Medical Implant Receiver System

by

Mark Stevens

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2012 by the
Graduate Supervisory Committee:

Sayfe Kiaei, Chair
James Aberle
Bertan Bakkaloglu
Hugh Barnaby

ARIZONA STATE UNIVERSITY

December 2012

ABSTRACT

The medical industry has benefited greatly by electronic integration resulting in the explosive growth of active medical implants. These devices often treat and monitor chronic health conditions and require very minimal power usage. Sometimes this minimal power usage facilitates usage of energy scavenging, while at other times maximizing battery life to tens of years.

A key part of these medical implants is an ultra-low power two way wireless communication system. This enables both control of the implant as well as relay of information collected.

This research has focused on a high performance receiver for medical implant applications. One commonly quoted specification to compare receivers is energy per bit required. This metric is useful, but incomplete in that it ignores Sensitivity level, bit error rate, and immunity to interferers.

In this study exploration of receiver architectures and convergence upon a comprehensive solution is done. This analysis is used to design and build a system for validation. The Direct Conversion Receiver architecture implemented for the MICS standard in 0.18 μm CMOS process consumes approximately 2 mW is competitive with published research.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

INTRODUCTION

Application

Progress in medicine and improvements in technologies has led to an explosion of implantable medical devices to address many debilitating conditions. The earliest application was to treat heart Arrhythmias, and is commonly known as the pacemaker. Since this time there have been many other similar systems developed to treat diabetes, Parkinson's disease, chronic pain, brain and nervous system injuries, and incontinence [1][2]. These implantable systems have even been applied for post-operative monitoring following surgery [3]. All these conditions are treated through a combination of sensory substitution, drug delivery, artificial organs, and neural interfaces.

Medical Implantable systems have four majors parts: sensors/actuators, wireless transceiver, controller, and a power system. The sensors monitor a physical parameter such as visual, chemical content, humidity, pressure, electrical, temperature, or position [1]. Actuators force a physical parameter often through micromechatronics [4][5][6]. Micropumps or electrical stimulation are some common actuator dynamics. A wireless transceiver is needed for communication to an external system. Most often the external system could be issuing commands or receiving data. The controller in the system is the brains of the system and directs the measurement, actions, and communication necessary. It would be a very low power digital signal processor (DSP), microprocessor or

microcontroller. The power system is used for supplying power to enable activity of all other systems. In these systems it is often a battery, but energy scavenging has also been explored.

Because a surgical procedure is required to place the medical implant, these devices have higher standards of scrutiny than consumer electronics products. Important considerations are making the devices reliable and power efficient so that removal is infrequent or unnecessary. Small form factor is critical for a physical feasible system. Low unit cost is required to make the overall solution economical.

Focus of this research

The research team focused on developing a low-power transceiver system for medical implants. The system aimed to allow trading off performance for power efficiency. Three smaller research projects were combined to construct the final transceiver system.

One effort was the reconfigurable front-end that would support multiple receiver and transmit modes. In receive mode this front-end supported both an ultra-low power Super-Regenerative Oscillator for wakeup as well as a high performance Common-Gate cross-coupled Low Noise Amplifier (LNA). The transmit mode enables configuring this same circuitry as either an Injection Locked Power Oscillator, or as a Push-Pull Power Amplifier.

A second effort was directed at construction of a highly programmable All Digital Frequency Locked Loop (ADFLL) subsystem. In the transceiver high performance receive mode of operation the ADFLL was used for generation of the local oscillator for down-conversion. In the transmit modes the ADFLL directly drove the power amplifier structure in the two different modes. The digital injection after the loop filter decouples the modulation bandwidth from the loop bandwidth, while also allowing pulse shaping for low Adjacent Channel Power Ratio (ACPR).

The third effort was design of a low power, high performance receiver. This Direct-Conversion receiver system interfaces to the reconfigurable front-end to use the LNA configuration, and relies on the ADFLL for quadrature mixing. This high performance receiver is capable of reliable low Bit Error Rates (BER), moderate data-rates, in the presence of interferers.

Most of this dissertation focuses on the Direct-Conversion Receiver System, my contribution to the transceiver system. A clearer picture of how this fits into the whole system is given in the remainder of the introduction.

The MICS transceiver IC that addresses the very important ultra low power consumption requirements is now described. This power management is accomplished by providing multiple systems that tradeoff power needed for required performance. On the receive side, power consumed versus bit error rate performance is traded off through wakeup and mission modes. The transmit side

provides savings through selection of either a traditional push-pull amplifier, or an injection locked mode, depending on data-rate. The cost for these options is minimized through the introduced reconfigurable front-end, providing significant reuse of devices. The LO frequency synthesizer has been optimized for the MICS band transceivers, being implemented as an ADFLL. When the frequency synthesizer is used as a modulator for the power amplifier, the bandwidth is made to be independent of the loop filter bandwidth.

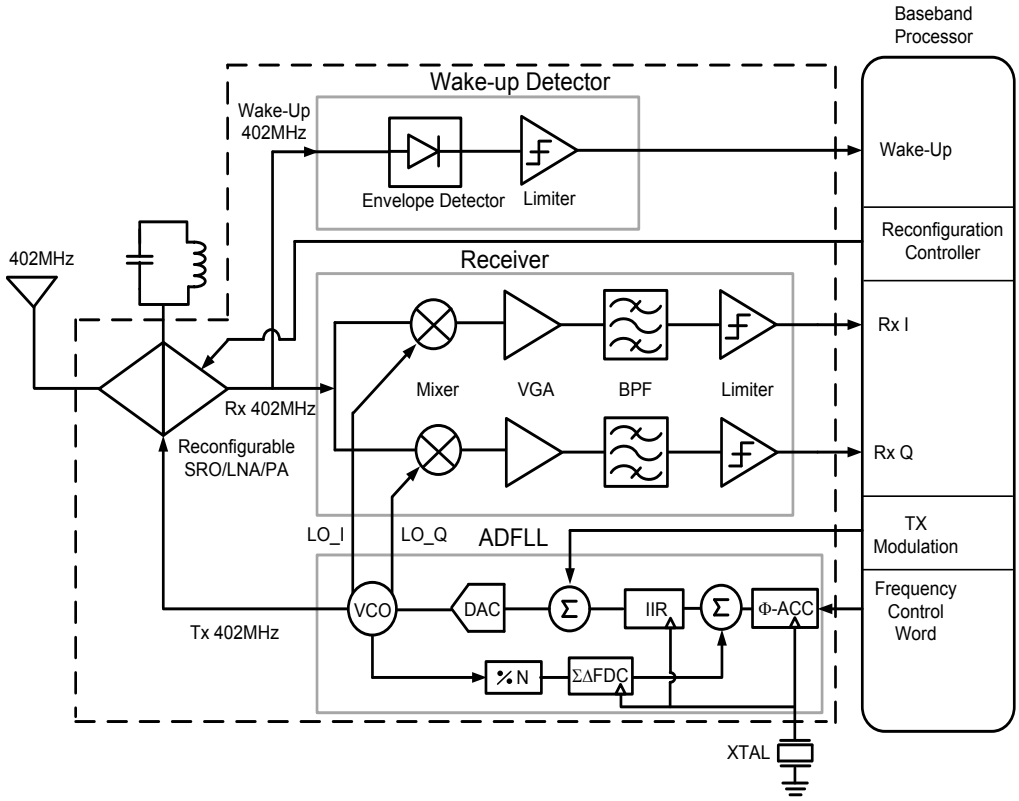


Figure 1 Overall transceiver System

Prior work

On the receiver side, one approach tried is Super-Regenerative architecture using On-Off Keying (OOK) modulation, because of the inherent low component count, enabling a low power operation. An OOK implementation was able to achieve -99dBm sensitivity with only 400 uW of power [9]. Other examples also show the efficiency of the Super-Regenerative architecture [10][11]. A problem identified in such systems is the need for a high Q filtering for acceptable performance in terms of sensitivity and selectivity [12]. An external tank having high Q implies higher cost components, in addition to required tunability for channel selection and manufacturability. Elimination of high Q components providing better integration is possible through a Q-enhanced filter topology, but it comes at the cost of extra power consumption, tunability circuitry, and initialization overhead [13]. The conflicting demands of low power and selectivity probably explains why Super-Regenerative receivers are not always applied in MICS systems.

Frequency Shift Keying (FSK) implementations using Direct Conversion or low-IF systems has also been tried [14] [15][16]. Both methods have the advantage of much lower IF-frequency, easing the channel-select filter requirements. Direct conversion is used with a high modulation index, attempting to minimize flicker noise impact [14]. Low-IF systems use image-reject poly-phase filtering providing channel selection [16] [17]. Elimination of the flicker noise comes at the cost of filter complexity, consuming more power and silicon die area.

A dual-receiver system supporting both Super-Regenerative OOK and BFSK has been also reported in [18]. The super-regenerative architecture is used as a wakeup receiver, while the quadrature mixing architecture is used for a high performance mission mode BFSK receiver.

Because MICS transmitters adopt simpler modulation schemes like OOK or BFSK, quadrature mixing for up-conversion is not required. Use of direct modulation through fractional-N synthesizers saves power, and has been implemented by many [9][14][16]. MICS transmitters adopt FSK modulation with high modulation indices, which allow low power implementations and minimize flicker noise impact.

RF band/MICS Specifications

A family of low power wireless standards(IEEE 802.15) exist to address the personal area network space (PAN) [19]. It would seem that these low power small distance standards would be ideal. More careful examination of these specs helps clarify why these are not best for medical implant systems.

First, 802.15 standards use the unlicensed Industrial Scientific Medical bands (890-950 MHz, 2.4 Ghz – 2.41 Ghz) frequencies, which are also heavily used by

other longer range, higher power systems with the WiFi 802.11 being the prime contender. 802.15 protocols avoid interference and achieve effective communication in these bands through using spread spectrum techniques. The many variants of the well known Bluetooth standard 802.15.1 uses frequency hopped spread spectrum (FHSS). Similarly, the Zigbee standard 802.15.4 uses direct sequence spread spectrum (DSSS). The added design complexity of both would translate to higher power consumption. A less complex, narrowband modulation technique should give lower consumption.

Second, lower attenuation through human tissue can be achieved at lower frequencies. This directly benefits the link budget as well.

Third, a lower frequency means that less power should be needed in the RF circuits.

In order to support medical implants more effectively a separate MICS band was approved in 1999, and is fully described in the reference [7].

A brief summary of the specifications is shown below

RF band	402-405 Mhz
Channel Bandwidth	300 KHz
Channels	10
Antenna power (EIRP)	-16 dBm or 25 μW
Operating range	0-2 m

Figure 2 MICS specifications

A fortunate advantage within the specification is the lack of many constraints. These allow much more freedom in the implementation.

For example, on the receiver side, there is no explicitly required sensitivity level, dynamic range, modulation type, interferer rejection, etc. And, on the transmitter side, there is no specific transmission mask requirement.

In the interest of designing the best transceiver, the standard concerns should nevertheless be addressed. The receiver should decode the weakest of signals reliably without being affected by adjacent channels and/or interferers. In addition the largest dynamic range should be supported through a high linearity receive chain. The transmitter should also minimally direct power towards non-intended channels.

RECEIVER ARCHITECTURE

The chosen architecture is selected after considering the several down-conversion IF alternatives, and other receiver system options. By carefully examining the many possibilities in wireless systems and weighing the benefits and downsides, the right tradeoffs can be made for the implantable wireless system.

Down-conversion IF

The first method used in radio systems, and still the most popular, is Super-Heterodyne.

Unfortunately one downside of this approach is mixing to the IF frequency can corrupt the channel through superimposing the image, as seen in Fig 3.

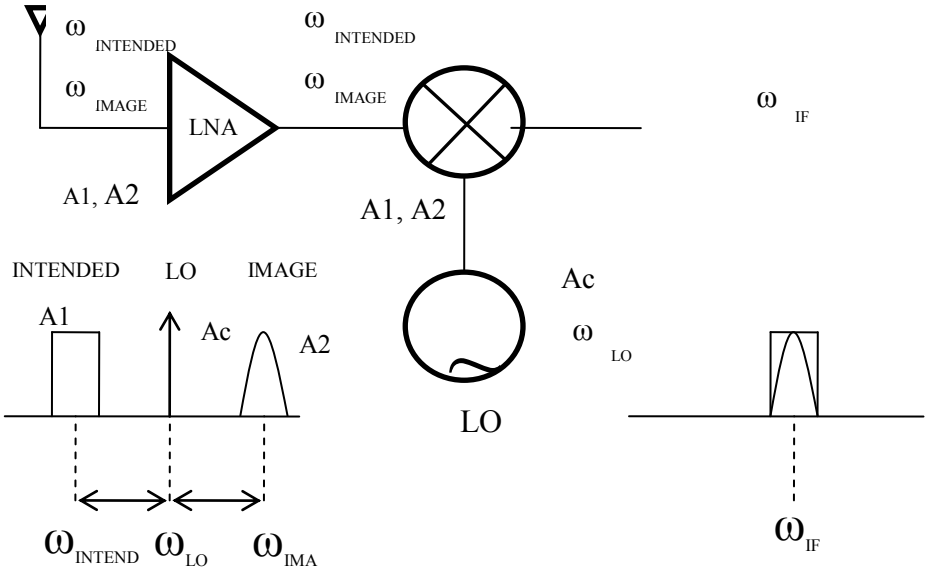


Figure 3 Corruption during down-conversion due to image

The mixing operation produces the following down-converted components

$$A_C\cos(W_{LO}t)\,[A_1\cos(W_{INTENDED}t+\emptyset_1)+A_2\cos(W_{IMAGE}t+\emptyset_2)] \rightarrow$$
$$\frac{1}{2}A_C[A_1\cos(|W_{INTENDED}-W_{LO}|t+\emptyset_1)+A_2\cos(|W_{IMAGE}-W_{LO}|t+\emptyset_2)] \quad (1)$$

If $|W_{INTENDED}-W_{LO}| = |W_{IMAGE}-W_{LO}| = W_D$ then

$$\frac{1}{2}A_C[A_1\cos(W_Dt+\emptyset_1)+A_2\cos(W_Dt+\emptyset_2)] \quad (2)$$

The resulting SNR (not including system noise) is

$$SNR = 20*log_{10}\left(\frac{A_1}{A_2}\right) \quad (3)$$

Use of a pre-selection filter and maximizing the attenuation of the image through a larger frequency span leads to the high IF approach. Fig 4 shows the intended frequency as tuned, with a low-IF of 2MHz (shown as the dashed lines), and with a high-IF of 40MHz (with the dotted lines). The RF tuned resonator has a Q of 80, giving only 5.5 dB attenuation for the low-IF case and 29.5 dB attenuation for the high-IF scenario.

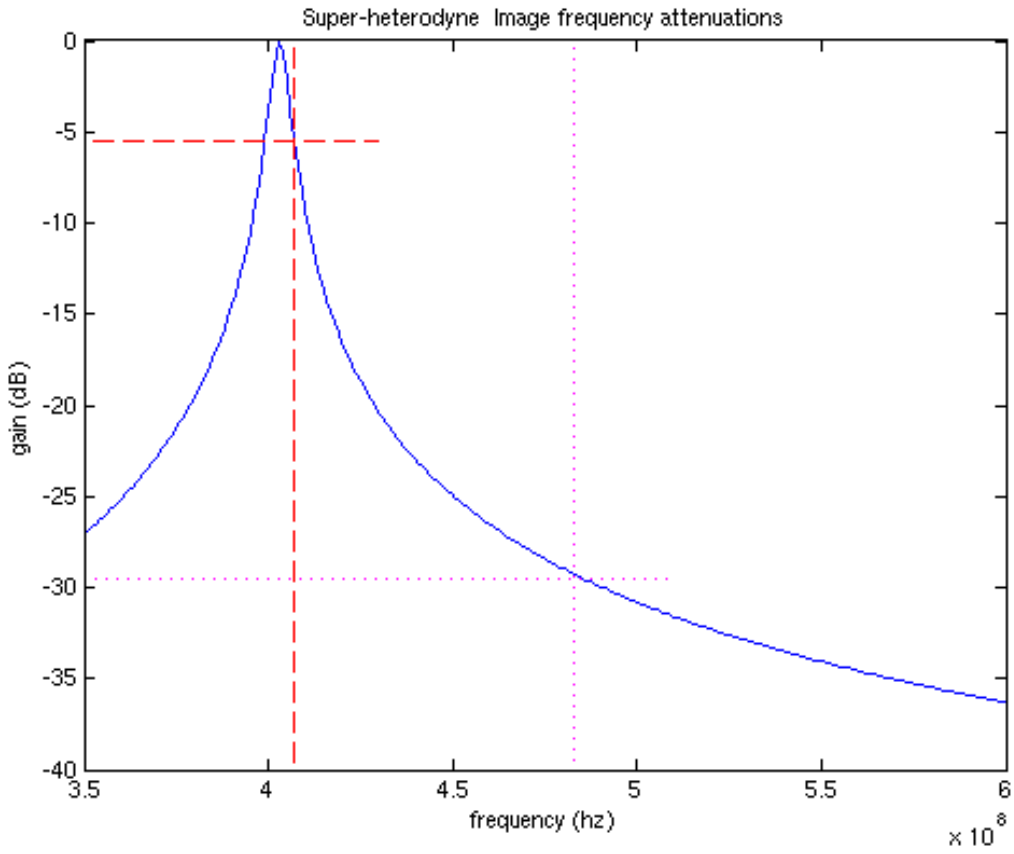
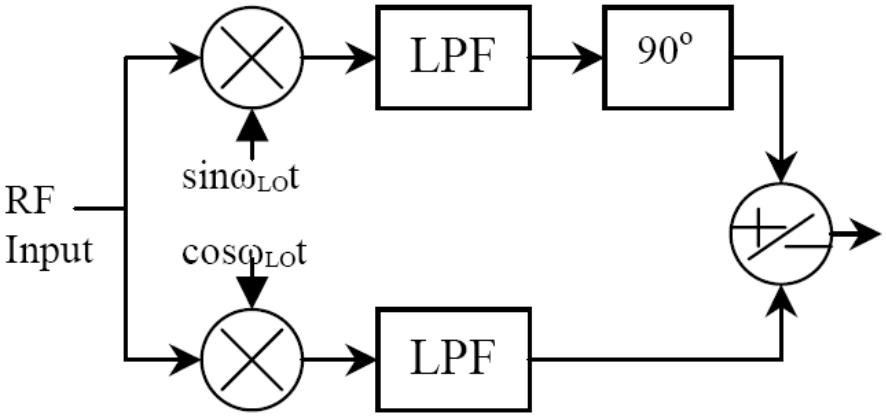


Figure 4 Image frequency attenuations

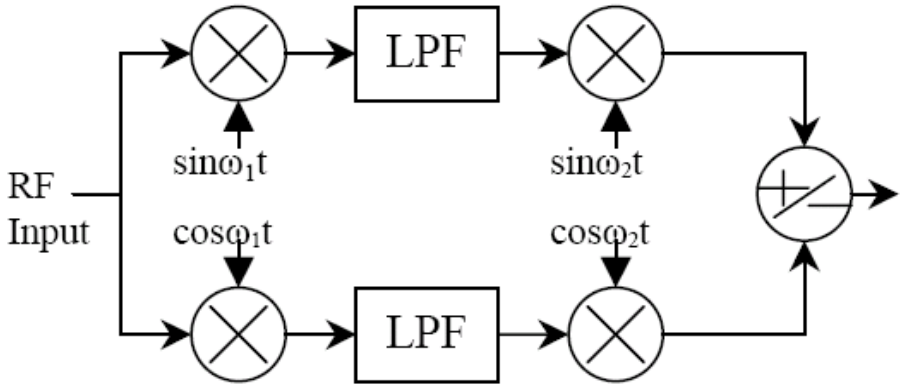
While a higher IF stage addresses the image issue, a higher power is needed for signal processing. Multiple such down-conversions would be needed in a complete receiver system.

More sophisticated and effective image rejection methods have also been used such as the Hartley architecture and the more contemporary Weaver architecture as shown in Fig 5 [20][21]. The cancellation achieved through Hartley and

Weaver architectures is limited through amplitude and phase mismatches. Literature has shown this to be limited to 30-35dB practically (through careful layout and matching techniques).



a)



b)

Figure 5 IMAGE REJECTION a) Hartley Architecture b) Weaver Architectures

The higher rejection enables less pre-filtering to be used to get the same image level. The LO can be moved much closer to the RF band, and a much lower IF can be generated without corruption from the image.

An even more recent approach to achieve image cancellation is through use of polyphase filters [22][23]. These filters operate on the down-converted quadrature components of the signal, and are able to distinguish between the image and intended RF bands. They have the advantage of only requiring one mixing stage, but have the drawbacks that these filters are complicated to design, and consume considerable area and power to operate.

These image cancellation techniques make low-IF a much more viable approach.

In summary, higher IF has bigger image rejection, but would require more power for processing at this higher baseband frequency. In contrast a lower IF would require less power for processing in baseband, but perhaps at the cost of more power to provide the image rejection.

A brute force approach of multiple stages to let the least image into the down-converted signal would lead to a higher performance, but also a much higher power system.

Direct Conversion, which had not used much in early radio systems, has been increasingly used in the past 30 years as device integration has progressed. The elimination of the image rejection problem is solved through this technique, but unfortunately is replaced by other problems. These new problems are Flicker noise, 2nd order distortion, and DC offset [24][25].

Second order distortion is a concern in the frequency translation step (mixers) and does not influence system behavior in high IF, very rarely in low IF, but is quite common with direct conversion. The equation below explains

$$x^2(t) = \alpha_2[\mathrm{A}_1\cos(w_1 t) + \mathrm{A}_2\cos(w_2 t)]^2$$
$$= \alpha_2[\mathrm{A}_1\mathrm{A}_2\cos(\{w_2 - w_1\}t) + \mathrm{A}_1\mathrm{A}_2\cos(\{w_2 + w_1\}t)]$$
$$\rightarrow \alpha_2[\mathrm{A}_1\mathrm{A}_2\cos(\{w_2 - w_1\}t)] \quad (4)$$

The two in-band RF frequency components can generate the new distortion term in the baseband, as also shown in Fig 6.

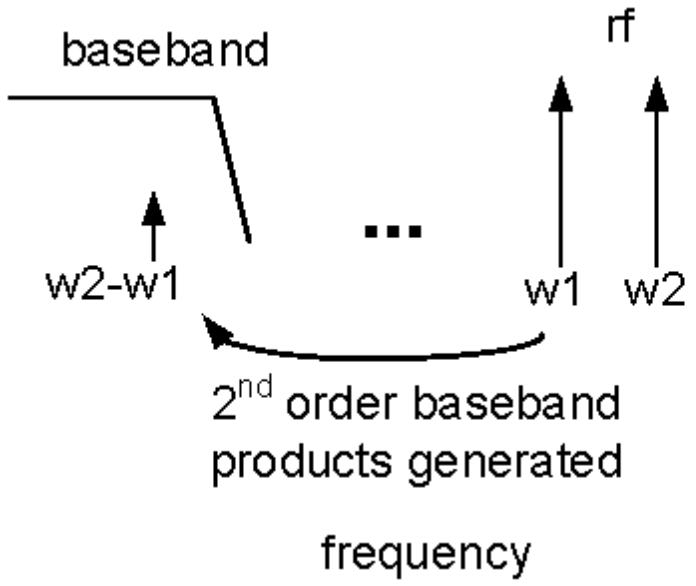


Figure 6 Illustration of 2nd order distortion

The traditional solution to this problem has been designing circuitry with very high IIP2s. More recently an approach tried is adaptive DSP processing to mitigate this non-linear effect [26].

A DC component is also caused through self mixing of LO components. The large signal strength of the LO mandates an extremely high level of isolation in order to protect the subsequent design stages from saturating. When $w_1 = w_2 = w$ the last equation translates to a DC term.

$$\begin{aligned} x^2(t) &= \alpha_2[\cos(wt) + \cos(wt)]^2 \\ &= \alpha_2[\cos(\{0\}t) + \cos(2wt)] \\ &\rightarrow \alpha_2[1] \end{aligned} \tag{5}$$

Flicker noise is caused by the transistors at low frequencies, and is governed by the equation:

$$\overline{i_n^2} = \frac{K}{f}\left(\frac{g_m^2}{WLC_{OX}^2}\right) \quad (6)$$

This effect is of primary concern of baseband subsystems, operating on the down-converted signals.

Two design choices can have a big impact on Flicker noise. First, gate design area can be increased greatly because the devices will be operating at low frequency. The added capacitances of these much larger devices causes very little problem with the frequency response characteristics. Second, low biasing currents, which result in low g_m, have the advantage of reducing the Flicker noise.

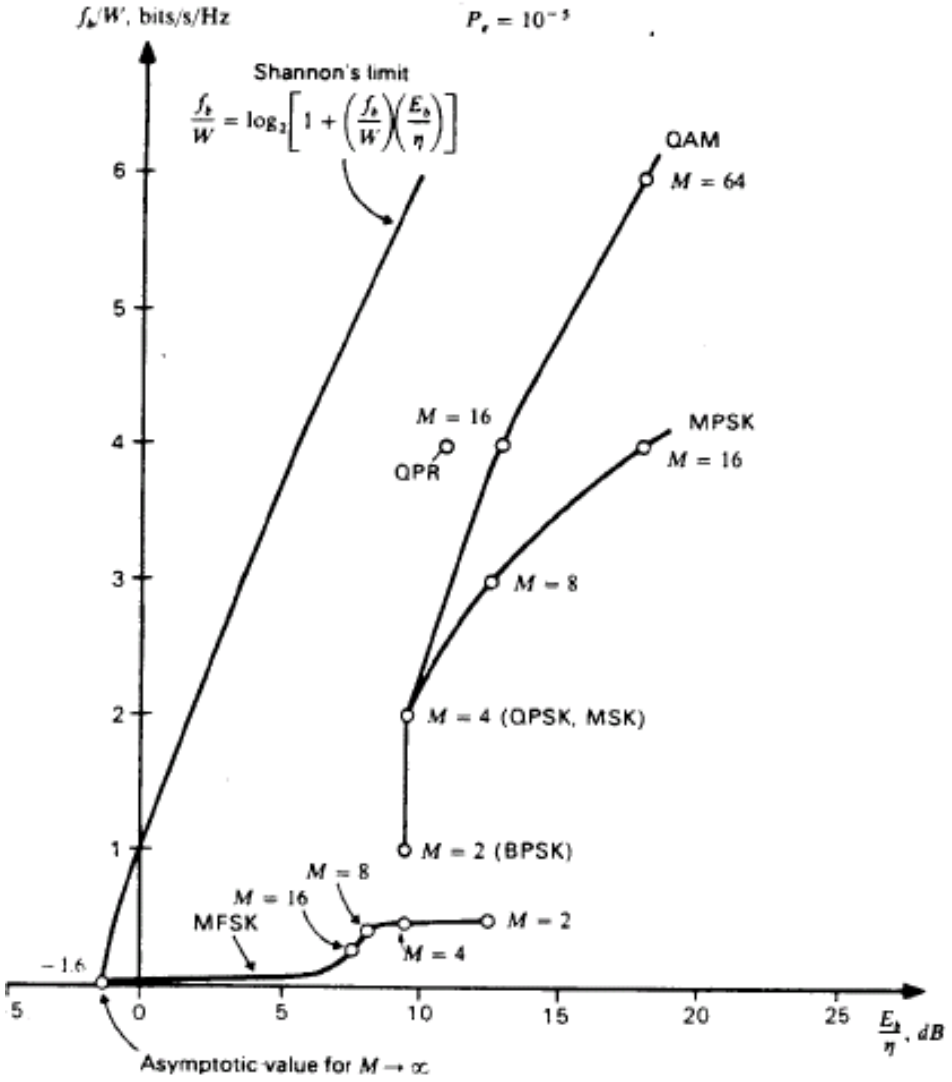
A solution to address both the Flicker noise issue and DC offset is AC coupling in the baseband stages. This would highly attenuate the lower frequencies, but could have impacts to the BER [24][25].

One clear advantage of a Direct Conversion system is it is always just a single stage. This makes the system smaller, consuming less power.

Signaling/modulation choices

Examination of bandwidth efficiency usage of the different modulation methods is done to understand how to get the most efficient transfer of information. Fig 7

shows that the highest bandwidth efficiencies are methods using multiple amplitude levels or phases to represent the data transferred.



bits/s/Hz vs. E_b/η for Probability of Error = 10^{-5}
taken from "Principle of Communication Systems"
Taub & Schilling, page 482

Figure 7 Data packing density per BW

Because the system is wireless and a huge unknown dynamic range exists this precludes use of an amplitude modulated method. Multiple phase information is possible, but comes at the cost of carrier synchronization so that sampling of the digital waveform is at the right time. Subsystems like Multiply-filter-and-divide or Costas Loop techniques as shown in Fig 8 can perform this function, but it is clear these add additional system complexity [27].

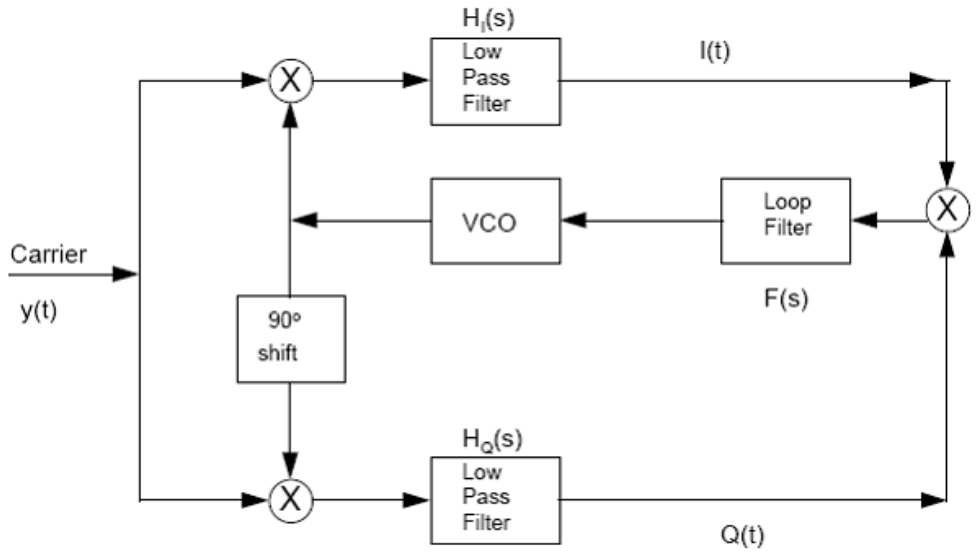


Figure 8 Carrier Synchronization Costas Loop (for BPSK)

If the coherency requirement is relaxed, the receiver system overhead could be reduced. The resulting modulations available with non-coherent methods are Frequency shift keying (FSK).

The specific frequency shift keying method chosen depends on a couple factors. First, looking at Fig 7 we can see that BFSK should have the highest bandwidth

efficiency of this family of modulations. This is due to less frequency spectrum wasted between each symbol frequency, as the FSK index increases. A second factor more related to the Direct Conversion Receiver (DCR) is using signal schemes that avoid putting the signal content around DC, so as to avoid both DC offsets and Flicker noise degradation. When the center frequency at baseband is DC, BFSK concentrates the symbol power at the frequency deviation offset. Although minimum-shift keying (MSK), a form of BFSK, maximizes the data-rate achievable for maintaining orthogonality an insignificant amount of power close to DC results. In contrast, a higher modulation index trades off bandwidth efficiency for immunity against DCR architecture weaknesses. This approach was first used in the early 1990s for pager designs [28].

Dynamic Range management

Many receiver systems have variable gain within the receive chain in the form of an automatic gain control amplifier (AGC). It adjusts the gain of the system based upon the input signal strength in order to expand the range of input powers over which the amplification is in the linear mode. This varying signal strength could be through natural attenuation or through fading effects.

These benefits come with the price of a complex system that consumes power and area, along with introducing noise and possibly have stability issues. Often the dynamic range is expanded by reducing the gain, to allow larger signal powers. If

higher dynamic range is not important or the costs are too high, this subsystem should not be included.

Multipath Fading

Because the system is wireless, multipath effects need to be considered.

The usage of these systems will usually have a relatively stationary patient (wearing the implant) and a stationary base station, removing the need to consider Doppler effects. Nevertheless, a changing environment around the transmitter/receiver could dynamically change the multipaths.

The signal range has a maximum diameter of 10m, with a resulting very small maximum time dispersion approximated as (D=20m / $3x10^8$ m/s) = 66.7 ns. This would result in a coherence bandwidth of 13 Mhz, which for the MICS RF of 402-405Mhz, would result in a flat fading scenario. In addition, because these systems will generally have a line of site path for transmission, Ricean fading model is most applicable [29].

Although mitigation techniques to address the multi-path dead zones are possible the practical solution is to either slightly move the patient or the base station to remedy the weakly received signal. Because we have a flat fading scenario the only alternative is adding spatial diversity (through multiple antennas). This would not only make the implant larger through the additional antenna, but also require an additional receiver front-end to process this second signal source.

RX system architecture & design specifications

Using the prior arguments in system design, a Direct Conversion Receiver architecture using non-coherent BFSK signaling is chosen. The receiver chain uses a quadrature mixer, followed by gain stages and filtering as shown. The VGA is for overall gain tuning, and not controlled through an automatic gain control system.

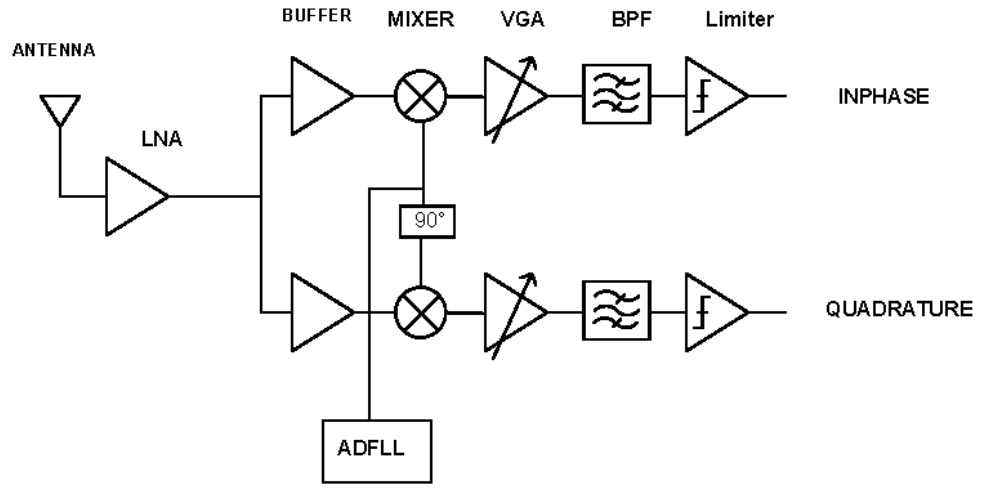


Figure 9 I/Q RX system

The baseband which is centered at DC will have positive and negative frequencies. In order to distinguish between these 100 KHz signals, quadrature mixing is used. The phase relationships between these quadrature mixed signals can then be used to indicate whether the positive or negative frequency had been received. This method of detection is called non-coherent symbol detection, because no time reference or phase locking information is needed in detection. As

mentioned earlier this has the advantage of making the system less complex and lowering overall power utilized.

A mathematical derivation of baseband decoding follows. The RF signal can be represented as a cosine signal with arbitrary phase.

$$\cos(\omega_{RF}t + \theta) \tag{7}$$

Quadrature mixing produces the following down-converted outputs

INPHASE:

$$A_{RF}cos(\omega_{RF}t + \theta)cos(\omega_{LO}t + \emptyset) \rightarrow \frac{1}{2}A_{RF}cos([\omega_{RF} - \omega_{LO}]t + \theta - \emptyset) \tag{8}$$

QUADRATURE:

$$A_{RF}cos(\omega_{RF}t + \theta)sin(\omega_{LO}t + \emptyset) \rightarrow -\frac{1}{2}A_{RF}sin([\omega_{RF} - \omega_{LO}]t + \theta - \emptyset) \tag{9}$$

When $\omega_{RF} > \omega_{LO}$ then the quadrature signal leads the inphase. In contrast when $\omega_{RF} < \omega_{LO}$ the quadrature signal lags the inphase. A "1" can be assigned to the case where $\omega_{RF} < \omega_{LO}$ and "0" otherwise.

Amplification of the baseband signal and limiting can produce digital signals on which this phase relationship can also be analyzed. The Inphase and Quadrature edge relationships easily give the detected state. A continuous time digital stream example along with detected value is shown in Fig 10.

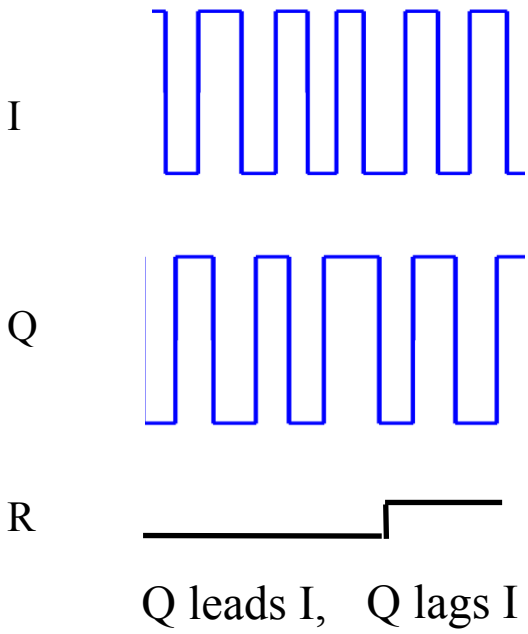


Figure 10 Inphase/Quadrature receive decoding

Link budget analysis

An understanding of the received signal strength is found in looking at the Link Budget Analysis in Fig 11 [27][32]. This pessimistic estimate gives the weakest signal powers received, although one or more of the path loss components could be less resulting in a stronger received signal.

Component	Path loss
Transmitter	-16 dBm
Free space loss (2m)	30 dB
Attenuation through human body to receiver antenna.	20 dB
Poor implant antenna gain	20 dB
Signal power	-86 dBm

Figure 11 Link Budget table

An extra margin of 10 dB is allotted for variation with a target sensitivity of -96 dBm. The system sensitivity is given by

$$P_s = -174\frac{dBm}{hz} + NF_{RX} + 10logB + SNR_{RX} \quad (10)$$

where SNR_{RX}=10.7 dB is the minimum signal to noise ratio required at the demodulator input for a BER of 1E-3 for non-coherent BFSK modulation at 75 Kbps. The channel bandwidth, as narrowed by the BPF, is B=80 KHz. This requires a noise figure no greater than 17.7 dB.

The link budget analysis aligns with findings in published literature [30].

NF

Noise figure for a cascaded system is found through the Friis equation [31]. The noise factor (F) and power gains (g) are unitless (not in dB).

$$F_{RX} = F_1 + \sum_{i=2}^{n} \frac{F_i - 1}{\prod_{j=1}^{i-1} g_j} \quad (11)$$

The equation above assumes conjugate matching, but it is possible a system that is constructed may have different input and output impedances. A mismatch of impedances can affect the noise figure calculation, and is handled as described below [32]. All numbers are still unit-less, but now we are using a voltage gain

rather than a power gain, and the resistive output and input impedances are included.

$$F_{t\_cascade} = F_1 + \sum_{m=2}^{n_1} \frac{F_m - 1}{\prod_{l=1}^{m-1} g_{v,i}^2 \left(\frac{R_{i,l}}{R_{o,l-1}+R_{i,l}}\right)^2 \left(\frac{R_{o,l-1}}{R_{o,l}}\right)} \quad (12)$$

In integrated chip designs many stages of the baseband and even some of the RF stages may not be conjugately matched, so this equation is very useful.

In the case of this design after the antenna to LNA interface all subsequent stages can use this equation, and with a further simplification. Because the input resistances are MOSFET gates which are MΩ and the output resistances driving these is on the order of 10s of KΩ, the following simplification can be used.

$$F_{t\_cascade} = F_{ANT} + \frac{F_{LNA} - 1}{g_{v,ANT}^2 \left(\frac{R_{i,LNA}}{R_{o,ANT}+R_{i,LNA}}\right)^2 \left(\frac{R_{o,ANT}}{R_{o,LNA}}\right)} + \sum_{m=3}^{n_1} \frac{F_m - 1}{\prod_{l=1}^{m-1} g_{v,i}^2 \left(\frac{R_{o,l-1}}{R_{o,l}}\right)} \quad (13)$$

Linearity (IIP2, IIP3)

Non-linearities in the system will cause out of band frequencies to be put in-band, corrupting the intended signal. These weak non-linearities are typically described through a Taylor series, which models it as a polynomial as shown below

$$y = \alpha_1 x^1(t) + y = \alpha_2 x^2(t) + y = \alpha_3 x^3(t) + \cdots \quad (14)$$

The two linearities of concern in a direct conversion receiver are 2nd and 3rd order. 2nd order had been described previously, and is a problem that can be seen during the mixing operation. 3rd order problems occur when there is no frequency translation, as occurs in most linear components. Expanding the third order term we see

$$x^3(t) = \alpha_3[\mathrm{A}_1\cos(w_1 t) + \mathrm{A}_2\cos(w_2 t)]^3 =$$

$$\alpha_3[\mathrm{A}_1\mathrm{A}_2\cos(\{2w_2 - w_1\}t) + \mathrm{A}_1\mathrm{A}_2\cos(\{2w_1 - w_2\}t)] \quad (15)$$

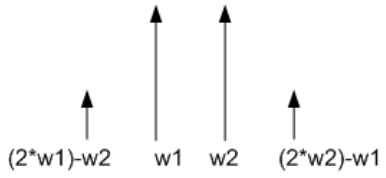


Figure 12 IIP3 illustrated

An example using the case of LNA shows how these non-linearities can introduce frequencies polluting the spectra. Channel 1 contains the intended signal, while channels 2 and 3 have strong interfering activity. This causes a spur to land in channel 1 due to 3rd order non-linearity.

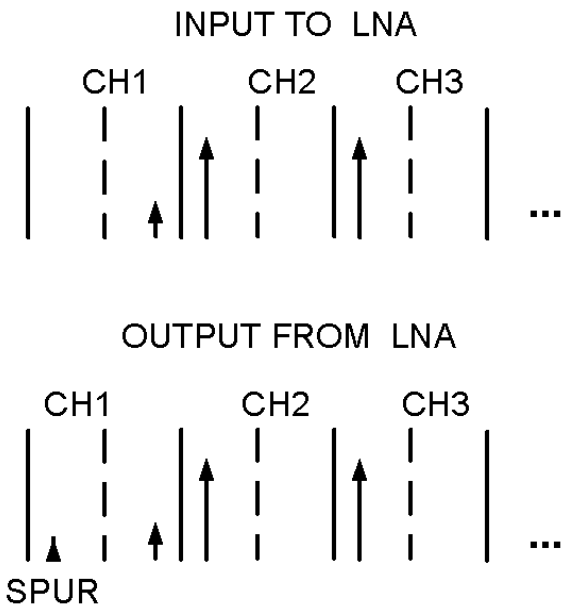


Figure 13 LNA 3rd order distortion

These non-linearities are especially damaging in the RF band because this is a relatively broadband spectra. Even with an external high Q tank of 100, the bandwidth that can interact to cause spurs is 4 Mhz for the MICS band ~400 Mhz

.

The estimation often used is below [32].

$$\frac{1}{IIP3^2_{TOTAL}} = \sum_{i=1}^{n} \frac{\prod_{j=1}^{i} G_j^2}{IIP3^2_{n-i}} \tag{16}$$

Circuit topology and resulting nonlinear regions will dictate most of the IIP3 and IIP2 limits. For IIP2, mismatch will lead to higher 2nd order distortion.

LO Phase Noise

Variation of LO frequency will cause existing channel(s) to blur their data in the frequency bands. Corruption can occur in the mixing process because the LO phase noise will cause other channels will be mixed into the intended channels. This process is called reciprocal mixing.

For example for the MICS technology where channel bandwidth is 300 KHz the phase noise offset at that value will be the superimposed adjacent channel interference magnitude.

The high modulation index used with BFSK to solve the flicker noise issue for Direct Conversion causes tighter constraints for phase noise requirements. Fig 14 shows how the unintended channel data can get placed in the baseband spectrum due to this phenomenon.

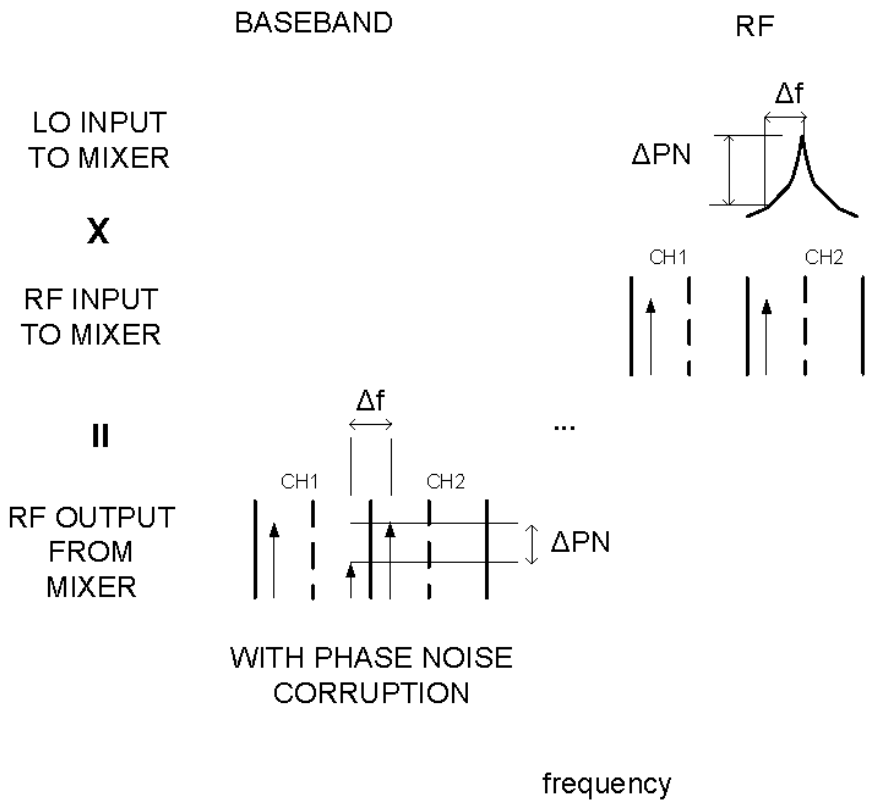


Figure 14 Phase Noise Corruption

The mathematical description is shown in the next equation [33]

$$L(f_m) = C(dBm) - S(dB) - I(dBm) - 10\log(B) \quad \text{(dBc/Hz)} \qquad (17)$$

Minimum Detectable Signal

The minimum detectable signal will be limited by two factors: Sensitivity and phase noise corruption. The sensitivity is solely based upon the noise floor at each stage in the system. The phase noise will cause adjacent channels to be mixed into the desired channel, during downconversion. If this phase noise is too high, it will limit the system performance. We would like the phase noise contribution to be just under the noise floor.

Dynamic Range and options

Dynamic range is limited by two factors

1) Linearity at high signal amplitudes (2^{nd} and 3^{rd} order distortions)
2) Minimum Detectable signals at low amplitudes

The dynamic range is the minimum of 2^{nd} or 3^{rd} order dynamic range. It really depends upon which dominates the system.

The common relation considering 3^{rd} order distortions is shown below [33].

$$DR_3 = \frac{2}{3}(OIP3 - N_0) \tag{18}$$

It is also rewritten for the 2^{nd} order distortion case.

$$DR_2 = \frac{1}{2}(OIP2 - N_0) \tag{19}$$

Immunity to interferers

Interferer/blocker immunity at RF typically is measured for a continuous wave tone, rather than as an actual traffic pattern. This single tone can cause system failure because of desensitization of the input stage, causing the much weaker intended signal to have a much reduced gain as described in the equation below

$$y(t) = \propto_1 A_D \left[1 + \frac{3\propto_3}{2\propto_1} A_D^2\right] cos(2\pi f_0 t) + \cdots \quad (20)$$

A second way system failure happens is if the interferer strength is such that the possible reduced front-end gains plus the rejection in the filtering is not enough to attenuate this signal to a level where the intended detectable signal SNR is acceptable.

Finally, reciprocal mixing of the interferer can rise above the noise floor causing signal quality degradation.

Breaking down the subsystem design

Knowing the overall specifications, we now need to define the cascade of components that will achieve that requirement. Initial design requires keeping several facts in mind. We need the highest gain and lowest noise in the first stages, with higher linearity in later stages (if possible).

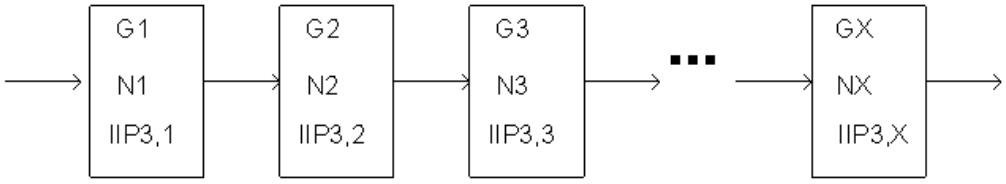


Figure 15 Noise Figure & Linearity of components

The component specifications used depend upon several factors: circuit topologies, technology for implementation, and power consumption. Several iterations were needed to arrive at acceptable specifications that were consistent between the high level architectural needs and the circuit level reality, as shown in Fig 16.

	POWER GAIN (dB)	VOLTAGE GAIN (dB)	Rin (Ω)	Rout (Ω)	NF(dB)	IIP3
LNA	16.5	26.5	200	2000	3	-26 dBm
MIXER	-5	-5	2000	2000	12	+11 dBm
PREAMP	-1	6	MOSFET GATE, Very high impedance >> Rout of previous stage	10000	14	-21 dBm
BPF	-10	10	MOSFET GATE, Very high impedance >> Rout of previous stage	100000	28	-47 dBm
LIMITER	58	38	MOSFET GATE, Very high impedance >> Rout of previous stage	10000	14	-33 dBm
OVERALL	58.5	75.5			17.7	-28.4 dBm

Figure 16 Circuit Block specifications

The goal of NF=17.7 dB was achieved with some margin, while -28.4 dBm IIP3 linearity is possible. Voltage gain is reported in the table above because after the

LNA most of the signal strength tracking is in the voltage domain. In order to be detected with some margin in the limiter hysteresis circuit we need the voltage wave to have an amplitude above 0.05V, requiring a total of 75.5 dB of system voltage gain.

The achievable sensitivity is reduced because the power gain is not so great, as can be seen in the Fig 16. This was a conscious choice in the design to reduce the critical overall power usage. With the much less available power in the baseband circuits this is an inevitable consequence.

The levels of noise and signal levels can be seen in Fig 17. Validation of signal and noise level growth as the received signal progresses through the receive chain can be seen. This information is used in planning for linearity of later stages as well as allocation of gain among subsystems. This specifically helps in the design of the BPF and limiter hysteresis blocks.

SUBSYSTEM		LNA	MIXER	PREAMP	BPF	LIMITER
NOISE (spot noise density)	dBm	-154.5	-158.8	-159.4	-155.8	-97.8
	Noise Density	7.10E-16 V^2/hz	2.28E-16 V^2/hz	1.29E-15 V2/hz	2.64E-13 V^2/hz	1.67E-9 V^2/hz
Minimum Sensitivity -96 dBm	dBm	-79.5	-84.5	-85.5	-95.5	-37.5
	V_{PEAK}	212 µV	119 µV	238 µV	752 µV	59.8 mV
+20 dB blocker -76 dBm	dBm	-59.5	-64.5	-65.5	-75.5	-17.5
	V_{PEAK}	2.12 mV	1.19 mV	2.38 mV	7.52 mV	598 mV
+30 dB blocker -66 dBm	dBm	-49.5	-54.5	-55.5	-65.5	-7.5
	V_{PEAK}	6.70 mV	3.77 mV	7.52 mV	23.8 mV	≥ 700 mV
+40 dB blocker -56 dBm	dBm	-39.5	-44.5	-45.5	-55.5	+2.5
	V_{PEAK}	21.2 mV	11.9 mV	23.8 mV	75.2 mV	≥ 700 mV

Figure 17 Subsystem signal and noise levels

In order for the BPF to reject interferers as designed, the signal in the filter path needs to stay within the linear range, and far enough away from the non-linear clipping limits to operate effectively. The filter has been designed for roughly 35 dB rejection for out-of-channel interferers. The linear range of the filter operational transconductance amplifiers (OTAs) needs to be comparable to this filter range. The gain budgeting is done in such a way that a low noise, and low amplitude signal is presented to the bandpass filter. The LNA through PREAMP stages provide enough gain with a low noise representation such that the very noisy Gm-C implementation does not impact the noise figure of the system significantly, while still supporting up to 35 dB stronger interferers than the noise floor. Fig 17 shows the interferer amplitudes at different powers, along with the corresponding voltage amplitude as seen internal to the Gm-C filter stages.

The limiter design is determined by the signal level and amplified noise level presented to the hysteresis comparator. The signal amplitude level must be great enough to exceed threshold levels, and the noise level must be below that level. Many hysteresis comparators operate effectively when the threshold range is between 5-15% of the supply range. Keeping this in mind the voltage gain of the limiter provides the gain needed for the weakest signal strength supported of -96 dBm. Five percent of the 1.8V supply gives a range of 90mV, that is, ± 45mV about the amplitude. The RMS voltage can be found for the 80 KHz BPF bandwidth with a constant spot voltage of -97.8 dBm + $10\log_{10}(80000)$ = -48.8 dBm, as given by the equation below.

$$V_{RMS} = \sqrt{\left[10^{\left(\frac{P_{AdBm}-30}{10}\right)}\right] R_{OUT,LIMITER}} \quad (21)$$

We see that the -96dBm signal has an amplitude of roughly 60mV, while the noise level has an RMS voltage of 11.5 mV.

In order to account for additional flicker noise selection of greater hysteresis values have been provided of ± 75 mV and ± 100mV.

The average current budget for the subblocks is seen in Fig 18.

Quadrature RX subblocks	Current
LNA	550 µA
Mixer	310 µA
Pre-amp	60 µA
Bandpass Filter	80 µA
Limiter	80 µA
TOTAL	1080uA

Figure 18 Subblock power breakdown

Performance

We define performance as a combination of data-rate achievable and resulting reliability. Data rate is in bit-per-second (bps), while reliability is measured in terms of bit-error-rate (BER). The reliability is for raw data-rate, not encoded data.

Theoretical BER predictions

The data-rate is limited by the spacing of the frequency deviation by the bit datarate. According to communication theory the upper bound is 1/Ts for non-coherent BFSK. Because we do not want much of the baseband signal near DC or as affected by Flicker noise the limit is not pushed to the theoretical maximum. Instead the baseband bandpass filter bandwidth is used to set the data-rate limit. We want this bandwidth large enough to support the desired data-rate without letting too much excess noise. The filter bandwidth is selected at 80 KHz, and the aim is to get as close to that limit as possible, while achieving good bit-error-rates.

From communications theory we expect a BER for non-coherent BFSK with Additive White Gaussian Noise (AWGN) as follows [27].

$$P_b = \frac{1}{2} e^{-\frac{E_b}{2N_0}} \tag{22}$$

Calculation of E_b/N_0 depends on knowing the SNR and the data-rate for the specified bandwidth [32]. That is

$$\frac{E_b}{N_0} = \frac{S}{N} \cdot \frac{B}{R_B} \tag{23}$$

where N_0 is the thermal noise density, B is the bandwidth, $N=N_0*B$ total integrated noise, S is signal power, and R_B is the signal rate.

If the data-rate decreases we will have more time that a bit or symbol holds a given value, giving more energy. Thereby we would expect a lower bit-error rate for increasingly slower data-rates. Alternatively, a better sensitivity is possible at the same bit-error-rate for a slower data-rate.

For a BER of 1E-3 we need E_b/N_0 of 12.4 according to equation 22. Varying the data-rate for the fixed bandwidth and noise in equation 23 the resulting SNRs needed are found for reasonable data-rates in medical applications.

DATA-RATE (Kbps)	20	40	60	75
SNR required (dB)	4.9	7.9	9.7	10.7

Figure 19 SNRs needed for BER=1E-3 for Designed BFSK Receiver Subsystem

The previous section addressed the partitioning of blocks and resulting specifications needed for the desired performance.

Practical/impractical BER validation

A simulation platform that has the gain, noise, and linearity information can be used as a cross-check against calculations. In addition other effects like filtering

and phase noise can be included. A frequency domain link budget analysis can be done, or a more detailed time-based simulation is possible.

For an accurate representation of all effects a time based simulation was done using ADS Ptolemy [34]. The testbench shown in Fig 20 compares what the receiver decodes against what was driven as stimulus. Because there is a delay in propagating through the system proper alignment in time is needed for verification. The BER can be found based upon the number of mismatching symbols for symbols sent. To verify to a BER of 1E-3, at least 1000 symbols need to be pushed through the system. For greater confidence at least an order of magnitude greater than that should be used, that is 10,000 symbols. Even more confidence in the bit-error-rate can be used by setting different seeds for the random data-bit generation.

The theoretical predictions do not account for interfering signals and other traffic, which could also change the reliability of the system. Nearby or strong signals could degrade the SNR of the signal seen, through either inadequate rejection or through non-linear effects.

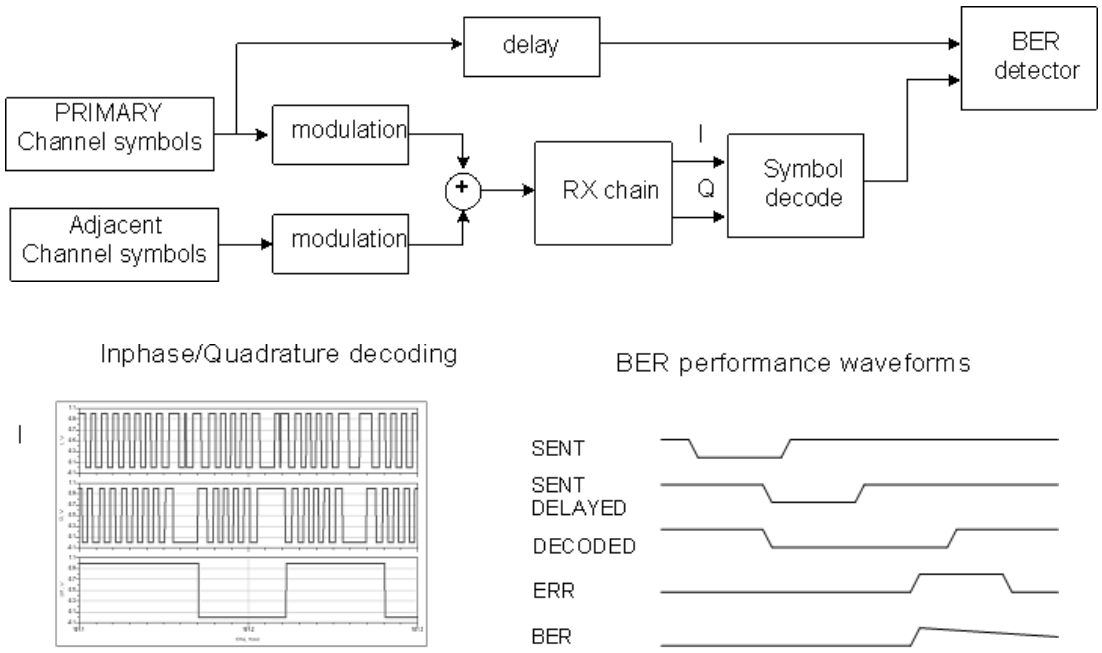


Figure 20 RX system level modeling

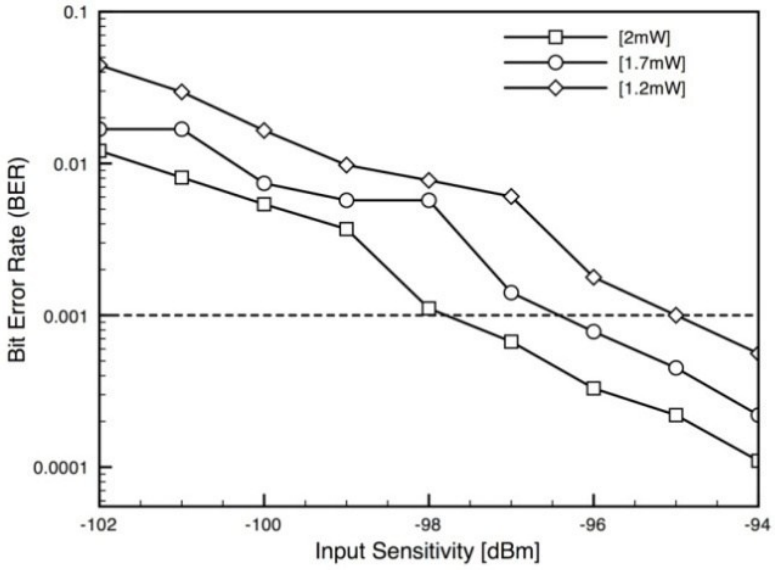
Another avenue to consider in getting even more accurate representations of the system is to go to a circuit level implementation to verification. Although exhaustive subsystem circuit simulations can be done, and a couple symbols can be simulated to verify general system operation, use of these models for full receiver system BER evaluation is extremely time consuming, and not practical.

BER simulation results

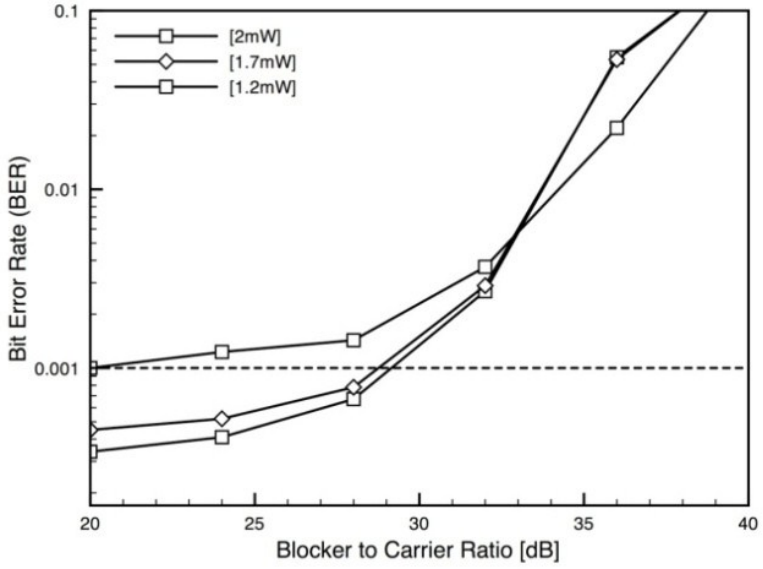
Two main BER simulations were run 1) evaluation of the minimum sensitivity 2) evaluation of tolerance of interferers. Data-rate used is 75 Kbps, and the results can be seen in Fig 21.

Minimum sensitivity for the receiver was checked for varying LNA powers. The higher power in the LNA led to a higher power gain, lower noise figure, and better sensitivity. Because the higher power (current) caused a reduction in input impedance the input tuning network needed to be adjusted in order to maintain the match and get the lower noise figure. The simulations matched closely to the link budget analysis.

The immunity to blockers was found for received signals that are 3dB greater than the sensitivity limit (for 1E-3 BER). The limiting factors are a combination of the linearity limits of the OTA and the channel select stopband attenuation, which are both roughly 35 dB. Phase noise corruption had a much less significant effect.



a) Without interferers



b) with interferers (blocker)

Figure 21 BER results

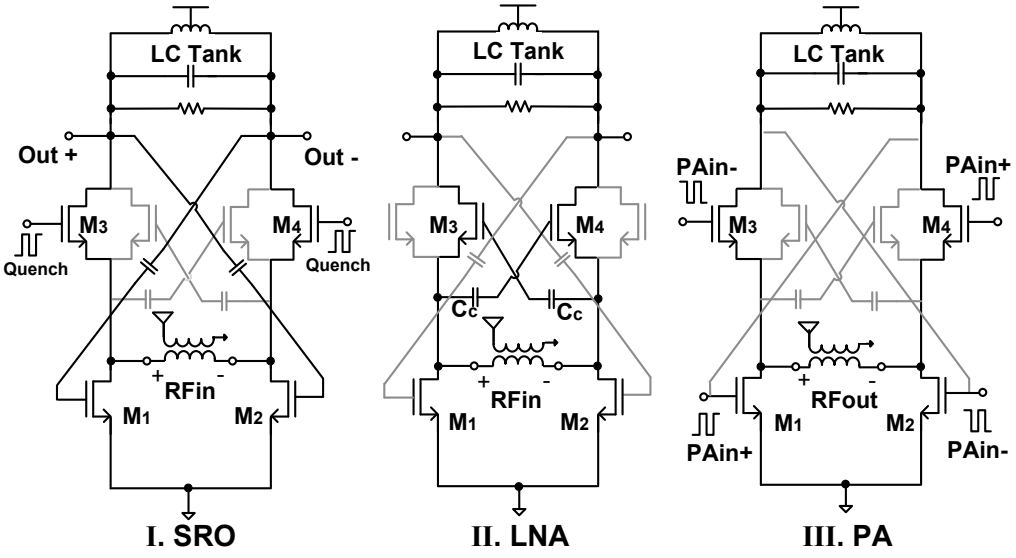
CIRCUIT DESIGN

General Design

Although the system is often portrayed as single ended, all circuits use differential design in order to reject even order harmonics and common-mode signals.

Reconfigurable Front-end

Although only one mode of the front-end is used in the receiver we describe all modes briefly next.



(a) wake-up (b) normal reception and (c) TX PA modes

Figure 22 Reconfigurable RF Front-End

Several reasons explain what allows this high reconfigurability. No requirements for simultaneous transmit and receive operation in the MICS standard, make half duplex operation acceptable. Therefore the same transistors used in receive mode as an LNA or SRO can be reconnected and biased appropriately to perform the PA transmit function. Because the transmit output power is very low, relative to what could be delivered in existing technology, a separate PA is not needed. The relatively low RF frequencies being used make the topology less sensitive to reconfiguration through switches and additional loading.

Overview of RX design choices

There are three subcircuits specifically chosen to target power saving in the receiver chain. First, the differential cross-coupled feedback in the common-gate LNA architecture facilitates best *gm* for the given current. Second, a passive mixer was chosen to lower power consumption, over an active mixer (such as a Gilbert topology). Third, Gm-C filter topology was selected in order to minimize current needed to implement the bandpass filter.

The second consideration in component design was highest possible linearity. Passive mixers in general will have higher linearity than active mixers. In the Direct Conversion architecture second order non-linearity must be minimized to avoid spectra pollution in the baseband during the down-conversion process.

LNA

A differential common-gate topology with cross coupled input feedback is the implementation used, as shown in Fig 23.

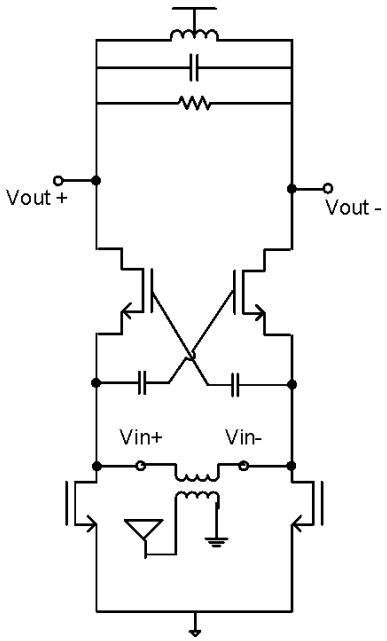


Figure 23 Differential Common-Gate Cross-Coupled LNA

Although the common-gate topology is inherently broadband, the circuit is made narrowband on both the input and output nodes. The input node requires a matching network (which is inherently narrowband) to maximize the received signal power. The output node is a parallel output tank which is tuned to restrict the amplified bandwidth. Although these networks are mostly independent they should both be tuned to the MICS band to maximize sensitivity.

Input matching can be done in many ways. To minimize power and get the best sensitivity lossless matching is used. For the frequencies of interest (which at ~400 MHz is a relatively low RF band), lumped component solutions tend to be the best solutions. The simplest lumped component solution is an L-match network, which uses two reactive components with one shunt and one series to perform matching. It has the weakness of only controlling 2 of the 3 important parameters: impedance transformation ratio, tuned frequency, and Q of the network [35]. The slightly more complicated networks with 3 reactive elements (pi and T networks) have 3 degrees of freedom allowing control of all 3 parameters. Practical considerations in the selection are using the least number of components, not requiring expensive high-Q components, and enabling some of the matching elements to be placed on-chip. Reasonably sized high-Q capacitors can be made on-chip, while high-Q inductors will need to be external components.

Impedance transformation between the antenna and common-gate input can reduce the required current to achieve a matched condition, saving precious power. A lower current will decrease the transconductance (*gm*), increasing the input resistance. From a practical perspective there are 2 problems with decreasing transconductance too much. First, the LNA Mosfet devices need to be minimum gate length for RF operation, which in turn leads to a lower output resistance. If the transconductance decreases too much it will no longer dominate the input resistance setting. Second, the higher the impedance transformation

required the resulting higher Q components that are needed for matching. This not only increases cost, but also makes the bandwidth of the impedance matching smaller.

The non-infinite resistance of the common-gate device means that the drain load will affect the overall input impedance. In order to minimally change the input resistance the smallest reasonable tank resistance should be chosen. Unfortunately this smaller tank resistance works against the maximum gain, which is g_m*R_d .

The input matching will be influenced by the body effect of the common-gate device, which will partially compensate for the additional resistance added in the drain.

The output tank does not need to be matched to the subsequent circuit, because distances involved are much smaller than a wavelength, but selectivity is nevertheless important. The Q of the tank should be large enough to cover the MICS bandwidth, but not much larger. A Q of roughly 80, can be found w0/BW = 406/5. For this high Q circuit an external inductor definitely will be required, although if the size is small enough the capacitor may be placed on-chip.

$$Q = \frac{R}{w_0 L} = w_0 RC \tag{24}$$

Knowing the inductor needs to be larger than packaging parasitic interconnect (of 1-2 nH), and that the capacitor should be no more than 10s of pF for silicon implementation, a reasonable solution is found. Because the tank resistance R_T is

much less than that of the resistance looking into the drain of the common-gate device of

$$R_{OUT} = r_o + [1 + (g_m + g_{mb})r_o]R_S \quad (25)$$

we can approximate the output resistance as R_T.

Knowing the output load, the resulting input resistance can then be found for matching.

The resistive component is

$$R_{in} = \frac{r_0 + R_D}{1 + (g_m + g_{mb})r_0} \quad (26)$$

It should not be forgotten that a reactive part to the input network exists in the form of input capacitances. Specifically the common-gate gate-to-source and the current source gate-to-drain capacitances will be the largest contributors. Capacitive estimates for these devices can be found using the following equations

$$C_{GS} = \frac{2}{3} C_{OX} W(L - 2L_D) + C_{OV} \quad (27)$$

$$C_{GD} = C_{OV} \quad (28)$$

Final input matching uses an L-Match network with an impedance transformation of 1:4 using a standard 50 Ω antenna and mapping to 200 Ω. The capacitor could be placed on-die with the inductor external. Similarly, the output tank network could have the capacitor on-die with the inductor external. The Smith Chart is shown below with the region highlight we expect to need to match.

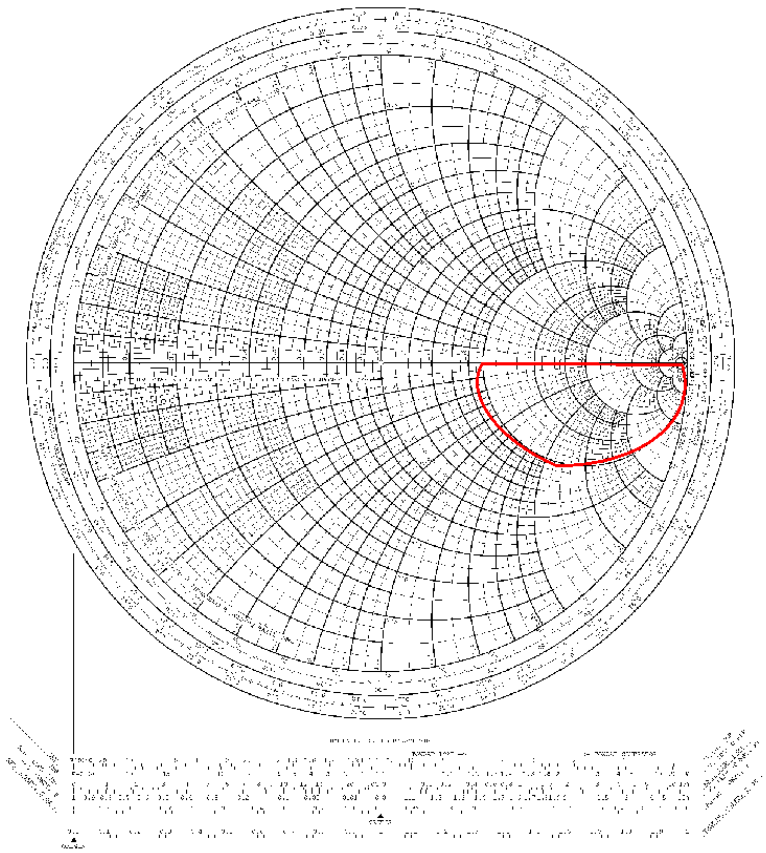


Figure 24 Smith Chart showing area for matching network

Although a calculation for exact component values is done, inevitable manufacturing deviations required tuning of matching components in the lab.

The matching is evaluated through simulation in Cadence as shown in Fig 25, giving an excellent match. Of course the S11 seen in the lab, even after tuning will not be so good. Nevertheless an S11 of less than -10 dB at the frequency of interest indicates a good match, with only 10% reflected power, equating to about 0.5 dB loss in system sensitivity.

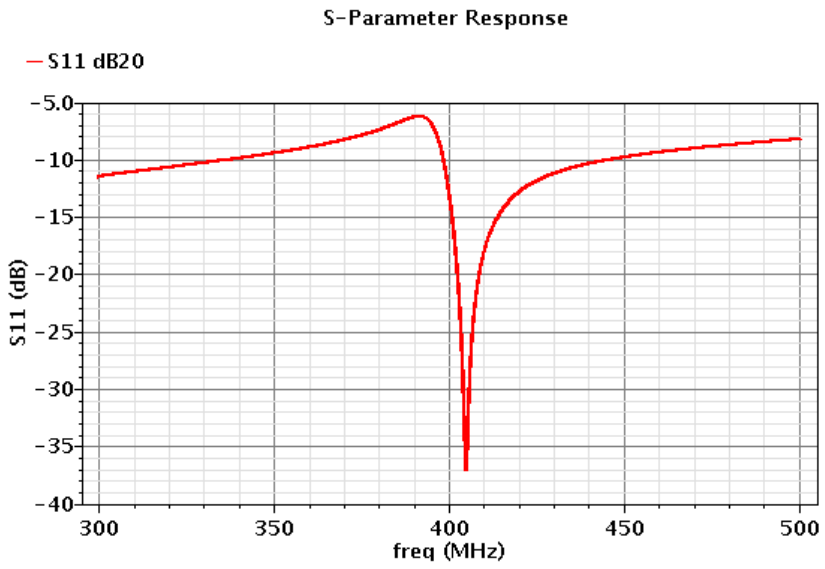


Figure 25 Simulated S11 for LNA

The forward gain (S21) has the peak value as predicted by the equation

$$P = \frac{V_P^2}{2R} = \frac{(g_m \cdot R_T)^2}{2R_T} = \frac{g_m^2 R_T}{2} \equiv 10log_{10}\left(\frac{g_m^2 R_T}{2}\right) \quad (29)$$

As with S11, S21 is simulated and the response is shown in Fig 26.

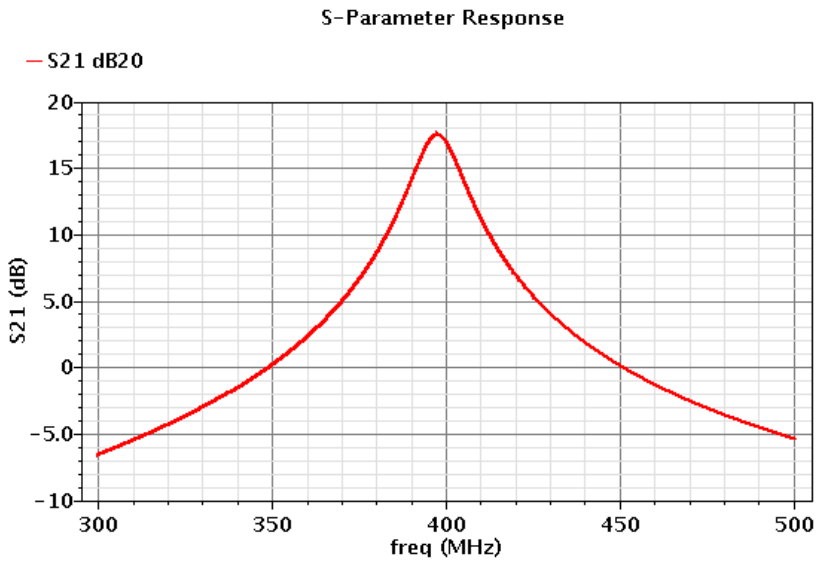


Figure 26 Simulated S21 for LNA

The isolation (S12) is mostly set by Cds, and is very low resulting in very little feedback. At the frequencies of interest it is negligible, and can be neglected. Finally, the S22 is not relevant, because matching is not required in interfacing to the next stage.

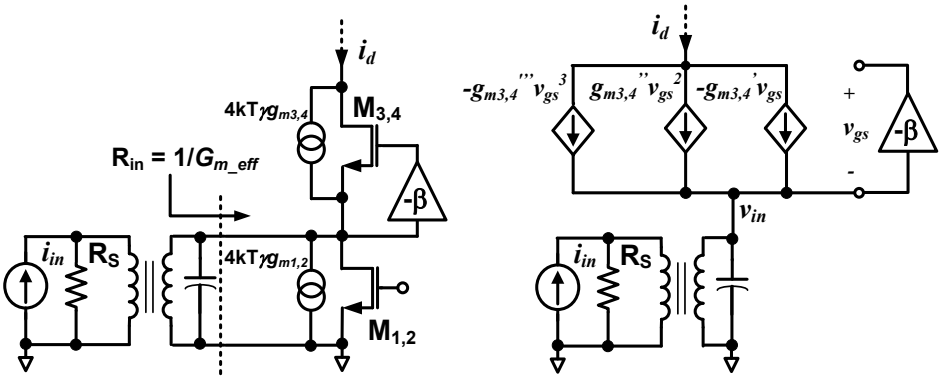
Cross-coupled input feedback increases the transconductance (Gm) for the same current [36], based upon this capacitive ratio (β) as shown below

$$G_{m,eff} = g_{m3,4}\left(1 + \frac{C_C}{C_{GS3,4}+C_C}\right) = g_{m3,4}(1+\beta) \qquad (29)$$

Even with the large devices of the common-gate devices it is not too difficult to make large capacitors of reasonable size to fit on die to accomplish the feedback. The biasing can be modeled as a resistor tied to a voltage, with the coupling

capacitor creating a first order high pass feedback network. A 5pF capacitor along with the 5 KΩ resistor gives a corner frequency of approximate 6MHz, while the Gm boosting factor α is roughly 5pF/(5pF+0.5pF)=0.91

The LNA half-circuit shown in Fig 27 can be used for determining the noise figure and linearity.



a) Noise model b) Linearity model

Figure 27 LNA half circuit model

The noise figure is given below where γ is the MOS excess noise factor, α is the self gain g_m/g_{do} , Rs is the source impedance (antenna), and $g_{m1,2}$ is the transconductance of the current sources.

$$NF = 1 + \frac{\gamma}{\propto G_{m,eff} R_S} + \gamma g_{m1,2} R_S \qquad (30)$$

Minimization of the noise figure is done by maximizing the common-gate transconductance $G_{m,eff}$, while minimizing the tail current source transconductance $g_{m1,2}$. Of course this common-gate transconductance

maximization assumes the input matching is still being done, so maximum gain is being achieved.

The noise figure versus frequency is shown in Fig 28.

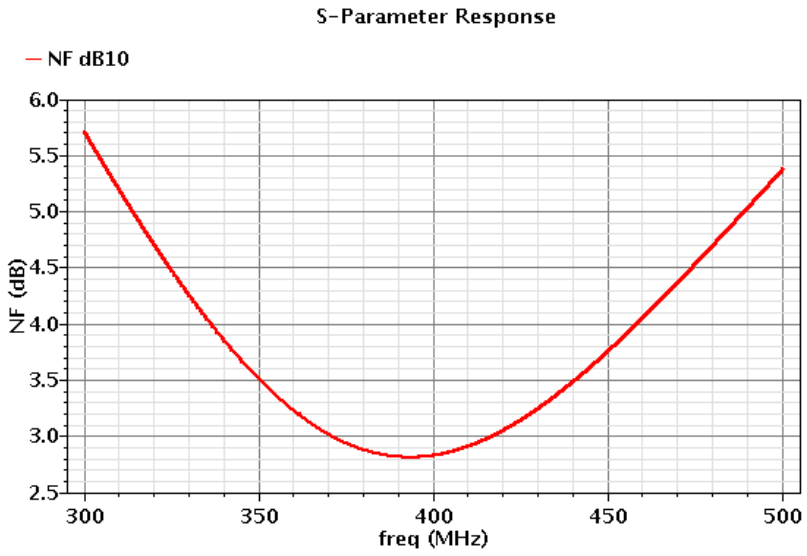


Figure 28 Simulated NF for LNA

Linearity was examined for different capacitive coupling feedback comparing simulation with theoretical expectations as seen in Fig 29. It makes sense that the larger gain accompanying a bigger capacitor ratio causes a lower input linearity point. From a power efficiency perspective it makes sense to have the highest ratio possible.

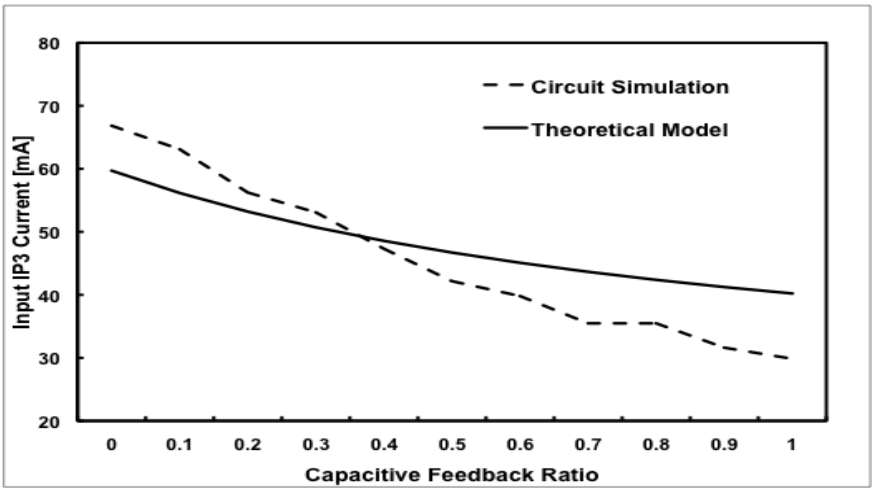


Figure 29 LNA linearity for varying capacitive feedback ratio

The theoretical IIP3 of the gm-boosted LNA is approximately

$$IIP3 = \sqrt{\frac{4}{3}\left(\frac{g_{m3,4}'}{g_{m3,4}'''}\right)}\left[\frac{1}{(1+\beta)R_S} + g_{m3,4}'\right] \quad (31)$$

where $g_{m3,4}$' and $g_{m3,4}$''' are the 1st and 3rd order transconductance coefficients of M3-M4 devices, respectively. Mismatch between the equation and simulations is due to limited filtering of higher order harmonics at the vin node of the circuit being present.

The final realized LNA subcircuit is summarized. C_C is chosen to be much larger than $C_{GS3,4}$ so that the effective transconductance is roughly twice the value without cross-coupling, which affects the gain, noise figure, and linearity. The LNA power gain (S21) is 16.5 dB, the noise figure is 2.8 dB, and IIP3=-26 dBm.

MIXER

The doubly balanced passive mixer is shown in Fig 30. These mixers offer several benefits.

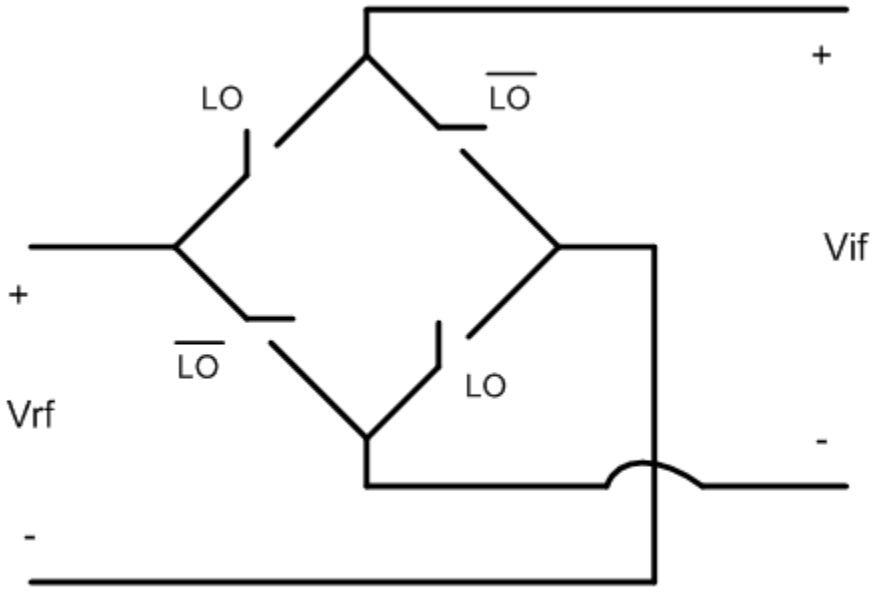


Figure 30 Passive Doubly Balanced Mixer

By being doubly balanced, as opposed to singly balanced there is more elimination of even order distortion. Even order distortion (especially 2^{nd} order distortion) is a big concern for direct conversion systems, where these non-linearities could place spurs in-band, degrading SNR. The differential LO and RF signals fed from the LNA make the system doubly balanced.

The "passive" description of the mixer is referring to the absence of a bias current in the circuit, which generally leads to less consumed power. Primary power

consumed in the subsystem is in driving the switches, for changing between conductive states and non-conductive states. The power consumed in the drivers is proportional to cv^2f [37]. For Direct Conversion operation the driving frequency is the RF channel centerpoint, and the voltage is the full supply range for the largest linearity, but the capacitance is a variable to be minimized for lowest power.

Direct conversion mixers have a noise figure that will be 3 dB lower than that of a Super-heterodyne system. Because the local oscillator is centered on the frequency to be down-converted, image signal bandwidth noise is not aliased into the mixing operation. That is we have double sideband noise, versus single sideband noise. This advantage effectively helps provide more margin in the design in the quest for a lower noise figure.

At the same time it is important to note that because the switches are driven by a square wave, odd order harmonics down-convert to the same baseband as for the intended fundamental component. These additional double sided components contribute unintended information. Fortunately, this contribution is not a significant degradation, and it is described below. The harmonics of the LO (n) will be of lower amplitude than the fundamental and is approximately proportional to π/n. Because the LNA is a tuned amplifier, those frequencies away from the tank resonance are attenuated. The two tuned stages form a second order BPF, where the form is

$$H_{BP}(s) = \frac{\frac{w_0}{Q_1}s}{s^2+\frac{w_0}{Q_1}s+w_0^2} \cdot \frac{\frac{w_0}{Q_2}s}{s^2+\frac{w_0}{Q_2}s+w_0^2} \tag{32}$$

The resulting degradation is adding two frequency bands on either side the harmonic, and is

$$\left\{\frac{\pi}{3}cos(3w_0t)\right\} \cdot 2 \cdot \left\{\frac{\frac{w_0}{Q_1}[3w_0]}{\sqrt{\left(w_0^2-[3w_0]^2\right)^2+\left(\frac{w_0}{Q_1}[3w_0]\right)^2}}\right\}\left\{\frac{\frac{w_0}{Q_2}[3w_0]}{\sqrt{\left(w_0^2-[3w_0]^2\right)^2+\left(\frac{w_0}{Q_2}[3w_0]\right)^2}}\right\} + \tag{33}$$

$$\left\{\frac{\pi}{5}cos(5w_0t)\right\} \cdot 2 \cdot \left\{\frac{\frac{w_0}{Q_1}[5w_0]}{\sqrt{\left(w_0^2-[5w_0]^2\right)^2+\left(\frac{w_0}{Q_1}[5w_0]\right)^2}}\right\}\left\{\frac{\frac{w_0}{Q_2}[5w_0]}{\sqrt{\left(w_0^2-[5w_0]^2\right)^2+\left(\frac{w_0}{Q_2}[5w_0]\right)^2}}\right\} + \cdots$$

A graphical representation is shown in Figure 31.

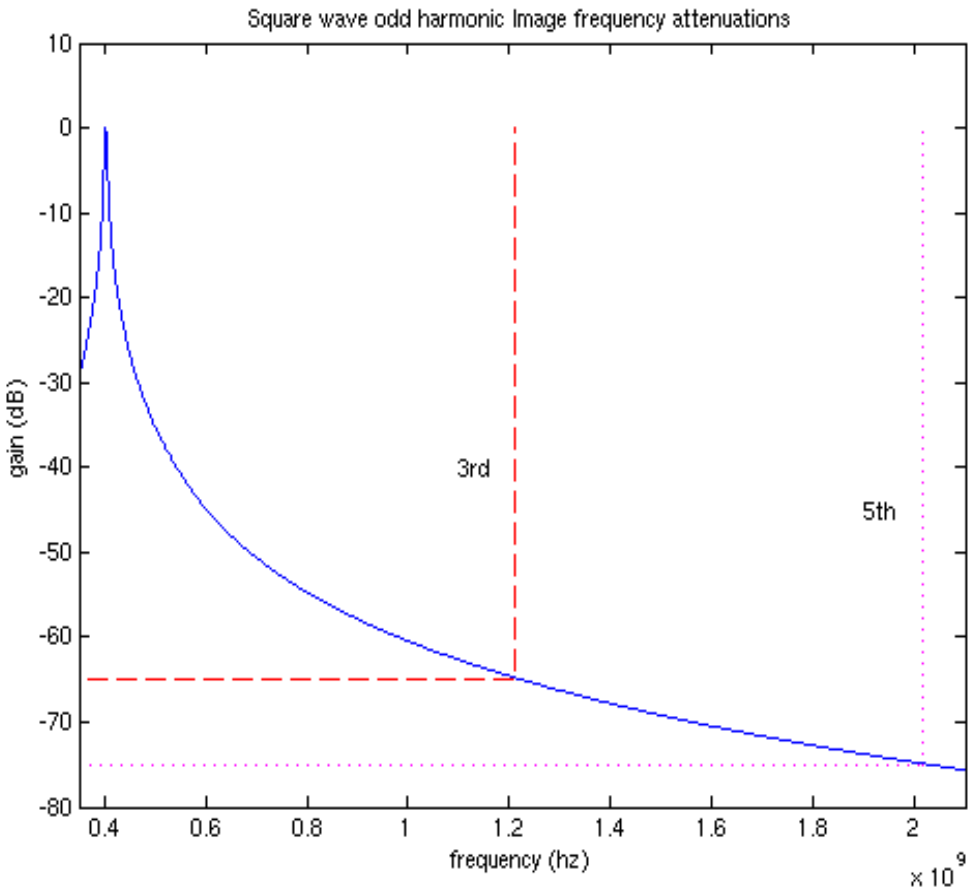


Figure 31 Odd Harmonic Image Frequency attenuations

The third and fifth harmonics are in the frequency range of 1206-1215 MHz and 2010-2025 MHz. For the tank Qs expected we have Q1=3,Q2=80 giving a magnitude attenuation of $\geq$ 65dB for the first two odd harmonics. Only if there are very strong signals at these frequencies of interest do they need to be considered as image noise components.

The passive switched circuit has a very large linear range of operation. In fact the range should be roughly 2 times a peak-to-peak voltage range roughly that of the

supply, because of differential operation, that is $20 * log_{10}\left(\frac{2*0.9}{\sqrt{2}}\right)$ =2.1 dB = 32.1 dBm. This assumes the voltage bias is half way between the power supply. The source follower buffer after the LNA performs the level shift from the supply level and also provides a low resistance source. Linearity starts to falls off once the rails are exceeded because the drain/bulk and source/bulk voltages of the mosfets will be clamped through the inherent forward biased PN junctions of the CMOS process.

These passive mixers also have drawbacks. The lack of gain in the passive system directly increases the noise figure. Also, because the mixer down-converts to DC, flicker noise is a significant concern. Finally a doubly balanced circuit has more switches to be toggled leading to higher driver power. The circuit is shown in Fig 32. For simplicity, a non-differential form is shown.

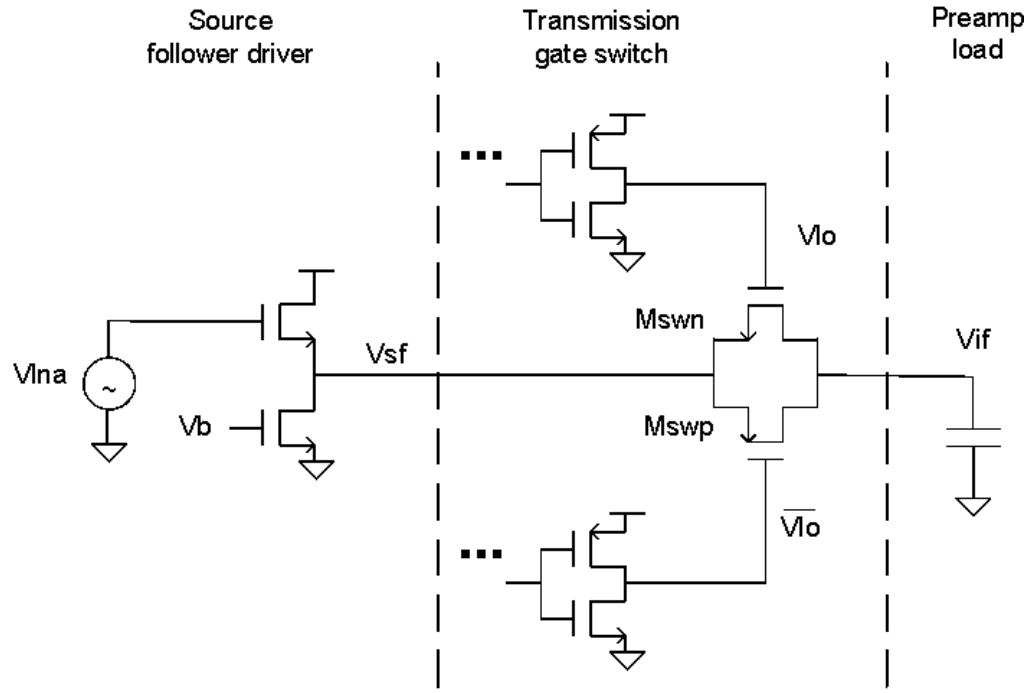


Figure 32 Circuit Diagram for Mixer switch path for conversion gain and LO leakage

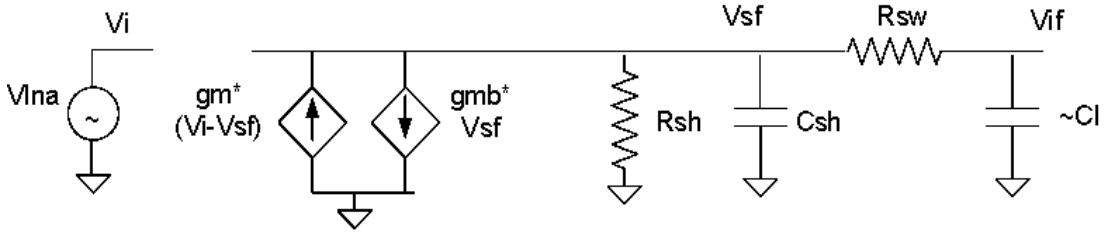


Figure 33 Mixer small signal model

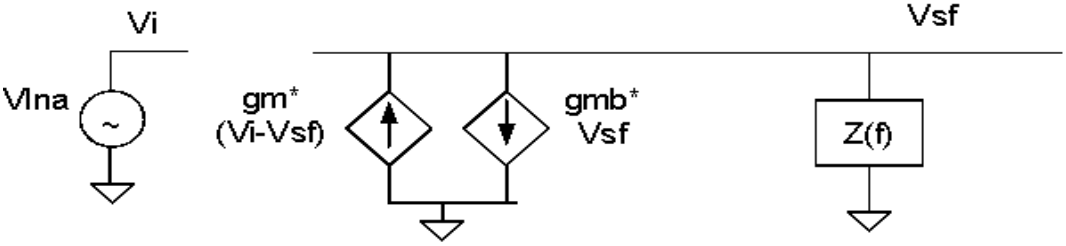


Figure 34 Mixer small signal Simplification

Looking at the small signal diagram in Fig 33 we get a general idea of the performance limitations. Simplifying the model by lumping the switches and next stage loading Fig 34 can give more insight.

$$g_m(V_I - V_S) = g_{mb}V_S + \frac{V_S}{Z_{EQ}(f)} \tag{34}$$

$$\frac{g_m}{\left[g_{mb}+g_m+\frac{1}{Z_{EQ}(f)}\right]} = \frac{V_S}{V_I} \tag{35}$$

The equivalent impedance is

$$Z_{EQ}(s) = \frac{s(R_{SH}R_{SW}C_L)+R_{SH}}{s^2R_{SH}R_{SW}C_LC_{SH}+s(R_{SH}C_L+R_{SW}C_L+R_{SH}C_{SH})+1} \tag{36}$$

By design Z_{EQ} is made to be a very high impedance at the down-converted frequencies, but low impedance at the mid to high frequencies. Therefore, the fundamental limit to conversion gain is due to the body effect (g_{mb}). The attenuation at higher frequencies requires taking into account the terms in equation 36.

Expanding out Fig 33 in Fig 35 reveals the contributors to the terms in equation 36, where the nodes V_{LO} and V_{LOB} are static values (ac grounds).

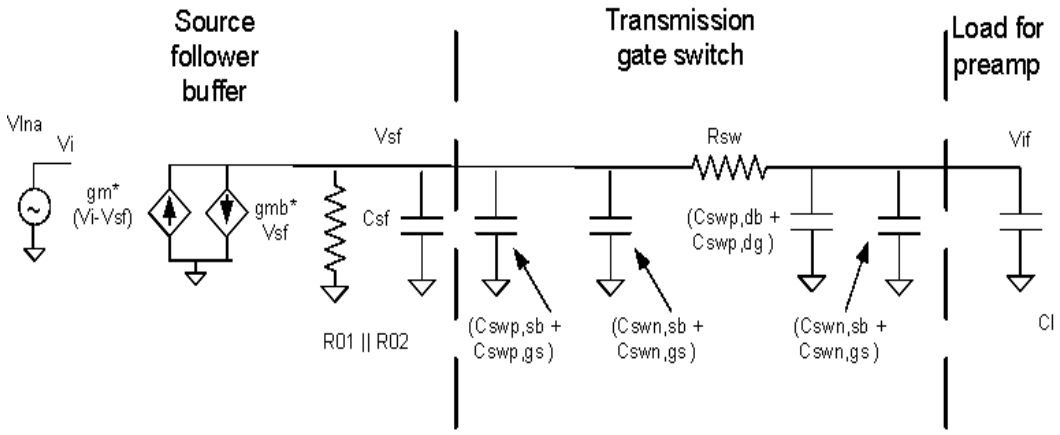


Figure 35 Detailed parasitics for Mixer switch path (conversion gain)

$$C_{SH} = C_{GD,2} + C_{DS,2} + C_{DS,1} + C_{SB,1} + C_{DB,2} + C_{SWP,SB} + C_{SWP,GS} + C_{SWN,SB} + C_{SWN,GS} \quad (37)$$

$$C_L = C_l + C_{SWN,SB} + C_{SWN,GS} + C_{SWP,DB} + C_{SWP,DG} \approx C_l \quad (38)$$

$$R_{SH} = R_{01} || R_{02} \quad (39)$$

Multiplying the two transfer functions we can get the final transfer function.

$$\frac{V_{IF}}{V_{LNA}}(s) = \frac{V_S}{V_I} \cdot \frac{V_{IF}}{V_S} = \frac{g_m}{\left[g_{mb} + g_m + \frac{1}{Z_{EQ}(s)}\right]} \cdot \frac{1}{\{sC_L R_{SW} + 1\}} \quad (40)$$

For this topology the conversion gain is found in evaluating the response by applying a square wave to the LO port and the received RF signal through the other port. The ideal conversion gain is simply mixing the fundamental of the LO

with the RF signal, and is $\frac{2}{\pi}$ [38]. Two loss mechanisms account for a realizable conversion gain less than the ideal predicted. First, the shunted load before the switch, reduces some of the current through the switch, reducing the voltage at the IF node. This loss can be found through this current ratio (evaluated at the RF frequency).

$$20log_{10}\left[\frac{I_{IF}}{I_{SH}+I_{IF}}\right] = 20log_{10}\left[\frac{\left(R_{SH}||\left[1/wC_{SH}\right]\right)}{\left(R_{SH}||\left[1/wC_{SH}\right]\right)+\left(R_{SW}+\left[1/wC_{L}\right]\right)}\right] \quad (41)$$

Second, the load after the switch at the IF node can reduce the voltage seen, if the corner frequency is either below or too close to the IF frequency. Equation 42 can be evaluated at the IF frequency to account for this loss.

$$\frac{V_{IF}}{V_S} = \frac{1}{\{sC_L R_{SW}+1\}} \quad (42)$$

One concern is that the resistance in the switch is variable, depending upon the bias point and instantaneous value. The change in resistance can change the 3dB point of the frequency response. In order to minimize the change in resulting resistance and provide more robust operation, complementary transmission gate structures are used. Figure 33 compares using single transistor devices versus complementary switch device resistance.

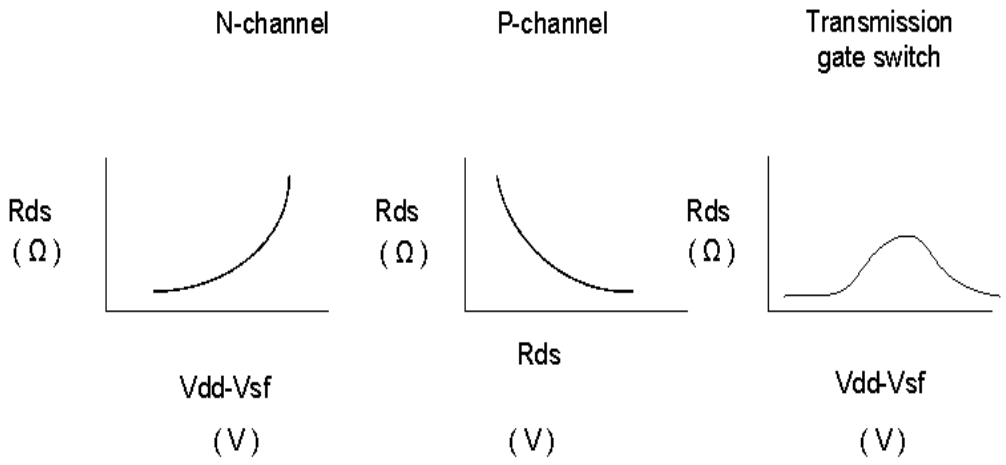


Figure 36 Switch resistance versus bias point

The equations are below describing the devices (for the linear region) [39].

$$R_{DS,N} = \left[K'_N \left(\frac{W_N}{L_N}\right)(V_{GS} - V_{TN})\right]^{-1} \; ; \; R_{DS,P} = \left[K'_P \left(\frac{W_P}{L_P}\right)(V_{SG} - V_{TP})\right]^{-1} \tag{43}$$

More specifically for the source follower circuit

$$R_{DS,N} = \left[K'_N \left(\frac{W_N}{L_N}\right)(\{V_{DD} - V_{SF}\} - V_{TN})\right]^{-1} \; ;$$

$$R_{DS,P} = \left[K'_P \left(\frac{W_P}{L_P}\right)(\{V_{SF} - V_{DD}\} - V_{TP})\right]^{-1} \tag{44}$$

The CMOS t-gate form is just

$$R_{DS,TGATE} = R_{DS,N} \,||R_{DS,P} \tag{45}$$

Plugging in the parasitics and using the derived equations can provide an expectation of the final conversion gain.

All structures will have a non-constant V_T , because of the body effect. This will increase the threshold voltage, increasing the resistance.

Cap parasitic from the source follower circuit are	24 fF
Switch parasitics may be	12 fF
Adding some margin for routing parasitic	14 fF
final shunt capacitance Csh	50 fF

Figure 37 Input shunt capacitance for mixer

A CMOS t-gate can have a peak resistance of 1 KΩ, with the source follower having a series resistance of 1 KΩ and the shunt source follower resistance is 50 KΩ .

Conversion gain optimization trades off lower shunt capacitance with higher resistance in the switch.

Doubly balanced conversion gain	-3.92 dB
Circuit loss at RF	-1.22 dB
Circuit loss at IF	-0.17 dB
final Conversion gain	-5.31 dB

Figure 38 Mixer Conversion Loss components

Measurement of the conversion gain gives about 1 dB less than the ideal. It is also observed the conversion gain decreases at high amplitudes due to gain compression.

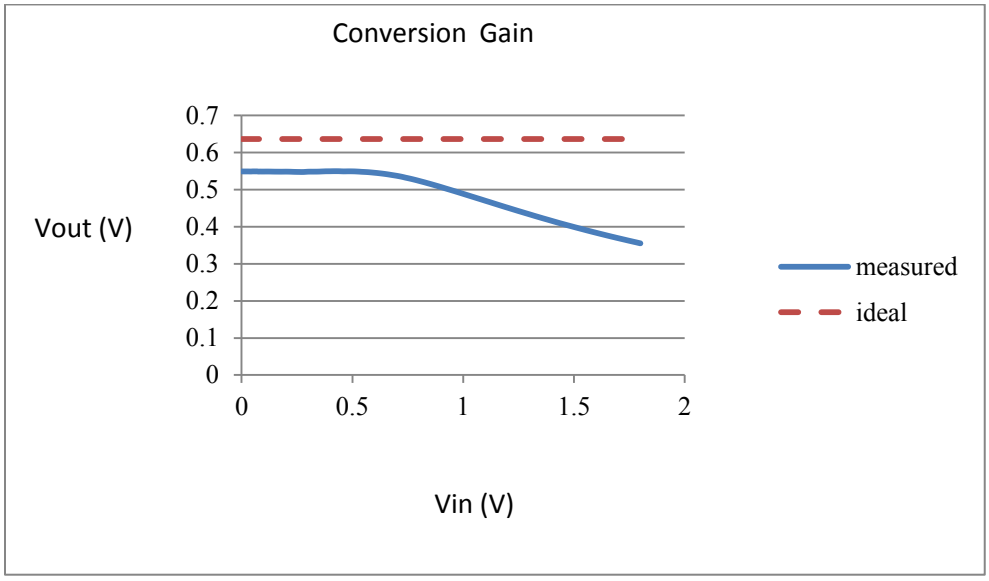


Figure 39 Mixer Conversion Gain

Mixer Linearity is shown in Fig 40 and quantified by doing a polynomial fit.

The resulting third order polynomial is

$$y(t)= \alpha_3*x^3(t) + \alpha_2*x^2(t) + \alpha_1*x(t) + \alpha_0 \quad (46)$$

$$=-0.0280*x^3(t) + (-0.2661)*x^2(t) + 1.2763*x(t) + (-0.0129)$$

Using these coefficients and the following relations

$$IIP3 = \sqrt{\frac{4\alpha_1}{3\alpha_3}},\ IIP2 = \frac{\alpha_1}{\alpha_2} \quad (47)$$

The following linearity measures are found IIP3 = 11.8 dBm ,IIP2 = 7.6 dBm

.

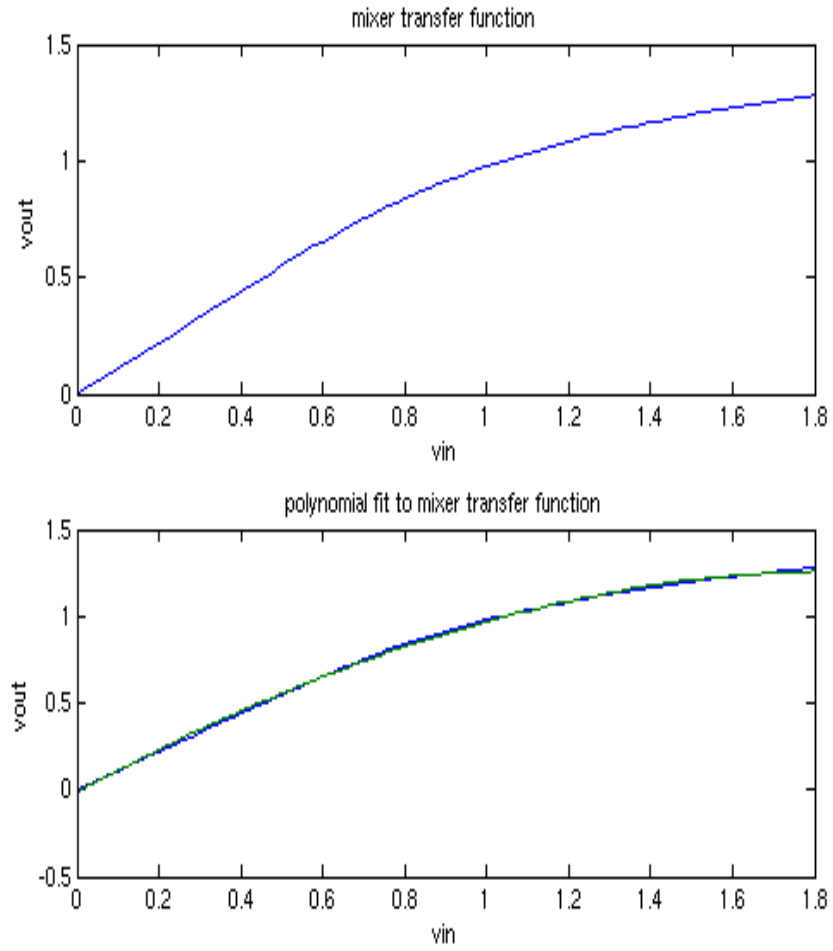


Figure 40 Mixer Linearity

Isolation between the ports is also an important issue, which can lead to desensitization in subsequent stages through propagation of the strong LO signal.

Ineffectiveness in addressing this issue can lead to self-mixing effects in the IF through weak LO-RF isolation, and feedthrough of the LO fundamental to the IF when having weak LO-IF isolation. In spite of the differential architecture and complementary CMOS switches, isolation is not automatically optimized.

The small signal model, where the LO is the dominant source, is shown in Fig 41.

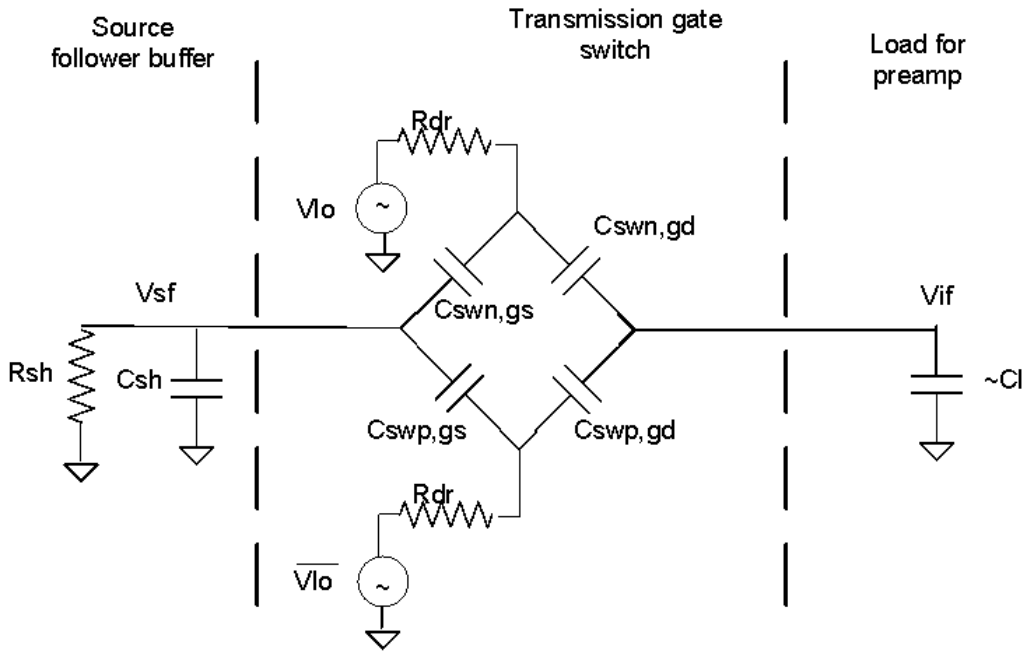


Figure 41 Detailed parasitics for LO transfer functions

A common choice in sizing the PMOS and NMOS transistors is building the switch is to make the PMOS transistor roughly 2-3x larger W/L ratio (based upon technology dependent mobility differences) in order to equate the resistance of each device. This also has the effect of minimizing the peak composite resistance and centering it between the supplies. The problem caused in selecting this sizing

is the capacitance ratio between the differential LO drivers is such that much of the LO magnitude will be transferred. This can be seen through superposition of the two differential inputs, using Sz factor.

In the case of LO-IF transfer we can see

$$V_{IF} = \left(\frac{C_S}{C_S+C_L}\right) V_{LO} + \left(\frac{S_Z \cdot C_S}{S_Z \cdot C_S+C_L}\right)(-V_{LO}) = V_{LO}\left[\frac{C_S}{C_S+C_L} - \frac{S_Z \cdot C_S}{S_Z \cdot C_S+C_L}\right] \quad (48)$$

If Sz=3, and using the assumption $C_L >> C_S$ then

$$V_{IF} = V_{LO}\left[\frac{C_S}{C_S+C_L} - \frac{3 \cdot C_S}{3 \cdot C_S+C_L}\right] \approx V_{LO}\left[-\frac{2 \cdot C_S}{C_L}\right] \quad (49)$$

For the best isolation an exact device capacitance match is desired, with Sz=1. Of course differences in biasing levels, process variations, and W/L variation will cause perfect isolation from being achieved, but to a first order much better cancellation is achieved as seen below.

$$V_{IF} = V_{LO}\left[\frac{C_S}{C_S+C_L} - \frac{1 \cdot C_S}{1 \cdot C_S+C_L}\right] = V_{LO}\left[-\frac{0}{C_S+C_L}\right] \quad (50)$$

Using Sz=1 and the conservative variation of 10%, a corresponding 10% of the original magnitude is driven through the isolation paths. That is, a 20dB additional attenuation is achieved.

$$V_{IF} = V_{LO}\left[\frac{C_S}{C_S+C_L} - \frac{1.1 \cdot C_S}{1.1 \cdot C_S+C_L}\right] \approx V_{LO}\left[-\frac{0.1 \cdot C_S}{C_L}\right] \quad (51)$$

For Cs=0.05 pF and Cl=10 pF with Sz=3 we get isolation of -20log10(1/100)=40 dB. With just 10% mismatch, for Sz=1, we see a much better isolation of -20log10(0.005)=66 dB .

LO-RF isolation is also an important issue because of the self mixing phenomena. That is, a DC term can be fed to the IF port. The importance of matched coupling capacitances is the same, except instead of a capacitive divider, for typical circuit parameters, the circuit behaves as a highpass network.

We see that in essence the voltage gain from LO to RF is

$$\frac{V_{RF}}{V_{LO}} \approx \frac{Z_{SF}}{Z_{SF}+Z_S} \tag{52}$$

Using the assumption that impedance is dominated by the capacitances, we get a capacitive divider of

$$\frac{V_{RF}}{V_{LO}} \approx \frac{R_{SF}}{R_{SF}+\frac{1}{sC_S}} = \frac{sC_S R_{SF}}{sC_S R_{SF}+1} = \frac{s}{s+\frac{1}{C_S R_{SF}}} = 20 \cdot log_{10}\left(\frac{2\pi f}{\sqrt{4\pi^2 f^2+\left(\frac{1}{C_S R_{SF}}\right)^2}}\right) \tag{53}$$

Corner frequency is at 3.18 Ghz. For Cs=0.05 pF, Rsf=1KΩ and f=400 Mhz we get -18 dB.

Of course isolation improves with more cancellation through differential action, as seen in the case of LO-IF isolation. With the reasonable 10% matching, additional 20 dB of isolation is added.

In summary the best overall performance is achieved by making the switches small. This gives small capacitances for the switch, which helps reduce LO driver power, isolation, and conversion gain. The limit on switch size must consider the higher switch resistance and more variation/mismatch that will result from the smaller device geometries.

VGA

After the mixer, signal strength needs to be increased prior to feeding into the bandpass filter. Otherwise, the noise would hide the weakest of signals, significantly degrading the noise figure.

The amplification is a cascade of differential self biased loads, and is programmable as shown in Fig 42. This variable gain is mostly intended to compensate for PVT variations, but could be used with a receive strength signal indicator to adjust the gain for a better dynamic range. In this system implementation, no receive signal strength indicator exists.

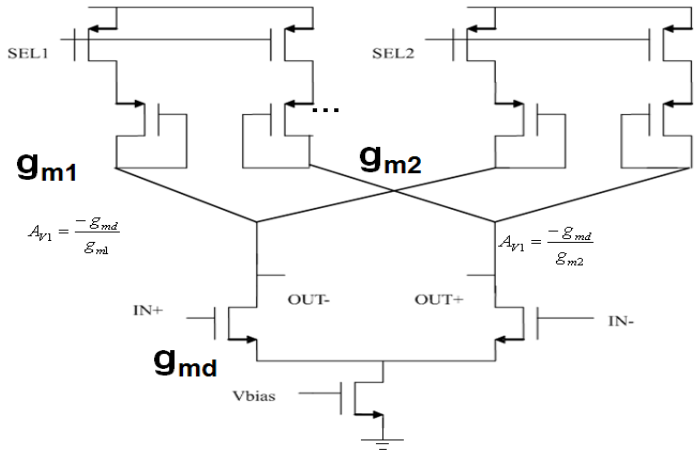


Figure 42 programmable variable gain amp

Large area devices used to minimize the introduced flicker noise also have the benefit of adding enough capacitance to help filter higher frequency downconverted signals, through the low pass transfer function. The corner frequency is high enough that even with PVT variation it is much higher than the intended down-converted bandwidth.

This reduction of bandwidth in the preamp helps condition the signal to the bandpass filter to have less interacting interferers thereby helping improve the performance of the filter in the system.

BPF

The filter required is of the bandpass type, where the center frequency of the filter should be the frequency deviation of the transmitted BFSK signal. Only a fraction of the total channel bandwidth is needed for the high modulation index demodulation. Excessive bandwidth simply degrades the SNR with no added value. Attenuation of lower frequencies reduces DC offset and flicker noise, from the mixer down-conversion. Attenuation of higher frequencies helps to suppress adjacent channel activity or other interference caused through 2^{nd} or 3^{rd} order distortion products from further up the receiver chain. The requirements balance the needs of having the highest modulation index, supporting a high data rate, while providing reasonable requirements for stopband attenuation. Fig 43 shows the requirements along with the elements addressed through each band.

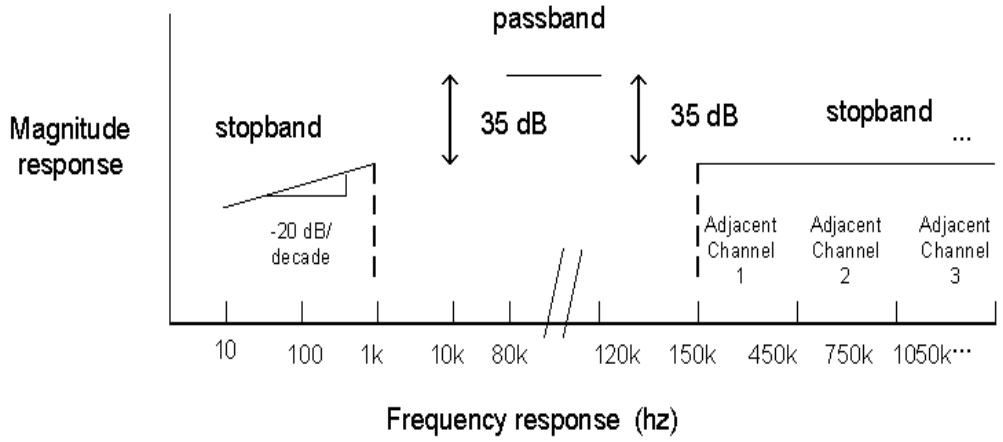


Figure 43 BPF requirements

Because flicker noise is proportional to 1/f, essentially increasing for each decade increase in frequency, the filter must at least counteract this effect.

As long as the interfering signals are not close to the IIP3 level, then no significant in-band spurs should be created. The interferers will then be reduced by the stopband attenuation (35dB). As long as the interferers do not rise above the noise floor (after attenuation), they will not impact performance. In the case of this system with a sensitivity of -95 dBm, the system should be able to tolerate up to -60 dBm interferers.

Because the filter attenuation is not flat in the stopband, but decreasing, even higher interference signals could be handled as seen through examining the magnitude transfer function.

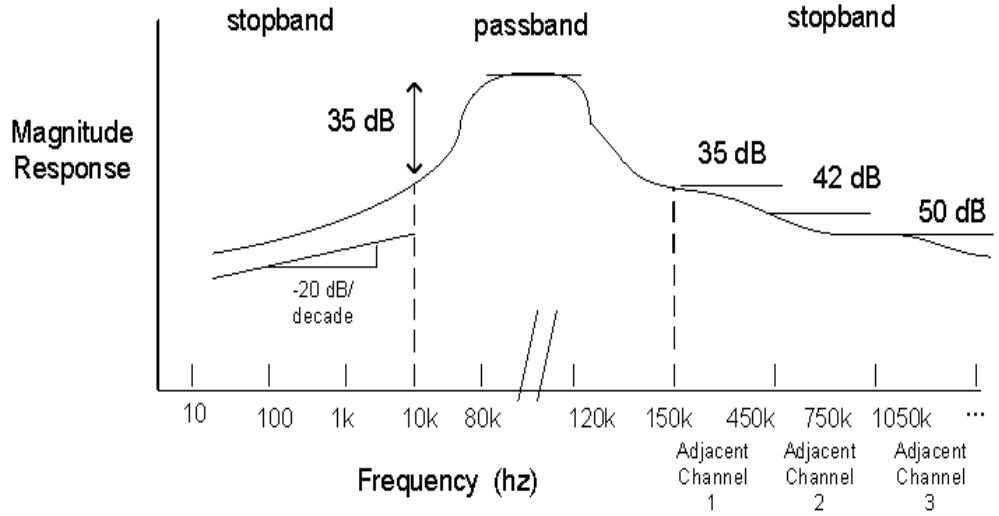


Figure 44 Refined BPF requirements

tandard BPF designs have symmetrical responses. That is, each passband typically has the same attenuation per decade away from the center frequency, for example a 2nd order BPF would have 40 dB attenuation per decade. This is acceptable for the low frequency rejection, but not for the higher frequency rejection, due to the fact that a couple of the adjacent channels will not be adequately attenuated. This is seen in Fig 45 for a Butterworth filter where the -40 dB/decade does not rolloff fast enough above the center frequency.

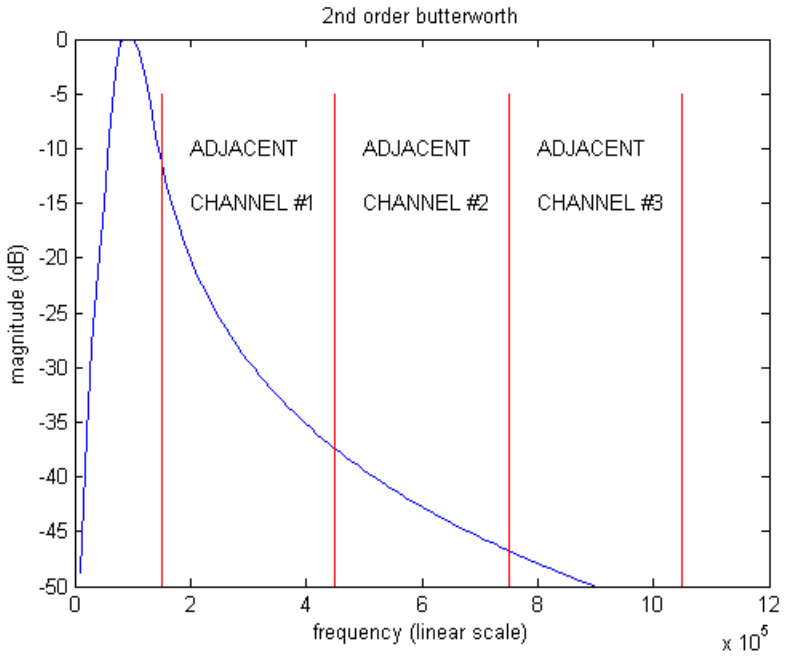


Figure 45 BPF transfer function (linear scale)

A straightforward approach would simply be to simply increase the order of the filter to provide the required rejection. A more targeted approach of increasing the

attenuation of immediately adjacent channels will similarly address the problem. This design takes the targeted approach to reduce overall filter size/complexity.

A common method increasing attenuation on the high frequencies is to incorporate zeros into the transfer function in the stopband region needing additional rejection [40][41]. Although canonical filter designs such as the Inverse Chebyshev or Elliptical can include the zeros, they are typically inserted symmetrically. Just as before the symmetrically placed zeros (about the center frequency) are extra overhead that make the design larger than necessary.

Although canonical filters are convenient in that all components have been calculated and the structure is in a ladder form, which has minimum frequency sensitivity to component variation, incorporation of the requirement cited in the last paragraph is not within the design methodology. Therefore, the filter design approach chosen is different, using a combination of independent cascaded stages to implement the desired transfer function. The first stages will implement a canonical Butterworth bandpass filter, while the second stages will add zeros to the design.

The next figure shows the discussed transfer functions to address the filtering requirement

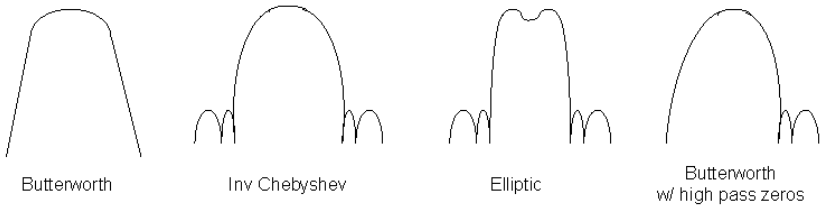


Figure 46 Various BPF frequency responses

A second order Butterworth filter is designed using commonly available tables for design coefficients. The filter form is

$$H_{BP}(s) = \frac{\frac{w_0}{Q}s}{s^2+\frac{w_0}{Q}s+w_0^2} \tag{54}$$

It meets the requirements on the low side of the center frequency. The high frequency attenuation requirement near the passband is met through addition of zeros.

The zeros or notch filters used are 2nd order and are described by the equation

$$H_N(s) = \frac{s^2+w_N^2}{s^2+\frac{w_N}{Q}s+w_N^2} \tag{55}$$

Both the notch frequency W_N and the Q of that notch are described by the above transfer function. Inspection of the transfer function reveals the gain is 1 away from the notch frequencies. This makes it easier to see the effect solely due to notch filter inclusion.

Placement of notch filters depends upon the desired attenuation in the stopband. Because most attenuation will be closest to the transition band/stopband boundary, generally this is the area to start placement. The next consideration in placement of notches is Q used. A smaller Q will provide a wider notch region, but with the negative consequence of also potentially distorting the frequency response of the passband region. A larger Q not only will be less effective in attenuating a larger range of frequencies, but may require more notches than lower Q ones for the same effectiveness. Bounded by minimally distorting the passband, the Q limits of the technology, and the desired adjacent channel attenuation the frequency placement and Q's of the corresponding notch filters can be determined.

A matlab procedure was written to explore the notch placement effect on stopband attenuation for the complete transfer function. That is, the transfer function is evaluated with bandpass filter and the one or more notch filters introduced. Once 2 or more notches have been included there will be lower attenuation sections where the peak or lobe might not be even to the others. For the optimal solution, no peak should be higher than the others, that is we are searching for an equiripple response in the stopband. Therefore, any responses that have too large a difference in peaks are discarded. The final guidance given to this procedure was to start with higher Qs, and use a lower Q for subsequent notch placements, with the reasoning that the farther the notches are from the passband, the less effect they should have on that response. Interestingly, this

same placement of zeros can be seen in the canonical filter forms, through visual inspection of the magnitude response.

Up to 35 dB in attenuation in the stopband was the goal, which was found to be achievable through the use of 2 notches in the adjacent band. The attenuation beyond these notches are caused by poles in the bandpass filter sections.

One effect of the several solutions is that in spite of using reasonably higher Q solutions, the passband response was still affected negatively. The previously flat passband response was made to droop at the higher passband frequencies by a couple dB. Equalization of the magnitude could be achieved through cascading a 2^{nd} order LPF section with a Q> 0.707 that would have the required amount of peaking.

$$H_{LP}(s) = \frac{w_0^2}{s^2 + \frac{w_0}{Q}s + w_0^2} \tag{56}$$

Magnitude equalization in the passband, also affected the stopband response, due to 2 additional poles in the response. As a result, a concurrent optimization accounting for all these factors was redone. Fig 47 shows without this correction factor, and then with it included.

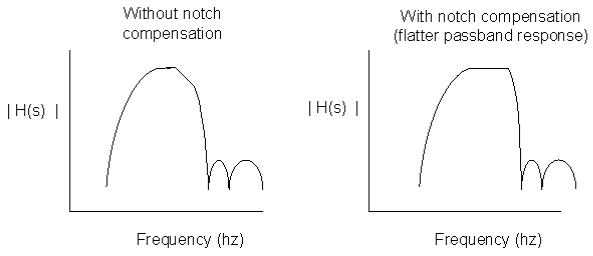


Figure 47 BPF notch compensation included

The final BPF structure is shown in Fig 48. Each block is a 2nd order function, with the BPF and LPF functions being all-pole systems. The notch filters as outlined earlier have poles and zeros.

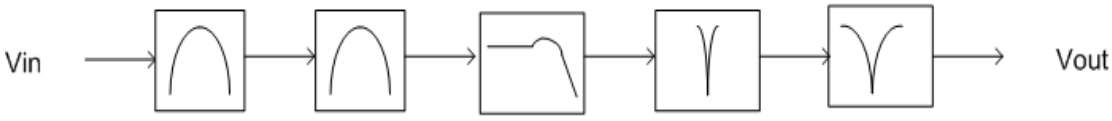


Figure 48 Cascaded structure of composite BPF

The resulting magnitude response is shown in Fig 49.

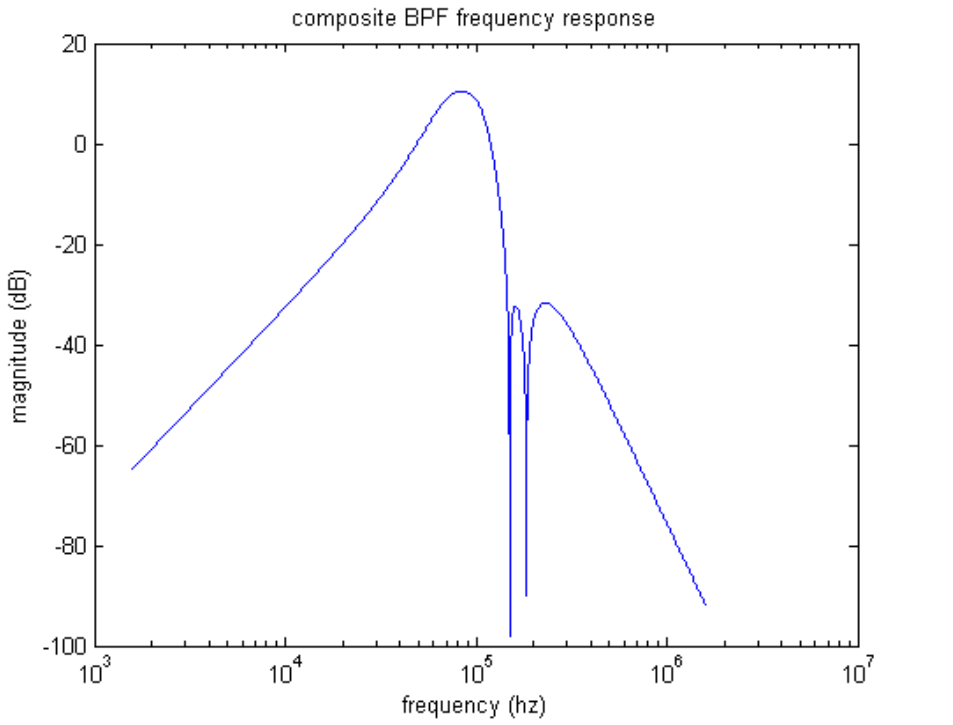


Figure 49 Magnitude frequency response for composite BPF

Implementation

The implementation technology for the filter is through use of Gm-C.

Many different transconductance amplifiers have been used in filter implementations, but the choice for this design is a differential amplifier with active load with common mode feedback as shown in Fig 50. This choice balances the need for low power and linearity along with tunability.

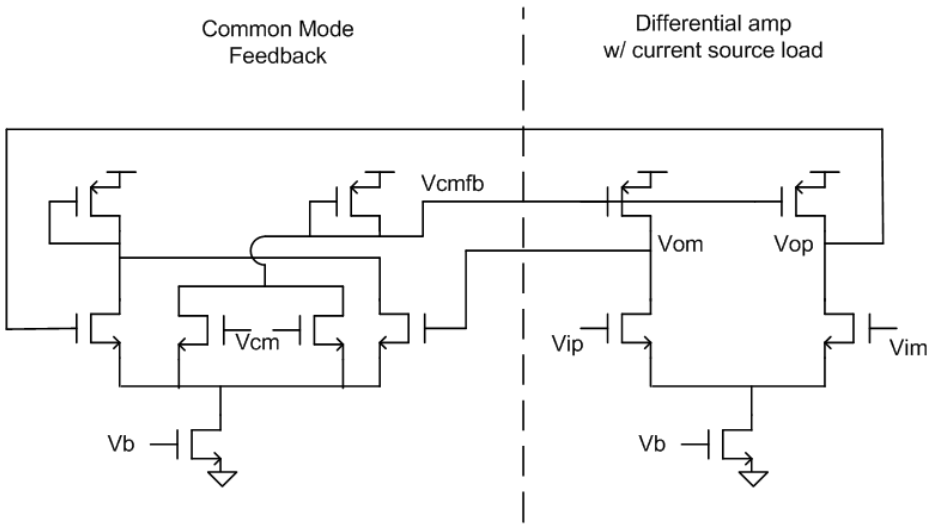


Figure 50 Transconductance Amplifier

The linearity of the OTA can limit the level of interferer that can be rejected. Specifically, the amplitude of the voltage at all stages cannot exceed the linear region of transconductance, otherwise distortion or harmonics will be introduced. For the OTA the region of linearity can be seen by examining the equation for the range.

$$V_{DI} = \sqrt{\frac{2 \cdot L \cdot I_{SS}}{KP_N \cdot W}} \tag{57}$$

It is clear the region depends upon the design parameters of differential device sizing and bias current. For Iss=1 μA and other design variables set appropriately, we can get V_{DI}= 0.1V as seen in Fig 51.

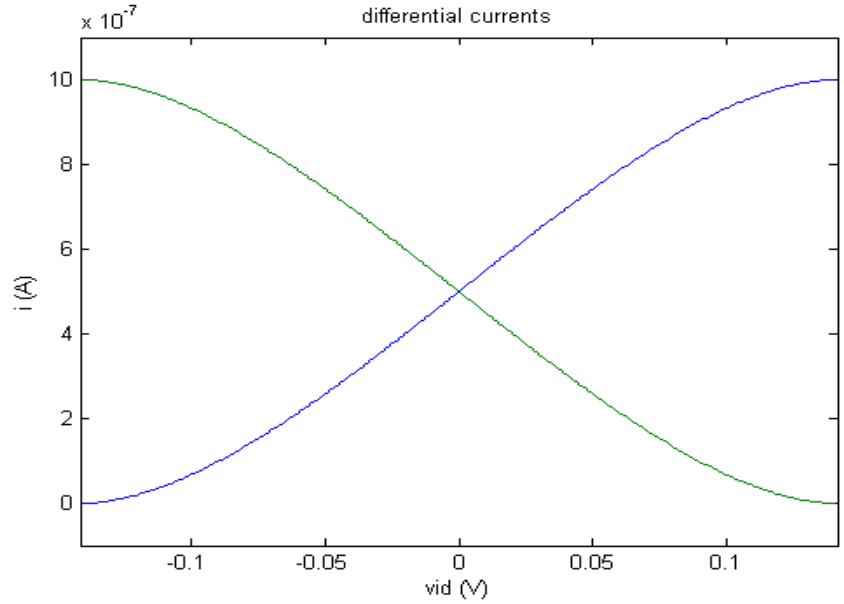


Figure 51 OTA Linearity

Using square-law saturation equations the two currents are

$$I_{D1} = \frac{I}{2} + \left(\frac{I}{V_{OV}}\right)\left(\frac{V_{ID}}{2}\right)\left(\sqrt{1 - \left(\frac{V_{ID}/2}{V_{OV}}\right)^2}\right) \quad (58)$$

$$I_{D2} = \frac{I}{2} - \left(\frac{I}{V_{OV}}\right)\left(\frac{V_{ID}}{2}\right)\left(\sqrt{1 - \left(\frac{V_{ID}/2}{V_{OV}}\right)^2}\right) \quad (59)$$

The resulting transconductance (g_m) is as follows

$$g_{m,D} = \sqrt{2I_D\left(W_D/L_D\right)K_N} \tag{60}$$

Often in differential amplifiers the transconductance (g_m) is maximized to maximize the voltage gain ($gm*R_{EQ}$) for a given bias current. Used alone this approach will have the effect of minimizing the acceptable input signal amplitude, before saturation occurs. A common method to extend the linearity is by including some form of feedback. This issue has been extensively studied and many forms of linearized OTAs have been published [38][40]. The cost paid for the benefit of greater linearity is higher circuit complexity and power.

In many cases these solutions are applied and the cost justifies the benefit. In contrast if the linearity range does not need to be very large and some small variation (weak non-linearities) in the transconductance can be tolerated both circuit complexity and power can be saved.

Through use of atypically small W/L ratios for the differential amplifier, and supplying very small input amplitudes it is seen the common differential amplifier OTA works wonderfully. The single ended amplitude supported is shown Fig 52. The signal strength supplied can be referenced through Fig 17, in the Link Budget Analysis discussion.

case	W (μm)	L (μm)	I_{SS} (μm)	KP_N ($\mu A/V^2$)		V_{DI} (V)	Comments
A	6.0	0.6	1.0	83		0.049	10/1 W:L ratio used for high gm
B	0.6	0.6	1.0	83		0.155	Ratio needed to extend linearity
C	3	3	1.0	83		0.155	Same ratio as "B", but increasing area to help minimize flicker noise

Figure 52 OTA Differential voltage range

Measurements using sizing for case C outlined are shown in Fig 53.

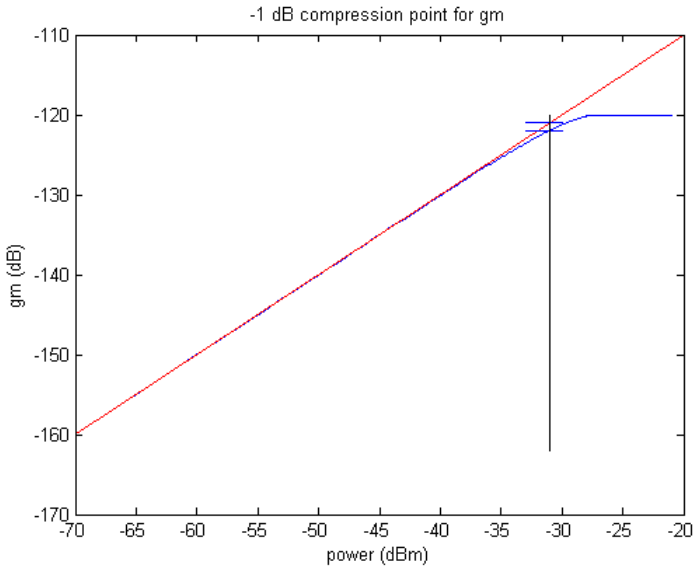


Figure 53 Gm-C Filter : OTA linearity

The overall filter linearity is more complicated than that of a single stage. In fact, each stage in the filter will have a different linearity depending upon the topology, loading, and frequency. The filter will always have a loaded OTA, which will be either a lossy or IDEALLY lossless integrating structure, as seen in Fig 54.

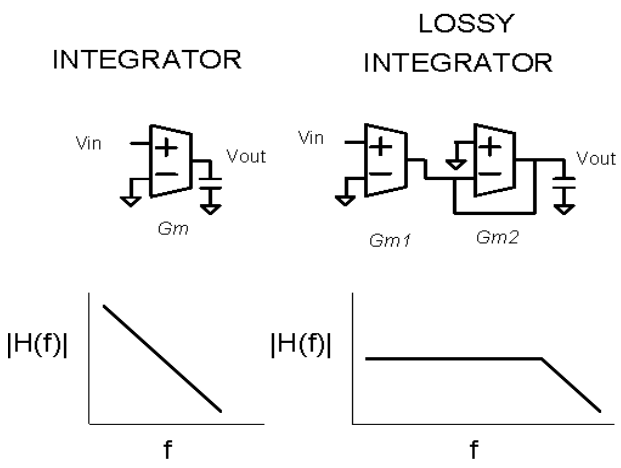


Figure 54 Integrator structures

Of course the higher the gain at the frequency evaluated, the lower the linearity, given that the linearity limit of the inherent structure has a constant range.

With just three devices between the rails the circuit has both good linearity and tunability [42][43][44].

Tunability is needed because due to process drift the frequency response could be off-center, degrading overall system performance considerably. The tunability is done by adjusting the transconductance of the differential pairs through adjusting the tail current source. In addition, adjustment to the common mode range for maximum can be done through this input (VCM).

The other parts needed to build these filters are the capacitors. On-chip capacitors are implemented in one of two ways, through MOSFET structures operated in the

inversion region, or through metal-oxide-metal structures. The MOSFET approach is part of the standard CMOS processing, but suffers from device linearity limits. Metal-oxide-metal structures do not introduce linearity problems due to voltage and current biasing, but require a process extension to make thinner oxide between the two metals to implement reasonably sized devices. Another key advantage of metal capacitors is both terminals can be connected to any voltage, while MOSFET capacitors require one terminal to be AC grounded.

In order to manage process drift, adjustment to the capacitance in the design is needed in addition to transconductance tuning. A typical method of providing this tuning is through selectable capacitor banks built up from a unit cell.

The conceptual diagram for capacitor tuning is shown in Fig 55.

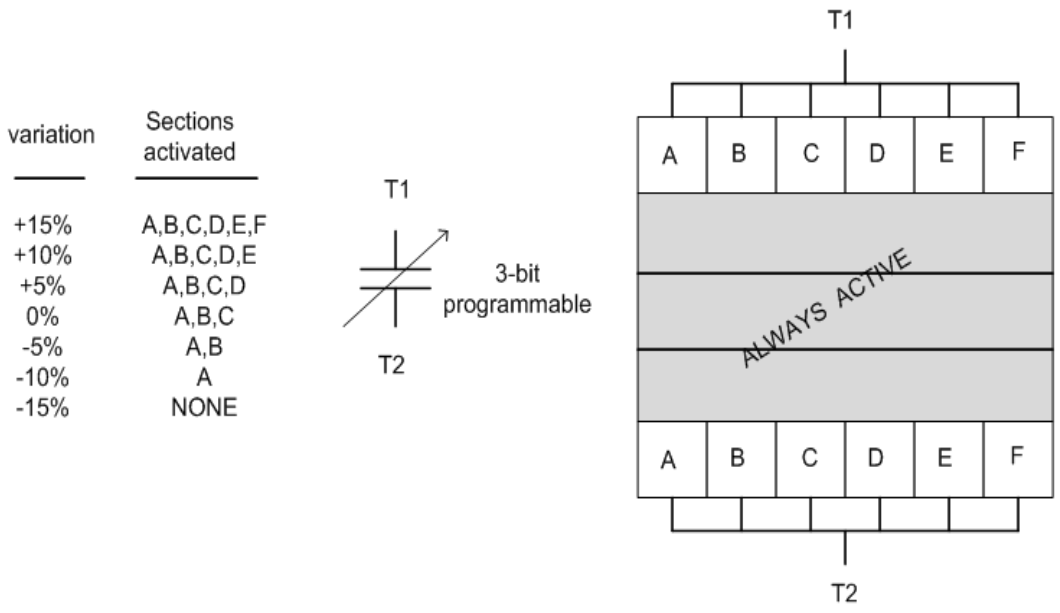


Figure 55 Capacitor Bank

The unit cell used in this design is 50 fF in capacity.

A more detailed example view of the variable capacitor implementation is shown in Fig 56.

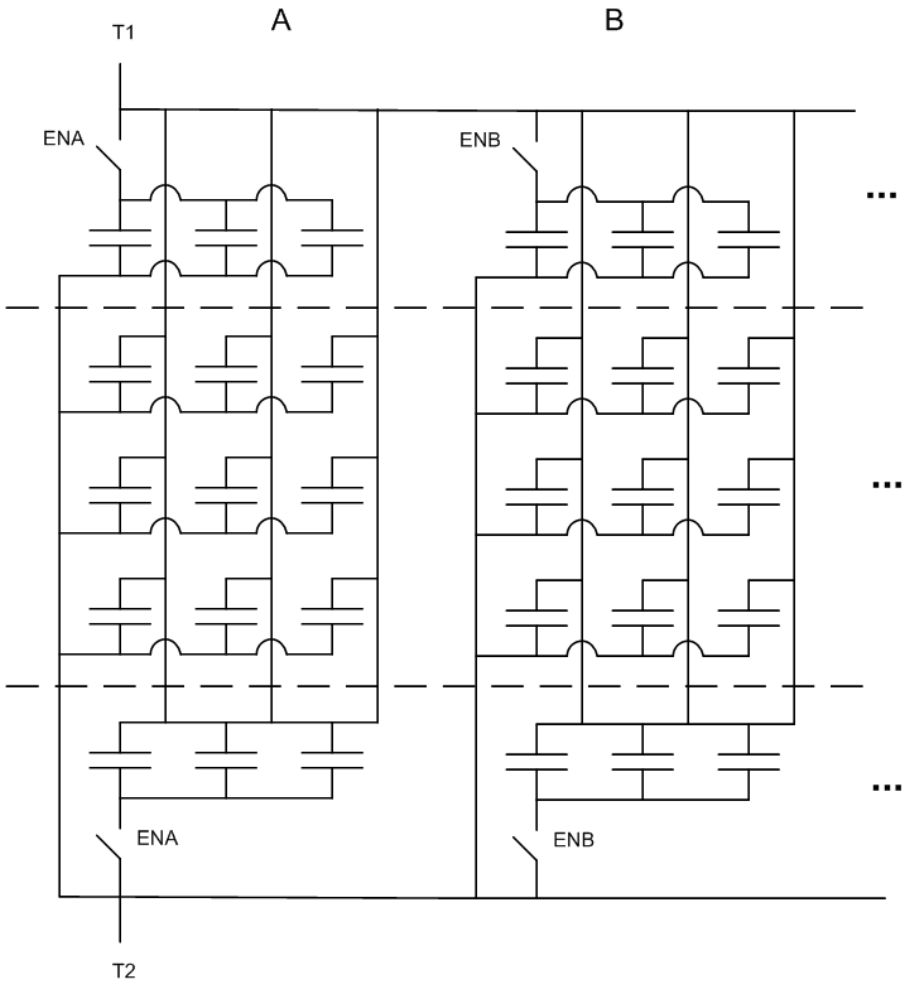


Figure 56 Detailed capacitor bank implementation example

Implementation of the transfer function depends both upon the transconductance values and capacitors. Because the different capacitor sizes are easier to change

than fixed transconductor cells, the different transconductor cells are minimized. In the limit of just one transconductor cell the dynamic range of capacitor values needed is the greatest. In order to both save area and get better matching, though more closely placed unit cells, 3 transconductor cells were used in the implementation. Even after this optimization there was still a great spread in the capacitance, but this was reduced through placing originally larger capacitors across the differential outputs (halving them), and doubling the smaller capacitors by having two of them in series.

A summary of all filter stages are shown in Fig 57. The nominal gm used is 6 μA/V.

Level	1	2	3	4	5
Filter type	BPF	BPF	LPF	NOTCH	NOTCH
Q	2.71	2.71	2	1	1
f0 (Khz)	71	106	114	150	216
C1 (pF)	10.3	6.9	8.4	6.4	5.2
C2 (pF)	17.6	11.8	8.4	6.4	5.2

Figure 57 Filter Subcomponents Summary

The 3 different second order filter topologies cascaded to build the design are in Fig 58. Each stage uses 4 OTAs and 2 capacitor structures, and is shown in single-ended form for illustration ease, although they are implemented differentially.

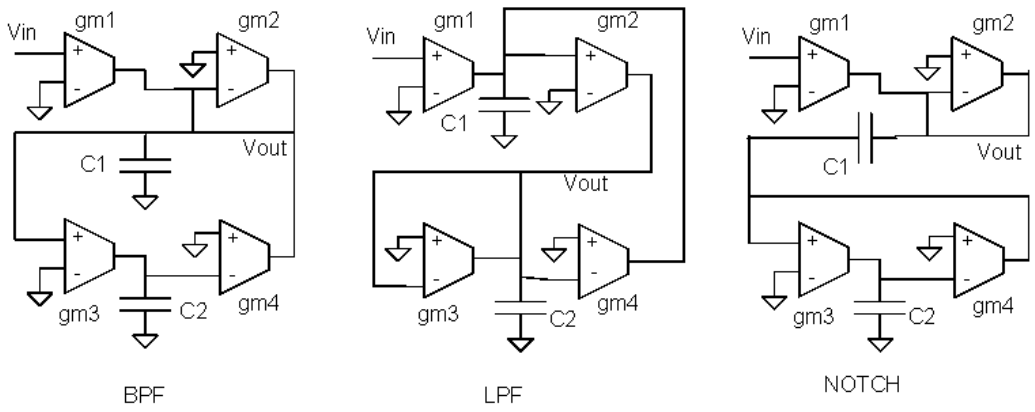


Figure 58 2nd order Gm-C building blocks for composite BPF

The transfer functions are below

$$H_{BP}(s) = \frac{sg_{m1}C_2}{s^2C_1C_2+sg_{m2}C_2+g_{m3}g_{m4}} \quad (61)$$

$$H_{LP}(s) = \frac{g_{m1}g_{m2}}{s^2C_1C_2+sC_1g_{m3}+g_{m2}g_{m4}} \quad (62)$$

$$H_{NOTCH}(s) = \frac{g_{m1}(s^2C_1C_2+g_{m3}g_{m4})}{s^2C_1C_2g_{m2}+sC_1g_{m3}g_{m4}+g_{m2}g_{m3}g_{m4}} \quad (63)$$

These filters work well for baseband, because a very low current helps to get the lower Gm. For the desired frequency response this gives the benefit of allowing the on-chip capacitances to decrease. One concern in sizing is making sure that the slew rate is large enough to support the largest signal amplitude.

For example on an OTA node the biasing current must be able to support the slew rate which is

$$SR_{provided} \equiv \frac{dv(t)}{dt} = \frac{I}{C} \quad (64)$$

The maximum slope a given sinusoid of v(t) = A*sin(wt) will be

$$SR_{required} \equiv \frac{dv(t)}{dt} = A * w * \cos(w * t) \quad (65)$$

$$SR_{provided} \geq SR_{required} \quad (66)$$

Plugging in some numbers used in the design we can quickly verify where we stand

First, the OTAs have a bias current of 1.5 uA, and the largest capacitance driven is 8 pF. Second, the maximum frequency of significant magnitude will be 150 Khz, giving w =2*π*(150E3) = 942E3 rad/s .

Solving for A, and plugging in numbers we get

$$A \leq \left(\frac{I}{C}\right)\left(\frac{1}{w}\right) = 0.212V \quad (67)$$

That is, the largest power before we see slew rate limiting is

$$20log_{10}(0.212) = -13.5\ dB = 16.5\ dBm \quad (68)$$

One other concern is that for really low currents the resulting Gm's will be very small, and hence possibly the output impedance could approach that value. Fortunately, as the current decreases the output impedance increases (because of the Early effect $r_o = |V_A|/I_D$), and the parasitic resistance does not affect the

design process. In order to keep the output resistance as large as possible, the length of all devices are much larger than minimum.

Total capacitance used for each composite filter is 76.8 pF with 80uA consumed total by the 40 OTAs.

LIMITER

The goal of this block is to amplify the output of the BPF to a level where it can feed to a 1-bit comparator with hysteresis stage for rejection of noise, as shown in Fig 59.

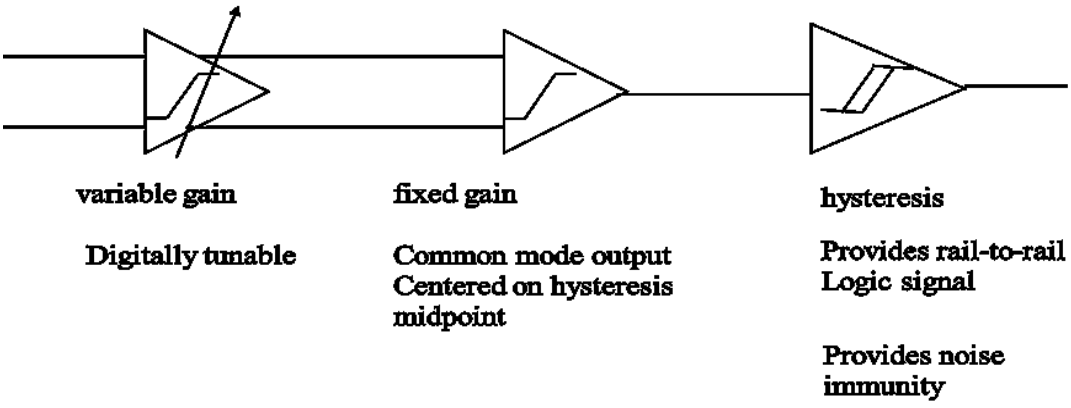


Figure 59 Limiter System

These linear gain stages are a cascaded sequences of differential amplifiers with diode connected loads, similar to that used for the variable gain amplifier block. Programmability is included that could adjust the gain for process variation, or signal strengths outside of the nominal range. The digital selection enables differently sized mosfet diodes to be switched in for various gains.

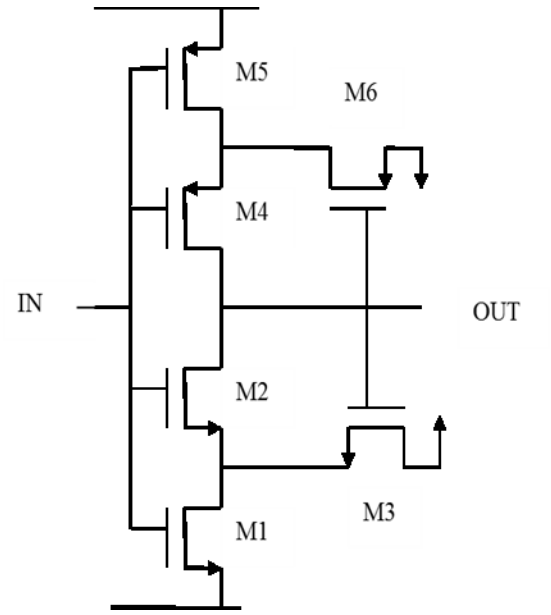


Figure 60 Hysteresis Circuit

The non-linear hysteresis circuit is based upon the 6-transistor implementation shown in Fig 60 [37]. The circuit looks like a CMOS inverter with state dependent feedback. As with all hysteresis circuits the trip points are set through feedback from the output. Conduction to the output requires Vin – Vs > V_{TH} for device M2, and manipulation of node Vs is key to deactivating and activating the hysteresis. When the output is low, M3 is in cutoff. Once the threshold voltage is exceeded on the gates of devices M1, M2 a current path to the output is established. In contrast when the output is high a higher input voltage is needed in order to get a current path to the output, because the feedback transistor M3 acts as a source follower along with the input current source M1. The source node Vs of M3 (and also M2) is set by the source follower drop from the high logic

voltage level. Device M2 stays in cutoff longer because Vs is higher in this configuration.

Symmetry in operation can be seen for the p-channel devices (M4-M6).

Generally the output will shut off one of the hysteresis switch biasing points, either between M1/M3 or M5/M6. In the case of output low M3 will be in cutoff, while output high puts M6 in cutoff. Therefore a current path for biasing one of the source follower circuits will occur, while the opposite leg will be connected to its respective power rail. The dynamic power dissipation of the circuit is similar to a standard CMOS inverter, but in addition a static power dissipation will be due to the source follower hysteresis.

Minimization of power in this circuit involves making the rail current source devices (M1/M5) as weak as possible, while keeping output node rise/fall times acceptable. Weakening these devices will keep the hysteresis bias current down, but will make the output transition slew rate limited.

If on average V_{in} is 0.9 (the biasing voltage) and this corresponds to 20 μA bias and the load of a driver is estimated at 100 fF, the slew rate is found

$$\frac{dV}{dT} = \frac{I}{C} = \frac{20\times10^{-6}A}{100\times10^{-15}F} = 0.2 \times 10^{9} V/s \qquad (69)$$

For the 1.8V voltage change required, the total time estimated is 9 ns. Although this is a rather slow transition time for this technology with typical times of

400ps, because the input signal is of 100 KHz frequency with a period of 10 μs, this larger transition time is not a problem.

The two devices connected to the output (M2 and M4) are sized as much stronger devices. This large aspect ratio is needed so that the voltage required to take the device out of cutoff and to large current carry capacity quickly is minimized.

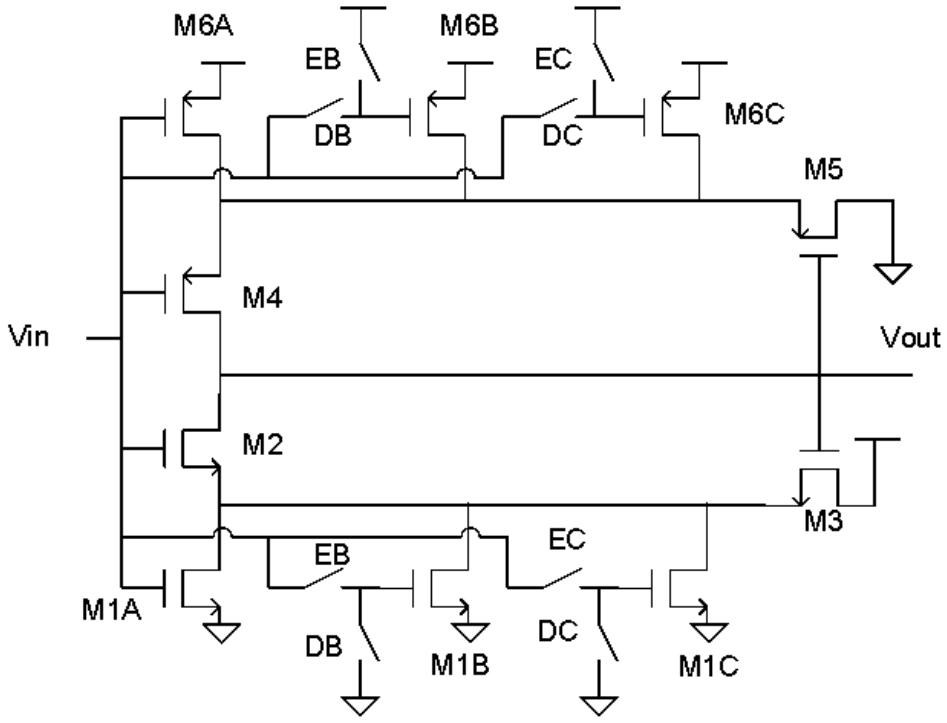


Figure 61 Programmable hysteresis

A small adjustment for hysteresis programmability is included in Fig 61. The switches are shown generically here, although CMOS t-gate switches should be used for inputs connecting to the Vin signal, while the pull-up/pull-down switches

can be single device switches of p-mos and n-mos devices respectively. The programmability modifies the trip points through addition of parallel current sources, increasing the drop across the source follower devices (M3,M6), as given by the standard square law saturation equation.

$$V_{GS} = \sqrt{\frac{2I}{K'(W/L)}} + V_{TH} \tag{70}$$

There are three settings for the hysteresis as shown in Fig 62. The signal names are two characters long. The first letter corresponds to enable or disable, and the values are always complementary. The second letter corresponds to the additional current source enable, for example “B” or “C”.

Hysteresis	EB	DB	EC	DC	Comments
0.05V	H	L	H	L	Minimum hysteresis
0.10V	H	L	L	H	Average hysteresis
0.15V	L	H	L	H	Maximum hysteresis

Figure 62 Hysteresis control

The average transfer function hysteresis transfer function is shown in Fig 63.

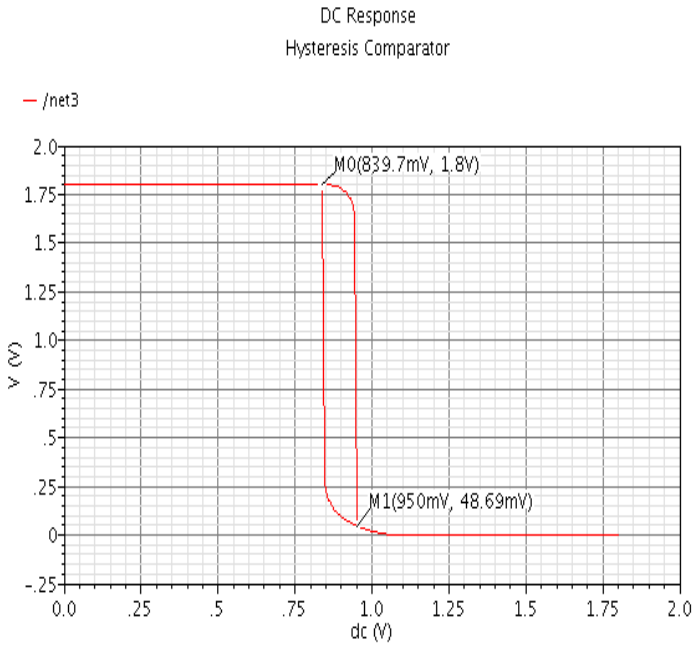


Figure 63 Hysteresis Transfer Function

RECEIVER SUBSYSTEM

Although circuit level simulation for BER evaluation is impractical, validation of system initialization and general operation is definitely possible. The Cadence Spectre simulator, which was also used for evaluating the subcircuits, was used at the top too [45]. Fastspice simulators can be used, but accuracy can often be a concern leading to questionable results. To rule out these uncertainties the same simulator with the same settings was used. Fig 64 shows the receive signal chain simulation for 120 µs, where the symbol value was changed once during the displayed time. The top signals are RF, moving to the mixers, bandpass filters, and eventually the limiter outputs.

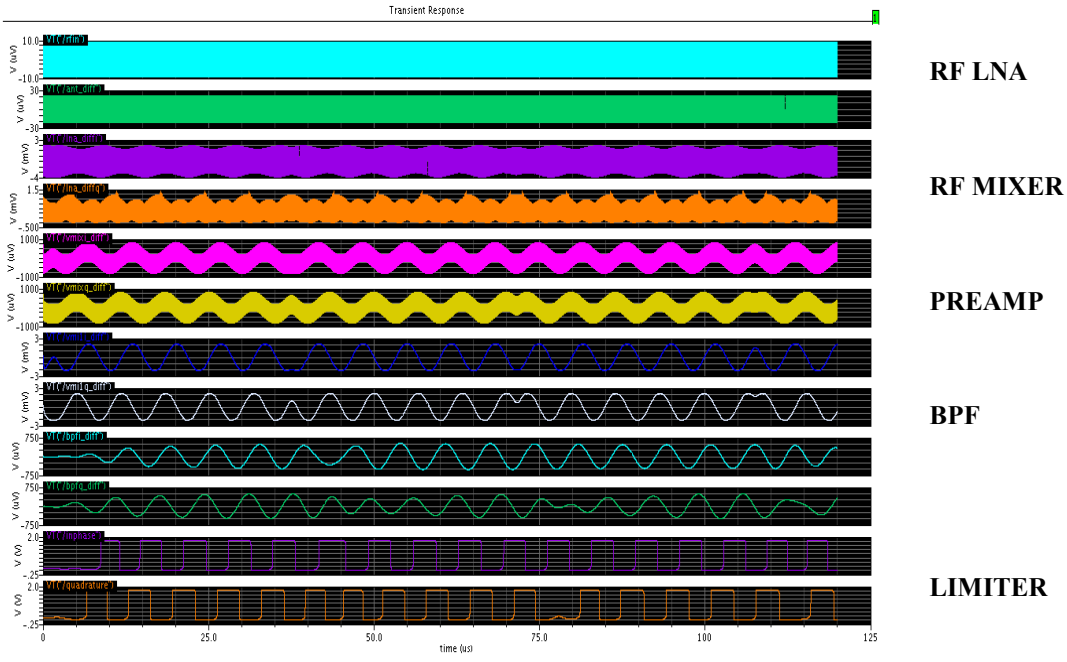


Figure 64 Receiver chain circuit level simulation

MEASUREMENTS and RESULTS

Sensitivity

The desired sensitivity of -96 dBm was not reached in the design. Fear that the noise would get amplified by too much, corrupting the input to the hysteresis, the level of gain in the receiver limiter stage was cut back from the original design. In retrospect, the design should have included observability at the I/Q hysteresis comparator inputs, to help verify more of the internal operation performance and limits.

Operation of the quadrature receiver chain was verified through application of the two symbols as continuous tones at the two offsets within the channel at higher power levels (-50 dBm), and monitoring the inphase and quadrature outputs. The lower frequency symbol in the channel would be indicated with the inphase leading the quadrature. The higher frequency would have the quadrature signal leading. Although for direct conversion the frequency offset should be symmetrical about DC, improper LO tuning or drift over time could give different down-converted frequencies for the two symbols. This would not only give different frequencies for the two symbol I/Q oscilloscope time domain plots, but also could attenuate the strength of one of the symbols, due to the symmetrical filtering. Therefore, care needed to be taken to place the LO frequency so that both offsets would be in the center of the bandpass filters for maximal gain.

A lower than desired sensitivity is seen on the oscilloscope through inphase and quadrature signals that no longer switch states. As the input power level is reduced the hysteresis gets a smaller and smaller amplitude signal. Eventually the input amplitude is smaller than the hysteresis resulting in no changes in limiter state. Because the symbol received is based upon comparison of phases, no such comparisons can be made for non-switching signals.

The receiver chain has some provisions for variable gain and limiter selection levels. The nominal gain should be 26-38dB with a tuning spread of 12 dB. The hysteresis supports a tuning range of 9dB. This gives the simulated range of sensitivity tuning of about 19 dB.

Measurements in the lab were made to find the minimum sensitivity, but precise failing points could not always be found. In the presence of a very strong input signal the inphase and quadrature triggered and the received symbol was clear. As the input signal weakened power near the failing point, the triggering was less regular and the pulses seen for the received symbols were not always happening. Finally, upon hard failure no pulses at all were seen on the inphase or quadrature. The range from clean signals to hard failure was as small as 1 dB up to a 3dB range.

For purpose of documenting measured sensitivity the lowest signal power where consistent, cleanly triggered waveforms for inphase and quadrature were seen on the oscilloscope was used. The lowest gain setting and largest hysteresis gave -53

dB sensitivity, while the largest gain with lowest hysteresis yielded -66 dB sensitivity. The difference in sensitivity tuning from simulation (19 dB) to measurement (13 dB) is accounted for by process variation.

Blocker/Interferer immunity

Out of channel rejection could be measured a couple ways.

First, just rejection of a continuous tone in an adjacent channel could be checked. The level out of channel filter rejection should be such that the output levels are comparable to the system noise. The filter stopband attenuation and linearity of the transconductor sets the limits of out-of-band rejection.

Second, blocking characteristics could also be measured by applying two tones: in-band and interferer. The in-band tone could be as low as the sensitivity level, up to the -1 dB level as desired. The interferer would typically start at the same level, and be increased until it causes failure to receive the intended signal.

Measurements are done for different interferer frequency offsets, f_i. The offsets would correspond to the different adjacent channel boundaries. In the MICS standard, the channel size is 300 KHz, so multiples of this size should be used.

Failure to provide rejection would be clear because the output frequency of the quadrature and inphase signals would correspond to the frequency f_i.

The lower sensitivity coupled with an unchanged linearity on the transconductors along with no visibility before the limiting operation makes checking interferer rejection very difficult. There is a small linear range for a detectable signal before the non-linearities of the transconductors of the active filters begin to cause distortion.

For measurements gain and hysteresis were set for maximum sensitivity. Applying intended 100 Khz offset signal of -65 dBm and using a variable frequency (f_i) and amplitude continuous wave interferer, Fig 65 was populated with measurements. To keep the number of measurements reasonable only 5dB resolution in power was used, and 5 frequency offsets were selected.

frequency amplitude	No Interferer	300 KHz	600 KHz	900 KHz	1200 KHz	3000 KHz
-40 dBm	OK	CORRUPT	CORRUPT	CORRUPT	CORRUPT	CORRUPT
-45 dBm	OK	CORRUPT	CORRUPT	CORRUPT	CORRUPT	OK
-50 dBm	OK	CORRUPT	CORRUPT	CORRUPT	CORRUPT	OK
-55 dBm	OK	OK	OK	OK	OK	OK
-60 dBm	OK	OK	OK	OK	OK	OK
-65 dBm	OK	OK	OK	OK	OK	OK
-70 dBm	DEAD	DEAD	DEAD	DEAD	DEAD	DEAD
-75 dBm	DEAD	DEAD	DEAD	DEAD	DEAD	DEAD

Figure 65 Interferer/blocker measurements

With no interferer the system has the earlier noted sensitivity limitation. The table indicates where addition of the interferer to adjacent channels begins to corrupt the decoded received signal.

The combination of the LNA LC tank along with the mixer cutoff frequency being in the MHz range contribute the extended range of interfere rejection for higher frequency offsets. Attenuation of these farther out-of-band interferers prior to signal propagation to the BPF filter accounts for this improvement in performance.

CONCLUSION

Evaluation of medical implant receiver requirements was done. A literature search of prior work in this area, as well as promising approaches to address the needs, were explored. The overall architecture and possible options common to receiver systems were examined. Having narrowed the general solution to use, convergence on the subblock specifications was done through theoretical analysis and system simulations. Knowing the specifications, circuit level explorations and choices were made. Design equations and SPICE simulations were run to evaluate the critical parameters.

Key among concerns for implantable receiver systems is high performance link reliability at an ultra-low power dissipation. The BFSK receiver implemented achieves a sensitivity of -97dBm at 75kbps for a power of 2mW, which is equivalent to 24nJ/b. It is capable of this high performance level even in the presence of much higher power interferers (30 dB greater).
A receiver summary is reported in Fig 66.

The MICS band transceiver IC has been fabricated in a bulk 0.18μm CMOS process and directly assembled on a FR4 test board. The active circuits occupy a total 3.8mm^2 die area and the die micrograph is shown in Fig. 67.

Architecture	Direct Conversion, Half duplex
Radio Frequency band	402-405 MHz
Bandwidth	< 300 KHz
Data rate	< 75 Kbps
Modulation	Non-coherent BFSK
BER	1E-3
NF	17.8 dB
IIP3	-34 dBm
IIP2	≥ 32 dBm
Sensitivity	-97 dBm
Phase noise requirement	-60 dBc/@ 100 KHz -84 dBc/hz @ 300 KHz
Blocker immunity	30 dB above minimum received signal. Limited by BPF design.
Multi-path fading mitigation	No receiver provisions. Re-position wearer of implant.
Dynamic Range extensions	No Automatic Gain Control Feedback system.
Total current	1100 μA

Figure 66 RX System summary

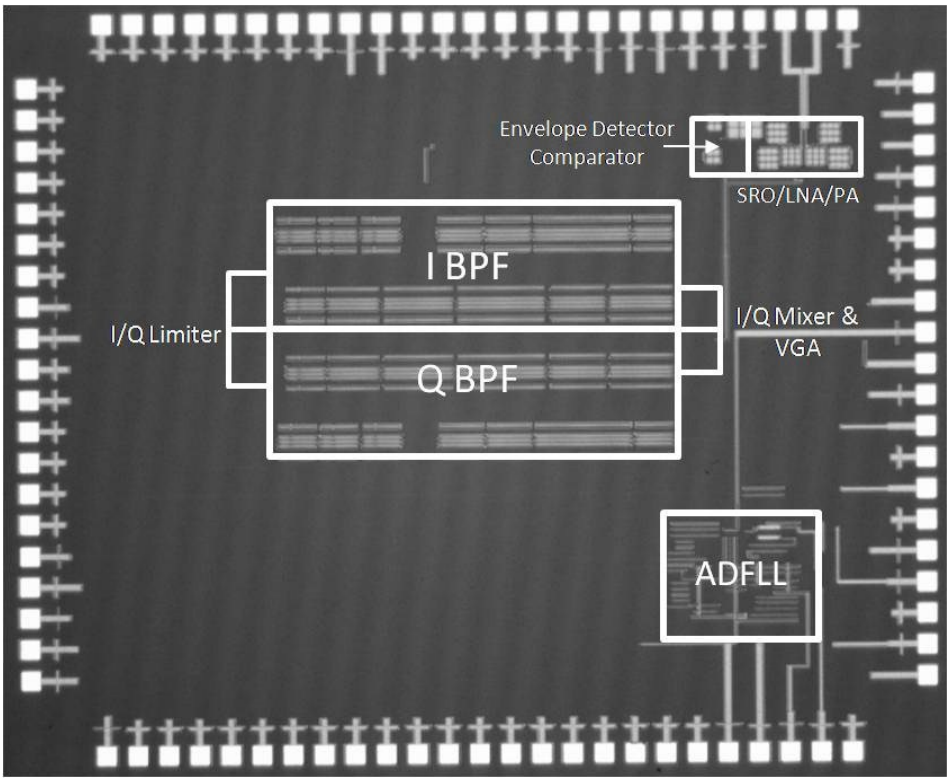


Figure 67 MICS Transceiver die photo

REFERENCES

[1] Hodgins, D.; Bertsch, A.; Post, N.; Frischholz, M.; Volckaerts, B.; Spensley, J.; Wasikiewicz, J.M.; Higgins, H.; von Stetten, F.; Kenney, L.; , "Healthy Aims: Developing New Medical Implants and Diagnostic Equipment," *Pervasive Computing, IEEE* , vol.7, no.1, pp.14-21, Jan.-March 2008

[2] Dario, P.; Carrozza, M.C.; , "Interfacing microsystems and biological systems," *Micro Machine and Human Science, 1995. MHS '95., Proceedings of the Sixth International Symposium on* , vol., no., pp.57-66, 4-6 Oct 1995

[3] Subbaiyan, D.; Pakula, L.S.; French, P.J.; Lange, J.F.; Jeekel, J.; Kleinrensink, G.J.; Sterenborg, H.J.C.; Robinson, D.J.; van Zaane, F.; Kaptein, J.G.; Tang, K.; Pandraud, G.; van Veen, J.; Draaier, M.; Vakalapoulos, K.A.; , "Post-operative wireless implants for monitoring, detection and treatment," *Solid-State Sensors, Actuators and Microsystems Conference (TRANSDUCERS), 2011 16th International* , vol., no., pp.2192-2195, 5-9 June 2011

[4] Szczesny, S.; Jetzki, S.; Leonhardt, S.; , "Review of Current Actuator Suitability for Use in Medical Implants," *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE* , vol., no., pp.5956-5959, Aug. 30 2006-Sept. 3 2006

[5] Doll, A.; Reimers, S.; Heinrichs, M.; Goldschmidtboeing, F.; Schrag, H.-J.; Hopt, U.T.; Woias, P.; , "A high performance bidirectional micropump for a novel artificial sphincter system," *Solid-State Sensors, Actuators and Microsystems, 2005. Digest of Technical Papers. TRANSDUCERS '05. The 13th International Conference on* , vol.1, no., pp. 188- 191 Vol. 1, 5-9 June 2005

[6] P. Dario; M.C. Carrozza; B. Allotta; E. Guglielmelli; , "Micromechatronics in medicine," *Mechatronics, IEEE/ASME Transactions on* , vol.1, no.2, pp.137-148, June 1996

[7] MICS Band Plan, *FCC Rules and Regulations*, Part 95, Jan. 2003

[8] Savci, H.S.; Sula, A.; Wang, Z.; Dogan, N.S.; Arvas, E.; , "MICS transceivers: regulatory standards and applications [medical implant communications service]," *SoutheastCon, 2005. Proceedings. IEEE* , vol., no., pp. 179- 182, 8-10 April 2005

[9] Bohorquez, J.L.; Chandrakasan, A.P.; Dawson, J.L.," A 350 μW CMOS MSK Transmitter and 400 μW OOK Super-Regenerative Receiver for Medical Implant Communications", IEEE Journal of Solid-State Circuits,vol. 44, Issue 4, Apr. 2009, pp. 1248-1259.

[10] Ching-Jen Tung; Yao-Hong Liu; Hui-Hsien Liu; Tsung-Hsien Lin; , "A 400-MHz super-regenerative receiver with digital calibration for capsule endoscope systems in 0.18-μm CMOS," *VLSI Design, Automation and Test, 2008. VLSI-DAT 2008. IEEE International Symposium on* , vol., no., pp.43-46, 23-25 April 2008

[11] Yao-Hong Liu; Ching-Jen Tung; Tsung-Hsien Lin; , "A low-power asymmetrical MICS wireless interface and transceiver design for medical imaging," *Biomedical Circuits and Systems Conference, 2006. BioCAS 2006. IEEE* , vol., no., pp.162-165, Nov. 29 2006-Dec. 1 2006

[12] Vouilloz, A.; Declercq, M.; Dehollain, C.; , "Selectivity and sensitivity performances of superregenerative receivers," *Circuits and Systems, 1998. ISCAS '98. Proceedings of the 1998 IEEE International Symposium on* , vol.4, no., pp.325-328 vol.4, 31 May-3 Jun 1998

[13] Jia-Yi Chen, Michael P. Flynn, John P. Hayes, "A fully Integrated Auto-Calibrated Super-Regenerative Receiver in 0.13um CMOS", IEEE Journal of Solid-State Circuits, vol. 42, no. 9, Sep. 2007, pp. 1976-1985.

[14] Tekin , A. , Yuce, M.R., Wentai, Liu, "A low power MICS band transceiver architecture for Implantable devices", IEEE Wireless and Microwave Technology Conference 2005, pp. 55-58.

[15] Feick, R. ; Rojas, O., "Modeling and Simulation of the Superregenerative receiver", IEEE Trans on Consumer Electronics, Vol 43, No. 2, May 1977, pp. 92-102.

[16] Cho, N.; Bae, J.; Yoo, H.-J , "A 10.8 mW Body Channel Communication/MICS Dual-Band Transceiver for a Unified Body Sensor Network Controller", IEEE Journal of Solid State Circuits, Vol 44, Issue 12, Dec 2009, pp. 3459-3468.

[17] Nezhad-Ahmadi, M.R.; Weale, G.; El-Agha, A.; Griesdorf, D.; Tumbush, G.; Hollinger, A.; Matthey, M.; Meiners, H.; Asgaran, S.; , "A 2mW 400MHz RF transceiver SoC in 0.18um CMOS technology for wireless medical applications," *Radio Frequency Integrated Circuits Symposium, 2008. RFIC 2008. IEEE* , vol., no., pp.285-288, June 17 2008-April 17 2008

[18] Bradley P.D., "An Ultra Low Power, High Performance Medical Implant Communication System (MICS) Transceiver for Implantable Devices", *IEEE BIOCAS 2006*, Nov 29 2006-Dec 1 2006, pp. 158-161.

[19] IEEE Std 802.15.4TM-2003-Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate WirelessPersonal Area Networks (LR-WPANs)

[20] R. Hartley, "Single-sideband modulator", U.S. Patent 1 666 206, Apr. 1928

[21] D.K. Weaver, “A third method of generation and detection of single sideband signals”, *Proc. IRE*, vol. 44, pp. 1703-1705, 1956

[22] J. Crols, M. Steyaert,”An Analog integrated Polyphase Filter for a High Performance Low-IF Receiver”, *Digest of Technical Papers, Symposium on VLSI Circuits*, June 1995.

[23] F. Behbahani, Y. Kishigami, J. Leete, A. Abidi,”CMOS Mixers and Polyphase Filters for Large Image Rejection”, *IEEE JSSC,* Vol 36, Issue 6, June 2001 pp. 873-887

[24] B. Razavi, ”Design Considerations for Direct-Conversion Receivers”, IEEE Trans. On Circuits and Systems II, Vol 44, No. 6, June 1997

[25] W. Namgoong, T. Meng, ”Direct-Conversion RF Receiver Design”, *IEEE Trans on Communications*, Vol 49, No. 3, Mar 2001

[26] Hsing-Hung Chen; Po-Chiun Huang; Chao-Kai Wen; Jiunn-Tsair Chen; , "Adaptive compensation of even-order distortion in direct conversion receivers," *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th* , vol.1, no., pp. 271- 274 Vol.1, 6-9 Oct. 2003

[27] J. Proakis, *Digital Communications*, McGraw-Hill, 2001

[28] J. F. Wilson et al, “A single-chip VHF and UHF receiver for radio paging”, *IEEE J. Solid-State Circuits*, vol 26, pp. 1944-1950, Dec. 1991.

[29] T. Rappaport, *Wireless Communications Principles and Practice*, Prentice Hall, 1991

[30] Johansson, A.J.; , "Performance of a radio link between a base station and a medical implant utilising the MICS standard," *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE* , vol.1, no., pp.2113-2116, 1-5 Sept. 2004

[31] H.T. Friis,”Noise Figure of Radio Receivers”, Proc. IRE, vol 32, pp. 419-422

[32] Q. Gu, *RF System Design of Transceivers for Wireless Communications*, Springer, 2010

[33] D. Pozar, *Microwave Engineering*, Wiley, 2005

[34] Advance Design Systems (ADS), Agilent Technologies, http://www.agilent.com

[35] T. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 2004

[36] W. Zhuo, X. Li, S Shekhar, S.H. K. Embabi, J. Pineda de Gyvez, D.J. Allstot, and E. Sanchez-Sinencio, “A capacitor cross-coupled common-gate low-noise amplifier”, IEEE Trans. On Circuits and Systems II, vol. 52, no. 12, Dec 2005, pp. 875-879

[37] R. Baker, *CMOS, Circuit Design, Layout, and Simulation*,Wiley, 2004

[38] B. Razavi, *RF Microelectronics*, Prentice Hall, 1998

[39] R. Pierret, *Semiconductor Device Fundamentals*, Addison-Wesley, 1996

[40] H. Song, *VLSI Analog Signal Processing Circuits*, Xlibris, 2009

[41] R. Schauman, M. Valkenburg , *Design of Analog Filters*,Oxford University Press, 2001

[42] H. Khorramabadi and P. R. Gray, "High-frequency CMOS continuous-time filters," *IEEE J. Solid-State Circuits,* vol.19, no.6, pp. 939-948, Dec. 1984

[43] C. Park, R. Schaumann,”Design of a 4-MHz Analog Integrated CMOS Transconductance-C Bandpass Filter”, IEEE JSSC, Vol. 23, No. 4, Aug 1988, pp. 987-995

[44] B. Nauta,”A CMOS Transconductance-C Filter Technique for Very High Frequencies”, IEEE JSSC, Vol 27, NO 2, Feb. 1992, pp. 142-153

[45] Spectre, Cadence, http://www.cadence.com/products/cic/spectre_circuit/pages/defaults.aspx