UNIVERSITÀ DI PISA

Facoltà di Ingegneria

Corso di Laurea Specialistica in Ingegneria Elettronica

Anno accademico 2003/2004

TESI DI LAUREA SPECIALISTICA

Metodologia per la caratterizzazione di amplificatori a basso rumore per UMTS

Relatori:

Prof. Bruno Neri

Ing. Fernando De Bernardinis

Prof. Rinaldo Castello

Candidato:

Simone Gambini

Contents

A mio Padre,
Diego Gambini

Ringraziamenti Senza l'aiuto, i consigli e il supporto dell'Ing.Federico Vincis, questo lavoro non sarebbe arrivato neanche a metá.
L'Ing.Vincis non é solo incapace di fare male il proprio lavoro , ma si adopera anche per migliorare quello di chiunque abbia bisogno di lui in laboratorio con un'assenza di egoismo che non mi era mai capitato di osservare prima. É stato un onore lavorare con te Fede!
Un ringraziamento anche ai Prof.Rinaldo Castello e Francesco Svelto dell'Universitá degli studi di Pavia, per la cortesia dimostrata nel fornire il design del ricevitore, e nelle visite effettuate presso il dipartimento di Elettronica dell'Universitá di Pavia stessa per discutere il progresso del lavoro.
Grazie anche a mia madre e a mio fratello Francesco(a.k.a. Jann). Non ci sono parole per spiegare quello che mi avete dato. Grazie a Fabione, a mio nonno Andrea, alle mie nonne Edra, Enrica e Clara, a tutti i miei zii e le mie zie, compresi Zio Fosco e Zia Gabriella e Zio Mauro.
Grazie ai miei amici di Via S.Martino Stefano, Luca, Marco e Gabriele a Marta e Rossella, per le molte ottime cene e la compagnia ancora migliore offerte senza parsimonia. Grazie anche a Lorenzo, a Federico Maria Franzé e a tutti i miei amici di Piombino ovvero il Cetto, il Guero, il Buddha, il Sire, Iggio, i Rain Day, ovvero Ballack, lo Zanna, il Ross, il Gegia; a Gregory, Fede, Sara e la Gianfa, Ezio e il Villa.
Grazie a Sanerio, ZonZibi, Steiner, al Ghignoletto, a Tiziano e al Mao per degnarmi ogni tanto di scendere da Bologna a considerarmi.
Infine, Grazie al Prof. Bruno Neri a all'Ing. Fernando De Bernardinis, che hanno avuto con me piú pazienza di quanto mi sarei aspettato in piú di una occasione.

Grazie

Chapter 1

Introduction

This thesis deals with the aged and traditionally prickly subject of Analog Design automation. As many other works, it is an attempt to capture and bridle into sets of equations and mathematical formalism that field of electronic design that, more than any other, has allowed whoever is part of it to define himself as "a lonely cowboy".
Despite many people think attempts in this sense are either intellectual crimes, or efforts with less probability of success than the fight of Lacoonte against the sea-snakes, this work was carried on with care, and interesting, although certainly non-definitive results were found. I hope readers will then temporarily forget their romantic idea of the figure of Analog Designer, and listen to the way this hopeless battle was fought.

1.1 A bit of Background

In recent years, the continuous scaling of digital CMOS technology has allowed an increasingly high number of functionalities to be implemented on a single chip; the increase in silicon potential has by far outpaced the corresponding evolution both of the designers capability to handle such an increased complexity and of the developed design aids; as a result, an increasingly wide "productivity gap" has appeared in the electronic industry. This is especially true for analog design, and most for what concerns system level design. Structured design patterns, which are the keystone to develop design automation, are all but established in the analog arena, so that even for a simple task such as sizing a two stage op-amp, unnumbered different choices and approximations may drive the design process toward different solutions. The approaches followed may be roughly summarized in equations-driven approaches, and simulation-driven approaches. The former approach is well represented by the works by M.Del MarHershenson([13] [14]), based on convex optimization theory, where a particular formulation of a 1μ, two stage op-amp circuit equations is introduced, that allows using efficient inner point algorithms to find the global optimum for any set of specifications. Although extremely interesting and powerful, this approach, as well as all of the known equation-based design automation patterns, has a critical point to cope with: the accuracy of the chosen device model, which is further constrained to generate equations posynomially representable. Developing models of this kind may be extremely time consuming, especially for deep submicron technologies and when small signal block nonlinear performance is a concern. On the other hand, simulation based approaches use the power of state of art transistor compact models to measure circuit performances. As a drawback, the optimization has to be performed on a generally nonconvex domain by means of expensive stochastic global algorithms(simulated annealing), so that the computational time increases exponentially with the number of considered degrees of freedom. Although

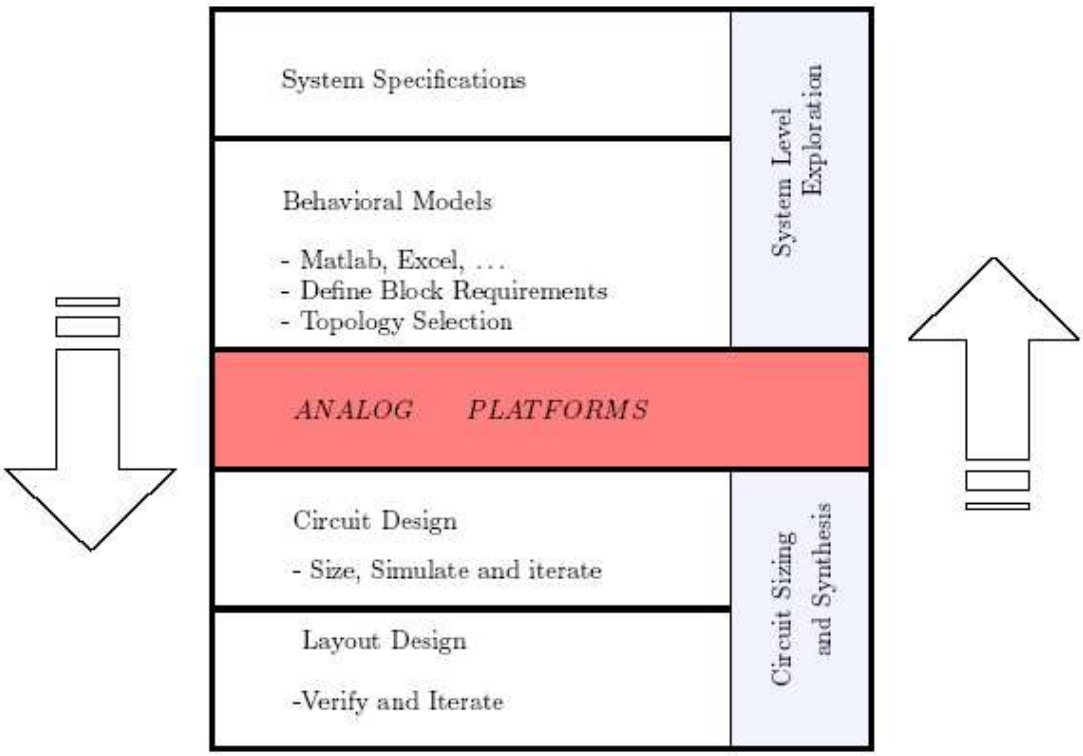

Figure 1.1: Analog Platform based design flow

useful to rapidly achieve high performance synthesis of small circuits, these methods fail both in treating large designs and in solving the theoretical problem of determining whether a given topology can or cannot satisfy a given set of specifications.
The above described scenario also has a detrimental influence on an higher level task such as architecture selection of a radio-receiver. The impossibility to claim rigorously claim feasible a set of system level performances for a given topology makes propagations of constraints to the different building blocks of any architecture based on previous design experience rather than on quantitative arguments. Experienced technical staff is therefore required in order to avoid system-circuit level design iterations that may increase the product time to market.
Even from the brief discussion above, it clearly appears that the key point in analog electronic system design is establishing a solid link between the architectural and the circuital level of abstraction, allowing high level design space exploration to be performed while enforcing blocks feasibility by construction. The platform paradigm is an attempt in this direction.
We may think an Analog Platform as an interface component between the system design world and the circuit design world. The system side of the analog platform is made up of a set of behavioral models. The circuit size is made up of class of ciruit instances described through an Analog Constraint Graph(ACG; see below).Simulation is used to relate this two sides, i.e. to extract from circuit instances satisfying the ACG, the parameters quantitatively describing the behavioral models(we'll call te set of

these parameters output space or O-space). Encapsulating a circuit in an analog platform means using the platform infrastructure to avoid giving unsatisfiable performance specifications to such a circuit, in that any attempt to specify circuit performances is now filtered by the platfrom, as clarified by figure (1.1). Once provided of sets of suitable behavioral models, both the above described equation based and the simulation based design automation approaches may be fit into this description. For a simulation-driven design automation engine, behavioral models are related to circuit sizings through equations. This comes at the expense of generality and accuracy.For simulation driven sutomation engines, mapping of sizings into higher-level performances happens through direct simulation. The main drawback , besides the scarce intellectual satisfaction deriving from nonglobality of the solution, is the exponential rise in computational time with increasing circuit dimension as, basing on the assumption of inaccuracy of equation-based approaches, only bounding boxes(i.e. $X \in [a, b]$) constraints are imposed on design variables.
As mentioned above, for the platform approach, circuit performances are still measured through simulation; on the other hand, circuit sizings may only be such that a set equations referred to in the following as Analog Constraint Graph(ACG) is satisfied. The set of equations "shapes" input space, consistently reducing its size. Another innovation of the platform paradigm comes from output space interpolation. The general idea is to feed a neural network(in [56] a statistical classifier) with the measured performances and to stop the sampling process when the machine learning process ends. This allows a further decrease in the number of simulations required to explore the performance space. Now , we summarily describe platform based design flow, which is graphically summarized in (1.1). Supposing to be designing a system such as an ADC or an integrated radio exploiting analog platforms, system level exploration and optimization may be performed using platform specific behavioral models. Thanks to the use of platform paradigm , these specifications are reachable BY CONSTRUCTION. At this point, the inverse of the performance evaluation function is used to derive optimal sizing of the building blocks themselves.
To transform the preceding discussion into an operating methodology, care must be put in two points:

1. Definition of blocks output space and behavioral models
2. Definition of the Analog Constraint Graph of each block, and of the necessary device model

1.2 UMTS receivers and goal of the work

The Universal Mobile standard for Telecommunication(UMTS) has been recently introduced into the consumer electronics market as an enabling technology for high data rate cellular communication. It is a Frequency-Duplex, code division multiple access(CDMA) based, direct sequence spread spectrum standard with operating frequencies in the $2GHz$ range. The details about chosen modulation format, despreading gain, and data frame structure are reported in [11] and won't be reproduced here. For the purpose of this work, operating bandwidths, signal multiplexing schemes and intermodulation tests are a sufficient background. The UMTS uplink bandwidth(Base station receives, mobile unit transmits) lies between 1.92 and $1.98GHz$, while the downlink is collocated between 2.11 and $2.17GHz$. Signals are partitioned in twelve nearby slots in the assigned 60MHz bandwidth, and further code-division multiplexed. Of this $5MHz$ overall channel bandwidths, only $3.84MHz$ have informative content, the rest being guard-band. Being UMTS a frequency duplex standard, receiver linearity is critical, as all incoming signals are superimposed to a transmitter leakage spaced only a few tens of megahertz from the signals itself, and probably much larger in amplitude. If a direct conversion architecture as that reported in [2] is chosen, linearity considerations must be applied to both second and third order distortion[54]. Role of high linearity design is further underscored if active mixers are chosen, which, introducing a flicker noise component that degrades the receiver noise figure, reducing system distortion budget. Research efforts have therefore been carried out in the past to understand correct ways to model intermodulation

distortion in RF building blocks([51], [50]), and to account for this distortion at a system level([55], [11], [2]), to enable aware architectural choices.When linearity constraints are conjuncted with noise constraints, radio receiver design space becomes extremely narrow (and in a real design one should account for I/Q imbalance, which was not made in this work).System level tradeoffs must be therefore accurate, and understanding the way the parameters describing blocks are related(e.g. relation between gain and linearity simultaneously achievable by a given block at a given current consumption) is a must. This work started from an exhisting design of direct conversion UMTS receiver, developed at the University of Pavia and described in ([11],[50],[2]), with the aim to develop and test a design space exploration tool for UMTS CMOS radio receivers, which, based on the platform paradigm described above, could constitute a solid system level framework act to confirm or improve results presented in [2], [11] and [50] and based more conventional approaches(i.e. design experience and simple behavioral models).

1.3 Roadmap

As the main critics that the design community usually addresses to works in the field of analog design automation are dependence on technology or application, and lack of a scientific approach(i.e. there is no quantitative way to "measure performances" of a design methodology, so that there is no quantitative way to "design" a design methodology), I'll apply the platform paradigm to the case study in such a way to make it most immune to these critics. Designing an analog platform quantitatively means understanding what properties must have the behavioral models used in the system optimization phase, and what accuracy is required by the modeled system for the various model components(linear and nonlinear responses, gain). We'll perform this analysis in the next chapter. Dependence on technology and application is imbued in the Analog Constraint Graph and in the behavioral model. An analog constraint graph based on rough device models, and conceived with in mind one precise application, will unavoidably be inaccurate when a different technology or a different application is modeled. The same holds for behavioral models as well. As a result, a considerable amount of time will be spent in studying accurate device models for use in the ACG, and in solving specific problems introduced by the use of so accurate models in the context of a platform paradigm. Also, a careful analysis of the sources of nonidealities(noise and distortion) in the modeled amplifiers will be performed to understand the accuracy of the proposed behavioral model. Particular emphasis will be on understanding third order nonlinearity arising at the LNA-Mixer interface port, a subject that, to the author's knowledge, has never been treated in depth in RF design literature. This is the subject of chapter 3 and 4. Pieces are put together in chapter 5. Here, a behavioral model is proposed for the low noise amplifier that keeps into account linear and nonlinear effects arising from communication with the mixer. This model is parametrized in terms of simultaneously selected LNA output space parameters. System optimization of the receiver based on the proposed set of behavioral model is also performed and discussed. Conclusions follow.

Chapter 2

System Level Modeling

2.1 Model Requirements

Real feasibility of system level design crucially depends on the chosen behavioral models. Models should be accurate enough to allow realistic estimates of system-level non-idealities, and yet simple enough to be practical for hand-exploration of system potentialities, or at least to achieve reasonable computation time if fed to a numerical optimizer. The accuracy and simplicity requirements obviously are contrasting, so that a tradeoff is unavoidable. This chapter deals with representations of mildly nonlinear RadioFrequency circuits and systems. Complexity arising from nonlinearity is in this case added to complexity deriving from the large differences in time constants which is typical of Radio Frequency communication systems.
Two different representations(Volterra Series and Describing Functions) are discussed and compared. A set of models based on Volterra series is proposed, and the complexity of this models is evaluated; finally, sensitivity analysis is used to relate block model parameters accuracy to receiver signal to noise plus interferer ratio accuracy predictions.

2.2 Input independent representations of nonlinear systems

Whereas linear systems excited with a purely sinusoidal input respond with a purely sinusoidal output of different phase and frequency, nonlinear systems exhibit totally different behavior: spectral components of the outputs always includes harmonics of the input tone; while subharmonics are sometimes encountered as well. Even if we only choose to deal with steady state response, initial conditions and input waveform play a critical role in determining the spectral component of the output. To avoid this kind of complications, we chose to model only asymptotically stable nonlinear systems which do not exhibit subharmonic lock. The minimum requisite of our representation is that it has to be input independent, i.e. the set of parameters appearing in the model equations must NOT depend on the input waveform. Examples of this kind of models are both describing functions and Volterra series based models.

2.2.1 Describing Functions and Generalized describing Functions

Describing functions theory is well known in the field of nonlinear control systems theory [7], where it has mainly been used for the stability analysis of nonlinear systems.They have also been used in local oscillator design [6]. Consider now a stationary SISO nonlinear system excited with a sinusoid of frequency f_1 and amplitude A. If the system is asymptotically stable,steady-state response y(t) will then be periodic with frequency fundamental frequency f_2.We define n_{th} order describing function of

the system the quantity

$$G_n(f_1, A) = \frac{\int_0^{\frac{1}{nf_2}} y(t)e^{(j2nf_2\pi t)}}{A} \tag{2.1}$$

i.e. the ratio of the n_{th} harmonic of the output spectrum to the amplitude of the input. If the system under exam is linear, then this function does not depend on A and reduces to the frequency response of the linear system. If on the other hand the system is a memoryless nonlinear system, frequency dependence disappears. This description may be used without any extra workload to deal with systems which exhibit subharmonic lock. If this does not happen anyway, $f_2 = f_1$ holds. To better understand the nature of this description,suppose that the nonlinear system of interest is excited by a single tone at frequency f_2, and followed by a band-pass filter of central frequency kf_2 and bandwidth $B \leq f_2$. If this happens the system output will contain only the k-th harmonic, and a k_{th} order describing function will be an exact representation of the system whenever a single tone excitation is provided(This also is a methodology to measure describing functions). The main problem with this kind of representation however occurs when the system is excited with arbitrary inputs. Two tone input waveforms with tones at frequency $f_1^1 and f_1^2$,and amplitudes $A_1 and A_2$ will give an output with frequency $f_2 = GCD(f_1^2, f_1^1)$. We may define a generalized second order describing function to cope with these situations. In general, m-tones input may be described by a n-th order generalized describing function, defined as

$$G_n(f_1...f_m; A_1...A_M) = \int_0^{\frac{1}{nf_o}} y(t)e^{2j\pi f_o nt}dt \tag{2.2}$$

$$f_o = GCDf_i i = 1...m \tag{2.3}$$

We remark here that

$$G_n(f_1, 0, 0, ..., 0; A_1, 0, ..., 0) = G_n(f_1, A_1)$$

so that the single tone describing functions may be derived from the multiple tone ones. It is also apparent that the most general representation of this kind for a nonlinear system is a an application G*: $(R \times C)^\infty \rightarrow (R \times C)^\infty$(i.e. an application form the space of non-periodic spectra to the space of non-periodic spectra),so that machine memory needs unavoidably put a limit on the attainable frequency resolution of any implementation of this method. Another major disadvantage of describing functions is the scarce modularity of this approach. Consider a memoryless nonlinear system with f_2 periodic output and linear feedback network with frequency response $\beta(f)$. Moreover, suppose $\beta(f_2) = -1, \beta(f), f \neq f_2 = 0$ (ideal bandpass feedback), and suppose the system is characterized by $G_1 = \frac{e^A}{A}$. Overall first order describing function could be found simultaneously solving

$$in(t) = A - e^{in(t)}$$

$$\int_0^{1/f_2} y(t)dt = e^{in(t)}$$

Which cannot be made analytically. These problems discourage the direct use of describing functions when efficient behavioral models are needed; describing functions remain however a useful conceptual tool to their exact nature.

2.2.2 Wiener-Volterra series

Wiener-Volterra series expansion of a nonlinear system may be derived in different ways, each highlighting one particular feature of the model.In [12], a methodology for deriving the expressions of Volterra kernels of a circuit containing nonlinear capacitors, resistors and current sources directly from nodal equations is shown. In [58] on the other hand, the relation between state-space representation of a

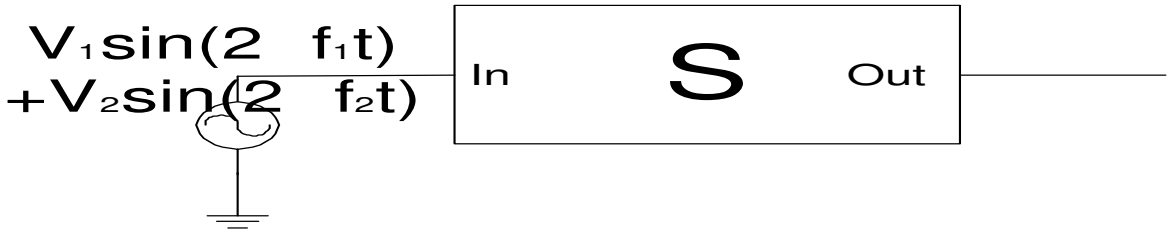

Figure 2.1: Intermodulation test

nonlinear system and its Volterra expansion is highlighted.Whatever derivation is followed, the output of a Volterra represented nonlinear system may be written as:

$$y(t) = h_1(t) \bullet x(t) + h_2(t_1, t_2) \bullet (x(t_1)x(t_2)) + ... + h_N(t_1...t_N) \bullet \Pi_1^N x(t_i) \tag{2.4}$$

Here, dots represent convolution operations, and quantities h_i are called generalized impulse response functions, or Volterra Kernels, of order i. If Fourier transform of both members is taken, equation (2.5) is obtained:

$$Y(f) = H_1(f)X(f) + \int_{f_1=-\infty}^{f_1=+\infty} H_2(f - f_1, f_1)X(f - f_1)X(f_1)df_1 + ... \tag{2.5}$$

Where quantities H_i are obtained as i-dimensional Fourier transforms of the i-th order Volterra Kernels. Volterra expansion has been proven to achieve uniform convergence over compact sets([58]); while global uniform convergence is only achieved for a particular class of systems , referred to in [21] as fading-memory systems. Major problems of Volterra representation come from system identification. Despite different nonlinear terms give spectrally superposed responses, a frequency domain approach based on (2.5) is the most natural choice; however, it requires accurate choice of input frequencies([20]). Moreover, the number of frequency points chosen for identification has a sensible impact on model accuracy [8] and cannot be optimally determined a priori. A different approach could be that of building Volterra models from circuit analysis and device models, as suggested in [12]. Although useful from a circuit design perspective, this approach cannot be accurate from a quantitative standpoint due to the lack of sufficiently accurate device models [35]. Volterra modeling anyway presents sensible advantages over describing functions, principally in the fact that the a fixed set of kernels may be used for describing the system under any kind of inputs; and in the greater modularity of the approach (See [12] for a description of the ways of deriving a Volterra representation for a system obtained from the arbitrary connection of Volterra described systems).

2.2.3 Relation between Volterra series and Describing function representation

In order to understand the connection between Volterra and describing function representation of a nonlinear system, consider figure (2.1) The system S is excited with two sinusoidal tones of amplitudes A_1 and A_2 and frequencies f_1 and f_2 such that $f_1 - f_2 = f^* = GCD(f_1, f_2)$.Therefore, $f_1 = kf^*$, $f_2 = (k+1)f^*$ holds. If we choose to adopt a describing functions based representation, third order intermodulation product evaluates to :

$$IMD_3^{1,DF} = G_{k+2}(A_1, A_2, f_1, f_2)$$

$$IMD_3^{2,DF} = G_{k-1}(A_1, A_2, f_1, f_2)$$

. If a third order Volterra representation is assumed on the other hand, we have:

$$IMD_3^{1,V} = \frac{1}{4}H_3(f_2, f_2, -f_1)A_2^2 A_1$$

$$IMD_3^{2,V} = \frac{1}{4}H_3(f_1, f_1, -f_{-2})A_1^2 A_2$$

If harmonic responses are now considered, we have

$$y(f_1)^V = H_1(f_1)A_1 + \frac{3}{4}H_3(f_1, f_1, -f_1)A_1^3$$

$$y(f_1)^{DF} = G_k(A_1, 0, f_1, f_2)$$

We suppose now to be dealing with a memoryless nonlinear system, so that y(t)=g[x(t)] with $g : R \to R$. If nonlinear system transformation g is assumed analytical and $A_1 + A_2 \leq r_g$ where r_g is the convergence radius of the Taylor series expansion of g, we may integrate per per series to arrive at

$$G_{k+1}(f_1, f_2, A_1, A_2) = \int_0^{\frac{1}{f^*}} g(A_1 \cos(2\pi f_1 t) + A_2 \cos(2\pi f_2 t)) \exp(2(k+1)\pi f^* t) dt = \sum_{i=0}^{\infty} \int_0^{\frac{1}{f^*}} \frac{a_i}{i!}(A_1 \cos(2\pi f_1 t) + A_2 \cos(2\pi f_2 t))^i \exp(2\pi(k+1)f^* t) dt \quad (2.6)$$

When the polynomial operator is applied to the sum of sinusoids, all harmonics of the GCD of these two sinusoids are generated. Due to the orthogonality property of sinusoidal functions, however, only contributions at frequency $(k+1)f^* = 2f_2 - f_1$ give to describing function a non zero contribution.We know one of these terms to be given by the cubic nonlinear term a_3. Under the hypothesis made of $f_1/f_2 = k/(k-1)$, it can be shown using elementary discrete mathematics [26] that other terms contributing to output at $(k+1)f^*$ occur for nonlinear coefficients $3 + n \cdot (2k-1)$ where $n \in N$. This is a very interesting result. Suppose in fact a maximum nonlinearity order M is to be considered(i.e. suppose to know that for values input signals occurring in practice, function g is well approximated by the first M terms of its Taylor expansion): in this a different number of nonlinear terms will contribute to nonlinear distortion depending on system input frequencies. Moreover we also see that if k is high the intermodulation response of the system will be third-order nonlinearity dominated, independently of the overall system nonlinearity. Repeating preceding calculations for the first order one input describing functions G_1 gives:

$$G_1(A_1) = \sum_{i=0}^{\infty} A_1^{2i} \frac{a_{2i+1}}{(2i+1)!} \quad (2.7)$$

so that effectively all of the odd nonlinear polynomial coefficients contribute to gain compression or expansion. We also see that whenever third order nonlinearity is assumed dominant, a tradeoff occurs between identifying gain compression and intermodulation distortion characteristics of the system.

2.2.4 Linear Periodically Time-Varying Systems

Linear Periodically time varying systems have been introduced in Electrical Engineering to deal with nonlinear systems exhibiting a strongly nonlinear behavior with respect to one input port and an almost linear behavior with respect to other excitations.The classical example of this situation is mixer RF port modeling([22],[4]), although great results have been obtained as well in [24],[3] for phase noise modeling. A linear Periodically time varying system is described by a an impulse response function,

$h(t, t-\tau) = h(t+T, t+T-\tau)$).Conversely to the LTI case, dependence on t must be explicitly included. Output to such a system at time t is calculated via superposition integral:

$$V_u(t) = \int_0^t h(t, t-\tau)u(\tau) \tag{2.8}$$

Due to $h(t, t-\tau)$ periodicity in t, we may expand it as a Fourier Series:

$$h(t) = \sum_{k=0}^{\infty} h_k(\tau) \exp\left(j * k * 2 * \pi\tau/T\right) \tag{2.9}$$

Substituing (2.9) into (2.8), we have

$$V_u(t) = \sum_{k=0}^{\infty} h_k(\tau) \exp\left(j * k * 2 * \pi\tau/T\right) \bullet u(\tau) \tag{2.10}$$

$$V_u(f) = \sum_{k=0}^{\infty} h_k(f)u(f - k/T) \tag{2.11}$$

This modeling methodology the basis of Spectre Periodic AC analysis, and is especially useful when response to the large tone may be efficiently and accurately calculated. Labeled $H(V, f)$ the voltage dependent system transfer function and $O(f_1, A_1)$ the operator result of the solution of steady state circuit equations with input frequency f_1 and amplitude A_1, then saying the circuit is approximated by a $PSS + PAC$ behavior is equivalent to saying that $H(O(f_1, A_1), f_2) \mid_2 A_2$ (where $\mid_2$ denotes second harmonic) is an expression of the second order system describing functions. As a result, conclusions drawn in the previous chapter about describing functions may be extended to the case of LPTV systems and $PSS + PAC$ representations.

2.3 Proposed continuous time model

Proposed models were built in a Volterra Series framework. Developed models are suited to all those systems which can be represented by a signal flow graph made of arbitrarily connected Linear Time Invariant and Polynomially Nonlinear systems. The major problem with modeling RF systems is the existence of narrow band signals centered around high frequency carriers. We deal with this kind of signals using a slight modification of classical complex-envelope theory [23]. The main difference with classical theory is that there could be the need of feeding a signal with a spectrum consisting of narrow band signals centered around harmonics of a fundamental carrier f_0. Consider for now the case of an input spectrum made up of N equally spaced tones, fed into a system such that its output y(t) is related to the input x(t) by $y(t) = x(t)^2$. If $c_0, c_1, ..., c_N$ are the tones amplitudes, and calling the output spectrum $o_0, ..., o_{2N}$, we may write:

$$o_i = \sum_{k=i}^{N} c_k c_{i-k} = \vec{c} \bullet \vec{c}(i) \tag{2.12}$$

$$c_{-i} = \bar{c_i} \tag{2.13}$$

$$i = 0...2N$$

Suppose now a higher nonlinearity order is present.We may see the operation of calculating the output of a such a nonlinear system built of two steps: in a first step, power of the input signal up to the order of nonlinearity K are calculated, later these are scaled by the polynomial coefficients and added.

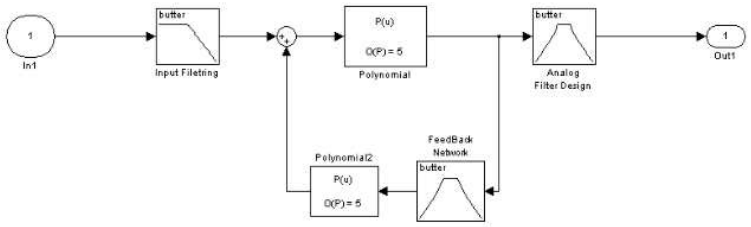

Figure 2.2: Example of system structure

In order to calculate the third order terms $o_i^{'} i = 0...3N$ once second order response $o_i(t) i = 0...2N$ is known, it is sufficient combining such a second order response with the input in a manner similar to the one described above:

$$o_i^{'} = \sum o_k c_{i-k} = \vec{o} \bullet \vec{c}(i) \tag{2.14}$$

Fastest way to do this is using fast convolution algorithms, which are present in MATLAB library. We further notice that all the information necessary to perform calculations is contained in the $c_i, i \geq 0$, as $c_{-i} = \bar{c_i}$ holds. The model of computation may be summarized as follows:

1. Given $\vec{c}$, compute $c_i, i < 0$ and define $\tilde{c} = [\vec{c}, c_i, i < 0]$
2. Put $o_0 = \tilde{c}$
3. Put $o_i = o_{i-1} \bullet \tilde{c}$.Repeat this step for i=1 to N
4. For $i < N - 1$(nonlinearity order),zero pad o_i so that$dim(o_i) = dim(o_N)$ holds $\forall i$
5. Compute $\sum_{i \geq 0} o_i$

Now, it is important to notice that even though this model of computation has been described for time-independent c_i, it is still valid as long as these coefficients are a slowly varying function of time. The signal is then a quasi-sinusoidal signal, or a multi carrier narrow-band signal. Mathematical condition for this to hold is that the narrow-band spectra centered around the harmonics at the output of the nonlinear block are disjoint, i.e.

$$B_{in} \leq \frac{f_{carrier}}{2K} \tag{2.15}$$

Where B_{in} is the maximum of the bandwidths of input signal complex envelopes, and K is the nonlinearity order.

2.3.1 Discrete-Time modeling

Even though the use of baseband equivalent representation saves substantial computational time for narrow band signals, the use of a variable step integration process for solving model equations may still be an overkill.This overhead may be reduced through a discrete time implementation. This transformation solely regards the frequency dependent part, as memoryless nonlinear blocks are left unchanged. Filters are converted to baseband equivalent representation and discretized by sampling their continuous time frequency response with step ΔF, and using a MATLAB optimization algorithm to find the optimal IIR representation.IIR representation has been preferred due the better stability properties of the MATLAB IIR filters nonlinear phase response synthesis tool with respect to the FIR filters one. The discretization frequency ΔF is related to the output signal bandwidth B , to the sampling frequency and to the execution time of the system by the relation:

$$\Delta F = \frac{2B}{T_{tot} f_{sample}} = \frac{4 * B^2}{T_{tot}} \tag{2.16}$$

If, as it is the case in the proposed model, frequency resolution may be chosen by the user, then (2.16) may be used to choose T_{tot} the model execution time. Discrete-time models introduce an extra-degree of freedom in the modeling effort, allowing to trade off frequency resolution for computational time as highlighted by equation(2.16).Anyway, neither this class of models solves the main disadvantage of building time-domain models, which arises when multiple nonlinear blocks are connected in series. In these circumstances, overall nonlinearity rises exponentially, exponentially increasing output signal bandwidth. If we suppose to model a radio as a cascade of three third order nonlinear blocks,representing LNA,mixer and baseband chain, an overall nonlinearity of twenty seventh order is obtained.Recalling that UMTS carrier is about 2Ghz, equation 2.15 stops being satisfied when $B_{in} \approx 37Mhz$.This reduced value does not allow, for example, to perform the UMTS intermodulation test.Furthermore it is worth noticing how, even for the standard, 1 channel UMTS bandwidth of 5Mhz, the maximum required computation time has decreased up to a value of 3.7nS, which is extremely short. If time-domain simulation remains the choice, this complexity increase may be dealt with by adding after each an anti aliasing filter with cutoff frequency $f_{max} \leq B_{in}$.
Obviously, this comes at the expense of accuracy.

2.3.2 Frequency Domain model

If now Fourier transform of 2.14 is taken, then one obtains

$$O_h^i(f) = \sum_{k=-M}^{k=M} O_k^{i-1} \bullet C_{h-k} \tag{2.17}$$

We remark that $\vec{C}$ is the Fourier transform of $\vec{c}$ and as such, is composed by N vectors(one for each of the input carrier harmonics). Each vector will have $K^{'} = \frac{B}{\Delta F}$ components. Equation(2.17) is quite more complex than (2.14). We see that elements of O^{i-1} and C (which are vectors) are convolved and then summed in a convolutional way. In order to get better understanding of this process, consider Kronecker Product[17] between two vectors $\vec{a}, \vec{b}$. This will be a matrix G such that:

$$G_{i,j} = a_i * b_j$$

Summing the elements on the k-th non principal diagonal of G:

$$S_k = \sum_{j=-M}^{M} a_j b_{k-j} = (\vec{a} \bullet \vec{b})(k) \tag{2.18}$$

If now we make the hypothesis $a_i \forall i$ is a $K^{'}$ elements row vector, and substitute multiplication in (**??**), with convolution, we get to (2.17). Further notice that not due to hermitian symmetry of spectra, it is not necessary to calculate all the terms of the Kronecker product: posing $dim(O^i) = M$,$dim(C) = N$, and recalling that $M \geq N$, we see that only the non-principal diagonals of index $i \geq -(M-N), i \leq N$ need to be considered. In this representation, linear filtering is equivalent to element-wise multiplication with a $NxK^{'}$ matrix, which represents the system frequency response sampled as described by (2.19):

$$H_{i,j} = H(i * f_c + j * \Delta F) \tag{2.19}$$

We further recall this model is inherently a continuous time model , so that there is no need for solving differential equations in this model. This eliminates the need for the tradeoff between frequency resolution and execution time proper of time-domain models. If calculations performed with this model have to be compared with a discrete time model, the aforementioned tradeoff is obviously reintroduced, as in order to obtain a frequency resolution of ΔF, a system with input bandwidth B and sampling frequency f_c has produce exactly $N = \frac{2B}{\Delta F}$ samples, which in takes us back to equation (2.16). Finally, we also see that the minimum computation time dependence on considered nonlinearity order here is absent(there is no execution time).However, equation (2.15) still has to hold if spectra corresponding to different harmonics do not have to overlap. If this equation should not hold for some block on the other hand, spectral superposition may easily be computed. This information may be easily be used to update the spectral description matrix(See [9]). Reader will anyway notice that if this happens, there is certainly less and less gain in computational complexity from using a multi-carrier baseband equivalent model. We'll make an example to see how this happens. Suppose input bandwidth is 80Mhz, and there is only one carrier with frequency is 200Mhz. For a nonlinearity of order 3, baseband equivalent spectrum around carrier fundamental becomes 240Mhz,while that around harmonic number 2 and DC is 160Mhz. Baseband equivalent spectra are superposed now, but is no matter if superposed zones are simplified before next nonlinear operations are performed.The transformation is shown in figure (2.3)

2.3.3 Tone Based Model

There are circumstances under which the proposed vector based model performs quite badly form a computational efficiency perspective. These may be generically classified as those where the input is made up of a very few tones, with strongly variable spacing. Consider for example the situation of figure (2.4).In order to obtain a vector based representation for this spectrum, a resolution of ΔF_1 has to be chosen.The spectrum is thus described by a vector of $2 + \frac{\Delta F_2}{\Delta F_1}$ elements, while if input frequencies were specified along with amplitudes, a 3×2 matrix would be enough. Even though complexity analysis of the models has not yet been performed, it is intuitive to understand how this situation becomes unfair when $\frac{\Delta F_2}{\Delta F_1} \gg 1$. To deal with these situations a list based input representation is chosen.Input is now described as a $K \times 2$ matrix, where K is the number of input tones, and only nonzero spectral components are represented.Nonlinear operations and filtering are performed only for those frequencies that included in the input and output spectra, saving a big deal of computation time in case of input sparse spectra.

2.4 Complexity evaluation for the proposed models

All of the proposed models are equivalent, in that they predict steady state behavior of system under exam, for any kind of inputs:therefore, use of one or another model solely depends on what has the lowest complexity and highest ease of use. Refer to the model shown in figure (2.2).Here we have discrete time execution, input signal bandwidth B around M harmonics.Filter frequency resolution is still ΔF,and nonlinearity order N. From (2.16) we find $\frac{T_{sim}}{T_{exec}} = \frac{2B*N}{\Delta F}$ which is the number of timesteps needed.For each timestep, output of 2 IIR filter and a nonlinear block has to be calculated. Supposing for the sake of

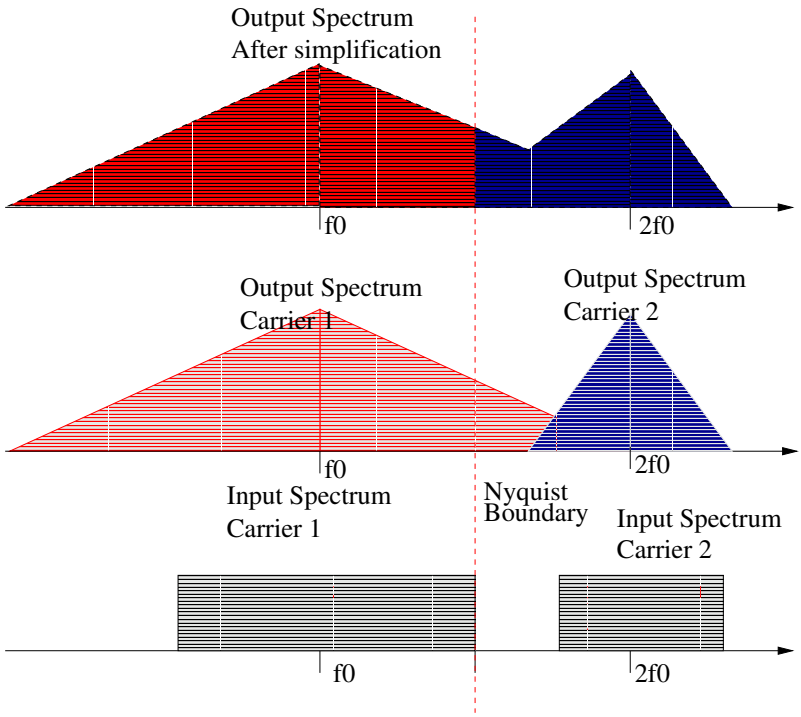

Figure 2.3: Spectral superposition and simplification for FD models if (2.15) doe not hold

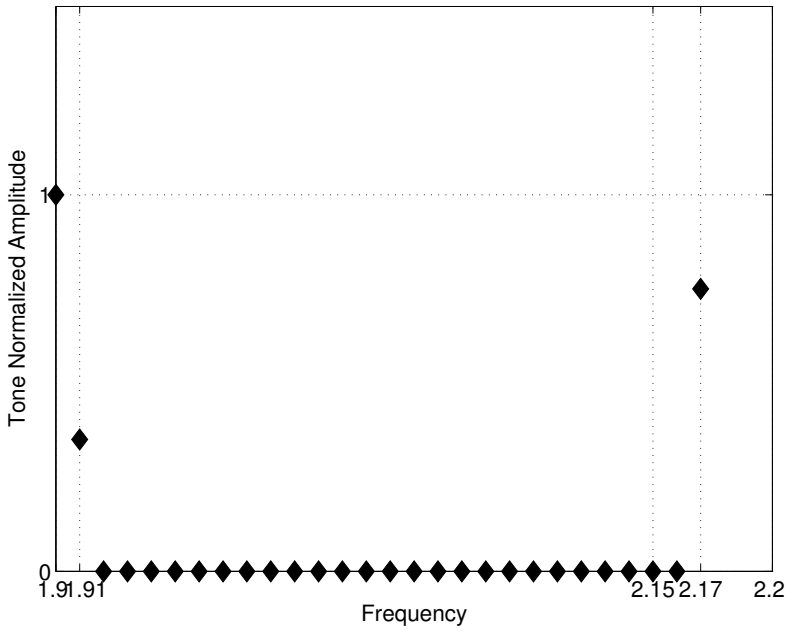

Figure 2.4: Example of spectrum ill-conditioned for a vectorial frequency domain model(frequency domain stiff system): the total bandwidth is imposed by $f_3 - f_1 = 280MHz$, while $\Delta F = f_2 - f_1 = 10MHz$: a 28-element vector is required for vectorial representation, despite only three tones are present.

simplicity this filters to have the same order s, then this takes $(N+1)s$ multiplications for each timestep, for each harmonic(The term N is due to the fact the as in $SIMULINK$ multirate model execution is supported, filter collocated after the nonlinear block will have an execution time N times higher than the one preceding it, as reflected by 2.16). Input harmonics are M, while output harmonics are NM. Then number of multiplications required to calculate filter outputs is $N \cdot (M+1) \cdot (N+1) \cdot s$. For the nonlinear block, N convolutions have to be performed. If these are calculated using a fast convolution algorithm([17]), each convolution requires calculating two FFTs,a product and a IFFT. Computational expense for computing the FFT of a j-elements vector is $j \log j$, so that complexity for each convolutional step may be written as:

$$3(j_1 + j_2 - 1)(1 + \log(j_1 + j_2 - 1))$$

For the case in exam, put $D_i = dim(o_i)$, $D_{in} = dim(C_{in})$:

$$N_{mult} = \frac{2 * B * N}{\Delta F}(N(M+1)(N+1)s + \sum_{i=1}^{N-1} D_i \log(D_i) + D_{in} \log(D_{in}) ... \tag{2.20}$$

$$+(D_{in} + D_i - 1) \log(D_{in} + D_i - 1) \tag{2.21}$$

$$D_{i+1} = D_i + D_{in} - 1 \tag{2.22}$$

$$D_1 = 2M + 1 = D_{in} \tag{2.23}$$

It is easy to prove that

$$D_i = 2 * M * (i) + 1$$

, so that the nonlinear terms contribution becomes:

$$N_{mult}^{NL} = \sum_{i=1}^{N-1} 3(2M \cdot (i+1) + 1) \log(2M \cdot (i+1) + 1)) \tag{2.24}$$

Linear dependence on input bandwidth is apparent through the factor $\frac{2B \cdot N}{\Delta F}$, which also shows an hyperbolic dependence on frequency resolution.Dependence on N and M is super linear, and in particular, dependence on N is stronger the higher is M. Consider now the frequency domain model.Here all computations are performed at the same time(there are no timesteps). We further suppose all of the channels to have the same bandwidth, and thus the same number of elements in each row.Filtering operations each require vector multiplication: if M channels are present, each with $\frac{2B}{\Delta F}$ channels are present at the input, at the output there will be M channels, each with $\frac{2N \cdot B}{\Delta F}$ components. Thus, total number of operations for the linear part is $\frac{2BM}{\Delta F}(N+1)$. For the nonlinear part, N convolutive steps have to be performed. For each of this steps, the non-principal diagonals of a Kronecker product matrix are considered.We remark that these matrices are obtained from vector products in the form $\vec{b} \diamond \vec{a}$, with $\dim(\vec{b}) \geq \dim(\vec{a})$. Moreover, we recall that only positive output harmonics are calculated.Consider than a $(2N+1) \times 2M+1$ matrix A, and define i-th non principal diagonal the set of coefficients $A_{(2N-i-k),+k)}$.Because of the imposed size constraints, one finds:

$$2N - i - k \in (0, 2N) \rightarrow k \in (-i, 2N - i) \tag{2.25}$$

$$k \in (0, 2M) \tag{2.26}$$

$$i = N - M - 1 ... 2N \tag{2.27}$$

Solving (2.27) yelds

$$k \in (0, 2M+1) i \geq 2(N - M) \tag{2.28}$$

$$k \in (0, 2N - i) i \prec 2(N - M) \tag{2.29}$$

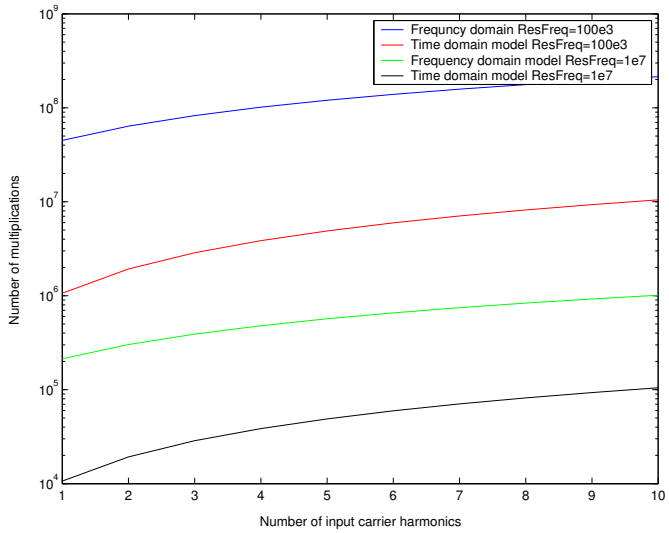

Figure 2.5: Comparison of complexity of time and frequency domain models

Thus, number of convolutions to be computed is

$$N_{conv} = (N-M+1)(2M+1) + \sum_{i=0}^{2(N-M)} 2N-i = (N-M+1)(2M+1)+2N(2(N-M)+1)-(N-M)(2(N-M)+1)$$

, and after rearranging terms:

$$N_{conv} = 2N^2 - 4M^2 + 2NM + 2M + 2N + 1 \tag{2.30}$$

Each convolution is made between a vector of length j_1 and one of length j_2, and finally complexity becomes:

$$\sum_{i=1}^{N} N_{conv}^i \cdot C_{conv}^i \tag{2.31}$$

$$N_{conv}^i = 2(2(M-i)+1)^2 - 4(2M+1)^2 + 2(2M+1)^2 + 4i(2M+1) + 2(2(2M+i)+2) + 1 \tag{2.32}$$

$$C_{conv}^i = 3(1 + \frac{2B(i+1)}{\Delta F}) \log (1 + \frac{2B(i+1)}{\Delta F}) \tag{2.33}$$

Evaluation of the preceding formulas show that time-domain model requires about one order of magnitude less multiplications than a frequency domain one.On the other hand, frequency domain model has been implemented in MATLAB, which is naturally a faster environment than Simulink: as a result,difference in execution times is not so bad. However, the frequency domain model was not used in this work, as the tone based model is better suited to the spectra of isgnals used [2] to perform UMTS tests. For what regards the aforementioned tone based model, computations are here performed directly. Suppose that the tone based represented model is excited with M tones.Input filter computations will then require M multiplications. Before taking care of the output filtering step, the nonlinear block is examined. We need to compute both the frequencies and the amplitudes of the tones which build up the output spectrum. The model used computes all of the output tones, and then simplifies eventual

outputs with the same frequency. For a signal made up of M tones,fed into a nonlinearity of order N, number of output tones is

$$N_{outTone} = \sum_{i=1}^{N}(2M+1)^i = \frac{(2M+1)((2M+1)^N - 1)}{2M}$$

Where positive and negative tones are considered in the final expression.As usual, the factor 2M+1 derives from the need to consider both positive and negative frequencies in these computations. Many of these tones will actually have the same frequency and thus will be recompacted after simplification. Moreover, it is the simplification process itself to give to each output tone the correct multiplicity. A different approach might rely on calculating output tones multiplicity symbolically, and calculating only the terms resulting different from symbolic analysis.Obviously, this approach does not eliminates the need for simplification to be made after computing nonlinear terms, but dramatically reduces the number of output tones. Supposing a nonlinear product of order j is considered, with M positive input tones, say $f_1 ... f_M$,plus DC, output frequencies may be written as

$$F_o = \sum_{\sum n_i = j} n_i f_i$$

. Coefficients for each of these output frequencies may be derived from elementary combinatorial calculus(see for example [15]) to be equal to

$$C_o = \frac{j!}{\Pi_{\sum n_i = j} n_i!}$$

Also notice that this approach actually requires performing operations on just M+1 tones, as one can use hermitian simmetry to eliminate the negative side of the spectrum. Although this approach might seem appealing, it requires coefficients C_o to be either calculated or stored, adding "hidden" complexity. For this reason, it is not practical to implement tone based models when either the nonlinearity order or the number of input tones is exceedingly high. Conversely, they represent a conceptually simple approach, and they are effective for low-complexity, sparse spectra as those used in UTRA-FDD tests.

2.5 Model Implementation

Time Domain Model All time domain models have been implemented in $Simulink^R$, and completed of MATLAB scripts used to set their parameters and execute them from command line ([10]). The blocks have been and parametrized in function of the input bandwidth, so that execution time of each block is automatically calculated to be the minimum such that conditions of 2.16 hold. As already mentioned, in the early phase of the work filters were converted to baseband and discretized using a general frequency domain approach and MATLAB algorithms for IIR filters synthesis.This choice was made in order to produce the maximum stability and generality of the approach, especially when nonlinear phase response filters are a concern. In light of the knowledge acquired in the course of work, we saw that most filters can be modeled as second order bandpass sections with a low quality factor; furthermore predicting phase response generally falls outside the scope of behavioral models. In light of these details, filter synthesis process could be re-evaluated at the purpose of replacing IIR filters with FIR filters and decreasing the computational effort associated with performing the Filter Synthesis. One of the problems encountered during the use of complex valued time domain models lies in evaluating intermodulation distortion terms. As previously stated, a two tone input of amplitude A_0, A_1 and frequency separation ΔF is represented in this model as $V_{in} = A_0 + A_1 \exp j2\pi\Delta F t$. Third order intermodulation products lie at $f_1 = f_0 - \Delta F, f_2 = f_0 + 2\Delta F$, so one of these terms is actually

spectrally superimposed on one of the signals in this complex representation. To separate these two contributions, complex processing is performed on the signal itself. In detail, third order contributions at frequency $f_0 \pm \Delta F$ become: $-\frac{A_1^3}{4}\exp(2\cdot\pi\cdot\Delta F t) + \frac{A_0^2 \hbar A_1}{4}\exp -(j\cdot 2\pi\Delta F t)$ Which are superimposed to the linear term $k_1 A_1 \exp(j\cdot 2\pi\cdot\Delta F)$).Separating signal real and imaginary part and considering their value at $t = t^*$

$$-\frac{\mid A^2 \mid B}{4}\cos(x+2\phi_a) + (\mid A \mid - \mid A^3/4 \mid)\cos(x+\phi_a) = Re(V_u(t))$$
$$+\frac{\mid A^2 \mid B}{4}\sin(x+2\phi_a) + (\mid A \mid - \mid A^3/4 \mid)\sin(x+\phi_a) = Im(V_u(t))$$

and finally

$$\mid IM \mid = \frac{\frac{Re(V_u)}{\cos(x+2\phi_a)} - \frac{Im(V_u)}{\sin(x+2\phi_a)}}{2} \tag{2.34}$$

$$\mid S \mid = \frac{\frac{Re(V_u)}{\cos(x+\phi_a)} + \frac{Im(V_u)}{\sin(x+\phi_a)}}{2} \tag{2.35}$$

Where $S = \mid A - A^3/4 \mid$ and $x = 2\cdot\pi\Delta F t^*$ and ϕ_a is known. Similar problems also arise in the definition of input signals, so that ad-hoc signal and noise sources have also been developed. It may be interesting to detail the implementation of a pulse generator and of thermal and flicker noise generators. Consider for example a PSK modulated signal, which may be written as $s(t) = \cos(2\pi f_c t + \Phi_K)$. When this signal is considered in presence of a second sinusoid of frequency $f_1 = f_c - \Delta F$ and amplitude A_1, representing transmitter leakage or another interferer, time varying phasor representa $(A_1 + A_2\exp(j2\pi\Delta F t + \Phi_K) = A_1 + A_2\cos(\Phi_K)\cos(2\pi\Delta F t) - jA_2\sin(\Phi_K)\sin(2\pi\Delta F t)$.Real and Imaginary part are generated separated and fed into the system.
For a white noise source, it is known ([23])that defined N_0 the flat power spectral density, its band-pass equivalent is given by

$$n_i(t) + j\cdot n_q(t) \tag{2.36}$$

with n_i and n_q gaussian uncorrelated processes and $S(n_i) = S(n_q) = \frac{N_0}{2}$. The in phase and in quadarature terms are egenrated through $MATLAB^R$ Band Limited White Noise block, and then phase shifted and added. Finally, flicker noise is generated through the algorithm reported in [46]. Notice that this generator is inherently real because carrier frequencies are supposed much higher than flicker noise corner frequencies of the modeled devices. Obviously all of random seeds of different noise sources are generated randomly at the beginning of the execution phase of the model itself, so as to minimize the correlation amongst sequences generated by the same source at different executions and samples of different source during the same execution.

Frequency Domain Model Because of the previously calculated high complexity, frequency domain model was never used in practice, despite the core scripts have been implmented and tested in MATLAB or simple input signals.

Tone Based Model The nonlinear tone based model implementation is due to F.Vincis and is described in ([1]). The use of this model resulted in a significant improvement in simulation time and ease of data access and interpretation with respect to the time domain model for the input signals assumed in [2] as test vectors, so that this model has been used as the engine of the optimization process descirbed in 5.9.2.

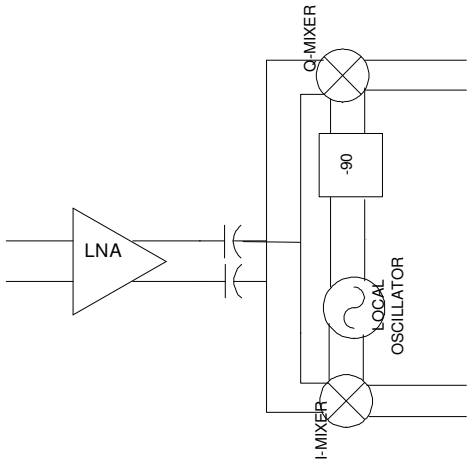

Figure 2.6: Receiver front end architecture

2.6 Sensitivity Analysis

In any attempt to build behavioral models of a class of circuit instances, model accuracy requirements have to be accurately evaluated. If the model is supposed to be used human operated system-level exploration, usually the accuracy requirements are very relaxed, especially if the intended operator is an experienced circuit designer, which can use behavioral models informations just as qualitative hints on circuits sizings.If on the other hand automated synthesis has to be performed, accuracy requirements are much more stringent. The effort reuired by the development of this kind of models is such that a quantitative way to trade-off accuracy on different circuit performances, given an application, must be envisioned. This section proceeds as follows: starting from receiver architecture, structure of the behavioral model of each block is chosen(i.e. block nonlinearity order, filter orders etc.) ; using the parameters of this model, expressions for the signal to noise plus distortion(signal to noise floor) ratio predicted by the model are given. Sensitivity analysis is finally applied to these equations, so as to relate global model accuracy requirements constraints to model parameters accuracy. In the following, we'll refer to a frequency domain description, however, because of the previously stated relation between time and frequency domain models, all of the following derivations also hold for time-domain models.

2.6.1 Building Blocks Models

For the remaining part of the section, we'll suppose the front end to be composed of the cascade connection of a LNA and a mixer.This structure is depicted in 2.6 The mixer local oscillator port is supposed to be driven by a local oscillator, characterized by the corner frequency of the Lorentzian portion of its phase noise spectrum f_c. As local oscillator was not modeled, this parameter is supposed constant and is not included in the given analysis. Low-noise amplifiers and mixer are both supposed to exhibit a third order nonlinear behavior, with gain and nonlinear coefficients showing for the moment unspecified dependence on frequency.Finally, notice that 2.6 underlines the AC coupling between the LNA and each of the I and Q mixers: as a consequence, LNA contributions to overall second order intermodulation distortion are neglected and LNA may be characterized by an odd function.Consider now the expression reported in [2], expressing compliance of a front end to UMTS standard system requirements.

$$\frac{S}{N+I} \geq 7 - SPG[dB] \tag{2.37}$$

Where SPG is the despreading gain, for the case of 12.2kBs UMTS transmission equal to 25 and the value of 7 is the SNR value required by a PSK modulation to achieve a BER of 10^{-3} First of all, notice that 2.37 is a conservative constraint as in fact, UMTS data frame includes a cyclic redundancy check code that contributes to SNR with an asymptotic decoding gain [23] of approximately $2.6dB$ which is not included in the derivation of 2.37. This equation has to hold both for the case of sensitivity test and for the third order intercept test. In both cases, we assume two sinusoidal inputs are applied to the system, characterized by amplitudes A_1 and A_2 and frequencies f_1 and f_2. Then equation 2.37 splits into

$$\frac{P_2^2}{IIP_2} + 4KTR_sBW * F + P_{Leakage}^{Tx} PN_{135Mhz} \leq -99dBm \tag{2.38}$$

$$\frac{P_2^2 P_1}{IIP_3^2}(1 + PN_{67.5}) + 4KTR_sBW * F + P_1 PN_{135Mhz} + P_2 PN_{67.5Mhz} \leq -96dBm \tag{2.39}$$

These equations may be seen as a generalization of the ones reported in [2], and are useful in a system design phase. Here, they will be used to derive accuracy requirements fro the chosen behavioral models, using sensitivity analysis. In order to simplify this task, we'll express these equations in terms of building blocks behavioral model parameters, rather than in terms of system performance indexes.We get:

$$\frac{P_2^2 k_2^2 G_{LNA}^2}{G_{MIX}^2} + 4KTR_sBW * F + P_{Leakage}^{Tx} G_{Mix}^2 PN_{135Mhz} \leq -99dBm \tag{2.40}$$

$$\frac{P_2^2 P_1 (k_3^{LNA} G_{Mix} + G_{LNA}^3 k_3^{Mix}))^2}{G_{LNA}^2 G_{Mix}^2} + 4KTR_sBW * F + P_{Leakage}^{Tx} PN_{135Mhz} \leq -96dBm \tag{2.41}$$

$$F = F_{LNA} + \frac{(\frac{IRN_{Mix}}{4 * KT * R_s BW} - 1)}{G_{LNA}^2} \tag{2.42}$$

These equations can be used to get some understanding of the LNA-Mixer cascade. Note that second order nonlinearity is introduced by mixer alone, while third order nonlinearity has contribution from both low noise amplifier and mixer. The relative magnitude of these contributions depends on the gain and nonlinear coefficients of the individual blocks. Neglecting the interdependence of IM_3^{LNA} and IM_3^{MIX} on gain,it is interesting to look at the curve in figure 2.7, representing the relative contribution of LNA nonlinearity overall third order intermodulation distortion. We see that, while this contribution is dominant at low values of LNA gain,it vanishes as gain is increased.For a real design, and assuming $G_{LNA} + G_{Mix}$ constant,LNA third order nonlinearity will increase as Gain is increased, while mixer nonlinear coefficient k_3 decreases(for a given power dissipation). As a result, a more realistic curve would probably show a floor for the relative LNA contribution to distortion for high gain values. Similarly, assuming a fixed overall gain G_T(e.g. 31[dB] as cited in [2]), relative contribution of LNA to noise factor of the cascade increases with increasing LNA gain, decreasing as a consequence mixer noise importance, while increasing second order distortion.As for an inductively degenerated topology LNA gain is largely determined by the tuned load, while Noise Figure is dominated by input stage parameters, it makes sense to study the way distortion and noise interact in determining the interference floor for the second order intermodulation test. If $k_2^{Mix}, NF_{LNA}, IRN_{Mix}$ and G_T are assumed constant, we it is found

$$N + I \approx \frac{P_2^2 k_2^2 G_{LNA}^4}{G_T^2} + 4KTR_sBW(F_{LNA} + \frac{F_{MIX} - 1}{G_{LNA}^2})$$

Where the term due to reciprocal mixing was neglected, being independent of LNA gain. Plots of this quantity are reported in figure 2.7. Specified all other parameters, an optimum Low Noise Amplifier exists, that minimizes the value of $N + I$.Using elementary calculus, we see that this occurs at

$$G_{LNA}^{Opt} = (\frac{2KTR_s \cdot BW \cdot (F_{Mix} - 1)G_T^2}{P_2 k_2^2})^{\frac{1}{6}} \approx (\frac{2IRN_{Mixer} \cdot BW \cdot G_T^2}{P_2^2 k_2^2})^{\frac{1}{6}} \tag{2.43}$$

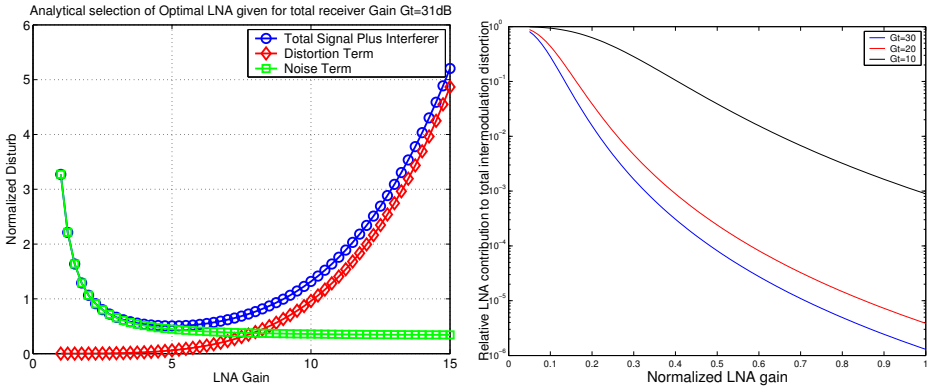

Figure 2.7: Present system design methodology examples: optimal gain selection(left) and third order intermodulation distortion partition computation(Computation performed for design in [2])

, where $F_{Mix} \gg 1$ was considered. For example, design reported in [2] reports $31dB$ gain, as well as $k_2 = 20 \cdot 10^{-3}$,$F_{Mix} = 24.465$. This leads to $G_{LNA}^{OPT} = 16.35dB$.
Although useful from a system design perspective, these observations cannot be accurate, as they neglect actual relations amongst circuit performances.For example, in the latter analysis we did not consider connections between mixer second order distortion, here summarized by parameter k_2, and mixer conversion Gain G_{mix}, or between any of these quantities and the input referred noise. These relations may hardly be expressed analytically for a relatively complex system as a downconverter; nonetheless an experienced designer may use its knowledge of circuits to interpret results from these oversimplified analysis in a critical fashion. For example, it easy to notice that downconverter linearity is expected to improve while gain decreases, while at the same time IRN will probably increase.These considerations lead to deriving a Low Noise Amplifier Gain specification slightly higher than the value predicted using 2.43(For the design in [2],$G_{LNA} = 18dB$). However, intuition does not always give optimal results, especially when designers are not skilled as skilled as authors of [2]. When put in a design automation perspective,equations 2.38, 2.39 may be used to derive performance constraints on the proposed behavioral model accuracy. We perform this task through sensitivity analysis. This step is crucial in our modeling efforts, allowing model accuracy on linear performance to be traded for accuracy on linear performances, which are much easier to predict, therefore reducing the modeling effort.Labeled $P_2 = \frac{P_2^2 k_2^2 G_{LNA}^2}{G_{MIX}^2}$ the nonlinear contribution to second order distortion test noise floor, and $P_3 = \frac{P_2^2 P_1 (k_3^{LNA} G_{Mix} + G_{LNA}^3 k_3^{Mix}))^2}{G_{LNA}^2 G_{Mix}^2}$ the nonlinearity contribution to noise floor in the third order intermodulation test; applying sensitivity analysis to equations(2.40-2.42), we obtain

$$\frac{\Delta N}{N} \approx 2\frac{\Delta G_{LNA}}{G_{LNA}}\frac{F_{Mix}}{F_{Receiv}} \tag{2.44}$$

$$\frac{\Delta P_2}{P_2} = -2\frac{G_{Mix}}{G_{Mix}} + 2\frac{\Delta G_{LNA}}{G_{LNA}} + 2\frac{\Delta k_2^{MIX}}{k_2^{MIX}} \tag{2.45}$$

$$\begin{aligned}\frac{\Delta P_3}{P_3} \approx \frac{\Delta G_{LNA}}{G_{LNA}}\frac{-2G_{MIX}k_3^{LNA} + 4G_{LNA}^3 k_3^{MIX}}{G_{MIX}k_3^{LNA} + G_{LNA}^3 k_3^{MIX}} + \frac{\Delta G_{MIX}}{G_{MIX}}\frac{-2G_{LNA}^3 k_3^{MIX}}{G_{MIX}k_3^{LNA} + G_{LNA}^3 k_3^{MIX}} \\ + \frac{\Delta k_3^{LNA}}{k_3^{LNA}}\frac{2G_{MIX}k_3^{LNA}}{G_{MIX}k_3^{LNA} + G_{LNA}^3 k_3^{MIX}} + \frac{\Delta k_3^{MIX}}{k_3^{MIX}}\frac{G_{LNA}^3 k_3^{MIX}}{G_{MIX}k_3^{LNA} + G_{LNA}^3 k_3^{MIX}}\end{aligned} \tag{2.46}$$

Substituting back into 2.38,2.39:

$$\frac{\Delta D_2}{D_2} = 2(-\frac{G_{Mix}}{G_{Mix}} + \frac{\Delta k_2^{MIX}}{k_2^{MIX}})(1 - \frac{N}{D}) + 2\frac{\Delta G_{LNA}}{G_{LNA}}((1 - \frac{N}{D}) - \frac{F_{Mix}}{F_{Receiv}}\frac{N}{D}) \leq \epsilon \tag{2.47}$$

$$\frac{\Delta D_3}{D_3} = \frac{\Delta G_{LNA}}{G_{LNA}}(2\frac{F_{Mix}}{F_{Receiv}}\frac{N}{D} + a_1) + \frac{\Delta k_3^{LNA}}{k_3^{LNA}}a_2 + \frac{\Delta k_3^{MIX}}{k_3^{MIX}}a_3 + a_4\frac{\Delta G_{MIX}}{G_{MIX}} \leq \epsilon \tag{2.48}$$

$$a_1 = \frac{-2G_{MIX}k_3^{LNA} + 4G_{LNA}^3k_3^{MIX}}{G_{MIX}k_3^{LNA} + G_{LNA}^3k_3^{MIX}}(1 - \frac{N}{D}) \tag{2.49}$$

$$a_2 = \frac{2G_{MIX}k_3^{LNA}}{G_{MIX}k_3^{LNA} + G_{LNA}^3k_3^{MIX}}(1 - \frac{N}{D}) \tag{2.50}$$

$$a_3 = \frac{G_{LNA}^3k_3^{MIX}}{G_{MIX}k_3^{LNA} + G_{LNA}^3k_3^{MIX}}(1 - \frac{N}{D}) \tag{2.51}$$

$$a_4 = \frac{-2G_{LNA}^3k_3^{MIX}}{G_{MIX}k_3^{LNA} + G_{LNA}^3k_3^{MIX}}(1 - \frac{N}{D}) \tag{2.52}$$

Where ϵ is the tolerable error and D_2 and D_3 respectively represent the total disturb component in the second and in the third order intermodulation performance tests. If now we assume error on model parameters to be independently distributed random variables, we may add variances.Resulting equations respectively represent an ellipsoid in R^3 and an ellipsoid in R^4.These ellipsoids describe the way accuracy on different system parameters may be traded. We will exploit this derivation to reduce to the minimum the accuracy specifications on non-linear responses, which are harder to measure and model.Considering for example performances extracted from design described in [2], we find $k_3^{LNA} = .75, k_3^{MIX} = 6.7, a_1 = 3, a_2 = .0011, a_3 = 1.5, a_4 = -1.5, F_{Mix}/F_{tot} = .47, N/D_2 = .5, N/D_3 = .25$. Notice that errors on the mixer second order nonlinearity parameter k_2 basically depend on simulation convergence issues, causing an inherent trade-off between accuracy and simulation time.Details about how this trade-off has been addressed in this work are exposed in [1]. However, accuracy within 4% is obtained in most cases.For LNA linear gain, accuracy obtainable with the proposed model is order of 10%.Same holds for the case of mixer. Now we can invert previous relations for $\epsilon = .5$ and find:

$$\frac{\Delta D_2}{D_2} \approx .12 \tag{2.53}$$

$$(\frac{\Delta k_3^{MIX}}{k_3^{MIX}})^2 2.25 \leq .04 \tag{2.54}$$

Where k_3^{LNA} was neglected due to extremely low sensititivity to this parameter. For a relative error on k_3^{LNA} of $\pm 50\%$, required relative error on mixer k_3 is about 15%. As shown later,such a low relative error value is not easily met when on a widespread number of designs as those generated from an ACG.

Chapter 3

Device Modeling

We introduced in a previous chapter how one of the main problems of platform based design is the dependence of the explored space on the chosen set of constraints imposed through the Analog Constraint Graph, and through these on the chosen device model. A model as accurate as possible should be chosen to allow re usability of the ACG,i.e. portability of the ACG itself on the maximum number of technologies or applications with minimal changes. Moreover, accurate device models are needed for the identification of effects that dominate the non-ideal behavior of a given circuit; and for getting insight on the modeling strategy itself.
As this works deals with low-noise amplifier modeling, noise performance and input matching reliability are the major concerns. Extrinsic and intrinsic capacitive effects in the active devices, integrated inductors and device noise sources have then been investigated.

3.1 Integrated Inductor Model

The crucial role played by integrated inductors in determining the performances of RF systems is well-known.As a result, numerous models have been developed through the years, together with methods to improve performance metrics of integrated inductors(quality factor Q, self resonant frequency SRF), while maintaining compatibility with CMOS technology. All of the prosed models have the topology reported in figure 3.1.All of the prosed models have the topology reported in figure 3.1. Despite its accuracy, exact expressions for the parameters of this model are not readily available, so that design-kit models are usually built extracting these value from accurate 3D electromagnetic simulators such as HFSS. As models of this kind obviously are not suited for automatic design, in this work we used a simplified model such as that depicted in 3.2. This model allows exact predictions of Q and Self Resonant Frequency. Parameters L, R and C may be expressed in terms of inductance value L,operating frequency ω_0,quality factor Q and self resonant frequency ω_{SRF} as

$$R_{series} = \frac{\omega_0 L}{Q} \tag{3.1}$$

$$C = \frac{1}{L(\omega_{srf}^2 + \frac{\omega_0^2}{Q^2})} \tag{3.2}$$

Another parameter of great importance for integrated inductors is die area occupation. An integrated inductor usually has an intrinsic area on the order of $100\mu \times 100\mu$, which must be added to an almost equally large area reserved by design kit spacing rules. An empirical formula to estimate inductors area

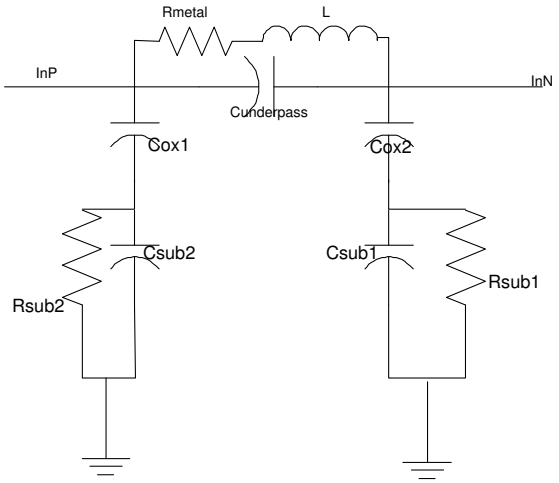

Figure 3.1: Complete equivalent circuit of integrated inductors

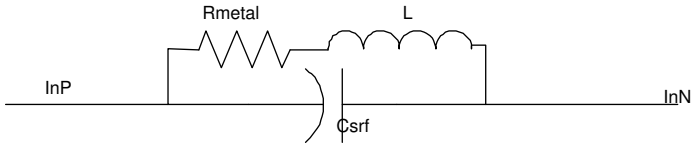

Figure 3.2: Simplified equivalent circuit of integrated inductors

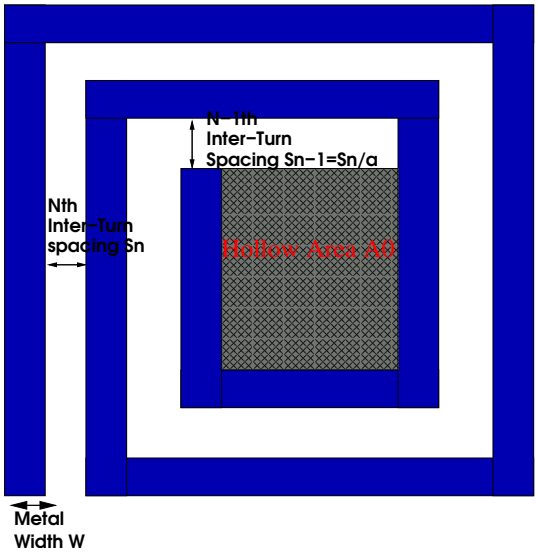

Figure 3.3: Physical meaning of the parameters in equation 3.5

given their inductance value([33]) is:

$$A_{tot} = (\frac{1.3 \cdot 10^{-7} A_m^{\frac{5}{3}}}{L W_m^{1.75} W_{spacing}^{.25}})^6 \tag{3.3}$$

Where A_m is the metal area, W_m the metal width and $W_{spacing}$ the distance between nearby tracks. In the initial phase of the work, this formula was not known and another expression was developed for square inductors by expressing the area as a function of the number of turns,spacing and hollow area.First notice that length of i_{th} turn is related to the length of the $(i-1)_{th}$ turn (see figure 3.3 by $L_i = L_{i-1} + 5(W_m + aS_{i-1}) = L_0 + 5 \cdot i \cdot W_m + \frac{a^{1+i}-1}{a-1} S_0$, where W_m is the metal width of each turn,$L_0 = 3\sqrt{A_0}$ the first turn perimeter(A_0 is the central hollow area [45]), and spacings between successive turns are assumed described by a geometrical sequence of reason a,i.e. $S_{i+1} = a \cdot S_i$. Then $A = L_N^2$ and

$$A = (3\sqrt{A_0} + 5(N-1)W_m + S_0 \frac{a^N - 1}{a-1})^2 \qquad if a > 1 \tag{3.4}$$

$$A = (3\sqrt{A_0} + 5(N-1)(W_m + S_0))^2 \qquad if a = 1 \tag{3.5}$$

N can be found by solving for the inductance value, assumed to grow linearly with the overall metal length ([6]) and thus expressed as

$$L = \alpha(\sum_{i=0}^{N} L_i) = \alpha(N(L_0) + \frac{N(N-1)}{2} W_m + \frac{a(a^N - 1) - N}{a-1} S_0 \qquad if a > 1 \tag{3.6}$$

$$L = N \cdot L_0 + \frac{N(N-1)}{2}(W_m + S_0) \qquad if a = 1 \tag{3.7}$$

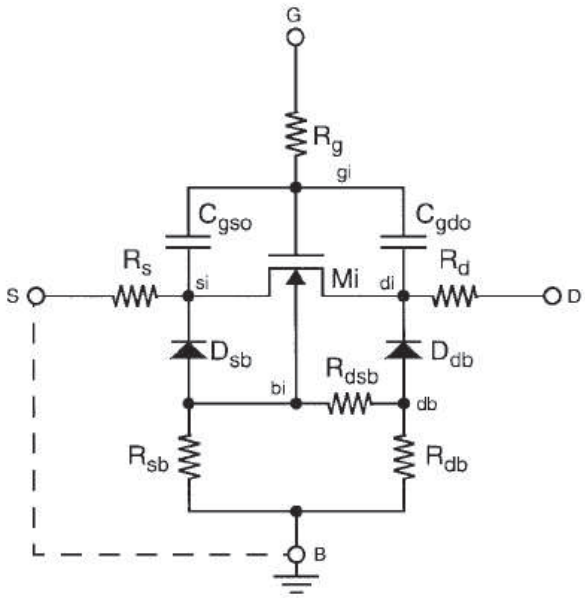

Figure 3.4: MOSFET equivalent circuit,including extrinsic components

3.2 Small Signals Models for transistor extrinsic parts

According to [37], an accurate model for a short-channel transistor at high frequencies should have the topology reported in figure 3.4. Besides the intrinsic section M_1,overlap capacitances C_{gso} and C_{gdo},parasitic junction Diodes D_{sb} and D_{db}, gate,drain and source access resistances R_G,R_D and R_S static feedback resistance and distributed substrate network should be added. In this work only the drain-bulk junctions and overlap capacitances were modeled, due to the crucial role they play in Radio-Frequency Low Noise Amplifier Design. For what regards the distributed gate resistance, the expression $R_g = \frac{R_{sq}W}{12N^2L}$ holds([27]), where R_{sq} is the polysilicon sheet resistance and N is the number of gate layout fingers(assumed contacted at both ends).It can be made negligible by increasing N(for example for the value of $10\Omega/sq.$ reported in [27] for R_{sq} and $\frac{W}{L} = 2500$, $R_G = 1\Omega$ for N=50) and has therefore been neglected in the following. If it has to be included, however, this may be simply done by increasing the value of device non-quasi static resistance r_{nqs} to $r'_{nqs} = r_{nqs} + R_g$. Finally, source and drain access and spreading resistances R_D and R_S and substrate network R_{db}, R_{sb}, R_{dsb} may be modeled as described in [37]. This effect was not included in the digital MM9 Model available for simulations, so it was not modeled.

3.2.1 Overlap Capacitances

Being due to the lateral penetration of Drain Diffusion into the channel, overlap capacitances may quite accurately be represented the a parallel plate capacitor having plates of area $A = WL_{ov}$ and capacitance per unit area C_{ox}. L_{ov} is a measure of the extension of S/D diffusions into the channel, is related to the transistor effective channel length, L_{eff}, by [30]

$$L_{eff} = L_{Drawn} - \Delta L_M - 2L_{ov} \tag{3.8}$$

Where ΔL_M is a zero-mean mask misalignment factor and will be neglected in the following. A direct measurement of L_{ov} obviously is not possible, so that L_{eff} is usually extracted by one of the methods described in [27], and equation 3.8 is used to calculate L_{ov}. As C_{gso} may be directly measured, a tradeoff occurs between L_{eff} and C_{gso} modeling accuracy, especially when fringing effects have a non negligible contribution in determining overlap capacitances, so that the parallel plate approximation results ineffective. To overcome this problem, PhilipsMM9 describes overlap capacitances according to the equation

$$C_{gxo} = C_{ol}W \tag{3.9}$$

where C_{ol} is a process constant.For the ST-Microelectronics HCMOS8D .18 μ process used in this work, $C_{ox} = 8.45e-3, C_{ol} = 3.47e-10, L_{ov} = 45e-9$, so that $C_{ox}L_{ov} = 3.8e-10$ only slightly different from the nominal process value.

3.2.2 Junction Diodes

The S/D-bulk junction diodes contribute a static, temperature dependent leakage current J_0,as well as a small signal resistance r_d and capacitance c_d.For this application, our main interest was capacitance c_d modeling.This capacitance may be described as the sum of three different contributions, respectively determined by the bottom plate area of the junction, the gate sidewall area and the gate edge area, respectively called $C_{jbr}, C_{jsr}, C_{jgr}$. To achieve the maximum accuracy, we directly implemented Philips JUNCAP model using MATLAB. The only difference between the implemented code and the original from Philips is that this version does not account for device interdigitation and junction sharing. Equations may be found in Philips JUNCAP model manual([29]). It is worth noticing that diodes are only modeled for applied voltage less than a process dependent constant with typical values close to zero: as a consequence this model does not give accurate results when either the drain or the source junction is forward biased, as for example happens in Dynamic Threshold or bipolar operation of MOS devices.Fortunately this is not a concern for the application in exam. Figure 3.5 reports the original JUNCAP model results, compared to results from the implemented version.

3.3 Intrinsic Transistor Modeling

3.3.1 Model Topology selection

As reported in [25], operating frequency specifications impact the choice of a device model topology to be used in design and verification. The classical, 5-capacitances model gives correct results for frequencies up to approximately $\omega_T/10$; a complete quasi-static model such as Philips MM9 instead, gives accurate results up to $\omega_T/3$. Exploring upper frequency regions requires non-quasi static effects to be introduced. For .18μ transistors, unity gain frequencies ω_T of the order of 10 to 35GHz are achievable for current levels of .5 $\rightarrow$ 3mA: as a result, the UMTS downlink bandwidth of 2. 11Ghz lies between $\omega_T/5$ and $\omega_T/15$. Notwithstanding the loss of accuracy of such a model, a five capacitances model described in the following subsections was used for imposing Analog Constraint Graph constraints. Details of the derivation follow.

3.3.2 Long Channel Model

Modeling of intrinsic section of long-channel MOS transistors is described in this section using a bulk-charge linearized model similar to the one reported in [31]. Current density equation valid in all operating regions is derived, and specialized for the saturation cases. Results hold for any inversion level, and are parametrized by a single physically meaningful constant I_0.

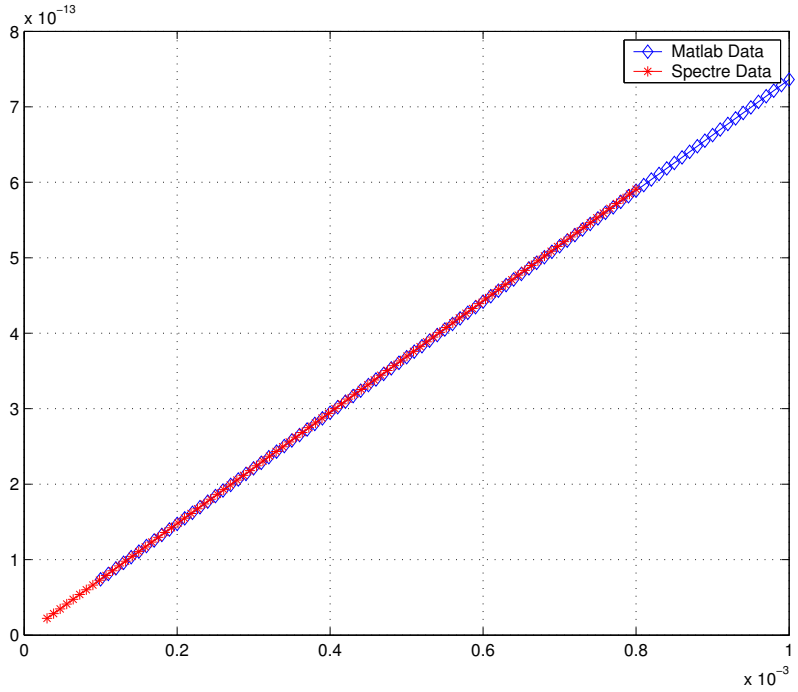

Figure 3.5: Comparison between simulated and calculated values of parasitic drain/source junction capacitances

When charge-sheet approximation holds for the inversion layer [41],current density in a long channel N-MOSFET satisfies the equation:

$$I(x) = W(\mu_n q n(x)\frac{\partial \Psi_s}{\partial x} - qD_n\frac{\partial n}{\partial x}) \tag{3.10}$$

Where Ψ_s is the surface potential. Substitution of Einstein relation for non degenerate semiconductors $D/\mu = \frac{KT}{q}$ yields:

$$I(x) = W(\mu_n q n(x)\frac{\partial \Psi_s}{\partial x} - q\mu_n V_{th}\frac{\partial n}{\partial x})$$

If now quasi static operation is assumed, we may integrate previous equation along the channel using continuity equation to obtain :

$$\int_0^L I(x)dx = I_D L = q\mu_n W(\int_0^L n(x)d\Psi_s - V_{th}(n(L) - n(0)))$$

This is the general equation for the charge-sheet model. In this work, we preferred to use a linearized bulk-charge model that allows much simpler expressions of small signal parameters to obtained with a scarce loss of accuracy. We substitute then $qn(x) = Q(x)$ and $dQ \approx nC_{ox}d\Psi_s$ to obtain:

$$I_D = \mu_n\frac{W}{L}(\frac{(Q(L)^2 - Q(0)^2)}{2nC_{ox}} - V_{th}(Q(L) - Q(0))) \rightarrow \mu_n\frac{W}{L}C_{ox}nV_{th}^2(f^2(L) - f^2(0) - 2(f(L) - f(0)))$$

where n is the subthreshold slope of the device, related to the body effect coefficient γ by $n = 1 + \frac{\gamma}{\sqrt{\Psi_0 + V_x}}$. V_x represents the value of surface potential in the point with respect to which the bulk charge is linearized. The choice of this point influences symmetry and accuracy properties of the model as described in [25], and should not be overlooked when very high accuracy is a concern. To minimize the number of parameters and simplify equations, n was however chosen as a constant in this work, and its value determined through simulations. As $n = 1 + \frac{C_{gb}}{C_{ox}WL + C_{gb}}$, this choice is equivalent to neglect gate-bulk capacitances variations with applied voltage. One immediate consequence is that this model does not predict any maximum in the gm/I_d ratio as a function of gate voltage V_g, failing the treetop test([25]). Finally, as $I_0 = \mu_n C_{ox} n V_{th}^2, i_F = f^2(L) - 2f(L), i_R = f^2(0) - 2f(0)$

$$I_D = I_0\frac{W}{L}(i_F - i_R)$$

. For the $.18\mu$ used in this work, $I0_n = 290nA, I0_p = 100nA, n_N = 1.5, n_P = 1.3$ were calculated, while better accuracy was obtained for $I0_n = 250nA, I0_p = 120nA$.
In the forward active region, $I_R = 0$(obviously this holds as velocity saturation in the channel was neglected) and the previous equation becomes:

$$I_d = I_0\frac{W}{L}(f^2 - 2f) \tag{3.11}$$

Integrating equation 3.10 from zero to a generic point x in the channel gives us the charge distribution as a function of position:

$$I_D = \frac{x}{L}I_0(f^2(0) - f^2(x) - 2(f(0) - f(x)))$$

and

$$f(x) = 1 - \sqrt{1 + (f^2(0) - 2f(0))\frac{(L - x)}{L}} = 1 - \sqrt{1 + i_F\frac{(L - x)}{L}} \tag{3.12}$$

Once relationship between surface potential and terminal voltages is known, small signal parameters may be written using the chain rule as $\frac{\partial I_d}{\partial f_S}\frac{\partial f_s}{V_x}$ with V_x generic terminal voltage. As only forward active

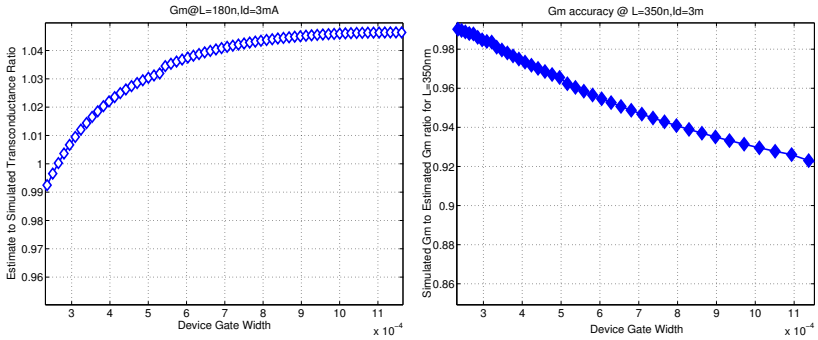

Figure 3.6: Model comparison for transconductance

region is modeled, and $\frac{\partial f_S}{\partial V_G} = \frac{1}{n}\frac{\partial f_S}{\partial V_s}$; $\frac{\partial f_S}{\partial V_B} = \frac{n-1}{n}\frac{\partial f_S}{\partial V_s}$, calculating $\frac{\partial f_S}{\partial V_s}$ is sufficient. Current Density may be written using quasi Fermi Levels as ([27])

$$I_D = \mu_n \frac{W}{L} \int_{V_S}^{V_D} Q(x) dV$$

Thus

$$g_{ms} = \frac{\partial I_D}{\partial V_S} = n g_m = -\mu_n C_{ox} \frac{W}{L} Q(0) = \frac{2I_0}{V_{th}} \frac{W}{L} f \tag{3.13}$$

.

Differentiation of equation 3.11 leads to

$$g_{ms} = \frac{I_0}{V_{th}} \frac{W}{L} 2(f-1) \frac{\partial \Psi_s}{\partial V_S}$$

Equating these two formulas yields

$$\frac{\partial \Psi_s}{\partial V_S} = \frac{f_S}{(f_S - 1)} \tag{3.14}$$

So that surface potential is found 3.14 independent of well potential at low inversion levels , while linearly increasing with in strong inversion.

3.3.3 Small signal parameters

The most important small signal parameter in analog design certainly is the device transconductance gm, or its counterpart gm/I_D ratio or gm efficiency. From previous equations, gm may be expressed as

$$g_m = \frac{2I_0}{nV_{th}} \frac{W}{L} (\sqrt{1 + i_F} - 1) \tag{3.15}$$

For low inversion coefficients levels, we find $\sqrt{1+i_f} - 1 \approx i_f/2$, and $gm \to \frac{I_d}{nV_{th}}$ as well known from subthreshold devices regional modeling. For high inversion levels instead, we have $\sqrt{1+i_f} - 1 \approx \sqrt{i_f}$ and $gm = \sqrt{2\mu_n C_{ox} W/L I_d}$ as known from strop inversion regional models. Figure 3.3.3 reports predictions of equation 3.15 and of PhilipsMM9 for a 1μ long transistor biased at $I_D = 1mA$. Relative accuracy obtained is within . 1. As $gm = \frac{\partial I_d}{\partial V_{gs}}$ is known, an implicit equation for the $I_d - V_{gs}$ device characteristic may be found through integration:

$$\frac{dI_d}{I_0 \frac{W}{L} (\sqrt{1 + IC} - 1)} = \frac{2dV_{gs}}{nV_{th}}$$

recognizing $\frac{dI_d}{I_0 \frac{W}{L}}$ and performing integration of both sides:

$$V_{ov} = nV_{th}(\sqrt{1+IC} + \log(\sqrt{1+IC} - 1)/2) \tag{3.16}$$

.

Capacitances

The subject of MOSFET intrinsic capacitance modeling is widely known in literature [16][47]. From a model design standpoint, most important choice in this field regards whether the developed model should be a charge conserving one or not. Charge conversing models are suited for large signal or transient analysis, and must make use of Ward-Dutton terminal charges partition or similar approaches. Moreover, models based on this approach naturally lend themselves to be integrated in a nine-capacitances small signal equivalent circuit. Not being interested in transient analysis, we did not develop terminal charges and capacitances expression based on charge-conserving schemes; as a result, following calculations suppose a 0-100 charge partition. Integrating equation 3.12 and then applying the methodology shown above for G_m yields:

$$C_{gs}\frac{2C_{ox}WL}{3}\frac{f(f-3)}{(f-2)^2} \tag{3.17}$$

Important notes have to be taken on this equation. Usually, strongly inverted devices are assumed to show a capacitance $C_{gs} = \frac{2}{3}C_{ox}WL$. We can see that in equation 3.17 this is verified as $f \to \infty$. So, classic strong strongly-inverted value is conserved. For weak inversion operation, classical texts report [28] $C_{gs} \approx 0$. Anyway, an important difference has to be made in this case. A MOST may be driven from strong inversion regime toward subthreshold regime in two different ways: the first one is forcing $I_d \to 0$ while holding W as a constant. This leads to (substituting in eq. 3.17) $Cgs \to 0$. On the other hand, inversion level also is decreased if current drive is held constant and transistor aspect ratio is increased. Evaluating then the limit for $W \to \infty$ in equation (3.17) one finds:

$$C_{gs}^{\infty} = \frac{I_d}{4I_0}C_{ox}L^2 \tag{3.18}$$

So that maximum intrinsic capacitance obtainable is linearly increasing with current drive. Figure **??** reports a comparison of intrinsic C_{gs} as predicted by equation (3.17) and by PhilipsMM9 for the a 1μ long transistor biased at a drain current of 1mA. Relative accuracy is of the order of . 3. Charge conservation may be used to find an expression for C_{gb}:

$$C_{gb} = \frac{n-1}{n}(C_{ox}WL - C_{gs}) \tag{3.19}$$

Figure 3.3.3 reports predictions of C_{gb} after equation 3.19 and after Philips MM9.

3.3.4 Noise

Noise behavior of MOSFETs has been widely studied. It is well known that at high frequencies, Drain Induced Gate Noise(DIGN), becomes significant and finally limits noise performance of RF circuits. In this section we will first characterize thermal noise from the channel, and then gate noise. Noise from the channel is known to be thermal in strong inversion and Shot in weak inversion[25]; experimental results have anyway proven that calculations made assuming purely thermal noise are also accurate. Experimental results have anyway proven that calculations made assuming purely thermal noise are also accurate. We will follow this approximation in the derivation of the equations for noise. For a section dx of channel,

$$dS_I(f) = 4KT\mu_n Q_n(x)W dx \tag{3.20}$$

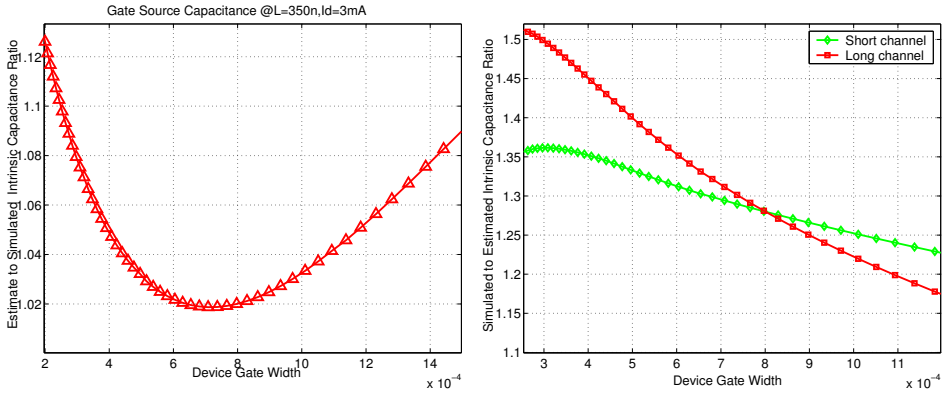

Figure 3.7: Small signal gate source capacitance after eq.3.17 and PhilipsMM9(Short Channel Model euqations are reported in the appendix)

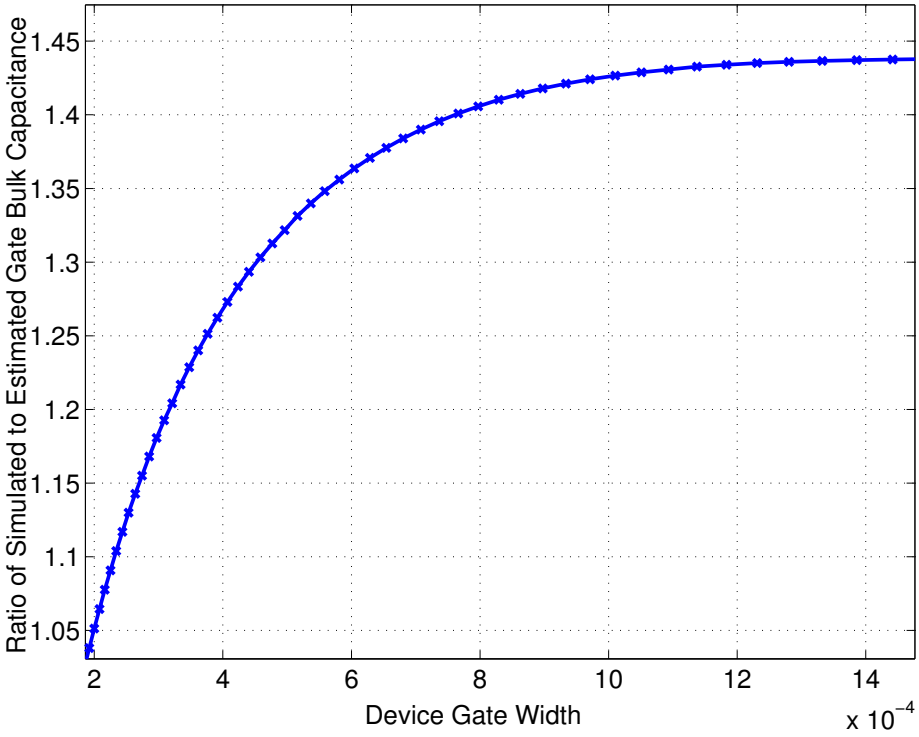

Figure 3.8: Small signal gate bulk capacitance after eq.3.19 and PhilipsMM9

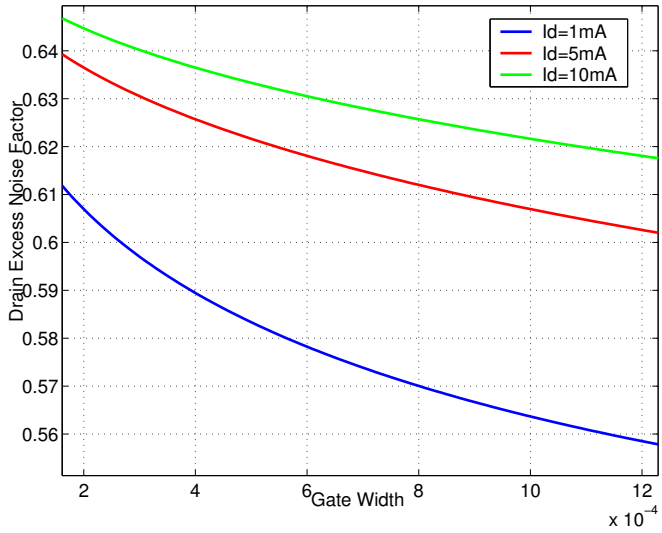

Figure 3.9: Drain excess noise factor versus gate width for variable current

holds. Combining with equation(3.12),integrating, and reporting to the input by dividing for g_m,

$$S_V(f) = 4KT\frac{\gamma}{g_m} \tag{3.21}$$

$$\gamma = \frac{2f-3}{3(-2+f)} \tag{3.22}$$

Notice how, for $f \to 0, \gamma \to \frac{1}{2}$ as predicted by the shot noise hypothesis, while for large f $\gamma \approx \frac{2}{3}$ as predicted by strong inversion formulas. Plots of γ as a function of gate width are reported in figure 3.9 for the aforementioned $.18\mu$ technology are reported below. Variations in drain potential couple to the gate through the insulator capacitance, giving rise to the so-called drain induced gate noise according to $\Delta I_g(x) = j\omega C_{ox} W \Delta V(x)$. According to [65], ΔV must be calculated by considering thermal noise in every section in the channel as a current source ΔI_d and solving for surface potential perturbations introduced by this source in the channel after imposing boudary conditions $\Delta V(0) = \Delta V(L) = 0$. This approach allows writing the noisy surface potential induced by the thermal noise current source ay x_0 ΔI_das

$$\Delta V(x) = -\frac{\Delta I_d x}{g(x)} \qquad if x \leq x_0$$

$$\Delta V(x) = \frac{\Delta I_d (x-L)}{g(x)} \quad if x_0 + \Delta x_0 \leq x \leq L \tag{3.23}$$

Where $g(x) = \mu_n Q(x) = \mu_n C_{ox} W n V_{th} (\sqrt{1 + IC\frac{x-L}{L}} - 1)$ is the differental conductance of the channel at abscissa x. In the frequency domain, gate noise current induced by drain noise thermal current at x_0 evaluates to

$$\Delta I_g(x_0) = \frac{j\omega \Delta I_d(x_0)}{\mu_n n V_{th}} \int_{x_0}^{L} \frac{L-x}{\sqrt{1+IC\frac{L-x}{L}} - 1} dx - \int_0^{x_0} \frac{x}{\sqrt{1+IC\frac{L-x}{L}} - 1} dx \tag{3.24}$$

This equation may be used to express both the power spectral density of the gate noise and the correlation coefficient c in all operating regions. Unfortunately, previous errors and unavailability of [65] up to a

few days ago did not allow going further in this derivation.As a result, throughout the work we used for gate noise quantities the classical strong inversion expression, which is found for example in [6]:

$$S_{ig} = 4KT\frac{\omega^2 C_{gs}^2}{5gm}\delta \tag{3.25}$$

$$\delta = \frac{4}{3} \tag{3.26}$$

$$c = j.385 \tag{3.27}$$

3.3.5 Non-Quasi Static Model

As charge distribution in the channel is known for both of the proposed models, solution of the time-varying continuity equation would allow the derivation a non-quasi static(NQS) model. For the frequency range of operation,however it was enough to introduce first order non-quasi static corrections r_{nqs}:

$$r_{nqs} = \frac{1}{5g_{m0}} \tag{3.28}$$

Non-quasi static input resistance can instead give significant contributions to overall input match and should then be included in input-sizing equations.

3.4 Large Signal Characteristics

Equation 3.15 may be used to calculate device nonlinear behavior using the differentiation chain rule: $\frac{\partial gm}{\partial v_{gs}} = \frac{\partial gm}{\partial IC}\frac{gm}{I_0 W/L}$([32]). Performing these calculations gives:

$$\frac{gm}{I_0 W/L} = \frac{2}{nV_{th}}(\sqrt{1+IC}-1) \tag{3.29}$$

$$\frac{gm_2}{I_0 W/L} = \frac{1}{(nV_{th})^2}\frac{\sqrt{1+IC}-1}{\sqrt{1+IC}} \tag{3.30}$$

$$\frac{gm_3}{I_0 W/L} = \frac{1}{(nV_{th})^3}\frac{\sqrt{1+IC}-1}{\sqrt{1+IC}^{\frac{3}{2}}} \tag{3.31}$$

Device nonlinear behavior versus operating point is shown in figure 3.10. Capacitance nonlinearity may also be extracted from these equations. At this purpose, we may use the chain rule and find $\frac{\partial^i C_{ij}}{\partial V_{gs}^i} = \frac{\partial^i C_{ij}}{\partial IC^i}\frac{\partial^i IC}{\partial V_{gs}^i} = \frac{\partial^i C_{ij}}{\partial IC^i} gm_{i-1}/(I_0 W/L)$It is interesting to see that i_{th} order distortion due to weakly nonlinear capacitors is related to $(i-1)_{th}$ order distortion due to transconductance. This result could be anticipated as device transconductance is related to is proportional to Source terminal charge, while small signal capacitances are derivatives of the same charge. Using long-channel formulas, we recall $C_{gs} = 2/3 C_{ox} W L \frac{Q(Q+3)}{(Q+2)^2}, Q = \sqrt{1+IC}-1$ and find

$$\frac{1}{C_{gs}^i}\frac{\partial C_{gs}}{\partial IC} = \frac{5+\sqrt{1+IC}}{2IC+2IC^2+4IC\sqrt{1+IC}} \tag{3.32}$$

$$\frac{1}{C_{gs}^i}\frac{\partial^2 C_{gs}}{\partial IC^2} = \frac{-8-3IC-20\sqrt{1+IC}}{4(1+IC)^{\frac{3}{2}}(1+\sqrt{1+IC})^2(-1+IC+\sqrt{1+IC})} \tag{3.33}$$

Where superscript í ís for intrinsic. In order to obtain the ratio of capacitance derivatives to total gate-source capacitance, these terms must further multiplied times $\frac{1}{1+C_{gso}/C_{gs}^i} = \frac{1}{1+\frac{3C_{ol}(1+\sqrt{1+IC})^2}{2C_{ox}(-1+\sqrt{1+IC})(2+\sqrt{1+IC})}}$ which can be sensibly lower than 1 at low IC.

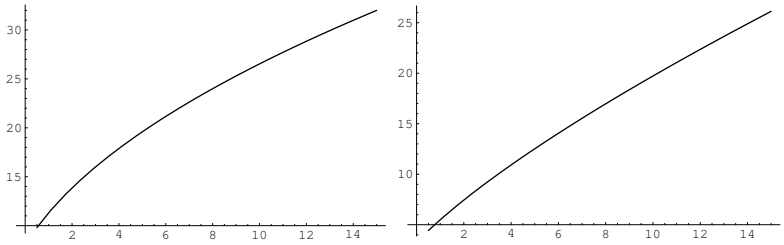

Figure 3.10: Normalized Device second(left) and third(right) order intercept point as a function of inversion coefficient when considering only transconductance nonlinearity

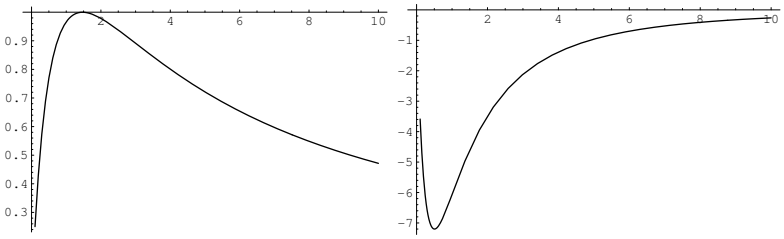

Figure 3.11: Nonlinear coefficients of gate source capacitance versus operating point

Chapter 4

Low Noise Amplifier Analysis

The Low-Noise amplifier is the first block of any receiver chain. It has to raise the signal strength to a level sufficiently high to make negligible the effect of noise from subsequent blocks, while at the same time introducing the least possible noise of its own and providing conjugate match to a 50 Ω antenna. Because of these facts, classical performance metrics of a low noise amplifier are gain,Noise Figure(NF) and S_{11}. A real low noise amplifier however, will also introduce nonlinearity. As in the modeled receiver the mixer is AC coupled to the LNA, only third order LNA distortion is important at a system level. Previous research ([5],[9]) has shown how a very well suited topology to perform the tasks above mentioned is the inductively degenerated transconductor reported in figure 4.1. The modeled receiver used this same topology, which is analyzed in detail afterward.

4.1 Input Matching

An accurate small signal model of this LNA topology is reported in figure4.2). Here C_{pad} is the pad plus protection diodes small signal capacitance while $C_{gb}, C_{gs}, C_{gd}, C_{db}, C_{sb}$ the transistor small signal capacitances(C_{sb} goes in parallel with C_{L_s} and thus in not labeled in fig. (4.2)). Notice that at the drain of M1 only one diode parasitic capacitance is shown due the junction sharing of M1 drain and M2 source. This solution reduces distortion([12],[11]), but requires $L_2 = L_1, W_2 = W_1$. Also recall

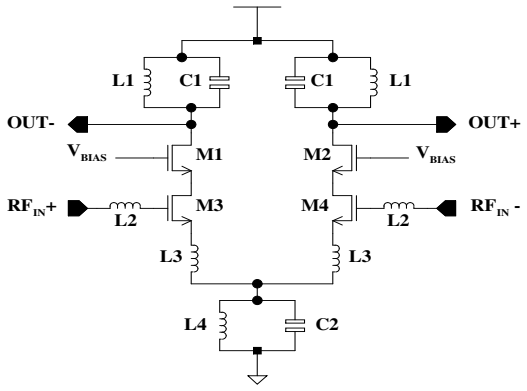

Figure 4.1: Inductively degenerated low noise amplifier amplifier schematic

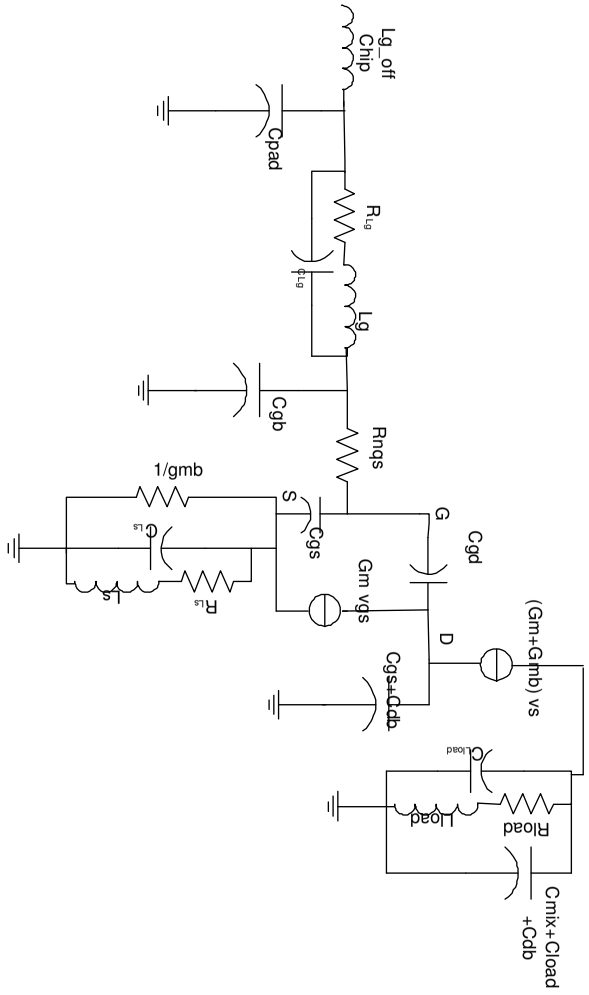

Figure 4.2: Inductively degenerated low noise amplifier amplifier small signal model

that due to unavailability of foundry models for inductors, we used a simplified inductor model. Hence, the capacitances C_{sub1} and C_{sub2} present in the π inductor model between each of the terminals and ground, as well as the resistor which shunt these capacitors cannot be included in our simulations and their effect on design is not included here. However, notice that actually C_{sub1} goes in parallel to C_{pad}, which is much larger, while C_{sub2} is shunted by C_{gb}(and comparable in value). If a first order estimate of C_{sub2} were known, however, this could be added to C_{gb} in the derived equations, leading to even more accurate results.
In order to write input matching equations, we need to calculate the impedance seen between node IN and ground. For the moment, we neglect the influence of C_{PAD}, as we are now interested in calculating the impedance seen between the gate of the transconducting device M1 and ground. Applying nodal analysis, the following expression is found:

$$Y_{in} = j\omega_0 C_{gs}(\frac{Y_s}{Y_s + j\omega_0 C_{gs} + gm_1}) + f(C_{gd}) \tag{4.1}$$

$$Y_s = \frac{1}{Z_s}$$

$$Z_s = \frac{1}{gm_b} \parallel \frac{j\omega_0 L_s(1 - j/Q)}{-\omega_0^2 L_s(C_{sb} + C_{srf}) + j\omega_0^2(C_{sb} + C_{srf})L_s/Q + 1} \tag{4.2}$$

$$f(C_{gd}) = j\omega_0 C_{gd}(1 - \frac{-gm_1 + j\omega_0 C_{gd}(Z_s(gm_1 + j\omega_0 C_{gs1}) + 1)}{(1 + (Z_s)(gm_1 + j\omega_0 C_{gs1}))(gms_2 + j\omega_0(C_{gs2} + C_{db} + C_{gd}))}) \tag{4.3}$$

Notice that $f(C_{gd})$, representing the effect of the gate-drain overlap capacitance on input admittance, may also be calculated through the Miller Theorem to be $j * \omega_0 * C_{gd}(1 - Av(j\omega_0)$, where $A_v(\omega)$is the gain of the cascode stage at the operating frequency. As a result ,we may also deduce that $G_{cascode} = \frac{-gm_1 + j\omega_0 C_{gd}(Z_s(gm_1 + j\omega_0 C_{gs1}) + 1)}{(1 + (Z_s)(gm_1 + j\omega_0 C_{gs1}))(gms_2 + j\omega_0(C_{gs2} + C_{db} + C_{gd}))}$we may simplify this expression by defining $Q_L = \frac{1}{\omega_0 C_{gs} R_s}$ and recalling that usually $\omega_0^2 L_s C_{gs} \ll 1$, $\frac{\omega_T L_s}{R_s} \approx 1$ to obtain

$$G_{cascode} = \frac{-gm/(1 + jQ_L) + j\omega_0 C_{gd}}{(gm(2n - 1)/n + j\omega_0(C_{gs} + C_{db} + C_{gd}))}$$

. Plots of this quantity , as well as of the exact formula,are reported in figure 4.3. Consider now the second term,Z_s. Expliciting inductor self resonant frequency $\omega_{srf} \approx \sqrt{\frac{1}{LC_{srf}}}$, it may be rewritten as

$$Z_s = \frac{j\omega_0 L_s(1 - j/Q)}{\omega_0^2/\omega_{srf}^2(1 + \frac{C_{sb}}{C_{srf}})(-1 + j/Q) + 1}$$

For real designs in $.18\mu$ CMOS, $L_s \approx 500pH$, while $\omega_{srf} \approx 10\pi Grad/sec.$ As a result $C_{srf} \approx 2pF$ while $C_{sb} \leq 400fF$. As a result, we may neglect this effect and consider $Z_s = \frac{1}{gm_b} \parallel j\omega_0 L_s(1 - j/Q) \approx \frac{j\omega_0 L_s(1 - j/Q}{1 + gm_b\omega_0 L_s}$ for the following. The effect of gm_b is essentially that of introducing additional series resistance, equal to $gm_b L_s^2 \omega_0^2$. Its effect is the stronger the higher the frequency, and the higher gm_b and L_s. For the very low values of L_s used in practice however, we see that this contribution is below 2Ω even at 5. 4GHz and will therefore be neglected. Now, we discuss the effect of C_{gd}. For a MOS transistor acting in saturation, effect of drain potential on channel charge is negligible, so that $C_{gd}^i \approx 0$,$C_{gd} = C_{gdo}$ holds. While in a long channel, strongly inverted design $C_{gs} \gg C_{gd}$ is always verified, in deep submicron technologies and at low inversion levels $C_{gd} \approx C_{gs}/2$, giving significant contribution(e. g. a $400\mu wide, .18\mu$ long transistor in ST HCMOS8D has $C_{gs} \approx 350fF, C_{gd} = 150fF$. In figure 4.4, we report plots of $f(C_{gd})$ and of the term previously considered in the design. This effect can also be intuitively undertstood in terms of series to parallel narrowband impedance transformations. Suppose small signal parameters have typical values $C_{gs} = 350fF, C_{gd} = 150fF, \omega_T L_s = 35\Omega$, and

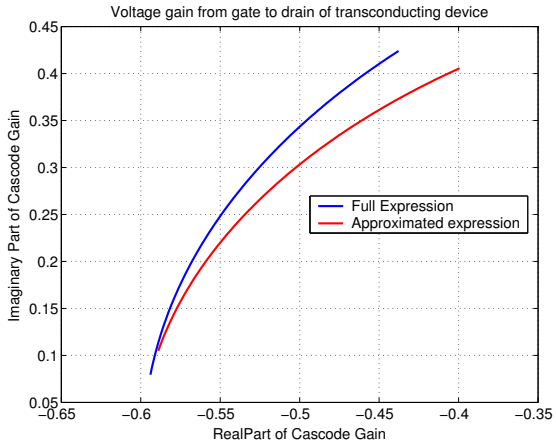

Figure 4.3: Voltage from the gate to the drain of M1

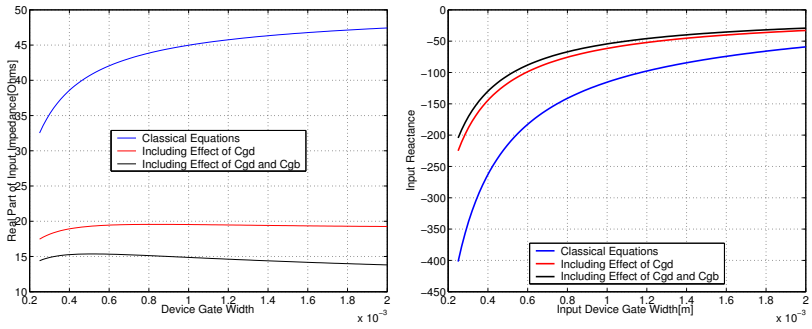

Figure 4.4: Comparison between values of input impedance for classical equations and equations including effect of C_{gd}

$G_{cascode} \approx -.7 + j.25$. The effect of C_{gd} is that of introducing a parallel branch $R = 1684\Omega$,$C = 255fF$(see that $Q = \omega_0 RC = \frac{1-Re(G_{cascode})}{Im(G_{cascode})} \approx 7$ holds). If a series to parallel conversion is performed on the impedance seen looking to the right of C_{gb}, we it is found $R_1 = 1377\Omega$,$C_1 = 350fF$. The parallel of these two RC branches yields a series equivalent $R_{series} = 20\Omega, C_{series} \approx 600fF$, which is well approximated by the empirical formula $R_s = \frac{\omega_T L_s}{1+\frac{2C_{gd}+C_{gb}}{C_{gs}}}$ which in this case gives $R_{series} = 19\Omega$. It is apparent how the overlap capacitance C_{gd} plays a very important role in determining both the real and the imaginary part of the total impedance seen looking into the gate of M1. We may see that the effect for C_{gd} is twofold: on one side, it decreases the resistance seen looking into the gate of M1 by a factor about than 50%,requiring larger degeneration inductors, on the other hand, it increases the input capacitance by a factor of approximately 1. 7 C_{gd}, resulting in a lower value of gate inductance required to achieve matching than predicted by classical equations. When **??-??** are explicited for L_s in terms of device parameters $C_{gs}, C_{gs}, C_{gb}, gm, C_{gd}, gm_b$, and source resistance R_s, a second order expression $aL_s^2 + bL_s + c = 0$is found. Once coefficients a, b and c have been calculated, L_s is found as

$$L_s = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{4.4}$$

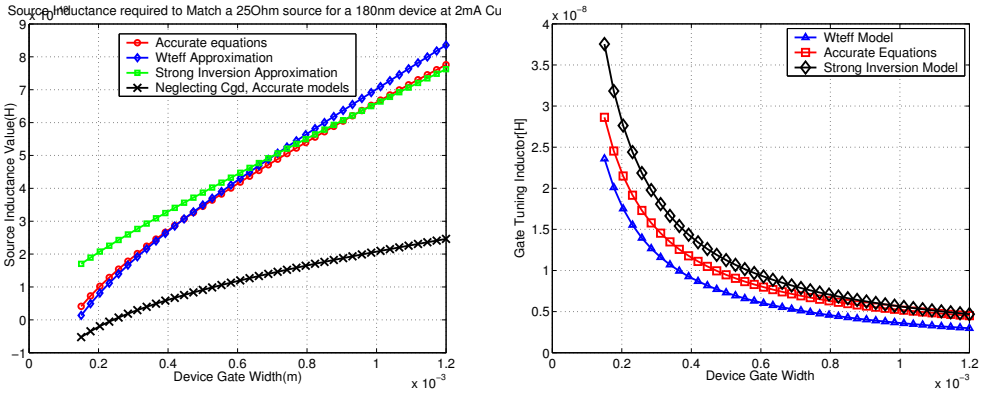

Figure 4.5: Comparison between several expressions for calculating degeneration inductance L_s(left) and gate detuning inductance(right) in ST HCMOS8D .18μ technology.

. Because of the extremely complex expressions for these parameter, however,using numerical methods to solve for L_s is much more practical(this is the approach we followed). Figure 4.5 reports a comparison of the values of L_s required to match a 25 Ωsource according to the classical equations, to the equation reported above and to the $\omega_{T_{eff}}$ approximation reported in ([11],[6])

$$L_s = \frac{R_s}{\omega_{T_{eff}}} \tag{4.5}$$

$$\omega_{T_{eff}} = \frac{gm}{C_{gs} + 2C_{gd} + C_{gb}} \tag{4.6}$$

The difference between the ideal and the two more exact curves is apparent. Notice that in this analysis, except for the approximation$\omega_T L_s \approx R_s$ made in developing the expression for $G_{cascode}$, all device second order effects,i. e. overlap and parasitic junction capacitances, as well as dependence of C_{gs} on the operating point are included at once in the design process, allowing very accurate prediction. For practical cases however, the value of L_s given by (4.5) is a very good approximation. It is also interesting the fact that for the given technology, L_s is well approximated by the expression $L_s = R_s \frac{3C_{ox}WL}{2gm}$ which corresponds to neglecting overlap capacitances and dependence of intrinsic gate-source capacitance on operating point. The approximation made in using for the degeneration inductance the value $L_s = \frac{R_s}{\omega_{T_{eff}}}$ may be more intuitively stated in saying that the circuit of figure 4.2 is equivalent to the one in figure **??** for input matching calculations, once $C_{gs}^{'} = C_{gs} + 2C_{gd} + C_{gb}$ is used for the active device gate-source capacitance. This is the basis of the simplified input matching equations used in the next section.

4.2 Noise

This section deals with Low-Noise Amplifier design for minimum noise figure. Starting from simplified input matching equations, an analysis of different contributions to LNA noise factor is performed. This analysis is more comprehensive than the one reported in literature, as both noise from the gate,source and load integrated inductors as well as noise from the cascode device are explicitly modeled.

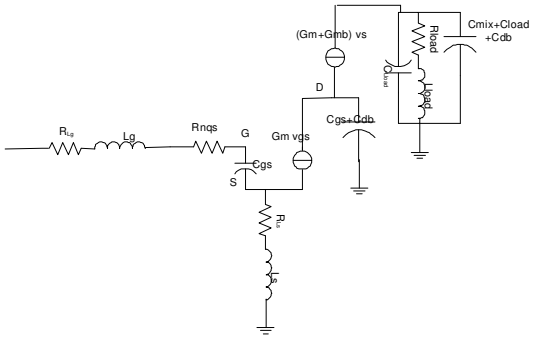

Figure 4.6: Small signal LNA core equivalent circuit for simplified input matching equations derivations

4.2.1 Simplified Input Matching Equations

To get to closed form expression for the Noise-Factor F as a function of device gate width and bias current simple enough to be intuitively meaningful , we develop input matching equations neglecting the effects of C_{pad} and C_{gd}. In this case, the amplifier small signal model becomes the one reported in 4.6 We allow gate and source inductors to have different quality factors Q_g and Q_s to model composite on-chip/off-chip gate inductors:assuming Q_{off} is the off chip discrete inductor quality factor, and $L_g^{OnChip}/L_g = F$ we find

$$1/Q_g = (1-F)/Q_{off} + F/Q_s \tag{4.7}$$

Then, we have $Z_{in} \approx \omega_{T_{eff}} L_s + r_{nqs} + \frac{\omega_0 L_s}{Q_s} + j(\frac{-1}{\omega_0 C_{gs}} + \omega_0 L_s)$. Input matching equation then becomes:

$$R_s = \omega_{T_{eff}} L_s + r_{nqs} + \frac{\omega_0 L_s}{Q_s} - \frac{(\frac{-1}{\omega_0 C_{gs}} + \omega_0 L_s)}{Q_g} \tag{4.8}$$

And may be solved for L_s and L_g to obtain

$$L_s = R_s \frac{(1 - r_{nqs}/R_s - Q_L/Q_g)}{\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)} \tag{4.9}$$

$$L_g = \frac{R_s Q_L}{\omega_0} \frac{(\omega_{T_{eff}} + \omega_0(1/Q_s - (1 - r_{nqs}/R_s)/Q_L)}{\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)} \tag{4.10}$$

$$Q_L = \frac{1}{R_s C_{gs} \omega_0} \tag{4.11}$$

As special cases of these formulas, we find that for F=1(fully on chip integrated inductor), we have$Q_s = Q_g$, and $L_s = \frac{R_s(1 - r_{nqs}/R_s - Q_L/Q_s)}{\omega_{T_{eff}}}$. In finally $r_{nqs} = \approx 0$ and $Q_s \approx \infty$, $L_s = R_s/\omega_{T_{eff}}$ as from classical LNA design equations.

4.2.2 Noise Analysis

Major contributor to noise figure for this topology remains transconducting device. To evaluate this contribution, we perform a return ratio analysis on the equivalent circuit in figure 4.7. Return ratio factor βA and input network functions $\alpha_{ng} = \frac{V_P}{i_{ng}}, \alpha_{nd} = \frac{V_P}{i_{nd}}$are found to be

$$\beta A(\omega) = -gm \frac{\omega L_s}{1 - \omega^2(L_g + L_s)C_{gs} + j\omega C_{gs}(R_s + R_{Lg} + R_{Ls})} \tag{4.12}$$

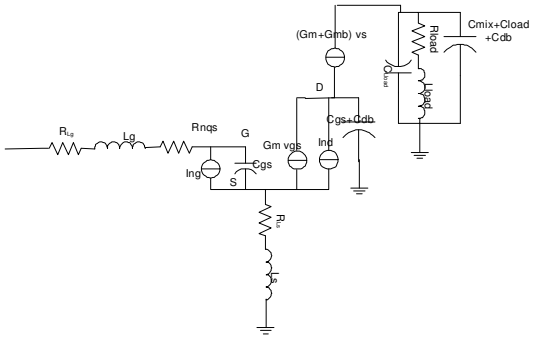

Figure 4.7: Small signal equivalent circuit for transconducting device contribution to noise figure calculations

$$\alpha_{ng} = \frac{R_s + R_{Lg} + R_{Ls}(1 + \frac{1}{j\omega C_{gs}(R_s+R_{Lg}+R_{Ls}})}{1 - \omega^2(L_g + L_s)C_{gs} + j\omega C_{gs}(R_s + R_{Lg} + R_{Ls})} \tag{4.13}$$

$$\alpha_{nd} = \frac{\omega L_s}{1 - \omega^2(L_g + L_s)C_{gs} + j\omega C_{gs}(R_s + R_{Lg} + R_{Ls})} \tag{4.14}$$

the DIGN and channel thermal noise contributions to output current to be:

$$i_u^{cc} = (\frac{\omega_{T_{eff}}}{j\omega_0}(1 + \frac{Q_L}{(2 - \omega_{T_{eff}}L_s/R_s - r_{nqs}/R_s)})i_{ng} + i_{nd})\frac{1}{(1 + \frac{\omega_{T_{eff}}L_s}{R_s}\frac{1}{(2-\omega_{T_{eff}}L_s/R_s - r_{nqs}/R_s)}} \tag{4.15}$$

And for the output power spectral density:

$$PSD_o = 4KT\frac{\frac{\omega_{T_{eff}}^2}{\omega_0^2}(1 + \frac{Q_L^2}{D^2})\frac{\omega_0^2 C_{gs}^2}{5gm} + \gamma gm + 2 \mid c \mid gm\sqrt{4/15})}{(1 + \frac{\omega_{T_{eff}}L_s}{R_s D})^2} \tag{4.16}$$

where $D = (2 - \omega_{T_{eff}}L_s/R_s - r_{nqs}/R_s)$. Substituting the simplified input matching equations, one finds:

$$i_u/I_{nd} = \frac{1}{1 + \chi\frac{\omega_{T_{eff}}(1 - r_{nqs}/R_s - Q_L/Q_g)}{(\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)}} \tag{4.17}$$

$$i_u/I_{ng} = \frac{\omega_{T_{eff}}}{j\omega_0}\frac{(1 + \frac{Q_L}{jD})}{1 + \chi\frac{\omega_{T_{eff}}(1 - r_{nqs}/R_s - Q_L/Q_g)}{(\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)}} \tag{4.18}$$

$$\chi = \frac{R_{Match}}{R_s} = \frac{S_{11} + 1}{-S_{11} + 1} \tag{4.19}$$

$$\eta = \frac{1}{1 + \frac{2C_{gd} + C_{gb}}{C_{gs}}} \tag{4.20}$$

$$D = \frac{-\frac{r_{nqs}}{R_s}(\omega_0(1/Q_s - 1/Q_g) + \omega_{T_{eff}}(1 + \frac{Q_L}{Q_g}) + 2\omega_0(1/Q_s - 1/Q_g)}{\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)} \tag{4.21}$$

$$PSD_o = 4KTgm(\gamma + \frac{4\eta^2}{5}(1 + \frac{Q_L^2}{D^2}) + 2\eta \mid c \mid \sqrt{4/15})(\frac{1}{1 + \chi \cdot \frac{\omega_{T_{eff}}(1 - r_{nqs}/R_s - Q_L/Q_g)}{D(\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)}})^2 \tag{4.22}$$

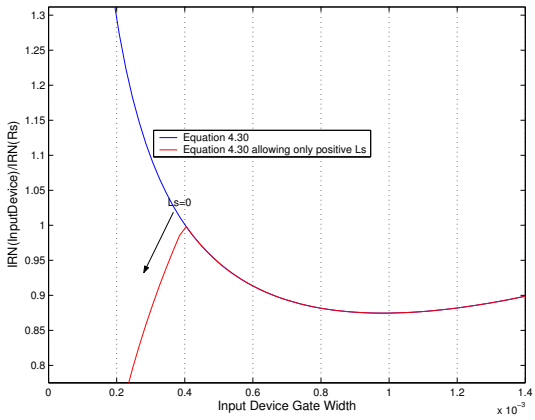

Figure 4.8: F_{input}-1 for Id=1mA,Q=7. 6. Red curve shows results restricting $L_s$to positive values.

It has been previously shown that this device contributes to the output noise power spectral density with both its channel noise and its drain induced gate noise(DIGN). An expression for the LNA noise factor neglecting other contributions is [5]:

$$F = 1 + \frac{(1+\chi)^2\omega_0}{\omega_{T_{eff}} Q_L (1 + \frac{\omega_{T_{eff}}(1 - r_{nqs}/R_s - Q_L/Q_g)}{D(\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g))^2}} (\gamma + \eta^2 \frac{4}{15}(1 + \frac{Q_L^2}{D^2}) + 2\eta \mid c \mid \sqrt{4/15}) \tag{4.23}$$

In [5] it is also shown that the presence of DIGN determines the presence of a minimum in F(W) where W is the device gate width, for any bias current I_d. Later Gatta et al. ([11]) included in the analysis the effect of γ factor variation with bias point. Notice that 4.23 is more general than the equations presented in [5] and [11] as it explicitly includes effects from non-quasi static resistance,inductor series resistance and unperfet input matching in the analysis. This effect corresponds to the factor $D \neq 1$ and to the terms $\chi, \frac{\omega_{T_{eff}} L_s}{R_s} \neq 1$. Notice that this analysis actually supposes that the simplified equivalent circuit 4.6, which was proven to be first-order equivalent to the full equivalent circuit **??**, is also equivalent at the purpose of calculating output noise power spectral densities. This assumption is not theoretically supported and will be verified through simulation. Figure 4.8 reports quantity in 4.23 for $Id = 1mA$,$\chi = 1$,$Q_g = Q_s = 7.6$. Noise from series resistance of integrated inductors may be easily included in the analysis: for what regards gate inductor, R_g goes in series with the source resistance R_s, contributing to the overall noise factor with a term

$$\omega_0 L_g/(QR_s) = \chi \frac{Q_L}{Q_g} \frac{\omega_{T_{eff}} + \omega_0(1/Q_s - (1 - r_{nqs}/R_s)/Q_L)}{\omega_{T_{eff}} + \omega_0(1/Q_s - 1/Q_g)} \tag{4.24}$$

. The situation is slightly more involved for noise from R_{sl} the series resistance of source degeneration inductance. Through nodal analysis, it can be shown that transfer function from V_{sl} to output short circuit current is

$$G_{sl} = \frac{gm}{2j/Q + 1 + \frac{\omega_0 L_s \omega_{T_{eff}} C_{gs}}{Q_s}}$$

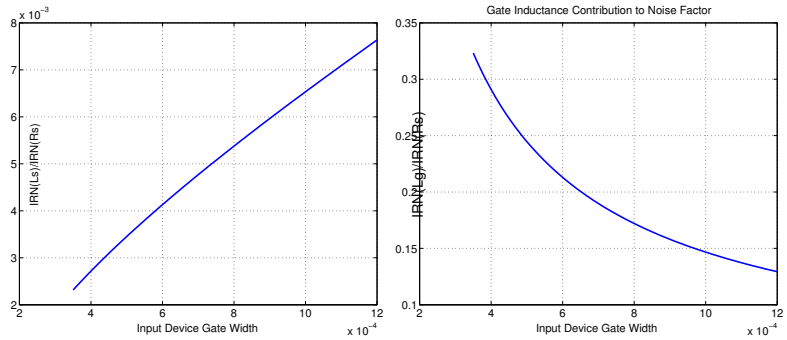

Figure 4.9: Contribution of degeneration inductance finite series resistance(left) and gate inductance(right) to the overall noise factor

Recalling that $R_{sl} = \omega_0 L_s/Q_s$, input referred noise contribution from L_s becomes $\frac{4KTR_{sl}}{(4+(Q_L+\frac{\omega_{T_{eff}}L_s}{Q_s R_s})} = \frac{4KT\omega_0 L_s}{Q_s(4+(Q_L+\frac{\omega_{T_{eff}}L_s}{Q_s R_s})^2)}$. Substituting approximate expression for L_s gives

$$4KTR_s\chi\frac{(1-r_{nqs}/R_s-Q_L/Q_q)(\omega_{T_{eff}}+\omega_0(1/Q_s-1/Q_g))Q_s\omega_0}{(4(\omega_{T_{eff}}+\omega_0(1/Q_s-1/Q_g))^2Q_s^2+\omega_{T_{eff}}^2(1-r_{nqs}/R_s-Q_L/Q_g)^2)} \tag{4.25}$$

A plot of 4.25 normalized to $4KTR_s$ is reported in figure 4.9 for $Q_g = 3Q_s, Q_s = 7.6, \omega_0 = 4.2\pi Grad/s, I_d = 1mA$. Contribution from source inductance appears to be negligible, so it will not be considered in the following.
Noise from the common-gate device also has some influence. First of all, notice from figure 4.11 that in this case GIDN and channel thermal noise generators act in parallel, so that a single transfer function may be used to refer both their contributions back to the input. If this noise generator is I_n^{CG}, we find from elementary small signal analysis:

$$I_u^{CC} = I_n^{CG}(1-G_{cg}) = I_n^{CG}(\frac{j\omega_0(C_{db}+C_{gs})}{gm+gm_b+j\omega_0(C_{db}+C_{gs}}$$

Where G_{cg} is the gain of the Source to Drain current gain of the common gate device. Note that for ideal operation (i. e. $G_{cg} = 1$) noise from the cascode device does not appear at the output(For the sake of precision, previous relation predicts this unrealistic situation for $\omega_0 arrow 0$; however this only happens because $\frac{1}{gm+gm_b} \parallel r_{ds} = \frac{1}{gm+gm_b}$ was considered,i. e. M1 was assumed to act as an ideal current source). Transconducting gain from the signal input to the output, including cascode device, may be written as

$$G_m = \frac{\omega_{T_{eff}}}{j2R_s\omega_0}\frac{gm+gm_b}{gm+gm_b+j\omega_0(C_{db}+C_{gs})} \tag{4.26}$$

Using 4.26, input referred noise from the cascode device may be written as:

$$IRN_{commonGate} = I_n^{CG}\frac{1}{(1+\chi)R_s(C_{db}+C_{gs})\omega_0^2}\omega_{T_{eff}}(gm+gm_b)$$

Substituting the expression of I_n^{CG} in terms of device parameters,

$$IRN_{commonGate} = 4KT(\gamma gm+\frac{4\omega_0^2C_{gs}^2}{15gm})(\frac{-(1+\chi)R_s(C_{db}+C_{gs})\omega_0^2}{\omega_{T_{eff}}(gm+gm_b)})^2 \tag{4.27}$$

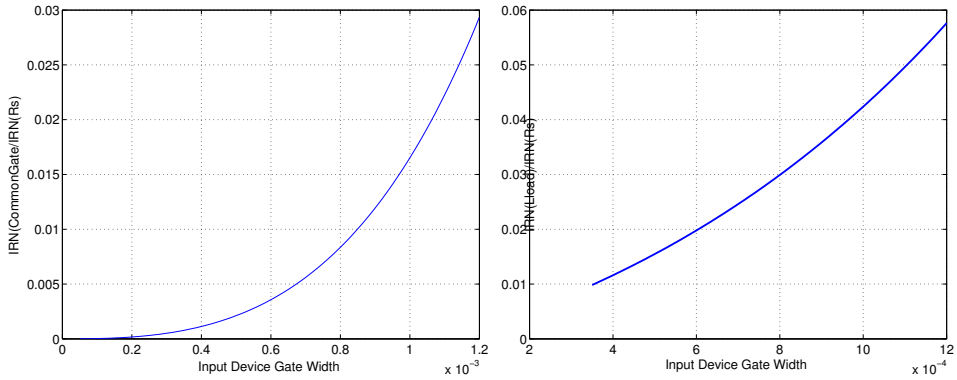

Figure 4.10: Cascode device(left) and Load Inductor contributions to noise factor

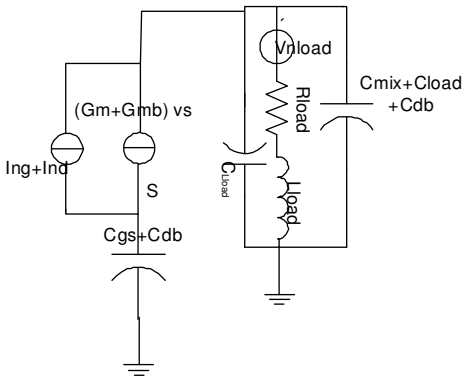

Figure 4.11: Small signal circuit for common gate and load inductor device noise contribution calculations

We report the magnitude 4.27 in figure 4.10. Finally, we consider noise contribution from the load integrated inductor. The circuit corresponding to this analysis is shown in figure **??**. Applying nodal analysis, we find:

$$F_{load} = \frac{(1+\chi)^2 \omega_0 R_s}{Q_s L_{load}(1+1/Q_s^2)\omega_{T_{eff}}^2}(1+(\omega_0/\omega_C)^2) \tag{4.28}$$

where $\omega_C = (gm + gm_b)/(C_{gs} + C_{db})$ is the common gate stage current gain -3dB frequency in radians per second. Notice that increasing L_{load} decreases F_{load} and at the same time increases gain. Thus, we see that from a small signal viewpoint, load inductors as large as possible are desired. Figure 4.10 reports values of 4.28 for $R_s = 25\Omega$, $L_{Load} = 2.8nH$ at various device widths. Notice that for any given current,increasing device width decreases both $\omega_{T_{eff}}$ and ω_C, explaining the monotonic behavior of F_{Load}.

4.2.3 Noise Optimization

Various analyses have been performed to validate the proposed design methodology and understand the effect of all those terms that had not been trated in detail in literature. First of all consider noise from the gate inductor. From matching analysis, it is known that the higher the gate width of the input

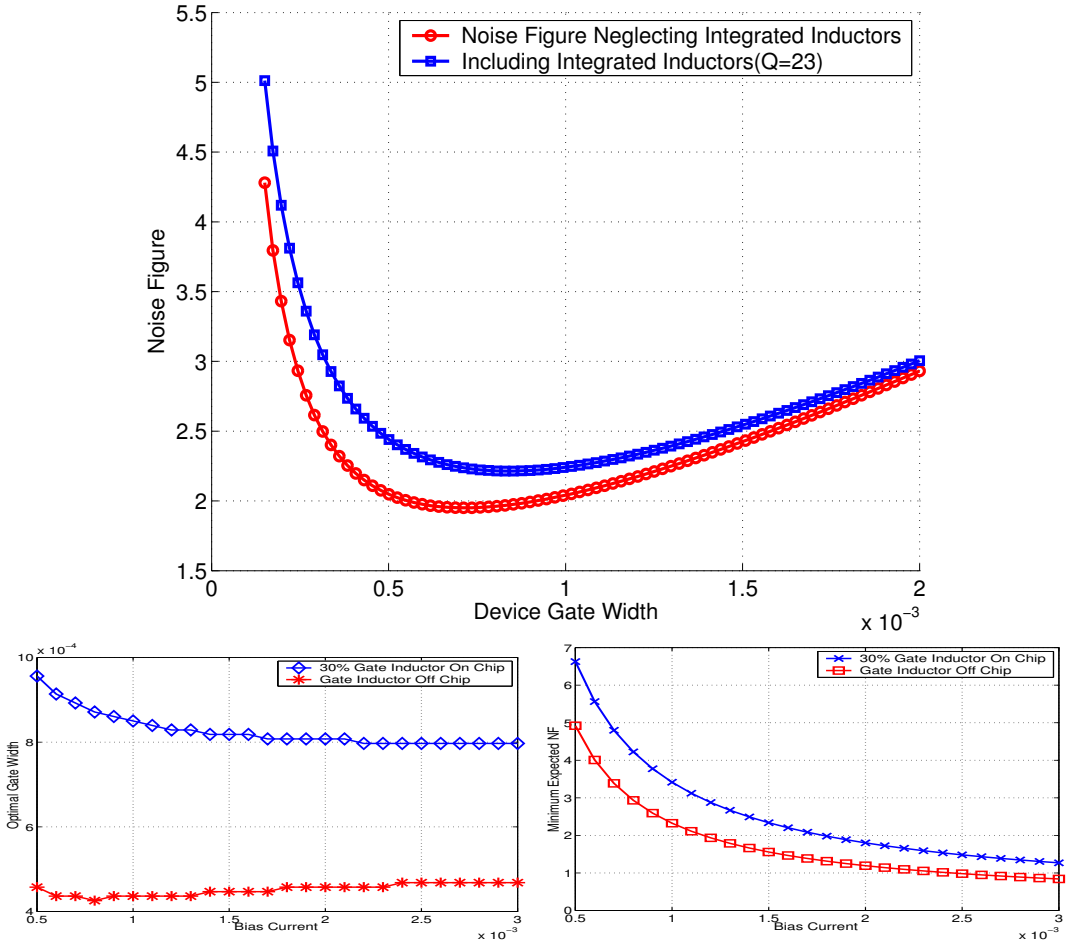

Figure 4.12: Effects of the finite quality factors of integrated inductors on low-noise amplifier noise performance: from the left, plots of NF versus gate width, optimal width versus bias current and Minimum Noise Figure versus power are reported

device, the lower the value of the gate inductor required, and thus, for a given power dissipation, the lower the noise figure. As a result, gate inductors introduce a term analogous to channel thermal noise in that it decreases for increasing gate width. Thus, we expect the introduction of this term to shift value of optimal width to higher values than predicted for example by equations reported in [5]. This is effect is noticed in figure 4.12, where, from left to right, Noise Figure versus gate width at $2mA$ bias current, optimal noise figure versus power dissipation and optimal gate width versus power dissipation are reported. The most sensible effect certainly is on optimal width prediction which passes from values around 500μ for noiseless inductors to values around 850μ for noisy inductors. It is important to notice that the effect of integration of gate passives becomes more pronounced as the scenario of an on-chip Duplexer([59]) becomes realistic. In this context the possibility to exit use high-Q off chip inductor is missing, so that the potential benefit in LNA noise performance from technology scaling due to increase in the term $\frac{\omega_T}{\omega_0}$ may be offset by the increased noise due to lower gate quality factor. The term χ is intended to account for finite residual S_{11}. It was introduced after noticing that an *Analog Constraint Graph* that performed noise optimization after imposing feasibility constraints on

Bandwidth	Bias current	L_s from 4.4	W_{opt}	L_s from 5.2	S_{11}	NF_{min}
$2.1GHz$	$1mA$	660pH	860μ	500pH	$-22dB@2.1GHz$	$4.55dB@1.9GHz$
$2.1GHz$	$2mA$	550pH	860μ	500pH	$-30dB@2.1GHz$	$2.6dB@2GHz$
$900MHz$	$1mA$	1. 3nH	$2.1mm$	1. 2nH	$-13dB@860MHz$	$4.1dB@800MHz$
$900MHz$	$2mA$	1. 2nH	$2.1mmm$	. 8nH	$-18dB@800MHz$	$2.55dB@800MHz$
$5.4GHz$	$2mA$	. 5nH	800μ	. 64nH	$-13dB@5.1GHz$	$3dB@3GHz$
$5.4GHz$	$3mA$	. 5nH	500μ	. 64nH	$-10dB@5.4GHz$	$2.5dB@3GHz$
$5.4GHz$	$2mA$	. 5nH	280μ	. 64nH	$-13dB@5.1GHz$	$4dB@4.7GHz$
$5.4GHz$	$2mA$	$.5nH$	280μ	$.64nH$	$-8.57dB@5.7GHz$	$3.4dB@4.67GHz$

Table 4.1: LNA noise optimization summary

the source inductors($L_s \geq .5nH$), was seen not to select transistors with minimum channel length for a bias current above $1.8mA$ per branch, as these configurations required too low values of source inductance for any value of device width W_n. As a result, we sought analytical expressions that allowed us to awarely trade S_{11} and NF. The effect of χ is twofold: on one side, low values of χ(say . 6), increase the transonducer gain thus decreasing input referred noise; on the other hand, they require low transit frequencies to be matched , and therefore push feasible solutions toward zones of the $NF - W$ plot dominated by gate-noise. Obviously, this effect will be worse the shorter the channel length, i. e. , the higher the transistor cutoff frequencies. As this tradeoff is being explored while writing, results are not yet stable. However some preliminary results are reported in figure 4.13. It is apparent that at high bias current, optimal noise figure point is obtained matching to a source resistance slightly higher than the source nominal resistance. It is also apparent that the source resistance chosen to be matched has great impact over noise-optimal device width, which decreases as χ is increased. Factor η is now examined. It was introduced to keep into account that, despite the fact that $C_{gs}^{'} = C_{gs} + 2C_{gd} + C_{gb}$ was used to derive simplified input matching equations and noise transfer functions, only $C_{gs}^{'}$ contributes to $DIGN$. Actually, it was found through simulation that a better accuracy is obtained in comparisons $SPECTRE^R$ RF circuit simulator when $\eta = 1$ is assumed. Reasons behind this result are not yet understood at the time of writing.

Table 4.1 summarizes tests performed in order to validate the noise-optimization process described above on bandwidths corresponding to different applications[1]. Noise optimization methodology gives very accurate predictions in the UMTS and GSM bandwidths. In the 802. 11 bandwidth,it tends to overestimate the device optimal gate width W_{opt}, shifting noise versus frequency minimum almost $700MHz$ below the operating frequency. Through simulation, it was verified that for example the $2mA$ circuit, optimized by this methodology at 280μ has a true optimal width of 200μ, giving a minimum noise figure of $3.7dB@5.4GHz$.

Input matching performance also is degraded, mainly due to active devices nonunilaterality at this operating frequencies. About his subject, note that all circuits were simulated excluding non-quasi static effects, due to unavailability of foundry models. Due to the decrease in optimal gate width, this is certainly an effect dominating input matching performance in the WLAN case, at least if the LNA is assumed to be 50Ω differential matched($r_{nqs} \approx 8\Omega$ for a 280μ wide ,$.18\mu$ long device biased at $2mA$ drain current).

[1]Noise optimization process for all WLAN examples did not include common gate device effects; moreover for the last two rows, optimization was performed without imposing source inductor feasibility constraints(i. e. design theoretically requires $L_s \leq 500pH$ to match a 50Ω differential source)

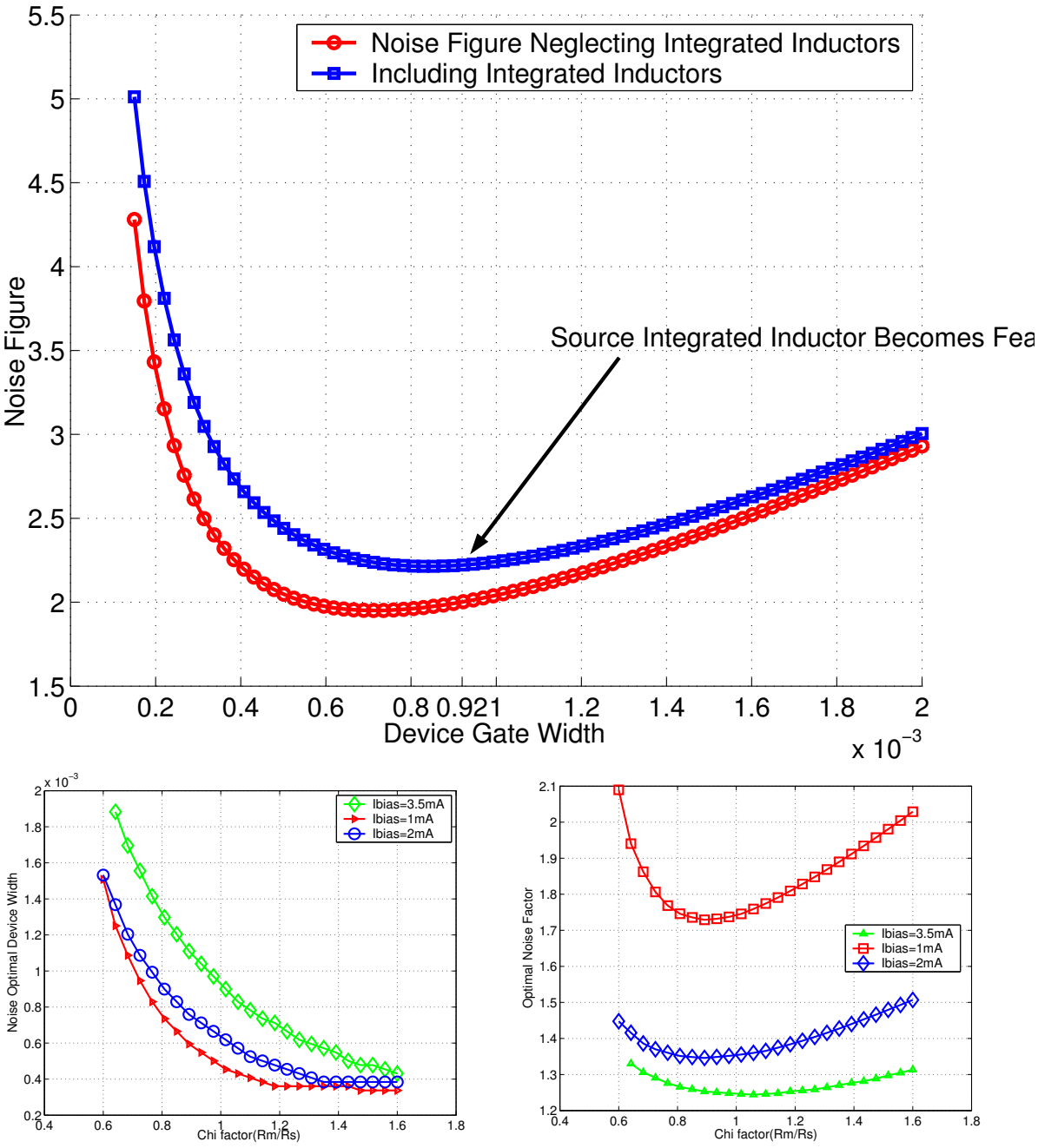

Figure 4.13: Trading off S_{11} for Noise: effect of feasibility range of inductors on selectable widths,optimal device width, and minimum NF as a function χ

4.3 Gain

The effects of overlap capacitances on Low-Noise Amplifiers transfer Function is now discussed. From a qualitative standpoint, classical low-noise amplifier gain equations([6]) neglect the presence of the gate-drain overlap capacitance C_{gd}, and consider device gate in series to source resistance. Finite admittance of the gate drain capacitance instead drains some current to the device transconductance action, reducing its gain. Moreover, a zero is introduced in the transmission path the gate-drain admittance equals the transconductance seen from gate to ground. These considerations may be translated into equations to give:

$$G_m = \frac{\omega_T}{j2R_s\omega_0} K(j\omega_o/s_Z + 1)\frac{1}{1+\frac{j\omega_0}{\omega_C}} Z_L(\omega_0) \tag{4.29}$$

$$\approx \frac{\omega_{T_{eff}}}{j2R_s\omega_0} K(j\omega_o/s_Z + 1)\frac{1}{1+\frac{j\omega_0}{\omega_C}} Z_L(\omega_0)$$

$$\omega_C = \frac{gm + gm_b}{C_{gs} + C_{gd} + C_{gb} + C_{db}} \approx \omega_{T_{eff}}/2 \tag{4.30}$$

$$s_z \approx \frac{gm}{C_{gd}} \tag{4.31}$$

$$K = \frac{Y_{gg}}{Y_{in}} = \frac{RC_2 j\omega_0}{-RR_2CC_2\omega_0^2 + (R(C+C_2) + R_2C_2)j\omega_0 + 1} \approx .5 \tag{4.32}$$

where C and R have been previously defined, while $C_2 = \frac{C_{gs}}{1-L_sC_{gs}\omega_0^2} approx C_{gs}$ and $R_2 = \omega_T L_s + R_{Ls}$. Gain is reduced by a factor of about 2 with respect to the case of ideal operation by the factor k. The effect of the zero is instead negligible for most design cases, and will be afterward neglected. This effect is also adequately kept into account by replacing $\omega_{T_{eff}}$ with ω_T in the callical equations, so that we may state circuit in figure (4.2) to be equivalent to circuit in figure(**??**) for what concerns small signal gain calculations as well. Gain is thus controlled by ratio of the operating frequency to the device effective cutoff frequency, and the value of load impedance Z_L. An LC tuned load is used and sized to resonate at the operating frequency ω_0, so that $Z_L)(\omega_0) \approx QL_{load}$. Maximizing L_{load} thus maximizes gain for any given sizing. Maximum selectable value of L_{load} is imposed by common-gate device parasitic capacitance on the output node, $C_{par} = C_{gd} + C_{db}$, inductor self-resonating frequency capacitance C_{srf} and mixer input capacitance C_{mix} to be $L_{load}^{Max} = \frac{1}{\omega_0^2(C_{mix}+C_{srf}+C_{par})}$. Values of C_{mix} ranging from a couple hundred femtofarads to about 1pF may be expected depending on circuit input stage sizing. Maximum values of L_{load} of the order of a $4nH$ are found from this analysis, leading to $Z_L^{MAX}(\omega_0) = 400\Omega$ and to $G_{max} \approx 20$.

4.4 Distortion

Being interested in modeling and automatically sizing an high performance low-noise amplifier, a deep understanding of the sources of third order nonlinear distortion of this circuit is necessary, so that we may estimate the accuracy of a model based on the structure described in chapter 2, as well as understanding suitable strategies for modeling the LNA-Mixer interface. We perform this analysis using Volterra series. As Volterra series needs models for the derivatives of small signal circuit elements, we make use of the detailed model developed in chapter 2. This does not include a model neither for r_{ds}, nor for its derivatives, as it is well known that no accurate physical model for these quantities exists [25]. As a result, we'll neglect this contribution in the following.
Nonlinear products are generated by all of the bias dependent elements in 4.2, i. e. device transconductances and intrinsic capacitances, parasitic junction diodes, and nonlinear load capacitance. We

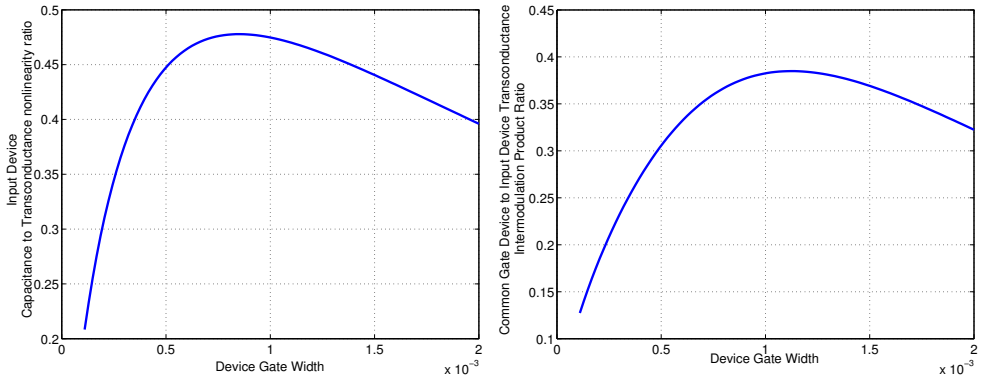

Figure 4.14: Ratio of Nonlinear Gate Source Capacitance(left) and Common Gate to Device transconductance contribution to overall nonlinearity for a $.18\mu$, $1mA$ design @$2.1GHz$

consider first contributions from transconducting device. Using the expressions reported in **??**,it is easy to prove that

$$\alpha_1 = v_{gs}/V_{in} = \frac{1}{1-\omega^2(L_g+L_s)C_{gs}+j2R_s\omega C_{gs}}$$

At this point, two different nonlinear terms come into play: transconductance nonlinearities inject a signal directly to the LNA output(supposing common gate to act as a buffer); while gate-source capacitance nonlinearities are represented by a current source connected in parallel to the gate-source capacitance itself. Note that these two current sources have respectively the same connections of the channel thermal noise generator and of the DIGN noise generator and thus experience the same transfer functions to the output. Third order nonlinear coefficients of intrinsic gate source capacitance $K_2^{C_{gs}}$ and of device transconductance gm_3,generate third order intermodulation distortion given by:

$$I_u^{IM3} = \frac{3}{4}\alpha_1^2(\omega_1)\alpha_1(\omega_2)V_1^2\bar{V_2}\left(\frac{\omega_{T_{eff}}}{j2\omega_1-\omega_2}(1-jQ_L/D)K_2^{C_{gs}}j(2\omega_1-\omega_2)+gm_3\right)\frac{1}{1+\frac{\omega_{T_{eff}}L_s}{DR_s}} \tag{4.33}$$

Nonlinear capacitance to transconductance relative contribution is

$$\frac{gm}{gm_3}\frac{K_2^{C_{gs}}}{C_{gs}}(1-jQ_L/D) \tag{4.34}$$

And depends on frequency through Q_L. We see from figure **??** that device transconductance contributes more than twice the nonlinear capacitance. Second order nonlinear coefficients also play a role in determining third order distortion. Using formulas reported in **??**, we can prove that ratio of this contribution to distortion to distortion arising from third order coefficients evaluates to:

$$R = \frac{1}{3}\frac{\omega_{T_{eff}}(K_1^{C_{gs}})^2(\omega_1-\omega_2)(R_s+r_{nqs}+R_{Lg}+R_{Ls})(1-\frac{jQ_L}{D})+gm_2^2R_{Ls}}{\omega_{T_{eff}}(1-\frac{jQ_L}{D})K_2^{C_{gs}}+gm_3}$$

And may be shown to be lower than . 05 in all practical cases. As a result, we'll neglect third order distortion due to interactions of second order nonlinear generators in the following.
Cascode device and transconducting device drain-bulk junction capacitance contributions will now be evaluated. At this purpose, recall that as $C_j(V) = \frac{C_{j0}}{(1+\frac{V}{\Psi_0})^X}$(where obviously C_{j0} is the zero bias junction

capacitance,$V \leq 0$ the total applied voltage and the grading coefficient X satisfies $X \approx .5$ for submicron technologies) to calculate capacitance nonlinear coefficients:

$$C_1^j = C_j \frac{-X}{\Psi_0(1+\frac{V}{Psi_0})} \tag{4.35}$$

$$C_2^j = C_j \frac{X(X+1)}{\Psi_0^2(1+\frac{V}{Psi_0})^2} \tag{4.36}$$

Voltage at drain of the transconducting device may be expressed as

$$V = \alpha_1 V_{in} \frac{gm}{gm + gm_b + j\omega(C_j + C_{gs} + C_{gd})} = G_1 V_{in} \tag{4.37}$$

This term gives a contribution to the overall distortion given by

$$I_u^{IM3_{CG}} = (C_2^{gs} + C_2^j) G_1(\omega_1)^2 G_1(\omega_2) V_{in}^3 (2\omega_2 - \omega_1) \frac{1}{1+\frac{j(2\omega_2-\omega_1)}{\omega_C}} \tag{4.38}$$

Where ω_C is the previously defined $3dB$ angular frequency of common gate device acting as a current buffer. To evaluate the relative contribution of nonlinear drain-bulk capacitance to overall distortion, we may take the ratio of 4.33 to 4.38 and find

$$\frac{gm_3}{(C_2^{gs} + C_2^j)(2\omega_1 - \omega_2)} (1 + j\frac{\omega_1}{\omega_C})^2 (1 + j\frac{\omega_2}{\omega_C}) \tag{4.39}$$

Equation 4.39 predicts values of the order of 100 when typical design values are substituted. As a result, distortion from nonlinear junction capacitance will be neglected afterward. Distortion from common gate device transconductance is known to be negligible as long as frequency of operation is much lower than $\omega_C \approx \omega_{T_{eff}}/2$. For low power designs this condition may fail to hold.
Third order nonlinearity from common gate device may be expressed as

$$G_1(\omega_1)^2 G_2(\omega_2) \frac{(gm_3) j\frac{2\omega_1-\omega_2}{\omega_C}}{1+\frac{j(2\omega_1-\omega_2)}{\omega_C}} \tag{4.40}$$

If ratio of 4.40 to 4.33 is taken,

$$R_3 = \frac{1}{(1+\frac{j\omega_1}{\omega_C})^2(1+\frac{j\omega_2}{\omega_C})} \frac{j(2\omega_1-\omega_2)}{\omega_C} \tag{4.41}$$

A plot of this quantity is reported in figure 4.14. Although insignificant for low current levels , its contribution rises to about . 4 times the input transconductor contribution for a 1. 5mA bias current. Preceding analysis shows that LNA input stage third order nonlinear distortion is generated by input device nonlinear capacitance and transconductance, and common gate device transconductance. When only these three effects are considered,equations 4.38 and 4.33 may be combined:

$$k_3(\omega_1,\omega_2) = \frac{3}{4}\alpha_1(\omega_1)^2\alpha_1(\omega_2)((\omega_{T_{eff}}(1-jQ_L/D)K_2^{C_{gs}} j + gm_3)\frac{1}{1+\frac{\omega_{T_{eff}} L_s}{DR_s}} + \tag{4.42}$$
$$(\frac{1}{(1+\frac{j\omega_1}{\omega_C})^2(1+\frac{j\omega_2}{\omega_C})})\frac{jgm_3(2\omega_1-\omega_2)}{\omega_C}))\frac{1}{1+j\frac{((2\omega_1-\omega_2)}{\omega_C}}$$

We may use this analysis to forecast accuracy properties of behavioral models based on the structure assumed in chapter 2. If only distortion due input device transconductance is considered, 4.43 gets

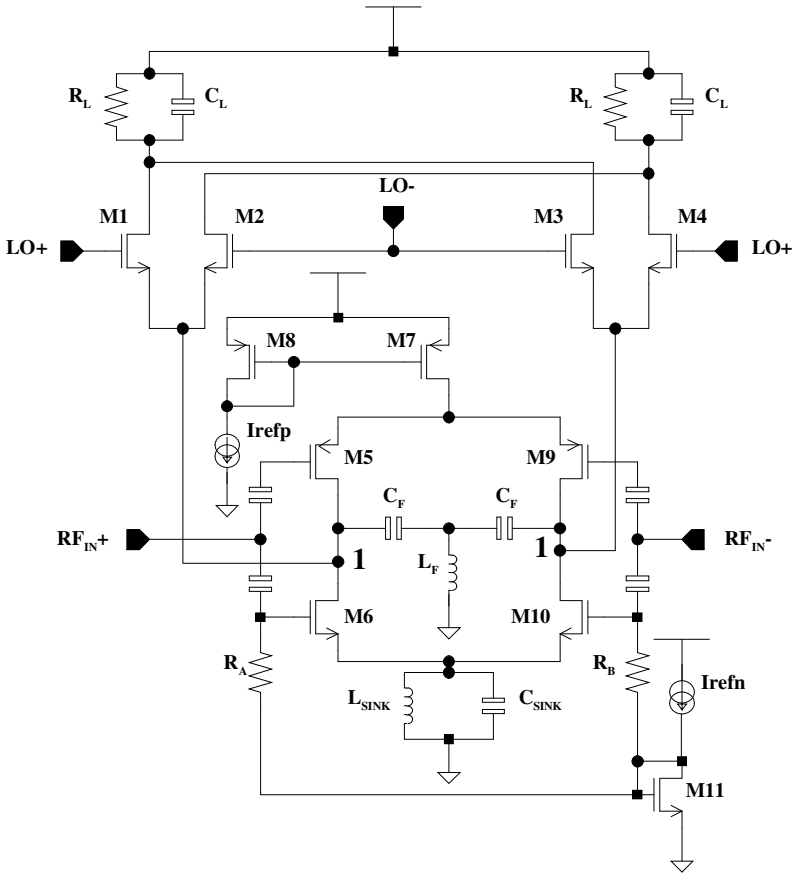

Figure 4.15: Simplified schematic of the mixer proposed in [2]

the appearance of a filter-memoryless nonlinearity-filter system. Contributions of common gate device and gate-source input device capacitance, through terms $1/(1+\frac{j\omega}{\omega_C})$,$Q_L$ and D, introduce frequency dependence in the otherwise purely polynomial nonlinear term. As these factors cause relative variations of up to $\pm 15\%$ in the Volterra Kernel, when frequency is varied in the $1.9GHz \rightarrow 2.1GHz$ range, we cannto expect models developed models to show accuracy better than this.
Final contributor to distortion from the LNA is in the nonlinear load tank, due to nonlinearity in Mixer input capacitance. Large signal LNA output acts to change the value of output tank nonlinear capacitance. As load inductor is assumed perfectly linear, this leads to a time dependent shift in output tuning frequency, and ,as a consequence, to a change in LNA gain. Being signal dependent, this term is a distortion contribution. Intuitively, we expect this distortion term to be higher the higher the LNA gain(higher signal amplitudes), and, for any given gain, the higher the output tank quality factor Q(steeper frequency response and higher output impedance at resonance). The intermodulation kernels of this tank may be expressed as in [49] in terms of the nonlinear capacitance coefficients. In this case, both C_1 and $C_2 \neq 0$, and the third order kernel,neglecting multiple interactions of second order terms, becomes([49])

$$k_3(\omega_1,\omega_2) = \frac{3}{4}C_2(2\omega_1-\omega_2)G_{LNA}(\omega_1)^2 G_{LNA}(\omega_2)Z_L(2\omega_1-\omega_2) \tag{4.43}$$

. Care must be used to get a reasonable expression for C_2. At this purpose, we need to get some greater insight in the load mixer architecture. Refer to the mixer schematic reported in figure (4.15). As

reported in [51],[50], this topology uses a complementary input stage to achieve a higher conversion gain for a given current dissipation (assuming square LO dive, gain may be expressed as $(gm_n + gm_p)\frac{2}{\pi}R_L$, as opposed to the value $gm\frac{2}{\pi}R_L$ valid for a standard Gilbert Cell), while at the same time setting a different current in the switching pairs and in the transconducting devices in order to minimize second order intermodulation distortion and flicker noise([50]). The drawback is that due to the presence of two different devices, an extra drain-bulk parasitic junction is introduced by the p-MOS, while the impossibility to interdigitate the switching pairs and the transconducting devices adds capacitance to the tail node of the switching pairs, potentially increasing distortion([52]). Moreover, due to increased input pair transconductance and decreased switching pair bias current, an higher gain G

$$G = -\frac{gm_n + gm_p}{n \cdot gm_{sp}} \frac{1}{1 + \frac{\omega_0}{\omega_C}} \tag{4.44}$$

,where $\omega_C \approx \frac{n \cdot gm_{11}}{C_{db_{10}} + 2C_{db_1} + C_{db_6} + C_{gs_6} + 2C_{gso_1} + C^i_{gs_1}}$)is experienced from gain to drain of M_{10}(see **??**) and the LO drive was again supposed ideal so that only one of the switching pairs devices conducts for each instant. This high value of gain increases Miller effect through the input stage, thus increasing mixer input capacitance and limiting LNA maximum gain. Moreover, it increases nonlinearity of the receiver chain as explained below. At this point, we assume the LO drive of switching devices may be assumed as an ideal square wave, and that no capacitance is present at the tail node of the quad(i. e. node 1 in figure **??**) so that switching instantaneously commutate and we may use for their nonlinear coefficients of capacitances and currents values derived from DC characteristics. Now,notice that the nonlinear portion of load capacitance may be re-expressed as (C^{CG}_{db} represents LNA common gate device parasitic junction capacitance)

$$C_{out} = C^N_{gg} + C^P_{gg} - (C^N_{gd} + C^P_{gd})real(G) + (C^N_{gdo} + C^P_{gdo})(1 - Real(G)) + C^{CG}_{db}$$

. Where usually the first and the third term dominate .
Due to the large input swings and gain, input stage transistors spend a considerable fraction of the period in nonstauration. This would require the use of physical models for capacitances valid in all regions of operation, including out of the saturation region. Although models of this kind exist in literature, they were not considered in this work. Instead, we supposed $C_{gx} = f_x(\frac{V_{gs}}{n}, V_{ds})$ where the subscript x may assume the value g or d for gate or drain,f is a smooth function and n the subthreshold slope. Circuit topology further imposes $V_{ds} = V^{DC}_{ds} + G * v_{gs}$, and $v^N_{gs} = v^P_{gs} = V_{in}$. Then we have $\Delta C_{gd} = (\frac{\partial f}{\partial V_{gs}}\frac{1}{n} + \frac{\partial f}{\partial V_{ds}} * G)v_{gs})$. Considering second order variations, and using constitutive equations, we find:

$$C_2 = C_{gdo}\frac{gm^N_3 + gm^P_3 + (-gm^{SP}_3 + j(2\omega_2 - \omega_1)(-C^{gs_1}_2 + C^{db_1}_2)) * G^3}{gm^{SP}n} + C^{db}_2 - \\ -\Delta(\frac{\partial^2 f_D}{\partial^2 v^2_{gs}} + G^2\frac{\partial^2 f_D}{\partial^2 v^2_{ds}}) + \Delta(\frac{\partial^2 f_G}{\partial^2 v^2_{gs}} + G^2\frac{\partial^2 f_G}{\partial^2 v^2_{ds}}) \tag{4.45}$$

Where symbol Δ is used to remark the third order nonlinear coefficient cancellation that happens between N and P device in the first stage. Equation 4.46 makes evident that distortion from load nonlinearity depends in a complex fashion on the design parameters. First of all, contributions from device capacitances appear as differences in the values of a same function f ,calculated for different values of its arguments. Assuming f is exactly the same for P and N transistors, cancellation is obtained when $V^N_{ov} = V^P_{ov}, V^N_{ds} = V^P_{ds}$ and at the same time $C^N_{gg} = C^P_{gg}$. Neglecting the V_{ds} term, we see that the first condition is equivalent to $(\frac{W}{L})_N = (\frac{W}{L})_P I_N I^P_0/(I^N_0 I_P)$, while second condition is expressed as $W^N L^N = W_P L_P$. If $L_N = L_P$ is assumed, symmetry constraints become $(\frac{W}{L})_N = (\frac{W}{L})_P, I_N I^P_0/(I^N_0 I_P) = 1$, so that input transistors must be sized so that their bias currents are in inverse ratio with respect to their

mobilities. Obviously there will unavoidably be some difference in f for N and P transistors, leading to inaccuracies in the previous condition. Moreover, equation 4.46 predicts a nonzero distortion floor, even when first order capacitive terms cancel out. This floor is due to the transconductance nonlinearities of input and switching devices which do not cancel, and act changing the small signal gain between gate and drain and thus the input referred Miller capacitance, and is anlyzed in the rest of the section.
Now,dependence of this contribution on design parameters is discussed.
For any given current drive,decreasing the gate width of input transistors results in both increasing the gate-to-drain small signal gain G and nonlinear coefficient gm_3, so that C_2 rises dramatically;increasing the switching devices gate width on the other hand increases their nonlinear coefficient gm_3, but at the same time reduces G. Depending on the particular sizing, this may lead to an increase or o decrease of distortion, according to what effect dominates. Relative magnitude of overall load distortion to distortion from the transconducting part will now be evaluated for a real design. Taking the ratio of 4.43 to 4.43 4.46 is found.

$$R_4 \approx \frac{gm^3 \cdot Z_L^2(\omega_1) Z_L(\omega_2) C_2(2\omega_1 - \omega_2)}{gm_3 \frac{1}{(1+\frac{\omega_{T_{eff}} L_s}{DR_s})(1+\frac{j\omega_0}{\omega_C})}} \tag{4.46}$$

Where we considered both cascode device and transconducting transistor nonlinear capacitance to contribute about 25% to overall distortion(this is a slightly pessimistic approximation) to simplify expressions. In order to get an estimate for the value of C_2, recall that in order to improve third order linearity of the voltage to current conversion, mixer input transistors are biased at overdrive voltages of approximately . 3 V, corresponding to inversion coefficients about 50. For a minimum channel length device and a $2mA$ driving current in the given technology, this corresponds to gate width of 30μ and gives a value of C_2^{gs} of $-1.8 \cdot 10^{-16}, gm_3 = 6.8 \cdot 10^{-3}, gm = 12mS$. For the PMOS pair, biased with $1mA$ current a similar inversion coefficient is obtained for $W_n \approx 40\mu$,resulting in $gm = 6mS$. The switching pairs typically have much lower inversion coefficients and current, and are not minimum sized devices(this for matching and flicker noise reduction reasons). For a $W = 400\mu L = .35\mu$ device biased at $1mA$ drain current, $gm_3 = .43, gm = 18mS$ giving $G \approx 1.3$. Notice that we consider only one of the switching pair devices to carry the whole difference in N and PMOS branch currents, supposing an ideal LO square wave drive. For an LNA on the other hand, we may suppose $I_{bias} = 2mA, W_n = 500\mu, L = .18\mu$ and find $gm = 33mS, gm_3 = 1$. Last, recall that for common gate device $V_d \approx V_{dd} = 1.8V$ to find $C_2^{db} = 1.5 \cdot 10^{-15}$. C_2 evaluates to $94 \cdot 10^{-15}$, with switching pairs nonlinearity contributing more than 90 % of the total, and the remaining almost completely due to transconductor nonlinear gain. Notice that due to the frequenct roll-off of G in real designs the relative contribution of the switching pair term will decrease. If now we suppose $L_{load} = 2.8nH \rightarrow Z_L \approx 280\Omega$, we find $R_4 \approx 3$. Thus, distortion due to the interface may be as high as distortion due to the transconducting load even if the mixer is designed to meet stringent linearity requirements as a stand alone-block. It has been verified experimentally that on randomly generated mixers, R_4 as large as 10 may be found. Thus, interface distortion is a critical component in integrated low-noise amplifier distortion analysis.

4.5 Current Reuse LNA

A topology similar to the one we discussed above is the current reuse,complementary topology discussed in [11] and reproduced in figure 4.17 . This topology has been reported to give better gain and linearity performances, while exhibiting noise figures similar to the ones produced by the NMOS only counterpart. The design of this low-noise amplifier topology presents a number of difficulties. First, the stability of the bias point itself is reduced with respect to the unipolar counterpart. This is due to the presence of high impedance nodes ($OUT-$,$OUT+$) that demand the use of a common-mode feedback loop. Second, a higher number of devices contributes to noise in this topology. Great care must then be taken in sizing the devices themselves in order to minimize their contributions. Design is further complicated

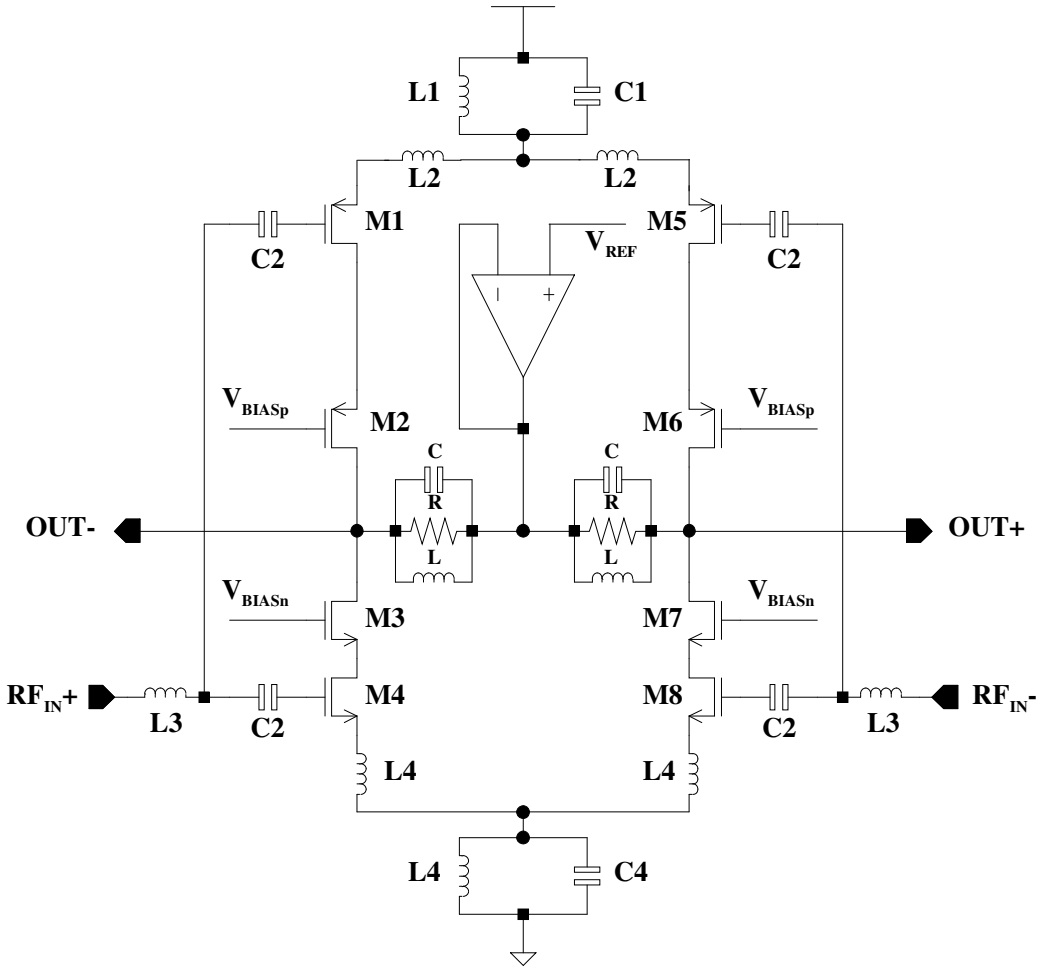

Figure 4.16: Current Reuse Low Noise amplifier Schematic

by the additional number of degrees of freedom presented by this topology(two degeneration inductors, two devices). Following the pattern developed for NMOS inductively degenerated LNAs, exact input matching equations are derived for this amplifier topology. Gain,noise and linearity performances are then analyzed.

4.5.1 Input Matching

Equivalent small signal circuit of current reuse low-noise amplifier is shown in figure **??**. Impedance seen from gate of device M_1 to may be seen as $Z_{in} = Z_N \parallel Z_P$. Each of Z_n and Z_P is in turn composed of a Miller RC branch and a component due to the intrinsic device only. For each RC branch the equations reported for inductively degenerated NMOS low noise amplifier continue to hold, when expressed in the full form(i. e. without the approximation $gm \cdot \omega_0 L_s \approx \omega_0 C_{gs} R_s$). We want to develop an analytical expression for one of the source degeneration inductors and the gate inductor whenever both the devices and one of the degeneration inductors have been sized. First of all consider the completely sized branch: we may calculate the exact input impedance of this branch,Z and then convert it into an admittance Y by taking its reciprocal. For all practical cases, Z will be a capacitive admittance, so that the sized branch may be represented by a shunt RC branch. At this point, look at the almost-sized branch. This branch is perfectly similar to the small signal equivalent circuit of an inductively degenerated NMOS amplifier, except that C becomes now a function of degeneration inductance itself(this happens because of the aforementioned lack of validity of the input independence approximation). An important observation may anyway be made: first of all, notice that as $R_s \approx Re(Z_N) \parallel Z_P$), real part of impedances synthesized by the single branches will necessarily be larger than in the unipolar case. This effect is strengthened by the increased capacitance which is likely to be seen between the transconducting devices gate and ground, which implies a reduction of the gate inductance value and, for any given inductor quality factor, an increase in the fraction of input resistance synthesized by the active devices through local feedback and non-quasi static effects. We'll return later on his argument. For now, we limit to suppose that the per-branch synthesized input resistance increases by a factor $H \approx \frac{R_{sized}}{-R_s + R_{sized}}$, and introduce the new expression

$$G^x_{casco} = \frac{-gm/(1 + jQ^x_L/H) + j\omega_0 C_{gd}}{(gm(2n-1)/n + j\omega_0(C_{gs} + C_{db} + C_{gd}))} \tag{4.47}$$

Where x may be either P or N depending what branch is supposed to have been sized, and R_{sized} is the gate to ground conductance of the sized branch. Now, after considering the parallel of say NMOS Miller branch(that has been calculated using the approximation of 4.47 to sized branch impedance, standard topology equations may be used. One thing is immediate: due the aforementioned increase in the per branch synthesized resistance, transistor may operate at higher transition frequencies, reducing for any given power dissipation the gate-drain overlap capacitance and the related decrease in gain. This may have also beneficial effects on noise, since the input matching constraint is now broken and a wider design space is allowed for optimization.

4.6 Gain

Now, consider small signal gain of low-noise amplifier acting as a transconductor. It is reported in literature[11] that this topology features a doubled gain per unit current consumption with respect to the single transistor counterpart. This is not the case. Analysis reported in [11] is based on the observation that the two transconducting devices operate in parallel for small signal, so that they have a $G^{eq}_m = gm_n + gm_p$. Unfortunately, input matching considerations are neglected in this work. Suppose perfect input matching is realized at the operating frequency ω_0, so that a $R_{in} = R_s$. Current flowing into the positive lead of signal source V_1 is thus $I_{in} = \frac{V_s}{2R_s}$. This current undergoes a partition at the

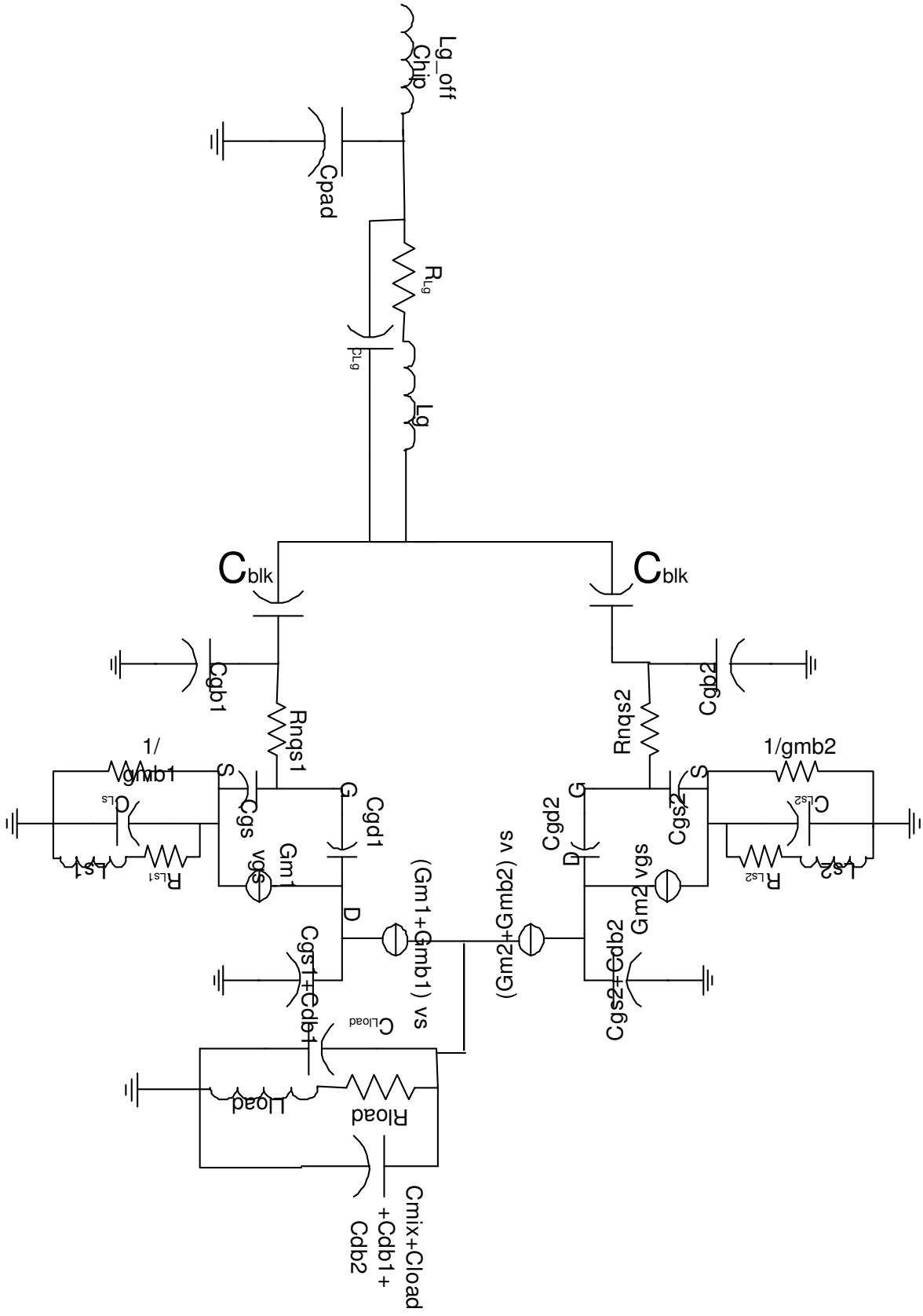

Figure 4.17: Current Reuse Low Noise Amplifier input stage equivalent circuit

common gate of M_1 and M_2, and is then recombined in the output node after each has been scaled respectively by $k_P\frac{\Omega_T^P}{\omega_0}$ and $k_P\frac{\Omega_T^P}{\omega_0}$, where K_N is another current division factor,analogous to the K defined in NMOS low noise amplifier schematic and accounting for the current division between active and Miller branch. Overall transconductive gain at the operating frequency may then be expressed as

$$G_m = \frac{1}{2R_s}(Z_pK_N\frac{\omega_{T_{eff}}^P}{j\omega_0} + Z_nK_P\frac{\omega_{T_{eff}}^N}{j\omega_0})\frac{1}{Z_n+Z_p} \tag{4.48}$$

In case perfect symmetry is obtained amongst the branches($Z_n = Z_p, \frac{\omega_{T_{eff}}^N}{\omega_{T_{eff}}^P} = 1$), then the previous expression reduces to:

$$G_m = \frac{1}{2R_s}K\frac{\omega_{T_{eff}}}{j\omega_0}$$

Which would be exactly equal to the single transistor counterpart, if degradation factor K were the same. However, as a result of higher allowed device transition frequencies(and thus reduced gate width, and decreased C_{gd}/C_{gs} ratio, devices gain is sligthly improved. Any modification is anyway much lower than the expected 6dB increase.

4.7 Noise

4.7.1 Simplified Input Matching Equations

As already done in the previous chapter, simplified matching equations are now derived for this topology of low noise amplifier in terms of passive quality factors and device parameters. At this consider the impedance seen looking between the gate of device M_1 and ground. From analogy with the NMOS case, this may be immediately written as: $Z_1 = R_1 + jX_1 = r_{nqs} + \omega_T^1L_s^1 + \frac{1-\omega_0^2L_s^1C_{gs}^1}{j\omega_0C_{gs}}$. In all practical cases, $\omega_0^2L_sC_{gs} = \ll 1$ and this may be simplified to $Z^1 \approx r_{nqs} + \omega_T^1L_s^1 + \frac{1}{j\omega_0C_{gs}}$. Performed approximation may be rewritten as $\frac{1}{Q_1}\frac{\omega_TL_s^1}{R_1}\frac{\omega_0}{\omega_T^1} \ll 1$, which supposing $Q \approx 5, \frac{\omega_TL_s^1}{R_1} \approx .9$ evaluates to $.25\frac{\omega_0}{\omega_T}$. The approximation gets worse as operating frequency is increased, but the scaling is not so bad, as usually ω_T is increased as well(recall that ω_T/ω_0 controls amplifier gain). Under these hypotheses, we may write $Z^1 = R^1(1 - j\frac{Q_1}{)}$ and perform a series to parallel impedance transformation to get

$$R_{shunt}^1 = R_1(Q_1^2+1) \tag{4.49}$$

$$C_{shunt}^1 = C_1\frac{(Q_1^2)}{Q_1^2+1} \approx C_1 \tag{4.50}$$

The equivalent circuit is now reported in figure 4.18. Impedance sen looking right from section $AA^{'}$ resonates at ω_0 for a value of L_g given by

$$L_g = \frac{R^2C}{1+\omega_0^2R^2C^2} \approx \frac{1}{\omega_0^2C} \tag{4.51}$$

with $1/R = 1/R_1 + 1/R_2$, $C = C_1 + C_2$. After substituting 4.51,input conductance of the signal source plus tuning inductor may be written as:

$$G_s = Q_g(Q_g+Q_i)/(2Q_gQ_i + Q_i^2 + Q_g^2(1+Q_i^2)))R_s)) \tag{4.52}$$

Where Q_g is the equivalent gate inductance quality factor, and $Q_i = \frac{1}{\omega_0R_sC}$. For any given transistor sizing Q_i is given, and thus input matching equation is found:

$$G_s = \frac{1}{R_1(Q_1^2+1)} + \frac{1}{R_2(Q_2^2+1)} \tag{4.53}$$

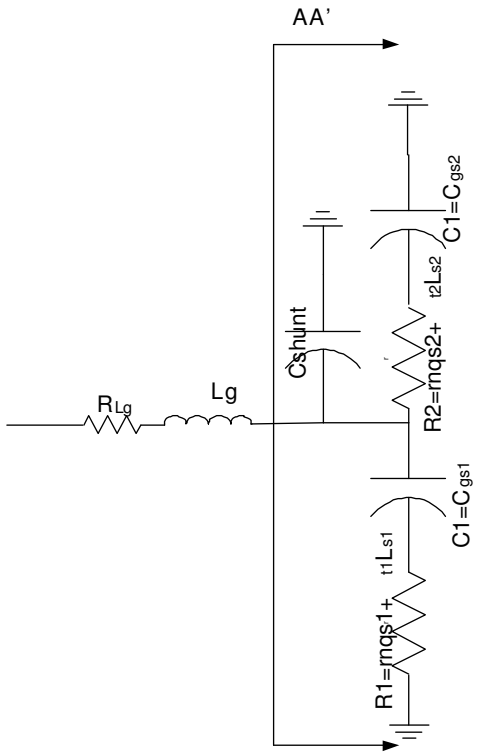

Figure 4.18: Equivalent circuit for simplified input matching equations:notice that the proposed matching equations account exactly for the capacitance shunting one branch, while for the other we may recall to $\omega_{T_{eff}}$ approximation

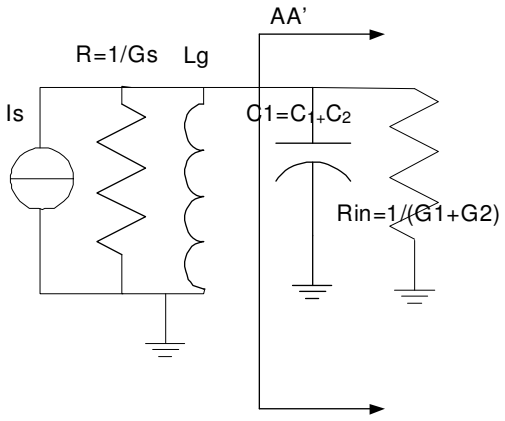

Figure 4.19: Equivalent circuit of the NPMOS amplifier after series to parallel conversion

After substituting expression for Q_2, and doing some algebra, we arrive at:

$$\omega_0^2 R^2 C_2^2 - \omega_0^2 R_2 C_2^2 R^* + 1 = 0 \tag{4.54}$$

Where $R^* = 1/(G_s - 1/R_1)$ is the residual source impedance to match. For the equation to have a real, positive solution , the following constraints have to be verified

$$R^* \geq 0 \tag{4.55}$$

$$\omega_0^2 C_2^2 (R^*)^2 \geq 4 \tag{4.56}$$

$$R_2 \geq \omega_T L_s^{Min} + r_{nqs_2} \tag{4.57}$$

The physical meaning of equations 4.55 and 4.53 immediately realized looking at the expression for the inverse of the real part of the admittance synthesized by a series RC branch for variable R. This is

$$R + \frac{1}{\omega^2 C^2 R}$$

and has a minimum of $2/(\omega C)$ for $R = 1/(\omega C)$. This accounts for 4.53. Finally 4.57 accounts for finite minimum value of inductance realizable in the given technology. It may be re

$$R_2 \geq \frac{I_0 Q}{nV_{th} W (C_{ox} L_{eff} \frac{2(Q(Q+3))}{3(Q+2)^2} + C_{ol}))} + \frac{nV_{th}(Q+2)}{10 I_d}$$

Recalling that $C_{gs2} = W \cdot (C_{ox} L_{eff} \frac{2(Q(Q+3))}{3(Q+2)^2} + C_{ol}) = C_2$,:

$$\frac{-\sqrt{5}\sqrt{C_2 {I_d}^2 n^2 (-4L_s^{min} + 5C_2 {R_2}^2) {V_{th}}^2}) + C_2 n V_{th}(5 I_d R_2 - 2nV_{th})}{C_2 n^2 {V_{th}}^2} \leq Q \tag{4.58}$$

$$Q \leq \frac{+\sqrt{5}\sqrt{C_2 {I_d}^2 n^2 (-4L_s^{min} + 5C_2 {R_2}^2) {V_{th}}^2}) + C_2 n V_{th}(5 I_d R_2 - 2nV_{th})}{C_2 n^2 {V_{th}}^2} \tag{4.59}$$

Recalling $Q = \sqrt{1 + IC} - 1$ one may state this condition to be equivalent to $IC \leq (Q_{max} + 1)^2 - 1$ or finally

$$\frac{I_d}{I_0 (Q_{max}^2 + 2Q_{max})} \leq \frac{W}{L} \leq \frac{I_d}{I_0 (Q_{min}^2 + 2Q_{min})} \tag{4.60}$$

So that for any fixed C_2, there are a maximum and a minimum allowed aspect ratios. This comes out of the fact that for any fixed capacitance, increasing Q decreases gm, thus decreasing the input resistance synthesized by local feedback; while at the same time it increases the non-quasi static contribution. Finally we need to relate Q to C_2 and to channel length L to get a full set of equations. It suffices to use $C_2 = \frac{I_d}{I_0 Q(Q+2)} (C_{ox} L_{eff}^2 \frac{2(Q(Q+3))}{3(Q+2)^2} + C_{ol} L_{eff})$, or

$$L_{eff} = \frac{-12 C_{ol} I_d - 12 C_{ol} I_d Q - 3 C_{ol} I_d Q^2 + \sqrt{3}\sqrt{I_d (2+Q)^3 (3 {C_{ol}}^2 I_d (2+Q) + 8 C_2 C_{ox} I_0 Q^2 (3+Q)^2)}}{4 C_{ox} I_d Q (3+Q)^2} \tag{4.61}$$

Constraints $4.55 - 4.57$ must be imposed when performing noise optimization.

4.7.2 Noise Analysis

Noise analysis of this stage is better performed using classical linear two-port noise theory and generalized Norton's theorem. We start finding the expressions for the input-referred noise sources V_n^1 and I_n^1, of one of the two sections. The circuit is reported in figure **??** for shorted input. Nodal analysis gives

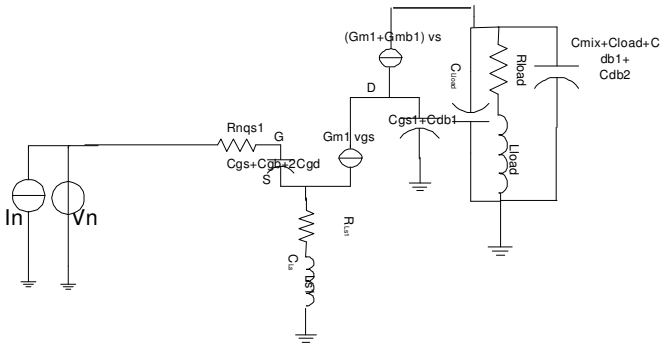

Figure 4.20: Small signal equivalent circuit used to calculate equivalent input noise sourced of each branch of the NPLNA

$$H = \frac{jgm_1\omega L_s^1}{1-\omega^2 L_s^1 C_{gs}^1 + gm_1 L_{sP} s} \approx \frac{jD_1/Q_1}{1+jD_1/Q_1}$$

$$I_{sc} = \frac{V_n}{Z_1}\frac{\omega_T^1}{j\omega_0} = I_{ng}H + I_{nd}(1-H)$$

$$I_{oc} = I_{in}\frac{\omega_T^1}{j\omega_0} = I_{nd} + I_{ng}\frac{\omega_T^1}{j\omega_0}$$

with $D_1 = 1 - \frac{r_{nqs}}{R_s}$. Expression for the input referred noise sources become then

$$V_n = \frac{jZ_P\omega_0}{\omega_T}(I_{nd}(1-H) + I_{ng}(H)) \tag{4.62}$$

$$I_n = I_{ng} + I_{nd}\frac{j\omega_0}{\omega_T} \tag{4.63}$$

and passing to power spectral densities:

$$x = \frac{\omega_0}{\omega_T} \tag{4.64}$$

$$S(V_n)(\omega) = 4KT(R_1^2(1+Q_1^2))x^2(\gamma gm_1\frac{Q_1^2}{Q_1^2+D_1^2} + \frac{\delta\omega_0^2 C_{gs}^2}{5gm}\frac{D_1^2}{Q_1^2+D_1^2} + 2 \mid c \mid \frac{D_1}{Q_1^2+D_1^2}\sqrt{\frac{\gamma\delta}{5}}\omega_0 C_{gs}) \tag{4.65}$$

$$S(I_n)(\omega) = 4KT(\frac{\delta\omega_0^2 C_{gs}^2}{5gm} + \frac{\omega_0^2}{\omega_T^2}\gamma gm + 2 \mid c \mid \frac{\omega_0}{\omega_T}\sqrt{\frac{\gamma\delta}{5}}\omega_0 C_{gs}) \tag{4.66}$$

$$Re(S_c)(\omega) = 4KT(\frac{Q_1(D_1-1)R_1(6C_{gs}^2\delta D_1\omega_0^2 - 30\gamma Q_1 gm^2 x^2 + 5C_{gs}\cdot gm\cdot\omega_0 x(-D_1 x + Q_1))}{30gm(D_1^2+Q^2)} \tag{4.67}$$

$$Im(S_c)(\omega) = 4KT(Re(S_c))\frac{D_1+Q_1^2}{Q_1(D_1-1)} \tag{4.68}$$

Which may be rewritten as:

$$S(V_n) = 4KT\frac{\gamma Q_1^2 + \frac{\delta D_1^2 x^2}{5} + 2D_1 \mid c \mid \sqrt{\frac{\gamma\delta}{5}}x}{Q_1^2\cdot gm_1}\frac{Q_1^2+1}{Q_1^2+D_1^2} \tag{4.69}$$

$$S(I_n) = 4KTgm_1x^2(\gamma + \frac{\delta}{5} - 2D_1 \mid c \mid \sqrt{\frac{\gamma\delta}{5}}) \tag{4.70}$$

$$Im(S_c) = \frac{D_1 + Q_1^2}{D_1^2 + Q_1^2}x^2R_1 \cdot gm(\delta/5D_1 - \gamma Q_1 + \frac{(-D_1x + Q_1)}{6}) \tag{4.71}$$

As contributions from N and P branch are uncorrelated, they add in power spectral density, so that overall PSD is the sum of terms from different branches. Source impedance may now be written as $Z_s = R_s(1 + \frac{Q_i}{Q_g} + jQ_i)$, so that noise factor F evaluates to:

$$F = 1 + \frac{Q_i}{Q_g} + \frac{S_{Vn}}{4KTR_s} + \frac{R_s^2((1 + \frac{Q_i}{Q_g})^2 + Q_i^2)S_{in}}{4KTR_s} + \frac{Re(S_c(\omega_0))(1 + \frac{Q_i}{Q_g}) - Im(S_c(\omega_0)(Q_i)}{4KT} \tag{4.72}$$

This equation may be used together with (4.55-4.57) to perform numerical input stage noise optimization as in [11]. Computations , show that noise factor is minimized when $\omega_T^P \approx \omega_T^N$ and $R_1 \approx R_2 = 2R_s$. These results will be used in the next section.

4.8 Distortion

A significant advantage shown by this LNA topology becomes apparent when third order intermodulation distortion is considered. Current division at the common gate of the transconducting devices acts in such a way that each device bears roughly of the input current. Supposing the average gate admittance of each device is as large at that synthesized by a unipolar counterpart, V_{gs} of each transistor is halved. As a result, contribution of each transistor to intermodulation distortion decreaser by a factor of about 8 and overall distortion goes down by a factor of 4. As a result, mixer distortion is even more important over thet total than in the NMOS case.

4.9 Conclusions

In this chapter, a detailed analysis of two different low noise amplifier topologies was performed, covering in detail problems associated with input matching, noise optimization and distortion. Proposed design equations were discussed and tested over different design cases. Important conclusions were also drawn about third order intermodulation distortion. Using Volterra Analysis, an expression for the nonlinear load due third order intermodulation distortion was derived and analyzed, showing that this contribution is significant in magnitude, and has complex dependence on detailed mixer sizing. This analysis represents the quantitative basis for the automated sizing and modeling effort that was developed in the subsequent chapter.

Chapter 5

Low Noise Amplifier Model

Low Noise Amplifier platform construction is now discussed. First, different analog constraint graphs used to size the NMOS and the current reuse topology are described, and their effectiveness in LNA design space exploration is discussed. Problems such as equations accuracy and conservativity are also faced. Next,problems introduced by LNA-mixer communication are described. Sensitivity analysis is used to calculate variations introduced in predicted linear response when load capacitance is varied; simulation is instead used to further demonstrate dependence of nonlinear response on mixer detailed sizing for a constant input capacitance. Afterwards, a behavioral model which is robust with respect to these effects is proposed and analyzed, and , finally, system level optimization setup and results are discussed.

5.1 Analog Constraint Graphs

As stated in the introduction, platform based design relies on Analog Constraint Graphs to reduce the size of input space I and speed up the convergence of the simulative process. An Analog Constraint Graph may be operatively seen as a set of equations(which for the graph branches and nodes), and an algorithm to be used for solving this set of equations. In general, the set of equations presents a number of unknowns higher than its own cardinality; as a result, some variables will be randomly chosen. As a result, the portion of circuit design space which is explored depends on the chosen set of equations, distribution functions and algorithm. As we'll see, a too tight set of constraints may easily result in biasing the design space exploration towards unintended zones, while a too weak set of constraints could not be useful in accelerating the simulation process.

5.1.1 Client Server Process

Before detailing the algorithms used to generate "valid" circuit configurations, we give an overview of the characterization process flow. As scheduling an Analog Constraint Graph , at least in principle, requires a number of equations and disequations to be solved and checked, as well as some random variable to be generated, the $MATLAB^R$ environment was envisioned to be well suited to this process. Actual simulation steps are instead performed on a Linux machine using the $OCEAN$ environment [19], a proprietary language of $CADENCE$ that allows simulations and data processing to performed in a batch fashion. Data transfer from MATLAB to Ocean was synchronized using a simple semaforic mechanism described in [60]. Software implementing these algorithms has been written in collaboration with F. DBernardinis and F. Vincis,and will be described elewhere([40],[1]).
To minimze characterization time, size of data blocks should be chosen in such a way that time spent simulating be roughly twice time spent generating configurations, so that, expect for the first latency

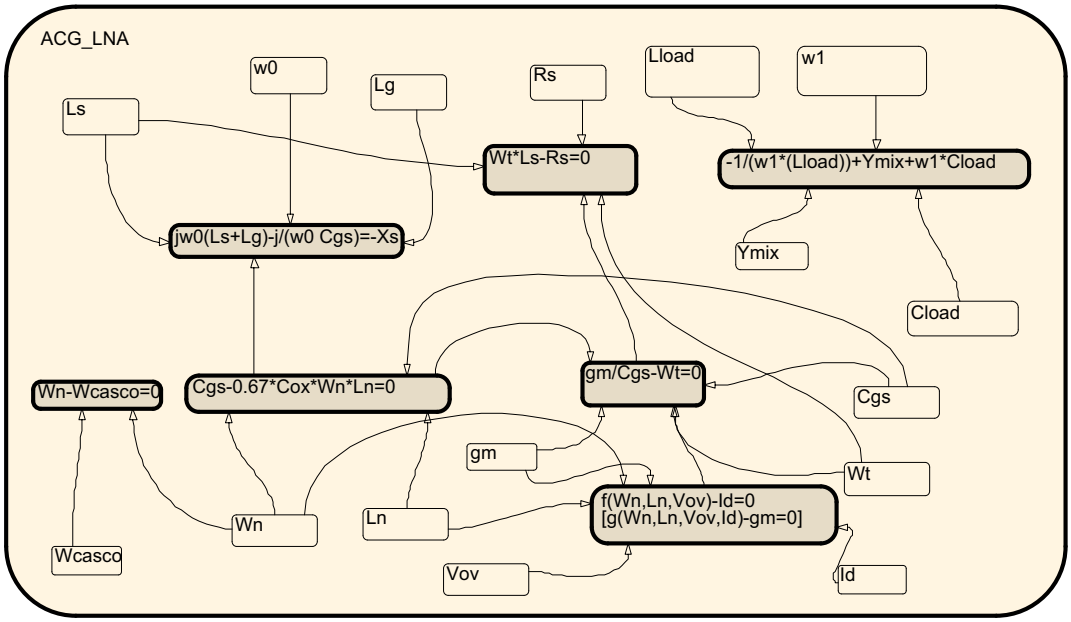

Figure 5.1: First NMOS low noise amplifier ACG

time, characterization process proceeds in pipeline. For very slow ACGs, this condition may be verified for data packets of as few as 10 elements. Data blocks of such a small size impairs statistical classifier convergence properties as described in [40], and thus should be avoided. As a result, speed in generating configurations is a very important metric for an Analog Constraint Graph.

5.1.2 NMOS low noise amplifiers

The first conceived NLNA analog constraint graph is reported in figure 5.1. It tries to impose input matching and and optimize noise figure of the amplifier using approximate equations and random perturbations. Input space for amplifiers characterized by using this ACG is made up of bias current I_d,input transistor channel lenght and gate width L_n and W_n, gate and source inductors L_g and L_s, and load tank passives values, L_{load} and C_{load}. I_d is chosen first, using a uniform distribution of extrema $I_d^{MIN} = .5mA$ and $I_d^{MAX} = 3mA$. Once I_d is known, device gate width must be chosen. As discussed in the previous chapter, an optimal value of this parameter exists , such that amplifier noise factor is minimized for the given current consumption. As observed in [11], this value biases the transistor in the moderate inversion region. The ACG proceeds then as follows:

1. Transistor channel length is sampled summing to the minimum allowed drawn lenght L_{min} of the process, an exponentially distributed random variable, with mean $\lambda = L_{min}/5$. Allowing channel lengths higher than minimum could decrease the gate width necessary to synthesize a given gate capacitance C_{gs}, improving linearity at the expense of noise figure

2. For the given current consumption, and using a strong inversion model, the device aspect ratio corresponding to $V_{ov} = V_{lim}$ s calculated as

$$(\frac{W}{L})^* = \frac{I_d(1 + V_{ov}/V_{sat})}{I_0(V_{ov}/(nV_{th}))^2} \tag{5.1}$$

Where V_{sat} is a model constant accounting for mobility reduction due to vertical and horizontal field chosen equal to 1V , $I_0 = k_n C_{ox}(2.5nVth^2)/2 = 100nA$. V_{lim} is a voltage value taken as the limit between strong and weak inversion, located around $78mV$. Both constant V_{lim} and random V_{lim} schemes were implemented. As explained later, the value of V_{lim} has a strong influence on the explored design space, as it controls the average aspect ratio of the transconducting devices and as a consequence their input capacitance.

3. To avoid restricting excessively the design space, random perturbations are superimposed on the found value. In particular, transi

4. Put $W_n = \frac{W}{L} \cdot L_n$, intrinsic transistor small signal parameters are calculated. For gm numerical derivative of the interpolative characteristic valid in all operating regions reported reported in [11] is used. For capacitance $C_{gs} = 2/3 C_{ox} W_n L_n$ is used. Source degeneration inductance is calculated to be equal to
$$L_s = R_s/\omega_T \cdot X \tag{5.2}$$
, with X gaussian random variable uniformly distributed in [. 9;1. 1]. If $L_s \leq L_s^{Min}$, a failure occurs. This failure is registered in a special variable; and the scheduling process restarts.

5. Gate degeneration inductance L_g is now calculated as
$$L_g = X_1 \cdot (\frac{1}{\omega_0^2 C_{gs}} - L_s) \tag{5.3}$$
, with X_1 random variable independent of X and uniformly distributed in [. 9,1. 1]. If $L_g > L_{gMax}$ occurs, a failure is registered in a special variable(different from the one dedicated to L_s) and the scheduling process restarts. Inspection of 5.3 immediately reveals the importance of V_{lim} in determining the scheduling process failure rate: the higher V_{lim}, the lower (on the average) C_{gs}, the higher L_g.

6. Load is now sized. An equivalent input mixer capacitance $C_{mix}$is sampled through an appropriate distribution(we'll return later on this point). Load inductor is chosen according to equation 5.4.
$$L_{load} = L_{load}^{min} + Max((1 - X_3) \cdot (L_{Loadmax} - L_{LoadMin}), 0) \tag{5.4}$$
Where $L_{Loadmax}$ and $L_{LoadMin}$ are the maximum and minimum allowed load inductors, and are statically determined;while X_3 is an exponentially distributed random variable with mean . 2. Notice the importance of the max function, which avoids the negative tails of the exponential distribution of $1 - X_3$. Choosing a high value of load inductor benefits noise figure and gain; and this is the reason this distribution privileges so much high value of load inductance.

7. As C_{mix} and L_{load} have been chosen independently, the output tank won'resonate at ω_0, unless a linear capacitor of value $C_{lin} = (\frac{1}{\omega_0^2 L_{load}} - C_{mix} - C_{srf}^{Lload} - C_{db}^{Device})$ is added in parallel. If $C_{lin} \leq 0$ results, a failure is claimed and Load sizing is restarted.

Once this step is completed, only the bias circuit must be sized before the LNA may be be simulated.

5.1.3 Bias Circuit

The bias circuit implemented in [11] for this low noise amplifier is a Wilson cascode current mirror. Its operating principle is easily explained: referring to 5.2,device M_4 creates a negative feedback loop that forces $I_d^{M_2} = I_{ext}$. Resistor R on the other hand, sets the value of cascode bias voltage. Source nodes of all the devices are tied to a common potential V_s, coinciding with the source potential of the

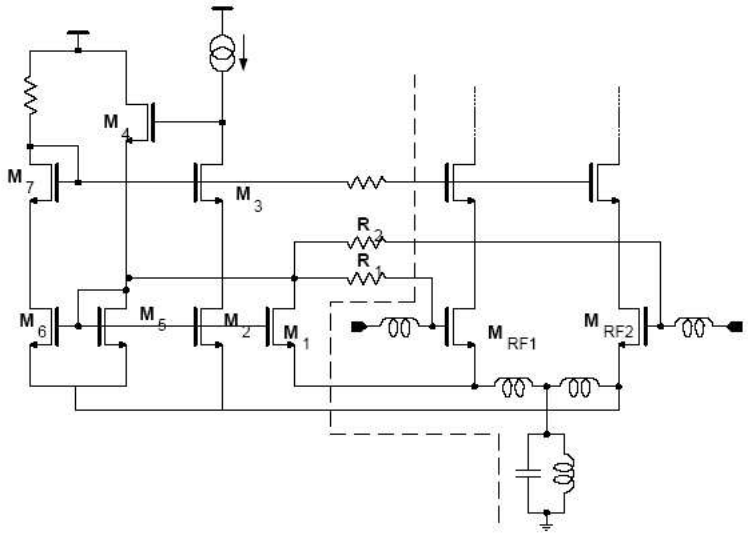

Figure 5.2: Bias circuit of the LNA

transconduting device. This choice increases the mirror accuracy by making it insensitive to DC drops caused by inductors finite quality factors. Supposing $(\frac{W}{L})_6 = (\frac{W}{L})_4 = (\frac{W}{L})_3 = (\frac{W}{L})_5 = (\frac{W}{L})_1/10$, current through R will be I_{ext}, leading to a cascode device gate voltage of $V_{dd} - RI_{ext}$. For the transconducting device drain node it may then be written $V_D = V_{DD} - RI_{ext} - V_{gs}^N$, or $R = \frac{V_{DD}-V_D-V_{gs}^N}{I_{ext}}$. Current in M_1 on the other hand, will be $10 \cdot I_{ext}$. To simplify ACG operation, this circuit has been parametrized in the CADENCE environment in function of transconducting device sizing, so that all of its parameters are implicitly set by the ACG. Value of R was chosen equal to $.55/I_{ext}$, to impose on the drain of the transconducting device an average voltage of $.3V$. Although this voltage is higher than that required for high swing operation, this choice avoids the bias circuit to accidentally put the transconducting transistor in the triode region as a result of equations inaccuracies.

5.1.4 Results from input space sampling

The Analog Constraint graph described above has been used in different characterization processes, leading to a deep understanding of its properties. The most critical parameters in the algorithm are the static bounds imposed on the inductor values L_{min}^g and L_{max}^g, and the average overdrive voltage of the transconducting device V_{lim}. These parameters are actually related through equation 5.3. In the early phase of the work, V_{lim} was chosen deterministically equal to 56mV. To motivate this choice, go back to consider 5.3. For a maximum gate inductance value of $10nH$, minimum C_{gs} needed to achieve series resonance at $2.1GHz$ is about $574fF$s. For a minimum length device this value corresponds to an aspect ratio of about 3126 or a gate width of 540μ. A transistor of such an aspect ratio biased at $V_{ov} = 78mV$, on the other hand,consumes about $4mA$ of current, higher than the allowed maximum bias. Changing the overdrive to $56mV$ gives a static consumption of $2mA$, which is exactly the nominal current per branch reported in [11]. The sheduling algorithm described above generates about ten configurations per minute, achieving a success rate of about 1:1000. Almost all of the failures reported are due to too high gate inductance. This is confirmed by analyzed input space samples. An histogram reporting reporting distribution of L_g for this schedule is reported in figure 5.3. Notice how the distribution is concentrated

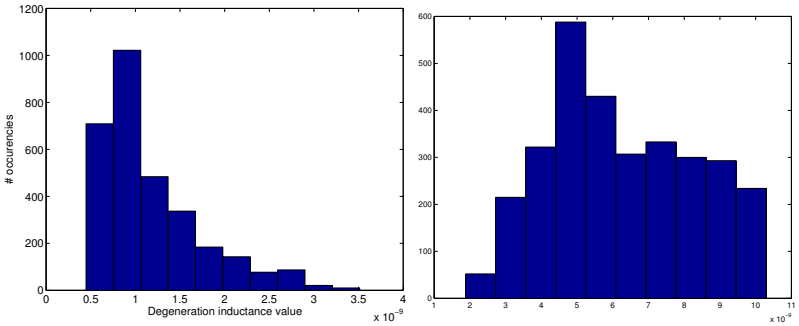

Figure 5.3: Gate(left) and source(right) degeneration inductors distribution after first reported ACG

in proximity of the maximum value of $10nH$. Distribution of source degeneration inductors values, on the other hand, is extremely concentrated around minimum values. When LNA performances are considered, the proposed algorithm achieves S_{11} around $-12dB$, at frequencies dispersed around and an average of $2GHz$ with a standard deviation of about $100MHz$. Considering the very simple model used, this is a success. Similar considerations also hold for the output tuning frequency. Interesting results also regard noise performance. If minimum Noise Figure in the $1.8GHz - 2.2GHz$ frequency range is reported versus power dissipation, an hyperbolic relation becomes apparent. Actually, $Pd \cdot NF$ product appears a random variable with estimated mean $8 \cdot 10^{-3}$ and standard deviation of $2 \cdot 10^{-4}$. Noise Figure minima however averagely appear around $1.8GHz$, leading to a significant loss of in-band noise performance. Recalling from previous chapter that NF , when input stage device alone is considered, is a function of $Q_{in} = \frac{1}{\omega_0 C_{gs} R_s}$ only, it is easily understood that this algorithm avergely overestimates noise-optimal device width by a factor of about $2.1/1.8 = 1.2$. To expand the design space exploration beyond these borders, the scheduling process was modified and a randomly variable V_{lim} was adopted. Namely,value of V_{lim} is now the sample of a uniform random variable, extrema $22mV and 180mV$. Static bounds on inductor values were kept the same. In doing this characterization, we mainly expected to get more information on the $P_d \cdot NF$ product distribution, and on the overdrives effectively allowed by the bound on inductor feasibility. In order to get this information, V_{lim} was enclosed in the output space this time. As V_{lim} is uniformly sampled right after the bias current and channel length, and as a failure-free Analog Constraint Graph would leave sampled variables distribution unchanged in the large numbers limit, observing the distribution of V_{lim} gives us information about the relative frequency of failure of inductor sizing equations for different bias voltages. Notice that increasing the overdrive voltage of the transconducting device not only reduces its input capacitance, requiring higher values of gate-inductance to achieve series resonance; but it also increases its ω_T, so that lower values of source inductance are needed. So, high V_{lim} configurations are more likely to encounter failures on both of the static bounds on the inductors values. The distribution is reported in figure 5.4. As expected, overdrive values higher than $70mV$ are seldom accepted by the ACG, while distribution is rather flat for values below $60mV$. Considering now $P_d \cdot NF$ product, we expected to draw the following information out of this process:

1. Wether a constant value of $P_d \cdot NF_{min}$ product could characterize any bias voltage V_{lim}
2. If the chosen V_{lim} of $56mV$ were the optimal one,i. e. supposing 1 holds, the overdrive with the lowest possible average value of $P_d \cdot NF_{min}$ product

The answer to these two questions is reported in figure 5.5. To make statistical measurements significant, we actually binned V_{lim} values into classes. Two classes are seen to give equivalent results, corresponding

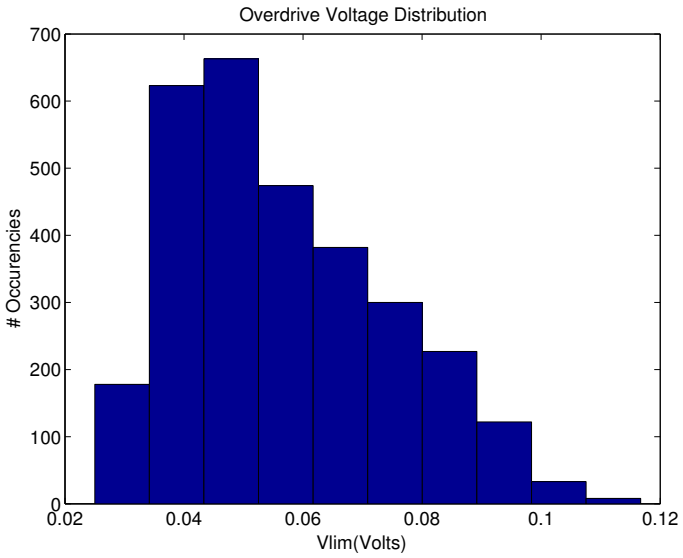

Figure 5.4: Distribution of allowed overdrives

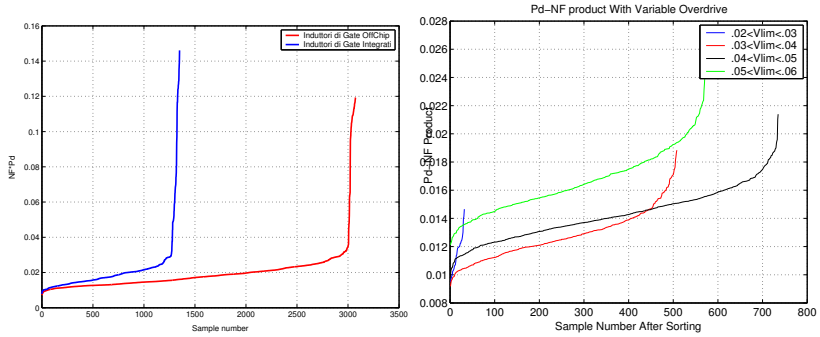

Figure 5.5: Value of $P_d \cdot NF_{min}$ product obtained from running the modified ACG(right) and from running the constant V_{lim} ACG(left)

to $V_{lim} \in [.03, .04]$ $V_{lim} \in [.04, , .05]$, while other design classes show inferior noise performance. Also notice that $P_d \cdot NF$ shows greater variations than for the single-valued V_{lim} case. This is due to the lower number of samples per class used in the variable-overdrive design with respect to the single overdrive case. Results obtained from this class of analog constraint graphs are overall satisfying. The main problem is in the overestimation of noise-optimal gate width, that results of satisfying NF_{min} but of too high $NF_{average}$. Correcting this error is very simple in a design refinement step; moreover, no serious overestimation of block performance is done when taking minimum noise figure as figure of merit as opposed to in band average noise figure, as for an hand optimized design this values are almost coincident(this is true only if the minimum occurs at a frequency close to the nominal operating frequency, otherwise, degradation due to ω_0/ω_T term in the expression of F reported in previous chapter must be accounted for). An alternative Analog Constraint Graph was anyway derived an implemented. Several reasons drove this effort:

- Notwithstanding what was said above, choosing noise-optimal gate width is the main concern in designing low-noise amplifiers. As a result, we did not consider r=satisfying results those leading to an overestimation of these parameters.
- Described Analog Constraint Graph is obtained by means of strong approximations. The first one regards the expression of C_{gs}, in which overlap capacitance, intrinsic portion of C_{gs} bias point dependence and gate-bulk and gate-drain extrinsic and intrinsic capacitances are completely neglected. For the chosen operating frequency and techonology, these effects seem to compensate, resulting in overall quite accurate L_g estimates. This may not happen for any given technology neither for any application. This is not a satisfying property.
- Noise optimization is based on the observation that noise-optimal device width depends on device DC current in such a way as to bias it at the edge of weak inversion region. This is true in the $2GHz$ bandwidth and in the given technology, but certainly is not true in a different scenario. Although the graph may be recalibrated by choosing an appropriate distribution for V_{lim},this random approach is not intellectually satsfying.

5.1.5 Alternative Analog Constraint Graph

The second Analog Constraint graph proposed is based on the noise-optimization process described in [5] and in [11], and extended in the previous chapter. It may be described as follows:

1. Bias current and channel lenght values are chosen exactly as in the preceding ACG
2. In order to choose optimal transistor gate width, equation 4.23 and similars are used together with constraints $L_s \geq L_s^{Min}$ and $L_g \leq L_g^{Max}$. Due to inavailability of foundry non-quasi static model, r_{nqs} is set to 0. Drain excess noise factor and transconductance are calculated using the accurate model described in chapter 2. Gate source small signal capacitance C_{gs} is instead estimated to be equal to
$$C_{gs} = C_{ol}W + \frac{2}{3}C_{ox}WL_{eff} \tag{5.5}$$
The reasons for this choice, that willingly overestimtes the transistor gate-source capacitance, are well understood in light of the results derived in previous chapter. Due to low inversion levels in the transconducting device, intrinsic C_{gs} value is much smaller than predicted from the strong inversion formula. If gate drain and gate bulk capacitance effects are not considered, this leads to both an overestimate of real part of impedance synthesized by the input branch and and to an underestimate of its capacitive admittance. As a result, predicted values of source inductance are very low, while corresponding gate inductance values are too high. This in turn impacts noise-optimal gate width, as we noise from the gate inductor is included in the noise

figure calculation. Equation 5.5 predicts higher values of capacitance and source inductance, reducing this problems. Effects of inaccuracies on gate inductor value prediction are further decreased at the simulative level as described below. Due to the complex dependence of small signal parameters on bias point, an analytical expression for W_{opt} was not found. Instead, it is extracted by numerically evaluating the noise factor $F \approx F_{input} + \frac{\omega_0 L_g}{Q}$ on 100 values of gate width between $W_{min} = 1e-4, W_{max} = 1e-3m$. Inaccuracy on optimal gate width due to discretization evaluates thus to $(W_{max} - W_{min})/200 = 4.5\mu$. This is done very efficiently in $MATLAB^R$ by using vectorized code. For each point, required values of source and gate inductors L_s and L_g are calculated using 5.2 and 5.3. Then, gate widths for which unfeasible inductor values were predicted are discarded, and optimal width is chosen amongst the remaining. If remaining widths are an empty set, a failure is claimed, and scheduling process restarts from bias current selection. Scripts implementing the ACG were written keeping in mind application and technolgy portability besides computational efficiency. As a result, process constants and static bounds on device widths,inductor values and quality factors and current levels are set using an external script; an the same holds for technological constants.

3. Load Sizing is performed as in the previously described ACG

5.1.6 Results from input and output space sampling

Theoretical advantages of this analog constraint graph over the previous are immediately understood. First of all, as an accurate noise optimization process is performed, we expect noise figure minimum frequency to lie averagely much closer to the nominal operating frequency, without need of hand-tuning the script. Second, as noise optimization is performed only after feasibility constraints have been imposed on inductors for all allowed values of gate widths, failures should be greatly reduced, increasing the speed of configuration generation. These expectations have been only partially confirmed by experiments. Noise performance is good. Average minimum noise figure frequency averagely lies at $2GHz$, so that even a conservative metric such as average in band NF is in acceptable ranges. Input matching however gives some problems. Even using 5.5, L_s required to achieve input matching is underestimated. Moreover, imposing feasibility constraints on the inductors before noise-optimization is performed results in giving a greater importance to input matching than to noise optimization, so that no $S_{11} - NF$ tradeoff is available. This is not a realistic situation, especially when an off-chip duplexer(i. e. present industrial standard) is used and reflections may be re-adjusted with the use of an external matching network at low cost. In fact, simulations of designs sized according to this ACG show extremely low S_{11} values(of the order of -20dB). When power levels are increased above $6mW$ however, gm/C_{gs} ratio of transconducting devices becomes so high that none of the allowed transistor gate widths accommodates for $L_s \geq L_s^{Min} = 500pH$ if minimum channel length is chosen. High levels of bias currents and minimum sized devices are therefore seldom generated by this algorithm. As channel lengths are almost fixed to the minimum value, and are chosen before noise optimization and matching constraints are imposed, ACG overall success rate decreases with this choice. Moreover, this upper limit on power consumption of minimum sized devices reflects on noise performances. Figure 5.6 reports a plot Noise Figure versus power dissipation and Channel length distribution versus power dissipation from the last schedule. While the relative frequency of minimum sized devices decreases while power is increased above the $6mW$ limit, measured noise flattens over a $3.2dB$ level while following a hyperbole-like dependence on power for lower current drives. To overcome this limitation, input matching must be transformed into a soft constraint, so that it may be traded off with noise. This can be done by randomly choosing the value of R_s at the beginning of each execution cycle. Assuming a maximum estimated S_{11} of $-12dB = .25$ is tolerated, we get for the synthesized resistance the bounds

$$.6 \leq R_{in}/R_s \leq 1.6 \tag{5.6}$$

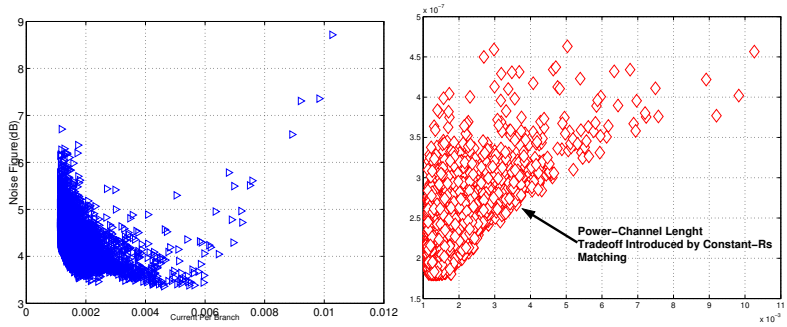

Figure 5.6: Noise-Figure versus power dissipation(left) and Channel Length versus power dissipation(right) from results of the second described analog constraint graph

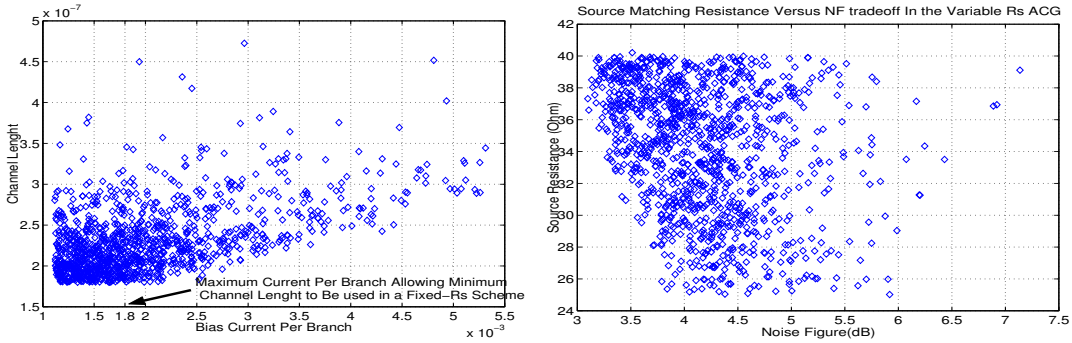

Figure 5.7: Matching resistance versus Noise Figure(Left) and Channel Length versus Power dissipation in a variable R_s scheme

Lower bound is ignored as in this context it may only worsen things. Characterization has then been repeated synthesizing input resistances chosen

$$X \cdot R_s$$

with X random variable uniformly distributed in the $[1; 1.6]$ interval.
Results from this schedule are being analyzed in the time of writing. Results appear promising. In fact, a lower minimum in band average NF is apparent in 5.7 ($\approx 3.05dB$ versus $3.4dB$). Also, maximum current per branch allowing a minimum channel length device to be used has passed from $1.8mA$ to $2.4mA$ with a $+50\%$ improvement with respect to previous solution.

5.1.7 Accuracy and conservativity of the used sets of equations

To understand what degree of correspondence exists between the estimated device parameters and the simulated ones, blocks of DC simulations were run. Notice that accuracy is not a major virtue for an Analog Constraint Graph, as the really important metric is conservativity(i. e. the ACG must be such that "good" configurations are not discarded, not that bad configurations are not generated). For example, the noise-optimization based analog constraint graph where input matching to a constant-value source resistance is imposed is an example of an accurate and yet not conservative ACG. Conservativity of an ACG can be ameliorated[?] by transforming equality constraints,i. e. constraints of the kind

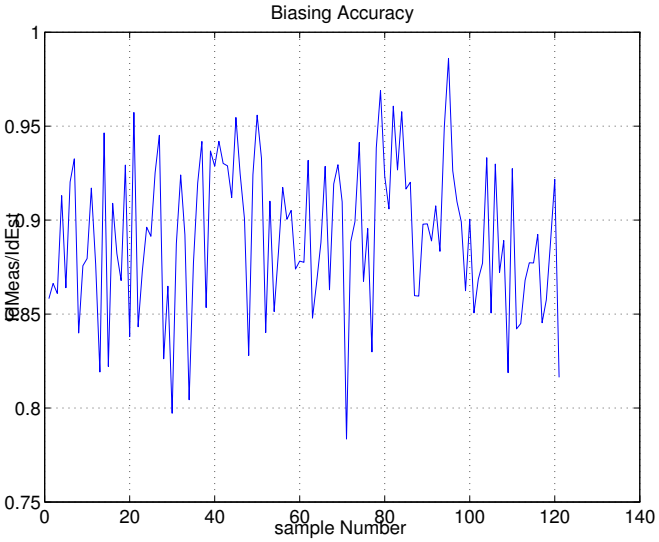

Figure 5.8: Simulated to estimated LNA bias current ratio

$f(\vec{x}) = 0$ and $\vec{x}$ is a vector of input space variables, into noisy-equality constraints, where an error is willingly introduced in the equation so tha mathematical formulation is changed to $f(\vec{x}) = n$, or $f(\vec{x}+\vec{N}) = 0$(here n is a sample of a random variable, while $\vec{N}$ is a k-dimensional random vector, and in order to get for these expression to be equivalent, $\nabla f \cdot \vec{N} = n$ must hold)[40]. Estimating the statistical properties of the committed errors may be useful for deriving the statistical properties of $\vec{N}$. In the case of this NMOS low noise amplifier, degree of accuracy of device small signal parameters has been analyzed reported in chapter 2;and we only need to understand the errors introduced by the bias circuit in the DC current imposed to the transconducting device. A plot of this quantity is reported in figure 5.8.

5.2 Current Reuse Amplifiers

5.2.1 Bias circuit

For the current-reuse amplifier, biasing is critical. This is both due to the high number of stacked transistors, and to the presence of a high impedance node where the two cascode sections are connected "head to head". This node must be stabilized through an appropriate common mode feedback loop;current consumption of this loop adds directly to LNA power consumption and should be kept minimal. Moreover, decoupling capacitors are needed to give separate DC values to both the gate of the transconducting NMOS and PMOS, which are shorted at DC. Feedforward bias circuit is a built of two coupled Sooch Cascode Current Mirror(see figure 5.9) [28], sized for maximum dynamic range. Notice that the N section of the circuit acts as a reference current for the P section,so that overall number of branches is minimized while matching between the N and P current is maximized. As in the N counterpart, source node of all the devices are not tied to ground but to the source of a reference input device to compensate for DC voltage drops due to inductors finite quality factor. This circuit is sized exactly as reported in [28], except for the level shifting devices, which have an aspect ratio equal to 1/6 of that of the common source devices(as opposed to the value of 1/3 reported in [28]) to account for body effect, and having a conservative level shift. Due to the relatively low output impedance of

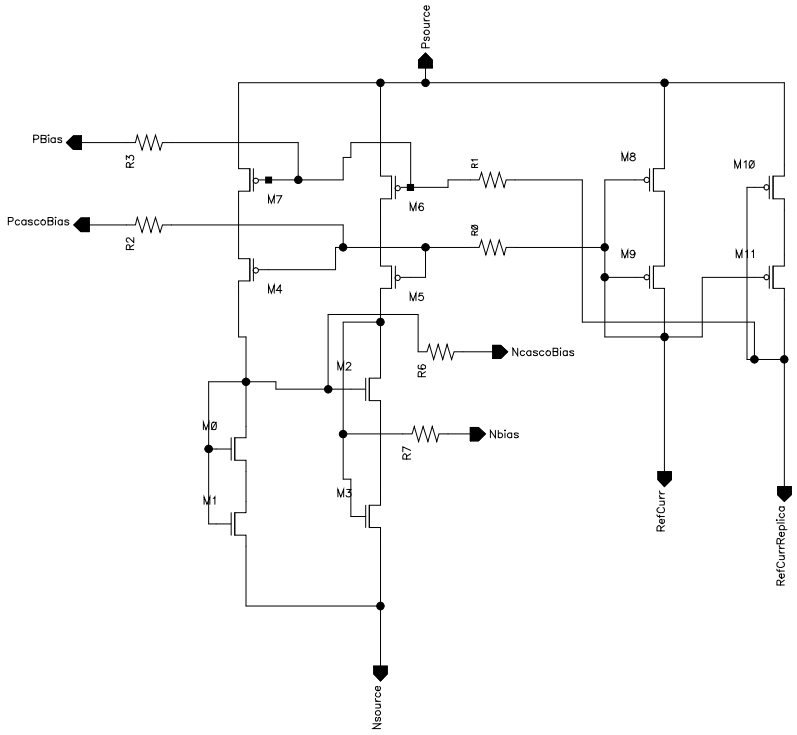

Figure 5.9: Current Reuse Amplifier Bias Circuit

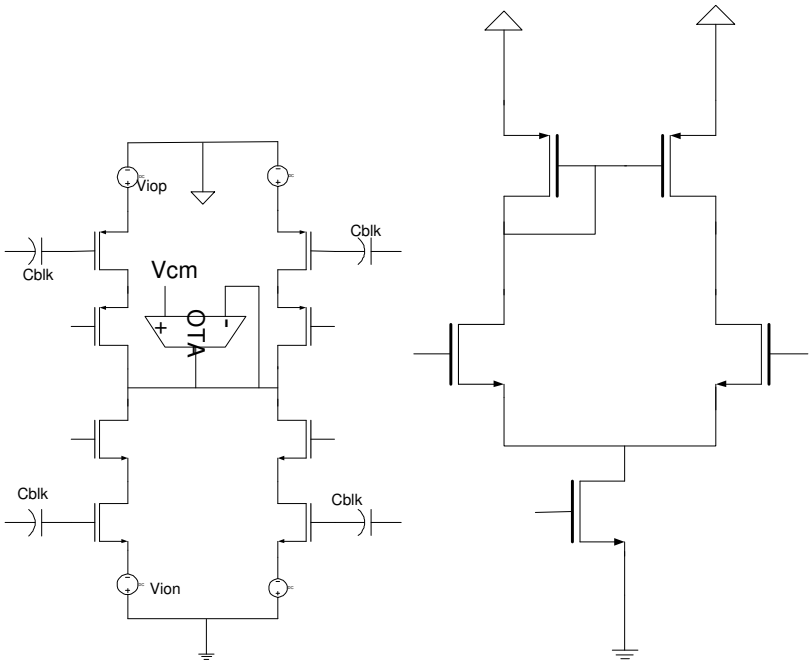

Figure 5.10: Common Mode Feedack Circuit, including bias decoupling capcitors offset generators shown with worst case polarity(left) and OTA schematic

the loaded amplifier, common mode feedback may be realized by simply adding a current mirror loaded OTA(see figure 5.10) closed in buffer configuration. The OTA furnishes the mismatch current between the two sections, so that it has to be biased with $I_{Tail} \geq \Delta I$. Supposing this condition holds, DC voltage of the common drain point will be given by $\frac{\Delta I}{Gm}$. To size the OTA, an estimate of the current mismatch between the two sections is needed. Threshold mismatch dominates errors in the current mirrors. From a DC perspective, worst case happens when the offsets are such that NMOS branch current drive is increased, while PMOS branch current drive is decreased. Assuming this configuration, ΔI is found by small signal analysis to be equal to:

$$\Delta I = gm_p V_{io}^P + gm_n V_{io}^N \tag{5.7}$$

As transistors are averagely biased at the same ω_T and same current, $gm_p \leq gm_n$. We may estimate gm efficiency in 15 for the PMOS amplifier and in 20 for the NMOS amplifier. Substituting these values, $\frac{\Delta I}{I_d} = 2(15V_{io}^P + 20V_{io}^N)$is found. Supposing $V_{io}^P = V_{io}^N = 2mV$, a relative mismatch of 15 % is found. Assuming a safety factor of 6, OTA tail current was chosen equal to half the main amplifier DC current. In order to miminize static error from the OTA finite gain, inversion coefficient in the transconducting pair was set equal to . 5(corresponding to an overdrive voltage of about $60mV$); mirror devices inversion coefficient is instead set to 10(corresponding to an overdrive voltage of about $140mV$). Channel lenghts are chosen to be $.18\mu$ for both devices to minimize area occupation. Under these conditions, a gm-efficiency of about 28 may be estimated for the tranconducting devices, and we find $V_{err} \approx \frac{\Delta I}{283\Delta I} \approx 10mV$ is expected, which is tolerable.
For what concerns AC performance, first notice that noise from the OTA acts as a common mode signal, and has therefore null influence as long as output impedances of the left and right sections of the LNA are perfectly balanced.
Giving specifications on OTA bandwidth may instead be more controversial. From baseband amplifier

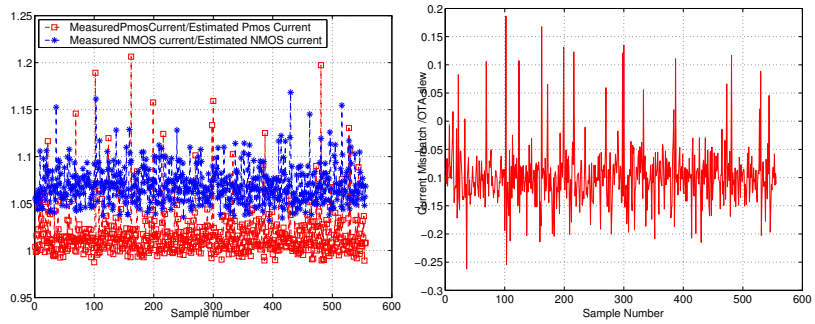

Figure 5.11: Ratio of estimated LNA bias current to simulated bias current(left), and ratio of mismatch current to maximum current drive of the common mode feedback OTA

design practice, it is known that common-mode feedback loops should have loop gain higher than one in the operating bandwidth of the main amplifier. However, one of the most established paradigms in RF design is that of avoiding to keep any feedback loop active in the GHz range, to avoid poorly modeled parasitic couplings to increase loop gain and drive the circuit into oscillation, so that actually

$$\omega_T \leq 2\pi 2.1 Grad/sec \tag{5.8}$$

should be imposed. As OTA current imposed from DC constraints explained above is almost always low enough that equation 5.8 holds, we do not need to include this last constraint into the CMFB ACG. The whole bias circuit, including common mode feedback amplifier, is sized using the $CADENCE$ environment to give a current gain of 10, without thus burdening the NPMOS input space. The accuracy of the bias circuit is reported in figure 5.11, where ratio of current mismatch to CMFB current drive is also reported.

5.2.2 Current Reuse Amplifier Analog Constraint Graphs

Due to the higher number of active devices, analog constraint graph of the current reuse amplifier is unavoidably more complex than the one proposed for the unipolar counterpart. Basic observations leading its development were reported in the previous chapter, along with [11]: to minimize noise factor, transition frequencies of the transconducting devices should be approximately equal; moreover , due to the higher input capacitance of this stage, overdrive voltages allowed are averagely larger. As for the NMOS counterpart, common gate device gate width and channel length should be kept equal to the ones of transconducting devices to allow interdigitation. As a result, degrees of freedom in sizing this circuit are represented by source degeneration inductors L_s^N and L_s^P, gate tuning inductance L_g,load parametersL_{load}, C_{load} and mixer input capacitance C_{mix}. The ACG of the NPMOS amplifier is reported in figure 5.12 Scheduling policy may be described as follows:

- Bias current is first sampled as a uniform random variable uniformly distributed in the $I_d^{Min} = .5mA$ $I_d^{Max} = 2mA$ range.

Active devices need now to be sized. Suppose for the moment we start from the N branch.

1. N device channel lenght L_n is randomly sampled through an exponential distribution analogous to the one used in the NMOS amplifier case

2. V_{lim} is assigned the value of a gaussian random variable of mean $V_{lim}^* = 80mV$ and standard deviation $V_{lim}^*/3$. Notice that a higher mean overdrive is used with respect to the N case in

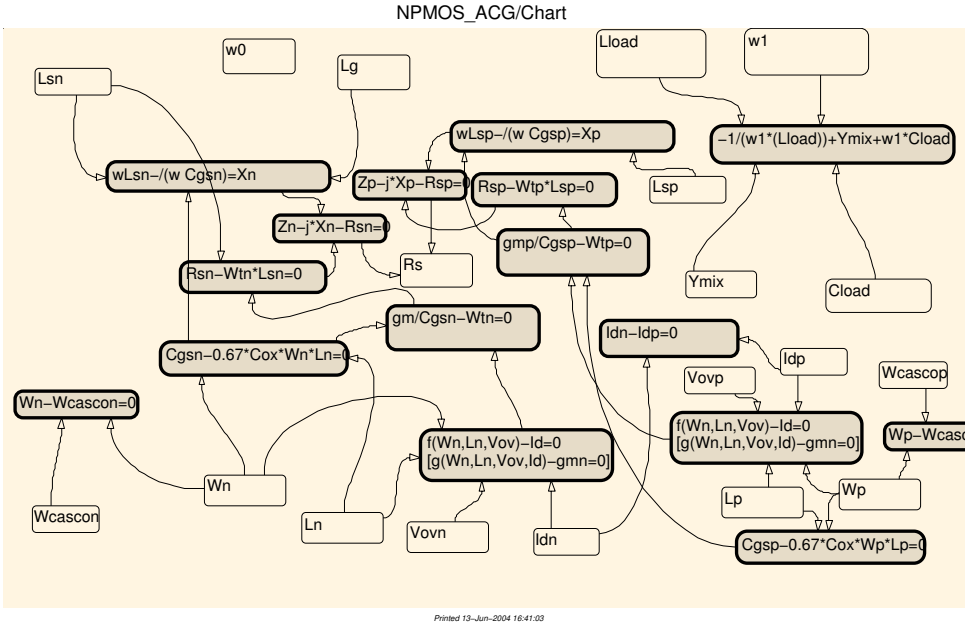

Figure 5.12: Current Reuse Amplifier Analog Constraint Graph

order to synthesize a larger impedance via local feedback. Standard deviation also is larger, as noise optimization process for this stage is not as documented as for the unipolar counterpart. Aspect ratio corresponding to this overdrive voltage and current is calculated using 5.1; a gaussian perturbation of mean 0 and standard deviation 1/6 is superimposed on this value, so that: $(\frac{W}{L})_N = (\frac{W}{L})^*(1 + N(0, 1/6))$. $W_n = (\frac{W}{L})_N \cdot L_N$ holds. If $W_n \geq W_n^{Max} = 600\mu$ a failure is claimed and scheduling process restarts. Small signal parameters ω_T^N, C_{gs}^N symmetrically to what was done in the NMOS amplfier case.

3. N-side degeneration inductance value L_s^N is sampled next. This variable is assumed to be uniformly distributed between $L_s^{Min} = .5nH$ and $L_s^{Max} = 5nH$.

4. P-side transconducting device is now sized. In order to ensure that for any given V_{lim} $E(\omega_T^P) = E(\omega_T^N)$ holds, a value $V_{lim}^{'} = V_{lim} \cdot N(1/3, 1/2)$ is sampled, and $(\frac{W}{L})_P^*$ is then determined by using equation 5.1. If corresponding $W_p \geq W_p^{Max} = 600\mu$, a failure is claimed. P-device small signal parameters are next calculated.

5. DC bias decoupling capacitors are sized next. Their values are taken as

$$C_i = Min(10 \cdot C_{gs}^i, 5pF) \tag{5.9}$$

6. Remaining degenaration inductor L_s^P is sized next. Its value is uniquely determined by the chosen parameters to be

$$L_{sp} = \frac{-b + \sqrt{b^2 - 4a \cdot c}}{2a} \tag{5.10}$$

$$a = a_1 - a_2 \tag{5.11}$$

$$b = b_1 - b_2 \tag{5.12}$$

$$c = c_1 - c_2 \tag{5.13}$$

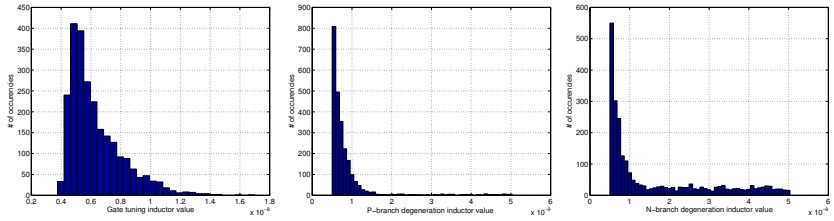

Figure 5.13: Distribution of source(first and second from the right) and gate(left) inductor values generated by "naive" NPMOS Analog Constraint Graph.

where parameters $a_1, a_2, b_1, b_2, c_1, c_2$ are defined in the appendix. This equation considers the effect of decoupling capacitors, transistor capacitances and finite quality factor of gate and source inductors(assumed equal). If 5.10 returns a complex value, or a value less than the minimum feasible inductance value, a failure is claimed, and sizing of the input pair restarts.

As one of the degeneration inductors is randomly sampled, while the other is determined solving an equation, an apparent source of bias is introduced. To overcome this cycle, an inner cycle alternates sampled inductors(i. e. on odd-numbered ACG invocations, N-side degeneration inductor is sized first; on even-numbered invocations P-side degeneration inductor is sized first).

1. Gate tuning inductance L_g is calcultated as

$$L_g = \frac{Im(Z_i)}{\omega_0}\delta \tag{5.14}$$

where Z_i is the impedance see between the common gate of the transconducting devices and ground, while δ is a corrective factor determined through simulation and equal to . 6 in the given techonology.

2. Load is sized as described for the NMOS amplifier

Configurations generated by this ACG show an increased average source inductance value, together with a decrease in the mean gate inductance(5.13). The $P_d \cdot NF$ product appears to be less flat than for the NMOS case, while mantaining an hyperbolic envelope(5.14). Source inductance distribution show the expected symmetry. The main drawback of this analog constraint graph is its very high failure rate. Averagely 1 configuration per minute is generated. This is a significant penalty, as time spent in generating a configuration becomes almost as high as time spent in evaluating its performances. A different Analog Constraint Graph was then conceived to overcome the problem. Simplified input matching equations reported in chapter 4 were used in the way reported below.

1. Drain Current is sampled as usual

2. One side of the amplifier is sized. In particular, channel lenght is first chosn to be equal to $L_{min}(1 + X_1)$ where X_1 is an exponential random variable of mean 1/9

3. Inversion coefficient IC_1 of active device is randomly sampled from a variable uniformly distributed in $[.5, 10]$. In terms of overdrive voltage this is equivalent to stating that V_{ov} lies between $16.7mV$ and $130mV$(althuogh it is not uniformly distributed).

4. Device gate width is obtained as $W_1 = \frac{I_d}{I_0 IC_1} L_1$. Small signal parameters of device M_1 are then calculated using the accurate model described in chapter 2.

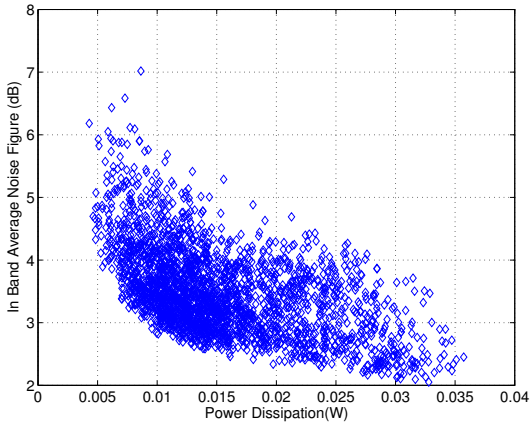

Figure 5.14: P_dNF plot for NPMOS amplfier configurations generated by this analog constraint graph

5. Corresponding degeneration inductor is randomly sampled according to

$$L_s^1 = L_s^{Min}(1 + X_1)$$

. Branch input admittance is calculated using the complete formula 4.1. For narrowband operation, this impedance may be represented by a parallel $R - C$ branch.

6. Capacitance of the remaining branch is then sampled. Code implementation is such that $\frac{C_N}{C_P} \in N(3, 1/2)$, so that $E\{\omega_T^N\} = E\{\omega_T^P\}$ if $E\{L_n\} = E\{L_p\}$. As total capacitance is now known, gate inductance L_g is derived using 4.51. If $L_g \geq L_g^{Max}$ a failure occurs and process restarts.

7. Input equivalent conductance G_s is calculated using 4.52. If 4.53 does not hold for this value of R_2, a failure is claimed, and process restarts. Otherwise, minimum and maximum allowed inversion coefficients are calculated using 4.58-4.59. Inversion coefficient of the transconducting device is chosen uniformly distributed between these two extrema.

8. L_P is finally calculated. If $L_P \leq L_{min} = .18\mu \parallel L_P \geq L_{Max} = .35\mu$, a failure is claimed and process restarts from the very the first step.

Theoretically,this ACG offers a great advantage over the one described above, as conditions for input matching that may be tested after only one device has been sized have been developed, allowing to stop the scheduling process much earlier. This ACG produces almost 200 configurations per minute on a $1GHzPentium3$ machine. Its potentialities still need to be verified in a real schedule.

5.3 Conclusions

In the preceding subsections, Analog Constraint Graphs describing the design process for NMOS and NPMOS LNAs were discussed together with their scheduling policies. Results obtained by different Analog Constraint graphs for each topology are reported in table 5.1. Parameter named "Explored to Total Space size ratio" is measured using a MonteCarlo approach as descirbed in [1], and is a measure of ACG effectiveness in reducing design space size. Parameters Configurations/Minute is a measure of the speed of the algorithm in generating valid configurations, as measured on a $1GHz$ clock,$256MbRAM$ personal computer. It is not reported for the advanced NP ACG as it has not yet been tested.

Topology	ACG	Explored to Total Space size ratio	Configurations/Minute
NMOS	"Naive"	10^{-4}	30
NMOS	Noise Optimized , Fixed R_s	10^{-6}	100
NMOS	Noise Optimized , variable R_s	10^{-5}	120
NPMOS	"Naive"	10^{-6}	3
NPMOS	Advanced	?	200

Table 5.1: Summary of performances of tested ACGs

Input stage design was mainly considered. In particular, it was demonstrated that if noise optimization is performed after that input matching to a fixed 25Ω impedance and inductor feasability bounds are imposed, a bias is introduced in the design space exploration process, because of neglected gate-shunting capacitive effects, and of unrealistic overestimation of importance of achieving a good input match with respet to optimizing noise performance. When capacitive effects are considered and matching to a randomly variable resistance source is imposed , noise optimization becomes feasible. For the current reuse Amplifier, direct noise optimization algorithms were not implemented. Instead, focus was on the development of an efficient Analog Constraint Graph, which corresponds to the second reported ACG. Tests performed in the preceding chapter on portability of equation set on different operating frequencies let hope this set of equations and scheduling strategy may reused almost unchanged when the application is changed. Representation assumed for mixer input branch in sizing the output tank of the LNA was a pure capacitance C_{mix}. As discussed in the following sections , this is an accurate representation only when linear response is considered.

5.4 Communication problems

Behavioral modeling efforts reported in literature up to present represent circuits as arbitrarily connected black-boxes,completely neglecting linear or nonlinear effects deriving from the physical connection of these blocks. Any automated system-level design effort, however, needs these effects to be addressed. For cascaded clocks,[39] suggests to treat this class of effects by including a parametrized model of the load block in the driver. Called $\vec{\lambda}\,|_{driver}, \vec{\lambda}\,|_{load}$ this interface modeling parameter, platform instances are labeled connectable iff $f(\vec{\lambda}\,|_{driver}, \vec{\lambda}\,|_{load}) = 0$ holds, where f is a function describing interface characteristics. In the present case , it has been shown in the previous chapter that for small signal analysis, mixer topology reported in [2] may be modeled as a parallel RC network, where

$$R = \frac{1}{\omega_0 C_{gd}(Re(G))} \tag{5.15}$$

$$C = C_{gg}^N + C_{gg}^P + C_{gd}(1 - Re(G)) \tag{5.16}$$

For well-designed circuit instances, $R \approx 2k\Omega$, $C \approx .5pF$. However, values of R and C depend on detailed circuit sizing through the nominal capacitance values and the parameter $G = \frac{gm_n+gm_p}{gm_{sp}+gm_{sp}+j\omega_0(C_{db}^N+C_{db}^p+C_{sb}^{sp}}$. Substituting a mixer instance with a different one results in variations in R and C, and thus in variations in the linear response of driving low noise amplifier. To estimate these variations, we make again use of sensitivity analysis. Supposing to model the LNA output load as a second order bandpass section,

$Z(\omega) = R_0 \frac{\frac{j\omega}{Q\omega_0}}{1-(\omega/\omega_0)^2+j\omega/(\omega_0 Q)}$, this results in

$$S_{\omega_0}^{|Z|} = \frac{Q^2(-1+x^4)}{(x^2+Q^2(-1+x^2)^2)} \tag{5.17}$$

$$S_{Q}^{|Z|} = -Q^2\frac{(-1+x^2)^2}{x^2+Q^2(-1+x^2)^2} \tag{5.18}$$

$$S_{R_0}^{|Z|} = 1 \tag{5.19}$$

Both these quantities depend on the value of $x = \frac{\omega}{\omega_0}$ and Q and evaluate to 0 when x=1. This is reasonable because this value coresponds to signal frequency equal to tank resonant frequency , so that first order variations are null. To get significative estimates, we refer to Q=6,x=1. 03, which corresponds to having a load nominally tuned at $f = 2.11GHz$ and incoming signal at $2.17GHz$(signal and load nominal resonating frequencies are at opposite ends of the UMTS downlink spectrum). In this circumstances, we find $S_{\omega_0}^{|Z|} = -3.9$,$S_{Q}^{|Z|} = -.1$. Now, notice that actually mixer input capacitance only contribute a fraction of total tank capacitance, as ,in the worst case, at least device parasitic capacitance loads the output node. In all other cases, a linear capacitance C_{lin} is connected in parallel to C_{mix}, so that $C_{tot} = C_p + C_{lin} + C_{mix} = 1/(\omega_0^2 L_{load})$. Then , neglecting variations of frequency response due to variations in Q, $\frac{\Delta Z}{Z} \approx -1.5\frac{\Delta C_{tot}}{C_{tot}}$. Morevore, for a mixer input resistance of $2k\Omega$, and a tank parallel resitance of 400Ω, we may evaluate the sensitivity of $R_{tot} = R_{tank} \parallel R_{Mix}$ to be

$$S_{R_{tot}}^{R_{mix}} = \frac{R_{mix}}{R_{mix}+R_{tank}} \approx .16$$

. Consider now the relation between R_{mix} and C_{mix}. For the design discussed in chapter 4, tail capacitance C of the switching pairs is mainly due to switching devices self loading source bulk parasitic. Substituing values in reported in chapter 4, $C \approx 600fF$ and $\omega_C \approx 30Grad/s$ are found ,so that at the operating frequncy gain form gate to drain of mixer input devices has a phase shift of -24 degrees. This phase shift leads to a Miller capacitive to resistive admittance ratio of $(1+\cos(.4))/(\sin(.4)) = 4.3$. Ratio of Miller capacitance to total input capacitance on the hand is readily expressed for a sinle device as

$$CR = \frac{(C_{gd})(1-Re(G))}{C_{gs}+C_{gb}+C_{gd}(1-Re(G))} =$$

$$\frac{(1-Re(G))C_{ox}L_{min}/4}{C_{ox}L_{min}/4+2/3C_{ox}L_{min}/2+1/9C_{ox}L_{min}/2+(1-Re(G))C_{ox}L_{min}/4} \approx \frac{1}{1+2.5/(1-ReG)} \approx .5$$

Where input stage devices were assumed to be minimum sized, so that $L_{eff} \approx L_{drawn}/2$ holds. We may therefore approximate $R_{mix} \approx \frac{9}{\omega_0 C_{mix}}$. This is a pessimistic estimate in case transconductiong devices have a channel lenght grater than the minimum, as in that case $CR \leq .5$. Variations in gain when different mixer configurations are connected to a same LNA may at this point be calculated to be

$$\frac{\Delta G}{G} = -.16\frac{\Delta C_{mix}}{C_{mix}} - 1.5\frac{\Delta C_{mix}+\Delta C_{lin}}{C_{tot}} \tag{5.20}$$

From 5.20 we see that variatios in LNA gain due to output tank resonant frequency variations with mixer capacitance evaluate to 0 if shunt linear capacitance is varied in such a way as to hold total capacitance constant. This is not true for variations in resistive load. Although this is a low-sensitivity parameter it may experience very large variations,limiting final gain accuracy. To maximize this accuracy, mixer input capacitance and linear resistance can be measured through simulations, defining thus interface composition condition as

$$R_{mix}^A = R_{mix}^B \tag{5.21}$$

$$C_{mix}^A + C_{lin}^A = C_{mix}^B + C_{lin}^B \tag{5.22}$$

Where ΔR is a constant ideally equal to 0, controlling allowed resistance mismatch. Although able to guarantee the highest accuaracy, this interafce composition rule was not used. The main reason for this choice is in the low sensitivity with respect to R_{mix} value that was proven above. The worst case estimated value is sow low that, if resistive mismatch is allowed to contribute half the total allowed gain accuracy of 10%, i. e. 5% , ±15% relative error on R_{mix} is tolerable. Actually, this error estimate may be very pessimistic. It is easy to obtain this conclusion looking back at the way this worst-case analysis was performed. Sensitivity to R_{mix} is in fact the higher the lower is R_{mix} itself. On the other hand, the linear approximation which is the basis of sensitivity analysis itself is much more accurate for downwards variations of R_{mix} than for upward ones. Variations in this sense happen the more rarely, the lower the chosen initial resistance value. These looser accuracy constraints naturally lend themselves to analytical modeling rather than parameter value extraction from simulation. Mixer input capacitance, on the other hand has shown in the example above a 1. 5 sensitivity. This value results in a ±1.6% required accuracy. This value cannot be obtained analytically,so it must be extracted through simulation. Proposed interface composition rule is finally expressed mathematically as follows:

$$C_{mix}^{LNA} + C_{lin}^{LNA} \geq C_{mix}^{Mixer} \tag{5.23}$$

To reduce error due to resistance variations, $\frac{\Delta Gain}{Gain} = \frac{R_{tank}}{R_{tank}+\hat{R_{Mix}}}\Delta(\hat{R_{mix}})/\hat{R_{mix}}$ is estimated and used to correct the maximum gain value with $\hat{R_{mix}} = \frac{1}{\omega_0 C_{gd}}(1 - Im)(G)$ is the estimated mixer input resistance. Next, consider the effect of the described communication problems on low-noise-amplifier nonlinear response. Suppose for the moment that distortion contribution due to nonlinearity in the gate capacitances is minimized by enforcing $C_{gg_N} = C_{gg_P}$, $IC_n = IC_p$. Second order derivative of mixer input capacitance may then be written recalling 4.46 as

$$C_{gdo}\frac{gm_3^N + gm_3^P + gm_3^{SP} * G^3}{gm^{SP}n} + C_2^{db} \approx \frac{C_{mix}}{4}\frac{gm_3^N + gm_3^P + gm_3^{SP} * G^3}{gm^{SP}n} + C_2^{db}$$

So that even for a constant input capacitance,when bias point of either switching pairs, or transconductor is changed, this term will change. Simulation and analytical results show that variations of the order of 100% may be expected. As the contribution of interface has been proven using 4.46 to be dominant in real low-noise amplifier designs, results that even imposing the interface composition rule derived for linear response distortion will vary.

5.5 Interface Model

To cope with the problem of modeling nonlinear interface effects in a way that is robust with respect to variations in the load mixer itself, we make the following observation: we are not interesting in modeling distortion on evey single node of the receiver chain, but only in predicting signal to noise plus distortion ratio at the output of the whole receiver. If distortion from the interface were for example considered to be an effect due to the Mixer, rather than due to the LNA, it would be reasonable for a behavioral model accounting for it only at the output of the mixer itself. In order to better understand this, consider figure 5.15. We may model the cascade of Low Noise Amplifier, Load Tank and Mixer as a nonlinear transconductor cascaded with a nonlinear load further cascaded with a nonlinear gain stage(Frequency conversion is implicit in the mixer model). If we assume each nonlinear source to be representable by a third polynomial $y = a_1 V + a_3 V^3$, overall distortion is expressed as:

$$D = V_{rf}^3(a_{3_{LNA}}a_{1_{Load}}a_{1_{Mixer}} + (a_{1_{LNA}}a_{1_{Load}})^3(a_{3_{Load}}a_{1_{Load}}a_{1_{Mixer}} + a_{3_{Mixer}})))$$

Now, recall that under the linear block composition rule, $a_{1_{LNA}}$ and $a_{1_{Load}}$ are constant with a 10 % accuracy. If we further suppose $a_{3_{LNA}}$ to be also constant and $a_{3_{Load}} = a_{3_{Load}}^{Mixer} + a_{3_{Load}}^{SelfLoading}$

$$D = V_{rf}^3(a'_{3_{LNA}}a_{1_{Mixer}} + (a'_{1_{LNA}})^3 a'_{3_{Mixer}}) \tag{5.24}$$

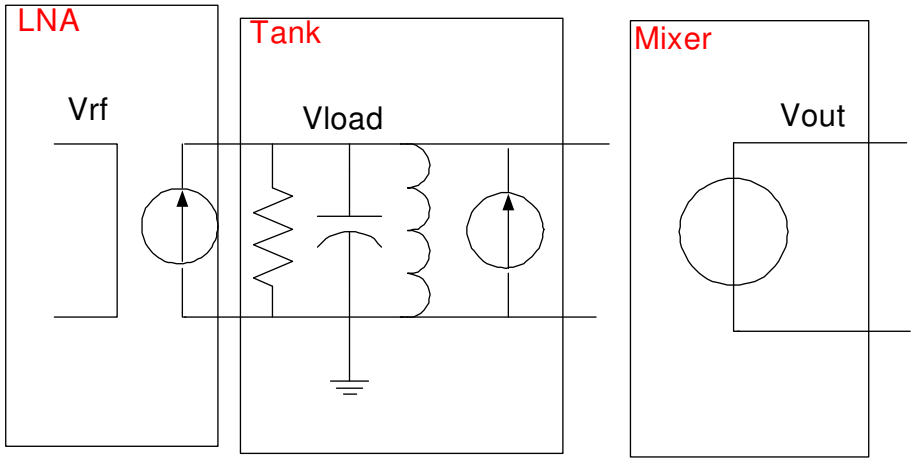

Figure 5.15: LNA-Mixer cascade connection highlithing different contributions to overall nonlinearity: each of the tank,the LNA and the Mixer is represented by a nonlinear voltage or current source

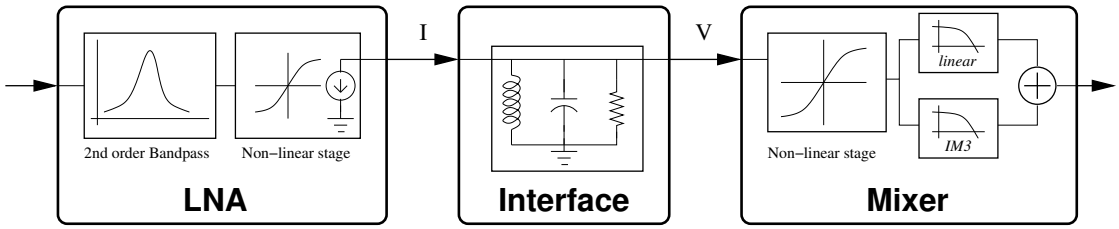

Figure 5.16: System equivalent to the one in figure 5.15 for total intermodulation distortion calculation: number of nonlinear geenrators has decreased to two, with coefficients given by 5.24

Where $a'_{3_{LNA}} = a_{3_{LNA}} a_{1_{Load}} + a^{SelfLoading}_{3_{Load}} a_{1_{Load}}, a'_{1_{LNA}} = a_{1_{LNA}} a_{1_{Load}}, a'_{3_{Mixer}} = (a^{Mixer}_{3_{Load}} a_{1_{Load}} a_{1_{Mixer}} + a_{3_{Mixer}})$. Equation 5.24 tells us that if linear response composition rule is enforced, then distortion at the output node is equivalent to that generated by an LNA closed on a perfectly linear load, connected to a mixer of nonlinearity augmented by a term $a^{Mixer}_{3_{Load}} a_{1_{Load}} a_{1_{Mixer}}$. This system is shown in figure 5.16. Obviously, intermodulation distortion at LNA-Mixer interconnection node has become an unobservable quantity. Now, take a deeper look at $a'_{3_{Mixer}}$. The term $a_{3_{Mixer}}$, intuitively due to nonlinearities in switching pairs and transconducting devices, must be more operatingly defined. Consider then a Mixer instance driven by two sinusoidal inputs of amplitudes A_1 and A_2, with $A_1 \gg A_2$ and frequencies f_2, f_1 such that $2f_1 - f_2 - f_{LO} = 1MHz$. Further assume that these inputs are applied to the mixer through ideal voltage sources with zero series resistance. Called A_{IM} the amplitude of mixer output spectrum at $1MHz$, we may define $a_{3_{Mixer}}$ as

$$a_{3_{Mixer}} \equiv \frac{A_{IM}}{A_1^2 A_2} \tag{5.25}$$

. $a'_{3_{Mixer}}$ on the other hand may be defined considering the same mixer, driven by a current source with output impedance Z_o such that $Z_o \parallel Z^{Mix}_{in} = a^{Load}_1$. Current source I_0 drives the mixer with two sinusoidal tones at the same frequencies f_1 and f_2 used in $a_{3_{Mixer}}$ measurement, and of amplitudes $I_1 = A_1/a_{1_{Load}}(f_1), I_2 = A_2/a_{1_{Load}}(f_2)$. Under these circumstances, call A'_{IM} the amplitude of mixer

output spectrum at $1MHz$, and define

$$a_{3_{Mixer}}^{'} \equiv \frac{A_{IM}^{'}}{A_1^2 A_2} \tag{5.26}$$

. The difference in these two terms is mixer input referred contribution of Low Noise Amplifier nonlinear load $a_{3_{Load}}^{Mixer} a_{1_{Load}} a_{1_{Mixer}}$. Notice that it is made of terms depending only on mixer instance detailed sizing except for $a_{1_{Load}}$, which is mostly determined by low-noise amplifier load inductance. To avoid introducing an additional interface composition rule, we may use 5.24, and arrive at

$$a_{3_{Mixer}}^{'}(a_{1_{Load}}) = a_{3_{Mixer}} + \frac{a_{3_{Mixer}}^{'} - a_{3_{Mixer}}}{a_{1_{Load}}^{*}} a_{1_{Load}} \tag{5.27}$$

So that given a Low-Noise Amplifier instance and a Mixer Instance such that rule of composition 5.23 holds, we only need to know the values of $a_{3_{Mixer}}$ and $a_{3_{Mixer}}^{'}$ for a given value $a_{1_{Load}}^{*}$ of tank impedance, as well as the value of Low Noise Amplifier instance loaded impedance $a_{1_{Load}}$ to calculate $a_{3_{Mixer}}^{'}$. Now, refer to the low-noise amplifier. We suppose it is driven by a voltage source with internal resistance equal to $R_{antenna} = 50\Omega$. Consider first linear response. Then, called i_{casco} the common gate-device intrinsic drain current, $a_{1_{LNA}} \equiv i_{out}/V_s$, while $a_{1_{Load}} \equiv V_{out}/i_{out}$ where V_s is the value of LNA input voltage, assumed small enough for nonlinear terms to be negligible. For a two-tone large signal input, characterized by frequencies f_1, f_2 and by amplitudes A_1, A_2, call i_{out}^{IM} the value of common gate device intrinsic drain-current at $f = 2f_1 - f_2$,I_{lna}^{IM} the common gate device drain current at $f = 2f_1 - f_2$ and define

$$a_{3_{LNA}} \equiv \frac{i_{out}^{IM}}{A_1^2 A_2}$$

$$a_{3_{Load}}^{LNA} \equiv \frac{I_{lna}^{IM}}{A_1^2 A_2}$$

Measuring these quantities requests accessing devices internal nodes, which is an unpractical task. A much more simple and accurate approach consists in directly measuring $a_{3_{LNA}}^{'}$ which under the same test circumstances, evaluates to

$$a_{3_{LNA}}^{'} \equiv \frac{V_{out}^{IM}}{A_1^2 A_2} \tag{5.28}$$

This completes the definition of interface model. Summarizing, the model is made up of the interface composition rule 5.23, and the strategy chosen to measure $a_{3_{LNA}}^{'}$ and $a_{3_{Mixer}}^{'}$. The model is based on including the contributions of load nonlinear portion into the mixer nonlinear response itself. Intermodulation distortion due to nonlinear load is therefore "saved" and added at the output of whole front-end. This makes intermodulation distortion predictions at the LNA-Mixer connection node inaccurate "by construction" while allows good accuracy to be obtained on overall intermodulation distortion prediction.

5.6 Low Noise Amplifier Linear and Nonlinear response model

In the previous equations, we have considered quantities $a_{1_{LNA}}, a_{1_{Mixer}}, a_{3_{Mixer}}, a_{3_{LNA}}$ to be frequency independent. This is not the case in practice, as UMTS performance tests involve input tones with freqeuncies ranging from 1. 9 to 2. 2GHz. Models for the frequency dependence of Low Noise Amplifier linear and nonlinear response need now to be evaluated. As mentioned in chapter 2, behavioral models investigated in this work have a filter-memoryless nonlinearity-filter structure. This structure was investigated as a lower number of parameters is required to model the linear response of a system

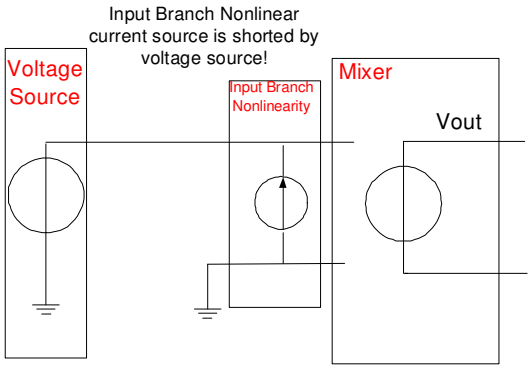

Figure 5.17: Measurement setup for $a_{3_{Mix}}$

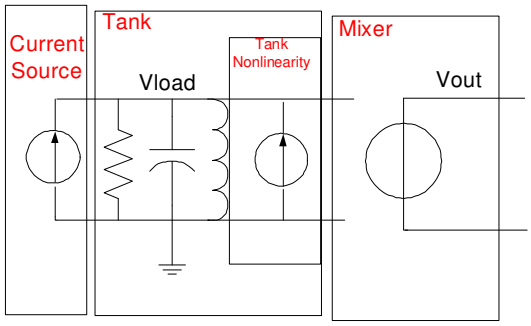

Figure 5.18: Measurements setup for $a'_{3_{Mix}}$

with respect to separate kernel measurement and identification. This obviously comes at the expense of accuracy. Consider first the quantities governing the linear LNA response, $a_{1_{LNA}}, a'_{1_{LNA}}$ and $a_{1_{Load}}$. Both from simulation and from the treatment developed in chapter 4, we know that $a_{1_{LNA}}$ may be well approximated by the cascade of an underdamped secind order low-pass filter(the quality factor of this filter is the quantity $Q_L = 1/(\omega_0 C_{gs} R_s)$ defined in chapter 4, which usually evaluates to 5-6)resonating at angular frequency ω_r, and a first order low-pass filter representing common gate device current gain which has bandwidth equal to the ω_C defined in chapter 4. As usually $\omega_0 \ll \omega_C$ this second component may be neglected without significantly degrading accuray. Overall current response is therefore estimated of the form $\frac{1}{1-(\omega_0/\omega_r)^2+j\omega_0/(Q\cdot\omega_r}$. Peak of the response is located at $\omega_{peak} = \omega_r\sqrt{1-\frac{1}{Q^2}} \approx \omega_r$, theoretically exactly at the operating frequency, and approximately evaluates to Q. Notice that this representation does not include the effects of the zero due to the overlap gate-drain capacitance, neither the effects of capacitive components shunting transconducting device gate to ground. It is probably because of these effects that LNA response, under certain circumstances, deviates from the behavior predicted by theory, exhibiting a notch. Conditions under which this happens were not well understood, however this event presents with nonegligible frequency, and avoids the aforementioned representation which, being physically based, looks appealing, to be usable. A much better conditioned problem is that of modeling the output tank response. From circuit analysis this response is found to be

$$Z(\omega) = \frac{j\omega L + R_L}{1 - LC\omega^2 + jRC\omega}$$

For the frequency range of interest however, term R in the numerator may be neglected, and load response is well approximated by a second order band-pass function. Similar considerations also hold for the overall voltage gain $a'_{1_{LNA}}$. Defined Q_Z, A_z, f_Z and Q_V, A_V, f_V the quality factor, maximum maximum absolute value and tuning frequency respectively of the LNA output impedance and voltage gain, accurate representations of linear response have been found to be:

- The cascade of a bandpass filter with unity gain, a memoryless block of gain A_v and a second filter. Both filters have tuning frequency f_V and quality factor $.7 \cdot Q_V$ as reported in [64].
- The cascade of a bandpass filter of tuning frequency f_V and quality factor Q_V, either preceding or following a memoryless gain block.
- The cascade of a second order high pass filter wih frequency response

  $$H(\omega) = \frac{1 - (\frac{\omega}{\omega_Z})^2 + j\frac{\omega}{\omega_Z Q_Z}}{1 - (\frac{\omega}{\omega_V})^2 + j\frac{\omega}{\omega_V Q_V}}$$

  , a memoryless block with $\frac{A_V}{A_Z}$ and a secondo order bandpass filter ection with frequency response $H'(\omega) = \frac{A_Z j\frac{\omega}{\omega_Z}}{1-(\frac{\omega}{\omega_Z})^2+j\frac{\omega}{\omega_Z Q_Z}}$

For each of the aforementioned models, a parameter extraction routine was conceived and implemeted in order to extract O-space parameter from simualtion data.Accuracy in predicting linear response was better than 5% for all of these models. The situation becomes more complicated when nonlinearitities are considered as well. The are introduced in the model by turning aforentioned memoryless linear blocks into memoryless nonlinear blocks. This happens as frequency dependance of Third Volterra Kernel is related to input and output filter frequency response by $H_3(\omega_1, \omega_2, \omega_3) = k_3 H_1(\omega_1) H_1(\omega_2) H_1(\omega_3) H_2(\omega_1 + \omega_2 + \omega_3)$. Accuracy of these representations were compared for both both LNA topologies, both for a sizing exhibiting a gain of 6 and for a sizing exhibiting a gain of 20. Best representation for a LNA loaded by an RC block was found to be constituted by a memoryless block followed by a second order banpass section with $Q = Q_V, f_0 = f_V$. If the LNA is instead loaded by a complete mixer[1], a filter-noninearity filter with parameters values assigned as discussed in the point 1

[1]Simulations in this case were taken considering the LO signal off

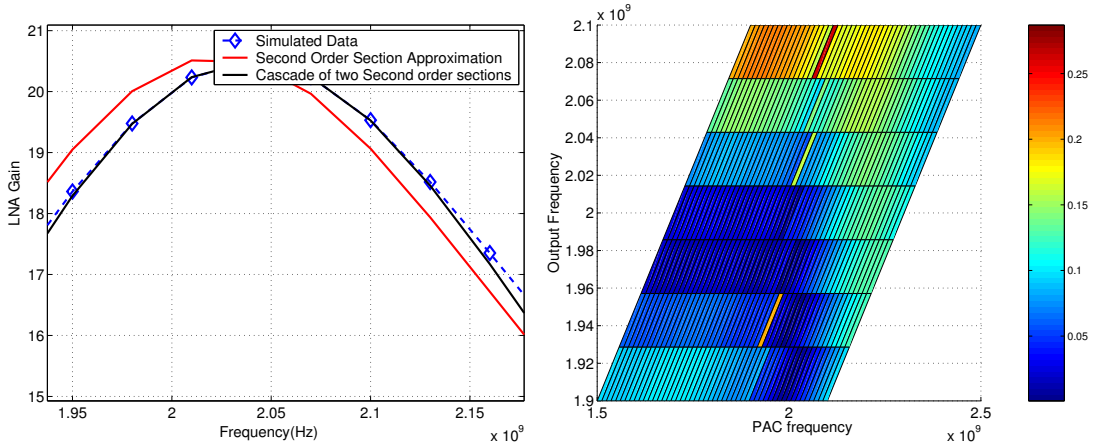

Figure 5.19: Example of model accuracy for linear(left) and nonlinear response when a load mixer is fixed

above is instead more indicated. Figure 5.19 reports results of the comparisons between $SPECTRE^R$ for what concerns linear and noninear response.

5.7 Platform Implementation

Different characterization processes were run, in order to test the validity of proposed models over a set of configurations as wide as possible. At the time of writing definitive results however are still lacking. Here we report some additional details, such as input and output space detailed definintion, and strategies used at the simulation level to achieve accurate input and output tuning. Latest available performance models were built buil using the noise optimized ACG with fixed source resistance for the NMOS amplifier, and the "naive" analog constraint graph for the NPMOS amplfier. Distortion is measured in the current domain, measuring total current flowing into the LNA load, constituted by the parallel connection of a mixer sized according to the ACG described in [1], and a linear capacitor and inductor. As a result, a whole 19 elements row vector describing mixer sizing is included as part of the LNA input space, which is summarized in tables 5.8. Targeted behavioral model has the high-pass filter/nonlinearity/band-pass filter structure described above; noise is propagated on a separate channel from signal by using Friis's formula. Low-Noise amplifier platform output space is then summarized in table 5.7, where parameters are cited together with their measurements strategy. During configuration generation process, capacitance of the generated mixer configuration is estimated using equation 5.16 and the advanced compact model described in chapter 3. This estimate is used to size LNA output load tank, but has not sufficient accuracy to be used when enforcing interface composition rule. As a result, at the begining of each performance evaluation step, an AC analysis is used to directly measure mixer input admittance and device parasitic capacitanceC_{db}. Linear capacitor C_{load} is then resized to guarantee output tank tuning at 2. 1GHz. If the required value of C_{load} is negative, $C_{load} = C_{load}^{Min} = 1aF$ is used [2]. In the same step, reactance X_i of impedance seen looking to the right of transconducting device is also measured and value of is adjusted so that $L_g = X_i/\omega_0$[3]Resulting distributions ofo input and output tuning frequencies are shown in figure 5.20. Output space parameters measurement is then

[2]This extremely small value is used instead of zero to avoid simualtor warnings)

[3]Notice that if estimated values of L_g are compared to values after this tuning step, the latter ones averagely appear to be almost half as large as the former, confirmning the results from analisys in chapter 4.

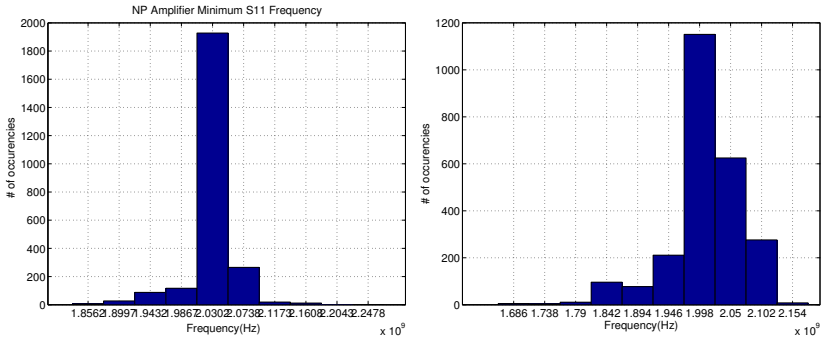

Figure 5.20: Distribution of Input and Output Tuning frequencies for NPMOS Low Noise Amplifier

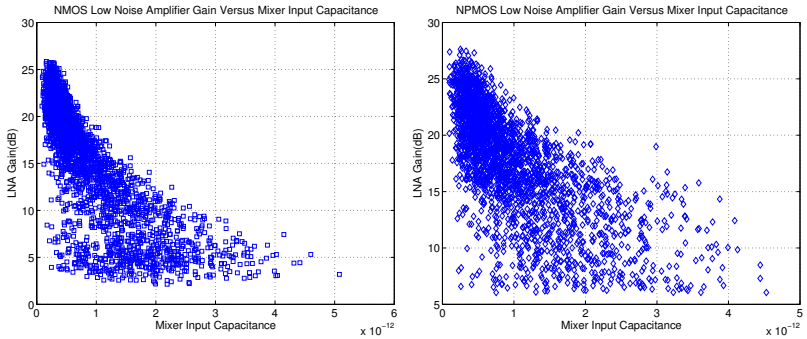

Figure 5.21: Mixer input capacitance versus gain for NMOS(left)and NPMOS(right) low noise amplfiers

performed. O -space elements for the Low Noise Amplifiers are reported in table 5.7. Notice that in bandwidth average noise figure is used as noise metric,while intermodulation distortion is measured using the same input frequencies as reported in [2] in order to maximize accuracy when the same tests are performed. Distribution of measured intermodulation distortion values is instead reported in figure 5.21 against gain of the amplifier. Area is instead estimated by approximating total LNA area with inductors area, which is calculated using equation 3.5. Figure 5.22 reports area values for instances of the NMOS amplifier(blue) and NPMOS amplifiers. Finally, C_{mix} has to be included in the output space in order to enforce 5.23. Characterization process characteristic times are of the order of 3 to five days simulation on $SunULTRA-60$ workstation. The process of building the platform is inherently time consuming, so that errors in the ACG generation and O -space parameters selection process have a high cost in terms of design time. This unfavorable situation is made worse by some form of characteristic

Performance	Input Signals
$Gain$	$f_1 = 2.101GHz$, $-30dBm$
$IM2$	$f_1 = 1.965GHz$, $-33dBm$;
	$f_2 = 1.964GHz$, $-33dBm$
$IM3$	$f_1 = 2.0325GHz$, $-40dBm$;
	$f_2 = 1.964GHz$, $-30dBm$

Table 5.2: UMTS tests.

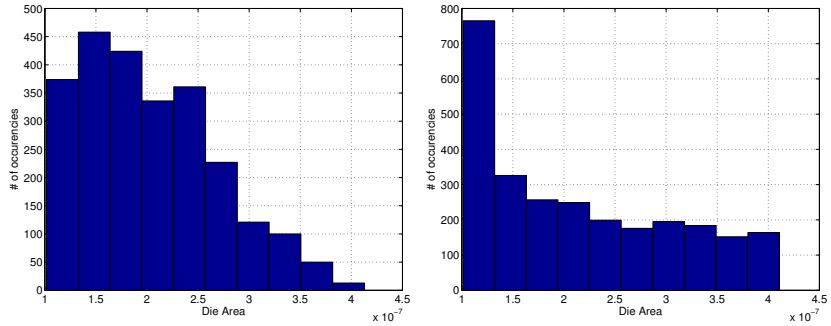

Figure 5.22: Estimated die area occupation for NMOS(right) and NPMOS(left) amplifier

Performance	Optimal	NN1	NN2	NN3
$Gain(dB)$	35. 48	35. 49	32. 9	35. 6
$Power(mW)$	13. 5	13. 5	15	15. 8
$IIP2(dBm)$	70	70	47	69. 5
$IIP3(dBm)$	-3. 1	-3. 1	-4	-3. 8
NF	7. 04	7. 04	7. 27	6. 8
NF_{min}	7. 16	7. 16	6. 72	7. 16

Table 5.3: Performance of optimal receiver configuration and nearest neighbors.

Performance	Optimal	NN1	NN2	NN3
$Gain(dB)$	27	27	27. 8	27. 2
$Power(mW)$	6. 4	6. 4	9. 9	8. 5
NF	3. 5	3. 5	3. 1	3. 3
$f_0(GHz)$	2. 08	2. 08	2. 08	2. 07
Topology	N-LNA	N-LNA	N-LNA	N-LNA

Table 5.4: Performance of LNA corresponding to optimal receiver configuration and nearest neighbors.

data structure of RF systems. Both for the LNA and the Mixer([1]), the software for output space interpolation and regression had numerous convergence problems ([40]), so that all but a few attempted characterizations stopped after their maximum number of iterations had been reached. Such a behavior was not reported for example in [63], where the same methodology was applied to baseband sections charaterization.

5.8 Model Accuracy

In order to test the proposed solution on a class of circuit instances as wide as possible, comparison of proposed behavioral model to $SPECTRE^{RF}$ transistor level simulator was undertaken on over 200 LNA-Mixer pairs automatically chosen from the respective performance models according to equation 5.23. Input signals were chosen according to [2], and are reported in 5.2. Mixer model is described in [1]. The results are summarized in table 5.6. A few comments are necessary. First,tested behavioral model did not include resistive mismatch effects compensation described in this chapter. Then notice that accuracy on $IM2_{Mixer}$ is significantly lower than what could be estimated from sensitivity analysis. Actualy, simulator numerical convergence problems may be blamed for this mismatch.

Performance	Optimal	NN1	NN2	NN3
$CG(dB)$	8. 62	8. 62	12	6. 23
$Power(mW)$	7. 22	7. 22	8	8. 5
$IIP2(dBm)$	94	94	75	71
$IIP3(dBm)$	12. 8	12. 8	7. 9	12. 1
$IRN(nV/sqrt(Hz))$	5. 16	5. 16	4	6. 75

Table 5.5: Performance of mixer corresponding to optimal receiver configuration and nearest neighbors.

5.9 System Level Optimization

5.9.1 Output Space Interpolation

Before introducing system optimization, a further insight on output space interpolation is needed. As briefly described in the introduction, Platform Based Design relies on the use of a statistical classifier(in [56] a Support Vector Machine) to estimate the perfomance model P from simulation data and reduce the number of simulations needed to complete a characterization process. There is therefore a difference between the estimated O space(i. e. the set of configurations labeled feasible by the statistical classifier), and the simulated O-space(i. e. the set of N-uples of output space parameters obtained from direct simulation of the configurations generated through the ACG). Figure 5.23 reports a projection of the estimated O-space on the $P_d - NF$ plane. Crosses(unfortunately not well visible), correspond to simulated points, while coloured zones must be interpreted according to the colorbar reported at the extreme right of 5.23: Red zones correspond to regions most probably feasible; Blue regions on the other hand, to regions least probably feasible. The effect of this estimation process may also be seen as an extension of the simulated O -space to reasonanly feasible results, so that a more efficient exploration is performed. This enlarged design space is used in the optimization phase to generate platform instances to be evaluated.

5.9.2 Optimazion

As a final step, system level optimization of a receiver platform built joining Low Noise Amplifier and Mixer platform according to 5.23 was performed. Receiver architecture is basically the same as the one reported in [2], with the difference that in that case, a capacitive divider was used at the interface between LNA and Mixer to lower LNA gain and thus reduce mixer distortion, while in this work, the capacitive divider was removed in order to explore the possibilities of LNA-Mixer co-design offered by the platform paradigm. Due to the non-convexity properties of the optimization domain, a stochastic algorithm derived from simulated annealing was used. Goal of the optimization process may be stated as

Performance	$\pm 1dB$	$\pm 2dB$	$\pm 3dB$	$\pm 6dB$	$\pm 10dB$
LNA out @2.101GHz	57%	93%	99%	100%	100%
LNA out @1.965GHz	56%	88%	98%	100%	100%
LNA out @1.964GHz	55%	88%	98%	100%	100%
LNA out $IM3$ @2.101GHz	5%	11%	21%	44%	65%
Mixer out $1MHz$	72%	99%	99%	100%	100%
Mixer out $IM2$ @$1MHz$	10%	24%	36%	65%	89%
Mixer out $IM3$ @$1MHz$	15%	24%	39%	80%	96%

Table 5.6: Comparison between receiver behavioural model and schematic simulator. Each column shows the number of samples providing an error lower than the reported one.

Parameter	Analysis
G_{LNA}, Q, f_0	AC $1GHz \rightarrow 3GHz$
NF	noise $1.5GHz \rightarrow 2.5GHz$
$IM3$	PSS @$f_1 = 2.03GHz$, $V_{in_1} = -40dBm$;
	PAC @$f_2 = 1.961GHz$, $V_{in_2} = -30dBm$
C_{Mix}	AC $1GHz \rightarrow 3GHz$
Z_{max}, f_0^Z, Q_Z	AC $1GHz \rightarrow 3GHz$

Table 5.7: LNA performance measurements setup.

NMOS	W_n	L_n	L_g	L_s	I_d	L_{Load}	C_{Load}	V_{io}	$\vec{k}_{Mixer}$			
NPMOS	W_n	L_n	W_p	L_p	L_s^P	L_s^N	L_g	I_d	L_{Load}	C_{Load}	V_{io}^N	V_{io}^P	$\vec{k}_{Mixer}$

Table 5.8: Summary of circuit Input Space parameters

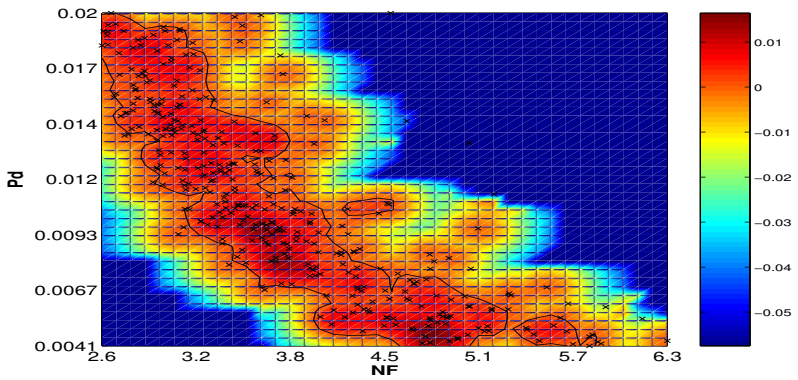

Figure 5.23: Projection on the $P_d - NF$ plane of the performance model of an N-MOS LNA . Crosses represent simulated points; regions labeled mosty probably feasible are coloured red, while regions labeled least probably feasible are in blue.

minimizing a weghted sum of die area consumption and power dissipation of the receiver, while ensuring compliance to UMTS standard. This condition is summarized by equations (2.38-2.39) . At the same time, $Gain$ should be kept as close as possible to the nominal value of $31dB$ reported in [2], in order to ensure compatibilty with the baseband sections. We now describe a step of the optimization algorithm, considering this pure Simulated Annealing. The detailed description of the pecularities of the used algorithm and of its implementation may be found in [40] or in [43]. At execution step N,define then quantities $T_n, C^*, x^*, \vec{a}$ as lattice temperature,current accepted cost, current accepted configuration, variance vector. Algorithm flows as follows.

1. LNA and Mixer instances are generated randomly using respective performance models. This consists in generating a random perturbation vector ΔC by sampling a vector random variable with variance $\vec{\sigma} = \vec{a}T_n$ and adding it to the current optimal configuration C^*. At this step, N and NPMOS Low Noise Amplifiers are undistinguishable, being represented by homogeneous sets of vectors of their O -space.

2. Performances of the receiver are evaluated by using the UMTS tests described in [2]. Receiver IIP2,IIP3,Power Dissipation, Area,Gain and Noise Figure are calculated according to

$$.G\mid_{dB} = V_u^{Gain}\mid_{dBm} +30 \tag{5.29}$$

$$IIP_2\mid_{dBm} = V_u^{IM2-1MHz}\mid_{dBm} - G\mid_{dB} \tag{5.30}$$

$$IIP_3\mid_{dBm} = V_{in}^{IM3} + \frac{V_u^{IM3-1MHz}\mid_{dBm} - V_u^{Lin}\mid_{dBm}}{2} \tag{5.31}$$

$$NF = 10\log F \tag{5.32}$$

$$F = F_{LNA} + \frac{\frac{IRN_{Mix}}{4\cdot KT\cdot BWR_s} - 1}{G_{LNA}^2} \tag{5.33}$$

$$P_d = P_d^{LNA} + P_d^{Mixer} \tag{5.34}$$

$$Area = Area_{LNA} + Area_{Mixer} \tag{5.35}$$

Where IRN_{Mixer} is the double sideband input referred noise of the Mixer circuit, defined in [6] as

$$O_{noise} = \frac{\int_{10KHz}^{1.92MHz} No(f)df}{1.92\cdot 10^6} \tag{5.36}$$

$$IRN_{Mixer} = \frac{O_{noise}}{CG_{mixer}^2 2} \tag{5.37}$$

As direct imposition of equations 2.38-2.39 and composition rule 5.23 as hard constraints would lead to an excessive number of failures and to unacceptably slow algorithm convergence rates(see 5.25, these conditions are included in the cost function itself by making use of the penalty funtion method [48]. As a result, cost function takes the form

$$C = a_1P_d + a_2Area + G(NF - NF_{min}) + H(G - G_{nom}) + U(C_{mixer}^{Mixer} - (C_{mixer}^{LNA} + C_{lin}^{LNA})) \tag{5.38}$$

Where G_{min} is the gain of $LNA+Mixer$ chain reported in [2],NF_{min} is the minimum noise figure calculated by inverting equations 2.38,2.39 for the above calculated values of and IIP2,IIP3; while G, U and H are nonlinear cost functions, which control the tolerance on receiver compliance to UMTS tests, block composability and baseband compatibility constraints. These tolerances should be chosen by taking in consideration model inaccuracies and optimization convergence time. In this work,

$$U = k_8(\frac{C_{mixer}^{Mixer} - (C_{mixer}^{LNA} + C_{lin}^{LNA})}{C_{mixer}^{LNA} + C_{lin}^{LNA}})^2 u(-C_{mixer}^{Mixer} + (C_{mixer}^{LNA} + C_{lin}^{LNA}) \tag{5.39}$$

$$H = 10a_1\tanh\left(a(NF - NF_{min})\right)^b u(NF - NF_{min}) \tag{5.40}$$

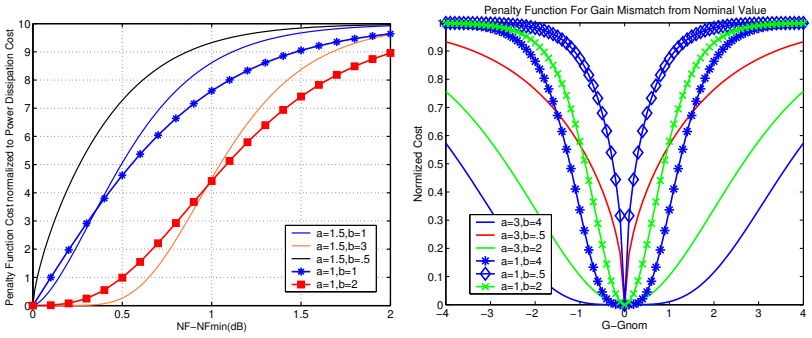

Figure 5.24: Shape of H(x) and of G(x) used in the optimization process

,where $u(x)$ is the Heaviside function or unity step and $a \geq 0, b \geq 0$, was chosen. Function H(x) is reported in figure 5.24 for different values of a and b. It does not give significant penalty for $NF - NF_{min} \leq x_L = \frac{1}{2a} \log \frac{(1+(.01)^{\frac{1}{b}})}{(1-(.01)^{\frac{1}{b}})} dB$, while gives a penalty nine times larger than cost of power if $NF - NF_{min} \geq x_H = \frac{1}{2a} \log \frac{(1+(.9)^{\frac{1}{b}})}{(1-(.9)^{\frac{1}{b}})} dB$. Since(recall chapter 1), we targeted to estimate Singnal to Noise Plus Distortion of the receiver(SINAD) with an accuracy of $\pm 50\%$ or $\pm 1.8dB$, a reasonable choice of parameters is $a = 1.6, b = 15$ which results in $x_L = .6, x_H = 1.8$. $G(x)$, on the other hand, must constrain gain to remain as close as possible to the nominal value G_{nom}. This is in contrast with optimization convergence speed, and a tradeof must be sought. We chose

$$H(x) = \mid (\tanh (\frac{G - G_{nom}}{a^{'}})^{b^{'}} \mid \tag{5.41}$$

with $a^{'} = 3, b^{'} = 4$. Plots of $H(x)$ are reported in figure 5.24 for different value of $a^{'}$ ad $b^{'}$.

3. If $C_n < C^*$, optimal point is updated $C^* = C_n$, $x^* = x_n$. If on the other hand $C_n \geq C^*$, a random variable X is sampled from a single-sided exponential deistribution of mean $k \cdot T_n$. Call this sample x. If $x \geq C_n - C^*$, then optimal point is updated, $x^* = x_n$, $C^* = C_n$.

4. According to a suitable algorithm, lattice temperature T_{n+1} is calculated

The process arrests after N_{stop} steps, or when more generally $T = T_{stop}$ is reached. Results from an optimization run are reported in tables 5.3-5.5. For each configuration, both the estimated optimal point (i. e. the optimal point calculated on the performance space interpolated by the statistical classifier as described in [56]) and its three nearest simulated neighbours are reported(distance is measured in R^N). It is interesting to compare results obtained through the platform based design flow with results reported in [2] and summarized in 5.9. First, a reduction in power dissipation of about $3.5mW(15\%)$ is predicted with respect to [2] while at the same time higher gain($35dB$) and IIP_2 are obtained. This increased IIP_2 results in increased noise immunity, so that a sensibly higher noise figure ($7.1dB$) is allowed to the receiver chain. IIP_3 is instead slightly worse ($1dBm$ lower). According to [62] the resulting increase in gain could actually positively impact the design of baseband sections, reducing their power consumption. This is therefore a welcome modification. On the other hand, the extremely high mixer IIP_2 ($94dBm$), is difficultly achievable in practice, although it was simulated using 3σ worst case random offset generators in the mixer as reported in [61] and in [1]. Authors of [2] found an $8dBm$ difference between simulated and measured values of IIP_2. Mismatch between load resistors is blamed for this performance degradation. As this contribution to IIP_2 was also neglected in this work, and considering the possible inaccuracie fo

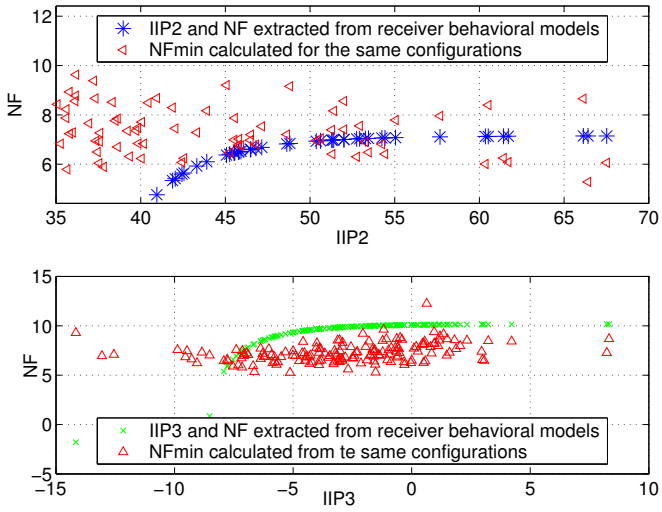

Figure 5.25: Exploring the size of the design space:results of equations 2.38 and 2.39, when calcultated on modeled receiver instances; feasible configurations have red triangles between both the gree and the blue patterns.

k_2 extraction, mixer optimal solution cannot be completely trusted. This feeling is further confirmed by large differences in IIP_2 values between this point and its nearest neighbours. At the same time LNA instance selected is probably penalized in term of noise optimal gate width by the ACG. In previous characterizations, noise figures as low as $3dB$ were obtained for the same power dissipation. If a more reasonable(though still extremely high) value of $70dBm$ is assumed for the mixer IIP_2(corresponding to NN3), in conjunction with the same Low Noise Amplifier used above we obtain receiver configuration $NN2$, which shows performances similar to that reported in [2], expect for having a noise figure $.5dB$ higher than the minimum allowed by equation 2.40. It is interesting to notice that, either due to the bad noise performances of the Low Noise Amplifiers characterized using the first version of the advanced Analog Constraint graph, or to the extremely narrow nature of the acceptable design space defined by equations 2.40, 2.39, the choice of extremely high linearity mixers seems unavoidable. Also, for the values of k_2 and G_T reported in optimal configuration, optimal value of gain predicted by equation 2.43 for the Low Noise amplifier was $20dB$. As expected, optimal configurations confirm the intuition that LNA gain values higher than the one predicted by 2.43 yeld better system level performance. Detailed circuit-level sizing for this configuration is reported in table 5.10. At present, results of the optimizer runs are still being interpreted and discussed.

Front End	
Noise Figure	$5.4dB$
IIP_2	$48dBm$
IIP_3	$-2dBm$
P_d	$17mW@1.8V$
LNA	
Gain	$18dB$
NF	$2.6dB$
Mixer	
IRN	$4.15nV/\sqrt{Hz}$
CG	$13.5dB$
IIP_2	$66dBm$
IIP_3	$0dBm$

Table 5.9: Performances extracted from [2] for the LNA+Mixer front end

I-Space parameter	NN1	NN2	NN3	LNA in [2]
L_g	$7.6nH$	$6.75nH$	$5.75nH$	$5.1nH$
L_s	$500pH$	$500pH$	$500pH$	$1.8nH$
L_n	$.18\mu$	$.19\mu$	$.18\mu$	$.18\mu$
W_n	570μ	690μ	780μ	248μ
$I_d(perbranch)$	$1.5mA$	$2.5mA$	$2mA$	$2mA$
L_{load}	$3.11nH$	$2.83nH$	$3.26nH$	$2.8nH$
C_{mix}	$435fF$	$412fF$	$514fF$	N/A
C_{Load}	$1.06pF$	$1.15pF$	$.943pF$	N/A

Table 5.10: Summary of LNA input space parameters for the optimized designs and the design in [2]

Chapter 6

Conclusions

System-level optimization of a UMTS front end was developed in this work using and expanding the concept of Analog Platforms([39]). Challenges posed by the application of the platform paradigm regard the design of Analog Constraint Graphs used to generate acceptable circuit configurations , and the development of accurate system-level behavioral models. These challenges require device, circuit and system-level knoledge to interact. To meet the first of these challenges, the study and development of physics-based accurate device models, the in depth analysis of noise and distortion behavior of CMOS Low-Noise Amplifiers, and the careful analysis of the effect of LNA and Mixer block performances on figures of merit of the whole receiver were undertaken. Based on this study, different algorithms to automatically generate CMOS low-noise amplifiers were implemented and their accuracy and effectiveness in describing the design space of the modeled blocks studied and discussed. The proposed solutions have been tested with success on different operating bandwidths without modifications, and possibilities exixst to export them unchanged to different techonolgies.
Behavioral modeling of a UMTS receiver was at the same time faced. Different representations of nonlinear systems and different models of computation have been investigated with regard to their composability, accuracy and complexity properties and succesively implemented. Then, sensitivity analysis was applied to equations assessing compliance of a direct conversion receiver to UMTS standard to derive accuracy requirements for building blocks' models from imposed accuracy on receiver SINAD. This effort allowed quantitatively exploring accuracy tradeoffs amongst blocks linear and nonlinear responses; to the author's knowledge, no previous behavioral modeling effort in the field of electronic system level design reports such an approach.
Particular effort was required by understanding problems arising from communication between LNA and Mixer. This communication influences both the linear and nonlinear response of the Low-Noise Amplifier, as well as of the overall receiver. Using small signal and Volterra analysis, equations describing mixer input impedance and capacitive nonlinearity at the LNA-Mixer interface were derived. For this particular topic, dependence on circuit parameters was individuated and studied, finally proposing a block composition rule for joining "compatible blocks" and an LNA-Mixer cascade behavioral model that should guarantee more accurate calculations of receiver linear and nonlinear response for all possible pairs of blocks labeled "compatible" than the ones previously proposed in literature.
Finally, numerical optimization of the receiver chain was performed, using a simulated annealing algorithm. Results of the optimization present an improvement of about 30% in power dissipation of the front end with respect to the solution proposed in [2], while still respecting the UMTS standard requirements. The presented work is the result of different iterations spent either in behavioral modeling, optimizer runs and in ACG design. First optimization runs were performed in december; since then much effort has been spent in understanding what effects were responsible for the model inaccuracies, and afterwards in modeling distortion due to the interface. Many simulations campaigns, along with a

PSS+PAC semi analytical interface model, were performed without success, and therefore have not been reported. Bugs in various parts of the developed tools were individuated and corrected. At the same time, also because of my personal interest in device modeling, the reasons of the success of the first developed "Naive" ACGs was investigated, to arrive at a formulation of LNA design equations accounting for all the interesting subthreshold-conduction related effects which had been only qualitatively understood in the early phase of the work, arriving at a more complete understanding of Low-Noise Amplifier design. Despite the efforts, some results are not yet stable, as numerous open issues remain, especially for what concerns blocks communication and behavioral model accuracy; on the other hand the developed methodology, although not yet mature for becoming an automated design tool, can certainly be satisfactorily used for performing the early stages of analog system-level design, obtaining a first circuit sizing to be later hand-refined. And even though this work, requiring to get confidence in the use of tools such as $MATLAB^R, SIMULINK$, Mathematica, $CADENCE$ and *Ocean*, produced its results with a much slower pace than a traditional approach would have done, a basis has been posed here , in [1], and in [40], that will allow future applications and developments of the methodology to be ready in a much shorter time, making its use in real design practice at a least reasonable choice.

Appendix A

Short-Channel MOS Model equations

The model we developed in section 3.3 neglects high field effects such as mobility degradation due vertical and horizontal field. As such, it is certainly well suited for the design of Low-Noise Amplifiers in the low GHz region, where we have seen operating points shift towards moderate or weak inversion regions. Higher operating frequencies (e.g. the WLAN case, but even more the emerging 20 and $60GHz$ bandwidths), or different circuits may require these effects to be included in the model. Moreover, initial errors in the derivation of Drain Induced Gate Noise power spectral density led to the misbelief that non-convergence of the integral expressing S_{ig} were due to the zero drain charge predicted by the long channel model. Then, an effort was made in deriving a self-consistent bulk charge linearized model using the methodology described in 3.3. The derivation is rigorous only for the lateral-field due mobility degradation, which was treated introducing for mobility the usual expression:

$$\mu_n = \frac{\mu_{eff}}{1 + \frac{E}{E_{sat}}}$$

Therefore, equation for current density changes into:

$$I_d(1 - \frac{\frac{\partial \Psi_s}{\partial x}}{E_{sat}}) = \mu_n(Q_n(x)\frac{\partial \Psi_s}{\partial x} - V_{th}\frac{\partial Q_n}{\partial x}) \tag{A.1}$$

Where $\Psi_s, Q_n(x) and V_{th}$ respectively represent surface potential mobile channel charge and thermal equivalent of voltage.Multipliyng both sides times dx and integrating one obtains:

$$I_F = \frac{(f_S - 1)^2 - (f_D - 1)^2)}{1 + \frac{f_S - f_D}{Q}_{sat}} \tag{A.2}$$

$$Q_{sat} = \frac{E_{sat} * L}{nV_{th}} \tag{A.3}$$

$$I_0 = \mu_n C_{ox}\frac{W}{2L}nV_{th}^2 \tag{A.4}$$

Unlike the case of long channel devices, the existence of finite maximum velocity for electrons in the channel makes so that $f_D \neq 0$.In particular, $f_D = \frac{I_D}{V_{sat}W*n*Cox*V_{th}}$ ([66]) is a the minimum value for f_D allowing drain to be velocity saturated and has been chosen as boundary condition in this work. Boundary condition sensibly complicates calculations. It is important to notice that defining

$I_2 = V_{sat}W * nC_{ox}V_{th}$, $I_2/(I_0\frac{W}{L}) = \frac{2V_{sat}L}{\mu_n V_{th}} = \frac{4E_{sat}L}{V_{th}} = 4nQ_{sat}$. Equation A.2 may be solved for f_S or IC yelding respctively

$$f_S = \frac{-2ICn + 4nQ_{sat} + \sqrt{IC^2(1-2n)^2 + 8ICn(1+2n(-1+Q_{sat}))Q_{sat} + 16n^2{Q_{sat}}^2}}{4nQ_{sat}} \tag{A.5}$$

and

$$IC = \frac{4nQ_{sat}(1 - 2nf_S + 2nQ_{sat} + \sqrt{(1-2n)^2 f_S^2 + (1+2nQ_{sat})^2 + f_S(-2+4n-8n^2Q_{sat})})}{-1+4n} \tag{A.6}$$

Integrating A.1 between 0 and the generic abscissa x allows to obtain the inversion layer charge at any point along the channel. The result is:

$$Q_{inv}(x) = \frac{-(ICL) + 2LQ_{sat} + \sqrt{(-(ICL)+2LQ_{sat})^2 + 4LQ_{sat}(ICLf_S - ICxQ_{sat} - 2Lf_SQ_{sat} + Lf_S^2Q_{sat})}}{2LQ_{sat}} \tag{A.7}$$

Now we can go the small signal parameters estimation, that still require knowledge of $\frac{\partial\Psi_S}{\partial V_S}$. Once again, we go back to IMREFs to develop calculations about g_m. The IMREF-based current density equation becomes now:

$$I_D(1 + \frac{\Psi_D - \Psi_S}{E_{sat}L}) = -\mu_n\frac{W}{L}\int_{V_S}^{V_D} Q_{inv}(V)dV \tag{A.8}$$

Taking derivatives with respect to V_S one finds

$$\frac{\partial I_D}{\partial V_S} - \frac{I_D}{E_{sat}L}\frac{\partial\Psi_S}{\partial V_S} = \mu_n\frac{W}{L}f_S \tag{A.9}$$

Exact evaluation of the previous expression requires solving for $\frac{\partial\Psi_s}{\partial(V_s)}$.This can be done by combining equation A.9 with the derivative of equation A.6 with respect to f_S.

$$G_{ms} = \frac{I_0}{V_{th}}\frac{W}{L}(\frac{4nQ_{sat}(-2n + \frac{1-f_S+n(-2+4f_S)-4n^2(f_S-Q_{sat})}{\sqrt{(1-2n)^2f_S^2+(1+2nQ_{sat})^2+f_S(-2+4n-8n^2Q_{sat})}})}{-1+4n})\frac{\partial\Psi_S}{\partial V_S}$$

Substituting this expression into equation A.9 allows accurate estimation of $\frac{\partial\Psi_S}{\partial V_S}$. However, resulting expressions are very complex and will not be reported.Small signal gate-source capcitance has been evaluated by using the formula

$$C_{gs} = \frac{\partial Q_{inv}}{\partial IC}\frac{\partial IC}{\partial V_S} = \frac{\partial Q_{inv}}{\partial IC}\frac{G_m}{I_0\frac{W}{L}} \tag{A.10}$$

We obtain Q_{inv} integrating equation A.7 between 0 and L after substituting A.5:

$$Q_{inv} = \frac{-(L(48IC^2Q_{sat}^2 - 96ICQ_{sat}^3 - \frac{((IC(-1+2n)-4nQ_{sat})^2)^{\frac{3}{2}}}{n^3} + \frac{(16ICn^2Q_{sat}^2+(IC(-1+2n)-4nQ_{sat})^2)^{\frac{3}{2}}}{n^3}))}{96ICQ_{sat}^3} \tag{A.11}$$

Differentiating this equation with respect to IC:

$$-(L(2 + (IC^2(1-2n)^2 + 2IC(1-2n)nQ_{sat} - 8n^2Q_{sat}^2)\cdot F \tag{A.12}$$

$$F = \frac{(-\sqrt{(IC(-1+2n)-4nQ_{sat})^2} + \sqrt{16ICn^2Q_{sat}^2 + (IC(-1+2n)-4nQ_{sat})^2}))}{48IC^2n^3Q_{sat}^3} \tag{A.13}$$

Appendix B

Input Matching equations used in "Naive" NPMOS amplifier

For the coefficients $a_1, a_2, b_1, b_2, c_1, c_2$ in 5.10, expressions reported below have been used(PMOS-branch is supposed sized):

$$C_i = \frac{C_{gs}^i C_{blk}}{C_{blk} + C_{gs}^i} \quad i = n, p \tag{B.1}$$

$$R_p = \omega_{T_p} L_{sP} \tag{B.2}$$

$$X_p = (\frac{-1}{\omega_0 C_p} + \omega_0 L_{sP}) \tag{B.3}$$

$$Z_p = R_p + jX_p a_1 = R_s Q_s(\omega_{T_n}^2 + \omega_0^2)(1 + 1/Q_g^2) \tag{B.4}$$

$$a_2 = (\omega_{T_n}^2 + \omega_0^2)(R_p - Q_g X_p)(1 + 1/Q_g^2) \tag{B.5}$$

$$b_1 = 2R_s Q(R_p(\omega_{T_n} + \omega_0/Q_g) - \frac{1}{C_n}(1 - \frac{\omega_{T_n}}{\omega_0 Q_g}) + X_p(\omega_0 - \omega_{T_n}/Q_g)) \tag{B.6}$$

$$b_2 = \frac{-2}{C_n}(1 - \frac{\omega_{T_n}}{\omega_0 Q_g})(Q_p R_p - X_p) + (\mid Z_p \mid)^2 \omega_{T_n} \frac{Q^2 + 1}{Q} \tag{B.7}$$

$$c_1 = (\mid Z_p \mid)^2 + (\frac{1}{\omega_0 C_n})^2 - \frac{2X_p}{\omega_0 C_n})R_s Q_g \tag{B.8}$$

$$c_2 = \frac{(\mid Z_p \mid)^2}{\omega_0 C_n} + (Q_g R_p - X_p)(\frac{1}{\omega_0 C_n})^2 \tag{B.9}$$

Appendix C

Effect of Pad Capacitance on input matching equations

As stated above, the effect of ESD protection diodes on input matching equations was not considered in simulations performed during this system-level study. For the sake of completeness, we report below some considerations regarding this effect. Supposing that the impedance seen looking right from the chip input is represented as a series RC branch as in figure C.1, impedance seen looking to into the chip becomes

$$R_{in} = \frac{R}{2x + x^2 + (1 + {C_2}^2 R^2 \omega^2)} \tag{C.1}$$

$$X_{in} = \frac{1 + x + xC_1^2 R^2 \omega^2}{\omega C_1 (2x + x^2 + (1 + {C_2}^2 R^2 \omega^2))} \tag{C.2}$$

Where $x = C_2/C_1$. C_1 is the residual equivalent input capacitance after on-chip portion of degeneration inductance has been added, so that if $L_g^{OnChip}/L_g^{Tot} = F$, $C_1 = C_g(1-F)$. For F=1, $x \to 0$, and classical formulas for a parallel to series conversion of an RC branch are found([57]). Magnitude of x factor is reported in figure C.2 assuming device gate widths ranging from 50μ to $1.2mm$ are used and input matching to a 25 Ω termination. Notice that in this case, assuming $R \approx 25$ and $C_2 \approx 1pF$,contribution from the $\omega_0^2 C_2^2 R^2$ is lower than .1 and may be neglected. Input Normalized input resistance and capacitance plots calculated under the same assumption are also shown in figure C.2. It can be readily seen how for narrow transistors(i.e. low C_1) a devastating effect is obtained. However, the situation gets rapidly better at higher gate widths. Also notice that this effect becomes a concern only for what regards the real part, as imaginary part may still be adjusted using bond wire and off-chip inductors once the input impedance is measured. A designer may then cope with this effect designing input stage to be matched to a $(x+1)^2$ times larger input resistance, using a suitable estimate of x.

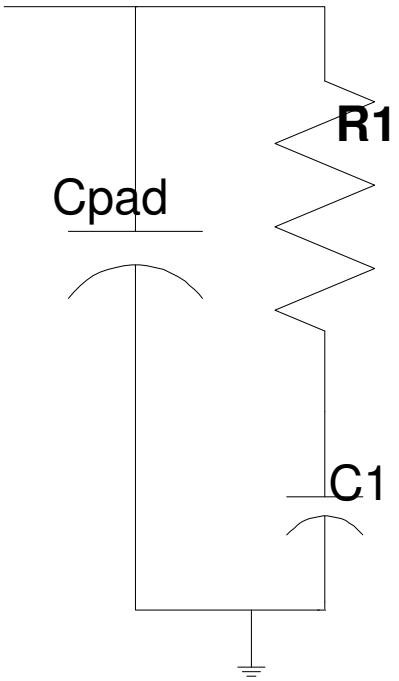

Figure C.1: Equivalent circuit for calculating Pad influence

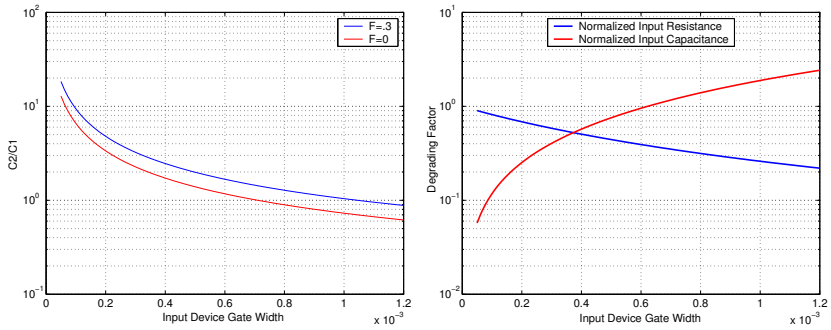

Figure C.2: C_1/C_2 ratio versus gate width(left) and degrading factors of chip input impedance(right)

Appendix D

Third order distortion arising from second order nonlinear terms

Second order nonlinear terms may interact with linear responses to produce third order intermodulation products. This effect is due to the local feedback introduced by inductive degeneration and is described in detail in [12]. To account for this effect(which we expect to be negligible, containing terms of the kind gm_2^2), $v_{gs}/I_{nd} = S_1$ and $v_{gs}/i_{ng} = S_2$ transfer functions are needed. These trasnfer functions have been previously derived and their expressions are $S_1 = \frac{\alpha_{nd}}{1+gm\alpha_{nd}}$
$S_2 = \frac{\alpha_{ng}}{1+gm\alpha_{nd}}$. For $\omega \approx 0$ and $\omega \approx 26Grad/s$, $gm\alpha_{nd} \ll 1$holds, so that $S_1 \approx \alpha_n d = R_{Ls}$,$S_2 \approx \alpha_{ng} = r_{nqs} + R_s + R_{Ls} + R_{Lg}$. Third order intermodulation distortion generated by second order nonlinearities may then be expressed as

$$Iu_{IM3} = \frac{1}{4}\alpha_1^2(\omega_1)\alpha_1(\omega_2)F \quad \text{(D.1)}$$

$$F = (\omega_{T_{eff}}(K_1^{C_{gs}})^2(\omega_1 - \omega_2)(r_{nqs} + R_s + R_{Lg} + R_{Ls})(1 - jQ_L/D) + gm_2R_{Ls}(gm_2 + 2j(2\omega_1 - \omega_2)K_1^{C_{gs}}))H$$

$$H = \frac{1}{1 + \frac{\omega_{T_{eff}}L_s}{DR_s}}$$

Bibliography

[1] F.Vincis *Metodologie per la caratterizzazione e la progettazione di circuiti e sistemi a radiofrequenza*Ph.D. Thesi, University of Pisa, To Appear in 2005

[2] F.Gatta,P.Rossi et F.Svelto *A Fully Integrated .18μ CMOS direct-conversion receiver front-end with on-chip LO for UMTS* IEEE Journal of Solid Sate Circuits, January 2004

[3] A.Demir e A.L.Sangiovanni Vincentelli *Analysis and simulation of noise in nonlinear elctronics circuits* Kluwer Academic Publishers

[4] K.Kundert *Introduction to RF simulation and its applications* IEEE Journal of Solid State Circuits, 1999

[5] D.K.Shaeffer *A 1.5V 1.5 Ghz CMOS Low Noise Amplifier* IEEE Journal of Solid State Circuits, June 1997

[6] T.H. Lee *The Design Of CMOS Radio Frequency Integrated Circuits,*2^{nd} *edition* Cambridge University Press, 2004

[7] K.Ogata *Modern Control Engineering* Prentice Hall,1973

[8] Iason Vassilio *Design Methodologies ...* Ph.D.Thesis Univeristy Of California at Berkeley

[9] Dennis Yee *A design methodology for highly integrated low-power receivers for wireless communications* Ph.D.Thesis,University of California At Berkeley.

[10] MATHWORKS *Model Construction Commands* in $MATLAB^R$ reference.

[11] F.Gatta *A 0.18μm CMOS UMTS Direct-Conversion Receiver* Tesi Di Dottorato, Universitá Degli Studi di Pavia

[12] P.Wombaq,W.Sansen *Distortion Analysis of Analog Integrated Circuits* Kluwer Academic Publishers 1999

[13] M.DelMar Hershenson,S.Boyd *Optimal design of CMOS Op-Amps via geometric programming*Applied and Computational control,signals,circuits and systems,Vol.2 Birkhauser 2000

[14] M.DelMar Hershenson,S.Boyd *Optimization of Inductor circuits via Geometric Programming* in Proceedings of the 36^{th} design automation conference, 1999

[15] A.Papoulis *Probability and Random Processes* McGraw-Hill

[16] F.DeGraff, D.B.M. Klassen *Compact Transistor Modeling for Circuit Design* Springer 1990

[17] MATHWORKS Inc. *Matlab 6.5 Reference*

[18] Wolfram Research *Mathematica 5.0*

[19] CADENCE Design Systems *Ocean Reference*

[20] S.Boyd O.L.Chua *Measuring Volterra Kernels* IEEE Transactions on Circuits And Systems,Vol.CAS-30,N.8,August 1983

[21] S.Boyd,O.L.Chua *Fading Memory and The Problem of Approximating Nonlinear Operators with Volterra Series* IEEE Transactions on Circuits And Systems,Vol.CAS-32,N.11,November 1985

[22] A.Sahani *Frequency Synthesis and Conversion in RF CMOS* Ph.D. Stanford Univeristy 1999

[23] S.Haykin *Communication Systems,* 4^{th} *edition* John Wiley And Sons, 2000

[24] A.Hajimiri *The Design Of Low-Noise Oscillators* Kluwer Academic Publishers, 1999

[25] Y.Tsividis *Operation and Modeling of The MOS Transistor* 2^{nd} edition, McGraw-Hill 1999

[26] R.Courant,H.Robbins *Che cosa é la matematica* Boringhieri

[27] Y.Taur,H.Ning *Modern VLSI devices* Oxford Univeristy Press, 1998

[28] P.Gray,R.Meyer et al. *Analysis and Design of Analog Integrated Circuits* John Wiley and Sons,2002

[29] Philips Semiconductors *JUNCAP Model*

[30] Philips Semiconductors *The MOSFET model, level 903*

[31] C.Galup Montoro et al. *An MOS transistor transistor model for analog circuit design* IEEE Journal of Solid State Circuits,Vol.33,n.10,October 1998

[32] C.Galup Montoro et al. *An MOS in MOSFET-C circuits* IEEE Journal of Solid State Circuits, 2000

[33] B.Razavi *RF Microelectronics* Prentice-Hall 1997

[34] C.C.Enz,E.Vittoz,F.Krummenacher *An analytical MOS transistor model in all regions of operation and dedicated to low-voltage and low-current operation* Analog Integrated Circuits And Signal Processing, Vol.8,no.1,1995

[35] Ronald van Langevelde,Francois M.Klaassen *An explicit, surface-potential-based MOSFET model for circuit simulation* Solid State Electronics 2000

[36] P.Woerlee,M.J.Knithel et al *RF-CMOS Performance Trends* IEEE Transactions on Electron Devices,Vol.48,NO.8,August 2001

[37] C.Enz *MOS Transistor Modeling for RF IC Design* IEEE Transactions on Solid State Circuits,VOL.35,No.2,February 2000

[38] I.Vassiliou et al. *A Top-Down Constraint Driven Design Methodology for Analog Integrated circuits* Kluwer Academic Publishers 1997

[39] L.Carloni,F.De Bernardinis e A.L.Sangiovanni Vincentelli *The art and science of integrated systems design* Proceedings of ESSCIRC 2002

[40] F.DeBernardinis *Design methodologies for analog RF and mixed signal circuits* Ph.D. Thesis, Univerisity of California at Berkeleey To Appear in 2005

[41] H.C.Pao, Sah *effect of diffusion current on characteristics of metal-oxide (insulator) semiconductor transistor* Solid State Electronics Oct.1966

[42] P.Arcioni,Rinaldo Castello,G.De Astis,E.Sacchi and F.Svelto *Design and Characterization of Si Integrated Inductors* In Proceedings of IEEE Instrumentation and measurement techonology conference, 1999

[43] L.Ingber *Adaptive Simulated Annealing* www.ingber.com

[44] P.Mendonca L.Calboa *New Simulated Annealing Algorithms* IEEE International Symposium on Circuits and Systems, 1997

[45] Jan Craninckx,M.Steyaert *A 1.8Ghz Low-Phase-Noise CMOS VCO Using Optimized Hollow Spiral Inductors* IEEE Journal of Solid State Circuits, May 1997

[46] G.Corsini,R.Saletti $1/f^{\nu}$ *power spectrum noise sequence generator* IEEE

[47] Chemming Hu et Al. *BSIM3v3 MOSFET Model users' Guide* Department of EECS, Univerisity of California at Berkeley, 1999

[48] F.Hillier,G.Lieberman *Introduction to Operations Research* McGraw-Hill

[49] R.G.Meyer *Distortion in variable capacitance diodes* IEEE Journal of Solid State Circuits, Feb. 1975

[50] D.Manstretta *A* $0.18\mu m$ *CMOS UMTS Direct-Conversion Receiver,Part II* Tesi Di Dottorato, Universitá Degli Studi di Pavia

[51] D.Manstretta,M.Brandolini e F.Svelto *Second Order Intermodulation Distortion In CMOS Downconverters* IEEE Journal of Solid State Circuits,Vol.38,Issue 3,March 2003

[52] M.Terrovitis,R.G.Meyer *Intermodulation Distortion In Current Commutating CMOS Mixers* IEEE Journal Of Solid State Ciruits,Vol.35,N.10,October 2000

[53] Y.Tsividis,G.Masetti *Problems in Precision Modeling of The MOS transistor for Analog Designs* IEEE Transactions on computer Aided Design of Integrated Ciruits and Systems,Vol.3,Issue 1,January 1984

[54] B.Razavi *Design considerations for Direct Conversion receivers* IEEE Transaction on Circuits and Systems II,Vol.44,June 1997

[55] F.Svelto,A.Bevilacqua *Nonlinear spectral analysis of wireless direct conversion receivers* in Proceedings of the 2002 IEEE Radio And Wireless Conference, RAWCON 2002

[56] F.DeBernardinis,M.Jordan and A.SanGiovanni-Vincentelli *Support Vector Machines for Analog Design Space Exploration* in Proceedings of IEEE Design Automation Conference,June 2003

[57] Krauss,B.Raabe *Solid State Radio Engineering* J.Wiley & Sons, 2000

[58] H.Nijemeier,A.Van Der Schaft *Nonlinear Dynamical Control Systems* Springer-Verlag, 1990

[59] D.Belot et al. *A MEMS over IC RF switch* in Proceedings of ISSCC 2003

[60] Fernando DeBernardinis *Heuristic for performance model generation* Report, April 2003

[61] P.Rossi, *Microlab Internal Report* May 2003

[62] R.Castello *Personal Communication*

[63] P.Nuzzo *Caratterizzazione dello spazio architetturale di un amplificatore transconduttivo* Tesi Di Laurea,Universitá di Pisa, Ottobre 2003

[64] A.Sedra,K.Smith *Microelectronic circuits* Ocford University Press, 2003

[65] A.Van Der Ziel *Noise in solid-state devices and circuits* John Wiley Interscience, 1985

[66] C.Galup-Montoro et Al. *The ACM Model for circuit simulation and equations for SMASH* available online at www.dolphin.fr(Dolphin Integration website)