TECHNICAL UNIVERSITY OF CRETE

School of Electrical and Computer Engineering

Design of a 5GHz Integrated Radio Frequency Power Amplifier in 90nm CMOS Technology

Diploma Thesis

Alexia Papadopoulou

Examinatory committee:

Associate Professor Matthias Bucher, Supervisor

Professor Konstantinos Kalaitzakis

Associate Professor Aggelos Bletsas

Chania, September 2016

Περίληψη

Ζώντας σε έναν συνεχώς αναπτυσσόμενο "ασύρματο" κόσμο, όλο και περισσότερος εξοπλισμός που επικοινωνεί ασύρματα με άλλες συσκευές, με ολοένα μεγαλύτερες ταχύτητες επικοινωνίας, προτείνεται στους σύγχρονους καταναλωτές. Η πληθώρα των συσκευών αυτών εκτείνεται από τα ακουστικά και εξοπλισμό για πιο εξιδεικευμένη χρήση, έως ηχεία, σταθμούς εργασίας και άλλες φορητές συσκευές που χρησιμοποιούν λ.χ.Bluetooth. Το φάσμα λοιπόν των 2.5GHz είναι φορτωμένο, λόγω της συνεχούς χρήσης των συσκευών αυτών αλλά και του εύρους φάσματος που χρησιμοποιούν. Για την επίλυση του προβλήματος του θορύβου που προκύπτει, η λύση ενός Wi-Fi δικτύου στη συχνότητα των 5GHz εμφανίζεται ως πιο βιώσιμη. Επιπρόσθετα, η ανάγκη για μικρού κόστους, υψηλής ολοκλήρωσης πομποδεκτών, οδήγησε στη χρήση της τεχνολογίας CMOS. Παρά το γεγονός ότι πολλά από τα δομικά στοιχεία ενός πομποδέκτη πραγματοποιήθηκαν σε τεχνολογία CMOS, το ζήτημα της σχεδίασης Ολοκληρωμένων Ενισχυτών Ραδιοσυχνοτήτων αποτελεί πρόκληση. Σε αυτή τη διπλωματική εργασία προτείνεται η σχεδίαση ενός Ενισχυτή Ραδιοσυχνοτήτων με συχνότητα λειτουργίας στα 5GHz για εφαρμογές που απαιτούν διαμόρφωση πλάτους, με μέγιστη ισχύς εξόδου 10dBm. Η μεθοδολογία της σχεδίασης και οι προκλήσεις της παρουσιάζονται, επίσης, ταυτόχρονα με την επιβεβαίωση της ορθότητας των αποτελεσμάτων μέσω πολλαπλών προσομοιώσεων.

Λέξεις κλειδιά: Ενισχυτής Ισχύος Ραδιοσυχνοτήτων, 5GHz, 90nm CMOS

Abstract

Living in an evergrowing "wireless" world, more and more electronic equipment communicating wirelessly with other devices at even higher communication speeds are presented to modern consumers. Plethora of devices from headsets, and other special-use equipment to speakers, docks and wearable devices, using e.g.Bluetooth. The 2.4GHz spectrum is getting crowded due to the bandwidth the devices use and the amount of time they're turned on. To help address the digital noise that comes with it, 5GHz Wi-Fi is presented as a more viable solution. In addition, the need for low cost, highly integrated transceiver building blocks led to the use of CMOS technology. Even though, many of the building blocks have been successfully integrated in CMOs technology, Power Amplifiers (PAs) are still a very challenging matter. This thesis proposes a PA design in 90nm CMOS technology for amplitude-modulated signal applications, with a maximum output power of 10dBm. The design methodology and related challenges are presented, together with the verification of the design via multiple simulations.

Key words: Radio Frequency Power Amplifier, 5GHz, 90nm CMOS

Acknowledgments

First and foremost, I would like to thank Professor Matthias Bucher for trusting me and giving me the chance to work on such an interesting and modern subject for my diploma thesis. Secondly, I would like to thank all members of the Electronics laboratory of Technical University of Crete and especially M.Sc. students Costas Papathanasiou and Nikos Makris, for their support, guidance and valuable advice.

Furthermore, I would like to thank my family for their support in all levels throughout the years. Last but not least, I would like to thank all of my friends for their help and moral support during our studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since the introduction of the first analog cellular phone in 1971 the communication industry in general and wireless industry in specific has followed a fast growth. More and more electronic equipment communicating wirelessly with other devices at ever higher communication speeds are presented to modern consumers. Some examples of this include, GSM, Bluetooth, WLAN, UMTS, ambient networks, femtocells, car-radio. With the advent of new evolutions in wireless communications such as 3G, 4G, the demand from the market on low-cost, miniaturized, low-power, and highly integrated wireless communication devices has increased. In order to meet these demands integration of RF building blocks into mainstream CMOS technology has been an ongoing trend. One of the most challenging RF blocks to implement in mainstream CMOS technology is the Power Amplifier (PA). So far, many RF building blocks of mobile communication devices have been implemented successfully in CMOS technology except for the PA, which is usually designed in a different technology. This is mainly because mainstream CMOS technology is optimized for low voltage, while power amplifiers operate at high voltages. Low breakdown voltages of CMOS transistors pose many reliability challenges for PAs. [1]

The Power Amplifier (PA) is the key building block in all RF transceivers. To lower the costs, by reducing board space and the number of components, it is highly desirable to integrate the entire transceiver and the PA in single CMOS chip operated with a single low "digital" supply voltage. Therefore, there is a need for highly efficient CMOS power amplifiers using a low supply voltage to achieve the goal of single-chip radio systems. A fully integrated CMOS solution would offer that and there would be no need of individual tuning of the different building blocks. To achieve this integration the PA must use the same fabrication process and supply voltage level as the digital circuits and secondly the matching networks must be on chip. The low quality factor of on-chip passives poses a big challenge. [2] Many researchers are focusing on the power amplifier design to achieve high efficiency and low cost with CMOS technologies.

Way back when Wi-Fi first came out, there were two protocol versions that you could choose from: 802.11a and 802.11b. From a consumer perspective, there wasn't much difference between the two. Devices based on 802.11b were generally less expensive and more readily available than those based on 802.11a, so the b specification quickly became the consumer standard. 802.11b operates in the 2.4GHz spectrum. These days, it's getting pretty crowded, and to help address the digital noise that comes with it, 5GHz Wi-Fi is making a comeback. 802.11a was a standard in 1999, which promised to bring network connections to devices delivered over the air instead of through copper cables. It was built around the 5GHz spectrum, but failed to gain much traction in the consumer market. Being the "first" Wi-Fi protocol, it faced a steep learning curve and deployment problems which delayed the deployment of 802.11a networks. Also, at the time, components that operated on 5GHz were generally more expensive and harder to come by than 2.4GHz components. When 802.11a was going through its initial phase, the 802.11b specification was being worked on. It offered basically the same features as 802.11a, but used less expensive and more readily available components. Due to these factors, 802.11b saw significant adoption amongst home and small-office users, whereas 802.11a only saw any level of "success" in enterprise network environments. Popularity of Wi-Fi began to grow, and the standards that backed it continued to improve. By 2003, a new standard had been ratified, though many devices were using the 802.11g draft specification prior to the date that it was made "official". This version of the Wi-Fi standard brought some of 802.11a's "stability" features and the inexpensive componentry of 802.11b, and the protocols were improved upon. All together the changes were able to increase speeds up to 54Mbp/s. Today it's still a viable option, but is giving way to 802.11n and 802.11ac. Unfortunately, 802.11g still uses the 2.4GHz spectrum, which is getting pretty crowded since all those Wi-Fi devices operate on the same frequency. Almost everyone has a microwave in their house. Some of them emit some of the radiation used to warm the food. Their "spurious emissions" cause bursts of noise around the 2.4GHz spectrum that can severely interfere with wireless signal. Bluetooth used to be limited to headsets and other special-use equipment, but as its feature-set increased, devices using Bluetooth increased too — and not just in number, but in the bandwidth they use and the amount of time they're turned on. Bluetooth speakers and docks are a good example of this, though wearables are quickly becoming more commonplace as well. Wireless keyboards, mice, trackpads, and trackballs can use Bluetooth to connect. Even those that use their own proprietary wireless hardware are typically still using the 2.4GHz spectrum. When 802.11n was introduced in 2009 it brought with it the ability to communicate at speeds up to 600Mbp/s. What's more, 802.11n also included the ability to work in either the 2.4GHz or 5 GHz spectra. Like the other standards before it, 802.11n was backwards compatible with its predecessors. Unfortunately, since most devices already on the market were already using 2.4GHz, most 802.11n wireless access points stuck to 2.4GHz as the primary operating frequency, and some devices didn't even include the hardware to use 5Ghz at all. 802.11ac was ratified in January 2014, but devices based on the draft specification were available for months prior. This standard brings the maximum data rates up to 1Gbp/s (almost double that of 802.11n). In most 802.11ac wireless access

points, both 2.4GHz and 5GHz hardware is included, though most segregate the traffic from each onto its own network. Finally, users can take advantage of the reduced noise available in the 5GHz spectrum. This generally provides faster data rates, fewer disconnects, and a more enjoyable experience. With a stronger signal and faster the throughput, less power is required to get your signal above the noise floor, which should result in better battery life in addition to better network performance. [3]

The goal of this diploma thesis is to present a single-ended Power Amplifier design in the 5GHz spectrum. This design was implemented in 90nm CMOS technology and we were able to achieve a maximum output power of 10dBm at the frequency of interest. The whole design process will be elaborately discussed and the verification will be provided via simulations with Virtuoso Analog Design Environment of Cadence®.

Chapter 2

Radio Frequency Power Amplifiers

The radio frequency (RF) power amplifier (PA) is a key component in modern telecommunication systems since its power consumption dominates the other parts in the system. The purpose of the RF PA is to amplify the radio signal to a necessary power level for transmission to the receiver. The transmission is performed through the air. It is important to handle the PA's conflicting behaviors of efficiency and linearity in this process. RF PAs are divided into different classes, i.e. A, AB, B etc., with respect to their power efficiency. In RF PAs there is a trade-off between efficiency (defined as the ratio of the generated and consumed RF and DC powers, respectively) and linearity. High efficiency and high linearity cannot be achieved at the same time. Class A RF PAs belong to the group of weakly nonlinear systems. The inherent nonlinearity of the PA causes interference with other transmitting channels and is something that one wants to avoid. Modern digital modulation techniques offer high data rates but use high bandwidths and peak-to-average ratios. Together with the nonlinear behavior of the PA, it enlarges the interference problem and leads to spreading of the transmitted spectrum, which is often referred to as spectral regrowth. The only solution to the above problem is the use of linear PAs. The design of linear and efficient RF PAs in modern radio telecommunication systems has been described in the literature as one of the most challenging design. A number of different techniques have been proposed over the years. Among the many promising techniques are analog and digital predistortion, crosscancellation, envelope elimination and restoration, and Chireix and Doherty amplifier techniques. All of these techniques have their pros and cons and none of them can be said to be the obvious choice for the future. One of the most promising techniques is the technique of digital predistortion. In order to work, digital predistortion requires knowledge of the nonlinear characteristics of the amplifier since it, in principle, applies the inverse of the raw amplifier to the signal prior to amplification. Accurate nonlinear characterization of the amplifiers is necessary for several of these techniques and for optimizing the amplifier design. Behavioral models, also denoted as black-box models, have attracted interest as a means for characterizing PAs. [4]

2.1 The place of Power Amplifiers in the transceiver chain

The objective of an RF transceiver is to transmit and receive information. We envision that the transmitter (TX) somehow processes the voice or data signal and applies the result to the antenna (Figure 2.1). Similarly, the receiver (RX) senses the signal picked up by the antenna and

processes it so as to reconstruct the original voice or data information. We must point out the fact that TX must drive the antenna with a high power level so that the transmitted signal is strong enough to reach far distances and that RX may sense a small signal and must first amplify the signal with low noise. On the receiver side, the signal is sensed by a "low-noise amplifier" (LNA) and on the transmitter side the signal is transmitted and drives the antenna through a "power amplifier" (PA). [5]

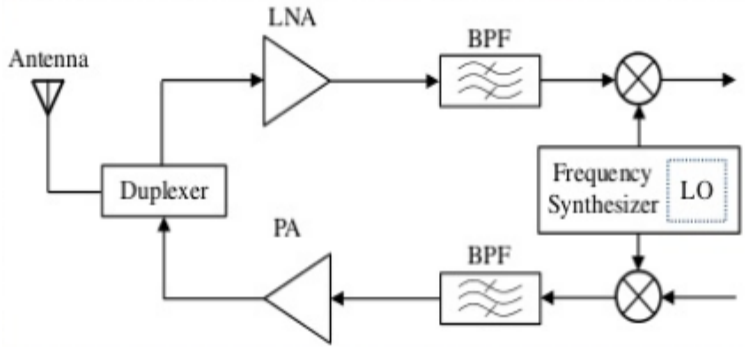


Figure 2.1. Generic diagram of a tranceiver front end. [6]

2.2. Basic power amplifier performance criteria

2.2.1. Characterization of Linear Circuits

Low-frequency circuits are usually analyzed in terms of transfer functions. This approach is seldom used at RF and Microwave frequencies. Analysis at these frequencies is usually in terms of one of many sets of single-frequency parameters.

The parameters most frequently used are the Y-, Z-, T-, and S-parameters. The first three sets of parameters relate the terminal voltages and currents in different ways, while the S-parameters are closely related to power incident to, and reflected from, a network. Because of the relative ease with which S-parameters can be measured and the useful information directly obtained from them, components are usually characterized by measuring their S-parameters and circuits are analyzed by calculating their S-parameters. The other parameters are often used to simplify the computation necessary for circuit analysis and synthesis. Also, the popularity of S-parameters stems from the fact that open and short circuits are not required in their definition or measurement, as is the case of all other parameters. Well behaved open and short circuits are difficult to obtain at high frequencies, and measured device stability may be compromised by open- or short- circuit terminations at its terminals, S-parameters are strictly small-signal parameters and are not suitable for describing non-linear operation. [7], [8]

2.2.1.1. S-parameters

S-parameters are defined in terms of incident and reflected components. In S-parameters theory, however, an incident component is defined as the component that would exist if the port under consideration were conjugately matched to the normalizing impedance at that port. The normalizing impedances are equivalents of the short-circuit and open-circuit terminations used to characterized a network in terms of its Y-, Z-, or T-parameters. They can be defined to have any arbitrary value (as long as the resistive part is positive and not equal to zero), but 50-Ω impedances are used in most cases.

In terms of the current and voltage at each terminal, the incident components (V_i, I_i, α) and the reflected components (V_r, I_r, b) of an N-port are defined by the following set of matrix equations:

$$a = \frac{1}{\sqrt{2}}\left[Z_0 + Z_0^{*} \right]^{\frac{1}{2}} I_i \quad (2.1)$$

$$b = \frac{1}{\sqrt{2}}\left[Z_0 + Z_0^{*} \right]^{\frac{1}{2}} I_r \quad (2.2)$$

$$I_i = \left[Z_0 + Z_0^{*} \right]^{-1} E_0 \quad (2.3)$$

$$E_0 = V + Z_0 I \quad (2.4)$$

$$I = I_i - I \quad (2.5)$$

$$V_i = Z_0^{*} I_i \quad (2.6)$$

$$V = V_i + V_r \quad (2.7)$$

$$Z_0 = \begin{bmatrix} Z_{01} & 0 & 0 & \dots & 0 \\ 0 & Z_{02} & 0 & \dots & 0 \\ 0 & 0 & Z_{03} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & Z_{0N} \end{bmatrix} \quad (2.8)$$

$$\frac{1}{\sqrt{2}}\left[Z_0 + Z_0^*\right]^{\frac{1}{2}} = \begin{bmatrix} \sqrt{R}_{01} & 0 & 0 & ... & 0 \\ 0 & \sqrt{R}_{02} & 0 & ... & 0 \\ 0 & 0 & \sqrt{R}_{03} & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \sqrt{R}_{0N} \end{bmatrix} (2.9)$$

With Z_{oj} the normalizing impedance at port j, Z_0 the corresponding matrix, Z_0^* the matrix with conjugate elements of those of Z_0, I_{ji} and V_{ji} the incident current and voltage at port j, I_{jr}, and V_{jr} the reflected current and voltage at port j, α_j the normalized incident component at port j, and b_j the normalized reflected component at port j. The reflected current is defined as the difference between the incident current and the actual current, while the reflected voltage is defined as the difference between the actual voltage and the incident voltage. The incident voltage is equal to the product of the conjugate of the normalizing impedance and the incident current; that is

$$V_i = Z_0^* I_i (2.10)$$

There are three different types of S-parameters, which are defined in the following way:

$$I_r = S^1 I_i (2.11)$$
$$V_r = S^V V_i (2.12)$$
$$b = S\alpha (2.13)$$

These sets of parameters are the current, voltage, and normalized S-parameters, respectively.

For a two-port network (Figure2.2), such as a power amplifier (single-stage or multi-stage) is considered, reduces to:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} (2.14)$$

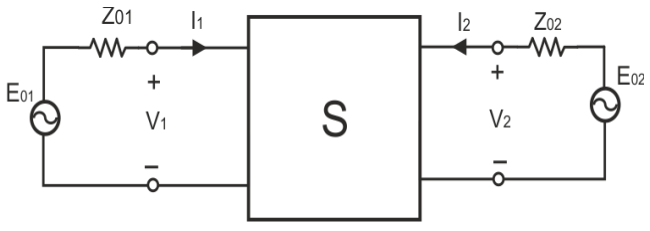


Figure2.2. The two-port network augmented to the normalizing impedances.

Generally, s_{ii} is the reflection coefficient at port i when all other ports are terminated in a 50-Ω impedance and s_{ij} is the transducer gain from port j to port i under the same conditions. Elaborately, the input reflection parameter s_{11} is defined by $s_{11} = \left.\frac{b_1}{a_1}\right|_{a_2=0}$ (2.15) and is the ratio of the reflected and incident waves at the input port when the reflection from the output impedance is zero. It represents the accuracy of the input matching. The forward transmission parameter s_{21} is defined by $s_{21} = \left.\frac{b_2}{a_1}\right|_{a_2=0}$ (2.16) and is the ratio of the wave incident on the load to that going to the input when the reflection from output impedance is zero. This parameter represents the gain of a circuit. The output reflection parameter is defined by $s_{22} = \left.\frac{b_2}{a_2}\right|_{a_1=0}$ (2.17)

and is the ratio of the reflected and incident waves at the output when the reflection from the input impedance is zero. This parameter represents the accuracy of the output matching. Also, s_{12} parameter is the ratio of the reflected wave at the input port to the incident wave into the output port when the input port is matched. In this case, the output port is driven by the signal source. This parameter characterizes the "reverse isolation" of the circuit (how much the output signal couples to the input signal) and is defined by:

$$s_{12} = \left.\frac{b_1}{a_2}\right|_{a_1=0} \quad (2.18)$$

The constraints on the current and voltage at the output terminals, when $a_2 = 0$ can be determined by:

$$0 = a_2 = \frac{V_2 + Z_{02} I_2}{2\sqrt{R_{02}}} \quad (2.19) \text{ , leading to } V_2 = Z_{02}\left[-I_2\right].$$

In order for a_2 to be equal to zero, the load impedance across the output port must therefore be equal to the normalizing at that port, and the electromotive force (EMF) must be equal to zero. [5], [7]

Finally, S-parameters are frequency-dependent complex values and are expressed in units of dB (decibels).

2.2.2.2. Z-parameters

Any of the voltages, currents, or power levels in a linear N-port network can be calculated in terms of the external signals (independent variables) when one of these sets of parameters is known at the frequency of interest.

The Z-parameters of an N-port network are defined by the expression V=ZI, which can be used to find the terminal voltages corresponding to any given set of terminal currents.

Where Z is

$$Z=\begin{bmatrix} z_{11} & z_{12} & \dots & z_{1N} \\ z_{21} & z_{22} & \dots & z_{2N} \\ \dots & \dots & \dots & \dots \\ z_{N1} & z_{N2} & \dots & z_{NN} \end{bmatrix} \quad (2.20)$$

, each element can be computed or measured by using the relationship

$$z_{ij}=\left.\frac{V_i}{I_j}\right|_{I_h=0}, h\in[1,2,3,...,N], h\neq j \quad (2.21)$$

That is, z_{ij} is the ratio of the voltages across the j[th] port (output signal) and the current at the i[th] port (input signal) with all the other ports idle (open-circuited).

Finally, V and I are defined by $V=\begin{bmatrix} V_1 \\ V_2 \\ . \\ . \\ V_N \end{bmatrix}, I=\begin{bmatrix} I_1 \\ I_2 \\ . \\ . \\ I_N \end{bmatrix}$, respectively.

Z-parameters are frequently used to find an equivalent set of parameters for two networks connected in series. When networks are connected in series, the terminal currents are the same, while the voltages add.

The Z-parameters of two networks connected (cascaded as most of PAs) in series are given in terms of the individual Z-parameters by $Z_T=Z_A+Z_B \quad (2.22)$.

At last, the relation that between the Z-parameters (impedance) and the Y-parameters (admittance) is as follows: [7] $Z=Y^{-1} \quad (2.23)$

2.2.2. Linearity

The need for linearity is one of the principal drivers in the design of modern PAs. Signals such as CW, FM, classical FSK, and GMSK (used in GSM) have constant envelopes (amplitudes) and, therefore, do not require linear amplification. Full-carrier amplitude modulation is best produced by high-level amplitude modulation of the final RF PA. Linear amplification is required when the signal contains both amplitude and phase modulation. Examples include SSB voice, vestigal-sideband television, modern shaped-pulse data modulation (QAM, QPSK, CDMA), and multiple carriers (OFDM). [9]

Considering a single-input, single-output system, for input x(t) the output is y(t). A system is considered to be linear if its output can be expressed as a linear combination (superimposition) of responses to individual inputs. More specifically, if the outputs in response to inputs $x_1(t)$ and $x_2(t)$ can be expressed as $y_1(t)=f[x_1(t)]$ and $y_2(t)=f[x_2(t)]$, respectively, then $\alpha y1(t)+by2(t)= f[\alpha x1(t)]+ f[bx2(t)]$ for arbitrary values of α and b. Otherwise, it is called non-linear. Furthermore, a system is called time invariant if a time shift in its input results in the same time shift in its output. A linear system can generate frequency components that do not exist in the input signal, if only is time-variant. In addition, a system is called memoryless if its output does not depend on the past values of its input (or the past values of the output itself).

For a memoryless linear system the input/output characteristic is:

$y(t) = \alpha x(t)$, where α is a function of time if the system is time-variant.

For a memoryless non-linear system, the input/output characteristic can be approximated with a polynomial:

$y(t) = \alpha_0 + \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) + ..., a_i \in \mathbb{R}$ or a_i may be a function of time, if the system is time-variant.

If a sinusoid is applied to a non-linear 3rd order system, the output exhibits frequency components that are integer multiples of the input frequency and if $x(t) = A\cos\omega t$, then

$$
\begin{aligned}
y(t) &= \alpha_1 A\cos\omega t + \alpha_2 A^2\cos^2\omega t + \alpha_3 A^3\cos^3\omega t \\
&= \alpha_1 A\cos\omega t + \frac{\alpha_2 A^2}{2}(1+\cos 2\omega t) + \frac{\alpha_3 A^3}{3}(3\cos\omega t + \cos 3\omega t) \qquad (2.24) \\
&= \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4}\right)\cos\omega t + \frac{\alpha_2 A^2}{2}\cos 2\omega t + \frac{\alpha_3 A^3}{4}\cos 3\omega t
\end{aligned}
$$

We can see that from the input of single frequency derives a dc quantity arising from second order non-linearity, a second and a third harmonic. A real system can never be though as fully linear and entails effects of non-linearity.

The power of a system is the total power, at the fundamental and harmonic frequencies. Most of the power is at the fundamental frequency. In some systems, the strongest harmonic is the second relative to the fundamental, without any filtering. Even-order harmonics can be canceled with a push-pull circuit, if desired; in that case, the strongest harmonic is the third and is of most concern to us. [10]

2.2.2.1. 1-dB compression point

The small-signal gain of circuits is obtained usually with the assumption that harmonics are negligible. Going back to the polynomial, with $\alpha_1\alpha_3<0$, the gain experienced by the input falls as A rises. We can quantify this effect by the 1-dB compression point, defined as the input signal level that causes the gain to drop by 1dB. On a log-log scale as a function of input level, the output level, A_{out} , falls below its ideal by 1dB at the 1-dB compression point (Figure 2.3). The input compression point proves relevant in the receive path. The same applies for the output.

To calculate the input 1-dB compression point, we equate the compressed gain, $\alpha_1+(3\alpha_3A^2/4)$ to 1-dB less than the ideal gain, α_1:

$$20\log\left|\alpha_1+\frac{3}{4}\alpha_3 A_{1-dB}{}^2\right|=20\log|\alpha_1|-1dB \quad (2.25)$$ and finally,

$$A_{1-dB}=\sqrt{0.145\left|\frac{\alpha_1}{\alpha_3}\right|} \quad (2.26)$$

The 1-dB compression point is typically in the range of -20 to -25 dBm and is widely used to characterize RF circuits. [5]

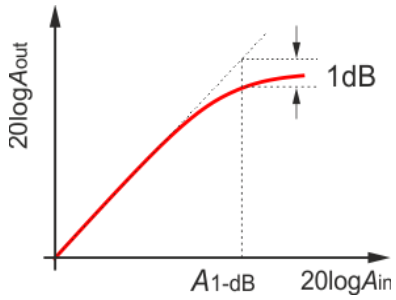


Figure.2.3. Definition of 1-dB compression point.

2.2.2.2. Intermodulation

Another scenario of interest in RF design occurs if two interferes accompany the desired signal. Such a scenario represents realistic situations and reveals non-linearities. If two interferers at ω_1 and ω_2 are applied to a non-linear system, the output generally exhibits components that are not harmonics of these frequencies. This phenomenon, called intermodulation, arises from mixing of the two components as their sum is raised to a power greater than unity. Let us assume $x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$. [5] Thus,

$$y(t) = \alpha_1(A_1 \cos \omega_1 t + A_2 \cos \omega_2 t) + \alpha_2(A_1 \cos \omega_1 t + A_2 \cos \omega_2 t)^2 + \alpha_3(A_1 \cos \omega_1 t + A_2 \cos \omega_2 t)^3 \quad (2.27)$$

Discarding the dc terms, harmonics and components at $\omega_1 \pm \omega_2$, we obtain the following

$$\omega = 2\omega_1 \pm \omega_2 : \frac{3\alpha_3 A_1^{\,2} A_2}{4} \cos(2\omega_1 + \omega_2)t + \frac{3\alpha_3 A_1^{\,2} A_2}{4} \cos(2\omega_1 - \omega_2)t (2.28)$$

$$\omega = 2\omega_2 \pm \omega_1 : \frac{3\alpha_3 A_1 A_2^{\,2}}{4} \cos(2\omega_2 + \omega_1)t + \frac{3\alpha_3 A_1 A_2^{\,2}}{4} \cos(2\omega_2 - \omega_1)t (2.29)$$

and these fundamental components:

$$\omega = \omega_1, \omega_2 : (\alpha_1 A_1 + \frac{3}{4}\alpha_3 A_1^{\,3} + \frac{3}{2}\alpha_3 A_1 A_2^{\,2}) \cos \omega_1 t + (\alpha_1 A_2 + \frac{3}{4}\alpha_3 A_2^{\,3} + \frac{3}{2}\alpha_3 A_2 A_1^{\,2}) \cos \omega_2 t (2.30)$$

So, the non-linearity of a system can be measured on the basis of generated spectra than on variations of the fundamental signal. The intermodulation components (IM) versus fundamental level change are equal to the order of non-linearity. Also, this phenomenon is trivial for systems with frequency modulation. In order to study and avoid intermodulation, two-tone tests can be performed, even on systems with small bandwidth.

The level of the harmonic and intermodulation products decreases stronger than the fundamental, with decreasing the input power. This deduces a crude method for linearization named Back-Off Power Optimization for Linearity.

2.2.2.3. Third intercept point (IP_3)

The principal difficulty in specifying the relative IM for a system is that it is meaningful only if the value A (amplitude) is given. Instead, a single measurement that captures the intermodulation behavior of a system can be performed and is called Third interception point IP_3 (Figure). The concept of IP3 originates from an observation that, if the amplitude of a tone rises, the amplitude of the output IM products increases sharply. Thus, is the A continues to raise, the amplitude of the IM products becomes equal to that of the fundamental at the output. The input level at which this occurs is the input third intercept point (IIP_3). The corresponding output is represented by the output third intercept point (OIP_3).

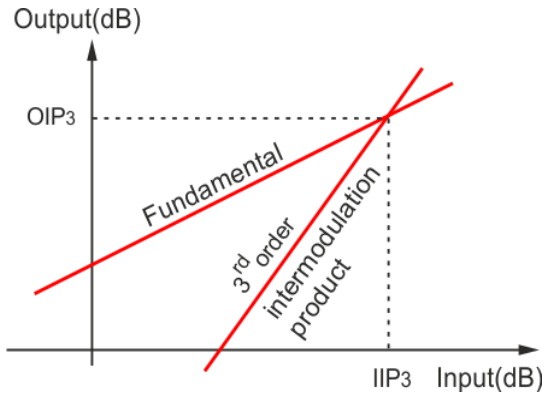


Figure 2.4. Definition of IP_3.

2.2.3. Power criteria

2.2.3.1. Saturated output power

Saturated output power, P_{sat}, is the output power of a power amplifier when the amplifier is saturated. Specifically, it is the output power where the P_{in} versus P_{out} slope goes to zero (Figure), meaning that the gain of the amplifier also goes to zero. Finally, when an amplifier is saturated means that beyond a specific value of input power, according to the design specifications, the output power will not exceed the saturated power.

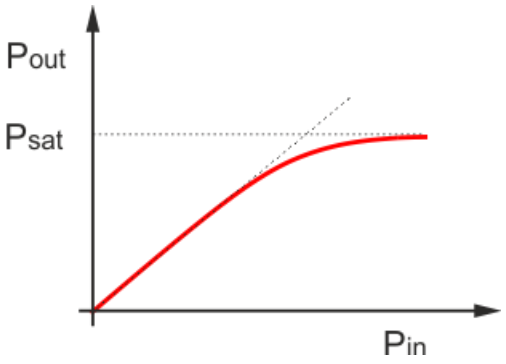


Figure 2.5. Saturated output power

Saturated output power is one of the standards of a power amplifier according to use, as defined by protocols seen in table 2.1.

Commonly, power amplifiers operate in a lesser output power, due to linearity, cost and space coverage.

Protocol	Frequency of Operation	Saturated Output Power
GSM	900 MHz,1.8 GHz	31 dBm
Bluetooth	2.4 GHz	16 dBm
802.11a	5 GHz	14-19 dBm
802.11p	5.9 GHz	30 dBm
802.15.3	60 GHz	10 dBm

Table 2.1.Protocols and saturated output power

2.2.3.2. Power gain

Power gain is defined as the ratio of the output power of an amplifier to the input power of the amplifier; it is normally expressed in dB and is defined as follows:

$$G = \frac{P_{out}}{P_{in}} (2.31)$$

2.2.3.3. Efficiency

Efficiency is a critical factor in PA design, because is usually the primary consumer of DC power in most wireless devices.

Drain efficiency (DE) is defined as the ratio of the average RF output power delivered to the load to the average dc input power drawn by the supply,

$$\eta = \frac{P_{out}}{P_{dc}} \times 100\% (2.32)$$

2.2.3.4. Power added efficiency

Drain efficiency measure has the drawback that the RF power delivered at the input of the amplifier is not accounted for. Furthermore, in some cases, the output stage may have a relatively low power gain, requiring a high input power, leading to the fact that DE overrates the actual efficiency. Power added efficiency (PAE) incorporates the average RF drive (input) power by subtracting it from the output power,

$$PAE = \frac{P_{out} - P_{in}}{P_{dc}} \times 100\% (2.33)$$

Thus, $PAE = (1 - \frac{1}{G}) \frac{P_{out}}{P_{dc}} = (1 - \frac{1}{G}) \eta (2.34)$

PAE gives a reasonable indication of PA performance when gain is high; however, it can become negative for low gains. [9]

2.3. Impedance matching

All systems as RF power amplifiers require that all available input power be transferred from one stage to the next. This requires that the stages be impedance matched to each other. The output impedance of a stage and the input impedance of the following stage must be matched to 50-Ω or be conjugates of each other. [8] Impedance matching or tuning is the process of matching one load to another of higher or lower value. The basic idea of impedance matching is illustrated in Figure 2.6 , which shows an impedance matching network placed between a load impedance and a signal source (input or output). The matching network is ideally lossless, to avoid unnecessary loss of power, and is usually designed so that the impedance seen looking into the matching network is Z_0 (Figure). Then reflections are eliminated to the left (right) of the matching network, although there will be multiple reflections between the matching network and the load. Impedance matching is important for the following reasons:

- Maximum power is delivered when the load is matched to the input/output impedance of a stage and power loss is minimized.
- Impedance matching sensitive receiver components (antenna, LNA etc.) improves the signal-to-noise ratio of the system.
- Impedance matching in a power distribution network (such as an antenna array feed network) will reduce amplitude and phase errors.

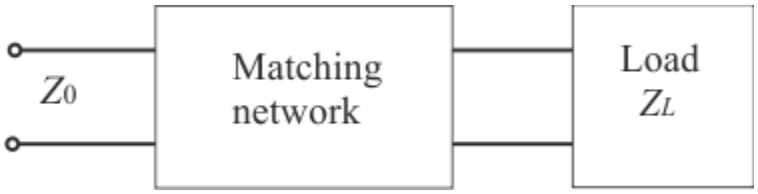


Figure 2.6. An impedance matching network placed between a load impedance and a signal source.

As long as the load impedance, Z_L, has some non-zero real part, a matching network can always be found. Factors that may be important in the selection of a particular matching network include the following:

- Complexity: the simplest design that satisfies the required specifications is the most preferable. A simpler matching network is usually cheaper, more reliable and less lossy than a more complex design.
- Bandwidth: any type of matching network can ideally give a perfect match (zero reflection) at a single frequency. In many applications, it is desirable to match a load over a band of frequencies. There are several ways doing this with a corresponding increase in complexity.
- Implementation: comparison between matching networks for a specific application should be made.
- Adjustability: the matching network may require adjustment to match a variable load impedance. Some types of matching networks are more amenable than others on this regard.[11]

2.3.1. L-matching networks

An L-matching network is a two-element matching network. There are four possible configurations as seen in Figure 2.7(a), (b), (c), (d).

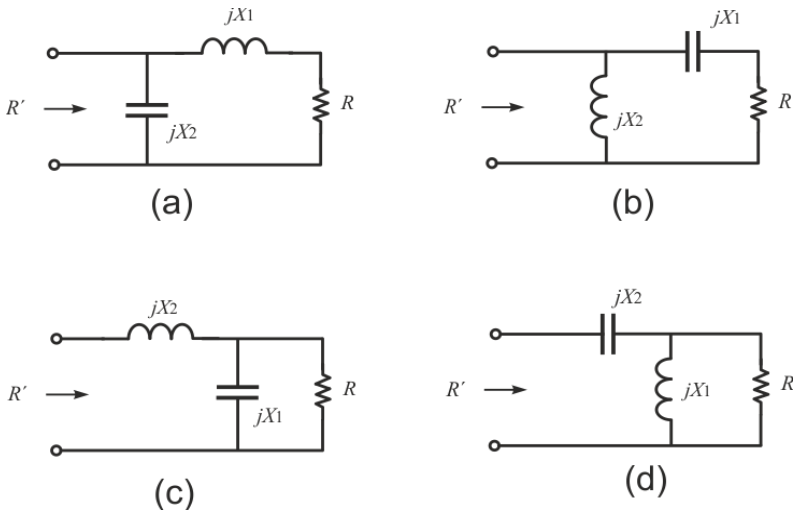


Figure 2.7. The four possible configurations for an L-network (a), (b), (c), (d).

Depending on the position of the first component (as viewed from the load), the load resistance can be transformed upwards or downwards with an L-network.

When the first reactive component is a series component, the transformation is upward; and when it is a parallel element, the transformation is downward. The second element in the L-

network is used to remove the residual reactance caused by the transformation element. This element is therefore the compensating element.

The basic principle used in narrowband impedance matching is that the resistance of a complex load is not the same when viewed in impedance or admittance form.

When a reactive element (X_1) is added in series with a resistor (R) and the equivalent parallel combination is considered (series too shunt transformation), the resistance increases with a factor $D_1 = 1 + Q_1^2$ (2.35), where $Q_1 = \frac{X_1}{R}$ (2.36)

When a reactive element (X_1) is added in parallel with a resistor (R) and the equivalent series combination is considered (parallel to series transformation), the resistance decreases with the same factor (D_1). In this case, however, the Q-factor is defined by $Q_1 = -\frac{R}{X_1} = \frac{B_1}{G}$ (2.37)

The ratios defined above are referred to as transformation Q_s.

The sign of the reactance or susceptance is carried over to the transformation Q. It follows that the transformation Q is positive when the effective series reactance is inductive or when the effective shunt susceptance is capacitive.

The reactance changes by a factor $E_1 = 1 + \frac{1}{Q^2}$ (2.38) in the transformation step. As is with the resistance, the reactance increases after a series to shunt transformation and decreases when a shunt to series transformation is considered.

The reactance of the first element used in L-networks is determined by the transformation Q required to transform the load resistance (R) to the value required (R'). The Q value can be calculated by using the relationship $R' = D_1 R = (1 + Q_1^2)R$ (2.39).

A positive or negative sign can be assigned to the transformation Q.

The second element in the L-network is used to achieve the desired reactance level. If a purely resistive input impedance is required, the reactance of this element is given by $X_2 = -X_1(1 + \frac{1}{Q_1^2}) = \frac{R'}{Q_1}$ (2.40) , if the first element is a series element by $X_2 = -\frac{X_1}{(1 + \frac{1}{Q_1^2})} = -R'Q_1$ (2.41), if the first element is a shunt element.

Equations (2.40) and (2.41) can be verified easily by using the relationships $Z = \frac{1}{Y}$ and $Y = \frac{1}{Z}$, respectively.

When the first element on the load side is a shunt element we can apply the following (Figure 2.8):

$$R' = \frac{R}{(1+Q_1^2)} \quad (2.42)$$

$$X' = \frac{X_1}{(1+\frac{1}{Q_1^2}} \quad (2.43)$$

$$X_2 = X_{in} + Q_1 R' \quad (2.44)$$

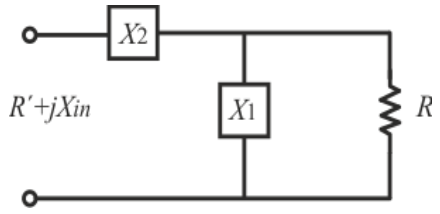


Figure 2.8. Transformation properties of an L-network when the first element on the load side is a shunt element.

When the first element on the load side is a series element we can apply the following (Figure 2.9):

$$R' = R(1+Q_1^2) \quad (2.45)$$

$$X' = X(1+\frac{1}{Q_1^2}) \quad (2.46)$$

$$B_2 = B_{in} + \frac{Q_1}{R'} \quad (2.47)$$

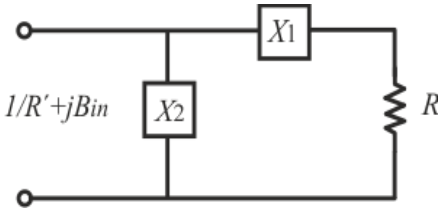


Figure 2.9. Transformation properties of an L-network when the first element on the load side is a series element.

If higher Q is needed back-to-back L-networks can be used, each one transforming down to center impedance that is lower than the source resistance. Two types of such networks exist

Pi-networks (Figure 2.11) and T- networks (Figure 2.10). The center impedance gives us control over the frequency response while still providing a perfect match at the center or design frequency. [12]

Because the cut off frequencies (3-dB) of L- Pi-, T- networks are known to good approximation $f_{-3dB} = f_0 \pm \frac{f_0}{Q_{\max}}$ (2.48), the exact bandwidth of these circuits can be determined easily.[7]

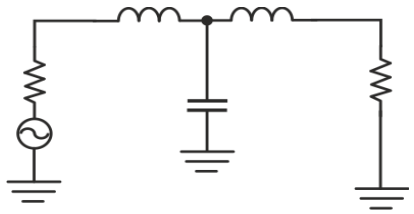


Figure 2.10. The T-network.

Loaded Q, the Q factor can be either high or low depending on the application. Narrowband filters have high loaded Q. Loaded Q is therefore not a measure of quality. Unloaded Q, however, which specifies the losses in components, is indeed a measure of quality since lowering component losses always increases circuit efficiency.

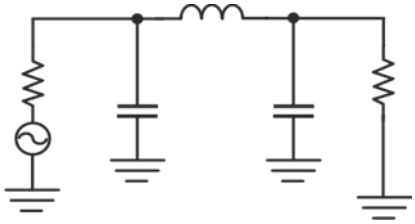


Figure 2.11. The Π-network.

2.3.2. Compensation network

The matching networks in the case of Power amplifiers may use a supplementary network called compensation network. Classic L-, Pi-, T- , networks are applied in cases when real load is matched to real load (in some cases the reactance part of the input/output impedance

is very small and can be “ignored”). When the reactance of the impedance to be matched is significant the matching networks include a compensation network formed by a capacitor in series with an inductor, as seen in Figure 2.12 and 2.13. [13]

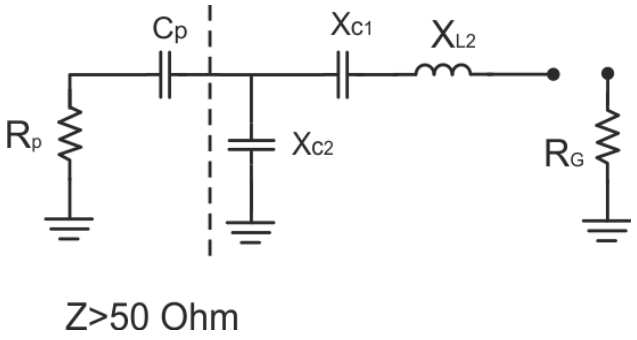


Figure 2.12. The reversed L-network with compensation.

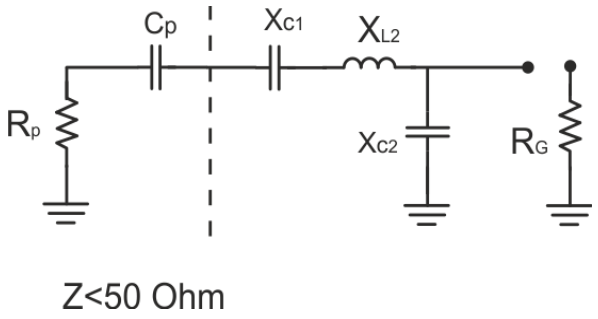


Figure 2.13. The L-network with compensation.

Parallel, also known as shunt or load capacitor is always on the high R side. Called “load” capacitor because it adjusts the real part of the load. Series capacitor sometimes called “tune” because it tunes out the reactive part of the load. By adding one reactance, exact impedance transformation is achieved at two frequencies. Better use can be made of large number of reactive elements using them all for load transformation. [14]

2.3.3. Smith Charts

Impedance and reflection coefficients are related quantities provided the reference impedance Z_0 is known (in most cases 50 Ohm, 30 or 75 Ohm). The relations can take the following form:

$$\Gamma = \frac{Z - Z_0}{Z + Z_0} \quad (2.49) \quad \text{and} \quad Z = \frac{1+\Gamma}{1-\Gamma} Z_0 \quad (2.50)$$

Γ= reflection coefficient (complex, dimensionless)

Z= impedance (complex, Ohm)

Z_0= reference impedance (real, Ohm)

Smith charts represent a polar depiction of the normalized reflection coefficient to a reference impedance. Most impedance matching problems can be analyzed as trajectories on the Smith chart, where the addition of a lumped element moves the total impedance along impedance or admittance constant circles. If we want to match our impedance to the reference impedance Z_0, the aim is to arrive at the center of the Smith chart by traveling along impedance and admittance arcs from starting point. The center of the Smith chart is situated on the unity impedance/or admittance circle. [8]

There are two Smith charts. The Impedance Z-Smith chart (Figure 2.14) and the admittance Y-Smith chart (Figure 2.15). On both charts there are circles of constant resistance and curved lines called constant reactance segments.

There are three passive elements to choose from (inductor, capacitor, resistor) and each could be set in series or in parallel. As a result, there are six different setups:

1. Series resistor: bigger value moves the current point to the right, along the Z-segments
2. Shunt resistor: bigger value moves the current point to the left, along the Y-segments
3. Series capacitor: smaller value moves the current point counter-clockwise, along the Z-circle
4. Shunt capacitor: bigger value moves the current point clockwise, along the Y-circle
5. Series inductor: bigger value moves the current point clockwise, along the Z-circle
6. Shunt inductor: smaller value moves the current point counter-clockwise, along the Y-circle

Resistors are frequency independent while capacitors and inductors are frequency dependent. Although, matching with resistors should be avoided because resistors consume power. The matching networks as mentioned in previous section include two or more elements. Even though increasing the number of elements leads to significant power losses, the choice depends on the matching problem. [15]

Furthermore, constant Q lines can be overlaid on Smith charts to estimate the matching network bandwidth. In general, the closer an impedance-matching trajectory comes to the edge of the Smith chart, the narrower the bandwidth.

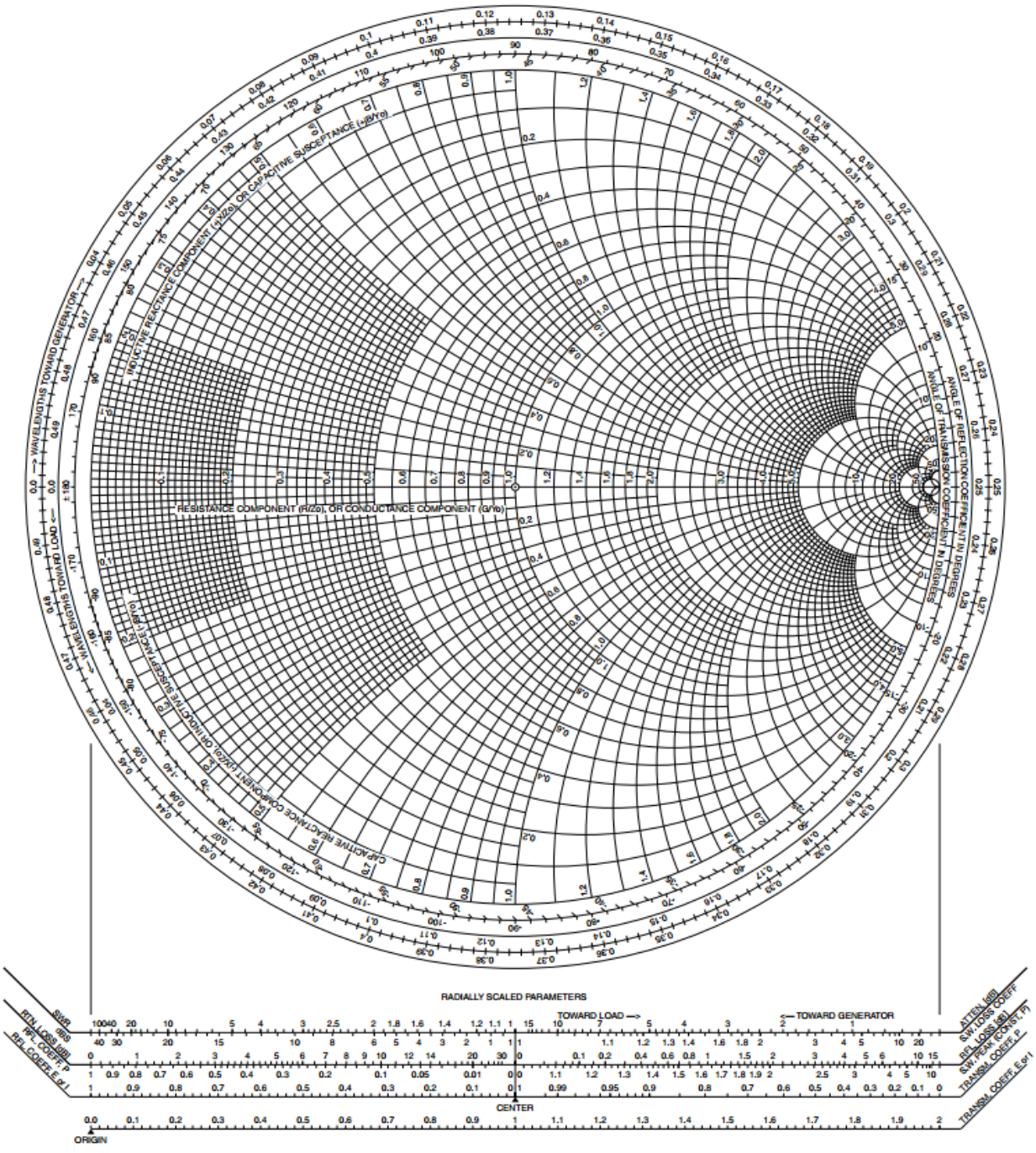


Figure 2.14. Impedance Smith Chart Black Magic Design ©.

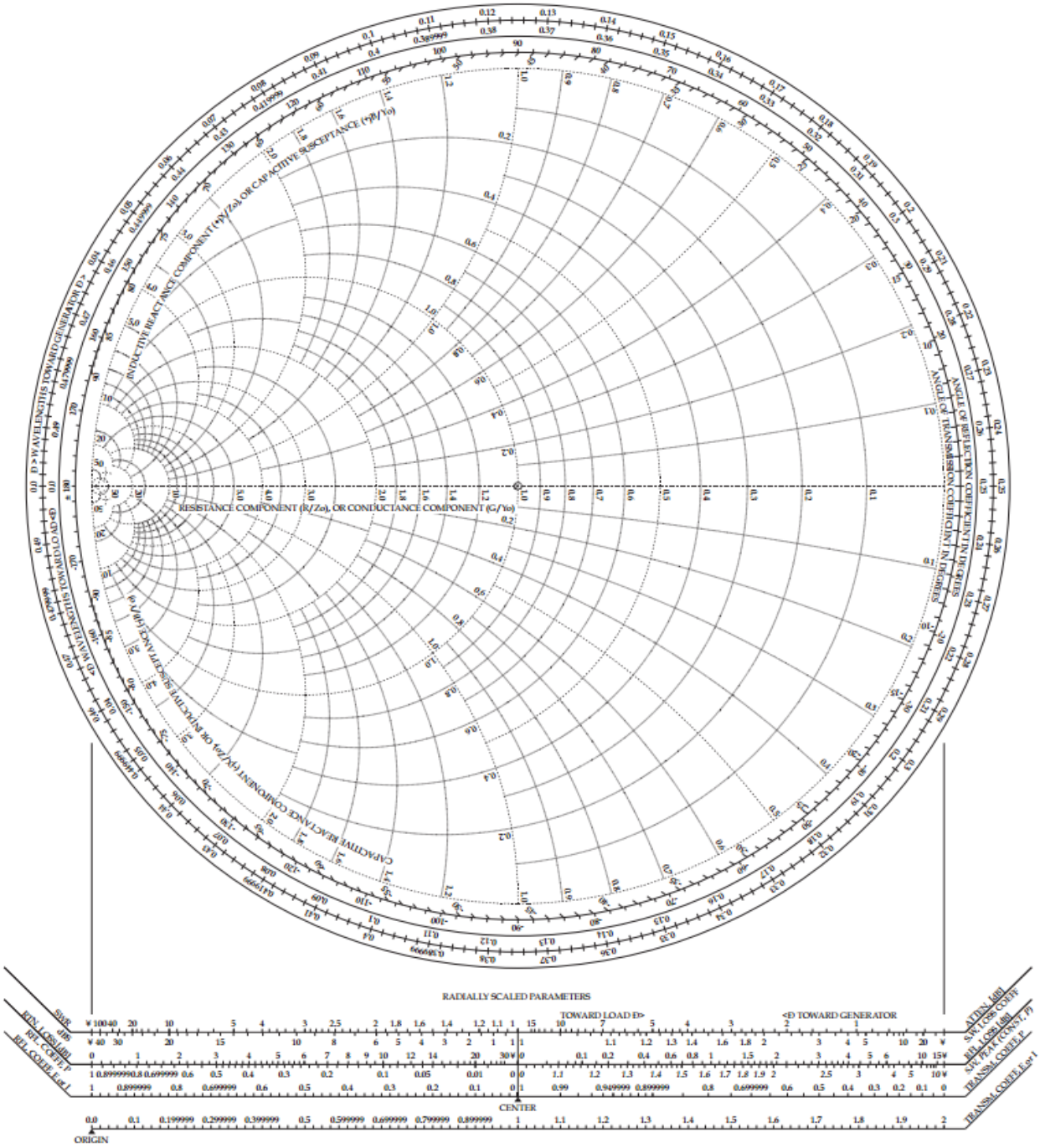


Figure 2.15. Admittance Smith Chart

2.4. Stability

A single stage transistor amplifier can be modeled by the circuit of Figure, where a matching network is used on both sides of the transistor to transform the input and output impedance Z_0 to the source and load impedances Z_S and Z_L.

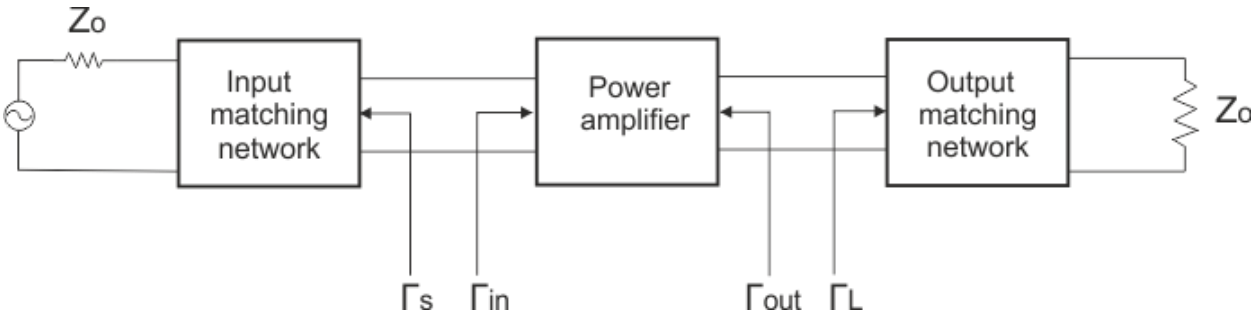


Figure 2.16. Single stage transistor amplifier.

In the circuit of Figure oscillation is possible if either the input or output port impedance has a negative real part; this would imply that $|\Gamma_{in}|>1$ or $|\Gamma_{out}|<1$. Because Γ_{in} and Γ_{out} depend on the source and load impedance matching networks, the stability of the amplifier depends on Γ_S and Γ_L as presented by the matching network. Thus, two types of stability can be defined:

1. Unconditional stability: the network is unconditionally stable if $|\Gamma_{in}|<1$ and $|\Gamma_{out}|<1$ for all passive source impedances.
2. Conditional stability: the network is conditionally stable if $|\Gamma_{in}|<1$ and $|\Gamma_{out}|<1$ only for a certain range of passive source and load impedances. The case is also referred to as potentially unstable.

The stability condition of a power amplifier is usually frequency dependent, since the input and output matching networks generally depend on frequency. It is possible for an amplifier to be stable at its design frequency but not in others. The rigorous general treatment of stability requires that the network S-parameters have no poles in the right half complex frequency plane, in addition to the conditions that $|\Gamma_{in}|<1$ and $|\Gamma_{out}|<1$.

Applying the above requirements for unconditional stability of Figure gives the following conditions that must be satisfied by Γ_S and Γ_L of the amplifier:

$$|\Gamma_{in}| = |s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L}| < 1 \quad (2.51)$$

$$|\Gamma_{out}| = |s_{22} + \frac{s_{12}s_{21}\Gamma_S}{1 - s_{11}\Gamma_S}| < 1 \quad (2.52)$$

If the device is unilateral (s_{12}=0), these three conditions reduce to the simple results that $|s_{11}|<1$ and $|s_{22}|<1$are sufficient for unconditional stability. Otherwise, the inequalities define a range of Γ_S and Γ_L where the amplifier is stable. Finding this range for Γ_S and Γ_L can be facilitated by using the Smith chart, and plotting the input and output stability circles. The stability circles are defined as the loci in the Γ_S (or Γ_L) plane for which $|\Gamma_{in}|=1$ (or$|\Gamma_{out}|=1$). The stability circles then define the boundaries between stable and potentially unstable regions of Γ_S and Γ_L. Γ_S and Γ_L must lie on the Smith chart (for passive matching networks).

When $|\Gamma_{in}|=1$ in the complex Γ plane, an equation of the form $|\Gamma - C|= R$ represents a circle with center at C (a complex number) and a radius R (a real number). Then, the output stability circle is:

$$C = \frac{(s_{22} - \Delta s_{11}^*)^*}{|s_{22}|^2 - |\Delta|^2} \quad \text{(Center of the circle)} \quad (2.53)$$

$$R = \left| \frac{s_{12}s_{21}}{|s_{11}|^2 - |\Delta|^2} \right| \quad \text{(Radius)} \quad (2.54), \text{ where}$$

$$\Delta = s_{11}s_{22} - s_{12}s_{21} \quad (2.55), \text{ the determinant of the scattering matrix.}$$

Similar results can be obtained for the input stability circle by interchanging s_{11} and s_{22}.

Given the S-parameters of the transistor, we can plot the input and output stability circles to define where $|\Gamma_{in}|=1$ and $|\Gamma_{out}|=1$. On the one side of the input stability circle we will have $|\Gamma_{out}|<1$, while on the other side we will have $|\Gamma_{out}|>1$.Similarly, for Γ_{in}. If the device is unconditionally stable, the stability circles must be completely outside (or totally enclose) the Smith chart. If the device is only conditionally stable, operating points for Γ_S and Γ_L must be chosen in stable regions, and it is good practice to check the stability at several frequencies near the design frequency. If it is possible to accept a design with less than maximum gain, a transistor can usually be made unconditionally stable by using resistive loading. [11]

For unconditional stability, there is also another measure called Rollet's stability factor (k) that can be described as follows: [8]

$$k = \frac{1 + |s_{11}s_{22} - s_{12}s_{21}|^2 - |s_{11}|^2 - |s_{22}|^2}{2|s_{12}s_{21}|} \quad (2.56) \text{ , where } k>1$$

When this is the case the device is stable no matter the Z_L at the output. Otherwise, the Z_L should be checked for introducing reflection Γ_L that drives the device to instability.

2.5. Classification of power amplifiers

Power amplifiers are categorized under many classes: A, B, C, AB, D, E, F, G, H, J etc. In this section classes A, B, C, AB, E and F will be discussed. The latter two also are referred to as high efficiency, switch-mode power amplifiers.

2.5.1. Class A

Class A amplifiers are defined as circuits in which the transistor (or cascode configuration) remains on and operates linearly across the full input and output range. The transistor bias current is chosen higher than the peak signal current I_p (Figure 2.17), to ensure that the device doesn't turn off at any point during the signal excursion. If linearity is required, then class A operation is necessary.

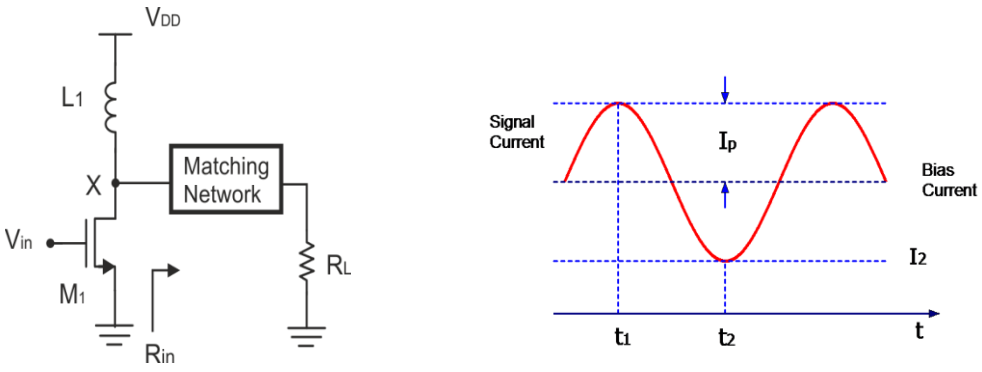


Figure 2.17. Class A stage.

Figure presents a simple model of a Class A single-stage power amplifier.

The circuit consists of an inductor L_1, whose inductance is high enough so it provides isolation toV_{DD} (dc voltage supply) from the ac current; L_1 is used for the drain biasing of transistor M_1. The matching network is such that the dc current is blocked from reaching R_L, in order to avoid power consumption.

The maximum drain efficiency of class A amplifiers derives from the fact that Vx is allowed to reach $2V_{DD}$ and nearly zero. Thus, the power delivered to the matching network is approximately equal to $\frac{(2\frac{V_{DD}}{2})^2}{2R_{in}} = \frac{V_{DD}^{\ 2}}{2R_{in}}$ (2.57) , which is also delivered to R_L, if the matching network is lossless. The inductive load carries a constant current of $\frac{V_{DD}}{R_{in}}$ from the supply voltage. Thus,

$$\eta = \frac{V_{DD}{}^2 / (2R_{in})}{V_{DD}{}^2 / R_{in}} = 50\% (2.58)$$

The other 50% of the power supply is dissipated by M1. However, since all components of the circuit have parasitic resistances, which waste power, (we consider all else, except from the R_L load resistor that the power is delivered to) the efficiency drops almost to 25%.

When no signal drives the amplifier, the quiescent power dissipation is:

$$P_{DC} = I_{DC} V_{DD} (2.59)$$

When a signal is present, the power dissipation of a transistor decreases because the transistor converts some of the quiescent power to signal power, but it is independent of the output signal amplitude, which is considered a sinusoid as the input.

In class A, the conduction angle is 360^o because the output transistor M_1 is always on, making this power amplifier the most linear of all (the output signal resembles the input signal). The designer usually tries to locate the Q point somewhere near the middle of the load line, so the signal can swing over the maximum possible range without saturating or cutting off the transistor, which would distort the signal.

The absence of harmonics in the amplification process, allows Class A to be used at frequencies close to the maximum capability (f_{max}) of the transistor. Small signal S-parameters can be used in simulations if the large signal is operating in class A, this explains that for large V_{in} non-linearity is introduced due to higher order harmonics.

They are used in applications requiring low power, high linearity, high gain, broadband operation, or high-frequency operation. [17], [18], [5]

2.5.2. Class B

The class B single-stage power amplifier employs two parallel stages each of which conducts only for 180^o of the ac cycle, thereby achieving a higher efficiency than class A. Figure 2.18 depicts the simple model of class B circuit. The drain currents of M_1 and M_2 are combined by a transformer T_1. The circuit is quasi-differential stage, with a balun driving the single-ended load. Class B operation requires that each transistor turns off for half of the period. The gate bias voltage of the transistors is chosen approximately equal to their threshold voltage.

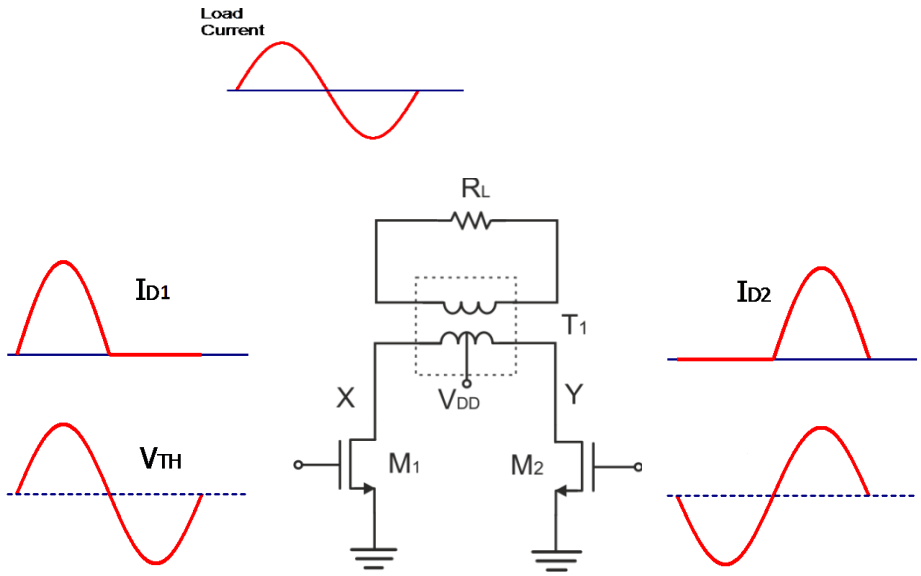


Figure 2.18. Class B stage

If the parasitic capacitances are small and the primary and secondary inductances are large, then V_X and V_Y (Figure 2.20) are also half-wave rectified sinusoids that swing around V_{DD}. The voltage swings V_X and V_Y in the presence of a resonant load in the secondary (or primary) can be approximated as:

$$V_X = V_p \sin \omega_0 t + V_{DD} \quad (2.60)$$
$$V_Y = -V_p \sin \omega_0 t + V_{DD} \quad (2.61)$$

, where V_p the peak output voltage.

V_X and V_Y resemble sinusoids that are 180^o out of phase and have a dc level of V_{DD} (Figure 2.19).

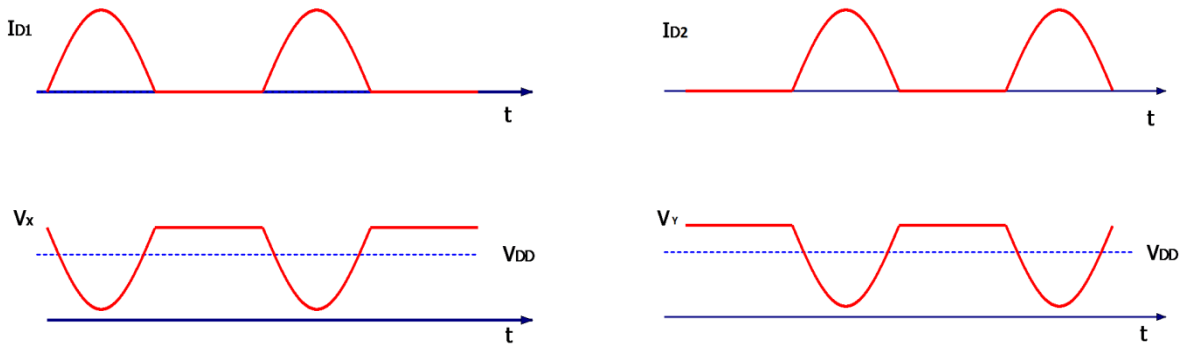


Figure 2.19. Current and voltage waveforms in a class B stage.

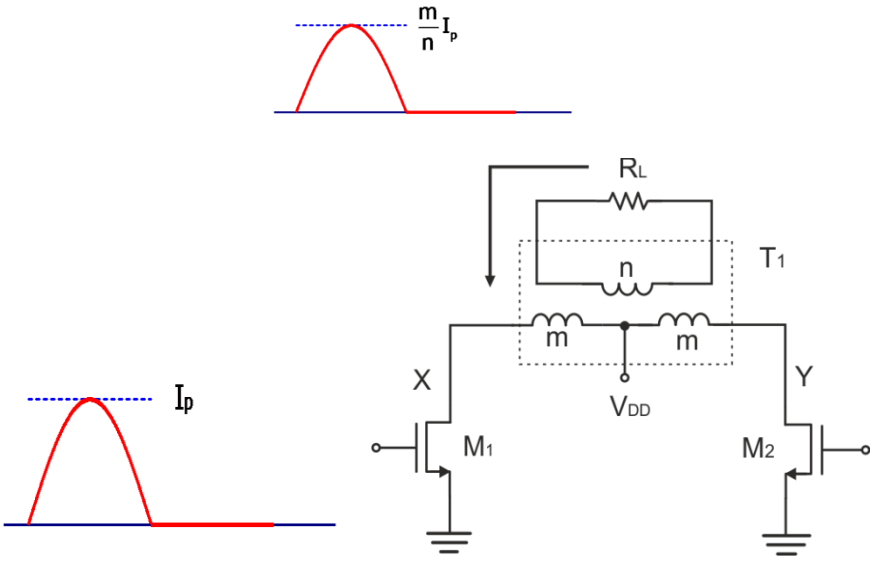


Figure 2.20. Class B circuit for efficiency calculation.

The average power delivered to R_L is equal to: $P_{out} = (\frac{m}{n})^2 \frac{R_L I_p^{\ 2}}{2} (2.62)$, where I_p the peak drain current, assuming the turn ratios shown in Figure 2.20.

The average power drawn by the nominal supply is equal to $P_{sup} = 2\frac{I_p}{\pi} V_{DD}$ (2.63)

When $V_{DD}=V_p$ (to maximize the efficiency) the drain efficiency becomes: $\eta = \frac{\pi}{4} \simeq 79\% (2.64)$

One of the problems that should be addressed with class B is that due to the conduction angle of each transistor of 180°, during the crossover from one transistor to the other unwanted distortion is introduced to the output signal. For this reason, it is called crossover distortion and to some degree depends on the amplitude of the output signal (the larger the amplitude the less significant it becomes). Also, the distortion will be less severe at high frequencies where the rate of change of the wave is much faster.

Class B operation often refers to half of the circuit shown in Figure, with the transistor still conducting for only half a cycle. The only requirement is that a resonant circuit must be placed in the output network of the transistor in order to reproduce the other half of the input signal. It is also quite non-linear but still has the same efficiency mentioned above. The non-linearity is caused by the harmonic distortion imposed to the amplified signal, which must be filtered.

The main disadvantage of a class B power amplifier is the use of transformers, which are bulky and expensive. Though, they are used more as an output stage compared to class A power amplifiers, due to their improved efficiency and its lower drain current. [5], [17], [18]

2.5.3. Class AB

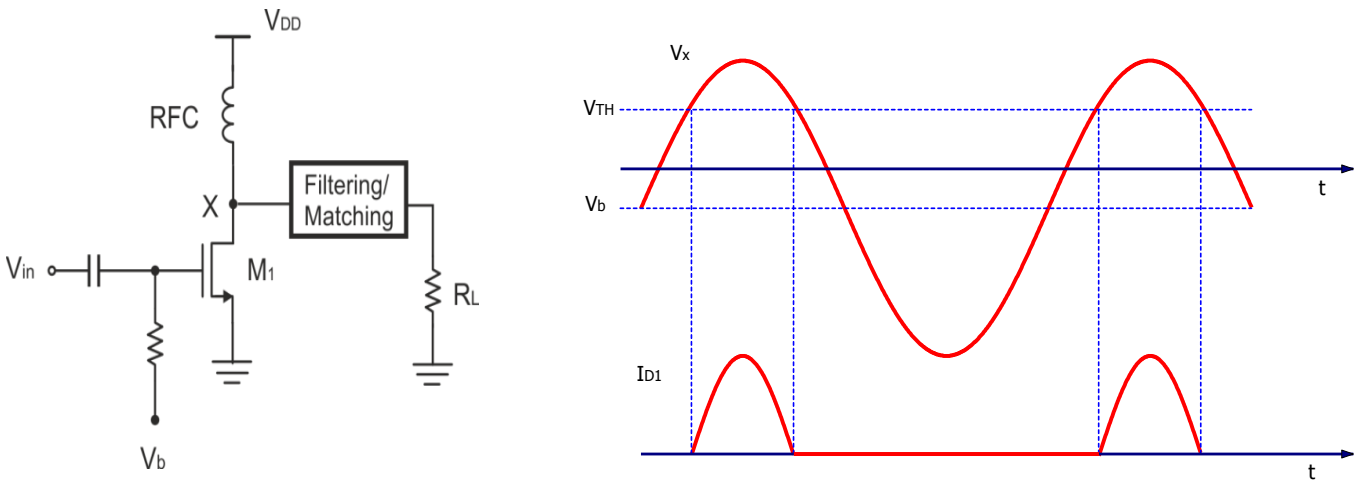
The class AB power amplifier is used to refer to a single-ended power amplifier whose conduction angle is between 180^o and 360^o, usually a Common-Source stage. It is less linear than a class A power amplifier and more than a class B and also a compromise of the two as far as the drain efficiency is concerned. It is not a linear amplifier; a signal with an amplitude–modulated envelope will be distorted significantly at this peak power level, because the conduction angle is a function of the drive level.

Conventional class AB operation incurs odd degree non-linearities in the process of improving efficiency. Theoretically efficiency can be increased up to 78.5%, if the transistor shall generate only even order non-linearities. Such a device will not generate undesirable close-to-carrier distortion. [5], [17], [18], [19]

2.5.4. Class C

A class C power amplifier is usually a narrowband amplifier (tuned power amplifier). The input signal in a class C stage is amplified to get large output power with an efficiency approaching 100%. The class C power amplifier is biased such that M_1 turns on if the peak value V_{in} raises V_x above V_{TH}, meaning that the Q point is placed below cut-off and so the transistor is cut-off for most of the ac cycle of the wave. V_x surpasses V_{TH} for only a fraction of the period, as if M_1 were stimulated by a narrow pulse as shown in Figure 2.21. Therefore, the device delivers a narrow pulse of current to the output every cycle. In order to avoid large harmonic levels at the antenna, the matching network must provide some filtering. The input impedance of the matching network is also designed to resonate at the frequency of interest, thereby making the drain voltage a sinusoid.

As the conduction angle decreases, the transistor is on for a smaller fraction of the period, thus dissipating less power. For the same reason, the transistor delivers less power to the load.



(a) (b)

Figure 2.21. Class C stage (a) and its waveforms (b).

If the drain current of M1 is assumed to be the peak section of a sinusoid and the drain voltage a sinusoid having a peak amplitude of V_{DD}, the efficiency can be calculated by:

$$\eta = \frac{1}{4}\frac{\theta - \sin\theta}{\sin\left(\frac{\theta}{2}\right) - \left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right)} \quad (2.65)$$, where θ the conduction angle.

The maximum efficiency of 100% is often considered a prominent feature of class C stages. The actual power delivered to the load is proven to be :

$$P_{out} \propto \frac{\theta - \sin\theta}{1 - \cos\left(\frac{\theta}{2}\right)} \quad (2.66)$$

Applying L'Hopital's rule, we can prove that P_{out} falls to zero as θ approaches zero.

A class C stage provides a high efficiency only if it delivers a fraction of the output power, which corresponds to full class A operation. It is consider to be a highly non-linear stage and requires large transistors. One of the major problems with utilizing Class C in solid-state applications is the large negative swing of the input voltage, which coincides with the drain output voltage peaks. This is the worst condition for reverse breakdown in any kind of transistor, and even small amounts of leakage current flowing at this point of the cycle have an important effect on the efficiency. In order to survive Class C operation, the transistor should have a voltage breakdown that is at least three times the active device's own dc voltage supply.[2]

Class C stage has a very improved efficiency in the cost of very heavy distortion of the output signal. It is however commonly used in high frequency RF amplifiers, where the pulses of current produced at the amplifier output can be converted to complete sine waves of a particular frequency by the use of LCR resonant circuits. [5], [17], [18]

2.5.5. Class E

Class E stages are switching- mode, tuned, nonlinear amplifiers that achieve high dc-to-ac conversion efficiencies approaching 100%, while delivering full power (Figure 2.22). The highest efficiency is obtained by minimizing the total power dissipated while the amplifier is delivering the desired output power. [5]

Class E stage consists of an output transistor M1, a grounded capacitor C1 and series network C2 and L1. One of the important advantages of the amplifier topology is that all parasitic shunt capacitances, including the output transistor capacitance C_{ds} and the parasitic capacitance of Lc, are absorbed into C1. [20]

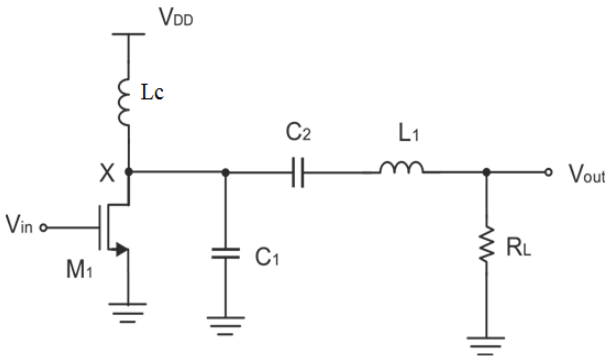


Figure 2.22. Class E stage.

Typically, class E power amplifiers can operate with power losses smaller by a factor of about 2.3, as compared with conventional Class B or C amplifiers using the same transistor at the same frequency and output power. [21]

The output transistor in this circuit operates as a switch rather than a voltage-dependent current source, ideally turning on and off abruptly. Called a “switching power amplifier”, such a topology achieves a high efficiency, in the actual design process, if:

1. M_1 sustains a small voltage when it carries current.
2. M_1 carries a small current when it sustains a finite voltage.
3. The transition times between the on and off states are minimized.

From the above 1 and 3 we conclude that the on-resistance of the switch must be very small and the voltage applied to the gate of M_1 must approximate a rectangular waveform. However, even with these two conditions, 2 may still be violated if M_1 turns on when V_x is high. Of course, in practice it is difficult to obtain sharp input transitions at high frequencies.

$R_{on} = \dfrac{1}{\mu_n C_{ox} \dfrac{W}{L}(V_{GS} - V_{TH})}$ (2.67), where μ_n is electron mobility, C_{ox} is gate capacitance per unit area, W and L are total gate width and length, respectively.

Class E amplifiers deal with the finite input and output transition times by proper load design. The values of C_1, C_2, L_1 and R_L are chosen such that Vx satisfies three conditions:

1. as the switch turns off Vx remains low long enough for the current to drop to zero, i.e., Vx and I_{d1} have non-overlapping waveforms (Figure 8(a))

2. Vx reaches zero just before the switch turns on (Figure 8(b)) and

3. dVx/dt is also near zero when the switch turns on.

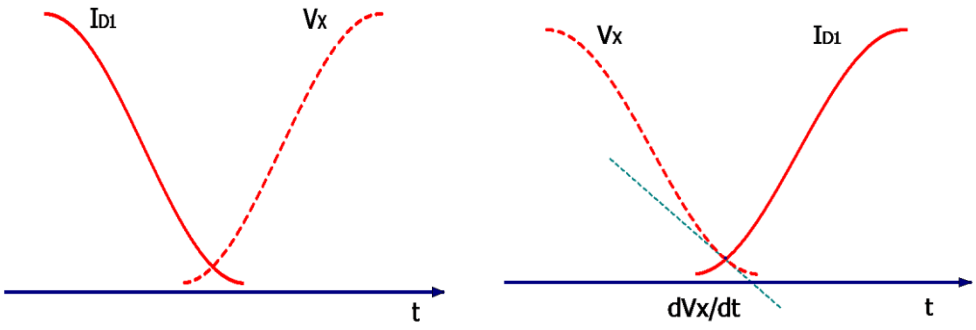


Figure 2.23.(a) condition to ensure minimal overlap between drain current and voltage, (b) condition to ensure low sensitivity to timing errors.

The first condition, guaranteed by C_1, resolves the issue of finite fall time at the gate of M_1. Without C_1, Vx would rise as V_{in} dropped, allowing M_1 to dissipate substantial power.

The second condition ensures that the V_{DS} and I_{D1} of the switching device do not overlap in the vicinity of the turn-on point, thus minimizing the power loss in the transistor even with finite input and output transition times.

The third condition lowers the sensitivity of the efficiency to violations of the second condition. That is, if device or supply variations introduce some overlap between the voltage and current waveforms, the efficiency degrades only slightly because dVx/dt=0 means Vx does not change significantly near the turn off point.

Class E stages are quite nonlinear and exhibit a trade-off between efficiency and output harmonic content. For low harmonics, the Q_L (Quality factor, loaded) of the output network must be higher than that typically required by the second and third conditions. In most standards, the harmonics of the carrier must be sufficiently small because they fall into other communication bands (note that low harmonic content does not necessarily mean that the PA itself is linear. The output transistor may still create spectral regrowth or amplitude compression).

Another property of the class E amplifiers is the large peak voltage that the switch sustains in the off state, approximately 3.56 $*V_{dd}$-$2.56V_{DS,sat}$, where $V_{DS,sat}$ is the minimum voltage across the transistor, raising serious device reliability or breakdown issues.[5]

Since the drain voltage of a class E PA ideally can reach such levels, a low supply voltage is needed to minimize the stress on the output stage transistors and to not exceed the critical gate oxide field of

$\sim 1\frac{V}{nm}$ for DC conditions. The damages in PAs are mainly due to hot carrier (HC) stress or Fowler-Nordheim (F-N) gate oxide wearout. Typically, in a class E PA, the drain voltage is high when the drain current is zero and therefore the HC stress is minimized and the transistor wearout will be dominated by the F-N gate oxide wearout.[21] The damages rather depend on the drain waveforms than on the dissipated power on the device. The electric field (or RMS voltage) is sufficiently high across the oxide to cause Time Dependent Dielectric Breakdown (TDDB). [22]

The class E amplifier includes an LC series band pass filter which is incorporated to create a short circuit at the desired switching frequency and blocks undesired harmonic components (Even though it causes harmonic distortion, because of the two different resonance frequencies of the filter). [23]

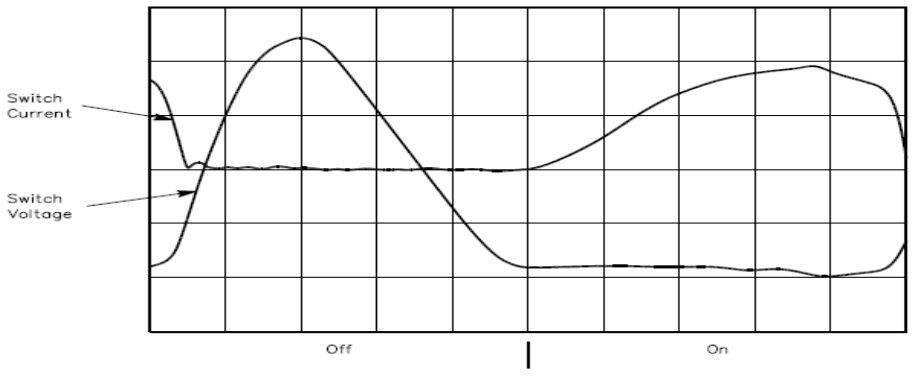


Figure 2.24. Actual transistor voltage and current waveforms in low-order Class E power amplifier. [24]

The values of R_L and V_{DD} are constrained by the requirement to deliver a specified power output to the load from the V_{DD} power supply, specifying one dictates the other. For highest efficiency, the highest possible V_{DD} should be used, within the V_{DD} limitation of the transistor, as seen above. The maximum allowable V_{DD} can be found from $3.56 * V_{DD} - 2.56 V_{DS,sat}$. The value of R_L can be calculated from the equation [5]

$$R_L = \left(\frac{2}{\frac{\pi^2}{4}+1}\right)\frac{(V_{dd}-V_{ds,sat})^2}{P_{out}} = 0.577\frac{(V_{dd}-V_{ds,sat})^2}{P_{out}} \quad (2.68)$$

Impedance transformation can be used if the load resistance is not equal to the value of R_L.The desired R_L may be chosen freely, according to design compromise to be made between efficiency and harmonic content of the power delivered to the load.

The amplifiers' output power P_{out} depends primarily (derivable analytically) on drain dc-supply voltage V_{DD} and the load resistance R_L, but secondarily (not derivable analytically) on the value of the chosen Q_L. [40]

Then, $L_2 = \frac{Q_L R}{2\pi f}(2.69)$, from the definition of Q_L.

To satisfy the conditions $V_{DS}=0$ and $\frac{dV_{DS}}{dt} = 0$ after switch turn off, given the chosen Q_L,

$$C_1 = \frac{1}{2\pi f R_L 5.447} \quad (2.70)$$

As the transistor becomes large, the drain capacitance increases and can be incorporated in the required capacitance C_1 for class E operation. However, it is important to minimize the voltage across the capacitance as the transistor is turned on, in order to minimize the energy losses. [22] On the other hand, power is lost due to the discharging of the shunt capacitor C_1 to the ground via the switch at the switch turn on moment. [25]

$$C_2 \approx \left(\frac{1}{(2\pi f)^2 L_2}\right)\left(1+\frac{1.42}{Q_L-2.08}\right) \approx C_1\left(\frac{5.447}{Q_L}\right)\left(1+\frac{1.42}{Q_L-2.08}\right)(2.71)$$

Note that L_2 is not resonant at the frequency of interest with C_2 or with the series combination of C_1 and C_2. Load reactance (if any) is accommodated be absorbing it into C_2 and L_2. The inductance of L_2 is decreased if the load series reactance is capacitative, or both are done if the load is a combination of inductance and capacitance. [25]

In order to complete the design equations, one should calculate the currents required for class E operation:

$$I_{DC} = \frac{P_{out}}{V_{dd}} \left[\frac{1}{1 - \frac{V_{ds.sat}}{V_{dd}}} \right] (2.72)$$

$$I_{ds.peak} = I_{DC} \left[1 + 1.862 (1 - \frac{0.5}{Q_L}) \right] (2.73)$$

Note that this is the minimum possible current required for an ideal case were the current fall time is zero. From this we know that the transistor needs to be designed to handle at least $I_{ds.peak}$. This could be achieved with a lot of different transistor sizes and combinations of parallel devices and number of fingers. [26]

A major obstacle in the design of class E CMOS PAs is the high peak drain voltage generated, and the low breakdown voltage of the MOS device, making it challenging to design a reliable class E PA.

The PAs shows slightly lower output power and efficiency, while delivering a sufficiently high output power to leave some margin for output matching network losses to have a certain power at the antenna. [27]

Due to the relationship between drain efficiency, $DE\,(2.74)$, switch on –resistance (R_{on}), and the load resistance (R_L) it is important to reduce the on-resistance for high output power and high power efficiency. Moreover, in order to achieve the same output power for a reduced power supply voltage, the load resistance needs to scale quadratically (eq.15). Since the on-resistance does not reduce as fast as R_L when technologies scale, a low-voltage high-efficiency PA require wider transistors in deep-submicron CMOS technologies.

$$DE \propto \frac{1}{1 + 1.14 \frac{R_{on}}{R_L}} (2.74)$$

In some designs deviating from the ideal proposed by the inventors a series inductance is used. When the switch is closed the conduction angle is 0^0 and when it is open it is 180^0, taking account that the duty cycle is 50%.So phase shifting is needed to meet the power dissipation requirements at the closing point. This is partially done by the series inductor introduced in the design.

$$L_s = \frac{1.1525 * R_L}{\omega_0} \quad (2.75)$$

Class E single stage power amplifier can be used also for energy harvesting, due to its signal output resembling a rectified SRRC pulse. [16]

2.5.6. Class F

Class F power amplifiers are another example of switching-mode power amplifiers (Figure 2.25).Their load network provides a high termination impedance at the second and third harmonics, the voltage waveform across the switch exhibits sharper edges than a sinusoid, thereby reducing the power loss in the transistor.

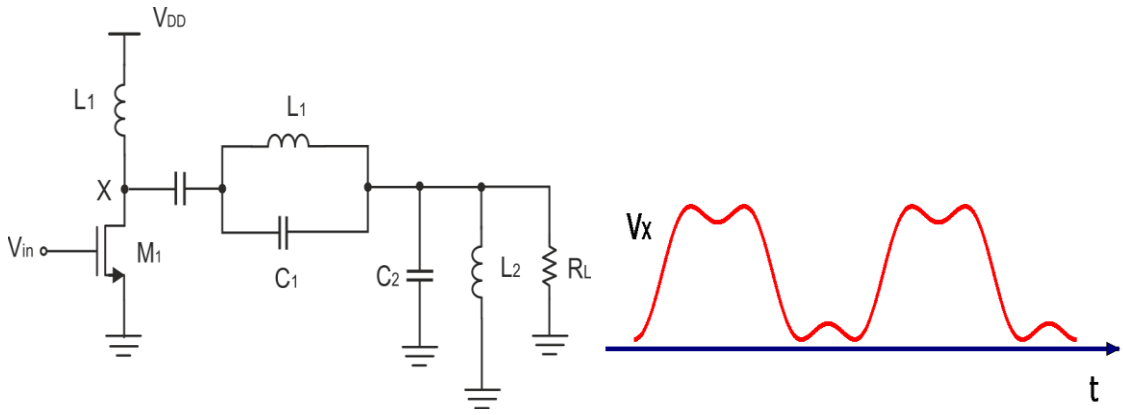


Figure 2.25. Class F stage.

A class F power amplifier consists of an inductor L1 and a capacitor C1 that form an LC tank, which resonates at twice or three times the input frequency, approximating an open circuit. Vx approaches a rectangular waveform with the addition of the third harmonic (Figure). This implies that third-harmonic peaking is viable only if the output transistor experiences heavy switching (its output current resembles a rectangular wave). This in turn requires that the gate voltage be driven by relatively sharp edges.

If the drain current of the transistor is assumed to be half-wave rectified sinusoid, the peak efficiency of class F amplifiers is equal to 88% from third-harmonic peaking. [5]

2.5.7. Summary

The quiescent DC operating point (Q-point) of an amplifier determines the amplifier classification. By setting the position of the Q-point at half way on the load line of the characteristic curve, the amplifier will operate as a class A amplifier. By moving the Q-point

lower down the load line changes the amplifier into a class AB, B or C amplifier. In a Class E and F power amplifier, the Q-point plays not so much a role. Then the class of operation of the amplifier with regards to its DC operating point (Figure 2.26) can be given as: [28]

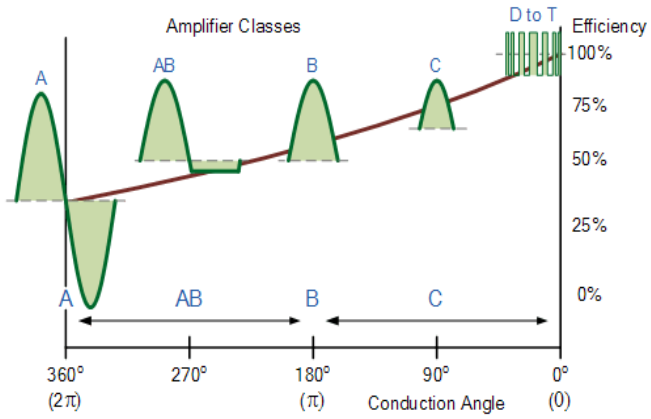


Figure 2.26. Amplifier classes and efficiency.

Although class F power amplifiers have excellent efficiency, ideally 100%, the amplifiers require harmonic termination circuits which occupy large area. Thus, class F amplifier is not suitable for integration. Class E power amplifiers are suitable for integration, since only three lumped element components are required to realize switching operation at the output port.

Conduction angle is defined as the percentage of the signal period during which the transistor remains ON multiplied by 360^{o}.

Chapter 3

Design of a 5GHz Integrated Radio Frequency Power Amplifier in 90nm CMOS Technology

3.1. 90nm TSMC CMOS Technology

In this chapter, the design process of the Power amplifier with a beneficial bandwidth of 800MHz around the operating frequency of 5GHz will be thoroughly discussed. The maximum output power of the amplifier was set to 10dBm and the maximum input power to 0 dBm. The design was made in Virtuoso® analog design environment of Cadence®. The design kit used is the 90nm RF Library from TSMC©. There are 4 MOSFET models in this release and are summed in Table.

Nominal Supply Voltage	Type	Number of fingers
1.2 V	NMOS (thin substrate)	≤ 64
1.2 V	PMOS (thin substrate)	≤ 64
2.5 V	NMOS (thick substrate)	≤ 64
2.5 V	PMOS (thick substrate)	≤ 64

Table 3.1.MOSFET models.

Models 1.2 V N/PMOS are BSIM4 models; in contrary 2.5 V N/PMOS are BSIM3. In addition, the maximum power of model valid range is limited to 0.21W and the nominal temperature is 25^0 C, while the valid temperature range is from -40^0 C to 125^0. The minimum length is 0.1um and the maximum width is 10um. The type of the substrate is a p-substrate. For the design, a 1P9M stack is available using ultra-thick top metal and an optional layer M10.The interconnection of this technology utilizes Low-K inter-metal dielectrics.

Design Flow:

1. Choosing the appropriate devices and dimensionalizing.
2. Building the design with ideal passive components.
3. Replacement of the ideal components with passive components provided by design kit.
4. Design rules check via simulations.
5. Performance criteria check via simulations.

Most of these five steps were repeated several times to achieve the requirements of the design.

3.2. RF chokes and Inductors

As the name implies, inductors for high-frequency circuits are used in the high-frequency band from 10MHz to several GHz. Such inductors (chokes) require a high Q (Quality factor) value, most have a non-magnetic core structure, and they are mainly used in the high-frequency circuits of mobile communications equipment, such as mobile phones, wireless LAN, and others. [29]

A "choke" is the common name given to an inductor that is used as a power supply filter element. They are typically gapped iron core units, similar in appearance to a small transformer, but with only two leads exiting the housing. The current in an inductor cannot change instantaneously; that is, inductors tend to resist any change in current flow. This property makes them good for use as filter elements, since they tend to "smooth out" the ripples in the rectified voltage waveform. A choke is used in place of a series resistor because the choke allows better filtering (less residual AC ripple on the supply, which means less noise in the output of the amplifier) and less voltage drop. An "ideal" inductor would have zero DC resistance. If you just used a larger resistor, you would quickly come to a point where the voltage drop would be too large, and, in addition, the supply "sag" would be too great, because the current difference between full power output and idle can be large, especially in a class AB amplifier. [30] The frequency selectivity of an RF choke is simply to pass DC and block everything else. Unfortunately, blocking very low frequencies requires a very large inductor. When designing a filter the Q, or quality factor, of an inductor is very important and the goal is a high Q component. When selecting an RF choke, a high Q inductor isn't always the best choice. High Q means low loss, but for an RF choke higher loss means better isolation. This means for a given inductance, a lower Q choke has more parasitic resistance than a high Q choke. Another consideration is that Q is frequency dependent, often heavily so. The Q-factor describes a ratio of an inductor's reactance to effective resistance. This value is frequency dependent, and therefore the test frequency is often specified.
Specifically, Q affects the sharpness of a resonant filter and the center frequency of an LC circuit. Because a high value of Q is preferred, datasheets specify a minimum Q value.

Another main factor that should be considered is the self- resonant frequency (SRF). SRF simply describes the frequency at which an inductor quilts working as an inductor. For RF designs, an SRF should be chosen with a minimum value that exceeds the operating frequency of one's circuit. [38]
Additionally, an inductor's tolerance should be considered too. [31] Tolerance percentages inform an engineer that the part they are using could have an inductive value different from what is listed in the datasheet by a certain percent.

3.3. MOS transistors

RF transistors are also called power transistors, due to the fact that they are used for designing power amplifiers or low noise power amplifiers. In Figure 3.1, a BSIM4 model NMOS transistor can be seen, all parasitic capacitances and resistances included. The "parasitics" of a MOS transistor correspond to all non-intentional and non-avoidable passive or active devices that exist around the MOS transistor. Namely, they are the parasitic capacitances between different regions of the MOS transistor, the resistances associated with the several terminals of the device, the p-n junctions that are the integral parts of the MOs transistor. These parasitic p-n junctions are normally reverse-biased such that their currents are negligibly small, also called substrate currents.

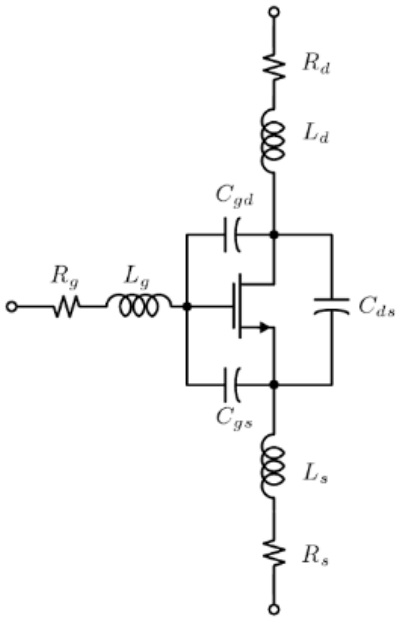


Figure 3.1.NMOs transistor (with respect to a BSIM4 intrinsic model).

These parasitic elements are geometry dependent. Some of them are also technology and bias dependent. C_{gs}, the total gate-source capacitance is the sum of the gate capacitance C_g corresponding to the carrier charge in the inversion layer that is induced and controlled by the gate-source voltage and the gate-source overlap capacitance C_{gso}. C_{gd}, the total drain-gate capacitance consists only of the drain-gate overlap capacitance. Although, this is a small capacitance, it has a very important influence on the high frequency performance of the MOS transistors (Miller effect). C_{ds}, is the total drain-source substrate capacitance, which is the total output capacitance of the MOS transistor.

An important definition for a MOS transistor that is related to the parasitic capacitances of the device is the "high frequency figure of merit", f_T. It is known that the low frequency input current of a MOS transistor is practically zero, and consequently the low frequency current gain

is infinite. At higher frequencies, on the other hand, the capacitive current that flows into the gate terminal becomes non-negligible. Hence, the current gain decreases at high frequencies. f_T, is defined as the frequency for which the small signal current gain is equal to unity (cut-off frequency). In RF design, a high value of f_T is critical. So, the smallest possible length of transistors should be used, due to the fact that $f_T \propto \frac{V_{DS,sat}}{L}(3.1)$. Another important, parameter is f_{max}, the maximum usable frequency. f_{max} , is the frequency at which the power gain is maximum and equal to unity (maximum oscillation frequency). It is known that under conjugate match, $f_{\max} = \frac{1}{2}\frac{f_T}{\sqrt{2\pi \times f_T \times C_{gd} R_g + \frac{R_g}{r_o}}}(3.2)$, where R_g is the parasitic gate resistance of the device and r_o the total intrinsic resistance of the device. This frequency depends on the magnitude of R_g and C_{gd}. So, if the transistor has to operate at a high frequency, a multi –finger structure must be used to decrease the parasitic resistance R_g and capacitance C_{gd} of the device (division of transistor W by N (number of fingers) and parallel connection). It is mandatory to evaluate this frequency together with the high frequency figure of merit, f_T. [32], [33]

Finally, it has been shown that, as a result of constant-field scaling, the peak $f_T(\sim 0.3mA/\mu m)$, peak $f_{\max}(\sim 0.2mA/\mu m)$ and optimum noise figure $NF_{\min}(\sim 0.15mA/\mu m)$ current densities are unchanged from one technology node to another, applying to technologies from 0.25 μm down to 90nm.[34]

For the design of the selected Power amplifier topology, NMOS transistors were used with the smallest channel length available (100nm). The threshold voltage in this 90nm process can reach up to 0.5 V.

3.4. The selected power amplifier topology

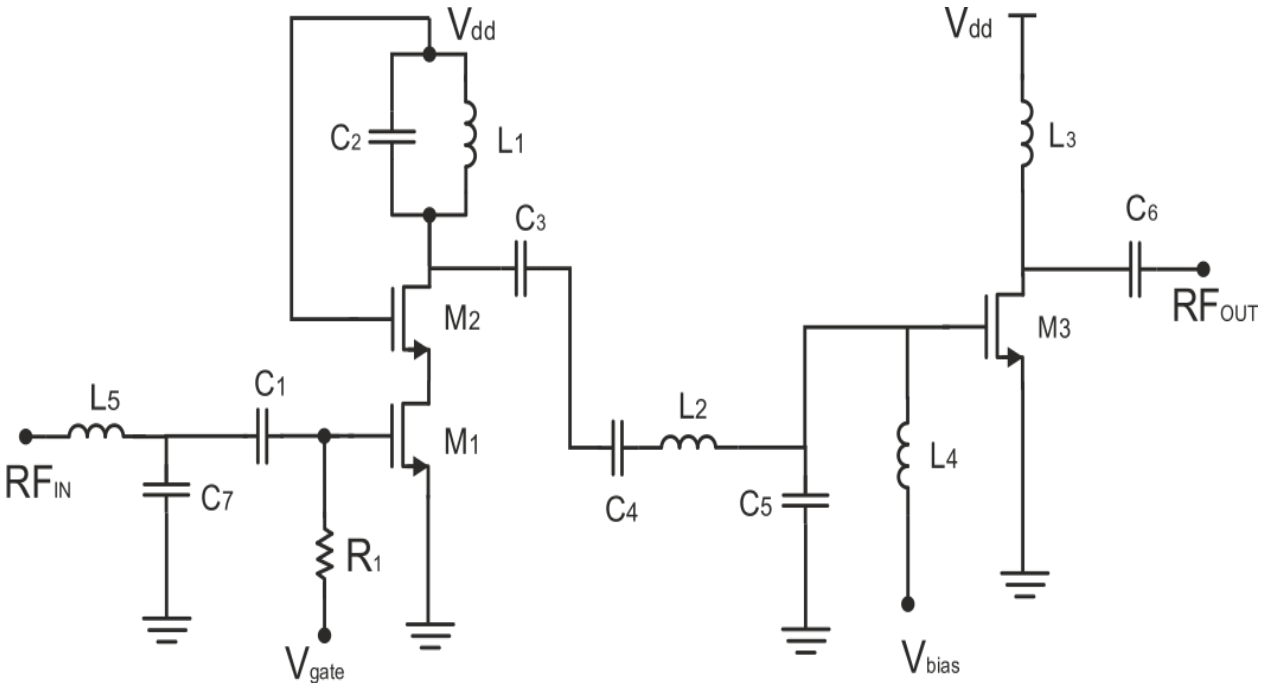


Figure 3.2. Selected Power Amplifier topology.

As mentioned in previous section, there are many topologies to choose from. The trade-off between efficiency and linearity is the most trivial one. One thought has been to design a high efficiency switching mode power amplifier combined with a linearization technique, such as Kahn's. This idea was rejected due to fact that reliable switching mode power amplifiers couldn't be designed in such a high frequency. So, the topology selected can be seen in Figure 3.2, owing to its higher linearity. The design of the Power amplifier consists of two stages, a cascode stage (gain stage) and a common source stage, in class A biasing. The transistors used are 1.2 V NMOS transistors because they show higher transconductance (g_m) compared to PMOS transistors of the same dimensions. Transistors M_1 and M_2 form the cascode pair. Transistor M_2 is called "cascode" transistor and works as a common-gate stage, when transistor M_1 works as a common-source. Transistor M_3 works as a common-source stage biased as class A. Furthermore, capacitors C_1, C_3 and C_6 are DC block capacitors, enabling each stage to be biased in its own conditions. Inductor L_1 and capacitor C_2 form an L-C resonator (L-C tank) for frequency selection. Inductor L_5 and capacitor C_7 are an L-network for input impedance matching to 50 Ohms. In addition, capacitor C_4, inductor L_2 and capacitor C_5 form a matching

network with compensation used for interstage matching. Finally, inductor L_4 is used for the biasing of the output stage in class A operation and inductor L_3 is used as a choke for drain biasing of the stage. The output impedance of the amplifier is considered to be 50 Ohms and the operating frequency is 5GHz.

3.5. Design methodology

In this section, the design methodology will be discussed. The design presented here is in schematic level.

3.5.1. Output stage

The design process always begins with the output stage. It is designed in such a manner that can support the desired output power. More elaborately, as mentioned in previous subsection, f_T and f_{max} play a significant role in choosing the biasing conditions of a device and transistor dimensions. So, all parasitic capacitances and resistances are bias controlled.

The output stage is a class A single stage power amplifier designed to support 10dBm output power. The ultimate goal is to bias the common-source transistor to have a current density approximately 0.2mA/μm, meaning that the total current would swing around this DC current point.

Going one step back, the maximum output voltage of this power amplifier is equal to $V_{max}=2V_{dd} - V_{DS,sat}$. Choosing $V_{DS,sat}= 0.25V$ and the appropriate current swing, multiple periodic steady state simulations were performed. Thus, the output power of the amplifier can be written as:

$$P_{out} = \frac{W \times I_{swing} \times (V_{dd} - V_{DS,sat})}{4} \quad (3.3)$$

From the above equation we can calculate the width of the device and was found to be W=320um. The DC current is chosen to be:

$$I_{DC} = 0.2(mA/um) \times W(um) \approx 62mA \quad (3.4)$$

With a DC simulation, the biasing for the common-source transistor was found to be 0.8 V and the dimensions of the device are 5um x 100nm x 64 (Figure 3.3).

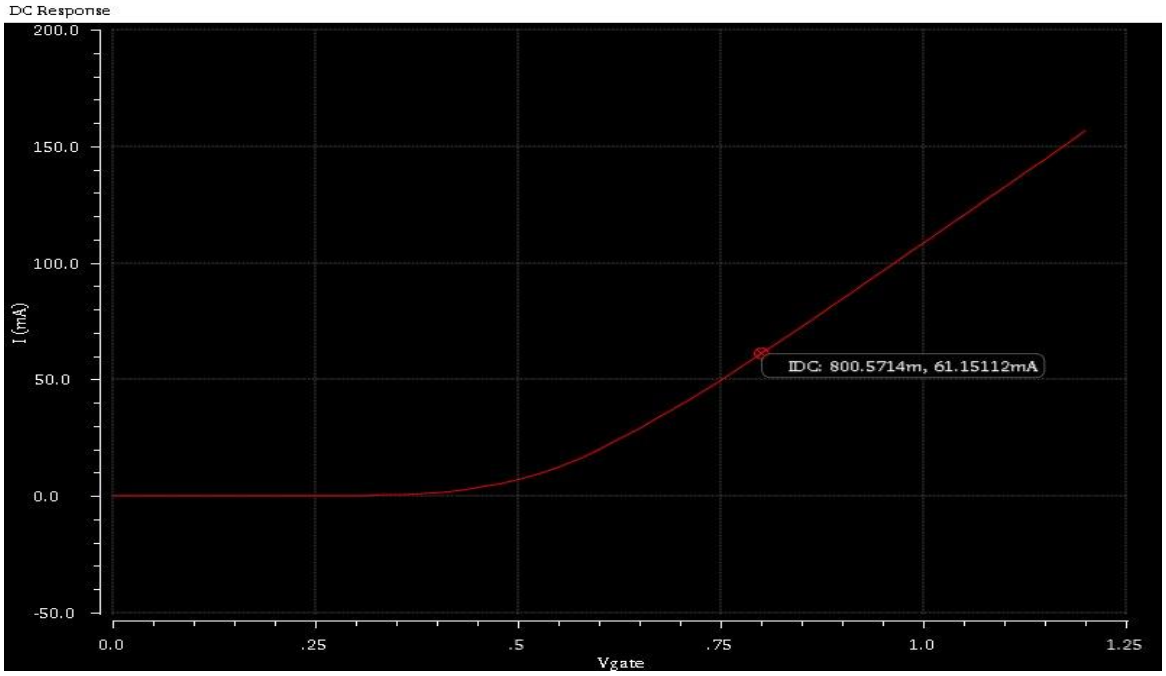


Figure 3.3.DC current point of output stage versus gate voltage.

In this stage, frequency selection is performed by a filter shown in Figure 3.4, including the parasitic capacitances of the device. The center frequency of the filter can be calculated with the following equation:

$$f_{5GHz} \approx \frac{1}{2\pi\sqrt{L_3 C_{ds}}} \quad (3.5)$$

The device dimensions that were chosen, given the biasing conditions, give us a C_{ds}= 48.3fF and the inductor needed is L_3= 900 pH, in order to resonate in the desired frequency. In this stage the phenomenon of electromigration is apparent due to the larger current value in contrary with the cascode stage, where the current is much less and will be discussed in the following subsection 3.5.2.

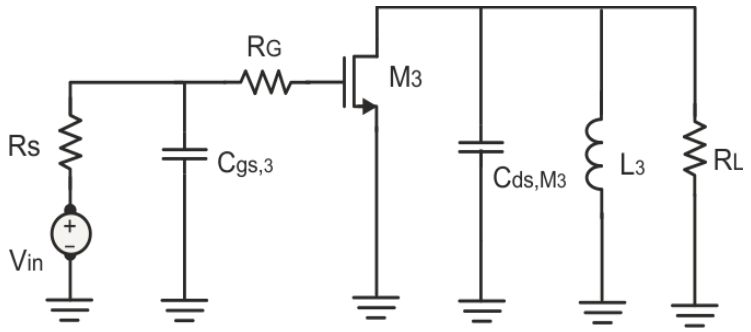


Figure 3.4. Small-signal high-frequency ac equivalent of output stage.

The final step of the design process for output stage is the choice of DC block capacitor C_6 and the value of inductor L_5 for the gate biasing. The DC block capacitor was chosen to be C_6=1 pF, in order not to affect the scattering parameters. Finally, inductor L_3 was chosen to be of 950 pH with an acceptable quality factor of Q=20. The inductor L_4=962pH was chosen as such, so it doesn't affect resonant frequency of the stage. In Figure 3.5 and Figure 3.6, the small-signal gain and the actual output power of the output stage are presented. At this point, we should mention that the power amplifier was designed to support a higher output power but because of the parasitic capacitances introduced by the passive elements required to conclude the design, the output power is limited to 10dBm. The passive elements used in the lower part of the frequency spectrum are much larger than in 30 GHz for example and they introduce significant non-linearities, despite the fact that a class A power amplifier is considered the most linear of them all.

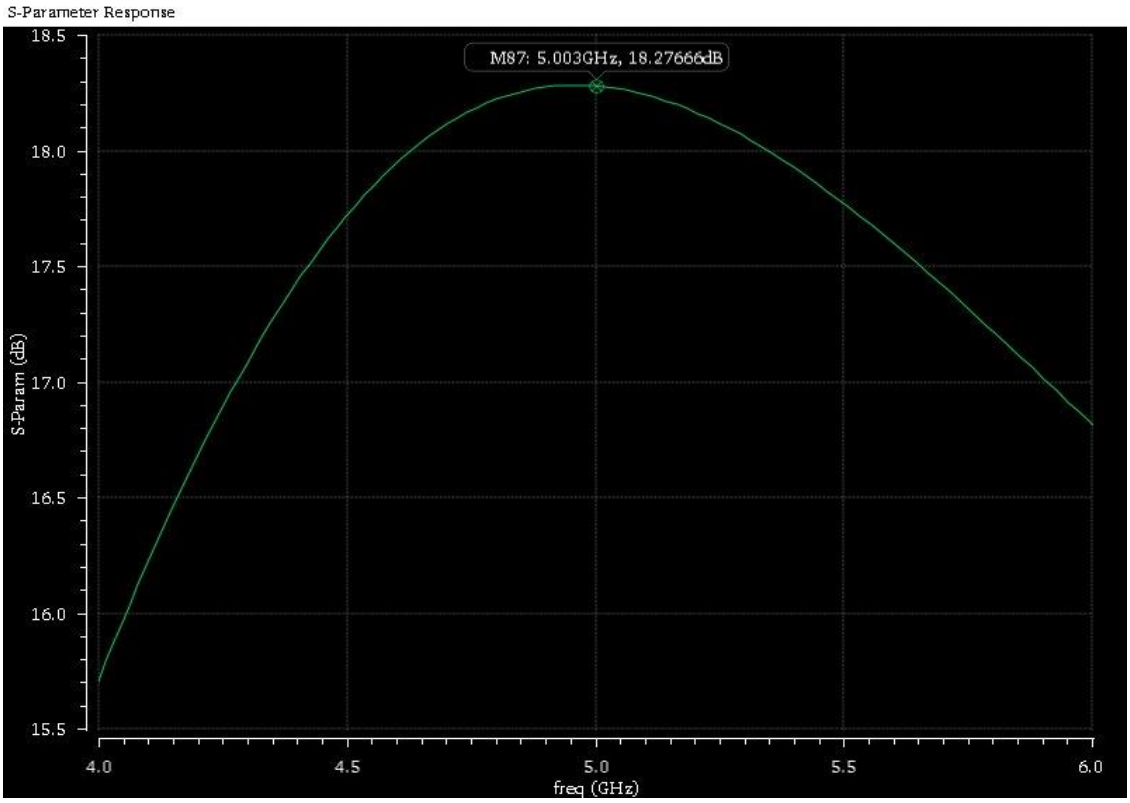


Figure 3.5. Small-signal gain of output stage.

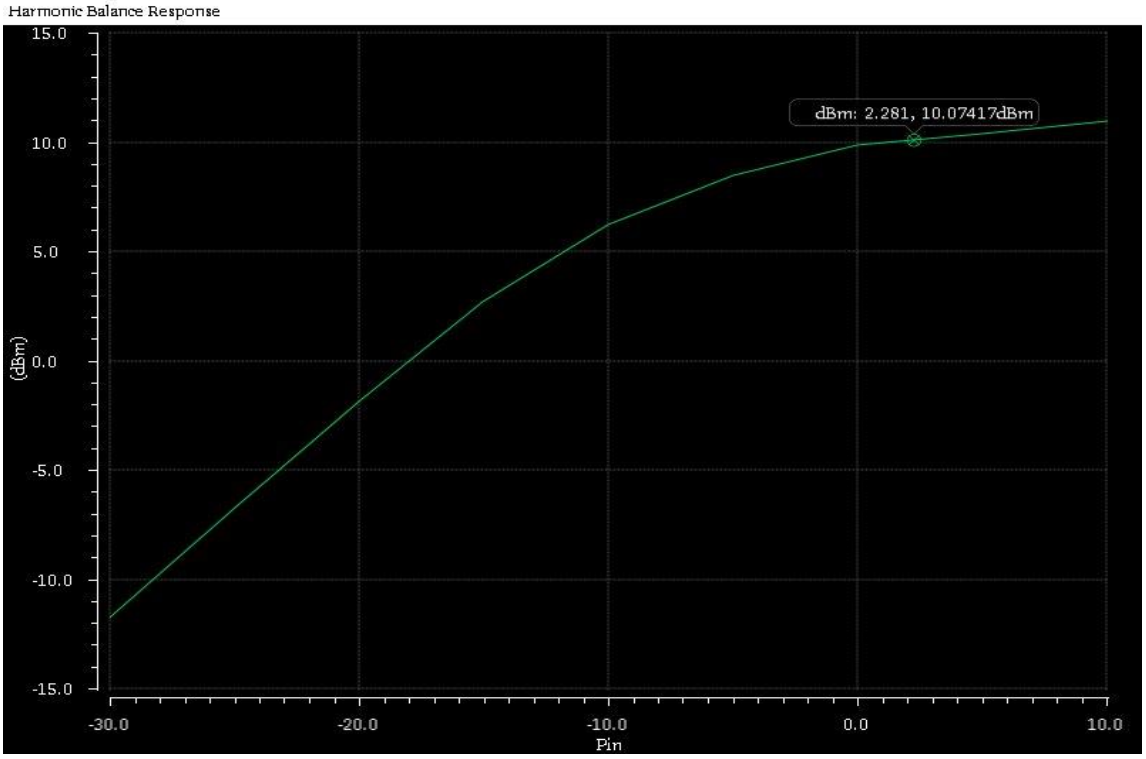


Figure 3.6. P_{out} versus P_{in} of output stage.

3.5.2. Cascode stage

From a cascode stage (Figure 3.7), the maximum small-signal gain that could be obtained can reach approximately 10dB. In our design, as seen in Figure 3.8 the small signal gain is 6.8dB and as a consequence it doesn't affect the overall gain compression of the design. Both transistors are in saturation and the dimensions of the transistors were chosen in respect to the electromigration phenomenon and the fact that the parasitic capacitances in cascode configurations are much greater. Of course, the length of the transistors is the smallest possible 100nm. And so, the dimensions of the transistors are 5um x 100nm x 10 fingers and 5um x 100nm x 50 fingers, for the common-source and common-gate transistor, respectively. Numerous pss analyses (periodic steady-state) were made to decide the gate biasing of each transistor. The common-gate or cascode transistor was biased with 900mV, equal to the DC supply voltage for ease and the common-source transistor was biased with 800mV. In Figure 3.8, the DC current of the cascode stage is presented and we can see that is much less than the DC current of the output stage. The common-gate transistor acts as a current buffer and a shield, as the output voltage rises, to the common-source transistor, where the current density is $\approx 0.2mA/um$.

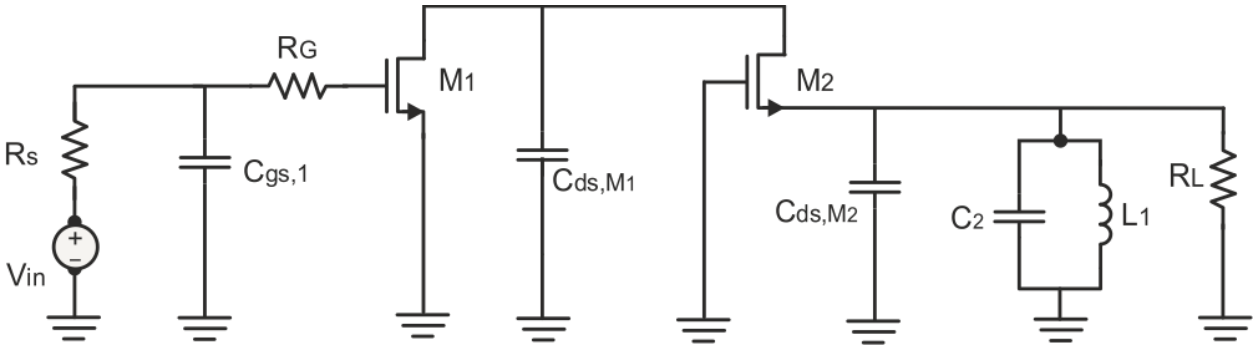


Figure 3.7. Small-signal high-frequency ac equivalent of cascode stage.

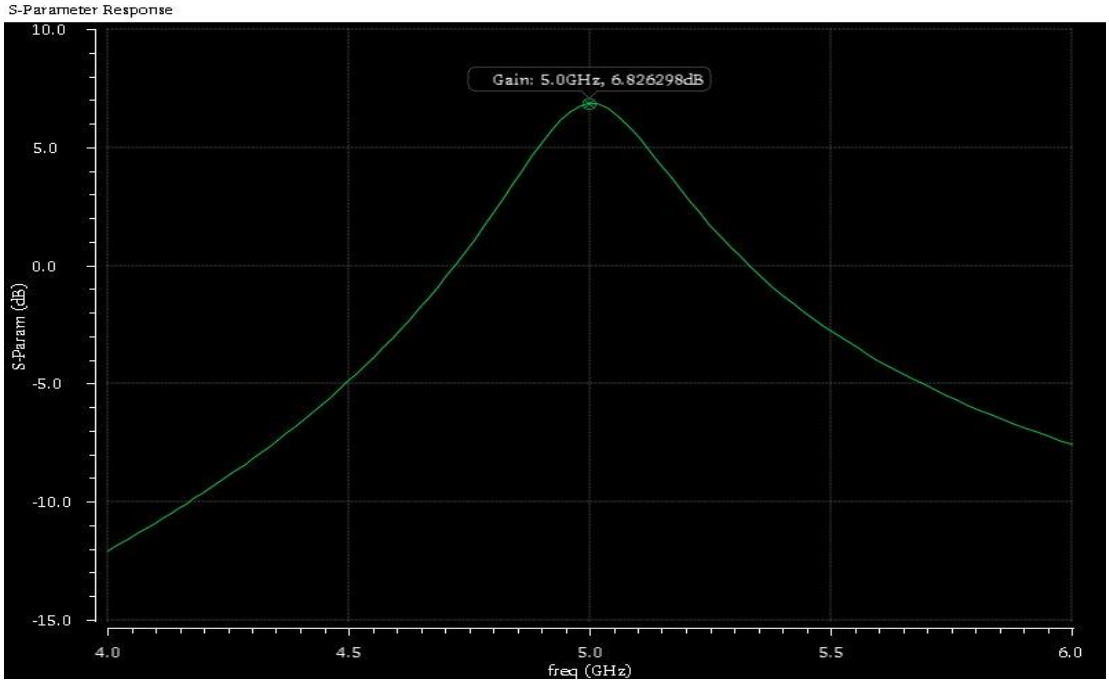


Figure 3.8. Small-signal gain of cascode stage.

Furthermore, as mentioned previously, the cascode stage topology chosen includes an L-C tank (resonator) for frequency selection. In Figure 3.9 the equivalent of a resonator can be seen. We should also mention that the quality factor of such a network is equal to:

$$Q = 2\pi \frac{The\ total\ energy\ of\ \ the\ system}{The\ energy\ lost\ in\ one\ period}$$

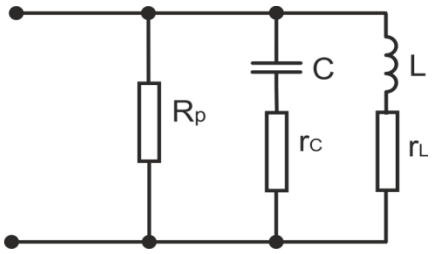


Figure 3.9.LC resonator equivalent, where Rp: represents the internal resistance of the signal source and other parallel losses, rL: represents all losses related to inductor, rc: represents all losses related to capacitor

The effective Quality factor reduces to the following:

$$Q_{eff} = \frac{R_{eff}}{\omega L} = \omega C R_{eff}$$

In respect to the above relation the natural frequency of the resonator is $\omega = 2\pi \times 5GHz(rad/\sec)$ and inductor L_1 has a value of 220pH (the smallest possible provided by the design kit) with a quality factor Q=34. The capacitor C_2 is of 4.7 pF and also the parasitic capacitance of the cascode transistor $C_{ds,M2}$ participates in the resonator. A small-signal high-frequency equivalent of the cascode stage can be seen in Figure 3.7. Resistor R_1 is equal to 109 Ohm. DC block capacitor C_1 has a value of 4pF and DC block –coupling capacitor C_3 a value of 1pF. Capacitor C_1 and resistor R_1 play part in the input matching of the design and they were chosen with care. Finally, C_3 capacitor was chosen as such of a big value in order not to have an effect on the scattering parameters.

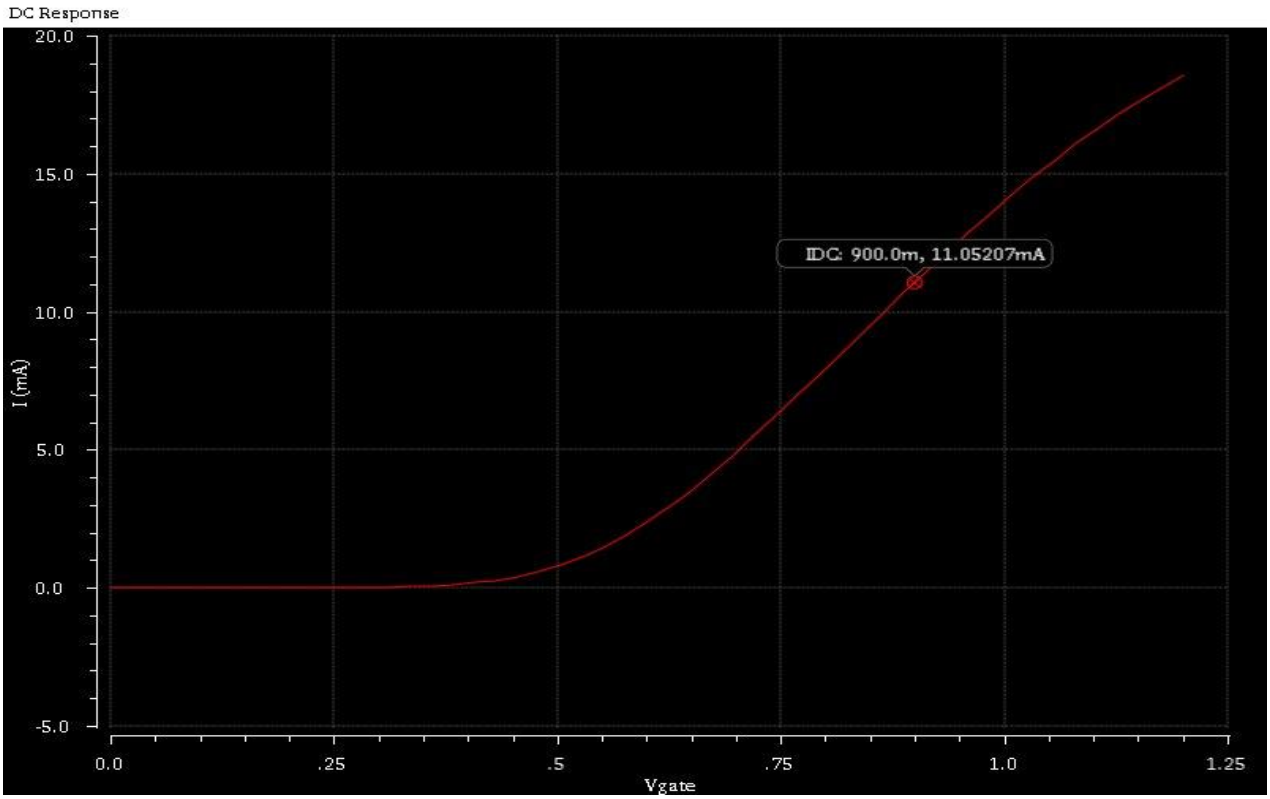


Figure 3.10.DC current point of cascode stage versus gate voltage.

3.5.3. Interstage matching

Any PA stage requires a driver. Before the details of circuit design are considered, it is important to decide how much drive power, and at what linearity, a particular case requires. Clearly, in applications where non-linear amplifiers are acceptable, and the PA stage is a saturated design, there is correspondingly no issue with regard to driver linearity. Driver chains become more challenging when overall linearity is important. The key issue is to control the generation on non-linear products in the driver stages while maintaining efficiency and power supply specifications.

A basic device chain obtained, it then remains to design the interstage matching network. In principle, the output power matching for a driver stage it is assumed to that the target impedance would be the output 50 Ohm termination. In a single-ended design, as ours, such matching networks are challenging for a number of reasons. First, the target or load impedance will have reactive component, being now the input impedance of another RF transistor. That gives it a frequency-dependent characteristic, a figurative moving target. Such is not the case when matching to 50 Ohm. There is also a serious stability issue. Even though a simple lowpass matching structure may appear, as usual, to do a decent job for the required inband response, it is

likely that the out-of-band response will place the low, highly reactive PA transistor input impedance directly across the driver output at some lower frequency, an odds-on stability conflict. Typically, stability can be restored only by the placement of lossy elements that will increase the driver power requirements. An additional complication of bias insertion is that an interstage matching network has to include both output and input bias insertion networks for the driver and PA devices, respectively, including DC block capacitors.

Troubleshooting is more difficult when the stage under scrutiny is embedded in an assembly with no 50 Ohm test points. The balanced amplifier approach has important attractions in removing, or bypassing most of the problems mentioned. Not only is each matching operation reduced t a simple 50 Ohm target, but the individual stages can be tested, either as an offline subcomponent development or even in-situ due to the provision of 50Ohm interface points. The driver output and PA bias networks are separated and can be designed at the 50 Ohm level. [39] This technique is used for our design, too. In our case, conjugate pair matching was impossible, due to restrictions in availability of certain inductor values.

The first step was to match the output impedance of the cascode stage to a 50 Ohm level as seen in Figure 3.11 and Figure 3.12, at the frequency of interest. In Figure 3.11, we can see the load transformation on a Smith chart and in Figure 3.12, the s_{22} parameter for confirmation that the target level was achieved.

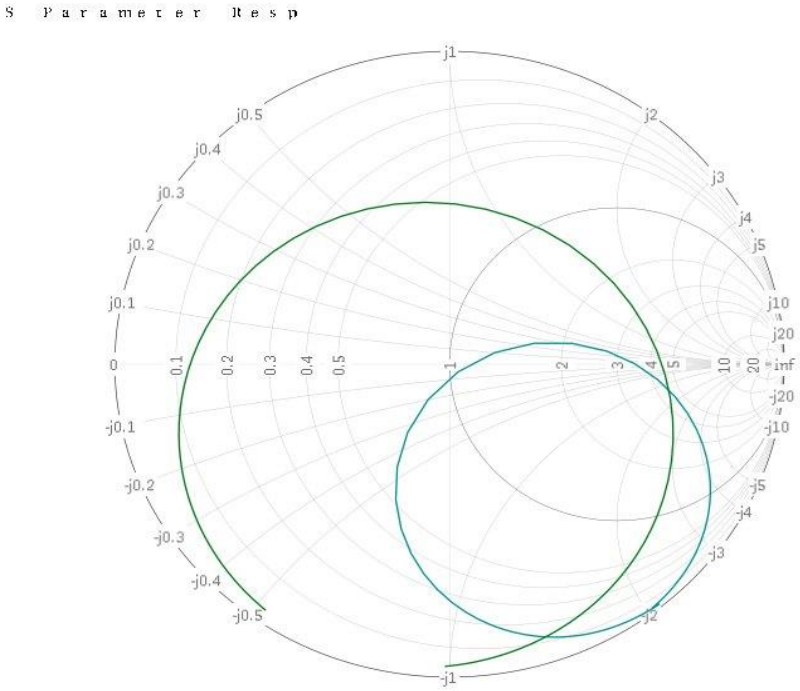


Figure 3.11. S_{22} of cascode stage after load transformation for the interstage matching Smith chart.

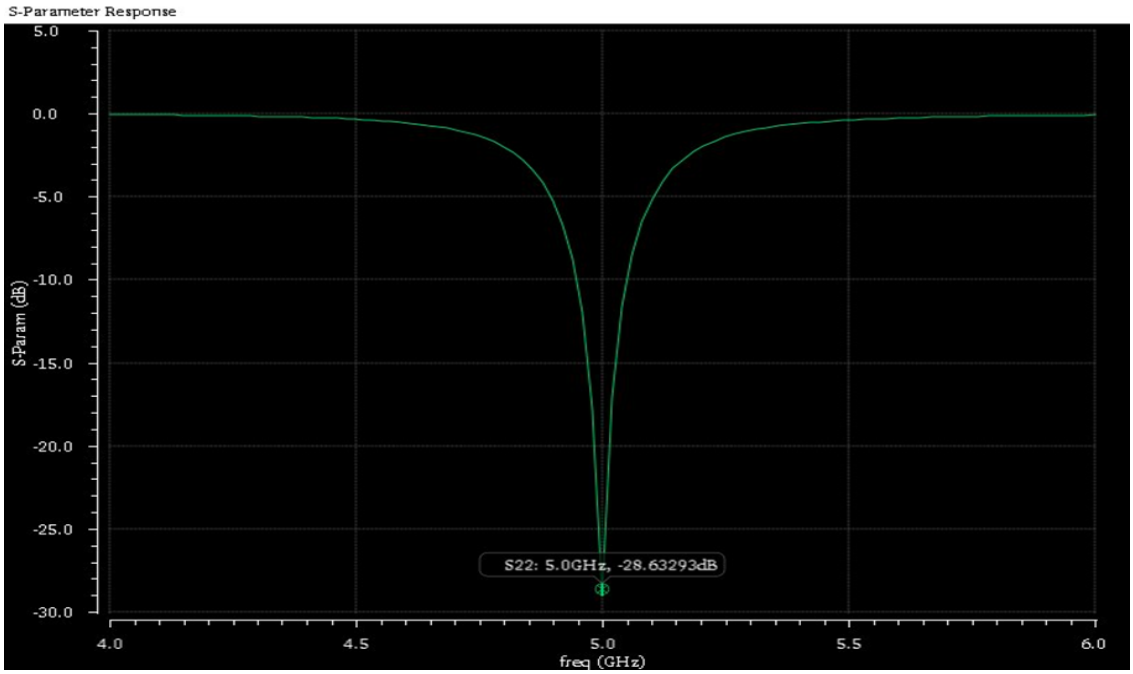


Figure 3.12.S_{22} parameter of cascode stage after matching.

Second step, was to match the input impedance of the output stage to 50 Ohm level, as seen in Figure 3.13 and Figure 3.14 at the frequency of interest.

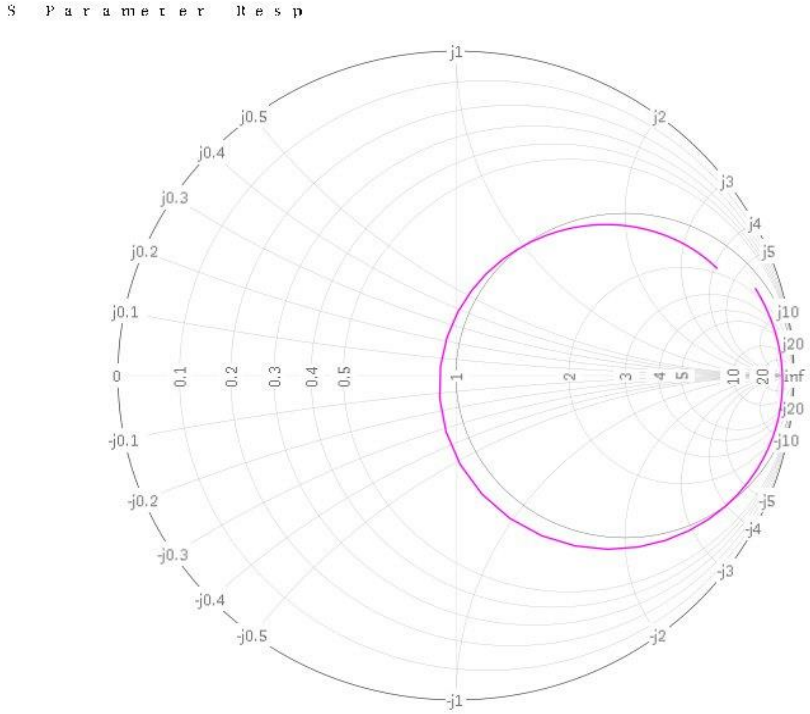


Figure 3.13.S_{11} of output stage after load transformation for the interstage matching Smith chart.

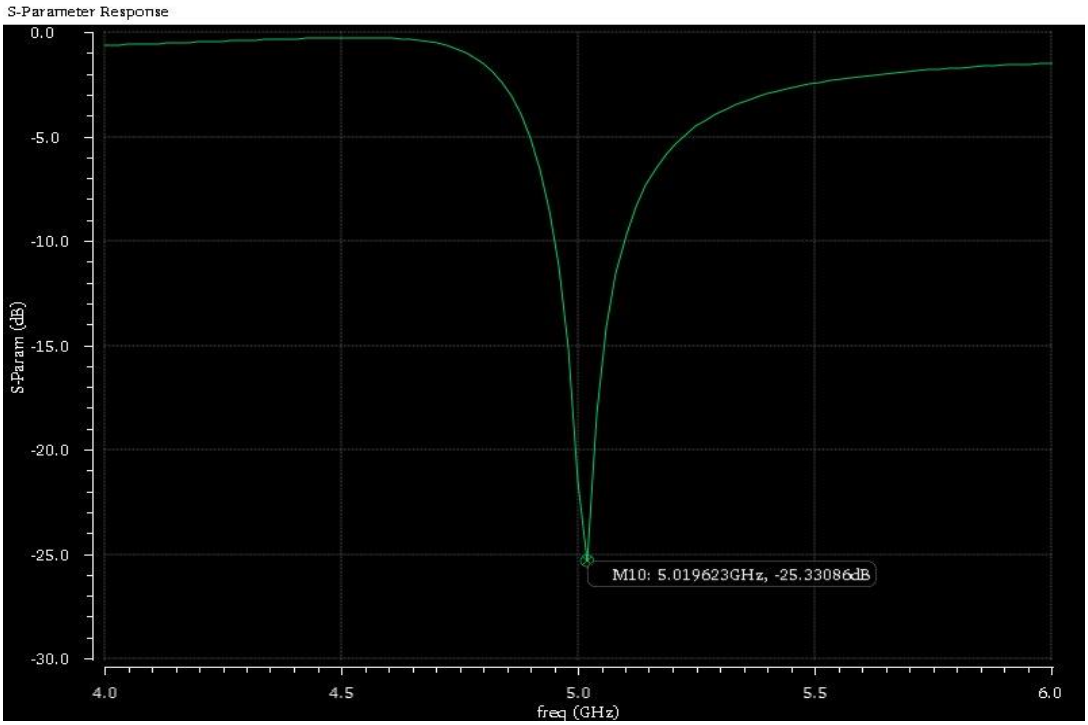


Figure 3.14. S_{11} parameter of output stage after matching.

The interstage matching was achieved with the use of an L-network with compensation. The components used for the network are C_4, C_5 and L_2 with values 500fF, 250fF and 7nH, respectively.

3.5.4. Input matching

For the completion of our design, we should calculate the load transformation of the input impedance to a 50 Ohm target level. Inductor L_5 and capacitor C_7 form an L matching network with values of 1.9nH and 200fF, respectively. We should also mention that the reactive part of C_1 DC block capacitor was taken into account for the process of the input matching of the design. In Figure 3.15 and Figure 3.16, the results of the process are presented. In Figure 3.15, we can observe the s_{11} scattering parameter of the overall design and in Figure 3.16, the load transformation on a Smith chart at the frequency of interest, where number 1 on the unity circle representing 50 Ohm, is reached.

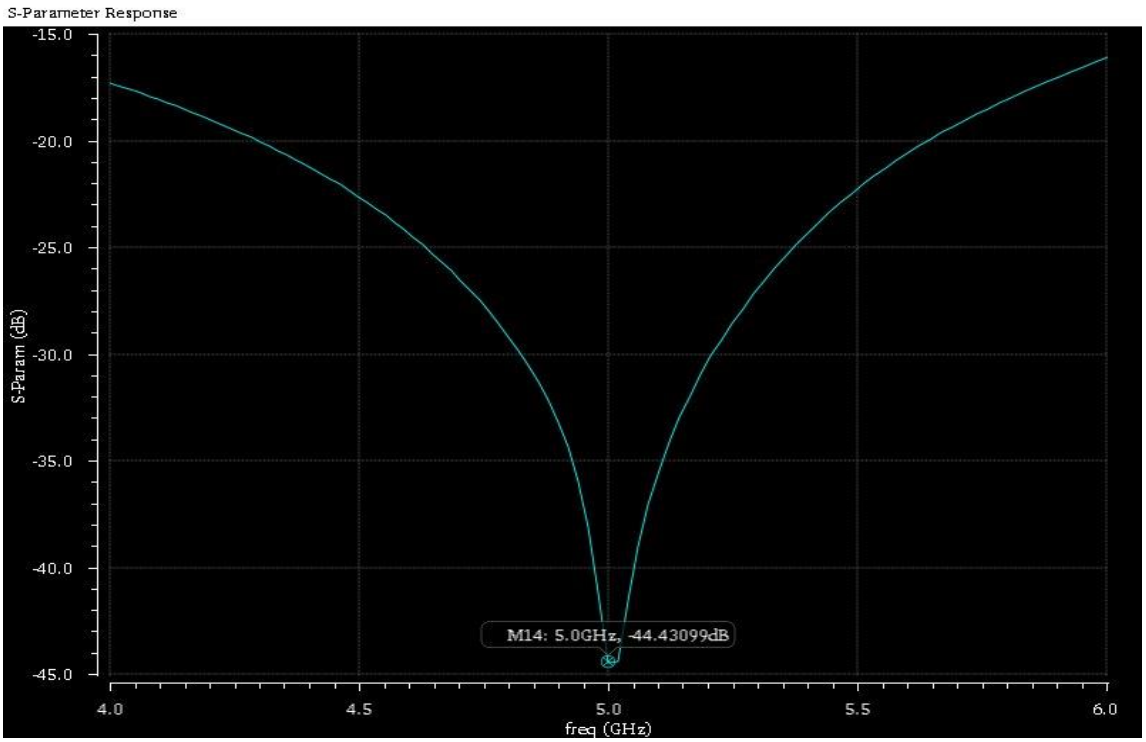


Figure 3.15.S_{11} parameter of power amplifier after input matching.

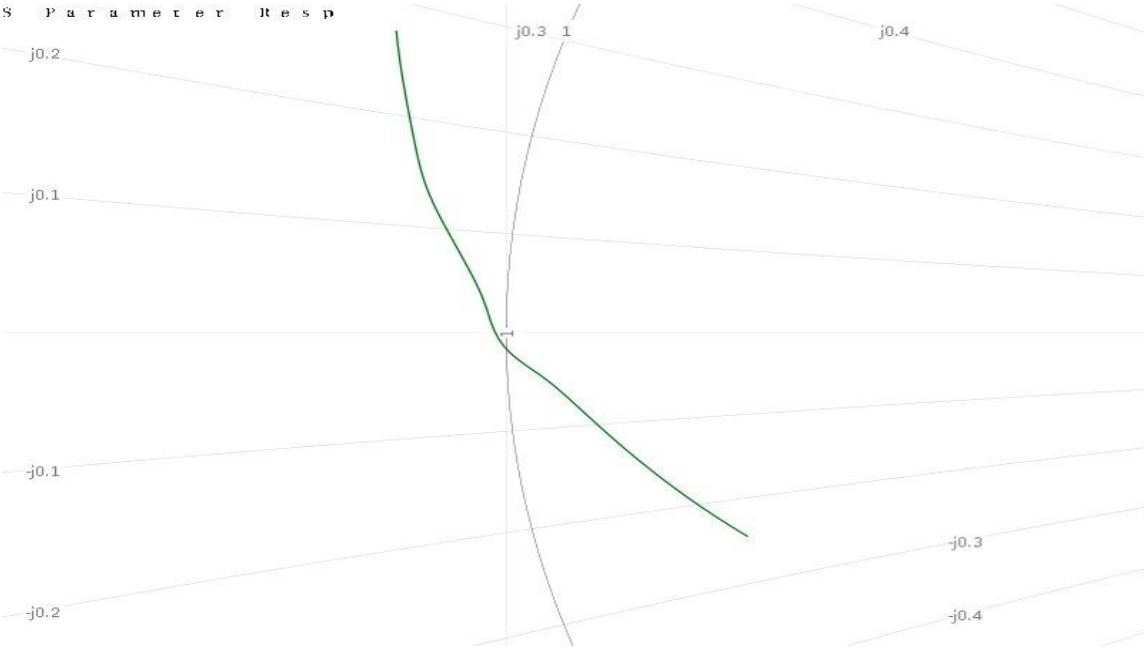


Figure 3.16.S_{11} parameter after load transformation of the power amplifier for the input matching Smith chart.

Chapter 4

Performance criteria

In this chapter, we will discuss in detail the performance of the designed Power Amplifier. As mentioned in previous subsections, the performance criteria of a Power Amplifier include linearity, stability, efficiency and several other trades. The performance of the Power Amplifier was verified via multiple and different simulations in Virtuoso Analog Design Environment of Cadence ®. The nominal temperature of all simulations were T=27^{o}C. In Figure 4.1 the actual design testbench is presented.

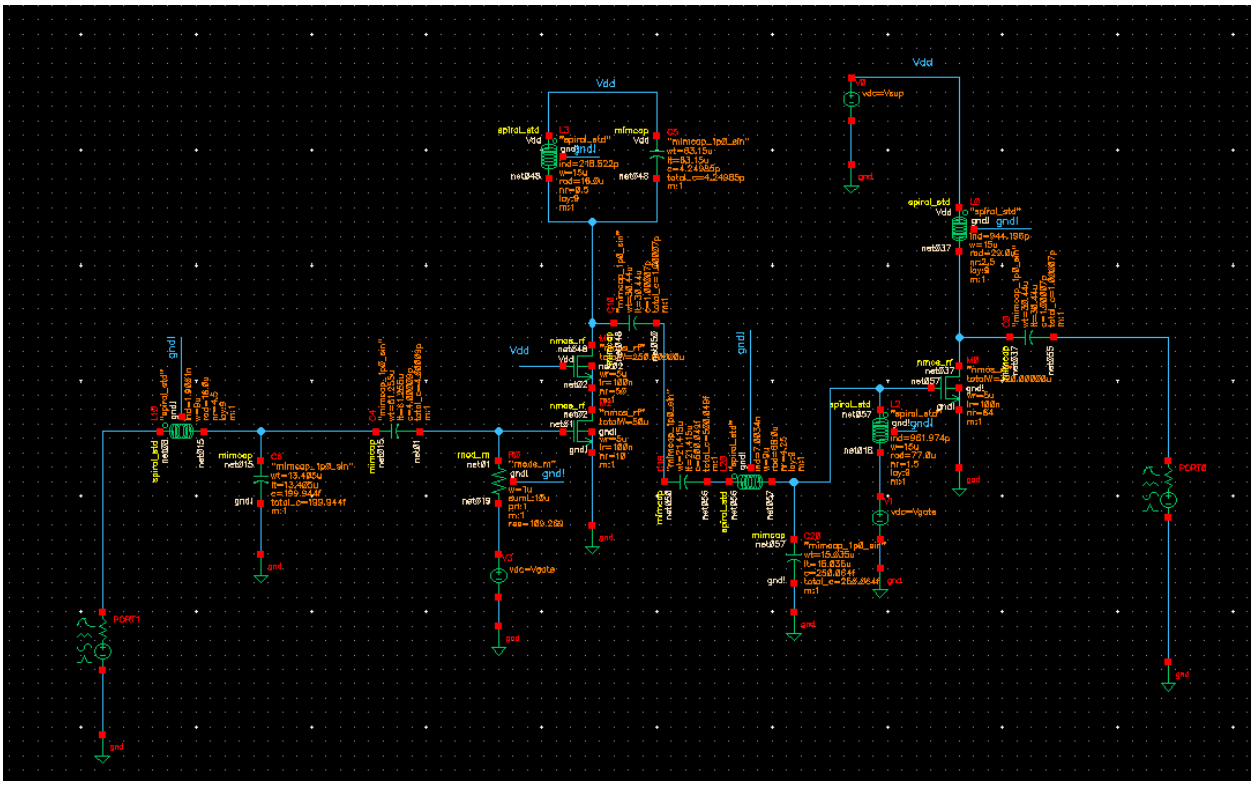


Figure 4.1. Testbench of actual design.

4.1 Power and linearity performance criteria

4.1.1 Input and output power characteristics

With the performance of harmonic balance (hb) simulations, we have extracted the following results. In Figure 4.2, the 1-dB compression point is shown. With the variation of input

power from -30dBm up to 10 dBm and the operating frequency being 5 GHz, the 1-db compression point is at 3.15dBm, with single point extrapolation at -30dBm, where the input power is -24dBm. The characteristic of input power of the amplifier versus the output power is shown in Figure 4.3. The input power being 0dBm we can see that the output power of the amplifier designed is 9.85 dBm, very close to the intended output power. And the saturated power of the amplifier is confirmed at input power greater than 2dBm, where it reaches almost 10dBm. This verifies that the designed Power Amplifier is considered to be a saturated PA. In Figure 4.4, the 1st order harmonic at 5GHz is plotted together with the 2nd order harmonic at 10GHz. We should stress the fact that the output power in the case of the 2nd harmonic with an input level of 0dBm is 2.6dBm, much less than the 1st harmonic as expected.

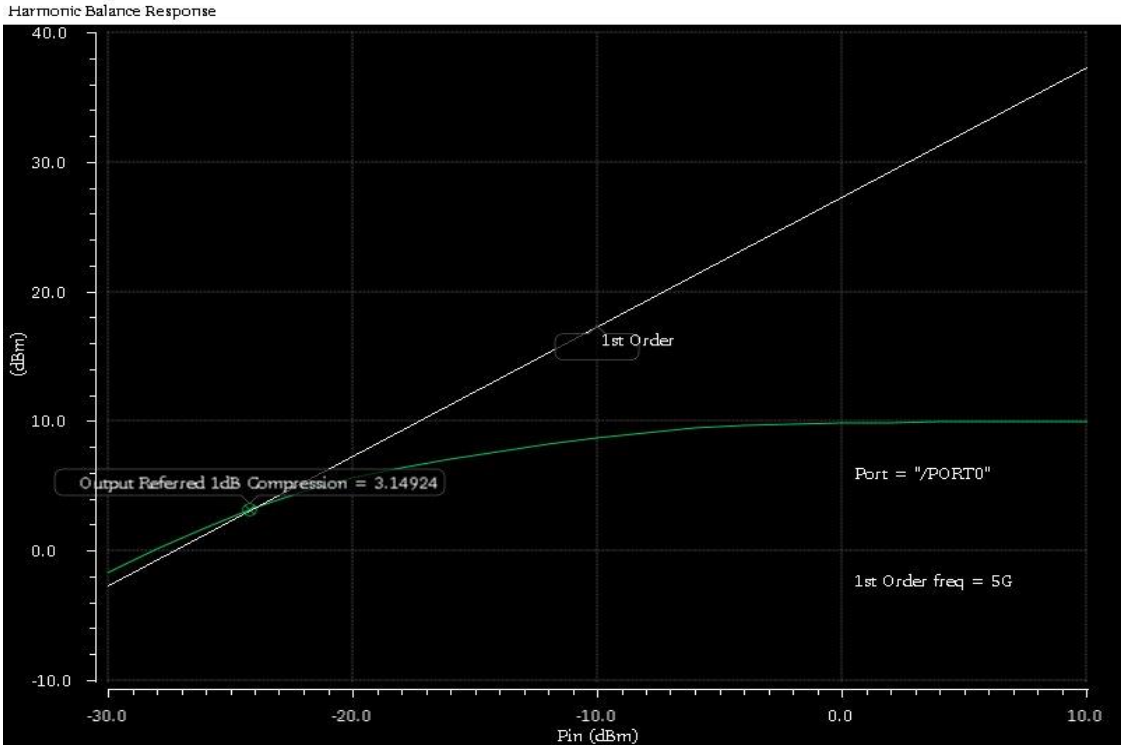


Figure 4.2.1-dB compression point of the amplifier.

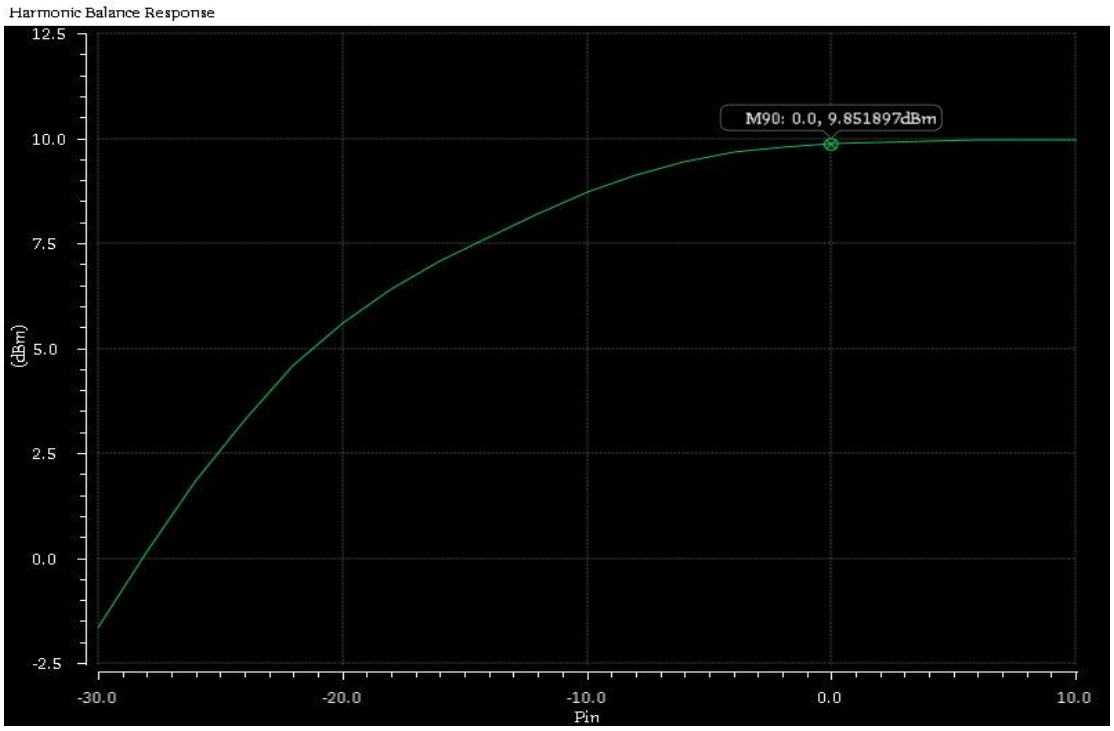


Figure 4.3. P_{in}-P_{out} simulation of the design.

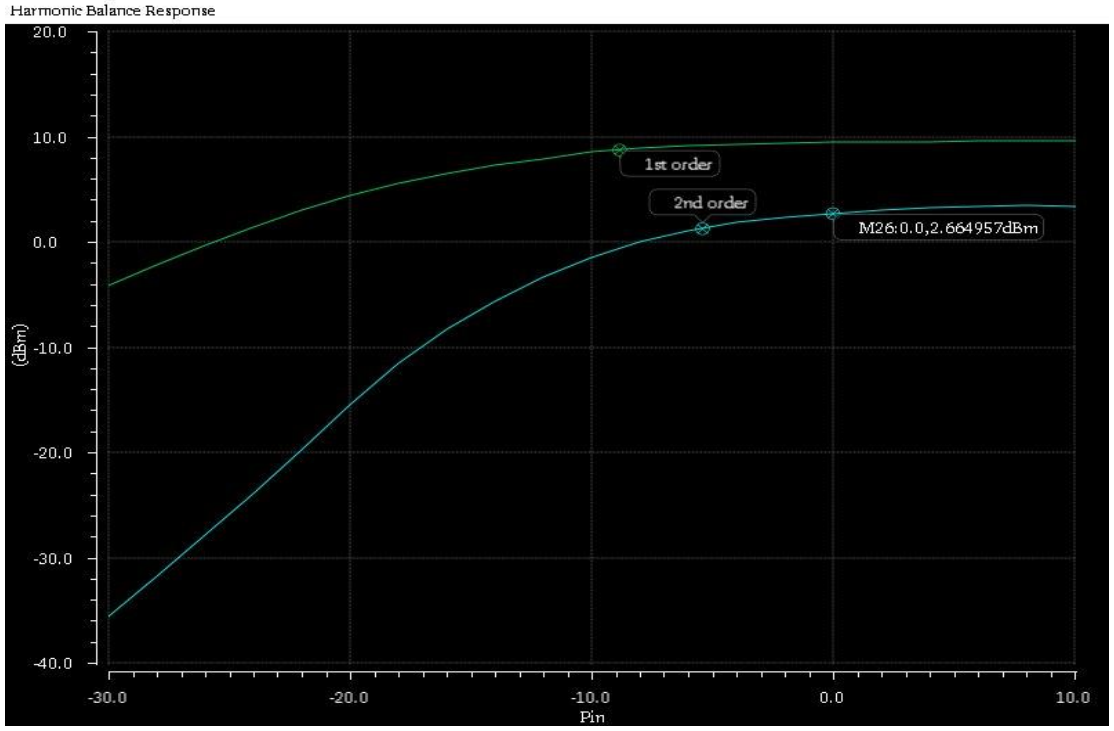


Figure 4.4. P_{in}-P_{out} simulation of the design, including the 2nd order harmonic at 10GHz.

4.1.2. S-parameter characterization

The scattering parameters are important to have an overview over the small signal characteristics of the designed power amplifier (Figure 4.5). Scattering parameter s_{21} represents the small signal gain, which is equal to almost 30dB at the operating frequency. While the 3-dB frequencies are 4.9 GHz and 5.12 GHz. The gain in the bandwidth of interest has a minimum of 22dB. In addition, scattering parameter s_{11} is equal to -44.4dB at the operating frequency, after the input matching, and remains below zero at all frequencies. This is also an indication of stability for the design. [37] Scattering parameter s_{12} is indicative of the reverse isolation of the design. The parameter is equal to -65dB in 5GHz and its low value mostly depends on the cascode stage.

4.1.3 Third order intercept point (IP3) measurements

In order to simulate the third order intercept point, a periodic steady-state (pss) together with a periodic ac (pac) simulation was performed. The results can be seen in Figure 4.6 and Figure 4.7. We can see that the third order intercept point (IP3) for the input is equal to -13.2dBm (IIP3) and for the output (OIP3) equal to 12.5dBm. There was made a single point extrapolation at -30dBm. The frequencies in question are chosen to be $\omega_2 = 5.001GHz$ and $2\omega_1 - \omega_2 = 4.999GHz$.

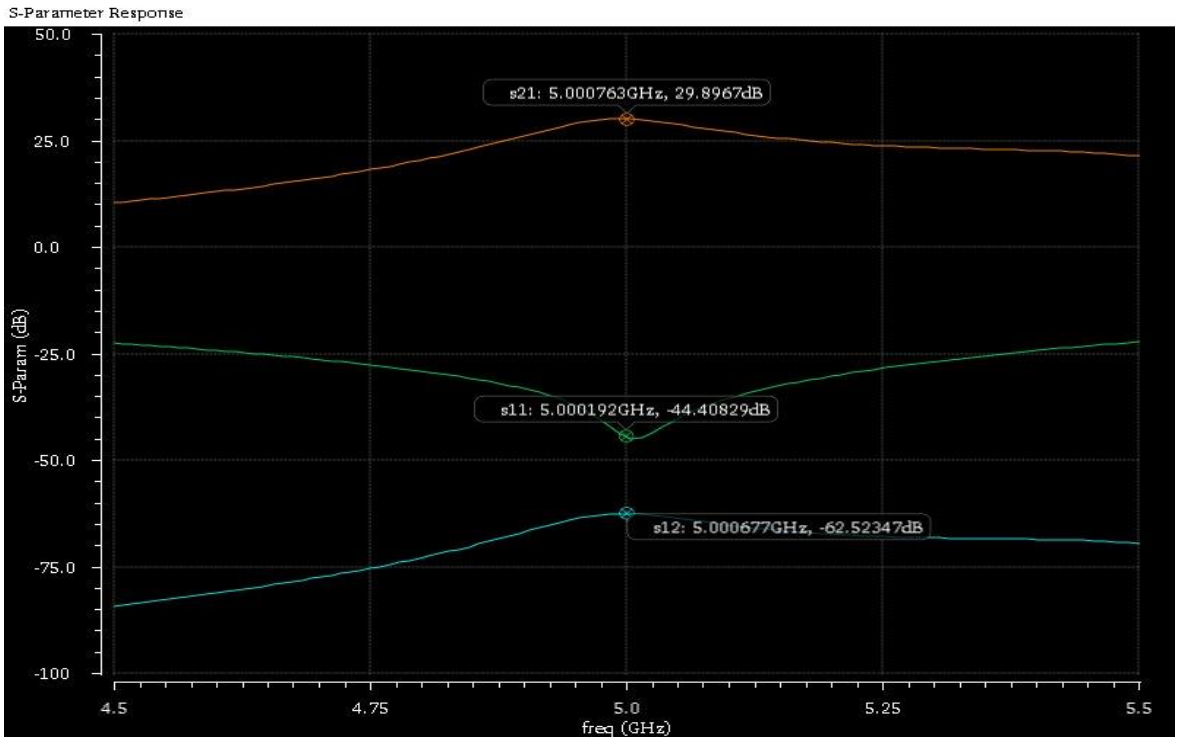


Figure 4.5. Scattering parameters of the overall design.

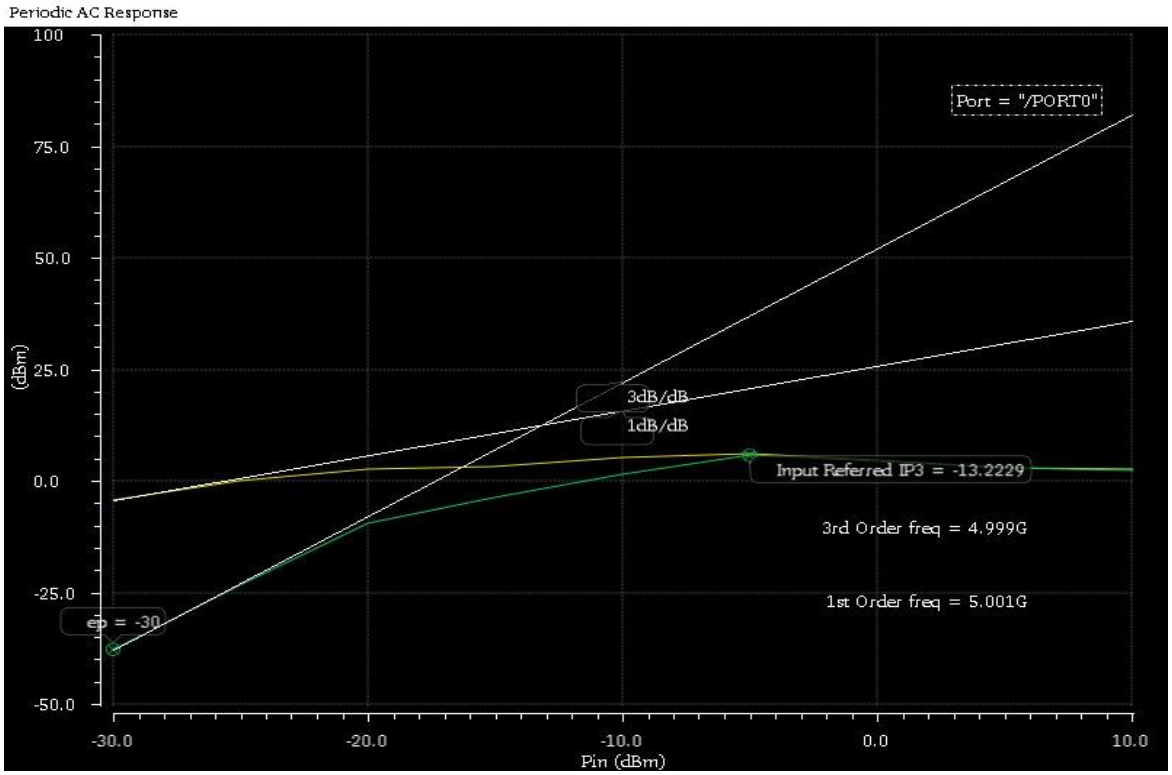


Figure 4.6. Input third intercept point (IIP3).

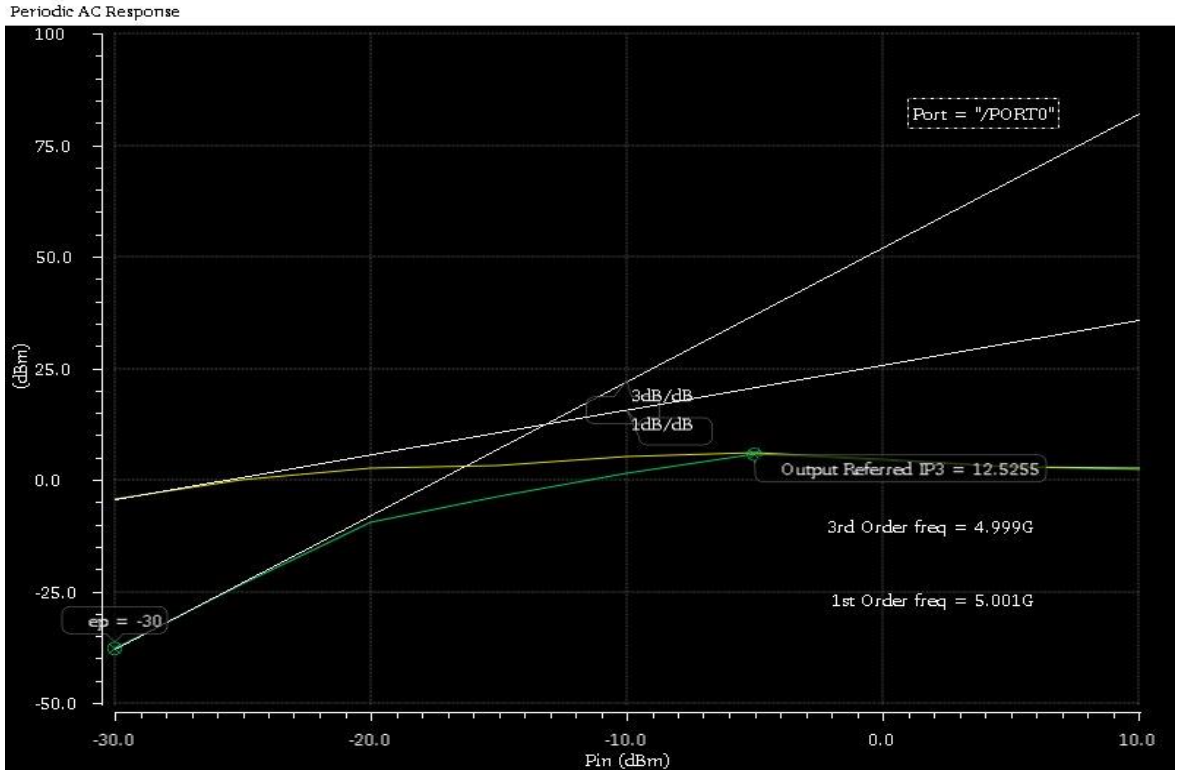


Figure 4.7. Output third intercept point (OIP3).

4.2. Efficiency measurements

As mentioned before in subsection 2.5.1, the efficiency of a class A power amplifier cannot exceed the limit of 25%. In Figure 4.8, the drain efficiency of the overall design is presented (22.9% in our case). Furthermore, a more accurate metric for the performance of the amplifier, as far as the efficiency is concerned, is the power added efficiency (PAE). PAE is 26.4% for 0dBm input power (Figure 4.9) for our design.

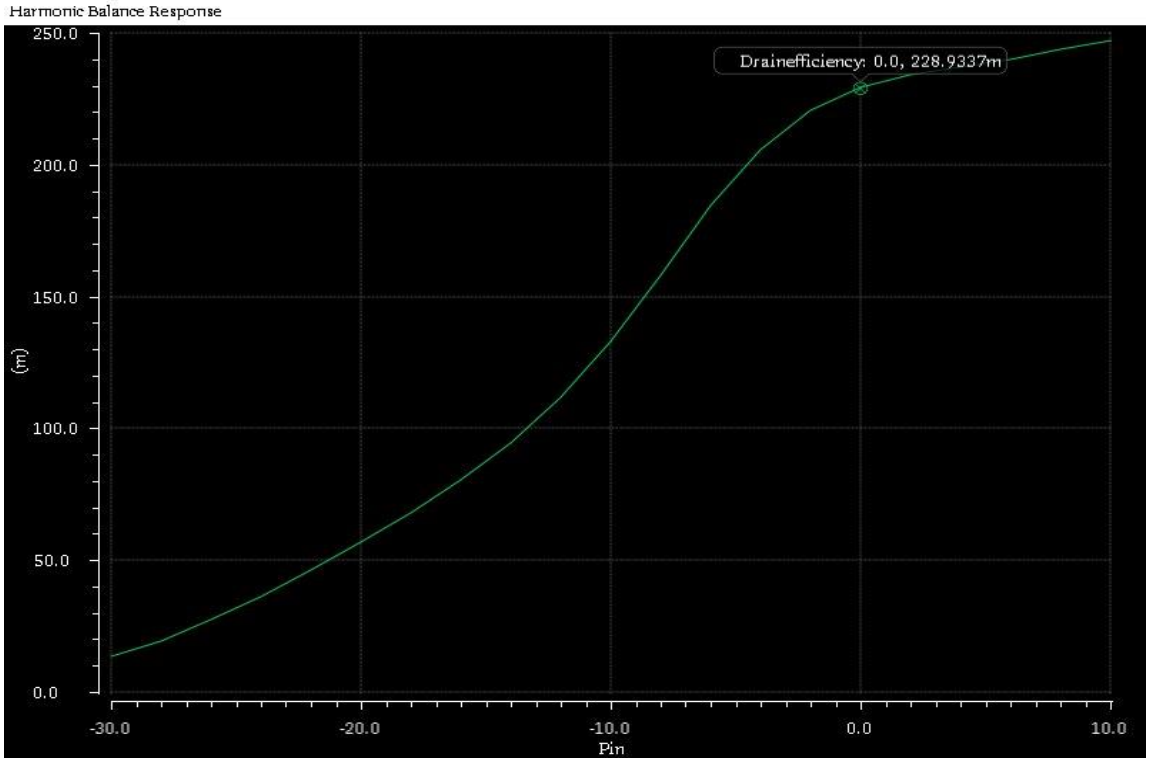


Figure 4.8. Drain efficiency.

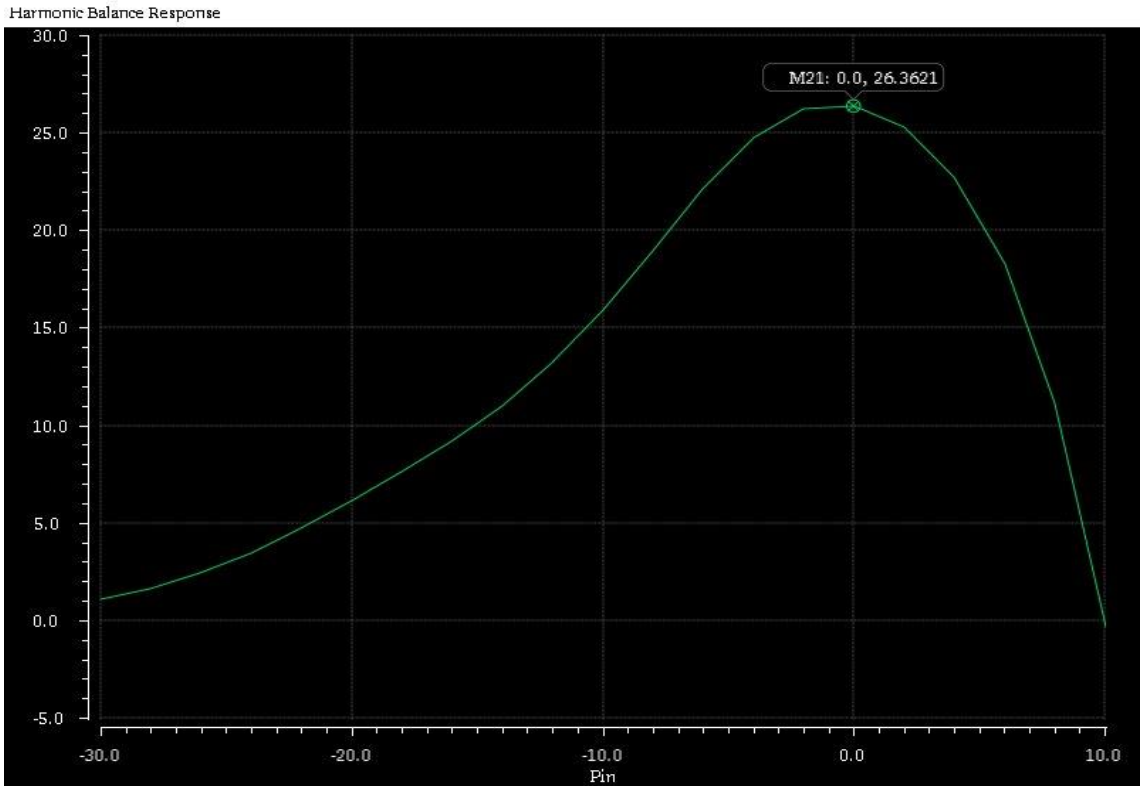


Figure 4.9. Power Added Efficiency.

4.3. Stability measurements

If a power amplifier is unconditionally stable, the stability circles must completely enclose the Smith chart. In Figure 4.10, we can see that at several frequencies in the bandwidth of interest the load stability circles are inside the unity circle of the Smith chart but not near the 50Ohm region. Despite the fact that all simulations where made considering a 50 Ohm termination, the power amplifier is conditionally stable and the stability circles are small. The stability circles around the frequency of 5GHz are larger than in other frequencies under test. In reality, in most cases the antenna that concludes the transmission path has an unpredictable, variable impedance that depends on reflection conditions. This phenomenon requires the unconditional stability of a circuit that could be ensured by the interval of a potential output matching network, which improves the stability factor (Figure 4.13). [35]
Unfortunately, this cannot be done without gain and power penalties.

An additional condition to ensure unconditional stability should be satisfied. One such parameter is the stability measure, B1f, which should be greater than zero. [36] Calculated as follows:

$$B1f = 1 + |S_{11}|^2 - |S_{22}|^2 - |\Delta|^2 > 0, \Delta < 1 (4.1)$$

In Figure 4.11, the plot of the alternative stability factor B1f is presented showing a value greater than zero without a 50 Ohm termination. The alternative stability factor of the circuit, is approximately 0.8. In Figure 4.12, the stability factor is shown without the 50 Ohm termination.

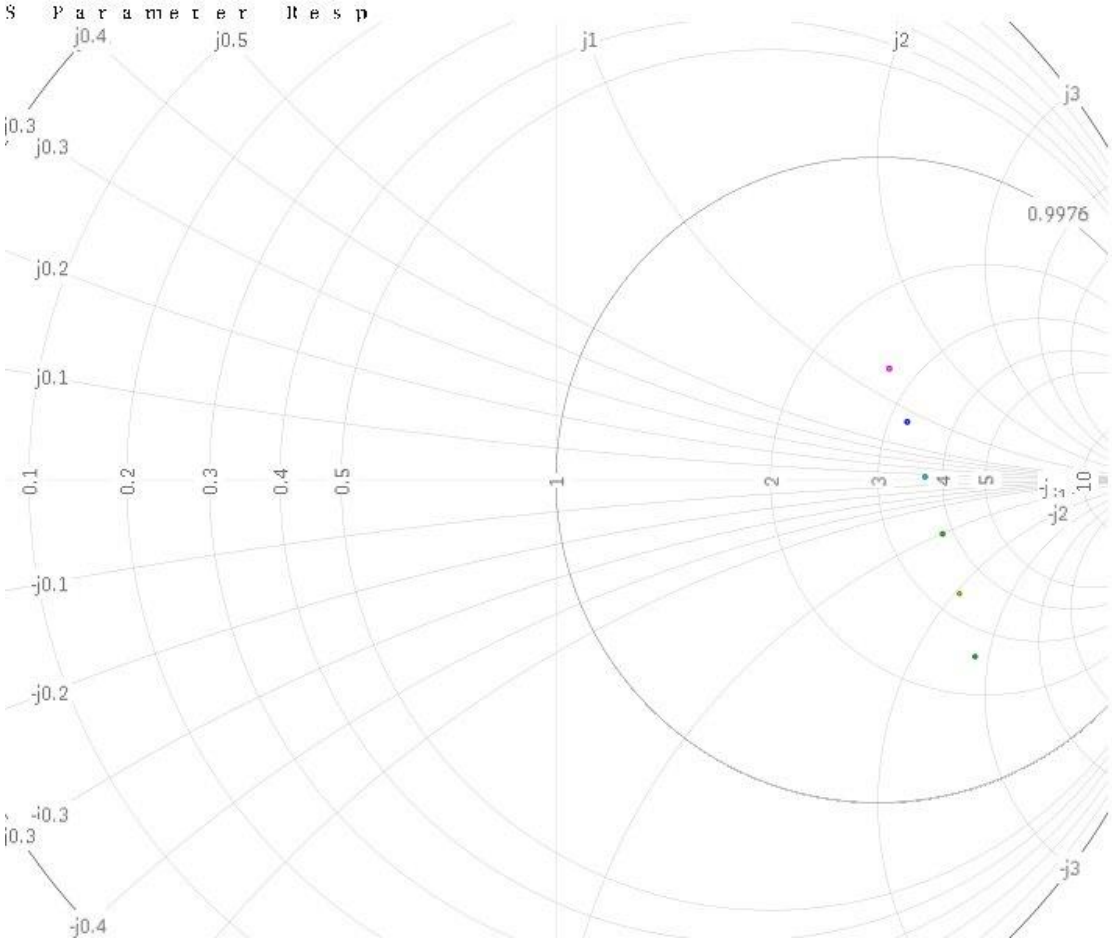


Figure 4.10. Load stability circles.

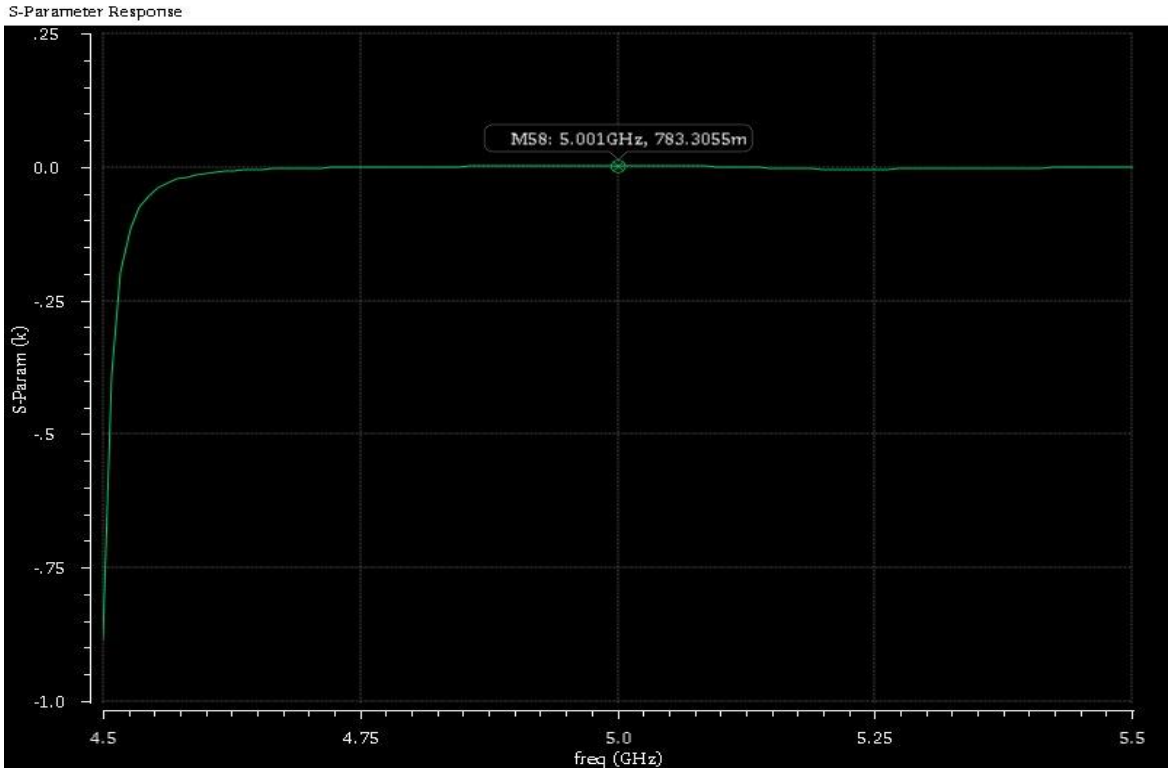


Figure 4.11. Alternative stability factor B1f.

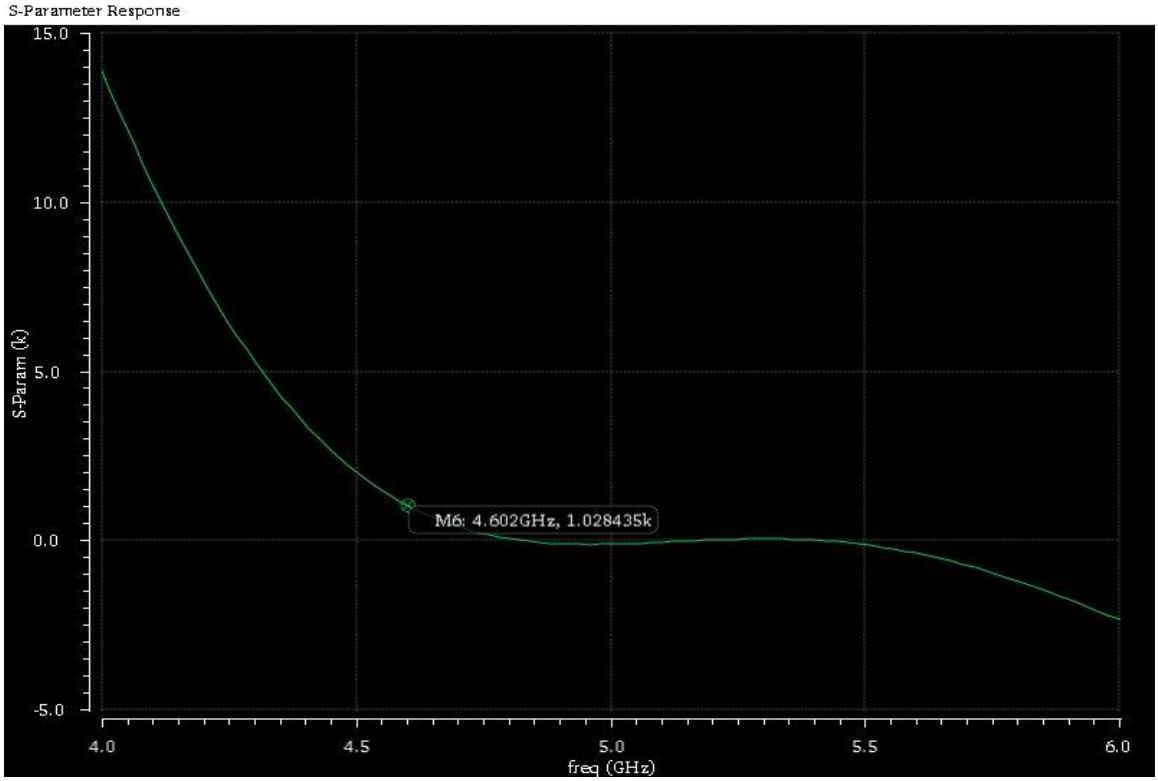


Figure 4.12. Stability factor k.

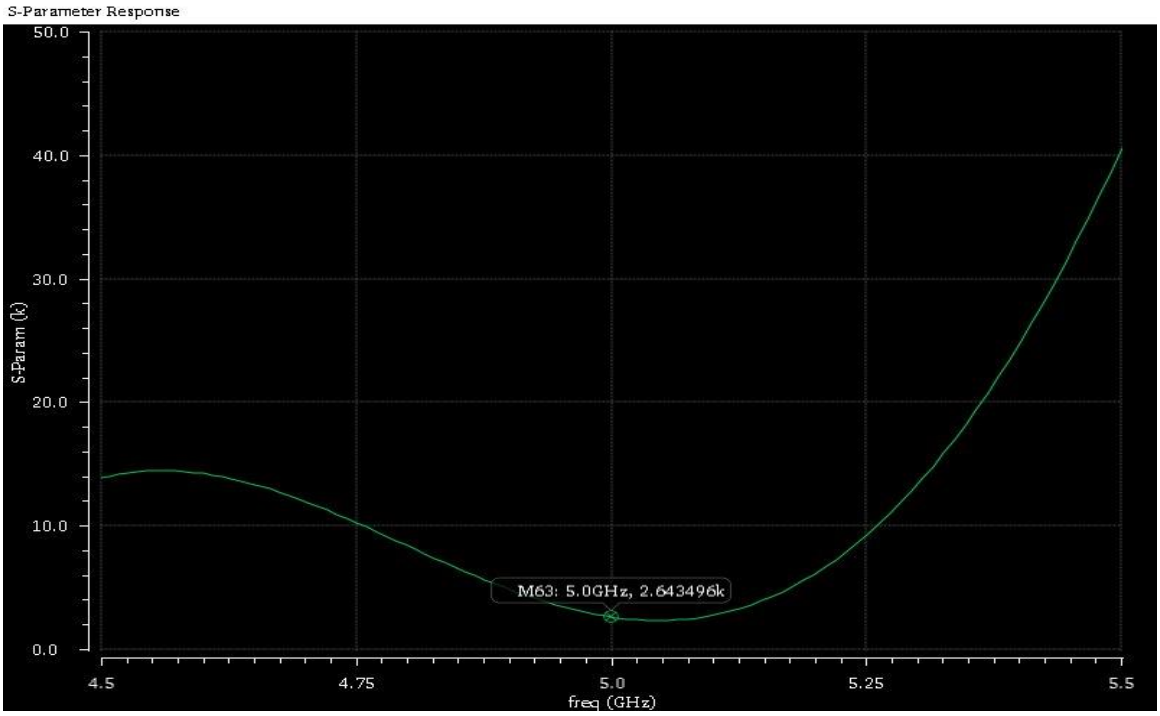


Figure 4.13. Stability factor after a potential 50 Ohm termination.

4.4. Design performance in different temperature and biasing conditions.

The Power Amplifier, as said before, is a part of a larger system, a transceiver. The real conditions of operation of such a system may vary, due to temperature variations and variations as far as the power supply is concerned. The nominal temperature of all simulations being 27^{o}C is close to ideal, so in order to verify the operation of the design, temperature sweep simulations were made to ensure that the power amplifier doesn't deviate from normal operation. Also, simulations were performed to capture the performance of the design in different biasing conditions.

Subjecting the power amplifier to 100mV lesser DC supply voltage and 100mV more DC supply voltage (+15% and -15%) in all pins used for biasing, the s_{11} of the overall design doesn't deviated from nominal, shown in Figure 4.14. In the contrary, s_{21} is affected severely when subjected to a lesser value of biasing, degrading small-signal gain (Figure 4.15).

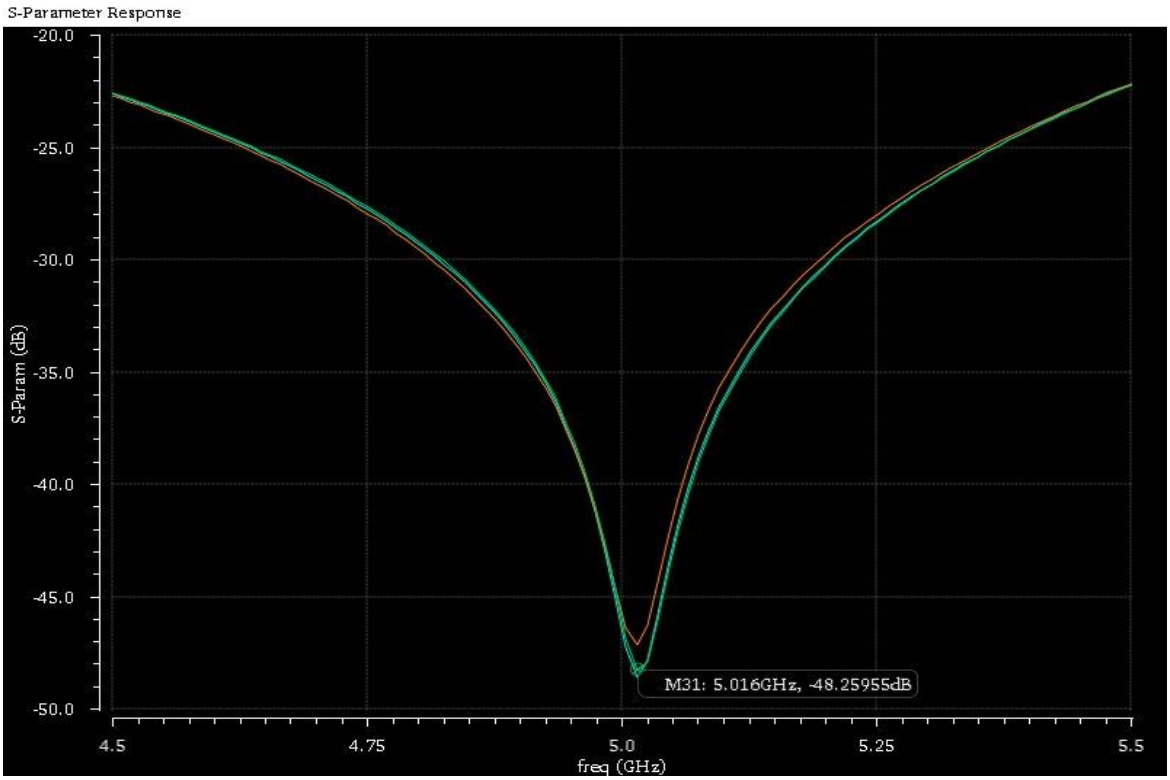


Figure 4.14. S_{11} parameter according to DC supply voltage variations.

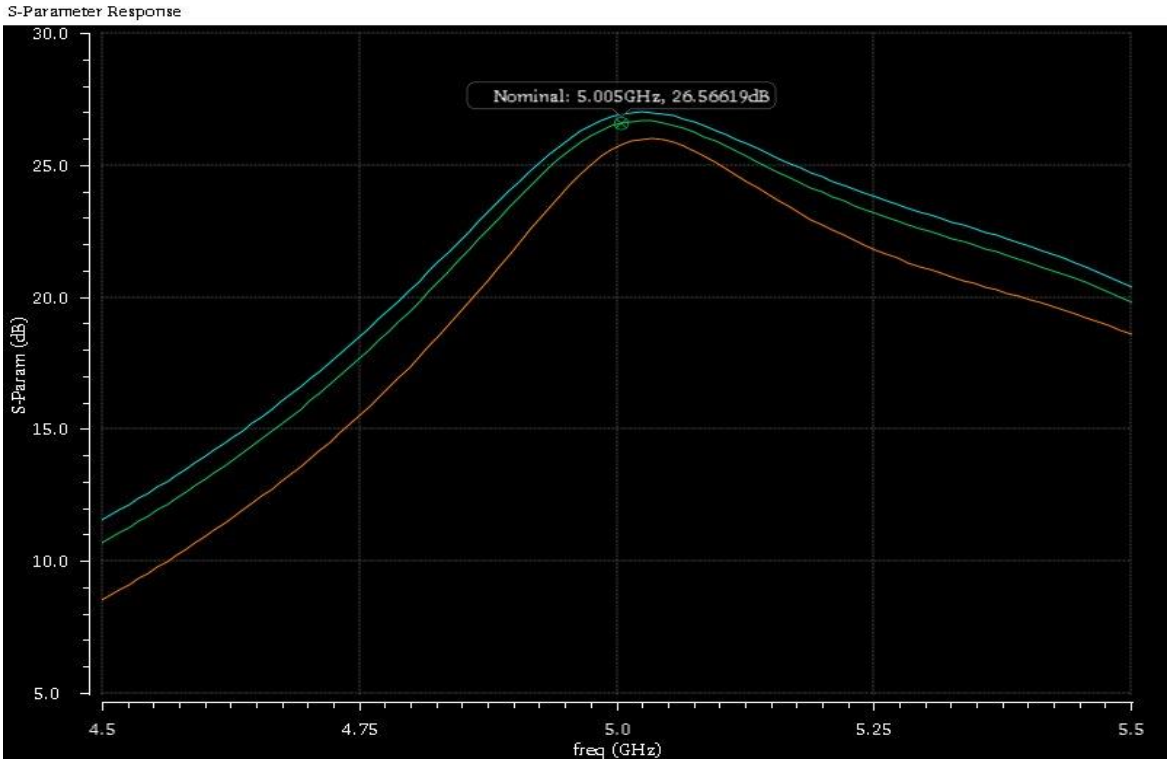


Figure 4.15. S_{21} parameter according to DC supply voltage variations.

The temperature range that the sweep was made is 0^o C up to 120^o C, a valid range for the design kit, which specifies degradation in performance beyond 125^o C. The results owing to temperature variations for s_{11} and s_{21} are shown in Figure 4.16 and Figure 4.17, repsectively.

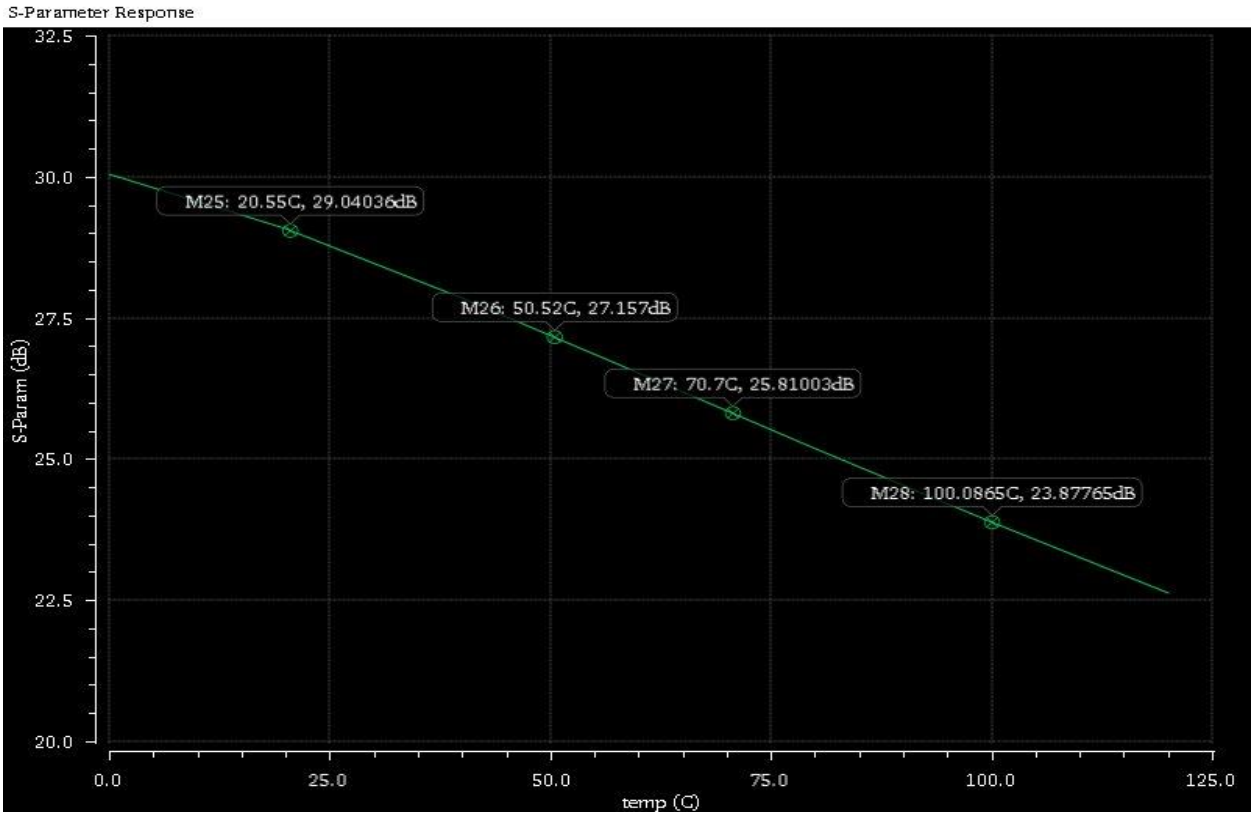


Figure 4.16.S_{21} parameter temperature sweep.

As expected, temperature changes the scattering parameters significantly. Both, s_{11} and s_{22} are degraded with the rise of temperature. Although, parameter s_{11} stays below zero at all times and s_{21} still provides a satisfactory small-signal gain.

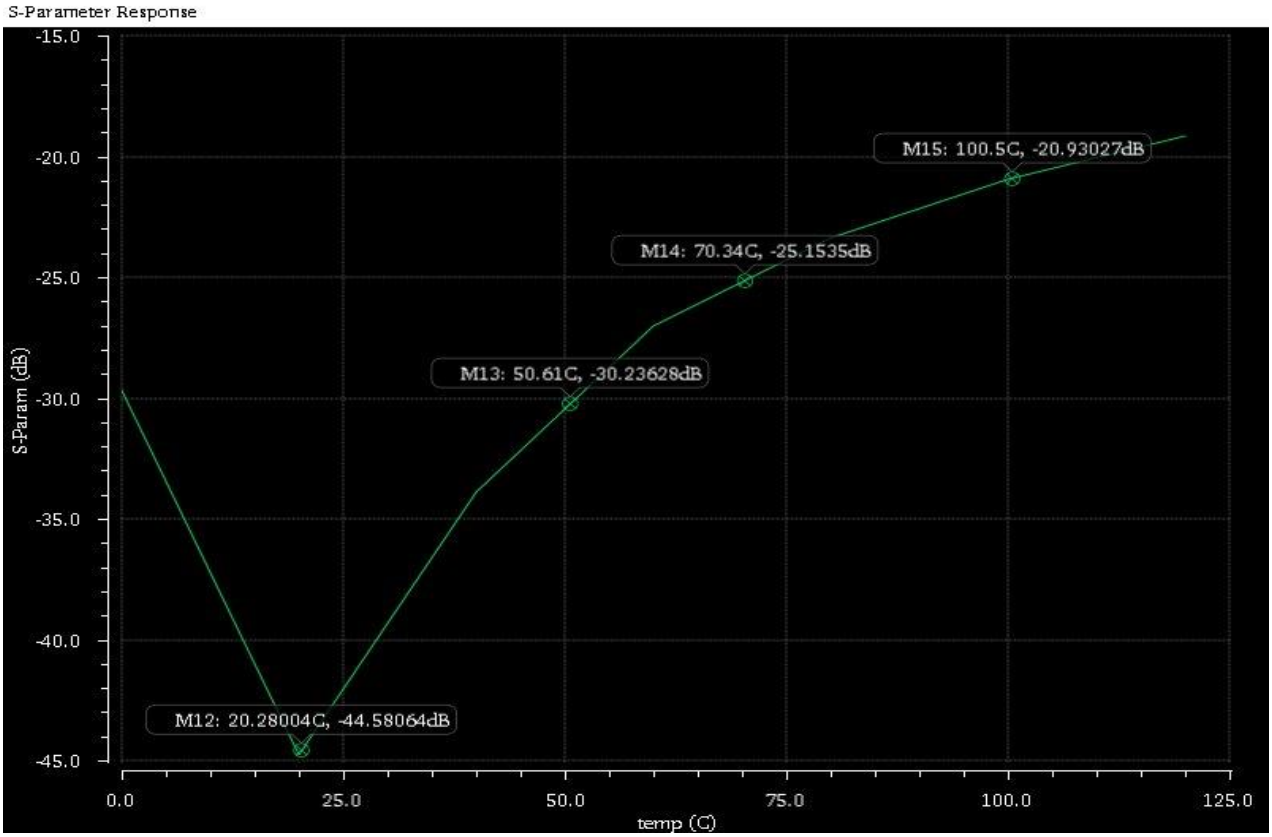


Figure 4.17.S11 parameter temperature sweep.

4.5.Comparison with other implementations.

In Table 4.1, two other implementations are compared with our work. At this point we should mention that the results of our work are based solely on schematic level design simulations and the implementations mentioned refer to post-layout simulations.

We can say that our design shows better drain efficiency and lesser power dissipation, while the saturated ouput power is comperable with the other designs. The 1-dB compression point of our design is much less than that of the other designs due to the fact our power amplifier is a saturated one. In order, to have a more conclusive comparison the layout of the design presented should be done.

Technology	P1-dB (dBm)	P_{sat} (dBm)	Gain (dB)	PAE (%)	Power Dissipation (mW)	DE (%)
[40] 90nm CMOS	20.5	16	N/A	N/A	610	6.7
[40] 90nm CMOS	20.5	11	N/A	N/A	530	2.4
This work	3.15	10	29.9	26	66	22.8

Table 4.1. Comparison with other implementations

References

[1] Mustafa Acar, Ph.D., *Power Amplifiers in CMOS technology: A Contribution to Power Amplifier Theory and Techniques*, CTIT, The Netherlands, 2011

[2] Haldi, P., Chowdhury, D., Liu, G., & Niknejad, A. M. (2007, June). A 5.8 GHz linear power amplifier in a standard 90nm CMOS process using a 1V power supply. In *2007 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*(pp. 431-434). IEEE.

[3] http://pocketnow.com/2014/01/23/5ghz-wifi

[4] Magnus Isaksson, Ph.D., *Radio Frequency Power Amplifiers: Behavioral Modeling, Parameter-Reduction, and Digital Predistorion*, KTH School of Electrical Engineering, 2007

[5] Behzad Razavi, *RF Microelectronics*, *Second edition*, Pearson, 2013

[6] http://www.slideshare.net/simenli/rf-ch8

[7] Pieter L. D. Abrie, *Design of RF and Microwave Amplifiers and Oscillators*, *Second edition*, Artech House, 2009

[8] Peter Vizmuller, *RF DESIGN GUIDE Systems,Circuits, and Equations*, Artech House, 1995

[9] Raab, F. H., Asbeck, P., Cripps, S., Kenington, P. B., Popovic, Z. B., Pothecary, N., ... & Sokal, N. O. (2003). RF and microwave power amplifier and transmitter technologies-Part 1. *High Frequency Electronics*, *2*(3), 22-36.

[10]Sokal, N. O. (2000, June). Class-E switching-mode high-efficiency tuned RF/microwave power amplifier: improved design equations. In *Microwave Symposium Digest. 2000 IEEE MTT-S International* (Vol. 2, pp. 779-782). IEEE.

[11]David M. Pozar, *Microwave Engineering, Third edition*, Wiley, 2005

[12] John B. Hagen, *Radio-frequency Electronics*, *Second edition*, Cambridge University Press, 2009

[13]http://www.engr.sjsu.edu/rkwok/Engr297/Val_Impedance%20Matching%20and%20Matching%20Networks.pdf

[14] http://www.nxp.com/files/rf_if/doc/app_note/AN721.pdf

[15]http://www.ece.ucsb.edu/Faculty/rodwell/Classes/ece218c/tutorials_etc/Impedance_Matching.pdf

[16]Acar, M., Annema, A. J., & Nauta, B. (2007, June). Variable-voltage class-E power amplifiers. In *Microwave Symposium, 2007. IEEE/MTT-S International*(pp. 1095-1098). IEEE.

[17]http://www.learnabout-electronics.org/Downloads/amplifiers-module-05.pdf

[18] Albert Malvino, David Bates, *Electronic Principles*, *Eighth edition*, McGraw-Hill Companies,Inc., 2016

[19]http://www.qsl.net/va3iul/RF%20Power%20Amplifiers/RF_Power_Amplifiers.pdf

[20] Kazimierczuk, M. K., & Puczko, K. (1987). Exact analysis of class E tuned power amplifier at any Q and switch duty cycle. *Circuits and Systems, IEEE Transactions on*, *34*(2), 149-159.

[21] Sokal, N. O. (1998, June). Class E high-efficiency power amplifiers, from HF to microwave. In *Microwave Symposium Digest, 1998 IEEE MTT-S International*(Vol. 2, pp. 1109-1112). IEEE.

[22] Fritzin, J., Sundström, T., Johansson, T., & Alvandpour, A. (2010, May). Reliability study of a low-voltage Class-E power amplifier in 130nm CMOS. In*Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on* (pp. 1907-1910). IEEE.

[23] Xu, X., Sun, Z., Xu, K., Yang, X., Kurniawan, T., & Yoshimasu, T. (2014, August). A 2.5-GHz band low-voltage class-E power amplifier IC for short-range wireless communications in 180-nm CMOS. In *Radio-Frequency Integration Technology (RFIT), 2014 IEEE International Symposium on* (pp. 1-3). IEEE.

[24] Sokal, N. O. (2001). Class-E RF power amplifiers. *QEX Commun. Quart*, *204*, 9-20.

[25]Sokal, N. O., & Sokal, A. D. (1975). Class E A new class of high-efficiency tuned single-ended switching power amplifiers. *Solid-State Circuits, IEEE Journal of*, *10*(3), 168-176.

[26]Tofte Røislien, N. (2009). 5.8 GHz, 1W high efficiency Power Amplifier in 90nm CMOS.

[27] Fritzin, J., & Alvandpour, A. (2009, January). Low voltage class-E power amplifiers for DECT and bluetooth in 130nm CMOS. In *Silicon Monolithic Integrated Circuits in RF Systems, 2009. SiRF'09. IEEE Topical Meeting on*(pp. 1-4). IEEE.

[28] www.electronics-tutorials.ws/amplifier/amplifier-classes.html

[29] www.murata.com/en-us/products/inductor/chip/feature/rf

[30] www.aikenamps.com/index.php/chokes-explained

[31] https://www.digikey.com

[32] Duran Leblebici, Yusuf Leblebici, *Fundamentals of High- frequency CMOS Analog Integrated Circuits*, Cambridge, 2009

[33] Nikolaos Makris, M.Sc., *Characterization and RF design aspects in advanced CMOS Technology*, School of Electronic and Computer Engineering, TUC, 2011

[34] Ali M.Niknejad, Hossein Hashemi, *mm-Wave Silicon Technology: 60GHz and Beyond*, Springer Science & Business Media, 2008

[35] Γεράσιμος Βλαχογιαννάκης, Σχεδίαση ολοκληρωμένου Ενισχυτή Ισχύος 10dBm συχνότητας λειτουργίας 60 GHz σε τεχνολογία CMOS, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, NTUA, 2011

[36]http://cp.literature.agilent.com/litweb/pdf/genesys200801/sim/measurements_linear/m_stability_factor.htm

[37] Wu, C. Y., & Hsiao, S. Y. (1997). The design of a 3-V 900-MHz CMOS bandpass amplifier. *IEEE Journal of Solid-State Circuits*, *32*(2), 159-168.

[38]http://www.electronicproducts.com/Passive_Components/Choosing_the_right_inductor.aspx

[39] Steve C. Cripps, *RF Power Amplifiers for Wireless Communications*, Artech house, 1999

[40] Palaskas, Y., Taylor, S. S., Pellerano, S., Rippke, I., Bishop, R., Ravi, A., ... & Soumyanath, K. (2006). A 5-GHz 20-dBm power amplifier with digitally assisted AM-PM correction in a 90-nm CMOS process. *IEEE journal of solid-state circuits*, *41*(8), 1757-1763.