantly reduce power consumption. Unfortunately, noise, linearity, and output power all require higher supply voltages to improve. Therefore, analog design gets harder with reduced supply voltages. This means we must be more careful when designing in CMOS and we must make the right architectural choices to enable circuit design at reduced supply voltages. Holistic optimization from architecture down to individual building blocks is a necessity.

At the same time, CMOS provides many advantages over other more exotic technologies. The first of these advantages comes from the low power consumption of digital signal processing in CMOS. This enables complete integration of mm-wave circuits with low frequency mixed-signal circuits and digital signal processing all on the same die, eliminating the need for complex and costly packaging of multiple dies. There is also an inherent operational efficiency in this integration since all signals are processed on the same die so no high frequency and/or high dynamic range IO is necessary.

High speed CMOS digital signal processing also allows built-in-self-test (BIST) capabilities to be integrated on the same die. This reduces test time and test complexity, one of the largest cost centers in semiconductor manufacturing. BIST allows transceivers to self-test and self-calibrate helping to quickly screen out faulty parts or debug problems to increase yield.

Finally, the small wavelength of 60GHz signals also provides advantages for integration of passive components. At 60GHz, the free-space wavelength is approximately 5mm, while the on-chip wavelength is approximately 2.4mm. These dimensions are on the order of typical die sizes for integrated circuits. On one hand this means that distributed effects must be taken into account, making circuit design more difficult. On the other hand, lumped passive components such as capacitors and inductors become very small and easy to integrate. At these frequencies, even antennas can be integrated on-chip thus removing any need to move mm-wave signals on- and off-chip.

1.3 60GHz Transceivers

The block diagram of a typical direct conversion transceiver is shown in Fig. 1.5. It consists of a transmitter path, a receiver path, and baseband which generates data to be transmitted and processes received data. The data generated by the baseband is first converted to an analog signal by the digital to analog converter (DAC) and upconverted to RF by a mixer. The upconverted signal is then amplified by the power amplifier (PA) and sent to

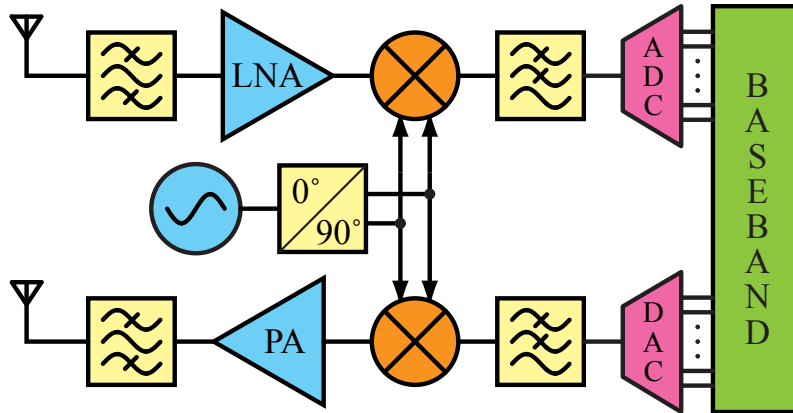


Figure 1.5: Direct conversion transceiver block diagram.

the antenna to be transmitted. On the receive side a low noise amplifier (LNA) amplifies the received signal from the antenna and a mixer converts the signal down to baseband. The resulting signal is then digitized by an analog to digital converter (ADC) and processed by the baseband to extract the information. Both transmit (TX) and receive (RX) paths require a high frequency Local Oscillator (LO) signal to be delivered to the mixer for up and down conversion respectively. This carrier signal is noisy and introduces phase variations to the modulation constellation. For example, Fig. 1.6 shows the effect of carrier phase noise on a QPSK constellation. If the carrier phase noise is too large, the transmitted data cannot be recovered without errors. Therefore, one of the main requirements on the LO generation is low phase noise.

Beside the LO there are other sources of noise which serve to corrupt the received signal. First, an antenna will receive noise from its environment proportional to the observed bandwidth, B .

$$N_{ant} = kTB \quad (1.2)$$

The filters in the RX are thus designed to limit the bandwidth to the desired signal bandwidth to limit this source of noise. Second, all circuits in the receiver chain contribute their own noise due to the presence of active devices and/or resistive losses. For any circuit, the noise factor, F , is defined as the ratio of SNR at its input to SNR at its output.

$$F = \frac{SNR_i}{SNR_o} \quad (1.3)$$

The noise figure, NF , is simply the noise factor converted to dB .

$$NF = 10 \log_{10} F \quad (1.4)$$

The noise factor is always greater than 1 so the noise figure is always greater than 0dB. For a cascade of multiple blocks, each with noise factor F_i and gain G_i , the noise figure of the

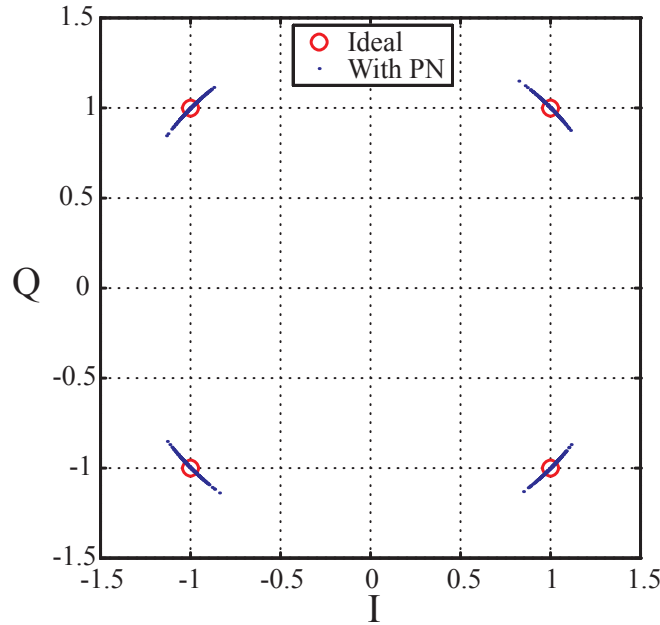


Figure 1.6: QPSK constellation with noisy carrier.

cascade is given by³

$$F_{tot} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \dots \quad (1.5)$$

1.3.1 Link Budget Analysis

The maximum communication distance for a transceiver is limited by the amount of power we can transmit, antenna gains, and noise. For a receiver using an antenna with gain G_r , the power received from a transmitter at a distance d sending power P_t into an antenna with gain G_t at frequency f can be found using the Friis free-space path loss equation [Friis46]

$$P_r = P_t G_t G_r \left(\frac{c}{4\pi f d} \right)^2 \quad (1.6)$$

where c is the speed of light. The path loss is simply the received power normalized to the transmitted power

$$L_{chan} = G_t G_r \left(\frac{c}{4\pi f d} \right)^2 \quad (1.7)$$

Notice that for a fixed antenna gain the path loss gets worse with increasing frequency. To get a feeling for how large this loss can be at 60GHz, we can solve (1.7) assuming 0dBi gain

³Original derivation by [Friis44] but can also be found in [Razavi98, pg. 45] or [Gonzalez97, pg. 298].

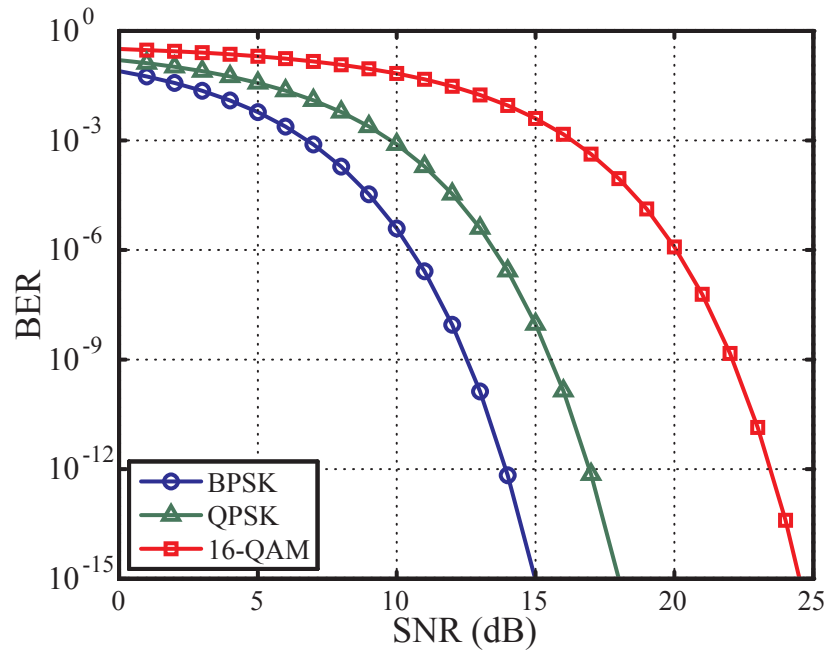


Figure 1.7: BER as a function of SNR for different modulation schemes.

antennas at either end of the link, resulting in a loss of 68dB at a distance of only 1m. This loss assumes a clear line of sight between the transmitter and receiver, so any obstacles or reflections could increase this number dramatically.

Antennas with high gain are expensive and bulky and therefore not compatible with low cost mobile applications. Increasing the maximum communications distance thus requires either increasing the transmitted power or reducing the minimum required received power by increasing the receiver sensitivity. The maximum power that can be transmitted by a PA is limited by fundamental process parameters such as supply voltage and passive component losses so there is a fundamental limit on the output power that can be achieved from a stand-alone transmitter.

On the receiver side, the minimum required input power is determined from the minimum required SNR. The SNR for a given BER depends on the particular modulation scheme. Fig. 1.7 shows the BER versus SNR for three different modulation schemes assuming an Additive White Gaussian Noise (AWGN) channel [Tse05]. A modulation scheme and required BER are first elected and the minimum required SNR is found. The minimum input power to the receiver can then be calculated from the signal bandwidth and the receiver noise figure using

$$P_{in} \geq SNR + NF + 10\log_{10}(kTB) \quad (1.8)$$

Therefore, for a given signal bandwidth and modulation scheme the only way to increase the sensitivity of a stand-alone receiver is to decrease the noise figure. Since there are fundamental limits to how low a noise figure can be, there are fundamental limits to a stand-alone receiver's sensitivity.

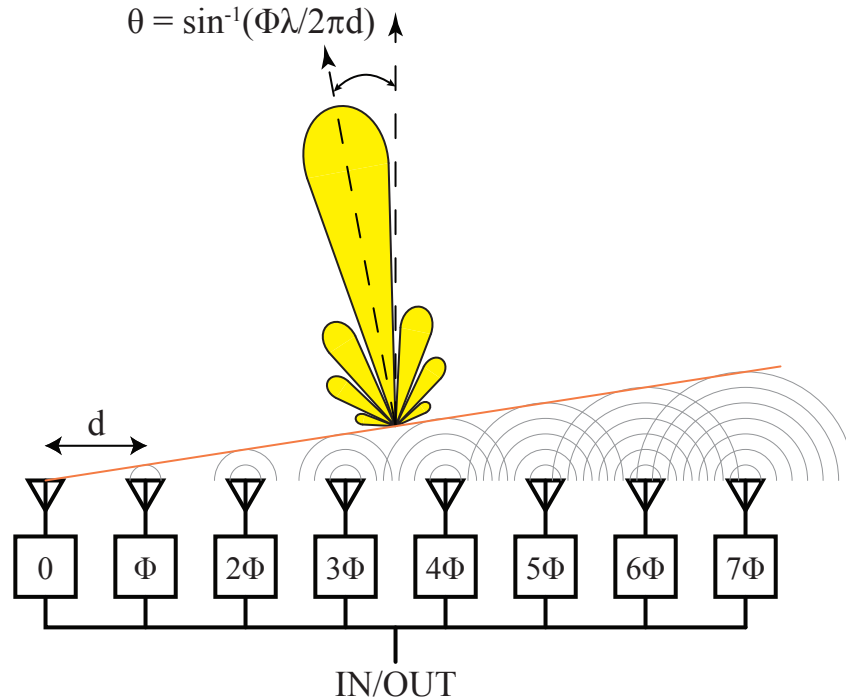


Figure 1.8: Uniform linear 8-element phased array transceiver block diagram.

1.3.2 Phased Arrays

A phased array transceiver (Fig. 1.8) can be used to overcome these fundamental limitations on transmit power and receive sensitivity. Such a transceiver consists of multiple elements, each with its own antenna and phase shifter, positioned in an array with spacing on the order of the wavelength of the carrier frequency. The most common type of array is a uniform linear array with $\lambda/2$ spacing. The antennas can be either directional themselves or omnidirectional, transmitting in all directions. For a low cost 60GHz phased array, planar antennas must be used which are roughly omnidirectional but require very little area.

In the TX, each element transmits the same signal shifted in phase by each element's phase shifter. The transmitted signals then add in space. The phase settings are used to align the transmitted signals from each antenna such that they add constructively only in one direction and cancel each other out in other directions. Most of the energy is thus transmitted in a very narrow beam whose direction can be changed by adjusting the phase shift in each element.

In a very similar way, in the RX, each element receives the same signal (but uncorrelated noise). The phase shifters are then used to realign the received signals such that only signals arriving from a certain direction add up constructively while signals coming from other directions cancel each other out. The array thus has very high sensitivity in one direction and very low sensitivity in all other directions.

A phased array thus behaves just like a mechanically steered, highly directional antenna. The

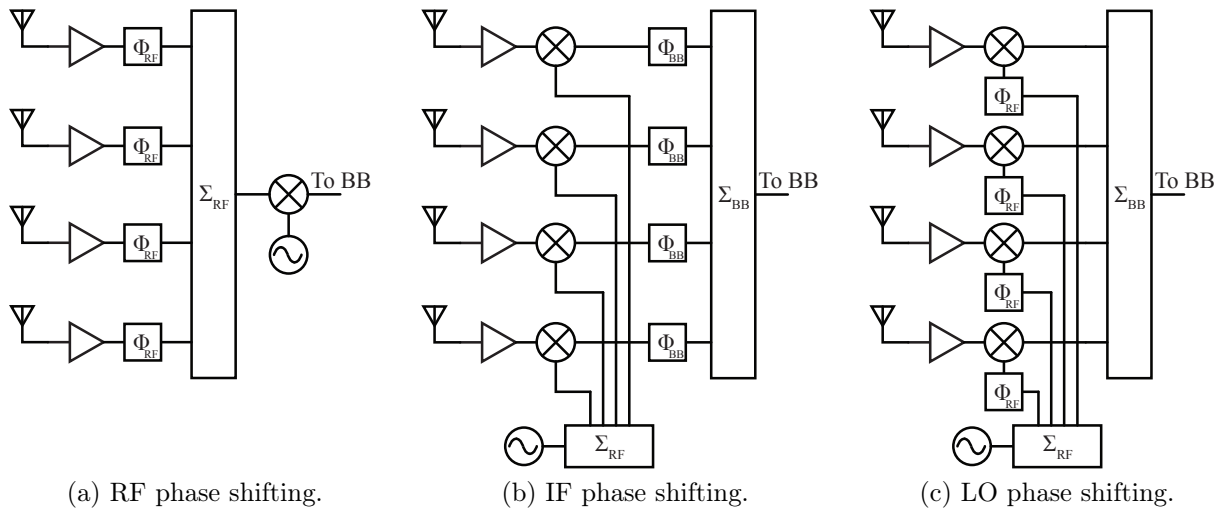


Figure 1.9: Phased array architectures.

advantage is that simple, low cost antennas can be used, and steering can be accomplished very quickly and accurately by adjusting the phase of each phase shifter. A phased array is used either to break the fundamental limitations on transmit power and receive sensitivity of a stand-alone transceiver or to relax the requirements on the performance of each element for a given level of overall array performance. In the TX, the equivalent isotropically radiated power (EIRP) increases with the square of the number of elements, N .

$$EIRP_{array} = N^2 \cdot P_{tx,el} \quad (1.9)$$

In the RX, the received signals are added together after being phase shifted appropriately leading to an N^2 improvement in received signal power from the desired direction. However, each receiver also receives the same level of uncorrelated noise which, when added together, results in an increase in received noise by a factor of N . The array SNR thus only improves by a factor of N over the SNR of each individual receiver element.

$$SNR_{array} = N \cdot SNR_{el} \quad (1.10)$$

Thus, if both the TX and RX use arrays with N elements the link budget is improved by a factor of (N^3) .

The phase shifting operation can be performed in one of three ways. As shown in Fig. 1.9 the phase shifter can be placed either in the RF path, the LO path, or in the baseband (also called IF phase shifting).⁴ Phase shifting in the RF path (Fig. 1.9a) [Natarajan07] requires wide band and low loss phase shifters. Wide bandwidth is required due to the wide signal bandwidth being transmitted, and low loss is required to maintain low noise figure in the RX or high efficiency in the TX. After initial amplification using an LNA, the

⁴Fig. 1.9 shows an RX array but the same ideas apply for a TX array.

signals are combined at RF which means the combiner must also be low loss and wideband. However, only one mixer and LO are needed. Unfortunately, RF phase shifters and power combiners are bulky and have significant loss which is directly in the sensitive signal path in this architecture. Furthermore, it is difficult to achieve well controlled high phase resolution for RF phase shifters.

Phase shifting at baseband (Fig. 1.9b) requires an LNA and mixer in each element and an identical LO to be fed to each element for downconversion. Phase shifting is performed on the downconverted signal using IF phase shifters. Luckily, IF phase shifters are very compact and can have very wide bandwidth and resolution while maintaining low power consumption as shown in [Marcu09]. After phase shifting, the individual signals must be combined. Since this combination is done at baseband simple current summation can be used very effectively. The advantage of this topology is that it allows for very flexible and low power phase shifters and combiners as well as a very modular architecture. Digital signal processing can also be utilized to perform more complex signal manipulations to achieve the desired array behavior. This is not possible with the RF architecture since the signal arriving at the baseband has already been combined. The disadvantages are that each element is in effect a full transceiver and a high frequency signal, the LO, must be split and distributed to each. The bandwidth of the LO path is no longer determined by the signal bandwidth but by the LO tuning range requirements which could be larger. However, the LO path is largely insensitive to amplitude variations so linearity and loss are much less of a concern here as opposed to the signal path (as long as the LO maintains sufficient signal swing).

The final architecture is phase shifting in the LO path (Fig. 1.9c) [Hashemi05, Babakhani06, Natarajan06]. This architecture also requires each element to have an LNA and mixer. However, each mixer is fed an appropriately phase shifted LO signal. The mixing action causes the phase of the LO to get transferred to the downconverted signal. This can be shown by multiplying the RF signal, a carrier with arbitrary amplitude modulation $V_m(t)$ and phase modulation $\theta_m(t)$, with the LO signal, a carrier with phase shift ϕ , as shown in (1.11). Applying a low pass filter removes the upper side-band at frequency $(\omega_{RF} + \omega_{LO})$, leaving only the desired modulation signal with phase shift ϕ , as shown in (1.12).

$$\begin{aligned} V_{IF} &= [V_m(t) \cos(\omega_{RF}t + \theta_m(t))] \cdot \cos(\omega_{LO}t + \phi) \\ &= \frac{1}{2} V_m(t) [\cos(\theta_m(t) + \phi) + \cos((\omega_{RF} + \omega_{LO})t + \theta_m(t) + \phi)] \end{aligned} \quad (1.11)$$

$$V_{IF} \Big|_{LPF} = \frac{1}{2} V_m(t) \cos(\theta_m(t) + \phi) \quad (1.12)$$

In this architecture high frequency phase shifting is required as in the RF architecture and LO distribution is required as in the IF architecture. The only advantage is that signal combination can be performed simply at baseband as in the IF architecture. Unfortunately, this topology has the disadvantages of both previous architectures without providing significant new advantages.

The foregoing analysis is summarized in Table 1.1. While both RF and IF phase shifting architectures provide similar performance, very little attention has been paid to the latter.

	Phase Shift in RF	Phase Shift in IF	Phase Shift in LO
LO distribution	No	Yes	Yes
Phase Shift Frequency	High	Low	High
Phase Shift Bandwidth	Wide	Wide (Scalable)	Narrow
Combiner Frequency	RF	IF	IF
Signal Combining	Early	Late	Late

Table 1.1: Phased array architecture comparison

With advances in CMOS technology allowing high speed and low power baseband phase shifters and signal processing, IF phase shifting is becoming more attractive.

1.4 Related Work

Significant progress has been made in the design of mm-wave circuitry in silicon leading to higher levels of integration in 60GHz transceivers. For example, Floyd et al. described an integrated mm-wave front-end implemented in a SiGe process in [Floyd06]. However, being in a SiGe process makes a fully integrated solution costly since baseband signal processing should be performed in a separate chip in a modern scaled CMOS process for low power consumption. Multiple CMOS 60GHz transceivers achieving Gb/s data rates have also been presented with varying levels of integration, including [Wang07], [Tanomura08], [Pinel08], [Tomkins09], and [Marcu09]. Unfortunately, most of these solutions (except [Marcu09]) are either not completely integrated, missing digital signal processing, and in some cases LO generation and distribution, or give little insight into the design challenges and trade-offs. As explained in the previous section, however, a phased array transceiver is needed for efficient communication at high data rates and the works listed above are single-element transceivers.

More recently there have been an increasing number of demonstrations of 60GHz phased array transceivers. [Reynolds10] and [Valdes-Garcia10b] presented a 16-element phased array RX and TX respectively in a $0.12\mu\text{m}$ SiGe process. While these arrays show excellent performance, they are targeted for wall-powered applications and consume too much power for mobile devices. CMOS based phased arrays have also been shown by [Cohen10] and [Emami11] both utilizing the RF phase shifting architecture. [Emami11] is another array targeted at wall-powered applications and thus consuming too much power for mobile devices. [Cohen10], on the other hand, presents a very low power phased array but at the expense of reduced performance (gain, noise figure, output power).

The RF phase shifting architecture is popular due to its simplicity. Once the signal is phase shifted and combined at RF, the remaining radio (mixer and IF/baseband processing) is identical to the standard, single-element, transceiver shown in Fig. 1.5. However, the baseband phase shifting architecture provides multiple advantages and should be explored

further. Some work has begun to address the particular design challenges of baseband phased arrays [Borremans09, Raczkowski10] but no fully integrated solutions have been shown and, more importantly, there has been no effort to systematically study and optimize the LO distribution. If left unchecked, a sub-optimal design of this subsystem could cause it to be one of the largest power consumers even in a single-element transceiver, as we have shown previously in [Marcu09]. A systematic optimization approach is needed.

1.5 Thesis Outline

This thesis examines the LO generation and distribution challenge for fully integrated, high data rate, 60GHz phased array transceivers utilizing a baseband phase shifting architecture. Strategies for optimal low power design are presented with a focus on holistic optimization from architectural choices down to block level design. Since mm-wave design greatly depends on the quality and performance of passive components, the details of passive design are described first in Chapter 2. Next, the design of low power voltage controlled oscillators (VCO) is explored in Chapter 3. LO generation using a Phase-Locked Loop (PLL) is then described in Chapter 4. Finally, optimal LO distribution strategies for fully integrated phased array transceivers are discussed in detail in Chapter 5.

1.5.1 Design Methodology

Throughout this thesis, the focus will be on reducing power consumption and cost. Therefore, only standard digital CMOS processes are used with no costly analog process options such as Ultra Thick Metal (UTM) layers or Metal-Insulator-Metal (MIM) capacitors. Due to the skin effect, currents at mm-wave frequencies are confined to the conductor surface and so, the additional metal thickness of UTM does not reduce losses significantly. Passive components can be either lumped or distributed (see Chapter 2) and no preference is assumed. Either choice will be shown to be valid depending on the particular circumstances of each design. Nevertheless, all interconnect and passive components are simulated using a full-wave 3D EM simulator such as Ansoft's HFSS in order to carefully account for all parasitic loading and distributed effects. Finally, all transistors use standard design kit models with layout parasitic extraction for extrinsic parasitic resistances and capacitances.

Chapter 2

Passive Design

It is often said that the performance of *mm*-wave circuits comes down to the quality of the passives rather than the active circuits. At low-GHz frequencies, passive design generally involves ensuring matching within arrays of components and proper shielding. Passives at these frequencies are simply lumped components with parasitics. At *mm*-wave frequencies, on the other hand, passives are an integral part of the design process at every step due to distributed effects that arise at such high frequencies. Every trace, every route, every component, both active and passive, must be carefully placed taking into account the resulting parasitics and distributed effects in order to ensure accurate, high performance designs. In this chapter we will describe the design of passive components at 60GHz. We will begin by describing the characteristics of lumped and distributed resonant tanks. Next, we will discuss the implementation of individual passive components used to build these tanks as well as the trade-offs involved in their design. We will focus specifically on the implications of integrating these components on silicon for high frequency applications.

2.1 Lumped Resonant Tanks

A resonant tank for an oscillator can be implemented either using lumped components such as inductors and capacitors, or in a distributed fashion, using transmission lines. In this section we will describe the design using lumped components. There are two types of tanks which are used with different oscillator topologies. The series *RLC* tank (Fig. 2.1a), has an impedance given by

$$Z_{ser} = j\omega L + \frac{1}{j\omega C} + R \quad (2.1)$$

where the resistor, R , represents the series losses present in the inductor and capacitor. We can now define three new terms critical in the following discussion: the resonance frequency,

ω_o , the series quality factor, Q_s , and the tank characteristic impedance, Z_o .

$$\omega_o \triangleq \frac{1}{\sqrt{LC}} \quad (2.2)$$

$$Q_s \triangleq \frac{\omega_o L}{R} \quad (2.3)$$

$$Z_o \triangleq \sqrt{\frac{L}{C}} \quad (2.4)$$

Note that Q_s can be written in multiple, equivalent ways if we make use of ω_o and Z_o .

$$Q_s = \frac{\omega_o L}{R} = \frac{1}{\omega_o C R} = \frac{\sqrt{L/C}}{R} = \frac{Z_o}{R} \quad (2.5)$$

Using the newly defined variables, ω_o and Z_o , (2.1) can be rewritten as (2.6).

$$Z_{ser} = R \left[1 + jQ_s \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) \right] \quad (2.6)$$

Immediately we can see that at the resonance frequency, the imaginary part goes to zero and the impedance is equal to the resistance, R . The impedance is plotted versus frequency, normalized to the resonance frequency, in Fig. 2.1b showing the minimum occurs at resonance as expected. To completely generalize this equation, we can also use Z_o , the characteristic impedance of the tank. Using (2.2) and (2.4), we can show that $Z_o = \omega_o L$, so that (2.6), can be written as (2.7)

$$Z_{ser} = Z_o \left[\frac{1}{Q_s} + j \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) \right] \quad (2.7)$$

Furthermore, some algebraic manipulation (reproduced in Appendix 2.A) can demonstrate that the bandwidth is inversely proportional to Q_s .

$$\Delta\omega_{3dB} = \frac{\omega_o}{Q_s} \quad (2.8)$$

The parallel RLC tank, on the other hand, (Fig. 2.1c), has an admittance equal to

$$Y_{par} = \frac{1}{j\omega L} + j\omega C + \frac{1}{R} \quad (2.9)$$

For the parallel tank, we define the parallel quality factor, Q_p , as

$$Q_p \triangleq \frac{\omega_o C}{G} \quad (2.10)$$

where $G = R^{-1}$. Similar to Q_s , we can also use (2.2) and (2.4) to rewrite Q_p in a number of equivalent ways.

$$Q_p = \frac{\omega_o C}{G} = \frac{R}{\omega_o L} = \frac{R}{\sqrt{L/C}} = \frac{R}{Z_o} \quad (2.11)$$

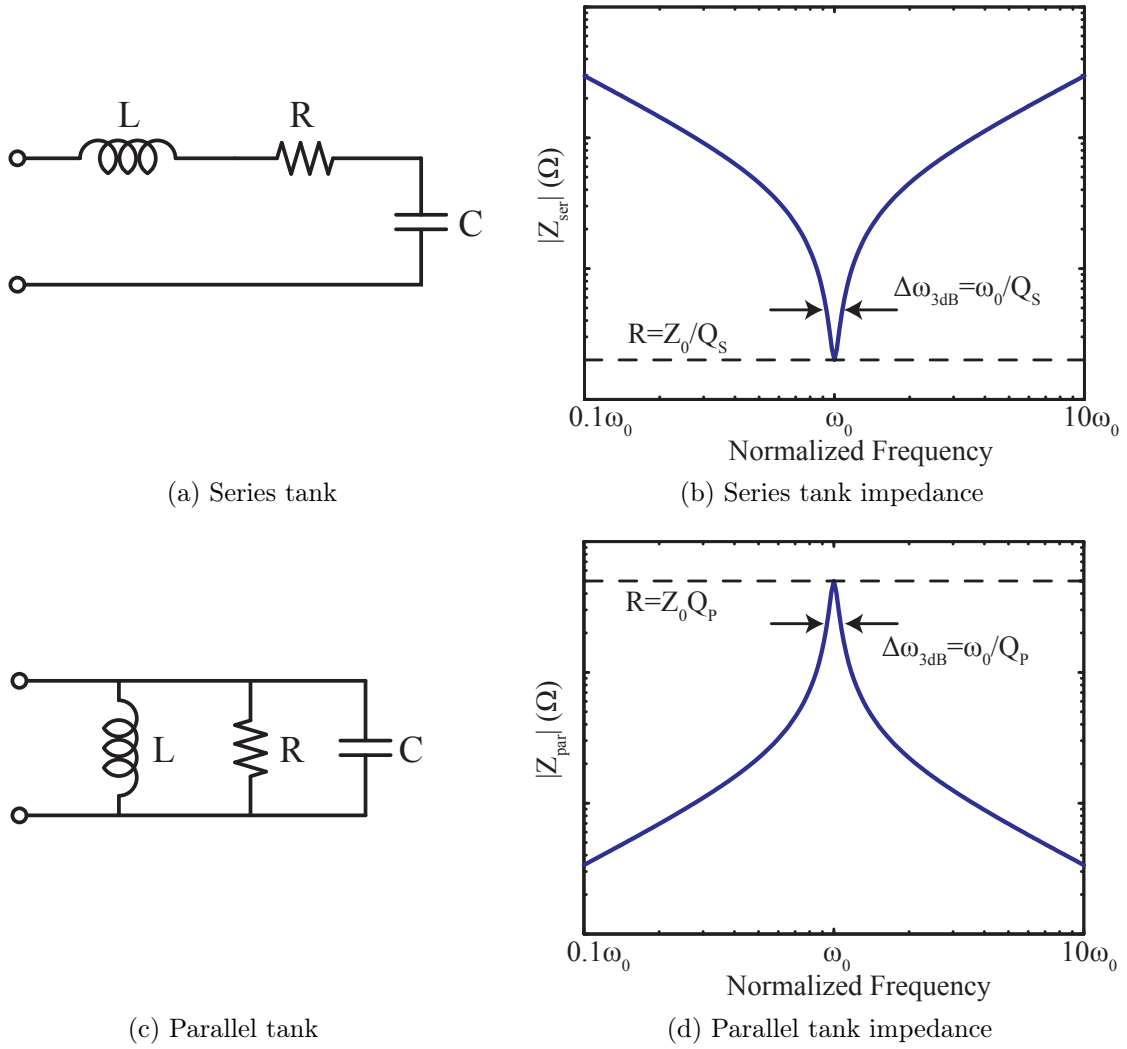


Figure 2.1: Lumped resonant tanks.

Using (2.2) and (2.10), (2.9) can be rewritten as (2.12).

$$Y_{par} = G \left[1 + jQ_p \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) \right] \quad (2.12)$$

Finally, the impedance is simply the inverse of (2.12)

$$Z_{par} = \frac{R}{1 + jQ_p \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)} \quad (2.13)$$

which can also be written in terms of the tank characteristic impedance given by (2.4).

$$Z_{par} = \frac{Z_o}{\frac{1}{Q_p} + j \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)} \quad (2.14)$$

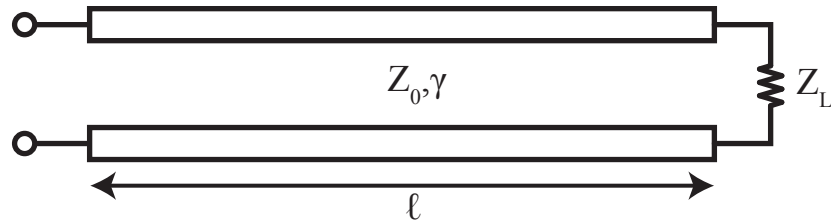


Figure 2.2: Transmission line with arbitrary load.

Once again, at the resonance frequency, the imaginary part goes to zero and the impedance is just equal to the resistance, R . However, in this case, this is the maximum impedance over all frequencies (Fig. 2.1d). Nevertheless, the bandwidth is still inversely proportional to the quality factor.

$$\Delta\omega_{3dB} = \frac{\omega_o}{Q_p} \quad (2.15)$$

In modern scaled CMOS processes, quality factors for integrated LC tanks near 60GHz are generally less than 20, significantly lower for highly tunable tanks due to the higher losses of variable capacitors.

2.2 Distributed Resonant Tanks

The two types of resonant tanks we have discussed thus far are made up of lumped components, however, the same behavior can be emulated with distributed components, namely transmission lines. A transmission line (Fig. 2.2) can be fully described by two complex numbers: the characteristic impedance, Z_o , and the propagation constant, γ . In general, Z_o can be assumed to be real for all practical purposes even for moderately lossy transmission lines. The propagation constant, on the other hand, is generally complex and is defined as

$$\gamma = \alpha + j\beta \quad (2.16)$$

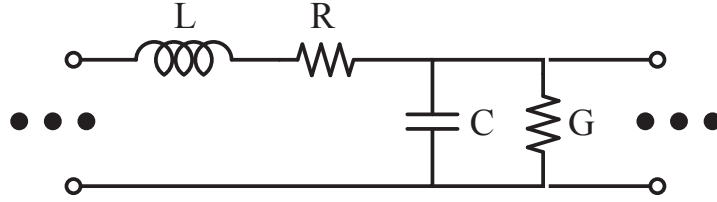
where α is the loss of the transmission line, given in nepers per meter (1 neper \approx 8.686dB), and β is the propagation constant, given in radians per meter, which is a function of frequency and the phase velocity of the transmission line, ν_p ,¹

$$\beta = \frac{\omega}{\nu_p} \quad (2.17)$$

Furthermore, a signal of frequency ω has a wavelength, λ , given by

$$\lambda = \frac{2\pi\nu_p}{\omega} \quad (2.18)$$

¹The phase velocity is the velocity of light in a given medium.

Figure 2.3: *RLGC* ladder representation of transmission line.

This allows us to rewrite β directly as a function of the wavelength.

$$\beta = \frac{2\pi}{\lambda} \quad (2.19)$$

Any transmission line can also be represented by a distributed *RLGC* ladder network shown in Fig. 2.3 which is made up of the series inductance and resistance, and shunt capacitance and conductance, per unit length. The well known Telegrapher's Equations [Pozar04, Collin00] can then be used to describe the voltage and current at any point x along the line.

$$\frac{dV(x)}{dx} = -(R(x) + j\omega L(x)) \cdot I(x) \quad (2.20)$$

$$\frac{dI(x)}{dx} = -(G(x) + j\omega C(x)) \cdot V(x) \quad (2.21)$$

The distributed parameters are related to Z_o and γ by (2.22) and (2.23) respectively.

$$Z_o = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \quad (2.22)$$

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} \quad (2.23)$$

For a lossless line ($R = 0, G = 0$) the above relations reduce to

$$Z_o = \sqrt{\frac{L}{C}} \quad (2.24)$$

$$\gamma = j\omega\sqrt{LC} \quad (2.25)$$

Using (2.17) and (2.25) we can then see that the phase velocity is given by

$$\nu_p = \frac{1}{\sqrt{LC}} \quad (2.26)$$

The input impedance of a loaded transmission line, like the one shown in Fig. 2.2, depends both on the load, Z_L , as well as its length, ℓ , [Niknejad07, pg. 300].

$$Z_{in} = Z_o \frac{Z_L + Z_o \tanh \gamma \ell}{Z_o + Z_L \tanh \gamma \ell} \quad (2.27)$$

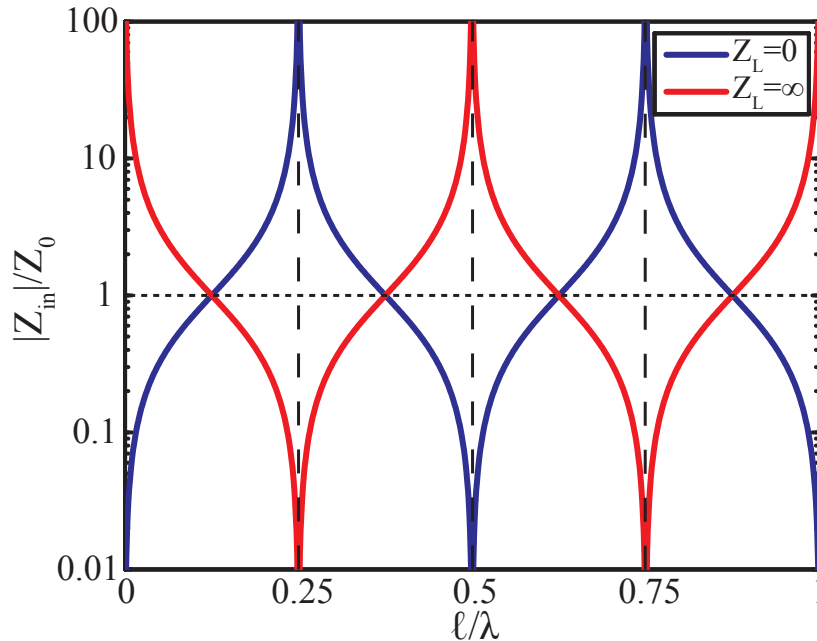


Figure 2.4: Ideal transmission line input impedance.

At the extremes, we can either make the load a short circuit ($Z_L = 0$) or an open circuit ($Z_L = \infty$). For a lossless transmission line γ becomes completely imaginary. So, using (2.19), the input impedance is then given by (2.28) and (2.29) for the two cases respectively (both shown in Fig. 2.4).²

$$Z_{sc,i}(\ell) = jZ_o \tan\left(2\pi\frac{\ell}{\lambda}\right) \quad (2.28)$$

$$Z_{oc,i}(\ell) = \frac{-jZ_o}{\tan\left(2\pi\frac{\ell}{\lambda}\right)} \quad (2.29)$$

From Fig. 2.4 we can see that the absolute value of the input impedance of a short circuited lossless transmission line goes to infinity if the length is equal to $\lambda/4$ (or any odd multiple of $\lambda/4$) and goes to zero for all even multiples of $\lambda/4$. The absolute value of the input impedance of an open circuited lossless transmission line, on the other hand, goes to infinity for lengths equal to even multiples of $\lambda/4$, and zero for lengths equal to odd multiples of $\lambda/4$. In order to keep the area and loss low, we will limit our study to only the smallest possible designs. Conceptually we can see that a shorted line of length $\lambda/4$, or an open line of length $\lambda/2$ looks like an ideal parallel LC tank. Similarly, an open line of length $\lambda/4$, or a shorted line of length $\lambda/2$ looks like an ideal series LC tank.

Unfortunately, the real world is not ideal and transmission lines do exhibit loss. Just like real LC tanks, transmission line tanks cannot provide zero or infinite input impedance. For

²The subscript i has been used to represent the ideal, lossless case.

a realistic, lossy transmission line tank we can define a resonant quality factor, Q_r .

$$Q_r \triangleq \frac{\beta}{2\alpha} \quad (2.30)$$

This definition leads to the same relationship between center frequency and bandwidth as the quality factor of an LC tank given in (2.15). The derivation can be found in Appendix 2.B. To enable comparisons to lumped tanks we will now derive equations for the input impedance of a lossy transmission line resonant tank as a function of frequency, ω , and the resonant quality factor, Q_r . We begin with the general equation for the input impedance of a lossy transmission line given by (2.27) and look at the two special loading cases, open and short.

$$Z_{sc}(\ell) = Z_o \tanh(\gamma\ell) \quad (2.31)$$

$$Z_{oc}(\ell) = \frac{Z_o}{\tanh(\gamma\ell)} \quad (2.32)$$

Let us assume that we are designing for a particular resonance frequency, ω_o , which has an associated wavelength, λ_o , given by 2.33.

$$\lambda_o = \frac{2\pi\nu_p}{\omega_o} \quad (2.33)$$

Next, the resonant quality factor, $Q_{r,o}$, for our transmission line at the design frequency, ω_o , is given by

$$\begin{aligned} Q_{r,o} &= \frac{\beta_o}{2\alpha} \\ &= \frac{\omega_o}{2\alpha\nu_p} \end{aligned} \quad (2.34)$$

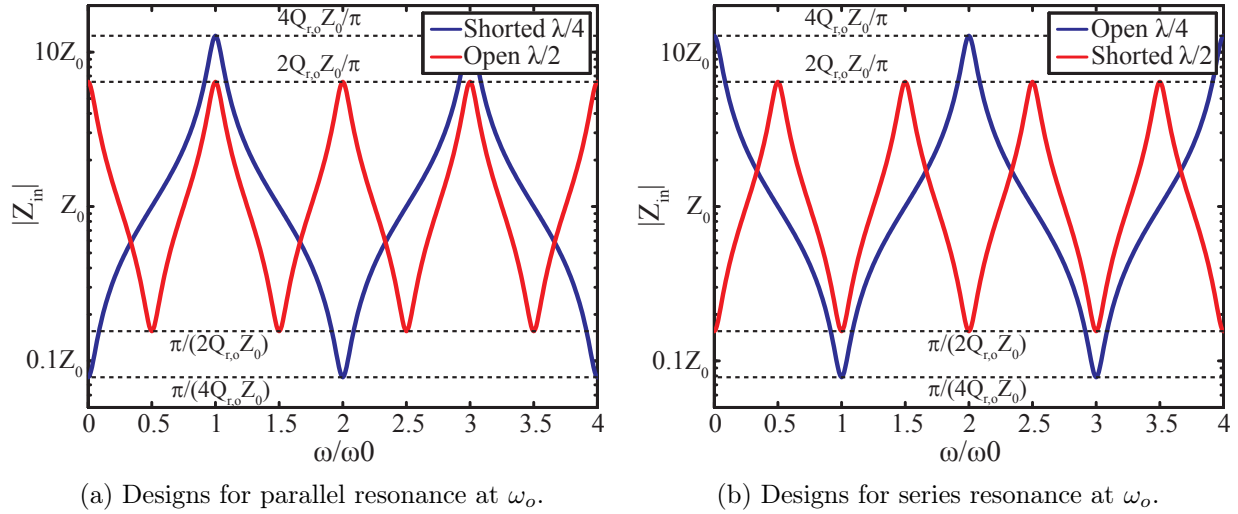
Finally, we normalize our transmission length, ℓ , to the design wavelength, λ_o , calling the result, ℓ_n , which leads to

$$\begin{aligned} \ell &= \lambda_o \ell_n \\ &= \frac{2\pi\nu_p \ell_n}{\omega_o} \end{aligned} \quad (2.35)$$

We can now use (2.17) and (2.34) to rewrite γ as

$$\begin{aligned} \gamma &= \alpha + j\beta \\ &= \frac{\omega_o}{2Q_{r,o}\nu_p} + j\frac{\omega}{\nu_p} \end{aligned} \quad (2.36)$$

assuming that the loss is constant across frequency. This is not strictly correct as will be described later but for our current purposes it is a reasonable and useful simplification. The

Figure 2.5: Lossy transmission line input impedance (plotted for $Q=10$).

$\tanh(\gamma\ell)$ term in (2.31) and (2.32) can now be rewritten using (2.35) and (2.36).

$$\begin{aligned}
 \tanh(\gamma\ell) &= \tanh\left[\left(\frac{\omega_o}{2Q_{r,o}\nu_p} + j\frac{\omega}{\nu_p}\right)\frac{2\pi\nu_p\ell_n}{\omega_o}\right] \\
 &= \tanh\left[2\pi\ell_n\left(\frac{1}{2Q_{r,o}} + j\frac{\omega}{\omega_o}\right)\right]
 \end{aligned} \tag{2.37}$$

Putting this term back into (2.31) and (2.32) yields our final simplified expressions (plotted in Fig. 2.5).

$$Z_{sc}(\omega) = Z_o \tanh\left[2\pi\ell_n\left(\frac{1}{2Q_{r,o}} + j\frac{\omega}{\omega_o}\right)\right] \tag{2.38}$$

$$Z_{oc}(\omega) = \frac{Z_o}{\tanh\left[2\pi\ell_n\left(\frac{1}{2Q_{r,o}} + j\frac{\omega}{\omega_o}\right)\right]} \tag{2.39}$$

There is, however, one important distinction between transmission line resonant tanks and their lumped equivalents. As we can see from (2.38) and (2.39), the \tanh term gives rise to a periodicity in the frequency response. Thus, a transmission line tank of a given length will exhibit multiple resonances at frequencies for which the line length is a multiple of $\lambda/4$. In practice this means that a given tank will resonate not just at the designed frequency but also at its harmonics (Fig. 2.5). However, this is generally not a concern for mm-wave designs since active devices exhibit very low gain at the higher harmonics. Furthermore, the loss mechanisms present in transmission lines designed on silicon tend to increase with frequency. Metal loss increases with the square root of frequency due to the skin effect, while dielectric loss, due to the loss tangent, increases proportional to frequency. In practice,

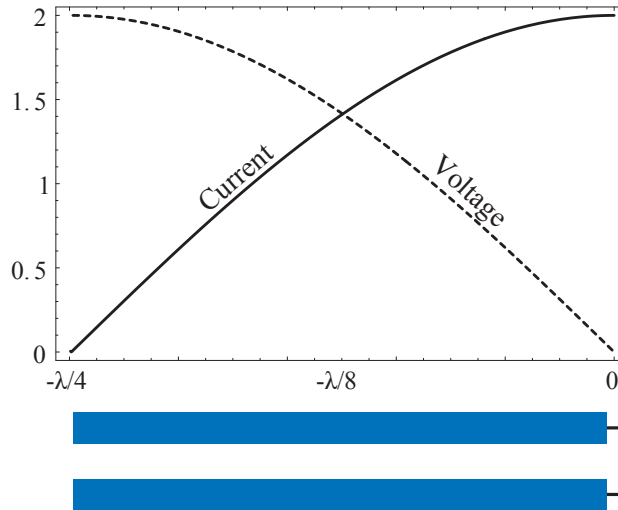


Figure 2.6: Current and voltage standing waves for a quarter-wavelength transmission line.

startup conditions will usually only be met at the fundamental frequency but not at its harmonics so no parasitic oscillations will occur. Nevertheless, this effect should be known to the designer and care must be taken to ensure that oscillation can only occur at the desired frequency by design rather than chance.

2.3 Tapered Transmission Line Resonators

Transmission line resonators exhibit another interesting property, the standing wave. This effect can be exploited to reduce loss through tapering of the transmission line. We will use the shorted $\lambda/4$ resonator as our example but the analysis can also be applied to other lengths and load conditions. Since many applications require a differential oscillator design, let us use a coplanar stripline (described below in Section 2.5.4).

When a shorted length of transmission line is driven by a sinusoidal signal, the signal travels down the transmission line to the load. Upon reaching the load it is completely reflected back towards the source due to the impedance mismatch caused by the short. The reflected wave then travels back toward the source and the total signal along the line is the superposition of the incident and reflected waves. This superposition forms a standing wave with wavelength equal to λ [Gonzalez97]. If we were to measure the voltage at any point along the transmission line we would observe a sine wave of constant amplitude. At different points along the line, the amplitude would vary but its phase would not.

A $\lambda/4$ line then holds exactly one quarter of the standing wave. To find the resulting current wave we can decouple the Telegrapher's equations (2.20) and (2.21) which gives us a second-order system of differential equations. Assuming low loss conditions ($R = G = 0$) the result is the differential equation representation of an arbitrary transmission line in the voltage

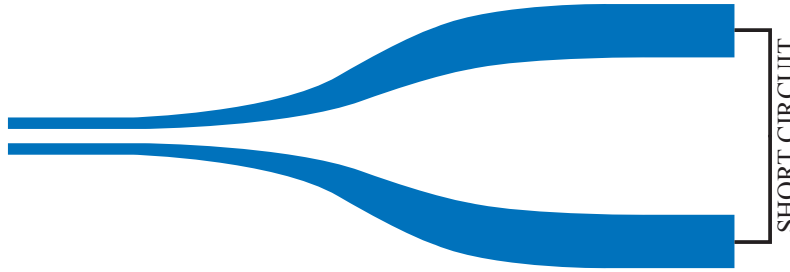


Figure 2.7: A tapered quarter wave transmission line utilizes wide width and large gap spacing when the current is high (voltage is low) and narrow width and small gap when the voltage is high (current is low).

domain [Womack62, Youla64]

$$\frac{d^2V(x)}{dx^2} - \frac{1}{L(x)} \cdot \frac{dL(x)}{dx} \cdot \frac{dV(x)}{dx} - \omega^2 L(x) C(x) V(x) = 0 \quad (2.40)$$

$$I(x) = -\frac{1}{j\omega L(x)} \cdot \frac{dV(x)}{dx} \quad (2.41)$$

As we can see from (2.41), the current is 90° out of phase with the voltage. The standing waves along our $\lambda/4$ resonator are shown in Fig. 2.6 for a coplanar stripline (CPS).³ At the shorted end of the line, the voltage is at a minimum and the current at a maximum, so the losses at this point come mainly from the series resistance of the metal lines. Conversely, at the driven end, the voltage is at a maximum and the current at a minimum, so the losses at this point come mainly from the shunt conductance between the differential lines. This phenomenon can be exploited to lower the losses of the resonator and thus raise the quality factor. At the shorted end of the line we would like to increase the conductor width to reduce the series resistance. Since the voltage is low at this point, the shunt conductance is not very important. Conversely, at the driven end of the line, where the voltage is at a maximum we would like to reduce the width and spacing of the conductors to reduce the shunt conductance due to substrate coupling. Since the current is low at this point the resulting large series resistance of the line is not important. Along the rest of the line, shunt conductance and series resistance can be traded-off to reduce overall loss. Conceptually, the resulting taper should have a shape similar to Fig. 2.7.

The preceding discussion has completely ignored the effects of such a taper on the characteristic impedance of the line at each point and thus the question must be asked: does the characteristic impedance have to be constant along the taper? The simple answer is no. The idea of tapering a transmission line has been studied and used extensively in the microwave community for many years as a way to provide a conjugate match between two sections of transmission line with different characteristic impedances. As early as the 1930s many authors had worked out solutions for the behavior of a tapered transmission line with

³The CPS transmission line is described in Section 2.5.4.

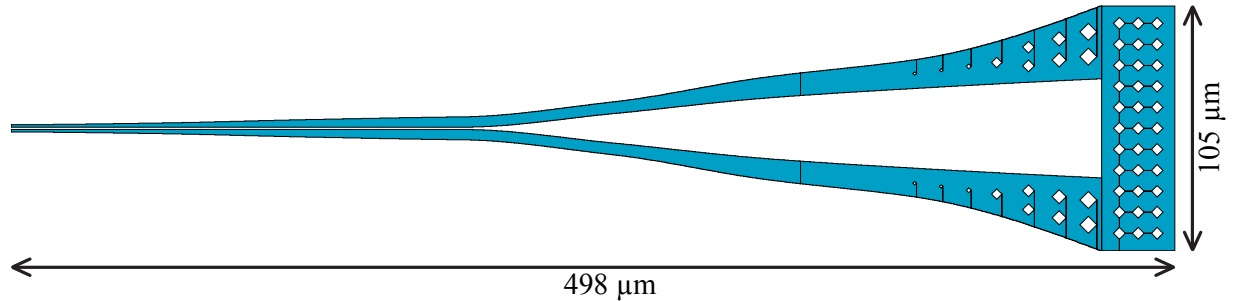


Figure 2.8: The layout of the optimized quarter wave line. The characteristic impedance, Z_o , is non-constant. Slotting is introduced to satisfy design rules.

characteristic impedance profiles that varied in a specified manner (e.g.: linear, exponential, Gaussian, etc.). However, it is not immediately obvious how these results can be applied to tapering resonators.

For simplicity let us begin by assuming constant characteristic impedance is maintained across the entire taper. Luckily, a CPS line's dimensions can be adjusted to achieve infinitely many combinations of series loss versus shunt loss while maintaining a constant characteristic impedance. In fact, [Andress05] showed that if the characteristic impedance is kept constant along the taper, the transmission line can be thought of as a piecewise construction of infinitesimally small uniform transmission line segments with the same characteristic impedance, Z_o , but different complex propagation constant, β . Using this construction, the voltage and current profiles along the line can be calculated in the phase domain which is related to the physical domain defined by the position, z , along the line by

$$\theta(z) = \int_0^z \beta(z') dz' \quad (2.42)$$

In the phase domain, the voltage and current profiles of a tapered resonator with constant characteristic impedance are always sinusoidal in shape just as shown in Fig. 2.6 for a uniform resonator in the physical domain. This closed form solution then allows a straightforward optimization of the series and shunt losses directly in the phase domain. The result is a quality factor improvement of ideally 60% over the best untapered resonator.

On the other hand, if the characteristic impedance is allowed to vary, the current and voltage profiles will no longer be sinusoidal even in the phase domain. Their solution can still be found analytically if the characteristic impedance profile is well behaved (e.g.: linear [Lu97], exponential [Womack62]) but for arbitrary lines, only numerical solutions are possible. An optimization method was developed in [Marcu08b] which optimized the transmission line taper for a 60GHz resonator, without constraints on the characteristic impedance profile, based on a numerical solver for the voltage and current profiles.⁴ The resulting taper (Fig. 2.8) achieved a quality factor of 15, a 70% improvement over the untapered resonator at 60GHz,

⁴Details of the optimization methodology can be found in [Marcu08a].

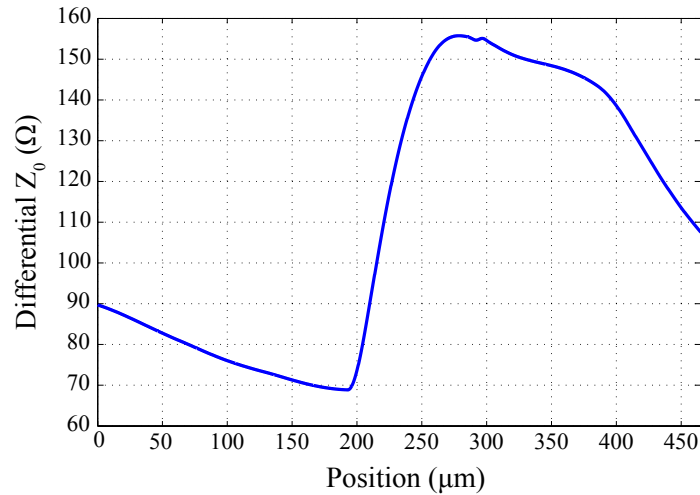


Figure 2.9: The optimum characteristic impedance profile.

and a 10% increase over the maximum achievable for a constant characteristic impedance taper. The resulting optimum characteristic impedance profile is shown in Fig. 2.9. Examining the layout and characteristic impedance profile we can immediately recognize that the optimization maximized the capacitance per unit length (low Z_o) at one end, and maximized the inductance per unit length (high Z_o) at the other. In effect, the optimization built the closest thing it could to an LC tank. In fact, at 60GHz the optimum tapered transmission line resonator achieves performance on par with an optimized lumped LC tank [Marcu08a]. Unfortunately, the LC tank also occupies less die area and is thus the better choice for most designs at 60GHz.

2.4 MEMS Resonators

MEMS resonators rely on mechanical resonances and thus their resonance frequencies are determined solely by their shape and design dimensions relative to the speed of sound in the material.⁵ They are very popular at low-GHz frequencies due to their extremely high quality factor, however they cannot be easily integrated on-chip. To achieve higher resonance frequencies, these devices must be made physically smaller so manufacturing tolerances and capabilities limit the maximum achievable frequency.

The following are the most common MEMS resonators in use today:

- Crystal (XTAL) resonators are crystals of piezoelectric material such as quartz. High frequency resonators are cut in the shape of a rectangular plate while low frequency

⁵In transmission lines energy is carried by electromagnetic fields which travel at the speed of light, while in MEMS resonators, energy is carried by mechanical vibrations which travel at the speed of sound in the material.

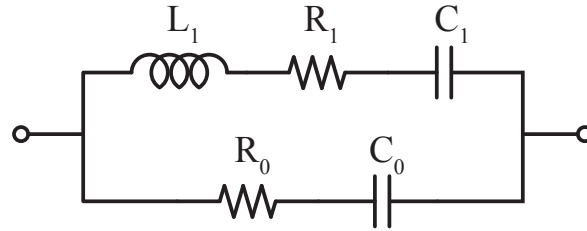


Figure 2.10: MEMS resonator model.

resonators are cut in the shape of a tuning fork. Electrodes are attached to couple energy into the resonator. The resonance frequencies range from a few kHz to 300MHz with quality factors on the order of 10^6 .

- Film Bulk Acoustic Resonators (FBAR) [Ruby01] utilize a thin slice of piezoelectric material such as, AlN or ZnO , sandwiched between two electrodes. The resonant frequency is determined by the thickness of the piezoelectric material which is accurately trimmed during manufacturing. Thus, resonators of different resonant frequencies cannot be manufactured together. The resonance frequency for FBARs ranges between 100MHz and 10GHz with quality factors on the order of 1000 in a high-volume, industrial production environment.
- Disc resonators [Nguyen07] are made of a disc of material surrounded by electrodes. They can be manufactured using either a piezoelectric material with directly attached electrodes, or the electrodes can simply couple energy to a freestanding disk electrostatically. In the latter case, the material of the disc is not piezoelectric. The resonance occurs in the plane of the disc so the resonance frequency is set by the lateral dimensions rather than the thickness of the disk. Unlike FBARs, this allows the selection of resonance frequency at design time by individually sizing different resonator discs as needed. Reported resonance frequencies currently range between 100 MHz and 1.5GHz with quality factors on the order of 10,000 in a research lab production environment. With continued scaling, however, these devices should be able to reach frequencies over 10GHz.

A simple model, shown in Fig. 2.10, can be used for MEMS resonators of all kinds. It was first introduced in [Pro57] and expanded in [Larson00] by adding resistor R_0 for a better fit between model and measurement. Called the Modified Butterworth-Van Dyke (MBVD) model, it exhibits both a series resonance at ω_s and a parallel resonance at ω_p .

$$\omega_s = \frac{1}{\sqrt{L_1 C_1}} \quad (2.43)$$

$$\omega_p = \frac{1}{\sqrt{L_1 \frac{C_1 C_0}{C_1 + C_0}}} = \omega_s \sqrt{1 + \frac{C_1}{C_0}} \quad (2.44)$$

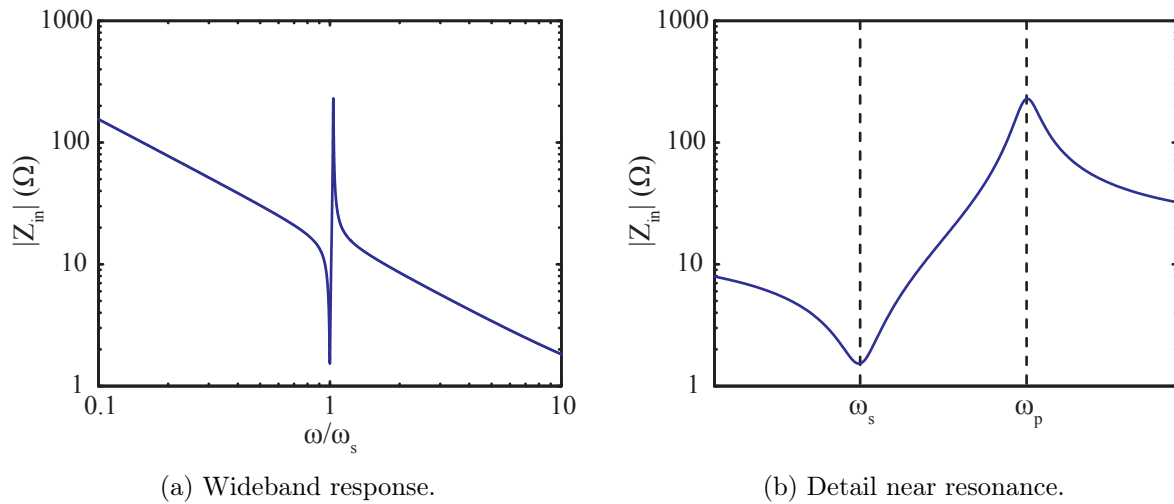


Figure 2.11: MEMS resonator impedance.

Even though $\omega_s < \omega_p$, the two resonant frequencies are actually very close to each other with the parallel resonance occurring just slightly above the series resonance since $C_1 \ll C_0$. The magnitude of the resonator impedance is plotted in Fig. 2.11a. Closer inspection of the impedance near resonance (Fig. 2.11b) shows that the resonator behaves like an inductor for $\omega_s < \omega < \omega_p$, and like a capacitor for frequencies outside this region. Oscillators based on MEMS resonators utilize this inductive behavior in a series configuration such as a Colpitts oscillator (see Section 3.3). The resonator can effectively take on any value of inductance to resonate out any value of capacitance that the oscillator presents and create oscillation at some frequency between ω_s and ω_p . On one hand, this means that the oscillation frequency is very well defined. On the other hand, it also means that the frequency cannot be tuned significantly even if that is desired.

The frequency-Q product is a commonly used figure of merit which allows a fair comparison between MEMS resonators at various frequencies. Some commonly reported values for the fQ product are given in Table 2.1. These numbers then allow us to extrapolate the quality factors that would be possible if these technologies could be extended to 60GHz, as shown in the rightmost column. As we can see, these values are significantly higher than what is achievable with on-chip resonators, however, MEMS resonators have not yet been manufactured anywhere near mm-wave frequencies. Furthermore, integration with silicon processes is crucial to their adoption at mm-wave frequencies since moving signals on and off the die incurs significant loss and will reduce the effective quality factor of an off-chip resonator.

Resonator Type	Frequency	Q	fQ Product	Extrapolated Q @ 60GHz
XTAL	16MHz	1,000,000	1.6×10^{13}	267
FBAR	1.9GHz	2,500	0.475×10^{13}	79
Disc	1.46GHz	15,248	2.22×10^{13}	370

Table 2.1: Frequency-Q product of MEMS resonators.

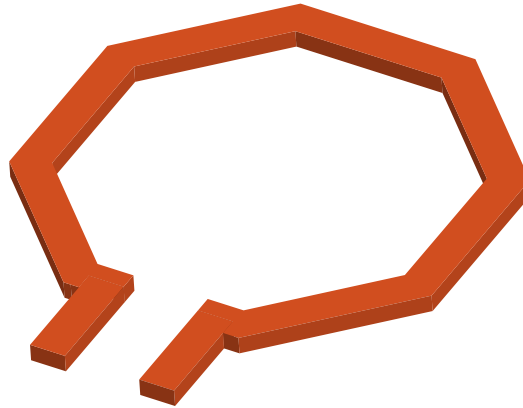


Figure 2.12: Single turn ring inductor.

2.5 Passive Components

In the previous sections we have explored resonant tanks for *mm*-wave frequencies. In this section we will go into further detail regarding the design of individual passive components which can be used either as standalone devices, or as part of the resonant tanks previously discussed.

2.5.1 Inductors

Inductances required for most 60GHz designs are in the range of 10pH-1nH. It is difficult to build inductors with values near or above 1nH so inductances for matching purposes or resonant tanks should be well below this value. Ring inductors (Fig. 2.12) are the easiest way to build inductors of moderate value (approx. 50-500 pH). For smaller values, a ring inductor will be difficult to design accurately but a short section of transmission line can be used instead. The input impedance of an ideal lossless transmission line with a short circuit load is given by (2.28). If we assume the length of the transmission line is very small (indeed much smaller than $\lambda/4$) we can use the first term of the Taylor series expansion of $\tan(x)$

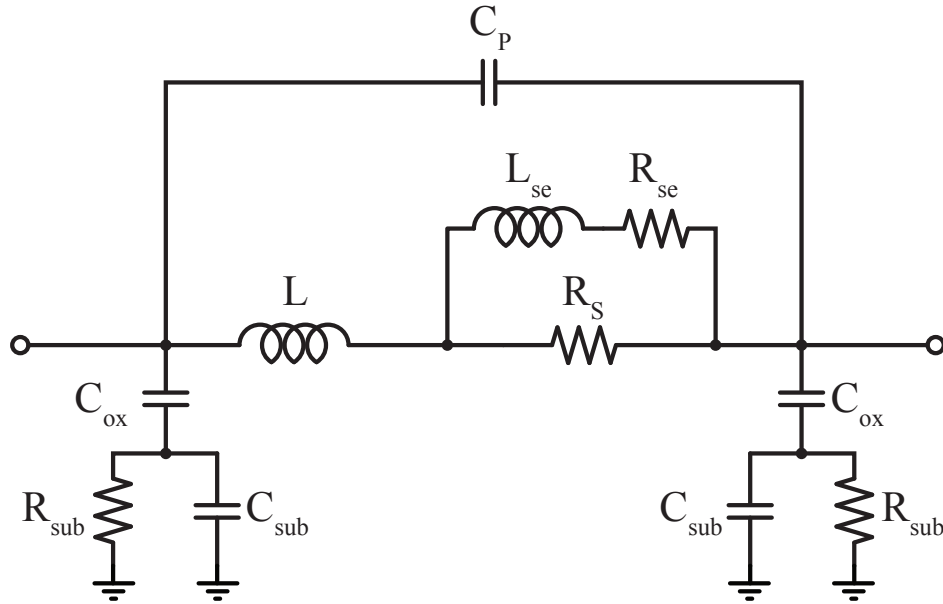


Figure 2.13: Wideband lumped element inductor model.

to find a very good approximation for the input impedance.

$$\begin{aligned}
 Z_{in} &= jZ_o \tan\left(2\pi\frac{\ell}{\lambda}\right) \\
 &= jZ_o \tan\left(\frac{\omega\ell}{\nu_p}\right) \\
 &\approx j\frac{Z_o\omega\ell}{\nu_p}
 \end{aligned} \tag{2.45}$$

The impedance looks like an inductor with an effective value of

$$L_{eff} = \frac{Z_o\ell}{\nu_p} \tag{2.46}$$

The length of the transmission line can be designed very accurately, making it very easy to design even very small values of inductance.

Fig. 2.13 shows a lumped element model of an inductor based on the models introduced in [Yue00, Cao03]. It takes into account many loss mechanisms and nonidealities and is accurate over a very wide range of frequencies. There are two loss mechanisms present in an inductor: metal losses and substrate losses. The finite resistance of the metal used to build the inductor introduces a series resistance, R_S , and leads to conductive losses within the metal. Due to skin effect, this resistance is frequency dependent, increasing with the square root of frequency. This can roughly be modeled with L_{se} and R_{se} . The semiconducting silicon substrate also introduces losses due to induced eddy currents as well as dielectric

resonance. The finite, non-zero conductance of the substrate leads to induced eddy currents that experience conductive loss, represented by R_{sub} . The dielectric loss tangent of the substrate gives rise to dielectric resonance losses which increase in direct proportion to frequency thus making R_{sub} frequency dependent. Finally, parasitic capacitance between the inductor leads, C_P , as well as the capacitance of the oxide and substrate, C_{ox} and C_{sub} respectively, lead to self-resonance of the inductor. That is, the inductor will only behave as an inductor well below the self-resonance frequency and as a capacitor above the self-resonance. The effective inductance seen across the leads of the inductor will also be frequency dependent even in the inductive region. To find an analytic expression for the effective inductance we can just assume a very simple model of self-resonance where we just have an ideal inductor, L , in parallel with an ideal capacitor, C . The admittance of this structure is given by

$$\begin{aligned} Y &= \frac{1}{j\omega L} + j\omega C \\ &= \frac{1}{j\omega L} (1 - \omega^2 LC) \\ &= \frac{1}{j\omega L} \left(1 - \frac{\omega^2}{\omega_{sr}^2} \right) \end{aligned} \quad (2.47)$$

where ω_{sr} is the self-resonance frequency. The effective inductance is then

$$L_{eff} = \frac{L}{1 - (\omega/\omega_{sr})^2} \quad (2.48)$$

As we can see from (2.48), L_{eff} is approximately equal to L for frequencies well below the self-resonance frequency. However, as we approach the self-resonance frequency, L_{eff} increases asymptotically approaching infinity at $\omega = \omega_{sr}$. Since the exact self-resonance frequency is hard to predict, the designer must always ensure that the operating frequency is well below the expected self-resonance frequency otherwise, the effective inductance will be hard to predict accurately. This is also why shielding of inductors at mm-wave frequencies is not nearly as popular as it is at lower frequencies, since the presence of a shield, while reducing substrate effects, decreases the self-resonance frequency, making it very difficult to build moderate to large inductors. Similarly, capacitance between multiple turns of a spiral inductor also limits practical designs at mm-wave frequencies to two turns.

The model shown in Fig. 2.13 takes all of the above described effects into account but is clearly too complex for hand calculation. At a particular design frequency, however, we can lump all the losses into one resistance, R_F , in series with the effective inductance at that frequency, L_F , which takes into account the effects of self-resonance.⁶ This simplified model is shown in Fig. 2.14. Clearly, this is an oversimplification of the actual loss mechanisms described above but is accurate as long as it is only used to model a real inductor over a narrow frequency range.

⁶The subscript F has been used to denote that these parameters are a function of frequency.



Figure 2.14: Simplified inductor model valid over a narrow frequency range.

It is now instructive to define an inductive quality factor which will represent its loss and will be useful in comparing different inductors. The quality factor, in general, is defined as the ratio of energy stored per cycle to energy dissipated per cycle.

$$Q \triangleq 2\pi \frac{W_s}{W_d} \quad (2.49)$$

The quality factor of the inductor shown in Fig. 2.14 is then given by (2.50) and is usually no higher than 20-30 at 60GHz.

$$Q_L = \frac{\omega L_F}{R_F} \quad (2.50)$$

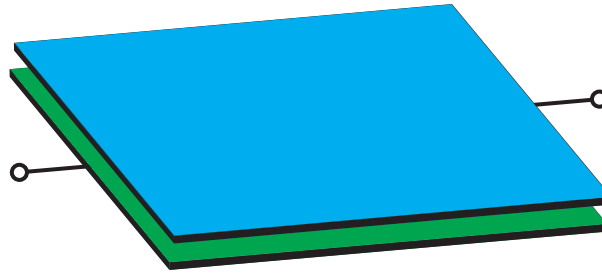
2.5.2 Capacitors

Capacitors are generally found in matching networks and resonant tanks but are also used as AC coupling elements between stages, allowing independent biasing of the outputs and inputs of cascaded stages, as well as DC decoupling elements to filter out high frequency noise on DC bias lines. The two most popular capacitor topologies are parallel plate, also called MIM (metal-insulator-metal) shown in Fig. 2.15a, and fringing, also called MOM (metal-oxide-metal) shown in Fig. 2.15b. Most CMOS processes offer a special MIM process option which uses extra metal layers and a dielectric with high permittivity, leading to very high capacitance density. Without this option, standard metal layers must be used and the capacitance density is thus very low.

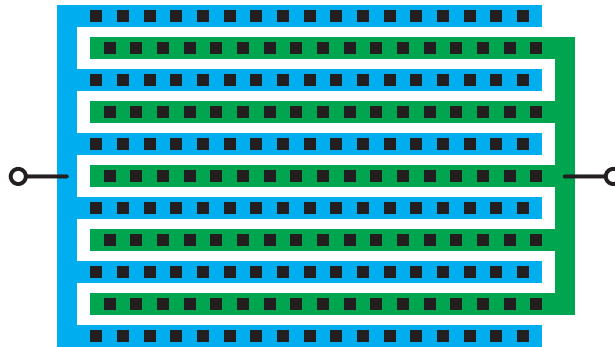
MOM capacitors on the other hand utilize standard process options. They are made up of many long and narrow fingers, made by strapping multiple metal layers together with vias. Minimum metal width and spacing is used to maximize capacitive density. MOM capacitors are found as standard cells in most modern processes, making them very popular. Furthermore, their use of both parallel plate and fringing effects leads to reasonable capacitance density, while using many narrow fingers somewhat mitigates the skin effect by distributing current better than a single plate of metal.

Similar to inductors, capacitors suffer from many nonidealities. A lumped element model of a capacitor, valid over a wideband of frequencies, is shown in Fig. 2.16. Self-resonance due to lead inductance as well as inductance within the structure of the capacitor itself (such as from long, narrow fingers),⁷ is represented by a single inductor, L_S . Metal losses, R_S ,

⁷Utilizing many fingers in parallel as well as alternating the direction of the fingers between layers (i.e.: 0° , 90° , 0° , etc.) helps to mitigate the latter contribution.



(a) MIM capacitor.



(b) MOM capacitor.

Figure 2.15: On-chip capacitor structures.

and substrate losses, R_{sub} , with their associated frequency dependencies are also present (R_{se} , L_{se}). Substrate coupling through C_{ox} and C_{sub} is especially important however. If a capacitor is used with one terminal grounded, as is the case for a bypass capacitor, C_{ox} and C_{sub} simply increase the effective capacitance and, in most cases, this is helpful. Adding a shield to the capacitor, which increases C_{ox} , does not harm us in this case. If, on the other hand, the capacitor is used in series as a decoupling capacitor, signal is actually lost through C_{ox} and C_{sub} . In this configuration we would like to reduce the parasitic capacitances as much as possible. However, noise coupling from the substrate may be a problem in some designs, making a shield necessary. In that case, the selection of a shield involves a trade-off between the negative effects of loss versus the negative effects of noise coupling and is completely design dependent. Finally, some capacitors may suffer from leakage through the dielectric which introduces a shunt resistance, R_P . MIM and MOM capacitors generally do not have any leakage ($R_P = \infty$), however, transistor gate leakage in deeply scaled CMOS processes could be high enough to appreciably reduce R_P and significantly affect the loss in varactors (described in Section 2.5.3).

Due to self-resonance, the effective capacitance seen across the leads of the capacitor will be frequency dependent. To find an analytic expression for the effective capacitance we can just assume a very simple model of self-resonance where we just have an ideal capacitor, C ,

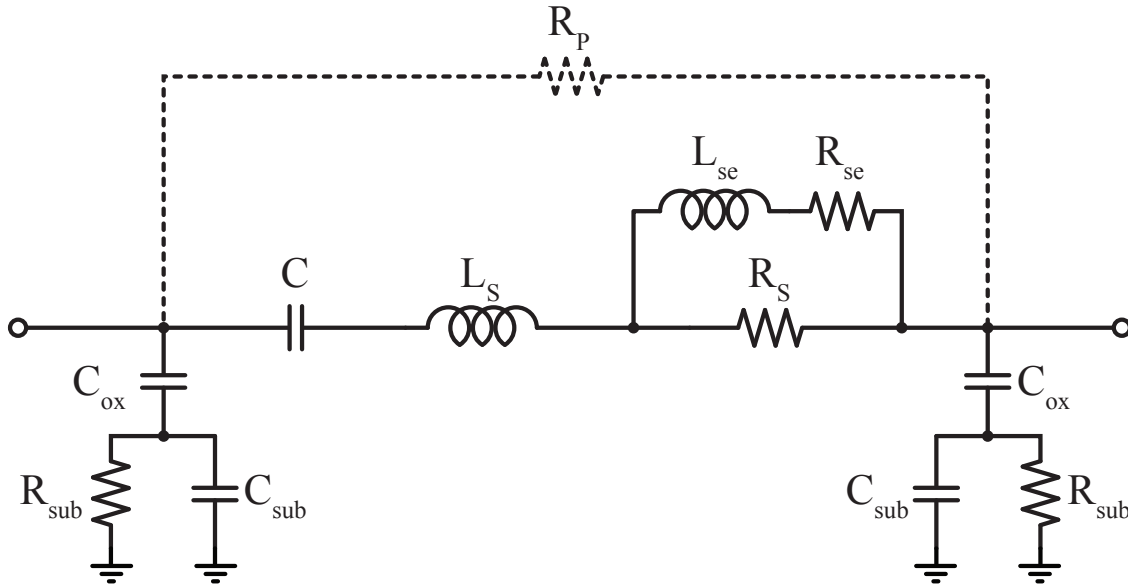


Figure 2.16: Wideband lumped element capacitor model.



Figure 2.17: Simplified capacitor model valid over a narrow frequency range.

in series with an ideal inductor, L . The impedance of this structure is given by

$$\begin{aligned}
 Z &= \frac{1}{j\omega C} + j\omega L \\
 &= \frac{1}{j\omega C} (1 - \omega^2 LC) \\
 &= \frac{1}{j\omega C} \left(1 - \frac{\omega^2}{\omega_{sr}^2}\right)
 \end{aligned} \tag{2.51}$$

where ω_{sr} is the self-resonance frequency. The effective capacitance is then

$$C_{eff} = \frac{C}{1 - (\omega/\omega_{sr})^2} \tag{2.52}$$

The model that takes all of the above effects into account (Fig. 2.16) is clearly too complex for hand calculation. At a particular design frequency, however, we can lump all the losses into one resistance, R_F , in series with the effective capacitance at that frequency, C_F , which takes into account the effects of self-resonance. This simplified model is shown in Fig. 2.17. Just

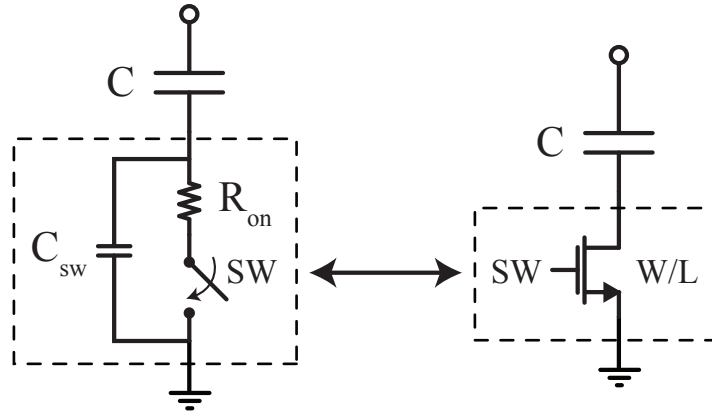


Figure 2.18: Switched capacitor.

as in the inductor case this is an oversimplification of the actual loss mechanisms accurate only over a narrow frequency range.

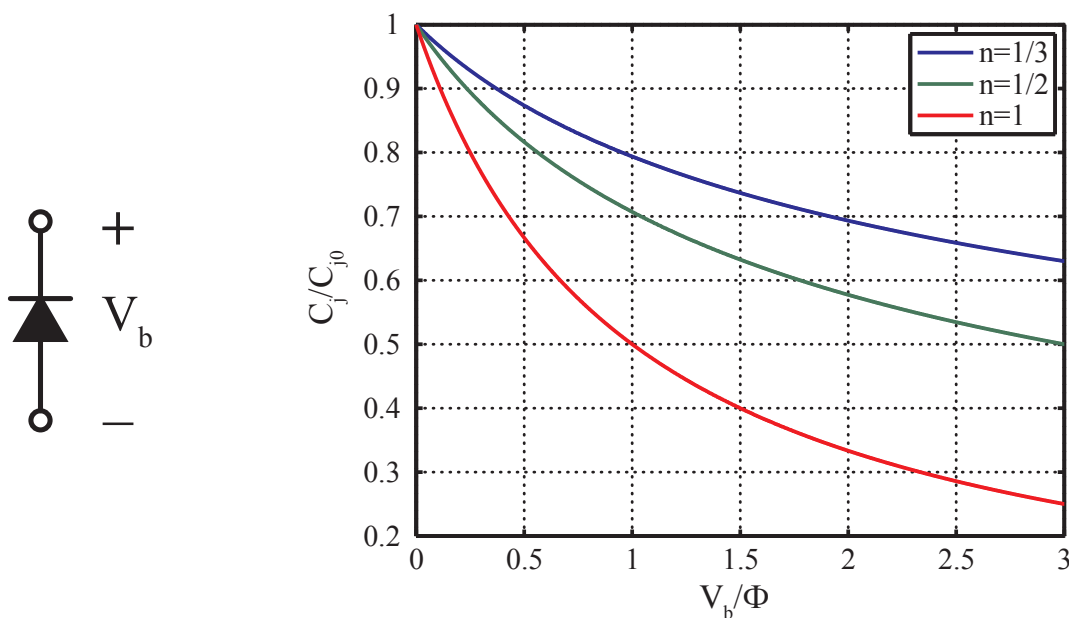
We can now define the capacitive quality factor which will represent its loss, just as we did for the inductor. The quality factor of a capacitor is given by

$$Q_C = \frac{1}{\omega R_F C_F} \quad (2.53)$$

Note that whereas the Q of an inductor increases with frequency, for a capacitor it decreases. This means that for low frequency designs the overall quality factor of a resonant network will generally be limited by the inductors. At mm -wave frequencies, however, this is not the case and capacitive quality factors, especially when varactors are used, will limit the overall network Q . This is not to say that inductor Q is no longer relevant as, in fact, inductive quality factors do not increase indefinitely. Frequency dependent losses and self-resonance ultimately create an upper bound on the Q of any given inductor.

2.5.3 Varactors

A varactor is a variable capacitor whose capacitance depends on a control voltage. It is difficult to change the value of an inductor in fine increments so instead, variable capacitors are used to change the resonance frequency of a resonant tank. The simplest form of variable capacitor, shown in Fig. 2.18, is simply a switched capacitor. When the switch is off, the capacitance seen at the input port is simply the parasitic capacitance of the switch, C_{sw} , in series with C . Thus, we would like C_{sw} to be very small such that the effective off-capacitance is also very small. When the switch is on, only the capacitance C is seen in series with the on resistance of the switch, R_{on} . We would thus also like for R_{on} to be very small in order for C_{sw} to be effectively shorted out and for the quality factor of C to remain high. The switch is implemented with a MOS transistor, however, so there is an inherent trade-off between



(a) Reverse biased diode.

(b) Capacitance versus tuning voltage.

Figure 2.19: Diode varactor.

C_{sw} and R_{on} .

$$C_{sw} = C_{ox}WL \quad (2.54)$$

$$R_{on} = \frac{1}{g_{ds,0}} = \frac{L}{kW(V_{gs} - V_T)} \quad (2.55)$$

The switch should have minimum channel length but the optimum width must be selected based on an acceptable trade-off between the capacitance ratio C_{on}/C_{off} and the quality factor of the on-capacitance. Unfortunately, C is generally small for practical mm-wave designs and so the switch must also be kept small to maintain a useful on-off capacitance ratio. This limits the achievable $g_{ds,0}$ of the switch and thus the quality factor of the switched capacitor. However, technology scaling does help in this regard so switched capacitors are becoming much more useful in *mm*-wave designs.

Another form of variable capacitance is the depletion capacitance of a reverse biased diode, shown in Fig. 2.19a. Unlike a switched capacitor, this type of varactor is continuously variable as a function of the analog biasing voltage as described by (2.56) where C_{j0} is the capacitance at zero bias, V_b is the reverse bias voltage, ϕ is the built-in potential (around 1V for silicon), and n represents the doping profile (assume 1/2 for diodes in a standard CMOS process).⁸ Representative curves of capacitance versus tuning voltage are shown in

⁸In general, the doping profile of the diode junction can be linear ($n = 1/3$), abrupt ($n = 1/2$), or hyperabrupt ($n = 1$) [Pierret96].

Fig. 2.19b for three different doping profiles.

$$C_v = \frac{C_{j0}}{\left(1 + \frac{V_b}{\phi}\right)^n} \quad (2.56)$$

Finally, a MOS transistor can also be used as a varactor by using the gate as one terminal and tying together the source, drain and bulk as the second terminal [Andreani00] as shown in Fig. 2.20a. Depending on the gate-source bias voltage, V_{gs} , this device can operate in one of three regions: accumulation, depletion, and inversion. Let us take an NMOS device as an example. If the gate voltage is brought down below the S-D-B voltage, holes are attracted from the substrate and accumulate at the oxide-semiconductor interface, eventually forming a conductive layer for low enough gate voltages. This is called the accumulation region. On the other hand, if the gate voltage is higher than the S-D-B voltage, holes are pushed away from the oxide-semiconductor interface, creating a depletion region whose depth increases with increasing V_{gs} . For $V_{gs} > V_T$, the threshold voltage, the device enters the inversion region where a channel of electrons is formed between the drain and source. The device is in strong inversion for $V_{gs} \gg V_T$. Since the source and drain terminals are tied together there is no current flowing in the channel, it simply acts as a conductive plate. Thus, the capacitance both in accumulation and strong inversion is equal to $C_{ox}WL$. In depletion, on the other hand, the capacitance reduces as the depletion region depth increases. The capacitance versus V_{gs} in all regions of operation is shown in Fig. 2.20b.

The quality factor can be derived by assuming a distributed RC network model for the gate capacitance and channel resistance [Andreani99] as shown in Fig. (2.21). The series impedance between the two terminals of the varactor is then given by

$$Z_C = \frac{1}{j\omega C_{ch}} + \frac{R_{ch}}{12} \quad (2.57)$$

where the lumped channel capacitance, C_{ch} , and channel resistance, R_{ch} , in strong inversion are given by

$$C_{ch} = C_{ox}WL \quad (2.58)$$

$$R_{ch} = \frac{1}{g_{ds,0}} = \frac{1}{k\frac{W}{L}(V_{gs} - V_T)} \quad (2.59)$$

This model assumes that the metal, gate, and contact resistances are negligible compared to R_{ch} . Due to the distributed nature of the channel resistance, its effective value is actually reduced by a factor of 12. The quality factor is then given by

$$\begin{aligned} Q_{MOS} &= \frac{12}{\omega C_{ch} R_{ch}} \\ &= \frac{12k(V_{gs} - V_T)}{\omega C_{ox} L^2} \end{aligned} \quad (2.60)$$

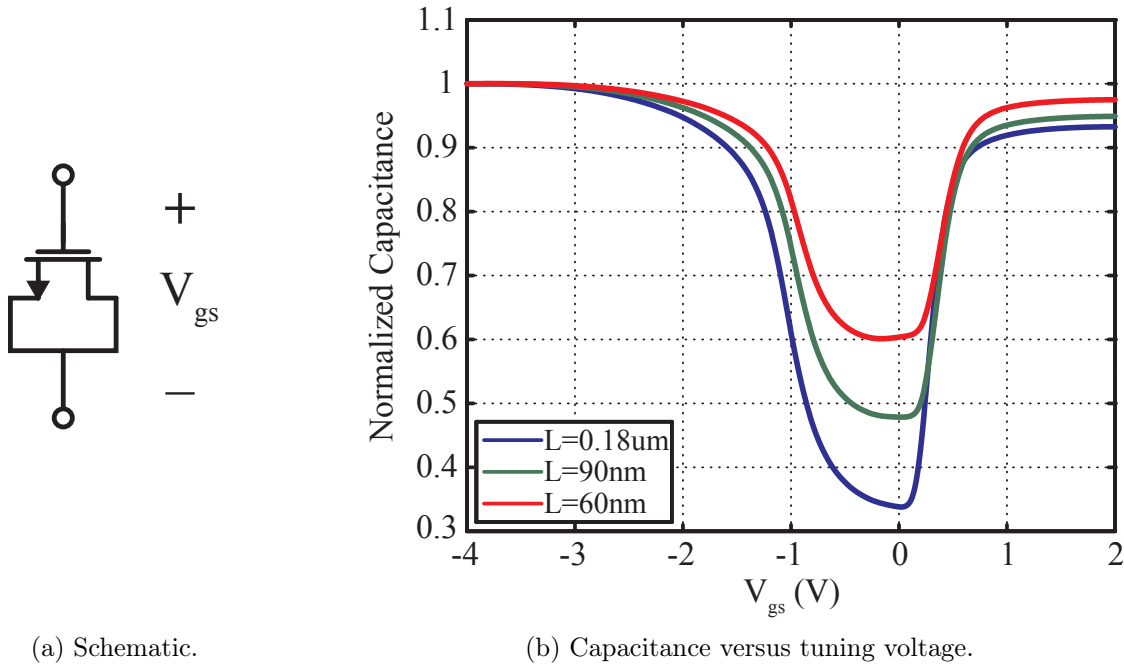


Figure 2.20: MOS varactor.

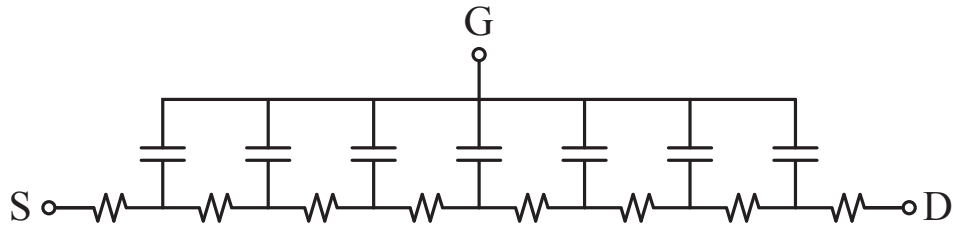


Figure 2.21: Distributed channel impedance model for MOS varactor.

As we can see from (2.60) and Fig. 2.20b, choosing the channel length of the device involves a trade-off between capacitive tuning range and quality factor.

Under large signal excitation, such as the large swings present in a VCO, the instantaneous capacitance of the MOS varactor varies over the signal period. The effective capacitance is then related to the average capacitance over the period [Hegazi03]. For large signal swings, the effective tuning range of the varactor will be reduced if the varactor enters both the accumulation and inversion regions of operation within one period. The fundamental problem here is the fact that the varactor tuning curve is nonmonotonic. The problem would be solved if the varactor could be prevented from entering either accumulation or inversion.

Accumulation MOS (AMOS) varactors [Castello98, Soorapanth98] utilize a structure very similar to an NMOS except that it is fabricated in an N-Well instead of the P substrate as shown in Fig. 2.22b. Compare this structure with the standard NMOS layout shown in

Fig. 2.22a. (A PMOS manufactured in a P-Well instead of an N-Well can also be used to form the P-type of this device.) At high frequencies this structure prevents minority carriers from being injected into the depletion region so that the channel cannot be inverted. High frequency operation is thus limited to the accumulation and depletion regions only, making the tuning curve both monotonic as well as nicely behaved with a smooth, shallow slope. Furthermore, this structure has higher Q due to the higher mobility of electrons compared to holes.

Inversion mode (IMOS) varactors [Andreani00], on the other hand, only utilize the depletion and inversion regions by inhibiting accumulation. This is very easily accomplished with the standard MOS varactors by tying the bulk to the lowest voltage in the circuit, in most cases GND as shown in Fig. 2.22c. The gate and S-D terminals are biased separately based on the particular oscillator topology. The varactor tuning curve has a much steeper slope in the inversion region than the accumulation region making the capacitance very sensitive to control voltage. However, with large signal swings the average capacitance of the varactor will be determined by the amount of time the signal spends on one side of the transition versus the other each period. This is similar to pulse-width modulation and effectively linearizes the large signal capacitance tuning curve without affecting the maximum and minimum achievable values.

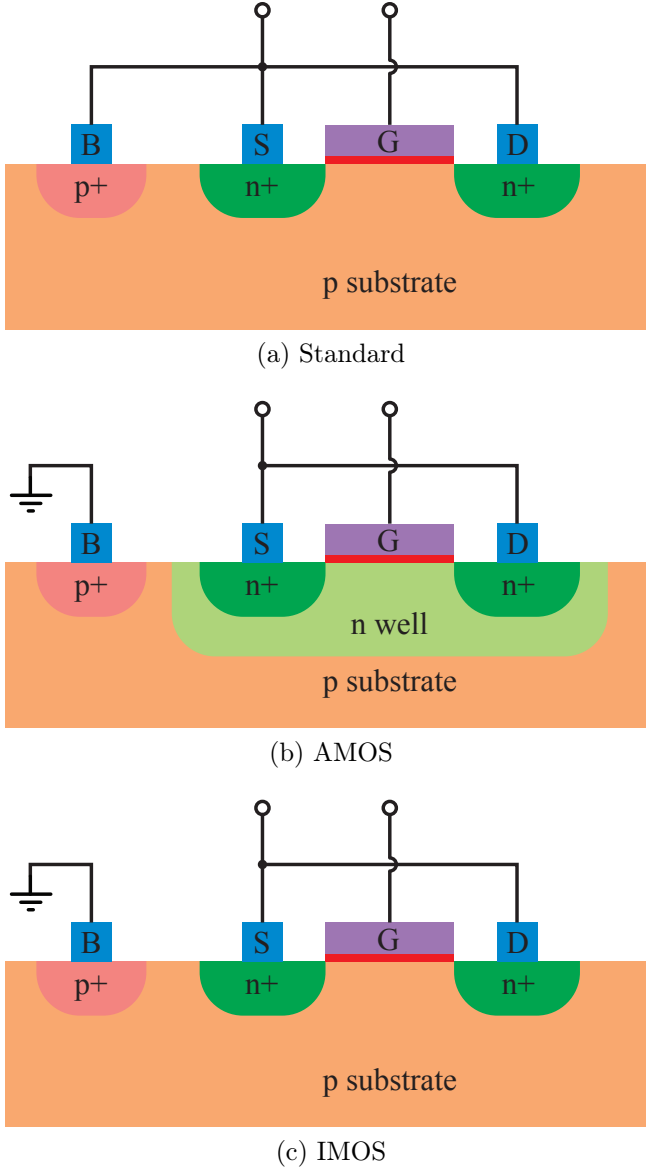


Figure 2.22: MOS varactor layout.

2.5.4 Transmission Lines

At lower RF frequencies (e.g.: 1-5GHz), transmission lines are generally not used on-chip since the required lengths are on the order of the wavelength, and thus are too large. As the frequencies increase into the *mm*-wave domain, wavelengths reduce to on the order of a few mm (as the name *mm*-wave implies). For example, the on-chip wavelength at 60GHz is only 2.4mm, a length that is easily realizable. On the other hand, as frequencies increase interconnect lengths begin to approach an appreciable fraction of the wavelength, no longer

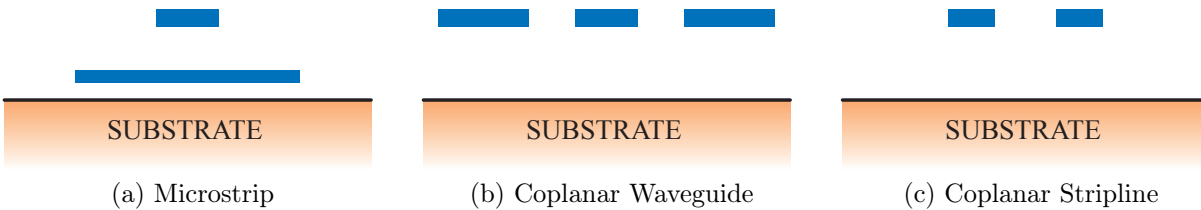


Figure 2.23: On-chip transmission lines.

acting as simple wires or RC delay lines. In other words, distributed effects begin to manifest themselves. At these frequencies transmission lines can be used in a matched environment for a robust interconnect design. This will be further explored in Chapter 5 where we will look at LO distribution strategies for large phased array transceivers.

The most popular transmission line topologies (shown in Fig. 2.23) are microstrip, coplanar waveguide (CPW), and coplanar stripline (CPS). The microstrip transmission line, first presented by [Grieg52] as a planar alternative to the three dimensional transmission structures of the day⁹, is made up of a single conductive strip atop a wider ground plane. Microstrip lines offer the most layout flexibility due to the use of a single conductor which allows for very compact turns. The ground plane between the signal conductor and substrate also shields the signal from substrate noise and losses if the metal thickness is much larger than the skin depth, making the microstrip especially attractive. However, adjacent microstrip conductors easily couple to each other if they are placed in close proximity. Thus, care must be taken in routing different signals to prevent unwanted coupling. The characteristic impedance of microstrip is approximately proportional to the ratio h/w . Since h is determined by our substrate stackup, we can only control w to achieve a desired characteristic impedance. However, the height of the top metal layer in deeply scaled CMOS processes is constantly reducing, making high characteristic impedance difficult to achieve. Furthermore, while the signal conductor can be placed in the highest, thickest metal layer, the ground plane must be placed in a lower, thinner metal layer leading to higher losses.

CPW, on the other hand, is a single layer transmission line first proposed by [Wen69] (with analysis for finite substrates by [Knorr75]), in which both signal and ground conductors occupy the same layer. This allows CPW lines to utilize only the topmost, thick metal layer. While this reduces metal losses, it also exposes the signal conductor to the lossy substrate, possibly increasing substrate loss. However, due to the presence of ground planes on either side of the signal conductor, coupling between adjacent CPW transmission lines is minimized [Riazat86]. The characteristic impedance of CPW is a complicated function of both w and s . In fact there are an infinite number of solutions for any given value of characteristic impedance. In order to select the best design we must take into account the

⁹Coaxial cable, waveguides, and 2-wire line were the standard transmission structures used in radio in the 1950s. The microstrip was much easier to manufacture due to its planarity and occupied much less space allowing for smaller radii.

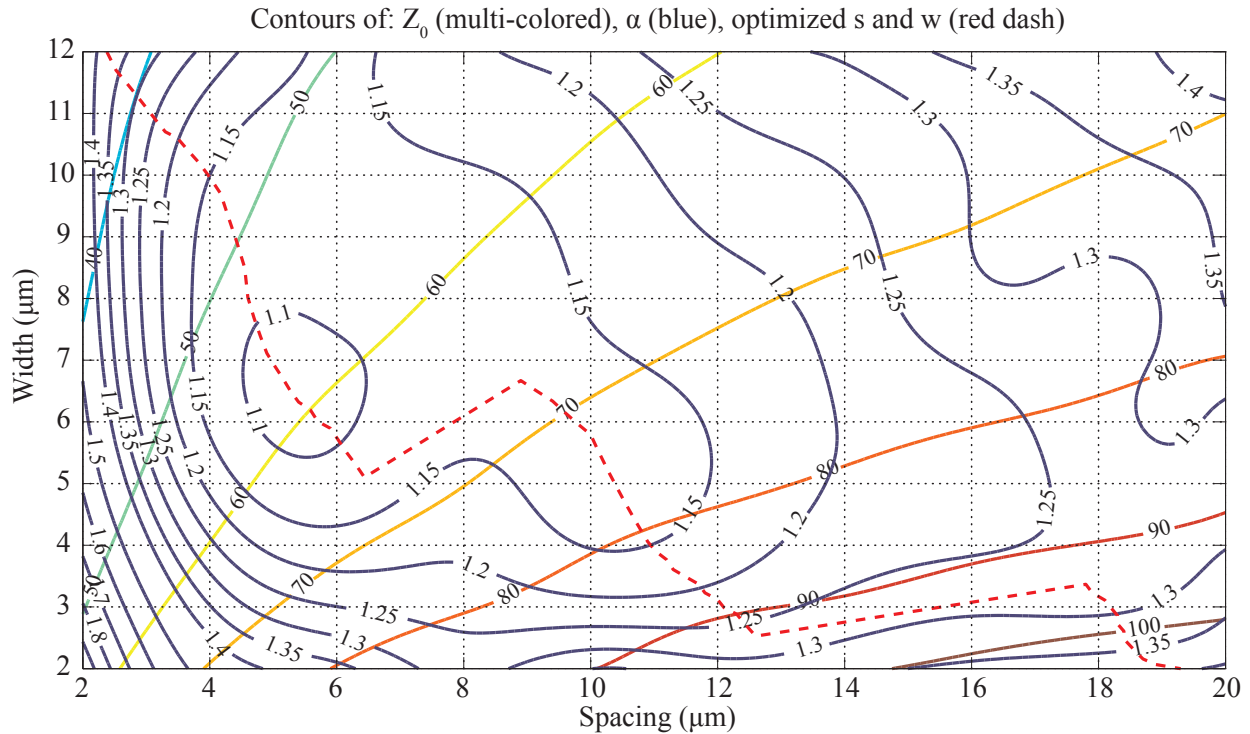


Figure 2.24: CPW design space.

loss. Conceptually we can imagine that if the width is made very narrow, metal losses will increase. We can also imagine that for very small spacing the electric field will be very closely constrained mostly between the conductors. Since very little electric field penetrates the substrate, there would be very little loss due to the semiconducting substrate. On the other hand, if the spacing is made very wide more of the electric field will penetrate the substrate leading to increased losses.

To quantify the above discussion we now employ a 3D EM simulator to simulate the effects of w and s on both characteristic impedance and loss. The results of a simulation in which w and s were both swept over a wide range are plotted in Fig. 2.24. The colored contours radiating out from the origin are contours of constant characteristic impedance, while the blue contours give the loss, plotted in dB/mm . As we can see, there are indeed an infinite number of combinations of dimensions for any particular value of characteristic impedance and our intuition regarding the effects of width on metal losses and spacing on substrate losses was correct. The loss contours now give us a way to choose the CPW dimensions that result in the lowest loss. For example, if we would like a 70Ω line, the lowest achievable loss is approximately $1.12dB/mm$ which occurs at $s = 9.6\mu m$ and $w = 6.1\mu m$. Repeating this procedure for every value of characteristic impedance gives the dashed red line which traverses the plot in Fig. 2.24 from the top left to the bottom right. This line thus represents the optimum CPW design for any value of characteristic impedance. Using this result we can then plot the minimum achievable loss versus Z_o in Fig. 2.25.

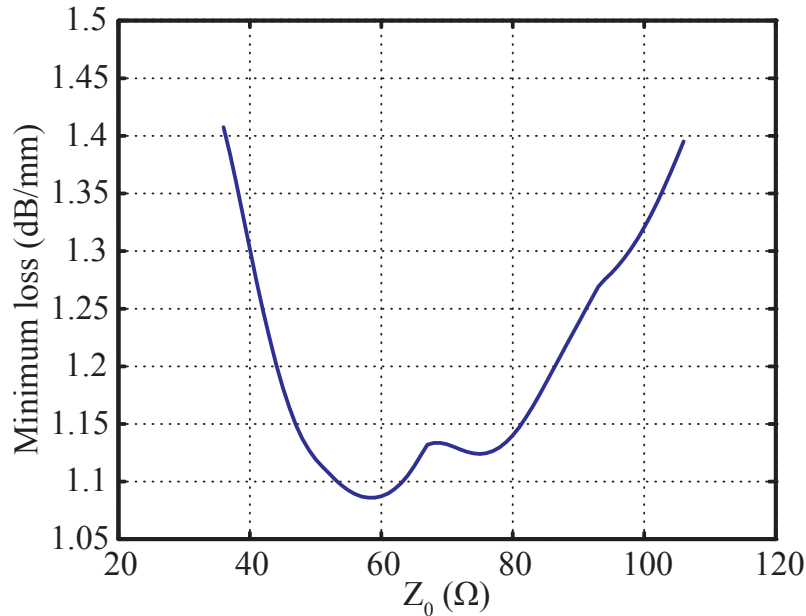


Figure 2.25: CPW design space.

The CPW transmission line has one other effect which must be taken into account. A slotline transmission line, first proposed for use as a transmission line in [Cohn69], is simply a narrow gap in a conductive plane. In such a structure the electric field lines extend across the gap causing a voltage difference between the conductors on either side of the gap. The standard CPW line can thus also be seen as two coupled slotlines (CSL). In the CPW mode the electric fields in the two slots have opposite polarities, while in the CSL mode, the electric fields have the same polarity. Thus, in the CSL mode a voltage difference exists between the left-most and right-most conductors, what we called the ground planes of the CPW. When driven in a CPW mode, these two ground planes are tied together at the source and the coupled slotline mode cannot exist. However, far from the driving point, the two ground conductors are no longer tightly coupled. Thus, if there is a discontinuity in the line, such as a turn or a junction, mode-mixing can occur and the CSL mode can arise. To prevent this from happening, a bridge can be placed in a lower metal layer, close to the discontinuity, shorting the two ground planes together at the same potential and killing the unwanted mode near the bridge [Riaziat86]. Unfortunately, the slotline mode is not eliminated but will actually grow as it moves away from the bridge [Ponchak05]. Thus, bridges must be placed at regular intervals even along straight sections of transmission line with no discontinuities. For designs at 60GHz it was found experimentally that placing bridges every $100\mu\text{m}$ sufficiently suppressed the unwanted mode. Adding these bridges will lower the Z_o of the CPW mode due to excess capacitance between the signal line and each bridge. Thus, in designs with many discontinuities, where many bridges are necessary, their effect on Z_o must be taken into account.

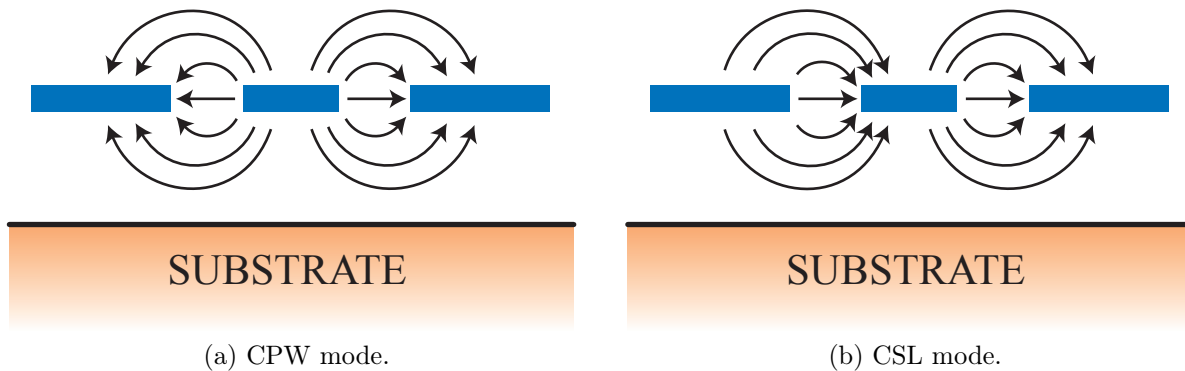


Figure 2.26: E-fields for the two modes present in the CPW structure.

The previous discussion also leads us to the grounded CPW (G-CPW) structure, which is a combination of microstrip and CPW. Conceptually, G-CPW is equivalent to adding more and more bridges along a CPW until the entire length is completely filled in, creating a continuous ground plane below the conductor as well as continuous walls connecting this lower ground plane with the the two CPW ground planes. G-CPW thus has the substrate shielding advantages of microstrip and the adjacent signal shielding of CPW. However, it has much higher capacitance per unit length than either microstrip or CPW for the same dimensions due to the increased ground plane area, leading to significantly reduced Z_o .

Finally, the CPS transmission line, made up of two coplanar metal strips on a dielectric substrate, is actually the dual to the CPW line and was thus also first analyzed by [Wen69] and [Knorr75]. In one mode of operation, it can be driven with a single ended source by connecting the other conductor to ground. However, unlike the other transmission lines presented above, the real advantage of CPS is that it can be driven differentially. Since most oscillator topologies of interest are differential, CPS can be used either to distribute the differential VCO output, or even in the VCO tank as will be described in Section 2.3. If no ground plane is present near the CPS line then the even mode has very high impedance and is also very lossy since the only return path is by capacitive coupling to the substrate. In practical designs a ground plane is present but is placed far enough away so as to not significantly decrease even mode rejection, a desirable characteristic in differential, or odd mode, systems.

2.A Derivation of Lumped Resonant Tank Bandwidth

In this Appendix we will derive the relationship between resonant tank bandwidth and its quality factor. We begin with the impedance of a series LC tank given by (2.6) and normalize to R , the impedance at resonance.¹⁰

$$Z_{ser,n} = \frac{Z_{ser}}{R} = 1 + jQ_s \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) \quad (2.61)$$

To find the bandwidth we must find the upper and lower frequencies at which the magnitude of the normalized impedance given by (2.61) increases by 3dB, or in other words when

$$\begin{aligned} |Z_{ser,n}|^2 &= 2 \\ 1 + Q_s^2 \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 &= 2 \\ Q_s^2 \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 &= 1 \end{aligned} \quad (2.62)$$

Taking the square root of both sides leads to two possible quadratic equations we must solve

$$Q_s \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) = \pm 1 \quad (2.63)$$

Taking the first case we have

$$\begin{aligned} Q_s \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) &= 1 \\ \frac{Q_s}{\omega_o} \omega^2 - Q_s \omega_o &= \omega \\ \omega^2 - \frac{\omega_o}{Q_s} \omega - \omega_o^2 &= 0 \end{aligned} \quad (2.64)$$

which can be solved using the quadratic equation to give two roots of (2.62).

$$\omega_{1,2} = \frac{\omega_o}{2Q_s} \pm \frac{1}{2} \sqrt{\left(\frac{\omega_o}{Q_s} \right)^2 + 4\omega_o^2} \quad (2.65)$$

Similarly, taking the second case, we have

$$\begin{aligned} Q_s \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) &= -1 \\ \frac{Q_s}{\omega_o} \omega^2 - Q_s \omega_o &= -\omega \\ \omega^2 + \frac{\omega_o}{Q_s} \omega - \omega_o^2 &= 0 \end{aligned} \quad (2.66)$$

¹⁰We can also perform an identical derivation for a parallel LC tank if we start with its admittance normalized to $G = R^{-1}$.

which can be solved to give the other two roots of (2.62).

$$\omega_{3,4} = -\frac{\omega_o}{2Q_s} \pm \frac{1}{2} \sqrt{\left(\frac{\omega_o}{Q_s}\right)^2 + 4\omega_o^2} \quad (2.67)$$

Out of the four roots of (2.62), two are positive and two are negative. To find the 3dB bandwidth we simply find the difference between the two positive roots, ω_1 and ω_3 .

$$\begin{aligned} \Delta\omega_{3dB} &= \omega_1 - \omega_3 \\ &= \frac{\omega_o}{2Q_s} - \left(-\frac{\omega_o}{2Q_s}\right) \\ &= \frac{\omega_o}{Q_s} \end{aligned} \quad (2.68)$$

2.B Derivation of Distributed Resonant Tank Bandwidth

In this Appendix we will derive the relationship between the bandwidth and quality factor of a uniform transmission line resonator. We begin with the input impedance of a shorted quarter-wave transmission line tank ($\ell_n = 1/4$) given by (2.38).

$$\begin{aligned} Z_{in} &= Z_o \tanh \left[2\pi\ell_n \left(\frac{1}{2Q_{r,o}} + j\frac{\omega}{\omega_o} \right) \right] \\ &= Z_o \tanh \left(\frac{\pi}{4Q_{r,o}} + j\frac{\pi}{2} \frac{\omega}{\omega_o} \right) \end{aligned} \quad (2.69)$$

First, we expand the $\tanh(\cdot)$ term using trigonometric identities

$$\begin{aligned} Z_{in} &= Z_o \frac{\tanh\left(\frac{\pi}{4Q_{r,o}}\right) + \tanh\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)}{1 + \tanh\left(\frac{\pi}{4Q_{r,o}}\right) \tanh\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)} \\ &= Z_o \frac{\tanh\left(\frac{\pi}{4Q_{r,o}}\right) + j \tan\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)}{1 + j \tanh\left(\frac{\pi}{4Q_{r,o}}\right) \tan\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)} \end{aligned} \quad (2.70)$$

Next, we multiply both the numerator and denominator by $-j \cot\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)$ yielding

$$Z_{in} = Z_o \frac{1 - j \tanh\left(\frac{\pi}{4Q_{r,o}}\right) \cot\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)}{\tanh\left(\frac{\pi}{4Q_{r,o}}\right) - j \cot\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)} \quad (2.71)$$

Assuming $Q_{r,o} \gg 1$, we can make the simplification that $\tanh\left(\frac{\pi}{4Q_{r,o}}\right) \approx \frac{\pi}{4Q_{r,o}}$

$$Z_{in} \approx Z_o \frac{1 - j \frac{\pi}{4Q_{r,o}} \cot\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)}{\frac{\pi}{4Q_{r,o}} - j \cot\left(\frac{\pi}{2} \frac{\omega}{\omega_o}\right)} \quad (2.72)$$

Since we are interested in the frequency region near resonance, we can now introduce a new variable, $\delta\omega$, which is a small deviation in frequency from ω_o .

$$\frac{\omega}{\omega_o} = 1 + \frac{\delta\omega}{\omega_o} \quad (2.73)$$

This new variable allows us to rewrite (2.72) and use $\tan\left(\frac{\pi}{2} \frac{\delta\omega}{\omega_o}\right) \approx \frac{\pi}{2} \frac{\delta\omega}{\omega_o}$ since we are assuming that $\delta\omega \ll \omega_o$

$$\begin{aligned} Z_{in} &\approx Z_o \frac{1 - j \frac{\pi}{4Q_{r,o}} \cot\left(\frac{\pi}{2} + \frac{\pi}{2} \frac{\delta\omega}{\omega_o}\right)}{\frac{\pi}{4Q_{r,o}} - j \cot\left(\frac{\pi}{2} + \frac{\pi}{2} \frac{\delta\omega}{\omega_o}\right)} \\ &= Z_o \frac{1 + j \frac{\pi}{4Q_{r,o}} \tan\left(\frac{\pi}{2} \frac{\delta\omega}{\omega_o}\right)}{\frac{\pi}{4Q_{r,o}} + j \tan\left(\frac{\pi}{2} \frac{\delta\omega}{\omega_o}\right)} \\ &\approx Z_o \frac{1 + j \frac{\pi}{4Q_{r,o}} \frac{\pi}{2} \frac{\delta\omega}{\omega_o}}{\frac{\pi}{4Q_{r,o}} + j \frac{\pi}{2} \frac{\delta\omega}{\omega_o}} \end{aligned} \quad (2.74)$$

Finally, remembering that $Q_{r,o} \gg 1$ and $\delta\omega \ll \omega_o$ we can make one further simplification to Z_{in}

$$\begin{aligned} Z_{in} &\approx \frac{Z_o}{\frac{\pi}{4Q_{r,o}} + j \frac{\pi}{2} \frac{\delta\omega}{\omega_o}} \\ &= \frac{\frac{4}{\pi} Q_{r,o} Z_o}{1 + j 2 Q_{r,o} \frac{\delta\omega}{\omega_o}} \end{aligned} \quad (2.75)$$

We now see that at resonance, $Z_{in} = \frac{4}{\pi} Q_{r,o} Z_o$. To find the 3dB bandwidth we would like to find $\delta\omega$ for which the input impedance drops by 3dB on either side. Or, in other words

$$\begin{aligned} \frac{|Z_{in}|^2}{\left(\frac{4}{\pi} Q_{r,o} Z_o\right)^2} &= \frac{1}{2} \\ \frac{1}{|1 + j 2 Q_{r,o} \frac{\delta\omega}{\omega_o}|^2} &= \frac{1}{2} \\ 1 + 4 Q_{r,o}^2 \left(\frac{\delta\omega}{\omega_o}\right)^2 &= 2 \\ \delta\omega &= \frac{\pm \omega_o}{2 Q_{r,o}} \end{aligned} \quad (2.76)$$

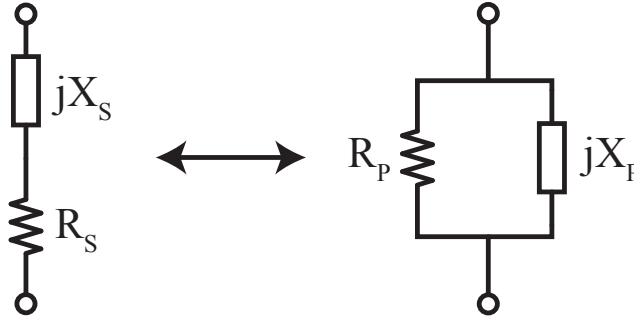


Figure 2.27: Series and parallel representations of a complex impedance.

To arrive at the final result, the 3dB bandwidth is the difference between the positive and negative values of $\delta\omega$. Notice that (2.77) is identical to (2.68).

$$\begin{aligned}\Delta\omega_{3dB} &= \delta\omega_p - \delta\omega_n \\ &= \frac{\omega_o}{Q_{r,o}}\end{aligned}\quad (2.77)$$

2.C Series-to-Parallel Transformation

In this Appendix we will derive a very useful transformation between series and parallel representations of a complex impedance (Fig. 2.27). This transformation is only valid at one frequency making it useful only for narrowband approximations. Nevertheless, most systems of interest are narrowband so this transformation is used extensively in matching network and resonant tank design. The transformation is performed by setting the impedance of the series representation equal to that of the parallel representation at one frequency.

$$\begin{aligned}R_s + jX_s &= \frac{jR_p X_p}{R_p + jX_p} \\ R_s + jX_s &= \frac{R_p X_p^2}{R_p^2 + X_p^2} + j \frac{R_p^2 X_p}{R_p^2 + X_p^2}\end{aligned}\quad (2.78)$$

Equating the real and imaginary parts of the right and left sides of (2.78) leads to

$$R_s = \frac{R_p X_p^2}{R_p^2 + X_p^2}\quad (2.79)$$

$$X_s = \frac{R_p^2 X_p}{R_p^2 + X_p^2}\quad (2.80)$$

Next, for the two representations to be equivalent, their Q must also be equal.

$$Q_s = Q_p = Q\quad (2.81)$$

The quality factor can be written in many equivalent ways for the series and parallel networks at the frequency of interest, ω_o .

$$Q_s = \frac{\omega_o L_s}{R_s} = \frac{1}{\omega_o C_s R_s} = \frac{X_s}{R_s} \quad (2.82)$$

$$Q_p = \frac{\omega_o C_p}{G_p} = \frac{R_p}{\omega_o L_p} = \frac{R_p}{X_p} \quad (2.83)$$

Thus, we can use (2.83) and (2.82) to simplify (2.79) and (2.80).

$$R_p = R_s (1 + Q^2) \quad (2.84)$$

$$X_p = X_s (1 + Q^{-2}) \quad (2.85)$$

Finally, for $Q \gg 1$, R_p and X_p simplify to

$$R_p \approx R_s Q^2 \quad (2.86)$$

$$X_p \approx X_s \quad (2.87)$$

Chapter 3

Voltage Controlled Oscillator

In this chapter, we will describe the design of voltage controlled oscillators (VCOs) at 60GHz. We will begin by introducing the concept of oscillators and their important metrics before delving into implementation challenges through some sample designs. First, we will discuss the design of a cross-coupled oscillator in Section 3.2, followed by an overview of other oscillator topologies of interest in Section 3.3. In these sections we will focus on the general design methodology, including ensuring startup and achieving tuning range and phase noise requirements. We will then explore the push-push architecture in Section 3.5. This architecture allows the use of a lower frequency core, enabling high frequency generation even in older technologies. Finally, in Section 3.6, we will present the design and measurement of both a fundamental mode and push-push VCO for 60GHz LO generation in 65nm and 90nm technologies, respectively.

3.1 A Short Introduction to Oscillators

An oscillator is a regenerative system which utilizes positive feedback to induce instability and sustain stable oscillation. A general oscillator can be analyzed using one of two methods: feedback and negative resistance. A generalized feedback system is shown in Fig. 3.1a and the gain from input to output is given by (3.1).

$$H(s) = \frac{A(s)}{1 - A(s)f(s)} \quad (3.1)$$

In order for this system to allow oscillation, it must provide an output with no input. In other words, the gain must be infinite. From (3.1), the only way for this to occur is if the loop gain, $A(s)f(s)$, at a given frequency is equal to 1 with a phase shift of 0° (or a multiple of 360°). This is called Barkhausen's criteria [Gonzalez97] and is a necessary but *not* sufficient condition for oscillation to actually occur. In fact, for oscillation to start up, the closed loop transfer function must have a pair of complex conjugate poles in the

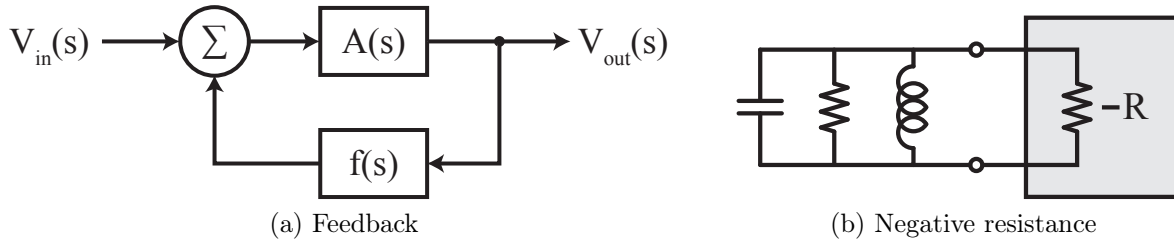


Figure 3.1: Mechanisms of oscillation.

Right-Hand Plane (*RHP*), leading to an exponentially increasing sinusoidal oscillation. As the oscillation amplitude grows, the compressive nonlinearities in the active devices reduce the effective gain, limiting the amplitude. At steady-state, the effective loop gain is exactly equal to 1, meeting Barkhausen's criteria. However, these criteria may be met at many frequencies so, generally, at mm-wave frequencies a resonant tank is used as a filter in the feedback path to set the frequency of oscillation accurately.

This also leads us to another equivalent way of looking at oscillators. Most topologies, especially those utilized at mm-wave frequencies, can be separated into two parts: an active negative resistance generator which, as its name implies, generates a negative resistance at its terminals, and a passive resonant tank, as shown in Fig. (3.1b). If the negative resistance generated is larger than the real part of the tank impedance, the system will oscillate at the resonant frequency of the tank.

3.2 Design of a Cross-Coupled Oscillator

The design of voltage controlled oscillators involves trade-offs between power consumption, phase noise, tuning range and output power. It is important to understand these trade-offs in order to optimize a design for a given application. We begin by studying a cross-coupled differential pair VCO whose schematic is shown in Fig. 3.2. The parallel LC tank provides a resonant load, setting the frequency of oscillation. The cross-coupled differential pair presents a negative resistance which cancels out the losses in the tank, represented by R_p , to allow sustained oscillation [Razavi98, p. 227]. The oscillation frequency is given by (3.2) where L_p is the tank inductance and C_p is the total tank capacitance made up of fixed and variable capacitors which will be described later in Section 3.2.2.

$$\omega_o = \frac{1}{\sqrt{L_p C_p}} = \frac{1}{\sqrt{L_p (C_{fixed} + C_{var})}} \quad (3.2)$$

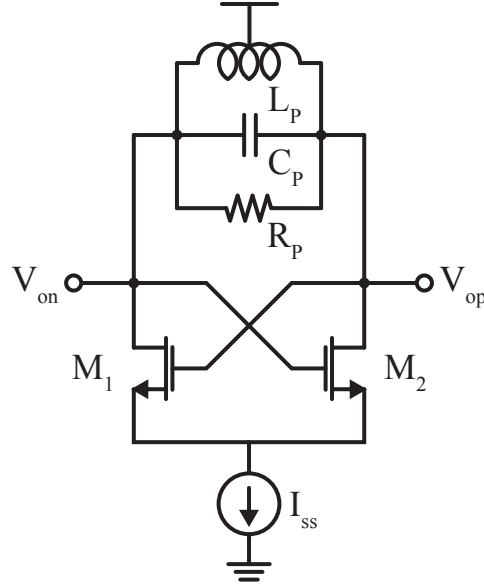


Figure 3.2: Cross-coupled differential pair VCO with tuning.

3.2.1 Startup Conditions

To determine the startup conditions necessary for oscillation, this structure can be analyzed either using feedback or negative resistance approaches. Using the feedback approach, we can start by assuming that all of the losses present in the tank are lumped into the resistor R_p . Furthermore, due to the differential operation we can build a differential mode half circuit by splitting the tank in half at the center tap and assuming the common source point is a differential ground. This leads to a load at the drain of each transistor of

$$Z_{t,H} = \left(\frac{j\omega L_p}{2} \right) \parallel \left(\frac{1}{j\omega 2C_p} \right) \parallel \left(\frac{R_p}{2} \right) \quad (3.3)$$

The loop gain is then directly given by (3.4) and must be greater than 1 at the resonance frequency to allow oscillation to start up.

$$A_l(\omega_o) = \left(g_m Z_{t,H} \Big|_{\omega=\omega_o} \right)^2 \geq 1 \quad (3.4)$$

At resonance $Z_{t,H} = R_p/2$ thus leading to (3.5).

$$\begin{aligned} \left(\frac{g_m R_p}{2} \right)^2 &\geq 1 \\ g_m &\geq \frac{2}{R_p} \end{aligned} \quad (3.5)$$

Another way of stating (3.5) is that the small-signal gain, $g_m R_p$ must be greater than 2.

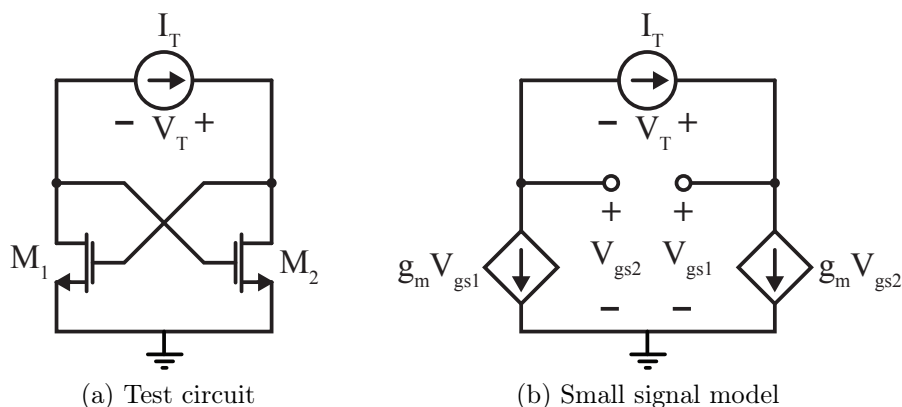


Figure 3.3: Cross-coupled differential pair input impedance.

We can also arrive at the same result by using a negative resistance approach. The total differential tank impedance is given by

$$Z_t = j\omega L_p \parallel \left(\frac{1}{j\omega C_p} \right) \parallel R_p \quad (3.6)$$

which, at resonance, reduces to $Z_t = R_p$. Next, we find the resistance presented by the cross-coupled differential pair by applying a test current source, I_T as in Fig. 3.3a. This circuit can then be solved using the small signal model shown in Fig. 3.3b. First, we can see directly that

$$V_T = V_{gs1} - V_{gs2} \quad (3.7)$$

Second, there is only one loop present in the circuit so

$$I_T = g_m V_{gs2} = -g_m V_{gs1} \quad (3.8)$$

which leads to

$$V_{gs2} = -V_{gs1}. \quad (3.9)$$

Substituting (3.9) into (3.7) and using the result along with (3.8) will now allow us to find the input resistance.

$$\begin{aligned} R_{in} &= \frac{V_T}{I_T} \\ &= \frac{V_{gs1} - (-V_{gs1})}{-g_m V_{gs1}} \\ &= \frac{-2}{g_m} \end{aligned} \quad (3.10)$$

In order for the oscillator to start up, the negative admittance presented by the cross-coupled differential pair must be larger than the positive admittance presented by the tank.

$$\begin{aligned} \left| \frac{1}{R_{in}} \right| &\geq \frac{1}{R_p} \\ \frac{g_m}{2} &\geq \frac{1}{R_p} \\ g_m &\geq \frac{2}{R_p} \end{aligned} \quad (3.11)$$

The result in (3.11) agrees with (3.5) showing the validity of both approaches. Finally, since g_m is proportional to the bias current of the device,

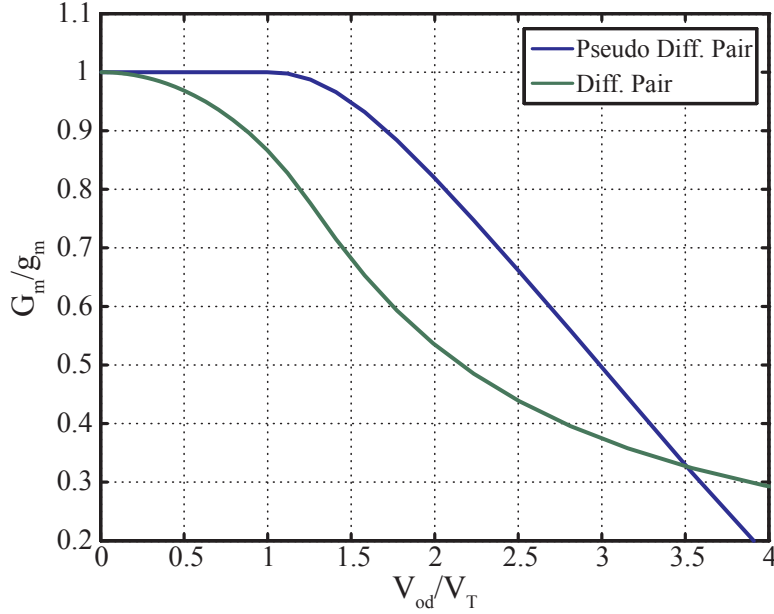
$$g_m = \begin{cases} \frac{2I_d}{V_{ov}} & \text{long - channel} \\ \frac{2I_d}{V_{sat}} & \text{short - channel} \end{cases} \quad (3.12)$$

the result in (3.11) implies that a higher R_p will lead to reduced power consumption by allowing the use of a smaller g_m . For example, in a differential pair with long channel devices

$$\begin{aligned} g_m &\geq \frac{2}{R_p} \\ \frac{I_{ss}}{V_{ov}} &\geq \frac{2}{R_p} \end{aligned} \quad (3.13)$$

As long as the above criteria for g_m are met, any noise present in the system near the resonance frequency will be amplified and grow exponentially due to the presence of two complex conjugate poles in the *RHP*. While the amplitude is small, the active devices behave like small signal amplifiers. As the amplitude grows, nonlinearities begin to come into effect. This is true even if the core consists of ideal square law transistors. As an example, let us examine a cross-coupled pseudo-differential pair like the one shown in Fig. 3.3a. As long as the output swing remains within one threshold ($V_{od} \leq V_T$) and less than $2V_{ov}$, M_1 and M_2 remain in saturation and the effective transconductance, G_m , is equal to the small-signal transconductance, g_m . When the output swing exceeds V_T , the transistors are pushed into the triode region and G_m drops with increasing amplitude. For $V_{od} > 2V_{ov}$ one device will be completely cutoff for part of the cycle while the current in the other device will continue to increase. If $2V_{ov} > V_T$, the devices will never enter the triode region and will go directly into cutoff.

In a true differential pair there is a current source setting the total bias current as in Fig. 3.2. In that case, for small input signals ($V_{od} \leq \sqrt{2}V_{ov}, V_T$) both devices remain in saturation

Figure 3.4: Large signal G_m .

and G_m is given by

$$\begin{aligned}
 G_m \Big|_{V_{od} \leq \sqrt{2}V_{ov}, V_T} &= \frac{k' W}{2 L} \sqrt{\frac{4I_{ss}}{k' (W/L)} - V_{od}^2} \\
 &= \frac{I_{ss}}{2V_{ov}^2} \sqrt{(2V_{ov})^2 - V_{od}^2}
 \end{aligned} \tag{3.14}$$

For very large V_{od} all of the bias current will be steered completely to one side then the other during the cycle. In this case, G_m is given by

$$G_m \Big|_{V_{od} \gg \sqrt{2}V_{ov}, V_T} = \frac{I_{ss}}{V_{od}} \tag{3.15}$$

The transition region again depends on the value of V_T relative to V_{ov} . If $\sqrt{2}V_{ov} \geq V_T$, as V_{od} increases above V_T , the devices will go into the triode region for part of each cycle. When V_{od} rises above $\sqrt{2}V_{ov}$ one device will be completely cutoff for part of the cycle, forcing the entire bias current to flow in the other device. For $\sqrt{2}V_{ov} > V_T$ the devices never enter the triode region, they go directly into cutoff when $V_{od} > \sqrt{2}V_{ov}$. The large signal G_m normalized to the small-signal g_m is plotted versus to the oscillation amplitude normalized to V_T for both a pseudo-differential pair core and a standard differential pair core in Fig. 3.4.

Based on the previous discussion we can now see that during startup the cross-coupled core starts off in the small-signal regime with maximum G_m . The amplitude then grows until G_m drops to the value where the loop gain is exactly equal to 1. In other words, when the large

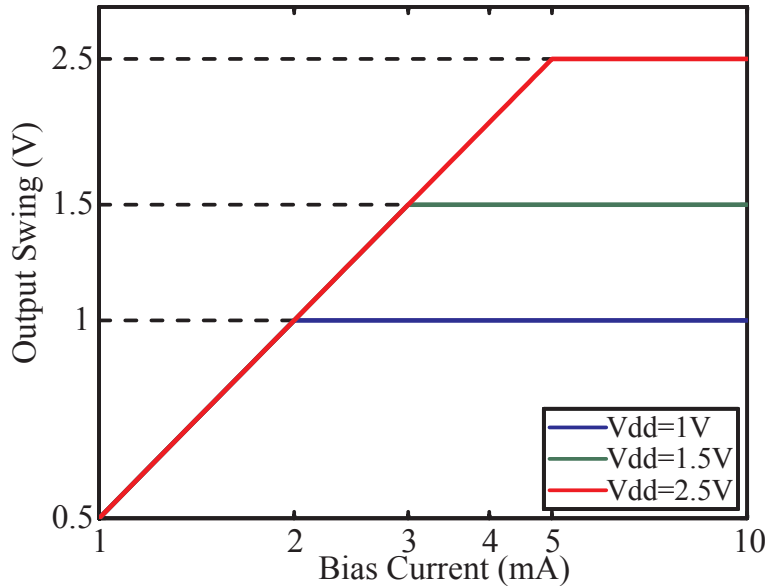


Figure 3.5: Current limited vs. voltage limited operation. ($Z_o = 50, Q_T = 10$)

signal transconductance, G_m , meets the condition

$$G_m = \frac{2}{R_p} \quad (3.16)$$

This is a stable point of operation. If the amplitude were to increase beyond this point, G_m would reduce forcing the amplitude to reduce. On the other hand, if the amplitude were to reduce below this point, G_m would increase, forcing the amplitude to increase. This nonlinear behavior acts like inherent negative feedback, stabilizing the amplitude of oscillation. The ratio g_m/G_m is also known as the safety factor, denoted η in this work, and is usually chosen to be at least 2-3. This ensures that there is sufficient gain in the system to ensure reliable startup even in the presence of excess loss.

With proper design, the voltage amplitude in steady-state should be high enough to completely switch the current from one side to the other of the differential pair with each cycle. At low frequencies this makes the current injected into the tank a square wave that switches between I_{ss} and $-I_{ss}$. The resonant tank then filters out and selects only the fundamental harmonic making the output voltage

$$\begin{aligned} V_{od,LF} &= I_{sw} \Big|_{\omega=\omega_o} Z_T(\omega_o) \\ &\approx \frac{4}{\pi} I_{ss} R_p \end{aligned} \quad (3.17)$$

Unfortunately, at high frequencies finite switching time and low device gain make the current waveform look more like a sinusoid with amplitude I_{ss} rather than a square wave, leading to

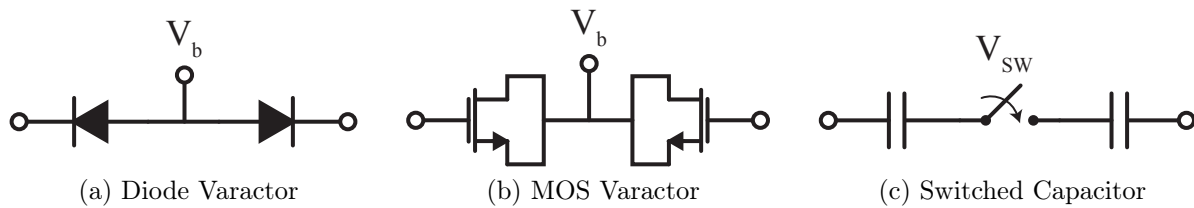


Figure 3.6: Variable capacitor architectures.

a smaller output voltage amplitude.

$$V_{od,HF} \approx I_{ss} R_p \quad (3.18)$$

Thus, increasing either the bias current or R_p will lead to higher output amplitude. However, there is a fundamental limit to the output voltage swing which is imposed by the supply voltage. Since the drain of each transistor is tied to the supply through an inductor, the drain voltage can swing above V_{dd} . Assuming a sinusoidal output voltage, the maximum the drain of each transistor can reach is $2V_{dd}$. If I_{ss} is increased beyond the point where this voltage swing is reached on the drain, then the output voltage will no longer increase. This gives rise to two regions of operation [Hajimiri99]: the current limited domain and the voltage domain. In the current limited domain the output amplitude is proportional to the bias current and is not a function of the supply voltage (ignoring the channel-length modulation effects). In the voltage limited domain the output amplitude is proportional to the supply voltage and is not a function of the bias current as long as that bias current is high enough to maintain operation within this domain. A representative plot of output amplitude versus bias current is plotted in Fig. 3.5 for different values of supply voltage.

3.2.2 Tuning the Tank

The total capacitance in the tank can be separated into fixed and variable components.

$$C_p = C_{fixed} + C_{var} \quad (3.19)$$

C_{fixed} is made up of the parasitic and layout capacitances from the core devices plus any fixed capacitance added to the tank, C_T , such as from a buffer or divider which loads the VCO output.

$$C_{fixed} = C_T + \left(\frac{C_{gs}}{2} + \frac{C_{ds}}{2} + 2C_{gd} \right) \quad (3.20)$$

C_{var} , on the other hand, is the variable capacitance which is added to tank for tuning the resonant frequency. This variable capacitance can be implemented as a varactor (Figs. 3.6a or 3.6b), switched capacitor bank (Fig. 3.6c), or some combination thereof. The two varactor structures (Figs. 3.6a or 3.6b) can be controlled with either an analog or digital

control signal. Analog control of V_b allows any value of capacitance between the minimum and maximum values of varactor capacitance to be achieved, allowing fine control of the VCO output frequency. Digital control on the other hand, setting V_b only to either V_{dd} or ground, provides only the minimum and maximum values of varactor capacitance and is thus very similar to the switched capacitor circuit in Fig. 3.6c. This allows for coarse tuning of the VCO frequency in steps determined by the size of the switched capacitance or, in the case of a switched varactor, the on-off ratio of capacitance. Generally, arrays of switched capacitor structures are used in conjunction with an analog varactor. The VCO frequency is first coarsely set with the digitally controlled capacitors and then finely set with the analog varactor. As we will see in Chapter 4, it is advantageous from a noise standpoint to use a large bank of switched capacitors along with a smaller analog varactor to cover a large tuning range, rather than one large analog varactor. Regardless of the actual implementation, C_{var} will have some minimum fixed capacitance, $C_{v,o}$, along with its variable component, ΔC_v .

$$C_{var} = C_{v,o} + \Delta C_v \quad (3.21)$$

For overall tuning range calculations, ΔC_v includes both the analog variable capacitance as well as any digitally switched capacitance while $C_{v,o}$ is the capacitance presented to the tank when all variable and switched capacitors are tuned to their lowest settings. $C_{v,o}$ is then just another fixed capacitance which is always present, just like C_{fixed} .

$$\begin{aligned} C_{min} &= C_{fixed} + C_{v,o} \\ &= C'_{fixed} \end{aligned} \quad (3.22)$$

$$\begin{aligned} C_{max} &= C_{fixed} + C_{v,o} + \Delta C_v \\ &= C'_{fixed} + \Delta C_v \end{aligned} \quad (3.23)$$

The tuning range of the tank can then be defined as

$$\begin{aligned} TR &= \frac{\omega_{max}}{\omega_{min}} \\ &= \sqrt{\frac{C_{max}}{C_{min}}} \\ &= \sqrt{\frac{C'_{fixed} + \Delta C_v}{C'_{fixed}}} \\ &= \sqrt{1 + \frac{\Delta C_v}{C'_{fixed}}}. \end{aligned} \quad (3.24)$$

To achieve high tuning range we would therefore like to maximize the ratio $\Delta C_v/C'_{fixed}$. Initially, this may seem trivial to accomplish by adding arbitrary amounts of variable capacitance while lowering the tank inductance to maintain the same ω_o . However, as seen from (3.5), we would like to maximize the tank impedance at resonance, R_p , in order to get high current efficiency. This impedance is a function of the losses of the inductor and capacitors

as defined by their respective quality factors. Thus, we can now derive an equation for the resistor, R_p , the aggregate of the losses in the tank, as a function of the component quality factors as shown in (3.25)

$$\begin{aligned}
 R_p &= \frac{1}{G_p} \\
 &= (G_{p,C} + G_{p,L})^{-1} \\
 &= \left(\frac{\omega_o C_p}{Q_C} + \frac{1}{\omega_o L_p Q_L} \right)^{-1} \\
 &= \sqrt{\frac{L_p}{C_p}} \left(\frac{1}{Q_C} + \frac{1}{Q_L} \right)^{-1} \\
 &= Z_o \left(\frac{1}{Q_C} + \frac{1}{Q_L} \right)^{-1}
 \end{aligned} \tag{3.25}$$

where we have used the series to parallel transformation, described in Appendix 2.C, under the assumption that $Q_L \gg 1$ and $Q_C \gg 1$. Thus, from (3.25) we can define a tank quality factor, Q_T , which is equal to the parallel combination of the component quality factors.

$$Q_T = \left(\frac{1}{Q_L} + \frac{1}{Q_C} \right)^{-1} \tag{3.26}$$

The total parallel resistance at resonance is then given by

$$R_p = Q_T Z_o \tag{3.27}$$

Notice that this agrees with the input impedance of a parallel tank at resonance given by (2.14). If the capacitor loss dominates the tank,¹ $Q_C \ll Q_L$ and (3.25) simplifies to

$$R_p \approx Q_C \sqrt{\frac{L_p}{C_p}} = Q_C Z_o \tag{3.28}$$

In order to maintain low power consumption we would like to maximize R_p . Clearly, increasing the component quality factors helps, but (3.25) and (3.28) also tell us that for low power consumption at a given resonant frequency, ω_o , we would like to maximize L_p and minimize C_p (in other words maximize Z_o). This leads to a trade-off between tuning range and power consumption since maximizing tuning range requires adding more capacitance while decreasing L_p to compensate. Furthermore, as the desired resonance frequency increases, the required capacitance and inductance values reduce and parasitics become a proportionally much larger part of the tank capacitance. At *mm*-wave frequencies, wide tuning range designs are thus much more difficult to achieve than at lower, traditional RF frequencies.

¹At high frequencies the quality factor of inductors tends to be higher than that for capacitors and much larger than for varactors.

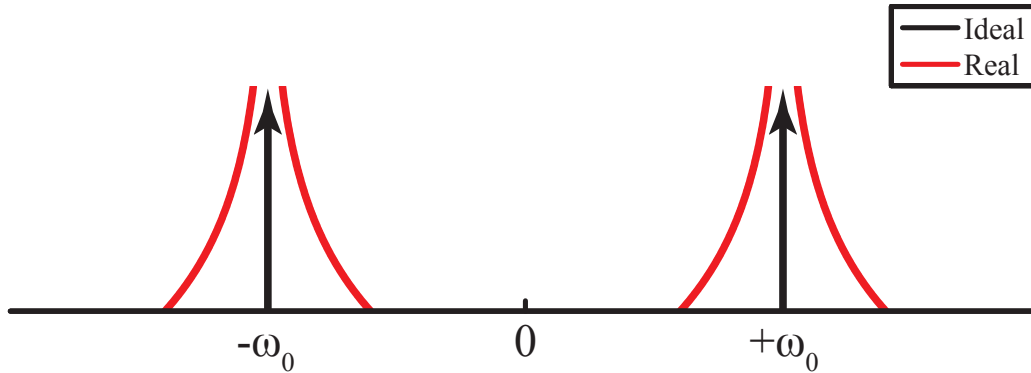


Figure 3.7: Oscillator output spectrum.

3.2.3 Phase Noise

An ideal RF oscillator produces an output sinusoid at a fixed frequency with a constant amplitude.

$$V_{out,i}(t) = A \cos(\omega_o t + \phi) \quad (3.29)$$

This ideal oscillator appears as two impulses at $\pm\omega_o$ in the frequency domain. Any real oscillator, however, will have active and passive devices which introduce noise into the system, disturbing the amplitude and phase of the oscillator output. Furthermore, the actual waveform of the output may be rich in harmonics, and will not necessarily be a sinusoid.² Thus, a general oscillator output looks like

$$V_{out}(t) = A(t) f(\omega_o t + \phi(t)) \quad (3.30)$$

where $A(t)$ and $\phi(t)$ are functions of time which represent amplitude noise and phase noise respectively, and $f(t)$ is a periodic function with a period of 2π . At *mm*-wave frequencies, the power that can be generated at higher level harmonics is very limited, so oscillator outputs can be approximated by a sinusoid instead of the more general function f . Furthermore, well designed oscillators will have amplitude limiting mechanisms which reduce the effect of noise on the output amplitude. This can be accomplished either through an explicit Automatic Gain Control (AGC) loop, or simply through negative feedback inherent in the device nonlinearity as described in Section 3.2.1. Unfortunately, there is no such mechanism to control the phase of the oscillator. Variations in phase, $\phi(t)$, look like variations in the instantaneous frequency of the oscillator. This causes spreading in the spectrum of the oscillator creating “skirts” as shown in Fig. 3.7.

We will now derive the phase noise spectrum due to noise sources present in an oscillator using the negative resistance model of Fig. 3.1b following a procedure similar to that found in [Lee00]. Out of all the passive components, only resistors can contribute noise. Ideal capacitors and inductors do not contribute any noise but their presence does shape the noise

²Ring oscillators, for example, will produce an output closer to a square wave.

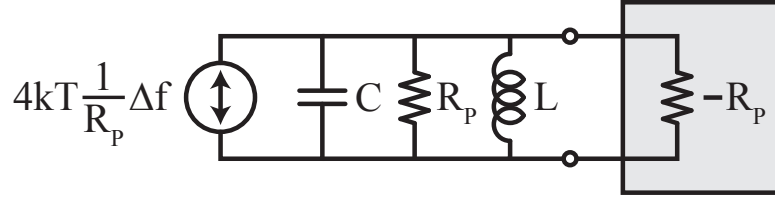


Figure 3.8: Oscillator LTI noise model.

spectrum by the filtering action they create. Active devices in the core do contribute noise which can be treated just like the noise from the tank resistance. Therefore, in the analysis to follow we will initially focus only on the noise due to the tank resistance and then expand to include active noise sources. Thus, the only noise source is white noise from R_p as shown in Fig. 3.8. To find the equivalent noise voltage density we multiply the noise current density by the square of the magnitude of the tank impedance.

$$\begin{aligned} \frac{\overline{v_n^2}}{\Delta f} &= \frac{\overline{i_n^2}}{\Delta f} |Z_T|^2 \\ &= 4kT \frac{1}{R_p} |Z_T|^2 \end{aligned} \quad (3.31)$$

Since the active devices generate an impedance exactly equal to $-R_p$, the effective impedance of the tank actually looks like an ideal LC tank.

$$\begin{aligned} Z_T(\omega) &= \frac{1}{j\omega C + \frac{1}{j\omega L}} \\ &= \frac{j\omega L}{1 - \left(\frac{\omega}{\omega_o}\right)^2} \end{aligned} \quad (3.32)$$

We are interested in the spectrum at an offset, $\Delta\omega$, from the resonance frequency, ω_o , so we can rewrite (3.32) as

$$\begin{aligned} Z_T(\omega_o + \Delta\omega) &= \frac{j(\omega_o + \Delta\omega)L}{1 - \left(1 + \frac{\Delta\omega}{\omega_o}\right)^2} \\ &= \frac{j(\omega_o + \Delta\omega)L}{1 - \left[1 + \frac{2\Delta\omega}{\omega_o} + \left(\frac{\Delta\omega}{\omega_o}\right)^2\right]} \\ &= \frac{-j\omega_o L \left(1 + \frac{\Delta\omega}{\omega_o}\right)}{\frac{\Delta\omega}{\omega_o} \left(2 + \frac{\Delta\omega}{\omega_o}\right)} \end{aligned} \quad (3.33)$$

However, since we are only interested in a very small offset relative to the resonance frequency,

$\Delta\omega \ll \omega_o$ and (3.33) can be simplified.

$$Z_T(\omega_o + \Delta\omega) \approx -\frac{j\omega_o L}{2\frac{\Delta\omega}{\omega_o}} \quad (3.34)$$

The definition of the parallel tank quality factor given by (2.11) now allows us to write (3.34) in terms of Q_T rather than the tank inductance.

$$Z_T(\omega_o + \Delta\omega) = -j\frac{\omega_o R_p}{2Q_T \Delta\omega} \quad (3.35)$$

The equivalent output noise voltage is then given by plugging (3.35) back into (3.31).

$$\begin{aligned} \frac{\overline{v_n^2}}{\Delta f} &= 4kT \frac{1}{R_p} \left(\frac{\omega_o R_p}{2Q_T \Delta\omega} \right)^2 \\ &= 4kT R_p \left(\frac{\omega_o}{2Q_T \Delta\omega} \right)^2 \end{aligned} \quad (3.36)$$

This noise will cause variations in both phase and amplitude but, as mentioned earlier, any real oscillator design will damp out or limit amplitude noise so we are only interested in the phase component. Assuming the output waveform is a sinusoid, the noise power will be equally split between the phase and amplitude according to the equipartition theorem of thermodynamics.

$$\left. \frac{\overline{v_n^2}}{\Delta f} \right|_{phase} = 2kT R_p \left(\frac{\omega_o}{2Q_T \Delta\omega} \right)^2 \quad (3.37)$$

Finally, we can find the phase noise, $\mathcal{L}\{\Delta\omega\}$, as the ratio of noise to signal power reported in dBc/Hz . It can be found by normalizing the rms noise voltage to the rms amplitude squared of the signal ($V_{o,rms}^2 = |V_o|^2/2$) and taking the \log as shown in (3.38). In this form we can see that reducing R_p or increasing $V_{o,rms}$ leads to a reduction in phase noise, setting up a trade-off between power consumption and phase noise.

$$\mathcal{L}\{\Delta\omega\} = 10 \log \left[\frac{2kT R_p}{V_{o,rms}^2} \left(\frac{\omega_o}{2Q_T \Delta\omega} \right)^2 \right] \quad (3.38)$$

To make (3.38) more general we can rewrite it in terms of the signal power, $P_{sig} = V_{o,rms}^2/R_p$, resulting in (3.39).

$$\mathcal{L}\{\Delta\omega\} = 10 \log \left[\frac{2kT}{P_{sig}} \left(\frac{\omega_o}{2Q_T \Delta\omega} \right)^2 \right] \quad (3.39)$$

From (3.39) we can see that phase noise near the carrier has a $1/(\Delta\omega)^2$ dependence on frequency. However, we have ignored other sources of noise in this analysis. In reality active devices will add excess white noise and also exhibit $1/f$ noise for very small offsets which in turn leads to a $1/|\Delta\omega|^3$ region in the phase noise response. Furthermore, the signal coming

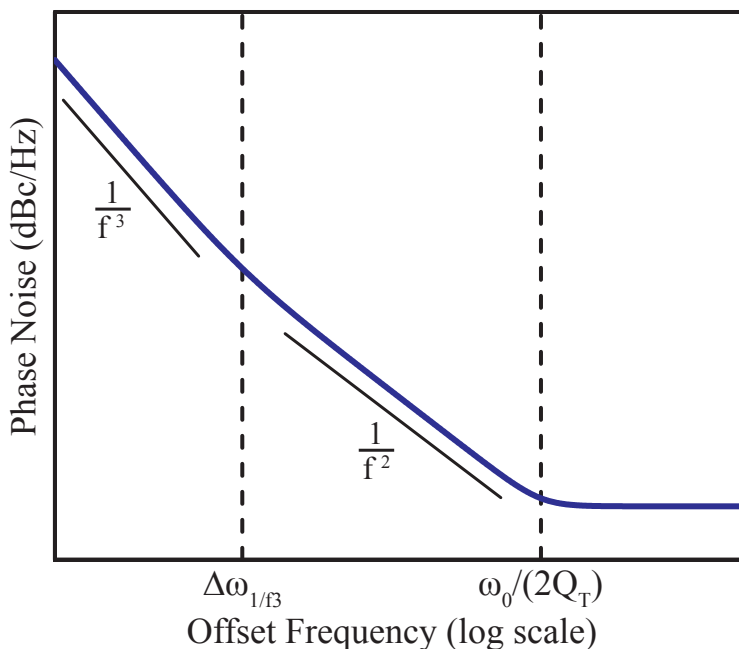


Figure 3.9: Phase noise: Leeson's model.

out of the oscillator usually must be buffered by an amplifier or at least routed to another block. In either case, the thermal noise floor of the amplifier or any lead resistance will introduce a white noise floor for large enough offsets from the carrier. These effects can be added with some minor modifications to (3.39), leading to (3.40).

$$\mathcal{L}\{\Delta\omega\} = 10 \log \left[\frac{2FkT}{P_{sig}} \left\{ 1 + \left(\frac{\omega_o}{2Q_T\Delta\omega} \right)^2 \right\} \left(1 + \frac{\Delta\omega_{1/f^3}}{|\Delta\omega|} \right) \right] \quad (3.40)$$

The factor F accounts for excess noise from active devices and other sources, while $\Delta\omega_{1/f^3}$ is the boundary between the $1/(\Delta\omega)^2$ and $1/|\Delta\omega|^3$ regions. It should be noted here that this boundary does not necessarily occur at the $1/f$ noise corner for the active devices but can in fact be much lower. $\mathcal{L}\{\Delta\omega\}$ is plotted in Fig. 3.9. This representation of phase noise is called Leeson's model, first presented in [Leeson66], and gives us an intuitive framework for making design decisions. Leeson's model shows us that in order to reduce phase noise we must increase the tank Q and the signal power. Both of these points make intuitive sense since increasing tank Q reduces the bandwidth, filtering out more noise, while increasing signal power can be seen as simply increasing SNR .

Unfortunately, this model does have some drawbacks. Leeson's model implies that the noise flattens out beyond $\Delta\omega = \frac{\omega_o}{2Q_T}$. In reality this is not necessarily the correct corner frequency. Furthermore, $\Delta\omega_{1/f^3}$ is an empirical fitting parameter which must be determined through measurements since the above LTI analysis cannot predict frequency translation. Finally, F can be determined for a given topology assuming the same amount of noise is applied by the

active devices at all times (e.g.: Colpitts [Huang98], cross-coupled pair [Rael00]). In reality, the noise is cyclostationary and is not constant throughout the period of oscillation.

A linear time-variant (LTV) approach [Hajimiri98] must be used to take into account the effect of this cyclostationary noise. In this approach, for each noise source, an impulse of current is applied at a given time within the oscillation period, affecting both the amplitude and phase of the output. After enough cycles have passed for the amplitude variation to die out, the phase shift versus an undisturbed oscillator is recorded. The location of the impulse is then varied across the entire oscillation period and its effect on the phase recorded at each point. This effectively constructs a time variant impulse response called the Impulse Sensitivity Function (ISF) which is represented by Γ and is periodic since it repeats along with the oscillator waveform. Another method to find the ISF, also given in [Hajimiri98], involves direct calculation on the oscillator output waveform, f , as represented in (3.30) with $A(t) = 1$ and $\phi(t) = 0$.

$$\Gamma(x) = \frac{f'}{(f')^2 + (f'')^2} \quad (3.41)$$

where f' and f'' are the first and second derivatives of the function f . This allows us to test a very simple yet instructive case: a sinusoidal output waveform. At mm -wave frequencies this is reasonably close to the actual output waveform of practical oscillators since high tank Q and low gain at higher frequencies tend to reject higher order harmonics. In this case

$$\begin{aligned} f &= \cos(x) \\ f' &= -\sin(x) \\ f'' &= -\cos(x) \end{aligned}$$

which results in

$$\begin{aligned} \Gamma(x) &= \frac{-\sin(x)}{\sin^2(x) + \cos^2(x)} \\ \Gamma(x) &= -\sin(x) \\ \Gamma(\omega t) &= -\sin(\omega t) \end{aligned} \quad (3.42)$$

This result, plotted in Fig. 3.10, shows that Γ reaches its peak during the transition region of the output waveform and is minimum during the peaks. The noise injected by active devices is thus most destructive during the transition region and has minimal effect near the peaks. Once we have found the ISF for each noise source in the circuit the phase noise in the $1/(\Delta\omega)^2$ region is calculated by summing together the noise of each source multiplied by the square of the rms value of its associated ISF, $\Gamma_{n,rms}$, and normalizing the result to the square of the maximum charge present in the tank, $q_{max}^2 = V_{tank}^2 C_{tank}^2$. This last step is equivalent to dividing by the signal power as we have done previously to arrive at (3.39). The resulting phase noise as derived in [Hajimiri98] is given by

$$\mathcal{L}\{\Delta\omega\} = 10 \log \left[\frac{1}{2(\Delta\omega)^2} \cdot \frac{1}{q_{max}^2} \cdot \sum_n \left(\frac{\overline{i_n^2}}{\Delta f} \cdot \Gamma_{n,rms}^2 \right) \right] \quad (3.43)$$

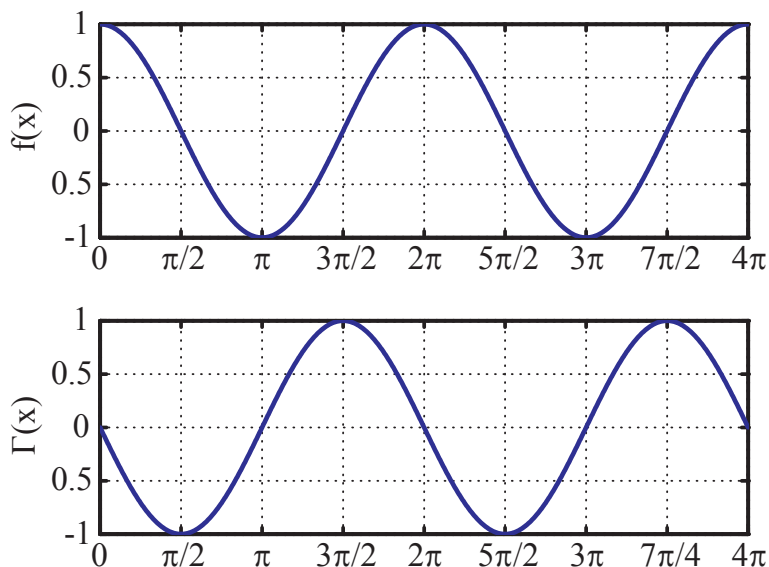


Figure 3.10: Output waveform and ISF of an ideal sinusoidal oscillator.

This approach allows for direct calculation of the phase noise profile and is a very powerful tool but is not always applicable to hand calculation due to Γ being a potentially complicated function. However, under the simplifying assumption of mostly sinusoidal waveforms, $\Gamma(x)$ is itself a sinusoid and $\Gamma_{rms}^2 = 1/2$. The ISF approach then becomes much easier to apply.

One other very important result of this work which should be mentioned involves the $\Delta\omega_{1/f^3}$ corner frequency. From [Hajimiri98, Lee00] we see that

$$\Delta\omega_{1/f^3} = \omega_{1/f} \left(\frac{\Gamma_{dc}}{\Gamma_{rms}} \right)^2 \quad (3.44)$$

where $\omega_{1/f}$ is the device $1/f$ noise corner frequency, Γ_{dc} is the mean value of the ISF, and Γ_{rms} is the rms value of the ISF. Since the rms value of a waveform is always greater than or equal to its average, (3.44) implies that the $\Delta\omega_{1/f^3}$ corner frequency is always less than or equal to the device $1/f$ noise corner frequency and may even be zero for $\Gamma_{dc} = 0$. This means that with proper design (e.g.: a waveform with perfectly symmetric rise and fall), the oscillator may not exhibit a $1/|\Delta\omega|^3$ region at all.

3.2.4 Design Optimization

The foregoing discussion has provided a framework which will allow us to consider optimization of the design parameters. High data rate transceivers require wide signal bandwidths, usually making the noise contribution from the $1/|\Delta\omega|^3$ region of the oscillator phase noise insignificant compared to the $1/(\Delta\omega)^2$ region and the flat wideband noise. Reducing the white noise far from the carrier is accomplished by increasing the size of the oscillator buffer.

On the other hand, reducing the phase noise in the $1/(\Delta\omega)^2$ region can only be accomplished by optimizing the design of the VCO.

To get started we can use the same procedure that allowed us to arrive at (3.38), one similar to, but more general than that presented in [Ham01]. In the previous discussion we have only included noise contributions due to the tank loss. Unfortunately, the active device noise is an equal or even dominant part of the overall noise. First, we must remember that in steady-state the average transconductance of the active devices is reduced just to the point where the loop gain is equal to 1. Thus, regardless of how high g_m is to ensure reliable startup, G_m is only a function of R_P . Conceptually, we can then imagine that the noise of the active devices is a function of G_m rather than the small-signal g_m . Using this simplistic assumption along with (3.16), the noise due to the active devices is then

$$\overline{i_{n,MOS}^2} = 4kT\gamma G_m \Delta f = 4kT\gamma \frac{2}{R_P} \Delta f \quad (3.45)$$

In the above equation, γ is equal to $2/3$ for long-channel devices and can be assumed to be 2 for short-channel devices. On the other hand, the noise current due to R_p is given by

$$\overline{i_{n,Rp}^2} = 4kT \frac{1}{R_P} \Delta f \quad (3.46)$$

The ratio of the two noise sources is then

$$\frac{\overline{i_{n,MOS}^2}}{\overline{i_{n,Rp}^2}} = 2\gamma \quad (3.47)$$

The overall noise is thus only a function of R_P regardless of the small-signal g_m of the device or, for that matter, regardless of the type of device (e.g.: MOS, BJT, Tube).

It may seem here that we have oversimplified the very complex situation of a time varying system and arrived at an incorrect result. In fact, more sophisticated analyses by [Rael00] (using mixer noise theory) as well as [Andreani05] and [Mazzanti08] (using the ISF approach) have shown that the ratio of active device noise to tank noise is indeed $1 : \gamma$ in the current limited regime. We can then use the more general result from [Rael00] for the factor F in (3.40) assuming that the oscillator waveforms are sinusoidal rather than square and that the tail noise is negligible by design.

$$F = 1 + \gamma \frac{I_{ss} R_P}{V_{od}} \quad (3.48)$$

Using these results we can rewrite (3.40) in the $1/(\Delta\omega)^2$ region

$$\begin{aligned} \mathcal{L}\{\Delta\omega\} &= 10 \log \left[\frac{\left(\overline{i_n^2} / \Delta f \right) F |Z_T|^2}{V_{od}^2} \right] \\ &= 10 \log \left[\frac{kT}{R_P V_{od}^2} \left(1 + \gamma \frac{I_{ss} R_P}{V_{od}} \right) (L\omega_o)^2 \left(\frac{\omega_o}{\Delta\omega} \right)^2 \right] \end{aligned} \quad (3.49)$$

In the current-limited domain $V_{od} \approx I_{ss}R_p$ for sinusoidal waveforms. Using this information along with (2.11) allows us to put (3.49) in terms of only a handful of design parameters.

$$\begin{aligned} \mathcal{L}\{\Delta\omega\} \Big|_{I-lim} &= 10 \log \left[\frac{kT(1+\gamma)}{R_P(I_{ss}R_p)^2} (L\omega_o)^2 \left(\frac{\omega_o}{\Delta\omega}\right)^2 \right] \\ &= 10 \log \left[\frac{kT(1+\gamma)}{I_{ss}^2 Q_T^3 Z_o} \left(\frac{\omega_o}{\Delta\omega}\right)^2 \right] \end{aligned} \quad (3.50)$$

If either I_{ss} or R_p are increased enough such that $I_{ss}R_p = V_{dd}$, the oscillator will enter the voltage limited domain where $V_{od} \approx V_{dd}$. Again, (3.49) can be rewritten for the voltage limited domain using (2.11).

$$\begin{aligned} \mathcal{L}\{\Delta\omega\} \Big|_{V-lim} &= 10 \log \left[\frac{kT}{R_P V_{dd}^2} \left(1 + \gamma \frac{I_{ss}R_p}{V_{dd}}\right) (L\omega_o)^2 \left(\frac{\omega_o}{\Delta\omega}\right)^2 \right] \\ &= 10 \log \left[\frac{kT Z_o}{V_{dd}^2 Q_T} \left(1 + \gamma \frac{I_{ss}Q_T Z_o}{V_{dd}}\right) \left(\frac{\omega_o}{\Delta\omega}\right)^2 \right] \end{aligned} \quad (3.51)$$

We can now summarize our results as follows

$$\mathcal{L}\{\Delta\omega\} \propto \begin{cases} \frac{(1+\gamma)}{I_{ss}^2 Q_T^3 Z_o} & (I - limited) \\ \frac{Z_o}{V_{dd}^2 Q_T} \left(1 + \gamma \frac{I_{ss}Q_T Z_o}{V_{dd}}\right) & (V - limited) \end{cases} \quad (3.52)$$

where the oscillator will operate in the current-limited domain for $I_{ss}Q_T Z_o < V_{dd}$, and in the voltage-limited domain for $I_{ss}Q_T Z_o > V_{dd}$. Generally, V_{dd} and Q_T are out of our control but we can use I_{ss} and Z_o to optimize for minimum phase noise. Since I_{ss} appears in the denominator of (3.52) in the current-limited domain but in the numerator in the voltage-limited domain, the optimum value for I_{ss} for minimum phase noise is that which places the oscillator at the boundary between the two domains.

$$I_{ss,opt} = \frac{V_{dd}}{R_P} = \frac{V_{dd}}{Q_T Z_o} \quad (3.53)$$

By plugging in the boundary condition given by (3.53) into (3.52) we can find the minimum value of phase noise.

$$\mathcal{L}\{\Delta\omega\} \Big|_{min} \propto \frac{Z_o(1+\gamma)}{V_{dd}^2 Q_T} \quad (3.54)$$

Unlike the result in [Ham01], (3.52)-(3.54) are applicable for a general resonant tank regardless of whether its loss is dominated by inductive or capacitive components, a crucial factor for *mm*-wave designs.

We can now revisit Fig. 3.5 and add the phase noise profile as a function of bias current (Fig. 3.11). This is a very powerful result. First, it shows that indeed the optimum phase noise occurs at the boundary between domains of operation. To minimize phase noise for

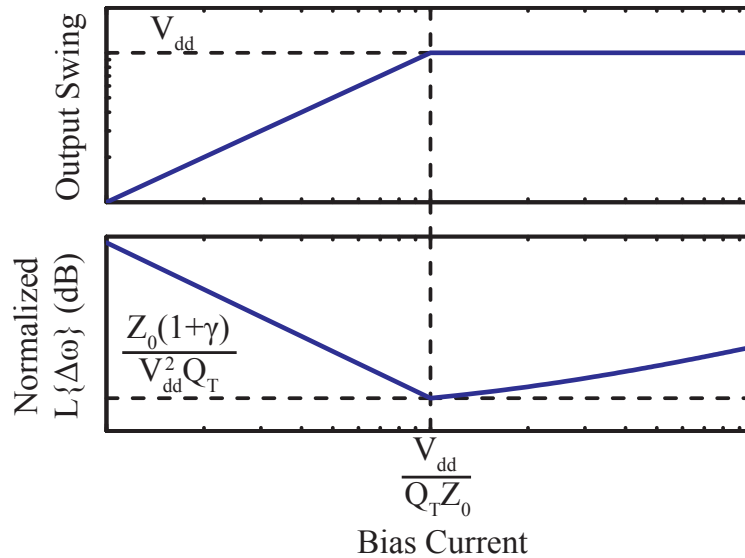


Figure 3.11: Phase noise optimization.

a given Z_o the bias current should be chosen to put the oscillator at this boundary. To decrease phase noise further we would need to decrease Z_o . However, doing so would also mean increasing the bias current in order to remain at the boundary. Without an associated increase in bias current, the phase noise would increase since the oscillator would be in the current-limited regime where phase noise, given by (3.50), is inversely proportional to Z_o . Once again, we find a trade-off between performance and power consumption. One final consideration is that, according to (3.28), Z_o also determines the startup conditions for the oscillator. Thus, if Z_o is decreased without an associated increase in bias current, the safety margin for startup will be reduced.

3.3 Other Fundamental Mode Oscillator Topologies

The cross-coupled oscillator is very popular due to its simplicity both in analysis and design. However, other topologies may provide benefits in certain applications.

3.3.1 Colpitts

The Colpitts oscillator, first presented in [Colpitts27] and shown in Fig. 3.12, utilizes a single transistor in common-gate configuration. The feedback from drain to source is provided through a capacitive divider. The inductance, L , is then used to tune out the total tank capacitance. Since the transistor gain in common-gate configuration is positive and the feedback does not invert the phase, the loop has overall positive feedback.

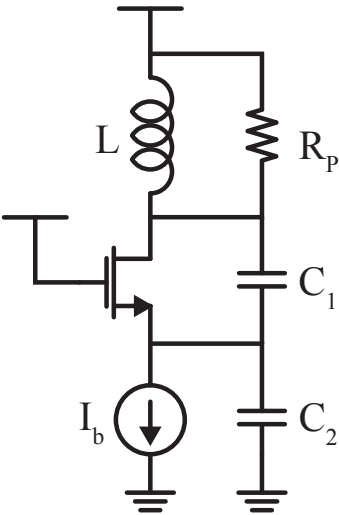


Figure 3.12: Colpitts oscillator schematic.

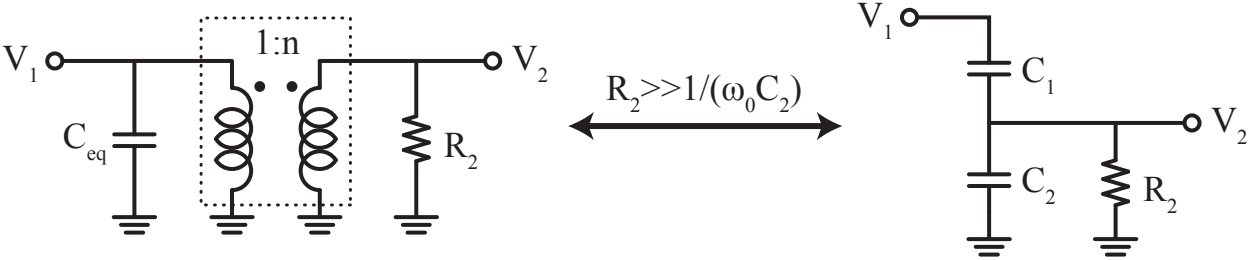


Figure 3.13: Capacitive divider as ideal transformer.

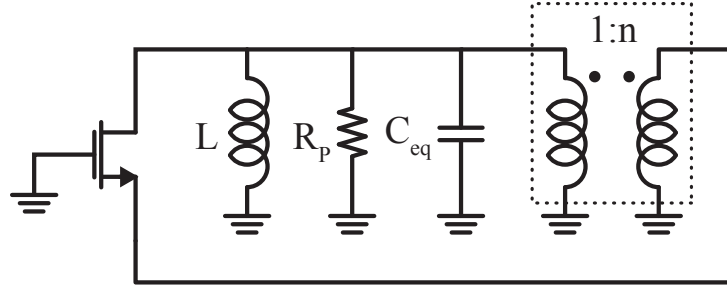


Figure 3.14: Colpitts oscillator effective model (biasing omitted).

A capacitive divider acts like an impedance transformer (Fig. 3.13) over a narrow bandwidth as long as high circuit Q is maintained (i.e.: $R_2 \gg 1/\omega_o C_2$). We can define the effective turns ratio, n , for the transformer as a function of the capacitance ratio from the voltage divider equation.

$$n \triangleq \frac{V_2}{V_1} = \frac{C_1}{C_1 + C_2} = \frac{1}{1 + C_2/C_1} \quad (3.55)$$

The capacitive divider then transforms the drain voltage down to the source but also transforms the source impedance back up to the drain, loading the output tank. The effective tank capacitance, C_{eq} , is given by the series combination of C_1 and C_2 .

$$C_{eq} = \frac{C_1 C_2}{C_1 + C_2} \quad (3.56)$$

For synthesis, (3.55) and (3.56) can be inverted to give equations for C_1 and C_2 .

$$C_1 = \frac{C_{eq}}{1 - n} \quad (3.57)$$

$$C_2 = \frac{C_{eq}}{n} \quad (3.58)$$

Parasitic capacitances from the transistor must also be taken into account: C_{gd} appears in parallel with the total effective tank capacitance, while C_{gs} appears in parallel with C_2 affecting n .

Let us now derive the start-up conditions for this oscillator. The effective model is shown in Fig. 3.14. The forward gain is the gain from the source of the transistor to its drain

$$A(\omega) = g_m Z_T = g_m \left(j\omega L \parallel R_T \parallel \frac{1}{j\omega C_T} \right) \quad (3.59)$$

while the feedback gain is simply the gain across the transformer (including the effect of C_{gs})

$$f = n = \frac{C_1}{C_1 + C_2 + C_{gs}} \quad (3.60)$$

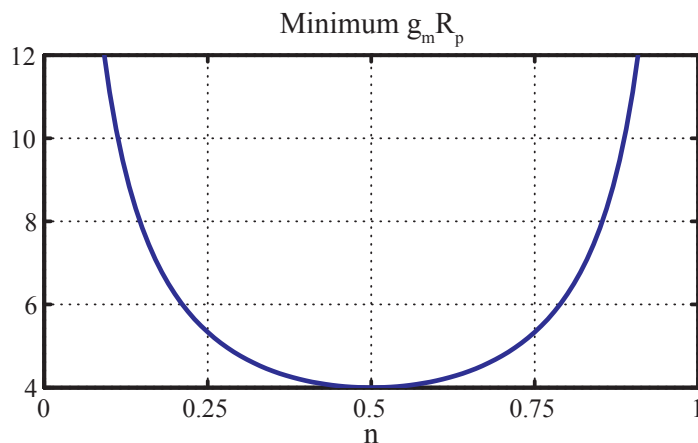


Figure 3.15: Colpitts startup constraint.

Due to the transformer, the source impedance loads the output tank so the total tank resistance is given by

$$R_T = R_P \parallel \frac{1}{n^2 g_m} = \frac{R_P}{1 + g_m R_P n^2} \quad (3.61)$$

where R_p represents the loss of the inductor and capacitors. On the other hand, the total tank capacitance (including the effects of C_{gs} and C_{gd}) is given by

$$C_T = \frac{C_1 (C_2 + C_{gs})}{C_1 + C_2 + C_{gs}} + C_{gd} \quad (3.62)$$

The inductance L must then be chosen to resonate out C_T at the required frequency of oscillation. At resonance, the tank impedance is equal to R_T and the loop gain, Af , simplifies to

$$Af \Big|_{\omega=\omega_o} = g_m R_T n = \frac{g_m R_P n}{1 + g_m R_P n^2} \quad (3.63)$$

For the oscillator to start up, this loop gain must be greater than 1

$$\begin{aligned} \frac{g_m R_P n}{1 + g_m R_P n^2} &> 1 \\ g_m R_P (n - n^2) &> 1 \\ g_m &> \frac{1}{R_P (n - n^2)} \end{aligned} \quad (3.64)$$

As before, R_P is just the loss of the passive components and is equal to $Q_T Z_o$ of the tank. As we can see from Fig. 3.15, the minimum possible value for g_m occurs for $n = 0.5$ (in other words $C_1 = C_2$)

$$g_{m,min} = \frac{1}{R_P (0.5 - 0.25)} = \frac{4}{R_P} \quad (3.65)$$

Thus, unlike the cross-coupled differential pair, for a Colpitts oscillator the minimum possible small-signal intrinsic gain, $g_m R_P$, must be greater than 4. Using the long-channel approximation in (3.12) as an example, the minimum bias current can then be found.

$$\begin{aligned} g_m &\geq \frac{4}{R_P} \\ \frac{I_b}{V_{ov}} &\geq \frac{2}{R_P} \end{aligned} \quad (3.66)$$

Notice that (3.13) and (3.66) are identical in terms of overall required bias current for startup for a given tank. However, the Colpitts oscillator only provides a single-ended output for this bias current.

In this topology, the drain current exhibits a pulsed behavior, injecting energy into the tank only for a short period of time. Sample drain voltage and current waveforms for a typical Colpitts oscillator (Fig. 3.16) show this Class-C operation (from PA nomenclature [Krauss80, p. 394], [Lee04b, p. 499]). The active device injects current into the tank in narrow and tall pulses for only a small fraction of the oscillation period and is off for the remainder. This current is rich in harmonics but the tank filters out and converts only the fundamental into voltage. The fundamental harmonic current can be well approximated by $I_{\omega_o} \approx 2I_b$ ([Krauss80, p. 396], [Lee04b, p. 501], [Andreani05]). Using (3.61) and (3.64) we can find the effective tank resistance in steady-state when the loop-gain is equal to 1.

$$R_{T,ss} = \frac{R_P}{1 + G_m R_P n^2} = \frac{R_P}{1 + n/(1-n)} = R_P (1-n) \quad (3.67)$$

giving an output voltage amplitude at the fundamental of

$$V_{\omega_o} = I_{\omega_o} R_{T,ss} = 2I_b R_P (1-n) \quad (3.68)$$

Thus, we can see that the source impedance loads the tank through the capacitive transformer, reducing the tank impedance at resonance and the output amplitude by a factor of $(1-n)$.

The selection of n thus affects startup (3.64) and the output amplitude (3.68) but also the phase noise. The noise injected into the tank due to R_P is just

$$\overline{i_{n,T}^2} \Big|_{R_P} = \overline{i_R^2} = 4kT \frac{1}{R_P} \Delta f \quad (3.69)$$

On the other hand, due to the capacitive transformer, the transistor is effectively degenerated by an impedance equal to $n^2 R_P$. Thus, the amount of drain noise that is injected into the

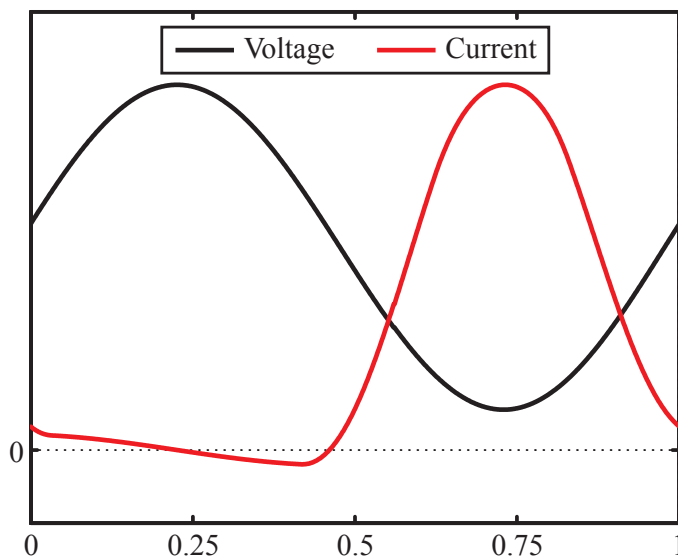


Figure 3.16: Colpitts oscillator waveforms.

tank is given by

$$\begin{aligned}
 \overline{i_{n,T}^2}|_{MOS} &= \overline{i_d^2} \left| \frac{1/G_m}{1/G_m + n^2 R_P} \right|^2 \\
 &= 4kT\gamma G_m \Delta f \left| \frac{1/G_m}{1/G_m + n^2 R_P} \right|^2 \\
 &= 4kt\gamma G_m \Delta f \left(\frac{1}{1 + n/(1-n)} \right)^2 \\
 &= 4kT\gamma G_m \Delta f (1-n)^2 \\
 &= 4kT\gamma \frac{1}{R_P n (1-n)} \Delta f (1-n)^2 \\
 &= \overline{i_R^2} \gamma \frac{1-n}{n}
 \end{aligned} \tag{3.70}$$

where we have used (3.64) and (3.69) to simplify (3.70). A much more rigorous derivation based on ISF theory is presented in [Andreani05] and arrives at the identical result. We can now see that decreasing n leads to an increase in noise contributed by the active device relative to that from the tank loss while increasing n reduces the output amplitude. There is an optimum value, n_{opt} , which minimizes phase noise and is a function of only γ , the transistor noise factor. From (3.40) and (3.68), n_{opt} is the value of n which minimizes $F/(1-n)^2$. We will call this factor Ψ .

$$\Psi = \frac{1 + \gamma(1-n)/n}{(1-n)^2} = \frac{1}{(1-n)^2} + \frac{\gamma}{n(1-n)} \tag{3.71}$$

γ	n_{opt}
2/3	0.3
1	1/3
2	0.38

Table 3.1: Sample values of n_{opt} for Colpitts oscillator.

To find n_{opt} we take the derivative of (3.71), set it equal to zero, and solve for n .

$$\begin{aligned}
\frac{d\Psi}{dn} &= \frac{\gamma(2n-1)}{n^2(n-1)^2} - \frac{2}{(n-1)^3} = 0 \\
0 &= 2(\gamma-1)n^2 - 3\gamma n + \gamma \\
n_{opt} &= \frac{3\gamma - \sqrt{9\gamma^2 - 8(\gamma-1)}}{4(\gamma-1)} \\
n_{opt} &= \frac{3 - \sqrt{1 + 8/\gamma}}{4(1 - 1/\gamma)} \\
n_{opt} &= \frac{2}{3 + \sqrt{1 + 8/\gamma}} \tag{3.72}
\end{aligned}$$

As an example, Table 3.1 shows n_{opt} for some common values of γ .

One potential advantage of this topology is that the active device only injects current into the tank during a fraction of the period and is off for the rest of the time due to the Class-C operation exhibited in Fig. 3.16. The active device is on and providing most of the current when the tank voltage is at a minimum. During the transitions and the peak in voltage, the active device is mostly off. In a standard cross-coupled oscillator the current waveforms are closer to a square wave (or at high frequencies, a sine wave). There is always at least one device on and injecting current and noise into the tank. More importantly, both devices are injecting current into the tank during the transition region. From the analysis in Section 3.2.3 it would seem that the Colpitts oscillator injects noise into the tank at the best possible time and should thus achieve superior phase noise performance. Unfortunately, it is not clear whether this is true in real designs especially at high frequencies where all waveforms begin to look sinusoidal.

3.3.2 Common-Drain Colpitts

The Colpitts style of oscillator can provide one significant advantage in low power designs. We begin by changing the topology slightly to create a common-drain Colpitts oscillator [Schlesinger45] as shown in Fig. 3.17a.³ In this structure, the transistor acts as a source

³Also called a Cathode/Emitter/Source-Follower oscillator but recognized early on by [Schlesinger45] and others as another form of a Colpitts oscillator.

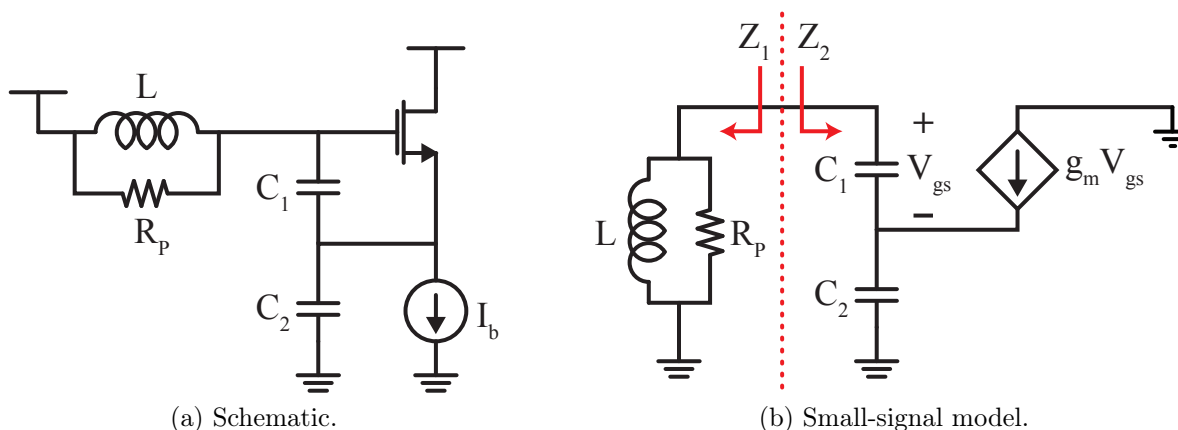


Figure 3.17: Common-Drain Colpitts Oscillator.

follower providing positive gain from gate to source, while the capacitive transformer creates the feedback path back to the gate closing the positive feedback loop. Unlike the common-gate topology however, C_{gs} here appears in parallel with C_1 . This should be easily mitigated by adjustment of C_1 and C_2 without affecting n or C_{eq} . While the capacitive loading in this case is slightly different, the overall operation is not.

We can rederive the startup conditions using the negative resistance approach represented by Fig. 3.1b. We first split up the circuit as shown in Fig. 3.17b, where R_P represents the losses of the tank and C_{gs} and C_{gd} have been omitted for clarity. We now apply a test voltage source at the gate and, using KVL and KCL, we can see that

$$\begin{aligned}
 V_T &= V_s + V_{gs} \\
 &= \frac{i_s}{sC_2} + V_{gs} \\
 &= \frac{i_T + g_m V_{gs}}{sC_2} + V_{gs} \\
 &= \frac{i_T + g_m i_T / (sC_1)}{sC_2} + \frac{i_T}{sC_1} \\
 &= i_T \left(\frac{g_m}{s^2 C_1 C_2} + \frac{1}{sC_1} + \frac{1}{sC_2} \right)
 \end{aligned} \tag{3.73}$$

The input impedance, Z_2 , is then given by the ratio V_T/i_T , simplified using (3.55)-(3.58).

$$\begin{aligned}
 Z_2(\omega) &= \frac{-g_m}{\omega^2 C_1 C_2} + \frac{1}{j\omega C_1} + \frac{1}{j\omega C_2} \\
 &= \frac{-g_m n(1-n)}{\omega^2 C_{eq}^2} + \frac{1}{j\omega C_{eq}}
 \end{aligned} \tag{3.74}$$

As expected, Z_2 has a negative real part and a capacitive part equal to C_{eq} . We now turn our attention to the left side of 3.17b which contains the tank inductor as well as all losses

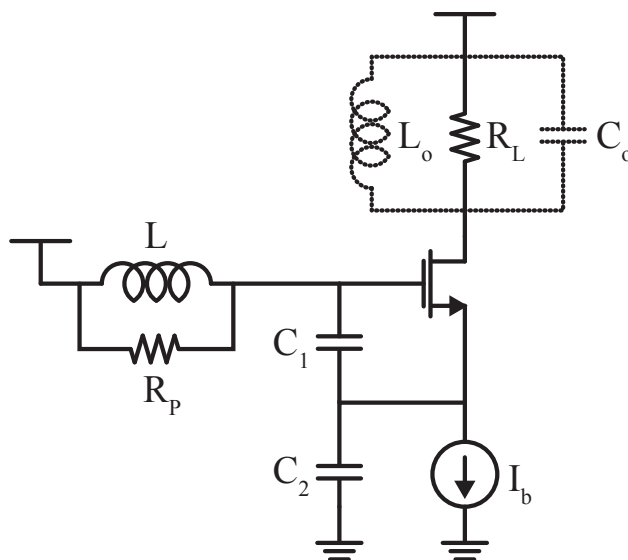


Figure 3.18: Common-drain Colpitts oscillator provides buffering for free.

associated with tank inductance and capacitance lumped into R_P . Assuming high Q_T we can perform a parallel to series conversion yielding

$$Z_1(\omega) = j\omega L_s + R_s \approx j\omega L + \frac{R_P}{Q_T^2} \quad (3.75)$$

We can now see that the inductor must be chosen to resonate out C_{eq} at the resonance frequency and, for oscillation to occur, the negative resistance provided by the core must be greater in magnitude than R_s .

$$\begin{aligned} \left| \mathcal{R}\{Z_2(\omega_o)\} \right| &\geq \mathcal{R}\{Z_1(\omega_o)\} \\ \frac{g_m(n - n^2)}{\omega_o^2 C_{eq}^2} &\geq \frac{R_P}{Q_T^2} \\ g_m(n - n^2) Z_o^2 &\geq \frac{R_P}{Q_T^2} \\ g_m &\geq \frac{R_P}{Z_o^2 Q_T^2 (n - n^2)} \\ g_m &\geq \frac{1}{R_P (n - n^2)} \end{aligned} \quad (3.76)$$

We can now return to discussing the advantage of this topology. Since the core device is in a common-drain configuration we simply connected its drain to the supply. This is not necessary since the operation of a common-drain amplifier is not affected by impedance at the drain as long as the device remains in saturation. If we place a resistor, R_L , at the drain

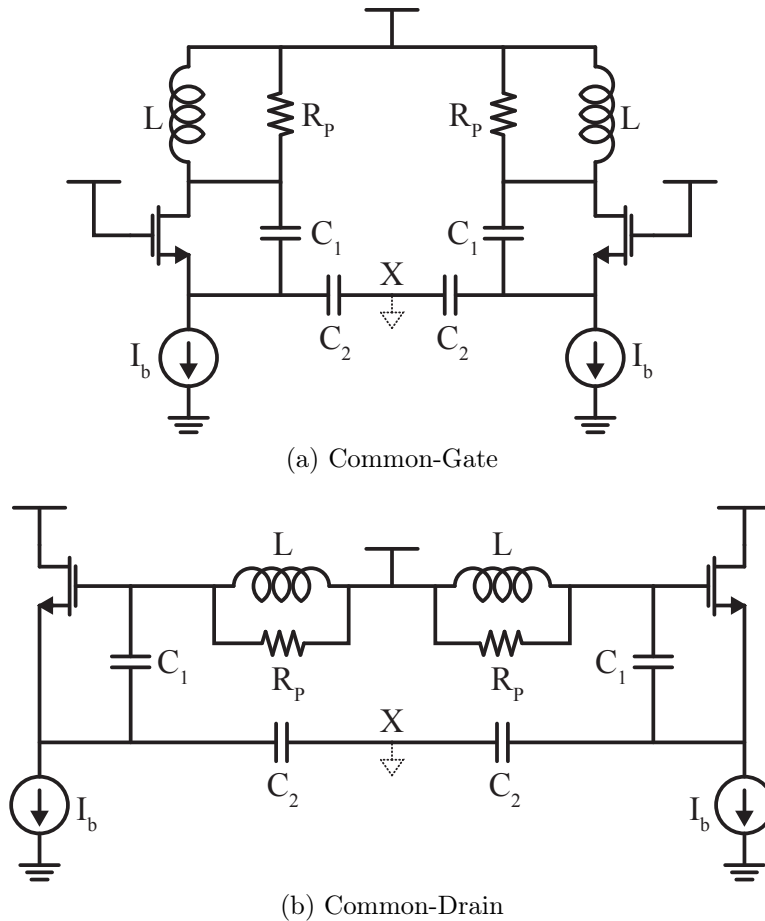


Figure 3.19: Differential Colpitts Oscillators

as proposed by [Dauphinee97] and shown in Fig. 3.18 (ignoring L_o and C_o for now) we have a free buffered output from our oscillator. By taking the output from the drain instead of the source we isolate the resonant tank from external loading allowing it to be independently optimized regardless of the load.

To allow for reduced supply voltages, a resonant tank could be used instead of a resistor (L_o and C_o in Fig. 3.18). Also suggested by [Dauphinee97], this would maintain a high impedance at the frequency of interest while presenting a very low impedance at DC, eliminating the overhead required due to the bias current running through R_L . [Voinigescu00] also proposed a shunt peaking approach by placing an inductor in series with R_L , tuning out some of the output capacitance to increase the output swing but doing nothing for the DC drop across the resistor.

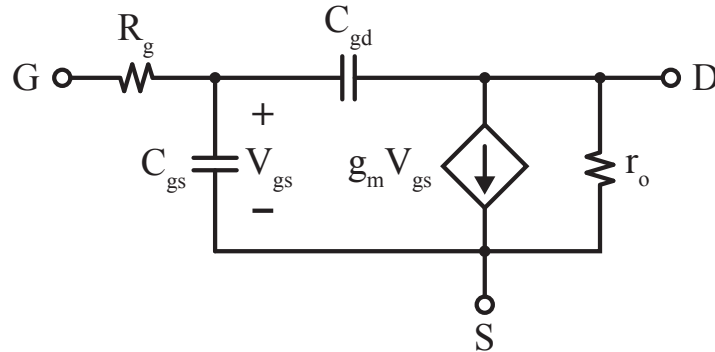


Figure 3.20: Transistor small-signal model.

3.3.3 Differential Versions

In the early days of radio, active elements were expensive so topologies with the fewest number of them were favored. Thus, early oscillators utilized only a single active element with positive feedback to generate oscillation.⁴ However, this comes at the expense of providing only a single-ended output. Since many applications require differential outputs and transistors today are cheap, this is no longer an advantage. Luckily differential versions of these oscillators are easily achievable and perform identically to their single-ended counterparts. [Rogers00] shows just such an implementation of a differential common-base Colpitts oscillator. The differential version of the common-collector Colpitts is presented in [Voinigescu00]. Both are integrated in Silicon bipolar processes. Representative schematics for their MOS counterparts are shown in Fig. 3.19.

In order to ensure that oscillation can only occur in the differential mode, startup conditions must not be met in the common-mode. In both cases this can be accomplished by leaving node X , the midpoint of capacitor C_2 , floating. In differential mode, this node is a virtual ground so operation is not affected. In common-mode, on the other hand, C_2 disappears making $n_{CM} = 1$. From 3.64 and 3.76 we can see that this condition would make the g_m required for startup equal to ∞ . Thus, this simple step ensures no common-mode oscillation can occur. For high frequency designs this node should be well isolated in layout to maintain a high impedance (low capacitance) to ground.

3.4 Cross-Over Frequency

In the derivations presented above, the oscillator cores can generate negative resistance at any frequency, seemingly without bound. Since there is a maximum frequency for which the active devices can generate power gain, f_{max} [Razavi94, Manku99], there should be an upper

⁴At the time, these were vacuum tubes, however, the operation and analysis is really no different with transistors, whether BJT or MOS.

frequency limit beyond which the oscillator core can no longer generate the required loop gain for oscillation. The reason we have not seen this limitation thus far is our simplistic transistor model. A more realistic transistor model is shown in Fig. 3.20. The most important addition is the series gate resistance, R_g , which is made up of two components [Shaeffer97]: the gate routing resistance, R_{poly} , and a resistance due the non-quasistatic nature of the channel, R_{NQS} .

$$R_g = R_{poly} + R_{NQS} \quad (3.77)$$

For a multi-finger layout, the physical gate resistance is equal to

$$R_{poly} = \frac{R_{\square} W}{3L N_F^2} \quad (3.78)$$

where R_{\square} is the sheet resistance of the gate material, W the total device width, L the channel length, and N_F the number of fingers.⁵ Thus, R_{poly} can be made negligible if the number of fingers is increased sufficiently or the sheet resistance reduced such as by using a metal gate. The second component arises at very high frequencies where the distributed nature of the the channel creates a phase shift in the gate impedance making it look less like a pure capacitor [vanderZiel70]. At frequencies well below $5f_T$, a reasonable limit for most applications of interest, this effect can be modeled as a series gate resistance [Shaeffer97]

$$R_{NQS} = \frac{1}{5g_m} \quad (3.79)$$

Thus, with proper device layout, at high frequencies the total gate resistance can be approximated as

$$R_g \approx \frac{1}{5g_m} \quad (3.80)$$

Due to this R_g as well as finite output conductance, any real active device can only provide power gain at frequencies below a certain frequency called f_{max} . This frequency can be approximated using any of (3.81)-(3.83) [Razavi94, Manku99]

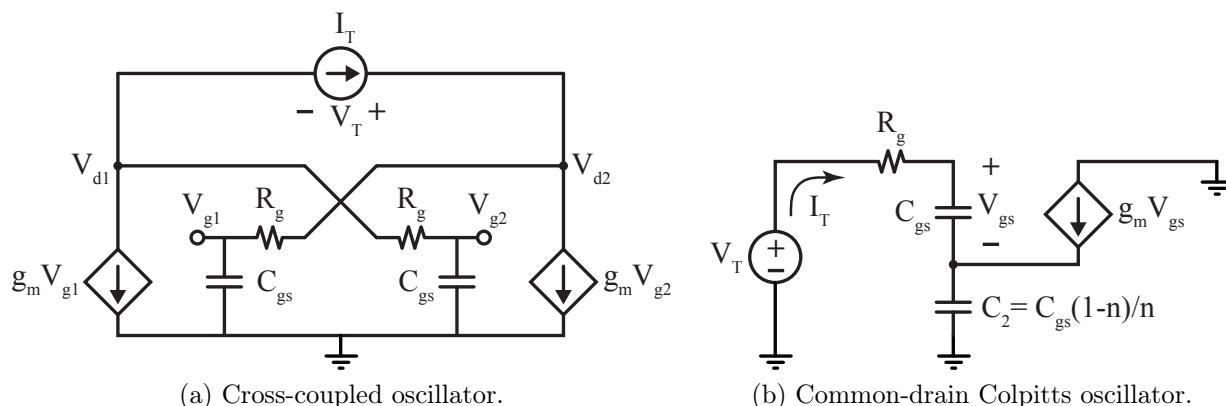
$$f_{max} \approx \frac{1}{2\pi} \sqrt{\frac{g_m^2 r_o}{4R_g (C_{gs} + C_{gd}) [C_{gs} + (1 + g_m r_o) C_{gd}]}} \quad (3.81)$$

$$f_{max} \approx \sqrt{\frac{f_T}{8\pi R_g C_{gd}}} \quad (3.82)$$

$$f_{max} \approx f_T \sqrt{\frac{r_o}{4R_g}} = f_T \sqrt{\frac{5}{4} g_m r_o} \quad (3.83)$$

We thus expect the maximum frequency of oscillation to be ultimately limited by f_{max} . Nevertheless, different topologies will have different limitations.

⁵The factor of 3 in the denominator comes from the distributed nature of the gate resistance [Razavi94] and is valid for single-sided gate contacts. For double sided gate contacts, it should be equal to 12.

Figure 3.21: Oscillator small signal models including R_g and C_{gs} .

We will begin our analysis by first introducing only R_g and, at least initially, ignoring C_{gd} and r_o . The small-signal model for the cross-coupled oscillator is shown in Fig. 3.21a. We see that R_g and C_{gs} form a low pass filter now from the drain of one transistor to the gate of the other. Following the same procedure from Section 3.2.1, the input admittance can readily be found [Razavi11]

$$Y_{in} = \frac{j\omega C_{gs} - g_m}{2(1 + j\omega R_g C_{gs})} \quad (3.84)$$

The real part of which is equal to

$$\mathcal{R}\{Y_{in}\} = \frac{R_g C_{gs}^2 \omega^2 - g_m}{2(1 + R_g^2 C_{gs}^2 \omega^2)} \quad (3.85)$$

Beyond a certain frequency, this conductance becomes positive. We will call that frequency the cross-over frequency, $\omega_{c,o}$.

$$\begin{aligned} g_m &= R_g C_{gs}^2 \omega_{c,o}^2 \\ \omega_{c,o}^2 &= \frac{5g_m}{C_{gs}^2} \\ \omega_{c,o} &= \omega_T \sqrt{5} \end{aligned} \quad (3.86)$$

We will now turn to the common-drain Colpitts oscillator, with the small signal model shown in Fig. 3.21b. To provide a fair comparison between the two topologies we will use the differential topology for the Colpitts oscillator and assume the transistors are identically sized and biased. Furthermore, since we are trying to find the maximum frequency of oscillation we would like to maintain minimal capacitive loading. We did not add any extra tank capacitance in the cross-coupled analysis. For the Colpitts analysis we will let $C_1 = C_{gs}$ and only add an explicit C_2 which, from (3.55), will be equal to

$$C_2 = C_1 \left(\frac{1-n}{n} \right) = C_{gs} \left(\frac{1-n}{n} \right) \quad (3.87)$$

Repeating the analysis from Section 3.3.2, the input impedance is

$$Z_{in} = \frac{\omega^2 C_{gs}^2 \left(\frac{1-n}{n}\right) R_g - g_m}{\omega^2 C_{gs}^2 \left(\frac{1-n}{n}\right)} + \frac{1}{j\omega C_{gs} (1-n)} \quad (3.88)$$

The cross-over frequency for the common-drain Colpitts, $\omega_{c,Colp}$, is then

$$\begin{aligned} g_m &= \omega_{c,Colp}^2 C_{gs}^2 \left(\frac{1-n}{n}\right) R_g \\ \omega_{c,Colp}^2 &= \frac{5g_m^2}{C_{gs}^2} \left(\frac{n}{1-n}\right) \\ \omega_{c,Colp} &= \omega_T \sqrt{5 \left(\frac{n}{1-n}\right)} \\ \omega_{c,Colp} &= \omega_{c,o} \sqrt{\frac{n}{1-n}} \end{aligned} \quad (3.89)$$

We can now compare the performance for the two topologies.

$$\begin{aligned} \omega_{c,Colp} &> \omega_{c,o} \\ \sqrt{\frac{n}{1-n}} &> 1 \\ n &> 1-n \\ n &> 0.5 \end{aligned} \quad (3.90)$$

From this simple result we can see that the cross-over frequency is higher for the common-drain Colpitts topology than for the cross-coupled topology only for $n > 0.5$ and can in fact be much lower for smaller values of n . From our previous results we have seen that n is usually not set higher than 0.5 so the cross-coupled topology may be a better choice overall.

Including the effects of C_{gd} and r_o can make the required algebra cumbersome and intractable. Furthermore, at high frequencies these simple models do not accurately reflect the behavior of the active devices. For these reasons we now turn to simulation in order to account for all transistor parasitics and nonidealities. As before, the cross-over frequency of the Colpitts oscillator is normalized to the cross-over frequency of the cross-coupled oscillator to enable comparison. The results are plotted in Fig. 3.22a for three different channel lengths. For values of n greater than approximately 0.4, the Colpitts oscillator has a higher ω_c than the cross-coupled oscillator. However, especially for short-channel devices, the difference is not substantial.

Examining the results from a different perspective gives further insight. At any given frequency below ω_c the input negative conductance of the Colpitts core is a function of n . There is thus an optimum n at each frequency for which the negative conductance is maximized for a given device size and bias current. We can then normalize this maximum achievable G_{in}

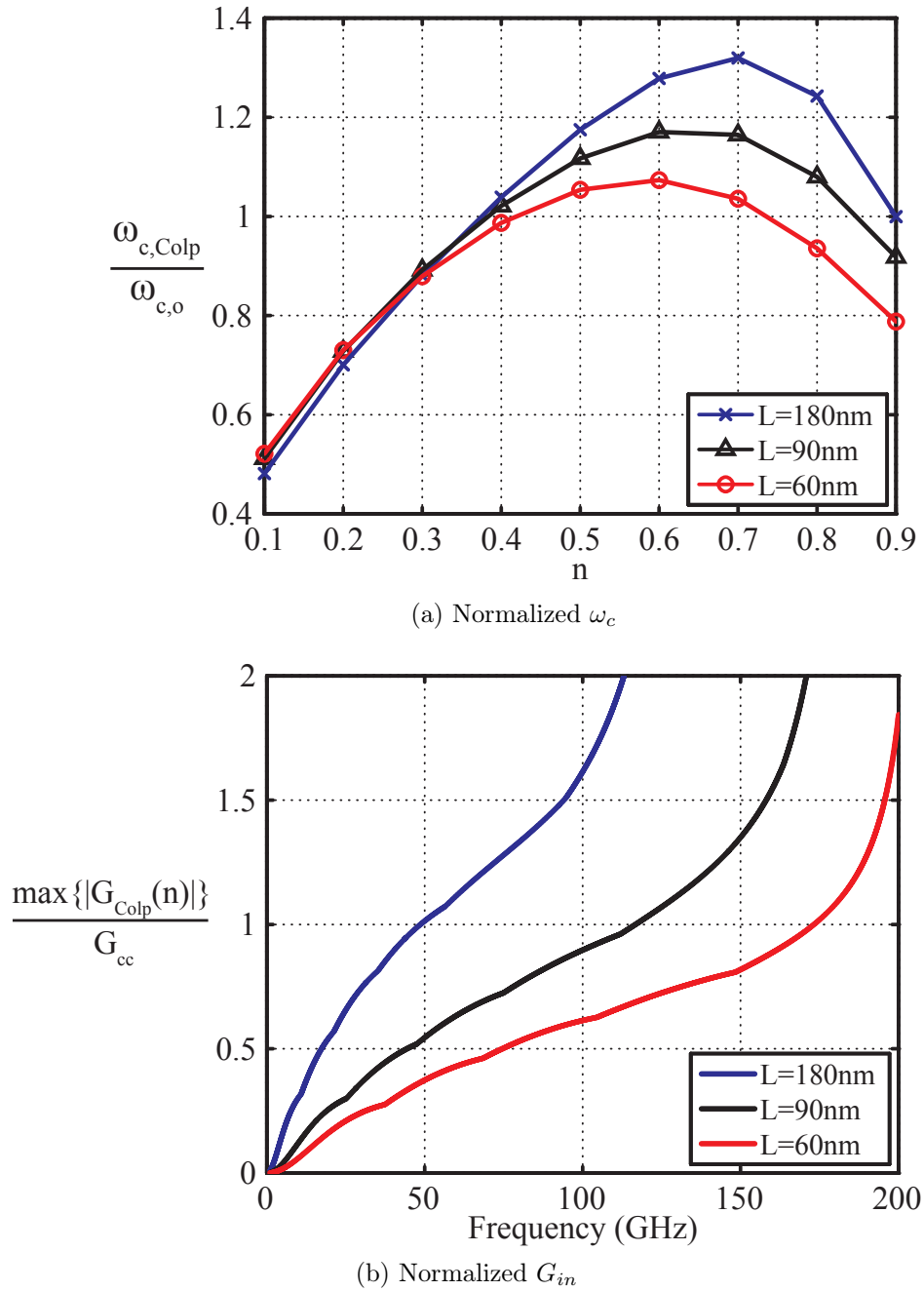


Figure 3.22: Performance comparisons of Colpitts versus cross-coupled core. All simulations are performed in a 65nm process node.

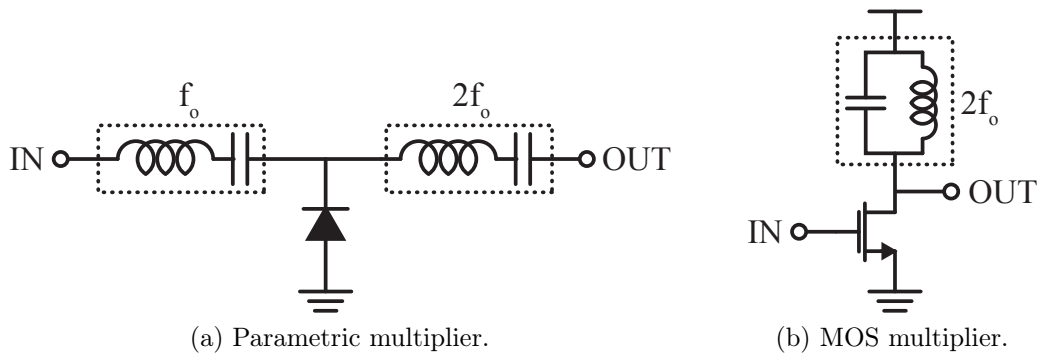


Figure 3.23: Frequency multipliers.

to the negative conductance provided by an identically sized and biased cross-coupled differential core at the same frequency and plot the result in Fig. 3.22b. The Colpitts topology is more power efficient than the cross-coupled topology at frequencies for which this ratio is greater than 1. It must be noted that Fig. 3.22b represents prelayout simulations so layout parasitics may shift the unity crossing frequency of the G_{in} ratio. Furthermore, this analysis is only valid for startup constraints. Phase noise and other considerations may ultimately result in opposing guidelines for topology selection.

3.5 The Push-Push Oscillator

The previous section has shown that all oscillators have an upper frequency limit beyond which the active devices do not provide sufficient gain for sustained oscillation. Attaining an output at a frequency beyond ω_c is thus not possible with fundamental oscillators. Furthermore, the higher gain available at lower frequencies and the losses of inductors and capacitors as a function of frequency may make a lower frequency VCO design more desirable in terms of phase noise or power consumption. Thus, we need a way to generate the desired 60GHz LO frequency from a lower fundamental frequency.

Frequency multipliers utilize a nonlinear element to produce harmonics of the input frequency and a filter to select only the desired harmonic while rejecting the fundamental. A general nonlinearity can be expressed by its Taylor series expansion. Assuming $V_{in} = A \cos(\omega t)$, the

output of this general nonlinearity is then given by

$$\begin{aligned}
 V_{out} &= a_0 + a_1 V_{in} + a_2 V_{in}^2 + a_3 V_{in}^3 + \dots \\
 &= a_0 + a_1 A \cos(\omega t) + a_2 A^2 \cos^2(\omega t) + a_3 A^3 \cos^3(\omega t) + \dots \\
 &= a_0 + a_1 A \cos(\omega t) + \frac{a_2 A^2}{2} [1 + \cos(2\omega t)] \\
 &\quad + \frac{a_3 A^3}{4} [\cos(3\omega t) + 3 \cos(\omega t)] + \dots \\
 &= \left(a_0 + \frac{a_2 A^2}{2} \right) + \left(a_1 A + \frac{3a_3 A^3}{4} \right) \cos(\omega t) \\
 &\quad + \frac{a_2 A^2}{2} \cos(2\omega t) + \frac{a_3 A^3}{4} \cos(3\omega t) + \dots
 \end{aligned} \tag{3.91}$$

As we can see, an n^{th} order nonlinearity driven by a sinusoidal signal generates sinusoids up to the n^{th} harmonic.

A passive example of this (Fig. 3.23a) is the parametric multiplier which uses the nonlinear capacitance of a diode (or MOS varactor) [Leenov59] as the nonlinear element. The diode has an exponential response rich in harmonic content. An input filter is placed at the fundamental frequency and an output filter at the required harmonic frequency. A simpler multiplier (Fig. 3.23b) uses the inherent nonlinearity of a BJT or MOS device biased near threshold and driven by a large fundamental signal [Ferndahl04]. Biasing the MOS device near threshold places it in weak inversion where it behaves very much like a BJT. Thus, both devices exhibit an exponential response. This type of multiplier is active, consuming DC power. Unfortunately, nonlinear multipliers are inefficient at converting the fundamental frequency to higher harmonics. Furthermore, the conversion gain, the amplitude at the desired harmonic divided by the amplitude of the input harmonic, is actually a function of the input amplitude, A . This can be seen by examining (3.91). By inspection, the conversion gains to the second and third harmonics respectively are

$$G_2 = \frac{a_2 A}{2} \tag{3.92}$$

$$G_3 = \frac{a_3 A^2}{4} \tag{3.93}$$

Multipliers must then be driven by as large a signal as possible for high conversion gain. This means that, generally, a driving amplifier is required both at the input for high conversion gain and at the output to increase the harmonic power [Emami07], making them very inefficient. Furthermore, substantial filtering must be provided at the output to ensure that the fundamental frequency as well as all other harmonics except the desired one are sufficiently suppressed.

There is however, another nonlinear system which we have overlooked: the oscillator itself. For example, as we saw from previous sections, while the output voltage is shaped by the

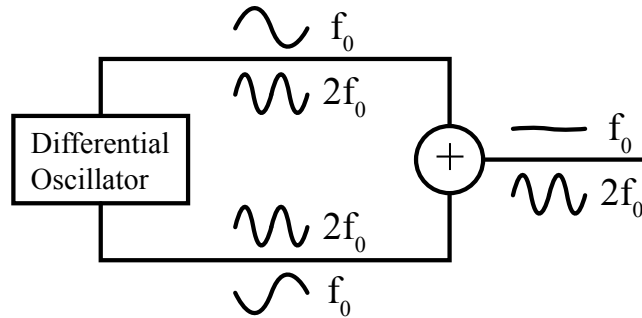


Figure 3.24: Push-push principle

high-Q tank resulting in a near sinusoidal signal, the driving current is in fact very nonlinear. In the case of the cross-coupled differential pair the current resembles a square wave which only contains odd harmonics. The Colpitts oscillator, on the other hand, exhibits a narrow and sharp pulse of current which contains every harmonic of the fundamental. How can we use these harmonic currents or other nonlinear behavior in the oscillator?

A push-push oscillator [Bender83] sums the outputs of two oscillators which are operating out of phase at the fundamental frequency. One way to ensure two oscillators operating perfectly out of phase is to use a differential oscillator. This idea is shown graphically in Fig. 3.24. Due to the squaring action of the second order nonlinearity, the second harmonic signals in both oscillators are actually in phase.⁶

$$\begin{aligned} \cos(\omega t) + \cos(\omega t + \pi) &= \cos(\omega t) - \cos(\omega t) \\ &= 0 \end{aligned} \quad (3.94)$$

$$\begin{aligned} \cos^2(\omega t) + \cos^2(\omega t + \pi) &= \frac{1}{2}[1 + \cos(2\omega t)] + \frac{1}{2}[1 + \cos(2\omega t + 2\pi)] \\ &= 1 + \cos(2\omega t) \end{aligned} \quad (3.95)$$

Thus, while the fundamental component cancels out, the second harmonic is added in phase creating an output only at the second harmonic. This idea can be extended to N oscillators summed at smaller relative phases of $2\pi/N$. The output of such an arrangement sums the N^{th} harmonic while cancelling out every other lower harmonic. This arrangement is called “ N -push” [Catli10] if the nonlinearity is generated in the oscillator itself or “linear superposition” [Huang08] if the oscillator is followed by a rectifier to generate the higher harmonics. The close-in phase noise of the N^{th} harmonic output is higher than the phase noise of the fundamental by $20 \log N$ (e.g.: for push-push the 2^{nd} harmonic phase noise is 6dB higher than the phase noise of the fundamental). Furthermore, the harmonic content of each individual oscillator or of a rectifier drops quickly with frequency, reducing the available power at higher harmonics to very small levels. For example, according to [Huang08], the conversion gain from fundamental to fourth harmonic using the linear superposition method with rectification is at best -15.4dB (assuming perfect rectification). That same work showed

⁶In fact, all odd-order harmonics will be out of phase, while all even-order harmonics will be in phase.

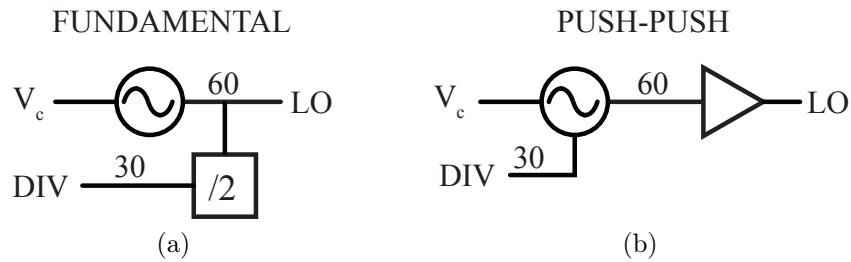


Figure 3.25: Fundamental vs. push-push.

a fourth harmonic linear superposition oscillator designed in $0.13\mu\text{m}$ CMOS with output frequency of 67GHz and output power of only -28dBm. These types of oscillators generally require significant amplification to bring their outputs to useful levels. Another limitation common to harmonic oscillators is that the output is single-ended.

The push-push, or by extension N -push, oscillator does have some potential advantages over fundamental mode oscillators. Since the quality factor of the tank will likely be limited by varactor Q , operation at a lower frequency could lead to higher tank quality factors and, by extension, improvements in phase noise. Also, moving the VCO design to lower frequencies enables higher fractional tuning range to be achieved due to the reduced contribution of parasitics to the overall tank capacitance as discussed in Section 3.2.2. Since the output frequency is N times the VCO frequency, the fractional tuning range of the output is equal to the fractional tuning range of the VCO and thus higher than what could be achieved in a fundamental design.

In a particular application the selection of push-push versus fundamental mode oscillator rests on the relative importance of phase noise, tuning range, power consumption, design complexity and output power. A harmonic oscillator removes the high frequency dividers that would be required with a fundamental mode design (Fig. 3.25a). However, as described above, the output power of a harmonic oscillator will always be lower than what can be achieved in a fundamental design, requiring extra buffering at the output for the harmonic design (Fig. 3.25b). We then turn our attention to the differences in the oscillators. Phase noise and power consumption advantages of the push-push versus the fundamental mode design depend on the scaling methodology used to arrive at the lower frequency. We begin from an optimized fundamental mode design at 60GHz with tank characteristic impedance $Z_{o,60}$, tank quality factor $Q_{T,60}$, and core current consumption $I_{b,60}$. We will assume that $I_{b,60}$ was chosen such that the design is operating at the intersection between the current and voltage limited regimes for optimal phase noise performance.

$$V_{od} \approx I_{b,60} Q_{T,60} Z_{o,60} \approx V_{dd} \quad (3.96)$$

To arrive at the push-push oscillator core operating at 30GHz we must scale the tank inductance and capacitance appropriately. In general, we will then arrive at a new tank design with characteristic impedance $Z_{o,30}$, and quality factor $Q_{T,30}$. We will assume that the bias

current, $I_{b,30}$, is adjusted so as to maintain the same signal swing and thus keep the design operating at the optimal point for phase noise.⁷

Since the voltage swing on the tank is approximately given by $V_{od} \approx I_b R_P$, the signal power can be written as

$$P_{sig} = \frac{V_{od}^2}{2R_P} = \frac{(I_b Q_T Z_o)^2}{2Q_T Z_o} = \frac{I_b^2 Q_T Z_o}{2} \quad (3.97)$$

Thus, using (3.39) and (3.97), the phase noise of the fundamental mode oscillator in the $1/(\Delta\omega)^2$ region is given by

$$\begin{aligned} \mathcal{L}\{\Delta\omega\}_{fund} &= 10 \log \left[\frac{2kT}{P_{sig,60}} \left(\frac{\omega_o}{2Q_{T,60}\Delta\omega} \right)^2 \right] \\ &= 10 \log \left[\frac{kT}{I_{b,60}^2 Q_{T,60} Z_{o,60}} \left(\frac{\omega_o}{Q_{T,60}\Delta\omega} \right)^2 \right] \end{aligned} \quad (3.98)$$

On the other hand, the phase noise of the push-push oscillator in the same region is given by

$$\begin{aligned} \mathcal{L}\{\Delta\omega\}_{pp} &= 10 \log \left[\frac{2kT}{P_{sig,30}} \left(\frac{\omega_o/2}{2Q_{T,30}\Delta\omega} \right)^2 \right] + 20 \log 2 \\ &= 10 \log \left[\frac{kT}{I_{b,30}^2 Q_{T,30} Z_{o,30}} \left(\frac{\omega_o/2}{Q_{T,30}\Delta\omega} \right)^2 \right] + 20 \log 2 \\ &= 10 \log \left[\frac{kT}{I_{b,30}^2 Q_{T,30} Z_{o,30}} \left(\frac{\omega_o}{Q_{T,30}\Delta\omega} \right)^2 \right] \end{aligned} \quad (3.99)$$

Let us now define a phase noise metric $\Delta\mathcal{L}$

$$\begin{aligned} \Delta\mathcal{L} &= \mathcal{L}\{\Delta\omega\}_{pp} - \mathcal{L}\{\Delta\omega\}_{fund} \\ &= 10 \log \left[\frac{I_{b,60}^2 Z_{o,60} Q_{T,60}^3}{I_{b,30}^2 Z_{o,30} Q_{T,30}^3} \right] \end{aligned} \quad (3.100)$$

Using this metric we can determine which oscillator will have better phase noise at the 60GHz output. If $\Delta\mathcal{L} > 0$, the fundamental oscillator phase noise will be lower. On the other hand, if $\Delta\mathcal{L} < 0$, the push-push oscillator phase noise will be lower. We can further simplify this expression by using the constant tank voltage swing constraint.

$$\begin{aligned} V_{od,60} &= V_{od,30} \\ I_{b,60} Q_{T,60} Z_{o,60} &= I_{b,30} Q_{T,30} Z_{o,30} \\ \frac{I_{b,60}}{I_{b,30}} &= \frac{Q_{T,30} Z_{o,30}}{Q_{T,60} Z_{o,60}} \end{aligned} \quad (3.101)$$

⁷Since the start-up gain of the oscillator is also proportional to $I_b Q_T Z_o$, maintaining constant signal swing by adjusting I_b also maintains constant start-up gain.

Substituting (3.101) into (3.100) we arrive at our final equation

$$\Delta\mathcal{L} = 10 \log \left[\frac{Z_{o,30} Q_{T,60}}{Z_{o,60} Q_{T,30}} \right] \quad (3.102)$$

Thus, the push-push oscillator solution will exhibit superior phase noise performance if

$$\frac{Q_{T,30}}{Q_{T,60}} > \frac{Z_{o,30}}{Z_{o,60}} \quad (3.103)$$

If we define Z_n to be the ratio on the right side of (3.103),

$$Z_n = \frac{Z_{o,30}}{Z_{o,60}} \quad (3.104)$$

then the push-push oscillator will be better if

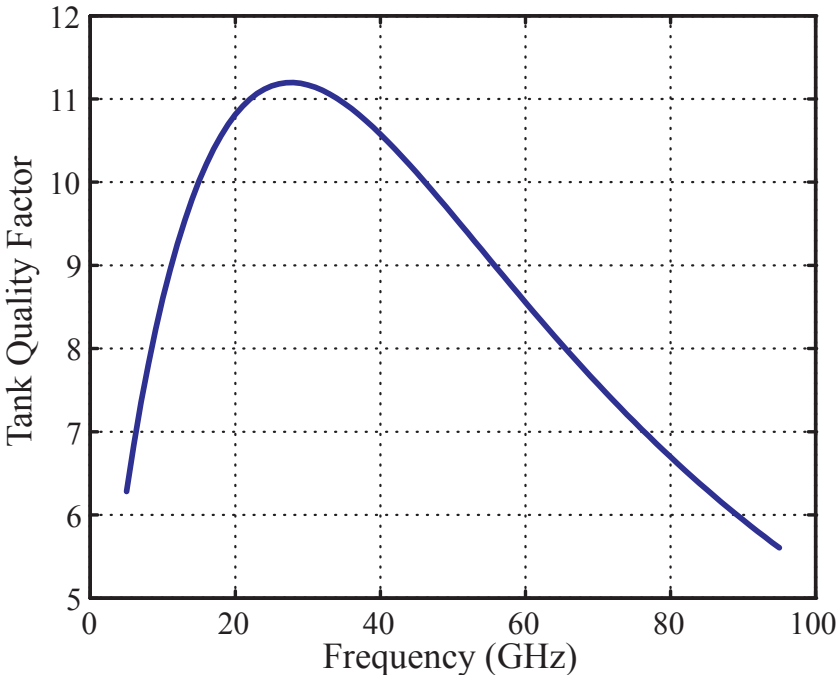
$$Q_{T,30} > Z_n Q_{T,60} \quad (3.105)$$

This is a general result for any possible scaling of Z_o but we can identify three specific cases of interest:

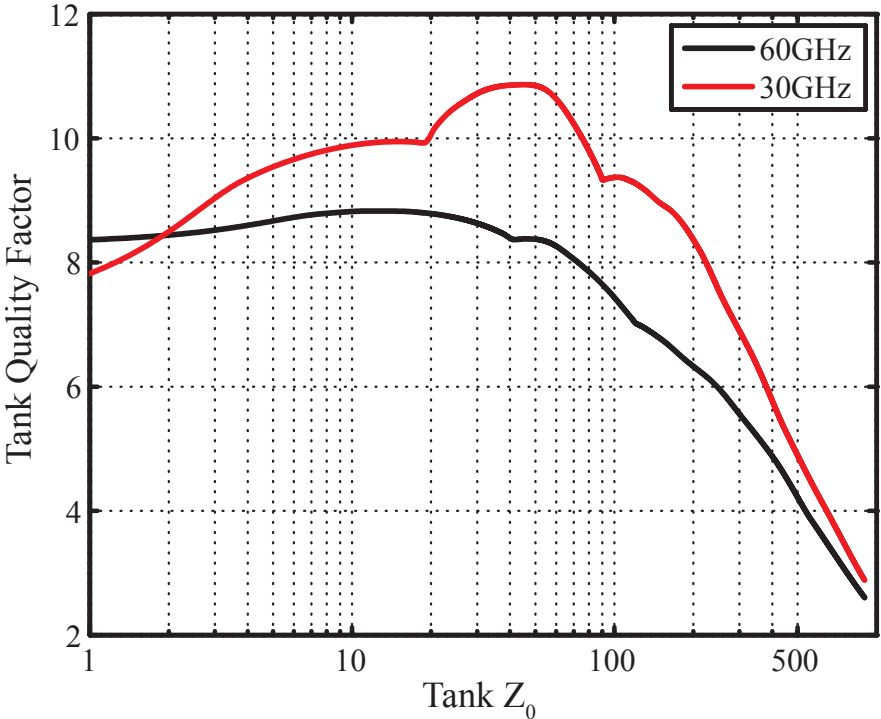
1. Constant C_T scaling. The tank inductor is scaled up by a factor of 4 while the tank capacitor is kept constant resulting in $Z_{o,30} = 2Z_{o,60}$, ($Z_n = 2$).
2. Constant Z_o scaling. Both tank inductance and capacitance are scaled up by a factor of 2 resulting in $Z_{o,30} = Z_{o,60}$, ($Z_n = 1$).
3. Constant L_T scaling. The tank capacitor is scaled up by a factor of 4 while the tank inductor is kept constant resulting in $Z_{o,30} = 0.5Z_{o,60}$, ($Z_n = 0.5$).

Assuming that the fixed capacitance from the core and buffers remains roughly constant, the first case results in no change in the tuning range. Under the same assumptions, the remaining two cases would allow an increase in tuning range since more variable capacitance can be added. Using (3.103) as well as (3.101) we can summarize the phase noise and power consumption results for these three cases in Table 3.2. Notice that as $Z_{o,30}$ reduces relative to $Z_{o,60}$, the power consumption for the push-push oscillator increases while the possible tuning range increases. The case of constant Z_o scaling provides a decent trade-off between these two competing requirements. In this case, if the achievable tank quality factor is higher at 30GHz than at 60GHz, the push-push oscillator will provide better phase-noise with lower power consumption and increased tuning range.⁸

⁸Scaling factors for Z_o outside of this range are generally not useful. Using $Z_{o,30} > 2Z_{o,60}$ would require a reduction in C_T leading to lower tuning range than what we started with. On the other hand, using $Z_{o,30} < 0.5Z_{o,60}$ would mean reducing L_T . While allowing larger increases in capacitance, significant increases in the power consumption required make this case undesirable as well.



(a)



(b)

Figure 3.26: Tank quality factor.

Case	Z_n	Push-push better if:	Power Consumption	Push-Push Tuning Range
1	2	$Q_{T,30} > 2Q_{T,60}$	$I_{b,pp} = \frac{1}{2}I_{b,fund}\frac{Q_{T,60}}{Q_{T,30}}$	-
2	1	$Q_{T,30} > Q_{T,60}$	$I_{b,pp} = I_{b,fund}\frac{Q_{T,60}}{Q_{T,30}}$	↑
3	0.5	$Q_{T,30} > 0.5Q_{T,60}$	$I_{b,pp} = 2I_{b,fund}\frac{Q_{T,60}}{Q_{T,30}}$	↑↑

Table 3.2: Push-push versus fundamental oscillator selection.

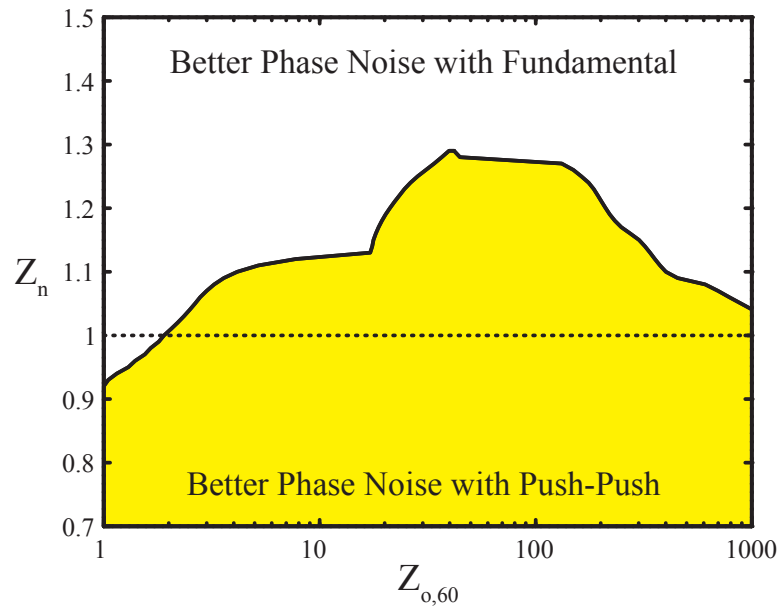


Figure 3.27: 60GHz push-push oscillator design space.

In general, capacitive quality factors decrease with frequency while inductive quality factors increase. Due to skin effect and self-resonance, the quality factor of inductors levels off at high frequencies. Since the tank quality factor is simply the parallel combination of the component quality factors (3.26), it should be dominated by the inductive quality factor at low frequencies, increasing with frequency, while at high frequencies it should be dominated by the capacitive quality factor, decreasing with frequency. There is thus a frequency at which the tank quality factor is maximum. Fig. 3.26a shows a representative plot of maximum achievable tank quality factor versus frequency for a standard 90nm CMOS process with no special RF options when using a single turn inductor. Indeed, we see that there is a peak near 30GHz. The location of this peak in general, however, will depend on the technology used as well as the scaling methodology. This implies that, with constant Z_o scaling, a push-push oscillator should be used to generate 60GHz.

Another way to approach this problem is through the tank Z_o , the selection of which is determined by the tuning range, phase noise and power consumption requirements. A larger

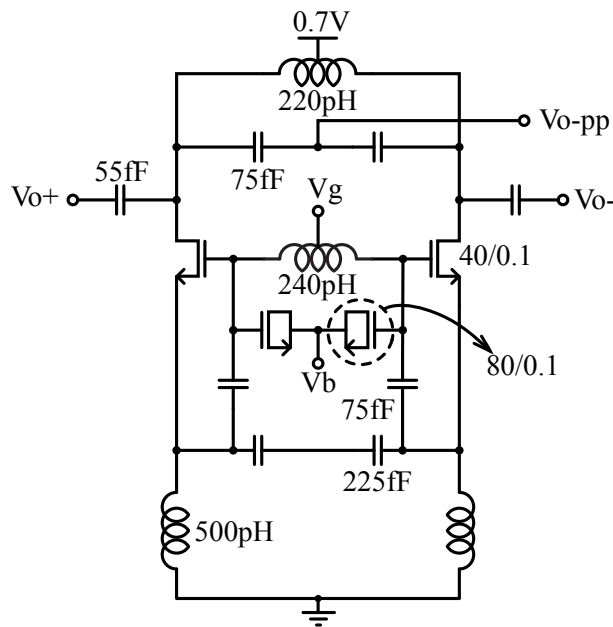
Z_o will lead to increased phase noise but decreased power consumption and tuning range. Let us construct a plot of maximum achievable tank quality factor versus tank Z_o at the two frequencies of interest in the same 90nm process in a more rigorous fashion. Both inductive and capacitive quality factors are modeled based on measured data. The actual inductance topology depends on the value of inductance required. For values less than approximately 100pH it was found that a short length of shorted transmission line provides higher quality factor than a lumped inductor. For values of inductance between approximately 100-400pH a single turn inductor is optimal. Beyond 400pH, a two turn inductor provides higher quality factor. The optimum dimensions for peak quality factor for each inductor value is simulated using a 3D EM field solver such as HFSS. The result (Fig. 3.26b) shows the tank quality factor versus Z_o for the two frequencies of interest. Using this data, for each value of $Z_{o,60}$ we can find the range of Z_n where $Q_{T,30} > Z_n Q_{T,60}$. The result is plotted in Fig. 3.27. The shaded area below the curve represents the design space for which a push-push oscillator will show better phase noise performance than a fundamental oscillator. Thus, we can see that for most values of $Z_{o,60}$, constant Z_o scaling ($Z_n = 1$) can be used to allow a good trade-off between power consumption and tuning range as discussed previously.

3.6 Design Case Studies

Two prototype oscillators were designed utilizing the theory presented above. The intended application for both prototypes is a direct conversion 60GHz transceiver. Thus, both oscillators must provide a strong 60GHz output to be used as the up- and down-conversion LO for the transmit and receive mixers respectively as well as provide sufficient tuning range to cover the 60GHz band over process, voltage, and temperature variations. First we will discuss the design of a push-push oscillator in 90nm CMOS. This prototype exhibits higher output power than other published push-push oscillators, however, this level is still not sufficient for low power 60GHz transceivers. Thus, a second prototype was designed in 65nm CMOS utilizing a fundamental mode topology. Due to the higher gain and smaller capacitive loading possible in the 65nm process, the fundamental mode oscillator achieves higher output power and tuning range while providing identical phase noise performance and consuming similar amounts of DC power.

3.6.1 Push-push Oscillator Prototype

It is well known that if a differential pair is driven by a signal at frequency f_o , the common-source node exhibits a strong response at $2f_o$. This is because each device acts as a source follower during one half of the f_o cycle causing the common source node to rise and fall two times for each period of f_o , generating a signal at $2f_o$. The behavior is also seen in cross-coupled differential pair oscillators making this an example of a simple push-push oscillator as presented in [Copani10]. However, the signal available at this node is very small, limiting the power that can be extracted at the second harmonic (in that work only -25dBm).



(a) Schematic.

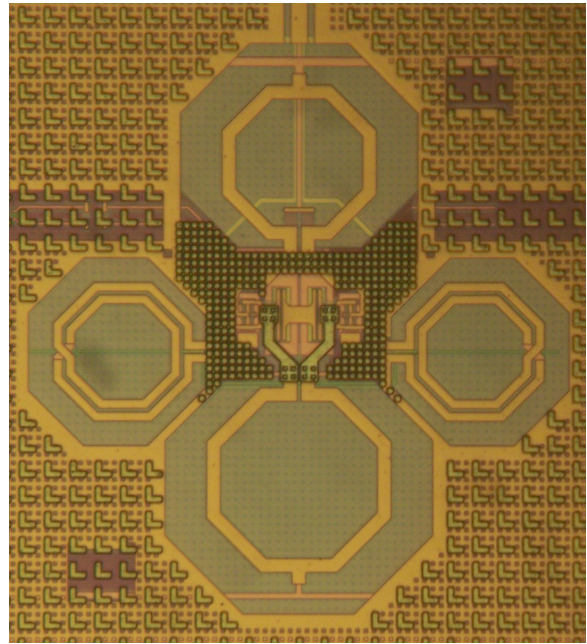
(b) Die photo (size: $320\mu\text{m} \times 340\mu\text{m}$).

Figure 3.28: Push-push oscillator prototype.

The differential common-drain Colpitts oscillator is another great candidate for the push-push topology. A prototype 60GHz push-push oscillator based on this topology was designed in a standard digital 90nm CMOS process with no special RF options. The schematic is shown in Fig. 3.28a and the die photo in Fig. 3.28b. The 30GHz core of the oscillator is a differential common-drain Colpitts topology with a secondary output tank at the drain [Dauphinee97] to generate a buffered 30GHz output. The core was sized and biased to produce strong second harmonic currents. The push-push output can then be taken from any common-mode point in the circuit. Based on earlier work by [Smith89] and [Kobayashi99], [Voinigescu00] suggests taking the output from the center of the gate inductor. However, in order to avoid disturbing the main 30GHz tank, the push-push output is taken from the output tank instead by using a center tapped capacitor.

This oscillator was implemented as part of the LO generation subsystem of a 60GHz direct conversion transceiver [Marcu09]. The single-ended push-push output was utilized as the 60GHz LO signal for both the transmitter and receiver, while the differential 30GHz output was used in the integrated PLL to lock the VCO to a stable external reference. The output tank of the VCO was thus sized to maximize the 60GHz output power while still providing sufficient differential signal swing at 30GHz to drive the following divider stages.

To reduce the phase noise of the VCO, chokes are utilized instead of current sources. However, this means that all three terminals of the MOS transistors in the oscillator core are connected to inductors, resulting in very large voltage swings above and below the supply

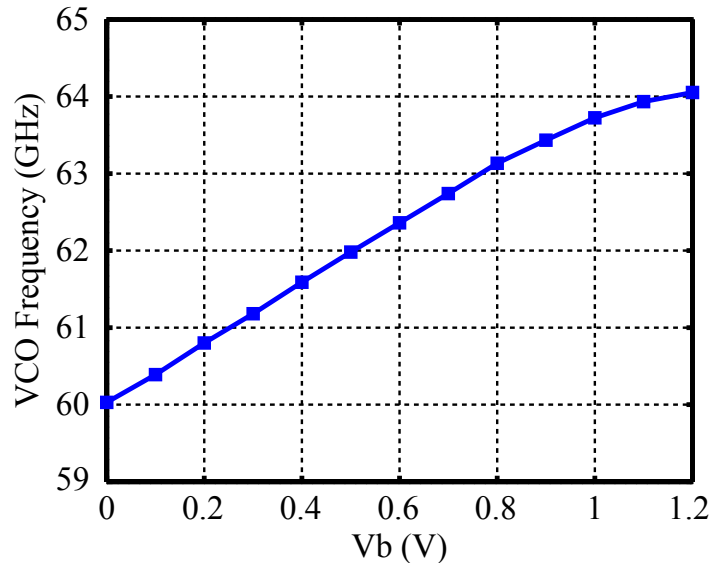


Figure 3.29: Push-push oscillator measured tuning range.

rails. To limit these swings for reliability, the VCO is operated from a 0.7V supply, drawing 12mA of DC current. IMOS varactors ($80\mu\text{m}/90\text{nm}$) are added in parallel with the capacitive divider to provide frequency tuning without affecting the feedback factor.

The VCO has a measured tuning range of 59.6-64GHz (7.1%) at the push-push output, as shown in Fig. 3.29. The tuning range is shifted up in frequency due to overcompensation of the expected parasitics in the design. The simulated output power at 60GHz is -9dBm, whereas the measured power at the designed bias point is only -20dBm. Adjusting the bias leads to a maximum measured output power of -14dBm. Simulations show that the reduction in measured versus simulated output power can be explained by a significant reduction in the loaded Q of the resonant tank. However, another likely culprit is incorrect prediction of the second harmonic generation due to the limited accuracy of the active device models at these high frequencies. Nevertheless, significant buffering is required at 60GHz to bring this signal up to the levels required by the mixers in the transceiver ($\sim 0\text{dBm}$). Finally, the measured phase noise (Fig. 3.30) at 10MHz offset from the 60GHz carrier is -112dBc/Hz.

3.6.2 Fundamental Oscillator Prototype

The low output power available from push-push designs is a significant stumbling block in direct conversion transceivers since it leads to excess power consumption from buffering. Furthermore, the output power is highly unpredictable leading to designs with large margins, again increasing power consumption in the overall system. A fundamental mode design, on the other hand, provides a simple, robust, and proven solution to 60GHz LO generation in deeply scaled CMOS. Thus, a fundamental mode cross-coupled pair oscillator, embedded in

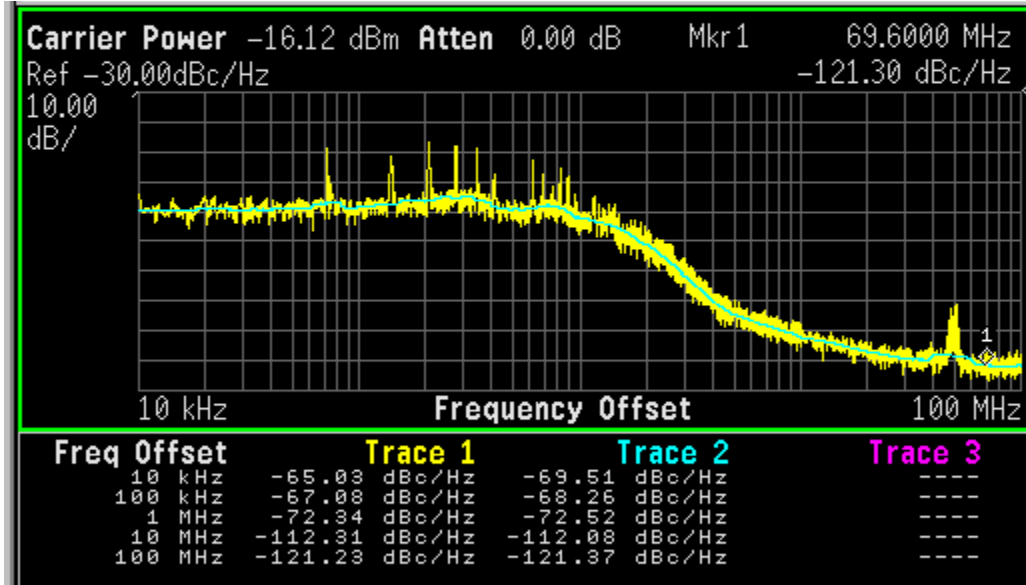


Figure 3.30: Push-push oscillator measured phase noise.

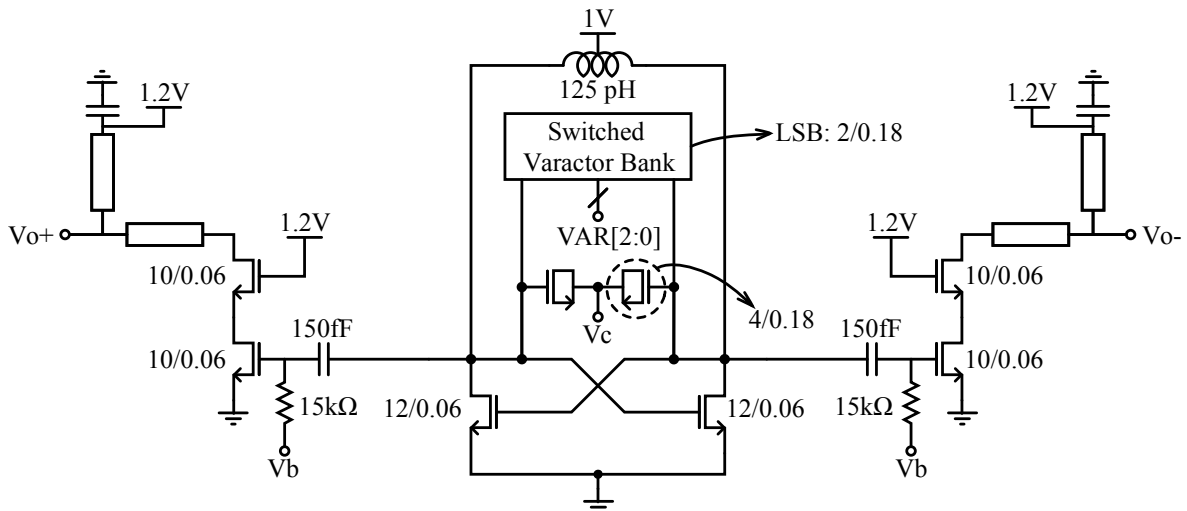


Figure 3.31: Cross-coupled oscillator schematic.

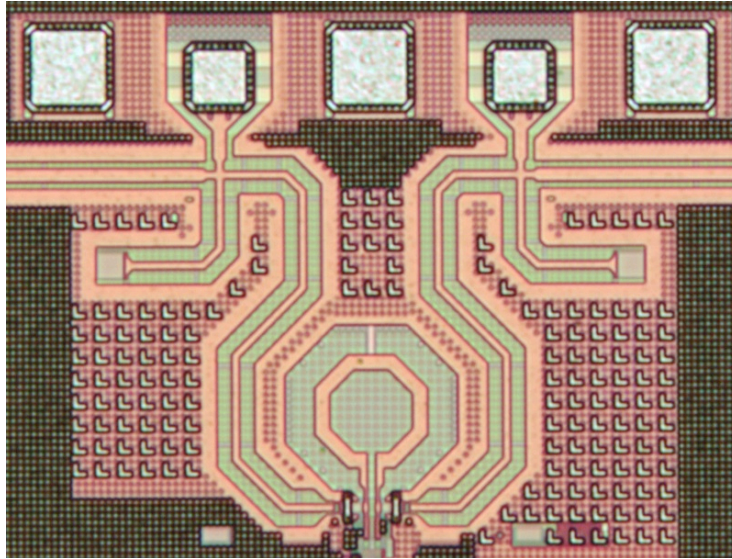


Figure 3.32: Cross-coupled oscillator die photo. (Area shown: $490\mu\text{m} \times 380\mu\text{m}$)

an integrated PLL, was chosen as the LO generator for a 60GHz 4-element direct conversion phased array transceiver in 65nm CMOS presented in [Tabesh11a]. The self-biased cross-coupled topology was chosen due to its simplicity and robustness, as well as the low supply voltage available. The core of this prototype is biased with a voltage in this design, nominally at 1V, for flexibility in testing. This voltage could be set with a programmable resistor in a final design. Alternatively, there is sufficient headroom available for a current source to be used instead at the expense of potentially higher phase noise. This selection must be made depending on the noise and bias stability requirements.

Fig. 3.31 shows the schematic of the oscillator and its buffers. The core consists of two cross-coupled $12\mu\text{m}/60\text{nm}$ NMOS transistors. The tank inductor is a single-turn 125pH octagonal inductor. Our previous work has shown that single-turn inductors are nearly optimal for low loss resonators at 60GHz [Marcu08b]. Tuning is achieved by a combination of an analog varactor and a 3-bit digitally switched binary varactor bank to achieve high tuning range while maintaining a low K_{VCO} for reduced noise sensitivity. All varactors are of IMOS type and are built using a unit cell with $W = 2\mu\text{m}$ and $L = 0.18\mu\text{m}$. This unit cell forms the LSB of the varactor bank. Matching is improved by using the same type and size of device for all varactors and using arrays to achieve the required sizes. To ensure overlap between tuning bands, the analog varactor is equal to 2LSB of the bank. The channel length of the varactor devices was chosen larger than the minimum of the process in order to increase the tuning range without significantly sacrificing tank quality factor. The final design achieves 12.4% tuning range with approximately 1.9GHz/V K_{VCO} per band. The oscillator has simulated phase noise of -115dBc/Hz @ 10MHz offset from the 60GHz carrier while consuming only 9mW from a 1V supply.

The output of the cross-coupled oscillator in this case must feed both the divider chain in the

PLL, as well as the buffers in the LO distribution network. In order to allow independent biasing, both the divider and the buffers are AC coupled to the core through coupling capacitors. Only the buffer path is shown in Fig. 3.31, however, another set of coupling capacitors is present in the design to connect the divider, exactly as is done for the buffers. Ideally, a very large coupling capacitor should be used to minimize signal loss due to the capacitive divider formed with the input capacitance of the buffer/divider. However there are two issues with a large coupling capacitor. First, the buffer must be made fairly large in order to drive the distribution network with sufficient power. With a large coupling capacitor, the entire buffer input capacitance would appear across the tank. Furthermore, such a large capacitor would have significant bottom plate capacitance, further loading the tank. However, if a small coupling capacitor is used, the capacitive divider reduces the effective capacitance seen by the tank to something smaller than both the input capacitance of the buffer or the coupling capacitor at the expense of a smaller signal reaching the buffer. Thus, the size of the coupling capacitor is chosen by trading off between tank loading and signal loss to the buffer.

It must be noted here that the process used in this design offers two flavors of devices with various threshold levels in order to address needs for either low power or high performance. Unfortunately, high frequency RF models were only available for the low power devices so these were used in the core of the oscillator in order to allow accurate simulation of both startup and phase noise. This does come at the expense of reduced performance so high performance devices were utilized in the buffers for higher efficiency. The models of the high performance devices had to be adjusted based on measurements by adding an additional gate resistance to account for non-quasistatic effects (*NQS*).

The buffer size was selected based on a trade-off between maximum achievable output power and loading of the oscillator tank. A matching network must then be utilized at the output of the buffer to transform the impedance presented by the distribution network to an optimum value for peak output power. A 50Ω transmission line distribution network was utilized. The selection of this impedance and its implications will be addressed in Chapter 5. In this case a buffer device size of $10\mu\text{m}/60\text{nm}$ with a conjugate output match provides the peak output power from the overall system. Each buffer consumes 3.6mW from the 1.2V global supply while providing 0dBm of nominal output power in simulation. The output match can be accomplished using a single stub topology with 80Ω CPW transmission lines. This allows for a very compact layout by meandering the transmission lines around the oscillator core's inductor as shown in the die photo, Fig. 3.32. The grounded stub provides drain biasing to V_{dd} and is grounded through a large AC bypass capacitor. Pads are added at the transition to the 50Ω transmission line for direct probing of the buffer output. The loading from the pad is mostly capacitive but must be taken into account when designing the matching network.

The measured VCO tuning range over all 8 tuning bands is 57.9-65.6GHz (Fig. 3.33), equivalent to a tuning range of 12.4%. The frequency shift between measurement and simulation is only 1.7GHz, a shift of less than 3%, showing the validity of the design approach. The VCO has a measured free running phase noise of -112dBc/Hz at 10MHz offset from the 60GHz

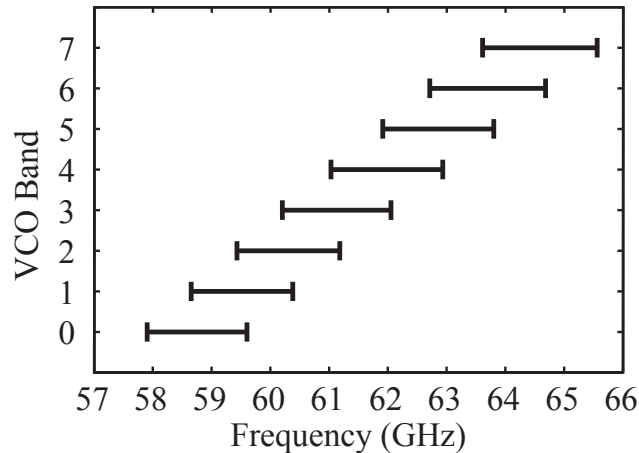


Figure 3.33: Cross-coupled oscillator measured tuning range.

carrier, a 3dB increase from simulation. The measured output power from one LO buffer is -1.8dBm, for a total differential power of +1.2dBm, less than only 2dB lower than predicted from simulation.

3.6.3 Performance Summary and Comparison

Two commonly used figures of merit for VCOs are FOM and FOM_T , the latter of which includes the tuning range.

$$FOM = \mathcal{L}\{\Delta f\} - 20 \log \left(\frac{f_{osc}}{\Delta f} \right) + 10 \log (P_{DC}) \quad (3.106)$$

$$FOM_T = FOM - 20 \log \left(\frac{FTR}{10} \right) \quad (3.107)$$

In the definitions above f_{osc} is the nominal output frequency, $\mathcal{L}\{\Delta f\}$ is the phase noise at an offset Δf from f_{osc} , P_{DC} is the DC power consumption in mW, and FTR is the frequency tuning range in %. Tables 3.3 and 3.4 summarize the measured results from the push-push and fundamental oscillators presented above respectively, and provide comparisons to the current state of the art at 60GHz. The figures of merit allow designs at different frequencies to be compared fairly and have been included in the tables.

The push-push oscillator presented here achieves the best FOM and FOM_T among 60GHz multi-push oscillators due to its low power consumption and competitive phase noise and tuning range. The output power is higher than all previously reported multi-push oscillators except [Liu04] which consumes 14 times more DC power. However, significant buffering would still be required for this push-push oscillator to be useful in most applications.

The fundamental oscillator, on the other hand, achieves competitive FOM and FOM_T among 60GHz fundamental oscillators. The output power is by far the highest reported

	Freq. (GHz) (TR)	P_{dc} (mW)	P_{out} (dBm)	PN (dBc/Hz) (@ 1MHz)	FOM	FOM_T
[Liu04]	62-64.5 (4%)	118	-6	-85	-160.3	-152.2
[Cho05]	52-52.5 (1%)	27.3	-16	-97	-177	-156.6
[Copani10]	52-58 (10.9%)	10	-25	-89	-173.8	-174.6
[Chiu10]	64.2-69.4 (7.5%)	27.5	-18	-76.23	-158.6	-156.2
[Catli10] (triple-push)	63.2-72.4 (13.5%)	18	-28	-95 (@ 10MHz)	-159.1	-161.7
This work	59.6-64 (7.1%)	8.4	-14	-112 (@ 10MHz)	-178.6	-174.8

Table 3.3: Push-Push oscillator performance summary and comparison

	Freq. (GHz) (TR)	P_{dc} (mW)	P_{out} (dBm)	PN (dBc/Hz) (@ 1MHz)	FOM	FOM_T
[Jimenez09]	51.7-61 (16.6%)	15 (buf: 9.6)	-9.8	-99.35	-182.6	-187
[Zhang09]	55.4-60.3 (8.5%)	15.6	-	-112 (@ 10MHz)	-175.3	-173.9
[Decanis11]	56-60.35 (7.5%)	22	-	-95	-176.8	-174.3
[LaRocca09]	58.3-63.8 (9.2%)	10.6 (buf: ??)	-5	-90.1	-175.6	-174.7
[Borremans08]	59-65.2 (10%)	3.9 (buf: 3.8)	-15	-95	-185	-184.9
[Li09]	61.1-66.7 (8.8%)	3.16 (buf: ??)	-14	-95	-186.1	-185
[Parvais10]	63-69 (9.1%)	28.6	-	-93	-174.8	-174
This work	57.9-65.6 (12.5%)	9 (buf: 7)	+1.2	-112 (@ 10MHz)	-178.3	-180.2

Table 3.4: Fundamental oscillator performance summary and comparison

among CMOS oscillators at 60GHz (including taking into account power consumption in the buffers). The differential output power is greater than 0dBm and is high enough to directly drive most mixers without requiring additional buffering. From simulation, phase noise was expected to be approximately 3dB lower than measured. While the current level is still competitive, future designs should focus on reducing this number further. At the same time, this design achieves the second highest tuning range reported near its frequency, covering the entire 60GHz band [IEE09]. However, to achieve sufficient margin over PVT variations, the tuning range should be increased to approximately 15%. This could be achieved with a larger varactor bank without significantly degrading power consumption or phase noise due to overdesign in the current version. Another solution is to use multiple VCOs with narrower tuning ranges to cover different parts of the band ([Parvais10]). However, this comes at the expense of excess area and power due to the extra VCOs and selection circuitry.

On first glance, [Jimenez09] seems to achieve the best solution to this problem by the judicious design of a single VCO. However, the K_{VCO} in that work varies between 1.5GHz/V at the band edges to 5.5GHz/V at mid-band. By comparison, the work presented here has a maximum K_{VCO} of 1.9GHz/V over the entire band. A higher K_{VCO} means that the oscillator is more sensitive to noise on the analog control line and could also cause instability in a PLL (see Chapter 4). Ideas such as the distributed DiCAD resonator presented in [LaRocca09] seem to provide the most promising solution to this problem at mm-wave frequencies. This type of resonator has been pushed to even higher tuning ranges in [Murphy10] while maintaining low phase noise and K_{VCO} below 1GHz/V.

Chapter 4

Low Power Phase Locked Loop Design

In order for radios to communicate with each other they must both be tuned to the same frequency. In general, standards and laws also dictate the absolute frequencies at which radios must operate. Unfortunately, the frequency accuracy of an integrated oscillator, such as the ones described in the previous chapter, cannot be guaranteed. Process variations cause the natural frequency between any two oscillators to be different. Furthermore, voltage and temperature variations will cause drifts in the instantaneous frequency of an oscillator. In many cases, it is also important for the oscillator to have a known phase but at startup the phase of any oscillator is completely random. Therefore, there needs to be a way to lock an oscillator to a known reference both in frequency and in phase. A Phase Locked Loop (PLL) is a feedback control system which performs just this function.

A general PLL is shown in Fig. 4.1. It consists of a VCO, a divider, a phase detector (PD), and a low pass loop filter (LPF). For now let us assume that $N = 1$ for simplicity, meaning that the LO is fed back unchanged to the PD. The phase detector then compares the phase of the LO to the phase of the reference and outputs a signal proportional to the phase difference, ϕ_{err} . The low pass filter then converts this signal to a voltage, V_C , and filters out any high frequency noise. The output frequency of the VCO is controlled by V_C thus closing the feedback loop. The components and topology should be designed such that

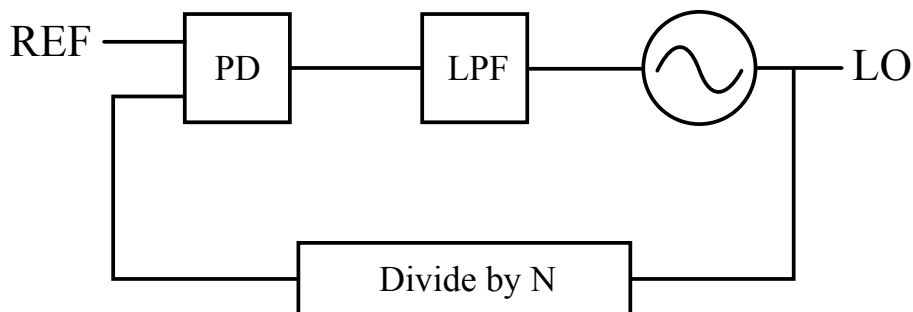


Figure 4.1: A general phase locked loop.

the overall feedback is negative and stable, locking the VCO phase to the reference phase.

Qualitatively we can say that on average the negative feedback loop will cause the VCO output to track the reference. This includes both the wanted reference signal as well as its unwanted noise. The low pass loop filter plays an important role here in that it determines the bandwidth of the feedback loop. Within the loop bandwidth, operation proceeds as described above. Outside this bandwidth, however, the phase error moves too quickly and V_C can no longer track it out. Thus, outside the loop bandwidth the VCO output is unaffected by the loop. The overall effect is that the average frequency of the LO and the close-in phase noise (within the loop bandwidth) will track the reference, while the far-out phase noise (outside the loop bandwidth) will be equal to the free-running VCO phase noise. The loop filter in high frequency PLLs is generally passive to reduce noise and power consumption. However, an active loop filter can provide gain higher than unity and is useful in certain circumstances. Active loop filters are beyond the scope of this work since our focus is on low power design.

The advantage of this system is that we can lock our inaccurate, noisy, on-chip oscillator to a known, calibrated, hopefully less noisy, off-chip reference. This ensures that separate systems can talk to each other and that they operate only at frequencies dictated by standards and applicable laws. However, to generate a 60GHz LO in our simple example we would need a high quality, off-chip 60GHz reference. Unfortunately, this is generally not available or is too big and power hungry to be used in a mobile device.

So far we have ignored the divider in this discussion. By dividing down the LO before comparing it to the reference, we reduce the required reference frequency. For the loop to achieve lock to the LO frequency, f_{LO} , must be equal to

$$f_{LO} = N \cdot f_{ref} \quad (4.1)$$

The operation of the loop does not change but the speed requirements on the phase comparison path can be greatly relaxed. Furthermore, the reference frequency can now be much lower than the LO, allowing us to use small and cheap low frequency sources such as crystal oscillators which have excellent frequency accuracy and stability (low noise).

The LO signal must be tuned to different channels. In this configuration, for a given reference frequency, the output frequency of the PLL can only be changed by adjusting the division ratio, N . Since N can only take on integer values, the output frequency can be adjusted in steps equal to f_{ref} . For narrower steps, the reference frequency can be divided down by M before being applied to the PLL. In that case the output frequency is given by

$$f_{LO} = \frac{N}{M} f_{ref} \quad (4.2)$$

For even better frequency resolution, a fractional divider must be used which, as the name suggests, allows for sub-integer division ratios. Fractional-N frequency synthesizers are beyond the scope of this work.

In this chapter, we will describe low power PLL design at 60GHz. We will begin with an overview of PLL dynamics, followed by analysis of noise contributors, and the design of

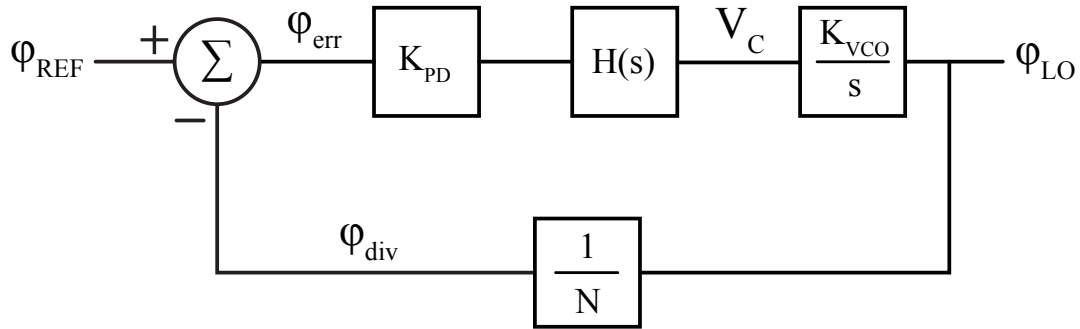


Figure 4.2: Phase domain model.

individual building blocks. In particular, we will focus on the design of frequency dividers which are the largest power consumers in a high frequency PLL, after the VCO. Finally, we will discuss trade-offs at the system level and present a sample design.

4.1 Phase Locked Loop Dynamics

Since a PLL is a control system whose purpose is to lock the phase (and frequency) of the VCO to the reference, the easiest way to analyze and design a PLL is using the linearized phase domain model shown in Fig. 4.2 in which each block has been replaced by its equivalent transfer function. A PLL and its components are very nonlinear, however, for small phase error (i.e.: when the loop is locked), a linearized model can be assumed. Furthermore, the purpose of this model is to allow analysis of noise contributions to the output phase noise. Since noise is a small signal phenomenon, the linear models can remain accurate.

4.1.1 The Linear Phase Domain Model

In the linearized phase domain the phase detector is modeled using two blocks: an ideal summer and a gain, K_{PD} , which depends on the particular phase detector topology. The simplest form of phase detector is a multiplier, which can be implemented as an XOR gate if the comparison frequency is sufficiently low ($< 1GHz$). With this type of PD the output, E , has a duty cycle which is set by the phase difference between REF and the divided clock, DIV, as shown in Fig. 4.3. The phase difference is also called the phase error, ϕ_{err} .

$$\phi_{err} = \phi_{REF} - \phi_{DIV} \quad (4.3)$$

The loop filter takes the average of the PD output, $E(t)$, to give $\langle E(t) \rangle$. To get the PD gain, K_{PD} , we simply take the derivative of $\langle E(t) \rangle$. For ϕ_{err} between 0 and π , the gain is

$$K_{PD} = \frac{1}{\pi} \quad (4.4)$$

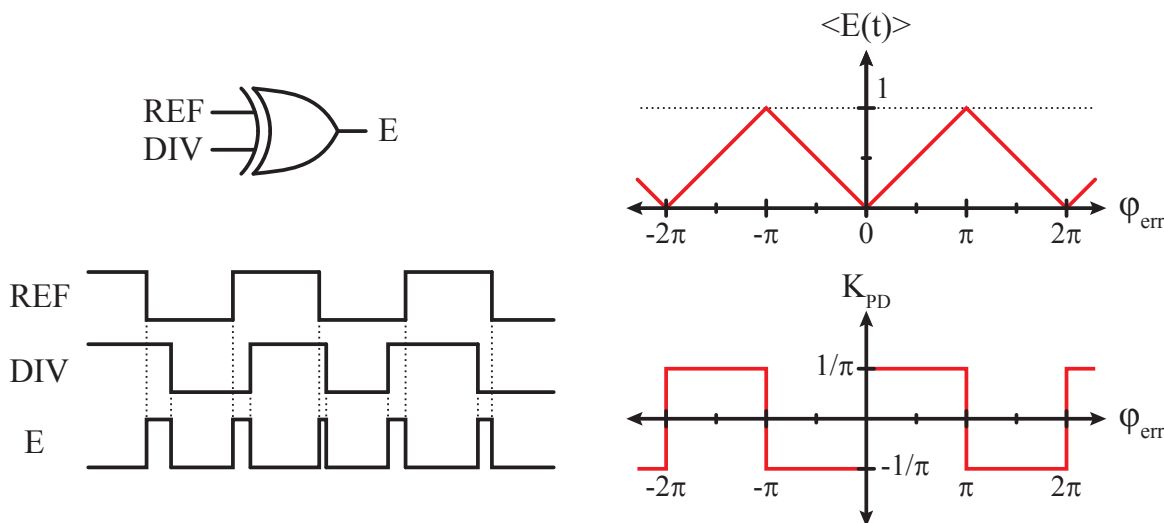


Figure 4.3: XOR phase detector.

This type of PD can be either single ended as shown or differential, providing $E^+(t)$ and $E^-(t)$. In the differential case, $\langle E(t) \rangle$ has the same sawtooth shape but the minima and maxima are at -1 and $+1$ respectively instead of 0 and $+1$ as in the single ended case. The differential PD gain is thus equal to $\pm 2/\pi$. In the locked condition, a PLL using this type of PD would lock the divided clock in quadrature with the reference (i.e.: $\phi_{err} = \pi/2$)¹. Notice that $\langle E(t) \rangle$ has even symmetry about the origin. If the phase error grows beyond the bounds $(0, \pi)$ the magnitude of K_{PD} does not change, but its sign becomes negative leading to positive feedback and instability in the loop.

The loop filter is simply modeled by its transfer function, $H(s)$, which is the Laplace transform of its impulse response, $h(\tau)$. The output frequency of the VCO is a function of the control voltage and the VCO gain, K_{VCO} , which has units of $[rads/V]$. Since phase is the integral of frequency, the VCO is modeled by an integrator with gain K_{VCO}/s .

$$\omega_{LO}(t) = K_{VCO}V_C(t) \quad (4.5)$$

$$\begin{aligned} \phi_{LO}(t) &= \int_{-\infty}^t \omega_{LO}(\tau) d\tau \\ &= \int_{-\infty}^t K_{VCO}V_C(\tau) d\tau \end{aligned} \quad (4.6)$$

$$\phi_{LO}(s) = \frac{K_{VCO}}{s} V_C(s) \quad (4.7)$$

¹In most cases coarse tuning of the VCO is used to bring the control voltage during lock near the middle of its range. This helps to linearize the PLL characteristics. In a Type-I loop either a single-ended or differential XOR PD can be used and the actual locked phase will depend on the desired VCO control signal to achieve lock, as well as the gain from the PD output to the VCO, if any. In a Type-II loop a differential XOR PD must be used and the locked phase will be exactly equal to $\pi/2$ since this type of loop can force the VCO control signal to any desired value while maintaining zero frequency error (see Section 4.1.4).

Finally, since the divider divides a frequency by a factor of N , it also divides phase by the same factor.

$$\omega_{div}(t) = \frac{\omega_{LO}(t)}{N} \quad (4.8)$$

$$\begin{aligned} \phi_{div}(t) &= \int_{-\infty}^t \omega_{div}(\tau) d\tau \\ &= \int_{-\infty}^t \frac{\omega_{LO}(\tau)}{N} d\tau \\ &= \frac{1}{N} \phi_{LO}(t) \end{aligned} \quad (4.9)$$

$$\phi_{div}(s) = \frac{1}{N} \phi_{LO}(s) \quad (4.10)$$

The divider is thus simply modeled in the phase domain as a gain of $1/N$.

We can now use standard feedback techniques using linear transfer functions to determine the loop phase response, stability, etc. The loop gain $A(s)$ is given by

$$A(s) = \frac{K_{PD}H(s)K_{VCO}}{Ns} \quad (4.11)$$

The closed loop gain from input to output is then

$$\begin{aligned} G(s) &= \frac{\phi_{LO}(s)}{\phi_{REF}(s)} \\ &= \frac{K_{PD}H(s)K_{VCO}/s}{1 + A(s)} \\ &= \frac{K_{PD}H(s)K_{VCO}N}{Ns + K_{PD}H(s)K_{VCO}} \end{aligned} \quad (4.12)$$

At DC the closed loop gain reduces to N ,

$$G(0) = N \quad (4.13)$$

This means the PLL multiplies low frequency phase noise from the input by a factor of N , the division ratio, to the output. Usually, for analysis, the closed loop gain is normalized by dividing by N in order to give unity gain at DC.

$$\begin{aligned} \frac{G(s)}{N} &= \frac{A(s)}{1 + A(s)} \\ &= \frac{K_{PD}H(s)\frac{K_{VCO}}{N}}{s + K_{PD}H(s)\frac{K_{VCO}}{N}} \end{aligned} \quad (4.14)$$

Next, the error function can be found using $A(s)$

$$\begin{aligned} E(s) &= \frac{\phi_{err}(s)}{\phi_{REF}(s)} \\ &= \frac{1}{1 + A(s)} \\ &= \frac{s}{s + K_{PD}H(s)\frac{K_{VCO}}{N}} \end{aligned} \quad (4.15)$$

Note that $E(s)$ can also be written as

$$E(s) = \frac{1}{1 + A(s)} = 1 - \frac{A(s)}{1 + A(s)} = 1 - G(s) \quad (4.16)$$

showing that $E(s)$ and $G(s)$ are always complementary. From (4.12) and (4.15) we can see that the closed loop gain is inherently low pass in nature (assuming the loop filter is low pass) while the error function, inherently complementary, is high pass in nature. This is further explored by examining cases with different loop filters.

4.1.2 First Order PLL

The simplest example is that in which the loop filter is completely removed (i.e.: $H(s) = 1$) giving

$$A_1(s) = \frac{K_{PD}K_{VCO}}{Ns} \quad (4.17)$$

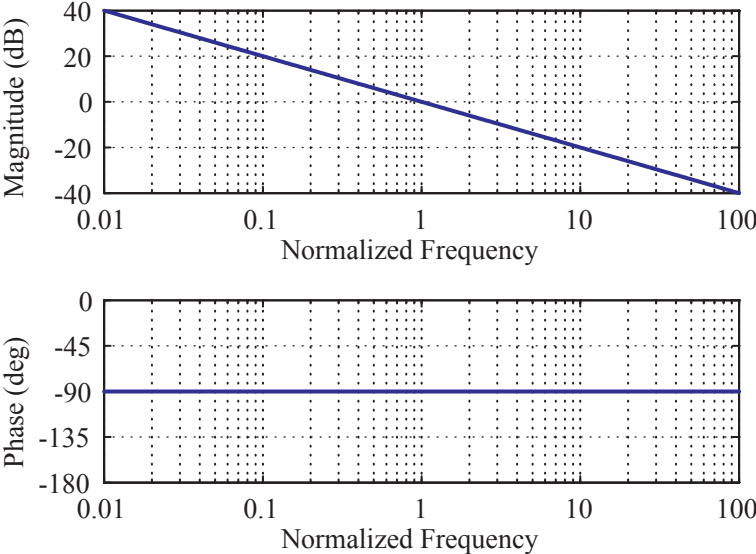
$$\frac{G_1(s)}{N} = \frac{K_{PD}\frac{K_{VCO}}{N}}{s + K_{PD}\frac{K_{VCO}}{N}} \quad (4.18)$$

$$E_1(s) = \frac{s}{s + K_{PD}\frac{K_{VCO}}{N}} \quad (4.19)$$

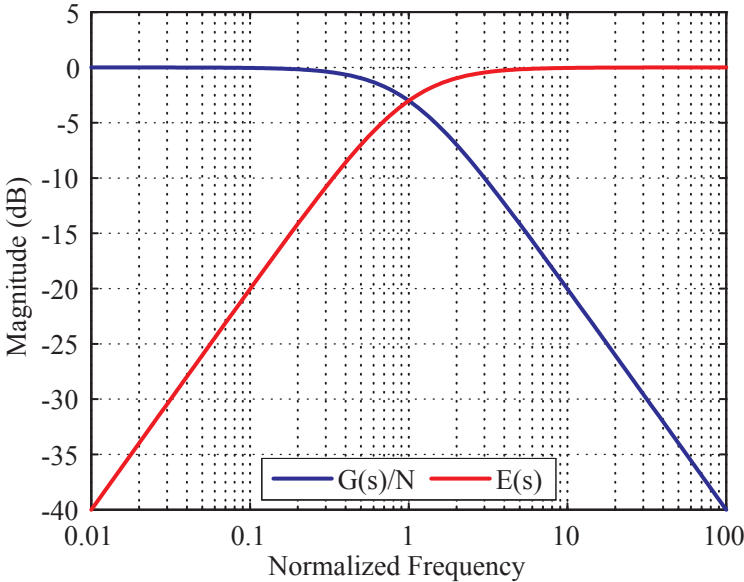
The Bode plot of $A_1(s)$ is shown in Fig. 4.4a, while the closed loop transfer functions are shown in Fig. 4.4b. This is called a first-order Type-I loop. The loop type is determined by the number of integrators, in this case 1 (the VCO). The loop order is the highest power of s in the denominator of the transfer function. Since this is only a first order loop, the phase margin is 90° and the loop is always stable. There are multiple definitions for loop bandwidth in the literature but for this work we will treat the cross-over frequency of the loop gain, ω_c , as the loop bandwidth (i.e.: the frequency at which $A(s) = 1$). The loop bandwidth in this case is simply determined by the gains of the components

$$\omega_C = \frac{K_{PD}K_{VCO}}{N} \quad (4.20)$$

From Fig. 4.4b, we can see that within the loop bandwidth the PLL passes any modulation on the input phase directly to the output unchanged (ignoring the multiplication by N),



(a) Bode plot of $A_1(s)$.



(b) Closed loop transfer functions.

Figure 4.4: First order loop response.

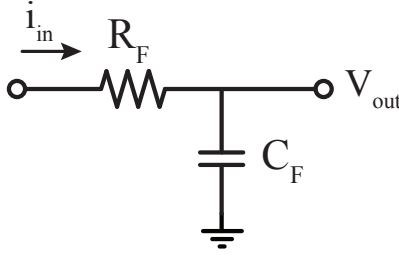


Figure 4.5: A first-order loop filter.

and suppresses ϕ_{err} . Outside the loop bandwidth, however, the error is no longer suppressed but the input is. Thus, any error generated outside the loop bandwidth passes directly to the output. In general this means any noise generated by the VCO is only suppressed inside the loop bandwidth, while noise coming from the reference is only suppressed outside the loop bandwidth (see Section 4.2 for details). This quantitative analysis confirms our earlier qualitative conclusions regarding the behavior of the loop inside and outside the loop bandwidth. Unfortunately, K_{PD} is determined by the topology and is fixed, while K_{VCO} is usually hard to change at will. The only way to change the loop gain would thus be to add an amplifier with gain A_V . The loop bandwidth could then be set using this amplifier.

$$\omega_{C,new} = \frac{A_V K_{PD} K_{VCO}}{N} \quad (4.21)$$

Unfortunately, the roll-off of the closed loop transfer functions is limited to $20dB/dec$ regardless of the amplifier gain.

4.1.3 Second Order PLL

Next we add the loop filter back into the loop. We begin with a simple, first order, RC loop filter as shown in Fig. 4.5 with the pole at ω_p .

$$H(s) = \frac{1}{1 + s/\omega_p} \quad (4.22)$$

$$A_2(s) = \frac{K_{PD} K_{VCO}/N}{s(1 + s/\omega_p)} \quad (4.23)$$

$$G_2(s) = \frac{K_{PD} K_{VCO} \omega_p / N}{s^2 + s\omega_p + K_{PD} K_{VCO} \omega_p / N} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (4.24)$$

$$E_2(s) = \frac{s(s + \omega_p)}{s^2 + s\omega_p + K_{PD} K_{VCO} \omega_p / N} = \frac{s(s + 2\zeta\omega_n)}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (4.25)$$

where $G_2(s)$ and $E_2(s)$ have been put into the standard form for second order systems by defining

$$\omega_n = \sqrt{\frac{K_{PD}K_{VCO}\omega_p}{N}} \quad (4.26)$$

$$\zeta = \frac{1}{2} \sqrt{\frac{\omega_p N}{K_{PD}K_{VCO}}} \quad (4.27)$$

The Bode plot of $A_2(s)$ is shown in Fig. 4.6a, while the closed loop transfer functions are shown in Fig. 4.6b. Notice that the presence of this second pole² reduces the phase margin and will introduce peaking in the closed loop transfer function (and ringing in the settling response) for $\zeta < 1$.

The rejection of the closed-loop transfer function, G_2 , is increased by the addition of the loop filter at ω_p as shown in Fig. 4.6b. However, the loop bandwidth still cannot be changed at will, while maintaining reasonable phase margin, without the addition of an amplifier (assuming K_{PD} , K_{VCO} , and N are out of our control). Furthermore, the roll-off of the error function, E_2 , is still only $20dB/dec$ and actually exhibits peaking near the loop bandwidth as the phase margin is reduced.

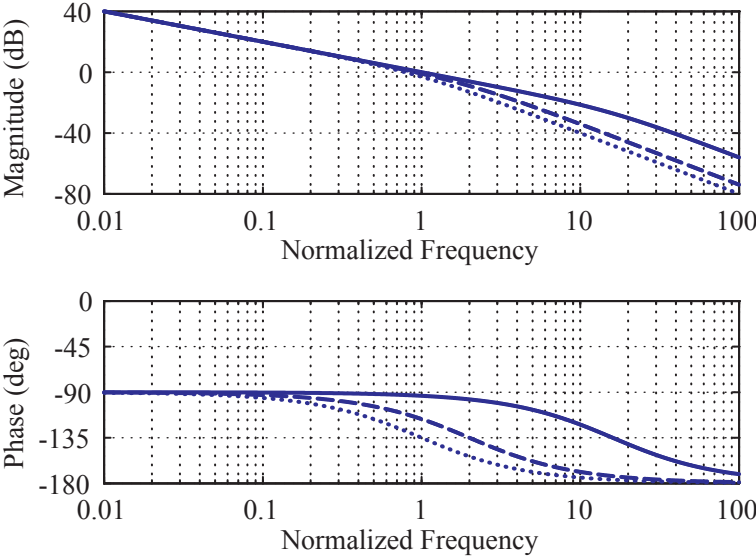
4.1.4 The Charge Pump and Phase Frequency Detector

We have found a number of problems with the simple first and second order loops. In addition to the ones we have already discussed, another problem is that this type of loop cannot maintain zero phase error over the full range of the VCO. This is because there is linear gain from the phase error to the control voltage, and thus, the output frequency³. As we saw in Fig. 4.3, the loop will tend to lock at a phase difference of $\pi/2$ between REF and the divided clock, leading to a control voltage near midrail. If a different VCO frequency is desired, the phase detector must output, and sustain, a different average value than midrail. This would imply that a phase error (a phase shift away from the $\pi/2$ locking point) must be maintained by the loop proportional to the frequency difference. However, if we add an integrator to the loop filter, the DC gain can be infinite. This would allow the control voltage (and thus the output frequency) to take on any value from 0 to V_{dd} irrespective of the average phase detector output. The integrator functionality is easily accomplished by the use of a charge pump and a capacitor as shown in Fig. 4.7a. The charge pump simply adds or removes charge from the capacitor, in proportion to the phase error, by pulsing a constant current with the appropriate polarity. In effect, this decouples the output of the phase detector from the VCO frequency as can be seen from Fig. 4.7b.

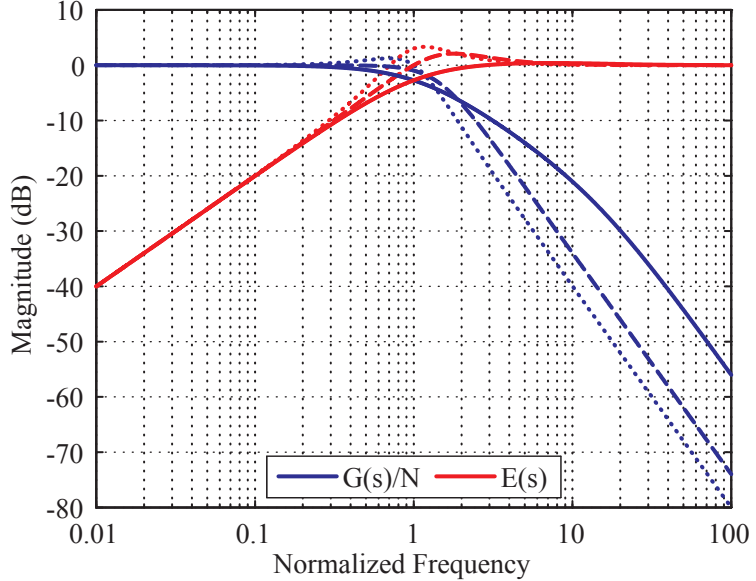
Another problem is that the XOR phase detector cannot always provide frequency acquisition. A constant frequency difference between the reference and divided clock would cause

²The first is a pole at DC introduced by the VCO which acts like an integrator.

³In fact, due to finite DC gain in the loop filter, the full VCO range may not even be accessible

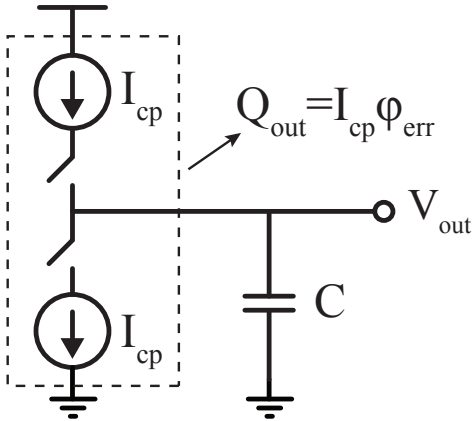


(a) Bode plot of $A_2(s)$.

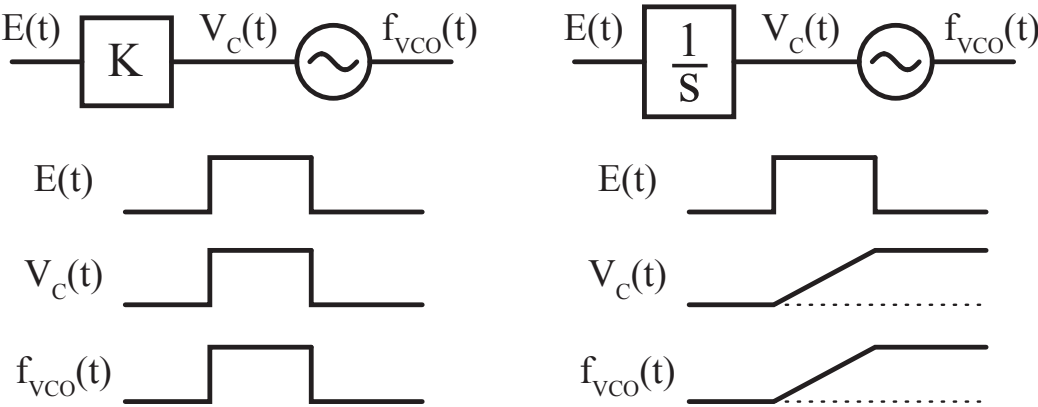


(b) Closed loop transfer functions.

Figure 4.6: Second order loop response (dotted: $\zeta = 1/2$, $PM = 52^\circ$; dashed: $\zeta = 1/\sqrt{2}$, $PM = 65^\circ$; solid: $\zeta = 2$, $PM = 86^\circ$).



(a) A charge pump integrator.



(b) Frequency control with gain block versus integrator.

Figure 4.7: Adding an integrator to the loop filter.

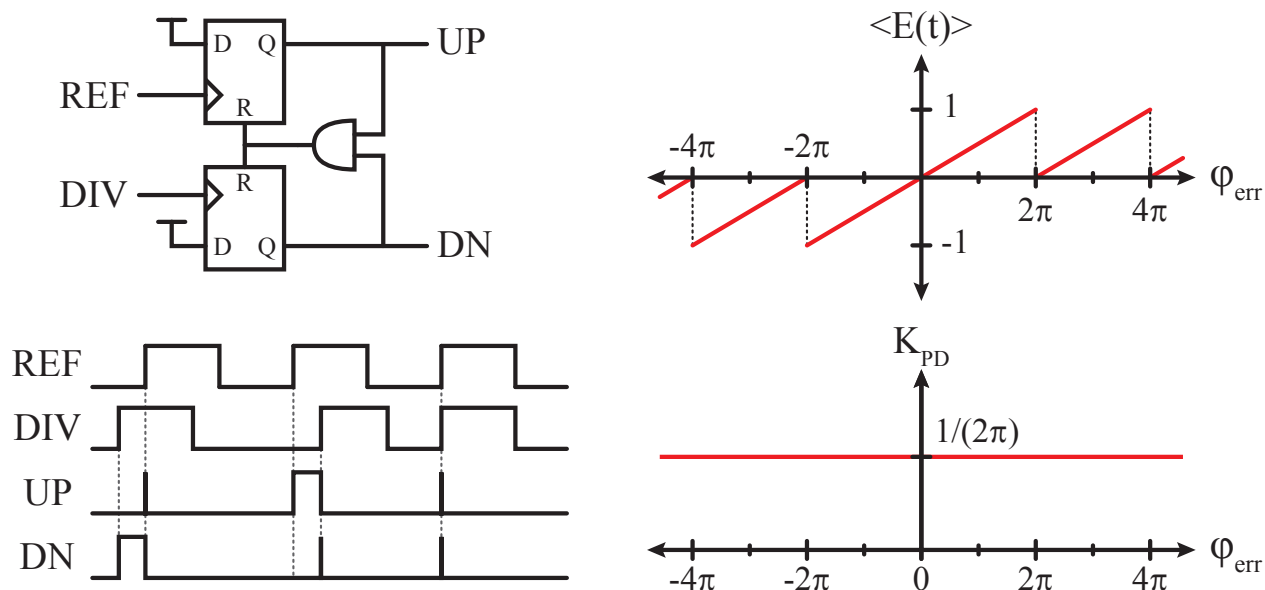


Figure 4.8: A flip-flop based Phase Frequency Detector (PFD).

the phase error to linearly ramp and grow beyond the bounds $(0, \pi)$ due to the integral relationship between frequency and phase. Thus, the phase detector gain would constantly flip sign going from a stable negative feedback to an unstable positive feedback. If the frequency error is small, the loop would eventually attain lock, although the settling time would be very long since the positive and negative excursions of the phase detector output would tend to cancel each other out. This process is called cycle slipping and for large enough frequency differences the loop may never attain lock. One solution would be to add a parallel frequency tracking loop which would only be activated to bring the frequency of the LO close enough to the desired frequency to avoid cycle slipping. Unfortunately this adds complexity to the system. A simpler solution is available, and it is called the Phase Frequency Detector (PFD).

A flip-flop based tri-state PFD was first presented by [Brown71] and is shown in Fig. 4.8. The inputs of both flos are tied high and the outputs are reset low. When an input arrives it clocks its respective flop changing its output to high. When both outputs have been set high by the arrival of their respective clocks both flos are then reset to low. In a real PFD design there must a minimum gate delay, T_{min} , before both flos are reset to ensure that even very short UP and DN pulses (e.g.: when the loop is close to lock) reach full amplitude since any switch has finite switching time. The effective output is the average of the difference

$$\langle E(t) \rangle = \langle UP(t) - DN(t) \rangle \quad (4.28)$$

and the effective gain is always

$$K_{PD} = \frac{1}{2\pi} \quad (4.29)$$

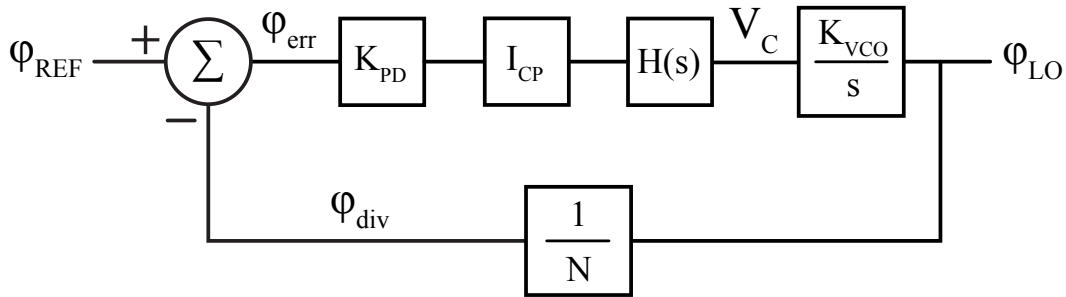


Figure 4.9: Charge pump based PLL.

Notice that the naming of the outputs of the PFD is the same as that of the inputs for the charge pump integrator. Indeed, a PFD is connected directly to a charge pump to control the addition or subtraction of current to the integrator cap. The operation of this combination in the loop is very straight forward to understand. If REF arrives before the divided clock, the PFD outputs larger UP pulses and only minimum DN pulses for reset. This charges up the capacitor, increasing the VCO control voltage. In turn, this should cause the VCO frequency to increase such that the period of the divided clock decreases, forcing it to catch up to REF. If, on the other hand, the divided clock arrives first, the PFD outputs larger DN pulses and only minimum UP pulses for reset. This discharges the capacitor and causes the divided clock period to increase, allowing REF to catch up. In the locked condition both REF and the divided clock arrive at the same instant causing both UP and DN to be enabled briefly before being reset. The length of these pulses is set by the gate delay in the reset path, T_{min} , which is critical to ensure proper operation of the PFD and charge pump. Both UP and DN currents should be matched by design such that no net charge is deposited or extracted from the loop filter and no change occurs on the control voltage.

Notice from Fig. 4.8 that the phase error characteristic has odd-symmetry about the origin and thus the gain is always constant at the same value and polarity regardless of phase error. There is no cycle slipping problem with this type of phase detector since as the phase error ramps, the output will always have the correct polarity, pushing the loop toward lock. A loop using a PFD will always regain lock in this way. ⁴

4.1.5 The Charge Pump PLL

Putting everything together we arrive at the most popular type of analog PLL in use today at high frequencies, the charge pump based, third order Type-II PLL with a PFD driving the charge pump. The complete phase domain model is shown in Fig. 4.9. Since the charge

⁴There is one other limitation we should mention here which must be taken into account. Due to the sampled nature of the loop, charge pump based PLLs must have a reference frequency that is at least 10 times the loop bandwidth to maintain good stability [Gardner80]. This is a separate consideration from the stability ensured by sufficient phase margin in the loop transfer function.

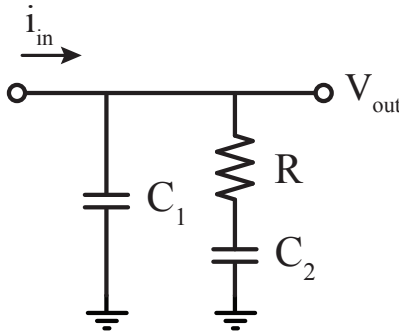


Figure 4.10: Second order loop filter for charge pump PLL.

pump integrator introduces an additional pole at the origin, a zero must be added to the loop filter to provide sufficient phase margin for stability. The most popular loop filter used for this type of design is the second order filter shown in Fig. 4.10, also called a lead-lag filter. Its transfer function is given by

$$H(s) = \frac{1 + s/\omega_z}{sC_{sum}(1 + s/\omega_p)} \quad (4.30)$$

where

$$C_{sum} = C_1 + C_2 \quad (4.31)$$

$$\omega_z = \frac{1}{RC_2} \quad (4.32)$$

$$\omega_p = \frac{C_1 + C_2}{RC_1C_2} \quad (4.33)$$

Notice that the pole in lead-lag filter is always located higher than the zero since it is set by the series combination of the two capacitors.

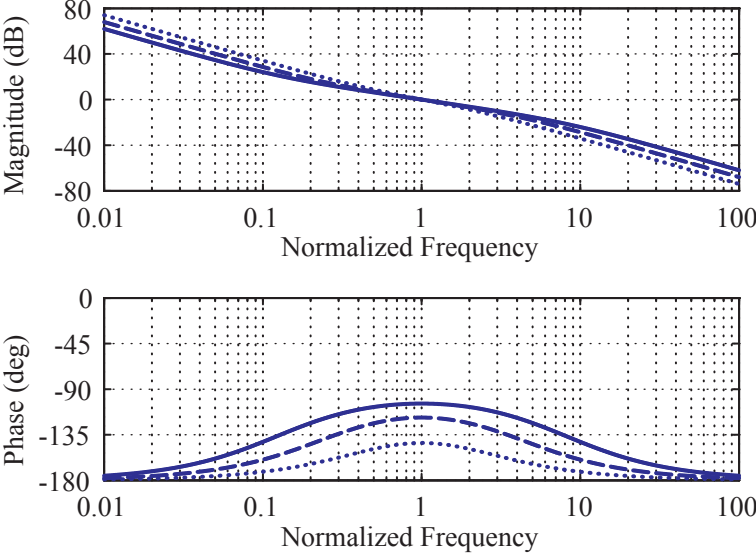
The charge pump simply converts the UP and DN pulses generated by the PFD to a current pulse with amplitude of $\pm I_{CP}$ which is then fed into the loop filter. At low frequencies the lead-lag filter looks like a capacitor equal to C_{sum} and, together with the charge pump, forms the integrator. The loop equations can be derived by inspection of Fig. 4.9.

$$A_3(s) = \frac{K_{PD}I_{CP}H(s)K_{VCO}}{Ns} \quad (4.34)$$

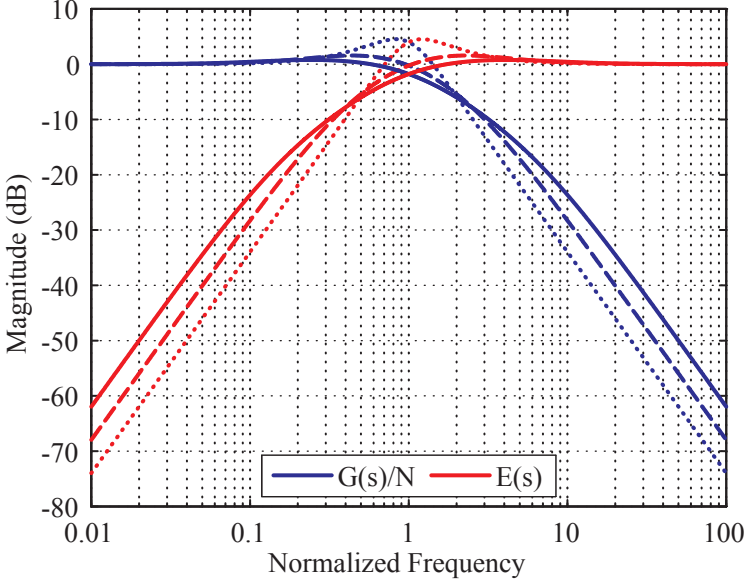
$$\frac{G_3(s)}{N} = \frac{K_{PD}I_{CP}H(s)K_{VCO}/N}{s + K_{PD}I_{CP}H(s)K_{VCO}/N} \quad (4.35)$$

$$E_3(s) = \frac{s}{s + K_{PD}I_{CP}H(s)K_{VCO}/N} \quad (4.36)$$

Notice that the loop gain can now be adjusted using the charge pump current, I_{CP} , allowing us to set the cross-over frequency of A_3 , and thus the closed-loop bandwidth, to any desired value.



(a) Bode plot of $A_3(s)$.



(b) Closed loop transfer functions.

Figure 4.11: Third order Type-II loop response (dotted: $f_p/f_z = 4$, $PM = 37^\circ$; dashed: $f_p/f_z = 16$, $PM = 62^\circ$; solid: $f_p/f_z = 64$, $PM = 76^\circ$).

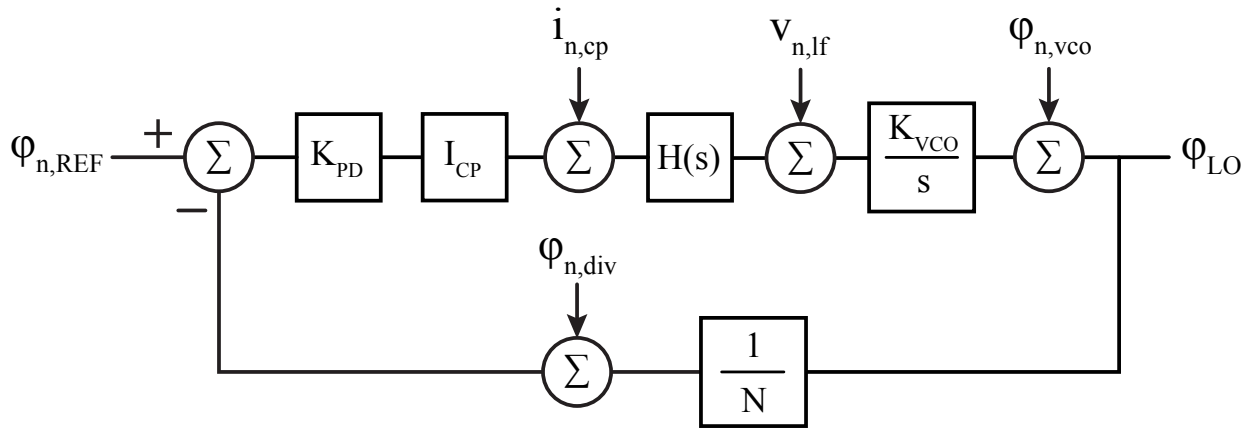


Figure 4.12: PLL noise sources.

The Bode plot of $A_3(s)$ is shown in Fig. 4.11a, while the closed loop transfer functions are shown in Fig. 4.11b all derived using the loop filter defined in (4.30). Both closed loop transfer functions show a $40dB/dec$ roll-off.⁵ For the best phase margin the loop bandwidth should be set to the geometric mean of the loop filter pole and zero. The actual phase margin and, thus, the peaking, settling time and overshoot are then set by the ratio of the pole to the zero as can be seen in Fig. 4.11a. Unfortunately, this type of loop shows peaking in both closed loop transfer functions, determined by the phase margin of A_3 . A more detailed analysis is beyond the scope of this work but can be found in many works on control systems or PLL design, including [Gardner05, pp. 25, 271].

4.2 Noise in Charge Pump Phase Locked Loops

In this section we will examine the effects of noise in the charge pump PLL using the filter shown in Fig. 4.10. The two main sources of noise in any well designed PLL are the phase noise of the reference and VCO. However, all other blocks will add their own noise contributions as shown in Fig. 4.12. Using the method of superposition we can find the contribution of each noise source to the output using the transfer functions derived above and add them up in power to find the total output noise of the PLL.

We start with the reference and VCO phase noise, $\phi_{n,REF}$ and $\phi_{n,VCO}$ respectively. These noise sources are modelled by adding the phase noise spectrum to the output in each case as shown in Fig. 4.12. To find the contribution of each to the total output phase noise we first have to determine the transfer function from each to the output. Luckily, these transfer functions have already been derived. The reference phase noise is at the input of the PLL so

⁵For additional high frequency rejection in G_3 , more poles can be added (e.g.: using an RC filter such as the one shown in Fig. 4.5). However, these poles must be placed much higher than ω_p of the lead-lag filter to ensure that they do not affect the phase margin and loop stability.

its transfer function to the output is simply the PLL's closed loop transfer function, $G(s)$. Thus the output noise contribution from the reference is given by

$$\phi_{n,out}\Big|_{REF} = \phi_{n,REF} \cdot |G(s)|^2 \quad (4.37)$$

By inspection we can see that the transfer function from the output of the VCO to the PLL output is equal to

$$\frac{\phi_{out}}{\phi_{VCO}} = \frac{1}{1 + A(s)} = E(s) \quad (4.38)$$

identical to the PLL's error function. The output noise contribution from the VCO is then

$$\phi_{n,out}\Big|_{VCO} = \phi_{n,VCO} \cdot |E(s)|^2 \quad (4.39)$$

Thus, if we refer all noise sources in the loop either to the reference or to the VCO output we only need to compute these two transfer functions, $G(s)$ and $E(s)$.

4.2.1 Noise Contributors

The phase detector and charge pump are lumped together to generate one noise source at the charge pump output. The charge pump current will be provided by MOS devices which exhibit both thermal and flicker noise. The total output referred noise current density of a device in the charge pump at frequency f is given by

$$\overline{i_{CP}^2} = 4kT\gamma g_m + \frac{K_f I_{CP}}{C_{ox} L^2 f} \quad (4.40)$$

where I_{CP} , g_m , and L are the charge pump current, effective transconductance, and channel length, respectively, of the transistor, while K_f and C_{ox} are process parameters. This noise current, however, is not present at all times. Recall from the PFD discussion in Section 4.1 that in the locked condition the UP and DN pulses are turned on at the same time for a short period of time, T_{min} , which is much less than T_{ref} .⁶ This performs a sampling action on the charge pump noise current. Generally, the reference frequency for the PLL is higher than the $1/f$ noise corner frequency so $1/f$ noise is oversampled while thermal noise is undersampled. We can define a duty cycle, D_{min} , of the charge pump in the locked condition as

$$D_{min} \triangleq \frac{T_{min}}{T_{REF}} \quad (4.41)$$

which is always $\ll 1$. The flicker noise is directly scaled by D_{min}^2 while the thermal noise is only scaled by D_{min} [Arora05] due to aliasing which causes high frequency thermal noise to

⁶Since the PFD outputs only digital levels, its amplitude noise is ignored. Its contribution, however, comes from generating these timed pulses so the timing jitter should be taken into account for a more accurate noise profile.

fold down into the PLL bandwidth. The effective output noise current of the charge pump is thus given by

$$\frac{\overline{i_{CP,eff}^2}}{\Delta f} = F \left[D_{min} 4kT\gamma g_m + D_{min}^2 \frac{K_f I_{CP}}{C_{ox} L^2 f} \right] \quad (4.42)$$

The noise current density is multiplied by a factor F to account for the total number of devices contributing noise to the output of the charge pump (e.g.: current mirrors). The loop filter only contains one noisy element, the resistor. We refer the single side band (SSB) noise from this element to the output of the loop filter.

$$\begin{aligned} \frac{\overline{v_{LF,eff}^2}}{\Delta f} &= 2kTR \left| \frac{\frac{1}{j\omega C_1}}{\frac{1}{j\omega C_1} + \frac{1}{j\omega C_2} + R} \right|^2 \\ &= 2kTR \left| \frac{C_2/C_1}{1 + C_2/C_1 + j\omega RC_2} \right|^2 \\ &= \frac{2kTR (C_2/C_1)^2}{(1 + C_2/C_1)^2 + (f/f_z)^2} \end{aligned} \quad (4.43)$$

The divider also adds its own phase noise, $\phi_{n,DIV}$. For a known topology this noise may be calculated but is usually just simulated. The divider phase noise is added to its output just as was done for the VCO and reference. From Fig. 4.12 it is clear that this noise is indistinguishable from the phase noise of the reference since these two sources are effectively added together. As we will see later, this has some important implications for mm-wave PLL designs. For simplicity we refer the charge pump noise back to the reference and the loop filter noise forward to the VCO output. The divider noise is already added directly to the reference so it does not need to be moved. Thus, to find the contribution of each noise source to the output noise we multiply the noise sources referred to the reference by $|G_3(s)|^2$ from (4.35), and the noise sources referred to the VCO output by $|E_3(s)|^2$ from (4.36).

$$\phi_{n,out} \Big|_{VCO} = \phi_{n,VCO} \cdot |E_3(j\omega)|^2 \quad (4.44)$$

$$\phi_{n,out} \Big|_{REF} = \phi_{n,REF} \cdot |G_3(j\omega)|^2 \quad (4.45)$$

$$\phi_{n,out} \Big|_{DIV} = \phi_{n,DIV} \cdot |G_3(j\omega)|^2 \quad (4.46)$$

$$\begin{aligned} \phi_{n,out} \Big|_{CP} &= \frac{\overline{i_{CP,eff}^2}}{\Delta f} \cdot \frac{|G_3(j\omega)|^2}{K_{PD}^2 I_{CP}^2} \\ &= \left[4kT\gamma g_m + \frac{2\pi K_f I_{CP} D_{min}}{C_{ox} L^2 \omega} \right] \cdot \frac{F D_{min}}{K_{PD}^2 I_{CP}^2} \cdot |G_3(j\omega)|^2 \end{aligned} \quad (4.47)$$

$$\begin{aligned} \phi_{n,out} \Big|_{LF} &= \frac{\overline{v_{LF,eff}^2}}{\Delta f} \cdot \left| \frac{K_{VCO}}{j\omega} \right|^2 \cdot |E_3(j\omega)|^2 \\ &= \frac{2kTR (C_2/C_1)^2}{(1 + C_2/C_1)^2 + (\omega/\omega_z)^2} \cdot \frac{K_{VCO}^2}{\omega^2} \cdot |E_3(j\omega)|^2 \end{aligned} \quad (4.48)$$

The total output phase noise is then given by

$$\phi_{n,PLL} = \phi_{n,out}\Big|_{REF} + \phi_{n,out}\Big|_{VCO} + \phi_{n,out}\Big|_{DIV} + \phi_{n,out}\Big|_{CP} + \phi_{n,out}\Big|_{LF} \quad (4.49)$$

4.2.2 Design Optimization

The selection of loop bandwidth and other parameters is determined by requirements on phase noise (or jitter) and settling time. A couple of observations can be made directly by examining (4.44)-(4.48). From (3.12) we know that for a constant current density $g_m \propto I_d$ in a MOS device. Thus, from (4.42) we can see that

$$\frac{\overline{i_{CP,eff}^2}}{\Delta f} \propto I_{CP} \quad (4.50)$$

Finally, using this result with (4.47) it is apparent that the output phase noise of the PLL due to the charge pump is inversely proportional to the charge pump current

$$\phi_{n,out}\Big|_{CP} \propto \frac{1}{I_{CP}} \quad (4.51)$$

To reduce its phase noise contribution we would like to use the maximum charge pump current possible. However, to maintain the loop dynamics (i.e.: the same loop bandwidth), C_{sum} must be increased to maintain a constant ratio of I_{CP} to C_{sum} , requiring more die area. This can be seen by examining (4.30) and (4.34) since K_{PD} , K_{VCO} , and N generally cannot be changed at will. Thus, area and power consumption are traded-off for phase noise performance. Furthermore, maintaining the same loop dynamics means the relative location of ω_z and ω_p must also be maintained. Thus, C_1 and C_2 are both scaled at the same rate (maintaining a constant ratio of C_1/C_2) while R is scaled in inverse proportion. From (4.48), the output noise due to the loop filter is proportional to R . Increasing I_{CP} reduces the noise due to the charge pump but requires an increase in C_1 and C_2 and an equal reduction in R thus also reducing the phase noise due to the loop filter.

The selection of loop bandwidth involves a trade-off between noise contribution of the reference versus the VCO⁷ since the former is rejected outside the loop bandwidth and the latter inside. (The loop bandwidth is equal to the cross-over frequency of $A(s)$, called ω_C .) The approach to designing the loop dynamics involves the following steps:

1. Select the ratio of ω_p to ω_z that leads to an acceptable phase margin for stability and acceptable peaking in the loop transfer function (and settling). Note that lower phase margin leads to higher peaking.

⁷assuming these are the dominant noise sources in the loop

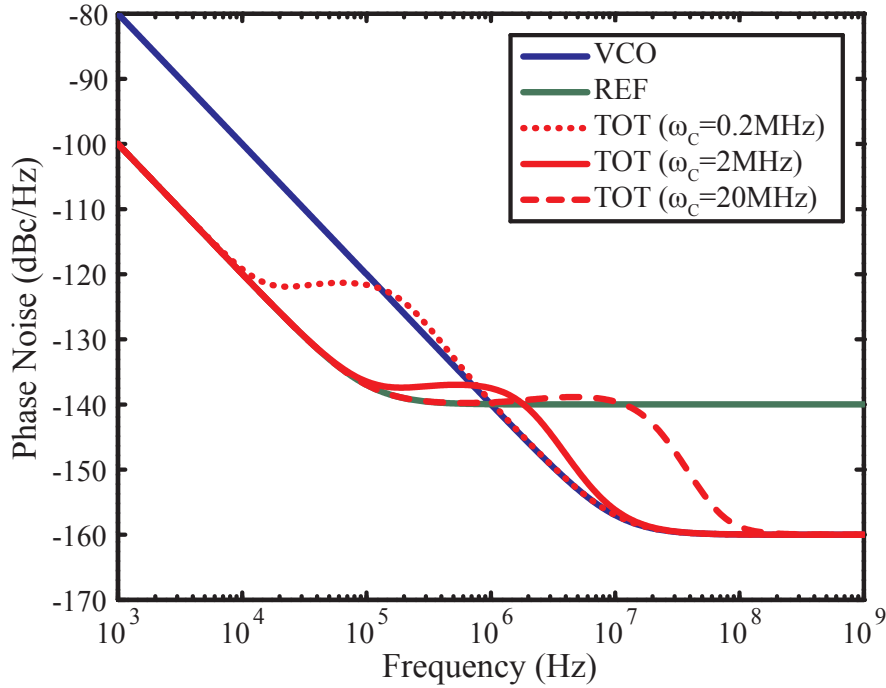


Figure 4.13: Selection of PLL bandwidth based on phase noise.

2. Select the desired loop bandwidth, ω_C , based on phase noise and/or settling constraints. Constraints on phase noise can arise either from a spectral mask based on a standard or in the form of spectral purity metrics such as the ones presented in Appendix 4.A.
3. Using (4.32) and (4.33) select R , C_1 , and C_2 to set ω_z and ω_p based on the ratio selected in step 1 and such that ω_C is at their geometric mean.

$$\omega_C = \sqrt{\omega_z \omega_p} \quad (4.52)$$

4. Using (4.30) and (4.34) select I_{CP} to set ω_C to the value chosen in step 2.
5. Scale I_{CP} and the filter components (maintaining constant ω_C , ω_p , and ω_z) to reduce the noise of the charge pump and loop filter until they are no longer dominant.

Generally, designs are constrained to a maximum power consumption and die area allowance. Thus, step 5 should continue until either the charge pump and loop filter noise contributions are no longer dominant, or the power consumption or area constraints are violated. In the case where a design is either power or area limited, the selection of loop bandwidth should be revisited and the design process (steps 2 through 5) repeated since the assumption that the reference and VCO phase noise are dominant is no longer true.

Fig. 4.13 shows a representative example of selecting the proper loop bandwidth for minimum phase noise in a Third order Type-II charge pump PLL. A common rule of thumb is to set

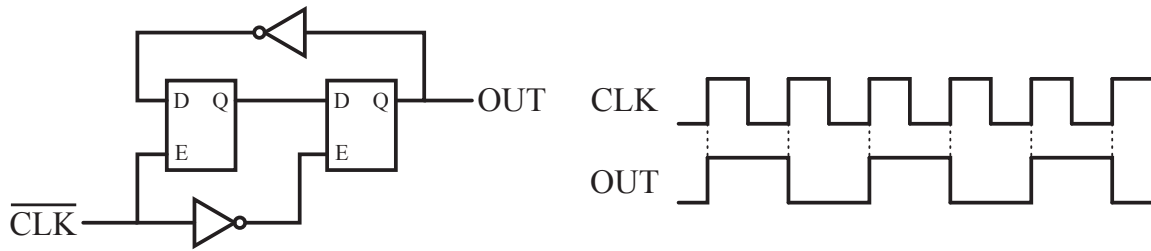


Figure 4.14: Master-slave flip-flop divide by 2.

ω_C approximately at the frequency where the reference and VCO phase noise profiles cross and Fig. 4.13 shows this to be a good approximation.

4.3 Frequency Dividers

Frequency divider topologies vary depending on the frequency of operation and the required divide ratio. We will begin by looking at simple circuits that divide by a factor of 2, and examining the limitations on operating frequency before briefly touching on more complex prescaler architectures.

4.3.1 Flip-Flop Dividers

The most common frequency divider, and the one on which many other architectures are based, is the master-slave flip-flop divide by 2 [Razavi98, pg. 290] shown in Fig. 4.14. A positive latch is transparent when its enable input is high, allowing the output to change with the input. When its enable input is low the positive latch is opaque and holds its output constant regardless of changes in the input. Inverting the polarity of the signal fed to the latch enable creates a negative latch which is transparent when the enable signal is low and opaque when the enable signal is high.

A master-slave flip-flop is made up of two latches connected in series and enabled by clocks of opposite polarity, CLK and \overline{CLK} . The master latch is transparent for half of the clock cycle, while the slave latch is transparent for the other half. A flip-flop is thus never fully transparent. The input is sampled by the master latch while the slave latch is holding the output steady. At the clock edge, the master latch begins holding the input present just before the clock edge, while the slave latch becomes transparent, presenting that signal to the output. A flip-flop is an edge triggered device that samples the input only at the clock edge. In a positive edge triggered flip-flop the master and slave latches are controlled by \overline{CLK} and CLK respectively while in a negative edge triggered flip-flop the clocks are swapped.

To make a divider, the inverted output of a flip-flop is fed back to its input. This causes the output to switch polarity on every rising edge of CLK (for a positive edge triggered

flip-flop). One period of the output is then equal to two periods of CLK achieving divide by 2 operation. This simple concept can be extended using extra logic circuits to create more complicated dividers which divide the input clock by any integer, or even half integers (e.g.: 1.5), by simply counting edges of the input clock.

Thus far we have discussed the latch and flip-flop only in general terms without specifying the underlying circuitry. A latch can be constructed using CMOS as shown in Fig. 4.15a. The power consumption of a CMOS based flip-flop divider, given by (4.53), is determined solely by the total capacitance of the circuit, the switching frequency, and the supply voltage.⁸

$$P_{dc} = C_{tot} V_{dd}^2 \frac{f_{in}}{2} \quad (4.53)$$

The maximum frequency of operation is determined by the delay of the latches which make up the flip-flop since the output of each latch must settle within half a period of the input clock.

$$f_{in} \leq \frac{1}{2t_{d,latch}} \quad (4.54)$$

Unfortunately, the CMOS flip-flop contains a lot of logic gates and the maximum frequency over PVT variations is limited to only a few GHz by this gate delay. Scaling CMOS processes into the deep submicron region reduces this gate delay and continues to improve the maximum speed of operation for CMOS based dividers.

An alternative logic family which reduces the total circuit capacitance and gate delay is True-Single-Phase-Clock (TSPC) logic first presented in [Ji-Ren87] and expanded in [Navarro Soares99]. A flip-flop in this logic family can be made with a minimum of only 9 transistors as shown in Fig. 4.15b. This is a dynamic logic family which stores values on parasitic capacitors rather than always driving nodes either to V_{dd} or GND through switches as in CMOS logic. Leakage currents discharge these parasitic capacitors over time so they must be refreshed periodically in order to maintain correct values. This sets a minimum frequency of operation for this type of logic. However, the reduced parasitics and logic depth significantly increase the maximum frequency of operation when compared to CMOS. Another advantage of TSPC logic is its reduced power consumption at a given frequency, also given by (4.53), due to the significantly reduced parasitics when compared to CMOS. Dividers of this type in a 65nm LP CMOS process have been shown to work up to 15GHz while consuming only $20\mu W/GHz$ [Deng10]. However, taking into account PVT variations, TSPC based designs are limited to approximately 10GHz in practice. Nevertheless, similar to CMOS, scaling into the deep submicron region continues to improve both speed and power consumption for these types of dividers.

For frequencies beyond 10GHz, Current-Mode Logic (CML) is required. A CML flip-flop is shown in Fig. 4.15c. This type of logic uses differential pairs to steer a bias current either to

⁸This ignores leakage current which, in high frequency operation, is a very small component of the total power consumption.

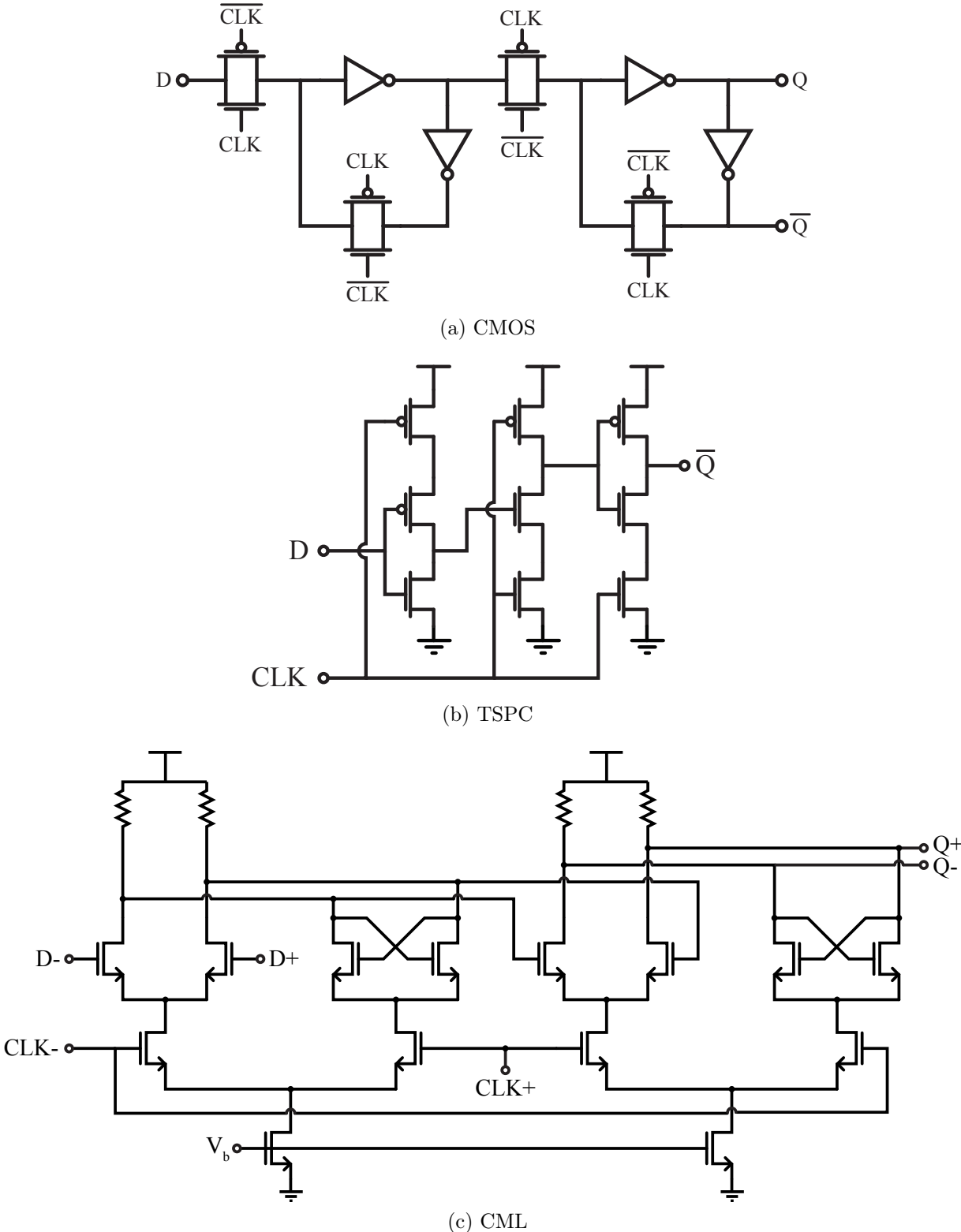


Figure 4.15: Common edge-triggered flip-flop topologies.

the positive or negative output to represent a logic '1' or '0' respectively. The load resistors and bias current set the maximum signal swing at the output,

$$V_{sw,max} = I_b R_L \quad (4.55)$$

which is generally much less than the full rail signal swing of digital logic. The reduced signal swing helps to speed up signal transition times. Unlike the logic styles presented above, for a given design the power consumption of CML logic is constant regardless of the frequency of operation and set by the bias currents of each stage. In order for this divider to operate properly it must be regenerative. In other words, the gain through the loop formed by the two latches must be greater than 1, similar to the startup condition of an oscillator.

$$(g_m R_L)^2 > 1 \quad (4.56)$$

The maximum frequency of operation of a CML divider is limited by two factors. The load resistance and the total capacitance present at the output of each stage (explicit and parasitic capacitance) creates a low pass filter

$$\omega_{3dB} = \frac{1}{R_L (C_L + C_{par})} \quad (4.57)$$

which eventually reduces the loop gain below the level required to sustain a signal. At the same time, the limited bias current can only charge and discharge a capacitor at a limited rate, called the slew rate, also limiting the maximum frequency of operation for a given minimum output voltage swing, $V_{sw,min}$. In a slew rate limited design, the output is approximately a triangular wave as the load capacitors are charged and discharged by a constant current equal to I_b .

$$f_{in} \leq \frac{I_b}{(C_L + C_{par}) V_{sw,min}} \quad (4.58)$$

The load resistance and bias current must thus be designed such that neither the low pass filter nor the slew rate limit the desired frequency of operation.

To maintain a constant signal swing, the bias current and load resistance can be scaled in opposite directions. Higher bias currents can be used to increase the slew rate for a given load capacitance. Since the resistance must be scaled down, the low pass filter pole will also move up in frequency. However, increasing bias currents means increasing device sizes and thus parasitic capacitance. Eventually, the parasitic capacitance dominates the total load capacitance and further increases in bias current will not improve the slew rate. At this point, further scaling will change the total capacitance and load resistance in proportion to each other and in opposite directions so the low pass filter pole will also remain unchanged. This condition thus sets the upper limit on the maximum frequency of operation for a CML flip-flop divider which, depending on the exact design, will either be limited by the slew rate, the low pass filter pole, or both. Scaling also helps in CML logic by reducing the parasitic capacitance relative to the driving strength of a transistor. The maximum

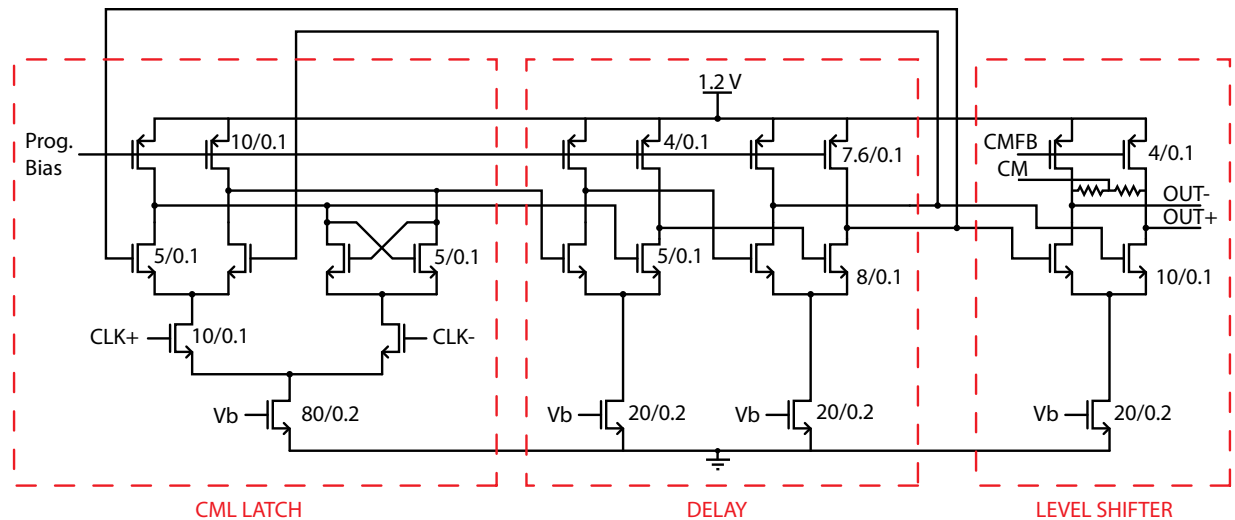


Figure 4.16: CML pulsed-latch divider.

frequency of operation of a standard CML flip-flop divider over PVT variations is limited to approximately 30GHz. Further improvements in speed can be achieved at the expense of die area by using inductive peaking techniques or altogether changing the load resistance to a load inductance which tunes out the capacitive parasitics. This latter option makes the divider a narrowband circuit and thus creates a minimum frequency of operation at which the loop gain drops below 1.

In order to speed up the maximum frequency of operation of a CML flip-flop even further we must reduce the delay around the loop. This can be accomplished by replacing the slave latch with a buffer as shown by [Kim05]. This creates a pulsed latch whose operation at high frequencies is the same as that of a flip-flop. When the master latch becomes transparent, the signal must travel through the master latch and return to its input only after it has become opaque. The slave latch is thus unnecessary if a buffer can provide enough delay to return the signal back to the input of the master latch after it has become opaque. Since a buffer will have less parasitic capacitance than a latch, the delay around the loop is reduced and the maximum frequency of operation is increased. If the frequency of operation is too low, however, the signal will be fed back to the input of the master latch while it is still transparent, creating a race-through condition. In this situation the pulsed-latch no longer operates as a flip-flop and the divider ceases to work. The minimum frequency of operation is thus set by the race-through delay. Another advantage of this topology at high frequencies is that the input clock must only drive one latch rather than two, reducing the loading on the input clock.

A sample schematic of a CML pulsed-latch divider in 90nm CMOS is shown in Fig. 4.16 [Marcu09]. Two buffer stages are used to increase the loop gain for proper operation at high frequencies. A buffer with common-mode feedback is included outside the loop to provide gain and level shifting in order to drive the succeeding division stage with a large enough

signal at the correct common-mode level. The load resistors in this design were implemented using triode region PMOS transistors, at the expense of increased capacitive loading, to allow tuning of the divider performance over PVT variations to ensure divider lock. Despite the increased capacitive loading, the divider was designed to achieve lock for input frequencies up to 35GHz over all PVT variations without any inductive peaking or tuning, saving die area. The output buffer reduces the loading of this divider stage and also provides gain and common-mode level shifting down to the level required by the next divider stage. The common-mode output of the buffer, CM , is sensed by a large resistor and set to the desired value by a common-mode feedback loop using a low-power and low-speed OTA (not shown) which controls the PMOS load bias voltage, $CMFB$. This divider, including output buffer, consumes a total of 8mA from a 1.2V supply.

4.3.2 Injection Locked Dividers

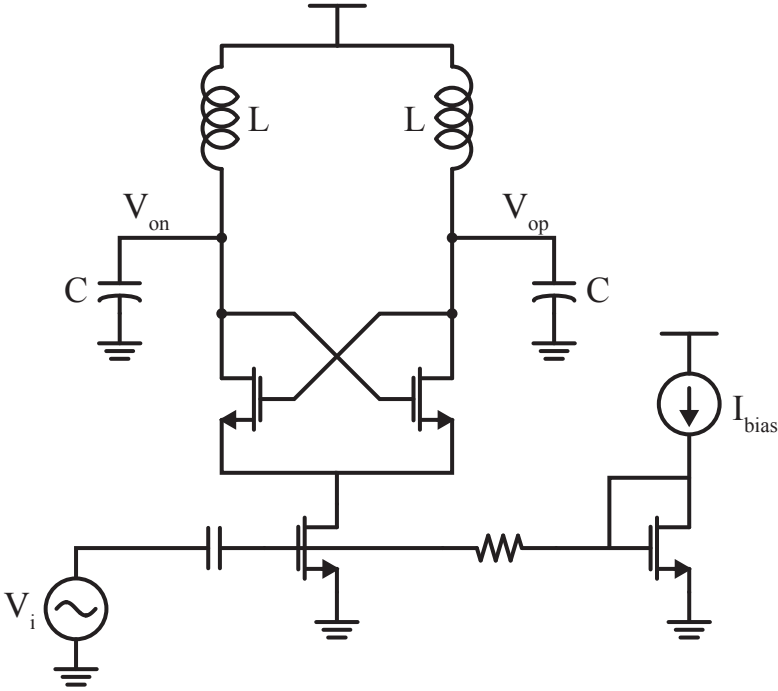
If a signal is injected into a free running oscillator with sufficient strength at a frequency near the oscillator's free-running frequency, the oscillator frequency and phase will lock to that of the injected signal. This effect is called injection locking. [Adler46] first derived the conditions necessary in order for locking to occur as

$$\frac{S_{inj}}{S_{osc}} > 2Q \left| \frac{\Delta\omega_o}{\omega_o} \right| \quad (4.59)$$

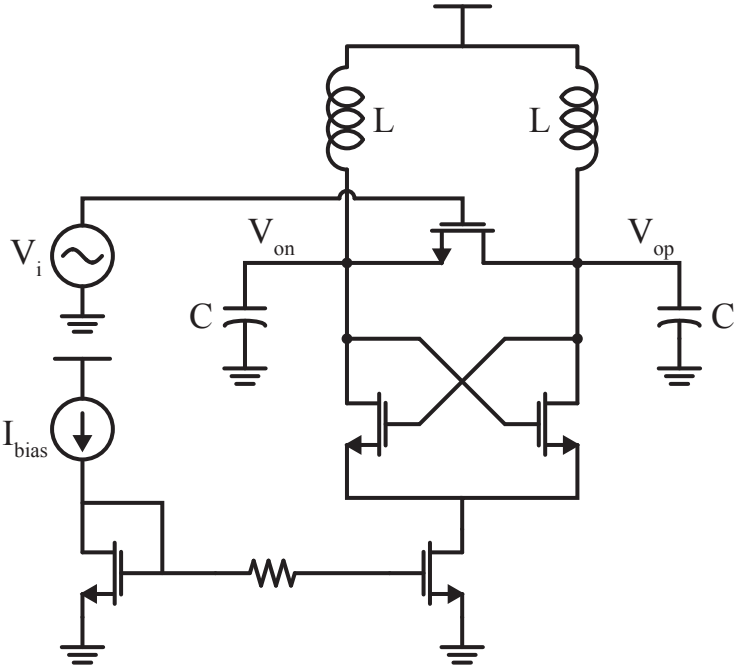
where S_{inj} and S_{osc} are the injected and oscillator signals respectively, Q is the oscillator's tank quality factor, ω_o is the oscillator's free running frequency, and $\Delta\omega_o$ is the offset frequency between the injected signal and ω_o . The locking range is defined as the largest $\Delta\omega_o$ which will still allow the oscillator to lock to the injected signal and is a function of the injected signal's strength as given by (4.59). We have used S_x to signify a general signal here which can be either a voltage or a current depending on the particular implementation. From (4.59) we can then see that for a given normalized injection signal strength, a higher Q leads to reduced locking range.

Within the locking range, the locked oscillator will track the injected signal's frequency and phase variations. Thus, the phase noise of the injection locked oscillator is equal to the phase noise of the injected signal at offset frequencies within the locking range and reverts to the free running oscillator's phase noise outside this range. This can be used to great advantage if the injected signal has lower phase noise than the free running oscillator. In other words, the power consumption of an injection locked oscillator can be reduced at the expense of its free running phase noise since the phase noise will be determined by the injected signal.

In Chapter 3 we discussed push-push oscillators in which a signal at $2\omega_o$ is extracted from an oscillator operating at ω_o . In fact, this process can be reversed using the principle of injection locking to form a divider. Injection locked frequency dividers are injection locked oscillators where the injection signal is at a harmonic frequency. One way to accomplish



(a) Tail-injection



(b) Direct-injection

Figure 4.17: Injection locked dividers.

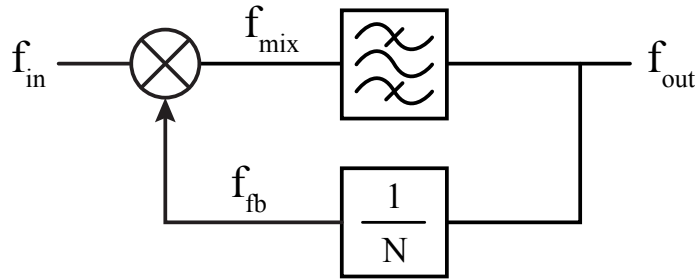


Figure 4.18: Regenerative (Miller) divider.

injection locking is to inject the harmonic signal at a node in the circuit which is itself generating that harmonic. For example, a signal at $2\omega_o$ can be injected into the common source node of a cross-coupled differential pair oscillator operating at ω_o as shown in Fig. 4.17a [Rategh99]. The major disadvantage of this topology is that the tail biasing transistor could be large and thus present a significant capacitive load to the driving oscillator. This can be remedied by using two parallel transistors if necessary, one providing the bias current, the other the injection current.

Another option for injection locking comes from the fact that the differential signal across the tank goes to zero twice each period. To perform injection locking at ω_o we can place a switch across the tank as shown in Fig. 4.17b driven by a signal at $2\omega_o$. A complementary PMOS/NMOS pair can also be used, driven by differential signals at $2\omega_o$. The direct injection topology generally achieves higher locking range at the expense of extra loading on the divider tank due to the injection transistors.

Injection locked dividers can be made to operate up to any frequency that an oscillator can be designed. However, their locking range is generally limited to less than 5% without tuning. This aspect will be addressed further in Section 4.4.

4.3.3 Regenerative Dividers

The last fundamental type of divider we will cover is the regenerative divider first proposed by [Miller39] and thus sometimes referred to simply as the “Miller” divider (Fig. 4.18). This type of divider is based on a mixer whose output is fed back to one of its inputs. Ignoring the $1/N$ block for now (i.e.: let $N = 1$), the mixer output is equal to

$$f_{mix} = f_{in} \pm f_{out} \quad (4.60)$$

A low pass (or band-pass) filter is added in the loop to select only the lower sideband for feedback.

$$\begin{aligned} f_{out} &= f_{in} - f_{out} \\ f_{out} &= \frac{f_{in}}{2} \end{aligned} \quad (4.61)$$

Thus, the only possible solution for the loop is if the feedback frequency is exactly equal to one half of the input frequency. If the upper sideband is selected by the filter, the loop has no solution. In general, we can also add a divider in the feedback loop such that $f_{fb} = f_{out}/N$. In that case, the filter must be a band-pass filter and select either the upper or lower sideband of the mixer output. The output frequency is then given by

$$\begin{aligned} f_{out} &= f_{in} \pm \frac{f_{out}}{N} \\ f_{out} \left(1 \pm \frac{1}{N}\right) &= f_{in} \\ f_{out} &= \frac{N}{N \pm 1} f_{in} \end{aligned} \quad (4.62)$$

where the denominator of the fraction is equal to $N + 1$ if the filter selects the lower sideband and $N - 1$ if the filter selects the upper sideband. Notice that selecting the upper sideband leads to $f_{out} > f_{in}$, a regenerative multiplier instead of a divider. Also, as N increases, the upper and lower sidebands (f_{mix}^+ and f_{mix}^-) converge to the same frequency, making it harder for the filter to select one while rejecting the other. With insufficient rejection of the unwanted sideband, the output either becomes distorted and may be unusable, or the divider/multiplier simply does not lock to the input.

The Rategh-type of injection locked divider (Fig. 4.17a) can also be seen as a Miller divider. The differential pair and input transistor form an active mixer with the RF port at the g_m transistor gate and the LO port at the differential pair gates. The drain current will then have frequency components at both $\omega_{in}/2$ and $3\omega_{in}/2$ and the output tank selects only the lower harmonic, converting it to a voltage which is then fed back to the LO port of the mixer. However, we can also provide the input to the differential pair transistors and feed back the output to the g_m transistor to create a pure regenerative divider. In this configuration the divider will not self-oscillate without an input present, unlike the cross-coupled topology, which, according to [Lee04a], should lead to reduced phase noise. Furthermore, [Lee04a] also shows that the locking range of the Miller divider can be made larger than a purely injection locked divider with equal power consumption.

Miller dividers, however, require large signal swings at the input to achieve good conversion gain from the mixer. Also, the mixer devices are made large for the same reason. Thus, a Miller divider presents a larger load to the preceding stage and also requires larger input signals for proper operation when compared with injection locked dividers. Despite the fact that the two topologies are very similar, the injection locked divider is used in practice more often due to these limitations of the Miller divider.

4.3.4 Prescalers

So far we have only discussed dividers with a fixed division ratio. However, in order to tune the output frequency of the PLL for different channels, the total division ratio of the divider

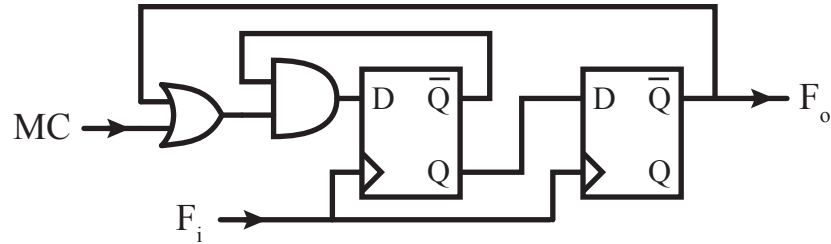


Figure 4.19: 2/3 prescaler.

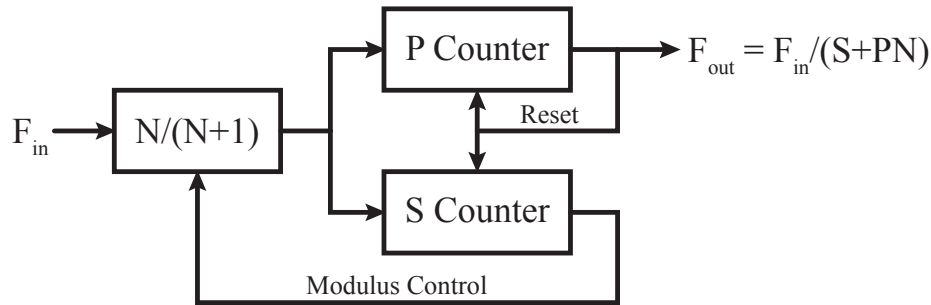


Figure 4.20: Program/Swallow Counter.

chain must be variable. Cascaded multi-modulus dividers are used to generate prescalers with any integer division ratio. A 2/3 prescaler based on a flip-flop divide by two circuit is shown in Fig. 4.19. The mode control signal, MC , controls the division ratio. When MC is set low, the circuit divides by 3. When MC is set high, the feedback path from the second flip-flop is disabled and the circuit divides by 2. This principle can be extended using more flip-flops and logic to create any $2^n / (2^n + 1)$ prescaler but the input clock must drive all the flip-flops leading to high clock load.

To build a prescaler with programmable division ratio we can use one of two topologies. The first is called the Program/Swallow Counter [Razavi98, pg. 270], shown in Fig. 4.20, which uses a multi-modulus $N/(N+1)$ divider and two counters. The prescaler operates as follows. The program counter (P) can be programmed to any value while the swallow counter (S) is programmable from 0 to P. Both counters are programmed with an initial value and count down to zero as they are both clocked by the output of the multi-modulus divider which is initially dividing by $N+1$. When the S counter reaches zero, it changes the modulus of the multi-modulus divider to N . The P counter continues to count down until it reaches zero as well. At this point the P counter provides the output signal and resets all three blocks to begin the next cycle. This means the output signal is sent after $S(N+1)$ input cycles plus $(P-S)N$ input cycles for a total division ratio, N_{tot} of

$$\begin{aligned}
 N_{tot} &= S(N+1) + (P-S)N \\
 &= SN + S + PN - SN \\
 &= S + PN
 \end{aligned} \tag{4.63}$$

If a small value of N is used, the counters must operate at very high speed making the design complicated and power hungry. On the other hand, if a large value of N is used, the multi-modulus divider itself will be power hungry and provide a large load to the input since many flip-flops would all be clocked by the input. Furthermore, the accumulated jitter in a long counter must be removed by retiming its output using a flip-flop clocked by, for example, F_{in} . A lower frequency may need to be used for practical purposes at the expense of reduced jitter performance on the final output.

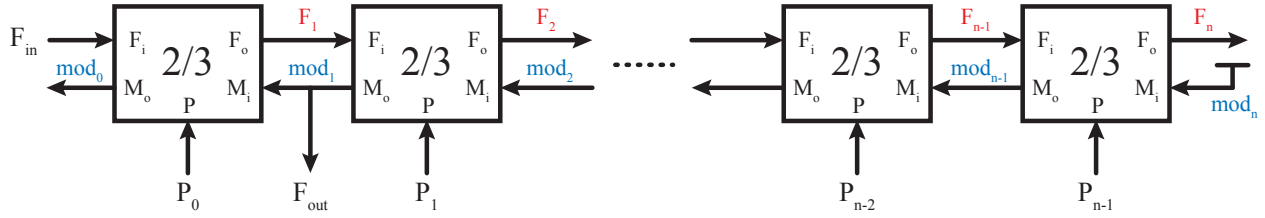
A modular prescaler topology introduced by [Vaucher00b] alleviates both of these problems. This topology, shown in Fig. 4.21a, uses modular divide by 2/3 stages with local connections between adjacent stages only. Since there are no global connections, this topology is much easier to design and layout even for high frequency operation. Each divide by 2/3 stage has three inputs (F_i , P_i and M_i) and two outputs (F_o and M_o). F_i and F_o are the main clock inputs and outputs of the cell, respectively. M_i and M_o are the pulsed mod signal inputs and outputs, respectively. The P_i input is the divide modulus setting. One possible implementations of the divide by 2/3 stage is shown in Fig. 4.21b. The upper half of the cell is simply a programmable divide by 2/3 block whose control signal is provided by the lower half, also called the End-Of-Cycle logic. The cell normally operates in the divide by 2 mode, flipping its output, F_o , for every other edge of F_i . When the mod signal, M_i , arrives, if P_i is asserted high, the cell swallows the next pulse of the input without changing its output, performing a division by 3. If instead P_i is low, the cell continues dividing by 2. In either case it then asserts its own M_o high for one period of its input clock, F_i . The last cell in the chain has its mod input, M_i , permanently asserted high. For a chain of n cells with the input clock having a period T_{in} , the final output clock has a period equal to

$$\begin{aligned} T_{out} &= T_{in}P_0 + 2T_{in}P_1 + 2^2T_{in}P_2 + \dots + 2^{n-2}T_{in}P_{n-2} + 2^{n-1}T_{in}P_{n-1} + 2^nT_{in} \\ &= (P_0 + 2P_1 + 2^2P_2 + \dots + 2^{n-2}P_{n-2} + 2^{n-1}P_{n-1} + 2^n) T_{in} \end{aligned} \quad (4.64)$$

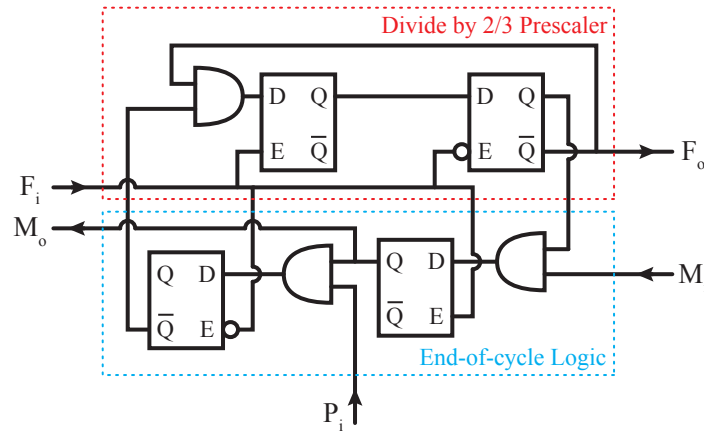
The programmable divide ratio can thus be set to any integer from 2^n (for all $P_i = 0$) to $2^{n+1} - 1$ (for all $P_i = 1$). To extend the maximum possible division ratio, a counter driven by the final clock output can be used to provide the mod signal of the final stage [Vaucher00a].

The timing diagram for a 4 stage Vaucher prescaler using the unit cell from Fig. 4.21b is shown in Fig. 4.21c. All P_i are initially set low leading to the lowest divide ratio (shortest T_{out}). Then, all P_i are set high leading to the highest divide ratio (longest T_{out}). The operation in either case proceeds as follows. As the *mod* signal progresses backwards down the chain, the individual values of P_i are locked in. After mod_0 is set, each individual block whose P signal was high swallows the next input pulse it receives before resuming normal operation as a divide by 2. Each individual block whose P signal was low simply operates as a divide by 2. When F_4 , the final stage output, has changed state twice, the cycle is complete and the next cycle begins. Notice that the control signals, P_i , must be ready before the first *mod* signal goes high. The easiest way to ensure this is to lock in all P_i using the falling edge mod_1 .

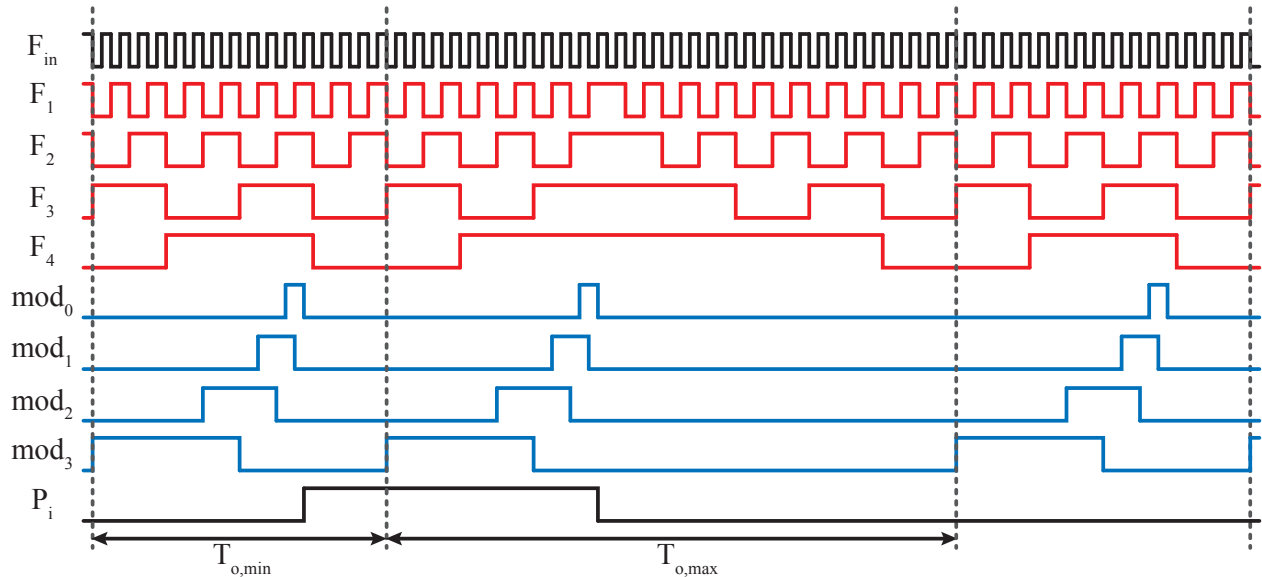
Another advantage of this architecture is the built-in retiming of the final divided clock.



(a) Block diagram.



(b) Latch based unit cell schematic.



(c) Timing diagram.

Figure 4.21: Vaucher modular prescaler.

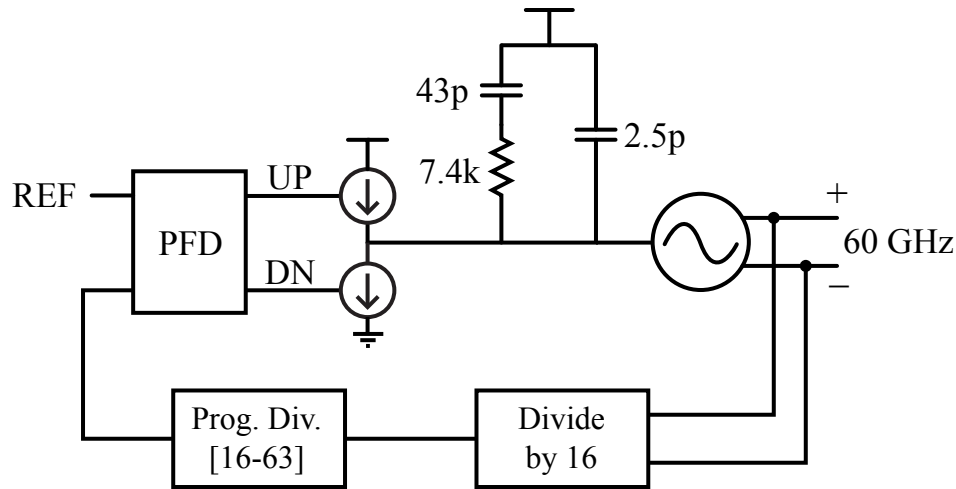


Figure 4.22: Integer-N PLL block diagram.

From Fig. 4.21a we can see that mod_1 , the final output of the prescaler, is generated directly by F_1 , the output clock of the first stage. The only way to improve on this would be to use mod_0 which is generated directly by F_{in} , giving the best possible jitter performance. This pulse has a width equal to one period of F_{in} so, in practice, the pulse width of mod_0 (or even mod_1 for that matter) may be too small for the phase detector to operate correctly. Instead, one of the wider mod pulses may be used at the expense of reduced jitter performance.

4.4 Sample Design

A low power 60GHz PLL was designed as part of a 4-element, baseband phase shifting, phased array transceiver in 65nm CMOS [Tabesh11a]. The transceiver was designed to utilize as much of the unlicensed 60GHz band as possible to transmit data at high rates over a distance of a few meters. The PLL provides the LO signal for both the TX and RX. It is a fully integrated charge pump based PLL that was optimized for minimum integrated output phase noise. The core of this 3rd order type-II integer-N PLL (Fig. 4.22) consists of a fundamental mode VCO, Fig. 3.31, (presented in Section 3.6.2) with directly coupled buffers to drive the LO distribution chain. The VCO was chosen to be a fundamental mode design since it provides significantly higher output power than a push-push or higher order harmonic multiplier [Wu09], easing the gain requirements in the distribution chain. Since the LO must be distributed to many elements, splitting losses will reduce the LO power, and using a high output power VCO significantly reduces the overall power consumption in a phased array transceiver. This can only be achieved by the design of a low power 60GHz VCO and 60GHz divider. As demonstrated next, the overall power consumption of this PLL is lower than the power consumed by a harmonic design.

The use of a fundamental mode oscillator requires a low power 60GHz divider, making an

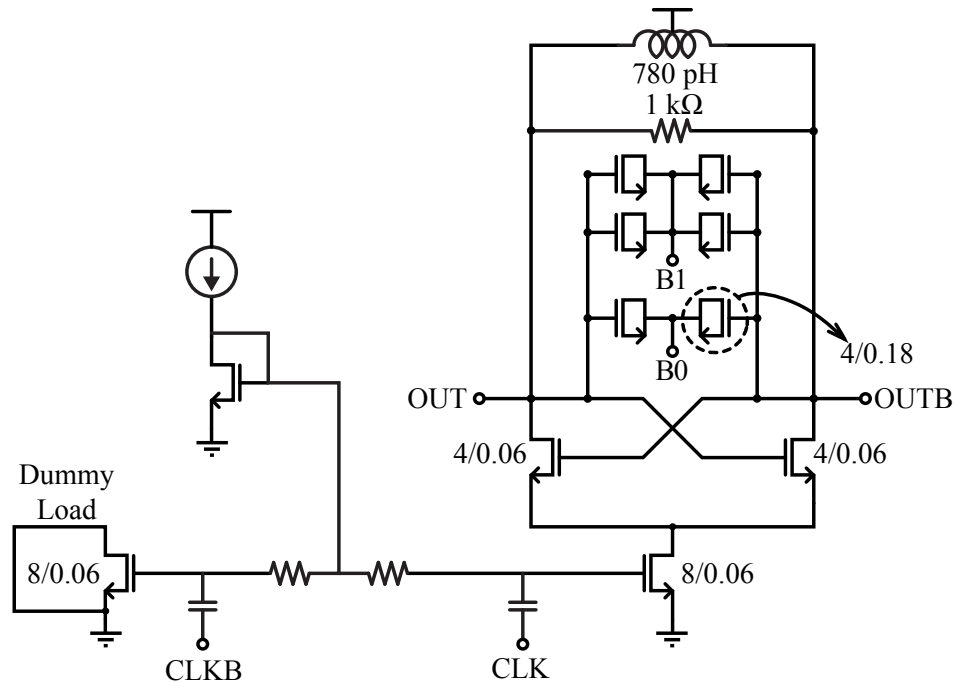


Figure 4.23: Schematic of injection locked divider.

injection locked design desirable [Rategh99]. Since the phase noise is set by the injected signal, coverage of 3 VCO bands can be ensured by adding a parallel resistor R_L to de-Q the tank (Fig. 4.23) without affecting the output phase noise. The total locking range is extended to cover the rest of the VCO's tuning range by using a 2-bit switched varactor bank (which could be tied to the MSBs of the VCO control bits), also maintaining 50% overlap between divider bands. Thanks to Ling kai Kong for the injection locked divider design used in this PLL and shown in Fig. 4.23. A dummy load is included to provide a fully balanced environment for the VCO.

The rest of the divider chain includes two current-mode-logic (CML) master-slave dividers and a TSPC flip-flop divider followed by a 16 to 63 programmable divider based on the modular Vaucher design [Vaucher98]. The total divide ratio can thus be set from 256 to 1008 in steps of 16. The injection locked divider consumes 3.6mW from the 1.2V global supply while the rest of the dividers consume 5mW.

The differential VCO outputs each drive a single-ended distribution network for the RX and TX portions of the phased array, respectively, through independent single-ended LO buffers. To maintain high impedance in the VCO tank for low power consumption, both the first divider and the LO buffers are capacitively coupled to the VCO core. The buffers – whose outputs are matched to the distribution network impedance with transmission lines (for compactness) – also provide gain and isolation from the distribution network. Each buffer consumes 3.6mW from the 1.2V global supply.

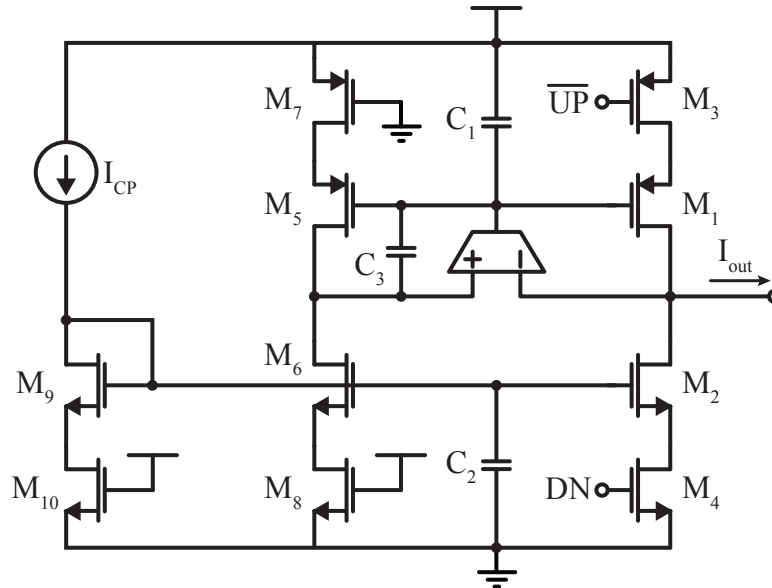


Figure 4.24: Simplified charge pump schematic.

The phase comparison path consists of a flip-flop based PFD [Brown71] followed by a charge pump and an on-chip 2^{nd} order loop filter. The charge pump is based on the design presented in [Temporiti04]. The charge pump consists of two current sources (M_1 and M_2) which are switched on by the \overline{UP} and DN signals from the PFD. One of the current sources (in this case the NMOS) is biased using a current mirror ($M_9 - M_{10}$). The other is biased using a replica bias path ($M_5 - M_8$) and a feedback OTA. The negative input of the OTA is directly connected to the charge pump output, forcing the drain of M_5 (M_6) to be the same as the drain of M_1 (M_2) through negative feedback to the PMOS current source gate. In order to achieve rail-to-rail input common mode range and high DC gain for the OTA a dual-input pair folded-cascode topology is utilized.

Capacitors C_1 and C_2 are bypass capacitors for the gate bias voltages while C_3 is used to stabilize the feedback path. With the OTA feedback configured as shown however, there is also a positive feedback path through the charge pump output which must be taken into account to ensure overall stability. In this design, the large loop filter at the output of the charge pump sufficiently reduces the gain of the positive feedback loop allowing the negative feedback to dominate. For a more robust design, the negative terminal of the OTA should be connected to a voltage that is derived from the charge pump output but low pass filtered to reduce the effect of the positive feedback path as much as possible. Such a design is easily achievable by connecting the negative input of the OTA to the midpoint of the RC leg of the loop filter. This voltage is exactly a low pass filtered version of the charge pump output.

The selection of charge pump current and loop filter components was optimized to minimize integrated output phase noise for the PLL as described in Section 4.2. The loop filter values are fixed as shown in Fig. 4.22. The nominal charge pump current is 1mA but is

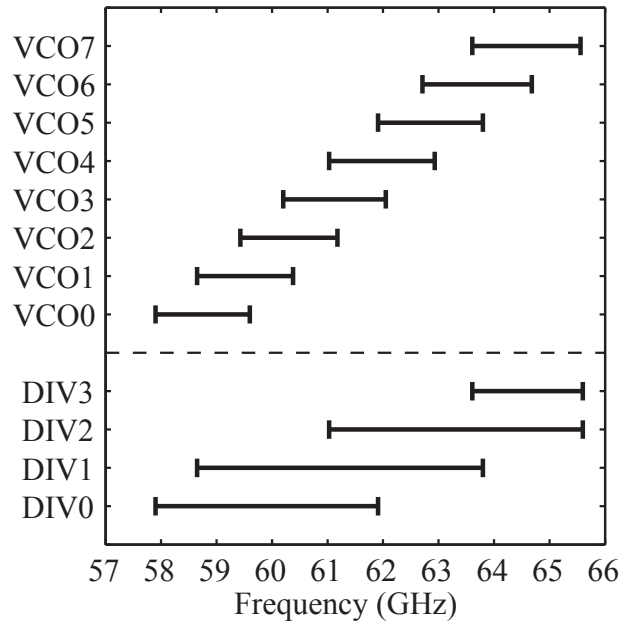


Figure 4.25: Measured VCO and injection locked divider tuning range. (Measurement of divider tuning range limited by VCO.)

programmable from $250\mu A$ to $2mA$ to allow tuning of the loop bandwidth.

The PLL output was directly probed on-chip to allow measurement and characterization. The measured VCO tuning range over all 8 tuning bands is 57.9-65.6GHz (Fig. 4.25), equivalent to a tuning range of 12.4%. The VCO has a measured free running phase noise of $-112dBc/Hz$ at 10MHz offset. The measured output power from one LO buffer is $-1.8dBm$, for a total differential power of $+1.2dBm$. The divider locking range can be roughly measured by varying the VCO frequency until the divider loses lock.

The lower part of Fig. 4.25 also shows the locking range of each of the 4 divider bands. The lower side of divider band 0 and the upper sides divider bands 2 and 3 extend beyond the range of the VCO and cannot be measured. However, it is apparent that the divider can lock over a much wider range than the VCO can tune and that each VCO band is completely contained within at least one divider band. This ensures that the divider chain can never lose lock while the PLL is trying to achieve lock as long as the correct divider band is selected.

The locked PLL spectrum (Fig. 4.26) and phase noise (Fig. 4.27) were measured using an external downconversion mixer to bring the LO in range of an Agilent E4440A Spectrum Analyzer. Since the measured response for both the TX and RX of the phased array transceiver this PLL was a part of is centered around 61GHz [Tabesh11a], the PLL reference for these tests was 119MHz with the programmable divider set to 32 in order to output an LO at 61GHz. For compliance to the IEEE 802.15.3c 60GHz single carrier standard PHY [IEE09], 135MHz would be used for the reference to allow tuning of the four standard

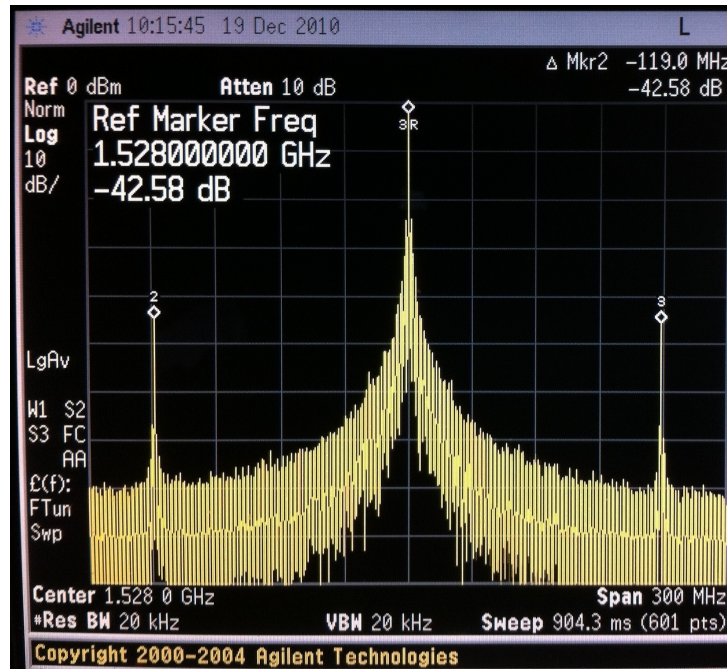


Figure 4.26: Spectrum of locked PLL at 61GHz downconverted with external mixer to allow measurement with Agilent E4440A Spectrum Analyzer. Reference spurs are less than -40dBc .

channels (58.32GHz, 60.48GHz, 62.64GHz, 64.80GHz) and setting the programmable divider appropriately (27, 28, 29, 30); all well within the capabilities of this system.

Measurements of the locked PLL show reference spurs to be lower than -40dBc (Fig. 4.26) while the best measured PLL phase noise (Fig. 4.27) is -82dBc/Hz in-band and -107dBc/Hz at 10MHz offset which was achieved for the lowest charge pump current setting. Due to incorrect sizing of the final divider stage and the buffer at the output of the divider chain, the divider chain output noise is dominated by these two stages and is much higher than expected. Since divider noise is indistinguishable from reference noise, this effect appears as excess reference noise in measurements of the PLL output noise. Thus, the lowest charge pump current, leading to the lowest loop bandwidth, shows the best phase noise performance since it reduces the effect of reference and divider noise. In simulation with an appropriately sized divider and buffer (which would only have consumed an additional 1mW of power), the in-band PLL output noise is approximately 12dB lower. Fig. 4.28 shows the effect of such a fix.

As manufactured, the divider noise dominates the PLL phase noise out to very large offsets and leads to a total integrated phase noise (from 100Hz to 100MHz) of -12.13dB . With the resized divider and buffer, system simulations show the integrated phase noise drops to -22.49dB . This level is sufficient for low BER communication but should be improved in future work to provide more margin for the system as a whole. Simulation shows the noise of

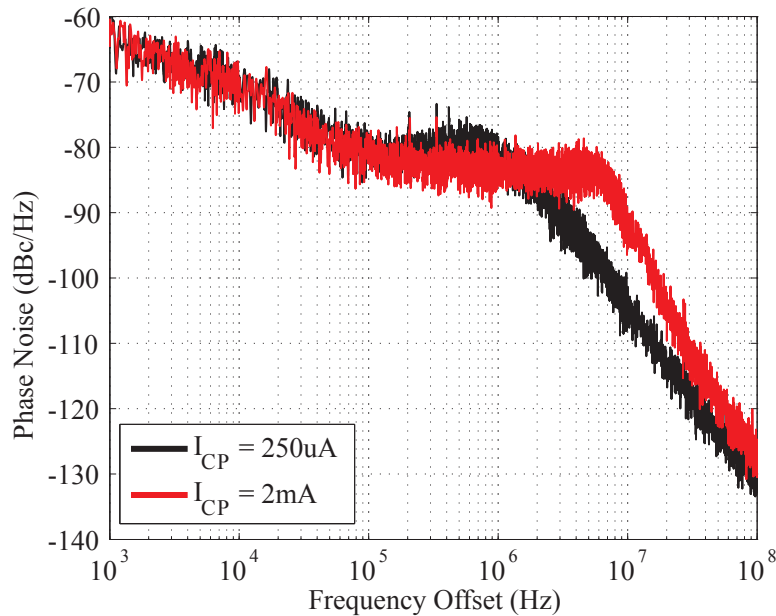


Figure 4.27: Measured phase noise at 61GHz for two loop bandwidth settings, measured using Agilent E4440A Spectrum Analyzer

the first, injection locked, stage of the divider chain to be sufficiently low even in the current design, 20dB below the total for the entire divider chain (as manufactured, without resizing). Thus, further improvements in divider noise would require a redesign of the programmable divider as well as the final stages of the fixed divider to decrease their noise contributions.

The entire PLL, including LO buffers, consumes only 29mW while achieving high tuning range and similar noise performance to previously reported low-power PLLs (Table 4.1). The power consumption is the lowest reported to date among PLLs designed for the 60GHz band, while providing the highest output power and maintaining competitive phase noise performance and low spur levels. An excellent design with very wide tuning range is presented by [Murphy10] which achieves very wide tuning range without sacrificing phase noise performance. However, the power consumption is almost 2.5 times that of the design presented here. Thus, the overall results show the design presented here to be a very competitive solution for 60GHz phased array transceivers.

It should be noted that this PLL was designed using a very good frequency reference. If a lower cost reference must be used, the reference noise will increase. Due to the large division ratio inherent in mm-wave PLLs, the output referred reference noise could be much higher than the noise due to the VCO. In that case, the loop bandwidth should be set as low as possible to take advantage of the low VCO phase noise (see Section 4.2.2). This is unlike low frequency PLLs where the output referred reference noise is almost always much better than the on-chip VCO.

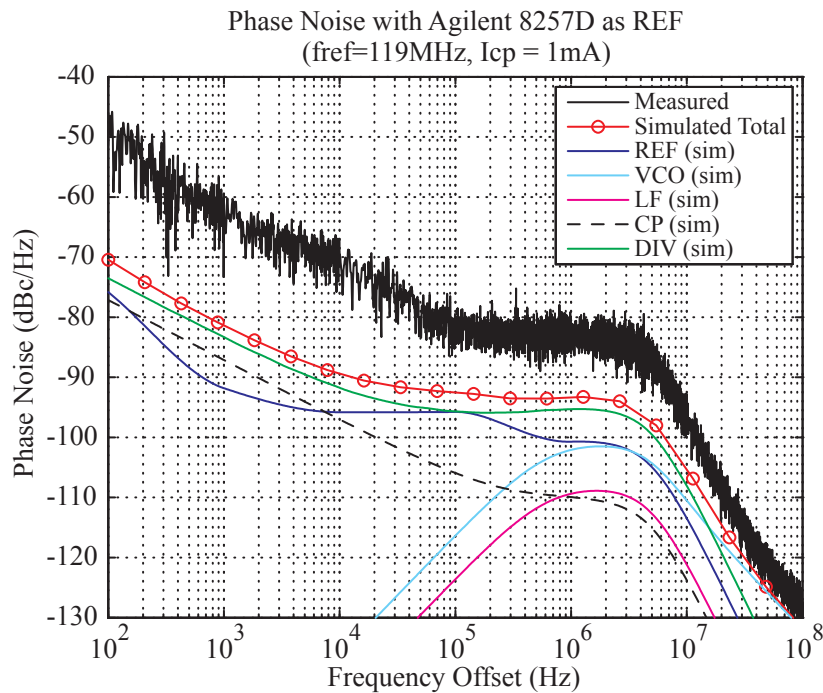
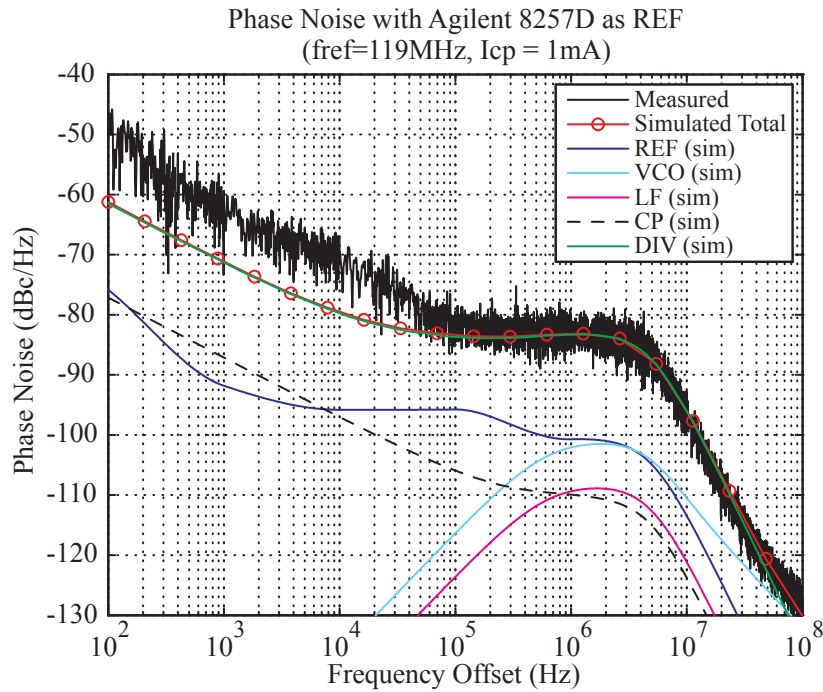


Figure 4.28: Comparison between measured PLL phase noise and system simulation (including individual contributors).

	[Murphy10]	[Wu09]	[Zhang09]	This Work
Technology	CMOS 65nm	CMOS 0.18 μ m	CMOS 65nm	CMOS 65nm
Frequency (GHz) (Tuning Range)	42.1-53 (22.9%)	53-58 (9%)	55.4-60.3 (8.5%)	57.9-65.6 (12.5%)
DC Power (mW)	72	35.7	46	29
Reference Spur (dBc)	-	-40	-35	-42
Output Power (dBm)	-	-37.85	-7	+1.2
PLL Phase Noise (dBc/Hz)	-81 (in-band) -84.5 @ 1MHz*	-85.2 (in-band) -90.9 @ 10MHz	-65 (in-band) -87 @ 1MHz	-82 (in-band) -107 @ 10MHz
VCO FOM (dBc/Hz)	-179	-157.5	-175.3	-178.3
VCO FOM_T (dBc/Hz)	-186.2	-156.6	-173.9	-180.2

* from phase noise plot of PLL locked to 51.84GHz

Table 4.1: PLL performance summary and comparison

4.A Spectral Purity Metrics

In order to quantify the quality of a signal with a complicated phase noise profile different metrics can be used depending on the requirements of the particular application. In some cases a spectral mask limits the phase noise at each offset frequency. To reduce the phase noise spectrum to a single number representing the purity of the carrier we compute its rms value.

$$\sqrt{2 \int_{\omega_{lo}}^{\omega_{hi}} 10^{\mathcal{L}\{\Delta\omega\}/10} df}$$

This involves integrating the single-sideband phase noise spectrum between the offset frequencies ω_{lo} and ω_{hi} .⁹ These upper and lower limits are set by the requirements of the particular application. ω_{hi} is usually set to the PLL reference frequency. ω_{lo} is set by the data packets length or, in the case of systems utilizing carrier recovery, the bandwidth of the carrier recovery loop. Using this integral we can then compute the following metrics:

$$IPN = 20 \cdot \log_{10} \left[\sqrt{2 \int_{\omega_{lo}}^{\omega_{hi}} 10^{\mathcal{L}\{\Delta\omega\}/10} df} \right] \quad (4.65)$$

$$\delta_{j,rms} = \frac{1}{\omega_o} \sqrt{2 \int_{\omega_{lo}}^{\omega_{hi}} 10^{\mathcal{L}\{\Delta\omega\}/10} df} \quad (4.66)$$

$$\theta_{n,rms} = \frac{180^\circ}{\pi} \sqrt{2 \int_{\omega_{lo}}^{\omega_{hi}} 10^{\mathcal{L}\{\Delta\omega\}/10} df} \quad (4.67)$$

⁹Since the phase noise spectrum is symmetric about the carrier, we simply multiply the integral of the single-sideband phase noise by 2 to get the integrated double-sideband phase noise.

The integrated phase noise in dB , labeled IPN , is equivalent to the signal to noise ratio (in dB) of the noisy carrier and is useful when the carrier is used as a downconversion LO. In the case of a phase modulated signal, the received SNR due to LO phase noise only is equal to the IPN of the LO. $\delta_{j,rms}$ is the rms jitter in seconds and represents the uncertainty in the time between zero-crossings (i.e.: the instantaneous period of the carrier). This metric is more readily applicable when the carrier is used to sample a signal as it represents the uncertainty in the sampling interval. Finally, $\theta_{n,rms}$ is the rms phase error of the carrier in degrees and is simply the rms jitter converted into the phase domain.

Chapter 5

LO Distribution

A phased array transceiver utilizing baseband or LO phase shifting requires the LO signal generated by a PLL (as described in Chapter 4) to be distributed to each individual element of the array. As the number of elements in the phased array increases, the power consumption of the LO distribution can quickly get out of control if not carefully managed. Aside from optimizing individual components for energy efficiency, appropriate architectural choices must be made to allow scalability to large numbers of elements while maintaining low per-element power consumption. In this chapter, we will describe the different ways we can distribute the LO to multiple elements of a phased array transceiver. We will discuss both architectural design choices as well as the optimization of individual building blocks leading to a sample design of an LO distribution subsystem for an 8 element baseband phased array transceiver. This transceiver uses quadrature direct conversion and was designed with high RF bandwidth in order to take advantage of as much of the 60GHz band as possible to send data at high data rates with simple modulation schemes (e.g.: QPSK, 16-QAM).

5.1 Mixer LO Requirements

To understand the requirements on the LO distribution we must begin at the mixer. The LO signal must be distributed to all mixers in the receive and transmit paths. Regardless of whether the mixer is active or passive, there is a minimum LO amplitude required for good conversion gain. As an example, two mixers were designed in a 65nm process, one passive and the other active. The mixers were sized to produce the same conversion gain at peak LO amplitude¹. The conversion gain was then plotted versus LO amplitude in Fig. 5.1. We can see that the active mixer conversion gain is relatively insensitive to LO amplitude above a minimum threshold. The passive mixer conversion gain, on the other hand, drops off quickly as the LO amplitude is reduced from its peak value. In either case, a minimum

¹Both mixers were simulated with the same peak LO amplitude.

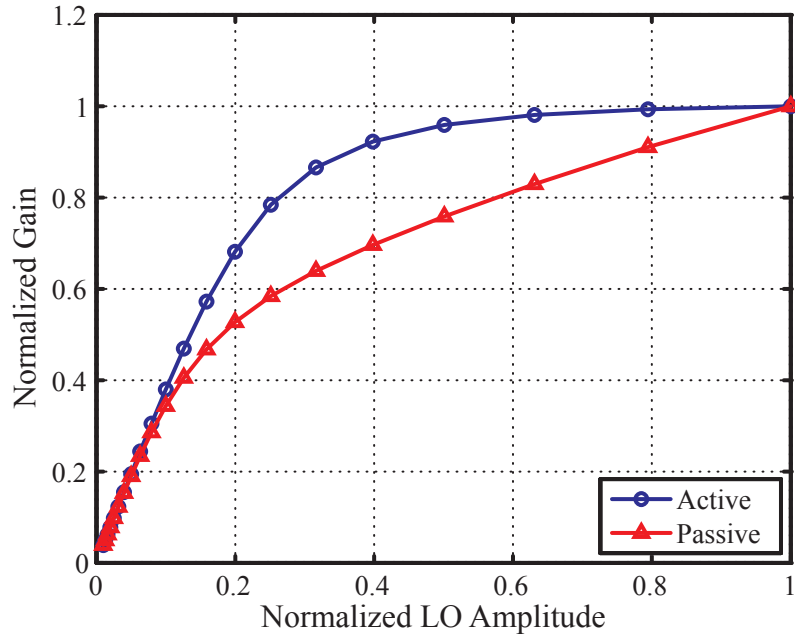


Figure 5.1: Comparison of mixer gain versus LO amplitude.

LO amplitude threshold is selected to give good conversion gain. Next, we must consider the input impedance of the mixer LO port which must be driven with this minimum amplitude.

The size of the mixer switches is determined by considerations such as linearity, output power, gain, and noise. The topology of the mixer also affects the total load impedance seen at the LO port. For example, a single gate mixer only has one switch, while a double-balanced Gilbert quad has four switches which must be driven in pairs with a differential signal. The single-balanced and double-balanced topologies are the most common since no signal summing is required before the mixer. At low frequencies, the input impedance of the mixer LO port is simply a capacitance which scales with the mixer switch size. Mixer switches can be very large leading to a large capacitive load for the LO path. At high frequencies, however, the gate resistance of the switches also introduces a real part to the input impedance, making it look like a capacitor with finite Q.² Without any loss of generality we can thus model the input admittance of the mixer LO port at any frequency as a shunt RC network (Fig. 5.3). As the switch size is increased the capacitance increases and the shunt resistance decreases. Due to the large voltage swing required, a buffer is always necessary to drive this impedance with sufficient power to achieve the required LO swing.³ The required LO amplitude along with the input impedance of the mixer LO port determines how much

²Due to C_{gd} , the load impedance of the mixer also affects the input impedance and should be taken into account.

³A good buffer also has low reverse gain and thus provides isolation between elements reducing possible coupling through the LO path.

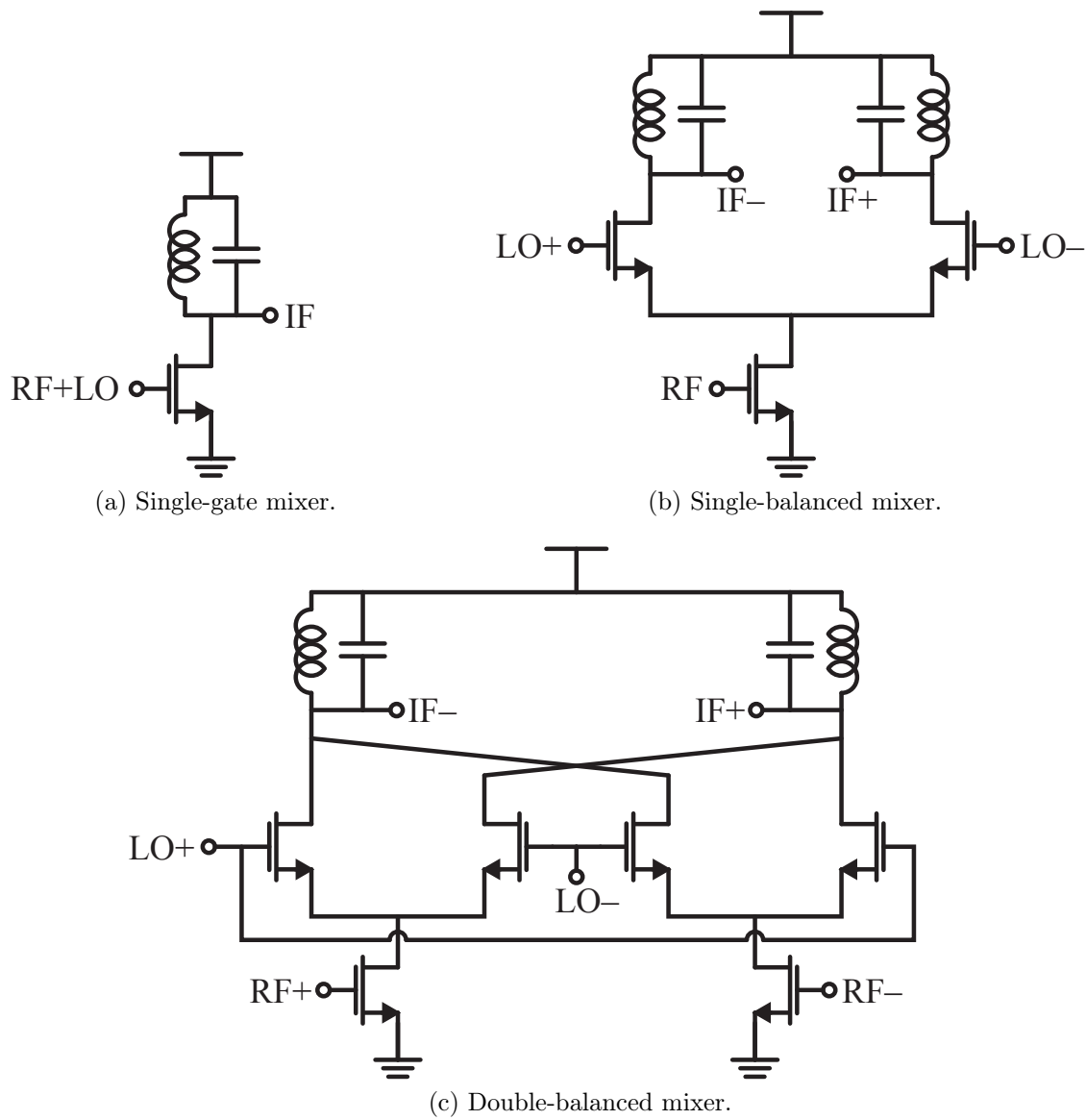


Figure 5.2: Common mixer topologies.

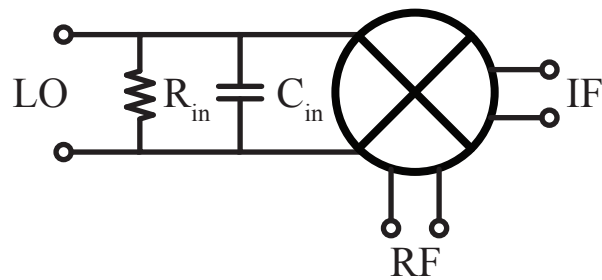


Figure 5.3: Mixer input admittance.

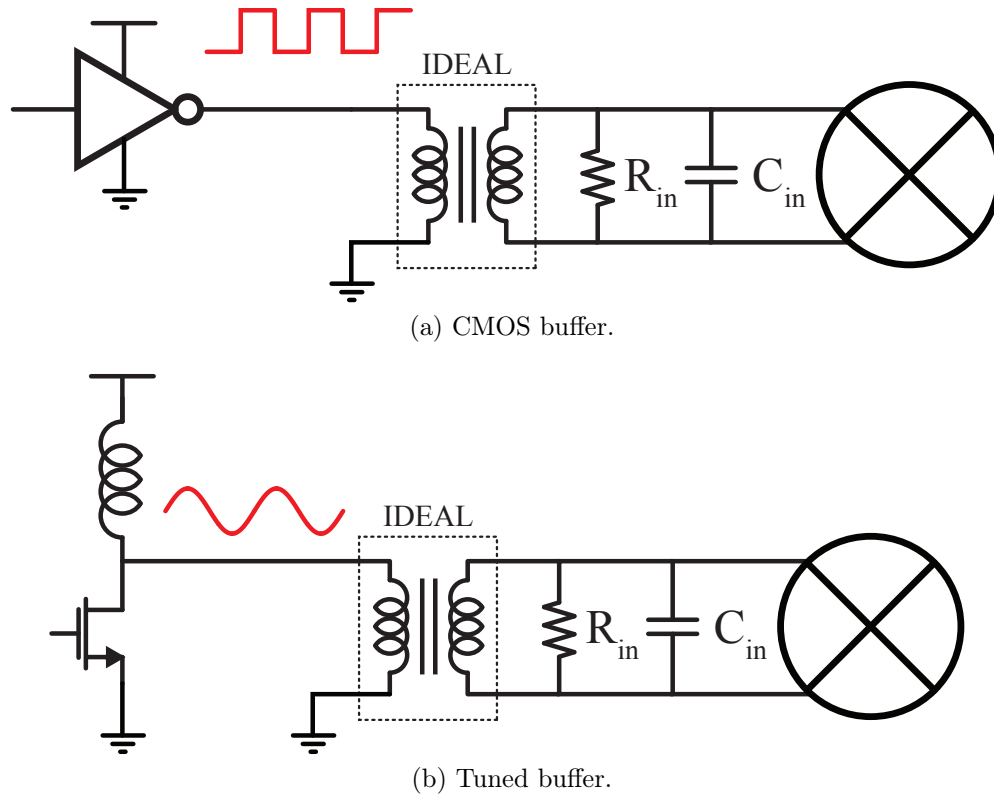


Figure 5.4: LO buffer topologies.

power must be delivered to each mixer with larger mixer switches resulting in higher power.

The simplest type of buffer is a CMOS inverter (Fig. 5.4a). Since most mixers require a differential LO signal, an ideal balun has been added to convert the single-ended output of the CMOS buffer to a differential signal. For now we will assume this balun is ideal but in the final design this balun will become an integral component. This type of buffer can provide the mixer LO port with a square wave LO signal with rail-to-rail swing. The dynamic LO power dissipated in charging and discharging the mixer input capacitance is equal to

$$P_{LO,CMOS} = C_{mix} V_{dd}^2 f_{LO} \quad (5.1)$$

There is no way to reduce this power without either reducing the LO swing or the mixer capacitance. Furthermore, we have assumed that the input impedance is dominated by the capacitor. At high frequencies, the resistance becomes a large part of the input impedance and also consumes LO power. Finally, CMOS buffers have a limited bandwidth beyond which they can no longer operate. At higher frequencies, a CML buffer with resistor load is used, trading off bandwidth for power consumption. Nevertheless, the power dissipated in simply driving the load with a square wave does not reduce.

At high frequencies we can instead utilize a resonant network both to increase the useful

frequency range of the LO buffer, as well as to reduce the LO power required. A tuned buffer can be created by using a common-source amplifier with an inductive load (Fig. 5.4b). Again, we have used an ideal balun to convert the single-ended output to a differential signal. Reflected through the balun, the inductor appears in shunt with the mixer input and allows us to tune out the mixer capacitance at the LO frequency. Since this is a tuned amplifier, the higher harmonics are rejected by the output tank and the resulting output is a sinusoidal signal. Another advantage of this type of amplifier is that the output can swing above V_{dd} leading to a maximum possible peak-to-peak LO swing of $2V_{dd}$, much higher than what is possible with either a CMOS buffer or a resistor loaded amplifier.

Assuming a high-Q inductor, the loaded Q of the resonant tank is approximately equal to the quality factor of the mixer input capacitance

$$Q_{mix} = \omega_{LO} C_{mix} R_{mix} \quad (5.2)$$

At resonance then, the buffer's load impedance is purely resistive and equal to R_{mix} . The power dissipated by the mixer input resistance driven by a sinusoidal LO signal with amplitude V_{LO} is equal to

$$\begin{aligned} P_{LO,tuned} &= \frac{V_{LO}^2}{2R_{mix}} \\ &= \frac{V_{LO}^2 \omega_{LO} C_{mix}}{2Q_{mix}} \\ &= \frac{\pi}{Q_{mix}} C_{mix} V_{LO}^2 f_{LO} \end{aligned} \quad (5.3)$$

To allow a fair comparison, the peak-to-peak output of the tuned amplifier should be equal to the peak-to-peak LO swing provided by the CMOS buffer (i.e.: $V_{LO} = V_{dd}/2$). For the same output swing, the tuned amplifier needs to provide LO power equal to

$$P_{LO,tuned} \Big|_{V_{LO}=V_{dd}/2} = \frac{\pi}{Q_{mix}} C_{mix} \left(\frac{V_{dd}}{2} \right)^2 f_{LO} = \frac{\pi}{4Q_{mix}} C_{mix} V_{dd}^2 f_{LO} \quad (5.4)$$

Thus, the required LO power is reduced approximately by the quality factor of the tank. It is therefore always beneficial to use a tuned buffer instead of directly driving the mixer input capacitance. The only way to reduce power further is to reduce the mixer switch size, the LO frequency, or the LO amplitude. Unfortunately, these factors are usually out of our control and we must focus on delivering the required power as efficiently as possible.

5.2 LO Generation Strategy

There are multiple strategies that can be used to create a local LO signal which provide trade-offs between design complexity, power consumption, and performance. The main goal of our

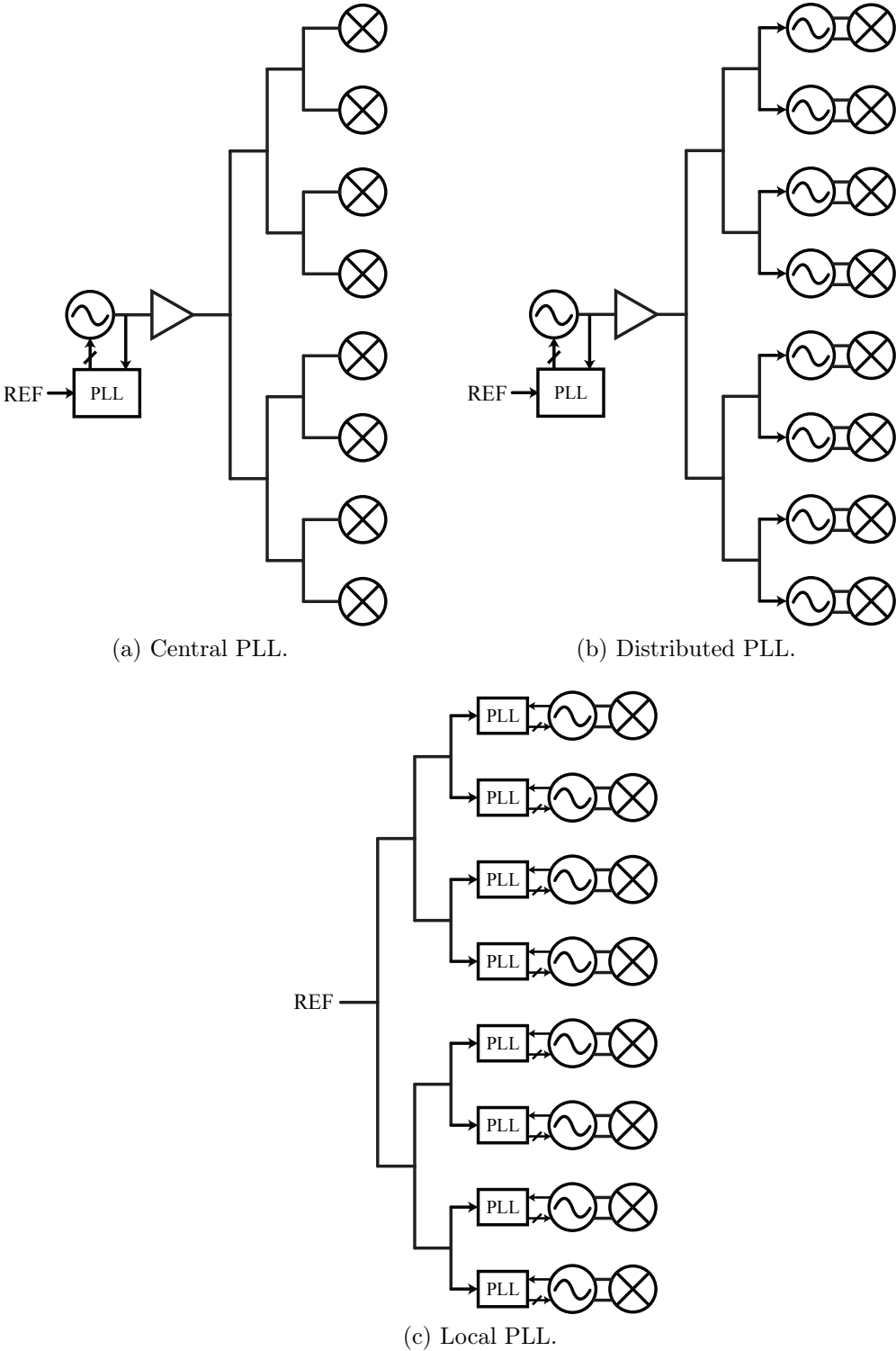


Figure 5.5: LO generation strategies.

distribution strategy, however, is reducing power consumption without affecting transceiver performance.

First, we must decide on the LO generation strategy. There is a spectrum of possibilities in this regard. On one end of this spectrum one central PLL is used to generate the LO signal which is then split and distributed among all the elements (Fig. 5.5a). The distributed signal in this case is the LO at mm-wave frequencies. The opposite of this strategy is to utilize a local PLL at each element (Fig. 5.5b). The only distribution required between elements then is a low frequency reference to which each PLL is locked. This distribution is trivial as the reference frequency is on the order of 10s to 100s of MHz. However, placing a PLL at each element requires very large area and power consumption as each element is essentially a standalone transceiver.

A central PLL is not without its challenges either. If there are only a small number of elements, they can be placed very close and distributing the LO from a central PLL is trivial despite the high signal frequency. If there are many elements, however, the LO must be split many times and distributed over very large distances that can approach the wavelength. We will address this issue later. As the signal is split and distributed, the power arriving at each local mixer reduces. Therefore, gain is required in the LO path to bring up the power to the level required by the mixers.

Between these two options there is also a hybrid possibility which we will call the distributed PLL. This option involves placing a local oscillator at each element and utilizing a central PLL to lock all the oscillators to the stable reference frequency. The remaining question in this case is the mechanism used to lock the local oscillators. Distributing a common VCO control signal seems like an attractive option since it is a low frequency signal, nominally at DC when the loop is locked. However, it is also a high impedance node and is extremely sensitive to coupling from other signals which would modulate the VCO. In a modern integrated transceiver there are many aggressors, including digital circuits and baseband signals which could very easily couple to these distributed VCO control lines. For this reason, the VCO control line must be kept as short as possible and well shielded to avoid any external coupling and maintain a pure LO signal. Furthermore, even if we could buffer this line well enough, mismatch between the different VCOs would lead to phase and frequency differences between the elements. Thus, using a single control line for multiple, widely separated, VCOs is not a robust solution.

Instead, local oscillators can be injection locked to a high purity central oscillator itself embedded in a PLL (Fig. 5.5b). The central oscillator can be optimized for low phase noise and consume more power since this will be amortized over the number of elements. Each local injection locked oscillator (ILO), on the other hand, can be optimized for low power consumption since its phase noise will track the phase noise of the injection locking signal from the central high purity VCO. The behavior of such a system is similar to the injection locked divider described in Section 4.3.2. The injection signal distributed to each ILO can potentially be very small but from (4.59) we can see that the injection locking range is directly proportional to the injection signal strength. Since the 60GHz band requires more

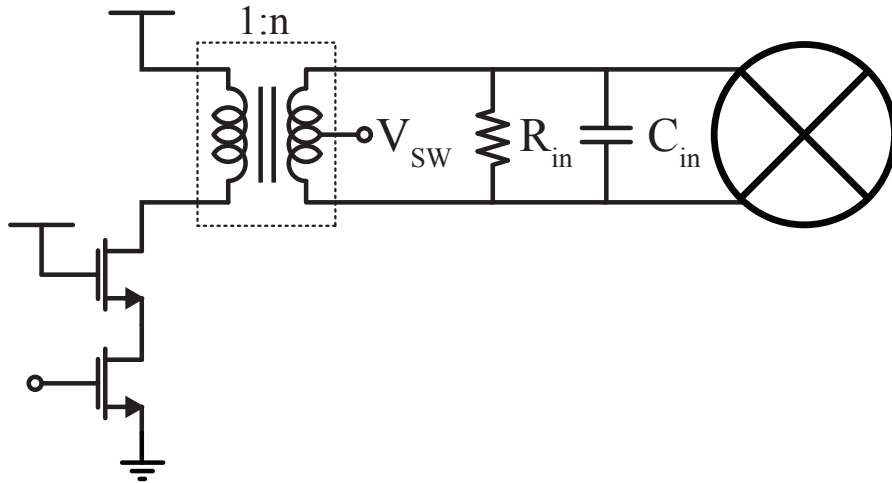


Figure 5.6: LO buffer schematic.

than 10% tuning range either the ILO tank quality factor must be made very low leading to high power consumption in the local oscillator itself, or the injection signal must be made very large leading to large power consumption in the distribution network. Another potential problem with this topology arises in a direct conversion transmitter. For high output power levels (+10dBm or higher), the modulated output signal could couple to and cause pulling in the ILO (it is after all designed to injection lock to an external signal). With a central oscillator, the coupling from each PA can be minimized to eliminate the pulling problem.

Of the three options discussed, a local PLL at each element would provide good performance and high flexibility in scaling an array up to many elements, at the cost of the highest power consumption and area usage. A central PLL, on the other hand, can be allowed to consume more power in order to achieve low phase noise since its power and area consumption is amortized over the array. A central PLL was thus chosen for this design.

The only remaining choice is then between a central or distributed PLL. In both cases, a strong signal must be distributed at the LO frequency to each element. Due to the trade-offs discussed above, the injection signal for an ILO cannot be very small. Therefore, the power consumption of the LO distribution will not be vastly different between the two choices.⁴ The choice then must be made on the performance of the local blocks.

5.3 Mixer LO Buffer Design Methodology

The design of the LO buffer begins with the mixer. The load impedance is determined based on the mixer topology and switch size. The LO amplitude is chosen based on the mixer requirements. The amplifier design should then focus on providing the required LO swing at the mixer with minimum power consumption. Therefore, a tuned buffer should be employed. Furthermore, the required LO signal is usually differential but it is much easier to split and distribute a single-ended signal. A balun should be included in the design of the buffer to convert the single-ended signal to differential. A single amplifier can then be used to drive the input of the balun directly.

The simplified schematic of the buffer described above is shown in Fig. 5.6. The balun in this design actually performs multiple functions. First, it converts the single-ended signal coming out of the buffer to a differential signal as required by the mixer. Second, its inductance tunes out both the mixer capacitance as well as the parasitic capacitance of the buffer. Third, it can perform an impedance transformation of the mixer load to bring the required voltage swing within the range the buffer can provide. Finally, it provides DC isolation and convenient biasing points. The grounded primary is used to provide the drain supply to the cascode buffer, while the center tap of the secondary is used to bias the mixer switches. In both cases, a bypass cap to ground is added in order to provide a low AC impedance. A matching network is added to the input of the buffer to provide a conjugate power match down to the distribution impedance. For simplicity we will use a single-stage L-match consisting of a shunt inductor at the transistor gate and a series capacitor. This type of matching network also conveniently provides DC isolation and a convenient bias point for the buffer input through the inductor. In general, we are not limited to this type of matching network but it provides a reasonable estimate of the loss one can expect from matching.

To maximize the efficiency of the buffer we must operate it close to its compression point, meaning we must use the maximum possible voltage swing. At mm-wave frequencies a cascode amplifier is preferred over a common-source topology since it provides higher gain and isolation (which leads to increased stability). This does, however, come at the expense of reduced maximum drain voltage swing and thus, lower efficiency.

The design procedure will proceed as follows. Scalable models for both the amplifier and transformer will first be introduced. Two design methods will then be described using these simplified scalable models. The first method is purely equation based and thus very fast but does not necessarily result in a globally optimum design. The second method involves holistic optimization of the buffer design including trade-offs between amplifier and transformer performance. While more time consuming, this method results in a globally optimum design.

⁴It is also possible to centrally generate and distribute a lower frequency signal which is then locally multiplied up to the LO frequency using either frequency multipliers or subharmonic ILOs. However, as discussed in previous chapters, these blocks are very inefficient at converting a lower frequency up to a higher frequency and would require additional local buffering, offsetting any power savings accrued from distributing a lower frequency signal.

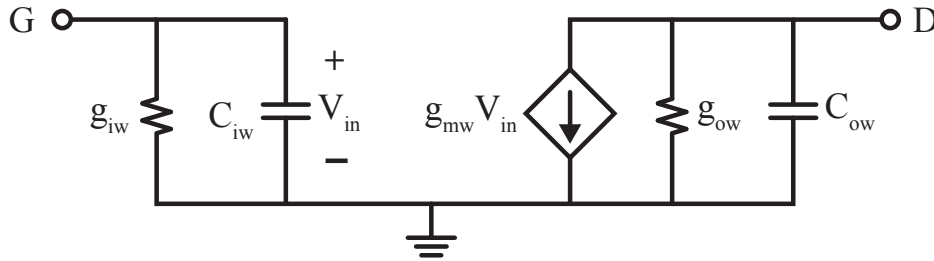


Figure 5.7: Width-scalable transistor model.

The results of these two methods is then compared over a wide range of mixer sizes. Finally, an alternative buffer topology using an injection locked oscillator is presented for comparison to the standard topology.

5.3.1 Scalable Amplifier Model

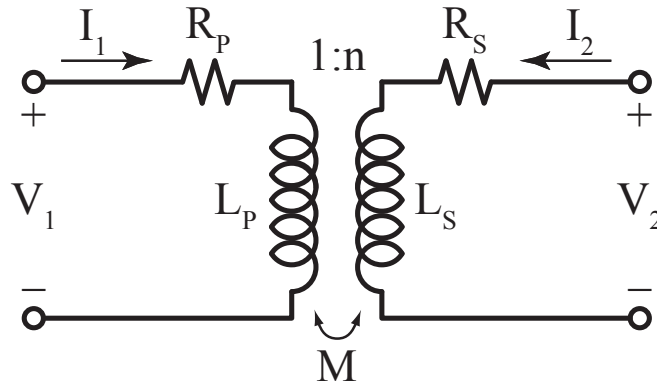
In order to facilitate the initial design process we will use simplified scalable models for both the amplifier and the transformer. The cascode buffer can be represented using the small signal model shown in Fig. 5.7. We will assume that both the common source and cascode devices are the same size allowing a shared-junction layout to reduce the parasitic capacitance at the common node. To maximize the buffer gain, we will maintain the current density which results in the maximum f_T by holding V_{gs} constant and scaling the device width as needed for the required bias current/transconductance. We will call this current density I_{dw} with units of $[A/m]$. The bias current, I_b , of a buffer having width W is then given by

$$I_b = I_{dw}W \quad (5.5)$$

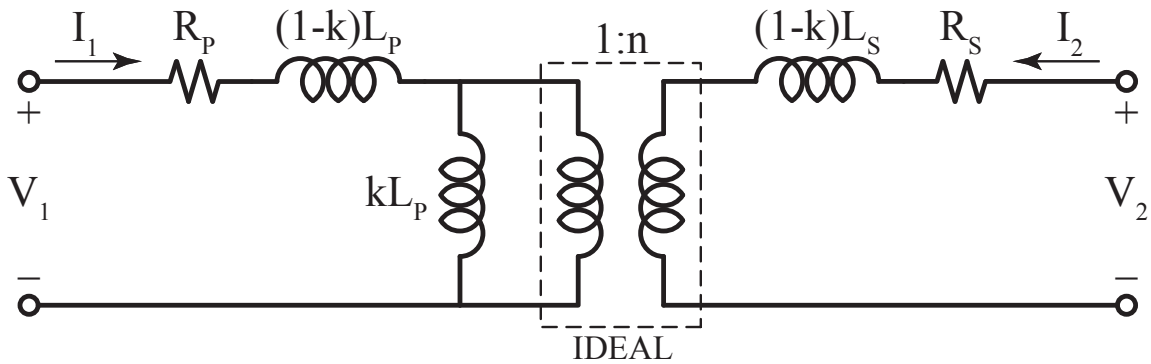
As a result, all model parameters in Fig. 5.7 are derived at this current density and given as a function of the device width, W (i.e.: $g_m = W \cdot g_{mw}$, $C_{in} = W \cdot C_{iw}$, $R_{in} = \frac{1}{W \cdot g_{iw}}$, etc.).

Note that both the input and output impedance of the amplifier are modeled as shunt networks which allows them to be easily absorbed into the input and output matching networks respectively. The output network consists of the output resistance and drain capacitance of the buffer and is commonly represented as a shunt network. The input network, on the other hand, consists of the gate resistance and gate to source capacitance of the common source device and is usually modeled as shown in Fig. 3.20. The shunt model in Fig. 5.7 is simply the result of a series-to-parallel transformation of this network. Beside the intrinsic device, however, we must also include the effects of layout parasitics. The final model parameters must thus be fit to the device performance after layout parasitic extraction in order to accurately represent its behavior.⁵

⁵Since we are using a cascode buffer, the reverse isolation is high so we can remove the input to output coupling capacitance from our model to greatly reduce computational complexity without significantly affecting the model performance.



(a) Mutual inductance model.



(b) Equivalent T-model.

Figure 5.8: Transformer models.

5.3.2 Scalable Transformer Model

Next, we require a model of the transformer. A transformer is simply a pair of coupled inductors, L_p and L_s , called the primary and secondary respectively. Each inductor has finite quality factor which is modeled as a series resistance for each inductor.

$$R_p = \frac{\omega L_p}{Q_p} \tag{5.6}$$

$$R_s = \frac{\omega L_s}{Q_s} \tag{5.7}$$

The coupling generates a mutual inductance M which is a function of the coupling factor k and the inductance values.

$$M = k\sqrt{L_p L_s} \tag{5.8}$$

The effective turns ratio n of the transformer is given by

$$n = \sqrt{\frac{L_s}{L_p}} \approx \frac{I_1}{I_2} \approx \frac{V_2}{V_1} \quad (5.9)$$

In an ideal transformer, the ratios of the primary to secondary voltage and current are equal to the turns ratio, n . However, in a real transformer loss and finite coupling factor make the turns ratio only an approximation of the voltage and current ratios. The transformer behavior is fully described by its 2-port Z-parameters which relate the port voltages to the port currents.

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} R_p + j\omega L_p & j\omega M \\ j\omega M & R_s + j\omega L_s \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (5.10)$$

This model, however, does not lend itself well to analysis so an equivalent model, valid at one frequency, is generally used. The equivalent T-model, shown in Fig. 5.8b [Aoki02], is based around an ideal $1 : n$ transformer with additional components to model losses and finite coupling. The finite coupling factor splits the total inductance into the magnetizing inductance kL_p and leakage inductances $(k - 1)L_p$ and $(k - 1)L_s$. The magnetizing inductance is the fraction of the inductance that actually performs the transformer function. The remainder is uncoupled inductance and appears as series lead inductance both on the primary and secondary. At low coupling factors the leakage inductance will significantly affect the behavior of the transformer. On-chip coupling factors are generally small, usually less than 0.9, since the oxide and silicon substrate act much like an air core.⁶ Finally, as before, the losses of both primary and secondary inductors are modeled by series resistors R_p and R_s given by (5.6) and (5.7) respectively.

The turns ratio of the transformer is chosen based on the ratio of the desired LO swing to the voltage swing that the buffer can provide. From the signal path standpoint, the transformer performs two functions: an impedance transformation to reduce the voltage swing required from the buffer, and single-ended to differential conversion. Instead of using a $1 : n$ transformer to perform both of these functions a standard lumped component matching network could perform the required impedance transformation followed by a $1 : 1$ transformer for single-ended to differential conversion. The advantage of using the $1 : n$ transformer is that its efficiency does not depend on the impedance transformation ratio, while the efficiency of a standard lumped component matching network decreases as the impedance transformation ratio increases [Aoki02]. This is the key reason for selecting a transformer based design.

5.3.3 Equation Based Buffer Design

We will now present the equation based design methodology for the buffer including sizing of the transformer and amplifier, as well as the input matching network. We begin with the

⁶Ferrite cores are used to increase coupling in low frequency transformers. At high frequencies, however, these cores self-resonate and are no longer useful.

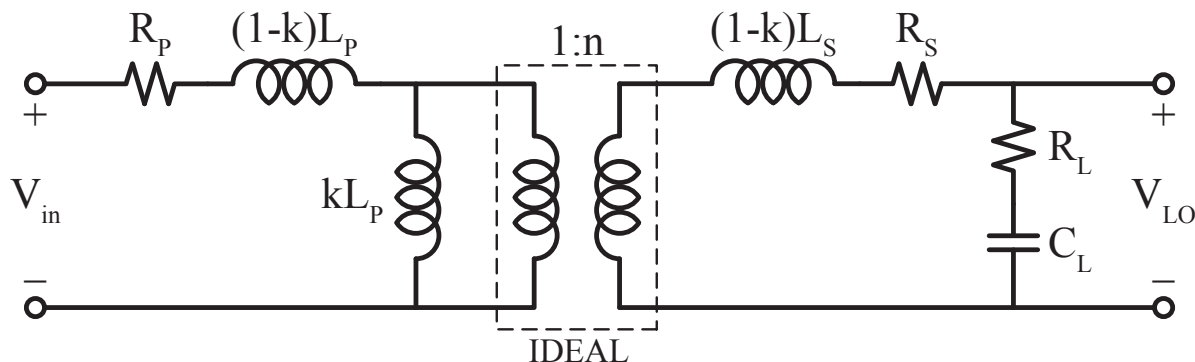


Figure 5.9: Model of transformer with effective mixer load.

selection of the transformer turns ratio. For peak power gain, a common source device should be biased at the current density that results in peak f_T . Similarly, in a cascode amplifier, the bottom (common-source) transistor should be biased near its peak f_T current density, while the top (common-gate) transistor gate should be biased at V_{dd} .⁷ The amplifier will only maintain high gain if both transistors remain in saturation. The output voltage swing of the cascode is thus limited to approximately one threshold voltage, V_T . Beyond this level, the top transistor is pushed into triode and the gain drops quickly. For the common-source, the output voltage can swing all the way down to V_{OV} , the overdrive voltage of the device, before it gets pushed into triode. The required transformer turns ratio can then be approximated using

$$n \approx \frac{V_{LO}}{V_{o,max}} \quad (5.11)$$

where V_{LO} is the desired differential LO amplitude and $V_{o,max}$ is the maximum buffer output voltage amplitude.

Next, we must select the transformer size. This problem is similar to the design of transformer coupled power amplifiers so we can use the procedure presented in [Aoki02] to find the transformer size that minimizes its loss when driving the mixer LO port impedance. The simplified model for this optimization problem is shown in Fig. 5.9, where the load impedance R_L and C_L are found by performing a parallel-to-series transformation on the mixer input impedance as shown in (5.12) and (5.13) respectively.

$$R_L = \frac{R_{mix}}{1 + Q_{mix}^2} = \frac{R_{mix}}{1 + (\omega_{LO} R_{mix} C_{mix})^2} \quad (5.12)$$

$$C_L = C_{mix} (1 + Q_{mix}^{-2}) = \frac{1 + (\omega_{LO} R_{mix} C_{mix})^2}{\omega_{LO}^2 R_{mix}^2 C_{mix}} \quad (5.13)$$

⁷A high swing cascode bias would increase the linear range of the output by reducing the cascode gate voltage but, due to severe channel length modulation in deeply scaled CMOS, the reduced drain bias voltage on the common source transistor would cause a significant drop in the linear gain and require more input power for the buffer.

Next, we can write the transformer efficiency as the ratio of the power delivered to the load to the total power delivered to the transformer.

$$\begin{aligned}\eta &= \frac{P_{load}}{P_{total}} \\ &= \frac{\left|I_2\right|^2 \frac{R_L}{n^2}}{\left|I_1\right|^2 R_p + \left|I_2\right|^2 \left(\frac{R_s}{n^2} + \frac{R_L}{n^2}\right)}\end{aligned}\quad (5.14)$$

To simplify this equation we must solve for I_1 as a function of I_2 . There is a current divider between the magnetizing inductance and the ideal transformer which can give us this ratio.

$$\left|I_2\right|^2 = \left|I_1\right|^2 \left| \frac{j\omega k L_p}{j\omega k L_p + \frac{Z_s}{n^2}} \right|^2 \quad (5.15)$$

where Z_s is the impedance present on the secondary of the ideal transformer

$$Z_s = R_s + j\omega(1-k)L_s + \frac{1}{j\omega C_L} + R_L \quad (5.16)$$

Using (5.15) and (5.16) we arrive at

$$\left|I_1\right|^2 = \left|I_2\right|^2 \frac{\left(\frac{R_s}{n^2} + \frac{R_L}{n^2}\right)^2 + \left(\frac{\omega L_s}{n^2} - \frac{1}{\omega n^2 C_L}\right)^2}{(\omega k L_p)^2} \quad (5.17)$$

Plugging (5.17) and (5.6)-(5.9) into (5.14) and simplifying yields

$$\begin{aligned}\eta &= \frac{R_L}{\left[\frac{(R_s + R_L)^2 + (\omega L_s - 1/\omega C_L)^2}{(\omega k L_p)^2} \right] \frac{R_p}{n^2} + R_s + R_L} \\ &= \frac{R_L}{\frac{(\omega L_s/Q_s + R_L)^2 + (\omega L_s - 1/\omega C_L)^2}{\omega L_s k^2 Q_p} + \frac{\omega L_s}{Q_s} + R_L}\end{aligned}\quad (5.18)$$

From (5.18) we can see that for a given turns ratio, there is an optimum L_s which gives the peak efficiency. To find this optimum inductance value we take the derivative with respect to L_s , set the result equal to zero, and solve.

$$L_{s,opt} = \frac{1}{\omega^2 C_L} \sqrt{\frac{1 + (\omega R_L C_L)^2}{1 + 1/Q_s^2 + k^2 Q_p/Q_s}} = \frac{\alpha}{\omega^2 C_L} \quad (5.19)$$

$$\alpha = \sqrt{\frac{1 + 1/Q_{mix}^2}{1 + 1/Q_s^2 + k^2 Q_p/Q_s}} \quad (5.20)$$

where R_L and C_L are the series transformed impedances of the mixer given by (5.12) and (5.13) respectively and Q_{mix} is the quality factor of the mixer input impedance.⁸ The maximum efficiency is then given by plugging (5.19) back into (5.18).

$$\eta_{max} = \left[\frac{1}{k^2 Q_{mix} Q_p \alpha} \left(1 + \alpha \frac{Q_{mix}}{Q_s} \right)^2 + \frac{(\alpha - 1)^2 Q_{mix}}{k^2 Q_p \alpha} + \alpha \frac{Q_{mix}}{Q_s} + 1 \right]^{-1} \quad (5.21)$$

Once the optimum transformer size has been selected, the buffer size must be chosen. Using the model in Fig. 5.9 and (5.16) we first calculate the load impedance seen by the buffer, $Y_{in,x}$,

$$Y_{in,x} = R_p + j\omega(1-k)L_p + \frac{j\omega k L_p \cdot \frac{Z_s}{n^2}}{j\omega k L_p + \frac{Z_s}{n^2}} \quad (5.22)$$

and the actual voltage gain of the transformer, A_{vx} .

$$A_{vx} = \frac{R_L + \frac{1}{j\omega C_L}}{R_L + \frac{1}{j\omega C_L} + R_s + j\omega(1-k)L_s} \cdot n \cdot \frac{(j\omega k L_p) \left\| \left(\frac{Z_x}{n^2} \right) \right.}{(j\omega k L_p) \left\| \left(\frac{Z_x}{n^2} \right) + R_p + j\omega(1-k)L_p} \quad (5.23)$$

The drain voltage swing that the buffer must provide is then

$$V_o = \frac{V_{LO}}{A_{vx}} \quad (5.24)$$

Based on the selection of transformer turns ratio, we know that V_o is approximately the maximum voltage swing the buffer can support and we will assume that the buffer is driven just into compression to achieve this output voltage. As in a typical Class-A design, the buffer bias current must then be chosen based on the output voltage swing and load impedance

$$I_b = \left| V_o (Y_{in,x} + Y_{o,buf}) \right| \quad (5.25)$$

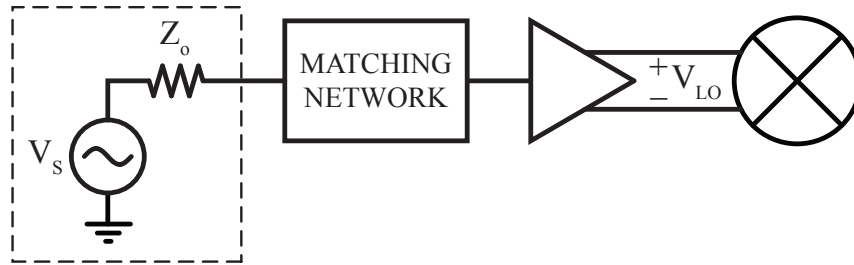
where $Y_{o,buf}$ is the output admittance of the buffer. We can then use (5.5) and the buffer model parameters from Fig. 5.7 to solve for the required buffer width.

$$\begin{aligned} (I_{bw} W_{buf})^2 &= |V_o (Y_{in,x} + W_{buf} (g_{ow} + j\omega C_{ow}))|^2 \\ \frac{(I_{dw} W_{buf})^2}{|V_o|^2} &= |\mathcal{R}\{Y_{in,x}\} + j\mathcal{I}\{Y_{in,x}\} + W_{buf} (g_{ow} + j\omega C_{ow})|^2 \end{aligned} \quad (5.26)$$

Expanding the right side of (5.26) and gathering terms leads to a second order polynomial which can be solved using the quadratic equation.

$$W_{buf}^2 \left[g_{ow}^2 + \omega^2 C_{ow}^2 - \frac{I_{dw}^2}{|V_o|^2} \right] + 2W_{buf} [g_{ow} \mathcal{R}\{Y_{in,x}\} + \omega C_{ow} \mathcal{I}\{Y_{in,x}\}] + |Y_{in,x}|^2 = 0 \quad (5.27)$$

⁸Alternatively, the primary inductance can be found by dividing the result from (5.19) by n^2 .

Figure 5.10: Matched buffer driven by source with impedance Z_o .

The final step in the buffer design is the input matching network which performs a conjugate match to the distribution network impedance, Z_o , for maximum power transfer. The selection of Z_o (addressed in Section 5.4) is actually under our control since all distribution is performed on-chip. The gate of the buffer could also be driven directly by a transmission line without a matching network. However, it is much more efficient to perform a conjugate match in order to maximize the power transfer from the distribution network to the buffer.

Once the buffer design is complete, the required input power can be found by first calculating the power required at the mixer and then subtracting the transducer power gain, G_T , of the buffer (including matching network losses). The real power required at the mixer is simply given by the the required LO voltage swing and the mixer input resistance

$$P_{mix} = \frac{|V_{LO}|^2}{2R_{mix}} \quad (5.28)$$

The transducer power gain of the buffer driven by a source with impedance Z_o (Fig. 5.10) is given by

$$G_{T,buf} = \frac{P_L}{P_{avs}} = |A_v|^2 \frac{4Z_o}{R_{mix}} \quad (5.29)$$

where A_v is the voltage gain from the source to the mixer (V_{LO}/V_S). The input power required for the buffer is then

$$P_{in} = \frac{P_{mix}}{G_{T,buf}} \quad (5.30)$$

5.3.3.1 Design Procedure Summary

1. Given a mixer, determine the required V_{LO} , its input impedance (R_{mix} and C_{mix}), and the resulting power that must be delivered for the required LO swing, P_{mix} .
2. Select an amplifier topology (eg.: common-source or cascode) and find its model parameters ($V_{o,max}$, g_{mw} , g_{ow} , C_{ow} , etc.).
3. Choose a balun topology (lateral/vertical, shape, etc.) and find the required turns ratio using (5.11).

4. Determine k , n , Q_p , and Q_s for the chosen balun type.
5. Using (5.19), find the optimum transformer size, L_s .
6. Using (5.27), find the optimum amplifier size, W_{buf} .
7. Given a distribution impedance, Z_o , design an input matching network for the buffer and calculate the total buffer power gain, $G_{T,buf}$, using (5.29).
8. Determine the power that must be delivered to the buffer by the distribution network by subtracting $G_{T,buf}$ from P_{mix} .

5.3.3.2 Limitations of Equation Based Design Method

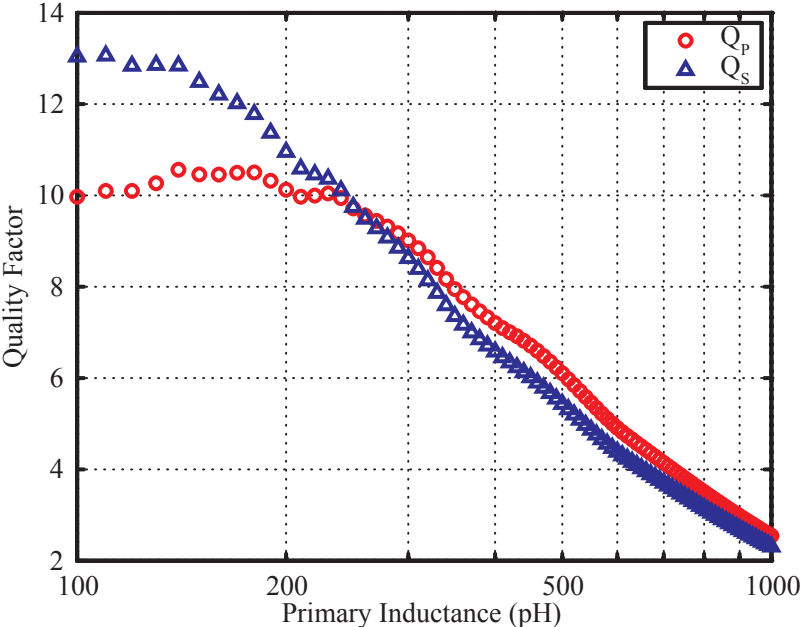
This simple analysis does not include the buffer output capacitance, or excess layout parasitics such as long leads, in the selection of the transformer size. Furthermore, this analysis ignores the fact that the transformer parameters n , k , Q_p and Q_s are all functions of the transformer size (i.e.: L_p and L_s). To show this effect, a 1 : 2 transformer with variable size was simulated using HFSS and the model parameters were extracted and plotted in Fig. 5.11. The size of the transformer will have to be adjusted to account for these factors but the above equations provide a good starting point for a first pass design.

One way to refine this first pass design into a more optimal design is to repeat the whole procedure in an iterative fashion. Begin by making a reasonable assumption of the transformer model parameters and selecting an optimum transformer size. Next, simulate that transformer and extract the actual model parameters. Using these new parameters, reoptimize the transformer size. Repeat this procedure until the optimum transformer size converges to a constant value from one step to the next. This method, however, may not arrive at a global optimum design, but rather a local optimum.

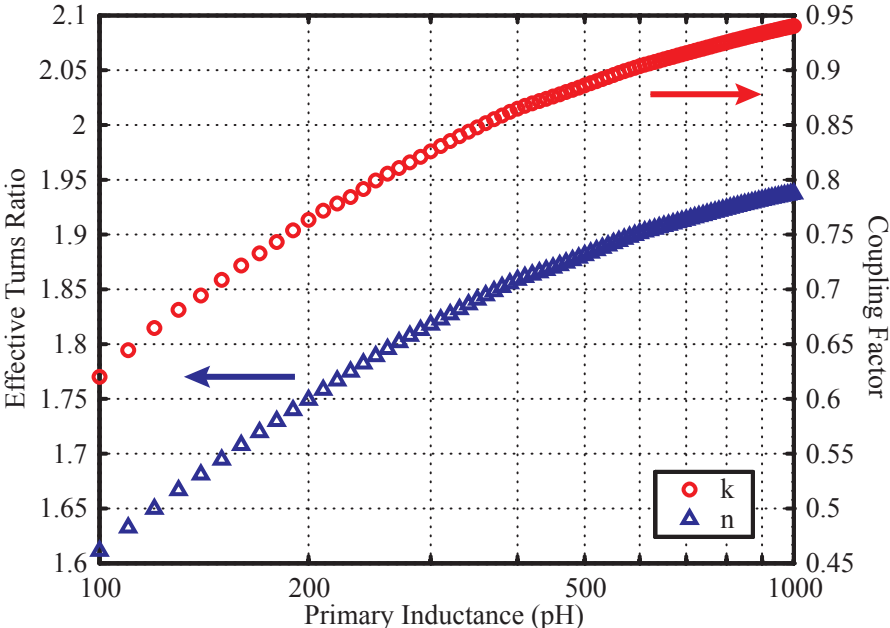
5.3.4 Optimization Based Buffer Design

A more rigorous, and time consuming, method can help us find the global optimum. First we must simulate a wide range of sizes for the transformer type of interest in order to extract the model parameters as a function of the transformer size. For example, Fig. 5.11 shows the model parameters versus L_p for a 1:2 lateral transformer. Doing so will allow us to exhaustively try all combinations of transformer and amplifier sizes to find the optimum for a given mixer size and required LO swing. The design procedure is as follows:

1. Given a mixer, determine the required V_{LO} , its input impedance (R_{mix} and C_{mix}), and the resulting power that must be delivered for the required LO swing, P_{mix} .
2. Select an amplifier topology (eg.: common-source or cascode) and find its model parameters ($V_{o,max}$, g_{mw} , g_{ow} , C_{ow} , etc.).



(a) Primary and secondary quality factors.



(b) Coupling factor and effective turns ratio.

Figure 5.11: Simulated 1 : 2 transformer parameters as a function of L_p

3. Choose a balun topology (lateral/vertical, shape, etc.) and find the required turns ratio using (5.11).
4. Using an EM simulator extract the transformer model parameters as a function of size, (i.e.: $n(L_s)$, $k(L_s)$, $Q_s(L_s)$, and $Q_p(L_s)$) using (5.6)-(5.10).
5. Calculate the loaded transformer voltage gain, A_{vx} , and input impedance, Y_{inx} , as a function of size using (5.23) and (5.22), respectively.
6. For each transformer size use the result from the previous step to determine the amplifier size needed to achieve the required LO swing using (5.27).
7. The previous step results in a range of design choices $[L_s, W_{buf}]$ which all give the required LO swing but have varying amounts of power consumption equal to $W_{buf}I_{dw}$. Select the value of L_s that results in the minimum amplifier size, W_{buf} , and therefore minimum buffer power consumption. This choice also has the highest power gain, reducing the required input power from the distribution network.
8. Given a distribution impedance, Z_o , design an input matching network for the buffer and calculate the total buffer power gain, $G_{T,buf}$, using (5.29).
9. Determine the power that must be delivered to the buffer by the distribution network by subtracting $G_{T,buf}$ from P_{mix} .

While resulting in a globally optimal design, this procedure involves significant investment in EM simulation up front which could be very time consuming.

5.3.5 Comparison Between Buffer Design Methods

Using the equations and methodology described above we will now design the optimum LO buffer for any arbitrary mixer size in a standard digital 1-poly, 7-metal, 65nm CMOS process with a 1.2V supply. We first need to extract the input impedance of the mixer using the shunt model shown in Fig. 5.3. For this purpose a single balanced mixer was designed with $20\mu m$ switching transistors. The differential input impedance, including extracted layout parasitics, was found from simulation to be $R_{mix} = 1.4k\Omega$ in parallel with $C_{mix} = 23fF$. Assuming the impedance will scale linearly with switch size, the input impedance of an arbitrary mixer with switch size W_{mix} can be estimated as

$$R_{mix} = \frac{28k\Omega \cdot \mu m}{W_{mix}} \quad (5.31)$$

$$C_{mix} = W_{mix} (1.15fF/\mu m) \quad (5.32)$$

$$Q_{mix} = \omega_{LO} R_{mix} C_{mix} = \frac{f_{LO}}{4.94GHz} \quad (5.33)$$

At 60GHz then $Q_{mix} \approx 12$. From simulation, the LO swing required for this mixer to have high conversion gain is 700mV differential amplitude. To be safe, we will design for an additional 1dB of LO swing: 780mV. The real input power required to drive the mixer to the designated LO swing can be calculated using

$$P_{mix} = \frac{|V_{LO}|^2}{2R_{mix}} = \frac{|V_{LO}|^2 W_{mix}}{56k\Omega \cdot \mu m} \quad (5.34)$$

To drive this mixer we select a cascode buffer for its high gain, good isolation, and its stability. For our 65nm process the threshold voltage of an NMOS transistor is on the order of 350-400mV so we must use a 1:2 balun to reduce the required output swing of the buffer within this range.

The balun can be either a lateral design with both primary and secondary placed in the top metal layer, or a vertical design with the primary and secondary stacked on top of each other in the top two metal layers. While an aluminum capping layer (AP) is available, there is no Ultra-Thick Metal (UTM) layer. The top two copper metal layers are thicker than the bottom five and provide the highest quality factor. The AP layer has a similar thickness but higher manufacturing variation so it is generally only used together with the top metal layer, to reduce the overall sheet resistance, or for cross-overs. Since we are using a 1:2 balun cross-overs and/or cross-unders are necessary. A lateral design allows cross-unders to be placed in the lower of the two thick metal layers to take advantage of the lower sheet resistance. A vertical design, on the other hand, must either place cross-unders in a thinner metal or cross-overs in the AP layer. Furthermore, the lower of the two metal layers presents a larger capacitance to substrate, reducing the self-resonance frequency and the quality factor. In either case, the vertical design will have lower Q but will result in higher coupling coefficient than the lateral design. For this design we choose a lateral balun to maximize the quality factor.

The sizes of the transformer and the buffer are chosen using the two methods described in Sections 5.3.3 and 5.3.4: equation and optimization. The equation based method described in Section 5.3.3 requires a reasonable estimate of transformer parameters. This method also assumes that the model parameters are not a function of transformer size. The estimated model parameters for our 1:2 balun for this method are:

$$\begin{aligned} k &= 0.7 \\ n &= 1.6 \\ Q_p &= 12 \\ Q_s &= 12 \end{aligned}$$

First, these parameters are used in (5.19) to find the optimal L_s for each mixer size. Next, the optimal buffer size for each mixer is selected using (5.27). The optimum bias point for maximum power gain of this buffer is $V_g \approx 0.7V$ for the common source device and $V_g = V_{dd}$

for the cascode device. The scalable buffer model parameters (Fig. 5.7) at this bias point are extracted from simulation, including layout parasitics:

$$\begin{aligned}
 g_{mw} &= 1.3mS/\mu m \\
 g_{ow} &= 50\mu S/\mu m \\
 g_{iw} &= 160\mu S/\mu m \\
 C_{iw} &= 1.8fF/\mu m \\
 C_{ow} &= 1.25fF/\mu m \\
 I_{dw} &= 310\mu A/\mu m
 \end{aligned}$$

The second method is the optimization based method described in Section 5.3.4 which involves first simulating and extracting the transformer model parameters over a wide range of sizes. For our chosen balun structure the extracted model parameters are plotted in Fig. 5.11. For each mixer size, the optimum design point $[L_s, W_{buf}]$ is found which results in the lowest buffer power consumption.

The final step is the input matching network design which is the same for both methods. For this step we assume the input is matched to a 50Ω system impedance using a lumped L-match. Models of quality factor versus size for the individual components are extracted from measurements and EM simulation and used to give a reasonable estimate of the losses for the input matching network. The transducer gain of the entire buffer, including the matching network, is then calculated using (5.29) and used along with (5.34) to provide the minimum required input power. This is the minimum power that must be delivered to the buffer in order for the correct LO swing to be provided to the mixer.

The results of both methods over a wide range of mixer switch sizes are plotted together for comparison in Fig. 5.12. The optimization based method provides the optimum design point but, as we can see from Fig. 5.12, using the simple equation based method with a reasonable estimate for the transformer model leads to a near optimal design over a very wide range of mixer sizes with a significant reduction in effort.

5.3.6 Injection Locked Oscillator As an LO Buffer

Instead of a buffer we could potentially use a small local oscillator, injection locked to a central reference, to directly drive the mixer. In this case, the oscillator itself should be designed to drive the mixer directly since any extra buffering would clearly make the power consumption higher than a standalone buffer. We can start by designing a simple cross-coupled oscillator. For reliability we have to bias the mixer core to a lower voltage to reduce the risk of oxide breakdown. We can easily reduce the core supply to 0.6V while still maintaining sufficient output swing to drive the mixer in its high gain region. As before, the mixer input impedance consists of a differential capacitance, C_{mix} , and a differential resistance, R_{mix} . The ILO core will also load the tank with parasitic capacitance and the

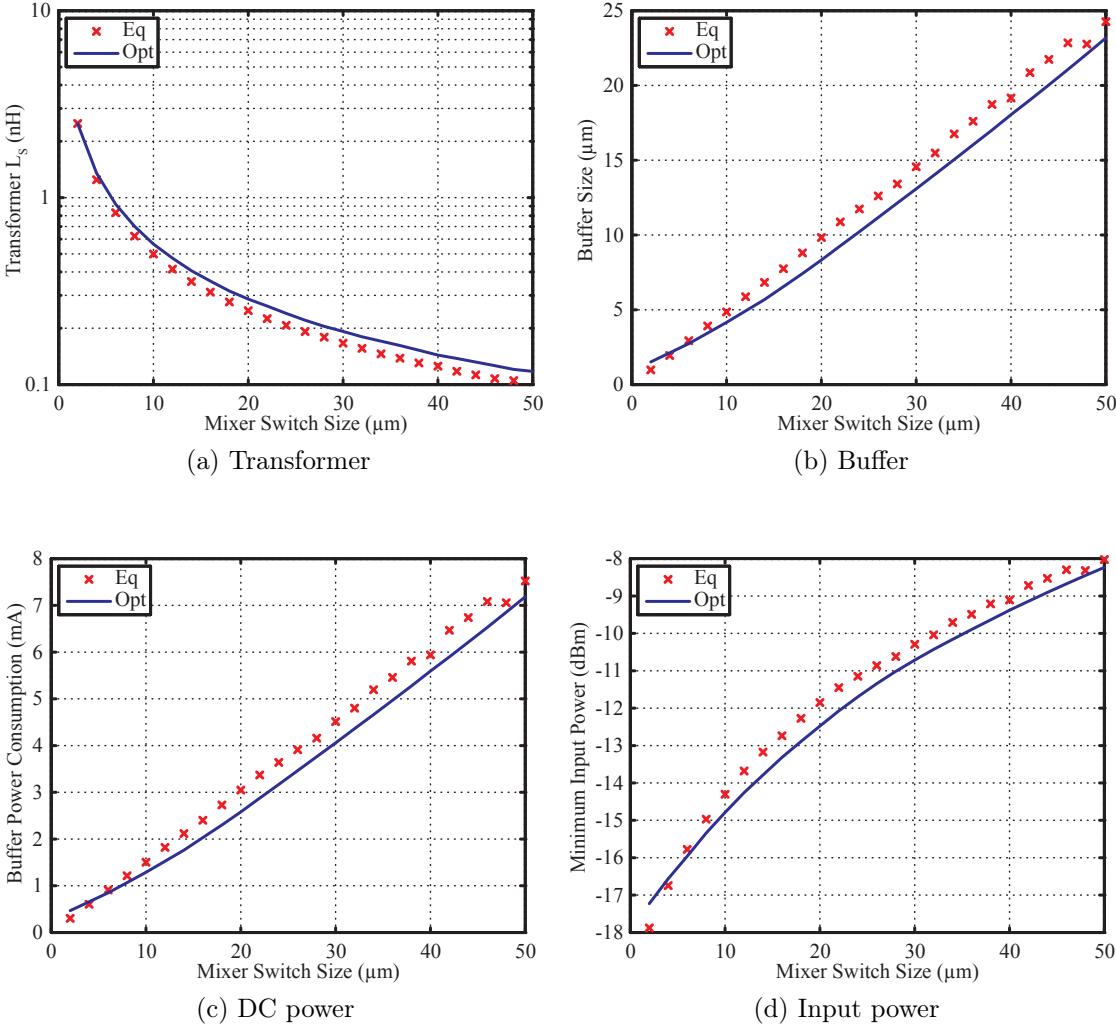


Figure 5.12: Optimum design of Mixer LO buffer versus mixer switch size.

device output resistance. As before, we will assume a very simple width-scalable model of a transistor as shown in Fig. 5.7. For a core size of W_{osc} , the total tank capacitance will be given by

$$C_T = C_{mix} + W_{osc} \frac{C_{iw} + C_{ow}}{2} \quad (5.35)$$

An inductor must then be chosen to resonate out this tank capacitance at 60GHz

$$L_T = \frac{1}{\omega_o^2 C_T} = \frac{1}{\omega_o^2 [C_{mix} + (C_{iw} + C_{ow}) W_{osc}/2]} \quad (5.36)$$

The core device input and output resistance will load the tank and the inductor will also have a finite quality factor, Q_L , which will generate a shunt resistance, R_p , in parallel with the tank

$$R_p = \omega_o L_T Q_L = \frac{Q_L}{\omega_o [C_{mix} + (C_{iw} + C_{ow}) W_{osc}/2]} \quad (5.37)$$

leading to a total tank resistance equal to

$$R_T = R_{mix} \left\| \left\| R_p \right\| \frac{2}{g_{ow} W_{osc}} \right\| \frac{2}{g_{iw} W_{osc}} \quad (5.38)$$

Finally, the core size must be chosen large enough such that the loop gain is greater than 1. For safety, the loop gain, A_l , is chosen much greater than 1 to ensure the oscillator starts up even with modeling errors or process variations.

$$A_l = \left(g_{mw} W_{osc} \frac{R_T}{2} \right)^2 \quad (5.39)$$

Putting everything together we can solve the above series of equations to give the minimum required core device size

$$W_{osc} = \frac{2\sqrt{A_l} \left(\frac{1}{R_{mix}} + \frac{\omega_o C_{mix}}{Q_L} \right)}{g_{mw} - \sqrt{A_l} \left[g_{ow} + g_{iw} + \frac{\omega_o (C_{ow} + C_{iw})}{Q_L} \right]} \quad (5.40)$$

Since there are two devices in the core, the total required bias current of the ILO is

$$I_{b,ILO} = 2W_{osc} I_{dw} \quad (5.41)$$

The total bias current required for an ILO is plotted versus mixer size in Fig. 5.13 for three different values of loop gain, A_l . The power required from an optimized buffer from the previous section is plotted as well for comparison.

In order to injection lock this oscillator to the central PLL, extra devices would be required to inject the locking signal into the tank. Furthermore, from (4.59) we know that in order to ensure robust injection locking over a wide band either the tank Q must be lowered or the injection signal strength must be increased. Another option is to tune the ILO to

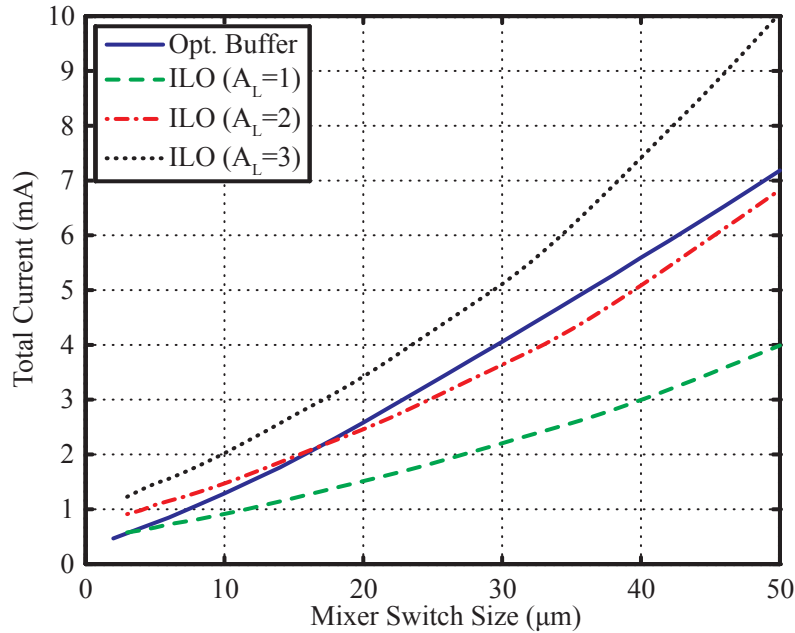


Figure 5.13: Current required for an ILO as a function of mixer switch size and loop gain (optimized buffer result added for comparison).

roughly track the central PLL frequency thus reducing the required injection lock range. The tank Q can be lowered artificially by adding a shunt resistance while tuning requires switched capacitors or varactors to be added to the tank. In fact, both techniques can be used as shown in the injection locked divider in Fig. 4.23. While (5.40) gives us the minimum required core sizing for a given loop gain, it does not take into account additional loading from the techniques described above and so (5.41) serves as an absolute minimum current consumption for the ILO.

From Fig. 5.13 we can see that for any reasonable safety factor the ILO consumes at least as much power as an optimized buffer. Again, this is a lower bound on ILO power consumption so the ILO would consume more power than an optimized buffer in a realistic scenario with a decent safety factor. The power consumption of the LO distribution subsystem is dominated by the power consumption of the block driving the mixer since it must be present at each element. This block should be an optimized buffer rather than an ILO since the ILO consumes more power and provides no benefits.

5.4 LO Distribution Strategy

The LO signal generated by the central PLL must now be distributed to each element in the array. We are designing the LO subsystem to support a quadrature direct conversion

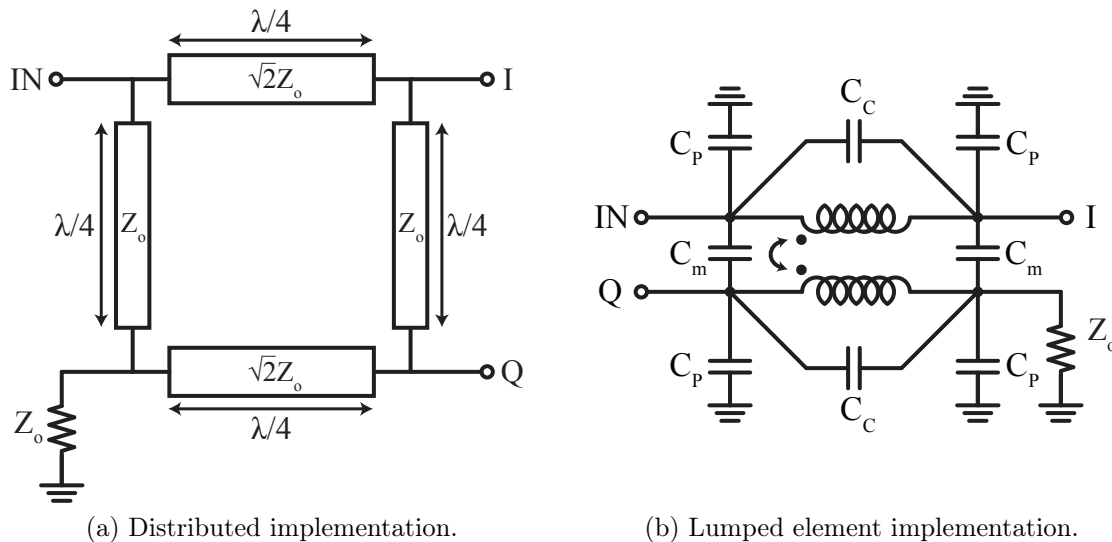


Figure 5.14: 90° Hybrids

transceiver with baseband phase shifting.⁹ Each element has a quadrature upconversion mixer for the transmitter (TX) and a quadrature downconversion mixer for the receiver (RX). If the TX and RX of each element share an antenna the LO signal can be shared. If different antenna arrays are used for transmitting and receiving, the RX and TX will likely not be collocated and a separate LO signal must be provided to each. Therefore, for an N -element array we must deliver $2N$ 60GHz LO signals with quadrature phases (i.e.: $I+$, $I-$, $Q+$, $Q-$). However, distributing and splitting a multi-phase signal requires large area and complicated routing and splitters.

To reduce the area, complexity, and power consumption of the distribution we should not distribute all four phases. Instead of generating the four phases centrally, we can generate them locally at each element. As we have already seen from the previous section, the conversion from single-ended to differential can easily be included in the design of the mixer LO buffer by using a balun transformer as part of its output matching network. Therefore, we only need two phases at each element, I and Q . Generating quadrature phases can be accomplished with a 90° hybrid made of either distributed [Marcu09] or lumped [Chin09] components. Both topologies, shown in Figures 5.14a and 5.14b respectively, offer similar performance in terms of loss from input to output and I/Q phase and amplitude imbalance, however, the lumped version occupies significantly smaller area making it a more attractive solution. Using these techniques means that only one phase must be provided to each element, simplifying splitting and distribution requirements.

Distribution of signals at low frequencies is straightforward since the interconnect can be

⁹The same techniques could be used for an LO phase shifting architecture with the addition of phase shifters. However, since LO phase shifting provides no real advantages, we will focus only on the baseband phase shifting architecture.

modeled as a capacitance. Longer interconnect lengths simply result in more capacitance. The scaling is predictable and well defined. As the frequency increases, the series resistance of the wires can also become important. Nevertheless, this scaling is also predictable and well defined. However, at mm-wave frequencies, the length of the interconnect can be a significant fraction of the signal wavelength and so distributed effects must be taken into account. At 60GHz, the on-chip wavelength is approximately 2.4mm which is on the order of the die size itself. This also means that transmission line lengths in the LO distribution network will be a significant fraction of the signal wavelength. Without a matched environment, the input impedance of the distribution network depends both on the load impedance as well as the physical layout of the entire distribution network. As the length of an unmatched transmission line is increased, its input impedance changes from capacitive to inductive and back. Therefore, an impedance matched environment is necessary in order to allow a priori prediction of the input impedance and enable the design of the driving amplifier. Without this, the entire radio would first have to be designed and the layout completed before the requirements on the LO generation could be determined, significantly increasing design time and complexity for the system as a whole. Also, an unmatched transmission line distribution gives rise to standing waves resulting in position dependent LO amplitude. Once again, this would mean that the layout would have to be predetermined and tap-off points selected a priori to ensure sufficient LO amplitude for each element, reducing design flexibility.

Furthermore, a signal distributed using a single wire does not have a well defined return path and therefore the high frequency performance cannot be predicted. A well defined return path is needed in order to allow modeling and prediction of losses and impedance levels. For this reason, transmission lines must be used for all signal routing at mm-wave frequencies. Another advantage of transmission line routing is a reduction in coupling between the signal being routed and adjacent structures. Due to the low loss and high isolation, a CPW structure is adopted.

For wire bonded chips, the elements of the phased array must be placed along the edges of the die. For flip-chip assemblies, the elements can be evenly distributed throughout the die. In either case, the LO signal must be distributed to each element with the same phase and amplitude, otherwise, a phase calibration mechanism must be used at each element. A tree structure (Fig. 5.15) can be used to equalize the distribution lengths and ensure that each element receives an identical LO signal, removing the need for phase calibration.

The distribution network is made up of transmission lines with characteristic impedance Z_o . In order to maintain a matched environment each signal split must be performed using a matched splitter. A 2-way Wilkinson splitter [Wilkinson60, Pozar04], shown in Fig. 5.16, splits power evenly between its two output ports (3dB ideal insertion loss) and also matching at all three ports and isolation between the outputs. This isolation is important in a phased array since any leakage between elements will limit the null depth of the array.

The spacing between elements is determined by the allowable pad spacing, as well as the physical size of the TX and RX circuits. Therefore, the only variable under our control is the system characteristic impedance. As we know from Chapter 2, the Z_o of on-chip

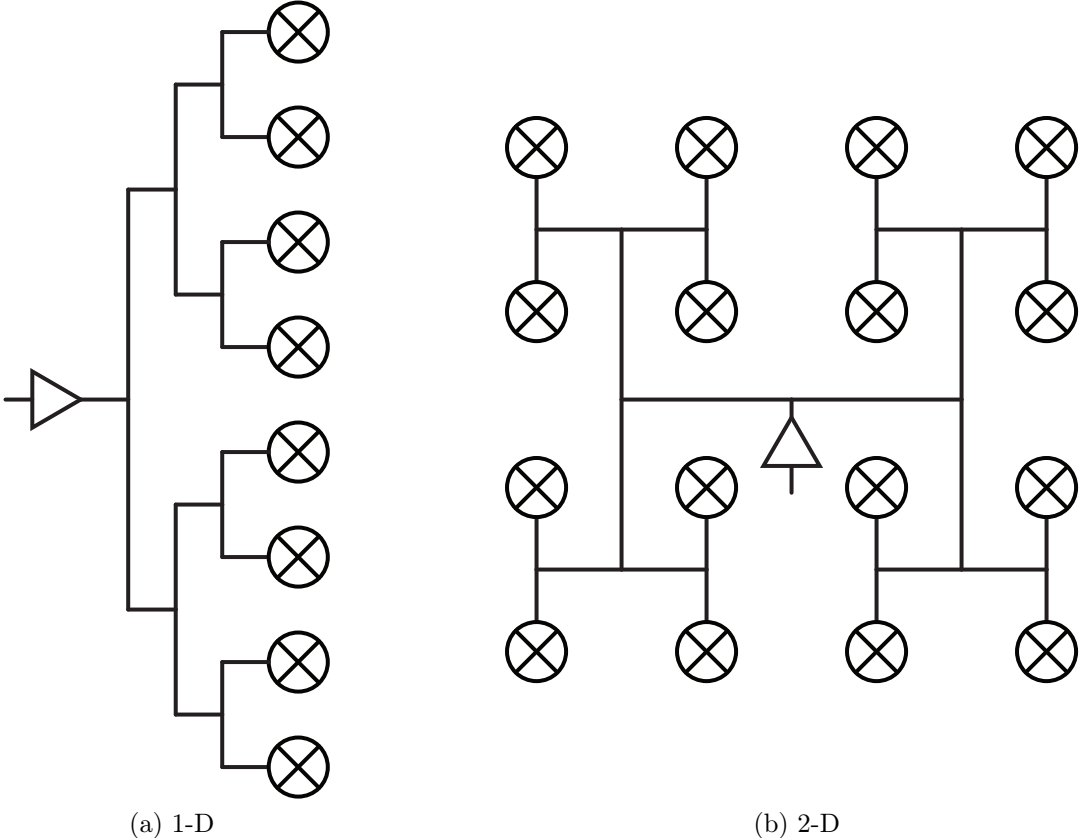


Figure 5.15: Tree distribution networks.

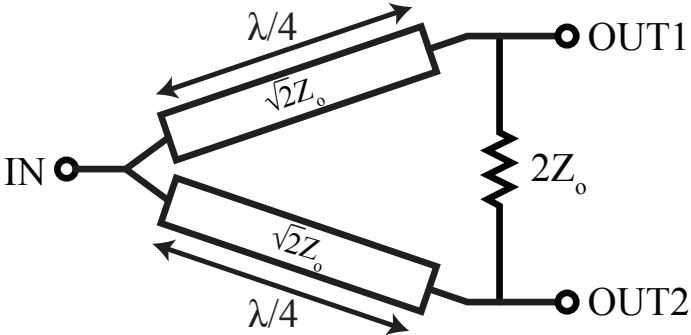


Figure 5.16: Wilkinson power splitter.

transmission lines is limited to a relatively narrow range without incurring significant loss. Since the 2-way Wilkinson power divider requires transmission lines with impedance equal to $\sqrt{2}Z_o$, the system impedance must be chosen carefully to reduce distribution losses.

For this optimization problem we will assume a linear array of N evenly spaced elements with pitch equal to d_{el} and a tree distribution network. To simplify the problem we will limit the number of elements to powers of 2 leading to $\log_2 N$ signal splits in each path plus distribution routing between splitters. Assuming the distribution dimensions are limited by the element spacing rather than the power splitters, the total routing length from the source to each element is equal to

$$l_r = d_{el} \frac{N - 1}{2} \quad (5.42)$$

The loss of the distribution transmission lines is then equal to

$$L_r = l_r \alpha_o \quad (5.43)$$

where α_o is the loss of a transmission line with characteristic impedance Z_o in dB/m. The loss of the Wilkinson splitter can be easily estimated since the length of the $\sqrt{2}Z_o$ arms must be equal to $\lambda/4$

$$l_W = \frac{\lambda_W}{4} = \frac{\pi}{2\beta_W} \quad (5.44)$$

where β_W is the propagation constant of a $\sqrt{2}Z_o$ transmission line in rad/m. The loss is then equal to 3dB (ideal splitting loss) plus the insertion loss due to the loss of the transmission lines making up the structure.

$$L_W = 3dB + l_W \alpha_W = 3dB + \frac{\pi \alpha_W}{2\beta_W} \quad (5.45)$$

where α_W is the loss of a $\sqrt{2}Z_o$ transmission line in dB/m. The total distribution loss is the sum of routing loss and $\log_2 N$ times the loss of a Wilkinson splitter.

$$L_{tot} = L_r + L_W \log_2 N = \alpha_o d_{el} \frac{N - 1}{2} + \left(3dB + \frac{\pi \alpha_W}{2\beta_W} \right) \log_2 N \quad (5.46)$$

Using the results from Chapter 2 (Fig. 2.25), we can then estimate the distribution losses as a function of N and d_{el} to find the optimum distribution Z_o . As an example, the loss as a function of Z_o for a 16 element array with $250\mu\text{m}$ pitch is plotted in Fig. 5.17. The optimal Z_o is equal to 55Ω . Below this value the Z_o distribution transmission line loss increases. Above this value, the $\sqrt{2}Z_o$ Wilkinson transmission line loss increases. However, there may be other considerations in selecting the system impedance. For example, if we want to be able to interface easily with external test equipment for debugging or characterization we should select a 50Ω system impedance. Luckily, from Fig. 5.17 we can see that this choice would only incur minimal additional loss.

This method is easily scalable to arrays with many more elements. However, as the loss of the distribution network increases, the required input power can become too high to be provided

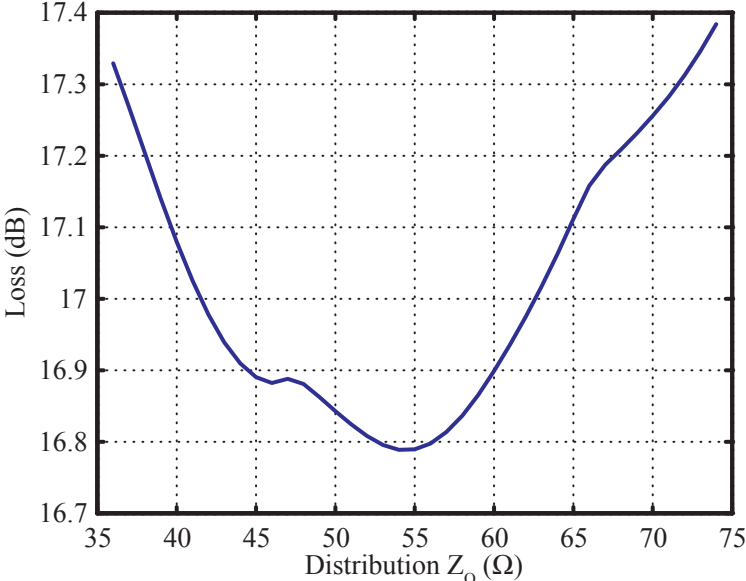


Figure 5.17: Loss of LO distribution network for a sample 16 element linear array with $250\mu\text{m}$ pitch.

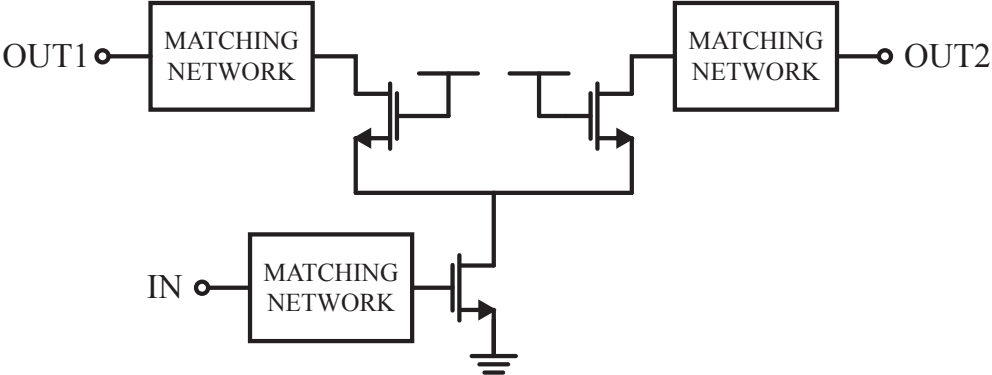


Figure 5.18: Active splitter.

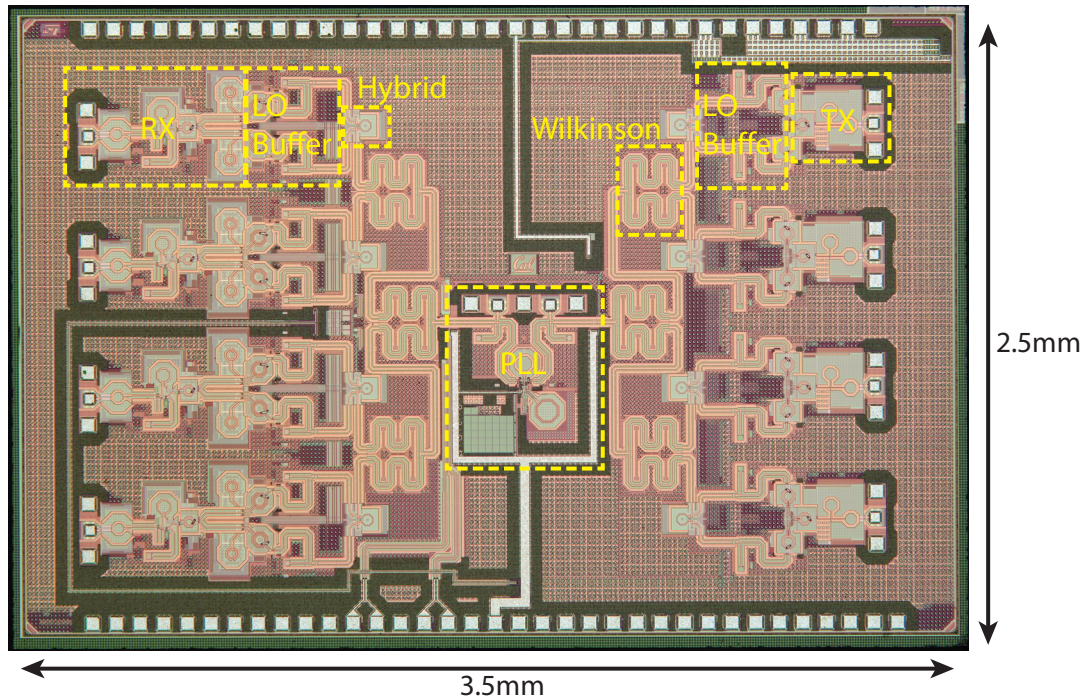


Figure 5.19: 4-element 60GHz phased array transceiver die photo.

with high efficiency. Therefore, additional buffering should be added as needed. In order to maintain low power consumption for the array, however, this buffering should be placed as close to the root of the distribution tree as possible. Another option is to replace one or more passive splitting stages with active splitters. A simplified schematic of such a splitter [Valdes-Garcia10a] is shown in Fig. 5.18. The common source transistor converts the input signal to a current which is then split equally into two cascode transistors. Each cascode transistor then has its own matching network to provide a match down to the distribution impedance. Also note, that these techniques could be used for RF signal distribution in an RF phased array.

5.5 Design Case Study

The above techniques were used to design an LO distribution network for a 4-element direct conversion phased array transceiver in a standard digital 65nm CMOS process [Tabesh11b]. The die photo of the transceiver is shown in Fig. 5.19. The LO is generated by the integer-N PLL presented in Section 4.4 which is placed in the center of the chip. The four RX elements are placed along the left edge with a pitch of $250\mu m$, while the four TX elements are positioned along the right edge with the same pitch. This placement allows each of the two single-ended LO signals generated by the PLL to drive the TX elements and RX

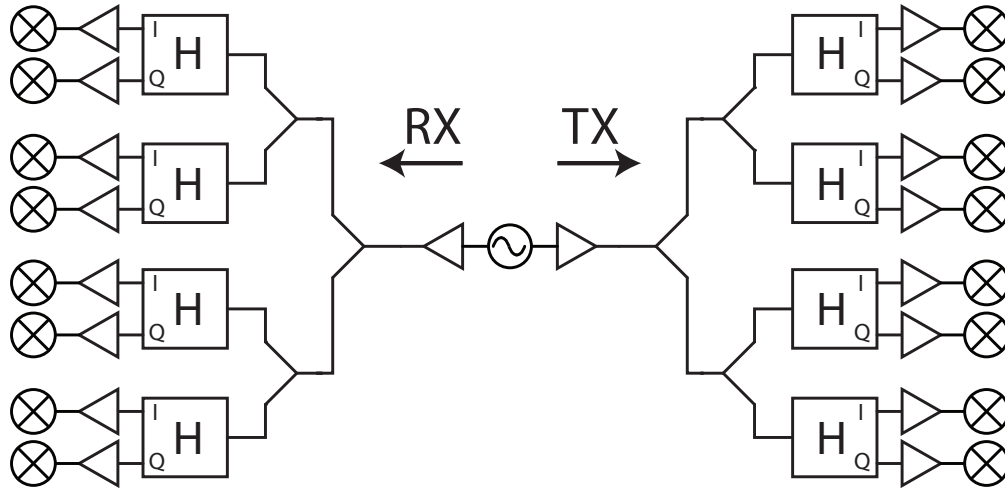


Figure 5.20: Phased array transceiver LO block diagram.

elements respectively using separate but identical distribution networks. The main goal of the LO distribution network design was minimizing overall power consumption.

A constant impedance LO distribution network using transmission lines and matched power splitters allowed arbitrary routing of the LO signal. However, to maintain phase and amplitude matching between elements, a fully balanced tree structure was utilized making all paths equal (Fig. 5.20). In-phase splitting is performed by Wilkinson dividers in two stages to deliver the single-ended LO to each of the 4 elements on either side. This is followed by local transformer-based hybrids which generate the quadrature LO for each element. The Wilkinson dividers (Fig. 5.21) utilize meandered 71Ω CPW transmission lines to reduce the area required as much as possible (Wilkinson area: $0.01\lambda^2$). The simulated insertion loss of this splitter is only 0.7dB. The hybrid on the other hand is a lumped transformer-based design which requires very little area by comparison (hybrid area: $0.002\lambda^2$) while still achieving only 0.7dB of insertion loss in simulation. Thanks to Maryam Tabesh for this hybrid design.

The detailed schematics of the entire LO path are shown in Fig. 5.22. At each mixer, a local buffer is designed to provide the required LO drive strength for high conversion gain. Each RX mixer is a single-balanced design with $20\mu m$ switches, while each TX mixer is a double-balanced design with $10\mu m$ switches. The input impedance of the mixers is thus roughly equal and both require approximately 700mV of differential LO amplitude for high conversion gain. The local LO buffer design proceeds as described in Section 5.3. Due to the large LO swing required at the mixers, a 1:2 transformer is used at the output of each LO buffer. This transformer also performs single-ended to differential conversion and provides impedance matching. Using the results in Fig. 5.12, the transformer secondary is set to approximately 300pH and the buffer size is chosen to be $10\mu m$, providing some margin of safety to ensure sufficient LO swing over process variations. Each buffer consumes 3mW from a 1.2V supply. By comparison, the total phased array power consumption is 34mW/element in either RX or TX mode, including LO generation and distribution.

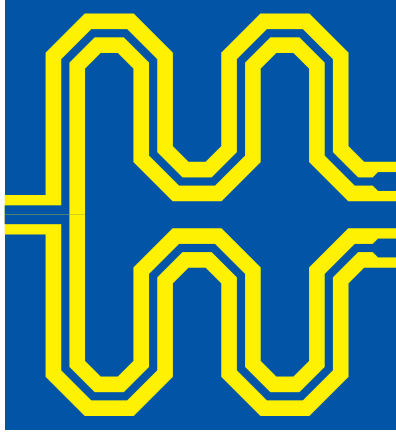


Figure 5.21: Compact 2-way Wilkinson power divider layout.

The center tap of the transformer is used to bias the mixer gates. The conversion from single-ended to differential, however, also creates a common-mode signal if there is a finite impedance on the center tap. A large bypass capacitor is usually placed at the center tap to provide a very small impedance to ground. However, the center tap has some inductance so instead of making this bypass capacitor as large as possible we size it to resonate out the center tap inductance at the LO frequency. This creates a series resonant network, minimizing the impedance to ground, and removing the common mode signal. The result is an efficient conversion from single-ended to differential and a very well balanced differential LO signal. A large resistor is used at the positive node of the capacitor to bias the mixer LO port through the center tap. Since there is nominally no gate current, this resistor can be made very large so as to not affect the low impedance created by the series resonant network.

Using the technique described in Section 5.4, the optimum distribution impedance is found to be approximately 55Ω . An impedance of 50Ω is selected since it provides nearly the same loss while allowing a better match to external testing equipment for debug and characterization. The input of each mixer LO buffer is matched to 50Ω using a single-stub transmission line matching network. To save area, 81Ω transmission lines are used. This is the highest impedance transmission line which still provides very low loss (Fig. 2.25). The minimum input power required for the buffer to provide the desired LO swing is -13dBm . The loss of the distribution network is approximately 12.5dB resulting in a required input power from the PLL of -0.5dBm . Since the PLL showed 0dBm of single-ended output power in simulation, no further buffering was deemed necessary. Despite the fact that measurements showed -1.8dBm of single-ended output power, the design included sufficient margin for proper operation. As discussed previously, for larger arrays or larger mixers, further buffering or active splitters would be required to increase the power delivered to each element.

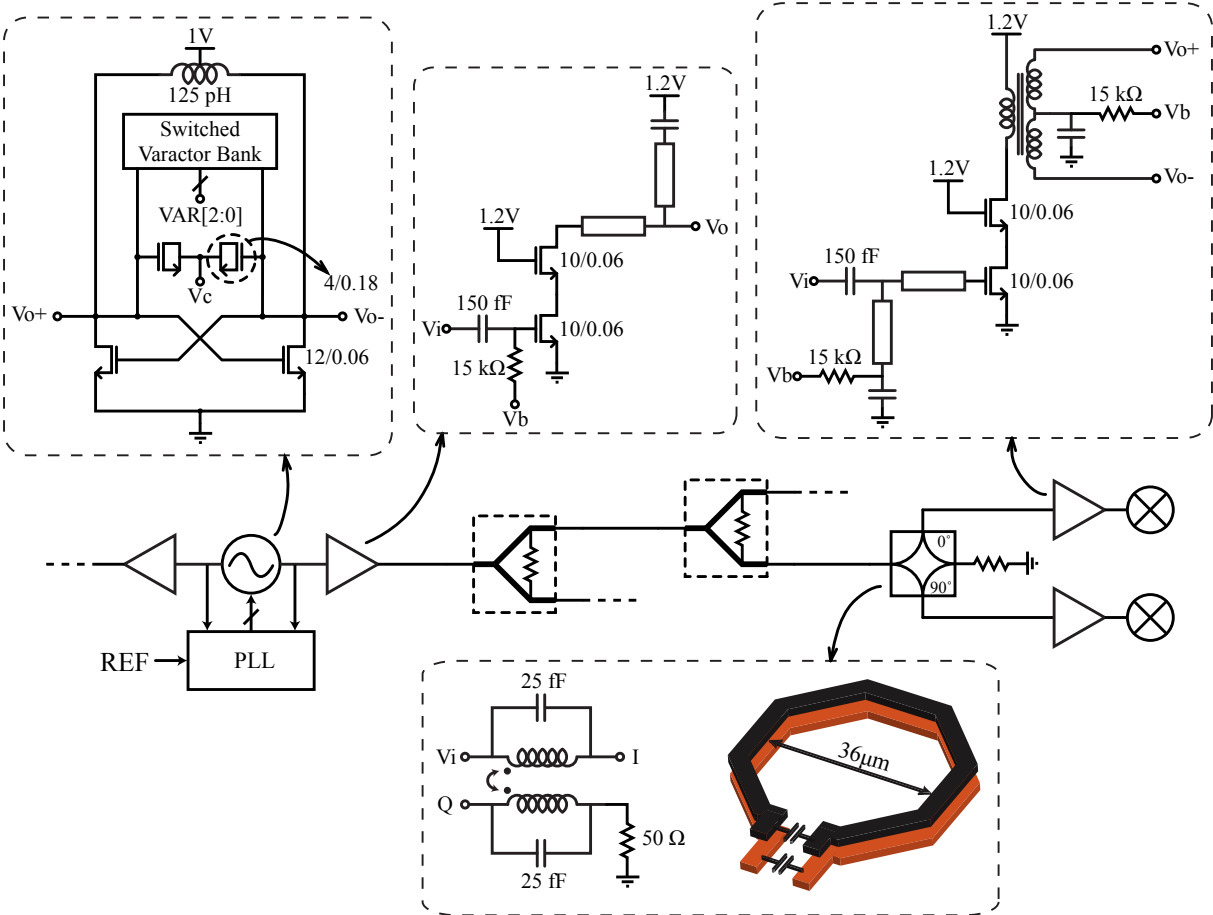


Figure 5.22: Schematic of VCO, LO buffers, and LO distribution chain including Wilkinson power splitters and transformer coupled lumped quadrature hybrid.

Chapter 6

Conclusion

Advances in technology and storage have led to increasing demands on wireless devices to achieve both high data rate communication and long battery life. The high bandwidth available in the unlicensed 60GHz band provides an excellent opportunity to meet these needs. Phased array transceivers can be used to break the fundamental bounds on performance of standard transceivers while also providing a simple means of beam steering. With advances in CMOS technology allowing high speed and low power baseband phase shifters and signal processing, IF phase shifting is becoming a more attractive phased array solution but very little attention has been paid to this architecture in the literature.

In this work we have identified the LO generation and distribution as one of the key bottlenecks of IF phased array design since the LO must be provided to every element in the array. Without an optimized design this subsystem could become the largest power consumer of the entire transceiver. In order to optimize the LO subsystem, we began by analyzing the trade-offs involved in LO generation. We presented methods of designing low power oscillators and also selecting between a fundamental or harmonic oscillator design based on tuning range and phase noise requirements. A push-push oscillator design in a 90nm CMOS process was presented which achieves the best FOM/FOM_T and the most efficient output power generation at the push-push port compared to previously reported multi-push oscillators. In order to complete the LO generation subsystem we then presented methods of optimizing PLL designs for low power consumption. This work resulted in a PLL design in a 65nm CMOS process which achieves record low power consumption while meeting phase noise and tuning range requirements for 60GHz radios.

Finally, we proposed LO distribution optimization methods for IF phased array transceivers. In order to provide the required LO swing for good conversion gain each mixer must have an LO buffer. This block was thus identified as the largest source of power consumption in the LO generation and distribution subsystem. Two methods were proposed for optimizing this block to minimize power consumption while providing the required LO swing to the mixer. While the comprehensive optimization method results in a globally optimal design, the simpler equation based method gives a nearly optimal design with significantly less effort.

Bibliography

- [Adler46] Robert Adler, "A Study of Locking Phenomena in Oscillators," *Proceedings of the IRE*, vol. 34, no. 6 pp. 351–357, Jun 1946.
- [Andreani99] Pietro Andreani and Sven Mattisson, "A 2.4-GHz CMOS Monolithic VCO Based on an MOS Varactor," *IEEE International Symposium on Circuits and Systems*, pp. 557–560, IEEE, 1999.
- [Andreani00] Pietro Andreani and Sven Mattisson, "On the Use of MOS Varactors in RF VCOs," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 6 pp. 905–910, Jun 2000.
- [Andreani05] Pietro Andreani, Xiaoyan Wang, Luca Vandi, and Ali Fard, "A Study of Phase Noise in Colpitts and LC-Tank CMOS Oscillators," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 5 pp. 1107–1118, May 2005.
- [Andress05] William F. Andress and Donhee Ham, "Standing Wave Oscillators Utilizing Wave-Adaptive Tapered Transmission Lines," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 3 pp. 638–651, Mar 2005.
- [Aoki02] Ichiro Aoki, Scott D. Kee, David B. Rutledge, and Ali Hajimiri, "Distributed Active Transformer-A New Power-Combining and Impedance-Transformation Technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 1 pp. 316–331, 2002.
- [Arora05] Himanshu Arora, Nikolaus Klemmer, James C. Morizio, and Patrick D. Wolf, "Enhanced Phase Noise Modeling of Fractional-N Frequency Synthesizers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 2 pp. 379–395, Feb 2005.
- [Babakhani06] Aydin Babakhani, Xiang Guan, Abbas Komijani, Arun S. Natarajan, and Ali Hajimiri, "A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Receiver and Antennas," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12 pp. 2795–2806, Dec 2006.

- [Bender83] John R. Bender and Colman Wong, "Push-Push Design Extends Bipolar Frequency Range," *Microwaves & RF*, vol. 22, no. 10 pp. 91–98, Oct 1983.
- [Borremans08] Jonathan Borremans, M. Dehan, Karen Scheir, M. Kuijk, and Piet Wambacq, "VCO Design for 60 GHz Applications Using Differential Shielded Inductors in 0.13 μm CMOS," *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 135–138, IEEE, Jun 2008.
- [Borremans09] Jonathan Borremans, Kuba Raczkowski, and Piet Wambacq, "A digitally controlled compact 57-to-66GHz front-end in 45nm digital CMOS," *IEEE International Solid-State Circuits Conference*, pp. 492–493,493a, IEEE, Feb 2009.
- [Brown71] J.I. Brown, "A Digital Phase and Frequency-Sensitive Detector," *Proceedings of the IEEE*, vol. 59, no. 4 pp. 717–718, 1971.
- [Cao03] Yu Cao, Robert A. Groves, Xuejue Huang, Noah D. Zamdmer, Jean-Olivier Plouchart, Richard A. Wachnik, Tsu-Jae King, and Chenming Hu, "Frequency-Independent Equivalent-Circuit Model for On-Chip Spiral Inductors," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3 pp. 419–426, Mar 2003.
- [Castello98] Rinaldo Castello, P. Erratico, S. Manzini, and Francesco Svelto, "A +/-30% Tuning Range Varactor Compatible with Future Scaled Technologies," *IEEE Symposium on VLSI Circuits*, pp. 34–35, IEEE, 1998.
- [Catli10] Burak Catli and Mona Mostafa Hella, "Triple-Push Operation for Combined Oscillation/Divison Functionality in Millimeter-Wave Frequency Synthesizers," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 8 pp. 1575–1589, Aug 2010.
- [Chin09] Ting-Yueh Chin, Jen-Chieh Wu, Sheng-Fuh Chang, and Chia-Chan Chang, "Compact S-/Ka-Band CMOS Quadrature Hybrids With High Phase Balance Based on Multilayer Transformer Over-Coupling Techn," *Transactions on Microwave Theory and Techniques*, vol. 57, no. 3 pp. 708–715, Mar 2009.
- [Chiu10] Hsien-Chin Chiu and Chih-Pin Kao, "A Wide Tuning Range 69 GHz Push-Push VCO Using 0.18 μm CMOS Technology," *IEEE Microwave and Wireless Components Letters*, vol. 20, no. 2 pp. 97–99, Feb 2010.
- [Cho05] Yi-Hsien Cho, Ming-Da Tsai, Hong-Yeh Chang, Chia-Chi Chang, and Huei Wang, "A Low Phase Noise 52-GHz Push-Push VCO in 0.18- μm Bulk CMOS Technologies," *IEEE Radio Frequency integrated Circuits Symposium*, pp. 131–134, IEEE, 2005.

- [Cohen10] Emanuel Cohen, Claudio Jakobson, Shmuel Ravid, and Dan Ritter, “A thirty two element phased-array transceiver at 60GHz with RF-IF conversion block in 90nm flip chip CMOS process,” *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 457–460, IEEE, 2010.
- [Cohn69] Seymour B. Cohn, “Slot Line on a Dielectric Substrate,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 17, no. 10 pp. 768–778, Oct 1969.
- [Collin00] Robert E. Collin, *Foundations for Microwave Engineering*, 2ns ed., John Wiley & Sons, Hoboken, NJ, 2000.
- [Colpitts27] Edwin H. Colpitts, *Oscillation Generator*, US Patent 1624537, 1927.
- [Copani10] Tino Copani, Hyungseok Kim, Bertan Bakkaloglu, and Sayfe Kiaei, “A 0.13-um CMOS Local Oscillator for 60-GHz Applications Based on Push-Push Characteristic of Capacitive Degeneration,” *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 153–156, IEEE, 2010.
- [Dauphinee97] Leonard Dauphinee, Miles A. Copeland, and Peter Schvan, “A Balanced 1.5GHz Voltage Controlled Oscillator with an Integrated LC Resonator,” *IEEE International Solids-State Circuits Conference*, pp. 390–391,491, IEEE, 1997.
- [Decanis11] Ugo Decanis, Andrea Ghilioni, Enrico Monaco, Andrea Mazzanti, and Francesco Svelto, “A mm-Wave Quadrature VCO Based on Magnetically Coupled Resonators,” *IEEE International Solid-State Circuits Conference*, pp. 280–282, IEEE, Feb 2011.
- [Deng10] Zhiming Deng and Ali M. Niknejad, “The Speed-Power Trade-Off in the Design of CMOS True-Single-Phase-Clock Dividers,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 11 pp. 2457–2465, Nov 2010.
- [Emami07] Sohrab Emami, Chinh H. Doan, Ali M. Niknejad, and Robert W. Brodersen, “A Highly Integrated 60GHz CMOS Front-End Receiver,” *IEEE International Solid-State Circuits Conference*, pp. 190–191, IEEE, Feb 2007.
- [Emami11] Sohrab Emami, Robert F Wiser, Ershad Ali, Mark G Forbes, Michael Q Gordon, Xiang Guan, Steve Lo, Patrick T McElwee, James Parker, Jon R Tani, Jeffery M Gilbert, and Chinh H Doan, “A 60GHz CMOS Phased-Array Transceiver Pair for Multi-Gb/s Wireless Communications,” *IEEE International Solid-State Circuits Conference*, pp. 164–166, IEEE, Feb 2011.

- [Ferndahl04] Mattias Ferndahl, Bahar M. Motlagh, and Herbert Zirath, "40 and 60 GHz Frequency Doublers in 90-nm CMOS," *IEEE MTT-S International Microwave Symposium*, pp. 179–182, IEEE, 2004.
- [Floyd06] Brian A. Floyd, Scott K. Reynolds, Ullrich R. Pfeiffer, Troy Beukema, Janusz Grzyb, and Chuck Haymes, "A silicon 60GHz receiver and transmitter chipset for broadband communications," *IEEE International Solid State Circuits Conference*, pp. 649–658, IEEE, 2006.
- [Friis44] Harald T. Friis, "Noise Figures of Radio Receivers," *Proceedings of the IRE*, vol. 32, no. 7 pp. 419–422, Jul 1944.
- [Friis46] Harald T. Friis, "A Note on a Simple Transmission Formula," *Proceedings of the IRE*, vol. 34, no. 5 pp. 254–256, May 1946.
- [Gardner80] Floyd M. Gardner, "Charge-Pump Phase-Lock Loops," *IEEE Transactions on Communications*, vol. 28, no. 11 pp. 1849–1858, Nov 1980.
- [Gardner05] Floyd Martin Gardner, *Phaselock Techniques*, 3rd ed., John Wiley & Sons, Hoboken, NJ, 2005.
- [Gonzalez97] Guillermo Gonzalez, *Microwave Transistor Amplifiers: Analysis and Design*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ, 1997.
- [Grieg52] D.D. Grieg and H.F. Engelmann, "Microstrip-A New Transmission Technique for the Kilomegacycle Range," *Proceedings of the IRE*, vol. 40, no. 12 pp. 1644–1650, Dec 1952.
- [Hajimiri98] Ali Hajimiri and Thomas H. Lee, "A General Theory of Phase Noise in Electrical Oscillators," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 2 pp. 179–194, 1998.
- [Hajimiri99] Ali Hajimiri and Thomas H. Lee, "Design Issues in CMOS Differential LC Oscillators," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 5 pp. 717–724, May 1999.
- [Ham01] Donhee Ham and Ali Hajimiri, "Concepts and Methods in Optimization of Integrated LC VCOs," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 6 pp. 896–909, Jun 2001.
- [Hashemi05] Hossein Hashemi, Xiang Guan, Abbas Komijani, and Ali Hajimiri, "A 24-GHz SiGe Phased-Array Receiver-LO Phase-Shifting Approach," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 2 pp. 614–626, Feb 2005.

- [Hegazi03] Emad Hegazi and Asad A. Abidi, "Varactor Characteristics, Oscillator Tuning Curves, and AM-FM Conversion," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 6 pp. 1033–1039, Jun 2003.
- [Huang98] Qiuting Huang, "On the Exact Design of RF Oscillators," *IEEE Custom Integrated Circuits Conference*, pp. 41–44, IEEE, 1998.
- [Huang08] Daquan Huang, Tim R. LaRocca, Mau-Chung Frank Chang, Lorene Samoska, Andy Fung, Richard L. Campbell, and Michael Andrews, "Terahertz CMOS Frequency Generator Using Linear Superposition Technique," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 12 pp. 2730–2738, Dec 2008.
- [IEE97] *802.11-1997 IEEE Standard for Information Technology- Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specific*, IEEE, 1997.
- [IEE09] *802.15.3c-2009 IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY)*, IEEE, 2009.
- [Ji-Ren87] Yuan Ji-Ren, Ingemar Karlsson, and Christer Svensson, "A True Single-Phase-Clock Dynamic CMOS Circuit Technique," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 5 pp. 899–901, Oct 1987.
- [Jimenez09] Jose Luis Gonzalez Jimenez, Franck Bade, Baudouin Martineau, and Didier Belot, "A 56GHz LC-tank VCO With 17% Tuning Range in 65nm Bulk CMOS for Wireless HDMI Applications," *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 481–484, IEEE, Jun 2009.
- [Kim05] Jaeha Kim, Jeong-Kyoum Kim, Bong-Joon Lee, Namhoon Kim, Deog-Kyoon Jeong, and Wonchan Kim, "A 20-GHz Phase-Locked Loop for 40Gb/s Serializing Transmitter in 0.13 μ m CMOS," *IEEE Symposium on VLSI Circuits*, pp. 144–147, IEEE, 2005.
- [Knorr75] Jeffrey B. Knorr and Klaus-Dieter Kuchler, "Analysis of Coupled Slots and Coplanar Strips on Dielectric Substrate," *IEEE Transactions on Microwave Theory and Techniques*, vol. 23, no. 7 pp. 541–548, Jul 1975.
- [Kobayashi99] Kevin W. Kobayashi, Aaron K. Oki, Liem T. Tran, John C. Cowles, Augusto Gutierrez-Aitken, Frank Yamada, Thomas R. Block, and Dwight C. Streit, "A 108-GHz InP-HBT Monolithic Push-Push VCO

- with Low Phase Noise and Wide Tuning Bandwidth,” *IEEE Journal of Solid-State Circuits*, vol. 34, no. 9 pp. 1225–1232, 1999.
- [Krauss80] Herbert L. Krauss, Charles W. Bostian, and Frederick H. Raab, *Solid-State Radio Engineering*, John Wiley & Sons, New York, NY, 1980.
- [LaRocca09] Tim R. LaRocca, Jenny Liu, Frank Wang, David Murphy, and Mau-Chung Frank Chang, “CMOS Digital Controlled Oscillator with Embedded DiCAD Resonator for 58-64GHz Linear Frequency Tuning and Low Phase Noise,” *IEEE MTT-S International Microwave Symposium*, pp. 685–688, IEEE, Jun 2009.
- [Larson00] John D. Larson, Paul D. Bradley, Scott Wartenberg, and Richard C. Ruby, “Modified Butterworth-Van Dyke Circuit for FBAR Resonators and Automated Measurement System,” *IEEE Ultrasonics Symposium*, pp. 863–868, IEEE, 2000.
- [Lee00] Thomas H. Lee and Ali Hajimiri, “Oscillator Phase Noise: A Tutorial,” *IEEE Journal of Solid-State Circuits*, vol. 35, no. 3 pp. 326–336, Mar 2000.
- [Lee04a] Jri Lee and Behzad Razavi, “A 40-GHz Frequency Divider in 0.18-um CMOS Technology,” *IEEE Journal of Solid-State Circuits*, vol. 39, no. 4 pp. 594–601, Apr 2004.
- [Lee04b] Thomas H. Lee, *The Design of Radio-Frequency Integrated Circuits*, 2nd ed., Cambridge University Press, New York, NY, 2004.
- [Leenov59] D. Leenov and A. Uhler, “Generation of Harmonics and Subharmonics at Microwave Frequencies with P-N Junction Diodes,” *Proceedings of the IRE*, vol. 47, no. 10 pp. 1724–1729, Oct 1959.
- [Leeson66] D.B. Leeson, “A Simple Model of Feedback Oscillator Noise Spectrum,” *Proceedings of the IEEE*, vol. 54, no. 2 pp. 329–330, 1966.
- [Li09] Lianming Li, Patrick Reynaert, and Michiel S.J. Steyaert, “A Low Power mm-Wave Oscillator Using Power Matching Techniques,” *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 469–472, IEEE, Jun 2009.
- [Liebe81] Hans J. Liebe, “Modeling attenuation and phase of radio waves in air at frequencies below 1000 GHz,” *Radio Science*, vol. 16, no. 6 pp. 1183–1199, 1981.
- [Liu04] Ren-Chieh Liu, Hong-Yeh Chang, Chi-Hsueh Wang, and Huei Wang, “A 63 GHz VCO Using a Standard 0.25 μm CMOS Process,” *IEEE International Solid-State Circuits Conference*, pp. 446–447, IEEE, 2004.

- [Lu97] Ke Lu, "An Efficient Method for Analysis of Arbitrary Nonuniform Transmission Lines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 45, no. 1 pp. 9–14, 1997.
- [Manku99] Tajinder Manku, "Microwave CMOS-Device Physics and Design," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 3 pp. 277–285, Mar 1999.
- [Marcu08a] Cristian Marcu and Ali M. Niknejad, *60 GHz Tapered Transmission Line Resonators*, Master of science, University of California, Berkeley, 2008.
- [Marcu08b] Cristian Marcu and Ali M. Niknejad, "A 60GHz High-Q Tapered Transmission Line Resonator in 90nm CMOS," *IEEE MTT-S International Microwave Symposium*, pp. 775–778, IEEE, Jun 2008.
- [Marcu09] Cristian Marcu, Debopriyo Chowdhury, Chintan Thakkar, Jung-Dong Park, Ling-Kai Kong, Maryam Tabesh, Yanjie Wang, Bagher Afshar, Abhinav Gupta, Amin Arbabian, Simone Gambini, Reza Zamani, Elad Alon, and Ali M. Niknejad, "A 90nm CMOS Low-Power 60GHz Transceiver With Integrated Baseband Circuitry," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 12 pp. 3434–3447, Dec 2009.
- [Mazzanti08] Andrea Mazzanti and Pietro Andreani, "Class-C Harmonic CMOS VCOs, With a General Result on Phase Noise," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 12 pp. 2716–2729, Dec 2008.
- [Miller39] R.L. Miller, "Fractional-Frequency Generators Utilizing Regenerative Modulation," *Proceedings of the IRE*, vol. 27, no. 7 pp. 446–457, Jul 1939.
- [Murphy10] David Murphy, Qun Jane Gu, Yi-Cheng Wu, Heng-Yu Jian, Zhiwei Xu, Adrian Tang, Frank Wang, Yu-Ling Lin, Ho-Hsiang Chen, Chewnpu Jou, and Mau-Chung Frank Chang, "A Low Phase Noise, Wideband and Compact CMOS PLL for Use in a Heterodyne 802.15.3c TRX," *European Solid-State Circuits Conference*, pp. 258–261, IEEE, Sep 2010.
- [Natarajan06] Arun S. Natarajan, Abbas Komijani, Xiang Guan, Aydin Babakhani, and Ali Hajimiri, "A 77-GHz Phased-Array Transceiver With On-Chip Antennas in Silicon: Transmitter and Local LO-Path Phase Shifting," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12 pp. 2807–2819, Dec 2006.
- [Natarajan07] Arun S. Natarajan, Brian A. Floyd, and Ali Hajimiri, "A Bidirectional RF-Combining 60GHz Phased-Array Front-End," *IEEE International Solid-State Circuits Conference*, pp. 202–597, IEEE, Feb 2007.

- [Navarro Soares99] J. Navarro Soares and W.A.M. Van Noije, "A 1.6-GHz Dual Modulus Prescaler Using the Extended True-Single-Phase-Clock CMOS Circuit Technique (E-TSPC)," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 1 pp. 97–102, 1999.
- [Nguyen07] Clark T.-C. Nguyen, "MEMS Technology for Timing and Frequency Control," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 54, no. 2 pp. 251–270, Feb 2007.
- [Niknejad07] Ali M. Niknejad, *Electromagnetics for High-Speed Analog and Digital Communication Circuits*, Cambridge University Press, New York, NY, 2007.
- [Parvais10] B. Parvais, Karen Scheir, V. Vidojkovic, R. Vandebriel, G. Vandersteen, C. Soens, and Piet Wambacq, "A 40 nm LP CMOS PLL for high-speed mm-wave communication," *European Solid-State Circuits Conference*, pp. 254–257, IEEE, Sep 2010.
- [Pierret96] Robert F. Pierret, *Semiconductor Device Fundamentals*, Addison-Wesley, Reading, MA, 1996.
- [Pinel08] Stephane Pinel, Saikat Sarkar, Padmanava Sen, Bevin Perumana, David Yeh, Debasis Dawn, and Joy Laskar, "A 90nm CMOS 60GHz Radio," *IEEE International Solid-State Circuits Conference*, pp. 130–131, 601, IEEE, Feb 2008.
- [Ponchak05] George E. Ponchak, John Papapolymerou, and Manos M. Tentzeris, "Excitation of Coupled Slotline Mode in Finite-Ground CPW With Unequal Ground-Plane Widths," *IEEE Transactions on Microwave Theory and Techniques*, vol. 53, no. 2 pp. 713–717, Feb 2005.
- [Pozar04] David M. Pozar, *Microwave Engineering*, 3rd ed., John Wiley & Sons, Hoboken, NJ, 2004.
- [Pro57] "IRE Standards on Piezoelectric Crystals-The Piezoelectric Vibrator: Definitions and Methods of Measurement, 1957," *Proceedings of the IRE*, vol. 45, no. 3 pp. 353–358, Mar 1957.
- [Raczkowski10] Kuba Raczkowski, Walter De Raedt, Bart Nauwelaers, and Piet Wambacq, "A wideband beamformer for a phased-array 60GHz receiver in 40nm digital CMOS," *IEEE International Solid-State Circuits Conference*, pp. 40–41, IEEE, Feb 2010.
- [Rael00] J.J. Rael and Asad A. Abidi, "Physical Processes of Phase Noise in Differential LC Oscillators," *IEEE Custom Integrated Circuits Conference*, pp. 569–572, IEEE, 2000.

- [Rategh99] Hamid R. Rategh and Thomas H. Lee, "Superharmonic Injection-Locked Frequency Dividers," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 6 pp. 813–821, Jun 1999.
- [Razavi94] Behzad Razavi, Ran-Hong Yan, and Kwing F. Lee, "Impact of Distributed Gate Resistance on the Performance of MOS Devices," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 41, no. 11 pp. 750–754, 1994.
- [Razavi98] Behzad Razavi, *RF Microelectronics*, Prentice Hall, Upper Saddle River, NJ, 1998.
- [Razavi11] Behzad Razavi, "A 300-GHz Fundamental Oscillator in 65-nm CMOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 4 pp. 894–903, Apr 2011.
- [Reynolds10] Scott K. Reynolds, Arun S. Natarajan, Ming-Da Tsai, Sean T. Nicolson, Jing-Hong Conan Zhan, Duixian Liu, Dong G. Kam, Oscar Huang, Alberto Valdes-Garcia, and Brian A. Floyd, "A 16-element phased-array receiver IC for 60-GHz communications in SiGe BiCMOS," *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 461–464, IEEE, May 2010.
- [Riaziat86] Majid Riaziat, Irene Zubeck, Steve Bandy, and George Zdasiuk, "Coplanar Waveguides Used in 2-18 GHz Distributed Amplifier," *IEEE MTT-S International Microwave Symposium*, pp. 337–338, IEEE, 1986.
- [Rogers00] John W.M. Rogers, Jose A. Macedo, and Calvin Plett, "The Effect of Varactor Nonlinearity on the Phase Noise of Completely Integrated VCOs," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 9 pp. 1360–1367, 2000.
- [Ruby01] Richard C. Ruby, Paul D. Bradley, Yury Oshmyansky, Allen Chien, and John D. Larson, "Thin Film Bulk Wave Acoustic Resonators (FBAR) for Wireless Applications," *IEEE Ultrasonics Symposium*, pp. 813–821, IEEE, 2001.
- [Schlesinger45] Kurt Schlesinger, "Cathode-Follower Circuits," *Proceedings of the IRE*, vol. 33, no. 12 pp. 843–855, Dec 1945.
- [Shaeffer97] Derek K. Shaeffer and Thomas H. Lee, "A 1.5-V, 1.5-GHz CMOS Low Noise Amplifier," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 5 pp. 745–759, May 1997.
- [Smith82] Ernest K. Smith, "Centimeter and millimeter wave attenuation and brightness temperature due to atmospheric oxygen and water vapor," *Radio Science*, vol. 17, no. 6 pp. 1455–1464, 1982.

- [Smith89] D.M. Smith, J.C. Canyon, and D.L. Tait, "25-42 GHz GaAs Heterojunction Bipolar Transistor Low Phase Noise Push-Push VCOs," *IEEE MTT-S International Microwave Symposium*, pp. 725–728, IEEE, 1989.
- [Soorapanth98] Theerachet Soorapanth, C. Patrick Yue, Derek K. Shaeffer, Thomas H. Lee, and S. Simon Wong, "Analysis and optimization of accumulation-mode varactor for RF ICs," *IEEE Symposium on VLSI Circuits*, pp. 32–33, IEEE, 1998.
- [Tabesh11a] Maryam Tabesh, Jiashu Chen, Cristian Marcu, Ling-Kai Kong, Shinwon Kang, Elad Alon, and Ali M. Niknejad, "A 65nm CMOS 4-Element Sub-34mW/Element 60GHz Phased-Array Transceiver," *IEEE International Solid-State Circuits Conference*, pp. 166–167,167b, IEEE, 2011.
- [Tabesh11b] Maryam Tabesh, Jiashu Chen, Cristian Marcu, Ling-Kai Kong, Shinwon Kang, Elad Alon, and Ali M. Niknejad, "A 65nm CMOS 4-Element Sub-34mW/Element 60GHz Phased-Array Transceiver," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 12, 2011.
- [Tanomura08] Masahiro Tanomura, Yasuhiro Hamada, Shuya Kishimoto, Masaharu Ito, Naoyuki Orihashi, Kenichi Maruhashi, and Hidenori Shimawaki, "TX and RX Front-Ends for 60GHz Band in 90nm Standard Bulk CMOS," *IEEE International Solid-State Circuits Conference*, pp. 558–635, IEEE, Feb 2008.
- [Temporiti04] Enrico Temporiti, Guido Albasini, Ivan Bietti, Rinaldo Castello, and Matteo Colombo, "A 700-kHz Bandwidth SigmaDelta Fractional Synthesizer with Spurs Compensation and Linearization Techniques for WCDMA Applications," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9 pp. 1446–1454, Sep 2004.
- [Tomkins09] Alexander Tomkins, Ricardo Andres Aroca, Takuji Yamamoto, Sean T. Nicolson, Yoshiyasu Doi, and Sorin P. Voinigescu, "A Zero-IF 60 GHz 65 nm CMOS Transceiver With Direct BPSK Modulation Demonstrating up to 6 Gb/s Data Rates Over a 2 m Wireless Link," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 8 pp. 2085–2099, Aug 2009.
- [Tse05] David Tse and Pramod Viswanath, *Fundamentals of Wireless Communications*, Cambridge University Press, Cambridge, UK, 2005.
- [Valdes-Garcia10a] Alberto Valdes-Garcia, Sean T. Nicolson, Jie-Wei Lai, Arun S. Natarajan, Ping-Yu Chen, Scott K. Reynolds, Jing-Hong Conan Zhan, and Brian A. Floyd, "A SiGe BiCMOS 16-element phased-array transmitter for 60GHz communications," *IEEE International Solid-State Circuits Conference*, pp. 218–219, IEEE, Feb 2010.

- [Valdes-Garcia10b] Alberto Valdes-Garcia, Sean T. Nicolson, Jie-Wei Lai, Arun S. Natarajan, Ping-Yu Chen, Scott K. Reynolds, Jing-Hong Conan Zhan, Dong G. Kam, Duixian Liu, and Brian A. Floyd, "A Fully Integrated 16-Element Phased-Array Transmitter in SiGe BiCMOS for 60-GHz Communications," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12 pp. 2757–2773, Dec 2010.
- [vanderZiel70] Albert van der Ziel, "Noise in Solid-State Devices and Lasers," *Proceedings of the IEEE*, vol. 58, no. 8 pp. 1178–1206, 1970.
- [Vaucher98] Cicero S. Vaucher and Dieter Kasperkovitz, "A Wide-Band Tuning System for Fully Integrated Satellite Receivers," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 7 pp. 987–997, Jul 1998.
- [Vaucher00a] Cicero S. Vaucher, "An Adaptive PLL Tuning System Architecture Combining High Spectral Purity and Fast Settling Time," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 4 pp. 490–502, Apr 2000.
- [Vaucher00b] Cicero S. Vaucher, Igor Ferencic, Marrhias Locher, Sebastian Sedvallson, Urs Voegeli, and Zhenhua Wang, "A family of low-power truly modular programmable dividers in standard 0.35- μm CMOS technology," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 7 pp. 1039–1045, Jul 2000.
- [Voinigescu00] Sorin P. Voinigescu and Miles A. Copeland, "A Family of Monolithic Inductor-Varactor SiGe-HBT VCOs for 20GHz to 30GHz LMDS and Fiber-Optic Receiver Applications," *IEEE Radio Frequency Integrated Circuits Symposium*, pp. 173–176, 2000.
- [Wang07] Chi-Hsueh Wang, Hong-Yeh Chang, Pei-Si Wu, Kun-You Lin, Tian-Wei Huang, Huei Wang, and Chun Hsiung Chen, "A 60GHz Low-Power Six-Port Transceiver for Gigabit Software-Defined Transceiver Applications," *IEEE International Solid-State Circuits Conference*, pp. 192–596, IEEE, Feb 2007.
- [Wen69] Cheng P. Wen, "Coplanar Waveguide: A Surface Strip Transmission Line Suitable for Nonreciprocal Gyromagnetic Device Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 17, no. 12 pp. 1087–1090, Dec 1969.
- [Wilkinson60] Ernest J. Wilkinson, "An N-Way Hybrid Power Divider," *IRE Transactions on Microwave Theory and Techniques*, vol. 8, no. 1 pp. 116–118, Jan 1960.
- [Womack62] Charles P. Womack, "The Use of Exponential Transmission Lines In Microwave Components," *IEEE Transactions on Microwave Theory and Techniques*, vol. 10, no. 2 pp. 124–132, Mar 1962.

- [Wu09] Chung-Yu Wu, Min-Chiao Chen, and Yi-Kai Lo, “A Phase-Locked Loop With Injection-Locked Frequency Multiplier in 0.18-um CMOS for V-Band Applications,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 7 pp. 1629–1636, Jul 2009.
- [Youla64] D.C. Youla, “Analysis and Synthesis of Arbitrarily Terminated Lossless Nonuniform Lines,” *IEEE Transactions on Circuit Theory*, vol. 11, no. 3 pp. 363–371, 1964.
- [Yue00] C. Patrick Yue and S. Simon Wong, “Physical Modeling of Spiral Inductors on Silicon,” *IEEE Transactions on Electron Devices*, vol. 47, no. 3 pp. 560–568, Mar 2000.
- [Zhang09] Ning Zhang and Kenneth K. O, “CMOS Frequency Generation System for W-Band Radars,” *IEEE Symposium on VLSI Circuits*, pp. 126–127, 2009.