Analytical Modeling of SiGe HBT Intermodulation Linearity

by

Yiao Li

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 07 , 2016

Keywords: SiGe HBT, intermodulation, linearity, IM3 cancellation

Approved by

Guofu Niu, Professor of Electrical and Computer Engineering
Fa Foster Dai, Professor of Electrical and Computer Engineering
Bogdan Wilamowski, Professor of Electrical and Computer Engineering

Abstract

Linearity is a key concern in RF systems, particularly an amplifier. Volterra series allow us to easily identify the contribution of various individual nonlinearities, as well as to identify the interaction between individual nonlinearities.

Nonlinear current source of Volterra series and direct derivation of Volterra series are two methods that analyze nonlinear distortion effectively. This thesis applied Volterra series to circuit linearity analysis and performed with the help of a matrix solver in Matlab. ADS used is to verify the validity of analytical expressions.

Supported by Volterra series approach, requirement of $IM3$ cancellation can be found at both low frequency and high frequency. ADS simulation is used to investigate conditions required for $IM3$ cancellation.

Acknowledgments

I would like to express my deepest thanks to my advisor, Dr. Guofu Niu, for the past two years' encouragement and guidance. I would like also to thank my other committee members, Dr. Fa Dai and Dr. Bogdan Wilamowski for their assistance.

Thanks to my groupmates, Zhen Li, Huaiyuan Zhang, Jiabi Zhang, Yili Wang and Anni Zhang. They help me for this two years study.

At last, I would like to thank my parents and families for their support and love.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Auburn Auburn University

LoA List of Abbreviations

SiGe: Silicon-germanium

HBT: Heterojunction bipolar transistor

HD2: Second-order harmonic distortion

HD3: Third-order harmonic distortion

IM2: Second-order intermodulation distortion

IM3: Third-order intermodulation distortion

HB: Harmonic balance

IP2: second-order intercept point

IP3: Third-order intercept point

Chapter 1

Introduction

As well known, most of electronic circuits are nonlinear. The linear assumption for most modern circuit theory is only an approximation in practice. The most commonly used elements in integrated circuit are transistors, resistors, capacitors, inductances, and diodes. All these elements mentioned above are always described as a nonlinear equivalent circuit. Thus analyzing such circuit is usually more complicated.

Contrast with "nonlinearity" and "distortion", "linearity" refers to the ability of a device or system to amplify input signals in a linear way. Silicon-germanium heterojunction bipolar transistors (SiGe HBTs) are in general nonlinear elements like other semiconductor devices. For instance, it has a very strong exponential $I_C - V_{BE}$ nonlinearity.

Although SiGe HBTs can be used in both nonlinear and linear circuits which depend on the the required application, nonlinearities in transistor still can't be avoided. These nonlinearities will create distortion in the signals we are interested in amplifying or transmitting. For instance: the intermodulation of two nearby strongly interfering signals caused by nonlinearity at the input of a receiver will influence the desired weak signal we are trying to receive.

Despite the strong $I - V$ and $C - V$ nonlinearities, SiGe HBTs also have excellent linearity in both small-signal (e.g., LNA) and large-signal (e.g., PA) RF circuit. It is clear that the overall circuit linearity strongly depend on the linear elements in the circuit, the source impedance, the load impedance and the interaction of $I - V$ and $C - V$ nonlinearities[1]. In this thesis, Volterra series approach will be used to analyze these nonlinearities.

1.1 Nonlinearity

When the input signal is sufficiently weak, the operation of a transistor circuit is linear. The response of a linear and dynamic circuit is characterized by an impluse response function in the time domain and a linear transfer function in the frequency domain[1]. Generally, using such approach requires v_{BE} to be much smaller than thermal voltage which is presented as $\frac{kT}{q}$ for a bipolar transistor.

For larger input signals, an active transistor circuit becomes a nonlinear dynamic system. Many of the nonlinearity concepts can be illustrated using simple power series, a concept which only applies to a memory-less circuit. In practice, even a linear circuit has memory elements (e.g., capactiors). Nevertheless, the use of power series simplifies the illustration of many commonly used linearity figures of merit[1].

The output voltage $V_{out}(t)$ of a nonlinear circuit can be illustrated as a power series when using a small-signal input $V_S(t)$[2]

$$V_{out(t)} = k_1 V_S(t) + k_2 V_S^2(t) + k_3 V_S^3(t), \tag{1.1}$$

where for simplicity we just give the third-order series. The concepts "harmonics", "intermodulation" are introduced below.

1.1.1 Harmonics

Substituting $V_S(t) = Acos\omega t$ into (1.1), the output voltage $V_{out}(t)$ can be obtained as[1]

$$\begin{aligned} V_{out}(t) =& k_1 Acos\omega t + k_2 A^2 cos^2\omega t + k_3 A^3 cos^3\omega t \\ =& \frac{k_2 A^2}{2} \\ &+ (k_1 A + \frac{3k_3 A^3}{4})cos\omega t \\ &+ \frac{k_2 A^2}{2} cos2\omega t \\ &+ \frac{k_3 A^3}{4} cos3\omega t. \end{aligned} \tag{1.2}$$

In (1.2), $\frac{k_2 A^2}{2}$ is dc shift, $(k_1 A + \frac{3k_3 A^3}{4})cos\omega t$ is a fundamental output at ω, $\frac{k_2 A^2}{2} cos2\omega t$ is a second-order harmonic term at 2ω and $\frac{k_3 A^3}{4} cos3\omega t$ is a third-order harmonic term at 3ω. Generally, we are interested in the harmonic of a given level respect to the fundamental output, which is so called harmonic distortion. The second-order harmonic distortion can be obtained from (1.2) as

$$HD2 = \frac{k_2 A^2}{2} / (k_1 A) = \frac{1}{2}\frac{k_2}{k_1} A, \tag{1.3}$$

where $\frac{3k_3 A^3}{4}$ is neglected compared to $k_1 A$.

$IHD2$ is defined as the extrapolation of the output at 2ω and ω intersect at a certain input level[1]. When setting $HD2$ =1 in (1.3), $IHD2$ can be obtained as

$$IHD2 = 2\frac{k_1}{k_2}. \tag{1.4}$$

It is clear that $IHD2$ is independent of the input signal level. Then using (1.4), $HD2$ can be calculated for any given small-signal input with amplitude A.

$$HD2 = \frac{A}{IHD2}. \tag{1.5}$$

$OHD2$, the output harmonic distortion is simply the product of small-signal gain and $IHD2$

$$OHD_2 = G \cdot IHD2 = 2\frac{k_1^2}{k_2^2}. \tag{1.6}$$

The third-order harmonic distortion $HD3$, the input and output intercept of the third harmonic distortion $IHD3$ and $OHD3$ can be defined similarly.

1.1.2 Intermodulation

Consider a two tone input voltage $V_S(t) = Acos(\omega_1 t) + Acos(\omega_2 t)$. The output not only has harmonics of ω_1 and ω_2, but also "intermodulation products" at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$. A full expansion of (1.1) using $V_S(t) = Acos(\omega_1 t) + Acos(\omega_2 t)$ shows that the output contains signals at ω_1, ω_2, $2\omega_1$, $2\omega_2$, $3\omega_1$, $3\omega_2$, $\omega_1 + \omega_2$, $\omega_1 - \omega_2$, $2\omega_2 - \omega_1$, $2\omega_2 + \omega_1$, $2\omega_1 - \omega_2$, and $2\omega_1 + \omega_2$ [1].

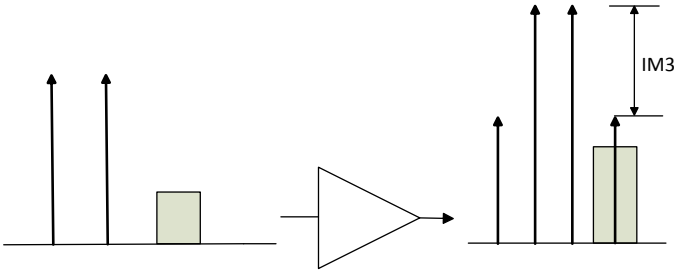

Figure 1.1: Intermodulation product of two strong interfers can corrupt the desired signals[1].

When ω_1 and ω_2 are closely spaced, the third-order intermodulation products at $2\omega_2 - \omega_1$ and $2\omega_1 - \omega_2$ are the major concerns, because they are close in frequency to ω_1 and ω_2, and

thus within the amplifier bandwidth, and cannot be filtered. Consider a weak desired signal near two strong interferers at the input. One intermodulation product falls in band, and corrupts the desired component, as illustrated in Fig. 1.1[1].

Substituting $V_S(t) = Acos\omega_1 t + Acos\omega_2 t$ into (1.1) leads to[1]

$$V_{out}(t) = (k_1 A + \frac{3k_3 A^3}{4} + \frac{3k_3 A^3}{2})cos\omega_1 t + \cdots + \frac{3k_3 A^3}{4} cos(2\omega_2 - \omega_1)t + \cdots . \tag{1.7}$$

Neglecting the higher order terms added to $k_1 A$, $k_1 Acos\omega_1 t$ and $\frac{3k_3 A^3}{4} cos(2\omega_2 - \omega_1)t$ are the fundamental terms at ω_1 and the intermodulation terms respectively.

The third-order intermodulation distortion ($IM3$) is defined as the ratio of the amplitude of the intermodulation product to the amplitude of the fundamental output.

$$IM3 = \frac{3k_3 A^3}{4} / k_1 A = \frac{3}{4}\frac{k_3}{k_1} A^2, \tag{1.8}$$

When setting $IM3$ equal to 1, the input third-order intercept ($IIP3$) can be obtained from (1.8), which is defined as the intersect of the fundamental output and the third-order output versus input at a given input level.

$$IIP3 = \sqrt{\frac{4}{3}\frac{k_1}{k_3}}. \tag{1.9}$$

It is clear in equation (1.9) $IIP3$ is independent on the input signal, thus it is more useful than $IM3$. Given $IIP3$, $IM3$ can be calculated with amplitude of desired small signal.

$$IM3 = \frac{A^2}{IIP3^2}. \tag{1.10}$$

Beacuse in equation (1.8) $IM3$ grows with A^2, thus $IIP3$ can be measured at a signal input level A_0,

$$IIP3^2 = \frac{A_0^2}{IM3_0}, \tag{1.11}$$

where $IIP3$ and A_0 are voltage, then transferring this equation to power level:

$$20log_{10}IIP3 = 20log_{10}A_0 - 10log_{10}IM3_0. \tag{1.12}$$

$20log_{10}IIP3$ is the power expression at intercept point in dB, and $20log_{10}A_0$ is the input power in dB. Then (1.12) can be rewritten as

$$P_{IIP3} = P_{in} + \frac{1}{2}(P_{out,1st} - P_{out,3rd}), \tag{1.13}$$

and

$$P_{OIP3} = P_{IIP3} + Gain = P_{out,1st} + \frac{1}{2}(P_{out,1st} - P_{out,3rd}). \tag{1.14}$$

Clearly, $IIP3$ is an important figure in RF low-noise amplifier. Because the interfering signals are often much stronger than the desired signal, thus generating strong intermodulation products will corrupt the weak but desired signal.

1.2 HBT Physical nonlinearities

Fig. 1.2 shows a typical SiGe HBT equivalent circuit. I_{CE} represents the collector current which is a nonlinearity controlled by V_{BE}. I_{BE} is the hole current representing hole injection into the emitter, and is a nonlinearity controlled by V_{BE} as well. I_{CB} represents the avalanche current which is a strong nonlinearity controlled by both V_{BE} and V_{CB}. C_{BE}

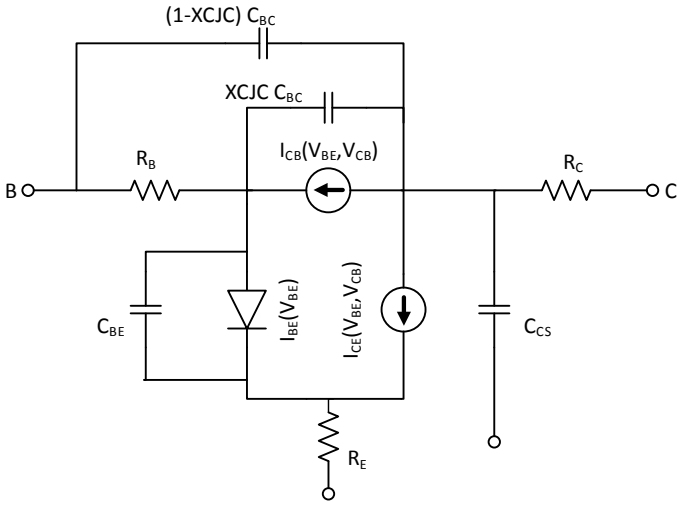

Figure 1.2: SiGe HBT equivalent circuit[1].

is the emitter-base junction capacitance, which includs the diffusion capacitance and depletion capacitance. Since diffusion charge is proportional to the transport current I_{CE}, C_{BE} becomes a strong nonlinearity controlled by V_{BE} when the diffusion capacitance dominates. C_{BC} is the collector-base junction capacitance.

1.2.1 I_{CE} nonlinearity

To first order, I_{CE} is controlled by V_{BE} which results in a nonlinear transconductance in weakly nonlinear circuit analysis. Assuming the nonlinear current $i(t)$ as a function of controlling voltage $v_C(t)$, one has[1]

$$\begin{aligned} i(t) &= f(v_C(t)) = f(V_C + v_c(t)) \\ &= f(V_C) + \sum_{k=1}^{\infty} \frac{1}{k!} \frac{\partial^k f(v(t))}{\partial v^k}\bigg|_{v=V_C} \times v_c^k(t). \end{aligned} \tag{1.15}$$

where $i(t)$ is the sum of the dc current and ac current, $v_c(t)$ is the ac voltage and V_C is the dc bias voltage.

Generally, it is enough to consider the first three terms of the power series for small $v_c(t)$. Thus the conductance nonlinear coefficients can be defined

$$\begin{aligned} g &= \left.\frac{\partial f(v)}{\partial v}\right|_{v=V_C}, \\ k_{2g} &= \left.\frac{1}{2!}\frac{\partial^2 f(v)}{\partial v^2}\right|_{v=V_C}, \\ k_{3g} &= \left.\frac{1}{3!}\frac{\partial^3 f(v)}{\partial v^3}\right|_{v=V_C}, \end{aligned} \tag{1.16}$$

$$k_{ng} = \left.\frac{1}{n!}\frac{\partial^n f(v)}{\partial v^n}\right|_{v=V_C}. \tag{1.17}$$

Then the ac current can be rewritten as

$$i_{ac}(t) = g \cdot v_c(t) + K_{2g} \cdot v_c^2(t) + K_{3g} \cdot v_c^3(t) + \cdots, \tag{1.18}$$

where g is the linearized small-signal transconductance element, K_{2g} and K_{3g} are the second-order and third-order nonlinearity coefficients respectively. The subscript g indicates that these coefficients are associated with the linearized transconductance "g". For an ideal SiGe HBT, I_{CE} is given by

$$I_{CE} = I_S \ \exp\left(\frac{qV_{BE}}{kT}\right) = I_S \ \exp\left(\frac{V_{BE}}{V_t}\right), \tag{1.19}$$

where $\frac{kT}{q}$ is the thermal voltage. The nonlinearity coefficients are

$$\begin{aligned} g_m &= \frac{I_{CE}}{V_t}, \\ k_{2gm} &= \frac{1}{2!}\frac{I_{CE}}{V_t^2}, \\ k_{3gm} &= \frac{1}{3!}\frac{I_{CE}}{V_t^3}, \end{aligned} \tag{1.20}$$

$$k_{ngm} = \frac{1}{n!}\frac{I_{CE}}{V_t^n}. \tag{1.21}$$

Thus the effective transconductance generate by $I_{CE} - V_{CE}$ nonlinearity can be obtained as

$$g_{m,eff} = \frac{i_c}{v_{be}} = g_m(1 + \frac{1}{2}\frac{v_{be}}{V_t} + \frac{1}{6}\frac{v_{be}^2}{V_t^2} + \cdots). \tag{1.22}$$

From (1.22) we can find that the nonlinear contributions to $g_{m,eff}$ increase with the voltage drop across the emitter-base junction. Generally, v_{be} decreases as biasing current increases, making $g_{m,eff}$ closer to constant. Thus the linearity of SiGe HBT can be improved by increasing biasing current, at the cost of power consumption. By decreasing v_{be} through increase emitter resistance or inductance can also help to improve its linearity, although at the expense of gain.

1.2.2 I_{BE} nonlinearity

The current gain β is constant when considering an ideal SiGe HBT. Then the base current I_{BE} represents as

$$I_{BE} = \frac{I_{CE}}{\beta}. \tag{1.23}$$

Therefore, the nonlinearity coefficients of the emitter-base conductance can be obtained similarly with the emitter-collector transconductance.

$$\begin{aligned} g_{be} &= \frac{g_m}{\beta}, \\ k_{2gbe} &= \frac{K_{2gm}}{\beta} = \frac{I_{CE}}{V_t^2 \beta}, \\ k_{3gbe} &= \frac{K_{3gm}}{\beta} = \frac{I_{CE}}{V_t^3 \beta}, \end{aligned} \tag{1.24}$$

$$k_{ngbe} = \frac{K_{ngm}}{\beta} = \frac{I_{CE}}{V_t^n \beta}. \tag{1.25}$$

1.2.3 I_{CB} nonlinearity

The I_{CB} current represents the avalanche current

$$\begin{aligned} I_{CB} &= I_{CE}(M-1) \\ &= I_{C0}(V_{BE}) F_{early}(M-1). \end{aligned} \tag{1.26}$$

where $I_{C0}(V_{BE})$ is I_C measured at zero V_{CB}, M is the avalanche factor and F_{early} is the Early effect factor.

At low current density, the avalanche factor M in SiGe HBT can be described as a V_{CB} function using Miller equation

$$M = \frac{1}{1-(V_{CB}/V_{CBO})^m}, \tag{1.27}$$

where V_{CBO} and m are fitting parameters.

The avalanche current I_{CB} is controlled by two voltages, V_{BE} and V_{CB}. Thus the power series of this current should be described by 2-D which including the cross-term contribution.

$$i_u = g_u \cdot u_c + K_{2gu} \cdot u_c^2 + K_{3gu} \cdot u_c^3 + \cdots , \tag{1.28}$$

$$i_v = g_v \cdot v_c + K_{2gv} \cdot v_c^2 + K_{3gv} \cdot v_c^3 + \cdots , \tag{1.29}$$

$$i_{uv} = K_{2gu\&gv} \cdot v_c \cdot u_c + K_{3_{2gu\&gv}} \cdot u_c^2 \cdot v_c + K_{3_{gu\&2gv}} \cdot u_c \cdot v_c^2. \tag{1.30}$$

The first two series i_u and i_v are 1-D nonlinear transconductance which contain one voltage, and the cross-term i_{uv} is 2-D nonlinear transconductance which is controlled by two voltage. Thus the nonlinear coefficient $K_{mjgu\&(m-j)gv}$ with $m > j$ is defined as

$$K_{mjgu\&(m-j)gv} = \frac{1}{j!}\frac{1}{(m-j)!}\frac{\partial^m f(u,v)}{\partial u^j \partial v^{(m-j)}}\bigg|_{u=U_C, v=V_C}. \tag{1.31}$$

1.2.4 C_{BE} and C_{BC} nonlinearity

Similarly to the nonlinearity of current, these capacitances are also nonlinear functions of the terminal voltage. For a small-signal distortion, the charge associated with nonlinear capacitor can be presented as[1]

$$Q(t) = f(v_C(t)) = f(V_C + v_c(t)) \tag{1.32}$$

$$= f(V_C) + \sum_{k=1}^{\infty} \frac{1}{k!}\frac{\partial^k f(v(t))}{\partial v^k}\bigg|_{v=V_C} \times v_c^k(t), \tag{1.33}$$

where *ac* part describes the power series of the stored charge. And nonlinearity coefficients are defined as

$$\begin{aligned} C &= \left.\frac{\partial f(v)}{\partial v}\right|_{v=V_C}, \\ k_{2C} &= \left.\frac{1}{2!}\frac{\partial^2 f(v)}{\partial v^2}\right|_{v=V_C}, \\ k_{3C} &= \left.\frac{1}{3!}\frac{\partial^3 f(v)}{\partial v^3}\right|_{v=V_C}. \end{aligned} \tag{1.34}$$

The *ac* charge can be written as

$$q_{ac}(t) = C \cdot v_c(t) + K_{2C} \cdot v_c^2(t) + K_{3C} \cdot v_c^3(t) + \cdots , \tag{1.35}$$

where C is the small-signal linear capacitance. The diffusion charge Q_D in SiGe HBT is proportional to I_{CE} through transit time τ_f

$$Q_D = \tau_f I_{CE} = \tau_f I_S \ \exp(\frac{qV_{BE}}{kT}). \tag{1.36}$$

Thus,

$$\begin{aligned} C_D &= \tau_f g_m = \tau_f \frac{qI_{CE}}{kT} = \tau_f \frac{I_{CE}}{V_t}, \\ K_{2C_D} &= \tau_f K_{2gm} = \tau_f \frac{q^2 I_{CE}}{2(kT)^2} = \tau_f \frac{I_{CE}}{2V_t^2}, \\ K_{3C_D} &= \tau_f K_{3gm} = \tau_f \frac{q^3 I_{CE}}{6(kT)^3} = \tau_f \frac{I_{CE}}{6V_t^3}, \end{aligned} \tag{1.37}$$

$$K_{nC_D} = \tau_f K_{ngm} = \tau_f \frac{q^n I_{CE}}{n!(kT)^n} = \tau_f \frac{I_{CE}}{n!V_t^n}. \tag{1.38}$$

In equation (1.37), C_D, K_{2C_D} and K_{3C_D} are proportional to g_m, K_{2gm} and K_{3gm} respectively. And the effective capacitance can be written as a function of v_{be},

$$C_{D,eff} = \frac{q_D}{v_b e} = C_D(1 + \frac{1}{2}\frac{v_{be}}{V_t} + \frac{1}{6}\frac{v_{be}^2}{V_t^2} + \cdots). \tag{1.39}$$

Similar to the effective transconductance $g_{m,eff}$, increasing bias current can make the effective diffusion capacitance more linear with expense of power consumption. A larger C_D itself also can improve the circuit linearity since higher I_C leads to a larger C_D which finally results in a smaller v_{be}.

The expression of the emitter-base and collector-base junction depletion capacitance are often given by

$$C_{dep}(V_f) = \frac{C_0}{(1 - \frac{V_f}{V_j})^{m_j}}, \tag{1.40}$$

where C_0, V_j and m_j are known model parameters. The nonlinearity coefficients can be analytical evaluated in this case. But more complicated in other models such as MEXTRAM model, the nonlinearity coefficients should be evaluated numerically.

Generally the collector-base depletion capacitance is much smaller than the emitter-base depletion capacitance, because the collector-base junction is under reserve bias for normal operation.

Through discussing these physical $I - V$ and $C - V$ nonlinearities in a SiGe HBT, we know how they influence the linearity of a SiGe HBT amplifier.

In order to find out which nonlinearity is dominant in the device, the analysis method used is Volterra series. This is an approach applies to small signal distortion such as that found in front-end low-noise amplifiers and mixers. In the next part, we will introduce one of Volterra series, the nonlinear current source method.

Chapter 2

Nonlinear current source approach

2.1 Introduction to Volterra series

Volterra series is a general mathematical approach for solving systems of nonlinear integral and integral-differential equations. Volterra series can be viewed as an extension of the theory of linear systems to weakly nonlinear systems. The essentials of Volterra series can be briefly summarized follows[1]:

Firstly, Volterra series describes a nonlinear system the same as a using Taylor series to approximate an analytic function. Similarly, the analysis is applicable only to weak nonlinearities.

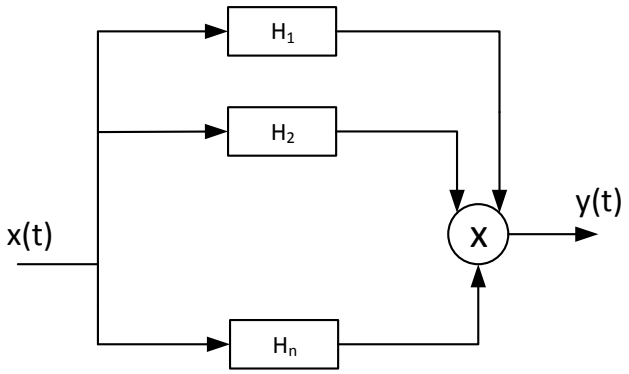

Figure 2.1: Illustrating basic process of Volterra series[3].

Secondly, the response of a nonlinear system to an input $x(t)$ is equal to the sum of the response of a series of transfer functions of different orders$(H_1, H_2, ..., H_n)$, as shown in Fig. 2.1

$$Y = H_1(x) + H_2(x) + H_3(x) + \cdots + H_n(x), \tag{2.1}$$

where in time domain, H_n is described as an impluse response $h_n(\tau_1, \tau_2, \cdots, \tau_n)$. As in linear circuit analysis, frequency domain representation is often more convenient. Thus nth order transfer function $H_n(s_1, s_2, \cdots, s_n)$ in frequency domain is obtained through a multidimensional Laplace transform of the time domain impluse response.

Thirdly, the first-order transfer function $H_1(s)$ is the transfer function at dc bias.

Thus solving the output of a nonlinear circuit is equivalent to solving the Volterra series $H_1(s)$, $H_2(s_1, s_2)$, and $H_3(s_1, s_2, s_3)$, $\cdots$

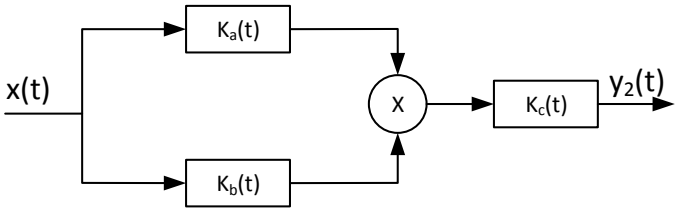

Figure 2.2: Illustrating the operation of a simple second-order system[3].

Fig. 2.2 shows a basic second-order system. It is clear that a second-order nonlinearity combines two signals to produce a second-order signal. With this simple system shown in Fig. 2.2, we can explain the operation of a second-order system in generally as follows: the two linear blocks represent the impulse response $k_a(t)$ and $k_b(t)$ which is fed from incoming signal $x(t)$, $z_a(t)$ and $z_b(t)$ are the two output respectively. Then these two signals are combined by a multiplier and the higher impulse response term is $k_c(t)$. $y_2(t)$ is the overall output of this linear system.

For this general system, it is not difficult to give the Volterra kernel equation[3]:

$$h_2(\tau_1, \tau_2) = \int_{-\infty}^{\infty} k_c(\sigma) k_a(\tau_1 - \sigma) k_b(\tau_2 - \sigma) d\sigma. \tag{2.2}$$

In this thesis we analyze nonlinear behavior in frequency domain, thus we can get equation in frequency domain using Fourier transforms. The Fourier transform of $h_2(\tau_1, \tau_2)$

is denoted as $H_2(j\omega_1, j\omega_2)$. The equation shows as follows[3]:

$$H_2(j\omega_1, j\omega_2) = K_a(j\omega_1)K_b(j\omega_2)K_c(j\omega_1 + j\omega_2), \tag{2.3}$$

where $K_a(j\omega)$, $K_b(j\omega)$ and $K_c(j\omega)$ denote the Fourier transfforms of the linear subsystems described by the impluse response $k_a(t)$, $k_b(t)$ and $k_c(t)$ respectively.

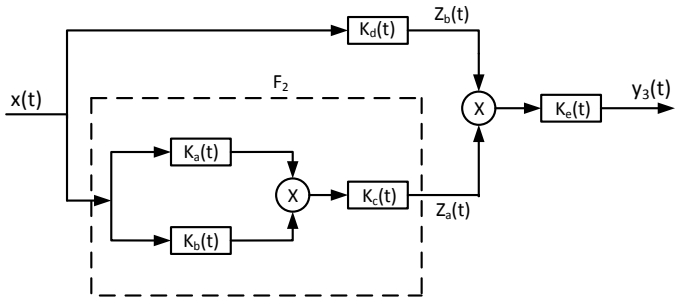

Figure 2.3: Illustrating the operation of a simple third-order system[3].

Similar to the second-order system, we can obtain a general block diagram to represent a third-order Volterra operator. It is clear that this third-order system will require two multipliers. In this way except the second-order signal, this system will combine three signals to produce a third-order signal. Fig. 2.3 shows a basic third-order system. $x(t)$ is the common input to a second-order system F_2 with the output $z_a(t)$, it's also the input to a linear system with impulse response $k_d(t)$ and output $z_b(t)$. The second-order system F_2 is defined in Fig. 2.2. Combining the output $z_a(t)$ and $z_b(t)$, we will obtain the impulse response $k_e(t)$ and total output $y_3(t)$.

The third-order kernel transform of this system in frequency domain is given by:

$$H_3(j\omega_1, j\omega_2, j\omega_3) = K_a(j\omega_1)K_b(j\omega_2)K_c(j\omega_1 + j\omega_2)K_d(j\omega_3)K_e(j\omega_1 + j\omega_2 + j\omega_3). \tag{2.4}$$

There are two approaches to solve nonlinear system based on Volterra series. One is nonlinear current source method[3], the other is direct derivation[11].

2.2 Nonlinear current source approach

In this section, the computation of first-order, second-order and third-order Volterra kernel of voltages and currents in a weakly nonlinear network will be explained.

2.2.1 First-order transfer function

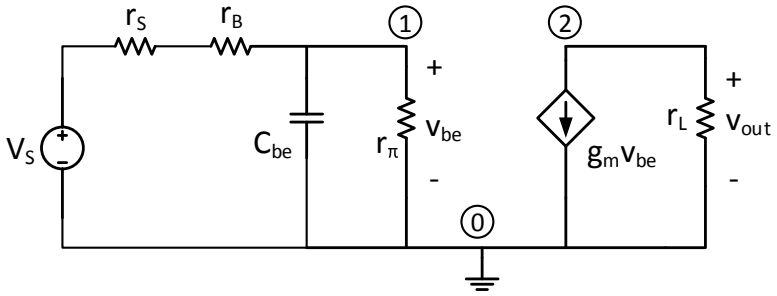

Figure 2.4: First-order small signal equivalent circuit.

At first, the response of the equivalent circuit to the external input should be transferred to a frequency domain function. In this way, every nonlinearity symbol is replaced by a linearized one. Considering the output voltage, all node voltage of this nonlinearity can be calculated. Fig. 2.4 shows the equivalent circuit of the first-order response. In general this calculation can be represented by the solution of the following matrix equation[3][4]:

$$Y(s)H_1(s) = IN_1, \tag{2.5}$$

where $Y(s)$ is the admittance matrix in this circuit, $H_1(s)$ is the vector of first-order Volterra kernel transforms of the node voltage 1 and 2, IN_1 is the vector of excitation. This admittance matrix can be calculated using Kirchoff's current law at every node. Thus in this equation the only unknown parameter is the Volterra transfer function. Applying Kirchoff's current law at node 1 in Fig. 2.4. we can obtain:

$$\frac{g_S \cdot g_B}{g_S + g_B}(V_1 - V_S) + g_\pi V_1 + s_1 C_{be} V_1 = 0. \tag{2.6}$$

At node 2 we can obtain the equation in the same way:

$$g_m V_{be} + g_L V_2 = 0. \tag{2.7}$$

In (2.6), g_S is the admittance of the source resistance and g_B is the admittance of the base resistance. C_{be} is the sum of diffusion capacitance C_π and depletion capacitance C_{je}. In (2.7), V_{be} equal to V_1.

Then combine (2.6) and (2.7) into one matrix equation:

$$\begin{bmatrix} \frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + sC_{be} & 0 \\ g_m & g_L \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} g_S V_S \\ 0 \end{bmatrix}. \tag{2.8}$$

In this matrix, V_1 and V_2 become the first-order Volterra transfer function when V_S is equal to 1. These transfer functions are denoted as $H_{1_1}(s)$ and $H_{1_2}(s)$. And in these two transfer function, the first subscript indicates the order of transfer function and the second subscript indicates the number of node. Then (2.9) can be rewritten as following:

$$\begin{bmatrix} \frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + sC_{be} & 0 \\ g_m & g_L \end{bmatrix} \begin{bmatrix} H_{1_1}(s) \\ H_{1_2}(s) \end{bmatrix} = \begin{bmatrix} g_S \\ 0 \end{bmatrix}. \tag{2.9}$$

This matrix equation can be solved using Cramer's rule. Then through defining the ratio of two determinants, $H_{1_1}(s)$ and $H_{1_2}(s)$ can be calculated.

$$det(s) = (\frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + sC_{be})(g_L). \tag{2.10}$$

Then H_{1_1} is obtained as

$$H_{1_1}(s) = \frac{\begin{vmatrix} g_S & 0 \\ 0 & g_L \end{vmatrix}}{det(s)} = \frac{g_S(g_L)}{dets(s)} = \frac{g_S}{\frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + sC_{be}}. \tag{2.11}$$

H_{1_2} is obtained as

$$H_{1_2}(s) = \frac{\begin{vmatrix} \frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + sC_{be} & g_S \\ gm & 0 \end{vmatrix}}{det(s)} = \frac{-g_m g_S}{dets(s)} = \frac{-g_m g_S}{(\frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + sC_{be})g_L}. \tag{2.12}$$

2.2.2 Second-order transfer function

In fact the first-order kernels are a linear transfer function. The method that was used for calculation of linear transfer functions described earlier can also be used to calculate the second-order Volterra kernels, but now with the so-called nonlinear current sources placed in parallel with the corresponding linearized circuit elements for each nonlinear element. In this way the node voltages that are found are equal to the second-order kernel transforms.

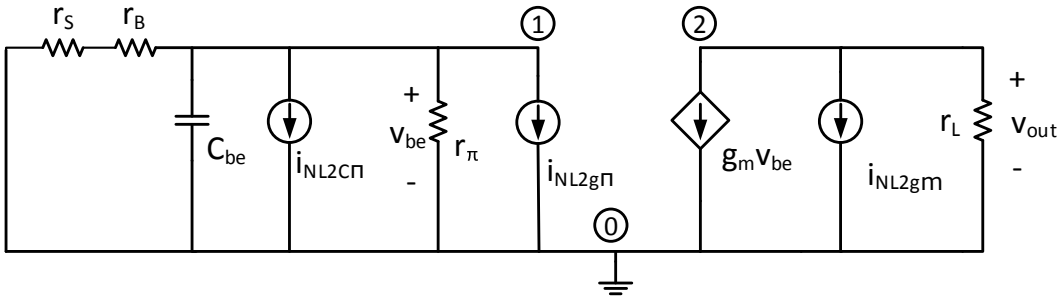

Figure 2.5: The computation of second-order kernels equivalent circuit.

Fig. 2.5 shows the circuit that has to be solved for the computation of $H_2(s_1, s_2)$. Every nonlinearity symbol in the original circuit gives rise to a nonlinear current source in the linearized circuit, which is placed parallel to each nonlinear element. And the direction of

these nonlinear current source is the same as the controlled current in the original nonlinear circuit.

In this case, we assume the base resistance r_B is a linear element. In Fig. 2.5, $i_{NL2_{g\pi}}$, $i_{NL2_{c\pi}}$ and $i_{NL2_{gm}}$ are the second-order nonlinear current sources.

type of basic nonlinearity	expression for nonlinear current source of order two
(trans)conductance	$K_{2_{g1}} H_{1k}(s_1) H_{1k}(s_2)$
capacitance	$(s_1 + s_2) K_{2_{C1}} H_{1k}(s_1) H_{1k}(s_2)$
2-D conductance(only cross term)	$\frac{1}{2}\left[K_{2_{g1\&g2}} H_{1k}(s_1) H_{1l}(s_2) + K_{2_{g1\&g2}} H_{1k}(s_2) H_{1l}(s_1)\right]$
3-D conductance(only cross term)	0

Table 2.1: Different basic nonlinearities of second-order nonlinear current source which is used for the calculation of the second-order Volterra series[3].

The value of these current sources depends on the second-order nonlinearity coefficient, on the type of the nonlinear element (ie:transconductance, capacitance) in original circuit, and on the first-order kernels of the controlling voltage. The second-order nonlinear current source are given in Table (2.1). In this table $H_{1_k}(s)$ is the first order transfer function of the node voltage which controls nonlinearity conductance and capacitance. $H_{1_l}(s)$ is the first-order transfer function the second controlled voltage. $K_{2_{g1}}$ and $K_{2_{C1}}$ are the seconder-order nonlinearity coefficient respectively. It is clear that the value of nonlinear current source is equal to the second-order nonlinearity coefficient multiply with the first-order transfer function of the voltage at different frequency.

Again matrix equation can be obtained as[3][4]:

$$Y(s_1 + s_2) H_2(s_1, s_2) = IN_2. \tag{2.13}$$

In this equation, second-order kernel transforms of the node voltage is represented as the vector $H_2(s_1, s_2)$. Vector IN_2 relate to the nonlinear current sources of order two. Comparing equation (2.5) and (2.13), we can find that they have the same admittance matrix, but with different frequency.

In this way, the second-order nonlinearity can be calculated through combining two first-order controlling voltages to produce a second-order signal. To calculate the second-order kernel transform, the circuit should be linearized and the nonlinear current source of order two are applied. Through using Kirchoff's current law at node 1 and node 2 in Fig., the matrix equation can be obtained as follows:

$$\begin{bmatrix} \frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + (s_1 + s_2)C_{be} & 0 \\ g_m & g_L \end{bmatrix} \begin{bmatrix} H_{2_1}(s_1, s_2) \\ H_{2_2}(s_1, s_2) \end{bmatrix} = \begin{bmatrix} -i_{NL2_{g\pi}} - i_{NL2_{c\pi}} \\ -i_{NL2_{gm}} \end{bmatrix}. \quad (2.14)$$

The leftmost matrix is similar with the one in linear circuit, but now evaluated at $(s_1 + s_2)$ instead of s_1. In (2.14) the unknown parameters are the second-order kernels of the two node voltage $H_{2_1}(s_1, s_2)$ and $H_{2_2}(s_1, s_2)$. The nonlinear current $i_{NL2_{g\pi}}$ is given according to Table (2.1):

$$i_{NL2_{g\pi}} = K_{2_{g\pi}} H_{1_1}(s_1) H_{1_1}(s_2). \quad (2.15)$$

Using $H_{1_1}(s)$ getting from (2.11), we can obtain (2.15):

$$i_{NL2_{g\pi}} = \frac{K_{2_{g\pi}} g_S^2}{det(s_1) det(s_2)}. \quad (2.16)$$

Similarly, we can find the value of $i_{NL2_{gm}}$ and $i_{NL2_{C_\pi}}$:

$$i_{NL2_{gm}} = K_{2_{gm}} H_{1_1}(s_1) H_{1_1}(s_2) = \frac{K_{2_{gm}} g_S^2}{det(s_1) det(s_2)} \quad (2.17)$$

$$i_{NL2_{C_\pi}} = (s_1 + s_2) K_{2_{C_\pi}} H_{1_1}(s_1) H_{1_1}(s_2) = \frac{(s_1 + s_2) K_{2_{C_\pi}} g_S^2}{det(s_1) det(s_2)}. \quad (2.18)$$

Applying Cramer's rule on (2.14) and using the value of second-order nonlinear current source, $H_{2_1}(s_1, s_2)$ can be obtained as:

$$H_{2_1}(s_1, s_2) = \frac{\begin{vmatrix} -i_{NL2_{g\pi}} - i_{NL2_{C_\pi}} & 0 \\ -i_{NL2_{gm}} & g_L \end{vmatrix}}{det(s_1 + s_2)} = \frac{-g_L}{dets(s_1 + s_2)}(i_{NL2_{g\pi}} + i_{NL2_{C_\pi}}). \tag{2.19}$$

Then using (2.16) and (2.18), (2.20) can be rewritten as:

$$H_{2_1}(s_1, s_2) = \frac{-g_L g_S^2(K_{2_{g\pi}} + (s_1 + s_2)K_{2_{C_\pi}})}{dets(s_1 + s_2)det(s_1)det(s_2)}. \tag{2.20}$$

For the second-order kernel transfer function at node 2:

$$H_{2_2}(s_1, s_2) = \frac{\begin{vmatrix} \frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + (s_1 + s_2)C_{be} & -i_{NL2_{g\pi}} - i_{NL2_{C_\pi}} \\ g_m & -i_{NL2_{g_m}} \end{vmatrix}}{det(s_1 + s_2)} \tag{2.21}$$

$$= \frac{-(\frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + (s_1 + s_2)C_{be})i_{NL2_{g_m}} + g_m(i_{NL2_{g\pi}} + i_{NL2_{C_\pi}})}{dets(s_1 + s_2)} \tag{2.22}$$

Using (2.10), (2.16), (2.18) and (2.17), we can obtain:

$$H_{2_2}(s_1, s_2) = \frac{-g_S^2(K_{2_{g_m}}(\frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + (s_1 + s_2)C_{be}) - g_m(K_{2_{g\pi}} + (s_1 + s_2)K_{2_{C_\pi}})}{det^3(s_1)det^3(s_2)det(s_1 + s_2)} \tag{2.23}$$

2.2.3 Third-order transfer function

In the next step, the third-order transfer functions are calculated. Similarly, the third-order must be solved as a mix of order one and two[3][4]:

$$Y(s_1 + s_2 + s_3)H_3(s_1, s_2, s_3) = IN_3 \tag{2.24}$$

type of basic nonlinearity	expression for nonlinear current source of order two
(trans)conductance	$K_{3_{g1}} H_{1_k}(s_1) H_{1_k}(s_2) H_{1_k}(s_3)$ $+\frac{2}{3}\left[H_{1_k}(s_1) H_{2_k}(s_2, s_3) + H_{1_k}(s_2) H_{2_k}(s_1, s_3)\right.$ $\left.+H_{1_k}(s_3) H_{2_k}(s_1, s_2)\right]$
capacitance	$(s_1 + s_2 + s_3) K_{3_{C1}} H_{1_k}(s_1) H_{1_k}(s_2) H_{1_k}(s_3)$ $+\frac{2}{3}(s_1 + s_2 + s_3) K_{2_{C1}}\left[H_{1_k}(s_1) H_{2_k}(s_2, s_3)\right.$ $\left.+H_{1_k}(s_2) H_{2_k}(s_1, s_3) + H_{1_k}(s_3) H_{2_k}(s_1, s_2)\right]$
2-D conductance(only cross term)	$\frac{1}{3} K_{2_{g1\&g2}}\left[H_{1_k}(s_1) H_{2_l}(s_2, s_3) + H_{1_k}(s_2) H_{2_l}(s_1, s_3)\right.$ $+H_{1_k}(s_3) H_{2_l}(s_1, s_2) + H_{2_k}(s_1, s_2) H_{1_l}(s_3)$ $\left.+H_{2_k}(s_1, s_3) H_{1_l}(s_2) + H_{2_k}(s_2, s_3) H_{1_l}(s_1)\right]$ $\frac{1}{3} K_{3_{2g1\&g2}}\left[H_{1_k}(s_1) H_{1_k}(s_2) H_{1_l}(s_3) + H_{1_k}(s_1) H_{1_k}(s_3) H_{1_l}(s_2)\right.$ $\left.+H_{1_k}(s_2) H_{1_k}(s_3) H_{1_l}(s_1)\right]$ $\frac{1}{3} K_{3_{g1\&2g2}}\left[H_{1_k}(s_1) H_{1_l}(s_2) H_{1_l}(s_3)\right.$ $\left.+H_{1_k}(s_2) H_{1_l}(s_1) H_{1_l}(s_3) + H_{1_k}(s_3) H_{1_l}(s_1) H_{1_l}(s_2)\right]$
3-D conductance(only cross term)	$\frac{1}{6} K_{3_{g1\&g2\&g3}}\left[H_{1_k}(s_1) H_{1_l}(s_2) H_{1_m}(s_3)\right.$ $+H_{1_k}(s_1) H_{1_l}(s_3) H_{1_m}(s_2) + H_{1_k}(s_2) H_{1_l}(s_1) H_{1_m}(s_3)$ $+H_{1_k}(s_2) H_{1_l}(s_3) H_{1_m}(s_1) + H_{1_k}(s_3) H_{1_l}(s_1) H_{1_m}(s_2)$ $\left.H_{1_k}(s_3) H_{1_l}(s_2) H_{1_m}(s_1)\right]$

Table 2.2: Different basic nonlinearities of third-order nonlinear current source which is used for the calculation of the third-order Volterra series transfer function[3].

The nonlinear current source of order three are shown in Table (2.2). $H_{1_k}(s)$ and $H_{2_k}(s)$ are the first-order and second-order transfer functions of the voltage respectively.

The nonlinear current source consist of two components. One is caused by the third-order nonlinearity, which combines three first-order signals into a third-order signal at frequency $(s_1 + s_2 + s_3)$. Another component is the average terms which acts upon a second-order signal and a first-order signal at controlling voltage. Then all these third-order signals propagate to the output.

Then the third-order transfer functions' matrix equation is given by:

$$\begin{bmatrix} \frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + (s_1 + s_2 + s_3) C_{be} & 0 \\ g_m & g_L \end{bmatrix} \begin{bmatrix} H_{3_1}(s_1, s_2, s_3) \\ H_{3_2}(s_1, s_2, s_3) \end{bmatrix} = \begin{bmatrix} -i_{NL3_{g\pi}} - i_{NL3_{c\pi}} \\ -i_{NL3_{gm}} \end{bmatrix}, \tag{2.25}$$

where $H_{3_1}(s_1, s_2, s_3)$ and $H_{3_2}(s_1, s_2, s_3)$ are the third-order transfer functions at node 1 and node 2 respectively. From (2.25), $H_{3_1}(s_1, s_2, s_3)$ and $H_{3_2}(s_1, s_2, s_3)$ can be found using rule

of Cramer

$$H_{3_1}(s_1, s_2, s_3) = \frac{g_L(i_{NL3_{g\pi}} + i_{NL3_{c\pi}})}{det(s_1 + s_2 + s_3)} \tag{2.26}$$

$$H_{3_2}(s_1, s_2, s_3) = \frac{-(\frac{g_S \cdot g_B}{g_S + g_B} + g_\pi + (s_1 + s_2 + s_3)C_{be})i_{NL3_{g_m}} + g_m(i_{NL3_{g\pi}} + i_{NL3_{c\pi}})}{det(s_1 + s_2 + s_3)}, \tag{2.27}$$

where $i_{NL3_{gm}}$, $i_{NL3_{g\pi}}$ and $i_{NL3_{c\pi}}$ can be calculated using Table (2.2).

$$\begin{aligned} i_{NL3_{gm}} &= K_{3_{gm}} H_{1_1}(s_1) H_{1_1}(s_2) H_{1_1}(s_3) \\ &+ \frac{2}{3} K_{2_{gm}} \left[H_{1_1}(s_1) H_{2_1}(s_2, s_3) + H_{1_1}(s_2) H_{2_1}(s_1, s_3) + H_{1_1}(s_3) H_{2_1}(s_1, s_2) \right] \end{aligned} \tag{2.28}$$

$$\begin{aligned} i_{NL3_{g\pi}} &= K_{3_{g\pi}} H_{1_1}(s_1) H_{1_1}(s_2) H_{1_1}(s_3) \\ &+ \frac{2}{3} K_{2_{g\pi}} \left[H_{1_1}(s_1) H_{2_1}(s_2, s_3) + H_{1_1}(s_2) H_{2_1}(s_1, s_3) + H_{1_1}(s_3) H_{2_1}(s_1, s_2) \right] \end{aligned} \tag{2.29}$$

$$\begin{aligned} i_{NL3_{c\pi}} &= (s_1 + s_2 + s_3) K_{3_{c\pi}} H_{1_1}(s_1) H_{1_1}(s_2) H_{1_1}(s_3) \\ &+ \frac{2}{3}(s_1 + s_2 + s_3) K_{2_{c\pi}} \left[H_{1_1}(s_1) H_{2_1}(s_2, s_3) + H_{1_1}(s_2) H_{2_1}(s_1, s_3) \right. \\ &\left. + H_{1_1}(s_3) H_{2_1}(s_1, s_2) \right] \end{aligned} \tag{2.30}$$

Substituting first-order transfer function and second-order transfer function into $i_{NL3_{gm}}$, $i_{NL3_{g\pi}}$ and $i_{NL3_{c\pi}}$, we can obtain $H_{3_2}(s_1, s_2, s_3)$.

In next part, direct derivation method is illustrated. It is more complex than nonlinear current method. The equation of nonlinear output must be the same in same conditions for two approaches.

Chapter 3

Direct derivation approach

3.1 First order derivation

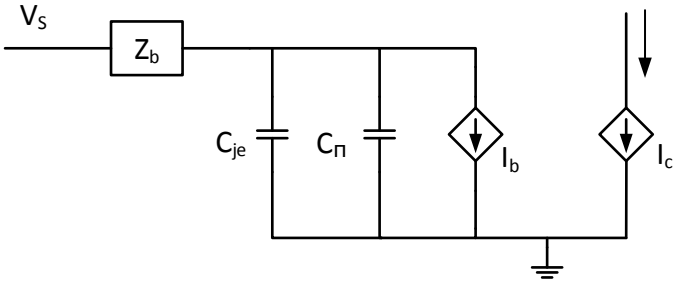

Figure 3.1: Circuit that has to be used for the analysis of distortion based on direct derivation approach[11].

Another approach is direct derivation. Fig. 3.1 shows the circuit that used to analyze distortion. This circuit ignores the effect of base-collector junction capacitance (C_{cb})as we do in nonlinear current source circuit. V_S is the voltage source. Z_b is the impedance between voltage source and base of transistor which set as the sum of source resistance (r_s) and base resistance of transistor (r_b). Q_π is the base-emitter diffusion charge, which is linearly proportional to the collector current I_c and the forward transit time τ_f. C_π is the base-emitter difussion capacitance, which is equal to $\frac{dQ_\pi}{dV_\pi} = g_m\tau_f$. C_{je} is the base-emitter depletion capacitance which is assumed to be a constant value in this model. I_b is the base current, which is equal to $\frac{I_c}{\beta}$, where β is the small signal low-frequency current gain.But at high-frequency I_b still have effect on the nonlinearity, thus it can't be ignored.

Using Kirchhoff's Voltage Law at Fig. 3.1, we can obtain an equation of the circuit[2][11]

$$V_S = (I_{C_{je}} + I_{C_\pi} + I_b)Z_b, \tag{3.1}$$

where $I_{C_{je}}$ is the current of C_{je} and I_{C_π} is the current of C_π. Then we can obtain

$$V_S = (sC_{je}V_\pi + s\tau_f g_m V_\pi + \frac{I_c}{\beta})Z_b + V_\pi, \tag{3.2}$$

where V_π is the base-emitter voltage drop accross C_π and C_{je}. For collector current I_c, a Volterra series expression is given by

$$I_c = A_1(s) \circ V_S + A_2(s_1, s_2) \circ V_S^2 + A_3(s_1, s_2, s_3) \circ V_S^3, \tag{3.3}$$

where V_S^n is the n'th power of the voltage source. $A_n(s)$ is the Volterra series coefficient. The operator $'\circ'$ in equation (3.3) indicates each component in V_S^n should multiply $A_n(s)$ at each frequency which result into a phase shifting by $A_n(s)$ of each frequency component in V_S^n.

$$\begin{aligned} I_c &= I_C + I_Q \exp(\frac{V_\pi}{V_t}) \\ &= I_C + I_Q \left[(\frac{V_\pi}{V_t}) + \frac{1}{2}(\frac{V_\pi}{V_t})^2 + \frac{1}{6}(\frac{V_\pi}{V_t})^3 + \cdots \right], \end{aligned} \tag{3.4}$$

where I_Q is the bias current of transistor, I_C is dc current and V_t is the thermal voltage.

$$V_\pi = C_1(s_1) \circ V_S + C_2(s_1, s_2) \circ V_S^2 + C_3(s_1, s_2, s_3) \circ V_S^3. \tag{3.5}$$

The way to calculate $A_n(s)$ is substituting equation (3.3) and (3.5) into equation (3.2). For the first order derivation,

$$\begin{aligned} V_S &= Ae^{j\omega_1 t}, \\ i_c &= I_Q(\frac{V_\pi}{V_t}), \\ V_\pi &= C_1(s_1) \circ V_S = C_1(s_1)e^{j\omega_1 t}, \end{aligned}$$

where we assuming A equal to 1, i_c is ac current. Then $C_1(s_1)$ can be calculated as

$$e^{j\omega_1 t} = \left[s_1 C_{je} C_1(s_1) e^{j\omega_1 t} + s_1 \tau f \frac{I_Q}{V_t} C_1(s_1) e^{j\omega_1 t} + \frac{I_Q}{V_t} C_1(s_1) e^{j\omega_1 t}/\beta_0 \right] \cdot Z_b(s_1) + C_1(s_1) e^{j\omega_1 t},$$

$$1 = C_1(s_1) \left[(s_1 C_{je} + s_1 \tau f \cdot g_m + \frac{g_m}{\beta_0}) \cdot Z_b(s_1) + 1 \right],$$

$$C_1(s_1) = \frac{1}{s_1 C_{je} Z_b(s_1) + s_1 \tau f \cdot g_m Z_b(s_1) + g_m Z_b(s_1)/\beta_0 + 1}. \tag{3.6}$$

Then $A_1(s_1)$ can be obtained as

$$\begin{aligned} A_1(s_1) = \frac{i_c}{N} &= \frac{g_m C_1(s_1) e^{j\omega_1 t}}{e^{j\omega_1 t}} \\ &= g_m C_1(s_1), \end{aligned} \tag{3.7}$$

where $N = e^{j\omega_1 t}$.

3.2 Second order derivation

For the second order derivation, the mixed frequency $\omega_1 + \omega_2$ occurs due to the square operator.

$$\begin{aligned} V_s &= e^{j\omega_1 t} + e^{j\omega_2 t}, \\ V_S^2 &= (e^{j\omega_1 t} + e^{j\omega_2 t})^2 = e^{2j\omega_1 t} + e^{2j\omega_2 t} + 2e^{j(\omega_1+\omega_2)t}, \\ i_c &= I_Q[(\frac{V_\pi}{V_t}) + \frac{1}{2}(\frac{V_\pi}{V_t})^2], \\ V_\pi &= C_1(s_1) \circ V_S + C_2(s_1, s_2) \circ V_S^2 \\ &= V_\pi = C_1(s_1) e^{j\omega_1 t} + C_1(s_2) e^{j\omega_2 t} + C_2(s_1, s_2) 2e^{j(\omega_1+\omega_2)t}. \end{aligned} \tag{3.8}$$

Since second-order $C_2(s_1, s_2)$ only considering the frequency at $\omega_1 + \omega_2$, thus neglect the influence of $e^{2j\omega_1 t}$&$e^{2j\omega_2 t}$. Then we can calculate V_π^2 as

$$V_\pi^2 = \left[C_1(s_1)e^{j\omega_1 t} + C_1(s_2)e^{j\omega_2 t} + C_2(s_1, s_2)2e^{j(\omega_1+\omega_2)t}\right]^2.$$

For the same reason the second-order only considering the frequency at $\omega_1 + \omega_2$, the square of each numerator in V_π^2 will lead to a even higher order. Thus neglect influence of these orders and we can rewrite V_π^2 as

$$V_\pi^2 = 2C_1(s_1)C_2(s_2)e^{j(\omega_1+\omega_2)t}. \tag{3.9}$$

Within (3.8) and (3.9), i_c can be obtained as

$$i_c = \frac{I_Q}{V_t}2C_2(s_1, s_2)e^{j(\omega 1+\omega_2)t} + \frac{I_Q}{2V_t^2}2C_1(s_1)C_1(s_2)e^{j(\omega_1+\omega_2)t}. \tag{3.10}$$

Substuiting (3.8), (3.9) and (3.10) into equation (3.2), $C_2(s_1, s_2)$ can be obtained as

$$\begin{aligned} V_S = &\Big\{(s_1 + s_2)C_{je}C_2(s_1, s_2)2e^{j(\omega_1+\omega_2)t} \\ &+ (s_1 + s_2)\tau f\left[\frac{I_Q}{V_t} \cdot 2C_2(s_1, s_2)e^{j(\omega_1+\omega_2)t} + \frac{I_Q}{2V_t^2} \cdot 2C_1(s_1)C_1(s_2)e^{j(\omega_1+\omega_2)t}\right] \\ &+ \frac{I_Q}{V_t} \cdot 2C_2(s_1, s_2)e^{j(\omega_1+\omega_2)t}/\beta_0 + \frac{I_Q}{2V_t^2} \cdot 2C_1(s_1)C_1(s_2)e^{j(\omega_1+\omega_2)t}/\beta_0\Big\} \cdot Z_b(s_1 + s_2) \\ &+ C_2(s_1, s_2) \cdot 2e^{j(\omega_1+\omega_2)t}, \end{aligned} \tag{3.11}$$

where $V_S = e^{j\omega_1 t} + e^{j\omega_2 t}$. Since no item relates to the frequency of $\omega_1 + \omega_2$ in this V_S, thus $V_S = 0$. Placing this requirement into equation (3.11)

$$\begin{aligned}0 = C_2(s_1,s_2)\Bigg[&(s_1+s_2)C_{je}Z_b(s_1+s_2)\cdot 2e^{j(\omega_1+\omega_2)t} + (s_1+s_2)\tau_f\frac{I_Q}{V_t}Z_b(s_1+s_2)\cdot 2e^{j(\omega_1+\omega_2)t}\\ &+\frac{I_Q}{V_t}Z_b(s_1+s_2)\cdot 2e^{j(\omega_1+\omega_2)t}/\beta_0 + 2e^{j(\omega_1+\omega_2)t}\Bigg]\\ &+ C_1(s_1)C_1(s_2)\left[(s_1+s_2)\tau_f\frac{I_Q}{2V_t^2}Z_b(s_1+s_2)\cdot 2e^{j(\omega_1+\omega_2)t} + \frac{I_Q}{2V_t^2}Z_b(s_1+s_2)\cdot 2e^{j(\omega_1+\omega_2)t}/\beta_0\right].\end{aligned} \tag{3.12}$$

The equation in first bracket is equal to $\dfrac{2}{C_1(s_1+s_2)}\cdot e^{j(\omega_1+\omega_2)t}$. Then (3.12) can be rewritten as

$$\begin{aligned}-C_2(s_1,s_2)\cdot\frac{2}{C_1(s_1+s_2)}e^{j(\omega_1+\omega_2)t} = C_1(s_1)C_2(s_2)\frac{I_Q}{2V_t^2}&\left[(s_1+s_2)\tau_f\cdot 2Z_b(s_1+s_2)\right.\\ &\left.+2Z_b(s_1+s_2)/\beta_0\right]e^{j(\omega_1+\omega_2)t}\end{aligned}$$

$$C_2(s_1,s_2) = -C_1(s_1)C_1(s_2)C_1(s_1+s_2)\frac{g_m}{2V_t}\left[(s_1+s_2)\tau_f Z_b(s_1+s_2) + \frac{Z_b(s_1+s_2)}{\beta_0}\right]. \tag{3.13}$$

Then $A_2(s_1,s_2)$ can be obtained as

$$A_2(s_1,s_2) = \frac{i_c}{N^2}, \tag{3.14}$$

where $N^2 = 2e^{j(\omega_1+\omega_2)t}$. Substuiting (3.10) into (3.14), we can obtain

$$A_2(s_1,s_2) = g_m\cdot C_2(s_1,s_2) + \frac{I_Q}{2V_t^2}C_1(s_1)C_1(s_2). \tag{3.15}$$

3.3 Third order derivation

For the third order derivation, frequency $\omega_1+\omega_2+\omega_3$ occurs at three first order frequency mixed or the first order and second order frequency mixed.

$$
\begin{aligned}
V_S &= e^{j\omega_1 t} + e^{j\omega_2 t} + e^{j\omega_3 t}, \\
i_c &= I_Q \left[(\frac{V_\pi}{V_t}) + \frac{1}{2}(\frac{V_\pi}{V_t})^2 + \frac{1}{6}(\frac{V_\pi}{V_t})^3 \right], \\
V_\pi &= C_1(s_1) \circ V_S + C_2(s_1, s_2) \circ V_S^2 + C_3(s_1, s_2, s_3) \circ V_S^3.
\end{aligned}
$$

Since third-order only considering the frequency at $\omega_1 + \omega_2 + \omega_3$, thus we can get the expression of I_C as

$$
\begin{aligned}
i_c &= \frac{I_Q}{V_t} \cdot 6C_3(s_1, s_2, s_3) e^{j(\omega_1+\omega_2+\omega_3)t} + \frac{I_Q}{6V_t^3} \cdot 6C_1(s_1)C_1(s_2)C_1(s_3) e^{j(\omega_1+\omega_2+\omega_3)t} \\
&= \frac{I_Q}{2V_t^2} \cdot 12\overline{C_1C_2} e^{j(\omega_1+\omega_2+\omega_3)t},
\end{aligned} \tag{3.16}
$$

where

$$
\overline{C_1C_2} = \frac{C_1(s_1)C_2(s_2, s_3) + C_1(s_2)C_2(s_1, s_3) + C_1(s_3)C_2(s_1, s_2)}{3}. \tag{3.17}
$$

Substituting V_π and (3.16) in (3.2)

$$
\begin{aligned}
V_S = \big\{ & (s_1 + s_2 + s_3)C_{je} \cdot 6C_3(s_1, s_2, s_3) e^{j(\omega_1+\omega_2+\omega_3)t} \\
& + \left((s_1 + s_2 + s_3)\tau_f + \frac{1}{\beta_0} \right) \left[\frac{I_Q}{V_t} \cdot 6C_3(s_1, s_2, s_3) e^{j(\omega_1+\omega_2+\omega_3)t} \right. \\
& \left. + \frac{I_Q}{6V_t^3} \cdot 6C_1(s_1)C_1(s_2)C_1(s_3) e^{j(\omega_1+\omega_2+\omega_3)t} + \frac{I_Q}{2V_t^2} \cdot 12\,\overline{C_1C_2} e^{j(\omega_1+\omega_2+\omega_3)t} \right] \cdot Z_b(s_1 + s_2 + s_3) \\
& + 6C_3(s_1 + s_2 + s_3) e^{j(\omega_1+\omega_2+\omega_3)t},
\end{aligned} \tag{3.18}
$$

where $V_S = e^{j\omega_1 t} + e^{j\omega_2 t} + e^{j\omega_3 t}$. This V_S has no item related to the frequency of $\omega_1 + \omega_2 + \omega_3$, thus it is equal to 0. Then the equation (3.18) can be rewritten as

$$\begin{aligned} 0 = {} & 6C_3(s_1,s_2,s_3)e^{j(\omega_1+\omega_2+\omega_3)t}\left[(s_1+s_2+s_3)C_{je} + (s_1+s_2+s_3)\tau_f\frac{I_Q}{V_t}\right. \\ & \left. + \frac{I_Q}{V_t\beta_0}\right]\cdot Z_b(s_1+s_2+s_3) + 1 \\ & + \frac{I_Q}{6V_t^3}\cdot 6C_1(s_1)C_1(s_2)C_1(s_3)e^{j(\omega_1+\omega_2+\omega_3)t}\left[(s_1+s_2+s_3)\tau f + \frac{1}{\beta_0}\right]\cdot Z_b(s_1+s_2+s_3) \\ & + \frac{I_Q}{2V_t^2}\cdot 12\ \cdot \overline{C_1C_2}e^{j(\omega_1+\omega_2+\omega_3)t}\left[(s_1+s_2+s_3)\tau f + \frac{1}{\beta_0}\right]\cdot Z_b(s_1+s_2+s_3). \end{aligned} \tag{3.19}$$

The expression in the first bracket is equal to $\dfrac{1}{C_1(s_1+s_2+s_3)}$. Then (3.19) can be rewritten as

$$\begin{aligned} -C_3(s_1,s_2,s_3)\frac{6}{C_1(s_1+s_2+s_3)} = {} & \left[(s_1+s_2+s_3)\tau_f + \frac{1}{\beta_0}\right]\cdot Z_b(s_1+s_2+s_3)\left[C_1(s_1)C_1(s_2)C_1(s_3)\frac{I_Q}{V_t^3}\right. \\ & \left. + \frac{2I_Q}{V_t^2}C_1(s_1)C_2(s_2,s_3) + \frac{2I_Q}{V_t^2}C_1(s_2)C_2(s_1,s_3) + \frac{2I_Q}{V_t^2}C_1(s_3)C_2(s_1,s_2)\right] \end{aligned} \tag{3.20}$$

$$\begin{aligned} C_3(s_1,s_2,s_3) = {} & -\frac{A_1(s_1+s_2+s_3)I_Q}{6V_t^3}\left[(A_1(s_1)A_1(s_2)A_1(s_3) + 6V_t\overline{A_1A_2})\right]\cdot\left[(s_1+s_2+s_3)\tau_f\right. \\ & \left. + \frac{1}{\beta_0}\right]Z_b(s_1+s_2+s_3), \end{aligned} \tag{3.21}$$

where

$$A_1A_2 = \frac{A_1(s_1)A_2(s_2,s_3) + A_1(s_2)A_2(s_1,s_3) + A_1(s_3)A_2(s_1,s_2)}{3}. \tag{3.22}$$

$A_3(s_1, s_2, s_3)$ can be obtained as

$$A_3(s_1, s_2, s_3) = \frac{i_c}{N^3}$$
$$= g_m\, C_3(s_1, s_2, s_3) + \frac{I_C}{6V_t^3} C_1(s_1)C_1(s_2)C_1(s_3) + \frac{I_C}{V_t^2}\overline{C_1 C_2}, \tag{3.23}$$

where $N^3 = 6e^{j(\omega_1+\omega_2+\omega_3)t}$[11].

We can use these transfer functions to determine which part of nonlineartity is more important, or to analyze the output of nonlinear systems.

In the next chapter, we will compare two Volterra series approach using Matlab and simulator to verify the way that we used to analyze distortion is correct. Based on these transfer functions, third-order intermodulation distortion cancellation analysis can be achieved.

Chapter 4

Third-order intermodulation distortion cancellation

4.1 Intercept point

As we discussed in chapter 1, second-order intercept point ($IP2$) and third-order intercept point ($IP3$) have significant influence on nonlinear device. In this section, we will use Volterra series method implemented in Matlab and Harmonic balance implemented in Advanced Design System (ADS) to analyze the output of nonlinear system. Volterra series include direct derivation method and nonlinear current source method.

4.1.1 IP2

The second-order intercept point which is known as $IP2$ including $IIP2$ (input intercept point) and $OIP2$ (output intercept point), is a second-order distortion generated by nonlinear systems and devices.

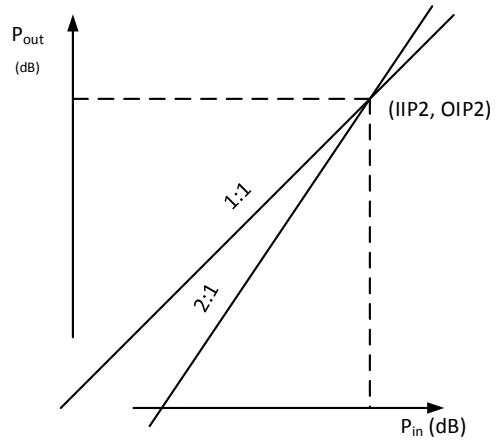

Figure 4.1: The definition of the second-order intercept point

At low power level, the fundamental output power has a one-to-one ratio with the input power, while the second-order output power is two-to-one to the input power. But if the input power is high enough to reach saturation, the out put power will flatten out at some point in both first-order and second-order case.

The $IIP2$ is the input power at the intercept point where the first-order and second-order line intersect with each other. Similarly to $OIP2$, which is the output power at that intercept point. Fig. 4.1 shows the magnitude response of a two-tone mixer as a function of input power.

The actual value of $IIP2$ and $OIP2$ of a device can be measured, being related by the small signal gain of the device or system. Power gain can be obtained as[4]

$$Gain = 10\log_{10}(\frac{P_{out}}{P_{in}}) = 10\log_{10}(\frac{V_{out}^2}{2r_L}\frac{8r_S}{V_S^2}) = 10\log_{10}(\frac{V_{out}^2}{r_L}4r_S), \tag{4.1}$$

where the unit of power gain is dB, r_S and r_L are source resistance and load resistance respectively. Output power $P_{out} = \frac{A_{IP2}^2}{2r_L}$, input power $P_{in} = \frac{A_{IP2}^2}{8r_S}$. A_{IP2} is the input voltage amplitude at the $IP2$ point which is equal to

$$A_{IP2} = \sqrt{\frac{H_{1_2}(s)\ g_{L2nd}}{H_{2_2}(s)\ g_{L1st}}}, \tag{4.2}$$

where g_{L1st} and g_{L2nd} are the load impedance of the first-order and second-order respectively. H_{1_2} is the first-order transfer function at node 2 and H_{2_2} is the second-order transfer function at node 2.

IIP2, the input power at the second order intermodulation intercept can be obtained as

$$IIP2 = 10\log_{10}(\frac{A_{IP2}^2}{2\cdot 2\cdot (r_L + r_S)}) + 30 = 10\log_{10}(\frac{A_{IP2}^2}{8r_S}) + 30, \tag{4.3}$$

where the unit is dBm. In (4.3) the first "2" is the average value, the second "2" is the available value. We set $r_S = r_L$ in (4.3), thus $IIP2 = \frac{A_{IP2}^2}{8r_S}$. $OIP2$ can be obtained as:

$$OIP2 = IIP2 + Gain. \tag{4.4}$$

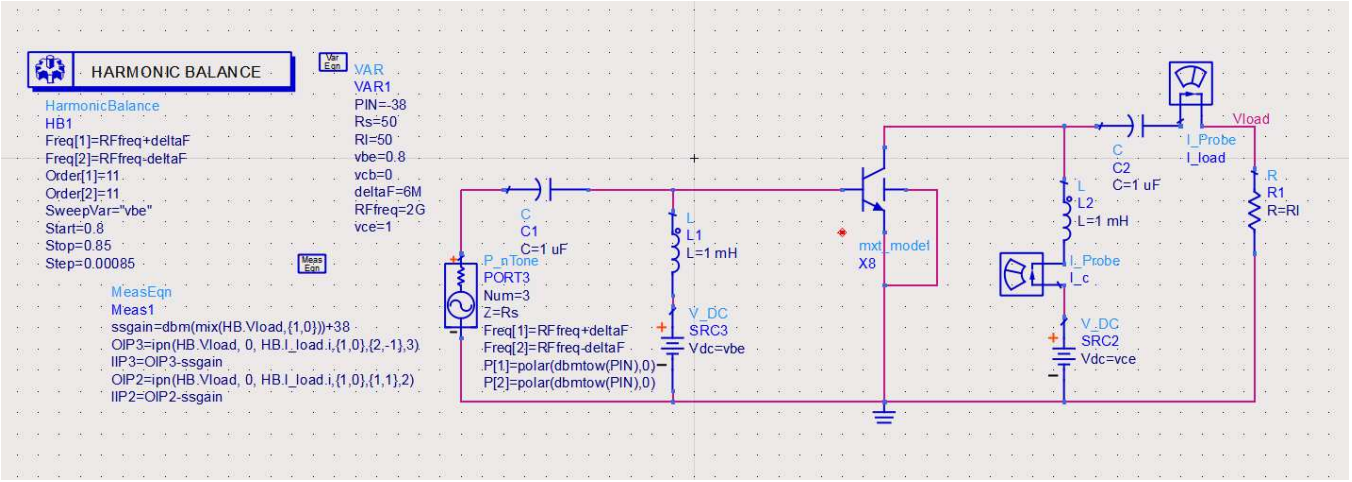

Figure 4.2: ADS simulation circuit and parameters setting.

parameter name	value
I_S	$139.146aA$
B_f	480
R_b	8.6Ω
C_t	$37.18fF$
τ_f	$9.36ps$

Table 4.1: Parameters used in Volterra series method.

Fig. 4.2 shows the ADS simulation circuit, parameters of BJT model and simulation equations. In this simulation, we set $Pin = -38dBm$, $R_S = R_L = 50\Omega$, $f_1 = 2GHz$ and $f_2 = 2GHz \pm 6MHz$. $V_{ce} = 1V$ and V_{be} is swept from $0.75V$ to $0.84V$. $X8$ in this circuit is a mextram model, the device size is $0.12{\times}18\mu m^2$. τ_f and C_t can be approximated from

(4.5) and the plot is shown in Fig. 4.3

$$C_{be} = C_t + g_m \tau_f, \tag{4.5}$$

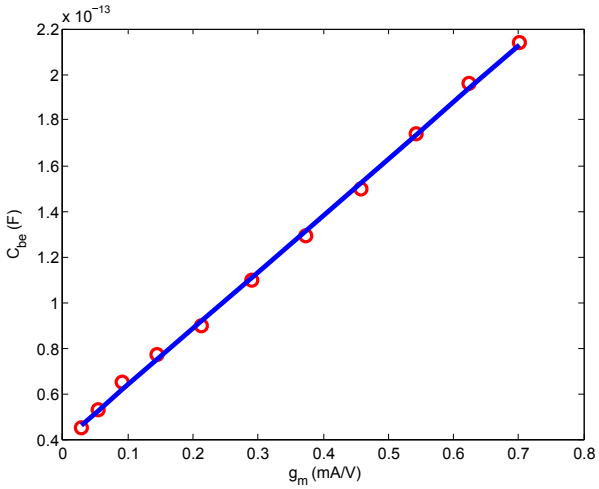

Figure 4.3: Using $C_{be} - g_m$ plot to obtain τ_f and C_t.

C_{be} is the base-emitter diffusion capacitance. We obtain C_{be} from operation point information of mextram Verilog-A. Put the value of $C_{be} - g_m$ in Matlab and use polyfit function to obtain its slope and zero bias point. The value of the slope of $C_{be} - g_m$ is τ_f and zero bias of $C_{be} - g_m$ is C_t.

Fig. 4.4 shows the plot of cutoff frequency versus I_C. I_C under peak f_T is around 20mA, thus we just considering the condition of I_C smaller than 20mA.

The plot of power gain is shown in Fig. 4.5. It is obvious that based on Volterra series, plots of nonlinear current source approach and direct derivation are exactly the same. But when comparing the result with Harmonic balance in ADS simulation, the plots can match only at low current. Because the model of Volterra series approach just considering the influence of I_C and I_B. g_m simplifies to $\dfrac{I_C}{V_t}$, where $V_t = 25.8mV$ is thermal voltage. Thus it is not accurate any more when I_C becomes higher.

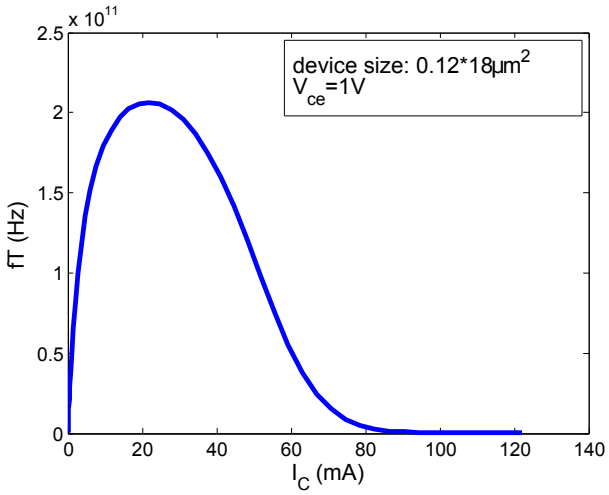

Figure 4.4: f_T versus I_C plot.

Fig. 4.6 and Fig. 4.7 show the plot of $IIP2$ and $OIP2$ respectively. Similarly we simplify I_C, I_B and g_m in the model of Volterra series method, thus Volterra series method cannot match well with Harmonic balance implemented in ADS simulation .

4.1.2 IP3

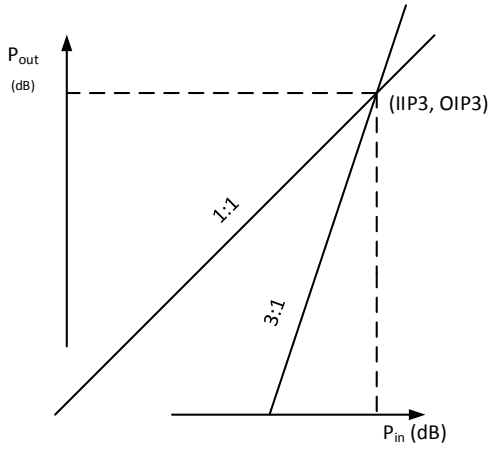

Figure 4.8: The definition of the third-order intercept point

The third-order intercept point ($IP3$), also based on the idea that the nonlinearity of a device can be modeled using a low-order polynomial and derived by Taylor series expression.

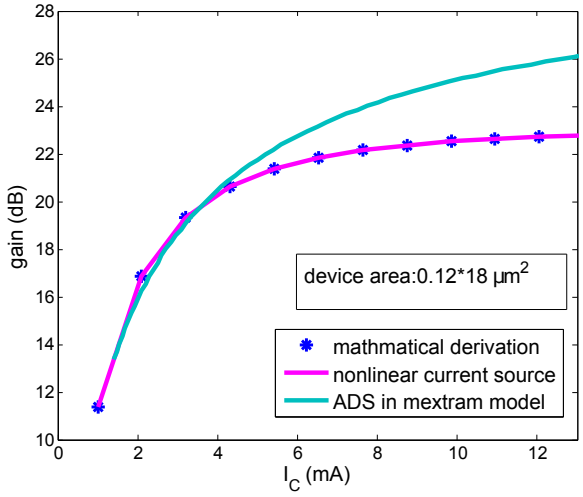

Figure 4.5: Comparing power gain versus I_c between Volterra series method in Matlab and Harmonic balance in ADS.

Similarly, Fig. 4.8 shows the $IP3$ point at the intercept of ratio one-to-one line and three-to-one line. Then the x-axis of $IP3$ is $IIP3$ and y-axis is $OIP3$.

In $IIP3$ case[4],

$$\frac{3}{4}A^3 H_{3_2}(s_1, s_1, -s_2) = AH_{1_2}(s_1), \tag{4.6}$$

A_{IP3} can be obtained as

$$A_{IP3} = \sqrt{\frac{4}{3}\frac{H_{1_2}\ g_{L3rd}}{H_{3_2}\ g_{L1st}}}, \tag{4.7}$$

where g_{L3rd} is the third-order load impedance and H_{3_2} is the third-order transfer function of voltage at node 2.

Using (4.7) and (4.1), we can obtain $IIP3$ and $OIP3$ as

$$IIP3 = 10log_{10}(\frac{A_{IP3}^2}{8r_S}) + 30, \tag{4.8}$$

$$OIP3 = IIP3 + Gain, \tag{4.9}$$

where the unit of $IIP3$ and $OIP3$ in these two equations are in dBm.

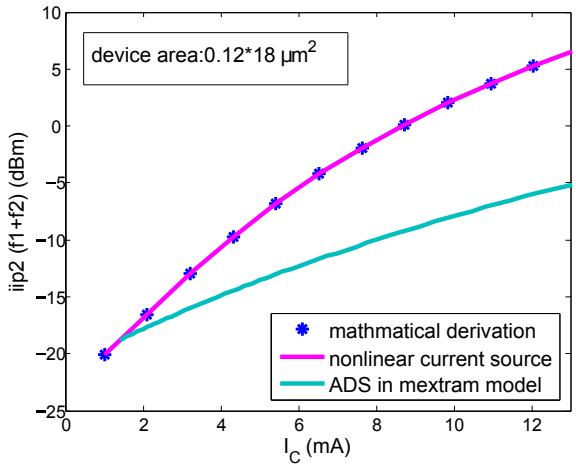

Figure 4.6: Comparing $IIP2$ versus I_C between Volterra series method in Matlab and Harmonic balance in ADS at $f_1 + f_2$.

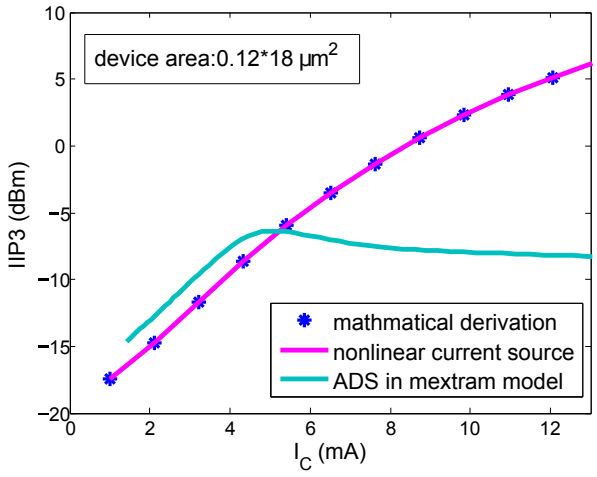

Figure 4.9: Comparing $IIP3$ versus I_C between Volterra series method in Matlab and Harmonic balance in ADS at $2f_1 - f_2$.

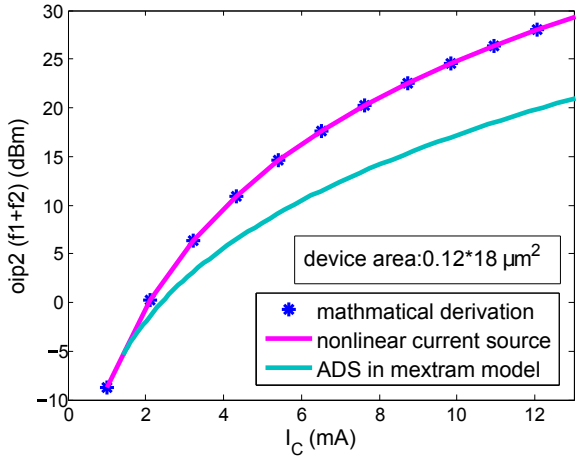

Figure 4.7: Comparing $OIP2$ versus I_C between Volterra series method in Matlab and Harmonic balance in ADS at $f_1 + f_2$.

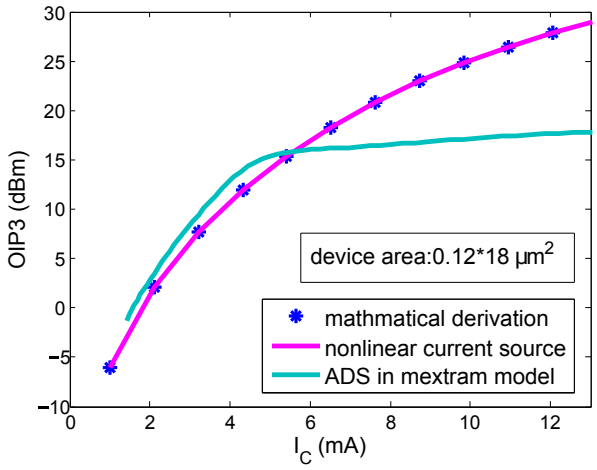

Figure 4.10: Comparing $OIP3$ versus I_C between Volterra series method in Matlab and Harmonic balance in ADS at $2f_1 - f_2$.

Fig. 4.9 and Fig. 4.10 show the plot of $IIP3$ and $OIP3$ *vs* I_C between calculation and ADS simulation respectively. $R_S = R_L = 50\Omega$, $f_1 = 2GHz$ and $f_2 = 2GHz + 6MHz$. We can find that two curves based on Volterra series can match well. Because of simplification of Volterra series' model, Harmonic balance implemented in ADS cannot match well with Volterra series method at higher I_C.

4.2 Third-order intermodulation distortion cancellation

One of the advantages to analyze $IP2$ and $IP3$ is that we can reduce this nonlinear distortion to meet the increasing demands on linearity made by todays mobile communication standards. Generally bipolar devices are strongly nonlinear due to their exponential terms. In order to achieve linearity, dc power consumption will be increasing since linearity is traded off against collector current. Thus it is important to find a way to facilitate linearity at low dc collector current. This $IM3$ cancellation approach is based on [5].

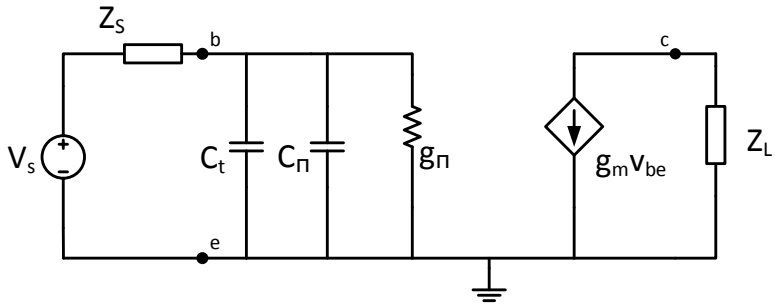

Figure 4.11: Simplified model of a bipolar transistor in common-emitter configuration.

First, we calculate the full expression of $IM3$ and decide the requirements for low-frequency cancellation. Fig. 4.11 shows the large-signal model of a bipolar transistor. We assume the base-emitter depletion capacitance C_t approximate to a constant value when V_{be} change at a very small range. The influence of C_{bc} can be ignored at low frequency because it only has a big influence at high frequency. Thus either using the approach of nonlinear current source or directly nonlinear circuit calculation, the first-order transfer function can be obtained as[5][6]

$$H_{1_2}(s) = \frac{-g_m Z_L(s)}{N(s)}, \tag{4.10}$$

where $N(s) = [1 + Z_S(s)g_\pi + s(C_t + C_\pi)Z_S(s)]$. $Z_S(s)$ is the source impedance which does not include base resistance. Because $R_B = 8.6\Omega$ can be neglected when comparing R_B with R_S. $Z_L(s)$ is the load impedance. We can find g_m, g_π and C_π in equation (1.21),(1.24) and (1.37).

Then the expression of the third-order Volterra kernel which is calculated at third-order intermodulation frequency $(2\omega_1 - \omega_2)$ is given by

$$H_{3_2}(s_1, s_1, -s_2) = \frac{-g_{m3} Z_L(2s_1 - s_2)\left[1 + (2s_1 - s_2)C_t Z_S(2s_1 - s_2)\right] T(s_1, s_1, -s_2)}{N(s_1)^2 N(-s_2) N(s_1 - s_2) N(2s_1) N(2s_1 - s_2)}, \quad (4.11)$$

where

$$\begin{aligned} T(s_1, s_1, -s_2) &= 1 - g_\pi Z_S(s_1 - s_2) - 2(g_\pi)^2 Z_S(2s_1) Z_S(s_1 - s_2) \\ &+ 2s_1(s_1 - s_2)\left[(C_t - 2C_\pi)(C_t + C_\pi)\right] Z_S(2s_1) Z_S(s_1 - s_2) \\ &+ (s_1 - s_2)\left[C_t - C_\pi(1 + 2g_\pi Z_S(2s_1))\right] Z_S(s_1 - s_2) \\ &+ 2s_1\left[C_t - g_\pi(C_t + 2C_\pi) Z_S(s_1 - s_2)\right] Z_S(2s_1). \end{aligned} \quad (4.12)$$

$IM3$ can be calculated as[5]

$$IM3 = \frac{3}{4} A^2 \left| \frac{H_{3_2}(s_1, s_1, -s_2)}{H_{1_2}(s_1)} \right| \quad (4.13)$$

where A is the amplitude of input voltage. Substituting (4.11) and (4.12) into (4.13), then we can rewrite $IM3$ as

$$IM3 = \frac{A^2}{8V_t^2} \left| \frac{\left[1 + (2s_1 - s_2)C_t Z_S(2s_1 - s_2)\right] T(s_1, s_1, -s_2)}{N(s_1) N(-s_2) N(s_1 - s_2) N(2s_1) N(2s_1 - s_2)} \right|. \quad (4.14)$$

The numerator in equation (4.14) relates two part, $(2s_1 - s_2)C_t Z_S(2s_1 - s_2)$ and $T(s_1, s_1, -s_2)$. In the first part, both C_t and Z_S work at the third-order intermodulation frequency $(2s_1 - s_2)$. Thus this term can be canceled when using an inductance as source impedance at this certain frequency. The second part $T(s_1, s_1, -s_2)$ is expanded into equation (4.12). It is clear that the linearity parameters and source impedance of this term only depend on the second-order intermodulation ($IM2$) frequency $(s_1 - s_2)$ and the second-order harmonic (H_2) frequency $(2s_1)$. In order to result in a cancellation of $IM3$, T should be zero at some value. At low

frequency, (4.12) can be rewritten as

$$T(s_1, s_1, -s_2) = 1 - g_\pi Z_S - 2g_\pi^2 Z_S^2. \tag{4.15}$$

When setting $T(s_1, s_1, -s_2)$ in (4.15) to zero, we can achieve the source impedance cancellation requirement at low-frequency

$$Z_S = \frac{r_\pi}{2} = \frac{\beta_f V_t}{2I_C}. \tag{4.16}$$

At high frequency, substituting $Z_S(s_1 - s_2) = Z_S(2s_1) = r_S$ into (4.12) and setting T equal to zero, $T(s_1, s_1, -s_2)$ in (4.15) can be rewritten as

$$\begin{aligned} 0 = {} & 2s_1(s_1 - s_2)\left[(C_t - 2C_\pi)(C_t + C_\pi)\right] r_S^2 \\ & +(s_1 - s_2)\left[C_t - C_\pi(1 + 2g_\pi r_S)\right] r_S \\ & +2s_1\left[C_t - g_\pi(C_t + 2C_\pi)r_S)\right] r_S, \end{aligned} \tag{4.17}$$

where $1 - g_\pi r_S - 2(g_\pi)^2 r_S^2$ can be ignored when comparing this part with terms relate to $s_1 - s_2$ or s_1. Similarly, the part of $2s_1$ and $s_1 - s_2$ can also be ignored when comparing them with terms relate to $2s_1(s_1 - s_2)$. Thus (4.17) can be rewritten as

$$0 = 2s_1(s_1 - s_2)\left[(C_t - 2C_\pi)(C_t + C_\pi)\right] r_S^2, \tag{4.18}$$

assuming s_1 and s_2 close with each other,

$$0 = 4s_1^2\left[(C_t - 2C_\pi)(C_t + C_\pi)\right] r_S^2, \tag{4.19}$$

$C_t = 2C_\pi$ and $C_t = -C_\pi$ are two solutions of (4.19). Because C_t cannot be a negative value, thus $IM3$ cancellation requirement at high frequency can be obtained as

$$C_t = 2C_\pi = 2\tau_f \frac{I_C}{V_t}. \tag{4.20}$$

Within these two requirements, a frequency independent cancellation of T can be achieved. We can observe that I_C at which cancellation occurs is fixed for a given C_t and τ_f. Thus (4.16) and (4.20) can be rewritten as a single requirement for Z_S at fixed I_C

$$Z_S(s_1 - s_2) = Z_S(2s_1) = r_S = \frac{\beta_f \tau_f}{C_t} \qquad (I_C = V_t \frac{C_t}{2\tau_f}). \tag{4.21}$$

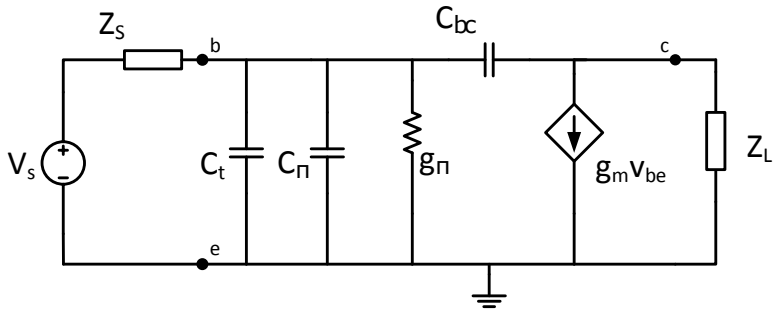

Figure 4.12: The circuit for high-frequency oip3 cancellation[5].

Fig. 4.12 shows the circuit used for high-frequency $IM3$ cancellation. Since capacitance C_{bc} added to our model, the expression of this circuit should be recalculated. Using nonlinear current source method in Fig. 4.12, the first-order transfer functions can be produced as[3][5]

$$Y(s) \cdot H_1(s) = IN_1,$$

$$\begin{bmatrix} Y_S(s) + g_\pi + s(C_t + C_\pi + C_{bc}) & -sC_{bc} \\ g_m - sC_{bc} & Y_L(s) + sC_{bc} \end{bmatrix} \begin{bmatrix} H_{1_1}(s) \\ H_{1_2}(s) \end{bmatrix} = \begin{bmatrix} Y_S(s) \\ 0 \end{bmatrix}, \tag{4.22}$$

where $Y(s)$ is the admittance matrix of the circuit, $H_1(s)$ is the first-order Volterra kernel transforms of the node voltage, $IN1$ is the vector excitation when source voltage (V_S) equal to 1 (V). Then by solving (4.22) using Cramer's rule, $H_{1_2}(s)$ can be given as

$$H_{1_2}(s) = \frac{-(g_m - sC_{bc})Z_L(s)}{1 + Z_S(s)g_\pi + sC_{bc}(1 + Z_L(s) + Z_L(s)(g_m - sC_{bc}))Z_S(s) + s(C_t + C_\pi)Z_S(s)}. \tag{4.23}$$

(4.10) can be obtained when setting C_{bc}=0 in (4.23).

Then the second-order Volterra kernel transforms of the node voltage can be solved in the same way

$$Y(s_1 + s_2) \cdot \begin{bmatrix} H_{2_1}(s_1, s_2) \\ H_{2_2}(s_1, s_2) \end{bmatrix} = \begin{bmatrix} -(inl2c\pi + inl2g\pi) \\ -inl2gm \end{bmatrix} = IN_2, \tag{4.24}$$

where $H_{2_1}(s_1, s_2)$ and $H_{2_2}(s_1, s_2)$ are the second-order ransfer functions. IN_2 is the matrix of the nonlinear current sources of order two which are parallel to their linear equivalents respectively in Fig. 4.12. The second-order nonlinear current sources are

$$\begin{aligned} inl2g\pi &= g\pi 2H_{1_1}(s_1)H_{1_1}(s_2) \\ inl2gm &= gm2H_{1_1}(s_1)H_{1_1}(s_2) \\ inl2c\pi &= (s_1 + s_2)c\pi 2H_{1_1}(s_1)H_{1_1}(s_2) \end{aligned} \tag{4.25}$$

Using Cramer's rule, $H_{2_1}(s_1 + s_2)$ and $H_{2_2}(s_1 + s_2)$ can be solved

$$\begin{aligned} H_{2_1}(s_1, s_2) &= \frac{-(Y_L(s) + (s_1 + s_2)C_{bc})(inl2g\pi + inl2c\pi) - (s_1 + s_2)C_{bc}inl2gm}{det(Y(s_1 + s_2))} \\ H_{2_2}(s_1, s_2) &= \frac{-(Y_S(s) + g\pi + (s_1 + s_2)(C_\pi + C_t + C_{b_c}))inl2gm}{det(Y(s_1 + s_2))} \\ &\quad + \frac{(gm - (s_1 + s_2)C_{bc})(inl2g\pi + inl2c\pi)}{det(Y(s_1 + s_2))}. \end{aligned} \tag{4.26}$$

The third-order transfer functions of node 2 can be obtained as

$$Y(s_1+s_2+s_3)\cdot\begin{bmatrix} H_{3_1}(s_1,s_2,s_3) \\ H_{3_2}(s_1,s_2,s_3) \end{bmatrix} = \begin{bmatrix} -(inl3c\pi + inl3g\pi) \\ -inl3gm \end{bmatrix} = IN_3, \tag{4.27}$$

where $H_{3_1}(s_1,s_2,s_3)$ and $H_{3_2}(s_1,s_2,s_3)$ are the third-order Volterra kernels. IN_3 is the nonlinear current sources of order three. These current sources are placed parallel with their equivalents in Fig. 4.12 which is given by

$$\begin{aligned} inl3g\pi &= g\pi 3H_{1_1}(s_1)H_{1_1}(s_2)H_{1_1}(s_3) + 2g\pi 2\overline{H_{1_1}H_{2_1}} \\ inl3gm &= gm2H_{1_1}(s_1)H_{1_1}(s_2)H_{1_1}(s_3) + 2gm2\overline{H_{1_1}H_{2_1}} \\ inl3c\pi &= (s_1+s_2+s_3)(c\pi 3H_{1_1}(s_1)H_{1_1}(s_2)H_{1_1}(s_3) + 2c\pi 2\overline{H_{1_1}H_{2_1}}), \end{aligned} \tag{4.28}$$

where

$$\overline{H_{1_1}H_{2_1}} = (\frac{H_{1_1}(s_1)H_{2_1}(s_2,s_3) + H_{1_1}(s_2)H_{2_1}(s_1,s_3)H_{1_1}(s_3)H_{2_1}(s_1,s_2))}{3})$$

.

By solving (4.27) with Cramer's rule and setting $s_1 = s_2 = jw1$, $s_3 = -jw2$, $C_{bc} = 0$, we can get equation (4.11).

Since for high-frequency capacitance C_{bc} would cause a voltage-current feedback to the base, this feedback influence the condition of $IM3$ cancellation at input. Thus the voltage drop across C_{bc} should be zero for second-order harmonic signals, which means the second-order voltage at the base ($V_{B,2}$) and collector ($V_{C,2}$) must be equal.

$$\begin{aligned} V_{C,2} &= (inl2gm - gmV_{B,2})Z_L(s_1+s_2) \\ V_{B,2} &= \frac{inl2g\pi + inl2c\pi}{gs(s_1+s_2) + g\pi + (s_1+s_2)(C_t+C_\pi)}, \end{aligned} \tag{4.29}$$

where g_s is the admittance of input impedance. Then by solving equation (4.29), we can obtain the value of $R_L(s_1+s_2)$ which is used for compensation the disturb of C_{bc}.

$$Z_L(s_1+s_2)=R_L=\frac{inl2c\pi+inl2g\pi}{inl2gm[gs+g\pi+(s_1+s_2)(C_t+C_\pi)]-gm(inl2g\pi+inl2c\pi)}, \quad (4.30)$$

substituting (4.25) into (4.30) we can obtain

$$\begin{aligned} Z_L(s_1+s_2)=R_L&=\frac{g_{\pi 2}H_{1_1}(s_1)H_{1_1}(s_2)+(s_1+s_2)C_{\pi 2}H_{1_1}(s_1)H_{1_1}(s_2)}{g_{m2}H_{1_1}(s_1)H_{1_1}(s_2)[g_S+g_\pi+(s_1+s_2)(C_t+C_\pi)-g_mg_{\pi 2}-g_m(s_1+s_2)C_{\pi 2}]}\\ &=\frac{g_{\pi 2}+(s_1+s_2)C_{\pi 2}}{g_{m2}[g_S+g_\pi+(s_1+s_2)(C_t+C_\pi)]-g_m[g_{\pi 2}+(s_1+s_2)C_{\pi 2}]}, \end{aligned} \quad (4.31)$$

where $g_\pi=\frac{g_m}{\beta_f}$, $g_{\pi 2}=\frac{g_m}{2\beta_f V_t}$, $C_\pi=\tau_f g_m$, $C_{\pi 2}=\tau_f\frac{g_m}{2V_t}$ and $g_{m2}=\frac{g_m}{2V_t}$. (4.31) can be rewritten as

$$\begin{aligned} Z_L(s_1+s_2)=R_L&=\frac{\frac{g_m}{2\beta_f V_t}+(s_1+s_2)\tau_f\frac{g_m}{2V_t}}{\frac{g_m}{2V_t}\left[g_S+\frac{g_m}{\beta_f}+(s_1+s_2)(C_t+\tau_f g_m)\right]-g_m\left[\frac{g_m}{2\beta_f V_t}+(s_1+s_2)\tau_f\frac{gm}{2V_t}\right]}\\ &=\frac{\frac{1}{\beta_f}+(s_1+s_2)\tau_f}{\left[g_S+\frac{g_m}{\beta_f}+g_m+(s_1+s_2)C_t\right]}, \end{aligned} \quad (4.32)$$

because $(s_1+s_2)\tau_f$ and $(s_1+s_2)C_t$ are much larger than other parts in (4.32), thus $\frac{1}{\beta_f}$, g_S, $\frac{g_m}{\beta_f}$ and g_m can be ignored. We can obtain R_L as

$$Z_L(s_1+s_2)=R_L=\frac{\tau_f}{C_t}=\frac{R_S}{\beta_f}. \quad (4.33)$$

$IM3$ cancellation depends on the proper ratio between fundamental voltage and second-order voltage, which means the linear feedback would not influence the cancellation when Z_S is fixed at a proper value.

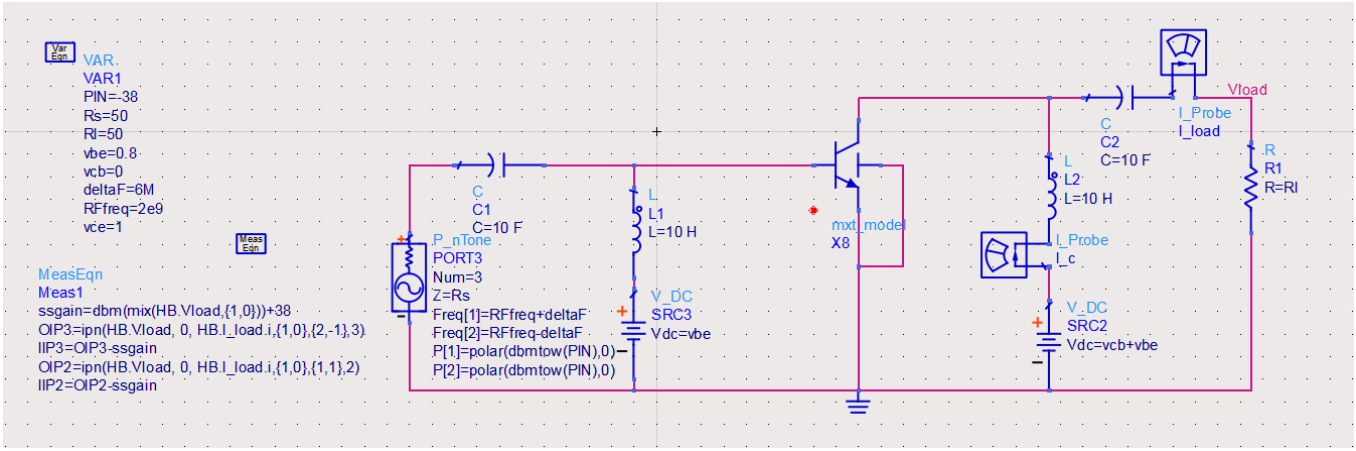

Figure 4.13: Equivalent circuit of a bipolar transistor in common-emitter configuration.

Fig. 4.13 shows the circuit of ADS simulation. $C1$, $C2$, $L1$ and $L2$ are capacitance and inductance of bias tee respectively. In order to meet requirement of (4.21) and (4.33), these capacitances and inductances should be very large which are set as $C1 = C2 = 10F$ and $L1 = L2 = 10H$. There has two DC voltage and one power source with N level frequency and power. $X8$ is a Mextram bipolar transistor, device size is $0.12{\times}18\mu m^2$.

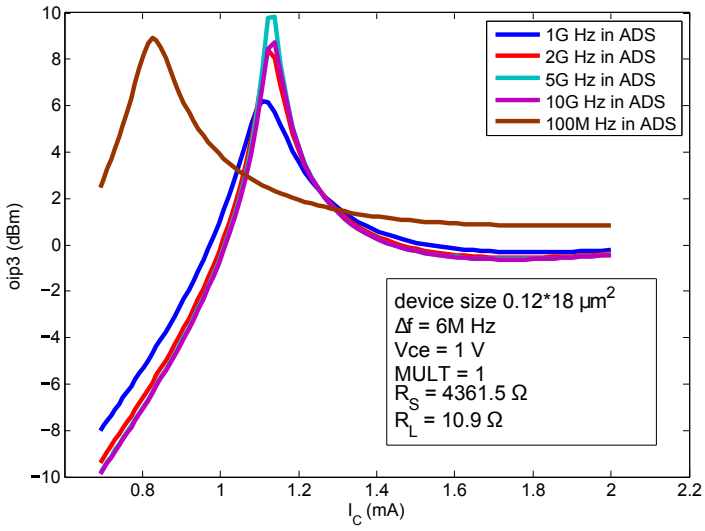

Figure 4.14: $IM3$ cancellation result in different fundamental frequency with MULT=1, where $R_S = 4361.5\Omega$ and $R_L = 10.9\Omega$

Fig. 4.14 shows the simulated $OIP3$ versus I_C at different fundamental frequency and a fixed $\Delta f = 6MHz$. We can get a approximate value of C_t and τ_f by using C_{be} versus gm curves which is similar to Fig. 4.3. When $MULT = 1$, $C_t = 52fF$ and $\tau_f = 0.63ps$. Thus $R_S = 4361.5\Omega$ can be calculated in (4.21), $R_L = 10.9\Omega$ can be calculated in (4.33). I_C at which cancellation occurs is equal to $V_t\dfrac{C_t}{2\tau_f}$, then $I_C = 1.1mA$ can be calculated. Using Harmonic balance in ADS simulation, we can find that I_C for peak $OIP3$ is around $1.18mA$. It means the calculated I_C is close to the simulated I_C. $\beta_f = 400$.

I_C at peak $OIP3$ under frequency of $1GHz$, $2GHz$, $5GHz$ and $10GHz$ are not far away from each other, which means the peak $OIP3$ is largely independent of frequency as expected from expression. But when comparing I_C for peak $OIP3$ at high frequency to the I_C at low frequency, there still exists some difference.

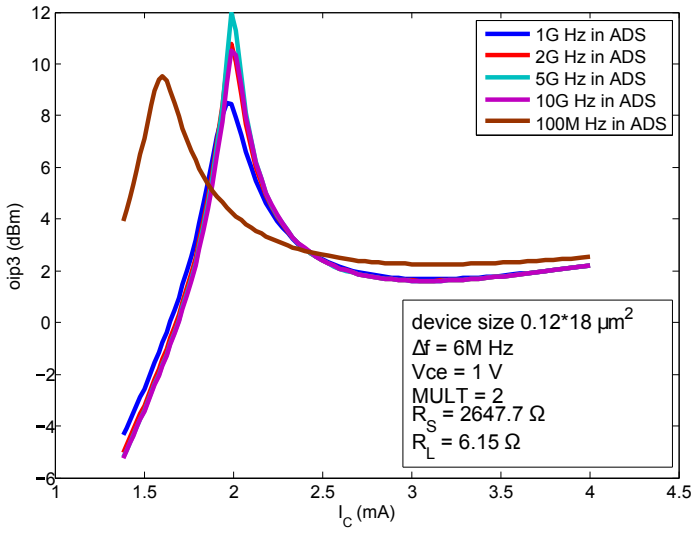

Figure 4.15: $IM3$ cancellation result in different fundamental frequency with MULT=2, where $R_S = 2647.7\Omega$ and $R_L = 6.16\Omega$

Fig. 4.15 shows $OIP3$ versus I_C when $MULT = 2$. Since the device size changed in a ideal way when changing $MULT$, C_t and τ_f should be recalculated. Similarly, C_t and τ_f can be obtained from $C_{be} - g_m$ plot, $C_t = 104.1fF$, $\tau_f = 0.641ps$. Within (4.21) and (4.33), $R_S = 2647.7\Omega$ and $R_L = 6.16\Omega$. $I_C = 2.1mA$ can be calculated from (4.21). The simulated I_C for peak $OIP3$ is around $2.05mA$, which is very close to the calculated I_C. $\beta_f = 430$.

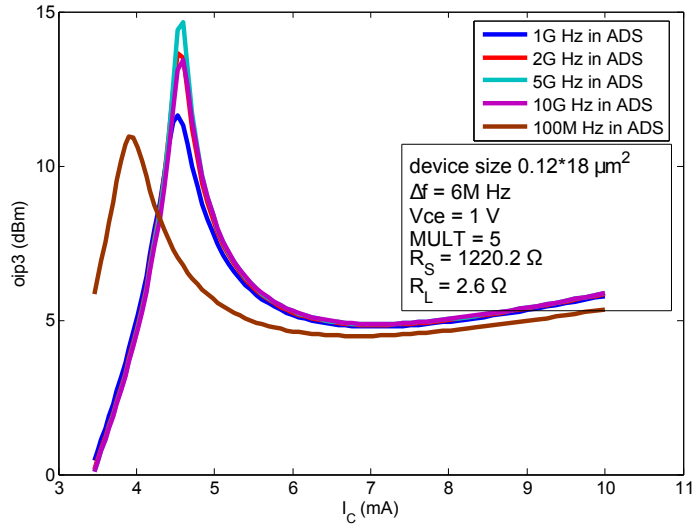

Figure 4.16: $IM3$ cancellation result in different fundamental frequency with MULT=5, where $R_S = 1220.2\Omega$ and $R_L = 2.6\Omega$

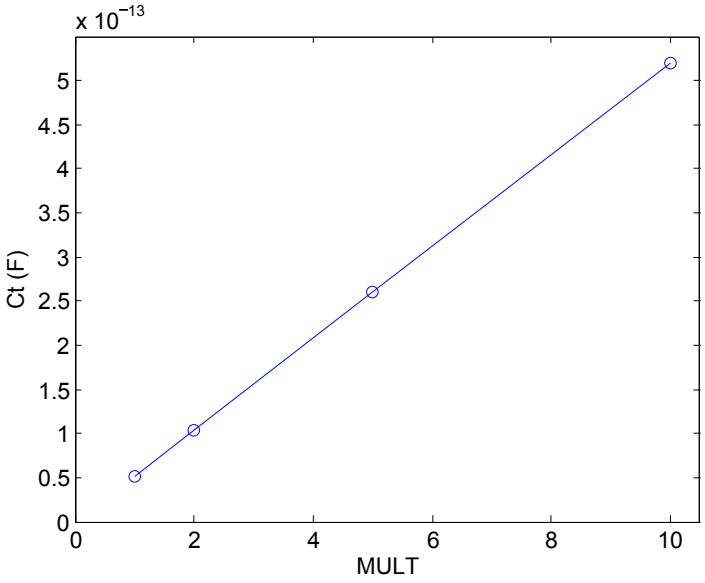

Figure 4.17: C_t changed with $MULT$.

Fig. 4.16 shows $OIP3$ versus I_C when $MULT = 5$. Similarly, C_t and τ_f are obtained from $C_{be} - g_m$ plot, $C_t = 260fF$, $\tau_f = 0.675ps$. $R_S = 1220.2\Omega$ and $R_L = 2.6\Omega$ are calculated

from (4.21) and (4.33) respectively. $I_C = 5mA$ can be calculated from (4.21). The simulated I_C in Fig. 4.16 is around $4.8mA$, which is close to the calculated I_C. $\beta_f = 470$.

We can find that I_C under peak $OIP3$ get closer when $MULT$ become larger. In this $MULT$ scale, C_t is proportional to $MULT$ which is shown in Fig. 4.17, current inversely proportion to $MULT$. τ_f doesn't change much.

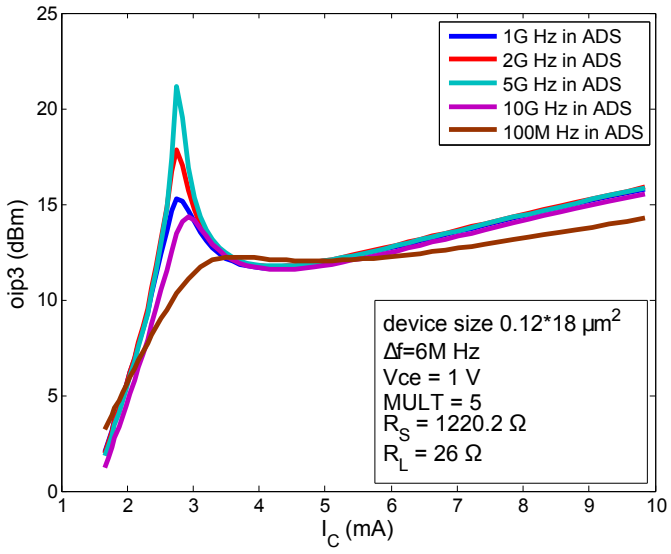

Figure 4.18: $IM3$ cancellation result in different fundamental frequency with MULT=5, where $R_S = 1220.2\Omega$ and $R_L = 26\Omega$.

Fig. 4.18 shows $OIP3$ versus I_C when $MULT = 5$. All parameters are the same as Fig. 4.16 except R_L. We change R_L 10 times larger than R_L we used in Fig. 4.16 and find that $I_C = 2.9mA$ under peak $OIP3$ in ADS simulation is no longer close to the calculated $I_C = 5mA$. $IM3$ cancellation even not occurs at $100MHz$. Thus R_L should be calculated from (4.33) and $IM3$ cancellation is not independent of R_L.

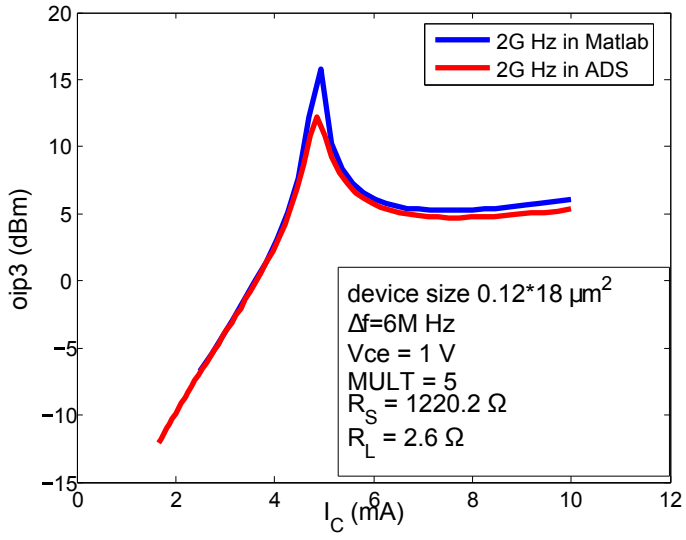

Figure 4.19: Comparing $IM3$ cancellation at fundamental frequency of 2GHz, MULT=5, $R_S = 1220.2\Omega$ and $R_L = 2.6\Omega$.

Fig. 4.19 shows the plot of $IM3$ cancellation in Matlab and ADS simulation. The Volterra series model use in Matlab does not include C_{bc} influence but Harmonic balance implemented in ADS simulation includes it. We can find that the value of I_C under peak $OIP3$ of Volterra series is around $4.8mA$, I_C for peak $OIP3$ of Harmonic balance is also around $4.8mA$, and the calculated I_C for peak $OIP3$ from (4.21) is $5mA$. They are very close at $2GHz$.

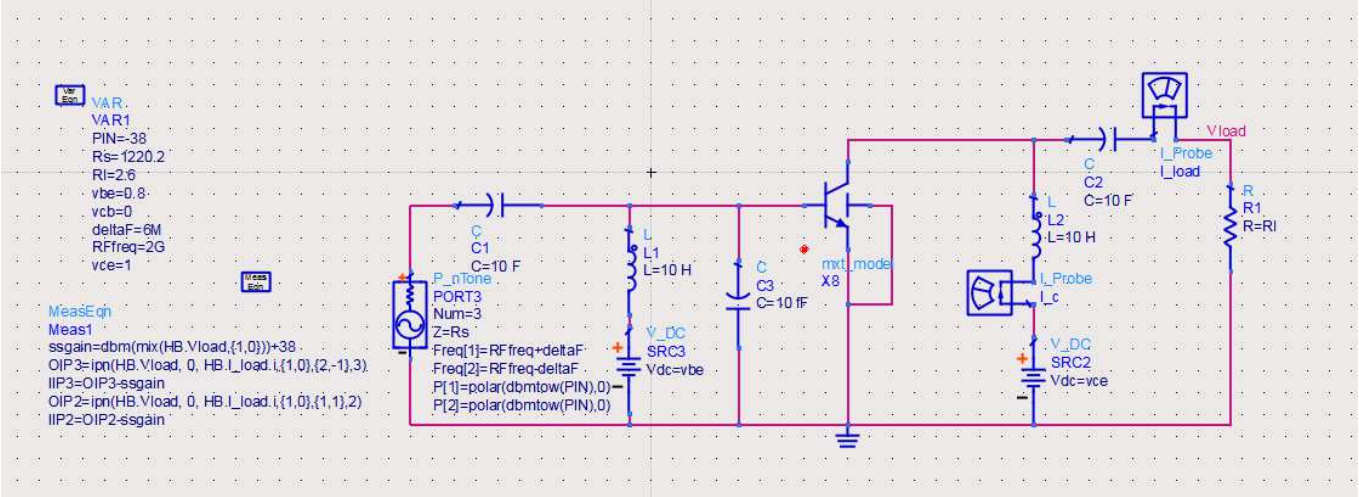

Figure 4.20: The circuit used for $IM3$ cancellation which placed a capacitance C_3 parallel with base-emitter junction.

Scaling $MULT$ is one way to change the position of I_C under peak $OIP3$, another way give more freedom choice of I_C is to place a capacitance C_3 parallel to base-emitter junction which is shown in Fig. 4.20.

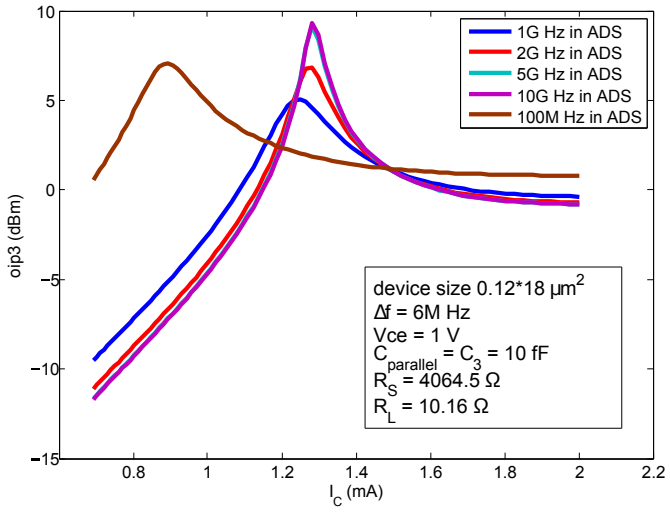

Figure 4.21: $IM3$ cancellation result in different fundamental frequency with $C_{parallel} = 10fF$, where $R_S = 4064.5\Omega$ and $R_L = 10.16\Omega$

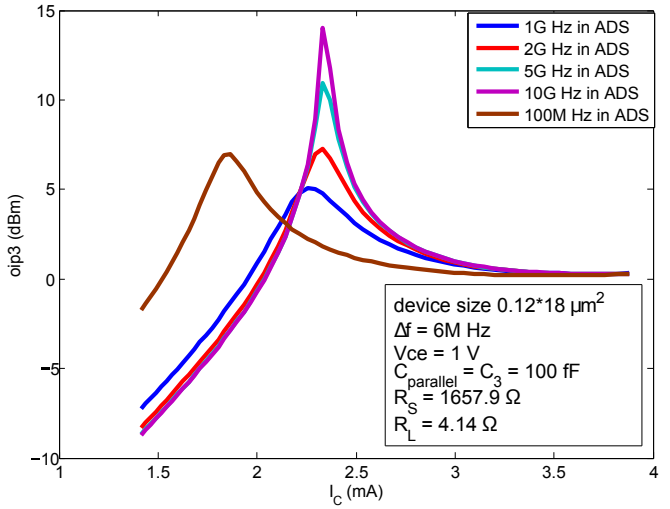

Figure 4.22: $IM3$ cancellation result in different fundamental frequency with $C_{parallel} = 100fF$, where $R_S = 1657.9\Omega$ and $R_L = 4.14\Omega$

Fig. 4.21 shows the $IM3$ cancellation when placed a capacitance with $C_3 = 10fF$ to base-emitter junction. Similarly, $C_t = 52fF$, $\tau_f = 0.63ps$ can be obtained from $C_{be} - g_m$

plot. $R_S = 4064\Omega$ and $R_L = 10.16\Omega$ can be calculated using (4.21) and (4.33) respectively. $I_C = 1.3mA$ is calculated from (4.21). The simulated I_C in Fig. 4.21 is around $1.3mA$ at four higher frequency, which is close to the calculated I_C. $\beta_f = 400$.

Fig. 4.22 shows the $IM3$ cancellation when placed a capacitance $C_3 = 100fF$ parallel to base-emitter junction. C_t, τ_f and β_f are the same as above. $R_S = 1657.9\Omega$ and $R_L = 4.14\Omega$ can be calculated in (4.21) and (4.33). $I_C = 3.1mA$ is calculated from (4.21). The simulated I_C in Fig. 4.22 is around $2.45mA$ at four higher frequency.

Chapter 5

Conclusion

In this thesis, nonlinearities of SiGe HBT are discussed and calculated by using volterra series and direct derivation approach.

Chapter 1 introduces the importance of nonlinearities in electric ciruit. Because these nonliearities will create distortion in the signals that we are interested in amplifying and transmitting. There are two types of nonlinearties, harmonic and intermodulation which is based on different frequency we used. Thorough understanding of HBT physical nonlinearties is very necessary.

Chapter 2 illustrates the nonlinear current source approach of obtaining Volterra series. Volterra series is a general mathematical approach for solving nonlinear systems, and nonlinear current source method is a way that put nonlinear current source parallel with each nonlinear element. The basic matrix is $Y(s) \cdot H(s) = I(s)$. Solving transfer function $H(s)$ at first-order, second-order and higher order we can obtain the output of nonlinear system.

Chapter 3 illustrate direct derivation approach.

In chapter 4, two approaches are compared with each other in Matlab. We also do Harmonic balance simulation in ADS with Mextram model and compare results with Volterra series approach. The plots of nonlinear current source and direct derivation approach are exactly the same, and they are very close to the simulation result. Based on the theory of Volterra series $IM3$ cancellation is discussed in this chapter. $Z_S = \frac{\beta_f \tau_f}{C_t}$ and $Z_L = \frac{\tau f}{C_t} = \frac{R_S}{\beta f}$ are the requirements of $IM3$ cancellation at both low frequency and high frequency. I_C under peak $OIP3$ in ADS simulation is very close to the calculated I_C at $1GHz$, $2GHz$, $5GHz$ and $10GHz$ with the same Δf, but not good at low frequency $100MHz$. $IM3$ cancellation is strongly dependent of R_L and R_S. C_{bc} will has a big influence of $IM3$ cancellation at

high frequency. From the simulation results, I_C under peak $OIP3$ get closer when $MULT$ become larger.

Bibliography

[1] J. D. Cressler, and G. Niu, Silicon-Germanium heterojunction bipolar transistors. Boston, MA: Artech House, 2002.

[2] K. Fong, and R. Meyer, "High-frequency nonlinear analysis of common-emitter and differential-pair transconductace stages," *IEEE Journal of SolidState Circuits.*, vol. 33, pp. 548–549, Apr. 1998.

[3] P. Wambacq, and W. Sansen, Distortion analysis of analog integrated circuits. Norwell, MA: Kluwer, 1998.

[4] S. A. Maas, Nonlinear microwave and RF circuit,2nd edition. Boston, MA: Artech House, 2003.

[5] M. van der Heijden, H. de Graaff, and L. de Vreede, "A novel frequency independent third-order intermodulation distortion cancellation technique for BJT amplifiers," *Bipolar/BiCMOS Circuits and Technology Meeting*, pp. 163–166, 2001.

[6] J. Reynolds, "Nonlinear distortion and their cancellation in transistors," *IEEE Trans. Electron Devices.*, vol. 12, pp. 595–599, 1965.

[7] G. Niu, Q. Liang, J. D. Cressler, C. S. Webster, and D. L. Harame, "RF linearity characteristics of SiGe HBTs," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 1558–1565, Sept, 2001.

[8] Q. Liang, Systematic analysis and optimization of broadband noise and linearity in silicon germanium heterojunction bipolar transistor. *Georgia Institute of Technology.*, 2004.

[9] B. M. Wilamowski, and G. Niu, Fundamentals of industrial electroncs (Electrical Engineering Handbook) 2nd ed. Edition. Taylor and francis group,LLC, 2011

[10] S. A. Maas, B. L. Nelson, and D. L. Tait, "Intermodulation in heterojunction bipolar transistors ," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 442–448, Mar, 1992.

[11] K. Fong, Design and optimization techniques for monolithic RF downconversion mixers. *University of California, Berkeley.*, 1997