A RADIO FREQUENCY VARIABLE-GAIN AMPLIFIER

by

Pramote Piriypoksombut

Memorandum No. UCB/ERL M95/34

22 May 1995
A RADIO FREQUENCY VARIABLE-GAIN AMPLIFIER

by

Pramote Piriypoksombut

Memorandum No. UCB/ERL M95/34

22 May 1995

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720
Abstract

A variable-gain amplifier at 1GHz has been designed. Various topologies have been investigated with the emphasis on noise and distortion performance as the gain changes, and the variable feedback architecture is chosen. A pre-amp is not needed because the input impedance has been designed to match the source impedance throughout the gain range. The gain varies from 20dB at 2.3V control voltage to 6dB at 3.3V control voltage, and the circuit can operate with supply voltage 2.7V to 5.5V. The performance has been verified with SPICE simulations using full device and package models. The circuit layout has been completed using CADENCE and will be housed in a SO8 package.
Acknowledgments

Thank you Professor Meyer for your valuable insights and guidance throughout this project. I am also thankful for your motivation and accessibility any time problems arise. Thank you Professor Gray for your support. Many thanks to Bill Mack for your help on all the layout work at Phillips. Thank you Cynthia for your encouragement from the beginning of this work. Thank you Sam H. Sheng for providing me with the system level knowledge and how this project fits in a bigger work. Thank you Ranjit “Substrate” Gharpurey and Darrin F. Young for sharing your intuition and for your kindness. Thank you Keith “Inventor” Onodera, for taking so much time off from your research to answer my questions. Thanks you Jeff Ou for discussing many ideas with me. Thank you Andy Abo for pointing out that my loop gain for the 2-stage double feedback is “feeble.” Thank you Lapoe, Sekhar, Tony Stratakos, and Dave Lidsky for your comments. I also would like to thank my fellow graduate students Tom, Edwin Y. Chan, Monico, Randy, and Keng R.B. Fong for their support.
# Table of Contents

Chapter 1: Introduction 1

Chapter 2: Design Considerations 4

2.1 Frequency Response 4

2.2 Gain Flatness 5

2.3 Noise Performance 6

2.4 Distortion 7

2.5 Input & Output Impedance Matching 10

2.6 Power Supply and Temperature Variation 13

Chapter 3: Choice of Architecture 14

3.1 Varying Transconductance 14

3.2 The Multiplier 16

3.3 Fixed Gain LNA's with Variable Attenuator 17

3.4 Two-Stage Double Feedback 19

3.5 Single Stage Feedback 23

Chapter 4: Analysis of Variable Feedback Amplifier 24

4.1 Input Impedance 24

4.2 Frequency Response 27

4.3 Noise Figure 27

4.4 Distortion 29

4.5 Output Stage 29

Chapter 5: Practical Considerations 31

5.1 Process Variation 31
5.2 Temperature Variation 31
5.3 Interconnects 32
5.4 Bias Circuit 34

Chapter 6: Conclusion 36

Appendix: Circuit Schematic, Simulation Results, and Layout a-f
Chapter 1

Introduction

Variable gain amplifiers are widely used in RF communication systems. The need for such a function can be seen by considering a cellular telephone, a wireless indoor communication network, or a cable TV system. The signal level that reaches the receiver is not constant, but decreases with distance from the transmitter. To illustrate, the AMPS cellular telephone system has a cell radius of about 5 miles and the dynamic range of the signal received by the mobile unit is 80dB to 90dB. For the Infopad project environment which confines communication to within a room, the signal strength varies as much as 40dB. The receiver has to compensate for this variation, whether the system is digital or analog. A block diagram of such a receiver is shown in Figure 1.1.

Con conventionally, most of the gain control is done in the Intermediate Frequency (IF), rather than the Radio Frequency (RF) section. Among the reasons for doing this are that the RF amplifier can be optimized for low noise, and we can use a higher impedance level in the IF, giving a higher voltage gain and wider gain range at lower power. The current swing in IF is lower, hence less distortion is introduced.

However, a variable gain amplifier at RF is desirable for many reasons. In a system that
the IF section cannot provide the gain control range to accommodate the full dynamic range of the signal, for example the Infopad AGC chain after the signal is undersampled, the RF section needs to give additional gain control. An RF VGA can also prevent mixer overload by reducing gain when the incoming signal is large, i.e., when the receiver is close to the transmitter. Another advantage of RF VGA is that the gain of a fixed LNA deviates from the designed value due to process variation, a VGA allows the system designer to adjust gain for optimal system performance.

A cable TV transmission consists of segments of coaxial cable and repeaters. The function of the repeater is to periodically amplify and condition the signal which is attenuated by propagation loss and splitting. The signal level to a repeater depends on the distance between repeaters and how many times line splitting occurs. In order to ensure that the output signal level of the repeater equals to a designed value regardless of the input signal level, an automatic gain control circuit is required.

An Automatic Gain Control (AGC) circuit block diagram is shown in Figure 1.2.

This work concentrates on the gain block. Typically, a level detector already exists in the
IF section of a communication system receiver unit and the control voltage can be generated there. If the control loop is implemented at the carrier frequency, as in a cable repeater, additional amplification is needed to boost the signal to be large enough for the level detector. A detailed discussion of a RF control loop can be found in [5]. The outline of this work is as follows.

Chapter 2: Discussion of design considerations for a variable gain amplifier. The specifications at maximum gain are very similar to those of a fixed gain amplifier but due to the changing characteristic of a VGA, we need to consider its performance at different gain settings.

Chapter 3: An overview of possible architectures for the VGA. Advantages and disadvantages of each architecture are considered.

Chapter 4: A detailed examination of the cascade of 2 single stage shunt feedback.

Chapter 5: A discussion of practical considerations in the production line. These include process variation, package parasitics and temperature variation.

Appendix: Circuit schematics and SPICE simulation results using a full device and package model.
Chapter 2

Design considerations

The design considerations of a RF variable-gain amplifier are somewhat different from those of a fixed gain LNA. The VGA must have the required frequency response, low noise, low distortion, input and output impedance matching. We will discuss the criterion for "good" noise and distortion performance, since noise figure and distortion change with gain. Input impedance needs to match the source impedance throughout the gain range. The specification to be met is summarized below:

- **Gain**: 0-20dB at 1GHz
- **Usable bandwidth**: 128MHz (for Infopad CDMA application)
- **Gain flatness**: 1dB across 128MHz
- **NF**: 3dB at maximum gain
- **Input & Output Reflection coeff.**: <-10dB
- **IP3 at input**: -10dBm
- **Supply Voltage**: 2.65V - 5.5V
- **Power**: 30mW at 2.65V
- **Operational temperature**: -40° to 125° C

### 2.1 Frequency response

The requirement for maximum gain at 1GHz is 20dB. For a single stage, single pole amplifier with midband gain G and -3dB frequency BW, the theoretical limit of the frequency response is

$$G \times BW = f_T$$  \hspace{1cm} (2.1)
Where $f_T$ is the unity current gain frequency of the device. The peak $f_T$ for the process considered here is 12GHz for minimum size devices and lower for larger devices. Due to noise and power considerations, the size and bias current chosen for the input device degrade $f_T$ to about 7 GHz. Therefore a single stage amplifier cannot meet this frequency response specification. A gain of 20dB at 1GHz is achieved with a 2-stage amplifier which has an ultimate limit of

$$\sqrt{G \times BW} = f_T$$

assuming 2 poles at 45° and no loading effects. In a real implementation, the bandwidth is narrower than that predicted by (2.2) due to parasitic capacitances and loading.

As the gain decreases, the bandwidth of the variable gain amplifier can increase or decrease depending on the architecture chosen, as explained in detail below. In a feedback circuit, the loop gain must be designed so that it does not increase significantly, or instability will result.

2.2 Gain Flatness

The gain flatness requirement dictates that the amplifier can not operate in the roll-off region. If the amplifier is ideal with one pole at $p$ and $|p| \ll 1$GHz, the magnitude of the transfer function near 1GHz is described by

$$|A(j\omega)| = \left| \frac{A_0}{1 + (j\omega)/p} \right| = \frac{A_0 p}{\omega}$$

(2.3)

The band edges of the 128MHz band centered at 1GHz are at 936MHz and 1064MHz. The gain variation $\delta$ in decibels of an amplifier that operates in the gain roll-off region across the band is then

$$\delta = 20\log\frac{1064}{936} = 1.11\, dB$$

(2.4)

which is already more than 1dB. A real amplifier will have additional poles which worsen the
gain flatness even further. Therefore, to meet the gain flatness requirement, the poles of the amplifier should exceed 1GHz or at least be at 1GHz.

2.3 Noise Performance

The noise performance is measured by noise figure which indicates how much the amplifier degrades the signal to noise ratio. The noise figure is defined by

\[ NF (\text{numerical}) = \frac{\text{Input SNR}}{\text{Output SNR}} \]

\[ = \frac{\text{Total output noise power}}{\text{Output noise from source resistance}} \]  \hspace{1cm} (2.5)

\[ NF (dB) = 10 \log (NF (\text{numerical})) \]  \hspace{1cm} (2.6)

For a noiseless amplifier, all the output noise comes from the source resistance and the noise figure is 0 dB. For a variable-gain amplifier, NF changes with gain. Under normal operation, the gain is highest when the input signal is small. The input SNR is small, so NF must be lowest at high gain to add as little noise to the signal as possible. Lower gain is used for larger input signals. To maintain equal or better output SNR, NF must not increase faster than the decrease in gain. The criterion for NF degradation for intermediate frequency AGC is 1dB for every dB of gain reduction. But since this VGA is at the receiver front end, we also have to consider noise from later stages. The criterion now is that the overall noise figure, including noise from subsequent circuits, can degrade 1dB when gain is reduced by 1dB.

Consider a cascade of 2 matched-impedance stages shown in Fig 2.1. Both stages could be amplifiers, or the first one an amplifier and the second a mixer. Thermal noise and shot noise
contribute to the total amplifier noise. Let the total noise of the $i^{th}$ amplifier referred to its input be $4kT\text{noise}_i$ (assume that the input-referred current shot noise can be put in this form). Thermal noise from the source impedance is equal to $4kTR_s$.

\[ NF_{overall} = NF_1 + \frac{NF_2 - 1}{A_1} \quad (2.7) \]

For example, the first stage is a VGA with NF 3dB (2 numerical) at power gain 100, the second stage is a mixer with NF 15dB (31.6 numerical). The NF$_{overall}$ in this case is 3.63dB. Suppose at VGA power gain of 4, the NF is 6dB, the overall NF becomes 10.7dB, a degradation of 7dB. But the input signal level increased by 14dB, resulting in 7dB higher overall output SNR in the latter case. If NF of the VGA degrades too much with gain, the output SNR can worsen.

### 2.4 Distortion

In RF amplifier application, the distortion we are mainly concern with is the third order intermodulation product (IM$_3$). The second order distortion can be filtered out, but the third order
interference is in the signal band and cannot be separated from the desired signal. Specifically, if the output \( S_0 \) is related to the input \( S_i \) by

\[
S_0 = a_1(\omega_1) \circ S_i + a_2(\omega_1, \omega_2) \circ S_i^2 + a_3(\omega_1, \omega_2, \omega_3) \circ S_i^3 + \ldots \tag{2.8}
\]

where \( \circ \) denotes a complex operator on the magnitude and phase defined by

\[
X(\omega_1, \omega_2, \ldots) \circ e^{j(\omega_1 + \omega_2 + \ldots)} = |X(\omega_1, \omega_2, \ldots)| e^{j(\omega_1 + \omega_2 + \ldots)} + \angle X(\omega_1, \omega_2, \ldots) \tag{2.9}
\]

If \( S_i \) contains 2 tones at \( \omega_1 \) and \( \omega_2 \), the interaction in the third order term in equation (2.8) produces third order intermodulation components at \( 2\omega_1 - \omega_2 \) and \( 2\omega_2 - \omega_1 \) as illustrated in Fig 2.2.

![Figure 2.2: Frequency Location of IM3 products](image)

The magnitude of the third order intermodulation product is given by

\[
IM_3 = \frac{\text{Output Component amplitude at } 2\omega_1(2) - \omega_2(1)}{\text{Amplitude fundamental } (\omega_1(2))}
\]

\[
= \frac{3}{4} \left| \frac{a_3(\omega_1(2), \omega_1(2), -\omega_2(1))}{a_1(\omega_1(2), \omega_1(2), -\omega_2(1))} \right| v_i^2
\tag{2.10}
\]

The analytical expression for \( IM_3 \) at high frequency is complicated even for a single stage amplifier without feedback. For feedback circuits or circuits with more than one stage, the
calculation is intractable, but there are guidelines for low distortion design. For a common emitter gain stage, the high frequency distortion is optimized near a bias current that sets $C_{\pi}$ equal to $C_{je}$.[3] If the circuit has multiple gain stages, the bias current should be staggered. The second stage should have more bias current than the first, and so on, as the signal swing in a later stage is larger. Use of feedback, reactive or resistive emitter degeneration, lowers distortion.

The figure of merit for distortion performance in an RF front-end amplifier is the input third order intermodulation intercept point, $IM_{3_{\text{int}}}$ or $IP_3$. This is usually given in power units of dBm. For a system impedance $50\Omega$, the power in dBm is

$$P(\text{numerical}) = \frac{v_{\text{rms}}^2}{R} = \frac{v_{\text{peak}}^2}{100} \quad (2.11a)$$

$$P(\text{dBm}) = 10\log \left( \frac{P(\text{numerical})}{1\text{mW}} \right) \quad (2.11b)$$

$IP_3$ is the point where the extrapolations of small-signal fundamental output and small-signal third order intermodulation product meet as illustrated in Fig 2.3. In this example, the amplifier gain is assumed to be 10dB. Two sinusoidal inputs with frequency $\omega_1$ and $\omega_2$ each at -40dBm produce fundamental output power -30dBm and $IM_3$ -60dB. Larger inputs at -30dBm produce output power -20dBm and $IM_3$ -40dB. The input $IP_3$ from extrapolation is -10dBm.

For a VGA, the circuit characteristics including the input $IP_3$ change as gain varies. Depending on the architecture, the input $IP_3$ may increase as gain decreases to accommodate large signal, or it may decrease with gain. A desirable feature of the VGA is to have higher input $IP_3$ as gain decreases.
2.5 Input & Output Impedance Matching

As the wavelength of the signal under consideration becomes comparable to the size of circuit elements, lumped circuit models do not apply very well any more and the electromagnetic wave nature has to be considered to ensure proper signal propagation. The wavelength of the signal at 1GHz is on the same order of magnitude as the size of the line that carries it. For example, the signal on a board with dielectric constant of 3 has a wavelength 10 cm, while the length of the microstrip line from the antenna to the amplifier can be several centimeters or more and the
line can not be treated as an ideal connection, but must be considered a transmission line. This transmission line has to be properly terminated by the input impedance of the amplifier, otherwise wave reflection from discontinuity in propagation constant (impedance) occurs and signal power is not efficiently transferred to the amplifier.

This design assumes characteristic impedance of transmission line $Z_0$ to be 50 $\Omega$. The input reflection coefficient, $S_{11}$, is the ratio of reflected wave voltage to incident wave voltage at the input of the amplifier. $S_{11}$ of a circuit with input impedance $Z_{in}$ is measured by

$$S_{11} (\text{numerical}) = \frac{Z_{in} - Z_0}{Z_{in} + Z_0}$$

$$S_{11} (dB) = 20\log|S_{11} (\text{numerical})|$$

Ideally, $Z_{in}$=50$\Omega$ and all of the incidental wave is transmitted to the amplifier. In this case, $S_{11}(dB)$ approaches a large negative number. In most applications $S_{11} < -10$ dB, which corresponds to 90% of the incidental power being transmitted, is sufficient. The output reflection coefficient, $S_{22}$, which has similar definition as $S_{11}$ with $Z_{in}$ replaced by $Z_{out}$, has the same requirement.

Since the common emitter is the most widely used configuration of a low noise amplifier, and is part of many architectures discussed in Chapter 3, its input impedance is examined here. Consider the small signal equivalent circuit below.
At high frequencies, we need to model inductances associated with interconnects. $L_b$ includes bond wire, lead inductance and any external inductance added off chip. $L_e$ models the bond wire to the ground plane. Neglecting $C_\mu$, the input impedance is found to be

$$Z_{in} = j\omega L_b + \frac{1}{j\omega C_1} + r_b + j\omega L_e + \frac{1 + j\omega g_m L_e}{j\omega C_\pi}$$  \hspace{1cm} (2.13)$$

To match $Z_{in}$ to 50Ω, we set the real part to 50 and the imaginary part to zero.

$$r_b + \frac{g_m L_e}{C_\pi} = 50$$  \hspace{1cm} (2.14)$$

$$\omega L_b - \frac{1}{\omega C_1} + \omega L_e - \frac{1}{\omega C_\pi} = 0$$  \hspace{1cm} (2.15)$$

$L_e$ is typically 2-3 nH, $C_\pi$ for the device size and bias current used in front end LNA is about 2pF. Without $L_b$, the quantity on the left hand side of (2.15) is negative at 1GHz. With no cascode, the Miller effect across $C_\mu$ makes the effective input capacitance larger, and only a small $L_b$ is needed. The disadvantage of this is that the -3dB bandwidth of the amplifier is small and the circuit operates in the gain roll-off region. As discussed in 2.2, the gain flatness requirement cannot be met. If we improve the bandwidth by adding a cascode, there is no Miller effect across $C_\mu$, but
a large $L_b$ is needed for impedance matching.

2.6 Power Supply and Temperature variation

The amplifier should be flexible enough to operate over a wide range of power supply and temperature change without significant change in performance. The ambient temperature can be $-40^\circ C$ to $80^\circ C$, and the thermal coefficient of the package is about $1^\circ C/mW$. These considerations affect the design of the bias circuit. This is discussed in detail in Chapter 5.
Chapter 3

Choice of Architecture

This chapter discusses the advantages and disadvantages of possible variable gain amplifier architectures. These include the varying transconductance, the multiplier, variable attenuator, and variable feedback.

3.1 Varying transconductance

The simplest possible variable-gain amplifier is shown below.

\[ g_m R_1 = \frac{V_{\text{control}} - V_{be}}{R_E V_T} R_1 \]  

(V.1)

The main advantage is the simplicity. \( Q_1 \) can be cascoded for wider bandwidth, although this will require a larger matching inductor in series with the base of \( Q_1 \) as discussed in section 2.5. \( V_{\text{control}} \) can be designed to compensate for temperature variation. Input impedance is constant with changing gain and is given by (2.13). The circuit uses no feedback, so it is low noise.
The noise figure is approximately

\[ NF = 1 + \frac{r_b + \frac{1}{2g_m}}{R_s} + \frac{2q\left( I_B + \frac{I_C}{|\beta|^2}\right) R_s}{4kT} \]  \tag{3.2} \]

\( g_m \) decreases at low gain, resulting in slightly higher noise figure.

This circuit has two important drawbacks. First, the frequency response degrades at low gain. To decrease the gain, the bias current \( I_c \) is reduced, and the unity gain frequency \( f_T \) of \( Q_1 \) decreases with the bias current. The device size needs to be large due to noise consideration, and the power constraint limits the current to several mA. Since the device is large, the parasitic capacitance \( C_{je} \) dominates and \( f_T \) decreases with \( I_c \).

Second, the distortion performance degrades as gain decreases (input signal level increases). Intuitively, this is because the bias current, \( I_c \), decreases when gain decreases but the output swing is the same. The low frequency transfer function from \( v_i \) to \( i_c \) is

\[ i_c = a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + \ldots \]

\[ = \frac{qI_c}{kT} v_i + \frac{1}{2} \left( \frac{q}{kT} \right)^2 I_c v_i^2 + \frac{1}{6} \left( \frac{q}{kT} \right)^3 I_c^3 v_i^3 + \ldots \]  \tag{3.3} \]

The expression for third order intermodulation product in terms of magnitude fundamental of output is

\[ IM_3 = \frac{3a_3}{4a_1} \left( \frac{i^2_{fund}}{i_c} \right) \propto \frac{i^2_{fund}}{I_c} \]  \tag{3.4} \]

When the input signal level increases, \( V_{control} \) decreases the bias current for constant output signal level. As can be seen from the above equation, \( IM_3 \) increases at low gain (large input). The inter-
cept point at input remains the same because IM3 in terms of input magnitude does not change with gain. The desired behavior is that input intercept point increases as gain decreases. Therefore this architecture is not chosen.

3.2 The Multiplier Circuit

![Multiplier Circuit Diagram]

This VGA configuration is the same as a multiplier or a mixer. Q1 provides the transconductance and input impedance matching. Q2-Q3 forms an emitter coupled pair and Vgc determines how much signal current in the collector of Q1 goes through Q3 to the output. In high gain mode, Vgc is high, and Q3 acts as a common base stage with a current gain of 1. To attenuate, Vgc is lowered, and only a portion of the signal passes up through Q3. This circuit is most often used as a 2-state variable gain amplifier. The transfer function from v1 to ic3 is given by

\[
i_{c3} = \frac{i_{c1}}{1 + e^{-v_t/v_T}}
\]  \hspace{1cm} (3.5a)

\[
i_{c1} = \frac{qI_{c1}}{kT}v_i + \frac{1}{2}\frac{q}{kT}I_{c1}v_i^2 + \frac{1}{6}\frac{q}{kT}I_{c1}v_i^3 + \ldots
\]  \hspace{1cm} (3.5b)

Figure 3.2: Multiplier as a VGA
Ideally, the transfer function from $i_{c1}$ to $i_{c3}$ is linear and only $Q_1$ contributes to distortion. In real devices, the base resistance and finite $\beta$ of the emitter coupled pair add distortion to the overall transfer function from $v_i$ to $i_{c3}$. The bias current in $Q_1$ does not change with gain, so the distortion due to $Q_1$ is fixed. At a lower gain, distortion performance worsens because the bias current in $Q_3$ is lower at low gain as the output swing remains the same.

No feedback is employed in the multiplier, giving very low noise figure at high gain. Since this circuit has attenuation, the noise performance at low gain is poor. As mentioned in section 2.1, one gain stage can not meet the frequency response requirement, therefore another stage is needed. The noise referred back to the input of $Q_1$ from the second stage is amplified in attenuation mode.

3.3 Fixed Gain LNA’s with Variable Attenuator

This architecture is the classical VGA in IF design. It has one fixed gain stage, followed by a variable attenuator, and another fixed gain stage. If the attenuator is placed in front, the SNR of small input will degrade too much and the signal can be buried in the noise floor before signal conditioning. If the attenuator is at the output, large input signals will experience large distortion before they are attenuated.

![Figure 3.3: Fixed gain amplifiers with variable attenuator](image-url)
The attenuator is a Π resistor network which is modeled as shown in Fig 3.4.

\[ \frac{V_{out}}{V_{in}} = \frac{R_{2b}}{R_1 + R_{2b}} \]  

\( (3.6) \)

Figure 3.4: Equivalent Π resistor network

The output of the attenuator, assuming no loading from the second amplifier, is simply a voltage divider.

In high gain mode, \( M_1 \) is fully on, \( M_{2a,b} \) are off and there is no attenuation. To attenuate, \( V_{g1} \) is decreased to increase the equivalent resistance of \( M_1 \) and \( V_{g2} \) is increased. Note that \( R_{2a} \) does not appear in eq. (3.6) but it shunts the output impedance of the first amplifier and the attenuation range is wider than an attenuator with only \( R_1 \) and \( R_{2b} \).

This architecture has many advantages. The input and output amplifiers are designed independently from the gain variation, and therefore can be optimized separately for input and output impedance matching, gain flatness, and distortion. The bandwidth, \( S_{11}, S_{22} \) are constant with gain. The noise at maximum gain is low, and the NF increases about 1dB for each dB of gain reduction due to noise referred back from the second amplifier. In IF application, the output amplifier may be designed to drive capacitive load or a different resistance.

However, this circuit is unsuitable for use at the RF front end. The noise from subsequent circuits: the mixer and the IF amplifiers, has to be taken into account as well. At low gain, the noise from the second amplifier referred back to the input is amplified due to the center attenuator.
If the lowest gain achievable is 0dB, the attenuation is 20dB, and noise from the second stage is amplified by 10dB when referred to the input. Spice simulation shows that at 21dB gain, NF is 3.7dB. At 3dB gain, the NF degrades to 15.3dB. Suppose the mixer noise figure is 15dB, we can find the overall noise figure according to equation (2.7). Overall NF is 4.1dB at 21dB gain and degrades to 17dB at 3dB gain. The improvement of the SNR at the output of mixer is only 5dB, leaving little room for further degradation by filters and IF amplifiers.

Another disadvantage is the loss due to loading. The variable resistors are realized by using MOSFET's in triode region. In full gain mode, M_{2a,b} can be turned off and the variable attenuator will have no loading effect. Unfortunately, as M_2's are gradually turned on, they are very nonlinear when they just begin to conduct because the triode region is very narrow. To avoid this region of operation, M_{2a,b} need to be always on. Simulations show that V_{gs}-V_t should be kept above 0.2V for linearity. The attenuator then loads the first amplifier. The input of the second amplifier can be buffered by an emitter follower to prevent loading the attenuator, but the gain stage connected to the emitter follower will reduce the input impedance. Therefore, there has to be excess gain to compensate for this loss.

The distortion characteristic is better than the varying transconductance or the multiplier. At larger signal level, the attenuator is activated, reducing the signal swing before gross distortion occurs.

3.4 Two-Stage Double Feedback

The three architectures above have a fixed front end which provides the advantage of a fixed input impedance. However, a VGA with a fixed front end has a short coming in that either the noise or distortion performance degrades more than desirable when the gain decreases. The distortion performance can be improved by making the circuit more linear at lower gain. Noise
The gain-bandwidth product is large because there are 2 gain stages inside the feedback loop. The circuit is more linear at low gain because the loop gain is higher and distortion is reduced as the loop gain increases.

![Figure 3.5: 2-stage double feedback circuit](image)

The gain calculated from current and voltage equations of the small signal model is

\[
\frac{v_o}{v_x} = \frac{\beta g_m (R_{e1} + R_{f2}) R_i'}{\beta g_m R_i' R_{e1} + R_{f2} (1 + g_m R_{e1})}
\]  
(3.7)

where

\[
\frac{1}{R_i'} = \frac{1}{R_i} + \frac{1}{R_{f2}}
\]  
(3.8)

If \(R_i >> R_{f2}\) this simplifies to a simple \(R_{f2}, R_{e1}\) voltage divider result

\[
\frac{v_o}{v_x} = \frac{R_{f2} + R_{e1}}{R_{e1}}
\]  
(3.9)
A large $R_j$ is desired because it makes the gain strongly dependent on $R_{f2}$. This enables $R_{f2}$ to control the gain. To achieve this, $V_o$ should be buffered by an emitter follower. The gain is made variable by making $R_{f2}$ variable. The input impedance varies as $R_{f2}$ changes and is equal to

$$Z_{in} = R_{f2} \cdot \frac{R_{e1}}{R_j} \cdot \frac{R_i}{R_{e2}}$$ \hspace{1cm} (3.10)

Impedance matching can be maintained by making $R_{f1}$ a variable resistor as well and keep the ratio $R_{f2}/R_{f1}$ constant. $R_{e1}$, $R_{e2}$, and $R_j$ can be fixed. The design equation for $Z_{in} = 50\Omega$ is

$$\frac{R_{f1}}{R_{f2}} = \frac{R_{e1}}{R_{e2}} \cdot \frac{R_i}{50}$$ \hspace{1cm} (3.11)

The impedance at the output depends on the output conductance of $Q_2$ and current source $I_2$, parasitic capacitance from $Q_2$ collector to substrate, and the input impedance of the emitter follower. These components vary because of process variations such as Early voltage, $\beta$, and resistor values. $R_j$ in (3.11) thus can vary over a wide range, and the input impedance can vary from the intended value.

The advantages of this feedback circuit are insensitivity to supply or temperature change, and better linearity. But as with any feedback circuit, it is noisier than the previous circuits. $R_{e1}$ adds a thermal voltage noise, and $R_{f1}$ adds a current noise to the input node which translates to a total voltage noise of

$$4kTR_{e1} + \frac{4kTR_s^2}{R_{f1}}$$

The total noise figure is approximately

$$NF = 1 + \frac{r_b + 1}{2g_m(Q_i)} + \frac{R_{e1}}{R_s} + \frac{2qI_B R_s}{4kT} + \frac{2qI_C R_s}{|\beta(j\omega)|^2 4kT} + \frac{R_s}{R_{f1}}$$ \hspace{1cm} (3.12)
This architecture is not practical for a variable gain amplifier because of the DC voltage constraints across $R_{f1}$ and $R_{f2}$. They are triode MOSFET's, denoted by $M_{f1}$ and $M_{f2}$, that need to have $V_{ds} = 0$. The current in $M_{f1}$ is the small base current of $Q_1$ and the condition $V_{ds}=0$ is easily met. However, this constraint is harder to meet for $M_{f2}$. If the connection is as shown in Fig. 3.6, the voltage at the emitter of $Q_1$ is a few tens of millivolts above ground, the collector of $Q_2$ is about $2V_{be}$ above ground if $V_{ce2}=0$ is assumed, and $M_{f2}$ can be in saturation region. To bias $M_{f2}$ in triode region, we can use an emitter follower buffer to shift the $Q_2$ collector voltage down by one $V_{be}$ and put a diode at the emitter of $Q_1$ to raise the voltage up by a $V_{be}$ to give $V_{ds}=0$, resulting in the circuit below.

![Figure 3.6: Modified 2-stage double feedback](image)

There are many interferences from the circuitry used to bias $V_{ds}$ of $M_{f2}$ to a small value. The diode introduces nonlinearity in the feedback path, the series resistance of the diode adds noise, and part of the signal is lost in $Q_3$ follower. From a power constraint, the current in $Q_3$ should be kept small, but this introduces signal loss from $1/gm^3-R_{f2}$ voltage divider. The feedback factor is now partly dependent on the bias circuitry, making $R_{f1}$ and $R_{f2}$ less effective in
changing the gain.

The signal swing and headroom are limited by the extra circuitry required. The collector of $Q_1$ is now $3V_{be}$ above ground, leaving only a small headroom for the current source $I_1$. The supply voltage can not go below $4V_{be}$ or about $3.3V$, so the circuit can not operate if the supply voltage drops to $2.65V$. Due to small headroom, a PNP current source with emitter degeneration which requires less headroom than a PMOS, but has the drawback of higher noise, must be used.

3.5 Single Stage Feedback

A single-stage feedback configuration has the same benefits of better distortion and noise performance at lower gain as the 2-stage circuit. But the implementation is much simpler, and there is no complication from the bias circuitry. The headroom is improved, allowing operation at lower supply voltage. Since there is no interference from the bias circuitry, changing the values of feedback resistors is more effective in changing the gain than in the 2-stage feedback circuit.

There is a minor disadvantage compared to the 2-stage circuit. As equations (2.1)-(2.2) demonstrate, the frequency response of a single-stage shunt feedback is worse than a 2-stage double feedback, but cascading two single-stage feedback can solve this problem. This is the architecture chosen and detailed analysis is provided next.
Chapter 4

Analysis of Variable Feedback Amplifier

The VGA with a variable feedback element is the chosen architecture. It has good noise and distortion performance across the gain range. At first observation, a pre-amplifier for input impedance matching seems necessary because the varying feedback element should change the input impedance of the variable gain circuit. But a pre-amp is undesirable since it increases power consumption and circuit complexity. Fortunately, more detailed analysis below shows that a pre-amp is not needed. Analysis of frequency response, distortion, and noise follows.

4.1 Input Impedance

The first hint on impedance matching comes from examining a local shunt feedback loaded with a current source and a buffer to the next stage. The high frequency small signal circuit shown below assumes that the impedance of the load is much higher the feedback.

![Local shunt feedback with high impedance load](image)

Figure 4.1: Local shunt feedback with high impedance load

To find $Z_{in}$ we put a voltage source $v_x$ and find the current flowing through it. Note that $R_f$ is in series with a current source. The value of the current $g_m v_1$ depends only on $g_m$ and voltage divider $r_b$, $C$, and $Z_e$. So $R_f$ should not affect the current drawn from $v_x$ and hence the input
impedance at all. Solving KCL at the emitter gives

\[ Z_{in} = \frac{1 + g_m Z_e + C_\pi (r_b + Z_e)s}{g_m + C_\pi s} \]  

(4.1)

which is independent of \( R_f \) as expected.

The voltage gain from \( v_x \) to the output is

\[ \frac{V_{out}}{V_x} = 1 - \frac{g_m R_f}{1 + g_m Z_e + C_\pi s (r_b + Z_e)} \]  

(4.2)

Equations (4.1) and (4.2) indicate that the gain can be changed by changing \( R_f \) while the input impedance remains constant. Unfortunately, the input impedance achieved this way is too low due to the following reasons. To minimize thermal noise from the base resistance, a large device (36x chosen) must be used. The bias current to maintain a high \( f_T \) for such a large transistor is high (2.5 mA chosen), resulting in a low input impedance. A series resistance or an L-C matching network is needed to match the impedance. If a series resistance is used, the noise performance will degrade and part of the signal power is dissipated in the resistor. An L-C network does not contribute to noise or signal loss, but the match is narrow-band and variation in the element values causes the input impedance to change significantly.

Another way to boost the input impedance is to reduce the current going into \( R_f \). This is achieved by putting the correct load at the collector. Note that \( v_1 \), the voltage across \( C_\pi \)

\[ v_1 = \frac{v_x}{1 + g_m Z_e + C_\pi s (r_b + Z_e)} \]  

(4.3)

is independent of \( R_f \) and load \( Z_l \). Therefore the current \( g_m v_1 \) is independent of \( R_f \) and \( Z_l \). If we let
$Z_i$ supply the right amount of current to $g_{mV_1}$, input impedance is matched.

The expression for input impedance for general $R_f$ and $Z_i$ is given below. The derivation involves no approximation.

$$Z_{in} = \left( \frac{1}{Z_i} + \frac{g_m}{1 + g_m Z_e + C_n s (r_b + Z_e)} \right)^{-1}$$

Little insight can be gained from the complicated expression above, although the ratio $R_f/Z_i$ provides some hint. Examining the circuit itself provides some intuition as to how the input impedance can be matched as $R_f$ is changing. If $Z_i$ is fixed, when $R_f$ decreases, the current in $R_f$ increases and the current from $Z_i$ decreases since the current $g_{mV_1}$ is fixed, resulting in lower impedance. If we want the current in $R_f$ to remain constant while $R_f$ is decreasing, $Z_i$ needs to decrease to provide the same current to $g_{mV_1}$. Implementation of $Z_i$ is another local shunt feedback stage, resulting in the circuit in fig. 4.2.

![Two stage local shunt feedback](image)

Figure 4.2: Two stage local shunt feedback

A SPICE simulation using a full device and package model shows that $S_{11}$ stays below $-10\,\text{dB}$ in the 128 MHz band centered at 1GHz throughout the gain range (see Appendix A).
4.2 Frequency Response

The unity gain frequency $f_T$ of the 36x device is about 7 GHz. A single stage amplifier can not achieve 20dB gain at 1GHz, but the theoretical gain-bandwidth product of the two stage amplifier is high enough to achieve that in practice. The gain under matched impedance condition is

$$A_v = \frac{1}{2} G_{m1} R_f = \frac{1}{2} G_{m2} R_f$$  \hspace{1cm} (4.5)$$

where $G_{m1} = \frac{g_{m1}}{1 + g_{m1} Z_{e1}} - \frac{1}{R_f}$ and $G_{m2} = \frac{g_{m2}}{1 + g_{m2} Z_{e2}} - \frac{1}{R_f}$

The gain is made variable by using triode MOSFET's as $R_{f1}$, $R_{f2}$ and controlling the gate voltages. The equivalent resistance, assuming a square law device and $V_{ds} = 0$, is

$$R_{eq} = \frac{1}{\mu C_{ox} \frac{W}{L} (V_{GS} - V_t)}$$  \hspace{1cm} (4.6)$$

Therefore, the gain as a function of control voltage is

$$A_v (V_{gc}) = \frac{1}{2} \left( \frac{g_{m1}}{1 + g_{m1} Z_{e1}} - \mu C_{ox} \left( \frac{W}{L} \right)_1 (V_{gc} - V_{be} - V_t) \right)$$
$$\mu C_{ox} \left( \frac{W}{L} \right)_2 (V_{gc} - V_{be} - V_t)$$  \hspace{1cm} (4.7)$$

It is worth noting that the gain is approximately inversely proportional to the control voltage. The gain range of this circuit is relatively small compared to that of a multiplier or an attenuator for this reason. The $1/V_{gc}$ characteristic is evident in the plot showing gain versus control voltage in Figure A1 in the Appendix.

4.3 Noise Figure

The major noise sources are the thermal noise from the base resistance and the feedback
resistor, collector and base current shot noise, and thermal noise from the current source \( M_1 \). The base resistance thermal noise is minimized by using a large device with \( r_b \) of only 11\( \Omega \). The collector current shot noise referred to input voltage noise is \( 1/2g_m \), thus should be minimized by running large bias current. However, power dissipation increases along with base current shot noise. SPICE simulation shows that noise figure does not improve significantly for bias current larger than 2.5 mA, which is the value chosen.

The current source \( M_1 \) is needed in place of a resistor due to power consideration as the supply voltage changes. A PMOS is used because it is less noisy than a PNP current source. The noise figure is approximately

\[
NF = 1 + r_b + \frac{1}{2g_m(Q_1)} + \frac{2qI_B R_s}{4kT} + \frac{2qI_C R_s}{|\beta(j\omega)|^2 4kT} + \frac{R_s}{R_{eq(M_p)}} + \frac{2}{3g_m(M_1) R_s g_m(Q_1)}^2
\]

(4.8)

The values used in this design are \( R_s \): 50 \( \Omega \), \( r_b \): 11 \( \Omega \), \( g_m(Q_1) \): 0.1 A/V, \( I_B \): 28\( \mu \)A, \( I_C \): 2.55 mA, \( |\beta| \): 1 GHz, \( R_{eq(M_f)} \): 340 \( \Omega \) at 20dB gain, \( g_m(M_1) \): 3.5mA/V. Substitution of these values gives

\[
NF = 1 + \frac{16}{50} + 0.027 + 0.05 + 0.147 + 0.0047 = 1.55 \text{ (numerical)}
\]

\[= 1.9dB\]

The noise from the second stage has a significant contribution because the collector of the first device is a low impedance node for the impedance matching requirement hence the voltage gain of the first device is low. SPICE simulation gives a total noise figure 2.6dB at 20dB gain.
4.4 **Distortion**

Distortion is dominated by the nonlinearity in the feedback resistors. As Figure 4.3 demonstrates, the linear range of a triode MOSFET is narrow at low $V_{gs}-V_t$ and widens as $V_{gs}$ increases, and the width of the linear region is approximately $V_{gs}-V_t$. To achieve lower gain, $R_f$ is reduced by raising the gate voltage, resulting in a larger linear range. Thus the linearity improves to accommodate larger input level at low gain, and the input $IP_3$ increases as gain decreases.

![Figure 4.3: $I_d$-$V_{ds}$ characteristic of a 60/1 NMOS](image)

4.5 **Output Stage**

The output impedance of the cascaded amplifier is a few hundred ohms, so an output stage for output impedance matching to a 50Ω load is needed. Either an emitter follower stage or a common emitter with emitter degeneration for linearity can be used. The real part of the output impedance of the emitter follower is low, and the overall output impedance is inductive due to the
β roll off at high frequency. Therefore, either a series resistance matching or an L-C matching network is needed. A series resistance introduces signal loss, and the gain prior to the output stage need to be higher than 20 dB to have an overall amplifier gain of 20 dB. Distortion is then higher because the signal swing at the internal nodes is larger. The L-C matching network alternative can be used, but the obtained output impedance is very sensitive to component variation and the match is narrow-band. In addition, the emitter follower still introduces loss due to the $1/g_m$ and $50\Omega$ impedance divider.

A common emitter with emitter degeneration allows unity gain and impedance matching simultaneously. Although the collector resistor needs to be 50 Ω for impedance matching, this configuration allows utilization of the gain of the device so that a unity gain can be achieved. The match is wide-band, and not very sensitive to component variations. This is the output stage chosen for this amplifier.
Chapter 5

Practical Considerations

The analysis in Chapter 4 captures most important aspects of the circuit. In actual fabrication, packaging, and placing the chip on a circuit board, there are many non-idealities that can alter the behavior of the VGA. There are deviations from nominal values of circuit components. The bond wires from the silicon to lead frames have inductance that affect the gain, input and output impedances. The metal line on board that brings supply voltage to the chip can present an inductive load to the internal supply node on chip and serves to couple signals in different stages to each other. Care must be taken to minimize the effects of these non-idealities on the circuit.

5.1 Process Variation

There is significant deviation from a nominal value in integrated-circuit fabrication. The absolute value of a resistor can be ±12% from nominal. As a result, the bias current, gain, and output impedance can vary. Doping concentrations of the emitter and the base can vary, resulting in deviation in the base transit time and base resistance. Longer transit time causes the bandwidth of the circuit to be lower than expected. A higher base resistance adds thermal noise directly at the input, causing a higher NF. The base-emitter and collector-substrate parasitic capacitances can be higher than nominal, degrading the frequency response further. To ensure proper production yield, we examined the performance at the worse possible case.

5.2 Temperature Variation

Since a PTAT current source is used in this circuit, the transconductance of the BJT’s is compensated against temperature variation. However, temperature change can still affect the gain through the temperature dependence of the equivalent resistance of the triode MOSFET’s. This can be seen by considering eq (4.6) repeated here for convenience:
$$R_{eq} = \frac{1}{\mu C_{ox} L (V_{GS} - V_t)}$$  \quad (5.1)

For a PTAT current source, the current increases linearly with absolute temperature. $V_{be}$ for a constant bias current decrease about 2 mV/C and slightly less for PTAT bias current. From SPICE simulation, the $V_{be}$ of the device size chosen changes from 0.885 V at -40°C to 0.617 V at 125°C, and $V_t$ decreases as temperature increases, thus the term $V_{GS} - V_t = V_{control} - V_{be} - V_t$ increases. The other temperature dependent parameter in (5.1) is $\mu$. Since the scattering mechanism is dominated by phonon scattering in this temperature range, the mobility $\mu$ decreases with increasing temperature. $\mu$ has a stronger temperature dependence than $(V_{gs} - V_t)$, so the overall effect is that for a constant gain control voltage, $R_{eq}$ and the gain increases with temperature.

### 5.3 Interconnects

The circuit fabricated on silicon needs to have ohmic contacts to the outside world for the input/output, power, and ground. The diagram below demonstrates an example of the interconnects inside and outside the chip.

![Interconnect Example](image)

Figure 5.1: Interconnect example

The interconnects affect the circuit because of the series inductance introduced. The inductance of a bond wire from silicon to the lead frame is about 1 nH/mm. At 1GHz, the imped-
The inductance from the external wire that bring supply voltage to the chip (Vcc wire in Fig. 5.1) needs to be considered. Ideally, the internal supply node (x) should be an ac ground, but with this inductance, it is not. Depending on where the external bypass capacitor which makes an ac ground (C_{ext} in Fig. 5.1) is placed, the inductance between the internal supply node (x) and the ac ground can be from a few nH to 50nH. This inductor allows a signal in a later stage to couple back to the previous stages, causing instability or distortion. The input and output impedances are also affected by the inductance. To minimize the effect of this large inductor, an on-chip passive bypass is employed. The ac current on the internal supply rail returns to ground via this low impedance bypass instead of through $L_{vcc}$, and the resulting voltage swing on the internal supply is greatly reduced. The placement of the on-chip bypass with respect to the complete circuit is shown in Fig. 5.2.

![Figure 5.2: Location of on-chip supply bypass](image)
The capacitor $C_{bp}$ and series resistance $R_{bp}$ form the passive bypass. The parasitic series resistance associated with the capacitor can be used to implement $R_{bp}$. $L_{vcc}$ is the inductor from internal supply to external ac ground. $L_{gnd1,2,3}$ are the bond wire inductance. The value of $C_{bp}$ is chosen such that around the operating frequency $\omega_0$, the impedance on the internal supply line is $R_{bp}$ instead of large inductive caused by $L_{vcc}$. This is achieved by setting

$$\omega_0 = \frac{1}{\sqrt{C_{bp}L_{gnd3}}}$$

(5.2)

The value of $R_{bp}$ so that the circuit exhibits no ringing to a step input is about 10Ω. Without the on-chip bypass, the impedance is $j\omega L_{vcc}$ which can be up to 300j. With the bypass, the impedance is $R_{bp}$ in parallel with this inductor. Since $R_{bp}$ is much lower, it dominates the impedance there.

5.4 Bias Circuit

The property of the bias current is chosen to meet the power supply voltage, and operational temperature specification. The circuit has to operate over a wide range of supply voltage (2.65 to 5.5 Volts) and temperature (-40°C to 125°C) without significant change in the characteristic. Supply independent current source is therefore required. The bias circuit must also have temperature compensation. The transconductance of a BJT is inversely proportional to temperature for a given collector current. If the collector current is made proportional to the temperature, then the transconductance is constant with temperature. A PTAT current source satisfies both of
the requirements.

\[ I_c = \text{const} \times T \]  

(5.3)

\[ g_m = \frac{I_c}{V_T} = \frac{I_c}{q kT} = \frac{q \cdot \text{const}}{k} \]  

(5.4)

The PTAT current generator is shown in Fig 5.3. For the parameters chosen,

\[ V_{be1} = V_{be2} + I_{c2}R \]

\[ I_{c1} = I_{c2} = \frac{V_T \ln 4}{R} \]  

(5.5)

The up-down helper corrects the \( \beta \) dependent of the bias current while maintaining \( V_{cb}=0 \), providing headroom for \( M_{\text{ptat1}} \). If only a simple npn helper is used, the headroom is not sufficient at 2.7V supply. \( \text{RE}_7 \) serves both as a current source to the pnp and startup path from \( V_{cc} \) to ground.

The base of \( Q_7 \) has a dominant pole that stabilizes the \( Q_7-Q_6-Q_4 \) feedback loop.

![Figure 5.3: PTAT Current Source](image-url)
Chapter 6

Conclusion

Various variable-gain amplifier topologies have been investigated. The variable local feedback circuit is chosen due to good noise and distortion performance across the gain range.

The characteristics from SPICE simulations are summarized below.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center Frequency</td>
<td>1 GHz</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>128 MHz</td>
</tr>
<tr>
<td>Gain</td>
<td>20dB @ ( V_{gc} ) 2.3 to 6dB @ ( V_{gc} ) 3.3</td>
</tr>
<tr>
<td>NF</td>
<td>2.6dB @ 20dB gain, 6.7dB @ 6dB gain</td>
</tr>
<tr>
<td>( IP_3 )</td>
<td>-15dBm @ 20dB gain, -2dBm @ 6dB gain</td>
</tr>
<tr>
<td>( S_{11}, S_{22} )</td>
<td>(&lt; -13dB)</td>
</tr>
<tr>
<td>Supply Voltage</td>
<td>2.7 to 5.5 V</td>
</tr>
<tr>
<td>Temperature Range</td>
<td>(-40^\circ) to (125^\circ) C</td>
</tr>
<tr>
<td>Power @ 2.7V, 25°C</td>
<td>32 mW</td>
</tr>
</tbody>
</table>
Appendix

Circuit Schematic and SPICE Simulation Results

Figure A1: Circuit Schematic
Component value listing

BJT's  
Q1a,Q1b,Q1c,Q2 : BNG12x  
Q3 : BNG 10x  

MOSFET's  
M1,M2 : 120/1  
Mf1 : 60/1  
Mf2 : 140/1  

Capacitors  
C1 : Large off chip  
C2* : 6pf. gate cap  
C3* : 4pf. gate cap  
C4* : 8pf. gate cap  
C5* : 4pf gate cap  
C bypass* : 15pf gate cap  
R bypass : 11Ω (parasitic series resistance of C bypass)  
C6 : Large off chip  

* All gate capacitors have 9% parasitic from the bottom plate to substrate.

Resistors  
RE3 : BN 12Ω  
RC3 : BN 50Ω  

Inductors  
All inductors are parasitic.  
Lgnd1,Lgnd3 : 2.35nH (corner bond wire plus lead frame)  
Lgnd2 : 1.72nH (center bond wire plus lead frame)  
Lgnd4 : 1.72nH (center bond wire plus lead frame)  
Lin : 2.72nH (bond wire, lead frame and 1nH series in C1)  
Lvcc : 2.35nH (corner bond wires plus lead frame)  
Lout : 1.72nH (center bond wire plus lead frame)  
Lvgc : 2.35nH (corner bond wire plus lead frame)  
Lcan : 1nH (bypass cap. parasitic series inductance)
Figure A2: Gain vs Control Voltage at 1GHz

Figure A3: Input Reflection Coefficient vs Frequency
Figure A4: Gain vs Frequency

Figure A5: IP₃ vs Gain
Figure A6: NF vs Gain
References


