Title
RF Front-End Circuits for FDD Radios: Duplexers and Receivers

Author
Shi, Kejian

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

RF Front-End Circuits for FDD Radios: Duplexers and Receivers

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical and Computer Engineering

by

Kejian Shi

2020

ABSTRACT OF THE DISSERTATION

RF Front-End Circuits for FDD Radios: Duplexers and Receivers

by

Kejian Shi
Doctor of Philosophy in Electrical and Computer Engineering
University of California, Los Angeles, 2020
Professor Asad Ali Abidi, Chair

Frequency-division duplexing (FDD) radios are widely used in wireless communication standards and multi-radio coexistence. It requires that the duplexer prevent both TX signal and noise from entering the receiver front-end, while operating under full TX power. It also requires the receiver front-end circuits to be highly linear when the TX leakage channel is adjacent to the receive channel.

This work describes the design of a dual-band electrical balance duplexer (EBD). The electrical balance duplexer supports dual-band TX-RX isolation for FDD operation at 5-7GHz. A network in the EBD can balance the antenna impedance (Z_{ANT}) in the TX channel and RX channel independently. The EBD provides >40dB isolation in an 80MHz channel bandwidth in the TX band (5-6GHz), for any $Z_{ANT}(f_{TX})$ with VSWR$\leq$2, and independently in the RX band (6-7GHz) when $Q_{ANT} \leq 4.3$. The EBD is designed for $\leq$4dB RX IL and $\leq$3.8dB TX IL.

Second, the properties of finite-Q LCR networks are derived. Simple results and derivations provide performance limits and design guidelines for the on-chip filter (two-port) and impedance synthesizer (one-port). General properties are given for the design of multi-band on-chip N-port LCR networks.

Third, second-order baseband filter circuits for mixer-first receivers are de-

signed for adjacent channel blocker tolerance. The design guidelines are given in simple expressions. The independent poles control by circuit elements is achieved. The prototype measurements verify the design. It can tolerate the maximum TX leakage from the dual-band EBD without gain compression and noise degradation.

Finally, the mixer switch nonlinearity is analyzed. Expressions for IIP3 and B1dB are given, showing the impact of supply voltage, FET size, clock waveform, operating frequency, and baseband load. The analysis matches the simulation very well.

The dual-band EBD and the mixer-first receiver with second-order baseband filters are fabricated in 65nm RF-SOI CMOS technology. They support the FDD operation for Wi-Fi 6/6E application between 5-7GHz.

The dissertation of Kejian Shi is approved.

Aydin Babakhani

Sudhakar Pamarti

Yuanxun Ethan Wang

Asad Ali Abidi, Committee Chair

University of California, Los Angeles

2020

To my parents . . .

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It has been a great journey for me to be a Ph.D. student under the advisory of Prof. Asad Abidi, and I would like to express my gratitude to him. His unique research perspective is what really attracts me: always trying to explain complicated phenomena using simple analysis that helps most engineers. He is a scientist that solves engineering problems. I really appreciate that he gives sufficient freedom for me to explore and choose my Ph.D. topic. I am very grateful for his insightful research guidance and strict standards, which I benefit a lot. He is without question a teacher and mentor for me.

I would like to thank Dr. Hooman Darabi, my supervisor at Broadcom Inc. that provided the essential support for my Ph.D. project. Without his support, the project cannot be conducted as smoothly.

I would also like to thank my committee members Prof. Sudhakar Pamarti, Prof. Yuanxun Ethan Wang, and Prof. Aydin Babakhani for valuable comments, questions, and feedback.

I appreciate the support, both academic and nonacademic, and inspiring discussions with students and researchers at UCLA: Hao Xu, Weiyu Leng, Jiyue Yang, Xiao Wang, Yu Zhao, Mahmoud Elhebeary, Liangxiao Tang, Shi Bu, Jiacheng Pan and other people. Without these friends, the Ph.D. journey could be very hard.

I am grateful to Dihang Yang, Milad Darvish, Edward Roth, Ali Afsahi, Guy Geshvindman, Hao Wu, Mohyee Mikhemar, David Murphy, Ken Wong, Rick Than, Richard Chen, Akshay Mahajan and other prople at Broadcom Inc. for their help during the design, tapeout and measurement of my chip, especially Milad, who helped a lot with the first tapeout in my life. I would like to express my special thanks to Dihang who gives helpful advice on my life and career.

At the end I would like to thank my parents and girlfriend for their continuous support, as always. My parents have offered me opportunities that they could not have when they were at my age. Finally, the dream is accomplished and this dissertation is dedicated to my parents.

VITA

2011–2015	B.S. in Physics, Peking University
2015–2017	M.S. in Electrical Engineering, University of California, Los Angeles
2018–2020	Ph.D. (Candidate) in Electrical and Computer Engineering, University of California, Los Angeles

CHAPTER 1

Introduction

1.1 Frequency-Division Duplexing

Nowadays mobile wireless communication applications mainly use two methods for duplexing: time-division duplex (TDD) and frequency-division duplex (FDD). The main purpose of duplexing is to reduce the number of antennas to save cost and space in mobile devices. In both TDD and FDD, the transmitter (TX) and receiver (RX) share the same antenna. In TDD, the TX and RX operate at different times and the antenna is connected either to the TX or the RX. While in FDD, the TX and RX operate at the same time, but at different frequencies.

There are many FDD applications. First, many standards are designed to use FDD, such as 3G CDMA, 4G LTE, and 5G NR. Second, there is another scenario that may use FDD, which is the so-called multi-radio coexistence. In multi-radio coexistence scenario, the TX and RX for different radio standards operate at the same time and different frequencies, while each standard may originally use TDD. Shared antenna or dedicated antennas may be used for different radio standards. When the TX and RX share the same antenna, it is FDD. An example of multi-radio coexistence includes transmitting Wi-Fi and receiving Bluetooth signals, or transmitting 5G NR and receiving Wi-Fi signals.

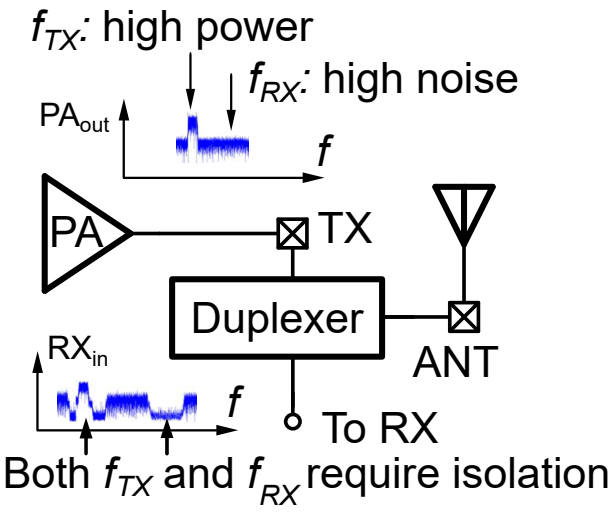


Figure 1.1: A duplexer connecting antenna (ANT), TX and RX ports.

1.2 Challenges in RF Front-End Circuits for FDD Radios

1.2.1 Duplexers

Antenna interface, or duplexer (Fig. 1.1), is the building block which connects between the antenna (ANT), the TX and RX ports. To enable FDD, dual-band TX-RX isolation must be provided by the duplexer in order to:

1. Prevent high power TX signal (at TX band f_{TX}) from leaking into the RX, so the RX does not saturate.

2. Prevent amplified TX noise (at RX band f_{RX}) from leaking into the RX, so the RX is not desensitized.

Nowadays, a duplexer uses two SAW filters for each TX and RX band pair (Fig. 1.2). The SAW filters reject the TX noise (at RX band) at the output of the TX, and the TX signal (at TX band) at the input of the RX. There are several disadvantages:

1. They are not tunable. Numerous SAW filters are required for nowadays multi-band FDD mobile applications.

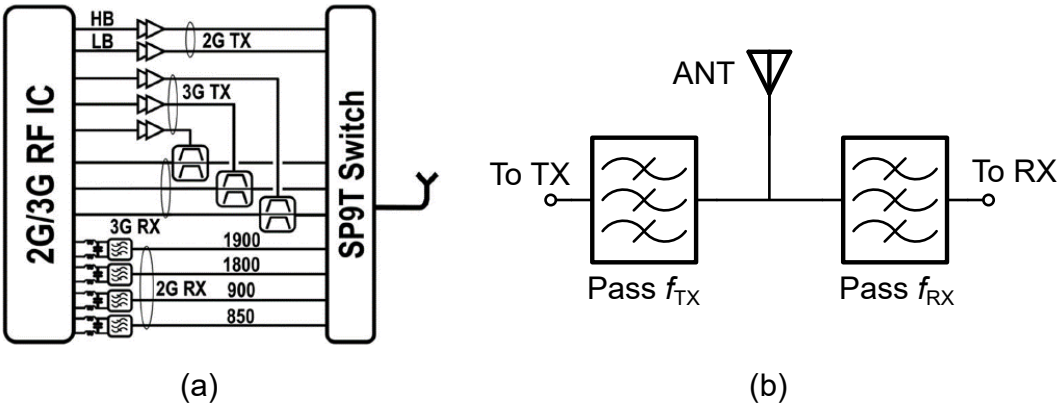


Figure 1.2: (a) A board block diagram in a 2G/3G cellular phone [1]; (b) A SAW duplexer.

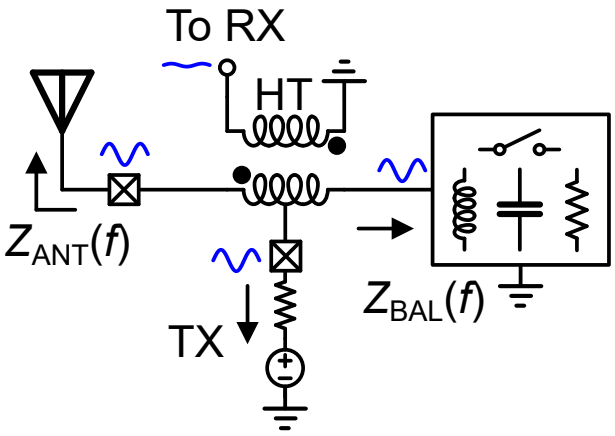


Figure 1.3: An electrical balance duplexer (EBD) consisting of a hybrid transformer (HT) and a balance network (BAL).

2. The SAW filters are off-chip components, which are bulky and costly.

There are many approaches [1, 3, 5, 6, 10–21] that have been studied earlier trying to replace the off-chip SAW duplexers with on-chip solutions. They can be summarized into three categories by how the FETs are used:

1. FETs used as amplifiers [19–21]:
   RF active cancellers fall into this category. The cancellers take a replica of the TX signal and/or TX noise, process the replica in RF domain, and

then subtract the TX leakage at the input of the RX. They are usually used together with other duplexing techniques due to the disadvantage of this approach:

- Bad linearity and poor power handling.

2. FETs used as periodic clock controlled switches [16–18]:
   The use of N-path filters or passive mixers falls into this category. The disadvantages of this approach are:

   - RX desensitization by clock phase noise. Due to reciprocal mixing, the clock phase noise can be amplified because of high TX operating power and raise the noise at RX input. At practical operating power level, the desensitization is not acceptable.
   - Poor power handling. The FETs will go into compression well below practical operating power level.
   - High power consumption. Since the duplexer connects the output of the power amplifier (PA), they usually consume comparable or higher power than the PA to drive the FETs.

3. FETs used as static switches [1, 3, 5, 6, 10–15]:
   RF passive cancellers including the electrical balance duplexer (Fig. 1.3) fall into this category. The FETs are used as static switches to select passive on-chip elements. This approach benefits from the large power handling, high linearity and low static power consumption. However, it suffers from:

   - Fundamental 3dB loss in both TX and RX paths. Because of the property of three-port passive networks, it is impossible to construct such a three-port lossless reciprocal network that is matched at all ports. The loss may be acceptable for some applications.

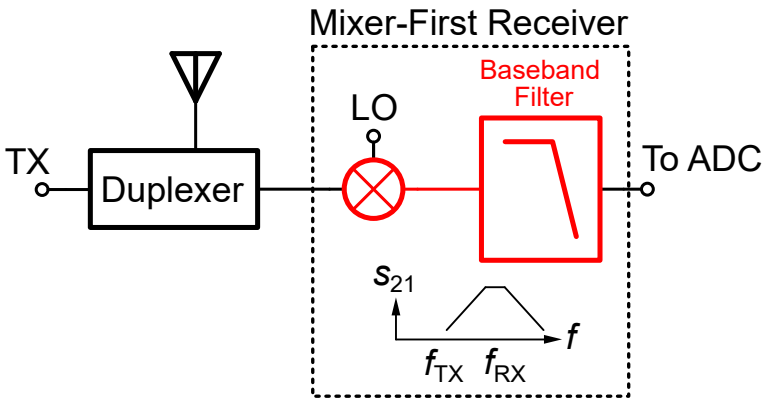


Figure 1.4: A mixer-first receiver.

- Poor dual-band isolation capability. Due to low-Q of on-chip elements, it is difficult for the tunable passive cancellers to track frequency response of the TX leakage channel for dual-band isolation.

Among all approaches, only the third category supports practical TX power level. Therefore, it is promising for actual use if the mentioned problems can be solved, especially the dual-band isolation capability.

1.2.2 Receiver Front-Ends

Multi-band receivers experience interference from blockers arising from a multitude of conditions, including leakage from their own transmitter when operating in FDD. The traditional use of SAW filters and SAW duplexers soon becomes impractical, which means that the receiver circuits must be designed to handle unfiltered blockers without entering gain compression that would corrupt the small wanted signal.

The mixer-first receiver [22–26] immediately shifts the received spectrum at the antenna to place the wanted channel at zero IF. An active baseband circuit follows the mixer as shown in Fig. 1.4. It amplifies the baseband spectrum—comprising the small wanted signal and nearby large blockers—and filters the

blockers relatively while amplifying the channel at zero IF (DC).

For close-in blockers, large filtering is needed for high dynamic range, where the baseband amplifier does not compress. For far-out blockers, which have been largely suppressed by filtering, the mixer dictates the dynamic range. The design of both the high dynamic range mixer and the sharp filtering baseband amplifier remains a challenge.

1.3 About the Dissertation

1.3.1 Contributions

This dissertation explores techniques that can overcome the challenges in the RF front-end circuits, both for the duplexer and the receiver front-end. The designs in this dissertation are not only described in detail, verified by silicon, but also justified by analysis throughout the dissertation. The contributions of the dissertation are summarized below:

- A state-of-the-art dual-band electrical balance duplexer with a novel independently tuned dual-band balance network is introduced for FDD applications.
- General properties of finite-Q LCR networks are analyzed for CMOS passive network design, and from the analysis a notch filter is implemented using CMOS passive elements achieving maximum $|S_{21}|$ transition rate.
- A simple and insightful second-order transimpedance amplifier (TIA) design methodology for mixer-first receivers is developed and verified by silicon.
- Small and large signal nonlinearities in a passive mixer are analyzed for passive mixer design.

1.3.2 Outline

This dissertation is organized as follows:

- Chapter 2 shows the design of an electrical balance duplexer that support independently tuned dual-band TX-RX isolation for FDD applications.
- Chapter 3 analyzes the properties of finite-Q LCR networks for CMOS passive network design.
- Chapter 4 shows the second-order transimpedance amplifiers (TIAs) and the mixer-first receiver design and analysis.
- Chapter 5 analyzes the large and small signal nonlinearity in passive mixers.
- Finally, possible future works and conclusions are included in Chapter 6.

CHAPTER 2

A Dual-Band Electrical Balance Duplexer for FDD Radios

2.1 Introduction: Electrical Balance Duplexer

The electrical balance duplexer (EBD), shown in Fig. 1.3, is an alternative to the SAW duplexer that can be integrated on a chip, at a lower cost and in a much smaller volume [1–6, 10–14]. It consists of a hybrid transformer (HT) and a variable impedance, the balance network (BAL). The HT connects the ANT, TX, BAL at its primary turn, and the RX to its secondary turn. To handle the high-power TX output, the EBD and associated balance networks are all passive (LCRs and high-voltage static switches).

Unlike the SAW duplexer, the EBD isolates the RX from the TX through the nulling action in the HT. When the TX port connects to the center of the HT primary turn and $Z_{BAL} = Z_{ANT}$, no TX power leaks into the RX due to symmetry. Therefore, the HT is balanced, and the TX-RX isolation is established.

When the HT is balanced, both TX-ANT and ANT-RX paths will suffer from 3dB fundamental loss. For TX-ANT path, due to symmetry, half of TX power will be dissipated in the BAL thus resulting in 3dB TX path insertion loss (TX IL). For ANT-RX path, due to symmetry as well, the BAL and ANT contribute equal noise to the RX, resulting in 3dB RX path insertion loss (RX IL). Note that the TX does not contribute noise because of the nulling action.

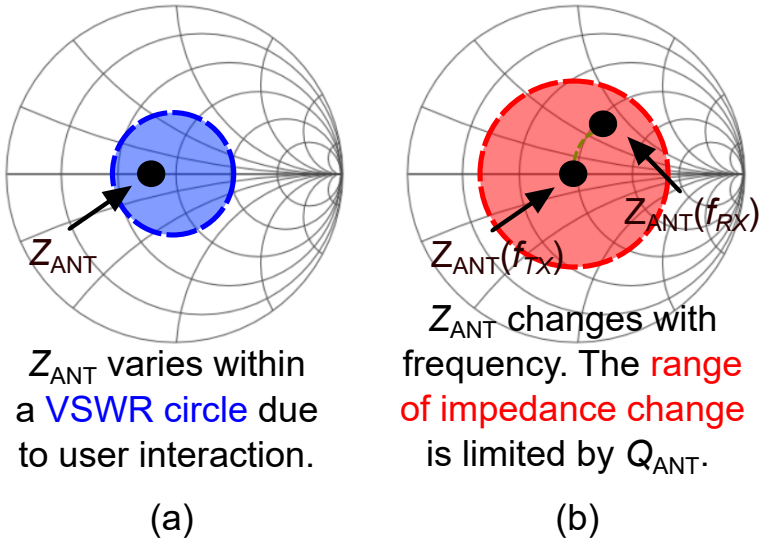


Figure 2.1: (a) Z_{ANT} changes due to user interaction; (b) Z_{ANT} is a function of frequency, and the impedance difference is limited by Q_{ANT}.

One challenge of the EBD design for FDD radios is the dual-band capability. For TX and RX to isolate at dual bands, $Z_{BAL} = Z_{ANT}$ must hold at both TX and RX bands. However, because the Z_{ANT} frequency response is usually different then Z_{BAL}, the dual-band capability remains a major challenge for EBD design.

2.2 Antenna Characteristics

The characteristics of the antenna highly impact the EBD performance, because sufficient TX-RX isolation depends on the balance condition $Z_{BAL} = Z_{ANT}$. Therefore, the antenna cannot be treated as a 50Ω resistor. The impedance of an actual antenna depends on many factors. The antenna impedance Z_{ANT} (or equivalently described by reflection coefficient $\Gamma_{ANT} \triangleq \frac{Z_{ANT}-50\Omega}{Z_{ANT}+50\Omega}$) can change over:

1. User interaction.

   Z_{ANT} changes with the orientation and proximity to the user's body, as specified by a range of antenna VSWR (Fig. 2.1). The VSWR=2 circle is

usually considered sufficient to cover most user scenarios [27–30]. Antennas in mobile devices are designed typically for a smaller VSWR, and an optional antenna tuning unit (ATU) [31] can lower it even more.

Adaptation algorithms [32,33] adjust EBD tuning element values in real time to maintain balance when antenna impedance changes over user interaction.

2. Frequency.

   Z_{ANT} is a strong function of frequency indicated by Q_{ANT}, which means that it changes considerably from f_{TX} to f_{RX}. As shown in Fig. 2.1, the impedance (reflection coefficient) change $\Delta\Gamma_{ANT}$ consists of magnitude and phase.

   - The impedance change in magnitude $|\Delta\Gamma_{ANT}|$ is limited by Q_{ANT}. Q_{ANT} is specified by the application to meet antenna matching requirement. For example, antennas for Wi-Fi operating between 5-7GHz require $Q_{ANT} = \frac{f_{center}}{\text{Bandwidth}} \leq 3$.
   - The impedance change in phase $\angle(\Delta\Gamma_{ANT})$, on the other hand, can be arbitrary. First, it changes when different antennas are used for the same application, thus cannot be predicted beforehand. Second, it changes with the length of the cable or PCB transmission line connecting the antenna and the EBD, thus unpredictable before the EBD is designed.

The EBD is viable if a circuit for Z_{BAL} is devised that tracks changes in Z_{ANT} with user interaction and frequency response, so it isolates in the TX channel (f_{TX}) and RX channel (f_{TX}) simultaneously for FDD.

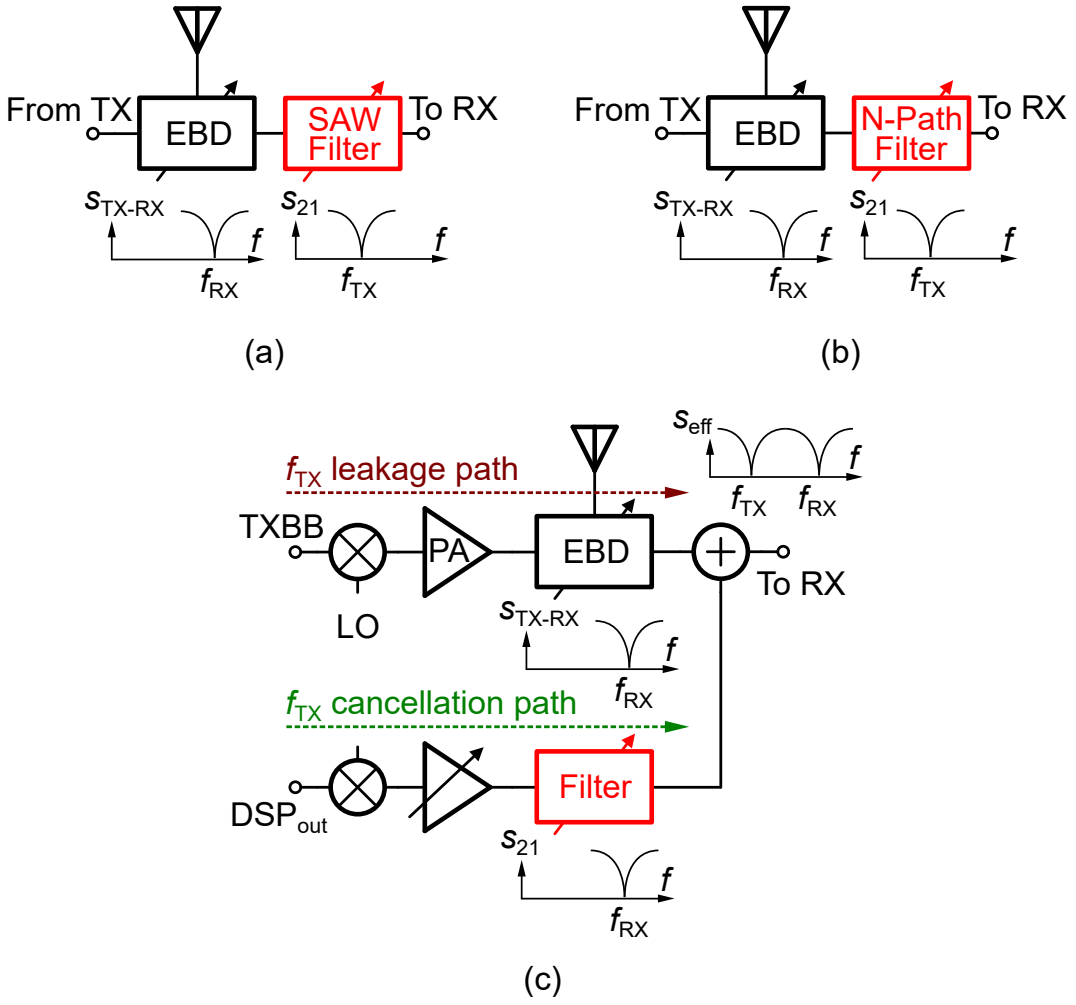


Figure 2.2: A single-band EBD and (a) a SAW filter [2]; (b) a N-path filter [3]; (c) an active cancellation path [4].

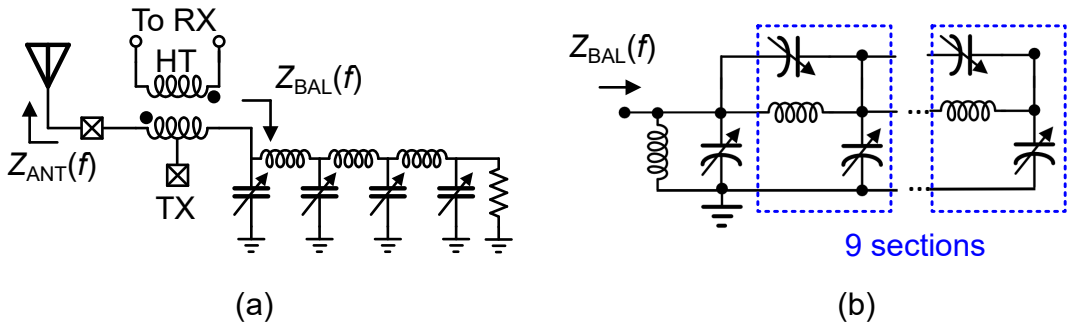


Figure 2.3: (a) A dual-band EBD in [5]; (b) A dual-band BAL in [6].

2.3 Prior Arts

There are works that enable dual-band isolation for the EBD [2–6, 11]. Their approaches fall into two categories:

1. A single-band EBD with additional filtering or cancellation circuits [2–4] (Fig. 2.2). [2] adds an additional SAW filter at the RX path to filter TX signal, while the EBD rejects the TX noise. Similarly, [3] replaces the SAW filter and inserts a N-path filter at RX path. [4] uses active cancellation to cancel TX signal at the input of the RX. Their disadvantages are:

    - [2] requires an off-chip SAW filter.
    - [3] cannot operate at practical TX power level, because of reciprocal mixing and compression of the N-path filter.
    - [4] requires an off-chip filter, in order to filter the cancellation path noise at the input of the RX.
    - Additional circuits will introduce additional loss or noise to the EBD.

2. A dual-band EBD [5, 6, 11] (Fig. 2.3). [5, 11] designs the balance network with four tuning elements, therefore the degree of freedom is sufficient to synthesize complex $Z_{BAL}(f_{TX})$ and $Z_{BAL}(f_{RX})$ simultaneously. [6] designs a more complicated balance network with more tuning degree of freedom. The major disadvantage of the two works is:

    - They are poor at tracking changes in Z_{ANT} with user interaction and frequency response simultaneously.

      Since tuning one capacitance does not only change $Z_{BAL}(f_{TX})$, but also $Z_{BAL}(f_{RX})$ at the same time, impedance at two bands cannot be adjusted independently. The tuning range of $Z_{BAL}(f_{RX})$ now highly depends on $Z_{BAL}(f_{TX})$. In other words, the ability of the EBD to track

antenna frequency response highly depends on the impedance, which changes due to user interaction. Therefore, the EBD can either track antenna user interaction, or the frequency response, but not both at the same time.

The prior arts on the dual-band EBD [2–6, 11] show that it is important to design a balance network that can synthesize impedance at two bands independently, in order to track Z_{ANT} change over user interaction and frequency.

2.4 Design Specifications

The application is the Wi-Fi 6/6E operating at 5-7GHz, where the TX is specified to operate at 5-6GHz and the RX at 6-7GHz. The duplexer will enable the simultaneous transmission and reception at different channels (FDD). In order to enable FDD, the TX-RX isolation is specified to be higher than 40dB over an 80MHz channel at dual-band. It must operate under high TX power level, up to +27dBm. Since there are 3dB fundamental losses for the EBD, the loss for both TX and RX paths is specified to be less than 4dB, with less than 1dB added loss from the implementation.

To summarize, the specifications are listed below:

1. Wi-Fi 6/6E (5-7GHz)
   - TX: 5-6GHz
   - RX: 6-7GHz
   - $Q_{ANT} \leq 3$
2. Sub-4dB IL
3. 40dB isolation

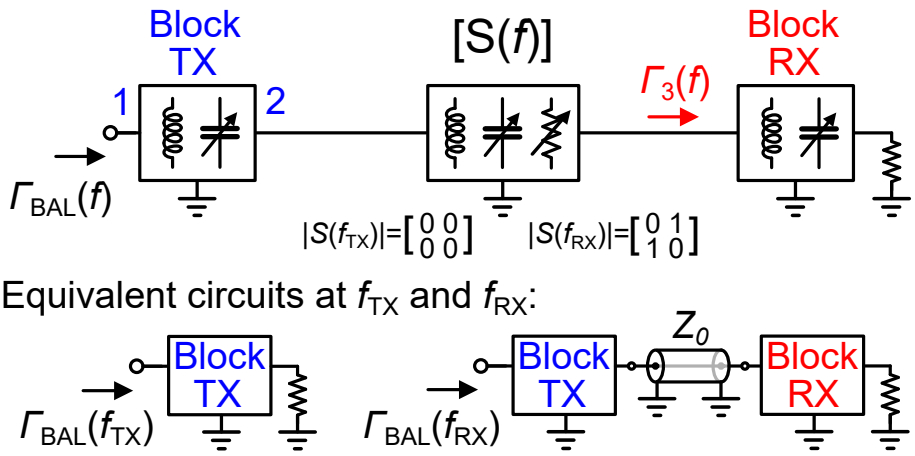


Figure 2.4: Proposed block diagram of the dual-band balance network and its equivalent circuits at f_{TX} and f_{RX}.

- 80MHz channel bandwidth (BW)

4. +27dBm TX power

Towerjazz 65nm RF-SOI CMOS technology is used to fabricate the chip.

2.5 Dual-Band Balance Network Design

2.5.1 Conceptual Principle of Operation

Fig. 2.4 shows the proposed balance network conceptual block diagram. The shown balance network can synthesize Z_{BAL} at both bands independently. It contains three sub-blocks: Block TX, Block RX, and the intermediate two-port filter $[S(f)]$ defined by its s-parameters.

At TX band, the intermediate two-port $[S(f)]$ rejects the transmission, and the Block RX does not affect the impedance $Z_{BAL}(f_{TX})$. The equivalent circuit is shown in Fig. 2.4. The Block TX is designed to set the impedance at f_{TX} solely:

$$\Gamma_{BAL}(f_{TX}) = s_{11,\text{B-TX}}(f_{TX}) \tag{2.1}$$

where $s_{11,\text{B-TX}}$ is the s-parameters of the Block TX. The tuning range $\Gamma_{BAL}(f_{TX})$

by the Block TX deals with antenna impedance changes over user interaction.

At RX band, however, the Block RX and Block TX load each other and both sub-blocks determine the impedance. The equivalent circuit is shown in Fig. 2.4. The impedance at RX band:

$$\Gamma_{BAL}(f_{RX}) = s_{11,\text{B-TX}}(f_{RX}) + \frac{s^2_{21,\text{B-TX}} \times \Gamma_3 \times e^{2j\theta}}{1 - s_{22,\text{B-TX}} \times \Gamma_3 \times e^{2j\theta}} \tag{2.2}$$

where $s_{21,\text{B-TX}}$ and $s_{22,\text{B-TX}}$ are the s-parameters of the Block TX, Γ_3 is the input impedance (reflection coefficient) of the loaded Block RX, and $e^{2j\theta}$ is the delay through the intermediate two-port. Simplifications can be made on equation 2.2 by assuming $s_{11,\text{B-TX}}(f_{RX}) \approx s_{11,\text{B-TX}}(f_{TX})$ (will discuss later in Block TX design section), $|s_{21,\text{B-TX}}|^2 \approx (1 - |s_{11,\text{B-TX}}|^2) > 0.88 \approx 1$ ($|s_{11,\text{B-TX}}| \leq 1/3$ for VSWR=2 and low loss) and $s_{22,\text{B-TX}} \times \Gamma_3 \approx 1/3 \times 1/2 \ll 1$ (negligible multiple reflections):

$$\Gamma_{BAL}(f_{RX}) \approx \Gamma_{BAL}(f_{TX}) + \Gamma_3 \times e^{2j\theta} \tag{2.3}$$

The impedance at the same terminal at f_{RX} is then set with unaltered Block TX, where the impedance difference $\Delta\Gamma_{BAL} \triangleq \Gamma_{BAL}(f_{RX}) - \Gamma_{BAL}(f_{TX})$ is solely set by the Block RX. The tuning range $\Delta\Gamma_{BAL}$ by the Block RX deals with antenna impedance changes over frequency.

The circuit now balances the HT at both TX and RX band independently.

2.5.2 Intermediate Filter $[S(f)]$ Design

The most critical sub-block of the balance network is the intermediate two-port filter $[S(f)]$, since its performance determines whether the independent balancing can be achieved.

The challenge of designing this sub-block is to realize a sharp enough $|s_{21}|$ transition from stopband (f_{TX}) to passband (f_{RX}) with the following constraints:

- The frequency spacing between the TX and RX band can be small ($\sim 10\%$).

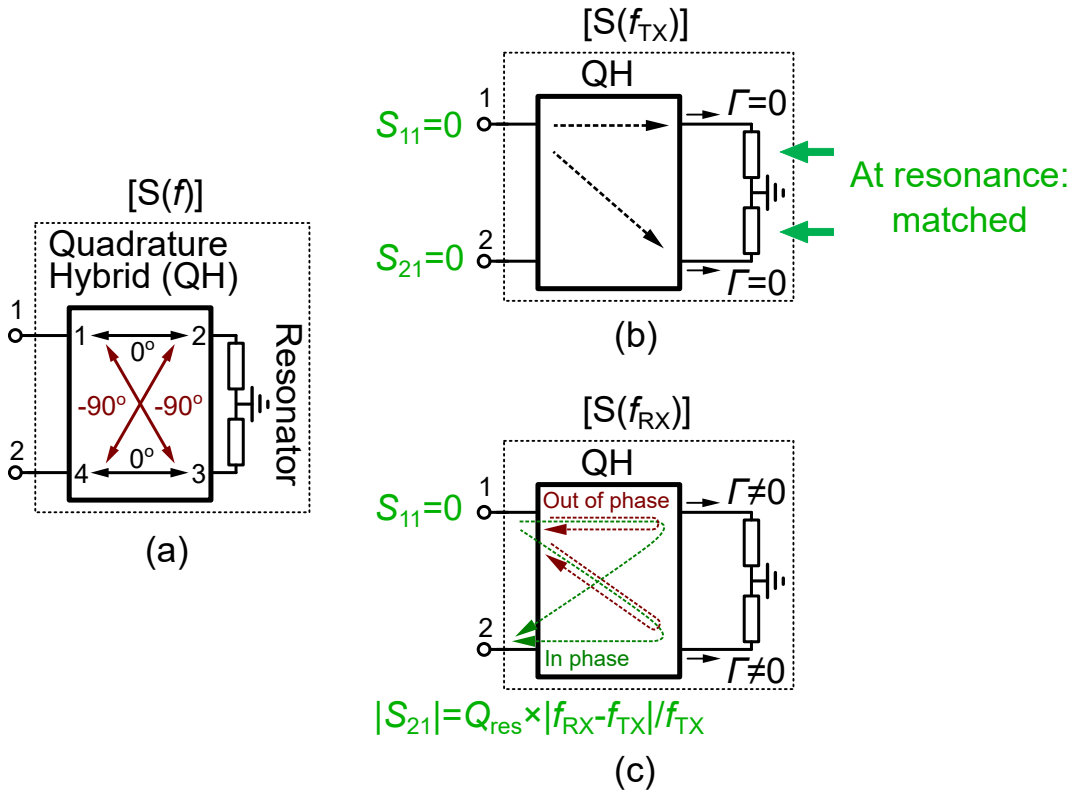


Figure 2.5: (a) The perfectly-matched absorptive bandstop filter [7]; (b) The operation at f_{TX}; (c) The operation at f_{RX}.

- It must be realized by low-Q ($\sim 10^1$) on-chip passives.

Therefore, an analysis on the maximum achievable $|\frac{\partial s_{21}}{\partial f/f}|$ using finite-Q passives and its design guidance is necessary. The detailed analysis is given in Chapter 3.

It is proved in Chapter 3 that the perfectly-matched absorptive bandstop filter introduced in [7] achieves the maximum $\left|\frac{\partial s_{21}}{\partial f/f}\right|_{max} = (Q_{L,max} \| Q_{C,max})$ using finite-Q passives, as shown in equation 3.28. The operations of the perfectly-matched absorptive bandstop filter are shown in Fig. 2.5. It consists of a quadrature hybrid (QH) and two identical lossy resonators loading the THRU and CPL ports of the QH.

At f_{TX}, incident power at port-1 is absorbed by two identical resonators at resonance without reflection ($\Gamma = 0$) by design, thus there are no reflection and

transmission ($s_{11} = s_{21} = 0$).

At f_{RX}, however, both resonators will appear as small but equal reactances. They will reflect some power entering port-1 ($\Gamma \neq 0$). The reflected waves reinforce at port-2 but cancel at port-1 ($s_{11} = 0$). s_{21} can be calculated:

$$\begin{aligned} |s_{21}(f_{RX})| &= |\Gamma(f_{RX})| \\ &= Q_{res} \times \frac{|f_{RX} - f_{TX}|}{f_{TX}} \\ &= Q_{res} \times (\Delta f / f) \end{aligned} \tag{2.4}$$

where Q_{res} is quality factor of the resonator, $\Delta f \triangleq |f_{RX} - f_{TX}|$. Equation 2.4 achieves the maximum $\left|\frac{\partial s_{21}}{\partial f/f}\right|_{max} = (Q_{L,max} \| Q_{C,max})$ shown in equation 3.28. Therefore, this filter gives the best performance on filter sharpness for our application.

A lumped element QH [8] with two lossy resonators shown in Fig. 2.6 realizes, at port-1 and 4 of the QH, the intermediate two-port filter with the s-parameters shown in Fig. 2.4. The QH four-port consists of two inductors and two tunable capacitors for different center frequencies. The resonator is designed in the following way: first, a parallel LC resonator with a tunable capacitor C_{res} determines the resonance frequency; second, coupled inductors coarsely transform the higher parallel resonance resistance R_P to the lower characteristic impedance Z_0; third, a tunable resistor is added to fine tune the resonance resistance for good matching at f_{TX}. The tunable C_{res} is designed to be 6-bit for sub-30MHz resonance frequency resolution between 5-6GHz, excluding $\sim$ 10% C_{res} tuning margin for variations.

The simulated s-parameters of the filter are plotted in Fig. 2.6. Two C_{res} settings are chosen for simulation, giving two resonance frequencies and thus two stopband frequencies. As predicted, the simulated s_{11} shows the wideband matching property in both stopband and passband. The simulated $|s_{21}|$ shows higher than 20dB stopband rejection and sub-2dB passband IL at $\sim$ 10% frequency offset.

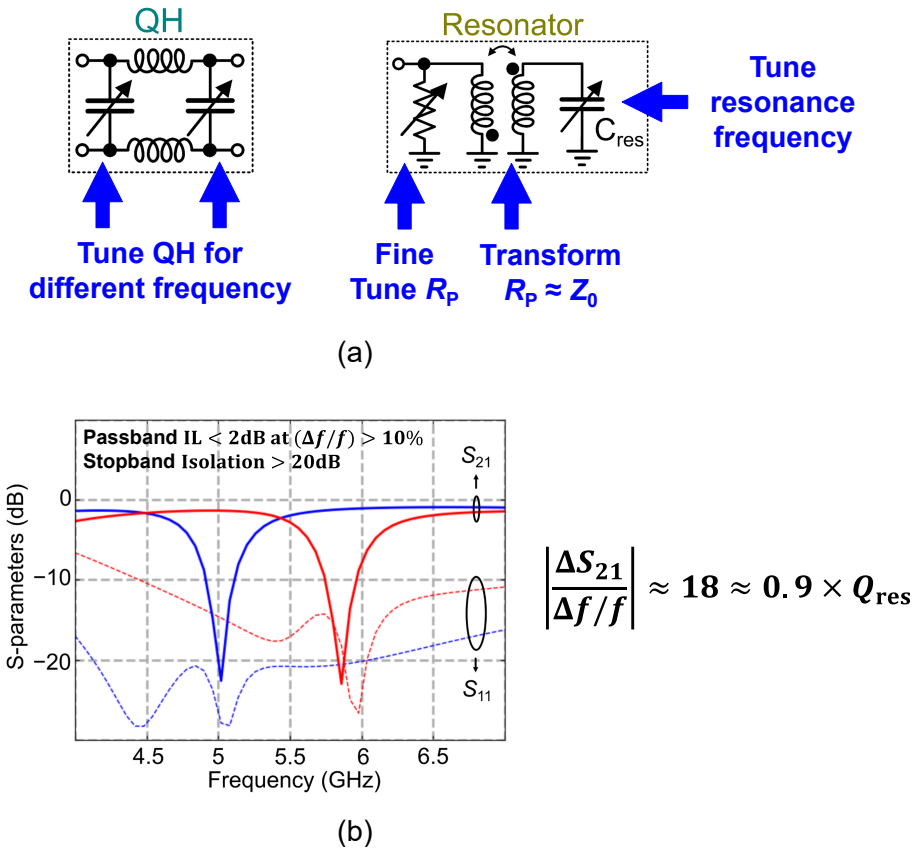


Figure 2.6: (a) The implementation of the QH [8] and the resonators in Fig. 2.5; (b) Simulated s-parameters of the two-port $[S(f)]$ at two resonance frequencies.

The $|\frac{\partial s_{21}}{\partial f/f}|$ is calculated from simulation:

$$\left|\frac{\partial s_{21}}{\partial f/f}\right| \approx 18 \approx 0.9 \times Q_{res} \tag{2.5}$$

Equation 2.5 indicates the implementation is almost the optimum that can be realized in this technology.

2.5.3 Block TX Design

The Block TX is designed to set the impedance at f_{TX} solely, indicated by equation 2.1. The $s_{11,\text{B-TX}}$ of the Block TX must be able to track the antenna impedance changes over user interaction, which is usually specified by the VSWR=2 circle.

On the other hand, the Block TX must not disturb the impedance at f_{RX}.

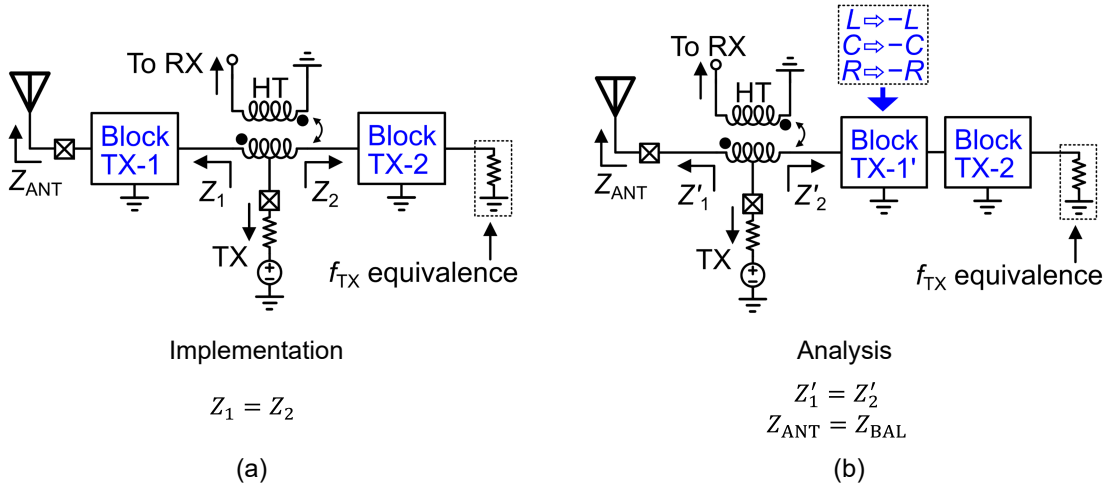


Figure 2.7: (a) The Block TX is split into two parts; (b) The equivalent circuit for analysis purpose without disturbing the established balance.

This requires $s_{11,\text{B-TX}}(f_{RX}) \approx s_{11,\text{B-TX}}(f_{TX})$. If above requirement is not met, equation 2.3 will be:

$$\Gamma_{BAL}(f_{RX}) \approx \Gamma_{BAL}(f_{TX}) + \Delta s_{11,\text{B-TX}} + \Gamma_3 \times e^{2j\theta} \tag{2.6}$$

where $\Delta s_{11,\text{B-TX}} = s_{11,\text{B-TX}}(f_{RX}) - s_{11,\text{B-TX}}(f_{TX})$. The second term introduced by the Block TX disturbs $\Delta\Gamma_{BAL}$, which ideally is solely determined by the Block RX for tracking antenna impedance changes over frequency. Therefore, $\Delta s_{11,\text{B-TX}}$ must be minimized.

The Block TX is implemented by splitting into two parts (Fig. 2.7), Block TX-1 and Block TX-2, at the antenna and balance network side of the HT, respectively. There are two explanations to how the $\Delta s_{11,\text{B-TX}}$ is minimized in this implementation:

1. The $(\partial s_{11,\text{B-TX-1}}/\partial f)$ and $(\partial s_{11,\text{B-TX-2}}/\partial f)$ of the Block TX-1 and TX-2 cancel each other partially, resulting in minimized frequency response of overall $\Delta s_{11,\text{B-TX}} = \Delta s_{11,\text{B-TX-2}} - \Delta s_{11,\text{B-TX-1}}$.

2. As Fig. 2.7 illustrates, the circuit transforms to an equivalent circuit such

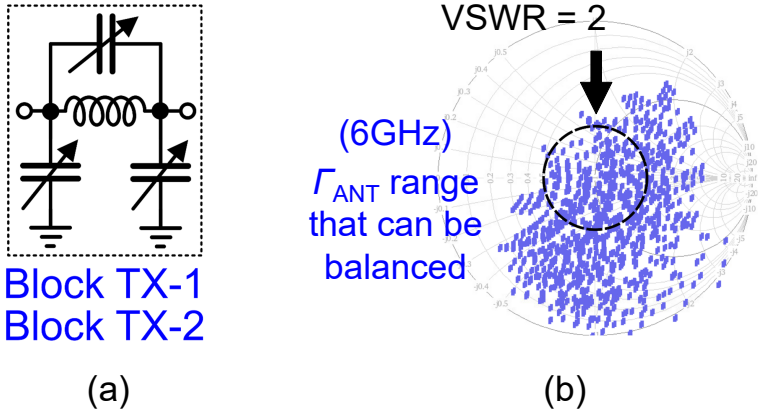


Figure 2.8: (a) The implementation of Block TX-1 and TX-2; (b) The simulated Γ_{ANT} coverage by the Block TX at 6GHz.

that the Block TX-1 is moved to the balance network side, while the balance condition is not disturbed. This can be done because adding series or parallel elements symmetrically at both sides of the HT do not disturb the balance condition. The transformed Block TX-1' now presents negative LCRs. The negative resistors, indicating negative loss in the Block TX-1', partially cancels the loss in the Block TX-2:

$$\text{Power loss}_{\text{B-TX}} = \text{Power loss}_{\text{B-TX-2}}(>0) + \text{Power loss}_{\text{B-TX-1'}}(<0) \tag{2.7}$$

From equation 3.38, the frequency sensitivity of a N-port s-parameters $|s_{21}|$ is limited by elements quality factor (Q_L, Q_C) and power lost in the two-port:

$$\left|\frac{\partial s_{11}}{\partial f/f}\right| \leq (Q_L \| Q_C) \times \underbrace{\left(1 - |s_{21}^2| - |s_{11}^2|\right)}_{\substack{\text{Power Loss \%} \\ \text{in the two-port} \\ \text{(source at port-1)}}} \tag{2.8}$$

Therefore, when the power loss is minimized, the $\Delta s_{11,\text{B-TX}}$ is minimized.

The Block TX-1 and TX-2 are implemented with one stage of LC π-network as shown in Fig. 2.8. The simulated Γ_{ANT} coverage is plotted in Fig. 2.8 with coarse resolution. A VSWR=2 circle can be covered to track the antenna impedance

changes over user interaction. The grounded tunable capacitors are 4-bit and the flying capacitors are 3-bit, providing sufficient Γ_{BAL} fine tuning resolution.

The Block TX-1 may introduce additional IL because it is in the signal path. The simulated average loss at 6GHz is $\mu = 0.78$dB, the standard deviation is $\sigma = 0.38$dB, and the minimum loss is 0.2dB.

2.5.4 Block RX Design

The loaded Block RX is a one-port impedance synthesizer, where the input impedance is Z_3, or given by Γ_3. As indicated in equation 2.3, a higher Γ_3 tuning range yields a higher covered range for the antenna frequency response. Therefore a higher Γ_3 tuning range is desired.

$|\Gamma_3| = 0.5$ design target is chosen in this design for two reasons:

1. $|\Gamma_3| = 0.5$ is sufficient to cover $Q_{ANT} = 3$ for moderate frequency offset ($\sim 10\%$). Because from equation 3.34:

$$\left|\frac{\Delta\Gamma_{BAL}}{\Delta f/f}\right| \approx \left|\frac{\Gamma_3}{\Delta f/f}\right| \sim 5 > 3 \geq \left|\frac{\Delta\Gamma_{ANT}}{\Delta f/f}\right| \tag{2.9}$$

   the balance network is sufficient to track the antenna impedance change over frequency response. More than 4dB losses from the Block TX and the intermediate filter are still tolerated, since $5/(10^{4/20}) > 3$.

2. When Γ_3 tuning range is too large, it becomes hard to cover the entire range uniformly.

The Block RX is implemented using two stages of LC π-network as shown in Fig. 2.9. The simulated Γ_3 coverage is plotted in Fig. 2.9 with coarse resolution. It can cover the $|\Gamma_3| = 0.5$ design target. The grounded tunable capacitors are 4-bit and the flying capacitors are 3-bit, providing sufficient Γ_3 fine tuning resolution.

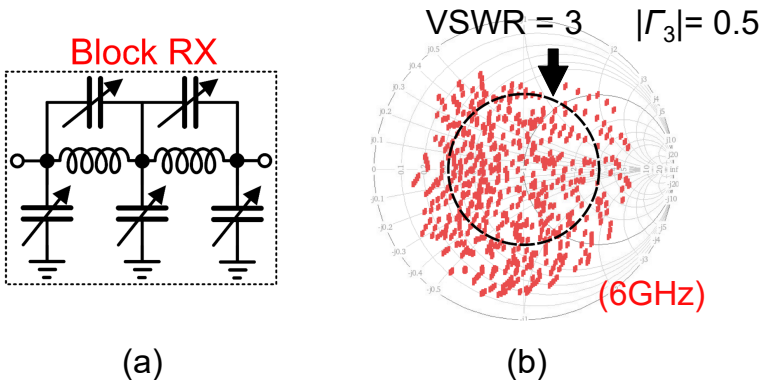


Figure 2.9: (a) The implementation of Block RX; (b) The simulated Γ_3 coverage by the Block RX at 6GHz.

2.5.5 High-Voltage Switch Design

FET switches are used to select the elements in tuning capacitors and resistors. The gate, body, drain, and source of the FET are biased with high impedance connection, so at RF they appear as floating nodes. The drain and source are the two terminals of the switch.

When the switch is on, it can be modeled by FET on-resistance R_{on}, as shown in Fig. 2.10. And when it is off, it can be modeled by FET capacitance $C_{off} = (C_{GD} \| C_{GS}) + (C_{DB} \| C_{SB})$. Although R_{on} and C_{off} change with different FET sizes, their product $R_{on} \times C_{off}$ is a constant, determined by technology.

A tuning capacitor, consisting a fixed capacitor C_{max} and a series FET switch, will suffer from loss due to R_{on} when the switch is on. The quality factor Q_C is dominated by $Q_C = 1/(\omega C_{max} R_{on})$. On the other hand, when the switch is off, the series capacitance ($C_{min} = C_{max} \| C_{off}$) is not zero due to non-zero C_{off}. The capacitance tuning range C_{max}/C_{min} and on-sate quality factor Q_C follows the relation below:

$$Q_C \times \left(\frac{C_{max}}{C_{min}} - 1 \right) = \frac{1}{\omega \times R_{on} C_{off}} \tag{2.10}$$

where $R_{on}C_{off}$ is determined by technology, and ω is the frequency of interest. Equation 2.10 shows that the loss and the capacitance tuning range trade with

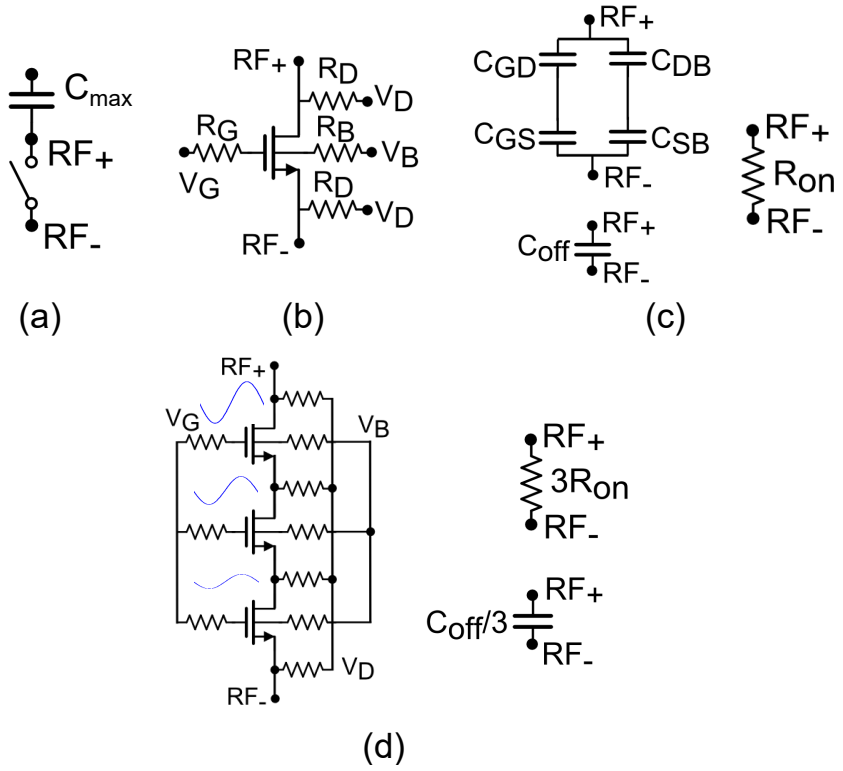


Figure 2.10: (a) The unit of tuning capacitors; (b) The switch realized by one FET; (c) Switch ON and OFF state equivalent circuits; (d) Stacking three FETs increases the voltage tolerance with identical $R_{on}C_{off}$ product.

each other, where the technology and operating frequency are the fundamental limitation.

The Towerjazz 65nm RF-SOI CMOS technology used in this design provides $R_{on}C_{off} \sim 150fs$ (parasitics extracted). The switches in the resonator tuning capacitor C_{res} (Fig. 2.6) are designed with low tuning range $(C_{max}/C_{min}) \approx 3$ but high $Q_C \sim 100$ (at 5GHz), for better Q_{res} and filter sharpness (equation 2.4). The switches in the rest capacitors are designed with moderate tuning range $(C_{max}/C_{min}) \approx 5$ and moderate $Q_C \sim 50$ (at 5GHz).

To operate at high TX power level, the switch must tolerate high voltage swing. This can be done by stacking multiple FETs as shown in Fig. 2.10. Because the gate, body, drain, and source of the FET are floating at RF, the voltage swing will

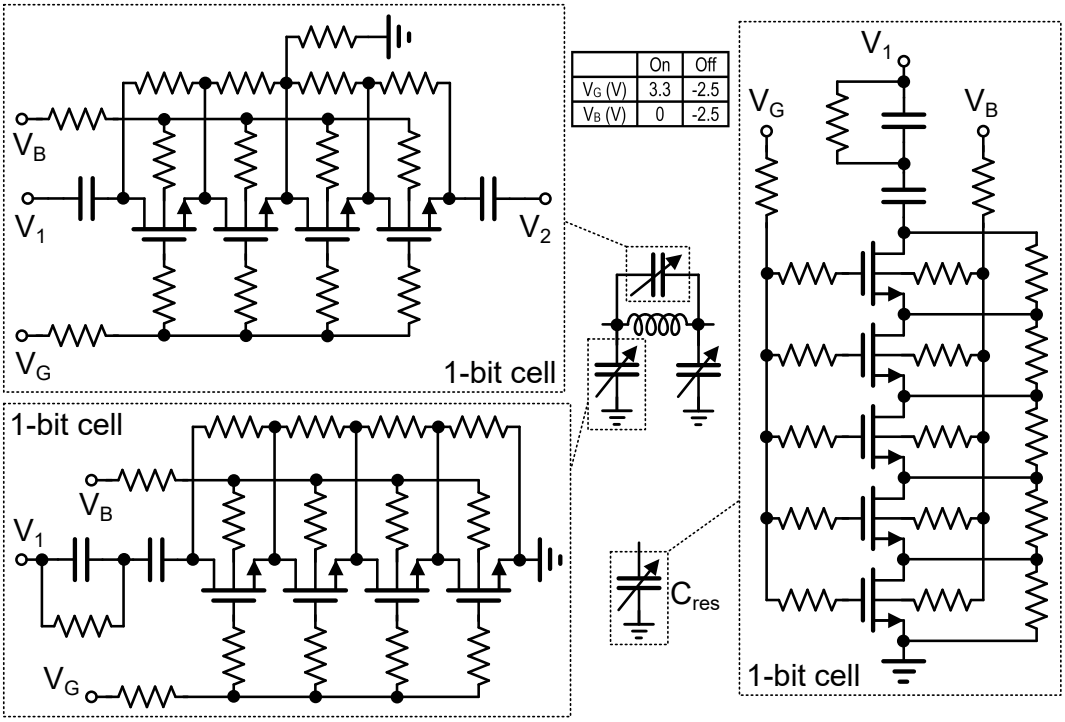


Figure 2.11: The switch design of the tuning capacitors.

be evenly distributed on all FETs. The maximum voltage swing is proportional to the number of FET stacks. As shown in Fig. 2.10, the $R_{on}C_{off}$ of the switch does not change when FETs are stacked.

The 2.5V FET in this technology can operate at $V_{DS} = 4$V without reliability problem (confirmed and measured by the foundry). At f_{TX}, half of TX power will be dissipated in the balance network, and most of it dissipated in the resonators. C_{res} in the resonators will suffer from the highest voltage swing. Stack-5 FETs are used for C_{res} that can tolerate the peak voltage of $\sim$ 20V for >27dBm TX power. Other tuning capacitors uses stack-4 FETs that can tolerate the peak voltage of $\sim$ 16V, which is sufficient for 27dBm TX power and a voltage gain of 2 (VSWR=2). The design implementations are shown in Fig. 2.11.

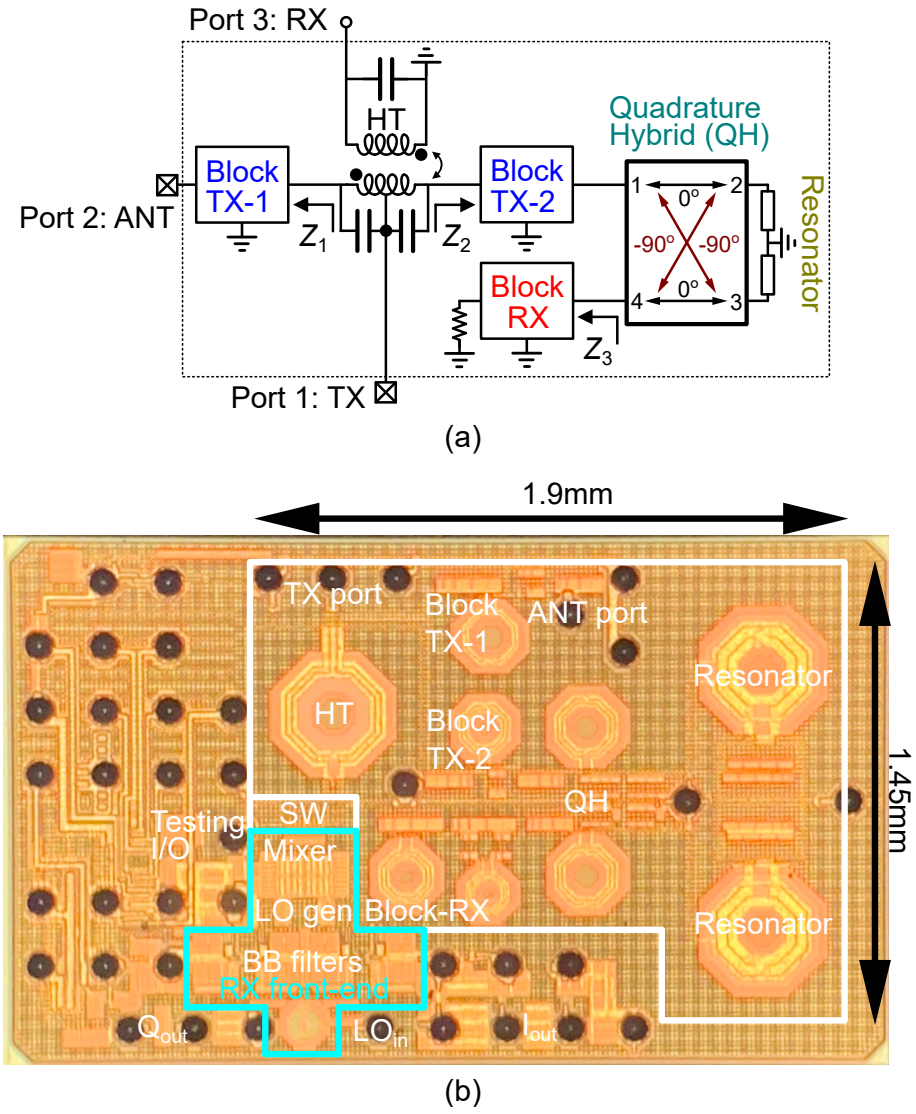


Figure 2.12: (a) The complete circuit diagram of the dual-band EBD; (b) Die micrograph of an EBD and a receiver.

2.6 Measurements

This duplexer is fabricated in Towerjazz 65nm RF-SOI CMOS technology. The die micrograph is shown in Fig. 2.12. The receiver shown in Fig. 2.12 will be introduced later. The EBD is measured without the receiver.

2.6.1 Dual-Band Balancing

The measured Γ_{ANT} coverage to track antenna impedance changes over user interaction and frequency response is plotted in this way: For each setting (of tunable elements), the s-parameters are measured. Then Γ_{ANT}, which will lead to perfect TX-RX isolation for this setting, is solved from the wave equations:

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix} \begin{pmatrix} 1 \\ \Gamma_{ANT} \times y \\ 0 \end{pmatrix} \tag{2.11}$$

where port-1, 2, 3 correspond to the TX, ANT, and RX ports, respectively, x is the outgoing wave at TX port, and y is the outgoing wave at the ANT port. From equation 2.11 the Γ_{ANT} is:

$$\Gamma_{ANT} = -\frac{s_{31}}{s_{32} \times s_{21} - s_{31} \times s_{22}} \tag{2.12}$$

Now for each setting, the corresponding Γ_{ANT} that balances the HT can be plotted on the Smith Chart. Finally, a range of covered Γ_{ANT} will emerge on the Smith Chart after going through all settings.

Fig. 2.13 shows the measured values of Γ_{ANT} (in blue color) that can be balanced at $f_{TX} = 6$GHz plotted on the Smith Chart for all Block TX settings. A VSWR=2 circle is covered for antenna impedance changes over user interaction in both cases. Therefore, this duplexer guarantees isolation at f_{TX} in all cases.

To check isolation in the RX band, a nominal 50Ω, as well as four corner cases that lie on the circle of VSWR=2 is considered. Since the antenna impedance will be different at $f_{RX} = 6.5$GHz, for each case it is explored that how large a difference $\Delta\Gamma_{ANT}$ can be balanced by selecting capacitors in the Block RX only. Fig. 2.13 plots the measured results (in red color). These plots show that the Block RX can balance an antenna with Q_{ANT} as high as 4.3, so long as $\Gamma_{ANT}(f_{TX})$ locates within VSWR=2. To cover 5 to 7 GHz for this Wi-Fi standard, Q_{ANT}

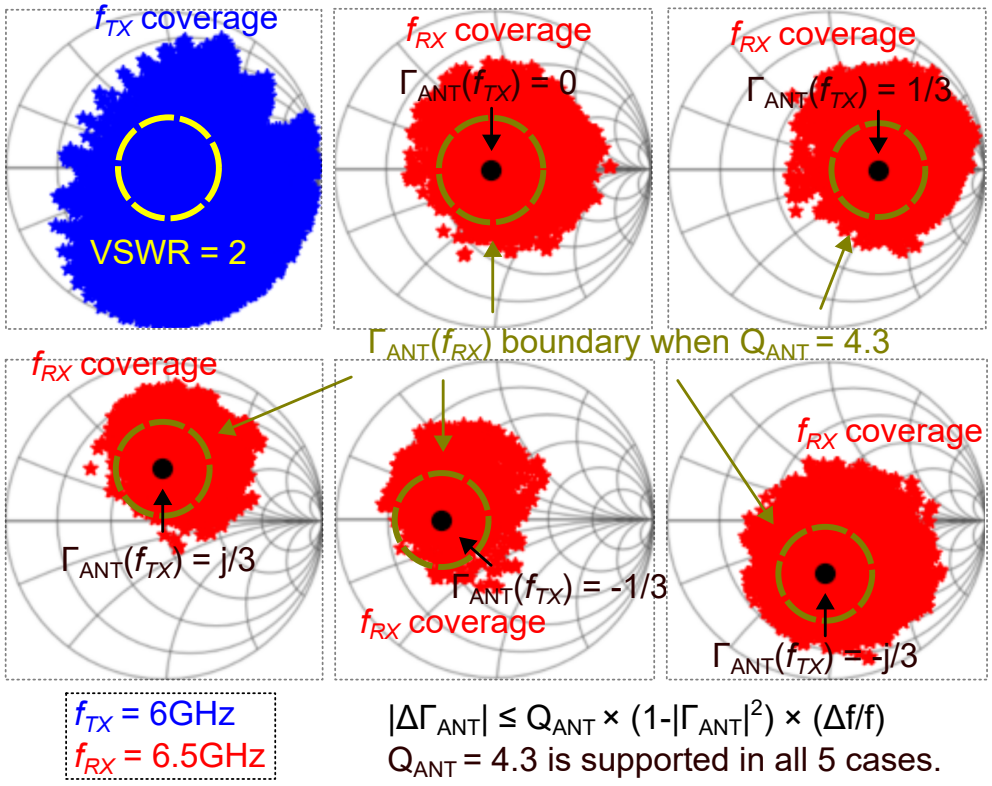


Figure 2.13: Measured $Z_{ANT}(f_{TX} = 6\text{GHz})$ that can be balanced by EBD (blue). For $Z_{ANT}(f_{TX})$ at five locations within VSWR=2, the measured range (red) of $Z_{ANT}(f_{RX} = 6.5\text{GHz})$ in the RX band that can be balanced.

must in any case be ≤ 3. Therefore, the EBD guarantees isolation of the TX noise at f_{RX} in all cases.

Fig. 2.14 shows another case where $f_{TX} = 5\text{GHz}$ and $f_{RX} = 6\text{GHz}$. At f_{TX}, a VSWR=2 circle is still covered. At f_{RX}, the covered $\Delta\Gamma_{ANT}$ range saturates because the frequency offset is now large (20%).

Fig. 2.15 illustrates the typical independent dual-band balancing. Three different settings in Block RX sweep across RX band, while maintaining isolation at fixed f_{TX}. Thanks to the high isolation of the intermediate filter at f_{TX}, the independent balancing is achieved.

The BW of 40dB isolation is also measured using a statistical approach. First, settings in the balance network are swept randomly for >500 points. Then the

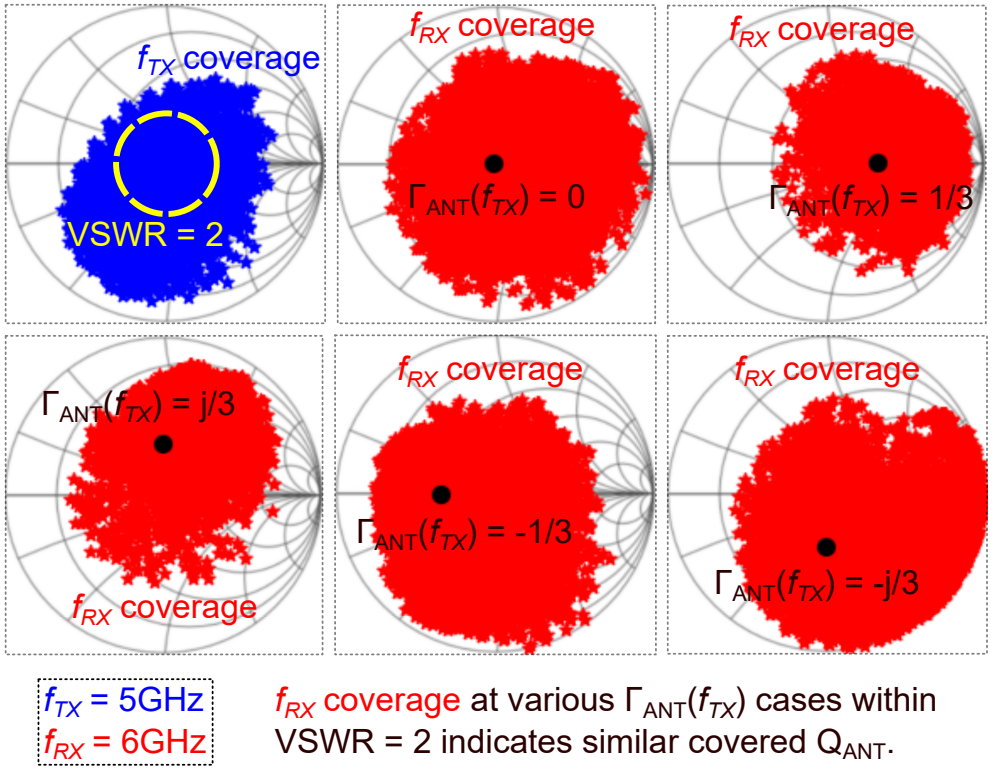


Figure 2.14: Measured $Z_{ANT}(f_{TX} = 5\text{GHz})$ that can be balanced by EBD (blue). For $Z_{ANT}(f_{TX})$ at five locations within VSWR=2, the measured range (red) of $Z_{ANT}(f_{RX} = 6\text{GHz})$ in the RX band that can be balanced.

isolation BW for each setting is measured if a notch with maximum isolation $\geq$43dB appears. The 3dB margin is used to filter out those which barely meet the 40dB isolation requirement and disturb the statistical results. The measured average 40dB isolation BW is $\mu_{\text{BW}} = 81\text{MHz}$, and the standard deviation is $\sigma_{\text{BW}} = 20\text{MHz}$. The measured results show that there is $\sim 50\%$ chance that a setting, which balances the HT at one frequency instead of the entire channel BW, achieves the isolation BW requirement (80MHz). However, there are multiple settings (>10) that can balance the HT thanks to the redundant tuning degree of freedom in the balance network. Therefore, the probability that all settings fail to meet the isolation BW requirement is very low. This 40dB isolation is guaranteed over an 80MHz channel BW.

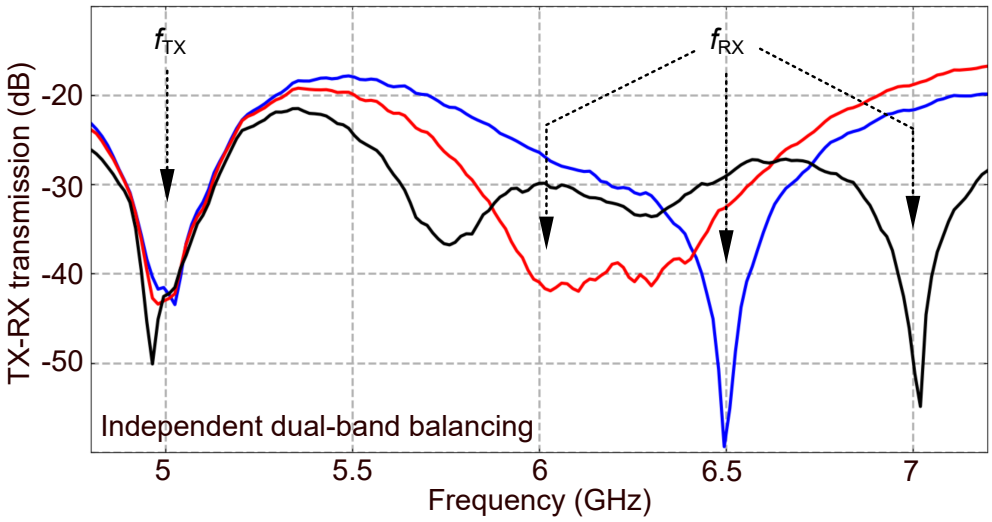


Figure 2.15: Measured dual-band TX-RX isolation demonstrating independent balancing.

2.6.2 Losses and Linearity

Fig. 2.16 shows the measured EBD loss in the TX path, loss in the RX path (=RX noise figure for this reciprocal circuit). Sub-4dB IL is achieved for both TX and RX paths. Fig. 2.16 shows the measured return loss at TX and RX ports. Good matching (< -10dB) is achieved at both ports. During the measurement, TX-RX isolation is set >40dB and the loss in the Block TX-1 is kept at a minimum (0.2dB at 6GHz).

The switch FETs in the EBD contributes a small nonlinearity. When a switch is turned on, its nonlinear on-resistance will contribute a small nonlinearity. On the other hand, when a switch is turned off, its nonlinear off-state capacitance (gate and body capacitance) will contribute a small nonlinearity. [14,34] mentioned that the resistors can introduce small nonlinearity due to self-heating. The resistor in Block RX does not introduce nonlinearity because the majority TX power has been filtered, while the tunable resistor in the resonators may introduce small nonlinearity. Fig. 2.16 shows the IIP3 measurement with all switches ON and OFF for EBD TX and RX paths.

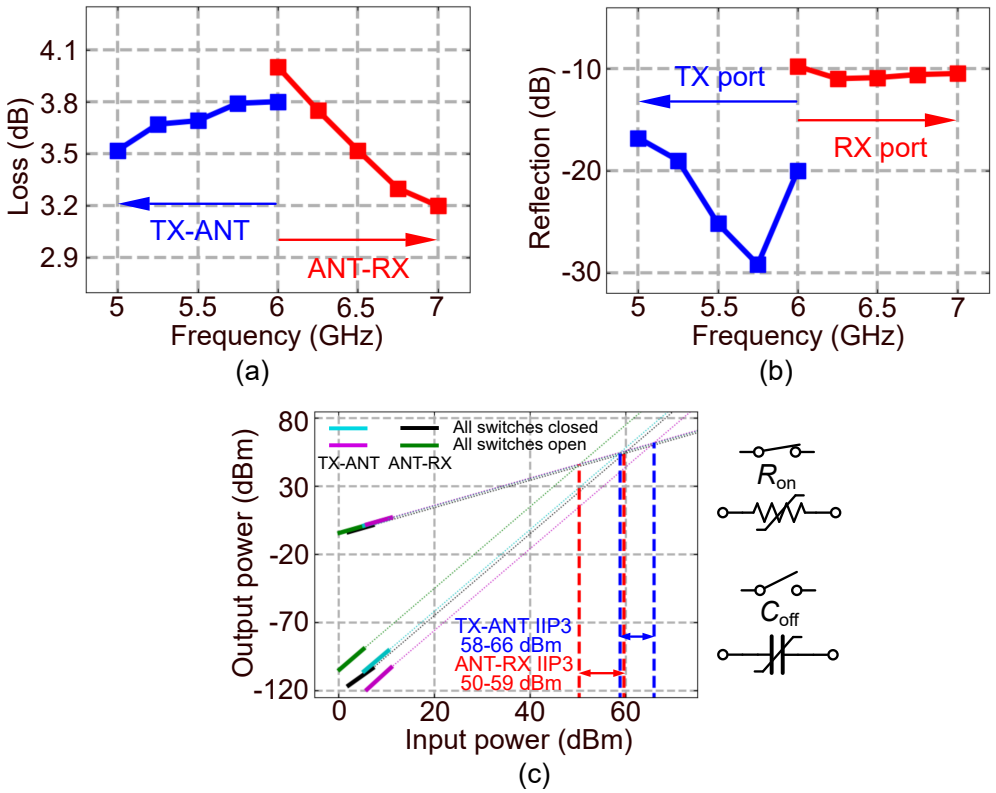


Figure 2.16: Measured EBD TX and RX paths (a) insertion loss (IL); (b) return loss; (c) IIP3.

2.6.3 Large Signal Measurement

The EBD must operate under high TX power level, therefore high power measurements are conducted to verify the EBD functionality.

Fig.2.17 shows the EBD TX path output power (ANT port) versus input power (TX port). The measured IL at high power level is consistent with the small-signal IL (3.5dB at 5GHz). The IL does not degrade under 27.5dBm TX power, or 24dBm at the antenna port. The power handling meets the specification (27dBm TX power).

Fig. 2.17 shows the high-power impact on the established isolation at small signal. The isolation degrades a small amount (<2dB) with $\sim$ 27dBm TX power. This is because the capacitance and resistance may change by a small amount due

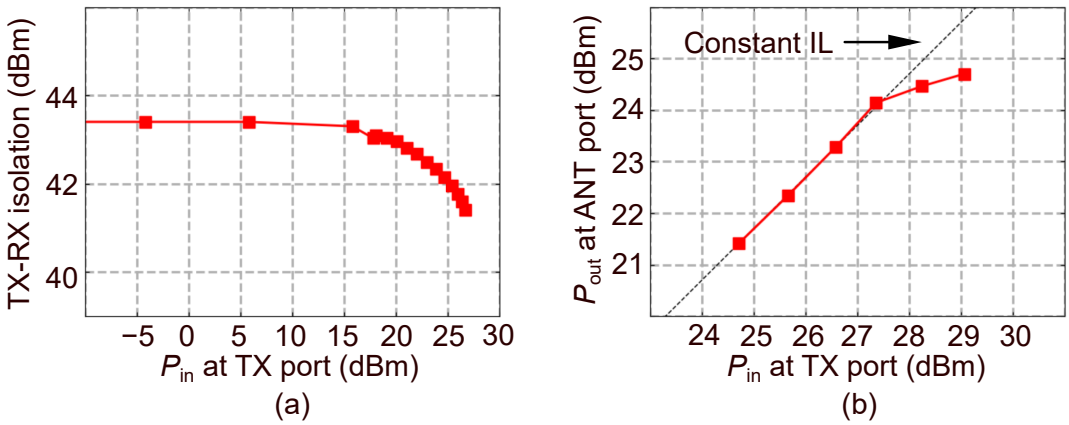


Figure 2.17: Measured EBD (a) TX power impact on established TX-RX isolation and (b) TX path power handling.

to large voltage swing and heating [14, 34]. The small isolation degradation can be easily resumed by adjusting the balance network settings.

2.7 Discussions and Conclusions

Table 2.1 summarizes the performance of this EBD. It meets all design specifications. It is the first fully integrated duplexer that supports ALL the requirements shown below for FDD applications:

1. It can track antenna impedance changes over user interaction.
2. It can track antenna impedance changes over frequency, which enables the dual-band capability. This duplexer supports actual antennas, whose impedance change with limited rate (Q_{ANT}) but arbitrary phase. Among all prior arts on dual-band EBD, only [6] supports this feature.
3. It provides sufficient dual-band isolation between the TX and the RX.
4. It operates at practical TX power level.

Table 2.1: EBD comparison table.

	[1]	**[13]**	**[5]**	**[3]**	**[6]**	**This work**
Architecture	EBD+LNA	EBD	EBD+RX	EBD+ N-path LNA	BAL	EBD
CMOS Technology	65nm	180nm SOI	65nm	180nm SOI	180nm SOI	65nm SOI
Frequency (GHz)	1.5-2.2	1.9-2.2	1.7-2.2	0.7-1	0.7-1	TX: 5-6/RX: 6-7
Area (mm^2)	0.2	1.75	2.2	14.4	8.28	2.3
Dual-Band Capability	No	No	Yes[1]	Yes	Yes	Yes
Supported ANT VSWR	1.3:1	1.5:1	2:1	2.3:1	1.1:1	2:1
Supported Q_{ANT}	-	-	-	-	4.8	4.3
Isolation (dB)	>50	>50	>50	>40	>40	>40
EBD TX path loss (dB)	2.5	<3.7	4.5	<3.5	-	<3.8
EBD RX path loss (dB)	5 (casc.)	<3.9	6.7 (casc.)	<9.2 (casc.)	-	<4
EBD TX path IIP3 (dBm)	-	>70	54	70	-	>58
EBD RX path IIP3 (dBm)	-	>72	-	26.7 (casc.)	-	>50
P_{max} at ANT (dBm)	-	27	27	15.5[2]	-	24

[1]: only for specific cases.

[2]: RX 1dB compression.

CHAPTER 3

Properties of Finite-Q LCR Networks

3.1 Introduction

The balance network is the most critical part of the dual-band EBD design. It is required to synthesize antenna impedance variations over user interaction and frequency response simultaneously. Failing to do so will disqualify the EBD for FDD applications. The balance network is required to work under high TX power level (e.g., +27dBm), therefore the only practical way to realize the balance network on a chip is using passive elements (i.e. LCRs).

Unlike lossless LC passive networks, where the s-parameters can be designed to change in arbitrarily high rate with respect to frequency, the s-parameters of the lossy passive network constructed with lossy LCs can only change with frequency with a limited rate. For dual-band designs such as the balance network in Chapter 2, it is necessary to study the properties of such passive networks focusing on their frequency responses. For example, the theoretical limit of tuning capability on $\frac{\partial \Gamma_{BAL}}{\partial f/f}$ for dual-band balancing is needed. Moreover, to realize the $\frac{\partial \Gamma_{BAL}}{\partial f/f}$ tunability, an intermediate filter $[S(f)]$ is introduced in Chapter 2, and a high $|\frac{\partial s_{21}}{\partial f/f}|$ is desired. Guidance is needed for the optimum design of the above passive networks.

Traditional filter synthesis design methodology works well using only high-Q elements. When it is applied to low-Q elements, the discrepancy between the design and the realization will fail its original purpose (e.g., Fig. 3.1). Because on-chip inductors and capacitors are fundamentally low-Q elements ($Q \sim 10^1$),

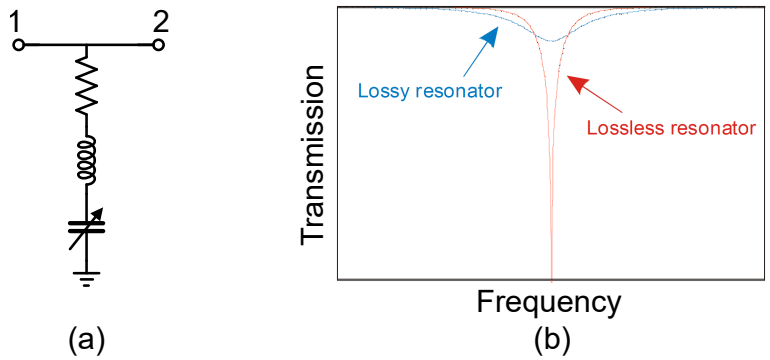


Figure 3.1: (a) Conventional bandstop resonator; (b) Effect of finite Q [7].

traditional filter synthesis will provide little help.

This chapter analyzes the properties of general finite-Q LCR networks focusing on the following questions:

1. What is the maximum $|\frac{\partial s_{21}}{\partial f/f}|$ of any LCR two-port that uses finite-Q elements? How to design such a filter with sharp $|s_{21}|$ transition?

2. What is the maximum $|\frac{\partial s_{11}}{\partial f/f}|$ of any LCR one-port that uses finite-Q elements? What is the dual-band capability limit of such a balance network?

3. What about a general N-port? What conclusions may be useful?

3.2 Electric and Magnetic Power in One-Port Networks

Before going into detail, an important statement must be proved first, which is used in the analysis later on. The statement is:

- The electric (capacitive) and magnetic (inductive) energy stored in an one-port LCR network is identical (at a given frequency f_0) if one of the following conditions is true:

    1. Input reflection coefficient Γ_1 is real.

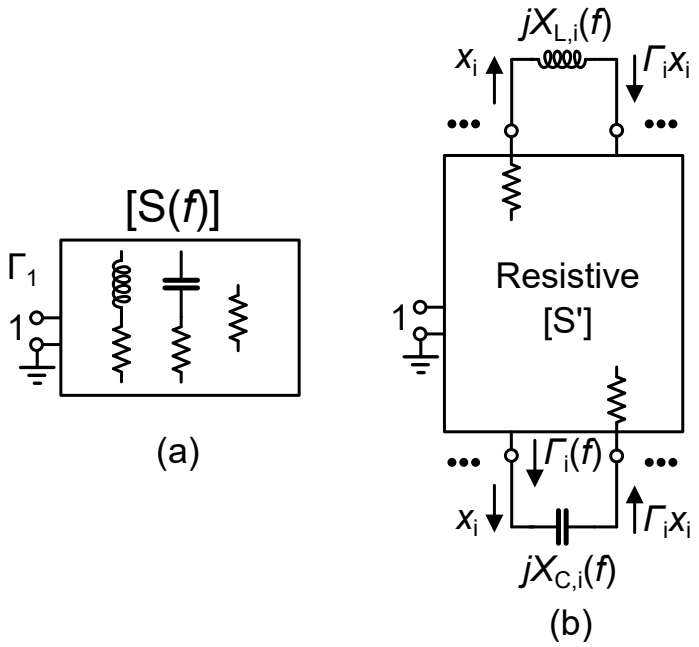


Figure 3.2: (a) An one-port finite-Q LCR network; (b) Transformed model for analysis purpose.

2. Imaginary part of input reflection coefficient $\Im(\Gamma_1)$ is much smaller than the stored power in the network.

3.2.1 Relation Between Impedance and Stored Power

In order to prove the statement, the original one-port LCR network is transformed into a new network, as shown in Fig. 3.2. The inductors and capacitors are moved out of the original network and load the new ports, and lossy resistors stay in the new network $[S']$. Thus s-parameters of the network $[S']$ are real. The characteristic impedance of each port is not important for this proof, and can be chosen arbitrarily. Note that the transformation does not change the circuit topology.

The incoming and outgoing waves can be related by s-parameters, shown in

equations:

$$\begin{pmatrix} \Gamma_1 \\ x_2 \\ ... \\ x_N \end{pmatrix} = \begin{pmatrix} s'_{11} & ... & s'_{1N} \\ s'_{21} & ... & s'_{2N} \\ ... & ... & ... \\ s'_{N1} & ... & s'_{NN} \end{pmatrix} \begin{pmatrix} 1 \\ \Gamma_2 x_2 \\ ... \\ \Gamma_N x_N \end{pmatrix} \tag{3.1}$$

where x_l is the outgoing wave at port-l, and Γ_l is the reflection coefficient at port-l. Equation 3.1 leads to the following relations:

$$\begin{pmatrix} x_2 \\ ... \\ x_N \end{pmatrix} = \begin{pmatrix} s'_{22} & ... & s'_{2N} \\ ... & ... & ... \\ s'_{N2} & ... & s'_{NN} \end{pmatrix} \begin{pmatrix} \Gamma_2 x_2 \\ ... \\ \Gamma_N x_N \end{pmatrix} + \begin{pmatrix} s'_{21} \\ ... \\ s'_{N1} \end{pmatrix} \tag{3.2}$$

$$\Gamma_1 = s'_{11} + \sum_{l=2}^{N} \left(s'_{1l} \times x_l \Gamma_l \right) \tag{3.3}$$

Note that $[S']$ is real, thus:

$$\begin{pmatrix} s'_{21} \\ ... \\ s'_{N1} \end{pmatrix} = \begin{pmatrix} s'_{22} & ... & s'_{2N} \\ ... & ... & ... \\ s'_{N2} & ... & s'_{NN} \end{pmatrix} \begin{pmatrix} \Gamma_2^* x_2^* \\ ... \\ \Gamma_N^* x_N^* \end{pmatrix} - \begin{pmatrix} x_2^* \\ ... \\ x_N^* \end{pmatrix} \tag{3.4}$$

where x_l^* is the complex conjugate of x_l. Applying equation 3.4, equation 3.3 can rewritten:

$$\Gamma_1 = s'_{11} + \sum_{l=2}^{N} \sum_{k=2}^{N} \left(|x_l|^2 \Gamma_l - s'_{lk} \times \Gamma_k^* \Gamma_l \times x_k^* x_l \right) \tag{3.5}$$

Note that the term $\sum_{l=2}^{N} \sum_{k=2}^{N} \left(s'_{lk} \times \Gamma_k^* \Gamma_l \times x_k^* x_l \right)$ in equation 3.5 is real. Thus, the imaginary part of equation 3.5 is:

$$\Im(\Gamma_1) = \sum_{inductors} \left(|x_l|^2 \times \Im(\Gamma_l) \right) + \sum_{capacitors} \left(|x_l|^2 \times \Im(\Gamma_l) \right) \tag{3.6}$$

where the sum is split into sums over all inductive ports and all capacitive ports. Notice that since $\Im(\Gamma_l) \geq 0$ is true for all inductor loaded ports and $\Im(\Gamma_l) \leq 0$ is

true for all capacitor loaded ports, the two terms in equation 3.6 presents opposite polarities. Thus, equation 3.6 can be rewritten:

$$\Im(\Gamma_1) = \sum_{inductors} \left(|x_l|^2 \times \sin\theta_l \right) - \sum_{capacitors} \left(|x_l|^2 \times \sin\theta_l \right) \tag{3.7}$$

where $|\Gamma_l| = 1$, and the $\theta_l \triangleq |\angle(\Gamma_l)|$ is the angle between Γ_l and the real axis, $0 \leq \theta_l \leq \pi$.

Now, the average power that is stored in reactance X_l is:

$$\begin{aligned} P_{av} = \frac{V_{peak}^2}{2X_l} &= \left| \frac{(V_{peak}^+)^2}{2Z_0} \times \frac{(1+\Gamma_l)^2}{\frac{1+\Gamma_l}{1-\Gamma_l}} \right| \\ &= |x_l|^2 \times |1-\Gamma_l^2| \\ &= 2|x_l|^2 \times \sin\theta_l \end{aligned} \tag{3.8}$$

where V_{peak} is the voltage at port-l, and Z_0 is the characteristic impedance at port-l and V_{peak}^+ is the outgoing wave voltage at port-l. Note that equation 3.8 is normalized to the incident power, thus dimensionless.

Thus, equation 3.7 can be expressed by the power stored in the inductors and capacitors:

$$\Im(\Gamma_1) = \frac{1}{2} P_{inductors} - \frac{1}{2} P_{capacitors} \tag{3.9}$$

where $P_{inductors}$ and $P_{capacitors}$ are normalized to incident power at port-1.

3.2.2 Discussions

For the input reflection coefficient Γ_1, there are two possibilities:

1. $\Im(\Gamma_1) = 0$.

   In this case, the electric power is equal to the magnetic power, according to equation 3.9:

   $$P_{capacitors} = P_{inductors} \tag{3.10}$$

2. $\Im(\Gamma_1) \neq 0$.

   In this case, the electric power is not equal to the magnetic power. However, it is still possible that $P_{inductors} \approx P_{capacitors}$, when the reflected power at port-1 is much smaller than the stored power in the network:

$$\begin{aligned} \Im(\Gamma_1) &\ll P_{inductors} \\ \Im(\Gamma_1) &\ll P_{capacitors} \end{aligned} \tag{3.11}$$

   For a network without standalone resistors (i.e. resistors only from lossy LCs), equation 3.11 is easy to satisfy when quality factor $Q_L \| Q_C$ is large:

$$\begin{aligned} \Im(\Gamma_1) &\ll (1 - |\Gamma_1|^2) \times \big(\min (Q_L \| Q_C) \big) \leq P_{inductors} \\ \Im(\Gamma_1) &\ll (1 - |\Gamma_1|^2) \times \big(\min (Q_L \| Q_C) \big) \leq P_{capacitors} \end{aligned} \tag{3.12}$$

   The proof is finished.

For most cases, where the networks are built from lossy LCs which $Q \gg 1$ and the input are well-matched, the magnetic power and electric power stored in the inductors and capacitors are equal.

3.3 Two-Port Filter Design Using Finite-Q LCR Elements

The intermediate two-port filter in the dual-band balance network (shown in Fig. 2.4) plays an essential role for independent impedance synthesis or balancing. It must reject the transmission between its two ports at TX band, and allow at RX band. This leaves the challenge of realizing a filter with sharp $|s_{21}|$ transition between TX band and RX band, which can be very close.

From the following analysis, two questions are answered in detail:

1. What is the maximum $\left| \frac{\partial s_{21}}{\partial f / f} \right|$ that can be achieved using finite-Q elements?
2. How to achieve the maximum $|s_{21}|$ transition rate $\left| \frac{\partial s_{21}}{\partial f / f} \right|_{max}$?

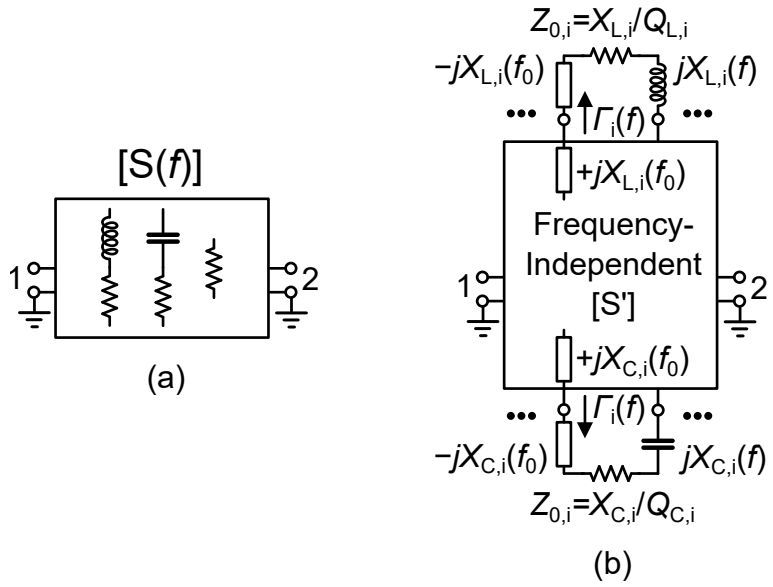


Figure 3.3: (a) A two-port finite-Q LCR network; (b) Transformed model for analysis purpose.

3.3.1 Maximum $\left|\frac{\partial s_{21}}{\partial f/f}\right|$ of Finite-Q LCR Networks

In order to analyze $\left|\frac{\partial s_{21}}{\partial f/f}\right|$, the original two-port is transformed into a new network as shown in Fig. 3.3. The inductors, capacitors, and their lossy resistors are moved outside of the original network, and new ports are defined for each inductor and capacitor. The characteristic impedance of each new port is defined to be equal to the lossy resistance of the reactor:

$$\begin{aligned} Z_{0,i} &= X_{L,i}/Q_{L,i} \\ Z_{0,i} &= X_{C,i}/Q_{C,i} \end{aligned} \tag{3.13}$$

where $Q_{L,i}$, $Q_{C,i}$ are the quality factor of each reactor. Then a pair of constant reactors $\pm jX_{L,i}(f_0)$ are added to each inductor loaded port (also the same for each capacitor loaded port), such that the port is impedance-matched (i.e. $\Gamma_i(f_0) = 0$) at f_0, and the circuit topology is not changed. Note that the new network $[S']$ is frequency-independent, since all frequency-dependent elements are outside of the network $[S']$.

The incoming and outgoing waves can be related by s-parameters:

$$\begin{pmatrix} s_{11} \\ s_{21} \\ x_3 \\ ... \\ x_N \end{pmatrix} = \begin{pmatrix} s'_{11} & ... & ... & s'_{1N} \\ s'_{21} & ... & ... & s'_{2N} \\ s'_{31} & ... & ... & s'_{3N} \\ ... & ... & ... & ... \\ s'_{N1} & ... & ... & s'_{NN} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \Gamma_3 x_3 \\ ... \\ \Gamma_N x_N \end{pmatrix} \tag{3.14}$$

where x_i is the outgoing wave at port-i and Γ_i is the reflection coefficient at port-i. Notice that $\Gamma_i \approx 0$ in the vicinity of f_0 because of the way the new network is constructed:

$$x_i|_{(|\Gamma_l| \ll 1, \forall\, 3 \leq l \leq N)} = s'_{i1} \tag{3.15}$$

Applying equation 3.15 to equation 3.14:

$$s_{21} = s'_{21} + \sum_{i=3}^{N} s'_{i1} \times \Gamma_i \times s'_{2i} \tag{3.16}$$

Take the derivative of equation 3.16 and note $[S']$ is frequency-independent:

$$\left| \frac{\partial s_{21}}{\partial f/f} \right|_{f_0} = \left| \sum_{i=3}^{N} s'_{i1} \times \left(\frac{\partial \Gamma_i}{\partial f/f} \right)_{f_0} \times s'_{2i} \right| \tag{3.17}$$

This is the expression that gives the $|s_{21}|$ transition rate. To find the maximum, several inequalities should be applied. First, Cauchy-Schwarz inequalities $|a+b| \leq |a| + |b|$ and $|2ab| \leq (|a|^2 + |b|^2)$ are applied:

$$\left| \frac{\partial s_{21}}{\partial f/f} \right|_{f_0} \leq \sum_{i=3}^{N} \left(\frac{|s'_{i1}|^2 + |s'_{2i}|^2}{2} \times \left| \frac{\partial \Gamma_i}{\partial f/f} \right|_{f_0} \right) \tag{3.18}$$

The inequality will hold as an equality when:

$$\begin{gathered} |s'_{i1}| = |s'_{2i}| \\ \angle \left(s'_{i1} \times \left(\frac{\partial \Gamma_i}{\partial f/f} \right)_{f_0} \times s'_{2i} \right) = \angle \left(s'_{k1} \times \left(\frac{\partial \Gamma_k}{\partial f/f} \right)_{f_0} \times s'_{2k} \right), \forall\, i, k \in [3, N] \end{gathered} \tag{3.19}$$

Notice that for capacitors and inductors:

$$\begin{aligned}\left|\frac{\partial\Gamma_i}{\partial f/f}\right|_{f_0} &= \lim_{\Delta X_{L,i}\to 0}\left|\frac{j\Delta X_{L,i}\times\frac{f}{\Delta f}}{2Z_{0,i}+j\Delta X_{L,i}}\right| = \left|\frac{X_{L,i}}{2Z_{0,i}}\right| = \frac{Q_{L,i}}{2}\\ \left|\frac{\partial\Gamma_i}{\partial f/f}\right|_{f_0} &= \lim_{\Delta X_{C,i}\to 0}\left|\frac{j\Delta X_{C,i}\times\frac{f}{\Delta f}}{2Z_{0,i}+j\Delta X_{C,i}}\right| = \left|\frac{X_{C,i}}{2Z_{0,i}}\right| = \frac{Q_{C,i}}{2}\end{aligned} \tag{3.20}$$

Applying equation 3.20 to equation 3.18:

$$\begin{aligned}\left|\frac{\partial s_{21}}{\partial f/f}\right|_{f_0} \le &\sum_{inductors}\left(\frac{|s'_{i1}|^2}{2}\times\frac{Q_{L,i}}{2}\right)+\sum_{capacitors}\left(\frac{|s'_{i1}|^2}{2}\times\frac{Q_{C,i}}{2}\right)\\ &+\sum_{inductors}\left(\frac{|s'_{i2}|^2}{2}\times\frac{Q_{L,i}}{2}\right)+\sum_{capacitors}\left(\frac{|s'_{i2}|^2}{2}\times\frac{Q_{C,i}}{2}\right)\end{aligned} \tag{3.21}$$

Equation 3.21 can be rewritten in terms of stored power:

$$\begin{aligned}\left|\frac{\partial s_{21}}{\partial f/f}\right|_{f_0} \le \frac{1}{4}\Bigg(&\sum_{Source\ at\ port\text{-}1}P_{inductors}+\sum_{Source\ at\ port\text{-}1}P_{capacitors}\\ &+\sum_{Source\ at\ port\text{-}2}P_{inductors}+\sum_{Source\ at\ port\text{-}2}P_{capacitors}\Bigg)\end{aligned} \tag{3.22}$$

From equation 3.10, the stored electric and magnetic power equals (will justify later):

$$\sum_{inductors}\left(|s'_{il}|^2\times Q_{L,i}\right)=\sum_{capacitors}\left(|s'_{il}|^2\times Q_{C,i}\right),\ l=1,2 \tag{3.23}$$

To calculate the maximum of power stored in the network, power conservation must be considered:

$$\left(\sum_{inductors}|s'_{il}|^2\right)+\left(\sum_{capacitors}|s'_{il}|^2\right)\le 1-|s'_{21}|^2-|s'_{ll}|^2,\ l=1,2 \tag{3.24}$$

The inequality will hold as an equality when the power only dissipates in the resistors of lossy LCs. Now equation 3.23 can be rewritten:

$$\left(\sum_{inductors}|s'_{il}|^2\right)\times Q_{L,ave}=\left(\sum_{capacitors}|s'_{il}|^2\right)\times Q_{C,ave},\ l=1,2 \tag{3.25}$$

where $Q_{L,ave}$ and $Q_{C,ave}$ are the average quality factor when weighted by stored power. With the constraints set by equation 3.24, the maximum stored power can be calculated:

$$\begin{aligned}
\left(\sum_{inductors}|s'_{il}|^2\right)\times Q_{L,ave} &= \left(\frac{Q_{C,ave}\times(1-|s'_{21}|^2-|s'_{ll}|^2)}{Q_{L,ave}+Q_{C,ave}}\right)\times Q_{L,ave} \\
&\leq \left(Q_{L,max}\|Q_{C,max}\right)\times\left(1-|s'_{21}|^2-|s'_{ll}|^2\right) \\
\left(\sum_{capacitors}|s'_{il}|^2\right)\times Q_{C,ave} &= \left(\frac{Q_{L,ave}\times(1-|s'_{21}|^2-|s'_{ll}|^2)}{Q_{L,ave}+Q_{C,ave}}\right)\times Q_{C,ave} \\
&\leq \left(Q_{L,max}\|Q_{C,max}\right)\times\left(1-|s'_{21}|^2-|s'_{ll}|^2\right)
\end{aligned} \tag{3.26}$$

where $l=1,2$, and $Q_{L,max}$ ($Q_{C,max}$) is the maximum inductor (capacitor) quality factor among all lossy inductors (capacitors). The inequality will hold as an equality when all power is stored in the elements with maximum Q. Note that $s'_{21}=s_{21}$, $s'_{11}=s_{11}$ and $s'_{22}=s_{22}$ at f_0. Applying equation 3.26 to equation 3.22:

$$\begin{aligned}
\left|\frac{\partial s_{21}}{\partial f/f}\right|_{f_0} &\leq \left(Q_{L,max}\|Q_{C,max}\right)\times\left(1-|s^2_{21}|_{f_0}-\frac{|s^2_{11}|_{f_0}+|s^2_{22}|_{f_0}}{2}\right) \\
&\leq \left(Q_{L,max}\|Q_{C,max}\right)
\end{aligned} \tag{3.27}$$

The last inequality will hold as an equality when $s_{11}=s_{22}=s_{21}=0$.

To conclude, when all the following design requirements are satisfied, the maximum $|s_{21}|$ transition rate can be achieved:

1. Matched: $s_{11}=s_{22}=0$.
2. Bandstop: $s_{21}=0$.
3. Symmetrical: $|s'_{i1}|=|s'_{i2}|$, $\forall\, i\in[3,N]$.
4. Power dissipated in elements with maximum Q.
5. All terms add in phase: $\sum_{i=3}^{N} s'_{i1}\times\left(\frac{\partial\Gamma_i}{\partial f/f}\right)_{f_0}\times s'_{2i}$.

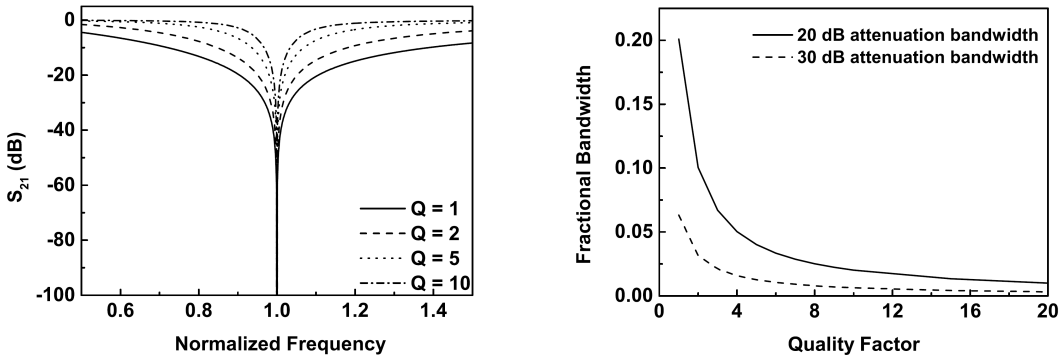


Figure 3.4: Isolation and isolation bandwidth of an absorptive bandstop filter in [9].

The requirements justify the step that derives equation 3.23 because $s_{11} = s_{22} = 0$. The maximum $\left|\frac{\partial s_{21}}{\partial f/f}\right|_{max}$ that can be achieved using finite-Q elements on a chip is:

$$\left|\frac{\partial s_{21}}{\partial f/f}\right|_{max} = \left(Q_{L,max} \| Q_{C,max}\right) \tag{3.28}$$

3.3.2 The Perfectly-Matched Absorptive Bandstop Filter

From equation 3.27, in order to achieve a high $|s_{21}|$ transition rate, the incident power must be dissipated in the 2-port filter with minimum transmission or reflection. This type of filter is known as the absorptive filter [7,9,35–38]. Common examples of such filters are the bridge-T notch filter and the twin-T notch filter. Unlike traditional reflective filters designed by the filter synthesis method, which reflect majority of the incident power at the stopband, the absorptive filters are designed to absorb the incident power at the stopband. Absorptive filter design methodologies are very different from traditional synthesized filters. Their operations, at stopband, are interpreted by passive cancellation of two or more paths, similar to the TX-RX isolation in the EBD, thus the rejection/isolation is not limited by the Q of the elements. However, Q will affect the isolation bandwidth, since different Q gives different $|s_{21}|$ transition rate. An example is shown in Fig. 3.4 [9].

Among many absorptive filters that can be implemented on a chip as the intermediate 2-port filter for the balance network design shown in Chapter 2, this candidate [7], called the perfectly-matched absorptive bandstop filter (Fig. 2.5), shows advantages over others for the following reasons:

1. It achieved the maximum $|s_{21}|$ transition rate.
2. It can be realized by lumped LCs with moderate number of Ls (moderate chip area).
3. It can be easily tuned without introducing other design problems.
4. The impedance of each reactor is realistic to be implemented with good Q.
5. Its s_{11} shows wideband-matching property, or in other words the s_{11} does not change drastically with frequency, which is a requirement for this design (Fig. 2.4).

The operations of the perfectly-matched absorptive bandstop filter are shown in Fig. 2.5. It satisfy all the five requirements to achieve the maximum $|s_{21}|$ transition rate, which is given by equation 3.28.

3.4 One-Port Impedance Synthesizer: Balance Networks

3.4.1 Maximum $\left|\frac{\partial s_{11}}{\partial f/f}\right|$ of Finite-Q LCR Networks

The balance network is a one-port impedance synthesizer, which is required to operate at dual-band. Thus, it is necessary to determine what is the maximum impedance change rate ($|\frac{\partial s_{11}}{\partial f/f}|$) that can be realized by finite-Q LCR networks.

The analysis is similar to the analysis of two-port networks, and the transformation is shown in Fig. 3.5. The only difference is that port-2 is now removed.

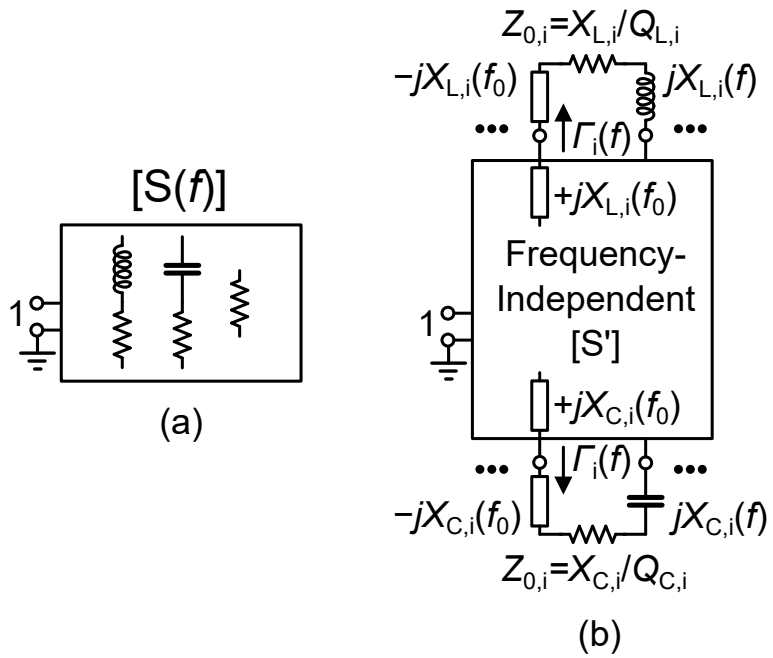


Figure 3.5: (a) An one-port finite-Q LCR network; (b) Transformed model for analysis purpose.

Equation 3.14 can be modified to:

$$\begin{pmatrix} s_{11} \\ x_2 \\ \dots \\ x_N \end{pmatrix} = \begin{pmatrix} s'_{11} & \dots & \dots & s'_{1N} \\ s'_{21} & \dots & \dots & s'_{2N} \\ \dots & \dots & \dots & \dots \\ s'_{N1} & \dots & \dots & s'_{NN} \end{pmatrix} \begin{pmatrix} 1 \\ \Gamma_2 x_2 \\ \dots \\ \Gamma_N x_N \end{pmatrix} \tag{3.29}$$

Note that $\Gamma_i \ll 1$:

$$s_{11} = s'_{11} + \sum_{i=2}^{N} {s'_{i1}}^2 \times \Gamma_i \tag{3.30}$$

Take the derivative of equation 3.30, and apply Cauchy-Schwarz inequalities, energy conservation, equation 3.26 and 3.20, and then follow the same derivation

as the two-port shown earlier:

$$\begin{aligned}\left|\frac{\partial s_{11}}{\partial f/f}\right|_{f_0} &= \left|\sum_{i=2}^{N} {s'_{i1}}^2 \times \left(\frac{\partial \Gamma_i}{\partial f/f}\right)_{f_0}\right| \\ &\leq \frac{1}{2}\left(\sum_{inductors}\left(|s'_{i1}|^2 \times Q_{L,i}\right) + \sum_{capacitors}\left(|s'_{i1}|^2 \times Q_{C,i}\right)\right) \\ &\leq \left(Q_{L,max} \| Q_{C,max}\right) \times \left(1 - |s_{11}^2|_{f_0}\right)\end{aligned} \tag{3.31}$$

where $s_{11} = s'_{11}$ at f_0, and $Q_{L,max}$ ($Q_{C,max}$) is the maximum inductor (capacitor) quality factor among all lossy inductors (capacitors). Note that now the energy conservation is:

$$\left(\sum_{inductors} |s'_{i1}|^2\right) + \left(\sum_{capacitors} |s'_{i1}|^2\right) \leq 1 - |s'_{11}|^2 \tag{3.32}$$

To achieve the maximum $|s_{11}|$ transition rate, several requirements must be met:

1. Power dissipated in elements with maximum Q.
2. All terms add in phase: $\sum_{i=2}^{N} {s'_{i1}}^2 \times \Gamma_i$.

3.4.2 Balance Networks and Limits on Q_{ANT} Coverage

Balance network (Γ_{BAL}) is a one-port network, thus its impedance change over frequency follows equation 3.31:

$$\left|\frac{\partial \Gamma_{BAL}}{\partial f/f}\right|_{f_0} \leq \left(Q_{L,max} \| Q_{C,max}\right) \times \left(1 - |\Gamma_{BAL}^2|_{f_0}\right) \tag{3.33}$$

Equation 3.33 gives the capability limit of the on-chip balance network to tolerate the antenna impedance frequency response. Now assuming the antenna can also be written in a similar form:

$$\left|\frac{\partial \Gamma_{ANT}}{\partial f/f}\right|_{f_0} \leq \left(Q_{ANT}\right) \times \left(1 - |\Gamma_{ANT}^2|_{f_0}\right) \tag{3.34}$$

where Q_{ANT} is the antenna quality factor. Then theoretically, the balance network should be able to cover Q_{ANT} up to the $Q_{L,max} \| Q_{C,max}$ of the on-chip elements:

$$\underbrace{Q_{ANT}}_{\text{covered}} \leq \left(Q_{L,max} \| Q_{C,max}\right) \tag{3.35}$$

Note that equation 3.34 may not be correct since Q_{ANT} is a lumped property which is not equal to maximum Q, in equation 3.31, of the reactive elements connected to the antenna port. The latter (e.g., the equivalent LCs in the cable connecting the antenna and the antenna port) may be extremely high-Q, and thus increase $\left|\frac{\partial \Gamma_{ANT}}{\partial f/f}\right|$. Thus, there is another requirement for the antenna in the EBD in Chapter 2:

- The antenna should be placed close to the EBD chip with a minimum length of cable or transmission line.

In the EBD design shown in Chapter 2, there are several factors that degrade the Q_{ANT} coverage ($Q_{ANT} \leq 4.3$) comparing to the maximum ($Q_{L,max} \| Q_{C,max}$) ≈ 20 given by equation 3.35:

1. Loss in the Block RX ($|\Gamma_3| \leq 0.5$, shown in Fig. 2.9) introduces $\sim$ 6dB loss on Q_{ANT} coverage.

2. Loss in the Block TX and the quadrature hybrid introduces $\sim 2 \times 3 = 6$dB loss on Q_{ANT} coverage. Note that the incident wave travel through the Block TX and the quadrature hybrid twice before being reflected by the balance network.

Clearly, there is still room to improve the Q_{ANT} coverage of the balance network.

3.4.3 Discussions on Balance Network Design

The balance network is the most important design block in the dual-band EBD. It dictates almost all the EBD performance metrics (except losses), which include occupied area, covered ANT VSWR, covered Q_{ANT}, isolation, IIP3 and power handling.

[1, 12] implement the simplest balance network possible using two tuning components. It is only able to balance the antenna at a single band, and achieves the minimum area. [13, 14] implement more complicated balance networks for single-band balancing. [5, 11] are essentially single-band balance networks since they do not provide the general dual-band capability discussed in Chapter 2. Those approaches all use resistors that dissipates power.

The analysis in this chapter reveals that the dissipated power in the network could be used more wisely that gives additional performance enhancement. The resistors can be replaced by networks that enable dual-band balancing. The power dissipated in the networks can then be utilized to improve Q_{ANT} coverage. The limits on Q_{ANT} coverage against the power dissipation in the balance network are given in equation 3.31 and 3.35.

[6] or [39] is the first and only work realizing the problem with resistors. However, they did not find the correct approaches that take advantage of the dissipated power for maximum dual-band coverage, which is given by the analysis in this chapter. They also did not find a network that implements the independent dual-band balancing, which is essential to realize both high ANT VSWR and Q_{ANT} coverage. The lack of design methodology results in large occupied area (ten inductors in the balance network) and poor ANT VSWR coverage for dual-band balancing. It also results in high voltage swing in the internal nodes when the resistors are absent, which limits the power handling of the balance network.

The dual-band balance network in Chapter 2 realizes the limitation and follows

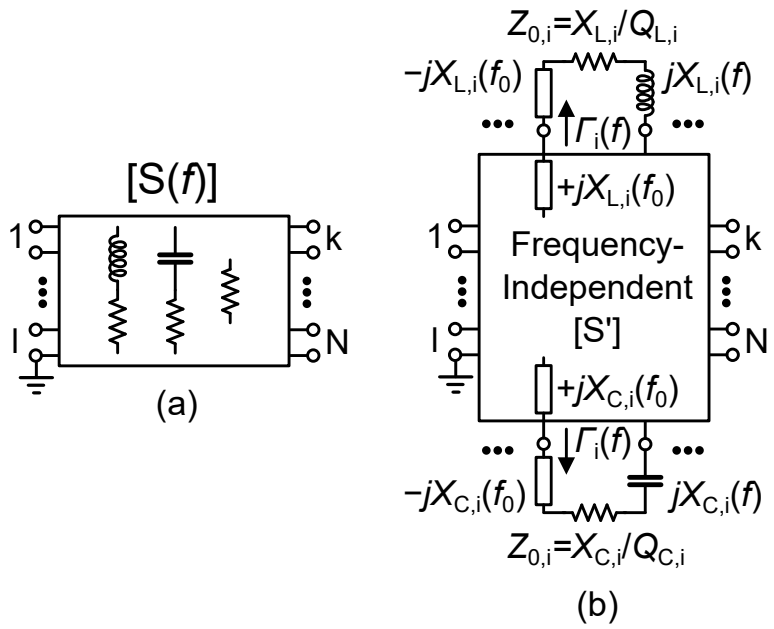


Figure 3.6: (a) A N-port finite-Q LCR network; (b) Transformed model for analysis purpose.

the design guidelines given by the analysis. It achieves both high ANT VSWR and Q_{ANT} coverage simultaneously for the first time. It uses less inductors than [6,39]. The voltage swing of internal nodes is well-defined and is confined to small amplification.

3.5 N-Port Networks

The analysis on the maximum $|s_{21}|$ transition rate for the two-port networks can be generalized to N-port networks with a few modifications. As shown in Fig. 3.6, the N-port is transformed into a similar network as in two-port analysis. The total number of ports in the new network is M, including the original N ports and the ports loaded by inductors or capacitors.

Consider a general case: the transmission from port-l to port-k, where $1 \leq$

$l, k \leq N$. The wave equations are:

$$\begin{pmatrix} s_{1l} \\ \dots \\ s_{ll} \\ \dots \\ s_{kl} \\ \dots \\ s_{Nl} \\ x_{N+1} \\ \dots \\ x_M \end{pmatrix} = \begin{pmatrix} s'_{11} & \dots & \dots & s'_{1M} \\ \dots & \dots & \dots & \dots \\ s'_{l1} & \dots & \dots & s'_{lM} \\ \dots & \dots & \dots & \dots \\ s'_{k1} & \dots & \dots & s'_{kM} \\ \dots & \dots & \dots & \dots \\ s'_{N1} & \dots & \dots & s'_{NM} \\ s'_{N+1,1} & \dots & \dots & s'_{N+1,M} \\ \dots & \dots & \dots & \dots \\ s'_{M1} & \dots & \dots & s'_{MM} \end{pmatrix} \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \\ \dots \\ 0 \\ \Gamma_{N+1} x_{N+1} \\ \dots \\ \Gamma_M x_M \end{pmatrix} \tag{3.36}$$

The transmission from port-l to port-k:

$$s_{kl} = s'_{kl} + \sum_{i=N+1}^{M} s'_{il} \times \Gamma_i \times s'_{ki} \tag{3.37}$$

Following the same steps in 2-port derivation, the $|s_{kl}|$ transition rate can be written as:

$$\begin{aligned} \left| \frac{\partial s_{kl}}{\partial f/f} \right|_{f_0} &= \left| \sum_{i=N+1}^{M} s'_{il} \times \left(\frac{\partial \Gamma_i}{\partial f/f} \right)_{f_0} \times s'_{ki} \right| \\ &\leq \sum_{inductors} \left(\frac{|s'_{il}|^2}{2} \times \frac{Q_{L,i}}{2} \right) + \sum_{capacitors} \left(\frac{|s'_{il}|^2}{2} \times \frac{Q_{C,i}}{2} \right) \\ &\quad + \sum_{inductors} \left(\frac{|s'_{ik}|^2}{2} \times \frac{Q_{L,i}}{2} \right) + \sum_{capacitors} \left(\frac{|s'_{ik}|^2}{2} \times \frac{Q_{C,i}}{2} \right) \\ &\leq \left(Q_{L,max} \| Q_{C,max} \right) \times \left(1 - \frac{\left(\sum_{i=1}^{N} |s_{il}^2|_{f_0} \right) + \left(\sum_{i=1}^{N} |s_{ik}^2|_{f_0} \right)}{2} \right) \end{aligned} \tag{3.38}$$

where $Q_{L,max}$ ($Q_{C,max}$) is the maximum inductor (capacitor) quality factor among all lossy inductors (capacitors). Note that now the energy conservation is:

$$\left(\sum_{inductors} |s'_{ih}|^2 \right) + \left(\sum_{capacitors} |s'_{ih}|^2 \right) \leq 1 - \left(\sum_{i=1}^{N} |s_{ih}^2|_{f_0} \right),\ h = l, k \tag{3.39}$$

Equation 3.38 is very useful for multi-band N-port passive network design, when using low-Q elements, e.g. on-chip elements. And it is also helpful for the feasibility verification of on-chip passive N-port design before going into design details.

- For example, the SAW duplexer (Fig. 1.2) is a 3-port network working at dual bands, connecting to ANT, RX, and TX ports. The specified s-parameters are different in TX and RX bands. According to equation 3.38, the s-parameters transition rate (e.g. $|\frac{\partial s_{ANT-RX}}{\partial f/f}|_{f_{TX}}$) of one signal path will trade with the insertion loss (e.g. $|s_{TX-ANT}|_{f_{TX}}$) of other signal paths. Now, if low-Q elements replace the SAW filters like in [36], indicating a fully-integrated alternative to the SAW duplexer's functionality. The estimation of the realistic insertion loss and s-parameters transition rate can be easily evaluated by equation 3.38.

CHAPTER 4

Design of a Second-Order TIA for Mixer-First Receivers

4.1 Introduction

As introduced in Chapter 1, both the high dynamic range mixer and the filtering active baseband circuit post design challenges. This chapter addresses the design methodology of the active baseband circuit that follows the mixer. In the next chapter, the passive mixer design for high dynamic range will be addressed.

The active baseband circuit amplifies the baseband spectrum, comprising the small wanted signal at DC as well as nearby large blockers, passing the wanted spectrum and filtering the blockers. As large filtering as is possible should be obtained with the fewest transistors, since the circuit amplifies the blocker before filtering it fully: the partly filtered blocker can force a transistor into gain compression. [40] showed that up to a second-order lowpass filter (arbitrary poles and zeros) can be constructed with a single op amp, using controlled positive feedback to create complex poles ($Q > 0.5$) for a maximally flat response. The circuit must be scaled for maximum dynamic range at the closest blocker frequency.

4.2 Prior Arts on Active Baseband Filters

Previous reports on this type of filter [22–26] do not discuss the context of the circuit topology, nor do they give a usable design methodology. They do not

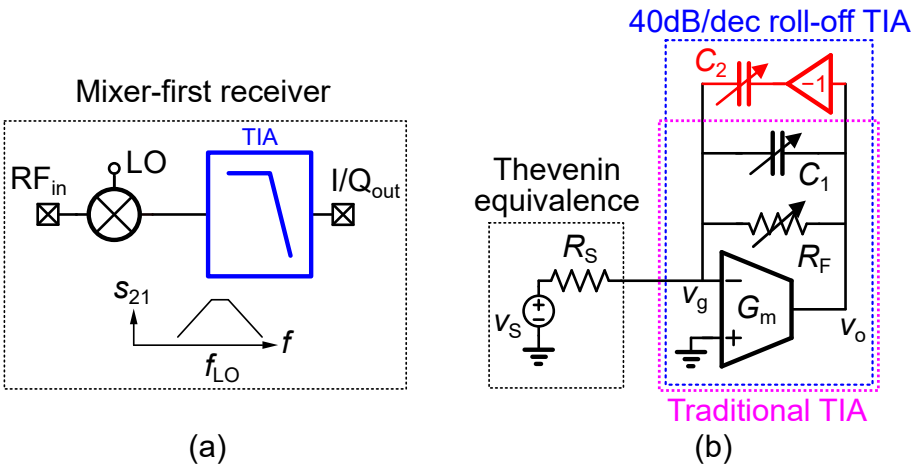


Figure 4.1: (a) A mixer-first receiver; (b) The traditional transimpedance amplifier (TIA) and a 40dB/decade filtering baseband TIA.

make clear if the circuit's dynamic range is optimum for its power consumption. Furthermore, in many cases the passive mixer enters gain compression first, in which case there is no point in striving for a greater dynamic range in the filter itself. The literature does not distinguish these effects. Specifically:

1. One or two low-frequency zeros prevent 2nd-order filtering of far-out blockers [22–26].

2. More than one amplifiers is used, wasting power [23–25].

3. A design methodology is absent, meaning complicated control of pole frequency and high power consumption [22].

4. Possible instability resulting in additional design constraints [25].

All these problems are solved in the design of this prototype chip.

4.3 Second-Order TIA Design

Fig. 4.1 shows a lowpass transimpedance (R_F, C_1) amplifier (TIA) that matches, in its pass-band, the impedance of the driving source (here this is the Thevenin

equivalent of the antenna with its matching circuit, and passive mixer). C_2 appears in feedback with a sign inversion indicating positive feedback. In a fully differential op amp, C_2 is connected to a non-inverting output. The y-parameters (port-1: v_g node to ground; port-2: v_o node to ground) of the TIA are:

$$\begin{aligned}\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} &= \begin{pmatrix} s(C_1+C_2)+\frac{1}{R_F} & -\left(s(C_1-C_2)+\frac{1}{R_F}\right) \\ G_m-\left(s(C_1-C_2)+\frac{1}{R_F}\right) & s(C_1+C_2)+\frac{1}{R_F}+\frac{1}{r_0} \end{pmatrix} \\ &\triangleq \begin{pmatrix} sC_F+\frac{1}{R_F} & -\left(s\Delta C+\frac{1}{R_F}\right) \\ G_m-\left(s\Delta C+\frac{1}{R_F}\right) & sC_F+\frac{1}{R_F}+\frac{1}{r_0} \end{pmatrix}\end{aligned} \tag{4.1}$$

where $A_0 = G_m r_0$ is the amplifier voltage gain, transconductance, and output resistance. Define $G'_m(s) \triangleq G_m - (s\Delta C + \frac{1}{R_F})$ combining the feedforward of both the amplifier and feedback network. In the y-parameter two-port equivalent, the capacitors appear in the parameters either as a sum $C_F = C_1 + C_2$ or a difference $\Delta C = C_1 - C_2$. This simplifies design.

Two transfer functions must be considered. The first is from the source v_S to the amplifier/filter output voltage, v_o. The second is to the amplifier input v_g. If a blocker voltage exceeds a certain limit at *either* node, it can drive the op amp into gain compression. In the y-parameter equivalent circuit (Fig. 4.2), the transfer functions of the shunt-shunt feedback circuit can be derived:

$$\begin{aligned}\frac{v_o}{v_S}(s) &= \frac{1}{R_S} \times \frac{v_o}{i_S} = \frac{1}{R_S} \times \frac{R_T(s)}{1+T(s)} \\ \frac{v_g}{v_S}(s) &= \frac{1}{R_S} \times \frac{v_g}{i_S} = \frac{1}{R_S} \times \frac{R_g(s)}{1+T(s)}\end{aligned} \tag{4.2}$$

where

$$\begin{aligned}R_T(s) &= \left(R_S \| R_F \| \frac{1}{sC_F}\right) \times G'_m(s) \times \left(r_0 \| R_F \| \frac{1}{sC_F}\right) \\ R_g(s) &= \left(R_S \| R_F \| \frac{1}{sC_F}\right) \\ T(s) &= R_T(s) \times y_{12}(s)\end{aligned} \tag{4.3}$$

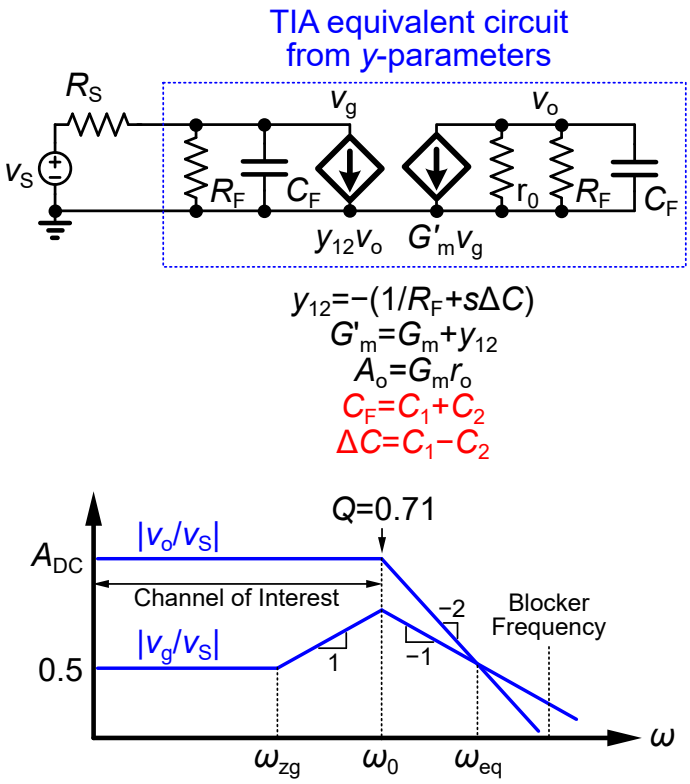


Figure 4.2: The equivalent circuit and the frequency response of the 40dB/decade filtering TIA.

The zeros and poles of the transfer functions are then calculated:

$$\begin{aligned}\frac{v_o}{v_S}(s) &= \frac{A_{DC}(1+\frac{s}{\omega_{zo}})}{1+\frac{s}{\omega_0 Q}+\frac{s^2}{\omega_0^2}} \\ \frac{v_g}{v_S}(s) &\approx \frac{0.5\times(1+\frac{s}{\omega_{zg}})}{1+\frac{s}{\omega_0 Q}+\frac{s^2}{\omega_0^2}}\end{aligned} \tag{4.4}$$

where the ω_0, Q, ω_{zo}, ω_{zg} are:

$$\begin{aligned}\omega_0^2 &\approx \frac{2G_m}{R_F C_F^2} \\ \underbrace{\frac{1}{Q}}_{\sim 1} &\approx \underbrace{\frac{\sqrt{2}(R_S+r_0)}{2\sqrt{R_S r_0}}}_{\sim 1} + \sqrt{\frac{R_S}{2r_0}}\underbrace{A_0}_{\gg 1}\underbrace{\frac{\Delta C}{C_F}}_{\ll 1} \\ \omega_{zo} &\approx -\frac{G_m}{\Delta C} \\ \omega_{zg} &\approx \frac{G_m}{A_0 C_F}\end{aligned} \tag{4.5}$$

ω_{eq} is then calculated when $|v_o/v_S(j\omega_{eq})| = |v_g/v_S(j\omega_{eq})|$:

$$\omega_{eq} \approx \frac{G_m}{C_F} \tag{4.6}$$

The design expressions in equation 4.5 and 4.6 assume:

1. Large amplifier gain $A_0 \gg 1$.

2. Good impedance match $R_F \approx R_S \times A_0$. Here the mixer conversion gain (i.e. $\frac{2\sqrt{2}}{\pi}$ for 25% duty-cycled clock [41, 42]) is assumed to be 1.

3. A good voltage amplifier $R_F \gg r_0$, or equivalently $G_m R_S \gg 1$.

4. $Q \approx 1/\sqrt{2}$ which leads to $C_F \gg \Delta C$ as shown in equation 4.5.

Fig. 4.2 illustrates the case when the blocker amplitude is larger at the amplifier input than at its output ($\omega > \omega_{eq}$). The literature seems to overlook this fundamental property: if the transfer function $v_o/v_S(s)$ consists *only* of a complex conjugate pole pair with $Q = 1/\sqrt{2}$ for a flat lowpass magnitude response, the transfer function $v_g/v_S(s)$ must consist of the same two poles, *and also* a low frequency zero. This means that at impedance match $|v_g/v_S| \approx 1/2$ near DC; but rises at 20 dB/decade towards the lowpass cutoff frequency because of a low frequency zero, then falls at 20 dB/decade at frequencies beyond. The transfer function to the output $|v_o/v_S|$ is flat until the filter cutoff, then falls at 40 dB/decade. The amplifier is a second-order filter to blocker channels at its output, but *only* first-order at its input!

Dynamic range is optimum when the major blocker at a specified frequency offset drives both input and output ports into compression at the same time. The two magnitude curves should intersect (ω_{eq}) at or lower than this blocker frequency.

Equation 4.5 indicates Q is sensitive to ΔC and is almost independent of C_F

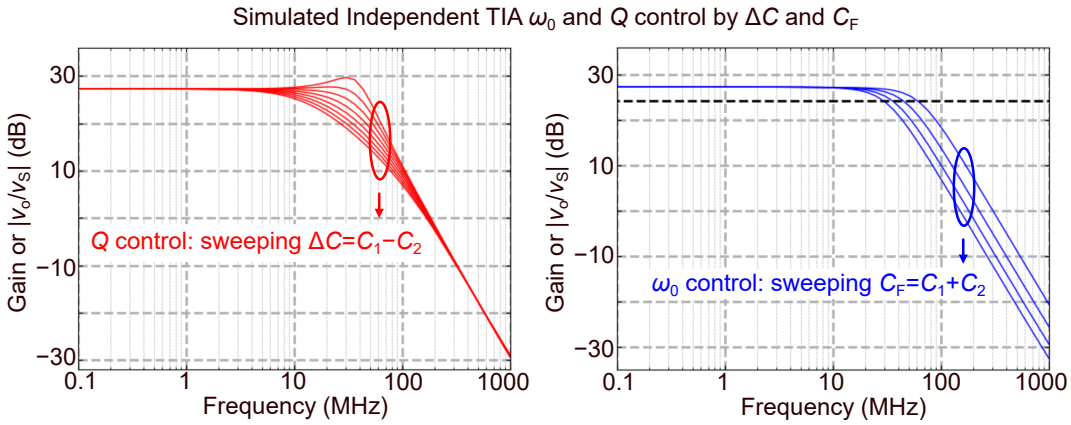


Figure 4.3: Simulated 2nd-order filtering TIA transfer functions demonstrating independent ω_0 and Q control.

because $C_F \gg \Delta C$, and ω_0 is sensitive to C_F but not ΔC.

$$\begin{aligned} \frac{\partial \omega_0}{\partial C_F} &\gg \frac{\partial \omega_0}{\partial \Delta C} \\ \frac{\partial Q}{\partial \Delta C} &\gg \frac{\partial Q}{\partial C_F} \end{aligned} \tag{4.7}$$

The simulated plots in Fig. 4.3 show the independent control of ω_0 and Q by C_F and ΔC. $Q = 1/\sqrt{2}$ gives a maximal flat lowpass characteristic, and ω_0 is the 3-dB filter bandwidth at baseband (BBBW).

To summarize, the element values are selected as follows:

1. R_F for impedance match.
2. C_F to tune ω_0 or BBBW.
3. ΔC to adjust Q.
4. G_m for power or noise.

The poles of the complete feedback circuit are well-defined in the left-half s-plane, guaranteeing stability.

G_m cannot be scaled down indefinitely to save power, because:

1. Its voltage noise rises.

2. Its r_0 may rise close to R_F, so ω_0 and Q can no longer be tuned independently.

3. The frequency of the zero ω_{zg} of transmission to the amplifier input is lowered, raising the amplitude there at the blocker frequency.

Therefore, $G_m R_S \gg 1$ is required, and ω_{eq} must be equal to or higher than the blocker frequency. For example, if $R_S = 50\Omega$, the total $G_m \gg 20$mS is required, so the expressions and properties of the TIA (equation 4.5, 4.6 and 4.7) are valid.

4.4 Passive Mixer and LO Design

The passive mixer can limit the receiver linearity and dynamic range, when the far-out blockers have been largely suppressed by the baseband filter. Both the mixer FET and clock waveform affect the mixer linearity. Their impacts on mixer linearity and dynamic range are analyzed in Chapter 5. Low FET on-resistance and sharp clock are desired.

The LO generation circuits for 25% duty-cycled clocks are shown in Fig. 4.5. The LO design is a challenge to support up to $f_{LO} = 7$GHz in this technology, since the cutoff frequency of the NFET is only $f_T \sim 60$GHz. In fact, CMOS inverter chain with fan-out of 1 fails at 14GHz (i.e., $2f_{LO}$ input) under slow-slow process corner, which means the buffers in Fig. 4.5 will fail at 14GHz under slow-slow process corner. This problem does not exist in advanced CMOS technology, where the FETs are much faster (e.g., NFET $f_T > 200$GHz in 65nm bulk CMOS technology)

To guarantee the functionality of the LO circuits for all $f_{LO} \leq 7$GHz, NMOS logic is used, which replaces all PFETs with resistors, as shown in Fig. 4.4. The NMOS logic increases the speed of the circuits, however, at the cost of the

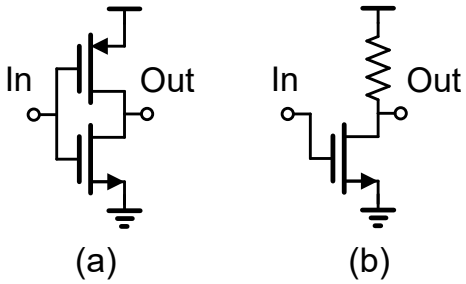


Figure 4.4: (a) A CMOS inverter; (b) A NMOS inverter is used for higher operating frequency.

static power consumption. Although the NMOS logic helps, clock waveforms still deteriorate at high frequencies (e.g., simulated clock swing $<$ 1V when $V_{DD} = 1.2V$ at f_{LO} =7GHz under typical corner).

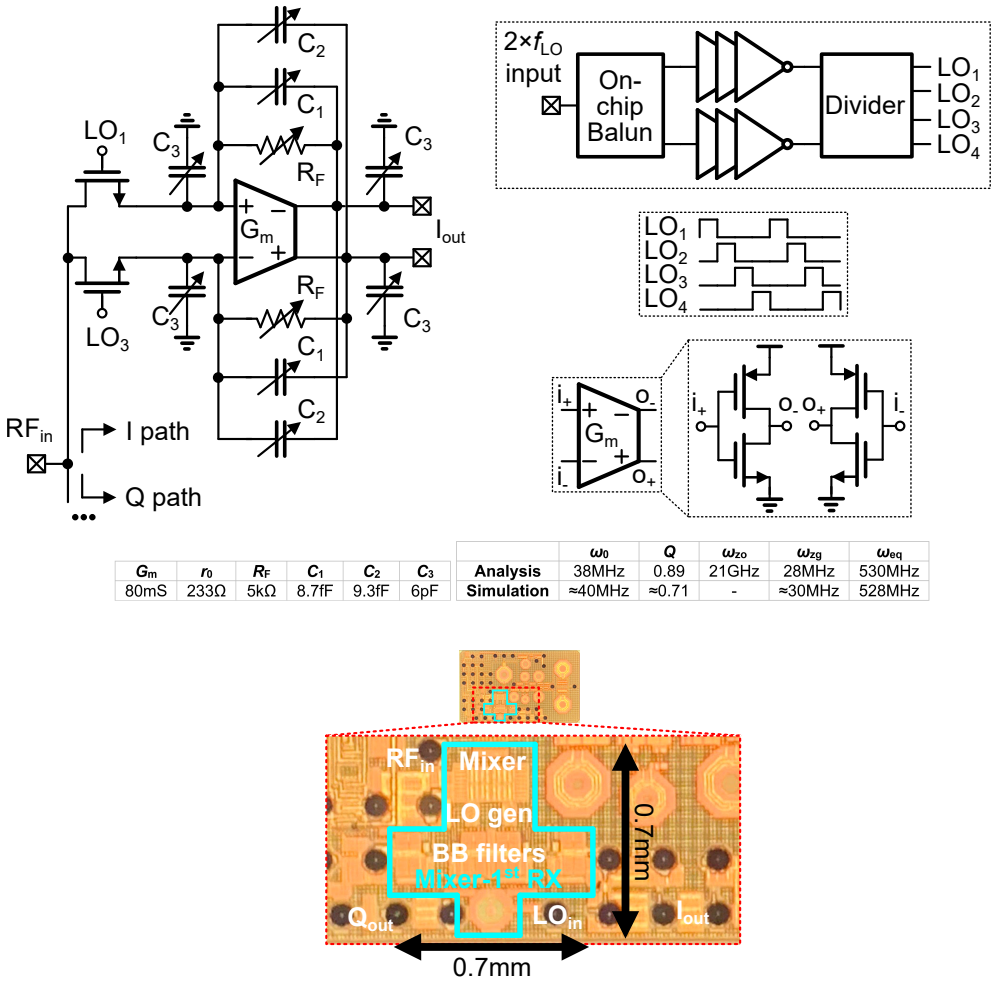
Both low FET on-resistance and sharp clock means higher power consumption. Therefore, the FET sizes in the mixer and LO generation circuits are chosen to maximize the IIP3 and B1dB (blocker 1dB compression point), while keeping below constant power budget $\sim$ 200mW.

4.5 Measurements

4.5.1 Receiver Measurement

A prototype mixer-first receiver is fabricated on 65nm RF-SOI CMOS technology (Fig. 4.5), as part of a monolithic duplexer. The process is not optimized for 1.2V FETs, so their cutoff frequency f_T is $\approx$60 GHz, much lower than in 65nm bulk CMOS. The waveforms in the LO generator, which produces a four phase clock at RF, degenerate from full-swing square pulses into lower amplitude Gaussian-like pulses. Both NF and linearity of the mixer suffer as a result.

The receiver operates from 1 to 7GHz. Capacitors C_3 at filter input and output contribute to C_F without changing ΔC. The filter passband is set to 40 MHz. Fig. 4.6 shows the measured frequency response for RF inputs, downconverted by



G_m	r_0	R_F	C_1	C_2	C_3
80mS	233Ω	5kΩ	8.7fF	9.3fF	6pF

	ω_0	Q	ω_{zo}	ω_{zg}	ω_{eq}
Analysis	38MHz	0.89	21GHz	28MHz	530MHz
Simulation	≈40MHz	≈0.71	-	≈30MHz	528MHz

Figure 4.5: Block diagrams and schematics for the fabricated mixer-first receiver and LO generation circuit; equation 4.5 and 4.6 compared with simulation with elements values; die micrograph.

3 GHz. The response is very close to an ideal second-order filter translated from zero IF to 3 GHz. There is no sign of a low frequency zero in $|v_o/v_S|$ up to 400 MHz offset, as expected. Fig. 4.7 shows the gain and noise figure (NF) measured for the mixer-filter cascade at different LO frequencies. The NF rise at high LO frequency is due to the mixers, as the clock waveform at their gate deteriorates in the LO generation chain.

The measured blocker NF rises by 3.5dB from reciprocal mixing with a 0dBm blocker. The measurement setup is as follows:

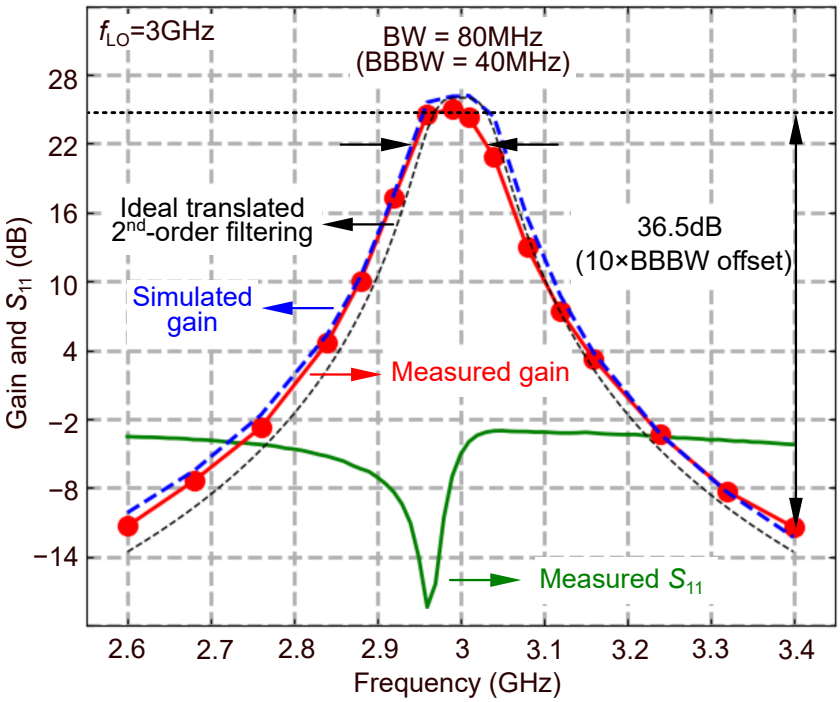


Figure 4.6: Measured gain of the fabricated mixer-first receiver, and comparison to simulation and ideal 40dB/decade roll-off BB filter; measured S_{11}.

- A signal generator serving as the blocker and a noise source are combined at the input of the device under test (DUT). The signal generator phase noise is filtered. Isolators are used to prevent the blocker from entering the noise source due to reflections.

- The output of the DUT connects to the spectrum analyzer. The NF in the RX band is then measured in the presence of the blocker using Y-factor method (built-in function of the spectrum analyzer).

The blocker 1dB compression point (B1dB) and two-tone OOB IIP3 are shown in Fig. 4.7 measured at f_{LO} = 3GHz. The blockers and/or signals are applied similarly to the blocker NF measurement. The blocker frequencies for the measurement are shown in Fig. 4.7. Both B1dB and IIP3 measurement 40 dB/decade improvement versus frequency offset f_{os} while compression is due to baseband TIAs, consistent with second-order filtering. At high frequency offset where the

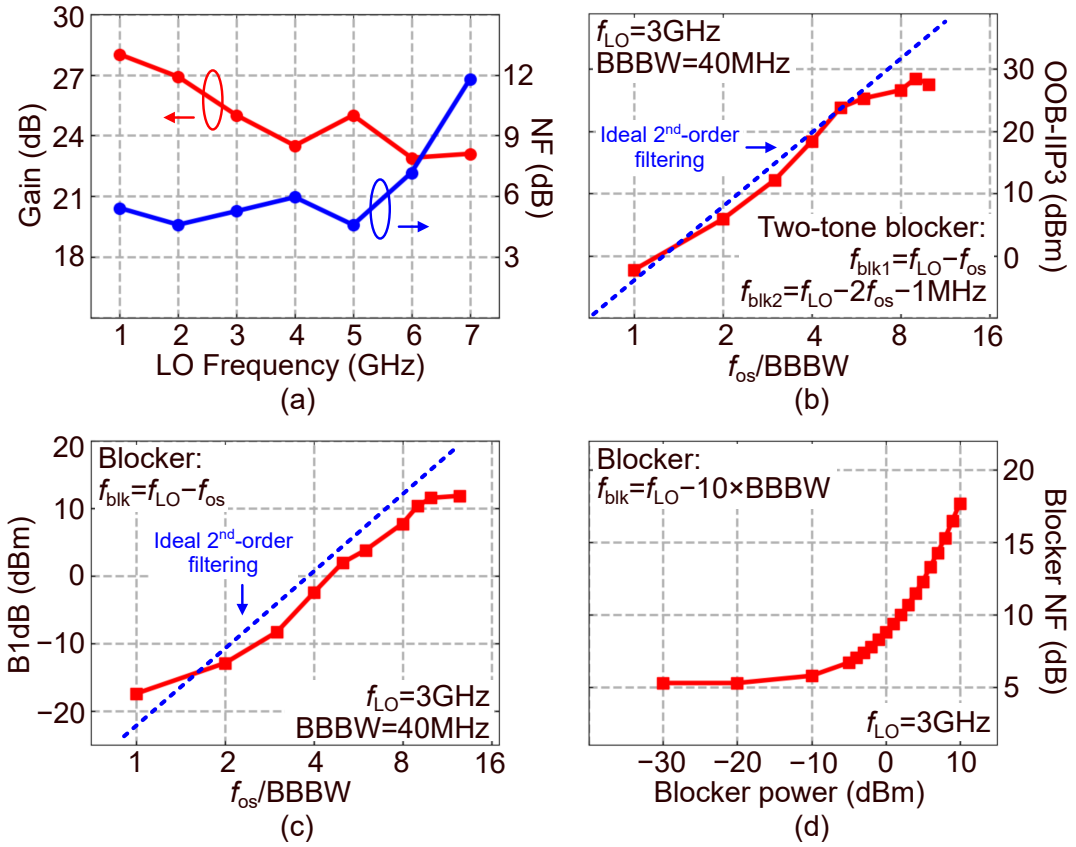


Figure 4.7: (a) Measured gain/NF versus LO frequency; (b) Measured OOB-IIP3 and (c) B1dB versus offset frequency; (d) measured block NF versus blocker power.

filtering has largely suppressed the blocker, the mixer dictates B1dB (11.8dBm) and IIP3 (28.5dBm). Mixer nonlinearity originates in modulation of the FET on-resistance by the blocker, and deteriorated clock waveform.

4.5.2 Cascaded Measurement with the EBD

The receiver is then connected to the RX port of the dual-band EBD in Chapter 2. They operate from 6 to 7GHz. The DUT input is now the EBD ANT port, while other setups are identical to the receiver standalone measurement.

The frequency response for f_{LO} = 6GHz is shown in Fig. 4.8. The response is very close to the receiver standalone measurement, and the translated ideal

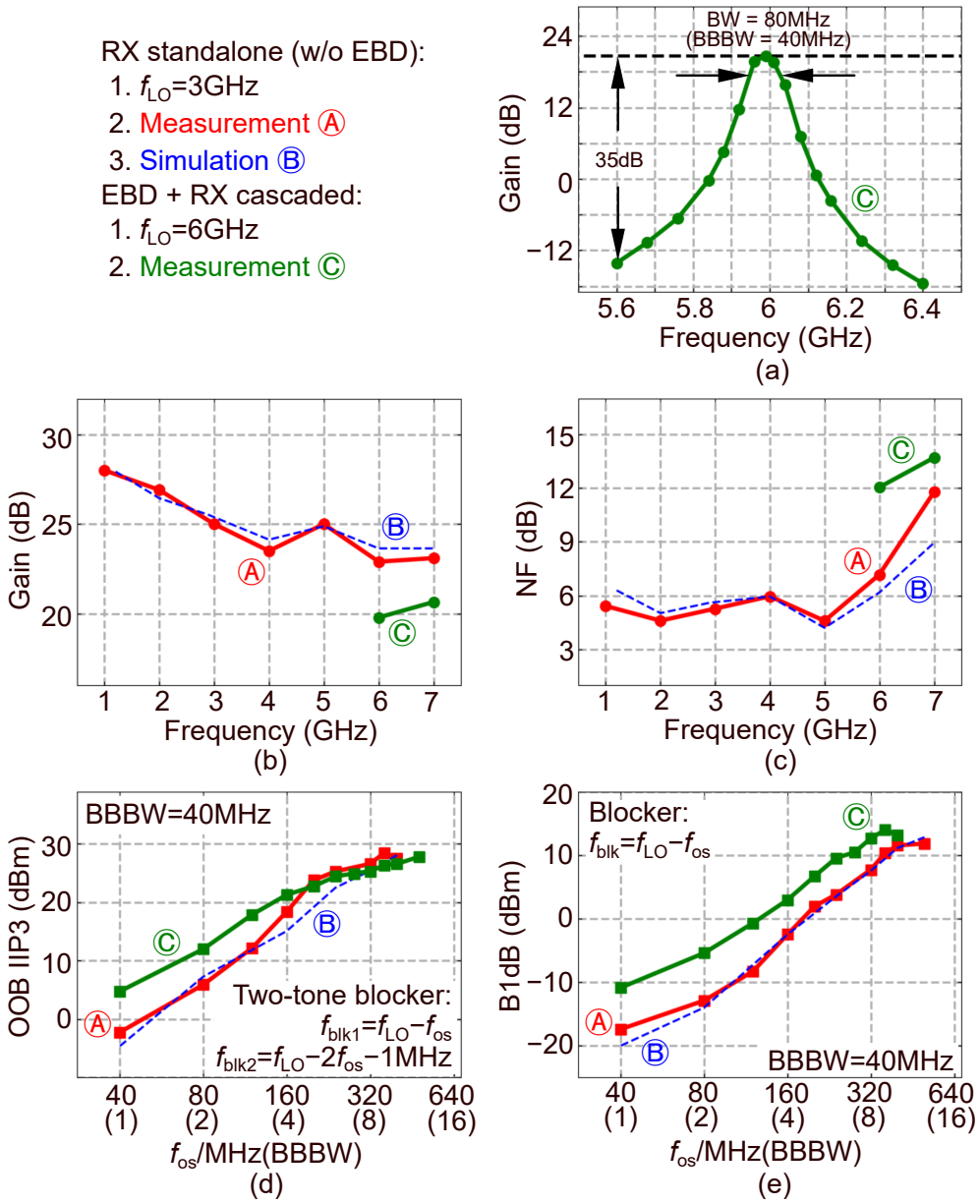


Figure 4.8: EBD and RX cascaded measurement and comparison to RX standalone measurement: (a) gain response; (b) gain and (c) NF versus LO frequency; (d) OOB-IIP3 and (e) B1dB versus blocker offset frequency.

second-order filter. Fig. 4.8 shows the gain and NF measurement at different LO frequencies. The changes in gain and NF are consistent with the EBD insertion loss ($\sim$ 3.2-4dB). Due to the deteriorated clock waveform, the NF is high. The measured blocker NF rises by 2.4dB from reciprocal mixing with a 0dBm blocker

at the EBD ANT port to mimic the TX leakage. When the blocker is applied at TX port, there is no impact on NF since it is attenuated by the EBD of $\sim$ 40dB.

The B1dB and IIP3 measurements are shown in Fig. 4.8 at $f_{LO} = 6$GHz. The blockers are applied at the EBD ANT port. Similar to the receiver standalone measurement, both B1dB and IIP3 versus blocker offset f_{os} are consistent with the baseband filter for close-in blockers, and the mixer dictates the B1dB (13.9dBm) and IIP3 (27.8dBm) for far-out blockers. Comparing to receiver standalone measurements, the B1dB and IIP3 are supposed to be higher by $\sim$ 5dB for close-in blockers due to lower gain, and $\sim$ 4dB for far-out blockers due to EBD insertion loss. The mixer linearity degrades at higher f_{LO} due to the deteriorated clock waveform.

4.6 Discussions and Conclusions

Table 4.1 compares this receiver with other similar works. [22] uses a differential topology, which lowers FET on-resistance modulation to improve IIP3 by $\sim$ 3dB, but at the cost of an off-chip balun. [24] uses a higher supply voltage that also lowers FET on-resistance modulation for higher IIP3. FETs of higher f_T as in [22, 25], will switch the mixer FETs crisply, boosting IIP3 until it is eventually limited by FET on-resistance modulation. In spite of the limits posed by the mixer, we have satisfactorily verified the methodology of the baseband TIA. The circuit consumes low power and offers second-order filtering with an optimized blocker dynamic range.

Table 4.2 summarizes the cascaded performance connecting the EBD and the receiver. The receiver works in 6-7GHz. Due to the poor clock waveform discussed earlier, the receiver does not perform as good in high frequencies (6-7GHz) as in low frequencies. This receiver does not experience gain compression and NF degradation at the TX leakage power level (e.g -13dBm).

Table 4.1: Mixer-first receiver comparison table.

	[22]	**[23]**	**[25]**	**[24]**	**This work**
BB Architecture	(differential) 2^{nd}-order 1-opamp TIA	shunting notch	2^{nd}-order Z_{BB}	2^{nd}-order NC TIA	2^{nd}-order 1-opamp TIA
Technology	45nm SOI	130nm SiGe BiCMOS	28nm	180nm	65nm SOI
Frequency (GHz)	0.2-8	2-11	0.2-2	0.2-1.2	1-7
Area (mm^2)	0.8	8.4	0.48	0.54	0.3
Baseband BW (MHz)	10	40-130	9	18	40
Gain (dB)	21	10-24	13	31.4	25
NF (dB)	2.3-5.4 (0.5-6GHz)	11±1 (5GHz)	4.3-7.6	3.4-4	4.6-7.2 (1-6GHz)
OOB-IIP3 (f_{os}/BBBW)	39dBm (8)	20dBm (1.75)	33.3dBm (8.8)	39.8dBm (16*)	28.5dBm (9)
B1dB (f_{os}/BBBW)	12dBm (8)	1.8dBm (1.75)	12dBm (6.6)	12dBm* (16*)	11.8dBm (12)
0dBm Blocker NF desense. (dB)	2.2	-	2	-	3.5
Supply (V)	1.2	4.5/2.5	1.2	1.8	1.2
Clock Power (mW)	30/GHz	1466-1494	3.6-36	45-135	211
RX Power (mW)	50	656	143	19.8	22

*: Estimated from figures.

Table 4.2: Cascaded (EBD + RX) receive path performance summary.

	This work: Cascaded receive path
Area (mm^2)	2.3 (EBD) + 0.3 (RX)
RF frequency (GHz)	6-7
Gain (dB)	20
Baseband BW (MHz)	40
OOB-IIP3 (f_{os}/BBBW)	27.8dBm (12)
B1dB (f_{os}/BBBW)	13.9dBm (9)
NF (dB)	12-13.7
0dBm Blk. NF desense. (dB)	2.4
Power (mW)	(BB: 22) + (LO: 211)
Supply (V)	BB & LO: 1.2

CHAPTER 5

Passive Mixer Nonlinearity Analysis

5.1 Introduction

CMOS passive mixer has been widely used [22–26, 43] in nowadays RF wireless receivers to perform the frequency translational function. It provides the unique bidirectional feature: it downconverts the RF signal to baseband (BB), at the same time it upconverts the BB signal to RF. Due to this feature, the BB impedance/filter can be translated to RF, so a very sharp RF filter can be realized [3, 22–26, 43–46] to filter unwanted signals.

Out-of-band (OOB) blockers, including the TX leakage, appear at the input of the RX. They can saturate the RX due to high power (B1dB), or desensitize the RX due to inter-modulation (IIP3). Therefore, filtering OOB blockers at the very input of the RX is crucial, especially for SAW-less receivers.

As shown in Fig. 1.4, a typical mixer-first receiver consists of a passive mixer and an active BB filter. The passive mixer translates the BB filter to the RF, and a very sharp RF filter then appears at the very input of the receiver. The OOB blockers, which appear at the input of the RX, are filtered before amplification. The method of BB filtering is discussed in Chapter 4. When strong filtering is available, the linearity of the receiver can be limited by the first stage of the receiver: the passive mixer. Therefore, it is necessary to study the nonlinearity in the passive mixer.

[47, 48] studied the passive mixer IIP3. However, the model is not com-

plete and the dependency on FET size either does not match the simulation very well [47], or is not given [48]. They also neglected the impact of clock rise and fall time, which is very important at high frequencies. In this chapter, comprehensive mechanisms of compression and inter-modulation in the passive mixer are discussed. B1dB and IIP3 are analyzed and predicted by expressions. Note that it is the first time that an analysis on the passive mixer B1dB is given.

5.2 OOB IIP3 Analysis

In passive mixers, the FET switch on-resistance R_{on} is considered the main contribution of the third-order nonlinearity in previous works [47, 48]. When the FET switch is ON, R_{on} is modulated by the drain and source voltage V_{ds}, or the channel voltage V_{ch}, and introduces third-order nonlinearity. However, [47] used an oversimplified model which assumes one terminal of FET is grounded, which results in an incorrect IIP3 dependency on FET size. [48] does not give the IIP3 dependency on FET size. In the following sections, a more comprehensive model is developed and the derived IIP3 shows agreement with the simulations.

The clock can also introduce third-order nonlinearity directly, which does not interact with R_{on} modulation, or indirectly, which interacts with R_{on} modulation. These effects are ignored in previous works. The finite clock slope can increase the R_{on} nonlinearity during clock transition, and reduce the overall R_{on}-introduced IIP3. This effect is referred as soft switching (SS) in the following sections. On the other hand, the moment that FET switches turn ON and OFF can change in the presence of a large blocker and finite clock slope. The FET switches ON and OFF timing are modulated by the blocker, and is referred as timing modulation (TM) in the following sections. These clock-related effects will be analyzed, and agreement with the simulations will be shown.

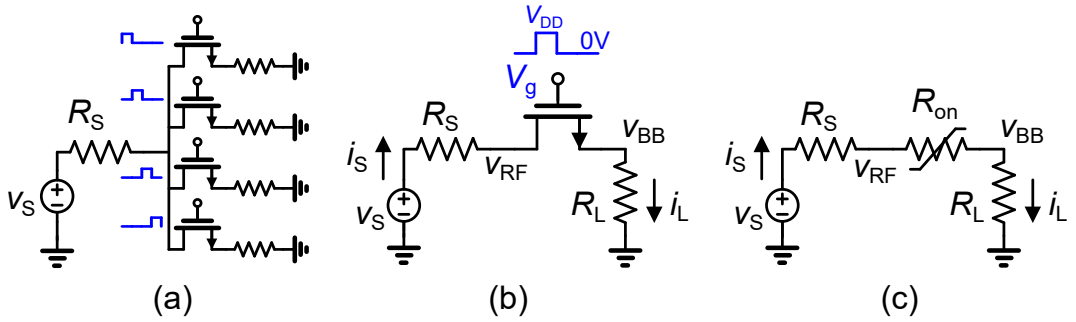


Figure 5.1: (a) A four-path mixer-first receiver with resistor load; (b) A circuit with only one signal path of the receiver; (c) The equivalent circuit.

5.2.1 R_{on} Nonlinearity

5.2.1.1 Non-Zero Baseband Load Resistance

Consider the circuit shown in Fig. 5.1. A FET switch (R_{on} or G_{on}) with time variant clock driving its gate, and a load resistor $R_L \neq 0\Omega$ are driven by a RF source v_S with source resistance R_S. The on-resistance and on-conductance of the FET switch is R_{on} and G_{on}, respectively. The current through the load resistor i_L can be calculated when considering a small variation on FET on-resistance ΔR_{on} or on-conductance ΔG_{on} due to non-zero v_S:

$$\begin{aligned} i_L &= \frac{v_S}{R_S + R_L + R_{on}} \\ &= v_S \times \frac{G_{on} + \Delta G_{on}}{(R_S + R_L) \times (G_{on} + \Delta G_{on}) + 1} \\ &\approx v_S \times \frac{G_{on}}{(R_S + R_L)G_{on} + 1} \times \left(1 + \rho\Big(\frac{\Delta G_{on}}{G_{on}}\Big) - \rho\alpha\Big(\frac{\Delta G_{on}}{G_{on}}\Big)^2\right) \\ &= v_S \times \frac{\alpha}{R_S + R_L} \times \left(1 + \rho\Big(\frac{\Delta G_{on}}{G_{on}}\Big) - \rho\alpha\Big(\frac{\Delta G_{on}}{G_{on}}\Big)^2\right) \end{aligned} \tag{5.1}$$

where $\rho \triangleq \frac{R_{on}}{R_{on}+R_S+R_L}$, $\alpha \triangleq \frac{R_S+R_L}{R_{on}+R_S+R_L}$. The i_L is expanded up to the third term in the Taylor series to study the IIP3, where ΔG_{on} is linearly related to v_S (will show later). Note that all variables, excluding R_S and R_L, are functions of time.

Since the second term does not contribute to IIP3, it can be ignored:

$$i_L \approx v_S \times \frac{\alpha}{R_S + R_L} \times \left(1 - \rho\alpha\Big(\frac{\Delta G_{on}}{G_{on}}\Big)^2\right) \tag{5.2}$$

Now if the clock driving the FET is infinitely sharp:

$$\begin{gathered} i_L \approx v_S \times \frac{\alpha}{R_S + R_L} \times \left(1 - \rho\alpha\Big(\frac{\Delta G_{on}}{G_{on}}\Big)^2\right), \text{ clock high} \\ i_L = 0\text{A}, \text{ clock low} \end{gathered} \tag{5.3}$$

When the FET is off, $G_{on} = 0$S and $\Delta G_{on}/G_{on} = 0$. When the FET is on, G_{on} and ΔG_{on} using EKV model [49] (ignoring the body effect and channel length modulation) is:

$$G_{on} = \beta(V_g - V_{th} - \frac{V_D + V_S}{2}) = \beta(V_g - V_{th} - V_{ch}) \triangleq \beta V_{ov} \tag{5.4}$$

$$\Delta G_{on} = \beta\frac{v_{RF} + v_{BB}}{2} \triangleq \beta v_{ch} \tag{5.5}$$

where $\beta \triangleq \mu C_{ox}W/L$ for the FET, V_{th} is the threshold voltage, V_g is the gate voltage, V_{ov} is the overdrive voltage, V_D and V_S are source and drain bias (DC) voltage, V_{ch} is the small-signal channel voltage, v_{RF} and v_{BB} are small-signal voltage at FET source and drain, and v_{ch} is the channel voltage change due to v_S. If $R_L > R_{on}$, v_{ch} and v_S are approximately linearly related:

$$\begin{aligned} v_{ch} &\approx \frac{R_L + R_{on}/2}{R_L + R_S + R_{on}} \times v_S \\ &\triangleq k \times v_S \end{aligned} \tag{5.6}$$

Using the results in equation 5.4, 5.5 and 5.6, equation 5.2 can be rewritten:

$$i_L \approx v_S \times \frac{\alpha}{R_S + R_L} \times \left(1 - \rho\alpha\Big(\frac{k}{V_{ov}}\Big)^2 \times v_S^2\right) \tag{5.7}$$

Therefore, the IIP3 of nonlinear current $i_L = a_1 v_S + a_3 v_S^3$ is:

$$\begin{aligned} V_{\text{IIP3, } R_L \neq 0\Omega} &\triangleq \sqrt{\frac{4}{3}\left|\frac{a_1}{a_3}\right|} \\ &= \sqrt{\frac{4}{3}\frac{V_{ov}^2}{k^2\rho\alpha}} \\ &\approx \sqrt{\frac{4}{3}\frac{V_{ov}^2}{k^2\rho}} \end{aligned} \tag{5.8}$$

where $\alpha \approx 1$ because usually $R_{on} \ll (R_S + R_L)$. Note that equation 5.8 is based on the assumption that $R_L > R_{on}$, so the nonlinear terms of $v_{ch}(v_S)$ can be ignored.

5.2.1.2 Grounded Baseband Node

Equation 5.8 gives the IIP3 with the assumption $R_L > R_{on}$. Therefore, it is necessary to look at the IIP3 when $R_L < R_{on}$. An extreme case is studied in this section, where $R_L = 0\Omega$.

Since v_{ch} now is non-linear function of v_S when $R_L = 0\Omega$, a new analysis method is used. When the FET switch is on:

$$\begin{aligned} i_L(v_{RF}) &= \frac{1}{2}\beta(V_{ov}^2) - \frac{1}{2}\beta(V_{ov} - v_{RF})^2 \\ &= \beta \times v_{RF} \times (V_{ov} - \frac{v_{RF}}{2}) \end{aligned} \tag{5.9}$$

$$\begin{aligned} v_S(v_{RF}) &= v_{RF} + i_L \times R_S \\ &= v_{RF} + \beta v_{RF}(V_{ov} - \frac{v_{RF}}{2}) \times R_S \end{aligned} \tag{5.10}$$

where v_{RF} is the RF node voltage. Applying implicit differentiation to equation

5.9 and 5.10:

$$
\begin{aligned}
i_L(v_S) &= \frac{d(i_L)}{d(v_{RF})} v_{RF}(v_S) + \frac{d^2(i_L)}{d(v_{RF})^2} \frac{v_{RF}^2(v_S)}{2} \\
&\approx \frac{d(i_L)}{d(v_{RF})} \times \underbrace{\left(\frac{d(v_{RF})}{d(v_S)} v_S + \frac{d^3(v_{RF})}{d(v_S)^3} \frac{v_S^3}{6} \right)}_{\substack{\text{Taylor expansion of } v_{RF}(v_S) \\ v_S^2 \text{ term not interested}}} \\
&\quad + \frac{d^2(i_L)}{d(v_{RF})^2} \times \underbrace{\frac{d(v_{RF})}{d(v_S)} \frac{d^2(v_{RF})}{d(v_S)^2} \frac{v_S^3}{2}}_{\substack{v_{RF}(v_S) \approx \frac{d(v_{RF})}{d(v_S)} v_S + \frac{d^2(v_{RF})}{d(v_S)^2} \frac{v_S^2}{2} \\ v_S^3 \text{ terms kept in } \frac{v_{RF}^2(v_S)}{2}}} \\
&= \underbrace{\beta V_{ov} \times \left(\rho v_S + \frac{(\beta R_S)^2 \rho^5}{2} v_S^3 \right) + (-\beta) \times \frac{\beta \rho^4 R_S}{2} v_S^3}_{\text{where } \frac{d(v_{RF})}{d(v_S)} = \rho,\ \frac{d^2(v_{RF})}{d(v_S)^2} = \beta \rho^3 R_S,\ \frac{d^3(v_{RF})}{d(v_S)^3} = 3(\beta R_S)^2 \rho^5} \\
&= (\beta V_{ov} \rho) v_S + \underbrace{\left(\frac{\beta^3 \rho^5 V_{ov} R_S^2 - \beta^2 \rho^4 R_S}{2} \right)}_{\rho \beta V_{ov} R_S = 1 - \rho} v_S^3 \\
&= (\beta V_{ov} \rho) v_S - \frac{\beta^2 \rho^5 R_S}{2} v_S^3
\end{aligned}
\tag{5.11}
$$

where only the first and third order terms are kept since other terms do not contribute to IIP3. The IIP3 of nonlinear current $i_L = a_1 v_S + a_3 v_S^3$ is:

$$
\begin{aligned}
V_{\text{IIP3},\ R_L = 0\Omega} &\triangleq \sqrt{\frac{4}{3} \left| \frac{a_1}{a_3} \right|} \\
&= \sqrt{\frac{4}{3} \frac{2 V_{ov}}{\beta \rho^4 R_S}} \\
&\approx \sqrt{\frac{8}{3} \frac{V_{ov}^2}{\rho^3}}
\end{aligned}
\tag{5.12}
$$

where $\rho \beta V_{ov} R_S \approx 1$ because $R_{on} \ll R_S$.

Comparing equation 5.8 and 5.12, the difference is the dependency on FET size:

- $R_L = 0\Omega$. $V_{\text{IIP3}} \sim \rho^{-1.5}$.

- $R_L \neq 0\Omega$. $V_{\text{IIP3}} \sim \rho^{-1.5} - \rho^{-0.5}$, depending on the resistance R_{on} and R_L. Note that $\lim_{R_L \to 0} k = \rho/2$, and equation 5.8 becomes:

$$V_{\text{IIP3, } R_L \neq 0\Omega} = \sqrt{\frac{16}{3}\frac{V_{ov}^2}{\rho^3}} \tag{5.13}$$

This is 3dB higher comparing to equation 5.12. In order for them to be consistent, k can be redefined:

$$k \triangleq \frac{R_L + (R_{on}/\sqrt{2})}{R_L + R_S + R_{on}} \tag{5.14}$$

where the difference is that the coefficient of R_{on} in the numerator is now $1/\sqrt{2}$ instead of $1/2$. Now, equation 5.8 can be used for all ranges of $R_L \geq 0\Omega$.

5.2.1.3 Parallel RC as the Baseband Load

In an actual RX, the baseband input impedance is usually modeled by a parallel RC instead of a resistor R_L, while the capacitor presents low impedance and filters the OOB blocker to avoid voltage gain at baseband. Therefore, the impedance $|Z_L|$ is small and capacitive when OOB blockers are applied.

In [47], the load impedance $Z_L = 0\Omega$ is used because for OOB blocker the baseband node $|Z_L| \ll R_S$ due to strong filtering. However, our analysis shows that $|Z_L|$ cannot be ignored unless $|Z_L| \ll R_{on}$, making $|Z_L|$ non-negligible in most cases. In this section, the IIP3 of the parallel RC load at baseband will be analyzed, and equation 5.8 is modified.

For the parallel RC baseband load shown in Fig. 5.2, the expression on v_{ch} shown in equation 5.6 is no longer correct. v_{ch} is now:

$$v_{ch} = i_L \times R_{on}/\sqrt{2} + v_{BB} \tag{5.15}$$

Note that $v_{ch} \propto v_S$ due to $v_{BB} \propto v_S$ and $i_L \propto v_S$. The IM3 terms are generated by the product $v_{ch}^2 v_S \times \text{clk}(t)$ shown in equation 5.1 and 5.5, where the clk(t) is

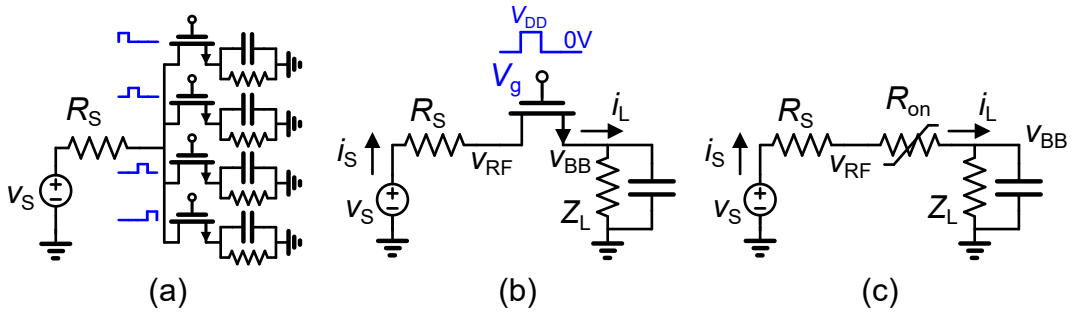


Figure 5.2: (a) A four-path mixer-first receiver with parallel RC baseband load; (b) A circuit with only one signal path of the receiver; (c) The equivalent circuit.

the 25% duty-cycled clock function (high=1, low=0):

$$\begin{aligned} v_{ch}^2 v_S \times \mathrm{clk}(t) &\approx \left(\frac{v_S \times R_{on}}{\sqrt{2}R_S} + v_{BB}\right)^2 v_S \times \mathrm{clk}(t) \\ &= \underbrace{k^2 \times v_S^3 \times \mathrm{clk}(t)}_{\text{where } R_L=0\Omega} + v_{BB}^2 v_S \times \mathrm{clk}(t) + \underbrace{2k v_{BB} v_S^2 \times \mathrm{clk}(t)}_{\text{where } R_L=0\Omega} \end{aligned} \tag{5.16}$$

For two-tone RF input at ω_1 and ω_2, clock at ω_0, the downconverted baseband voltage only contains $(\omega_1 - \omega_0)$ and $(\omega_2 - \omega_0)$ components:

$$\begin{aligned} v_S &= v_{S,p}(\cos\omega_1 t + \cos\omega_2 t) \\ \mathrm{clk}(t) &= \sum_{n=0}^{\infty} a_n \cos n\omega_0 t \\ v_{BB} &\approx v_{BB,p}\big(\sin(\omega_1 - \omega_0)t + \sin(\omega_2 - \omega_0)t\big) \end{aligned} \tag{5.17}$$

where a_n is the n-th Fourier coefficient of 25% duty-cycled clock, $v_{S,p}$ is the peak source voltage, and $v_{BB,p}$ is the peak voltage at baseband. The IM3 component of interest at $(2\omega_1 - \omega_2 - \omega_0)$ (identical coefficients for $(2\omega_2 - \omega_1 - \omega_0)$) is derived:

$$\begin{aligned} v_{ch}^2 v_S \times \mathrm{clk}(t) = k^2 v_{S,p}^3 \times \frac{3a_1}{8}\big(\cos(2\omega_1 - \omega_2 - \omega_0)t\big) \\ \underbrace{+ A_{v,cg}^2 v_{S,p}^3 \times \left(\frac{a_0}{2} - \frac{a_2}{8}\right)\big(\cos(2\omega_1 - \omega_2 - \omega_0)t\big)}_{\text{cancelled by differential output}} \\ + 2kA_{v,cg} v_{S,p}^3 \times \left(\frac{a_0}{2} - \frac{a_2}{8}\right)\big(\sin(2\omega_1 - \omega_2 - \omega_0)t\big) + \ldots \end{aligned} \tag{5.18}$$

where $A_{v,cg} \triangleq \frac{v_{BB,p}}{v_{S,p}}$ is the conversion gain, and $k = \rho/\sqrt{2}$ is evaluated when $R_L = 0\Omega$. Comparing with equation 5.8, the IIP3 can be calculated:

$$\begin{aligned} V^2_{\text{IIP3, }|Z_L|\geq 0\Omega} &\approx \sqrt{\frac{4}{3}\frac{V_{ov}^2}{\left(\left(\frac{8a_0-2a_2}{3a_1}\right)\frac{\rho}{\sqrt{2}}A_{v,cg} + \left(\frac{\rho}{\sqrt{2}}\right)^2\right)\rho}} \\ &\approx \sqrt{\frac{4}{3}\frac{V_{ov}^2}{\left(\sqrt{2}\rho A_{v,cg} + \left(\frac{\rho}{\sqrt{2}}\right)^2\right)\rho}} \\ &\triangleq \sqrt{\frac{4}{3}\frac{V_{ov}^2}{k_{RC}^2\rho}} \end{aligned} \tag{5.19}$$

where k_{RC} is defined:

$$k_{RC}^2 \triangleq \left(\sqrt{2}\rho A_{v,cg} + \left(\frac{\rho}{\sqrt{2}}\right)^2\right) \tag{5.20}$$

The IIP3 dependency on FET size is between $\propto \rho^{-1} - \rho^{-1.5}$.

5.2.2 Clock-Introduced Nonlinearity

In reality, the slope of the clock driving the gate of the FET is finite. This will change the IIP3 derived, because equation 5.3 and 5.15 will no longer be accurate. The finite clock slope will change the IIP3 in two aspects:

1. The ON and OFF switching of the FET are no longer instantaneous. During clock transition, the IM3 contribution may change drastically due to smaller instantaneous V_{ov} and higher instantaneous R_{on}.

2. The ON and OFF timing of the FET switch may be modulated by the blocker, therefore effectively the clock itself contains blocker frequency components.

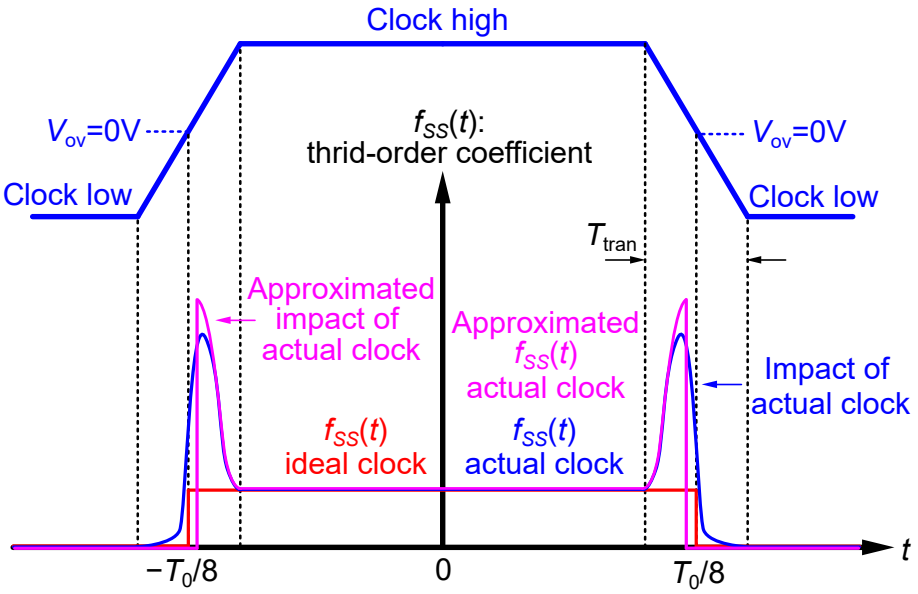


Figure 5.3: The third-order coefficient $\propto f_{SS}(t)$ increases during clock transition. This phenomenon is referred as soft switching effect.

5.2.2.1 IIP3 with Soft Switching

Considering equation 5.2, where G_{on}, α, ρ are periodic functions of time controlled by a soft-switching clock. Expand equation 5.2:

$$\begin{aligned} i_L(t) &\approx v_S\left(\frac{1}{R_S+R_{on}(t)} - \frac{R_S}{(1+R_SG_{on}(t))^3}\Delta G_{on}^2\right) \\ &\approx v_S\left(\frac{1}{R_S+R_{on}(t)} - \frac{R_S}{(1+R_SG_{on}(t))^3}\left(\frac{\partial \Delta G_{on}}{\partial v_{ch}}(t)\right)^2 v_{ch}^2\right) \\ &= v_S\left(\frac{1}{R_S+R_{on}(t)} - f_{SS}(t)v_{ch}^2\right) \end{aligned} \tag{5.21}$$

where $f_{SS}(t) \triangleq \frac{R_S}{(1+R_SG_{on}(t))^3}\left(\frac{\partial \Delta G_{on}}{\partial v_{ch}}(t)\right)^2$ is a periodic function with fundamental frequency ω_0, and $R_S \gg |Z_L|$. $f_{SS}(t)$ serves as the down-converter for the IM3 terms generated by $v_{ch}^2 v_S$. Now, the conversion gain for IM3 is different compared to infinitely sharp clock. On the one hand, the finite clock slope changes the conversion gain of IM3 terms in equation 5.21, because of the overshoot action of $f_{SS}(t)$ shown in Fig. 5.3 during clock transition $G_{on} \to 0$. This indicates a high nonlinearity contribution during clock transition. On the other hand, the finite

clock slope only changes the linear term $\frac{1}{R_S+R_{on}(t)}$ by a small amount since there is no overshoot during clock transition, thus ignored.

Since equation 5.21 involves all strong, moderate, and weak inversion region of the FET, a simplification for $\frac{\partial \Delta G_{on}}{\partial v_{ch}}(t)$ is needed. In strong inversion (FET is on), $G_{on} = \frac{I_{sat}}{V_{ov}/2} = \beta V_{ov}$. In weak inversion (FET is off), $G_{on} = \frac{I_{sat}}{V_T} \approx 0$, where $V_T \triangleq \frac{kT}{e}$ is the thermal voltage. The expression changes at moderate conversion region where $V_{ov} \approx 2V_T$. Therefore, $\frac{\partial \Delta G_{on}}{\partial v_{ch}}$ can be approximated in this way:

$$\begin{aligned} \frac{\partial \Delta G_{on}}{\partial v_{ch}} = \frac{\partial G_{on}}{\partial v_{ch}} &= \beta,\ V_{ov} \geq 2V_T \\ &= 0,\ V_{ov} < 2V_T \end{aligned} \tag{5.22}$$

The approximated $f_{SS}(t)$ is shown in Fig. 5.3. Then the fundamental term of $f_{SS}(t)$, or the IM3 conversion gain, can be calculated with clock shown in Fig. 5.3:

$$\begin{aligned} a_{1,\mathrm{SS}} &= \frac{2}{T_0}\int_{-\frac{T_0}{2}}^{+\frac{T_0}{2}} f_{SS}(t)\times\cos(\omega_0 t)dt \\ &\approx \frac{2}{T_0}\Bigg(\int_{-\frac{T_0}{8}}^{+\frac{T_0}{8}} f_{SS}(0)\times\cos(\omega_0 t)dt \\ &\quad + 2\times\int_{\left(\frac{T_0}{8}-\frac{V_{DD}}{2m_{clk}}\right)}^{\frac{T_0}{8}} f_{SS}(t)\times\cos(\omega_0\times\frac{T_0}{8})dt\Bigg) \\ &= \frac{2}{T_0}\Bigg(\int_{-\frac{T_0}{8}}^{+\frac{T_0}{8}} f_{SS}(0)\times\cos(\omega_0 t)dt \\ &\quad + 2\times\int_0^{\frac{V_{ov}-2V_T}{m_{clk}}}\left(\frac{\beta^2 R_S}{\left(1+R_S\beta(V_{ov}-m_{clk}t)\right)^3}\right)\times\cos(\omega_0\times\frac{T_0}{8})dt\Bigg) \\ &\approx \frac{2}{T_0}\times\left(\frac{T_0}{4}\times\frac{2\sqrt{2}}{\pi}\times\frac{\rho(0)}{R_S}\right) \\ &\quad\times\left(1+\frac{\pi}{4}\frac{1}{\rho(0)^2}\frac{2V_{ov}}{m_{clk}\times T_0}\frac{1}{(1+2R_S\beta V_T)^2}\right) \\ &= a_{1,\mathrm{SS}}|_{m_{clk}\to\infty}\times\left(1+\frac{\pi}{4}\frac{1}{\rho(0)^2}\frac{V_{ov}}{2V_{DD}}\frac{T_{\mathrm{tran}}}{T_0/4}\frac{1}{(1+\frac{2}{\rho(0)}\frac{V_T}{V_{ov}})^2}\right) \\ &\approx a_{1,\mathrm{SS}}|_{m_{clk}\to\infty}\times\left(1+\frac{\pi}{4}\frac{V_{ov}}{2V_{DD}}\frac{T_{\mathrm{tran}}}{T_0/4}\left(\frac{V_{ov}}{2V_T}\right)^2\right) \end{aligned} \tag{5.23}$$

where T_{tran} is the clock transition time, V_{DD} is the clock high voltage, $m_{clk} \triangleq \frac{V_{DD}}{T_{\text{tran}}}$ is the clock transition slope, $T_0 = \frac{2\pi}{\omega_0}$ is the clock period, $a_{1,\text{SS}}|_{m_{clk}\to\infty}$ is the IM3 conversion gain with infinitely sharp clock. The last step assumes $\rho(0) < \frac{2V_T}{V_{ov}}$, which is true when FET W/L is large and V_{ov} is small. The increase of IM3 conversion works for both resistive load and RC load, since $f_{SS}(t)$ and v_{ch} are uncorrelated.

Define factor k_{SS}:

$$k_{SS} \triangleq \left(1 + \frac{\pi}{4}\frac{V_{ov}}{2V_{DD}}\frac{T_{\text{tran}}}{T_0/4}\left(\frac{V_{ov}}{2V_T}\right)^2\right) \tag{5.24}$$

Then the IIP3 with soft switching can then be calculated from equation 5.23 and 5.24:

$$V_{\text{IIP3,SS}} = V_{\text{IIP3, }|Z_L|\geq 0\Omega} \times (k_{SS})^{-\frac{1}{2}} \tag{5.25}$$

When $T_{\text{tran}} = 0\text{s}$ (infinitely sharp clock), $V_{\text{IIP3,SS}} = V_{\text{IIP3, }|Z_L|\geq 0\Omega}$.

5.2.2.2 IIP3 with Timing Modulation

The last section discusses the clock impact on IM3 conversion gain, while assuming $R_{on}(t)$ is a periodic function with fundamental frequency ω_0. However, the finite clock slope also changes ON and OFF timing ($V_g - V_{th} - v_{ch}(v_S) = 0\text{V}$) at the presence of blockers, so the $R_{on}(t)$ not only contains ω_0 components, but also the blocker frequency components. The timing modulation can be considered as the phase-modulation $\theta(v_{ch})$ of the clock.

Considering the resistive load R_L at baseband, equation 5.1 can be rewritten

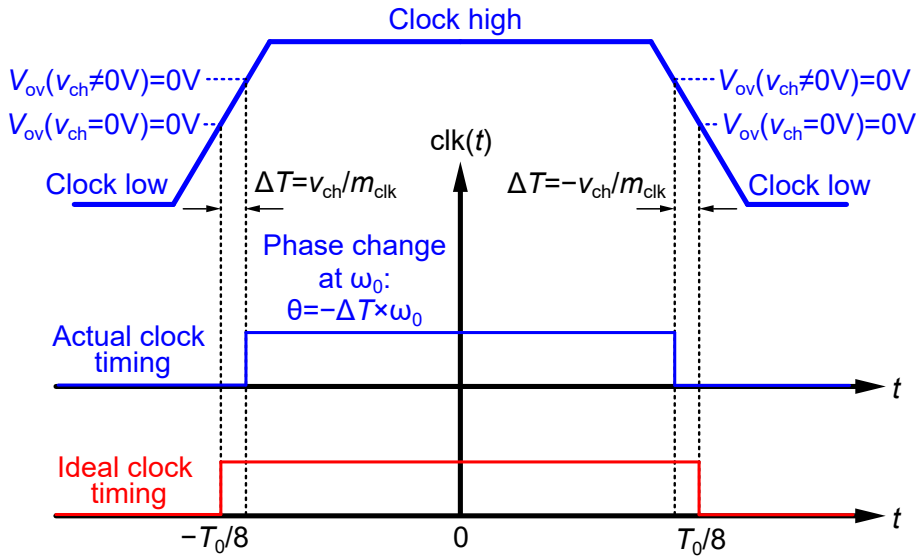


Figure 5.4: The FET ON/OFF timing function clk(t) changes with v_{ch} due to finite clock slope.

as:

$$\begin{aligned} i_L(t) &= a_1(t)v_S + a_2(t)v_S^2 + a_3(t)v_S^3 \\ &= \left(a_1(0)v_S + a_2(0)v_S^2 + a_3(0)v_S^3\right) \times \mathrm{clk}(t) \\ &\approx v_S \times \frac{\alpha(0)}{R_S} \times \left(1 + \rho(0)\left(\frac{\Delta G_{on}}{G_{on}}\right) - \rho(0)\alpha(0)\left(\frac{\Delta G_{on}}{G_{on}}\right)^2\right) \times \mathrm{clk}(t) \\ &\approx \frac{\alpha(0)}{R_S} \times \left(v_S + \rho(0)\left(\frac{k}{V_{ov}}\right)v_S^2 - \rho(0)\alpha(0)\left(\frac{k}{V_{ov}}\right)^2 v_S^3\right) \\ &\quad \times A_{clk}\cos\left(\omega_0 t + \theta(t)\right) \end{aligned} \tag{5.26}$$

where $\mathrm{clk}(t) \approx A_{clk}\cos\left(\omega_0 t + \theta(t)\right)$ contains the component around fundamental frequency ω_0 of the phase-modulated 25% duty-cycled square wave clock, which down-convert the RF signal to baseband. The effect of IM3 conversion gain change is ignored, and they can be considered individually.

The phase modulation of the clock due to finite clock slope is illustrated in Fig. 5.4:

$$\theta(t) = \frac{v_{ch}(t)}{m_{clk}} \times \omega_0 \times f_{\pm 1}(t) \tag{5.27}$$

where $f_{\pm 1} = -1$ for clock rising edge and $f_{\pm 1} = +1$ for clock falling edge. Note that now $\theta \propto v_{ch} \propto v_S$. Then equation 5.26 can be rewritten:

$$\begin{aligned}
i_L(t) &= \big(a_1(0)v_S + a_2(0)v_S^2 + a_3(0)v_S^3\big) \times A_{clk}\cos\big(\omega_0 t + \theta(t)\big) \\
&\approx \big(a_1(0)v_S + a_2(0)v_S^2 + a_3(0)v_S^3\big) \\
&\quad \times A_{clk}\left(\left(1 - \frac{\theta^2(t)}{2}\right)\cos\omega_0 t - \theta(t)\sin\omega_0 t\right) \\
&= A_{clk}\Bigg(a_1(0)v_S\cos\omega_0 t \\
&\quad - \left(a_2(0)\theta(t)v_S^2\sin\omega_0 t + \left(\frac{a_1(0)\theta^2(t)v_S}{2} + a_3(0)v_S^3\right)\cos\omega_0 t\right)\Bigg) \qquad (5.28) \\
&\approx \frac{A_{clk}\alpha(0)}{R_S} \times \Bigg(v_S\cos\omega_0 t - \Bigg(\frac{\rho(0)k^2\omega_0}{V_{ov}m_{clk}}f_{\pm 1}(t)\sin\omega_0 t \\
&\quad + \left(\frac{1}{2}\big(\frac{k\omega_0}{m_{clk}}\big)^2 + \rho(0)\big(\frac{k}{V_{ov}}\big)^2\right)\cos\omega_0 t\Bigg)v_S^3\Bigg) \\
&\approx \frac{A_{clk}\alpha(0)}{R_S} \times \left(v_S\cos\omega_0 t - \left(\frac{1}{2}\Big(\frac{k\omega_0}{m_{clk}}\Big)^2 + \rho(0)\Big(\frac{k}{V_{ov}}\Big)^2\right)\cos\omega_0 t v_S^3\right)
\end{aligned}$$

where the first step uses $\cos(x + dx) \approx \cos x - (\sin x)dx - \frac{1}{2}(\cos x)dx^2$, which is the Taylor expansion up to the second order. The last step used the relation:

$$\frac{1}{2}\Big(\frac{k\omega_0}{m_{clk}}\Big)^2 + \rho(0)\Big(\frac{k}{V_{ov}}\Big)^2 \geq \sqrt{\frac{\rho(0)}{2}}\frac{k^2\omega_0}{V_{ov}m_{clk}} \gg \frac{\rho(0)k^2\omega_0}{V_{ov}m_{clk}} \qquad (5.29)$$

where $\rho(0) \ll 1$. Now in equation 5.28, term $\rho(0)(\frac{k}{V_{ov}})^2$ is introduced by R_{on} modulation, while term $\frac{1}{2}(\frac{k\omega_0}{m_{clk}})^2$ is introduced by timing modulation. They do not interact with one another, thus can be calculated separately. The IIP3 with only timing modulation can be defined:

$$V_{\mathrm{IIP3,TM},R_L} = \sqrt{\frac{8}{3}\frac{V_{DD}^2}{k^2(\omega_0 T_{\mathrm{tran}})^2}} \qquad (5.30)$$

When the resistive load R_L is changed to RC load Z_L, the IIP3 is similar, where the difference is $k \to k_{RC}$ as analyzed in equation 5.20:

$$V_{\mathrm{IIP3,TM}} = \sqrt{\frac{8}{3}\frac{V_{DD}^2}{k_{RC}^2(\omega_0 T_{\mathrm{tran}})^2}} \qquad (5.31)$$

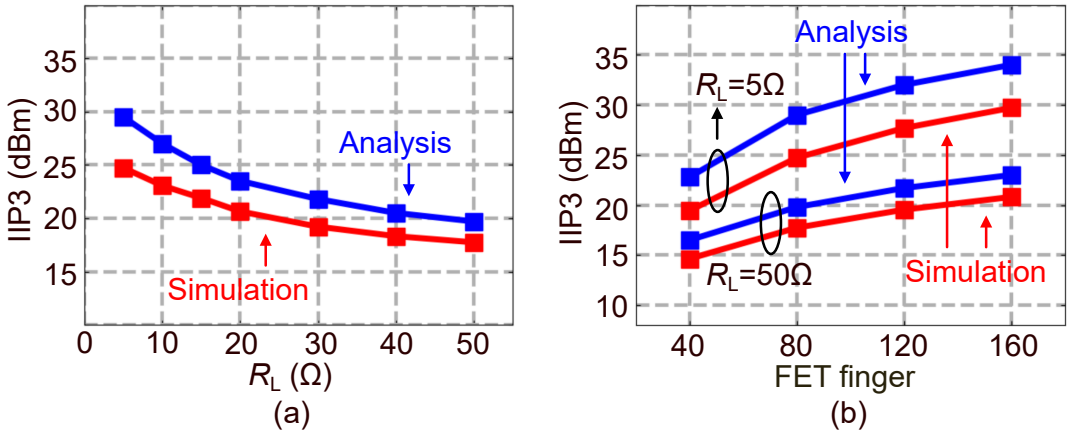


Figure 5.5: (a) The comparison between analyzed and simulated IIP3 versus load resistance R_L, where FET finger is 80; (b) The comparison between analyzed and simulated IIP3 versus FET sizes.

5.2.3 Complete IIP3 Expression

When a common parallel RC load is considered, the final IIP3 expression can be derived from equation 5.19, 5.25 and 5.31:

$$\begin{aligned} V^2_{\text{IIP3,tot}} &= V^2_{\text{IIP3,TM}} \| V^2_{\text{IIP3,SS}} \\ &= V^2_{\text{IIP3,TM}} \| (k_{SS}^{-1} V^2_{\text{IIP3, } |Z_L| \geq 0\Omega}) \end{aligned} \tag{5.32}$$

where $V^2_{\text{IIP3,TM}}$ is from timing modulation, $V_{\text{IIP3, } |Z_L| \geq 0\Omega}$ from R_{on} modulation, and k_{SS} is from soft switching.

To summarize, both the FET nonlinear R_{on} and finite slope clock contribute to the IIP3. The R_{on} modulation and timing modulation both contribute to the total IIP3 independently without interaction, and the soft switched FET during clock transition degrades R_{on} modulation IIP3.

5.2.4 Analysis Verification

Simulations are performed to verify the correctness of the analysis in 65nm RF-SOI CMOS technology.

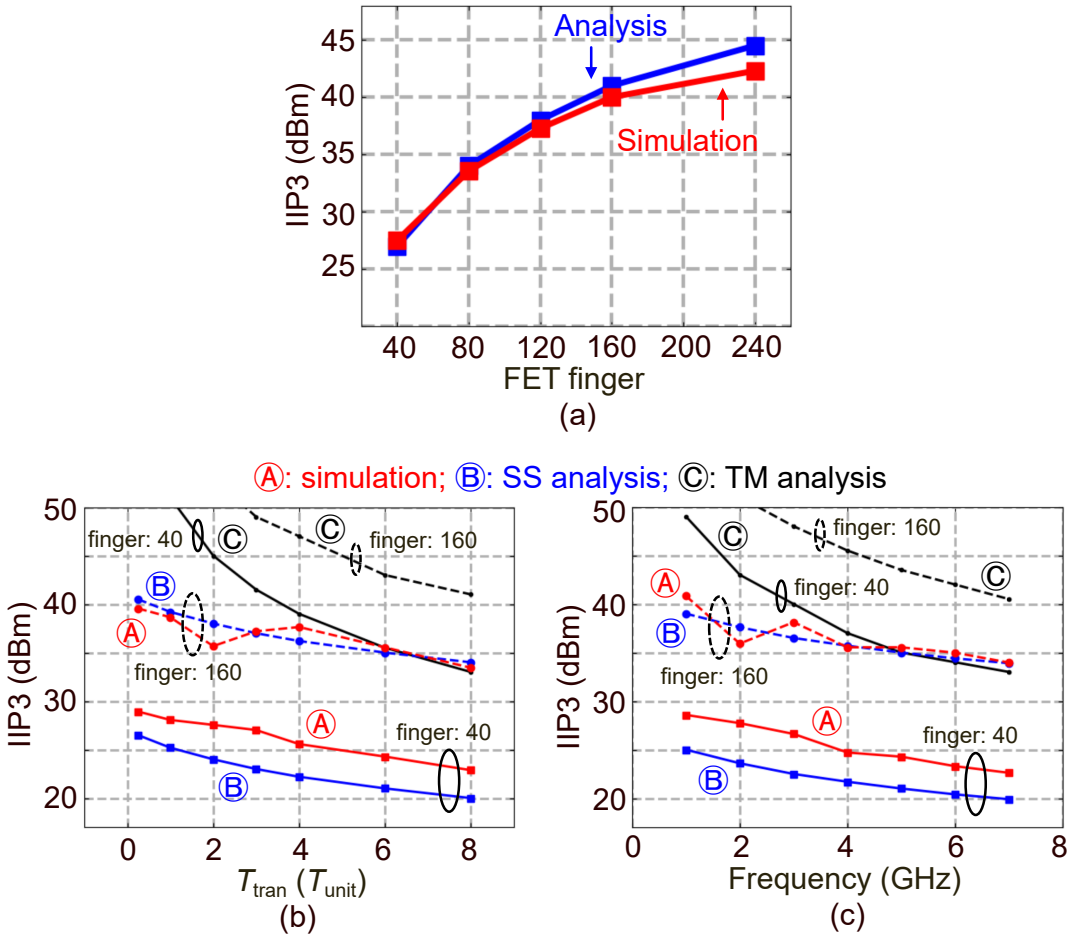


Figure 5.6: (a) The comparison between analyzed and simulated OOB-IIP3 versus FET sizes with ideal clock; (b) The comparison between analyzed and simulated OOB-IIP3 versus clock transition time T_{tran} at $f_0 = 5$GHz, where $T_{\text{unit}} = 3.25$ps; (c) The comparison between analyzed and simulated OOB-IIP3 versus clock frequency f_0, where $T_{\text{tran}} = 6T_{\text{unit}}$.

First, the resistive loaded mixer IIP3 with R_{on} modulation only (equation 5.8) is verified. The k factor defined in equation 5.14 depends on both R_L and R_{on}. Therefore, the IIP3 with resistive load depends on both R_L and R_{on}. Fig. 5.5 shows the analysis versus simulation. The FET finger is swept from $40 - 160$, which results in $2\Omega \leq R_{on} \leq 8\Omega$. The simulation shows agreement with the analysis with $\sim$ 5dB error in the worst case.

Second, the RC loaded mixer IIP3 with R_{on} modulation only (equation 5.19) is verified. Fig. 5.6 shows the analysis versus the simulation, where $A_{v,cg} = 3/50$, $V_{ov} = 0.4$V. They match well.

Third, the impact of deteriorated clock on mixer IIP3 (equation 5.25 and 5.31) is verified. Both soft switching and timing modulation change IIP3. The analysis indicates that the mixer IIP3 is functions of clock transition time T_{tran} and clock frequency f_0. Fig. 5.6 shows the analysis versus simulation when T_{tran} and f_0 are swept. The results show that the contribution from timing modulation can be ignored. The soft switching effect explains how the clock waveform and operating frequency affect the IIP3.

To conclude, the mixer IIP3 analysis matches the simulation very well. It reveals how IIP3 changes with FET size, clock waveform, frequency and baseband load.

There are multiple ways to improve IIP3:

1. Increase V_{DD}, or bias the FET gate at a higher voltage. The former results in higher power consumption.

2. Fast clock generation circuit. This leads to higher power consumption.

3. Large FET W/L and small on-resistance. This leads to higher power consumption if the clock generation is scaled accordingly.

5.3 OOB B1dB Analysis

Many works [22–26] have reported the measured B1dB of the passive mixer does not follow the $\sim$ 10dB difference with the measured IIP3. This indicates that the B1dB point does not originate from third-order nonlinearity alone. Higher-order nonlinearity contributes to the B1dB, therefore it is necessary to conduct large-signal instead of small-signal analysis.

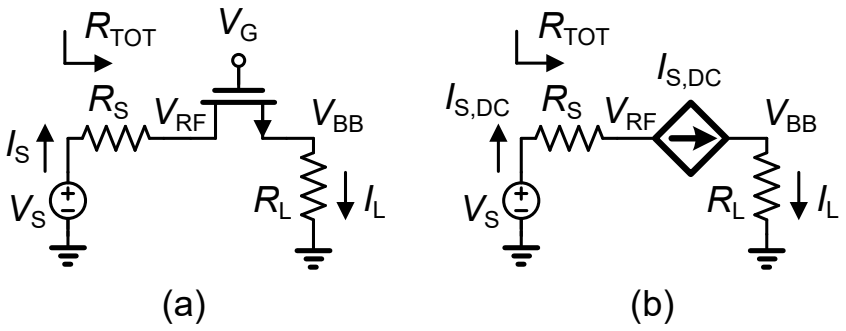


Figure 5.7: (a) DC analysis of a circuit similar to one signal path of the mixer-first receiver; (b) The FET is at the onset of the saturation.

Similar to IIP3 analysis, both FET (size, overdrive voltage) and clock (slope) affect the B1dB. When blocker power increases, the R_{on} modulation increases expansively instead of compressively, resulting in B1dB closely related to the onset of saturation. Both effects will be analyzed in the following sections, and the agreement between simulation and analysis will be shown.

5.3.1 R_{on} Nonlinearity

5.3.1.1 Compression of DC I-V Characteristics

Before going into details of B1dB analysis, it is helpful to look at the DC I-V characteristics with a resistor load R_L (Fig. 5.7). The voltage gain V_{BB}/V_S can compress when V_S is too large due to FET on-resistance R_{on} modulation. It is equivalent to the total series resistance $R_{TOT} = V_S/I_S$ expansion, because $V_{BB}/V_S = R_L \times (I_S/V_S)$.

The total resistance $R_{TOT} = V_{RF}/I_S + R_S$, where the V_{RF} and I_S satisfy the equation:

$$\begin{aligned} I_S &= \frac{\beta}{2}(V_{ov} - V_{RF})^2 - \frac{\beta}{2}(V_{ov} - V_{BB})^2 \\ I_S &= \frac{\beta}{2}(V_{ov} - V_{RF})^2 - \frac{\beta}{2}(V_{ov} - I_S R_L)^2 \end{aligned} \tag{5.33}$$

where $V_{ov} = V_G - V_{th}$. V_{RF} can be derived in terms of I_S:

$$V_{RF} = V_{ov}\left(1 - \sqrt{1 - \frac{2(R_{on} + R_L)}{V_{ov}} I_S + \left(\frac{R_L}{V_{ov}}\right)^2 I_S^2}\right) \tag{5.34}$$

The total resistance will increase drastically when the FET switch goes into saturation, therefore 1dB compression of DC I-V characteristics happens either at the onset or before the FET switch goes into saturation.

The source current of Fig. 5.7 at the onset of saturation $I_{S,DC}$ can be solved by the equation:

$$I_{S,DC} = \frac{\beta}{2}(V_{ov} - I_{S,DC}R_L)^2 \tag{5.35}$$

Then $I_{S,DC}$ is solved:

$$I_{S,DC} = \frac{V_{ov}}{R_L}\left(\left(1 + \frac{R_{on}}{R_L}\right) - \sqrt{\left(1 + \frac{R_{on}}{R_L}\right)^2 - 1}\right) \tag{5.36}$$

where $R_{on} = 1/(\beta V_{ov})$ is the small-signal on-resistance of the FET. The source power can be calculated accordingly.

The gain compression can also happen before saturation. However, the DC analysis will not help the upcoming analysis, therefore skipped.

5.3.1.2 Gain Compression with a RF Blocker

The discussion on R_{on} modulation will be divided into two cases:

1. The 1dB compression happens when FET is in triode region.
2. The 1dB compression happens when FET is in saturation region.

First, the B1dB assuming FET in triode region is analyzed.

Now consider large-signal scenario in equation 5.1. The type of load does not

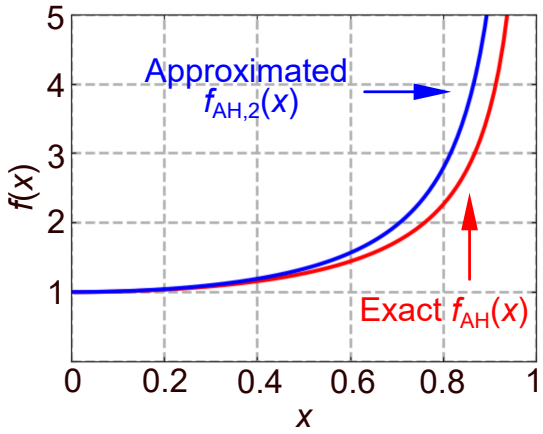


Figure 5.8: The effective third-order coefficient amplification factor $f_{AH}(x)$ for gain compression analysis, and its approximation $f_{AH,2}(x)$.

matter in this discussion as long as $R_S \gg |Z_L|$. The full expression of i_L is:

$$\begin{aligned} i_L &\approx v_S \times \frac{G_{on} + \Delta G_{on}}{R_S \times (G_{on} + \Delta G_{on}) + 1} \\ &\approx \frac{v_S}{R_S + R_{on}} \times \left(1 + \rho \sum_{n=0}^{\infty} (-\alpha)^n \left(\frac{\Delta G_{on}}{G_{on}}\right)^{n+1}\right) \\ &= \frac{v_S}{R_S + R_{on}} \times \left(1 + \rho \sum_{n=1}^{\infty} (\alpha)^{2n-1} \left(\frac{\Delta G_{on}}{G_{on}}\right)^{2n}\right) + \ldots \end{aligned} \tag{5.37}$$

where i_L, ρ, α, R_{on}, G_{on} and ΔG_{on} can be functions of time or DC variables. Only even-order terms of $\Delta G_{on} \propto v_S$ (resulting in odd-order terms of v_S) contribute to gain compression, thus kept. For equation 5.37 to be valid, $\Delta G_{on} \geq -G_{on}$. Note that saturation happens at some point when $G_{on} + \Delta G_{on} > 0$, therefore for some small ΔG_{on} equation 5.37 will be invalid.

For RF input $v_S = v_{S,p} \cos \omega_{RF} t$ and resistive load R_L at baseband, $\Delta G_{on} = \beta v_{ch} = \beta k v_S$, where k is given by equation 5.14. Equation 5.37 is used to calculate the gain compression from all higher-order (odd-order) terms. Then their effects are lumped into an amplified third-order coefficient, as if only the third-order term

contributes to the gain compression. The steps are shown below:

$$
\begin{aligned}
i_L &\approx \frac{v_{S,p}}{R_S+R_{on}} \times \Bigg(\cos\omega_{RF}t \\
&\quad + \rho \sum_{n=1}^{\infty} (\alpha)^{2n-1} \left(\frac{k v_{S,p}}{V_{ov}}\right)^{2n} \left(\cos^{2n+1}\omega_{RF}t \right) \Bigg) \\
&= \frac{v_{S,p}}{R_S+R_{on}} \times \Bigg(\cos\omega_{RF}t \\
&\quad + \rho \sum_{n=1}^{\infty} (\alpha)^{2n-1} \left(\frac{k v_{S,p}}{V_{ov}}\right)^{2n} \left(\frac{\binom{2n+2}{n+1}}{2^{2n+1}} \cos\omega_{RF}t \right) \Bigg) + ... \\
&= \frac{v_S}{R_S+R_{on}} \times \left(1 + \rho \frac{2}{\left(\frac{k v_{S,p}}{V_{ov}}\right)^2 \alpha^3} \sum_{n=2}^{\infty} \binom{2n}{n} \left(\frac{\left(\frac{k v_{S,p}}{V_{ov}}\right)^2 \alpha^2}{4} \right)^n \right) + ... \qquad (5.38) \\
&\triangleq \frac{v_S}{R_S+R_{on}} \times \left(1 + \frac{\frac{3}{4}\rho\alpha\left(\frac{k v_{S,p}}{V_{ov}}\right)^2}{\frac{3}{8}x^4} \sum_{n=2}^{\infty} \binom{2n}{n} \left(\frac{x^2}{4} \right)^n \right) + ... \\
&= \frac{v_S}{R_S+R_{on}} \times \left(1 + \frac{\frac{3}{4}\rho\alpha\left(\frac{k v_{S,p}}{V_{ov}}\right)^2}{\frac{3}{8}x^4} \left(\frac{1}{\sqrt{1-x^2}} - 1 - \frac{x^2}{2} \right) \right) + ... \\
&\triangleq \frac{v_S}{R_S+R_{on}} \times \left(1 + \frac{3}{4}\rho\alpha\left(\frac{k v_{S,p}}{V_{ov}}\right)^2 \times f_{AH}(x) \right) + ... \\
&= \left(a_1 v_{S,p} + \frac{3}{4} a_3 v_{S,p}^3 \times f_{AH}(x) \right) \cos\left(\omega_{RF} - \omega_0\right)t + ...
\end{aligned}
$$

where $x \triangleq \frac{\alpha k v_{S,p}}{V_{ov}}$ is proportional to the signal strength, $f_{AH}(x) \triangleq \frac{8}{3x^4}(\frac{1}{\sqrt{1-x^2}} - 1 - \frac{x^2}{2})$ is the amplification factor lumping all higher-order impact on gain compression into third-order coefficient a_3 and is shown in Fig. 5.8. Fundamental frequency terms of $\cos^n \omega_{RF}$ are kept in the expression to calculate the gain compression. $f_{AH}(x)$ is derived analytically due to the fact:

$$
\sum_{n=0}^{\infty} \binom{2n}{n} x^n = \frac{1}{\sqrt{1-4x}} \qquad (5.39)
$$

$\lim_{x\to 0} f_{AH}(x) = 1$ indicates that the gain compression is caused only by the third-order term for small signal. $\lim_{x\to 1} f_{AH}(x) = \infty$ indicates the gain compression is caused by all higher-order nonlinearity terms for large signal, and it diverges at

$x = \frac{\alpha k v_{S,p}}{V_{ov}} = 1$. Define:

$$\begin{aligned} v_{S,x=1} &\triangleq \frac{V_{ov}}{\sqrt{2}\alpha k} \\ f_{AH,2}(x) &\triangleq \sum_{n=0}^{\infty} x^{2n} = \frac{1}{1-x^2} \\ \mathrm{dB}(V_{\mathrm{B1dB},R_L \neq 0\Omega}) &\approx \mathrm{dB}(V_{\mathrm{IIP3},R_L \neq 0\Omega}) - 9.6\mathrm{dB} - 3\mathrm{dB} \end{aligned} \tag{5.40}$$

where $v_{S,x=1}$ is the effective voltage where $x = 1$, and 3dB in $V_{\mathrm{B1dB},R_L \neq 0\Omega}$ is from the difference of P1dB (one-tone signal gain compression) and B1dB (small-signal gain compression under large blocker). $f_{AH,2}$ is the approximation of f_{AH} assuming $\frac{\binom{2n+2}{n+1}}{2^{2n+1}} \approx \frac{\binom{2+2}{1+1}}{2^{2+1}} = \frac{3}{4}$, $\forall\ n \in \mathbb{Z}^+$, indicating $\cos^{2n+1} t \approx \frac{3}{4}\cos t + ...$, $\forall\ n \in \mathbb{Z}^+$, and the comparison is shown in Fig. 5.8. Now if f_{AH} can be approximated by $f_{AH,2}$ for simplicity:

$$i_L \approx \left(a_1 v_{S,p} + \frac{3}{4} a_3 v_{S,p}^3 \times f_{AH,2}(x) \right) \cos\left(\omega_{RF} - \omega_0\right) t + ... \tag{5.41}$$

Then the B1dB can be calculated by equation 5.40, 5.41 and 5.8:

$$\begin{aligned} V_{\mathrm{B1dB,triode}}^2 &= V_{\mathrm{B1dB},R_L \neq 0\Omega}^2 \left(1 - \frac{V_{\mathrm{B1dB,triode}}^2}{v_{S,x=1}^2} \right) \\ \Rightarrow V_{\mathrm{B1dB,triode}}^2 &= V_{\mathrm{B1dB},R_L \neq 0\Omega}^2 \| v_{S,x=1}^2 \end{aligned} \tag{5.42}$$

For RX baseband load, make changes in equation 5.42 that $k \rightarrow k_{RC}$ and $V_{\mathrm{B1dB},R_L \neq 0\Omega} \rightarrow V_{\mathrm{B1dB},|Z_L| \geq 0\Omega}$:

$$V_{\mathrm{B1dB,triode}}^2 \approx V_{\mathrm{B1dB},|Z_L| \geq 0\Omega}^2 \| v_{S,x=1}^2 \tag{5.43}$$

where $\mathrm{dB}(V_{\mathrm{B1dB},|Z_L| \geq 0\Omega}) \approx \mathrm{dB}(V_{\mathrm{IIP3},|Z_L| \geq 0\Omega}) - 12.6\mathrm{dB}$.

The discussion on FET in triode region is finished. Next, the B1dB assuming FET in saturation region is analyzed.

When the FET goes into saturation, large compression will kick in. Therefore, approximately the 1dB compression happens at the onset of saturation. For a RF input, this means at worst case the FET is at the onset of saturation.

Note that the source looks like a current source since $R_S \gg R_{on}$ and $R_S \gg |Z_L|$. Thus, for RF input source current $i_S = i_{S,p} \cos \omega_{RF} t$, the $i_{S,p}$ can be estimated assuming it drives the FET into saturation when the baseband reaches the peak voltage $v_{BB,p}$.

$$\begin{aligned} i_{S,p} &\approx \frac{\beta}{2}(V_{ov} - v_{BB,p})^2 \\ &= \frac{\beta}{2}(V_{ov} - i_{S,p} \times A_{v,cg} \times R_S)^2 \end{aligned} \tag{5.44}$$

where $A_{v,cg} \triangleq v_{BB,p}/v_{S,p}$ is the conversion gain. Equation 5.44 is similar to equation 5.35, therefore equation 5.36 can be used:

$$i_{S,p} = \frac{V_{ov}}{A_{v,cg} R_S} \left(\left(1 + \frac{R_{on}}{A_{v,cg} R_S}\right) - \sqrt{\left(1 + \frac{R_{on}}{A_{v,cg} R_S}\right)^2 - 1} \right) \tag{5.45}$$

Then the B1dB is calculated from $i_{S,p}$:

$$V_{\text{B1dB,sat}} = \frac{i_{S,p} R_S}{\sqrt{2}} \tag{5.46}$$

By assuming the operation region of the FET switch, the B1dB is calculated above. The remaining question is: what is the FET operation region at 1dB compression?

1. $V_{\text{B1dB,sat}} > V_{\text{B1dB,triode}}$:

   In this case, 1dB compression happens before FET entering saturation.

2. $V_{\text{B1dB,sat}} \leq V_{\text{B1dB,triode}}$:

   In this case, 1dB compression is caused by FET entering saturation.

Therefore,

$$V_{\text{B1dB},R_{on}} \approx \min\left(V_{\text{B1dB,sat}}, V_{\text{B1dB,triode}}\right) \tag{5.47}$$

5.3.2 Clock-Introduced Nonlinearity

The nonlinearity analysis for IIP3 (small signal) is not applied here for two reasons:

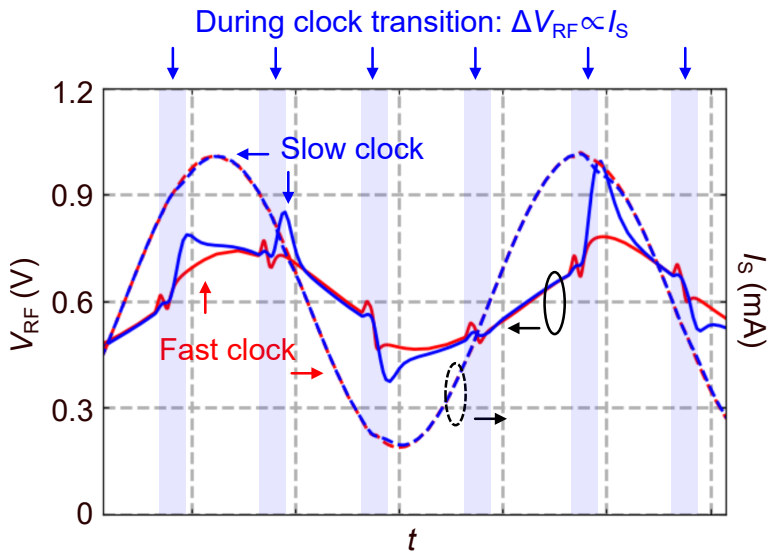


Figure 5.9: The simulated V_{RF} and I_S waveforms with fast and slow clocks. The V_{RF} node accumulates charge during clock transition and discharge during rest of the time.

1. During clock transition, the FET will be most likely in saturation, instead of triode region which is used in small-signal analysis. Therefore the soft switching analysis for IIP3 is invalid for B1dB.

2. The timing modulation usually plays a minor role (Fig. 5.6) in IIP3, for simplicity it is ignored here.

Here, a simple large signal model is developed based on observations from the simulation (Fig. 5.9). The model focuses on the charge at v_{RF} node during clock transition (non-zero transition time T_{tran}) and clock being high. Because 1dB compression happens at FET entering saturation (when the clock is high), the FET is mostly in saturation during clock transition. Note that $|Z_L| \ll R_S$ and $R_{on} \ll R_S$, thus the i_S is nearly a current source (with at most 1dB change due to compression) with negligible harmonic components. The saturation current is not sufficient to sink all the source current i_S during transition, thus accumulated at v_{RF} node. This accumulated charge must be recovered/discharged during this cycle in the worst case, such that the overall waveform is not disturbed. To

discharge, the current through the FET increases, so does the nonlinearity. At 1dB compression, the increased FET current is the B1dB current derived in the last section.

The charge equation in one clock cycle can be written at the B1dB point:

$$(i_{\mathrm{B1dB},R_{on}} - i_{\mathrm{B1dB,tot}}) \times (\frac{T_0}{4} - T_{\mathrm{tran}}) \approx i_{\mathrm{B1dB,tot}} \times T_{\mathrm{tran}} \tag{5.48}$$

where $i_{\mathrm{B1dB},R_{on}} = \frac{V_{\mathrm{B1dB},R_{on}}}{R_S}$, and $i_{\mathrm{B1dB,tot}}$ is the B1dB considering both the R_{on} and the clock-introduced nonlinearity. The B1dB point is then calculated from equation 5.48:

$$\begin{aligned} i_{\mathrm{B1dB,tot}} &\approx \frac{T_0/4 - T_{\mathrm{tran}}}{T_0/4} i_{\mathrm{B1dB},R_{on}} \\ V_{\mathrm{B1dB,tot}} &\approx \frac{T_0/4 - T_{\mathrm{tran}}}{T_0/4} V_{\mathrm{B1dB},R_{on}} \end{aligned} \tag{5.49}$$

The analysis on clock-introduced nonlinearity is finished. Note that this is a simplified model, which lumps the complicated transient response into a linear process.

5.3.3 Analysis Verification

Simulations are performed to verify the correctness of the analysis.

First, the RC loaded mixer B1dB with R_{on} modulation only (equation 5.47) is verified. Equation 5.43 and 5.46 are evaluated so the FET region of operation can be determined when 1dB compression happens. Fig. 5.10 shows the analysis versus the simulation, where $A_{v,cg} = 3/50$, $V_{ov} = 0.4$V. The FET finger is swept from $40 - 160$, which results in $2\Omega \leq R_{on} \leq 8\Omega$. The predicted B1dB by equation 5.43 and 5.46 are almost identical, indicating the 1dB compression happens near the onset of FET saturation, and can be explained by both approaches. The simulation shows agreement with the analysis.

Second, the impact of deteriorated clock on mixer B1dB (equation 5.49) is verified. The analysis indicates the mixer B1dB is functions of clock transition

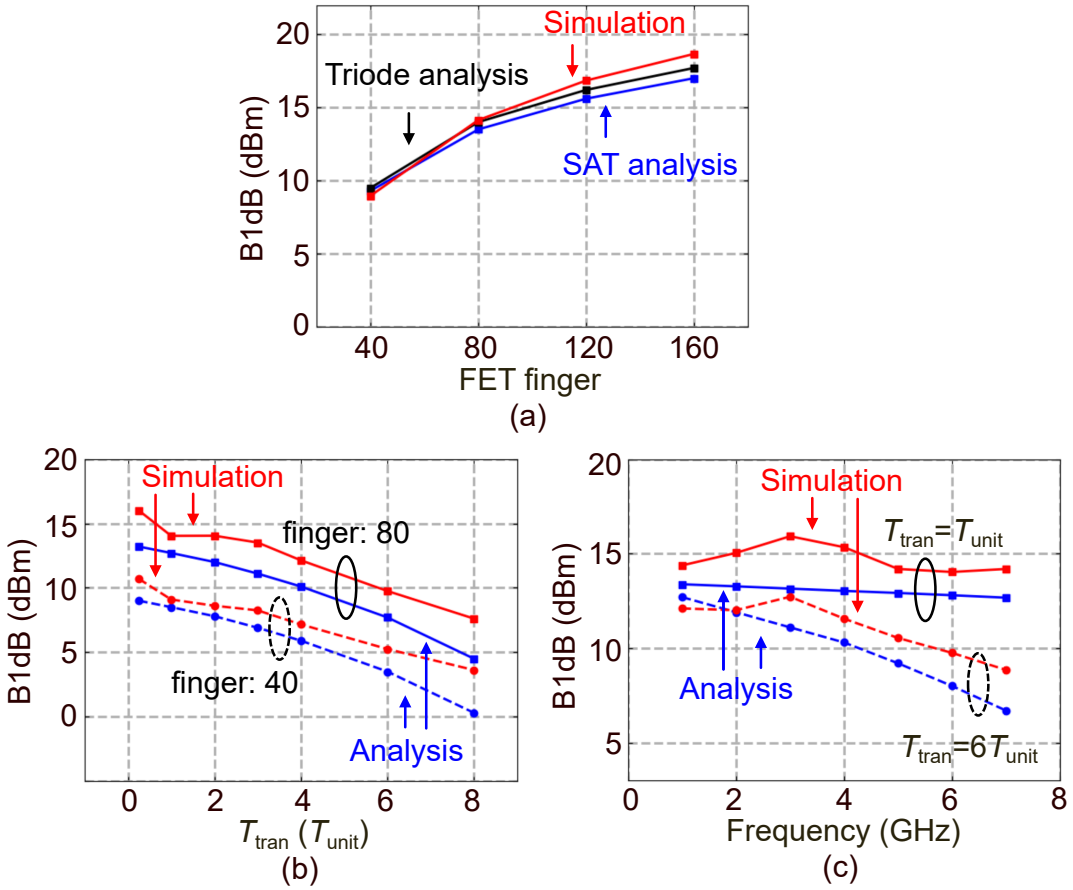


Figure 5.10: (a) The comparison between analyzed and simulated B1dB versus FET sizes with ideal clock; (b) The comparison between analyzed and simulated B1dB versus clock transition time T_{tran} at f_0 = 6GHz, where T_{unit} = 3.25ps; (c) The comparison between analyzed and simulated B1dB versus clock frequency f_0, where the FET finger is 80.

time T_{tran} and clock frequency f_0. Fig. 5.10 shows the analysis versus simulation when T_{tran} and f_0 are swept. The simple model successfully explains how the clock waveform and operating frequency affect the B1dB.

To conclude, the mixer B1dB analysis matches the simulation very well. It reveals how B1dB changes with FET size, clock waveform, frequency and baseband load.

There are multiple ways to improve B1dB:

Table 5.1: Analysis versus mixer-first receiver measurement in Chapter 4.

	Analysis	Measurement
IIP3 (f_{os}/BBBW)	29.8dBm* (10)	28.5dBm (9)
B1dB (f_{os}/BBBW)	11.1dBm* (10)	11.8dBm (12)
* $T_{\text{tran}} = 6T_{unit}$, finger = 80, f_0 = 3GHz.		

Table 5.2: Analysis versus cascaded (EBD + RX) measurement in Chapter 4.

	Analysis	Measurement
IIP3 (f_{os}/BBBW)	27.5dBm* (10)	23.8dBm** (12)
B1dB (f_{os}/BBBW)	8dBm* (10)	9.9dBm** (9)
* $T_{\text{tran}} \approx 6T_{unit}$, finger = 80, f_0 = 6GHz.		
** EBD IL $\approx$ 4dB subtracted.		

1. Increase V_{DD}, or bias the FET gate at a higher voltage. The former results in higher power consumption.

2. Fast clock generation circuit. This leads to higher power consumption.

3. Large FET W/L and small on-resistance. This leads to higher power consumption if the clock generation is scaled accordingly.

4. Use CMOS switch instead of NMOS switch. [50, 51] shows that the use the CMOS switch instead of NMOS switch can improve the B1dB. That is because the CMOS switch is always ON regardless of the blocker power, while the NMOS switch can go into saturation causing compression.

5.4 Analysis Verification with Measurement

Although the analysis on mixer IIP3 and B1dB has been validated with simulations, it is more solid to compare with measurements. In Chapter 4, IIP3 and

B1dB measurements are preformed for the fabricated mixer-first receiver as well as the EBD and receiver in cascade. For far-out blockers, the IIP3 and B1dB of the receiver are dominated by the mixer. They can be used to verify the mixer analysis.

Table 5.1 compares the receiver measurements and the analysis for far-out blockers. The clock transition time is determined by simulation. The analysis matches the measurement results remarkably well for both IIP3 and B1dB with $\sim$ 1dB accuracy.

Table 5.2 compares the cascaded measurements and the analysis for far-out blockers. The EBD RX IL is 4dB from the EBD measurement shown in Fig. 2.16, and is subtracted from the cascaded measurements for comparison. The analysis reasonably matches the measurement results with $\sim$ 3dB accuracy.

CHAPTER 6

Conclusions

This dissertation has shown designs and analysis that can help overcome the challenges in the RF front-ends, including the duplexer and the receiver.

The tunable integrated duplexer is of interest because it may replace the bulky and costly SAW duplexers for FDD radios. This dissertation explores a new passive approach to realize dual-band TX-RX isolation for actual antennas without deteriorating functionality over practical TX power. The new design approach is a result of interdisciplinary research between the analog/RF integrated circuit community and the microwave circuit community.

The key contributions of the duplexer include:

- A state-of-the-art dual-band EBD with a novel independently tuned dual-band balance network.
- An analysis on finite-Q LCR networks for CMOS passive network design. The analysis predicts the design limits, and gives guidelines for the balance network design as well as other CMOS passive network design.
- The perfectly-matched absorptive bandstop filter is realized on a chip. It achieves theoretical limits on $|s_{21}|$ transition rate predicted by the analysis.

The finite-Q LCR network analysis also indicates that there is still room for improvement. The possible future works include:

- A novel design on the balance network of EBD that increases Q_{ANT} coverage.

There is still plenty of room to improve Q_{ANT} coverage from the theoretical limits indicated by equation 3.31.

- A new topology to reduce the occupied chip area by reducing the number of inductors and tunable capacitors.

- Instead of the passive cancellation duplexer (e.g. EBD), a tunable passive filter-based integrated duplexer (e.g. similar to SAW duplexer) may also be implemented.

The blocker-tolerant receiver is also desired for SAW-less RF front-ends. This dissertation studies ways to improve the mixer-first receiver dynamic range in the presence of blockers.

The key contributions of the receiver include:

- The design methodology of a second-order baseband TIA for the mixer-first receiver is shown. Independent ω_0 and Q control is realized. The second-order TIA increases the receiver dynamic range at the present of OOB blockers.

- Comprehensive passive mixer IIP3 and B1dB analysis is given. The analysis reveals quantitatively that how the FET sizes, clock waveform, operating frequency, and baseband load contribute to the IIP3 and B1dB. The effects of the deteriorated clock waveform and operating frequency are analyzed for the first time. And the entire B1dB analysis is given for the first time as well. The analysis provides design guidelines to optimize the passive mixer for maximal linearity.

Possible improvements can be made on the receiver design as well:

- The independent ω_0 and Q control of the TIA puts a requirement on minimum G_m and this limits the further power reduction for the inverter-based

transconductance amplifier. An op amp, or a voltage amplifier (i.e., high A_0, low r_0) may be used to replace the transconductance amplifier for higher equivalent G_m with lower power consumption.

- A more advanced technology with faster FETs can improve the passive mixer noise, linearity and power performance.
- The mixer switches can be implemented by CMOS switches to improve B1dB.

REFERENCES

[1] M. Mikhemar, H. Darabi, and A. A. Abidi, "A Multiband RF Antenna Duplexer on CMOS: Design and Performance," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 9, pp. 2067–2077, 2013.

[2] B. van Liempd, A. Visweswaran, S. Ariumi, S. Hitomi, P. Wambacq, and J. Craninckx, "Adaptive RF Front-Ends Using Electrical-Balance Duplexers and Tuned SAW Resonators," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 11, pp. 4621–4628, 2017.

[3] G. Qi, B. van Liempd, P. Mak, R. P. Martins, and J. Craninckx, "A SAW-Less Tunable RF Front End for FDD and IBFD Combining an Electrical-Balance Duplexer and a Switched-LC N-Path LNA," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 5, pp. 1431–1442, 2018.

[4] L. Laughlin, C. Zhang, M. A. Beach, K. A. Morris, J. L. Haine, M. K. Khan, and M. McCullagh, "Tunable Frequency-Division Duplex RF Front End Using Electrical Balance and Active Cancellation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 12, pp. 5812–5824, 2018.

[5] S. H. Abdelhalem, P. S. Gudem, and L. E. Larson, "Tunable CMOS Integrated Duplexer With Antenna Impedance Tracking and High Isolation in the Transmit and Receive Bands," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 9, pp. 2092–2104, 2014.

[6] B. Hershberg, B. van Liempd, X. Zhang, P. Wambacq, and J. Craninckx, "20.8 A dual-frequency 0.7-to-1GHz balance network for electrical balance duplexers," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, 2016, pp. 356–357.

[7] A. C. Guyette, I. C. Hunter, R. D. Pollard, and D. R. Jachowski, "Perfectly-matched bandstop filters using lossy resonators," in *IEEE MTT-S International Microwave Symposium Digest, 2005.*, 2005, pp. 4 pp.–520.

[8] R. C. Frye, S. Kapur, and R. C. Melville, "A 2-GHz quadrature hybrid implemented in CMOS technology," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3, pp. 550–555, 2003.

[9] T. Lee, B. Kim, K. Lee, W. J. Chappell, and J. Lee, "Frequency-Tunable Low- Q Lumped-Element Resonator Bandstop Filter With High Attenuation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 11, pp. 3549–3556, 2016.

[10] M. Mikhemar, H. Darabi, and A. Abidi, "A tunable integrated duplexer with 50dB isolation in 40nm CMOS," in *2009 IEEE International Solid-State Circuits Conference - Digest of Technical Papers*, 2009, pp. 386–387,387a.

[11] B. van Liempd, J. Craninckx, R. Singh, P. Reynaert, S. Malotaux, and J. R. Long, "A Dual-Notch +27dBm Tx-Power Electrical-Balance Duplexer," in *ESSCIRC 2014 - 40th European Solid State Circuits Conference (ESSCIRC)*, 2014, pp. 463–466.

[12] M. Elkholy, M. Mikhemar, H. Darabi, and K. Entesari, "Low-Loss Integrated Passive CMOS Electrical Balance Duplexers With Single-Ended LNA," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 5, pp. 1544–1559, 2016.

[13] B. van Liempd, B. Hershberg, K. Raczkowski, S. Ariumi, U. Karthaus, K. Bink, and J. Craninckx, "2.2 A +70dBm IIP3 single-ended electrical-balance duplexer in 0.18um SOI CMOS," in *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*, 2015, pp. 1–3.

[14] B. van Liempd, B. Hershberg, S. Ariumi, K. Raczkowski, K. Bink, U. Karthaus, E. Martens, P. Wambacq, and J. Craninckx, "A +70-dBm IIP3 Electrical-Balance Duplexer for Highly Integrated Tunable Front-Ends," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 12, pp. 4274–4286, 2016.

[15] T. Zhang, A. R. Suvarna, V. Bhagavatula, and J. C. Rudell, "An Integrated CMOS Passive Self-Interference Mitigation Technique for FDD Radios," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 5, pp. 1176–1188, 2015.

[16] H. Yüksel, T. Tapen, Z. Boynton, E. Enroth, A. Apsel, and A. C. Molnar, "An FDD/FD Capable, Single Antenna RF Front End from 800MHz to 1.2GHz w/ Baseband Harmonic Predistortion," in *2018 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2018, pp. 120–123.

[17] T. Tapen, H. Yüksel, Z. Boynton, A. Apsel, and A. Molnar, "An Integrated, Software-Defined FDD Transceiver: Distributed Duplexing Theory and Operation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 1, pp. 271–283, 2020.

[18] L. Calderin, S. Ramakrishnan, A. Puglielli, E. Alon, B. Nikolić, and A. M. Niknejad, "Analysis and Design of Integrated Active Cancellation Transceiver for Frequency Division Duplex Systems," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 8, pp. 2038–2054, 2017.

[19] J. Zhou, P. R. Kinget, and H. Krishnaswamy, "20.6 A blocker-resilient wideband receiver with low-noise active two-point cancellation of >0dBm TX leakage and TX noise in RX band for FDD/Co-existence," in *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 2014, pp. 352–353.

[20] D. Montanari, G. Castellano, E. Kargaran, G. Pini, S. Tijani, D. De Caro, A. G. M. Strollo, D. Manstretta, and R. Castello, "An FDD Wireless Diversity Receiver With Transmitter Leakage Cancellation in Transmit and Receive Bands," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 7, pp. 1945–1959, 2018.

[21] K. Chu, M. Katanbaf, T. Zhang, C. Su, and J. C. Rudell, "A broadband and deep-TX self-interference cancellation technique for full-duplex and frequency-domain-duplex transceiver applications," in *2018 IEEE International Solid - State Circuits Conference - (ISSCC)*, 2018, pp. 170–172.

[22] Y. Lien, E. A. M. Klumperink, B. Tenbroek, J. Strange, and B. Nauta, "Enhanced-Selectivity High-Linearity Low-Noise Mixer-First Receiver With Complex Pole Pair Due to Capacitive Positive Feedback," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 5, pp. 1348–1360, 2018.

[23] E. C. Szoka and A. Molnar, "Circuit Techniques for Enhanced Channel Selectivity in Passive Mixer-First Receivers," in *2018 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2018, pp. 292–295.

[24] P. K. Sharma and N. Nallam, "Breaking the Performance Tradeoffs in N-Path Mixer-First Receivers Using a Second-Order Baseband Noise-Canceling TIA," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 11, pp. 3009–3023, 2020.

[25] S. Krishnamurthy and A. M. Niknejad, "Design and Analysis of Enhanced Mixer-First Receivers Achieving 40-dB/decade RF Selectivity," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 5, pp. 1165–1176, 2020.

[26] Q. Nehal, "Low power mixer-first receiver with second-order baseband filtering TIA," in *2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2020, pp. 117–120.

[27] M. A. Jensen and Y. Rahmat-Samii, "EM interaction of handset antennas and a human in personal communications," *Proceedings of the IEEE*, vol. 83, no. 1, pp. 7–17, 1995.

[28] K. Ogawa and T. Matsuyoshi, "An analysis of the performance of a handset diversity antenna influenced by head, hand, and shoulder effects at 900 MHz .I. Effective gain characteristics," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 3, pp. 830–844, 2001.

[29] C. Picher, J. Anguera, A. Andújar, C. Puente, and S. Kahng, "Analysis of the Human Head Interaction in Handset Antennas with Slotted Ground Planes," *IEEE Antennas and Propagation Magazine*, vol. 54, no. 2, pp. 36–56, 2012.

[30] G. A. Casula, A. Michel, P. Nepa, G. Montisci, and G. Mazzarella, "Robustness of Wearable UHF-Band PIFAs to Human-Body Proximity," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 5, pp. 2050–2055, 2016.

[31] Yichuang Sun and J. K. Fidler, "High-speed automatic antenna tuning units," in *1995 Ninth International Conference on Antennas and Propagation, ICAP '95 (Conf. Publ. No. 407)*, vol. 1, 1995, pp. 218–222 vol.1.

[32] L. Laughlin, C. Zhang, M. A. Beach, K. A. Morris, and J. L. Haine, "Electrical Balance Duplexer Field Trials in High-Speed Rail Scenarios," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 11, pp. 6068–6075, 2017.

[33] T. Vermeulen, B. van Liempd, B. Hershberg, and S. Pollin, "Real-time RF self-interference cancellation for in-band full duplex," in *2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2015, pp. 275–276.

[34] S. H. Abdelhalem, "Integrated Tunable Duplexer in CMOS Technology for Multiband Cellular Transceivers," Ph.D. dissertation, UC San Diego, 2013.

[35] D. R. Jachowski, "Passive enhancement of resonator Q in microwave notch filters," in *2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No.04CH37535)*, vol. 3, 2004, pp. 1315–1318 Vol.3.

[36] I. Hunter, A. Guyette, and R. D. Pollard, "Passive microwave receive filter networks using low-Q resonators," *IEEE Microwave Magazine*, vol. 6, no. 3, pp. 46–53, 2005.

[37] D. R. Jachowski, "Compact, frequency-agile, absorptive bandstop filters," in *IEEE MTT-S International Microwave Symposium Digest, 2005.*, 2005, pp. 4 pp.–.

[38] M. D. Hickle and D. Peroulis, "Theory and Design of Frequency-Tunable Absorptive Bandstop Filters," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 6, pp. 1862–1874, 2018.

[39] B. W. van Liempd, "Tunable RF Front-End Circuits," Ph.D. dissertation, Universiteit Twente, 2017.

[40] J. Friend, C. Harris, and D. Hilberman, "STAR: An active biquadratic filter section," *IEEE Transactions on Circuits and Systems*, vol. 22, no. 2, pp. 115–121, 1975.

[41] C. Andrews and A. C. Molnar, "Implications of Passive Mixer Transparency for Impedance Matching and Noise Figure in Passive Mixer-First Receivers,"

IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 57, no. 12, pp. 3092–3103, 2010.

[42] T. Iizuka and A. A. Abidi, "FET-R-C Circuits: A Unified Treatment—Part II: Extension to Multi-Paths, Noise Figure, and Driving-Point Impedance," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 9, pp. 1337–1348, 2016.

[43] D. Murphy, H. Darabi, A. Abidi, A. A. Hafez, A. Mirzaei, M. Mikhemar, and M. F. Chang, "A Blocker-Tolerant, Noise-Cancelling Receiver Suitable for Wideband Wireless Applications," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 12, pp. 2943–2963, 2012.

[44] C. Luo, P. S. Gudem, and J. F. Buckwalter, "A 0.4–6-GHz 17-dBm B1dB 36-dBm IIP3 Channel-Selecting Low-Noise Amplifier for SAW-Less 3G/4G FDD Diversity Receivers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 4, pp. 1110–1121, 2016.

[45] Y. Lien, E. A. M. Klumperink, B. Tenbroek, J. Strange, and B. Nauta, "High-Linearity Bottom-Plate Mixing Technique With Switch Sharing for N -path Filters/Mixers," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 2, pp. 323–335, 2019.

[46] P. Song and H. Hashemi, "RF Filter Synthesis Based on Passively Coupled N-Path Resonators," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 9, pp. 2475–2486, 2019.

[47] D. Yang, C. Andrews, and A. Molnar, "Optimized Design of N-Phase Passive Mixer-First Receivers in Wideband Operation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 11, pp. 2759–2770, 2015.

[48] H. Khatri, P. S. Gudem, and L. E. Larson, "Distortion in Current Commutating Passive CMOS Downconversion Mixers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 11, pp. 2671–2681, 2009.

[49] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog integrated circuits and signal processing*, vol. 8, no. 1, pp. 83–114, 1995.

[50] Y. Xu and P. R. Kinget, "A Switched-Capacitor RF Front End With Embedded Programmable High-Order Filtering," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 5, pp. 1154–1167, 2016.

[51] Q. Nehal, "Wideband Reconfigurable Blocker Tolerant Receiver for Cognitive Radio Applications," Ph.D. dissertation, UCLA, 2018.