ANALOG CIRCUIT DESIGN

Volt Electronics; Mixed-Mode Systems; Low-Noise and RF Power Amplifiers for Telecommunication

Edited by

Johan Huijsing
Delft University of Technology, Delft, The Netherlands

Rudy van de Plassche
Eindhoven University of Technology, Eindhoven, The Netherlands

and

Willy Sansen
Katholieke Universiteit Leuven, Belgium

KLUWER ACADEMIC PUBLISHERS
BOSTON / DORDRECHT / LONDON

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 0-7923-8400-8

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in North, Central and South America
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands

Table of Contents

Preface

This book contains the revised contributions of 18 tutorial speakers at the seventh AACD '98 in Copenhagen, April 28-30, 1998. The conference was organized by Ole Olesen, of the Technical University of Denmark. The program committee consisted of Johan H. Huijsing from Delft University of Technology, The Netherlands, Willy Samsen from the Katholieke Universiteit Leuven, Belgium and Rudy J. van de Plassche, Philips Research, The Netherlands.

The program was concentrated around three important topics in analog circuit design. Each of these three topics has been covered by six papers. Each of the three chapters of this book contains the six papers of one topic. The three topics are:

1-Volt Electronics
Design and implementation of Mixed Modes Systems.
Low-Noise and RF power Amplifies for the communication.

Other topics, which have been covered in this series before are:

1992	OpAmps ADC's Analog CAD.
1993	Mixed-Mode A/D design Sensor Interfaces Communication circuits.
1994	Low-Power low-Voltage Integrated Filters Smart Power.
1995	Low-Noise, Low-Power, Low-Voltage Mixed Mode with CAD Tirals Voltage, Current and Time References.

1996 RF CMOS circuit design
Bandpass Sigma Delta and other Converters
Translinear circuits.

1997 RF A-D Converters
Sensor and Actuator Interfaces
Low-noise Oscillators, PLL's and and Synthesizers.

We hope to serve the analog design community with these series of books and plan to continue this series in the future.

Johan H. Huijsing

Preface
1 – Volt Electronics

The strive for more electronics on a chip requires smaller transistor dimensions. This in turn results in lower breakdown voltages, in the order of 3 or 2 Volt for mainstream CMOS processes. Hence, together with low-cost battery use in wireless application, the supply voltages are forced in the direction of 1.8 Volt and 0.9 Volt.

At such low supply voltages not only digital signal processing but also analog signal processing has to be performed. The latter imposes a challenge on the analog circuit designer who finds the dynamic range of analog signals squeezed between the supply voltage roof and the noise floor.These circuit design challenges are met by the following 6 papers on 1- Volt electronics.

In the first paper by Wouter Serdijn, TU Delft, the Netherlands, a general approach for 1- Volt translinear circuit design is presented. The second paper by Christian Enz and Manfred Punzenberger, Rockwell, USA, evaluates this approach for 1- Volt log – domain filter design.

The third paper by Andrea Baschirotto and Rinaldo Castello, Univ. of Pavia, Italy, describes how a switched capacitor filter can be made functioning at a supply voltage of 1 Volt using a switched opamp technique. The fourth paper by Vincenzo Peluso and Michiel Steyaert, KU. Leuven, Belgium, shows how a sigma delta modulator can be designed for 1 Volt.

As lowcost battery use for pagers is desired, the design of a 1 Volt complete RF front end is presented by Ed Callaway, Motorola, USA, in the fifth paper .

Finally to breake away from the low-voltage battery supply problem, the sixth paper by Carel Dijkmans, Philips Research, the Netherlands, describes the design of DC – DC upconverters that may function at a primary voltage of 1 Volt.

Johan H. Huijsing

Dynamic Translinear Circuits

Wouter A. Serdijn, Jan Mulder,
Paul Poort, Michiel Kouwenhoven,
Arie van Staveren and Arthur H.M. van Roermund
Delft University of Technology,
Faculty of Information Technology and Systems/DIMES
Electronics Research Laboratory
Mekelweg 4, 2628 CD Delft, The Netherlands
phone: +31-15-2781715, fax: +31-15-2785922,
e-mail: W.A.Serdijn@its.tudelft.nl

Abstract

A promising new approach to shorten the design trajectory of analog integrated circuits without giving up functionality is formed by the class of dynamic translinear circuits. This paper presents a structured design method for this young, yet rapidly developing, circuit paradigm. As a design example, a 1-V 1.6-μA class-AB translinear sinh integrator for audio filter applications, is presented.

Introduction

Electronics design can be considered to be the mapping of a set of mathematical functions onto silicon. For discrete-time signal-processing systems, of which the digital signal processors (DSPs) today are by far the most popular, this comes down to the implementation of a number of difference equations, whereas for continuous-time signal-processing systems, often denoted by the term analog,

differential equations are the starting points. In mixed analog-digital systems, the analog parts, however, often occupy less than ten percent of the complete, i.e., the mixed analog-digital circuitry, whereas their design trajectory is often substantially longer and therefore more expensive than of their digital counterparts. Where does this discrepancy arise from? This can be partially explained by the fact that, at circuit level, for analog circuits far more components play an important role; various types of transistors, diodes, resistors and capacitors, to mention a few; sometimes also inductors, resonators, and others. Whereas for digital circuits, the complete functionality is covered by transistors only[1].

From the above, it automatically follows that, if we restrict ourselves to the use of as few different types of components as possible, without giving up functionality, we can shorten the analog design trajectory considerably, in the same way as this is done for digital circuits. One successful approach, as we will see in this paper, is given by the class of circuits called *dynamic translinear circuits*.

Dynamic translinear (DTL) circuits, of which recently an all-encompassing current-mode analysis and synthesis theory has been developed in Delft [1–3], are based on the DTL principle, which can be regarded as a generalization of the well-known 'static' translinear principle, formulated by Gilbert in 1975 [4]. The first DTL circuit was originally introduced by Adams in 1979 [5], being a first-order lowpass filter. Although not recognized then, this was actually the first time a first-order linear differential equation was implemented using translinear (TL) circuit techniques. In 1990, Seevinck introduced a 'companding current-mode integrator' [6] and since then the principle of TL filtering has been extensively studied by Frey [7–16], Punzenberger and Enz [17–31], Toumazou et al. [32–51], Roberts et al. [52–57], Tsividis [58–62], Mulder and Serdijn [63–84] and others [85,86].

However, the DTL principle is not limited to filters, i.e. linear differential equations. By using the DTL principle, it is possible to

[1]It must be noted that, for higher frequencies or bit rates, also the interconnects come into play. However, their influence is considered to be equally important for analog as well as digital systems.

implement linear *and* nonlinear differential equations, using transistors and capacitors only. Hence, a high functional density can be obtained, and the absence of large resistors makes them especially interesting for ultra-low-power applications [76].

DTL circuits are inherently companding (the voltage swings are logarithmically related to the currents), which is beneficial with respect to the dynamic range in low-voltage environments [87,88]. In addition, DTL circuits are easily implemented in class AB, which entails a larger dynamic range and a reduced average current consumption. Further, owing to the small voltage swings, DTL circuits facilitate relatively wide bandwidth operation. At high frequencies though, considerable care has to be taken regarding the influence of parasitic capacitances and resistances, which affect the exponential behavior of the transistor.

DTL circuits are excellently tunable across a wide range of several parameters, such as cut-off frequency, quality factor and gain, which increases their designability and makes them attractive to be used as standard cells or programmable building blocks.

The DTL principle can be applied to the structured design of both linear differential equations, i.e. filters, and non-linear differential equations, e.g., RMS-DC converters [89–91], oscillators [92–103], phaselock loops (PLLs) [80–82] and even chaos. In fact, the DTL principle facilitates a direct mapping of any function, described by differential equations, onto silicon.

Application areas where DTL circuits can be successfully used include audio filters, high-frequency filters, high-frequency oscillators, demodulators, infra-red front-ends and low-voltage ultra-low-power applications.

This paper aims to present a structured design method for DTL circuits. The static and dynamic TL principles are reviewed in Section 2. The general class of DTL circuits contains several different types. In Section 3, the correspondences and differences of log-domain, tanh and sinh circuits are treated. Finally, Section 4 presents the design method, applied to the design of a DTL integrator, starting from a dimensionless differential equation that describes the integrator behavior in the time domain. After four

hierarchical design steps, being dimension transformation, the introduction of capacitance currents, TL decomposition and circuit implementation, a complete circuit diagram results. Measurement results of the thus obtained DTL integrator, are presented.

2 Design principles

TL circuits can be divided into two major groups: static and dynamic TL circuits. The first group can be applied to realize a wide variety of linear and non-linear static transfer functions. All kinds of frequency-dependent functions can be implemented by circuits of the second group. The underlying principles of static and dynamic TL circuits are reviewed in this section.

2.1 Static translinear principle

TL circuits are based on the exponential relation between voltage and current, characteristic for the bipolar transistor and the MOS transistor in the weak inversion region. In the following discussion, bipolar transistors are assumed. The collector current I_C of a bipolar transistor in the active region is given by:

$$I_C = I_S e^{V_{BE}/V_T}, \tag{1}$$

where all symbols have their usual meaning.

The TL principle applies to loops of semiconductor junctions. A TL loop is characterized by an even number of junctions [4]. The number of devices with a clockwise orientation equals the number of counter-clockwise oriented devices. An example of a four-transistor TL loop is shown in Figure 1. It is assumed that the transistors are somehow biased at the collector currents I_1 through I_4. When all devices are equivalent and operate at the same temperature, this yields the familiar representation of TL loops in terms of products of currents:

$$I_1 I_3 = I_2 I_4. \tag{2}$$

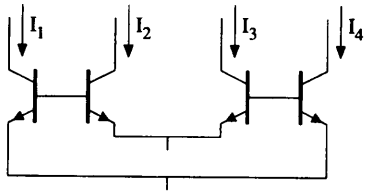


Figure 1: A four-transistor translinear loop.

This generic TL equation is the basis for a wide variety of static electronic functions, which are theoretically temperature and process independent.

2.2 Dynamic translinear principle

The static TL principle is limited to frequency-independent transfer functions. By admitting capacitors in the TL loops, the TL principle can be generalized to include frequency-dependent transfer functions. The term 'Dynamic Translinear' was coined in [89] to describe the resulting class of circuits. In contrast to other names proposed in literature, such as 'log-domain' [5], 'companding current-mode' [6], 'exponential state-space' [7], this term emphasizes the TL nature of these circuits, which is a distinct advantage with respect to structured analysis and synthesis.

The DTL principle can be explained with reference to the subcircuit shown in Figure 2. Using a current-mode approach, this circuit is described in terms of the collector current I_C and the capacitance I_{cap} flowing through the capacitance C. Note that the dc voltage source V_{const} does not affect I_{cap}. An expression for I_{cap} can be derived from the time derivative of (1) [6,89]:

$$I_{cap} = C V_T \frac{\dot{I}_C}{I_C}, \tag{3}$$

where the dot represents differentiation with respect to time.

Equation (3) shows that I_{cap} is a non-linear function of I_C and its time derivative $\dot{I}_C$. More insight in (3) is obtained by slightly

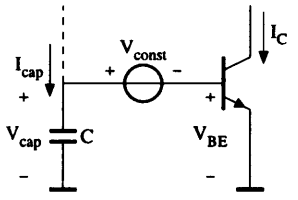


Figure 2: Principle of dynamic translinear circuits.

rewriting it:

$$CV_T\dot{I}_C = I_{cap}I_C. \tag{4}$$

This equation directly states the DTL principle: *A time derivative of a current can be mapped onto a product of currents.* At this point, the conventional TL principle comes into play, since the product of currents on the right-hand side (RHS) of (4) can be realized very elegantly by means of this principle. Thus, the implementation of (part of) a differential equation (DE) becomes equivalent to the implementation of a product of currents.

The DTL principle can be used to implement a wide variety of DEs, describing signal processing functions. For example, filters are described by linear DEs. Examples of non-linear DEs are harmonic and chaotic oscillators, PLLs and RMS-DC converters.

3 Classes of dynamic translinear circuits

In all DTL circuits, the voltages are logarithmically related to the currents. Therefore, these circuits are in some way instantaneous companding. Figure 3 shows the general block schematic of an instantaneous companding integrator [6]. In DTL circuits, the internal integrator is a linear capacitance. The expander E expands the output voltage of this integrator into a current, exploiting the exponential V-I transistor transfer function. Several types of DTL circuits can be distinguished within the general class of DTL circuits based on the particular implementation of E. Next to the most prevalent class of log-domain circuits, the two classes of tanh

and sinh circuits have been proposed by Frey [12]. In this section, we describe their characteristics, which can be derived from the generic output structures, depicted in Figure 4.

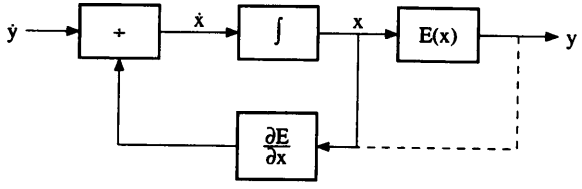


Figure 3: General block schematic of an instantaneous companding integrator.

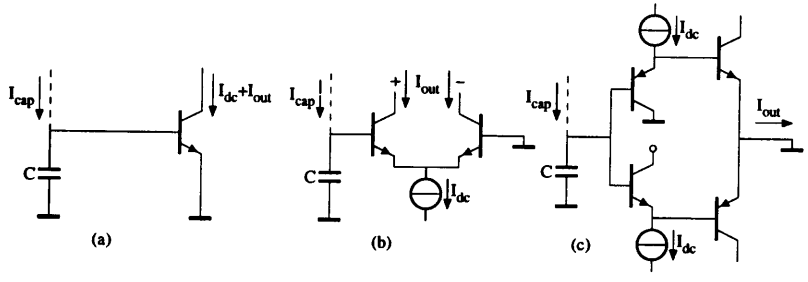


Figure 4: Generic output structures of (a) log-domain, (b) tanh, and (c) sinh circuits.

3.1 log-domain circuits

Most published DTL circuits are based on the common-emitter (CE) output stage shown in Figure 4(a), characteristic for the class of log-domain circuits. The transfer function from the capacitance voltage V_{cap} to the output current I_{out} is given by the well-known exponential law (1). In other words, E equals $\exp x$. The companding characteristics of a DTL circuit can be derived from the second order derivative of E with respect to x, denoted by E''. Without loss of generality, $x = 0$ is considered to be the quiescent point of the integrator shown in Figure 3. Figure 5 displays E'' for the output

stages shown in Figure 4. Applying a strict definition of companding, E'' should be strictly positive for $x > 0$ and strictly negative for $x < 0$. For log-domain circuits, a comparison of $E'' = \exp x$ with the strict definition of companding reveals that these circuits are indeed companding for $x > 0$; however, for $x < 0$ the exponential function constitutes a compression instead of an expansion. For a symmetrical output current, the overall behaviour of the CE output stage implies a compression rather than an expansion of the peak-to-peak signal swings [86].

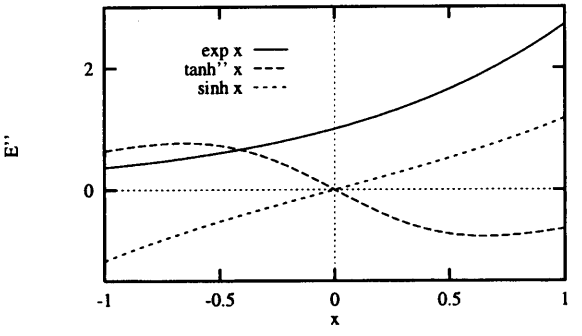


Figure 5: The second-order derivatives of the V-I transfer functions of the output stages shown in Figure 4.

From a current-mode point of view, the most important characteristic of a DTL output structure is the current-mode expression for the capacitance current I_{cap}. For log-domain filters, I_{cap} is given by Equation (3), where $I_C = I_{\text{dc}} + I_{\text{out}}$. As shown in Section 2, a linear derivative $\dot{I}_{\text{out}}$ is obtained by multiplying I_{cap} by $I_{\text{dc}} + I_{\text{out}}$.

A favorable property of log-domain circuits is that a linear damping term can be implemented by the connection of a dc current source I_o in parallel to a capacitance. This can be explained from Equation (4). If instead of I_{cap}, $I_{\text{cap}} + I_o$ is multiplied by $I_{\text{dc}} + I_{\text{out}}$, an additional term $I_o \cdot (I_{\text{dc}} + I_{\text{out}})$ is generated. The first term $I_o I_{\text{dc}}$ represents a dc offset current. The second term $I_o I_{\text{out}}$ results in a finite negative pole.

Typically, log-domain circuits operate in class A. The actual ac signal I_{out} is superposed on a dc bias current I_{dc}. As a consequence,

the output signal swing is limited to $I_{\text{out}} > -I_{\text{dc}}$. Note that this limitation is single sided, which is advantageous if a-symmetrical input wave-forms have to be processed. This characteristic can be exploited to enable class AB operation [6, 9]. Using a class AB set-up, see Figure 6, the dynamic range can be enlarged without increasing the quiescent power consumption. Using a current splitter, the input current I_{in} is divided into two currents I_{in1} and I_{in2}, which are both strictly positive, and related to I_{in} by: $I_{\text{in}} = I_{\text{in1}} - I_{\text{in2}}$. The current splitter impresses a constant geometric or harmonic mean on I_{in1} and I_{in2}. Next, I_{in1} and I_{in2} can be processed by two class A log-domain circuits. It is important to note that class AB operated log-domain circuits do satisfy the strict definition of companding due to the fact that only positive currents are processed, i.e., x is never negative.

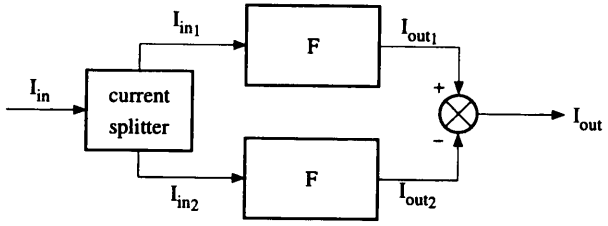


Figure 6: Set-up for class AB operation.

3.2 tanh circuits

Instead of a single transistor in CE configuration, the class of tanh circuits is characterized by a differential pair output structure [12], see Figure 4(b). The name of this class of circuits is derived from the well-known hyperbolic tangent V-I transfer function. The second-order derivative E'' is shown in Figure 5 and demonstrates that tanh circuits are not companding at all [48]. The differential pair implements a *compression* function.

The tail current of the differential pair is a dc current I_{dc}, and therefore, tanh circuits also operate in class A. The output current I_{out} is the difference of the two collector currents. The output swing

is limited to $-I_{dc} < I_{out} < I_{dc}$. Since this interval is symmetrical, the class AB set-up shown in Figure 6 cannot be applied to tanh circuits.

From Figure 4(b), the capacitance current I_{cap} is found to be:

$$I_{cap} = CV_T \left(\frac{\dot{I}_{out}}{I_{dc} + I_{out}} - \frac{-\dot{I}_{out}}{I_{dc} - I_{out}} \right). \tag{5}$$

A linear derivative $\dot{I}_{out}$ is obtained by multiplying this equation by $(I_{dc} + I_{out})(I_{dc} - I_{out})$:

$$2CV_T I_{dc} \dot{I}_{out} = I_{cap}(I_{dc} + I_{out})(I_{dc} - I_{out}). \tag{6}$$

Comparing Equations (4) and (6), we can see that the RHS of (6) is third-order, whereas the RHS of (4) is only second-order. Consequently, in general, TL loops of a higher order are required to implement a tanh circuit, resulting in a more complex circuit. In addition, a linear loss cannot be implemented by a dc current source connected in parallel to a capacitance. This leads us to the conclusion that tanh circuits do not seem to have any advantages over log-domain circuits.

3.3 sinh circuits

The third class of DTL circuits proposed in literature is formed by the sinh circuits [12]. The output structure, shown in Figure 4(c), is a complete second-order TL loop. It implements the geometric mean function $I_{dc}^2 = I_{out1} I_{out2}$. The actual output current I_{out} is the difference of I_{out1} and I_{out2}. Since both I_{out1} and I_{out2} are always positive, the sinh output structure operates in class AB, which is beneficial with respect to the dynamic range. The V-I transfer function of the output structure is a hyperbolic sine function. Figure 5 displays $E'' = \sinh x$ and shows that the sinh output stage implements a genuine expansion function.

The current-mode expression for the capacitance current I_{cap} is

given by:

$$I_{cap} = CV_T \frac{\dot{I}_{out_1}}{I_{out_1}}, \tag{7}$$

$$= -CV_T \frac{\dot{I}_{out_2}}{I_{out_2}}, \tag{8}$$

$$= CV_T \frac{\dot{I}_{out}}{\sqrt{4I_{dc}^2 + I_{out}^2}}, \tag{9}$$

$$= CV_T \frac{\dot{I}_{out}}{I_{out_1} + I_{out_2}}. \tag{10}$$

A linear derivative $\dot{I}_{out}$ is obtained by multiplying I_{cap} by the sum $I_{out1} + I_{out2}$. It is interesting to note that the voltage V_{cap} and the current $I_{out1} + I_{out2}$ are related through a hyperbolic cosine function; the first-order derivative of E with respect to x.

4 Structured design of a class-AB dynamic translinear integrator

Synthesis of a dynamic circuit, be it linear or non-linear, starts with a DE or with a set of DEs describing its function. Often, it is more convenient to use a state-space description, which is mathematically equivalent. The structured synthesis method for DTL circuits is illustrated here by the design of a first-order integrator, described in the time domain by:

$$\frac{dy}{d\tau} - x = 0 \tag{11}$$

This equation describes the integrator output signal y as a function of the input signal x. τ is the dimensionless time of the integrator.

4.1 Transformations

In the pure mathematical domain, equations are dimensionless. However, as soon as we enter the electronics domain to find an

implementation of the equation, we are bound to quantities having dimensions. In the case of TL circuits, all time-varying signals in the DEs, i.e., the input signals, the output signals and the tunable parameters, have to be transformed into currents. For the above expression, x and y can be transformed into the currents $I_{\text{in}} = x \cdot I_o$ and $I_{\text{out}} = y \cdot I_o$, I_o being the DC bias current that determines the absolute current swings.

Subsequently, the dimensionless time τ, can be transformed into the time t with its usual dimension $[s]$, using the equivalence relation given by:

$$d/d\tau = CV_T/I_o \cdot d/dt. \tag{12}$$

From this expression it can be deduced that the integrator will be linearly frequency tunable by means of control current I_o.

Applying the mentioned transformations, the resulting differential equation becomes:

$$CV_T\dot{I}_{\text{out}} - I_oI_{\text{in}} = 0 \tag{13}$$

4.2 Definition of the capacitance current

Conventional TL circuits are described by multivariable polynomials, in which all variables are currents. The gap between these current-mode polynomials and the DEs can be bridged by the introduction of capacitance currents, since the DTL principle states that a derivative can be replaced by a product of currents.

The capacitance currents can be introduced simply by defining them. To this end, several equivalent expressions for the capacitance current I_{cap} associated with the generic output stage of (class-AB) sinh circuits, depicted in Figure 4(c), can were obtained in Section 3. These equations all have two important characteristics in common. First, the denominators on the RHS are collector currents. This implies that these currents have to be strictly positive. Second, the numerators on the RHS are the time derivatives of the denominators.

With these characteristics in mind, we can define the capacitance current for the sinh integrator. As the capacitance current

will be used to eliminate the derivative from the DE, in the definition of this current, the derivative present in the DE has to be used. Using (10), the differential equation transforms into:

$$I_{\text{cap}}\left(I_{\text{out}_1} + I_{\text{out}_2}\right) = I_oI_{\text{in}}. \tag{14}$$

The current I_{cap} to be supplied to the capacitance C is thus given by:

$$I_{\text{cap}} = \frac{I_oI_{\text{in}}}{I_{\text{out}_1} + I_{\text{out}_2}}. \tag{15}$$

From this point on, the synthesis theory for static TL circuits can be used [104], since both sides of the above DEs are now described by current-mode multivariable polynomials.

4.3 Translinear decomposition

The next synthesis step is translinear decomposition. That is, the current-mode polynomial has to be mapped onto one or more TL loop equations that are characterized by the general equation:

$$\prod_{\text{CW}} J_{C,i} = \prod_{\text{CCW}} J_{C,i} \tag{16}$$

$J_{C,i}$ being the transistor collector current densities in clockwise (CW) or counter-clockwise (CCW) direction.

A two-quadrant multiplier/divider is required to implement the Right-Hand Side (RHS) of Equation (15). Since a class-AB implementation is pursued, this two-quadrant multiplier/divider has to be realized by two one-quadrant multiplier/dividers. This is realized by splitting the input current into two strictly positive signals I_{in_1} and I_{in_2}, the difference of which equals I_{in}. Rewriting Equation (15) yields:

$$I_{\text{cap}} = \frac{I_oI_{\text{in}_1}}{I_{\text{out}_1} + I_{\text{out}_2}} - \frac{I_oI_{\text{in}_2}}{I_{\text{out}_1} + I_{\text{out}_2}}. \tag{17}$$

Equation (17) is the basis for the block schematic of the sinh integrator depicted in Figure 7. At the input, a current splitter

generates I_{in_1} and I_{in_2} from I_{in}. Subsequently, the currents I_{in_1} and I_{in_2} are divided by $I_{out_1} + I_{out_2}$ in two separate circuits. The current $I_{out_1} + I_{out_2}$ is obtained from the sinh output stage. The output currents of the two multiplier/dividers are denoted by I_{cap1} and I_{cap2} and are respectively equal to the first and the second term on the RHS of Equation (17). Hence, the current supplied to the capacitance equals $I_{cap1} - I_{cap2}$. The use of a single capacitor is an advantage over the class-AB integrator proposed in [6] as it eliminates the necessity of matched capacitors. Finally, the capacitance voltage V_{cap} is applied to the sinh output stage via a voltage buffer to prevent interaction between the capacitance and the output stage.

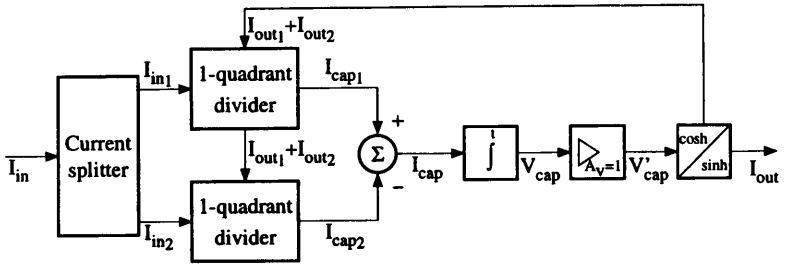


Figure 7: Block schematic of the class-AB translinear integrator.

4.4 Circuit implementation

The last synthesis step is the circuit implementation. The TL decomposition has to be mapped onto a TL circuit topology and the correct collector currents have to be forced through the transistors. Biasing methods for bipolar all-NPN TL topologies are presented in [104]. Additional implementation methods include the use of (vertical) PNPs, compound transistors or (simple) nullor implementations. If subthreshold MOSTs are used, some additional possibilities are the application of the back gate [105] and operation in the triode region [106].

The system blocks can be implemented by TL circuits, except of course the voltage buffer. To facilitate low-voltage operation, only folded TL loop topologies are allowed. A bipolar IC technology is used to implement the individual blocks.

4.4.1 Design of the input current splitter

A current splitter generates the currents I_{in_1} and I_{in_2} at the input of the integrator. In principle, the type of current splitter to be used at the input is not dictated by Equation (17). As the output stage is a geometric mean current splitter, the same function was chosen for the input current splitter.

The TL loop equation to be implemented is $I_{dc}^2 = I_{in_1} I_{in_2}$. Figure 8 depicts a 1 volt realization of this equation. The core of the circuit is the TL loop formed by Q_1 through Q_4. Transistors Q_1 and Q_3 are biased at a dc current I_{dc_1}. Transistor Q_2 conducts I_{in_2}. This current is inverted by a PNP current mirror and added to I_{in}. The resulting current I_{in_1} is conducted by Q_4, which is enforced by the Common-Collector (CC) stage Q_5. Biasing of Q_5 by means of a dc tail current source of the differential pair Q_2-Q_3 requires a relatively high dc current. This is disadvantageous with respect to the quiescent current consumption. A solution is dynamic biasing. The tail current of Q_2-Q_3 is generated by Q_6, Q_7 and Q_9, and equals $3I_{dc_1} + I_{in_2}$. Hence, Q_5 is biased at a dc current equal to only $2I_{dc_1}$.

The voltage source V_{dc_1} is necessary to ensure that the Q_7 does not saturate. Note that this voltage source has no effect on the TL loop. A convenient value for V_{dc_1} is 200 mV.

4.4.2 Design of the multiplier/divider

Once the bipolar input current I_{in} is decomposed into two positive currents $I_{in_{1,2}}$, such that the difference of these currents equals I_{in}, the two-quadrant multiplication of I_{in} can now be performed by the individual division of I_{in_1} and I_{in_2} by $I_{out_1} + I_{out_2}$, by means of two one-quadrant multiplier/dividers. The output currents of the one-quadrant multiplier/dividers satisfy:

$$I_{cap_{1,2}} = \frac{I_o I_{in_{1,2}}}{I_{out_1} + I_{out_2}}. \tag{18}$$

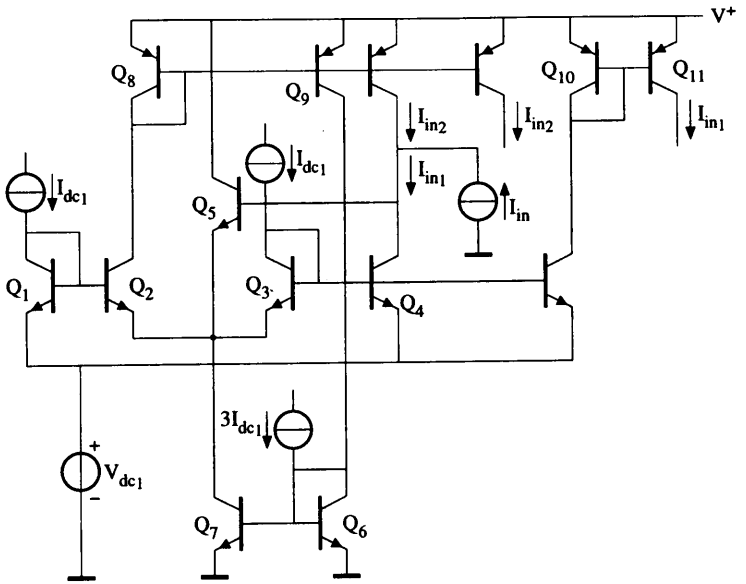


Figure 8: Implementation of the input current splitter.

As all linear factors in Equation (18) are strictly positive, it is a valid TL decomposition.

The 1 volt implementation of Equation (18) is shown in Figure 9. The second-order TL loop comprises Q_{12}–Q_{15}. Transistors Q_{13} and Q_{14} are biased by supplying respectively the currents $I_{\text{out}_1} + I_{\text{out}_2}$ and $I_{\text{in}_{1,2}}$ to the emitters of these devices. The collector current I_o of Q_{12} is enforced by the CC stage Q_{16}, which is biased by a dc current I_{bias1}.

A voltage source V_{dc_2} is necessary to ensure that the base voltages of Q_{13} and Q_{14} are always positive. Again, 200 mV is a convenient value.

The output of the multiplier/divider is the collector current of Q_{15}. Subtraction of I_{cap_1} and I_{cap_2} is performed by a PNP current mirror inverting I_{cap_2}.

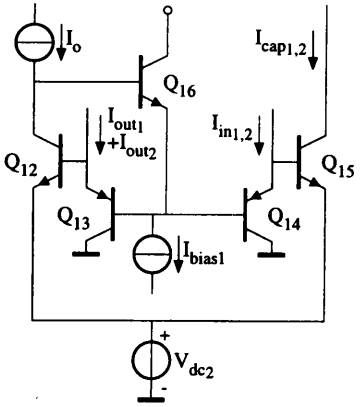


Figure 9: Implementation of the one-quadrant multiplier/divider.

4.4.3 Design of the voltage buffer

The current $I_{\text{cap}_1} - I_{\text{cap}_2}$ is supplied to the capacitor resulting in the voltage V_{cap}. A voltage buffer is used to minimize the interaction between the capacitor and the sinh output stage. The principle of the buffer amplifier is depicted in Figure 10(a). Ideally, the buffering is performed by the nullor. A level-shift between the input and the output of the buffer, represented by the voltage source V_{dc_3}, is necessary to avoid saturation of Q_{15} in the first multiplier/divider circuit. The output voltage is denoted by V'_{cap}.

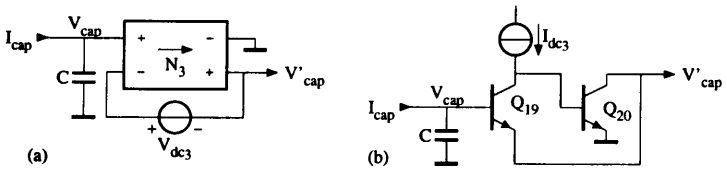


Figure 10: (a) Principle and (b) implementation of the voltage buffer.

The practical implementation of the nullor and the voltage source V_{dc_3} is shown in Figure 10(b). The nullor is implemented by two Common-Emitter (CE) stages, Q_{19} and Q_{20}. The level-shift is realized by the base-emitter voltage of Q_{19}. The output transistor Q_{20} must be able to sink the input current of the sinh output stage.

4.4.4 Design of the sinh output stage

The output stage has two functions. First, it enforces a geometric mean relation between the two output currents I_{out_1} and I_{out_2}. Secondly, it must provide the current $I_{\text{out}_1} + I_{\text{out}_2}$ to each of the multiplier/dividers, as shown in Figure 7.

The 1 volt realization of the output stage is depicted in Figure 11. The TL loop comprising Q_{21}–Q_{24} implements the sinh function given by:

$$I_{\text{out}} = 2I_{\text{dc}_2} \sinh \frac{V'_{\text{cap}} - V_{\text{dc}_4}}{V_T}, \tag{19}$$

where I_{dc_2} is a dc current. Note that Equation (19) is equivalent to the geometric mean function $I^2_{\text{dc}_2} = I_{\text{out}_1} I_{\text{out}_2}$.

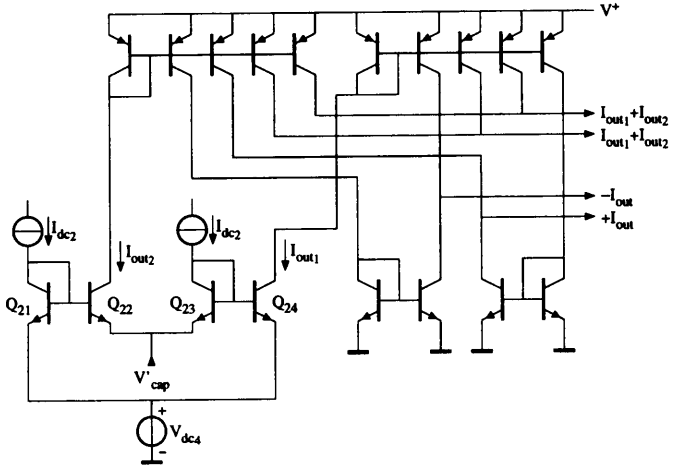


Figure 11: Implementation of the sinh output stage.

The current $I_{\text{out}_1} + I_{\text{out}_2}$ is supplied to the multiplier/dividers by means of PNP current mirrors. The output current I_{out} is generated by additional NPN current mirrors. The inverted output current $-I_{\text{out}}$ is added to easily enclose the integrator in a unity-feedback configuration by connecting $-I_{\text{out}}$ to the input of the integrator, which results in a first-order low-pass filter.

The voltage source V_{dc_4} is necessary to ensure that the emitter voltages of Q_{22} and Q_{23} are always positive. Once again, 200 mV is a convenient value.

4.5 Measurement results

Now that all the individual system blocks have been designed at circuit level, the sub-circuits can be linked together to form the integrator as depicted in Figure 7. For biasing purposes, the integrator is enclosed in a unity-feedback configuration, as discussed previously. This results in a first-order low-pass filter. Application of this filter in a hearing instrument was pursued. This leads to the required filter specifications shown in Table 1 [107]. For measurement purposes, the biasing current sources I_{dc_1}, I_{dc_2}, I_{dc_3} and I_{bias1} are realized by simple current mirrors and high-valued resistors. The frequency control current I_o is realized with a PTAT current source.

Table 1: Filter requirements.

Quantity	Value	Comment
Supply voltage	down to 1 V	
Current consumption	$< 5\ \mu$A	$I_{\text{in,max}} = 180\ \text{nA}_\text{p}$
Cut-off frequency (f_c) range	1.6–8 kHz	controllable
Dynamic range	68 dB	100 Hz–8 kHz
Total harmonic distortion	< 2 %	$f = 1$ kHz, $f_c = 1.6$ kHz, $I_{\text{in}} < 130\ \text{nA}_\text{p}$
	< 7 %	$f = 1$ kHz, $f_c = 1.6$ kHz, $I_{\text{in}} > 130\ \text{nA}_\text{p}$

To verify the integrator operation in practice, a semi-custom version of the active circuitry of the complete filter has been integrated in a standard 2-μm, 7-GHz process, fabricated at the Delft Institute of Microelectronics and Submicron Technology. Typical transistor parameters are: $h_{\text{fe,NPN}} \approx 100$, $f_{T,\text{NPN}} \approx 7$ GHz, $h_{\text{fe,LPNP}} \approx 80$ and $f_{T,\text{LPNP}} \approx 40$ MHz. The dc currents are set to $I_{\text{dc}_1} = I_{\text{dc}_2} = I_{\text{dc}_3} = 45$ nA, and $I_{\text{bias1}} = 135$ nA.

The capacitor has a value of 100 pF and is connected externally. The voltage sources $V_{dc1,2,4}$ equal 200 mV and are implemented by a resistive voltage divider.

The measurement results are summarized in Table 2 and are in good agreement with the expectations.

Table 2: Filter specifications.

Quantity	Value	Comment
Minimal supply voltage	0.95 V	
Supply current	1.9 μA	I_{in} = 180 nA_p
Quiescent supply current	1.6 μA	
Cut-off frequency range	1– > 8 kHz	
Maximal signal-to-noise ratio	63 dB	100 Hz–8 kHz
Dynamic range	73 dB	100 Hz–8 kHz
Max. total harmonic distortion	2.7 %	f_{in} = 1 kHz, f_c = 1.6 kHz, I_{in} = 180 nA_p

5 Conclusions

In this paper, it was shown that dynamic translinear circuits constitute an exciting new approach to the structured design of analog signal processing functions, using transistors and capacitors only. The presented design methodology was elaborated into the design of a class-AB translinear sinh integrator for audio filter applications. Measurements on a semi-custom version of the integrator illustrate the attractive properties of dynamic translinear circuits for low-power and low-voltage applications.

Acknowledgments

This research was partially funded by the Dutch Technology Foundation (STW), project DEL33.3251.

References

[1] Mulder, J., van der Woerd, A.C., Serdijn, W.A., van Roermund, A.H.M., General current-mode analysis method for translinear filters, IEEE Transactions on Circuits & Systems I-Fundamental Theory & Applications, Vol. 44, No. 3, Mar. 1997, pp. 193-197.

[2] A.C. van der Woerd, J. Mulder, W.A. Serdijn and A.H.M. van Roermund, Recent trends in translinear integrated circuits, proc. Electronics - ET'96, Sozopol, Bulgaria, 1996.

[3] Mulder, J., Serdijn, W.A., van der Woerd, A.C., van Roermund, A.H.M., A current-mode synthesis method for translinear companding filters, proc. ICECS'97, Cairo, 15-18 December 1997, pp. 1419-1422.

[4] B. Gilbert, Translinear circuits: a proposed classification, Electronics Letters, Vol. 11, No. 1, January 1975, pp. 14-16.

[5] R.W. Adams, Filtering in the log domain, Preprint No. 1470, presented at the 63rd AES Conference, New York, May 1979.

[6] E. Seevinck, Companding current-mode integrator: a new circuit principle for continuous-time monolithic filters, Electronics Letters, 22nd November 1990, Vol. 26, No. 24, pp. 2046-2047

[7] Frey, D.R., General class of current mode filters, Proceedings - IEEE International Symposium on Circuits and Systems, Vol. 2, 1993, pp. 1435-1437.

[8] Frey, D.R., Log-domain filtering: an approach to current-mode filtering, IEE Proceedings, Part G, Circuits, Devices & Systems, Vol. 140, No. 6, Dec. 1993, pp. 406-416.

[9] Frey, D.R., Current mode class AB second order filter, Electronics Letters, Vol. 30, No. 3, Feb. 3, 1994, pp. 205-206.

[10] Frey, D.R., A 3.3 volt electronically tunable active filter usable to beyond 1 GHz, Proceedings - IEEE International Symposium on Circuits and Systems, Vol. 5, 1994, pp. 493-495.

[11] Frey, D.R., On log domain filtering for RF applications, Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting 1995, IEEE, Piscataway, NJ, USA, pp. 12-18.

[12] Frey, D.R., Exponential state space filters: A generic current mode design strategy, IEEE Transactions on Circuits & Systems I-Fundamental Theory & Applications, Vol. 43, No. 1, Jan. 1996, pp. 34-42.

[13] Frey, D.R., Log filtering using gyrators, Electronics Letters, Vol. 32, No. 1, Jan. 4, 1996, pp. 26-28.

[14] Frey, D.R. Steigerwald, L., Adaptive analog notch filter using log filtering, IEEE International Symposium on Circuits and Systems, Vol. 1, 1996, pp. 297-300.

[15] Frey, D.R., Log domain filtering for RF applications, IEEE Journal of Solid-State Circuits, Vol. 31, No. 10, Oct. 1996, pp. 1468-1475.

[16] Frey, D.R, State space synthesis of log domain filters, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 481-484.

[17] Enz, C., Low-power log-domain continuous-time filters: An introduction, Proc. Low-power low-voltage workshop at ESSCIRC'95, Lille, France.

[18] Punzenberger, M., Enz, C., Low-voltage companding current-mode integrators, Proceedings - IEEE International Symposium on Circuits and Systems, Vol. 3, 1995, pp. 2112-2115.

[19] Gilles van Ruymbeke, Filtres continus integres programmables, These no. 1397, EPFL, Lausanne, 1995.

[20] Fried, R., Python, D., Enz, C.C., Compact log-domain current mode integrator with high transconductance-to-bias current ratio, Electronics Letters, Vol. 32, No. 11, May 23, 1996, pp. 952-953.

[21] Punzenberger, Manfred., Enz, Christian., New 1.2-V BiCMOS log-domain integrator for companding current-mode filters, IEEE International Symposium on Circuits and Systems, Vol. 1, 1996, pp. 125-128.

[22] Manfred Punzenberger, Christian Enz, A 1.2-V BiCMOS companding current-mode integrator for Sigma-Delta-modulators, Proc. ICECS '96, Rhodos, Greece, Oct. 1996.

[23] Dominique Python, Rafael Fried, Christian Enz, A 1.2 volt companding current-mode integrator for standard digital CMOS processes, Proc. ICECS '96, Rhodos, Greece, Oct. 1996.

[24] van Ruymbeke, Gilles. Enz, Christian. Krummenacher, Francois. Declercq, Michel, BiCMOS programmable continuous-time filter using voltage-companding, Proceedings of the Custom Integrated Circuits Conference 1996, IEEE, Piscataway, NJ, USA, 96CH35886, pp. 93-96.

[25] Yang, Fuji. Enz, Christian. van Ruymbeke, Gilles., Design of low-power and low-voltage log-domain filters, IEEE International Symposium on Circuits and Systems, Vol. 1, 1996, pp. 117-120.

[26] Enz, C.C. Punzenberger, M. Python, D., Low-voltage log-domain signal processing in CMOS and BiCMOS, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 489-492.

[27] M. Punzenberger, C. Enz, A 1.2-V BiCMOS class AB log-domain filter, IEEE Proc. ISSCC, San Fransico, USA, Feb. 1997.

[28] Punzenberger, Manfred. Enz, Christian C., Noise in instantaneous companding filters, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 337-340.

[29] Punzenberger, Manfred. Enz, Christian C., A low-voltage power and area efficient BiCMOS log-domain filter, proc. ESSCIRC '97, pp. 256-259

[30] Punzenberger, Manfred. Enz, Christian C., A 1.2-V low-power BiCMOS class AB log-domain filter, IEEE JSSC, Vol. 32, No. 12, pp. 1968-1978, December 1997.

[31] Van Ruymbeke G., Enz C.C., Krummenacher F., Declercq M., A BiCMOS programmable continuous-time filter using image-parameter method synthesis and voltage-companding technique, IEEE Journal of Solid-State Circuits, Vol. 32, No. 3, Mar. 1997, pp. 377-387.

[32] Ngarmnil, J., Toumazou, C., Micropower analogue filter, IEE Colloquium (Digest), No. 185, 1994, IEE, Stevenage, Engl., pp. 6/1-6.

[33] Toumazou, C., Ngarmnil, J., Lande, T.S., Micropower log-domain filter for electronic cochlea, Electronics Letters, Vol. 30, No. 22, Oct. 27, 1994, pp. 1839-1841.

[34] Ngarmnil, Jitkasame. Toumazou, Chris, Fully tuneable micropower log-domain filter, IEE Colloquium (Digest), No. 122, 1995, IEE, Stevenage, Engl., pp. 9/1-9/4.

[35] Jitkasame Ngarmnil, Chris Toumazou, T.S. Lande, A fully tunable micropower log-domain filter, Proc. ESSCIRC Lille, France, 1995, pp. 86-89.

[36] Eskiyerli, M.H., Payne, A.J., Toumazou, C., State space synthesis of integrators based on the MOSFET square law, Electronics Letters, Vol. 32, No. 6, Mar. 14, 1996, pp. 505-506.

[37] Eskiyerli, M.H., Payne, A.J., Toumazou, C., State-space synthesis of biquads based on the MOSFET square law, IEEE International Symposium on Circuits and Systems, Vol. 1, 1996, pp. 321-324.

[38] J. Mahattanakul, C. Toumazou, Independent control of transconductance gain and input linear range in a MOS linear transconductance amplifier, Electronics Letters, Vol. 32, No. 18, Aug. 29, 1996, pp. 1629-1630.

[39] Mahattanakul J., Toumazou C., Pookaiyaudom S., Low-distortion current-mode companding integrator operating at f_T of BJT, Electronics Letters. Vol. 32, No. 21, Oct. 10, 1996, pp. 2019-2021.

[40] J. Mahattanakul, C. Toumazou, Instantaneous companding in ASP, Proc. IEE colloquium on Analog Signal Processing, Oxford, United Kingdom, 1996.

[41] Ngarmnil, J., Toumazou, C., Micropower log-domain active inductor, Electronics Letters. Vol. 32, No. 11, May 23, 1996, pp. 953-955.

[42] Payne, Alison. Toumazou, Chris., Linear transfer function synthesis using non-linear IC components, IEEE International Symposium on Circuits and Systems, Vol. 1, 1996, pp. 53-56.

[43] Toumazou, Payne, Pookaiyaudom, High frequency log & square-root domain analog circuits: from companding to state-space, Tutorial at ICECS '96, Rhodos, Greece, Oct. 1996.

[44] Toumazou C., Payne A., Analogue electronics - log domain processing and its applications, Electronic Engineering, Vol. 68, No. 840, Dec. 1996, pp. 45-46.

[45] Jitkasame Ngarmnil, Micropower log-domain circuits: theory and applications, Ph.D. thesis, Imperial College, London, December 1996.

[46] Drakakis, E.M., Payne, A.J., Toumazou, C., Log-domain filters, translinear circuits and the Bernoulli cell, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 501-504.

[47] Drakakis, E.M., Payne, A.J., Toumazou, C., Bernoulli operator: a low-level approach to log-domain processing, Electronics Letters, Vol. 33, No. 12, June 5, 1997, pp. 1008-1009.

[48] J. Mahattanakul, C. Toumazou, Instantaneous companding and expressing: A dual approach to linear integrator synthesis, Electronics Letters Vol. 33, No. 1, Jan. 2 1997, pp. 4-5.

[49] Mahattanakul, Jirayuth. Toumazou, Chris., Non-linear design approach for high-frequency linear integrators, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 485-488.

[50] Mahattanakul, Jirayuth. Toumazou, Chris. Akbar, Atif Azim., DC stable CCII-based instantaneous companding integrator, IEEE International Symposium on Circuits and Systems, Vol. 2, 1997, pp. 821-824.

[51] Mahattanakul, J., Toumazou, C., Modular log-domain filters, Electronics Letters, Vol. 33, No. 13, June 19, 1997, pp. 1130-1131.

[52] D. Perry and G.W. Roberts, Log-domain filters based on LC ladder synthesis, Proc. IEEE ISCAS, Seattle, 1995, pp. 311-314.

[53] Perry, D., Roberts, G.W., The design of log-domain filters based on the operational simulation of LC ladders, IEEE Transactions on Circuits & Systems II-Analog & Digital Signal Processing, Vol. 43, No. 11, Nov. 1996, pp. 763-774.

[54] El-Gamal, Mourad. Roberts, Gordon W., LC ladder-based synthesis of log-domain bandpass filters, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 105-108.

[55] El-Gamal, Mourad. Leung, Vincent. Roberts, Gordon W., Balanced log-domain filters for VHF applications, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 493-496.

[56] Hematy, A. and Roberts, G.W., A fully-programmable analog log-domain filter circuit, proc. IEEE-CAS Region 8 Workshop on Analog and Mixed IC Design, Baveno, Italy, 12-13 September 1997, pp. 91-96.

[57] Leung, Vincent W. El-Gamal, Mourad. Roberts, Gordon W., Effects of transistor nonidealities on log-domain filters, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 109-112.

[58] Tsividis, Yannis, On linear integrators and differentiators using instantaneous companding, IEEE Transactions on Circuits & Systems II-Analog & Digital Signal Processing, Vol. 42, No. 8, Aug. 1995, pp. 561-564.

[59] Tsividis, Y., General approach to signal processors employing companding, Electronics Letters, Vol. 31, No. 18, Aug. 31, 1995, pp. 1549-1550.

[60] Tsividis, Yannis. Li, Dandan., Current-mode filters using syllabic companding, IEEE International Symposium on Circuits and Systems, Vol. 1, 1996, pp. 121-124.

[61] Tsividis Y., Externally linear, time-invariant systems and their applications to companding signal processors, IEEE Transactions on Circuits & Systems II-Analog & Digital Signal Processing, Vol. 44, No. 2, Feb. 1997, pp. 65-85.

[62] Tsividis, Y., Instantaneously companding integrators, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 477-480.

[63] Mulder, J., van der Woerd, A.C., Serdijn, W.A., van Roermund, A.H.M., Current-mode companding square-root domain integrator, Electronics Letters, Vol. 32, No. 3, Feb. 1, 1996, pp. 198-199.

[64] J. Mulder, W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund, Design of current-mode companding square-root domain dynamic circuits, Proc. IEE colloquium on Analog Signal Processing, pp. 8/1-5, Oxford, United Kingdom, 1996.

[65] Mulder, J., Serdijn, W.A., van der Woerd, A.C., van Roermund, A.H.M., Syllabic companding translinear filter, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 101-104.

[66] Mulder, J., Kouwenhoven, M.H.L., van Roermund, A.H.M., Signal x noise intermodulation in translinear filters, Electronics Letters, Vol. 33, No. 14, July 3, 1997, pp. 1205-1207.

[67] Mulder, J., Serdijn, W.A., van der Woerd, A.C., van Roermund, A.H.M., Dynamic translinear circuits: an overview, proc. Symposium on IC Technology, Systems & Applications, Singapore, 10-12 September 1997, pp. 31-38.

[68] Mulder, J., Serdijn, W.A., van der Woerd, A.C., van Roermund, A.H.M., Dynamic translinear circuits: an overview, proc. IEEE-CAS Region 8 Workshop on Analog and Mixed IC Design, Baveno, Italy, 12-13 September 1997, pp. 65-72.

[69] L.P.L. van Dijk, A.C. van der Woerd, J. Mulder and A.H.M. van Roermund, An ultra-low-power, low-voltage electronic audio delay line for use in hearing aids, IEEE Journal of Solid-State Circuits, Vol. 33, No. 2, February 1998, pp. 291-294.

[70] M.H.L. Kouwenhoven, J. Mulder and A.H.M. van Roermund, Noise analysis of dynamically nonlinear translinear circuits, Electronics Letters, February 1998.

[71] J. Mulder, W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund: An instantaneous and syllabic companding translinear filter, IEEE Transactions on Circuits and Systems I, Vol. 45, No. 2, February 1998, pp. 150-154.

[72] J. Mulder, M.H.L. Kouwenhoven,W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund: Analysis of noise in translinear filters, IEEE International Symposium on Circuits and Systems, 1998.

[73] J. Mulder, M.H.L. Kouwenhoven,W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund: Non-linear analysis of noise in static and dynamic translinear circuits, accepted, IEEE Transactions on Circuits and Systems - II, 1998.

[74] J. Mulder, M.H.L. Kouwenhoven,W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund: Noise considerations for translinear filters, accepted, IEEE Transactions on Circuits and Systems - II, 1998.

[75] J. Mulder, Static and dynamic translinear circuits, Ph.D. thesis, Delft University of Technology, 1998.

[76] Serdijn, W.A., van der Woerd, A.C., van Roermund, A.H.M., Chain-rule resistance: a new circuit principle for inherently linear ultra-low-power on-chip transconductances or transresistances, Electronics Letters, Vol. 32, No. 4, Feb. 15, 1996, pp. 277-278.

[77] W.A. Serdijn, M. Broest, J. Mulder, A.C. van der Woerd and A.H.M. van Roermund, A low-voltage ultra-low-power current-companding integrator for audio filter applications, Proc. ESSCIRC'96, Neuchatel, Switzerland, 1996, pp. 404-407.

[78] W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund, Dynamic-range optimization in ultra-low-power fully integratable companding filters, proc. IEEE ProRISC, Mierlo, the Netherlands, 1996, pp. 299-302.

[79] W.A. Serdijn, M. Broest, J. Mulder, A.C. van der Woerd, A.H.M. van Roermund, A low-voltage ultra-low-power translinear integrator for audio filter applications, IEEE Journal of Solid-State Circuits, Vol. 32, No. 4, Apr. 1997, pp. 577-581.

[80] Serdijn, W.A. Mulder, J. van Roermund, A.H.M., Dynamic translinear circuits, proc. Low-Voltage Low-Power workshop, ESSCIRC'97, Southampton, United Kingdom, 19 September 1997.

[81] Serdijn, W.A., Mulder. J., van Roermund, A.H.M., Shortening the analog design trajectory by means of the dynamic translinear principle, proc. ProRISC CSSP'97, Mierlo, 27-28 November 1997, pp. 559-566.

[82] W.A. Serdijn, J. Mulder and A.H.M. van Roermund: Dynamic translinear circuits, proc. AACD'98, Copenhagen, 28-30 April 1998.

[83] W.A. Serdijn, M.H.L. Kouwenhoven, J. Mulder and A.H.M. van Roermund: The design of high-dynamic-range fully integratable translinear filters, accepted, Analog Integrated Circuits and Signal Processing (Kluwer), 1998.

[84] W.A. Serdijn and J. Mulder (guest editors): Special issue on dynamic translinear circuits, to appear, Analog Integrated Circuits and Signal Processing.

[85] Liu, Shen Iuan. Liao, Yu-Hung., Table-based log-domain linear transformation filter, Electronics Letters, Vol. 32, No. 19, Sep. 12, 1996, pp. 1771-1772.

[86] Fox, Robert. Nagarajan, Mahalingam. Harris, John., Practical design of single-ended log-domain filter circuits, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 341-344.

[87] W.A. Serdijn, A.C. van der Woerd, A.H.M. van Roermund and J. Davidse: "Design principles for low-voltage low-power analog integrated circuits," Analog Integrated Circuits and Signal Processing 8, July 1995, pp. 115-120.

[88] B. Gilbert: "Current-mode circuits from a translinear viewpoint: a tutorial," in: C. Toumazou, F.J. Lidgey and D.G. Haigh (editors): Analogue IC design: the current-mode approach, Peter Peregrinus, London, 1990.

[89] J. Mulder, A.C. van der Woerd, W.A. Serdijn and A.H.M. van Roermund, An RMS-DC converter based on the dynamical translinear principle, Proc. ESSCIRC'96, Neuchatel, Switzerland, 1996, pp. 312-315.

[90] Mulder J., Serdijn, W.A., van der Woerd, A.C. van Roermund, A.H.M., Dynamic translinear RMS-DC converter, Electronics Letters. Vol. 32, No. 22, Oct. 24, 1996, pp. 2067-2068.

[91] Mulder, J., van der Woerd, A.C., Serdijn, W.A., van Roermund, A.H.M., RMS-DC converter based on the dynamic translinear principle, IEEE Journal of Solid-State Circuits, Vol. 32, No. 7, July 1997, pp. 1146-1150.

[92] W.A. Serdijn, J. Mulder, A.C. van der Woerd and A.H.M. van Roermund, A wide-tunable translinear second-order oscillator, pp. 295-298, proc. IEEE ProRISC, Mierlo, the Netherlands, 1996.

[93] Serdijn, Wouter A., Mulder, Jan., van der Woerd, Albert C., van Roermund, Arthur H.M., Design of wide-tunable translinear second-order oscillators, IEEE International Symposium on Circuits and Systems, Vol. 2, 1997, pp. 829-832.

[94] Serdijn, W.A., Mulder, J., van der Woerd, A.C., van Roermund, A.H.M., A wide-tunable translinear second-order oscillator, proc. ESSCIRC'97, Southampton, United Kingdom, 16-18 September 1997, pp. 224-227.

[95] W.A. Serdijn, J. Mulder, A.C. van der Woerd and A.H.M. van Roermund: A wide-tunable translinear second-order oscillator, IEEE Journal of Solid-State Circuits, Vol. 33, No. 2, February 1998, pp.195-201.

[96] J. Mulder, W.A. Serdijn, A.C. van der Woerd and A.H.M. van Roermund: A 3.3 V current-controlled square-root-domain oscillator, Analog Integrated Circuits and Signal Processing, Vol. 16, No. 1, April 1998, pp. 17-28.

[97] Pookaiyaudom, Sitthichai. Mahattanakul, Jirayuth., A 3.3 volt high-frequency capacitorless electronically-tunable log-domain oscillator, Proceedings - IEEE International Symposium on Circuits and Systems, Vol. 2, 1995, pp. 829-832.

[98] Sitthichai Pookaiyaudom, Jirayuth Mahattanakul, High-frequency low-distortion electronically-tunable log-domain oscillators, IEE Colloquium (Digest), 1995, England.

[99] Thanachayanont, A., Pookaiyaudom, S., Toumazou, C., State-space synthesis of log-domain oscillators, Electronics Letters, Vol. 31, No. 21, Oct. 12, 1995, pp. 1797-1799.

[100] J. Ngarmnil, A. Thanachayanont, C. Toumazou, On the design of micropower active resonators and oscillators using log-domain processing, Proc. ICECS '96, Greece, 1996.

[101] Thanachayanont, A., Payne, A.J., Pookaiyaudom, S., State-space synthesis of oscillators based on the MOSFET square law, Electronics Letters, Vol. 32, No. 18, Aug. 29, 1996, pp. 1638-1639.

[102] Pookaiyaudom, Sitthichai., Log domain oscillators, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 497-500.

[103] Thanachayanont, Apinunt. Payne, Alison. Pookaiyaudom, Sittichai., Current-mode phase-locked loop using a log-domain oscillator, IEEE International Symposium on Circuits and Systems, Vol. 1, 1997, pp. 277-280.

[104] E. Seevinck: "Analysis and synthesis of translinear integrated circuits," Elsevier, Amsterdam, 1988.

[105] J. Mulder, A.C. van der Woerd, W.A. Serdijn and A.H.M. van Roermund: "Application of the back gate in MOS weak inversion translinear circuits," IEEE Transactions on Circuits and Systems – I, Vol. 42, No. 11, 1995, pp. 958-962.

[106] A.G. Andreou and K.A. Boahen, Translinear circuits in subthreshold MOS, Analog Integrated Circuits and Signal Processing, Vol. 9, April 1996, pp. 141-166.

[107] W.A. Serdijn, A.C. van der Woerd, J. Davidse and A.H.M. van Roermund: "A low-voltage low-power fully-integratable front-end for hearing instruments," IEEE Transactions on Circuits and Systems – I, Vol. 42, No. 11, 1995, pp. 920-932.

1-V Log-Domain Filters

Christian Enz[1], Manfred Punzenberger[2]

(1) Rockwell Semiconductor Systems, Newport Beach, CA 92660, U.S.A.

(2) Electronics Laboratory, Swiss Federal Institute of Technology (EPFL) CH-1015 Lausanne, Switzerland

Abstract

The basic principle of log-domain circuits is introduced and its suitability for the realization of low-voltage (LV) and low-power (LP) continuous-time filters is established. The main limitations of practical implementations including noise are discussed and the advantage of the combination of companding and class AB operation is highlighted. Examples of BiCMOS and standard digital CMOS realizations operating at supply voltages as low as 1 V are presented. A comparison between classical g_m-C and log-domain continuous-time filters is drawn on the basis of the power consumption per pole and edge frequency normalized to the dynamic range. This comparison shows that log-domain filters are more power saving than more traditional filters.

I. INTRODUCTION

The continuous decrease of the supply voltage used for integrated circuits is mainly imposed by the scaling of modern technologies and the power consumption reduction of digital circuits. It has been well established that this voltage reduction does unfortunately not help to lower the power consumption of analog circuits [1]-[6]. On the contrary it generally even leads to an increase of the power consumption for maintaining the same signal-to-noise ratio (SNR) and bandwidth [1]-[6]. A possible way to accommodate for this supply voltage reduc-

tion without prohibitively increasing the power consumption of analog circuits is to use either syllabic [9]-[11] or instantaneous voltage companding [12]-[15]. As discussed in [12]-[15] this technique requires the input signal to be predistorted in order to avoid any distortion due to the nonlinear operation inherent to the instantaneous companding. The log-domain approach uses instantaneous companding where the currents, having basically an unlimited dynamic range (DR), are compressed logarithmically when transformed into voltages and expanded exponentially when converted back to currents [16]-[47]. The voltage swings are strongly reduced making them almost independent of the supply voltage, which can be reduced to the minimum required for a proper operation of the circuit. The log-domain technique is therefore well suited to the implementation of LV continuous-time analog signal processing circuits such as continuous-time filters (CTFs) [6][53].

In g_m-C CTFs the transconductor has often to be linearized in order to achieve the desired DR. This unavoidably degrades the g_m/I ratio proportionally to the increase of the linear range, resulting in a poor current efficiency [16]. Since log-domain circuits exploit the exponential characteristic of a single device, they do not require any linearization and hence preserve the maximum g_m/I ratio even for large signal operation [16]. Therefore, in addition to the LV operation, the log-domain technique also provides the most efficient use of the available current for implementing a given transconductance, which makes it also attractive for LP integrated circuit design.

The recently published measured results demonstrate the feasibility and the potential of this technique for LV and LP CTFs realized either in bipolar [18]-[32], BiCMOS [33]-[43] and even in CMOS technologies [43]-[47]. The combination of companding and class AB operation also opens the door to a more power efficient analog signal processing where the DR can be extended without increasing the maximum SNR nor the quiescent power consumption.

Section II briefly recalls the principle of log-domain filters. The fundamental limitations of the log-domain filtering principle for practical implementations in bipolar, BiCMOS and CMOS technologies are discussed in Section III. Some practical examples of LV bipolar/BiCMOS and CMOS implementations are described in Section IV and Section V respectively, highlighting their major differences. In Section VI, the

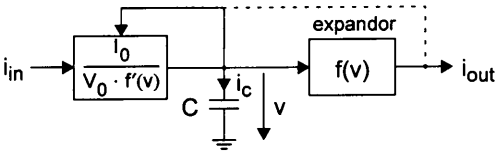


Fig. 1: General structure of an instantaneous companding integrator.

power efficiencies of a simple g_m-C and a log-domain integrator are derived and compared with respect to the supply voltage. A comparison between the power efficiency of integrated log-domain and more traditional CTFs published in the literature is also given.

II. Principle of Log-Domain Integrators

The general concept of an instantaneous companding current mode integrator has been extensively discussed in [12] and [14]. It is illustrated by the block diagram shown in Fig. 1 [12][14][18]. The companding current-mode integrator has to implement a globally linear integrator defined by

$$\frac{di_{out}}{dt} = \pm\frac{i_{in}}{\tau} \tag{1}$$

and this despite of the locally nonlinear expanding function of the output expandor (a transconductor in this case) relating the output current to the capacitor voltage

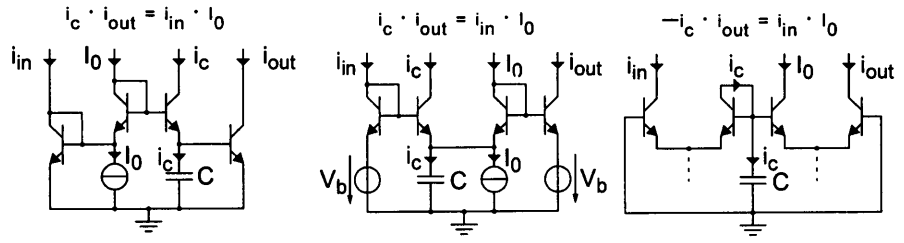
$$i_{out} = f(v). \tag{2}$$

Even if this function $f(v)$ could be either compressing or expanding, here it is always assumed as being an expanding function.

This overall linear operation can only be obtained if the current injected into the capacitor (which is assumed linear) is predistorted according to

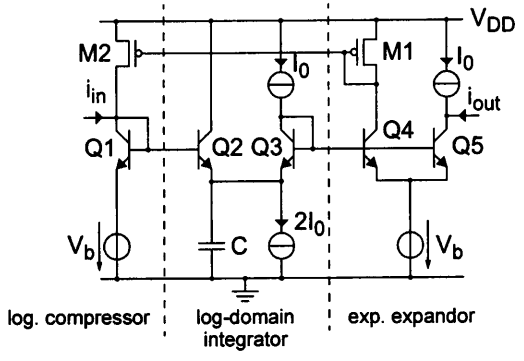
$$i_c = \pm C \cdot \frac{1}{f'(v)} \cdot \frac{i_{in}}{\tau} \tag{3}$$

where $f'(v)$ is the derivative of $f(v)$ with respect to voltage v. This predistortion can be implemented by dividing the input current i_{in} by

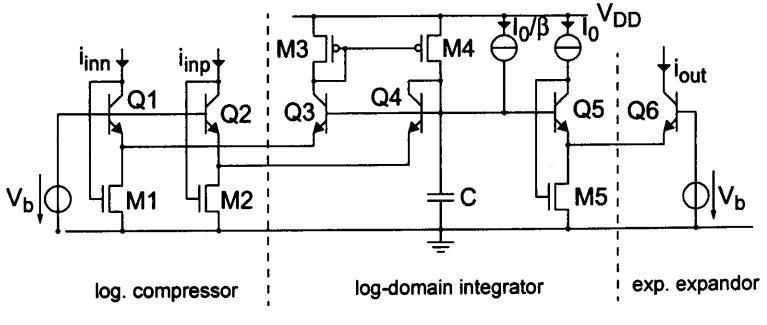


a) Stacked junctions. *b) Folded junctions with capacitor connected to emitter.* *c) Folded junctions with capacitor connected to base*

Fig. 4: Basic implementations of LV log-domain filters in bipolar or BiCMOS (bias not shown).



a) Capacitor connected to emitter.



b) Capacitor connected to base (and collector for bias purpose).

Fig. 5: Two examples of LV BiCMOS log-domain integrators.

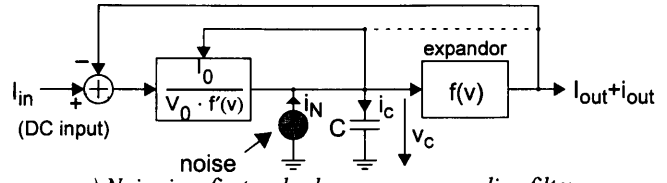
Q_2 and C are forming the log-domain integrator. Transistor Q_3 is used as an adjustable level shifter and Q_5 accomplishes the exponential expansion. Transistor Q_4 and the current mirror M_1-M_2 are needed for canceling the losses introduced by discharging the capacitor by a constant current. Note that the effect of the base currents can be reduced by inserting voltage followers between the collector and the base nodes of Q_1 and Q_3, respectively.

In order to build high-order LV filters by coupling log-domain integrators, the input and output dc voltage levels have to be compatible. This is accomplished by canceling the base-emitter voltage drop of Q_2 by the adjustable up level shifter Q_3. The latter also provides the required tuning of the integrator time constant by simultaneously changing the collector and emitter bias currents. Note that any mismatch between these two current sources is injected as an error current into the capacitor, resulting in a poor dc gain and a high distortion.

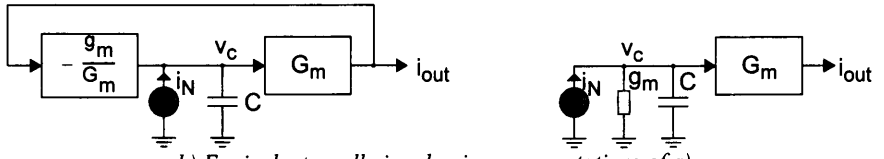
The basic building block used to build high-order filters requires both a non-inverting and an inverting input. The log-domain integrator of Fig. 5 a) can be extended to include also the inverting input by inserting two additional current mirrors [32]. Another way to include the inverting input is to extend the circuit of Fig. 5 a) to a differential class AB integrator.

Fig. 5 b) shows an example of a LV log-domain integrator with its capacitor connected to the base node of a translinear loop [34]-[40]. Unlike the integrator of Fig. 5 a), for this circuit it is straightforward to add an inverting input by simply inserting a single current mirror M_3-M_4 in order to invert the collector current of Q_3 . The level shifter Q_5, M_5 is now driven from its base and consequently the absolute error resulting from a mismatch between the base current and the cancellation current I_0/β can be expected to be approximately β times smaller, resulting in an improved linearity compared to the previous circuit.

The single-ended version of this integrator has been used to implement a 3rd-order low-pass Chebyshev filter with 1 dB passband ripple and having a nominal cut-off frequency of 320 kHz for a 1 µA I_0 bias current [39][40]. As can be seen from the schematic of the filter presented in Fig. 6, the circuit is very simple and the layout is therefore extremely compact [39][40]. The measured frequency responses of this



a) Noise in a first-order low-pass companding filter.



b) Equivalent small-signal noise representations of a)

Fig. 2: Noise in a first-order low-pass instantaneous companding filter.

noted that there are two kinds of noise contributions: noise arising from transistors that have a constant bias current (as for example the level shifter Q5 shown in Fig. 5) and noise sources that are due to transistors that have a collector or drain current that depends on the signal. The first kind of noise sources have a constant PSD, whereas the second may be non-stationary. In this analysis, for the sake of simplicity, it is assumed that the noise source $i_N(t)$ is stationary and has a PSD given by (7) with a constant noise factor.

A further assumption is that the noise bandwidth is much larger than the signal bandwidth. In this analysis the signal is assumed to be a DC input current I_{in}. In this case the noise analysis is identical to classical circuits. Since the noise can be assumed to be a small-signal, the circuit of Fig. 2 a) can be linearized at its operating point resulting in the diagram shown in Fig. 2 b). The feedback loop implements a conductance of value equal to g_m and therefore the noise injected on the capacitor is low-pass filtered with a cut-off frequency $g_m/(2\pi C)$. The power of the noise voltage on the capacitor is therefore equal to

$$V_N^2 = \frac{\gamma \cdot kT}{C}. \tag{8}$$

This noise voltage appears at the output as a noise current of power

$$I_{Nout}^2 = G_m^2 \cdot V_N^2 = \frac{I_{tot}^2}{(nU_T)^2} \cdot \frac{\gamma \cdot kT}{C} = \frac{\gamma}{n^2} \cdot \frac{q^2}{kTC} \cdot (I_0^2 + I_{in}^2) \tag{9}$$

where I_{tot} is the total rms current flowing from the expandor. According to (9), the output noise power depends on the sum of the square of the bias current and the signal power, which in this case is simply equal to the input signal power I_{in}^2. For sinusoidal input current with a frequency much below the filter cut-off frequency, I_{in}^2 is simply equal to half the square of the amplitude. A distinction has to be made whether the circuit can operate in class A or class AB. For a class A operation, the signal power is restricted by the bias current and the noise stays almost independent of the signal. For a class AB operation, the noise stays constant as long as the signal rms current is much smaller than the bias current and it increases proportional to the signal power when the signal rms current becomes much larger than the bias current

$$I_{Nout}^2 = \begin{cases} \frac{\gamma}{n^2} \cdot \frac{q^2}{kTC} \cdot I_0^2 & \text{for: } I_{in} << I_0 \\ \frac{\gamma}{n^2} \cdot \frac{q^2}{kTC} \cdot I_{in}^2 & \text{for: } I_{in} >> I_0 \end{cases} \tag{10}$$

The output noise power is plotted in Fig. 3 for both the class A and class AB filter assuming they both have the same bias currents and integrating capacitors and therefore identical cut-off frequencies. The signal-to-noise ratio can easily be evaluated from (9) as

$$SNR \cong SNR_{max} \cdot \frac{I_{in}^2}{I_0^2 + I_{in}^2} = \begin{cases} SNR_{max} \cdot \left(\frac{I_{in}}{I_0}\right)^2 & \text{for: } I_{in} << I_0 \\ SNR_{max} & \text{for: } I_{in} >> I_0 \end{cases} \tag{11}$$

where SNR_{max} is given by

$$SNR_{max} = \frac{n^2}{\gamma} \cdot \frac{kTC}{q^2} \tag{12}$$

SNR_{max} represents the maximum SNR that can be reached with a class AB log-domain filter. According to (12), it is proportional to the

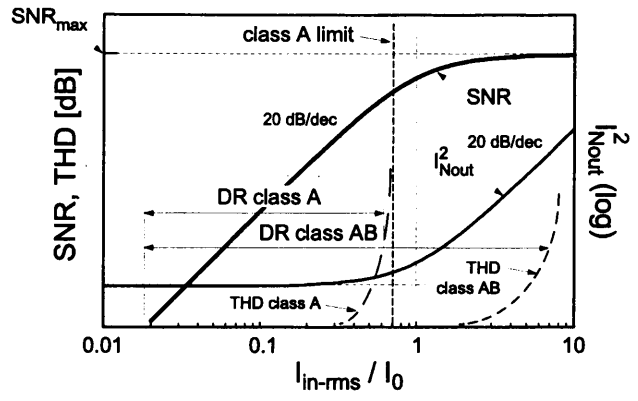


Fig. 3: Output noise power, signal-to-noise and total-harmonic-distortion versus the output signal rms current normalized to the bias current.

integration capacitance. It is also proportional to the temperature assuming that the bias current is constant and independent of the temperature. It should be noticed that, according to (12), the SNR for a CMOS implementation is n^2 larger than that obtained with a bipolar circuit both having the same capacitance and noise factor. The SNR is plotted together with the output noise power and the total-harmonic-distortion (THD) in Fig. 3 for a class A and a class AB filter having both the same capacitance and noise factor. The dynamic range (DR) is defined on the lower side by the minimum signal for which the SNR is equal to a given value (typically 0 dB) and on the other side by the maximum signal for a given distortion (or intermodulation). The minimum signal is identical for both class A and class AB circuits, assuming that they have equal bias currents, capacitors and noise excess factors. For class A circuits, the maximum signal is ultimately limited by the bias current, whereas for class AB circuits it might be limited to a higher value depending on the linearity performance which is only limited by nonideal effects (such as those resulting in a deviation of the I-V characteristic from the ideal expanding law, base currents in bipolar implementations, mismatches,...) but not by the nonlinearity of the device characteristic itself. Fig. 3 clearly shows that the DR range of class AB companding circuits can be substantially extended compared to that obtained with class A companding (or even non-companding) circuits without increasing the maximum SNR and therefore without

increasing the power consumption. This suggests that in order to exploit the larger DR available, instantaneously companding filters should be operated strongly in class AB while maintaining an overall good linearity, which is not an easy task. In practical case the class AB implementation requires two signal paths and an additional input signal class AB conditioner (or current splitter [21]) which might result in higher noise floor and therefore in a smaller DR extension than expected. Measurements have shown that the dynamic range of the same filter having the same capacitor could be extended by as much as 23 dB.

B. Device nonidealities and mismatch

Since log-domain filters are a dynamic extension of translinear loops, the device nonidealities have similar effects on the circuit's performance and similar measures should be taken in order to cope with them [29]. The most important device nonidealities are discussed below.

Base currents: One of the most important nonidealities in bipolar implementations of translinear loops and log-domain filters is the finite current gain of the transistors. This can result in losses of the individual log-domain integrators and in limited quality factors of the whole filter. But even more severe are the nonlinear effects such as signal distortion and intermodulation particularly for class AB implementations, where large signal current peaks can occur. An easy way to improve this situation is the use of translinear loops with alternating directions of the base-emitter junctions rather than a balanced ordering (cf. Fig. 5), since the former implementation tends to be less sensitive to finite current gains of the transistors [33]. Therefore, the low-voltage implementations presented hereafter are most preferable even for applications where higher supply voltages are available.

Saturation current and area mismatch: The design of translinear loops is normally based on the simplifying assumption of perfectly matched devices. In reality however, the unavoidable mismatch of the emitter area of bipolar junction transistors or of the aspect ratio of weakly inverted MOSFETs as well as the variation of their saturation currents introduces a constant error factor k into the current equation.

The output current of equation (6) affected by mismatch becomes therefore

$$\tilde{i}_{out} = i_{out} \cdot k = \pm\left(\frac{i_{in} \cdot I_0}{i_c} \cdot k\right). \quad (13)$$

The effect of this mismatch is an error on the time constant of the integrator and for the whole filter this causes inaccuracy on the cut-off frequency, pass-band gain errors and limited quality factors. It has to be noticed, that the mismatch does not introduce any distortion in class A filter implementations.

Parasitic capacitances: Since all the device transconductances in the filter are tuned proportionally with the cut-off frequency, the phase error due to parasitic capacitances is almost independent of the frequency tuning and does therefore not restrict the tuning range. In addition to these effects on the linear transfer function of the filter, parasitic capacitances are also causing intermodulation of the signal.

Series resistances: Base and emitter series resistances of BJTs and source resistances of MOSFETs cause a deviation from the ideal exponential device characteristic and consequently a degradation of the intermodulation distortion of the filter. For reducing this effect, the transistors have to be properly sized, which in turn increases the parasitic capacitances. Therefore, there is a trade-off between intermodulation and high-frequency operation.

There are many other device nonidealities such as high injection, body effect, early effect, self heating, nonlinear capacitors, etc... that may influence the operation of the log-domain filters. As already mentioned earlier, the nonlinear nature of log-domain circuits renders the analysis of these effect rather difficult and therefore they will not be discussed further here.

C. Noise coupling

Like for thermal noise, any type of perturbation of the voltage across the filter capacitors will be modulated by the signal. This causes on one hand an improved immunity against noise coupling from the digital part for low level signals, but on the other hand, the frequency spectrum of these disturbances may be folded down into the band of interest. When the digital noise is injected onto the capacitor of the integrator it is low-pass filtered. The sum of this filtered noise voltage u_{nf} and the signal voltage v_C are expanded to result the noise affected output current $\tilde{i}_{out}$.

$$\tilde{i}_{out} = I_S \cdot \exp\left[\frac{v_C + u_{nf}}{U_T}\right] \cong i_{out} \cdot \left(1 + \frac{u_{nf}}{U_T}\right). \quad (14)$$

In addition to the signal, the whole digital noise spectrum appears centered around each component of the signal at the output of the integrator.

IV. Low-Voltage Bipolar/BiCMOS Implementations

Log-domain integrators are generally based on translinear loops that implement (6). A straightforward realization of equation (6) is shown in Fig. 4 a) [18]. As explained in Section II, log-domain filtering is especially useful for the implementation of circuits that have to operate at low-voltage and therefore stacked junctions as shown in Fig. 4 a) should be avoided. They should be replaced with folded translinear loops as illustrated in Fig. 4 b) and Fig. 4 c), where the bias circuits have been omitted for the sake of simplicity. The critical branch of this kind of circuit typically comprises a diode connected transistor biased by a current source connected to the collector and another connected to the emitter so that the minimum supply voltage corresponding to one junction and two saturation voltages is at least 1 V. In such LV log-domain filters based on bipolar npn-only translinear loops, two different basic circuit concepts can be found: integrators with the capacitor connected to the emitter node as in Fig. 4 b) or to the base node as in Fig. 4 c). An example of each of these basic log-domain integrators will be described in more details below.

In BiCMOS processes, the MOS transistors can advantageously be used for implementing current mirrors and voltage followers, since they don't suffer from finite current gain and they allow a lower saturation voltage than bipolar transistors. Fig. 5 a) shows an example of a LV BiCMOS log-domain integrator with its capacitor connected to the emitter node. The signal is logarithmically compressed by Q_1, while

the derivative of the expanding function as illustrated in Fig. 1. The bias current I_0 sets the time constant τ according to

$$\tau = C \cdot \frac{V_0}{I_0} \tag{4}$$

where V_0 is a scaling voltage that will be defined below.

In the particular case of log-domain integrators, the expanding function is an exponential which has the two following fundamental advantages: a) it can be implemented very accurately using either a bipolar transistor or a MOS transistor biased in weak inversion and b) the derivative of the exponential function is still an exponential that can be realized with the same components as the expander insuring a good match between the expanding function $f(v)$ and the predistortion function $f'(v)$. An additional advantage is that the division required for the predistortion as stated by (3) can easily be realized by means of translinear loops [8] which can be implemented either with bipolar or MOS transistors biased in weak inversion. The derivative $f'(v)$ appearing in (3) is then nothing else than the transconductance which is given by

$$f'(v) = G_m(i_{out}) = \frac{i_{out}}{V_0}. \tag{5}$$

In a log-domain realization, the scaling voltage V_0 is typically equal to nU_T where $U_T = kT/q$ is the thermodynamic voltage and n is a factor close to unity representing the junction nonideality factor for a bipolar transistor or corresponding to the slope factor of the MOS transistor [48].

Substituting $f'(v)$ and τ given respectively by (5) and (4) into (3) results in

$$i_c \cdot i_{out} = \pm i_{in} \cdot I_0 \tag{6}$$

which corresponds to a basic four transistor translinear loop. Examples of log-domain integrators implementing (6) using bipolar transistors are described in Section IV whereas Section V presents different realizations using MOS transistors biased in weak inversion.

III. Fundamental Limitations of Instantaneous-Companding Integrators

Like in traditional linear circuits, the performance of log-domain filters are limited by the non-idealities of the devices such as noise and mismatches. At high signal level, the linearity of log-domain filters is essentially limited by the deviation of the I-V characteristic of the devices from the ideal exponential law. Another important limitation is due to coupling with noise sources such as interferers generated by the digital part of the circuit. All these aspects are briefly discussed below.

A. Noise

The strong logarithmic compression of the input current when it is injected onto the capacitor results in a very small voltage swing of the order of a few U_T. One may then think that reducing the signal amplitude may result in a degraded signal-to-noise ratio (SNR). Fortunately this is not the case thanks to the exponential expanding function which requires the signal to be only slightly larger than the noise voltage on the capacitor. In fact the noise limitations of log-domain circuits are similar to linear circuits as long as the circuit operates in class A. Things are changing when the signal currents may become much larger than the bias current as is the case for circuits operating in class AB.

The exact calculation of the noise in instantaneous companding circuits is made difficult due to the internal nonlinear behavior of the circuit and the nonstationary nature of the noise sources [14]. Nevertheless, some insight can be gained from a simplified analysis of the 1st-order low-pass companding filter shown in Fig. 2 a). For this analysis, it is assumed that the noise can be represented by a stationary noisy current source $i_N(t)$ having a constant power spectral density (PSD)

$$S_N = 4kT \cdot \gamma g_m = \frac{2\gamma}{n} \cdot 2qI_0 \tag{7}$$

where γ is the noise excess factor that accounts for additional noise sources that contribute to source $i_N(t)$ [7]. For a single device the noise excess factor γ is typically equal to $1/2$ for a BJT and $n/2 \approx 0.7$ for a MOST biased in weak inversion and at room temperature. It should be

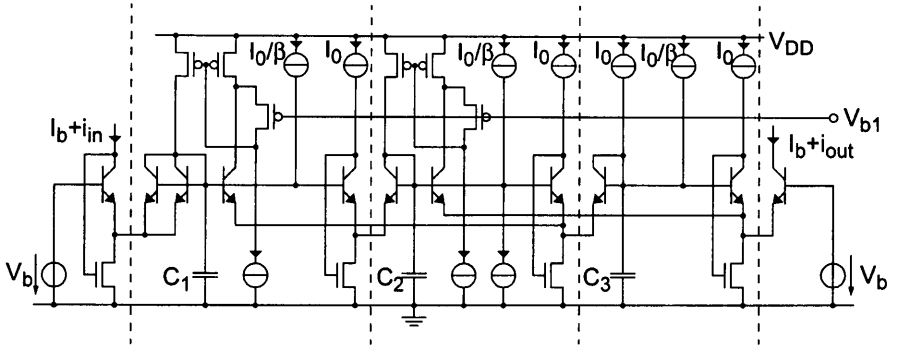


Fig. 6: Schematic of the class A 3 rd-order low-pass Chebyshev filter.

filter are shown in Fig. 7 for different bias conditions. The large tuning range observed in Fig. 7 is a fundamental property of log-domain filter resulting from the combination of having a transconductance proportional to the current on one hand and of the companding of the voltage on the other hand. Indeed, even if traditional linear bipolar transconductor also show the property of having their transconductance proportional to the bias current, they cannot exploit a similar wide tuning range due to the limited voltage swing inherent to linear transconductors. Since in log-domain the voltage swing on the capacitors is only a logarithmical function of the maximum to minimum current ratio, the latter can be increased without significantly affecting the voltages.

Note that the transfer functions of Fig. 7 have been measured at signal peak-to-peak currents equal to the bias current I_0. The dc gain error is explained by the fact that the supply voltage has been adjusted to the minimum acceptable value for each individual curve. This dc gain error vanishes when maintaining the same sufficiently large supply voltage for all the characteristics as can be seen in the inset of Fig. 7.

The total harmonic distortion (THD) has been measured at 1 kHz for two cut-off frequencies and is plotted in Fig. 8 versus the input current amplitude. Both measurements show that this filter can achieve good linearity ($THD < 1\%$) even for an input signal amplitude close to the bias current (corresponding to the hard clipping limit). The filter linearity has also been checked with a duotone input signal of equal amplitude. The fundamental and the 3rd-order intermodulation product

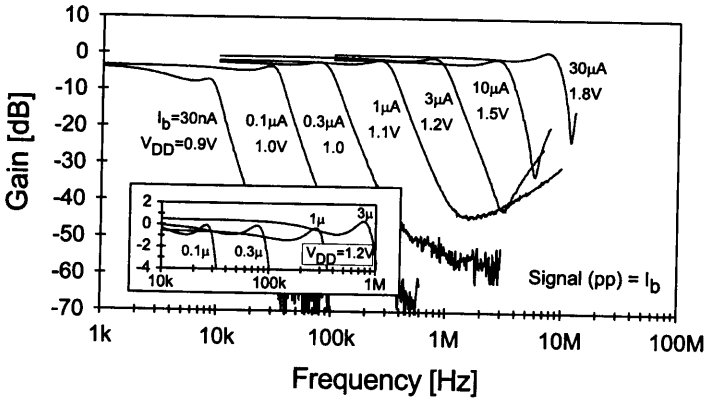


Fig. 7: Measured frequency response of a 3 rd-order Chebychev class A log-domain filter [39][40].

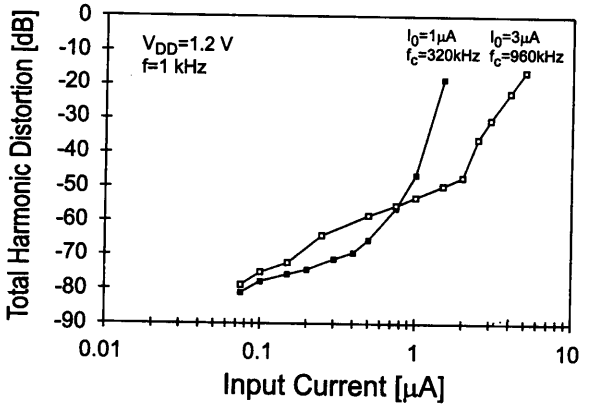


Fig. 8: Measured THD of a 3 rd-order Chebychev class A log-domain filter [39][40].

have been measured and are plotted versus the input signal amplitude in Fig. 9. The input referred intercept point (IIP3) is equal to $8\,\mu A_{rms}$, corresponding to 8 times the bias current. The noise floor has also been plotted in Fig. 9 and is about $1\ nA_{rms}$. This results in a DR of 57 dB for a 1 % THD or a spurious-free dynamic range (SFDR) of 47 dB.

Both of these measurements demonstrate that, despite their nonlinear internal operation, log-domain filters can achieve similar linearity than more traditional continuous-time filters and this at very low-voltage.

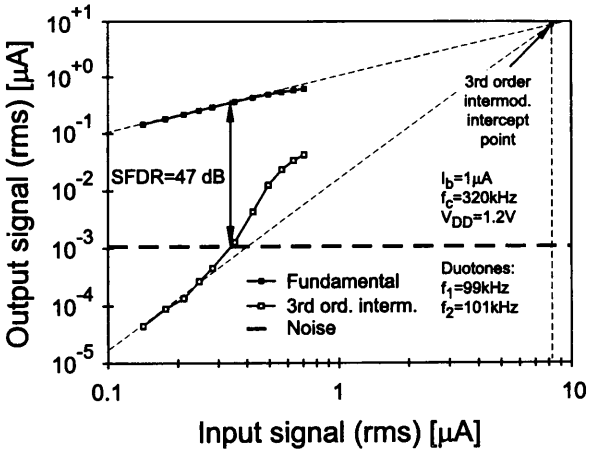


Fig. 9: Measured 3[rd]-order intermodulation intercept point of a 3 [rd]-order Chebychev class A log-domain filter [39][40].

One drawback of this class A implementation is its limited attenuation in the stop-band. This is due to a systematic mismatch between its inverting and its non-inverting input by a factor of $2/\beta$ due to the base currents of Q_3 and Q_4 [39][40]. The high frequency attenuation is also limited by the cut-off frequency of the MOS devices M_3 and M_4 forming the current mirror [39][40].

These problems can be circumvented by using the differential class AB structure presented in Fig. 10 and based on the same circuit concept [34][36][37]. The input and output signal are now represented by $i_{in} \equiv i_{inp} - i_{inn}$ and $i_{out} \equiv i_{outp} - i_{outn}$ respectively. Class AB implementations require an additional input signal conditioner that insures that all currents always stay positive. Although there are several ways to satisfy this constraint, a possible implementation imposes the product $i_{inp} \cdot i_{inn}$ to be equal to a constant (for example I_0^2) [21][37]. The input signal conditioner should be designed carefully in order to avoid adding any extra distortion due to mismatches within the conditioner [37].

The supply voltage required for the class AB integrator shown in Fig. 10 is typically equal to the V_{GS} voltage of the p-channel MOSFET M_9 plus the sum of the saturation voltages of Q_7 and M_3. The supply voltage may be reduced by replacing the current mirror M_9, M_8, M_6 (M_{10}, M_7, M_5) by a low-voltage folded cascode mirror [37]. Such a modified low-voltage version of the class AB integrator has been used to implement a 3[rd]-order low-pass Chebyshev filter with the same nominal 320 kHz cut-off frequency and 1 μA bias current as the class A realization discussed above [37]. The measured transfer functions of this filter are shown in Fig. 11 [36][37]. The same large tuning range can be observed for the class AB filter. The measured THD is plotted on the right y-axis of Fig. 12 versus the amplitude (upper x-axis) of a 1 kHz signal with the bias current I_0 set to 1 μA and the supply voltage to 1.2 V. The class AB operation allows the maximum peak value for a 1 % THD to be equal to 15 μA, corresponding to a modulation index (defined as the ratio between the current amplitude and the bias current) as high as 15. The left y-axis of Fig. 12 presents the measured SNR versus the signal rms value (lower x-axis). These measurements are difficult to achieve and therefore they could only be performed in the case of a dc input signal. The result confirms the saturation of the SNR occurring at signal levels much larger that the bias current as predicted by the theory given in Section III A. In this particular case, the maximum SNR value of 52.5 dB can be estimated by multiplying the contribution of the most noisy integrator among all three (the one having the smallest capacitor) by a factor three, considering that the noise factor of that individual integrator is about 5 [37]. The noise current is about 6.3 nA_{rms} resulting in a DR of 65 dB for a 1% THD. As shown in Fig. 12, the class AB operation allows to extend the DR by about 23 dB compared to what would be achievable with a class A filter hav-

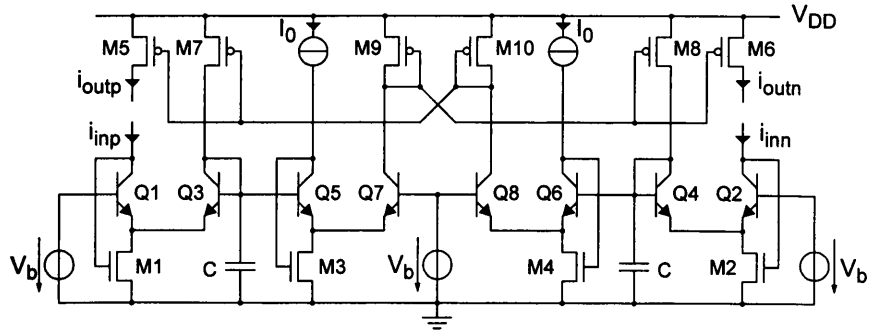


Fig. 10:LV BiCMOS class AB integrator [34][36][37].

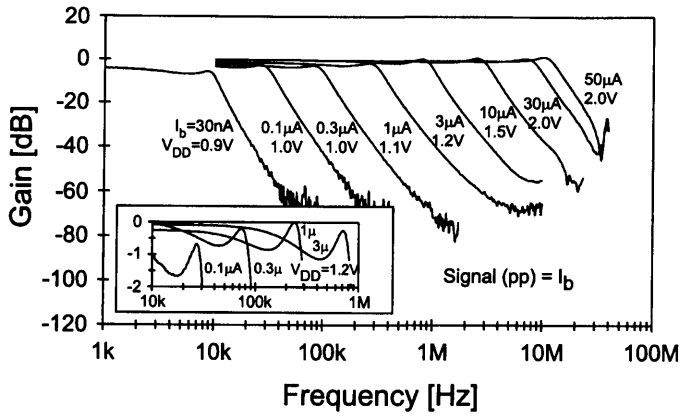


Fig. 11:Measured frequency response of a 3 rd-order Chebychev class AB log-domain filter [36][37].

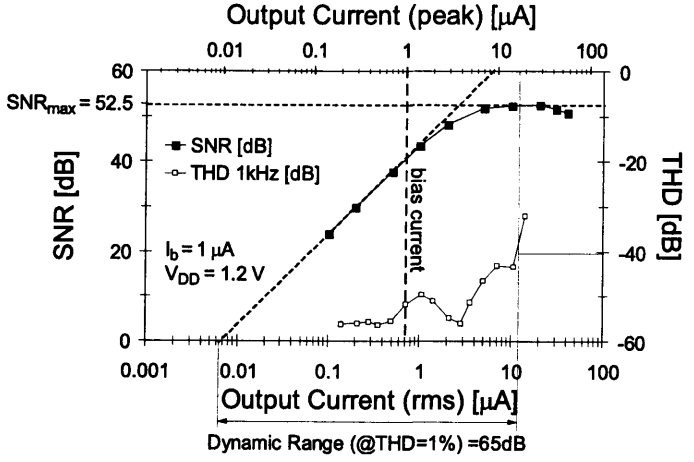


Fig. 12:Measured SNR and THD of a 3 rd-order Chebychev class AB log-domain filter [36][37].

ing the same noise floor. In reality the gain in terms of DR with respect to the class A implementation discussed above is only 8 dB. This is due to the fact that the class AB filter is about 6 times more noisier than the class A filter (the noise floor of the class AB filter is about 15 dB higher than the class A filter). A duotone measurement has also been performed and the measured amplitudes of the fundamental and the 3rd-order intermodulation product are plotted versus the input signal amplitude in Fig. 13. The SFDR is about 46 dB.

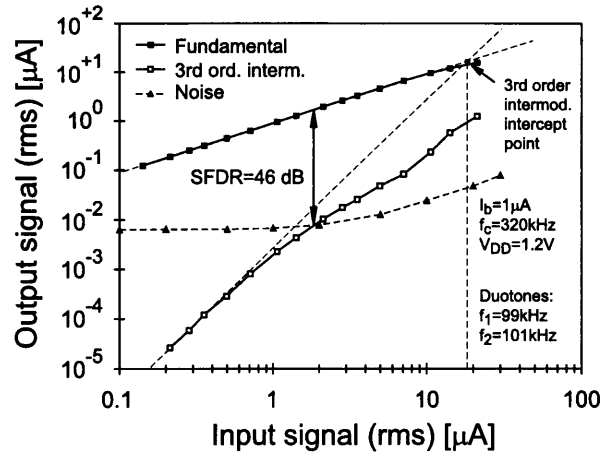
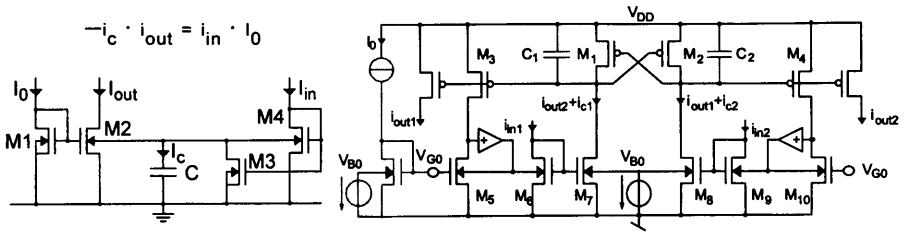


Fig. 13:Measured 3rd-order intermodulation and intercept point of a 3 rd-order Chebychev class AB log-domain filter [36][37].

V. CMOS Implementations

An alternative way to implement log-domain filters is to use MOS transistors biased in weak inversion. Although the circuits could be derived from the already existing bipolar implementations by simply substituting the BJTs with weakly inverted MOSTs, a better and more innovative approach is to account for the differences existing between these two devices and take advantage of the MOST properties. One of the important differences between the MOST and the BJT is that the transconductance from the gate is n times smaller than the transconductance from the source (or the transconductance of a BJT), where the slope factor n is related to the body effect [48]. Furthermore, this slope factor varies with respect to the gate-to-bulk voltage and therefore the g_m/I_D characteristic is not flat in weak inversion but shows a maximum value before decreasing when entering the moderate inversion region (see the right part of Fig. 15 a). Also, the current range where the approximately exponential I-V characteristic can be used, is limited to typically three decades compared to more than six decades for the BJT. This will have an impact on how to choose and optimize the operating point within the weak inversion region.

Also, the operating frequency is limited to rather low-frequencies



a) Class A implementation. *b) Differential class AB implementation.*

Fig. 14: Examples of CMOS log-domain integrators.

(typically below 100 kHz) due to the wide transistors required in order to move the operating point into the weak inversion region.

Another fundamental limitation related to the use of MOSTs in weak inversion is the poor matching of the threshold voltages, which as will be shown later, also affects the distortion performance of CMOS log-domain filters.

Unlike the BJT, which is a three terminal device, the MOST is basically a four terminal device, the fourth terminal being the well (or the substrate). The well can be advantageously used in order to implement given translinear functions in a very compact and efficient way [56], [47], [49]-[52]. Two examples of CMOS log-domain integrators that use the well as an active terminal are shown in Fig. 14. The left one (Fig. 14 a) is a very compact class A implementation whereas the right one is a differential class AB integrator based on a compact bulk-driven multiplier [57]. The MOS transistors used in these kind of circuits very often have their sources tight together. In this particular case it is useful for analysis to use the source potential as a common reference. In such condition, the drain current of a MOST biased in weak inversion and in saturation is controlled by the two voltages V_{GS} and V_{BS} according to [5][7][48]

$$I_D = I_{D0} \cdot \exp\left(\frac{V_{GS}}{n \cdot U_T}\right) \cdot \exp\left(\frac{n-1}{n} \cdot \frac{V_{BS}}{U_T}\right) \qquad (15)$$

where I_{D0} is the leakage current (defined as the drain current for $V_{GS} = V_{BS} = 0$) which is proportional to the aspect ratio W_{eff}/L_{eff} and to $\exp(-V_{TO}/U_T)$ [5][48].

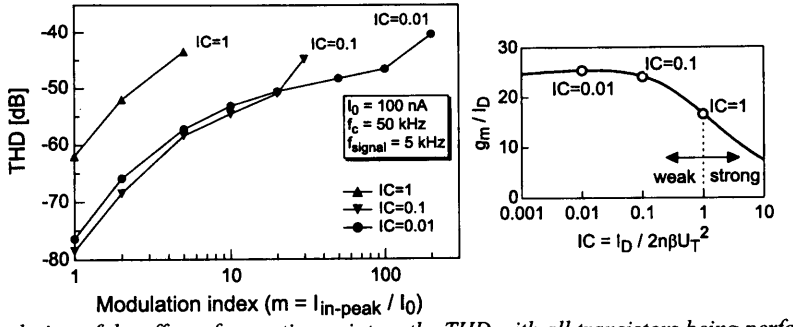
Using (15) and assuming that all the transistors of the circuit shown in Fig. 14 a) are matched, it can be shown that this circuit also satisfies (6) (with the minus sign) and therefore conforms to the log-domain integrator principle. Since i_{in} must be always positive, this circuit is only able to discharge the capacitor. This drawback can be circumvented and the circuit can be extended to include a non-inverting input by adding a p-channel and an n-channel current mirror as described in [47]. A fully differential implementation of the same circuit can also be found in [47].

The integrator of Fig. 14 b) is a fully differential class AB implementation based on a compact bulk-driven multiplier [57]. Note that thanks to the small voltage swing on the integrating capacitors (not exceeding 4 to $5nU_T$), C_1 and C_2 can be implemented by the nonlinear parasitic capacitances C_{gs} and C_{gb} of the rather large pMOS devices. This makes this kind of integrator well suited for an integration in a standard CMOS digital process [46].

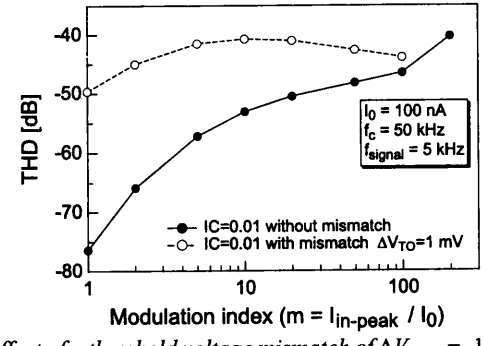
The integrator of Fig. 14 b) has been connected as a 1st-order low-pass filter and simulated using the EKV MOST model [48]. The THD versus the modulation index m (defined as the ratio of the peak input current i_{in} to the bias current I_0) is plotted in Fig. 15 for different values of the inversion coefficient $IC = I_0/(2n\beta U_T^2)$ with $\beta \equiv \mu C_{ox} W_{eff}/L_{eff}$. The right plot of Fig. 15 a) shows the position of the operating points on the g_m/I_D characteristic corresponding to the different ICs.

Fig. 15 a) indicates that large modulation index can be reached when setting the quiescent operating point on the flat part of the g_m/I_D characteristic. It is also interesting to note that the THD remains surprisingly low although some transistors work already in the moderate or even strong inversion region. Indeed, for a THD of about -45 dB, the maximum current corresponds to an inversion coefficient of a few units (typically between 3 and 5).

The degradation of the THD with a threshold voltage mismatch between M_1 and M_2 is illustrated in Fig. 15 b). It is clearly the dominant source of distortion at low current, whereas the deviation from the ideal exponential function when progressively leaving the weak inversion region becomes dominant at high current. This tends to demonstrate that despite the high threshold voltage mismatches, the log-



a) Simulation of the effect of operating point on the THD with all transistors being perfectly matched.



b) Simulation of the effect of a threshold voltage mismatch of $\Delta V_{TO} = 1mV$ on transistors M_1 and M_2 assuming the other transistors are perfectly matched.

Fig. 15: Simulation of the THD for different operating points (inversion coefficients *IC*) and for a threshold voltage mismatch.

domain technique can still be used when implemented in CMOS using weakly inverted MOSTs. Nevertheless, this still has to be demonstrated experimentally. Other CMOS implementations of log-domain filters can be found in the literature [44][45].

VI. Power Efficiency of Low-Voltage Continuous-Time Filters

In this Section, the power consumption required for implementing a given cut-off frequency f_c for a given dynamic range (DR) and total harmonic distortion (THD) will be derived as a function of the supply voltage V_{DD} for both a linear g_m-C and a log-domain 1st-order low-pass filter. The two structures will then be compared using the figure of merit defined below. The same figure of merit will then be calculated and compared for other representative low-voltage and low-power CTFs already published in the literature.

A. Definition of a factor of merit

The power efficiency of CTFs can be compared by means of the ratio p of the power-per-pole and edge frequency normalized to the dynamic range

$$p = \frac{power\ per\ pole\ and\ per\ edge\ frequency}{dynamic\ range}. \qquad (16)$$

It can be shown that the power required for implementing a single pole with a 100 % current efficient transconductor operating at rail-to-rail is given by [1]-[5]

$$P = 8kT \cdot f \cdot DR. \qquad (17)$$

The power-per-pole and edge frequency normalized to the dynamic range of this ideal filter is therefore equal to $8kT = 32\times10^{-21} J$ at room temperature. Comparison can be done by defining a factor of merit as

$$\xi = \frac{8kT}{p} = \frac{32\times10^{-21}}{p} < 1 \qquad (18)$$

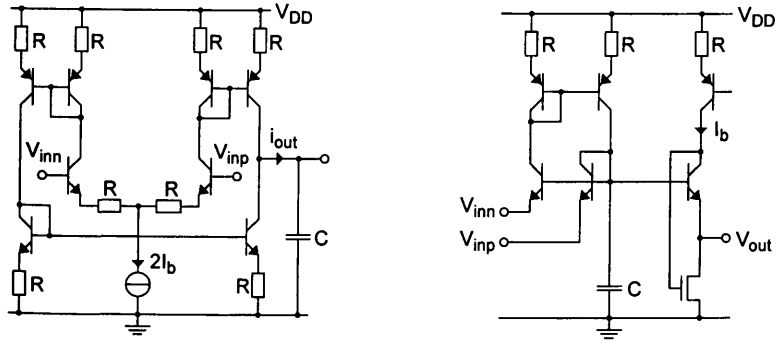


Fig. 16: Comparison of (a) linearized g_m-C integrator and (b) log-domain integrator.

B. Power efficiency of a linear g_m-C filter

The key problem of implementing a linear integrator is addressed by different approaches [53], which have all their individual advantages and problems. The g_m-C integrator with emitter-degeneration shown in Fig. 16 a) is chosen as an example for demonstrating the major design constraints, since it is rather power efficient and widely used.

This integrator is assumed to be connected as a first-order low-pass filter and driven by a sinusoidal input voltage having a frequency $\omega/(2\pi)$ much lower than the filter cut-off frequency f_c and an amplitude V_p equal to half the maximum swing defined as

$$V_p = (V_{DD} - V_{min})/2 \tag{19}$$

where $V_{min} = V_{bias} + 2U_R$. The voltage V_{bias} is an incompressible voltage, typically equal to one junction voltage plus two saturation voltages

$$V_{bias} = V_{BE} + 2 \cdot V_{CEsat} \cong 1V \tag{20}$$

and $U_R = R \cdot I_b$ is the quiescent voltage drop across the degenerating resistors.

The power consumption of this 1st-order low-pass filter is given by

$$P = 4 \cdot I_b \cdot V_{DD} \tag{21}$$

where I_b is the bias current required to implement the pole $f_c = G_{m0}/(2\pi \cdot C)$ with

$$G_{m0} = \frac{I_b}{U_T + U_R} \tag{22}$$

being the small-signal transconductance. The capacitance C is determined by the thermal noise power and the required dynamic range according to

$$C = \frac{\gamma_{lin} \cdot kT}{V_p^2/2} \cdot DR \tag{23}$$

where γ_{lin} is the noise excess factor of the linearized transconductor defined as the product of the equivalent input thermal noise resistance

and the transconductance G_{m0} [5][7]. Although γ_{lin} is a function of the bias current, it will be considered as constant in the following analysis.

From (22) and (23), the bias current can be rewritten as

$$I_b = G_{m0} \cdot (U_T + U_R) = 2\pi \cdot f_c \cdot C \cdot (U_T + U_R) \tag{24}$$

$$= 4\pi \cdot f_c \cdot \gamma_{lin} \cdot kT \cdot DR \cdot \frac{U_T + U_R}{V_p^2}$$

where U_R is set in order to meet the linearity requirements. It is therefore a function of the signal amplitude and the desired THD. The THD of the transconductor can be estimated from the large-signal transfer characteristic of the transconductor given by

$$i_{out} = I_b \cdot \tanh\left(\frac{v_{in} - R \cdot i_{out}}{2U_T}\right) \tag{25}$$

where $v_{in} = v_{inp} - v_{inn}$ is the differential input voltage. For low distortion, the term $R \cdot i_{out}$ in (25) can be approximated by $G_{m0} \cdot R \cdot v_{in}$ and the tanh function can replaced by its third-order Taylor power series. The following rule of thumb can then be found, which is useful for THD values up to a few percent

$$THD \cong \frac{1}{12} \cdot \left[(1 - G_{m0} \cdot R) \cdot \frac{V_p}{2U_T}\right]^2. \tag{26}$$

Replacing G_{m0} by (22), results in a condition for the voltage drop required to process a sinusoidal input voltage of amplitude V_p at a maximum THD

$$U_R \geq \frac{V_p}{2\sqrt{12 \cdot THD}} - U_T. \tag{27}$$

Equations (19) and (27) can then be solved with respect to V_p and U_R which become functions of V_{DD} and THD and can be replaced in the bias current equation (24). The power consumption can then be rewritten as a function of V_{DD}, DR, f_c and THD as

$$P = \frac{2\pi}{3} \cdot f_c \cdot \gamma_{lin} \cdot kT \cdot DR \cdot \frac{V_{DD}}{V_{DD} - V_{bias} + 2U_T} \cdot \frac{1 + 2\sqrt{12 \cdot THD}}{THD} \tag{28}$$

The related factor of merit of the g_m-C integrator is then given by

$$\xi = \frac{12}{\pi} \cdot \frac{1}{\gamma_{lin}} \cdot \frac{V_{DD} - V_{bias} + 2U_T}{V_{DD}} \cdot \frac{THD}{1 + 2\sqrt{12 \cdot THD}} \tag{29}$$

According to (29), the factor of merit is inversely proportional to the noise factor γ_{lin} and approximately proportional to the ratio of the available signal swing over the supply voltage [3][6]. It is also approximately proportional to the acceptable total harmonic distortion.

It has to be mentioned that this circuit has to be modified for practical implementations to allow frequency tuning. This can be achieved by inserting a variable voltage shift between the emitters of the PNP current mirrors [54]. For maintaining linear operation these two current mirrors can no longer be degenerated. Consequently, their noise contribution is much higher and the power efficiency of the transconductor becomes much lower.

C. Power efficiency of a log-domain filter

The power consumption of the log-domain integrator shown in Fig. 16 b) can be estimated in a similar way. In contrast to the previous circuit, the linearity of the log-domain integrator is only limited by parasitic elements and by device non-idealities such as series resistances and output conductances of the BJTs but no longer by the values of the components. Therefore, it is assumed below that the required linearity can be achieved without any additional power consumption. As can be seen from the presented measured results, this is a reasonable assumption as long as the required THD is in the range of -40dB to -60dB. The aim of this derivation is an estimation of the noise generated per pole by a high-order filter, therefore the contribution of the compressor at the filter's input and of the expandor at the filter's output are neglected.

For a better comparison with the g_m-C integrator, the current mirror and the bias current source are both using resistively degenerated bipolar transistors, causing a voltage drop U_R across these resistors.

The integrator is again assumed to be connected as a first-order low-pass filter and driven by a signal of much lower frequency than its cut-off frequency. The DR of the integrator is estimated at the output, after the expandor, in the domain of uncompressed currents. The signal's peak value is then limited by the bias current I_b and consequently the DR is given by

$$DR = \frac{I_b^2/2}{G_m^2 \cdot (\gamma_{log} \cdot kT)/C} = \frac{U_T^2 \cdot C}{2 \cdot kT \cdot \gamma_{log}} \tag{30}$$

where γ_{log} is the noise excess factor of the log-domain integrator which rigorously is a function of the bias current, but which can be assumed constant for a first-order analysis. The required capacitance is determined from (30) as in the case of the linearized integrator

$$C = \frac{2 \cdot kT \cdot \gamma_{log}}{U_T^2} \cdot DR \tag{31}$$

From (4) and the expression of the cut-off frequency $f_c = 1/(2\pi \cdot \tau)$ the required bias current can be found with $I_0 = I_b$ and $V_0 = U_T$

$$I_b = 2\pi \cdot f_c \cdot C \cdot U_T = 4\pi \cdot f_c \cdot \gamma_{log} \cdot kT \cdot DR/U_T. \tag{32}$$

The power dissipation and the factor of merit of the log-domain integrator are then given by

$$P = 3 \cdot I_b \cdot V_{DD} = 12\pi \cdot f_c \cdot \gamma_{log} \cdot kT \cdot DR \cdot V_{DD}/U_T \tag{33}$$

$$\xi = \frac{2}{3\pi} \cdot \frac{1}{\gamma_{log}} \cdot \frac{U_T}{V_{DD}} \tag{34}$$

According to (34), the factor of merit of the log-domain integrator is inversely proportional to V_{DD} and inversely proportional to the noise factor γ_{log}. The minimum supply voltage is still limited by the bias voltage V_{bias} given by the sum of a junction voltage plus two saturation voltages. The factor of merit given by (34) reflects the fact that the available signal swing of log-domain filters is almost independent of the supply voltage. Of course larger supply voltage can be used for

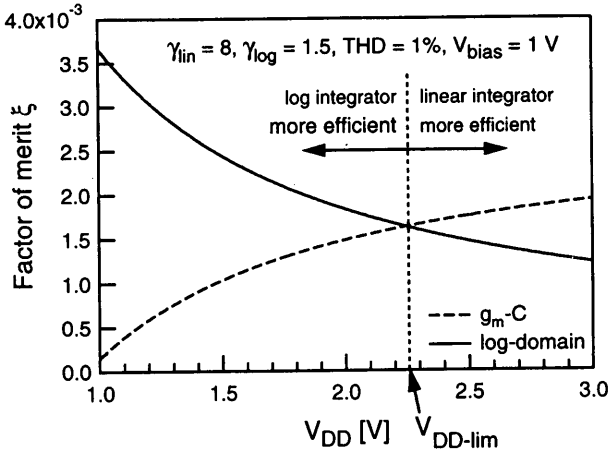


Fig. 17:Power efficiency comparison.

degenerating the current mirrors and therefore reducing their noise contribution and therefore the overall noise factor γ_{log}.

Both factors of merit are plotted in Fig. 17 versus the supply voltage V_{DD} for $\gamma_{lin} \cong 8$, $\gamma_{log} \cong 1.5$, $V_{bias} \cong 1V$ and $THD = 1\%$. It can be seen that below a certain supply voltage $V_{DD\text{-}lim}$, defined as the voltage for which both integrators are equally efficient and given by

$$V_{DD\text{-}lim} = V_{bias} + U_T \cdot \left[\frac{\gamma_{lin}}{\gamma_{log}} \cdot \frac{1 + 2 \cdot \sqrt{12 \cdot THD}}{18 \cdot THD} - 2 \right], \qquad (35)$$

the log-domain integrator becomes more power efficient than the g_m-C integrator. According to (35), $V_{DD\text{-}lim}$ is mainly determined by the acceptable distortion and by the ratio of the noise excess factors. In the case considered here, this limit is typically 2.25 V for the examples discussed above. But as shown in Fig. 18, $V_{DD\text{-}lim}$ quickly raises as the THD is decreased. For $THD = 0.1\%$, $V_{DD\text{-}lim}$ becomes as high as 10 V and the log-domain integrator would be about 37 times more power efficient than its equivalent g_m-C counterpart for $V_{DD} = 1.2V$. This simple analysis clearly demonstrates that log-domain filters are more power efficient than their g_m-C equivalent.

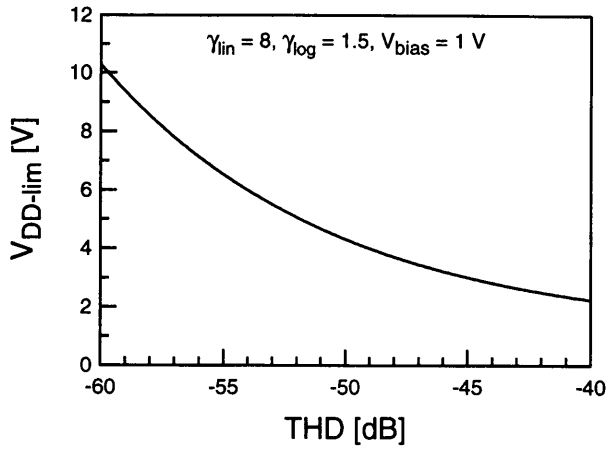


Fig. 18:Supply voltage limit corresponding to equal power efficiencies for both the g_m-C and the log-domain integrators versus the required THD.

D. Comparison with other published LV CTFs

The factor of merit defined above has been computed from available data published in the literature and the results are presented in Table I. The latter clearly demonstrate that the continuous-time current mode filters (including log-domain filters) corresponding to references [37][39] and [59] show the best factor of merit. It is interesting to note that the class AB log-domain filter presented in [37] is about 260 times better than the other 1 V bipolar filter presented in [65] but is still 669 times lower than the factor of merit of the ideal filter (by definition equal to unity). This suggests that there is still a lot of room for improvement.

Part of this difference may be attributed to the fact that even if the voltage swing is reduced to a few U_T, the circuit still needs an incompressible bias voltage V_{bias} in order to operate correctly which degrades its ξ.

TABLE I: COMPARISON OF THE FACTOR OF MERIT OF SEVERAL CTFS PUBLISHED IN THE LITERATURE

Process	Supply voltage	Power cons.	Cut-off freq.	Order	Power per pole and edge freq.	DR	ξ	Ref.
-	[V]	[W]	[Hz]	-	[J]	[dB]	-	-
BiCMOS	1.2	65.0E-6	320.0E+3	3	67.7E-12	65	**1.495E-3**	[37]
CMOS	1.5	75.0E-6	525.0E+3	1	142.9E-12	67	**1.122E-3**	[59]
BiCMOS	2.5	13.0E-3	600.0E+3	7	3.1E-9	79	**8.197E-4**	[61]
BiCMOS	1.2	23.0E-6	320.0E+3	3	24.0E-12	57	**6.711E-4**	[39][1]
BiCMOS	1.2	59.0E-6	950.0E+3	3	20.7E-12	52	**2.451E-4**	[39][2]
BiCMOS	3	11.5E-3	30.0E+6	2	191.7E-12	60.5	**1.873E-4**	[60]
bipolar	2.5	2.5E-3	192.0E+3	3	4.3E-9	69	**5.848E-5**	[63]
bipolar	1.3	6.0E-6	8.0E+3	2	375.0E-12	57	**4.274E-5**	[62]
CMOS	3	12.6E-6	945.0E+0	3	4.4E-9	63	**1.437E-5**	[64]
bipolar	1	100.0E-6	9.0E+3	5	2.2E-9	56	**5.747E-6**	[65]

1) With a bias current of 1 μA
2) With a bias current of 3 μA

VII. CONCLUSION

It has been shown that the log-domain technique is well suited for the implementation of LV CTFs, since the voltage swings become almost negligible with respect to the supply voltage required to properly bias the circuit. It also offers a high current efficiency since the design is based on the nonlinear device characteristics and therefore no additional linearization scheme is required.

The SNR of log-domain is very similar to classical linear circuit as long as the circuits operate in Class A. For class AB log-domain circuits, the noise starts to increase proportionally to the signal when the current becomes larger than the bias current. As a result the SNR starts to saturate to a maximum value. This fundamental property of class AB log-domain filters allows to extend the dynamic range without further increasing the SNR and the quiescent power consumption, since the latter is proportional to the SNR. Any device nonideality that make the device I-V characteristic deviate from the original exponential law will affect the performance of the circuit by introducing distortion. These nonidealities include mainly series resistance, high current effects and device mismatch. In addition, base currents will also affect the performance of log-domain circuits in a similar way they affect the operation of translinear loops.

Although most of the current implementations have used bipolar transistors, log-domain circuits can also be built with MOS transistors biased in weak inversion. CMOS implementations can take advantage of the differences existing between the BJT and the MOST to obtain very compact structure. Examples of BiCMOS and standard digital CMOS realizations operating with supply voltages as low as 1 V have been presented and demonstrate the feasibility and the potential of log-domain circuits for LV and LP applications.

The power consumption of a 1st-order low-pass gm-C filter and the equivalent log-domain filter is derived. Both filters are then compared with respect to a figure of merit defined as the power consumption per pole and cut-off frequency normalized to the dynamic range. This comparison shows that for low-voltage (< 2 V) log-domain filters become more power efficient than their traditional gm-C counterpart. This is mainly due to the fact that gm-C filter require the transconductor to be linearized in order to achieve the desired linearity resulting in a decrease of their power efficiency. The power consumption of log-domain filters is proportional to the product of the cut-off frequency, the dynamic range and the supply voltage. The same figure of merit is then used to compare different continous-time filters published in the literature including the two BiCMOS log-domain filters presented in this chapter. The class AB log-domain filter presented herein shows the best factor of merit.

The log-domain technique seems therefore very promising for the realization of continuous-time filters that have to operate at low-voltage and consume as little power as possible while maintaining reasonable dynamic range. However, the design of log-domain circuits is still at its beginning and yet a lot of work has to be accomplished to better understand the limitations in order to fully exploit this new technique and clearly demonstrate its advantage over the more classical approaches.

ACKNOWLEDGMENTS

This work was part of the TIBIA II ESPRIT IV project and is extended in the ALPINS ESD-Low-Power project, both in association with Siemens HL. The contribution to these projects described herein this paper is funded by the Swiss Office for Sciences and Education.

REFERENCES

[1] E. A. Vittoz, "Future of Analog in the VLSI Environment," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 1372-1375, May 1990.

[2] E. A. Vittoz, "Low-Power Design: Ways to Approach the Limits," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 14-18, May 1994.

[3] E. A. Vittoz, "Low-Power/Low-Voltage Limitations and Prospects in Analog Design," in *Advances in Analog Circuit Design Workshop*, Eds. R. S. van de Plassche, W. M. C. Sansen and J. H. Huijsing, Boston, Kluwer Academic Publishers, 1994, first published in Eindhoven, The Netherlands, March 29-31, 1994.

[4] R. Castello and P. R. Gray, "Performance Limitation in Switched-Capacitor Filters," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 865-876, Sept. 1985.

[5] C. C. Enz and E. A. Vittoz, "CMOS Low-power analog circuit design, " in Emerging Technologies, Tutorial for 1996 International Symposium on Circuits & Systems, Edited by R.Cavin and W.Liu, IEEE Service Center, Piscataway, NJ, 1996, chapter 1.2, pp. 79-133, ISBN 0-7803-3328-4.

[6] R. Castello, F. Montecchi, F. Rezzi and A. Baschirotto, "Low-Voltage Analog Filters," *IEEE Trans. Circuits Syst. I*, vol. CAS-42, Nov. 1995.

[7] E. A. Vittoz, *Micropower techniques*, in *Design of MOS VLSI circuits for telecommunications*, edited by J. Franca and Y. Tsividis, Prentice-Hall, 1994.

[8] B. Gilbert, "Translinear circuits: a proposed classification," *Electron. Lett.*, vol. 11, pp. 14-15, 1975.

[9] Y. Tsividis, V. Gopinathan, and L Tóth, "Companding in signal processing," *Electron. Lett.*, vol. 26, pp. 1331-1332, Aug. 1990.

[10] E. M. Blumenkrantz, "The analog floating point technique," *IEEE Symp. on Low-Power Electronics Dig. of Tech. Papers*, pp. 72-73, Oct. 1995.

[11] Y. Tsividis and D. Li, "Current-mode filters using syllabic companding," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 121-124, May 1996.

[12] Y. Tsividis, "On Linear Integrators and Differentiators Using Instantaneous Companding", *IEEE Trans. Circuits Syst. II*, vol. CAS-42, pp. 561-564, Aug. 1995.

[13] Y. Tsividis, "General approach to signal processors employing companding," *Electron. Lett.*, vol. 31, pp. 1549-1550, Aug. 1995; errata, *Electron. Lett.*, vol. 32, p. 857, Apr. 25, 1996.

[14] Y. Tsividis, "Externally Linear, Time-Invariant Systems and Their Application to Companding Signal Processors," *IEEE Trans. Circuits Syst. II*, vol. CAS-44, pp. 65-85, Feb. 1997.

[15] Y. Tsividis, "Instantaneously Companding Integrators," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 477-480, June 1997.

[16] C. C. Enz," Low-Power Log-Domain Continuous-Time Filters: an Introduction", ESSCIRC'95 WorkShop on LP and LV, Sept. 1995.

[17] R. W. Adams, "Filtering in the log domain," Preprint 1470, presented at 63rd Audio Eng. Soc. Conf., New York, May 1979.

[18] E. Seevinck, "Companding current-mode integrator: A new circuit principle for continuous-time monolithic filters," *Electron. Lett.*, vol. 26, pp. 2046-2047, Nov. 1990.

[19] D. R. Frey, "A general class of current mode filters," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 1435-1438, May 1993.

[20] D. R. Frey, "Log domain filtering: An approach to current mode filtering," *IEE Proc.*, pt. G, vol. 140, pp. 406-416, Dec. 1993.

[21] D. R. Frey, "Current mode class AB second order filter," *Electron. Lett.*, Vol. 30, pp. 205-206, Feb. 1994.

[22] D. R. Frey, "A 3.3 V electronically tunable active filter usable to beyond 1 GHz," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 493-496, May 1994.

[23] D. R. Frey, "Exponential State space filters: A generic current mode design strategy," *IEEE Trans. Circuits Syst. I*, vol. CAS-43, pp. 34-42, Jan. 1996.

[24] D. R. Frey, "Log-Domain Filtering for RF Applications," *IEEE Journal of Solid-State Circuits*, vol. SC-31, pp. 1468-1475, Oct. 1996.

[25] D. Frey, "State Space Synthesis of Log-Domain Filters," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 481-484, June 1997.

[26] D. Perry and G. W. Roberts, "Log-domain filters based on LC Ladder synthesis," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 311-314, May 1995.

[27] J. Mulder, W. A. Serdijn, A. C. van der Woerd, A. H. M. van Roermund, "A Syllabic Companding Translinear Filter," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 101-104, June 1997.

[28] M. El-Gamal and G. W. Roberts, "LC Ladder-Based Synthesis of Log-Domain Bandpass Filters," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 105-108, June 1997.

[29] V. W. Leung, M. El-Gamal and G. W. Roberts, "Effects of Transistor Nonidealities on Log-Domain Filters," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 109-112, June 1997.

[30] M. El-Gamal, V. W. Leung and G. W. Roberts, "Balanced Log-Domain Filters for VHF Applications," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 493-496, June 1997.

[31] E. M. Drakakis, A. J. Payne and C. Toumazou, "Log-Domain Filters, Translinear Circuits and the Bernouilli Cell," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 501-504, June 1997.

[32] R. Fox, M. Nagarajan and J. Harris, "Practical Design of Single-Ended Log-Domain Filter Circuits," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 341-344, vol. 1, June 1997.

[33] M. Punzenberger and C. C. Enz, "Low-Voltage Companding Current-Mode Integrators", *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 2112-2115, May 1995.

[34] M. Punzenberger and C. C. Enz, "A new 1.2 V BiCMOS log-domain integrator for companding current-mode filters," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 125-128, May 1996.

[35] M. Punzenberger and C. C. Enz, "A 1.2 V BiCMOS Companding Current-Mode Integrator for Σ-Δ Modulators, " *Proc. IEEE Int. Conf. on Electronics, Circuits, and Syst.*, vol. 1, pp. 235-238, Oct. 1996.

[36] M. Punzenberger and C. C. Enz, "A 1.2 V BiCMOS Class AB Log-Domain Filter," *Int. Solid-State Circ. Conf. Dig. of Tech. Papers*, pp. 56-57, Feb. 1997.

[37] M. Punzenberger and C. C. Enz, "A 1.2 V Low-power BiCMOS Class AB Log-Domain Filters," *IEEE Journal of Solid-State Circuits*, vol. SC-32, pp. 1968-1978, Dec. 1997.

[38] M. Punzenberger and C. C. Enz, "Noise in instantaneous companding filters," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 337-340, vol. 1, June 1997.

[39] M. Punzenberger and C. C. Enz, "A Low-Voltage Power and Area Efficient BiCMOS Log-Domain Filter," *Proc. of the European Solid-State Circ. Conf.*, pp. 256-259, Sept. 1997.

[40] M. Punzenberger and C. Enz, "A Compact Low-Power BiCMOS Log-Domain Filter," to be published in the *IEEE Journal of Solid-State Circuits*, July 1998.

[41] G. van Ruymbeke, C. Enz, F. Krummenacher and M. Declercq "A BiCMOS Programmable Continuous-Time Filter using Voltage-Companding", *Proc. IEEE Custom Integrated Circuits Conf.*, pp.93-96, May 1996.

[42] G. van Ruymbeke, C. C. Enz, F. Krummenacher and M. Declercq, "A BiCMOS Programmable Continuous-Time Filter Using Image-Parameter Method Synthesis and Voltage-Companding Technique, " *IEEE Journal of Solid-State Circuits*, vol. SC-32, pp. 377-387,

March 1997.
[43] C. C. Enz, M. Punzenberger and D. Python, "Low-Voltage Log-Domain Signal Processing in CMOS and BiCMOS," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 489-492, vol. 1, June 1997.
[44] C. Toumazou, J. Ngarmnil, and T. S. Lande, "Micropower log-domain filter for electronic cochlea," *Electron. Lett.*, vol. 30, pp. 1839-1841, Oct. 1994.
[45] C. Toumazou, J. Ngarmnil, and T. S. Lande, "Micropower log-domain filter for electronic cochlea," *Proc. of the European Solid-State Circ. Conf.*, pp. 86-89, Sept. 1995.
[46] D. Python, R. Fried, C. C. Enz, "A 1.2 V Companding Current-Mode Integrator for Standard Digital CMOS Processes, " *Proc. IEEE Int. Conf. on Electronics, Circuits, and Syst.*, vol. 1, pp. 231-234, Oct. 1996.
[47] R. Fried, D. Python and C. C. Enz, "Compact log-domain current-mode integrator with high transconductance-to-bias current ratio," *Electron. Lett.*, vol. 32, pp. 952-953, May 1996.
[48] C. C. Enz, F. Krummenacher and E. A. Vittoz, "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications," Analog Integrated Circuits and Signal Processing, 8, pp. 83-114, July 1995.
[49] R. Fried and C. C. Enz, "Bulk driven MOST transconductor with extended linear range," *Electron. Lett.*, vol. 32, pp. 638-639, March 1996.
[50] R. Fried and C. C. Enz, "CMOS parametric current amplifier," *Electron. Lett.*, vol. 32, pp. 1249-1250, July 1996.
[51] R. Fried and C. C. Enz, "MOST implementation of the Gilbert sin(x) Shaper," *Electron. Lett.*, vol. 32, pp. 2073-2074, Oct. 1996.
[52] R. Fried and C. C. Enz, "Simple and Accurate Voltage Adder/Subtractor," *Electron. Lett.*, vol. 33, pp. 944-945, May 1997.
[53] Y. Tsividis and J. O. Voorman, Eds., Intregrated Continuous-Time Filters. Piscataway: IEEE Press, 1993.
[54] J. O. Voorman, W. H. A. Brüls, P. J. Barth, "Integration of Analog Filters in a Bipolar Process," *IEEE Journal of Solid-State Circuits*, vol. SC-17, no. 4, pp. 713-722, Aug. 1982, also reprinted in [53].
[55] J. Mulder, A. C. van der Woerd, W. A. Serdijn, A. H. M. van Roermund, "General Current-Mode Analysis Method for Translinear Filters," *IEEE Trans. Circuits Syst. II*, vol. CAS-44, pp. 193-197, March 1997.
[56] J. Mulder, A. C. van der Woerd, W. A. Serdijn and A. H. M. van Roermund, "Application of the Back Gate in MOS Weak Inversion Translinear Circuits," *IEEE Trans. Circuits Syst. I*, vol. CAS-42, Nov. 1995.
[57] M. V. D. Gevel and J. C. Kuenen, "$\sqrt{x}$ Circuit Based on a Novel, Back-Gate-Using Multiplier," *Electron. Lett.*, vol. 30, pp. 183-184, March 1994.
[58] F. Yang, C. C. Enz and G. van Ruymbeke, "Design of Low-Power & Low-Voltage Log-Domain Filters," *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 117-120, May 1996.
[59] R. H. Zele and D. J. Allstot, "Low-Power Continuous-Time Filters," *IEEE Journal of Solid-State Circuits*, vol. SC-31, no. 2, pp. 157-168, Feb. 1996.
[60] F. Rezzi, A. Baschirotto and R. Castello, "A 3V 12-55 MHz BiCMOS Continuous-Time Filter with Pseudo-Differential Structure," , pp. 74-77, Lille, 1995.
[61] F. Yang and C. C. Enz, "A Low-Distortion BiCMOS 7th-Order Bessel Filter Operating at 2.5V Supply," *IEEE Journal of Solid-State Circuits*, vol. SC-31, pp. 321-330, March 1996.
[62] W. A. Serdijn, M. Broest, J. Mulder, A. C. van der Woerd, A. H. M. van Roermund, "A Low-Voltage Ultra-Low-Power Translinear Integrator for Audio Filter Applications," *IEEE Journal of Solid-State Circuits*, vol. SC-32, no. 4, pp. 577-581, April 1997.
[63] M. Koyama, T. Arai, H. Tanimoto, and Y. Yoshida, "A 2.5V Active Low-Pass Filter Using All-n-p-n Gilbert Cells with a 1-Vp-p Linear Input Range," *IEEE Journal of Solid-State Circuits*, vol. SC-28, pp. 1246-1252, Dec. 1993.
[64] A. Kaiser, "A Micropower CMOS Continous-Time Low-Pass Filter," *IEEE Journal of Solid-State Circuits*, vol. SC-24, no. 3, pp. 736-743, June 1989.
[65] H. Tanimoto, M. Koyama and Y. Yoshida, "Realization of 1V Active Filter Using a Linearization Technique Employing Plurality of Emitter-Coupled Pairs," *IEEE Journal of Solid-State Circuits*, vol. SC-26, pp. 937-945, July 1991.

1V switched-capacitor filters

A. Baschirotto, R. Castello

Dipartimento di Elettronica - Universita' di Pavia
Via Ferrata, 1 - 27100 Pavia - Italia
Tel. +39 - 382 - 505.22
FAX +39 - 382 - 422.583
E-mail: a.baschirotto@ele.unipv.it

1. Introduction

In the last years the interest toward low-power low-voltage integrated systems has consistently grown due to the increasing importance of portable equipment and to the reduction of the supply voltage of modern scaled-down technology IC [1, 2]. In such systems, a number of analog functions can be efficiently implemented with SC technique which ensures performance accuracy and large dynamic range. At low supply voltage these features have to be maintained. In particular the large output swing (rail-to-rail) must be targeted to achieve a sufficient dynamic range. This is because with the supply voltage reduction the noise level remains to first order constant, while the signal swing decreases more than linearly.

From the technological point of view CMOS technology seems to be more feasible than BiCMOS, for the following two reasons. First, scaled-down CMOS technologies feature threshold voltages of about 0.5V-0.6V which makes the MOS devices able to operate with a bias voltage (V_{GS}) comparable to that of bipolar devices (V_{BE}=0.7V). As a consequence, bipolar devices are no more clearly advantageous in term of voltage compliance. Second, SC circuits are usually included in mixed-signal systems in which the digital part is much larger than the analog one. This forces, for cost reduction, the use of a standard CMOS process. This second aspect does not usually allow the use of low threshold devices, which requires costly additional steps in the CMOS process [3, 4]. From these considerations, only standard CMOS technology will be considered in the following.

Fig. 1 shows the classical structure of an SC integrator preceded and followed by a similar stage. It is composed by an input branch, consisting of four switches and a capacitor, and by an integrator, made up of an opamp with a feedback capacitor.

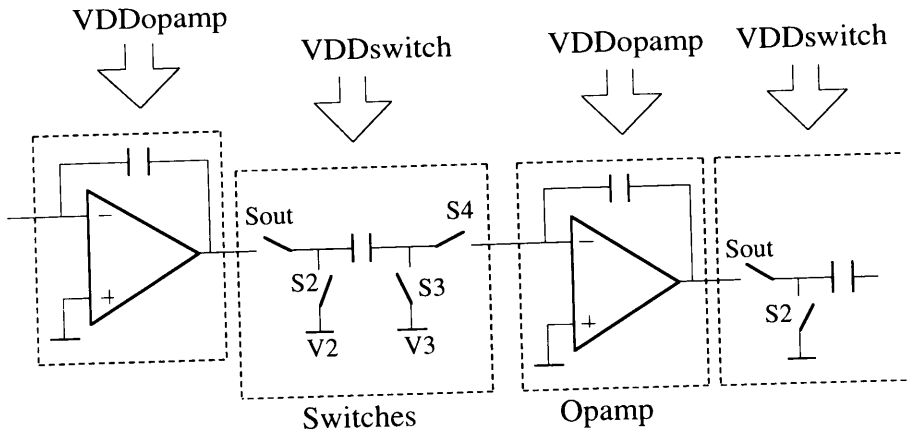


Fig. 1 - Typical SC integrator

Capacitor properties are not strongly affected by supply voltage reduction. On the other hand, at low supply it is difficult to properly operate the MOS switches and the opamp.

With the supply voltage reduction the MOS switches overdrive voltage is lowered inhibiting proper operation of classical transmission gate (complementary switches). The switch conductance for different input voltages changes depending on the supply voltage V_{DD} as shown in Fig. 2.a (V_{DD}=5V) and in Fig. 2.b (1V). For V_{DD}=1V, there is a critical voltage region centered around $V_{DD}/2$ for which both switches are not conducting. In a SC circuits to achieve rail-to-rail swing, the output of the opamp must necessary cross this critical region. Thus the switches connected to the opamp output node will not properly operate.

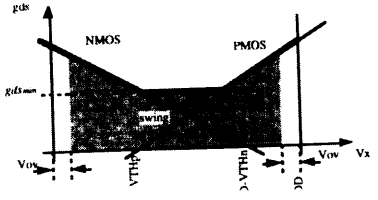


Fig. 2.a - Switch conductance with VDD=5V

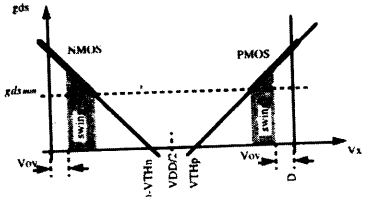


Fig. 2.b - Switch conductance with VDD=1V

On the other hand, opamp operation can be achieved with proper design using a supply voltage as low as $V_{TH}+2\cdot V_{ov}$, as described in the following. However such an opamp can only be used with some modifications at system level.

As shown in Fig. 1, switches and opamp sections may use different supply voltage. Three different approaches have been presented in literature for operating at low-supply voltage: the on-chip supply voltage multiplier, the on-chip clock multiplier and the switched-opamp technique. Only the last approach uses a single low-voltage supply for all the blocks.

This presentation is organized as follows. In Sec. 2 the different approaches at system level are described and discussed. Sec. 3 deals with circuit solutions for the basic building blocks. Experimental results from a design example are proposed in Sec. 4, while in Sec. 5 open problems and future developments are indicated.

2. System level considerations

In this section the three design approaches are presented and compared. They are different with respect to the supply used for the switch section and for the opamp section (as shown in Fig. 1). Depending on this choice, switches and opamp operation are guaranteed in different ways, as described in the following.

Each different approach suffers from particular problems which limits its applications. In all cases, however, it appears mandatory for the SC filter to be able to achieve rail-to-rail output swing. This forces to set the opamp output dc-voltage to $V_{DD}/2$.

On-chip supply voltage multiplier

If the SC designer wants to re-use all his know-how, the only possible design approach is to generate on-chip an auxiliary supply V_{DDmult} to be used to power the complete SC filter. In this way the SC filter is designed using the already available analog cells for opamp and switch, able "as-they-are" to operate from the multiplied supplied voltage (i.e. with $V_{DDswitch}=V_{DDopamp}=V_{DDmult}$).

For the switch operation, the higher supply voltage allows to use complementary transmission gates which ensure a certain minimum conductance for all possible voltages they are connected to (between 0 and V_{DD}). On the other hand, the opamp powered with a higher supply voltage can manage a larger signal swing, with a consequential larger dynamic range. Using this approach Nicollini et al. [5] have achieved a -80dB THD with a 4Vpp signal amplitude using a 1.5V supply. In addition a dynamic range larger than 80dB was also obtained.

This approach suffers from some critical limitations. First of all, the stressing of the technology. In fact in a scaled-down technology the

maximum acceptable electric field between gate and channel (for gate oxide break-down) and between drain and source (for hot electrons damage) must be reduced. This puts an absolute limit to the value of the multiplied supply voltage. In addition the need to supply a dc-current from the multiplied supply forces to use an external capacitor, which is an additional cost and could not be possible for other system considerations. Finally the conversion efficiency of the charge-pump cannot be 100% and this could limit the application of this approach in battery operated portable systems.

For these reasons this approach appears the least feasible for future applications and it will not be discussed any further.

On-chip clock voltage multiplier

A more feasible alternative to operate SC filters at 1V is the use of on-chip clock multiplier [6-12] to drive the switches, while the opamp operates from the low-supply voltage. In this case the voltage multiplier does not supply any dc-current but drives only the capacitive load due to the switch gates. Therefore no external capacitor is needed and the SC filter can be fully integrated. In this case the switches operate as in standard SC design while the opamp has to be properly designed to operate from the reduced supply voltage. In particular it becomes necessary to use V_{in_DC}=0 (as it will explained later), while maintaining V_{out_DC}=V_{DD}/2 for rail-to-rail output swing, which forces to implement a voltage level shift. This can be efficiently done using SC technique, taking advantage of the full functionality of the switches at any input voltage due to the multiplied clock supply. With reference to Fig. 1, choosing V_2=V_{out_DC} and V_3=V_{in_DC} gives the proper DC voltage at the input and at the output of the amplifier.

This solution, like the previous one, must not exceed the limit of the technology associated to the gate oxide break-down. A further potential problem arising from using V_{in_DC} equal to ground is the possible charge loss during the transient. This is explained with the help of Fig. 3.a.

The inverting opamp input node is always in parallel with a pn junction due to the NMOS switch connected to it (switch S4 in Fig. 1). Such a pn junction is due to the bulk-to-source (drain) of the MOS transistor as explained in Fig. 3.b, and it is normally reverse biased.

A negative charge transfer and the clock feedthrough in conjunction with the finite speed of response of the opamp cause negative voltage spikes at the inverting input (see Fig. 3.a). Such spikes, if sufficiently large, can forward bias the bulk diode, creating charge leakage, and a consequent voltage error. As an example a 500mV spike of 5ns

duration results in a 0.5mV voltage error on a 1pF integration capacitor at 100 °C and assuming a minimum size switch.

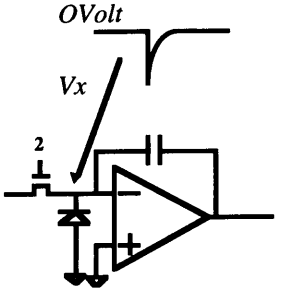


Fig. 3.a - Excursion of node X during operation

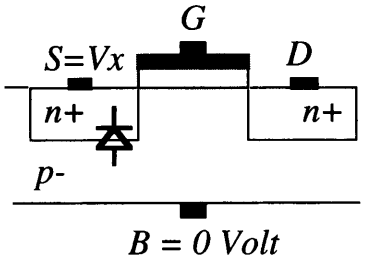


Fig. 3.b - NMOS switch structure

The size of the spike is proportional to the signal level (which is limited by the opamp supply) and to the clock amplitude (i.e. the clock supply). As a consequence this problem is difficult, if not impossible, to manage for a classical 5V SC filter opamp supply voltage, while its importance becomes progressively lower as the opamp and clock supply voltage are reduced. Although the spike due to the charge transfer tends to become negligible for low opamp supply voltage, the clock feedthrough effect remains important due to the large clock signal amplitude which use clock voltage multiplication. For a supply voltage larger than about 1.5V and a clock supply larger than about 3V, an error sufficiently large to be not acceptable can occur, if precautions are not taken.

Nonetheless this approach is largely used also because it allows the filter to operate at higher sampling frequency. Using this approach reached a 20Ms/s sampling frequency in a pipeline A/D converter has been reached [11]. In an improved solution related to this approach all the switches are driven with a fixed V_{ov} which ensures constant switch conductance and reduces also signal-dependent distortion. This however requires a specific charge-pump for each switch, increasing area, power consumption and noise injection.

Switched-opamp technique

In order to avoid any kind of voltage multiplier, a third approach called 'Switched-OpAmp' (SO) was proposed [12]. Three basic considerations leads to the SO technique.

1- The best condition for the switches driven with a low supply voltage is to be connected either to ground or to V_{DD}. Since switch S4

in Fig. 1 is connected to virtual ground, it follows that the opamp input dc-voltage has to be either ground or V_{DD}.

2- The opamp input node has to be bias at ground or V_{DD} also to minimize the required opamp supply voltage. On the other hand, the opamp dc output voltage has to be at $V_{DD}/2$ in order to have rail-to-rail output swing.

3- The switch Sout connected to the opamp output node cannot operate properly as the supply voltage is reduced below some minimum value, as explained before. Its functionality has then to be guaranteed in some different way.

Using the 'Switched-opamp' approach all the previous points can be satisfied at the same time. The proposed solution uses an opamp which can operate in a tri-state mode. In this way the critical output switch Sout is no more necessary and it can be eliminated, moving the critical problem to the opamp design. The resulting SO SC integrator is shown in Fig. 4 [14-16] where the function of the eliminated critical switch Sout is implemented turning on and off the op–amp through Sa, which is connected to ground.

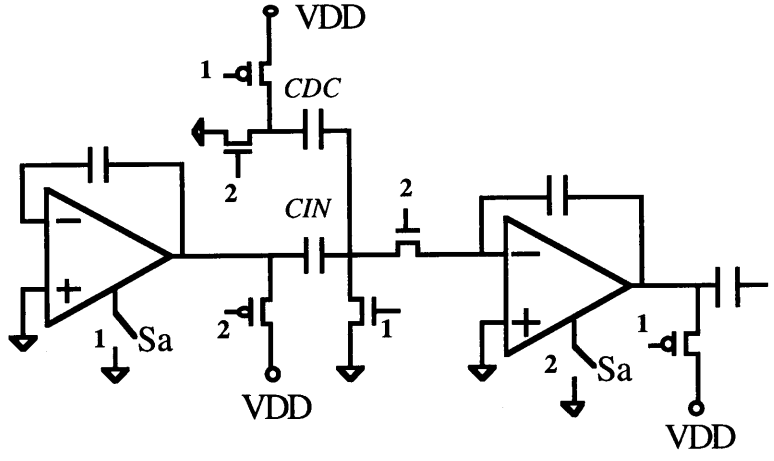


Fig. 4 - Switched-Opamp SC integrator

If the input dc-voltage is set to ground, all the switches are connected to ground (and realized with a single NMOS device) or to V_{DD} (and realized with a PMOS device). Therefore all the switches are driven with the maximum overdrive, given by V_{DD}-V_{TH}. This means that the minimum supply voltage required for proper operation of the switches is given by:

$$V_{DDmin} = V_{TH} + Vov \qquad (1)$$

where V_{TH} is the larger of the two threshold voltages (N–type and P–type). Notice that this is the same as the minimum supply voltage required for the operation of digital CMOS circuits.

A key feature of the scheme of Fig. 4 is the presence of switched-capacitor C_{DC}. Capacitor C_{DC} gives a fixed charge injection into virtual ground, producing a voltage level shift between V_{in_DC} and V_{out_DC}. The amount of the level shift can be evaluated writing the charge balance at the opamp input node, assuming no signal is present. This is given by:

$$- C_{IN} \cdot V_{out\_DC} - C_{DC}\, V_{DD} = C_{IN} \cdot (V_{in\_DC} - V_{DD}) + C_{DC} \cdot V_{in\_DC} \qquad (2)$$

Since V_{in_DC} is set to ground, the opamp output dc-voltage becomes:

$$V_{out\_DC} = V_{DD} \cdot (C_{IN} / C_{DC}) \qquad (3)$$

Using $C_{DC}=C_{IN}/2$ gives $V_{out\_DC} = V_{DD}/2$.

Capacitor C_{DC} influences the bias condition, while has no effect on the signal transfer function which is given by:

$$H(z) = (C_{IN} / C_{DC}) \cdot z^{-1/2} / (1 - z^{-1}) \qquad (4)$$

The SO approach suffers from the following problems.

Only the non-inverting and delayed integrator has been proposed up to now. Thus, to build high-order filters it is necessary to implement a sign change to close the basic two-integrators loop. This remains an open problem for the single ended structure (a part from using an extra inverting stage [13]). In addition, any error in the C_{DC} size results in an extra offset at the output node which limits the output swing. Furthermore all the noise present on V_{DD} is injected into the signal path.

All of these problems can be greatly alleviated using a fully differential structure. A fully-differential architecture provides both signal polarities at each node, useful to build high order structures without any extra elements (e.g. inverting stage). In addition any disturbance (offset or noise) injected by C_{DC} results in a common mode signal which is largely rejected by the fully-differential operation. Fully–differential structures, however, need a Common Mode FeedBack (CMFB) circuit, which could become critical at low supply voltage as discussed in the section on circuit design aspects.

Even if fully-differential structures solve the previous critical aspects of a SOA filter, there are other problems which still need a solution.

1- SOA structure operates with an opamp which is turned on and off. Its turn on time becomes the main limitation in the possible maximum sampling frequency.

2- The possibility of charge loss at the virtual ground node due to spikes still exists. The SOA structure, however, suffers less than Approach no. 2 since the clock amplitude is lower.

3- In a SOA structure the output signal is available only during one clock phase, while during the other clock phase the output is set to zero (return-to-zero). If the output signal is read as a continuous-time waveform the return to zero has two effects: a loss of 6dB in the transfer function, and an increased distortion due to the large output steps. On the other hand when the SOA integrator is used in front of a sampled-data system (like an ADC) the output signal is sampled only when it is valid and both the above problems are cancelled.

4- The switch connecting the sampling capacitor with the output of the previous stage still cannot be properly turned-on and off in xxx case. Only ac-coupling through a capacitor appears a good solution, until now.

A comparison between the three different design approaches is given in Table I.

	Supply multiplier	Clock multiplier	Switched opamp
$V_{DDswitch}$	V_{DDmult}	V_{DDmult}	V_{DD}
$V_{DDopamp}$	V_{DDmult}	V_{DD}	V_{DD}
New opamp design	No	No/Yes	Yes
New switch design	No	No	Yes
Output swing	+	-	-
Voltage spike charge loss	+	- -	-
Gate Break-down (V_{GS} limit)	-	-	+
Hot electron (V_{DS} limit)	-	+	+
Sampling frequency limitation	+	+	-
Power consumption	--	-	+
External component	Yes	No	No
Continuous waveform			
Gain loss	1	1	1/2
Return-to-zero distortion	+	+	- -
Input coupling	+	+	- -

3. Circuit level considerations

In this section the design of the basic blocks necessary to build low-voltage SC filters is discussed. We refer to the clock multiplier and to the SO approach, since for the approach no. 1 the cells can be designed in a standard way as if a large V_{DD} were used.

Opamp design

In scaled-down technologies the MOS transistor output impedance decreases and, as a consequence, also the achievable gain-per-stage decreases. Furthermore at low supply, stacked configuration (like cascode) must be avoided. To achieve a sufficiently large opamp gain it is then necessary to adopt multistage structures which, however, for stability reason, tend to have a relatively small bandwidth as compared with single stage structures. In addition multistage topologies suffers from increased turn-on time when used in conjunction with the SOA approach.

Multistage structures allow to separately optimize the different stages. In this way it is possible to operate with V_{in_DC} set to ground and with V_{out_DC} set to $V_{DD}/2$.

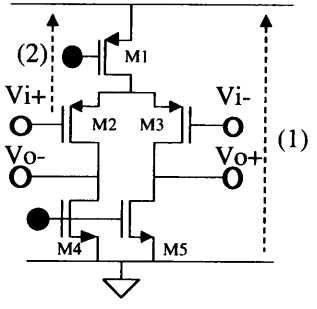


Fig. 5 - Differential input stage

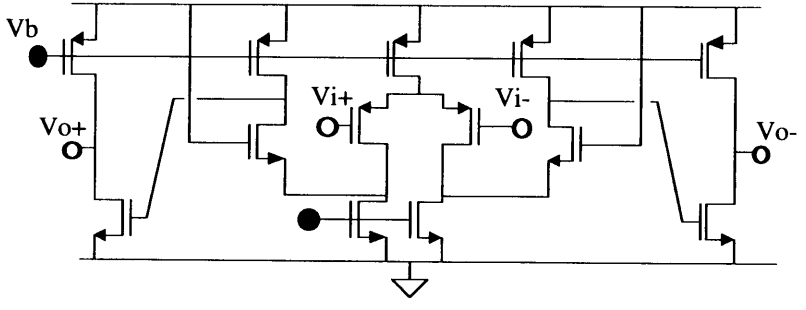


Fig. 6 - Complete two-stage amplifier

For the input stage, the differential structure appears the most feasible in order to reduce noise coupling and offset. Fig. 5 shows the simplest differential input stage. The case of PMOS input transistors with NMOS load devices has been used as an example. This corresponds to use NMOS input switches. The minimum supply voltage for proper operation (V_{DDmin}) is given by:

$$V_{DDmin} = \max \{3\, V_{ov} + V_{outpp},\ V_{in\_DC} + V_{TH\_P} + 2\, V_{ov}\} \quad (5)$$

where V_{outpp} is the peak-to-peak voltage swing at the output. The first condition must be satisfied to guarantee that the three stacked devices M3, M1, M5 operate in the saturation region (i.e. $V_{DS}>V_{ov}$). The second condition derives from the sum of the V_{GS} drop across the input device plus the V_{DS} of the PMOS current source. If V_{TH} is larger than $V_{ov}+V_{outpp}$ and choosing $V_{in\_DC}=0$, the theoretical minimum supply voltage V_{DDmin} is obtained i.e.:

$$V_{DDmin} = V_{TH} + 2\, V_{ov} \quad (6)$$

A complete differential amplifier that satisfied both above conditions is shown in Fig. 6. The folding structure ensures that the V_D of M4, M5 is biased slightly higher than one V_{ov} above ground. This is sufficient to maintain all the devices (with the possible exception of the cascode device itself) in the saturation region at all time. This is because both the source of the cascode device and the gate of the NMOS device in the output stage experiences a very small voltage swing. For this circuit the V_{DDmin} is still given by eq. (6), where V_{TH} is the largest between the NMOS and the PMOS threshold voltages.

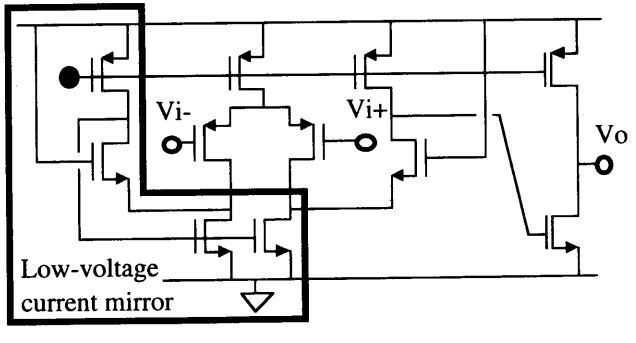


Fig. 7 - Single-ended opamp with low-voltage current mirror

The structure of Fig. 6 corresponds to a fully differential amplifier. If a single–ended output is required a current mirror must be implemented in some way. Classical current mirror topologies use a

diode connected MOS in the signal path. This increases the minimum supply voltage. A possible circuit capable to operate with the V_{DDmin} given by eq. (6) is shown in Fig. 7, where a low-voltage current mirror is used (indicated with bold line).

In alternative to the above class-A input stage, a class-AB input stage which is still able to operate with the V_{DDmin} given in eq. (6) has been proposed in literature [18]. This stage can be used when a large capacitive load is present and the power consumption must be reduced. However, as it is usually the case in a class AB stage, its structure is more complicated, resulting in an increased noise and offset.

If the opamp has to be used in a filter based on the SO concept, an other aspect in addition to the above ones has to be studied. This is the minimization of the opamp turn-on-time. For the solution proposed in literature up to now, this gives the limit to the maximum achievable sampling frequency. The turn-on time is limited by two factors: the turn-on time of current sources and the time required for charging the capacitor associated with all the node to their correct value. In particular the longest delay comes from re-charging the compensation capacitance.

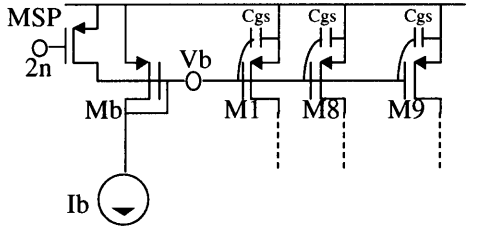


Fig. 8.a - Current source turn-off scheme

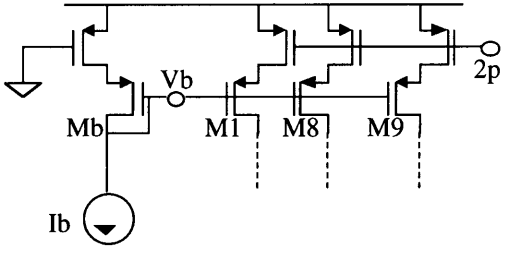


Fig. 8.b - Improved current source turn-off scheme

In the early SO realizations [12], the opamp was fully turned-off by shorting to V_{DD} the gate of all current source devices, as shown in Fig.

8.a. This approach increases the turn-on time since all the capacitances of the gates connected to Vb must be charged and discharged with the constant bias current Ib. This delay can be alleviated by turning-off the current source through MOS series switches, as shown in Fig. 8.b. In this case, the gate capacitance is kept charged.

Another technique useful to reduce the turn-on time is turning-off only part of the opamp, maintaining the rest of the structure active and ready to operate at the turn-on instant. In the opamp shown in Fig. 9 [15-16] only the second stage is turned-off, while the first stage is kept active even during the off phase.

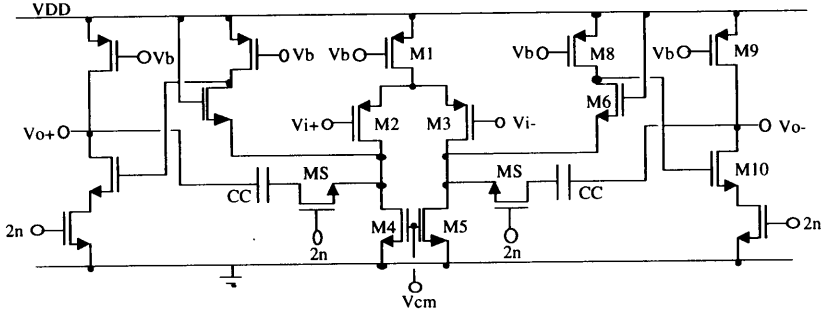


Fig. 9 - Switched-opamp structure

Technology	CMOS
Minimum gate size	0.5μm
Power supply	1V
Power consumption	80μW
Open loop gain	>75dB
Unity gain bandwidth	30MHz
Switching time	< 100ns

Table II - Opamp performance

The other major source of delay is associated to the re-charging of the compensation capacitance. In fact when the output voltage is tied to V_{DD} during the off-phase, the capacitance changes its stored charge and it must be re-charged during the next on-phase. This problem can be reduced by maintaining C_C charged during the off phase. One way to do this is shown in Fig. 9 [15-16]. Here C_C is connected to the source of M6, via switch MS. When the output stage is turned off, MS disconnects C_C from M6. This operation conserve the charge on C_C. MS can operate with the same supply voltage required by the opamp

(V_{DDmin}) by placing the source of M6 one V_{ov} voltage higher than ground.

CMFB design

For a fully differential topology an additional CMFB circuit is required to control the output dc voltage. This can be done using a continuous–time or a sample–domain (dynamic) approach.

For a continuous–time solution, the key problem is due to the fact that the inputs of the CMFB circuit must be connected to the opamp output nodes which are located around $V_{DD}/2$. For low supply voltage $V_{DD}/2$ is lower than V_{TH}. No MOS gate can therefore be directly connected to the opamp output node. Fig. 10.a shows a circuits which uses a passive level shift to circumvent this limitation. In this case a trade–off exists between the amount of signal present at the CMFB circuit input and the amount of level shift. In addition the resistive level shift decreases the gain.

On the other hand, for a dynamic CMFB circuit the key problem is the turning on and off of the switches. This could be easily done using a voltage multiplier and for this case no further discussion is necessary. On the other hand for the SO approach, a possible CMFB circuit is shown in Fig. 10.b. This implementation uses a single-ended switched-opamp integrator.

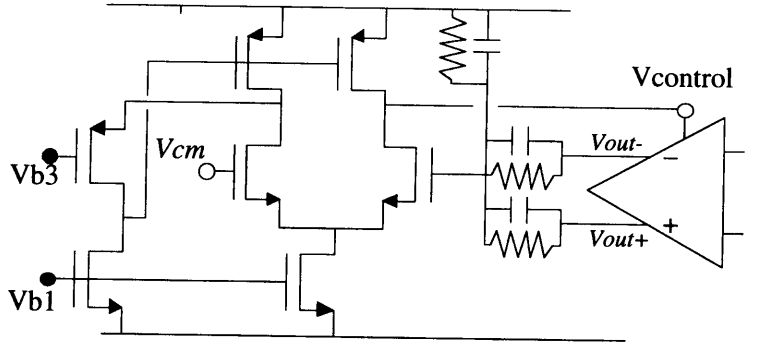


Fig. 10.a - Low-voltage continuous-time CMFB circuit

The opamp used in the integrator has a PMOS input stage with an input common-mode voltage equal to ground. Furthermore all the switches are connected to ground or to V_{DD}. The CMFB steady state occurs when the following condition is satisfied:

$$C_P\cdot(V_{DD}-V_{out+})+C_M\cdot(V_{DD}-V_{out-})+C_{CM}\cdot(0-V_{DD}) = 0 \quad (8)$$

With $C_{CM}=C_P=C_M$ the opamp common mode output voltage is set to $V_{DD}/2$.

The CMFB integrator adds a pole in the CM loop. To ensure CM loop stability, continuous-time feed-forward capacitors C_{FF} are included in the CM loop to create a zero. Capacitors C_{FF} are disconnected during the reset phase in this way preserving their charge.

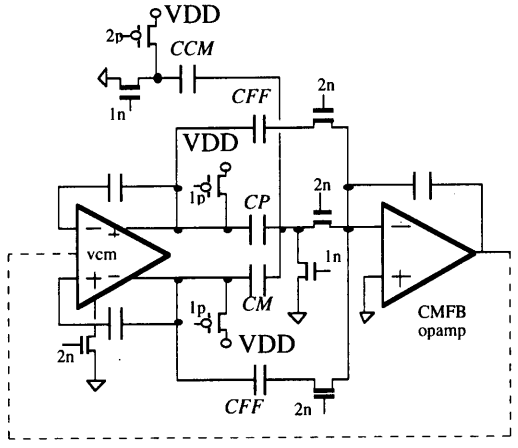


Fig. 10.b - Low-voltage SOA sampled-data CMFB circuit

Reduction of the charge loss due to voltage spikes

As explained before setting V_{in_DC} to ground can create a charge loss during the turn on transient. At the turn-on instant (while the opamp is not active, due to its finite bandwidth) two voltage steps are applied to the injecting capacitors C_{DC} and C_{IN}, as shown in Fig. 12.b. Capacitor C_{DC} is driven with a negative step (it is switched from V_{DD} to ground) and tends to decrease the voltage Vx, while capacitor C_{IN} is driven with a positive step (it switched from V_{out} to V_{DD}) and tends to increase voltage Vx. The negative spike can then be reduced ensuring that the positive charge contributed by C_{IN} is injected before the negative charge contributed by C_{DC}. This can be done clocking C_{DC} with a delayed phase (2D), as shown in Fig. 12.a.

On the other hand, assuming that the two capacitors are connected simultaneously, the voltage time variation of Vx can be described in the time domain by the equation:

$$V_x(t) = \frac{2\,R_{IN} - R_{DC}}{2\,(R_{DC} + R_{IN})} \cdot V_{DD} \cdot \exp\left(-\frac{R_{DC} + R_{IN}}{3\,R_{DC}\,R_{IN}\,C_{DC}}\,t\right) \quad (9)$$

where R_{IN} and R_{DC} are the resistance of the switches connected to C_{IN} and C_{DC} respectively. Therefore choosing the size of the switches in such a way that $2 \cdot R_{IN} > R_{DC}$, V_x can be made to move in the positive direction . Notice that the use of large size for these switches is not critical for the charge injection because they are not connected to the virtual ground.

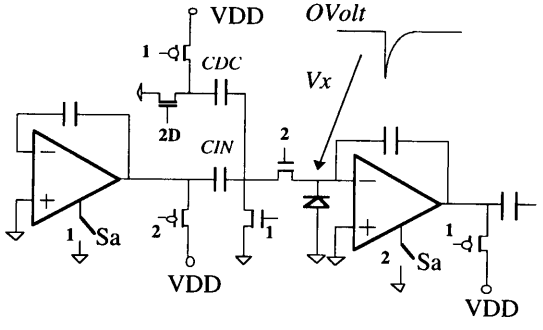


Fig. 11 - Switched-opamp integrator

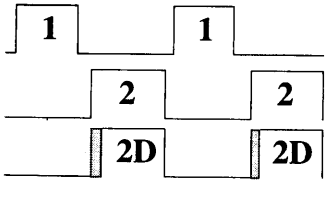


Fig. 12.a - Proposed clock phase timing

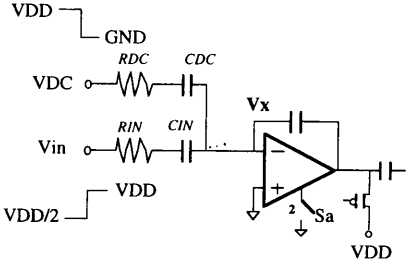


Fig. 12.b - Continuous time analysis of the integrator after turn on edge

4. Design example

As a design example, a bandpass switched-opamp filter is reported. The bandpass response has been chosen since it avoids the problem associated with the input switch. In fact this is still an open problem for the SOA approach. The bandpass biquad cell has been realized using a 0.5μm CMOS technology (V_{THP}=0.7V, V_{THN}=0.65V). The filter center frequency is one fourth of the sampling frequency with a

quality factor Q=7. The filter structure is shown in Fig. 13. It is made up of two integrators of the kind shown Fig. 4, operating during opposite clock phases. The chip area is about 0.15 mm^2. A unit capacitance of 0.25pF has been used.

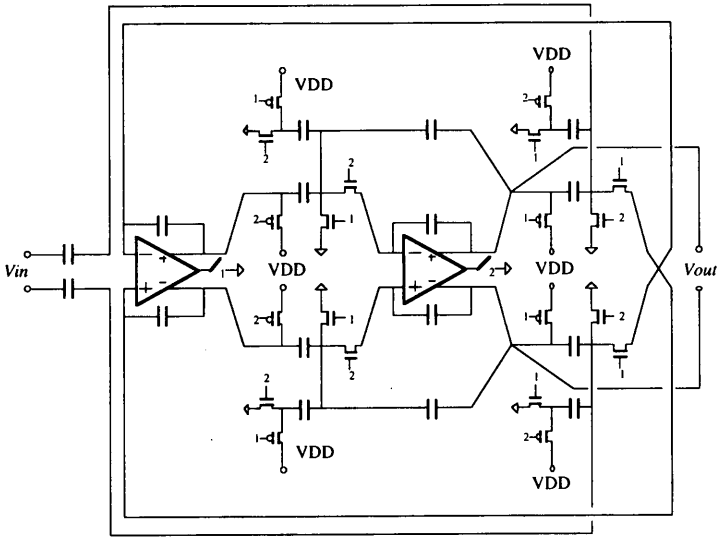


Fig. 13 - Prototype bandpass filter architecture

The prototype has been fully characterized with V_{DD}=1V, which corresponds to V_{TH}+2·Vov (using V_{ov}≈150mV). In this condition, the filter consumes about 160µW. Fig. 14 shows the filter frequency response at different clock frequency from 1.8MHz to 4.2MHz. Up to 3.4MHz a Q peaking of less than 2dB results. For higher sampling frequencies, this effect becomes more pronounced due to non-complete opamp settling (limited in particular by the opamp turn-on time). Using f_s=1.8MHz, the total output noise is about 800µV_{rms}. The immunity to the noise on the supply has been measured in terms of power supply rejection ratio (PSRR) in the passband, and in terms of power supply rejection (PSR) up to 2·f_s. The PSRR in the passband (300kHz-600kHz) is shown in Fig. 15 and is better than -46dB. In addition, for frequency up to 2·f_s the PSR always stays lower than -45dB.

Fig. 16.a-b shows the single-ended and the differential output signals, respectively. From the single-ended waveform it is evident the switched-opamp mechanism which is based on the fact that during one clock phase the output signal is tight to the power supply. On the other

hand it can be see that the signal is settling even for a signal amplitude as large as 1.6V_{pp}.

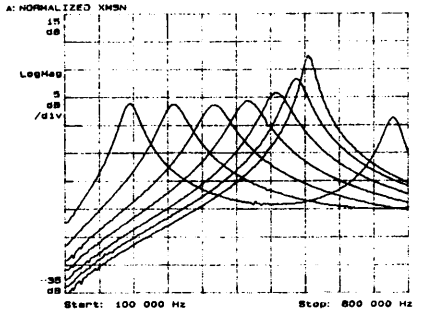


Fig. 14 - Frequency response (VDD=1V)

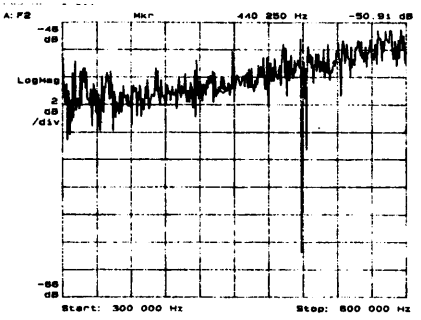


Fig. 15 - PSSR measurement

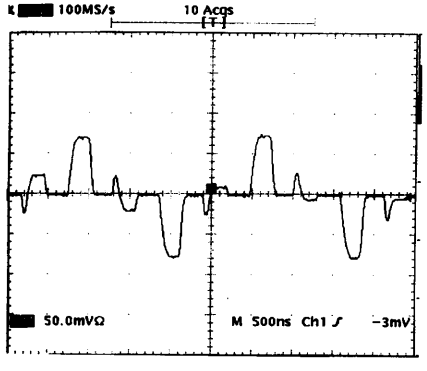


Fig. 16.a - Output single-ended waveform with 1.6Vpp differential input signal

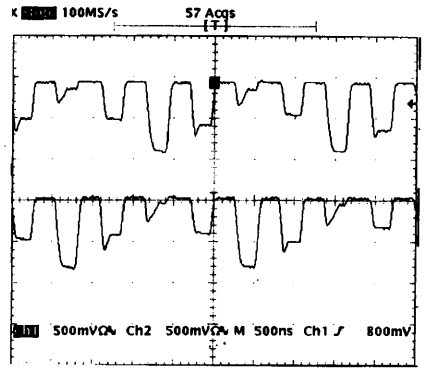


Fig. 16.b - Output differential waveform with 1.6Vpp differential input signal

A 3% intermodulation (IM) is measured for two input signals of 500mV_{pp} each, i.e. a total input signal of 1V_{pp} differential. On the other hand the 1% IM corresponds to a total input signal of 800mV_{pp}, as shown in Fig. 17. The dynamic range for 3% IM is about 50dB. The THD is characterised injecting a signal with a frequency f_{in} that falls in the passband. This results in the folding of the third harmonic at f_s-3·f_{in}. The 1% THD corresponds to a 725mV_{pp} input signal and the 3% THD to a 1.1V_{pp} input signal, as shown in Fig. 18.

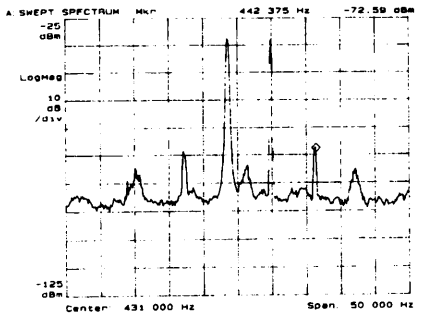


Fig. 17 - 1% IM measurement

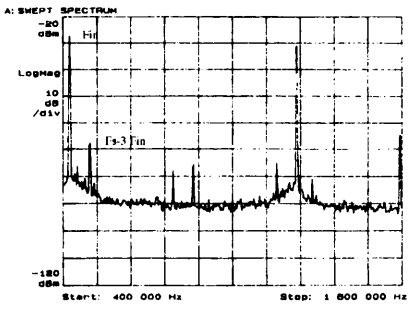


Fig. 18 - 1% THD measurement

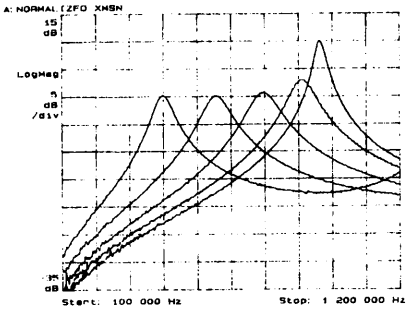


Fig. 19 - Frequency response with VDD = 0.9V

Technology	0.5µm CMOS
Supply voltage	1 V
Power consumption	160 µW
Sampling frequency	1.8 MHz
Q	6.6
fo	435 kHz
Max. output swing	1.6 Vpp
Total output noise	800 µVrms
THD 1%	725 mVpp
THD 3%	1.1 Vpp
IM3 1%	800 mVpp
IM3 3%	1 Vpp
Dynamic range (IM3 3%)	50 dB
PSRR	-45 dB
Filter chip area	0.15 mm²

Table III - Filter performance

The above linearity results are measured for the overall output waveform and are worse than those achievable if the output signal is sampled during its "true" phase. In fact during the reset phase very large output steps (that could be almost equal to V_{DD}) are forced at the opamp output resulting in slew-rate and glitches. These phenomena degrade the linearity of the continuous-time waveform, but they have no effect on the sampled version of the output signal, provided that it settles within the clock phase. For the case of two 700mV$_{pp}$ signals, the output signal has been sampled and an FFT analysis has been performed. In this case a 1.8% IM results, while for the continuous-time waveform the IM is 4%.

Finally the filter is still fully functional with a supply voltage of 0.9V (i.e. leaving only 200mV for 2·Vov) for f_s up to 3.2MHz. Fig. 19 shows the frequency response for sampling frequencies between

500kHz and 3MHz with 500kHz step for this reduced supply voltage. The filter performance is summarized in Table III.

5. Open problems and future development

Even if several implementations have demonstrated the possibility of realizing SC systems operating at 1V supply, some draw-backs that limit the achievable performance of an SOA filters still exist. In particular there is room for improvements in the turn-on time of the opamp and in the sensitivity to the supply noise. Moreover the SOA technique, as it stands, can realize precise transfer function only if an input ac-copling is acceptable by the system specifications. A possible topology that circumvents this problem, extending the SOA approach to any kind of transfer function is the switched-opamp series-switch discussed next.

Active switched-opamp series switch

The main lack of the SOA technique is the implementation of the series switch to be connected at the input signal. Up to now no solution for this problem has been proposed which offers rail-to-rail signal swing .

A possible solution to this problem is to connect to the input signal a passive impedance to be connected to a some kind of virtual ground.

The conceptual scheme of a possible implementation, shown in Fig. 20, consists in a switched-buffer, implemented with a switched-opamp in inverting configuration. If V_{batt}=0V, V_X is set to ground as previously described, and V_{out_DC} is fixed to $V_{DD}/2$ only if V_{s_DC} is set to $-V_{DD}/2$. This is however not a feasible value for the input stage operation voltage. On the other hand, if $V_{batt}=V_{DD}/2$, node X acts like a virtual ground set to $V_{DD}/2$. Using $V_{s\_DC}=V_{DD}/2$ for rail-to-rail input swing of the preceding stage, V_{out_DC} is set to $V_{DD}/2$. In this way only signal current flows through R1 and R2. With R1=R2, V_o follows V_s with negative unity gain.

The actual circuit is shown in Fig. 21, where the battery V_{batt} is implemented with switched-capacitor techniques. The circuit operates as follows. During phase 1 capacitor C1 is charged to V_{DD} while capacitor C2 is discharged. During phase 2 capacitors C1 and C2 are connected in parallel. Since C1=C2, a voltage equal to $V_{DD}/2$ appears across the capacitors. In this way C1 and C2 act like a battery from the opamp input node (set to ground) to node X which is set to $V_{DD}/2$, as required. During phase 2 the feedback is active, and $V_o(2)$ follows $V_s(2)$. The value of $V_o(2)$ is sampled on Cs which is the input capacitor for the following stage in which it injects its charge during the following phase 1.

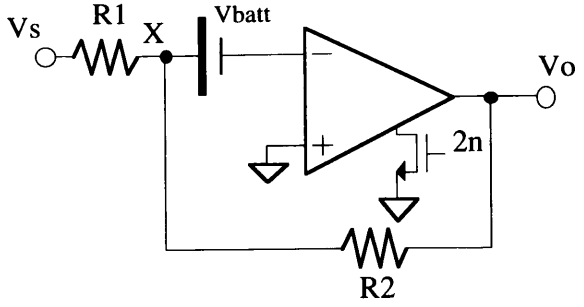


Fig. 20 - Conceptual scheme of the series switch

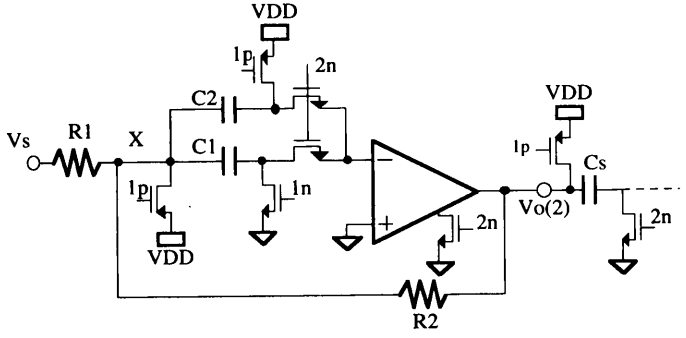


Fig. 21 - Complete switched-opamp buffer

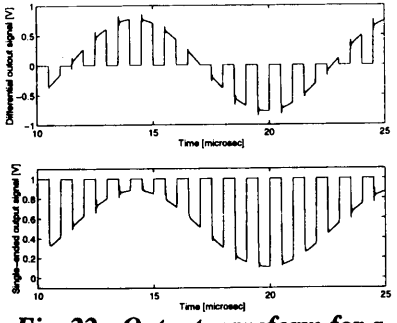


Fig. 22 - Output waveform for a 1.8Vpp input signal (upper trace: differential signal, lower trace: single-ended signal)

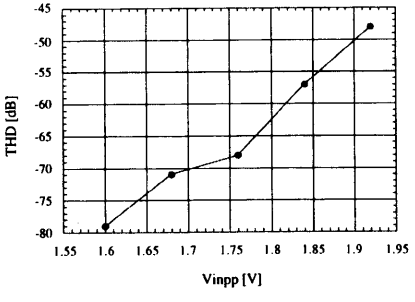


Fig. 23 - THD vs. differential signal amplitude

Using a 1V supply, the circuit has been simulated for a 87.9kHz input signal sampled at 1MHz. Fig. 22 shows the output waveform for a 1.8Vpp output signal amplitude. An FFT analysis has been performed on the output samples. Fig. 23 shows the achieved THD vs. differential input signal amplitude. A THD better than -60dB is achieve for signal up to 1.8Vpp differential signal.

Conclusions

The possibility of realizing SC circuits operated down to 1V supply in a standard CMOS technology has been discussed. The possible approach alternatives are described and compared. Between the three possibilities, the switched-opamp approach seems to be the most feasible for low-voltage and low-power systems. However the SOA approach requires the re-design of the entire SC system and still suffers from some open problems (dc-coupling, limited opamp turn-on time, CMFB implementation). In any case it seems possible that 1V SC filters achieve performance comparable to those achieved with circuits operating at higher supply voltages. This, however, may require a higher power consumption.

References

[1] A. Matsuzawa, "Low-voltage and low-power circuit design for mixed analog/digital systems in portable equipment", IEEE J. of Solid-State Circuits, Vol. SC-29, no.4, pp.470-480, April 1994.

[2] R. Castello, F. Montecchi, F. Rezzi, and A. Baschirotto, "Low-voltage analog filter", IEEE Transaction on Circuits and Systems - II - Nov. 1995 - pp. 827-840.

[3] Y. Matsuya, and J. Tamada, "1V power supply low-power consumption A/D conversion technique with swing suppression noise shaping", IEEE J. of Solid-State Circuits, pp.1524-1530, Dec. 1994.

[4] T. Adachi, A. Ishikawa, A. Barlow, K. Takasuda, "A 1.4V switched-capacitor filter", IEEE 1990 Custom Integrated Circuits Conference, pp. 8.2.1-8.2.4.

[5] G. Nicollini, A. Nagari, P. Confalonieri, and C. Crippa, "A -80dB THD, 4Vpp switched-capacitor filter for a 1.5V battery-operated systems", *IEEE J. of Solid-State Circuits*, Vol. SC-31, no.8, pp. 1214-1219, August 1996

[6] J. F. Dickson, "On-chip high-voltage generation in MNOS integrated circuits using an improved voltage multiplier technique", IEEE J. of Solid-State Circuits, Vol. SC-11, no.3, pp.374-378, June 1976.

[7] Y. Nakagone, H. Tanaka, K. Takeuchi, E.Kume, Y. Watanabe, T. Kaga, Y. Kawamoto, F. Murai, R. Izawa, D. Hisamoto, T. Kisu, T. Nishida, E. Takeda, K. Itoh, "An experimental 1.5V 64Mb DRAM" IEEE JSSC, Vol. 26, no. 4, pp. 465-472, April 1992.

[8] F. Krummenacher, H. Pinier, A. Guillaume, "Higher sampling frequency in SC circuits by on-chip clock voltage multiplier" 1983 European Solid State Circuits Conference, pp. 123-126.

[9] R. Castello, and L. Tomasini "A 1.5V high-performance switched-capacitor filters in BiCMOS technology", IEEE J. of Solid-State Circuits, Vol. SC-26, no.7, pp.930-936, July 1991.

[10] Jieh-Tsorng, Yueh-Huang Chang, Kuen-Long Chang, "1.2V CMOS Switched-Capacitor circuits", IEEE International Solid State Circuits Conference - February 1996 - pp. 388-389.

[11] T.B. Cho, and P.R. Gray, "A 10b, 20Ms/s 35mW pipeline A/D converter", IEEE J. of Solid-State Circuits, Vol. SC-30, no. 3, pp. 166-172, MARCH 1995.

[12] J. Crols, and M. Steyaert, "Switched-Opamp: an approach to realize full CMOS switched-capacitor circuits at very low power supply voltages", IEEE J. of Solid-State Circuits, Vol. SC-29, no.8, pp.936-942, August 1994.

[13] V. Peluso, M. Steyaert, and W. Sansen, "A switched-opamp 1.5V-100μW ΔΣ modulator with 12 bits dynamic range", Europ. Solid State Circ. Conf. (ESSCIRC 96) - Neuchatel, Sept. 1996, pp. 256-259.

[14] A. Baschirotto, R. Castello, and F. Montecchi, "Design strategy for low-voltage SC circuits", IEE Electronics Letters - 3rd March 1994, Vol. 30, pag. 378-379.

[15] A. Baschirotto, R. Castello, "A 1V 1.8MHz CMOS Switched-opamp SC filter with rail-to-rail output swing", IEEE Journal of Solid State Circuits - December 1997 - pag. 1979-1986

[16] A. Baschirotto, R. Castello, "A 1V 1.8MHz CMOS Switched-opamp SC filter with rail-to-rail output swing", IEEE International Solid-State Circuits Conference (ISSCC '97) - San Francisco, February 1997 - Digest of technical papers - pag. 58-59.

[17] R. Castello, A. Baschirotto, and A. Nagari, "Low voltage, large swing, switched-capacitor circuit employing switched operational amplifier", European patent (published) 94830318.5 - Japan patent (pending) 7-183434 - USA patent (pending) 326,107.

[18] V. Peluso, P. Vancorenland, A. Marques, M. Steyaert, W. Sansen, "A 900mV 40μW Switched-opamp DS modulator with 77dB dynamic range" IEEE International Solid State Circuits Conference - February 1998, pp. 68-69

[19] G. Nicollini, F. Moretti, M. Conti, "High-frequency fully-differential filter using operational amplifiers without common-mode feedback" *IEEE J. of Solid-State Circuits*, Vol. SC-24, no.3, pp.803-813, June 1989.

$1V$ $\Delta\Sigma$ A/D CONVERTERS

V. Peluso, M. Steyaert, W. Sansen

ESAT-MICAS
Katholieke Universiteit Leuven
Leuven, Belgium

ABSTRACT

A design approach for $1V$ Switched Capacitor $\Delta\Sigma$ modulators is presented. The chosen method of implementation is the Switched Opamp Technique, which is shortly revisited. Dedicated low voltage circuit blocks including a class AB OTA and comparater are treated. A rearranged $\Delta\Sigma$ modulator topology with half delay integrators is discussed. The influence of very low supply voltage on power consumption is illustrated. Measurement results of a $900mV$ $\Delta\Sigma$ modulator with $77dB$ Dynamic Range are discussed.

1 Introduction

In this era where the market for portable electronic systems such as wireless communication devices, hearing aids, etc. is continuously expanding, there is a growing need for the development of low voltage and low power circuit techniques and system building blocks. Circuits and systems using Switched Capacitor techniques (SC) in CMOS have become so widespread because they combine high quality signal processing to robustness of operation and efficiency of implementation. For these reasons it is desirable to have Switched Capacitor design techniques available for really low voltage operation. Classic Switched Capacitor techniques fail to operate for a supply lower than about $2.5V$, for the main reason that the switches cannot properly operate. It would further be desirable to do this without the use of expensive multi threshold processes that offer a low threshold device. The

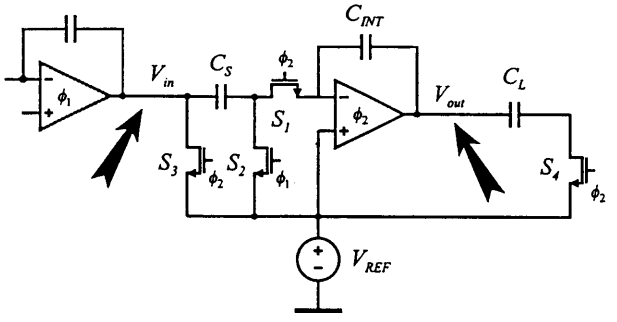


Figure 1: *The Switched Opamp integrator which is preceded by another integrator.*

prevalent solution nowadays is to use voltage multiplication to drive the switches. In this respect it is important to realize that there exists another driving force towards low voltage operation of CMOS circuits in general. It is the continuing trend to down scaling of the transistor dimensions. For that reason the future deep sub-micron MOST devices will not be able to withstand the multiplied voltage. The SIA roadmap predicts that the maximum supply voltage will drop to $1.5V$ in 2004 and to $0.9V$ in 2010. It will thus become imperative to avoid the application of voltage multipliers. The Switched Opamp Technique offers a solution to this problem.

2 The Switched Opamp Technique revisited

The Switched Opamp Technique allows to design Switched Capacitor circuits without the use of voltage multipliers or low V_T devices. [Ste 93][Cro 94] It is inspired by the observation that in a classic Switched Capacitor integrator, switches are distinguished in two classes: switches that have one terminal fixed to a reference level V_{REF} and switches that have to pass the entire signal range, which are typically found at the output of the opamp. The former can always be turned on. The latter are identified as the problem switches. The core idea of the Original Switched Opamp Technique boils down to leaving the switches at the output of the opamp out. The integrator thus obtained is shown in Fig. 1. Leaving the mentioned switch out, has

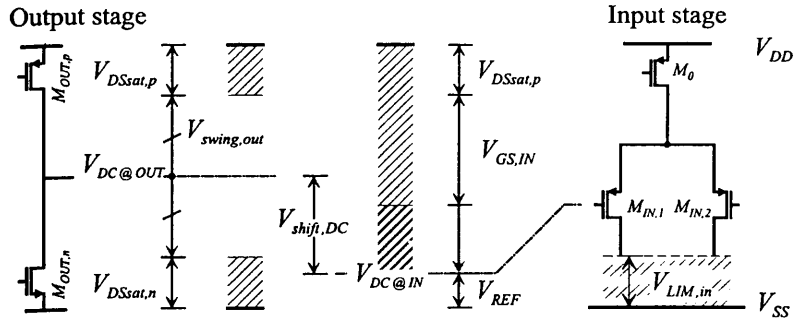


Figure 2: *The signal swings in the Modified Switched Opamp Technique.* $V_{DC@OUT} = (V_{DD} + V_{SS})/2$ *and* $V_{DC@IN} = V_{SS}$.

a few consequences. During the integration phase of the next integrator, the output of the opamp is shorted to ground. This calls for inactivation of the opamp, by switching it off, hence the nomenclature of the technique. A second consequence is that the basic Switched Opamp integrator cell now has a half delay.

The Original Switched Opamp Technique further makes the implicit choice to take the DC input level of the opamp equal to the DC output level and both at the (ground) reference level V_{REF}. As a consequence the output signal swing is less than the available swing and the switch overdrive voltage is less than maximally possible. The Modified Switched Opamp Technique enhances these aspects. [Bas 94][Bas 97a] There the DC level at the opamp output is taken in the middle of the full output swing, in practice: $(V_{DD} + V_{SS})/2$. Furthermore, the DC level at the opamp input is set at V_{SS}. This implies a reference level V_{REF} equal to V_{SS}. These choices are illustrated by figure Fig. 2. The DC offset of $V_{DD}/2$ which exists between input and output of the opamp is to be removed. This is done capacitively by C_{CM} in the Modified Switched Opamp integrator cell, which is shown in Fig. 3 in differential form (although only half of the capacitors are shown). Proper operation also requires a high reference level, taken equal to V_{DD}. Since two reference levels are now present, they are distinguished by the terminology $V_{REF,hi}$ and $V_{REF,lo}$, taken equal to V_{DD}

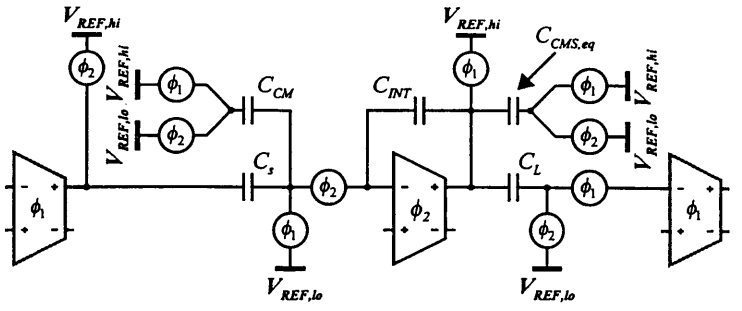


Figure 3: *The Differential Modified Switched Opamp integrator cell.*

and V_{SS} resp. The capacitor C_S and C_{INT} are the sampling and integrating capacitors and $C_{CMS,eq}$ represents the capacitive load of the common mode feedback system.

3 System level

3.1 $\Delta\Sigma$ modulator using half delay integators

A classic $\Delta\Sigma$ modulator makes use of full delay integrators. As mentioned in the above section, the basic Switched Opamp integrator cell presents only a half delay to the signal.

$$H(z) = \frac{z^{-1/2}}{\left(1 - z^{-1}\right)} \tag{1}$$

If a full delay integrator is really required, the integrator can be preceded by an (analog) half delay block. [Cro 94][Pel 97a] This can be implemented by a Switched Capacitor amplifier with unity gain, implemented with the Switched Opamp Technique. However, an extra opamp is required, causing increased power and area allocation. For the purpose of a $\Delta\Sigma$ modulator these additional amplifiers can be avoided by making use of a rearranged topology. In the architecture of Fig. 4 the half delay elements have been shifted to the feedback path. [Pel 98]. There they are in the digital domain.

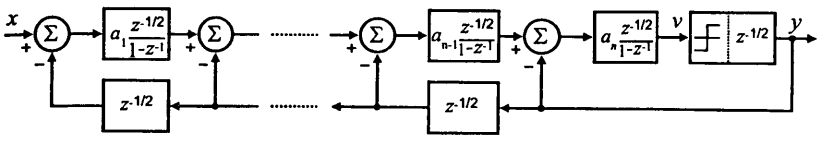


Figure 4: *n^{th} order single loop $\Delta\Sigma$ modulator topology using half delay integrators, with minimum number of independent parameters.*

They can be implemented with a half delay latch. Compared to the analog implementation, their power consumption is now negligible.

3.2 Considerations for Implementation

Before starting an implementation the question must be answered, as to which $\Delta\Sigma$ modulator architecture is the most suited for low voltage implementation. Single loop topologies, such as in Fig. 4 have the distinctive advantage over cascaded topologies that they are much less insensitive towards linear circuit non-idealities, such as GBW, and gain and switch resistance. This is a great asset, especially since it is difficult to achieve high gain at $1V$ supply voltage. Because of the reduced sensitivity is possible to use amplifiers with only one gain stage. Power consumption benefits a lot from the insensitivity for several reasons. Single gain stage amplifiers have less current branches but are also more power efficient than two stage amplifiers. Reduced GBW requirements also mean lower power drain.

4 Circuit level: $1V$ building blocks

To build an integrator with a Switched Capacitor type technique an amplifier is needed. These are usually constructed by deploying the basic circuit elements, such as the differential pair, the current mirror, the cascode transistor and the common source amplifying stage. When designing for $1V$ operation, some difficulty is encountered. Not enough headroom is left to freely make use of combinations of the mentioned circuits. It is practically impossible to use cascode transitors to increase gain. With at least four transistors stacked in $1\ V$ supply leaves no room for output signal swing. In

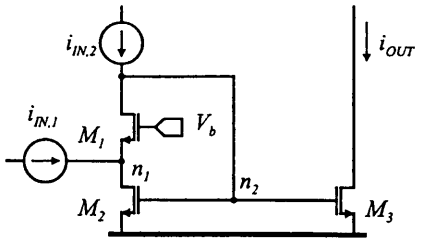


Figure 5: *Low voltage current mirror*

order to combine high gain and large output swing the designer must resort to two stage amplifiers. These require Miller compensation, which causes a much higher power consumption than a single stage amplifier which is frequency compensated by its load capacitance. Next the low voltage building blocks made use of in the implementation of [Pel 98] are discussed.

4.1 Current Mirror

The normal current mirror has one distinct disadvantage for low voltage operation. Since the input side is diode connected, the voltage drop required is V_{GS} at the input side. This makes the classic current mirror inadequate for 1 V operation. The circuit of Fig. 5 only needs a V_{DSsat} at the input node. The low voltage current mirror is actually a current feedback loop. It's input impedance is very low. For low frequencies it is

$$z_{in} \approx \frac{1}{g_{m1} A_{loop}} \approx \frac{1}{g_{m1} g_{m2} r_{oB}} \tag{2}$$

This property is explicitly made use of in the input stage which is presented next.

4.2 A differential Input Stage

In the circuit of Fig. 5 the input node can be considered virtually fixed in voltage for low frequencies. In the current mirror configuration V_b is used to fix the voltage level at the current input node. If an extra transistor is added and connected to the input of the current mirror with its source,

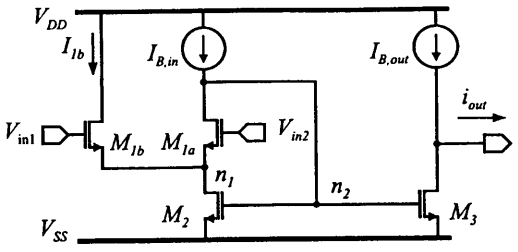


Figure 6: *A differential input structure based on the low voltage current mirror.*

that source can be considered fixed. If a voltage signal is applied at the input of the added transistor, a large current is generated and injected in the current mirror, and mirrored to the output. The resulting circuit is shown in Fig. 6. The added transistor is called M_{1b}. It is not necessary to bias M_{1a} with a constant voltage: its gate can be used as an extra input (here V_{in2}). The current feedback loop keeps the current constant through M_{1a}, which acts as a source follower. Consequently, if input signals V_{in2} and V_{in1} are applied, node n_1 undergoes a voltage excursion, equal to V_{in2}. Now v_{GS1b} equals $V_{in1} - V_{in2}$, so a larger current is generated through M_{1b}. In practice differential inputs are applied. Node n_1 therefore sees the opposite excursion such that the full differential input signal is applied between the gate and source terminals.

4.3 Class AB OTA

The differential input stage of Fig. 6 and the low voltage current mirror of Fig. 5 are the cornerstones of the class AB OTA shown in Fig. 7. It deploys two p-type input stages and two current mirrors. If $v_{in1} - v_{in2}$ is positive, a large current is generated in one input stage, and mirrored to either branch, of which one is sourcing and the other is sinking current. The other input stage is pinched off, and the current through it drops to zero. This OTA is particularly suited to operate with the Modified Switched Opamp integrator cell due to the fact that its DC input range includes V_{SS}. In combination with the Modified Switched Opamp Technique, which sets the DC level at

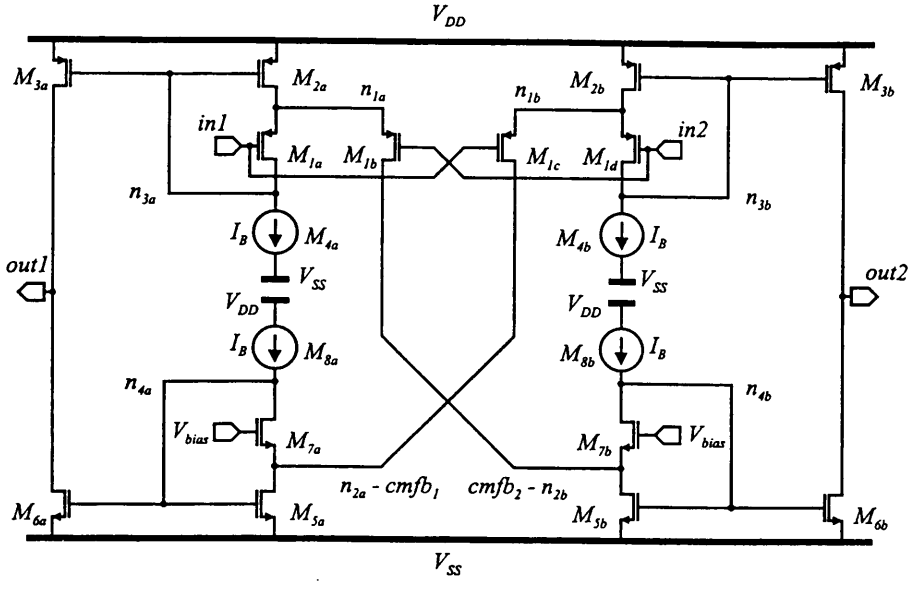


Figure 7: *The fully differential low voltage class AB OTA.*

the input of the amplifier at V_{SS}, this OTA can operate at a supply voltage even below $1V$. Also the amplifier is a single stage amplifier, which offers sufficient gain for implementation in a single loop $\Delta\Sigma$ modulator.

4.4 Common Mode Feedback

The classic Switched Capacitor common mode feedback cannot be used in the Switched Opamp methodology. It requires the output signal to be sampled with a passing type switch, connected to the output of the amplifier. These are the problematic switches at low voltage operation.

A suitable CMFB system consists of a capacitive sampler and error amplifier, as shown in Fig. 8. The capacitive sampler, shown on the left, also performs a division of the sample such that it can be applied to the input stage of an error amplifier which is biased close to strong inversion. [Cra 95]. The sampler is reset during the off phase of the OTA. The error amplifier shown on the right consists basically of a differential pair and a low voltage

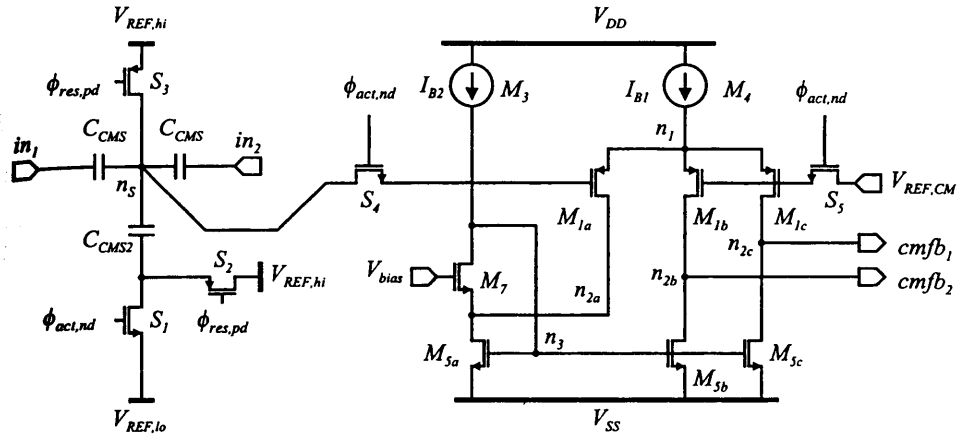


Figure 8: *An error amplifier based CMFB*

current mirror such that it has a single ended output. The feedback points in the OTA are in the signal path (n_{2a} and n_{2b} in Fig. 7). In order not to short these two nodes together when injecting the CMFB signal, the output branch of the error amplifier is split in two, such that it has two identical but isolated outputs.

4.5 Comparater

For very low supply voltage, it doesn't suffice to force an existing circuit to operate with the low supply. It requires the development of dedicated circuits. It enhances reliability to keep these as simple as possible. In the issue of low voltage comparators, the most difficult to do is the reset. A $\Delta\Sigma$ modulator doesn't need a comparator with high resolution, but it requires low hysteresis, so a good reset is important. A regenerative comparater is usually reset in its metastable state by shorting the outputs to eachother with a switch. Again, a switch that can handle a level in between the power supplies is not available in the Switched Opamp framework.

Fig. 9 shows an adequate comparater. It consists of a differential pair with input DC level at V_{SS} feeding into an n-type regenerative latch. The reset is not done to the metastable state for the reason explained. Instead

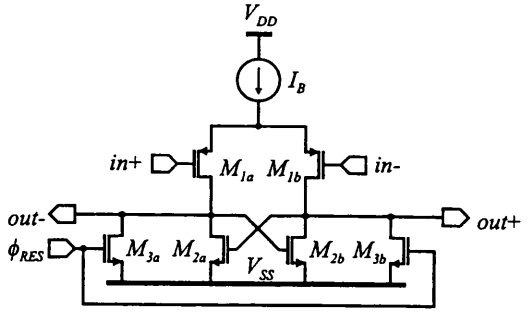


Figure 9: *The regenerative comparator with reset to to V_{SS}.*

the outputs are reset to V_{SS}. The operation of this comparator is as follows. The active phase starts when ϕ_{RES} goes low. The switches $M_{3a,b}$ become a high impedance. The output nodes rise simultaneously to the meta-stable point, from which point on the two outputs diverge. The divergence only occurs from the meta-stable point on, because only then the regeneration transistors $M_{2a,b}$ are properly biased and have their full transconductance, which determines the comparison speed. The low output is V_{SS}. The high output doesn't rise completely to V_{DD}, because the full I_B continues to flow through the current source and the input transistor of the high side. When the ϕ_{RES} goes low, the output nodes are shorted to ground and a reset occurs.

4.6 Switching Method

In the reported Switched Opamp realizations, the switching of the amplifiers is done by turning off the bias currents. [Cro 94] [Pel 97b] [Bas 97a] [Bas 97b] This is done with a switch that shorts the gate-source bias voltage with a switch. This method is quite slow turning on. Here a completely new way of switching the amplifiers is developed: it is to interrupt the current path from the power supplies to the circuit with a switch. This has two distinct advantages. First, the amplifier turn-on time is shorter than in the case of current switching. Second, all switches are driven by a buffer and therefore their turn-on and turn-off speed is under better control than in

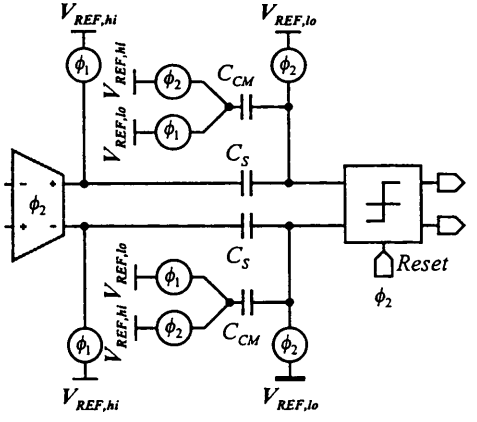


Figure 10: *The level shift network shifts the common mode level of the last integrator output signals from $V_{DD}/2$ to V_{SS}, which is the common mode input level of the comparator of Fig. 9.*

the current switching case.

Two types of supply rails are distinguished internaly: $V_{DD,BIAS}$ and $V_{SS,BIAS}$ are the internal supply nodes delivering a constant bias current and $V_{DD,DYN}$ and $V_{SS,DYN}$ are the internal supply nodes delivering a dynamic current. They are depicted in Fig. 11 Both supply nods have a separate switch. This has several advantages. The main advantage is that the bias currents are not influenced by the internal dynamic supply voltage which may be changing during the settling process. Also the total (dynamic) switch size is half of the size needed if each branch woud have a separate switch. A dynamic current flows each phase, either to one side or the other.

In principle all circuits in a Switched Opamp implementation can be switched. This is so for the differential integrating opamps but also for the CMFB amplifiers. Even input stages to comparators can be switched if desired. As a final note it is pointed out that it is good practice not to switch the bias circuits. Compared to the whole system their current drain is minimal. Even more, turning the bias off is exactly what makes the original switching method slow and is what is avoided in the new method.

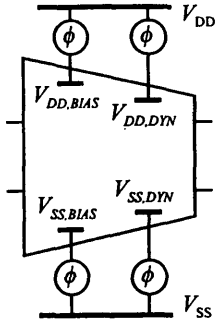


Figure 11: *The amplifiers are switched on and off by interrupting the current paths from the supplies by a switch.*

5 Power Consumption

Low supply voltage in analog circuits generally speaking causes higher power consumption. This is not so awkward. A smaller input voltage must be used. A certain dynamic range requirement puts a specification on the thermal noise level. The latter requires a minimum sampling capacitor size. The smaller the input voltage, the larger the capacitor must be. The capacitor is a load to the opamp in the integrator. The integrator must settle to an error set by the dynamic range. Due to the higher load, a higher transconductance is needed, which is set by the bias current. Hence the higher power consumption. This process is qualitatively given by the following expression [Mar 98]

$$P \propto kTDR^2BW\frac{V_{ov}}{V_{DD}} \tag{3}$$

where BW is the signal bandwidth, DR is the dynamic range and V_{ov} is the V_{GS}-V_T of the input stage transistors. This expression shows that the power is inversely proportional to the supply voltage.

The above described process is intrinsic to low voltage design and inevitable. Besides it, there are issues of more practical nature. Simply making circuits and systems work at low supply voltage causes increased power consumption. This fact is illustrated with a selection of effects below.

The settling speed of an integrator is determined by its GBW. The GBW of the Modified Switched Opamp integrator is given by

$$GBW_{int} = \frac{g_{mop}}{C_{L,eff}} \tag{4}$$

with

$$C_{L,eff} = C_S + C_{CM} + C_{FB} + \frac{C_L + C_{CMS,eq}}{\mathcal{F}_{dc}} \tag{5}$$

$$\mathcal{F}_{dc} = \frac{C_{INT}}{C_S + C_{FM} + C_{CM} + C_{INT}} \tag{6}$$

The meaning of the capacitors has previously been explained; C_{FB} is the feedback capacitor in a $\Delta\Sigma$ modulator. These expressions show that all the capacitors deployed with the purpose of enabling $1V$ operation, cause a largely increased capacitive load: there are more capacitors and also the feedback factor becomes smaller, which increases the effective load capacitance. Increased effective load capacitance causes increased power consumption.

The GBW of the error amplifier based CMFB system is given by

$$GBW = \frac{g_{meff,CM}}{2\pi C_{L,eff,CM}} \tag{7}$$

$$C_{L,eff,CM} = \frac{C_{CMS2}}{2} + \frac{C_L + C_{BP} + \frac{(C_S + C_{FB} + C_{CM})C_{INT}}{(C_S + C_{FB} + C_{CM} + C_{INT})}}{\mathcal{F}_{dc,CM}} \tag{8}$$

$$\mathcal{F}_{dc,CM} = \frac{C_{CMS}}{C_{CMS} + \frac{C_{CMS2}}{2}} = \alpha \tag{9}$$

The low supply voltage has required the deployment of an error amplifier. The CMFB amplifier has to drive a large capacitive load. To make things worse, the division factor α, required to apply the common mode sample to the input stage, constitutes a very small feedback factor which even boosts the effective CMFB load capacitance.

A third example is the low voltage current mirror of Fig. 5. Its low voltage operation has come at the cost of a an extra current branch: $I_{IN,2}$ is usually a bias current.

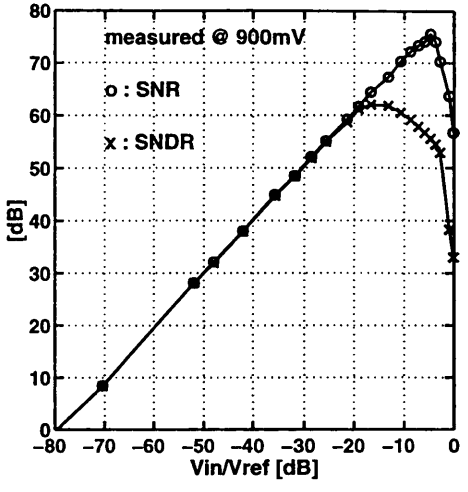


Figure 12: *The measured SNR and SNDR of the 900 mV $\Delta\Sigma$ modulator.*

On the other hand, there are a number of things that can be done that are beneficial to low power consumption. These are mainly at the system level. Already explained is the application of a $\Delta\Sigma$ modulator topology using half delay integrators in combination with the Switched Opamp Technique. Another important property of a $\Delta\Sigma$ modulator that must be made be made use of is the suppression of noise sources inside the loop. Noise injected at the internal summing nodes is suppressed so much by the large gain of the preceding integrators, that it can be neglected. This allows to scale the sampling capacitors of the integrators inside the loop down. The amplifiers have smaller load and hence require less power drain.

6 Implementation

The discussed design techniques and circuits are applied in the design of a third order single loop $\Delta\Sigma$ modulator of which the second and third integrator sampling capacitors are scaled 10 times smaller than the first. If the SO technique is used to implement a baseband processing circuit an input sampling switch cannot be avoided. It cannot be removed because it is not

Signal Bandwidth	200-16000	Hz
Sampling Frequency	1.538	MHz
max. diff. input level	500	mV_{pp}
DR	77	dB
peak SNR	76	dB
peak $SNDR$	62	dB
Supply voltage	900	mV
Power consumption	40	μW
chip core area	ca. 0.85	mm^2
Techology	0.5	μm std. CMOS

Table 1: *Performance summary of 900mV $\Delta\Sigma$ modulator.*

preceded by an opamp. The available input swing is thus much reduced. In this implementation it has been maximally exploited which has resulted in a peak-to-peak signal swing of $470mV$. The C_{CM} has also been eliminated in the first integrator by sampling the input signal to its common mode voltage. It saves on power consumption, because a large load capacitance and the noise associated with its switches are eliminated. The $\Delta\Sigma$ reference voltage is $200mV$ and it is sampled capacitively. The input sampling capacitor is $4pF$ and the smallest used is $200fF$. All transistors are biased in the moderate inversion region. The 3^{rd} order modulator is realized in a $0.5\mu m$ double poly standard analog CMOS process, with of V_{Tp} =550mV and V_{Tn} =620 mV. The total chip area excluding bonding pads is 0.85 mm^2. The 1^{st} integrator including CMFB consumes 60% of the total power. The 2^{nd} and 3^{rd}, each consume 13%, and the comparater and digital logic including clock drivers also consume 13%. The measured total core analog and digital power consumption is $40\mu W$. The sampling frequency is 1.538 MHz and the output bitstream is processed in Matlab with a combfilter for an oversampling ratio of 48. Fig. 12 shows the measured SNR and the $SNDR$. The signal bandwidth is thus $16kHz$ and the inband noise is considered from $200Hz$ on. The applied input signal frequency is 1.2 kHz. The measured peak SNR is $76dB$ and the measured input dynamic range is $77dB$. The measured peak $SNDR$ is $62dB$. The measured performance is summarized in Table 1.

7 Conclusion

Design of systems for very low supply voltage operation requires a holistic approach. The problems encountered in making systems work at $1V$ must be solved on the system-, implementation- and circuit level. On the system level a rearranged $\Delta\Sigma$ modulator topology, which uses half delay integrators allows to avoid the addition of extra amplifiers. On the implementation level dedicated circuits are required. A class AB OTA, CMFB and comparater, which can operate below $1V$ supply, are discussed. It is found that low supply voltage causes higher power consumption due to intrinsic and practical reasons. The design techniques are illustrated by means of the measurement results a $900mV$ $\Delta\Sigma$ modulator. The measured results allow to conclude that good performance is possible with Switched Opamp techniques operating at $1V$ supply.

References

[Bas 94] A. Baschirotto, R. Castello, and F. Montecchi, "Design strategy for low-voltage SC circuits", *Electronics Letters*, vol. 30, no. 5, pp. 378–379, Mar. 1994.

[Bas 97a] A. Baschirotto and R. Castello, "A 1V 1.8MHz CMOS Switched-Opamp SC filter with rail-to-rail output swing", in *Proceeding International Solid-State Circuits Conference*, Feb. 1997, pp. 58–56.

[Bas 97b] A. Baschirotto and R. Castello, "A 1V CMOS fully-differential Switched-Opamp bandpass $\Sigma\Delta$ modulator", in *Proceedings of the 23rd European Solid-State Circuits Conference*, Sept. 1997, pp. 152–155.

[Cra 95] S. Crapanzano and D. A. Johns, "2V fully-differential SC integrator in standard CMOS", *Electronics Letters*, vol. 31, no. 23, pp. 1995–1996, Nov. 1995.

[Cro 94] J. Crols and M. Steyaert, "Switched-Opamp: An approach to realize full CMOS switched-capacitor filters at very low power supply", *IEEE Journal of Solid-State Circuits*, vol. 29, no. 8, pp. 936–942, Aug. 1994.

[Mar 98] A. Marques, V. Peluso, M. Steyaert, and W. Sansen, "Bandwidth-resolution-power trade-off in $\Delta\Sigma$ analog to digital converters", Tech. Rep., K.U.Leuven, 1997.

[Pel 97a] V. Peluso, M. Steyaert, and W. Sansen, "A Switched Opamp $1.5V$-$100\mu W$ $\Delta\Sigma$ modulator with 12 bits dynamic range.", in *Proceedings of the 22nd European Solid-State Circuits Conference*, Neuchâtel Switzerland, Sept. 1996, pp. 256–259.

[Pel 97b] V. Peluso, M. Steyaert, and W. Sansen, "A 1.5V-$100\mu W$ $\Delta\Sigma$ modulator with 12-b dynamic range using the Switched-Opamp technique.", *IEEE Journal of Solid-State Circuits*, vol. 32, no. 7, pp. 943–952, July 1997.

[Pel 98] V. Peluso, P. Vancorenland, A. Marques, M. Steyaert, and W. Sansen, "A 900 mV - 40 μW $\Delta\Sigma$ modulator with 77dB dynamic range", in *Proceeding International Solid-State Circuits Conference*, SF, Feb. 1998.

[Ste 93] M. Steyaert, J. Crols, and S. Gogaert, "Switched-Opamp, a technique for realising full CMOS switched-capacitor filters at very low voltages", in *Proceedings of the 19th European Solid-State Circuits Conference*, Sept. 1993, pp. 178–181.

1-VOLT RF CIRCUIT DESIGN FOR PAGERS

Ed Callaway
Motorola Paging Products Division
Boynton Beach, Florida, U.S.A.

ABSTRACT

The motivation for 1-Volt operation of pager RF circuits is discussed. General principles covering low-voltage RF design are covered, followed by specific circuit examples of RF amplifiers, mixers, and VCOs. Finally, possible future directions of low-voltage RF design are examined.

1. INTRODUCTION

The pager RF environment can be severe. Unlike many wireless communication systems, the amplitudes of signals on the adjacent and alternate channels of a paging system are in general not controlled by protocol or international standard, and may be greater than the desired signal amplitude by more than 60 dB. The channel placement and amplitude of possible interferers is also not controlled in general, leading to intermodulation, blocking, and spurious response concerns. The desired on-channel signal may vary in electric field amplitude from 5 uV/m to 3 V/m--a range of 115 dB--and suffer from multipath effects that cause symbol-to-symbol amplitude variation of 40 dB or more.

In addition to these troubles, the pager RF engineer has other design constraints. The market often values small product size, so the antenna must be internal and extremely small, implying poor efficiency; an antenna gain of -17 dBi is not uncommon at VHF. To achieve the required sensitivity with such an antenna requires the lowest possible receiver noise figure. The market also values long battery life, as well as the use of small batteries, so

every effort is made to reduce the average current drain from the battery. Finally, the market constantly requires low product cost.

The paging industry has, of course, developed technologies over the years to meet these challenges. Modern paging receivers have system noise figures below 3 dB, and employ sophisticated AGC systems to address the dynamic range issues present in the RF environment. To achieve long battery life and low product cost, receivers powered from a single cell battery have been developed; in addition, paging protocols have been developed that provide for discontinuous receiver operation--the receiver is turned off the majority of the time the pager is in operation.

This paper examines 1-Volt RF circuit design in paging applications. Section 2 of this paper covers some general principles. Section 3 describes some specific circuit examples of RF amplifiers, mixers, and VCOs, and Section 4 concludes with a discussion of possible future directions in low-voltage RF design.

2. GENERAL PRINCIPLES

2.1 Battery Operation

It is often said that, while digital circuits obtain direct power savings from operation at lower voltage, the same does not apply to analog circuits: Operation at lower voltage often requires proportionately more current, so little or no net power savings result. This analysis overlooks the importance of the energy source in practical applications [1]. If a voltage conversion must take place to supply the circuit, power loss due to inefficiencies in the voltage conversion process may be significant. This is especially true in low-current applications where the converter overhead current (for bias, references, etc.) may not be negligible.

Switching voltage multipliers are of special concern in paging applications. Their waveforms typically have rapid transitions for good conversion efficiency; the resulting high-frequency harmonic content can cause receiver interference. (Recall that in many modern pagers no circuit board component is more than 2 cm away from the antenna.) Further, voltage converters cost precious board and/or IC die area, may add to component count, and increase costs. For all of these reasons, it is desirable to operate as much of the pager as possible directly from the battery.

2.2 Resonant Circuits

At baseband frequencies, active device load resistances must be either physical resistances or active loads (Figs. 1a and 1b). It is well known that, in the case of physical resistors, the supply voltage sets an upper bound on bipolar amplifier voltage gain [2]:

$$A_{V(\max)} = -\frac{V_{CC} - V_{sat}}{V_T}, \tag{1}$$

where V_{CC} is the supply voltage, V_{sat} is the saturation voltage of the active device, and V_T is the so-called thermal voltage ($\approx$ 25.9 mV at 300 K). Thus, even if $V_{sat} = 0$, with a 1-Volt supply the maximum possible voltage gain is $A_{v(max)} = 1/V_T = 38.6$.

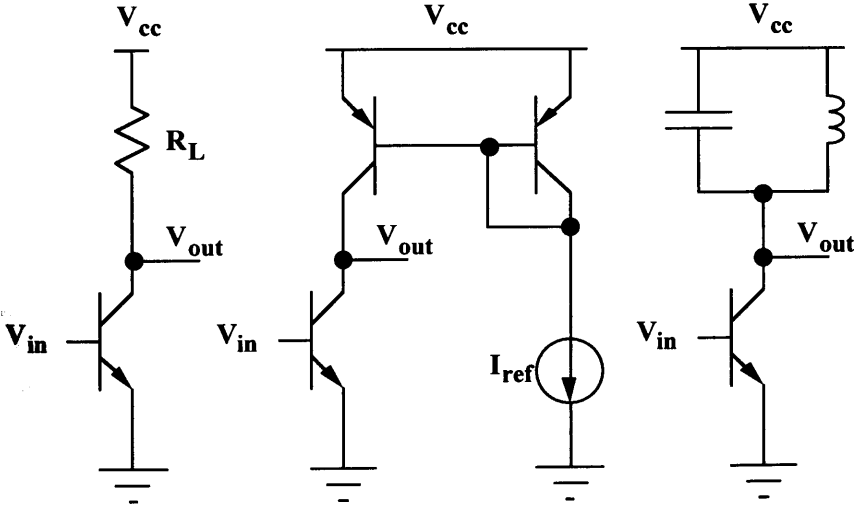


Fig. 1a. Amplifier with resistive load.

Fig. 1b. Amplifier with active load.

Fig. 1c. Amplifier with resonant load.

However, at RF frequencies one must consider the load impedance, rather than just the load resistance, and the effects of parasitic capacitance on the load can be severe. (The effect of this capacitance on current drain has been described elsewhere [3].) The capacitive reactance of 1 pF at 1 GHz, for example, is 159 ohms. Should parasitic capacitive reactance of this value exist across the load resistance, RF gain will be severely affected.

If the active device is a bipolar transistor with a collector current of 1 mA, its g_m at 1 GHz is at most 0.0386 ohm^{-1} (and probably much less, unless f_T is very high), so the amplifier voltage gain is at most $A_v = |g_m R_L| = (0.0386\ ohm^{-1})(159\ ohms) = 6.14$, or 15.8 dB.

Active loads are difficult to use in RF design because most IC processes are not truly complementary: The f_T of the PNP device, often a lateral design, may be two orders of magnitude less than that of the NPN device. The resulting amplifier load impedance is not the rather large r_o of the PNP, but instead its rather low parasitic capacitive reactance.

Being unable to develop significant load impedance due to parasitic capacitance, is the fundamental reason why physical resistances and active loads are rarely used at RF frequencies, especially in pagers and other power-sensitive applications. Instead, one turns a liability into an asset by adding an inductor to the load impedance, and employing parasitic capacitance as part of the total capacitance in a resonant tank circuit (Fig.1c). The great advantage of this structure in low-voltage design is that the entire supply voltage V_{CC} is now applied across the active device(s), enabling the use of a lower supply voltage than would otherwise be possible [4].

Use of resonant structures also implies that the resulting design will be relatively narrowband in nature. This fact may be used to advantage to improve receiver spurious response without the use of separate filter stages.

2.3 Topology Alternatives

The primary significant difference between 1-volt RF design, and RF design at higher voltages, is the limited number of circuit topology alternatives. Due to the voltage stackup, in bipolar design one is limited to one V_{be} and one or two V_{ce}s, depending on (among other factors) device size, current, required circuit linearity, and desired output voltage swing. This means many types of cascoding, for example, are not possible; the circuit designer must use ingenuity. The use of PNP transistors in low-frequency circuits is inevitable [5]; the RF designer awaits the time when typical integrated PNP devices become useful above 100 MHz. Typically the greatest design challenges are in support circuitry, such as biasing, base current compensation, and control circuits, and direct-conversion mixers, where

the loads are forced to be physical resistors or active devices, since the output is at baseband.

2.4 Biasing and the Power Supply

A well-designed supply and bias arrangement is critical for 1-Volt RF circuits. The terminal voltage variation of a single cell is not suitable for precise RF design (especially for oscillators); consequently, in pagers these circuits are typically supplied from a linear 1-Volt regulator. The design of this regulator is non-trivial: Besides operating itself from a single cell battery with as little overhead current as possible, the regulator must have a low dropout voltage (typically less than 50 mV), so the maximum life is obtained from the battery; it must be quiet both at RF frequencies, where it may interfere with received signals, and at baseband, where low frequency noise on the output may push (modulate) an oscillator; and it must have good power supply rejection (on the order of 40 dB or more) to reject sudden changes in battery terminal voltage caused by changing battery loads.

In addition, the regulator must have good absolute voltage tolerance as manufactured, since variations may affect RF circuit linearity, and the output voltage temperature coefficient (T.C.) must be well-controlled. Typically, the output voltage is made to track transistor V_{be} (i.e., T.C. $\approx$ -1 to -2 mV/C).

One of the (few!) advantages one has in pager RF design is that, since the product is usually designed for use on the body, the required operating temperature range is limited--usually -10 to +50 or +55 C.

2.5 Active Device Design

Until recently, the device of choice has been a bipolar transistor, for its high g_m/I_C ratio and good noise performance. However, recent MOS amplifier designs have been approaching (although not yet equaling) bipolar performance, and it seems clear that very small L_{eff} MOS devices soon will be very competitive.

For low-voltage operation, one wants large, ring-type bipolar devices for low V_{be}. As always, however, there are tradeoffs: For a given process technology, the larger device will have higher C_π, C_μ, and C_{cs} (and

therefore a lower f_T), forcing higher current drain. On the positive side, since its f_T is lower, the larger device typically has lower input impedance and therefore better intermodulation performance--always a consideration in low-voltage, low-current design--and a lower $V_{ce(sat)}$. This tradeoff of current drain vs. RF performance is not unique to 1-Volt, low-current design, and the emphasis placed on the lowest possible capacitance has been discussed elsewhere [1]. Other features of good RF transistors, such as low base resistance, are also not unique to 1-Volt design.

Similarly, for low-voltage operation, one wants RF MOS devices with low threshold voltage. The device designer has an extra degree of freedom with MOS, since the threshold voltage can be controlled independently of the device size, perhaps enabling the circuit designer to arrive at a better capacitance compromise. Shaeffer and Lee [6] have recently published an excellent guide to designing MOS RF LNA transistors.

Since paging receivers are turned off a majority of the time they are in operation, device leakage current is especially important. This leads one to the simultaneous desire for low threshold voltage devices for RF and high threshold voltage devices for biasing and control. Since the author is a receiver engineer, rather than an IC process engineer, the optimum solution to him seems to be a dual-threshold process.

3. CIRCUIT EXAMPLES

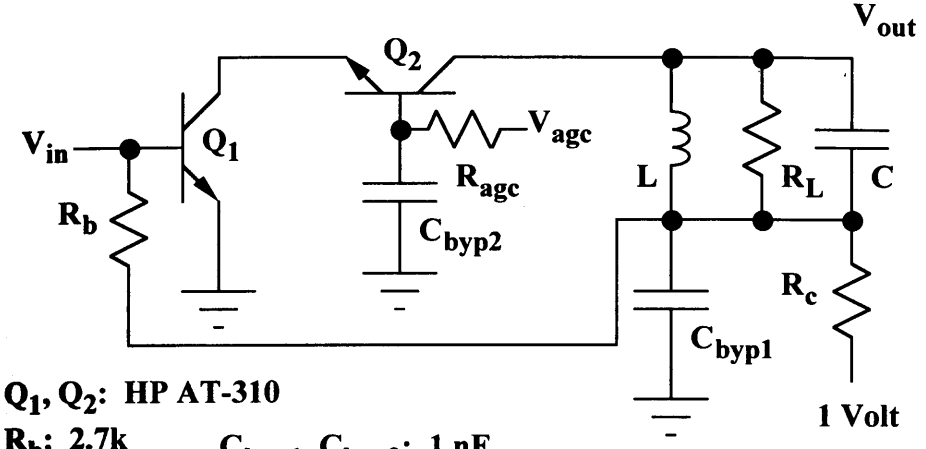
3.1 RF Amplifiers

A single-ended pager RF amplifier is shown in Fig. 2. It is a cascoded bipolar amplifier supplied by a 1-Volt regulator, and is typical of discrete SMD designs in production today. Cascode designs are favored due to several pleasing properties [7]:

1. High output impedance, so that the loaded Q of the output tank circuit remains high, improving both gain and selectivity;
2. Low s_{12} (reverse transmission), encouraging stability; and,
3. Q_1 voltage gain = 1, since its transconductance is g_m and the input impedance of Q_2 (the load impedance of Q_1) is $1/g_m$.

This last point is important in 1-Volt design because the base of Q_2 is nominally biased at 1 Volt, so its emitter voltage is a V_{be} less than this, or

approximately 300 mV. Thus, Q_1 V_{ce} is nominally 300 mV, and little collector voltage swing could be tolerated without saturating the device. Since Q_1 voltage gain is unity, however, this is not a problem unless extremely strong signals are encountered.



Q_1, Q_2: HP AT-310
R_b: 2.7k
R_c: 180
R_L: 1k
R_{agc}: 680
C_{byp1}, C_{byp2}: 1 nF
C: 3.9 pF
L: 6.8 nH

Fig. 2. Practical 1-Volt, 930 MHz RF amplifier.

R_b and R_c form a negative feedback bias circuit. Should the amplifier current increase, the voltage across R_c increases, reducing the base bias to Q_1 through R_b and reducing the amplifier current. R_c is used to set the amplifier current.

For best IM performance and stability C_{byp1} and C_{byp2} should provide as broadband bypass as is possible. Multiple capacitors, including values up to 1 µF, may be necessary.

R_L is added to the output tank circuit for two reasons. The first is to control amplifier gain by controlling the real part of the tank impedance: Without R_L this would be solely a function of the Q of the tank inductor L [8], a value that is not always well controlled. Secondly, R_L enables the designer to decouple amplifier current drain and gain. This may be desirable, for example, when it is desired to have moderate gain, but good

IM performance. In this case, the designer may opt to increase current drain for improved IM performance, by decreasing R_c, but reduce the amplifier gain by loading the output tank circuit with a lower value of R_L.

For improved IM3 performance at the cost of some increase in noise figure and loss of gain, a small amount of Q_1 emitter degeneration, in the form of a 10- to 50-ohm resistor, may be used. For improved IM2 performance, the low-frequency (below 40 MHz) gain of the amplifier may be reduced by employing a bypassed emitter resistor at Q_1 [9].

The amplifier's AGC control point is the base voltage of Q_2. The nominal, maximum gain potential at this point is 1 Volt. The AGC functions by lowering this voltage, thereby lowering the Vce of Q_1, eventually putting it into saturation. R_{agc} functions as an RF blocking resistor. The advantage of this gain reduction method over other methods of reducing the amplifier current (such as by reducing Q_1 base voltage), is that the amplifier's IIP3 remains substantially unaffected as the gain is reduced. This produces better overall receiver IM performance.

With a total current drain from 1 Volt of 1.1 mA, at 930 MHz this amplifier is capable of 15 dB gain, 2 dB noise figure, and an IIP3 of -22 dBm. A similar amplifier, tuned for 150 MHz, may have a current drain of 600 μA, 18 dB gain, 1.4 dB noise figure, and an IIP3 of -22 dBm.

A pseudo-differential version of this amplifier, suitable for integration, is shown in Fig. 3. The bias method has been changed to a current mirror approach. With a supply of 1 Volt, however, a true cascoded differential amplifier is not possible, due to the resulting low V_{ce} on the current source transistor. To circumvent this problem, the input devices Q_1 and Q_2 are current mirror biased from diode Q_5. The bias circuit is isolated from the amplifier input by resistors R_{blk1} and R_{blk2}, whose values are determined by the simultaneous need for RF isolation, and concern for Q_1 and Q_2 beta variation effects. The resulting amplifier is essentially two single-ended amplifiers fed out of phase.

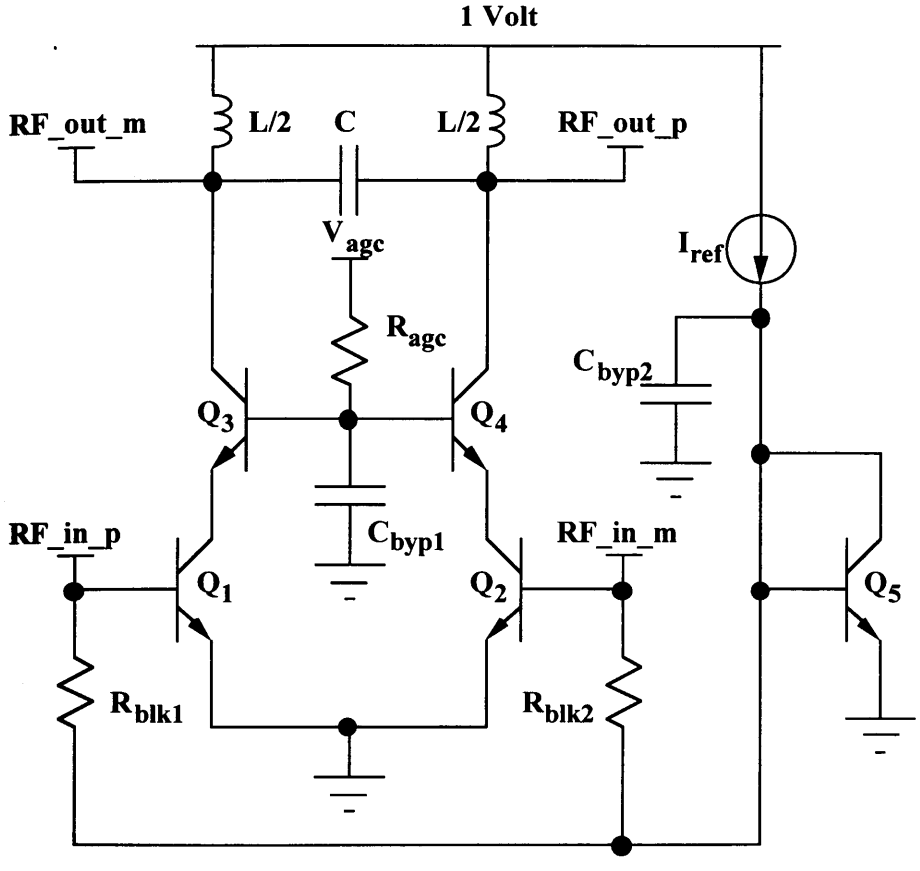


Fig. 3. Pseudo-differential 1-Volt RF amplifier.

3.2 Mixers

Discrete mixers in superheterodyne pagers are crude, inexpensive, very simple, very current-efficient, and hard to analyze (Fig. 4). Both the RF and LO inputs are injected across the emitter-base junction of a bipolar transistor; the output is taken from a tank at the collector, resonant at the IF. Since there is no RF-LO diplexer per se, tuning of the preselector and injection filter tank circuits is critical to successful operation. It has been found experimentally that an LO injection level of approximately 50 mV_{rms} is optimum; compared to the DC collector current without injection this corresponds to an increase in current of about twelve percent.

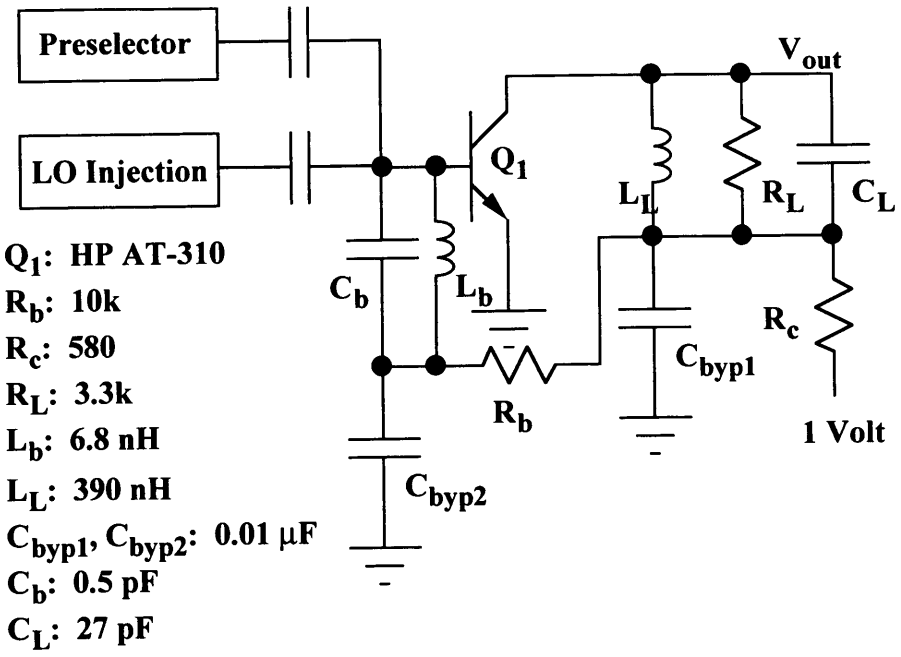


Fig. 4. Practical 1-Volt, 930 MHz mixer with 45 MHz IF.

As with all mixers, to ensure stability it is important that the transistor load impedance is small at the RF and LO frequencies, and that the source impedance is low at the IF frequency. Keeping the source impedance low at the IF frequency also reduces any noise at the IF that would otherwise degrade the mixer noise figure. Thus it is important in this design that the RF and IF be significantly separated in frequency, so that the reactance of C_L may be small at the RF and LO frequencies, and the reactance of L_b may be small at the IF frequency.

With a total current drain from 1 Volt of 400 µA, at 930 MHz this mixer is capable of 11 dB gain, 7 dB noise figure, and an IIP3 of -20 dBm. A similar design, tuned at 150 MHz, is capable of similar performance while drawing only 250 µA.

The commercial design of integratable mixers has, in large measure, centered around the need for high-performing mixers in zero-IF receivers.

There is a rich literature on the subject (e.g., [13]) and many designs have been patented; however, few of these operate from a 1-Volt supply.

As IC processes improve more 1-Volt superheterodyne mixer designs are expected to emerge; mixer circuits such as that described in [10] may find use in superheterodyne receivers.

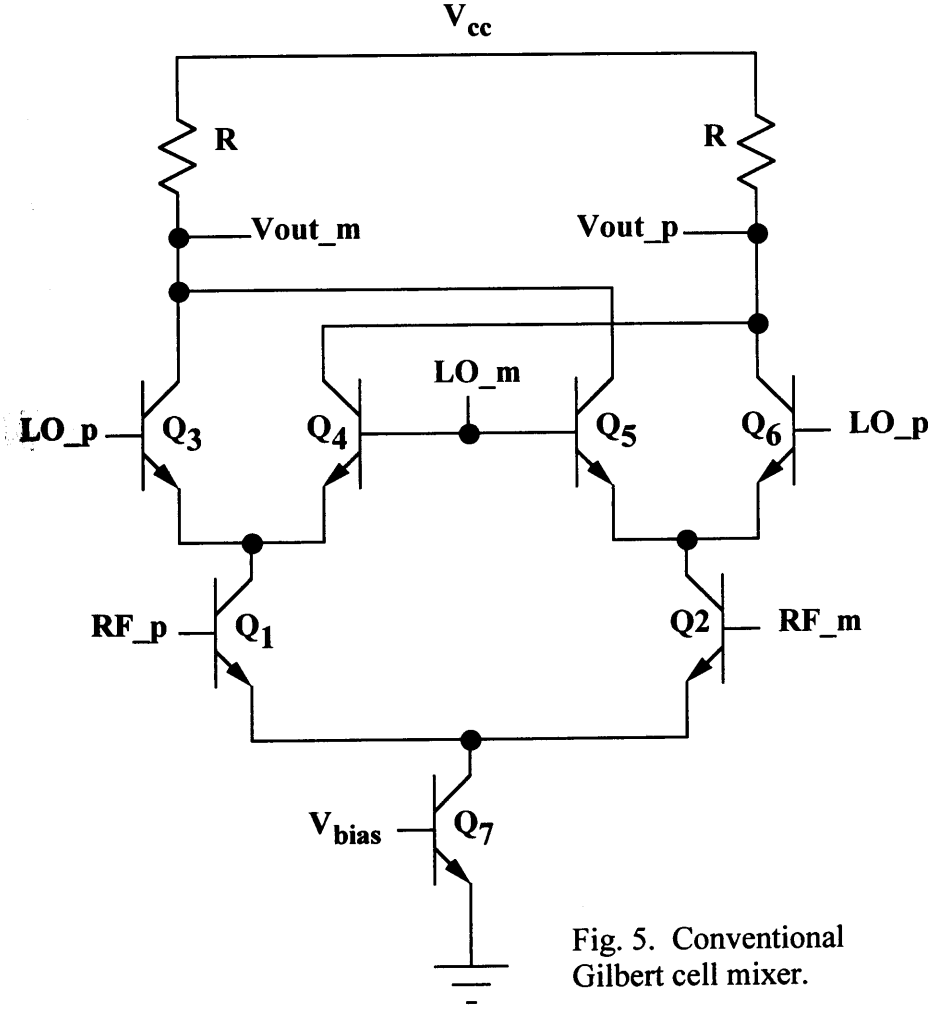


Fig. 5. Conventional Gilbert cell mixer.

The starting point for most 1-Volt zero-IF mixer designs is the well-known Gilbert cell [14] (Fig. 5). As it stands, it will not operate with a 1-Volt

supply. To surmount this problem, current source bias transistor Q_7 is usually deleted, and transistors Q_1 and Q_2 are rebiased. One method is analogous to the method used in the pseudo-differential RF amplifier (Fig. 6).

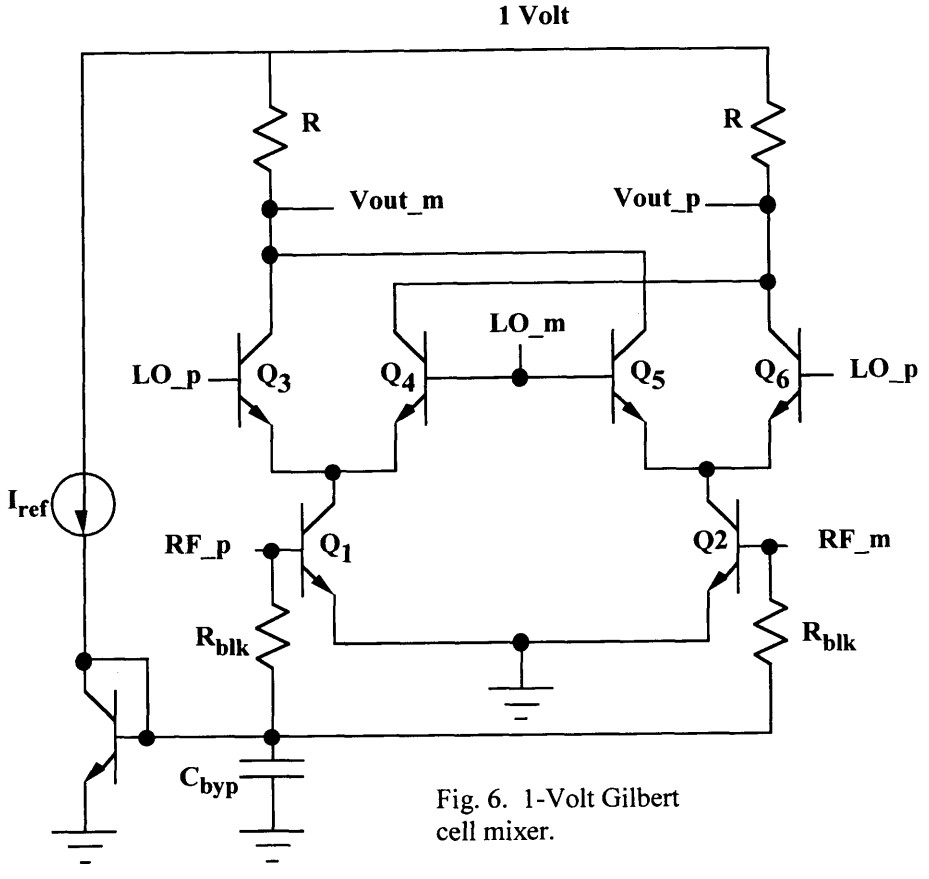


Fig. 6. 1-Volt Gilbert cell mixer.

With these mixers, one is always in search of the optimum tradeoff between current drain, gain, noise figure, IIP2, and IIP3. One method of achieving a better compromise is to realize that Q_1 and Q_2 need not be in the common-emitter configuration. The RF input to these mixers has been placed at the base [15], collector [16], and emitter [13] of Q_1 and Q_2.

An important, yet often-overlooked requirement in 1-Volt mixer design is the LO power requirement, for if the mixer operates from 1-Volt, the system probably requires that the LO source must, too. Many have

suggested the use of passive diode-ring mixers in pagers, but the LO power requirement makes them a poor choice.

3.3 VCOs

When 1-Volt LNAs and (especially) 1-Volt mixers struggle with dynamic range, one may expect 1-Volt VCOs to struggle with their related parameter, sideband noise, and this is indeed the case. Paging VCOs are typically allotted approximately 600 μA from 1 Volt, and must achieve a single-sided sideband noise specification of approximately -115 dBc/Hz at 25 kHz offset (20 kHz in Europe). This must be achieved with miniature components of moderate Q and low cost. Due to the sideband noise constraint, paging VCOs remain discrete bipolar circuits.

Q_1: HP AT-310
L_1-L_4: 2.1 μH
R_1: 100
R_2: 22
R_3: 510
L_5, C_1-C_3: Set by application

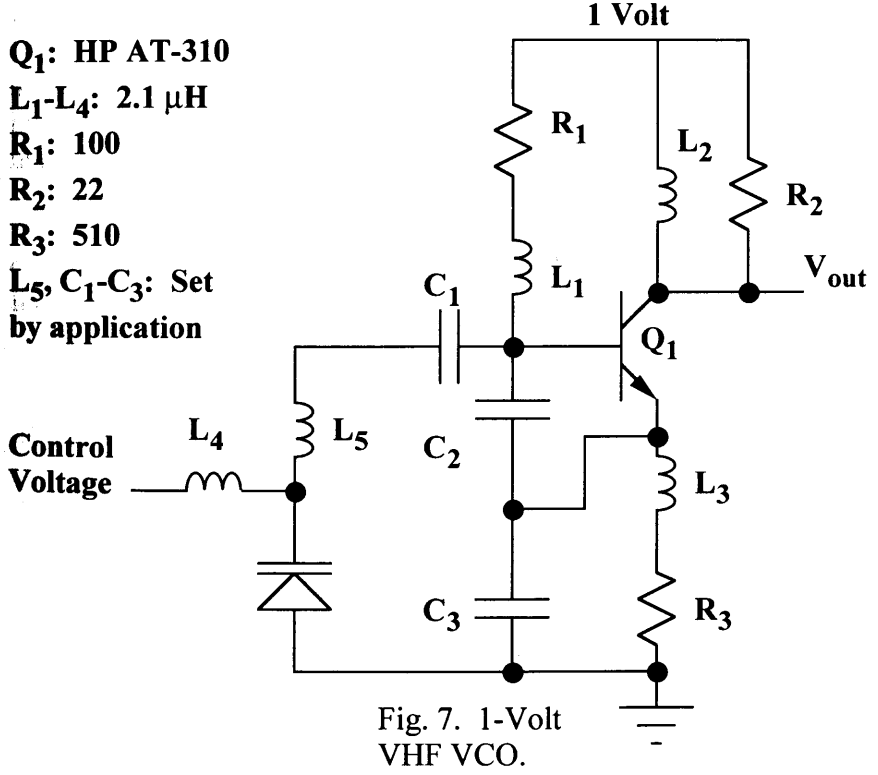


Fig. 7. 1-Volt VHF VCO.

A typical VHF pager VCO is shown in Fig. 7. It is a modified Colpitts design, in that the collector is raised slightly above ground, and the output

voltage is taken at that point. For good sideband noise performance, as with all VCOs it is vital that the supply noise be reduced to a minimum. The control voltage, too, must be quiet and, while its range will depend on product requirements, it unfortunately will have to exceed 1 Volt in most applications due to the limitations of available varactors. The Q of most varactors drops significantly below 500 mV bias, further limiting available tuning range. Most pagers employ a control voltage range of 0.5 to 2.5 Volts or so.

L_1 is an RF choke which passes the 1 Volt DC bias to the base of Q_1 while isolating the tank from ground. R_1 is used to reduce the Q of L_1, preventing spurious oscillations resulting from L_1 resonating with the capacitance at the base of Q_1. L_2 provides DC bias to the collector of Q_1. R_2 reduces the Q of L_2, lowering and controlling the collector impedance to ground. L_3 passes the DC emitter current to R_3 while keeping the relatively low impedance of R_3 from appearing across the VCO tank. R_3 sets the VCO emitter current, since the base of Q_1 is fixed at 1 Volt. L_4 is an RF choke that isolates the control voltage line from the VCO tank. L_5, C_1-C_3, and the varactor choice all depend on the intended application, and are affected by the desired tuning range, phase noise, startup time, and output level.

The output voltage of this oscillator is about 50 mV_{rms}. This is sufficient to drive a single-transistor mixer, and is low enough to avoid serious EMI concerns.

A paging receiver incorporating some of the above 1-Volt RF circuits is shown in Fig. 8.

4. FUTURE DIRECTIONS

After decades of relative obscurity, 1-Volt RF design is beginning to receive considerable attention. It is instructional, therefore, to consider future trends so that this attention is efficiently used.

The attention 1-Volt design is now receiving is due in large part to the continuing supply voltage reduction of digital systems. Since RF circuits are becoming a part of many digital systems, in the form of handheld wireless products, their supply voltages are inexorably tied to those of the digital systems. These supply voltages will continue to fall below 1 Volt, since

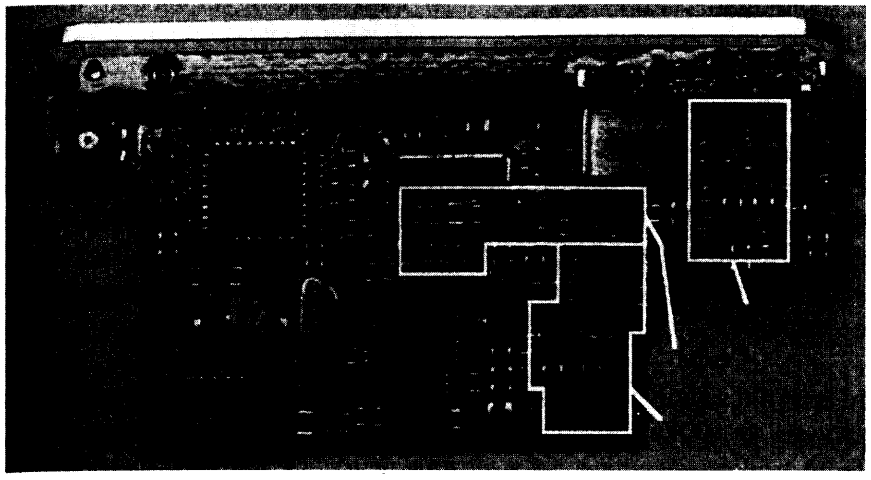


Fig. 8. A paging receiver incorporating a 1-Volt RF Amplifier, mixer, and VCO.

digital circuits continue to offer power savings as their supplies are reduced, and the trend to lower voltages shows no signs of stopping (e.g., [17]). How are RF designers to keep pace?

4.1 The Emergence of RF MOS

RF design with silicon bipolar transistors seems limited to a supply voltage just below 1 Volt by the bipolar V_{be} voltage. This barrier is made somewhat flexible by the emergence of SiGe HBTs, which offer a respite of 300 mV or so. The analogous MOS threshold voltage, however, is controllable to lower values. This implies that the use of MOS devices in RF design may be the path for 0.5-V solar cell applications, for example.

Fortuitously, MOS lithography is now reaching dimensions small enough to provide devices with useful RF performance. In addition to the MOS mixers described in [11], [12], [20], and [26], MOS LNAs have been reported in [6] and [18]-[23], and entire CMOS RF Front Ends in [24] and [25]. Small-geometry MOS RF devices offer several advantages for low-voltage wireless use. The primary advantage is that the devices improve in f_T, gain, and noise factor with each L_{eff} reduction. RF performance is thus tied to Moore's Law, and the RF designer can look forward to several generations of ever-improving devices, operating at ever-lower voltages, motivated by industry-wide economics. This is a far

more rapid (and certain) advance than that occurring in bipolar device design.

Additional process modifications that improve digital circuit performance, such as lowering substrate capacitance, employing interconnect dielectrics with lower dielectric constant for reduced crosstalk, adding more interconnect metal layers, etc., also will improve RF performance in general. This leads one to the conclusion that RF and digital design will continue along the present system integration path, making the mixed-signal IC the standard RF design problem of the future.

There is an old saying in IC design to the effect that one should "take what the process gives you," meaning that one should modify system and circuit design to exploit advantages of a given process, in order to overcome its disadvantages. The disadvantages of mixed-signal design in digital processes (suboptimal or non-existent analog components, noise, etc.) are certainly well-known; what advantages are available to overcome them?

One advantage to exploit is the density of digital circuits. Since digital MOS devices are so small, exquisitely detailed trimming, tuning, and control algorithms of RF circuits becomes possible, eliminating large external tuning components while making "smarter" RF devices and circuits. This may be the salvation of applications that could otherwise not use a digital MOS process--balanced mixers with a high IIP2 specification, for example, that may otherwise not be able to tolerate the matching performance of digital MOS could employ a matching improvement algorithm upon initialization. Since time is measurable to greater accuracy than almost any other physical parameter, a suitably clever circuit designer can use digital clocks to measure other parameters with great precision (e.g., the dual-slope voltmeter).

Another liability to be turned into an asset is the low supply voltage itself. As it happens, due to the excess noise generated by hot-electron effects in short-channel MOS devices, device minimum noise factor is obtained with low V_{ds} [6]. Thus, for low-noise operation, operation below 1 Volt may in fact be a requirement for MOS RF circuits.

Further, as MOS threshold voltages drop, more "higher voltage" techniques appear in the 1 Volt realm. For example, with V_t = 300 mV, a true Gilbert cell mixer becomes a possibility with a 1-Volt supply. Even the Darlington configuration--a distant dream to present-day 1-Volt circuit designers--becomes a reality.

And, while fast switching devices will always produce RF noise, the lower digital supply voltages and currents will at least help reduce digital noise that may couple to the RF and analog circuits through the chip's supplies, package, or substrate.

4.2 Reality

Despite the cleverness of circuit designers present and future, it is a sad fact of the physical world that the maximum attainable signal-to-noise ratio (SNR) of analog circuits falls with the square of the supply voltage [1]. While this is a compelling argument to move the signal from analog to digital representation as quickly as possible, one still has this limitation in the ADC. All else being equal (which, granted, is rarely the case), one loses 8.6 dB in theoretically attainable SNR moving from a 2.7-V supply to a 1-V supply, and an additional 6 dB moving from there to a 0.5-V supply. This is happening, of course, not because the noise is increasing (in fact, there may actually be less digital noise present in practical designs), but because the maximum attainable signal is reduced, since it is limited by the supply. Translated into RF receiver terms, sensitivity may not be affected, but performance is strong signal environments may suffer. This may be addressed two ways: By adapting the receiver to its environment, thereby limiting the instantaneous dynamic range required, or by controlling the maximum signal level present in the environment. An optimum system will probably involve both.

The dynamic range of all receivers is, of course, finite. When this dynamic range becomes less than that of signals present in the RF environment, steps must be taken to adapt the receiver. These steps are usually referred to as Automatic Gain Control (AGC) or, in general, adaptive signal processing, and they are a cornerstone of successful 1-Volt RF design. It is a complicated subject because, due to the filtering present in all receivers, the receiver's dynamic range is a function of frequency. And, since different receiver circuits affect signal quality in different ways when their

dynamic range is exceeded (intermodulation, analog filter collapse, blocking, ADC overrange, etc.), receiver stages must be protected in different ways. The traditional analog type of AGC, in which received signal strength is averaged and the signal path gain adjusted accordingly in an analog feedback loop, will be supplanted by other, more sophisticated methods of adaptive signal processing in the future.

Due to the cost, complexity, and design cycle time associated with sophisticated adaptive signal processing, even when performed with DSP, in the future the circuit designer will receive help from the wireless system designer, who will design the wireless system with the limited dynamic range of 1-Volt (and lower) receivers in mind. Adjacent channel and blocking signal strength limitation is already being done in cellular systems today (DECT, GSM), and greater emphasis will be placed on this in the future.

REFERENCES

[1] E. Vittoz, "Low-Power Low-Voltage Limitations and Prospects in Analog Design," in *Analog Circuit Design*, edit. R. van de Plassche, W. Sansen, and J. Huijsing, Kluwer, 1995.

[2] P. Gray and R. Meyer, *Analysis and Design of Analog Integrated Circuits, third ed.*, John Wiley and Sons, Inc., New York, 1993, p. 287.

[3] E. Vittoz, "Physical Limitations on CMOS Analog ICs," Mead Microelectronics: Advanced Analog CMOS IC Design Techniques Course, San Francisco, CA, April 1996.

[4] A. Abidi, "Low-Power Radio-Frequency IC's for Portable Communications," Proc. of the IEEE, vol.83, no. 4, pp. 544-569, April 1995.

[5] M. Fonderie and J. Huijsing, *Design of Low-Voltage Bipolar Operational Amplifiers*, Kluwer, Boston, 1993, p. 4.

[6] D. Shaeffer and T. Lee, "A 1.5-V, 1.5-GHz CMOS Low Noise Amplifier," IEEE J. Solid-State Circuits, vol. 32, No. 5, pp. 745-759, May 1997.

[7] P. Gray and R. Meyer, *op. cit.*, pp. 511-512.

[8] H. Krauss, C. Bostian, and F. Raab, *Solid State Radio Engineering*, John Wiley and Sons, Inc., New York, 1980, ch. 3.

[9] R. Meyer and A. Wong, "Blocking and Desensitization in RF Amplifiers," IEEE J. Solid-State Circuits, vol. 30, No. 8, pp. 944-946, August 1995.

[10] C. Chan, H-S Ling, and O. Choy, "A One Volt Four-Quadrant Analog Current Mode Multiplier Cell," IEEE J. Solid-State Circuits, vol. 30, No. 9, pp. 1018-1019, September 1995.

[11] J. Crols and M. Steyaert, "A 1.5 GHz Highly Linear CMOS Downconversion Mixer," IEEE J. Solid-State Circuits, vol. 30, No. 7, pp. 736-742, July 1995.

[12] P. Kinget and M. Steyaert, "A 1-GHz CMOS Up-Conversion Mixer," IEEE J. Solid-State Circuits, vol. 32, No. 3, pp. 370-376, March 1997.

[13] J. Wilson, et al., "A Single-Chip VHF and UHF Receiver for Radio Paging," IEEE J. Solid-State Circuits, vol. 26, No. 12, pp. 1944-1950, December 1991.

[14] B. Gilbert, "A Precise four-Quadrant Multiplier with Subnanosecond Response," IEEE J. Solid-State Circuits, vol. SC-3, pp. 365-373, December 1968.

[15] NEC, "μPC7349GB Direct Conversion FSK Data Receiver Preliminary Data Sheet," July, 1995.

[16] C. Takahashi, *et al.*, "A 1.9 GHz Si Direct Conversion Receiver IC for QPSK Modulation Systems," 1995 IEEE International Solid State Circuits Conference Digest of Technical Papers, pp. 138-139.

[17] F. Assaderaghi, *et al.*, "A Dynamic Threshold Voltage MOSFET (DTMOS) for Ultra-Low Voltage Operation," 1994 IEEE International Electron Devices Meeting Technical Digest, pp. 809-812.

[18] J. Y.-C. Chang and A. A. Abidi, "A 750 MHz RF amplifier in 2-μm CMOS," 1992 Symp. VLSI Circuits Digest of Technical Papers, pp. 111-112.

[19] J. Ou, "CMOS Low-Noise Amplifier Design for Monolithic RF Modems," http://kabuki.eecs.berkeley.edu/~jjou/slides/rf.11.95/

[20] A. Karanicolas, "A 2.7-V 900-MHz CMOS LNA and Mixer," IEEE J. Solid-State Circuits, vol. 31, No. 12, pp. 1939-1944, December 1996.

[21] C.-Y. Wu and S.-Y. Hsiao, "The Design of a 3-V 900-MHz CMOS Bandpass Amplifier," IEEE J. Solid-State Circuits, vol. 32, No. 2, pp. 159-168, February 1997.

[22] R. Johnson, *et al.*, "A 2.4 GHz Silicon-on-Sapphire CMOS Low-Noise Amplifier," IEEE Microwave and Guided Wave Letters, vol. 7, no. 10, pp. 350-352, October 1997.

[23] M. Harada and T. Tsukahara, "Monolithic 2 GHz and 4-6 GHz Low Noise Amplifiers Fabricated on a SIMOX Wafer," 1997 International Conference on Solid State Devices and Materials Extended Abstracts, pp. 132-133.

[24] A. Rofougaran, *et al.*, "A 1 GHz CMOS RF Front-End IC for a Direct-Conversion Wireless Receiver," IEEE J. Solid-State Circuits, vol. 31, No. 7, pp. 880-889, July 1996.

[25] A. Shahani, *et al.*, "A 12mW Wide Dynamic Range CMOS Front-End for a Portable GPS Receiver," IEEE J. Solid-State Circuits, vol. 32, No. 12, pp. 2061-2070, December 1997.

[26] M. Borremans and M. Steyaert, "A 2-V, Low Distortion, 1-GHz CMOS Up-Conversion Mixer," IEEE J. Solid-State Circuits, vol. 33, No. 3, pp. 359-366, March 1998.

DC/DC Conversion, the key to low power consumption.

E.C.Dijkmans, F.J.Sluijs.

Philips Research Labs, Eindhoven,
The Netherlands.

Abstract.

Currently the development of rechargeable batteries is concentrated on lithium-type chemistries, which provide a maximum open output voltage of 4.2 V, while the voltage limit in silicon VLSI technology is about 2 Volt as a maximum usable supply voltage. Optimization towards low power consumption will require a multitude of supply voltages, depending on the application. High efficient DC/DC converters will bridge the gap, using either a coil or a capacitor for intermediate energy storage. Efficiency of capacitive converters is strongly dependent on the ratio between the switched capacitor and its parasitics, as well as the discharge ratio of the capacitor. Inductive converters are preferred as they allow an arbitrary ratio of input and output voltage. Both types can show about 90% efficiency.

Introduction.

Battery operated portable systems like notebooks, GSM phones, hearing aids etc. require signal processing functions at minimum power consumption with a strong variation in speed requirements. In digital signal processing a clear trade-off can be made between speed and supply volt-

age, and supply voltages as low as 0.5 Volts can be expected. The supply voltage requirements for analog circuitry are often dictated by external signal sources and loads. Battery development has changed from Nickel-type cells, with an maximum cell voltage of 1.6 V during charging, to Lithium-ion and in future the Lithium-polymer cells which show maximum cell voltages up to 4.2 V during charging. Switched Mode DC/DC converters are essential for efficient conversion of the battery voltage to various supply voltages, needed to perform every function with minimum power drain. Both capacitors and coils are used in switched mode DC/DC converters for temporary energy storage. This paper will focus mainly on configurations suited for low power and low voltage operation.

Batteries.

The open voltage of battery cells is strongly dependent on the materials used, as shown in table 1. However the most important parameters for the choice of materials are the energy-to-weight ratio and the energy-to-volume ratio. In most portable equipment the size and weight of the battery is still dominating, especially when rechargeable or so-called secondary batteries are used.

Table 1: Various battery systems

Battery type	System	Nominal voltage	Maximum voltage	Energy density* Wh/kg	Energy density* Wh/ltr
Primary	$ZnMnO_2$	1.2 V	1.65 V	100-140	300-400
Primary	Zn-air	1.2 V	1.65 V	200-350	650-1050
Primary	Li-(CF)n	2.5 V	3.2 V	100-300	400-700
Primary	Li-MnO2	2.8 V	3.3 V	150-250	400-500
Secondary	SLA	2 V	2.3 V	30	60
Secondary	NiCd	1.2 V	1.6 V	32-55	86-180

Table 1: Various battery systems

Battery type	System	Nominal voltage	Maximum voltage	Energy density* Wh/kg	Energy density* Wh/ltr
Secondary	NiMH	1.2 V	1.6 V_	40-80	110-330
Secondary	Li-ion	3.6 V	4.2 V	60-110	110-300
Secondary	Li-polymer	3.6 V	4.2 V	50-100	100-250

The most widely used battery cells at the moment are the non-rechargeable or primary cells, having an open circuit potential of 1.65 V. From rechargeable or secondary cells the most popular are NiCd or its successor, the NiMH cells, with basically comparable characteristics having a nomonal cell voltage of 1.2 V, and at present the Li-ion battery, which differs significantly from the two nickel based batteries. Because battery characteristics are strongly dependent on the battery type (each battery technology is delivered in a wide range of sizes and types) and manufacturer, only the most common characteristics are described.

The NiCd battery displays the lowest energy density per unit volume and weight of the three technologies, a nominal cell voltage of about 1.2 V, an ohmic resistance of about 20 mΩ (AA cell) and a rather high self-discharge rate of 1-2% per day (strongly dependent on temperature). It has the advantage of robustness with respect to overcharging, which allows the use of relatively cheap chargers, as well as robustness to overdischarging. They also have the advantage of the possibility to draw large currents from the battery.

The NiMH battery has the negative cadmium electrode replaced by a Metal-Hydride electrode with the ability to absorb large quantities of atomic hydrogen, especially resulting in a larger energy content per unit volume than the NiCd battery while having the same nominal cell voltage, ohmic resistance and self-discharge rate. The NiMH battery shows less robustness with respect to overcharging and overdischarging, resulting in a more expensive charger. The latter results from the fact that

less adequate effects in the voltage curve appear during charging and temperature measurement becomes required for charge termination in case of fast charging.

However, growing demand in energy density has moved developments towards more exotic materials resulting in Li-ion and the future Li-polymer cells. Those materials show nominal cell voltages of 3.6 V.

The Li-ion cell differs from the nickel based systems. Next to the higher output voltage it has a larger energy content per unit weight and a lower self-discharge rate of about 0.3% / day (again strongly dependent on temperature). Overcharging and overdischarging must be avoided at all times because there are no controlled overcharge or discharge reactions in the cell. In fact, overcharging without any protection will result in fire or explosion, leading to the necessity to include a safety IC inside the Li-ion pack, interrupting the charge/discharge current when the battery voltage exceeds 4.2 V or drops below 2.3 V. During normal operation the total ohmic resistance for a single-cell battery pack, including the battery (100 mΩ), can amount to about 400 mΩ. The possibility to draw large currents from a Li-ion cell is lower than for Ni-type cells.

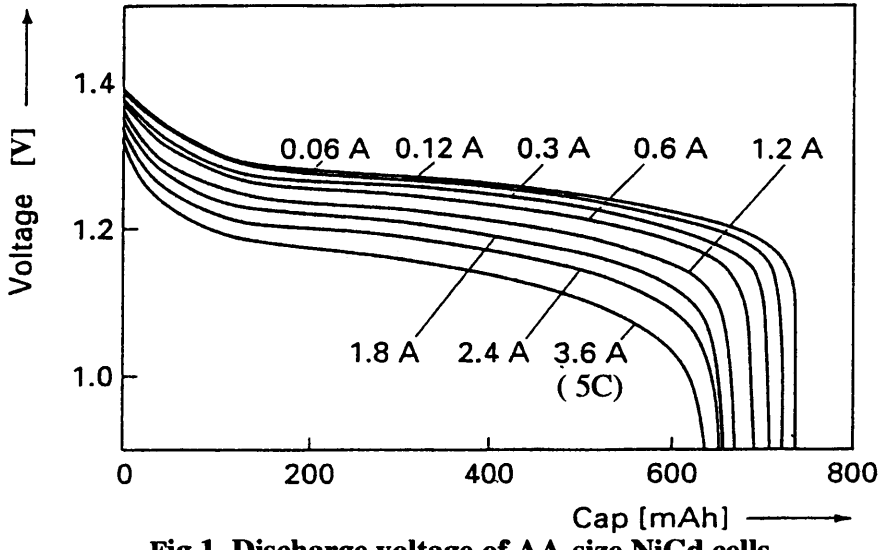


Fig.1. Discharge voltage of AA-size NiCd cells for various discharge rates.

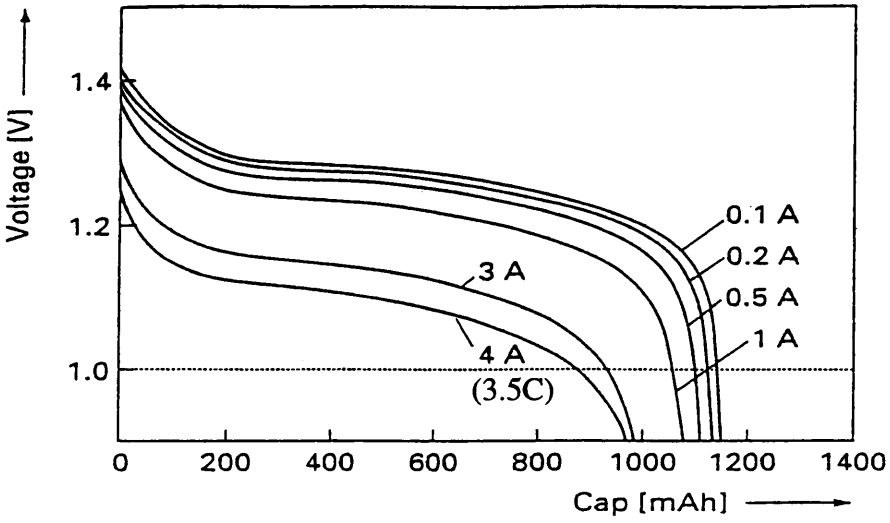


Fig. 2. Discharge voltage of AA-size NiMH cells for various discharge rates.

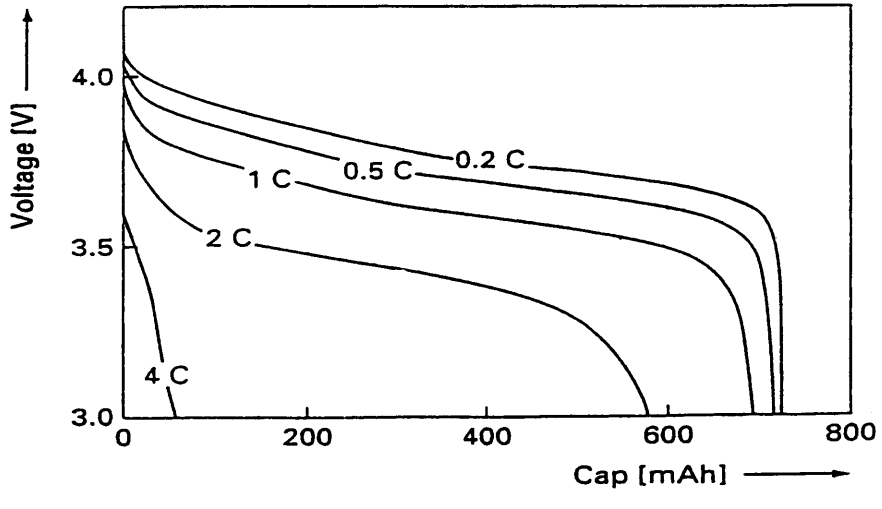


Fig. 3. Discharge voltage of a Li-ion cell at various discharge rates.

Fig. 1, 2 and 3 display discharge curves of an AA size NiCd battery, a NiMH battery and a Li-ion battery. It clearly shows the decrease of accessable energy as a function of discharge current. Often these curves are normalised to the capacity C, as used in Fig. 3, in which 1 C is the current at which the battery is discharged in one hour.

Market development.

To get an impression of which way the global market is moving we have to distinguish the power consumption per catagory.Applications with extreme low power or long battery refresh times, like watches, pagers, hearing aids etc., are expected to use primary cells for quite some time. The laptop and notebook market is changing in 1998 almost completely to Li-ion as is shown in the graph of Fig.4.

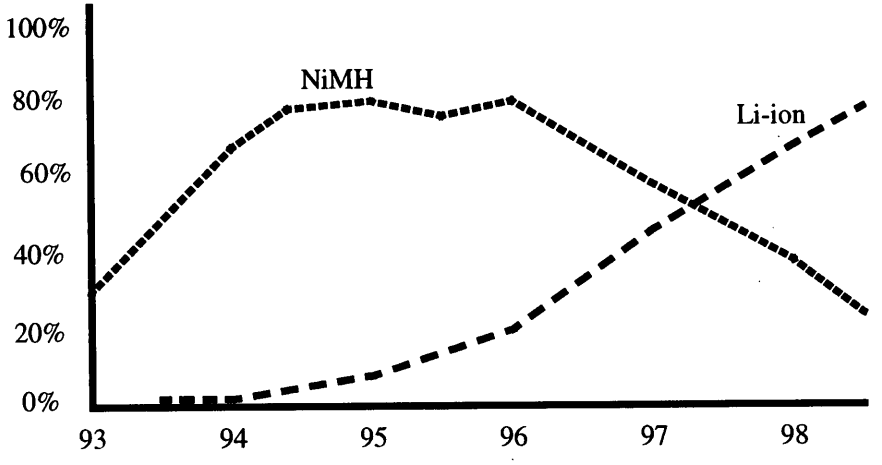


Fig 4.Notebook market share for NiMH and Li-ion

A similar development is going on in the cellular phone market, as shown in Fig.5.

For Li-polymer no hard data is yet available, but the general conclusion must be that future rechargeable portable equipment has to operate from a supply voltage of up to 4.2 V. Only equipment working with primary cells will continue to operate on around 1 V supply voltages. Those curves can be found in various market surveys like Dataquest and Nomura.

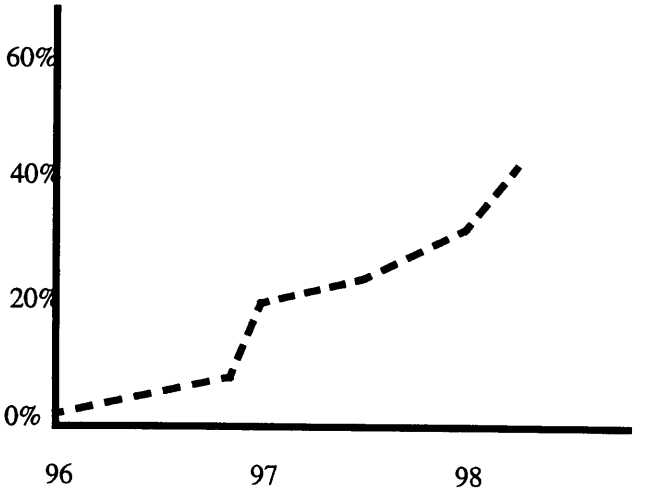


Fig. 5. Penetration of Li-ion in EU and third market

Silicon technology.

A similar roadmap prediction for silicon technology is published by the Silicon Industries Association Ref.[1], which learns us which maximum supply voltages can be applied on future sub-micron IC's (table. 2.). The dominating limiting factor here is the gate oxide breakdown voltage. Although the SIA roadmap only gives the maximum allowed voltages for logic, the values which are of interest for analog functions, can be derived from maximum field strength and oxide thickness. Here we see that usable power supply voltages will stay in the order of 1.5 to 2 V, even for technology below 0.1 Micron.

Table 2: Extraction from the SIA roadmap

Year of shipment	1997	1999	2001	2003	2006	2009	2012
Technology	250 nm	180 nm	150 nm	130 nm	100 nm	70 nm	50 nm
Vdd. logic	2.5-1.8	1.8-1.5	1.5-1.2	1.5-1.2	1.2-0.9	0.9-0.6	0.6-0.5
T_{ox} equiv.	4-5 nm	3-4 nm	2-3 nm	2-3 nm	1.5-2	1.5	1
Max. E-field.	5 MV/ cm	5 MV/ cm	5 MV/ cm	5 MV/ cm	>5 MV/ cm	>5 MV/ cm	>5 MV/ cm

Extrapolation and simulation of deep submicron devices, predict a usable gate-source voltage of 1.5 V, as can be seen in Fig. 6.

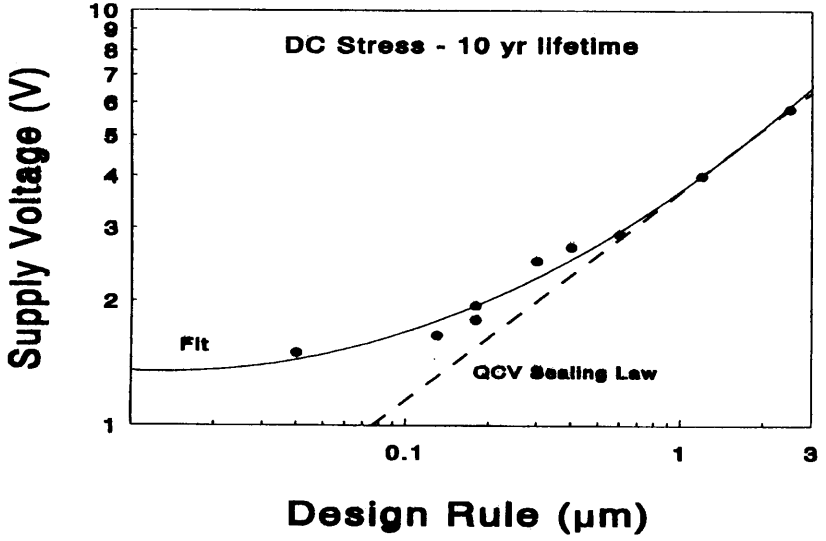


Fig.6. Maximum allowed gate to drain/source voltage.

In the SIA roadmap it is assumed that around 2005 a replacement for SiO_2 as a gate isolator has to be found which shows a higher ε. The

maximum field strength properties are supposed to increase, but to what extent is still uncertain.

It is clear that due to mismatch between battery voltage and required supply voltages DC/DC conversion is a necessity, and will be performed either by introducing a separate IC with sufficient high-voltage capability, or by the use of high voltage options. As most equipment requires the on-chip OTP or MTP-ROM functionality, it is likely that devices which can withstand higher voltages will always be available, and can also be used for high performance analog functions. Whether they can withstand battery voltages including some overshoot however is more uncertain.

The new materials used in batteries are the reason for extension of the simple task of energy conversion into battery management systems, which also monitor the state of charge, depending on the rate of charge/ discharge etc. Given this situation, one might expect the system partitioning such that all critical analog functions are implemented in the same silicon. In this presentation about 1 Volt DC/DC conversion no voltage limitations due to technology are taken into account.

DC/DC conversion.

Basically there are 3 topologies commonly used for power supply adaptation. The most simple one is linear regulation, with a dissipating series power device used as power bleeder. More energy efficient solutions use either a capacitor as an intermediate energy storage device (capacitive charge transportation), or by using a coil for inductive voltage conversion. Each topology has unique properties which makes it best suited for certain applications.

The linear regulator is constructed using a control circuit, driving an output transistor which takes the excess supply voltage. When a P-type output transistor is used, mostly a large, external capacitor is needed to obtain a low output impedance at high frequencies, and to obtain a stable control loop.Fig.6

When supply has to be generated for a large digital core, a low output impedance at high frequencies is essential, while the load impedance can vary between fully capacitive in sleep mode, to a load requiring clocksynchronous large peak currents in parallel with a capacitance.

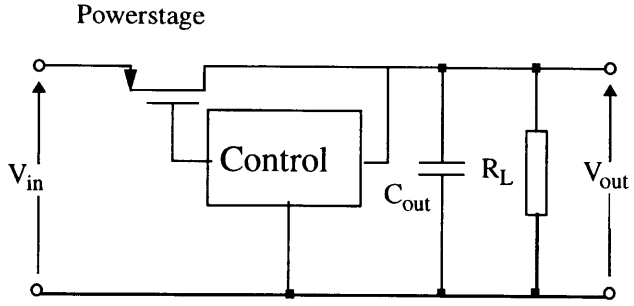


Fig. 7. Linear regulator with external buffer capacitor.

An other approach is to use an N-type voltage follower output stage. This configuration shows a relatively low output impedance at high frequencies but requires a voltage which is at least one Vgs above the output voltage to bias the output transistor properly. It also has stability problems with a purely capacitive load. Both problems have been solved (Ref. [2].), by using a bootstrap configuration to generate the gate bias voltage, and by using a separate source follower in the control loop, Fig. 8. This allows operation over a wide voltage range without external capacitor, which makes it better suited for integration.

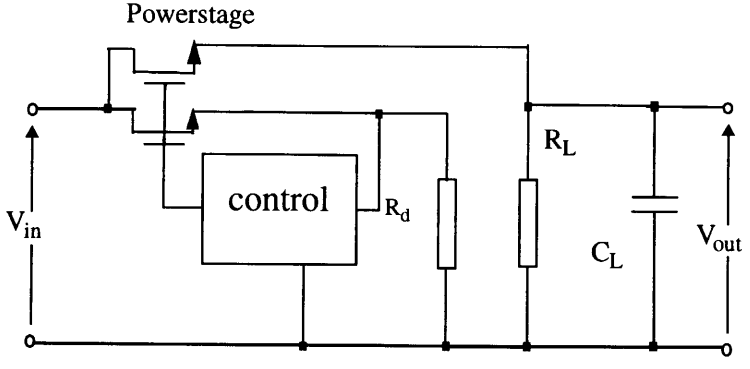


Fig. 8. Linear regulator with source follower output.

Capacitive power conversion.

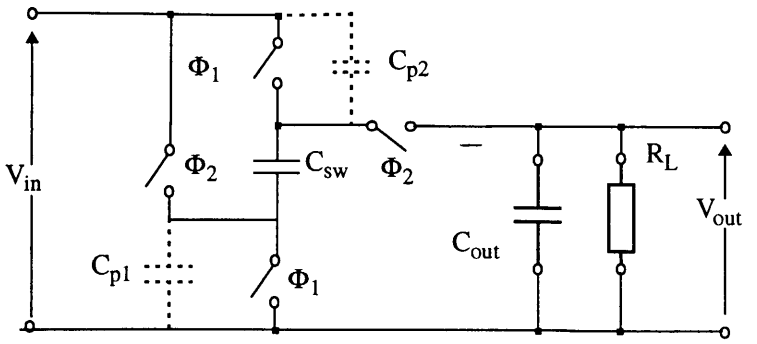


Fig. 9. Capacitive voltage doubler.

The classical capacitive converter is the voltage doubler as shown in Fig. 9. During Φ_1 capacitor Csw is charged to Vin, while during Φ_2 it is put in series with Vin and partially discharged into the buffer capacitor Cout. When the capacitor Csw is made on-chip, efficiency is already limited by the charging current through the parasitic capacitors Cp1 and Cp2, which in practice are about 10% or more of the capacitor Csw, excluding the parasitic charges in the switches. For high efficiency an external capacitor must be used Ref.[3]. In this case efficiency is depending on the charged/discharged voltage ratio of the switched capacitor.

If Vout = (2-Δ) Vin

then η=Pout/Pin = 1-Δ/2

Fig. 10 gives the efficiency of a 10 nF capacitor, switched at 1Mhz with a total series resistance (battery and switches) of 10 Ohms and a parasitic capacitance of the switches of 50 pF.

However with rechargeable batteries the voltage divider shown in Fig. 11 will be more appropriate. During Φ1 capacitor Csw and Cout are charged in series, while during Φ2 Csw and Cout are switched in parallel. Average output voltage over Cout will be Vin/2. However as the output voltage of the battery is not constant a converter with variable voltage ratio is needed.

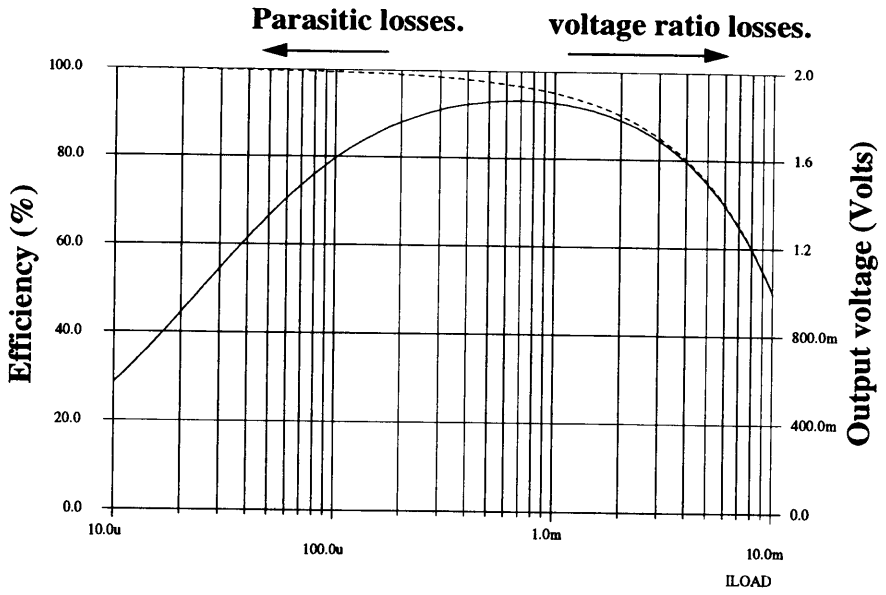


Fig. 10. Efficiency of a capacitive voltage doubler.

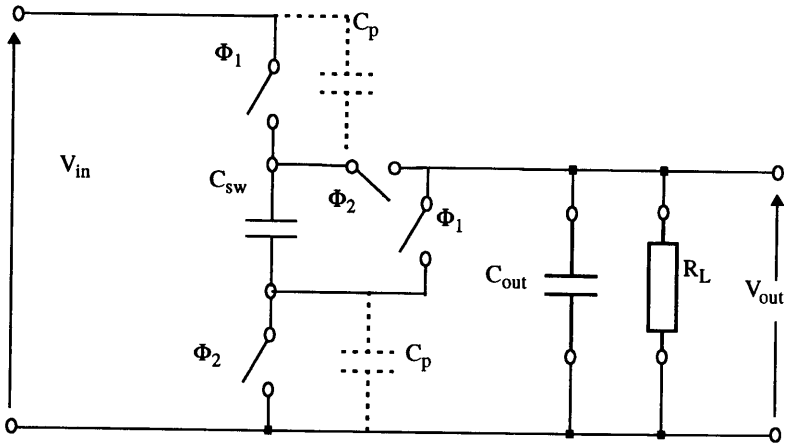


Fig. 11. Capacitive voltage divider.

Variable voltage ratios can be obtained by using a feedback configuration as shown in Fig. 10. When used with a clocked comparator its operates in a feedback-loop with a passive integrator Cout, and Csw acting as switched capacitor feedback. The output voltage over Cout is controlled by pulse density modulation. Deviation from the 2 to 1 voltage ratio will increase the voltage variation over Csw, so conversion efficiency will decrease.

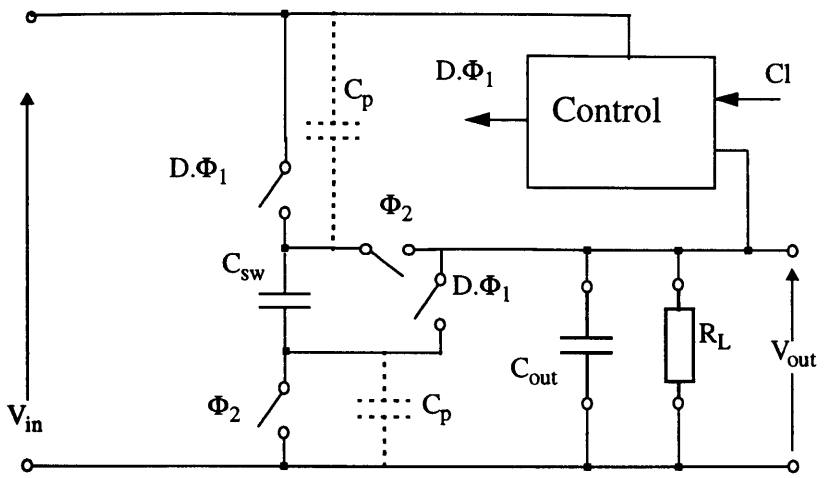


Fig. 12. Voltage regulator with pulse-density control.

In conclusion we can say that capacitive conversion has a limited efficiency when used with varying input voltages and varying loads, which makes it less suited for battery conversion applications.

Inductive DC/DC up-converter

The basic circuit is shown in Fig. 13. Energy from the input source is stored in the coil L_1 via switch S_1 and transferred to the output capacitor C_{out} via switch S_2 or diode D_1.
This converter can be used for up-conversion from a 1 V battery. With this converter, parts of the system which require a higher supply voltage can be supplied. The maximum voltage is limited by the used IC-process since the switches operate at that voltage. For switch S1 an NMOS transistor can be used, for switch S2 a PMOS transistor.

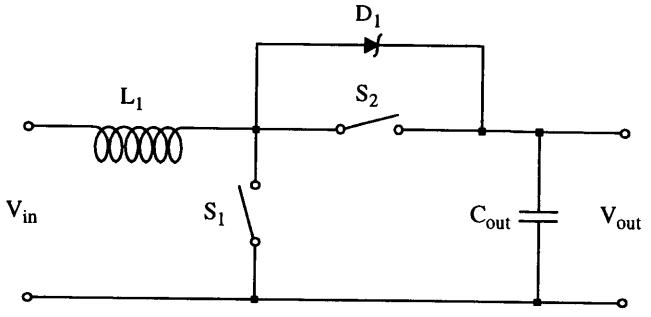


Fig. 13. Basic circuit DC/DC up-converter

Inductive DC/DC down-converter

For 1 V applications on a 3 V battery a DC/DC down-converter is required. The principle is equivalent to the up-converter, input and output are exchanged however. Figure 14 shows the circuit.

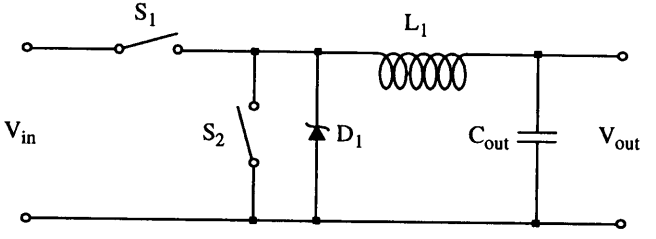


Fig. 14. Basic circuit DC/DC down-converter

Energy from the input source is stored in the coil L_1 and supplied to the output capacitor C_{out} via switch S_1. After that, remaining energy in the coil is transferred to the output capacitor C_{out} via switch S_2 or diode D_1. This converter can be used for down-conversion from a 3 V battery. With this converter, parts of the system can be supplied with a lower voltage. The battery voltage should not exceed the maximum voltage in the used IC-process since the switches operate at that voltage. For switch S1 a PMOS transistor and for switch S2 an NMOS transistor can be used.

Maximum output power

An important issue is the low series resistance of the switches. Since a low battery voltage is used a low total series resistance of the converter is necessary for an efficient DC/DC conversion at high output power. The maximum available output power of the DC/DC-converter is:.

Maximum output power: $$P_{out,\,max} = \frac{V_{in}^2}{4R}$$

This relation can be derived by modelling the DC/DC-converter as a voltage source V_{in} with a total series resistance R and applying Thevenin's laws to it. At maximum output power the efficiency is only 50%. For an output power of 1 W from a 1 V battery the maximal total series resistance R is 0.25 Ohm. This includes the losses in the battery, the coil, the output capacitor and the switches. For high efficiency this total series resistance should be small.

Discontinuous mode and continuous mode conversion

By means of the waveforms of the inductor current, which is also the battery current in case of up conversion, two often used discontinuous mode techniques are described: Pulse Frequency Modulation (PFM) and Pulse Width Modulation (PWM). Further continuous mode conversion with Pulse Width Modulation is described. Equivalent principles are valid for down conversion.

Pulse Frequency Modulation in discontinuous mode

The on-time, during which energy is stored as magnetic flux in the inductor, or the maximum coil current is fixed. The switching frequency depends on the load and the input voltage. As an example fig. 15 shows that the switching frequency increases for increasing load. Because in PFM the frequency is fully variable, solving possible EMI may be complicated.

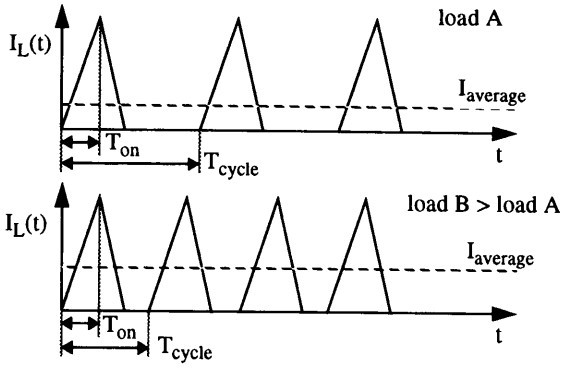


Fig. 15. Inductor current waveforms for PFM converter

Pulse Width Modulation in discontinuous mode

The switching frequency is fixed. The duty cycle $\{T_{on}/T_{cycle}\}$ depends on the load and the input voltage. As an example, fig 16. shows that the duty cycle increases as the load increases. As the switching frequency remains constant, also the capacitive switching losses do. Therefore PWM usually causes bad efficiency at light loads.

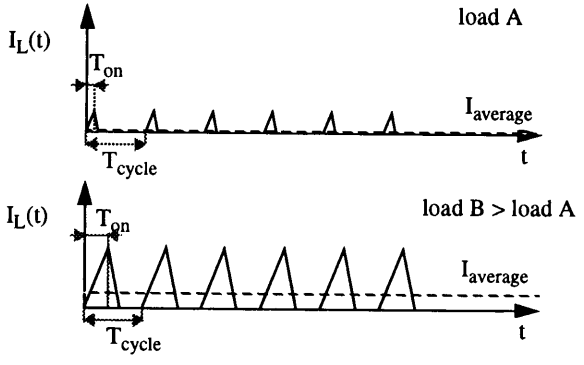


Fig. 16. Inductor current waveforms for PWM converter.

Pulse Width Modulation in continuous mode

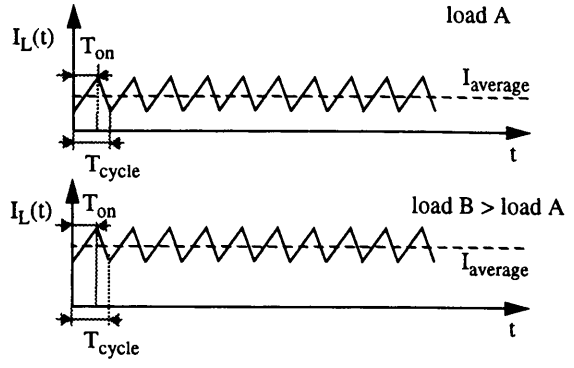


Fig. 17. Inductor current waveforms for converter in continuous mode.

In continuous mode conduction the switching frequency is free running or fixed. The duty cycle $\{T_{on}/T_{cycle}\}$ depends on the load and the input voltage. The duty cycle increases as the load increases or the input voltage decreases. With zero load the relation between dutycycle and input and output voltage for an upconverter is:

$$Dutycycle = \frac{T_{on}}{T_{cycle}} = \frac{V_{out} - V_{in}}{V_{out}}$$

Power losses

For a high efficiency it is important to analyse the losses in the converter. The main resistive $I^2 \cdot R$ losses are caused by the core and resistive losses of the coil, the MOS switches and the ESR of the output capacitor. Switching $f \cdot C \cdot V^2$ losses are caused by the parasitic capacitance. The circuit in fig. 18 shows the losses in the DC/DC up-converter.

The $I^2 \cdot R$ losses in continuous mode are substantially lower compared to $I^2 \cdot R$ losses in discontinuous mode. For the ease of explanation we project the resistive losses in the converter into one resistive element R.

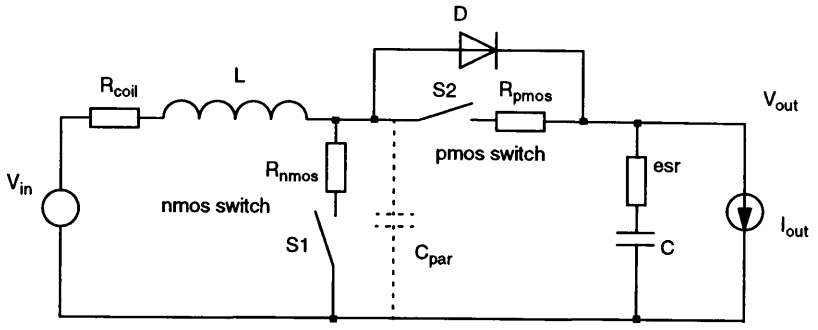


Fig. 18. Power losses in the DC/DC up-converter

For calculation of $I^2 \cdot R$ losses in discontinuous mode and continuous mode fig. 18 is used. In order to compare the losses between discontinuous and continuous mode operation of the DC/DC-converter, fig. 19 shows waveforms with the same mean current.

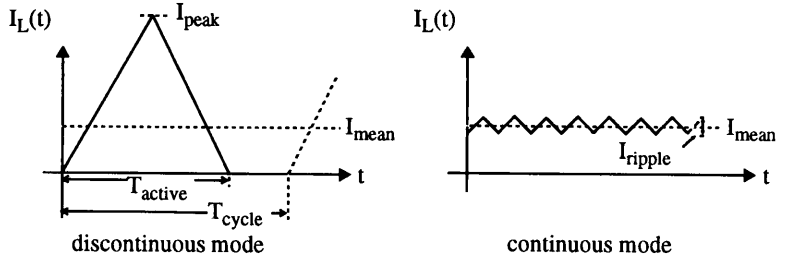


Fig. 19. Discontinuous mode and continuous mode compared

The $R \cdot I^2$ losses in discontinuous mode are:

$$P_{losses,\,discontinuous} = R \cdot \frac{4}{3} \cdot I_{mean}^2 \cdot \frac{T_{cycle}}{T_{active}}$$

The $R \cdot I^2$ losses in continuous mode are:

$$P_{losses,\,continuous} = R \cdot \left((I_{mean})^2 + \frac{1}{3} \left(\frac{I_{ripple}}{2} \right)^2 \right)$$

In general, continuous mode operation of the DC/DC-converter implicates higher switching frequencies than discontinuous mode operation. Resulting $f \cdot C \cdot V^2$ losses may possibly overrate the lower $I^2 \cdot R$ losses in continuous mode. Especially for small loads, when $I^2 \cdot R$ losses are small, discontinuous mode operation is often more appropriate.

Multi-mode operation

In case of an application where small loads and large loads are used a converter should be able to switch between discontinuous mode operation and continuous mode operation. This so-called multi-mode operation can guarantee a high efficiency over a large power range Ref.[4]. For the decision between the two modes information from the application can be used or the output current can be measured. In fig. 20 for a typical application the efficiency is plotted as function of the load current. In this example a battery voltage of 1.2V is boosted up to 2.4V. The switching frequency is 200kHz, the coil inductance is 330uH, the total series resistance of coil and switches is 1.5 Ohm and the parasitic capacitance is 200pF.

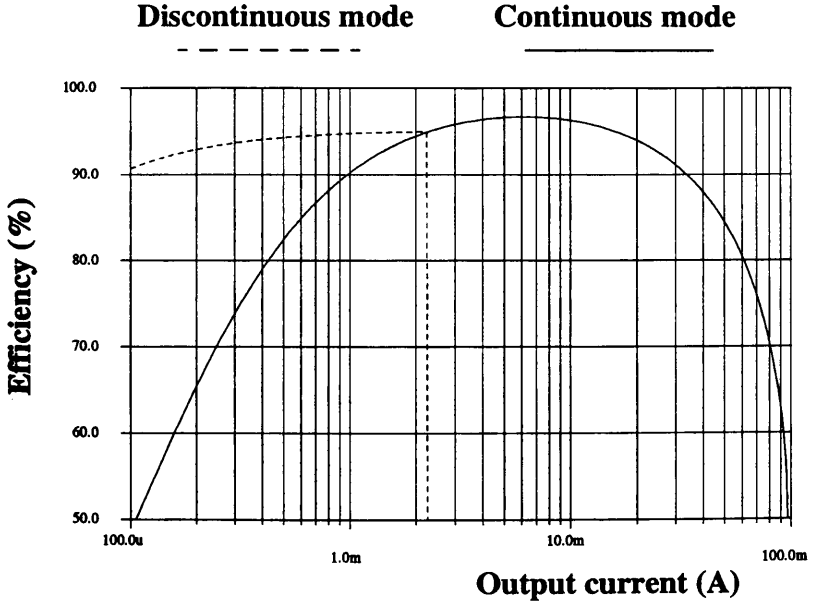


Fig. 20. Efficiency of inductive converter in multi-mode operation

In discontinuous mode Pulse Frequency Modulation is used. For the curve discontinuous the pulse amplitude equals the current ripple in continuous mode. The maximum output current is 2 mA in discontinuous mode and 100 mA in continuous mode. From the plot it is clear that for a high efficiency in this application discontinuous mode should be used for load currents from 0 to 2 mA, for load currents from 2 mA to 100 mA continuous mode should be used.

A higher maximum current in discontinuous mode can be realized by decreasing the inductance of the coil or increasing the switching period. Both result in a higher peak current and a lower efficiency over the complete range.

Example: combination of inductive and capacitive conversion

In this example we have an application which requires two output voltages, 0.9V for digital circuits and 1.8V for analog circuits. This application should run on a single battery with a voltage rage of 0.7-1.6V. This means that we need up-conversion for the 1.8V output and down-conversion for the 0.9V output.

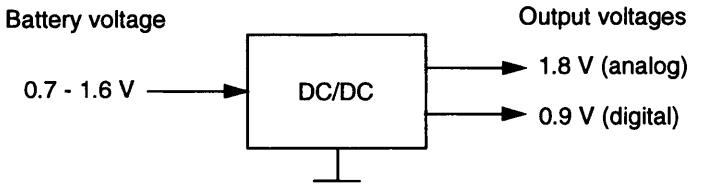


Fig. 21. Voltage requirements in application example

Due to the large voltage range of the battery voltage a capacitive converter gives a poor efficiency. A capacitive voltage doubler for the 1.8V output would result in an output voltage range of 1.4-3.2V. This causes unacceptable losses. For the 0.9V output capacitive down-conversion is not possible, this leaves voltage regulation down to 0.9V. This also causes unacceptable losses.

An inductive converter can give a high efficiency at the complete battery voltage range. An inductive up-converter can be used for the 1.8V output and an inductive down-converter for the 0.9V. With this solution high conversion efficiency can be realized. A disadvantage is that 2 coils are used.

A third option is a two step solution. First an inductive up-converter is used to boost up the battery voltage to 1.8V which is the analog output voltage. This stabilized 1.8V output voltage can be used as input voltage for a capacitive divider to create the 0.9V output. Since the division factor for this capacitive divider is exactly 2 a high conversion efficiency can be achieved. Fig. 22 shows this solution.

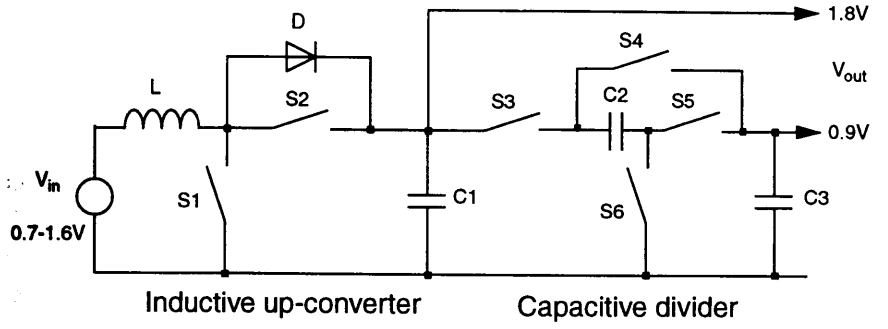


Fig. 22. Two step solution for 2 output voltages

For a battery voltage above 0.9V this solution can also be used in the opposite direction. First an inductive down-converter can be used for the 0.9V output. A capacitive voltage doubler can use this stabilized 0.9V to create the 1.8V output voltage. A disadvantage of the low intermediate voltage is that the local currents are higher which implies a lower series resistance of the switches to get the same conversion efficiency as the solution with inductive up-converter and capacitive divider.

Conclusions

For battery driven applications which need efficient use of energy a DC/DC converter is a must. As the output voltage of a battery varies strongly, only the inductive type shows sufficient efficiency. They can be used both as up or down converter from very low voltages onwards. Once a stable output voltage is obtained, capacitive converters can be used to derive various other supply voltages. Modern battery materials show output voltages which are not within the range of future VLSI processes. Either special options or a separate chip are needed, which also allows the analog designer to design his function for minimum power consumption.

References.

[1] SIA Semiconductor Industry Association, "The National Technology Roadmap for Semiconductors" 1997, p 46.
[2] G.W. den Besten and B. Nauta, " Embedded 5 V-to-3.3 V Voltage Regulator for Supplying Digital IC's in 3.3 V CMOS Technology.", Proceedings of ESSCIRC'97, pp.52-55, Sept. 1997.
[3] H. Neuteboom, B.M.J. Kup,and M. Janssens,"A DSP-Based Hearing Instrument IC.", IEEE Journal of Solid-State Circuits,Vol. 32, No. 11, Nov. 1997.
[4] F.J. Sluijs, K. Hart, W. Groeneveld, S. Haag, "Integrated DC/DC Converter with Digital Controller.", to be published on ISLPED 98 Conference.

Design and implementation of mixed-mode systems

Willy Sansen

Introduction

All integrated circuits are bound to contain both digital and analog subblocks in the future. Many different problems emerge. The involve substrate bounce and other substrate coupling effects, all deteriorating signal integrity. They require different design methodologies. Both aspects of mixed-signal design have been addressed in this session.

The first presentation, given by Bram Nauta, gives a number of design procedures to reduce the effect of substrate bounce. Apparently minor design issues such as the connection of a n-well of a pMOS differential pair, are shown to have a large impact on the performance of common analog circuits such as two band-gap circuits.

It is followed by a presentation by François Clément on substrate noise as well. He focuses on the electromagnetic effects of substrate coupling. He analyzes all coupling mechanisms and derives conclusions with respect to highly/lowly doped epi-layers and die-attachment techniques.

Tallis Blalack then gives a follow-up on design techniques to reduce the substrate coupling effects. Both "quieting the talker" as well as "desensitising the listener" are discussed. A nice example of enhanced isolation performance is added for sake of illustration.

The fourth paper is on design itself. It is given by Laurent Dugaujon and discusses the design issues and performance evaluation of a mixed-signal design combining a 1.2 GHz analog PLL with a 1M CMOS processor. Extreme power reduction leads to noise coupling, which can be successfully reduced by an appropriate layout style.

Design methodologies are as important in mixed-mode design as the design itself. This is why the two last presentation address such issues. The first one is brought by Nishath Verghese. He reviews modelling techniques for noise coupling in mixed/RF integrated circuits. They require modelling of the IC environment including the substrate, the package and the chip interconnects. The impact of this noise coupling is modelled using periodic transfer functions. Some real-life examples are added.

Finally Georges Gielen provides a top-down design design strategy of mixed analog-digital systems. Now attention has to be paid to system-level architectural design before proceeding to the detailed block design. He also gives an overview of the current techniques for analog circuit synthesis and mixed-signal layout assembly.

It can be concluded from these presentations that plenty of problems have to be solved towards mixed analog-digital design. On the other hand careful layout techniques embedded in a hierarchical design methodology allow to cope with most of them..

Substrate Bounce in Mixed-Mode CMOS ICs

Bram Nauta* and Gian Hoogzaad.

Philips Research Laboratories, Eindhoven, The Netherlands.

* Now with University of Twente, Enschede, The Netherlands

1 ABSTRACT

Substrate noise is one of the key problems in mixed analog/digital ICs. Although measures are known to reduce substrate noise, the noise will never be completely eliminated since this requires larger chip area and thus higher cost. Analog circuits on digital ICs simply have to be resistant to substrate noise. A general strategy is given which can be summarized as: the *supply* of the analog circuits must be referred to the substrate and the analog *signals* must be referred to a clean analog ground. Furthermore several design constraints are given to minimize the effect of substrate noise on analog. Two bandgap circuits are discussed and it is shown that apparently minor design issues, such as the connection of an n-well of a PMOS differential pair, can have large impact on the substrate sensitivity of this circuit. This has been verified by measurements.

2 INTRODUCTION

One of the major challenges in mixed signal IC design is to deal with substrate noise and crosstalk problems. Especially in CMOS technology the problems are serious, since both analog and digital share the same - low ohmic - substrate. As the switching speed and packing density of digital CMOS increases, and the supply voltage drops, it's more and more difficult to design analog modules with good performance. For example it is not trivial to design a 10 bit video ADC (1LSB = 1mV) or to design a low-jitter PLL while there is 300mV substrate noise.

This paper[1] describes briefly the origin of substrate noise (section 3), followed by an example of a simple current source to illustrate the problems in analog (section 4). In section 5a general substrate strategy for analog is given and in section 6 a practical example is given on what can go wrong with the design of a simple CMOS bandgap reference. Finally conclusions are drawn.

3 SUBSTRATE NOISE

Figure 1 shows a well known cross section of a standard CMOS IC. The substrate is very low-ohmic (say 0.01 Ohm cm) on top of which an epi-layer is grown. This epi-layer is several micrometers thick and has a higher resistance (say 10 Ohm cm). On this low-ohmic substrate both analog and digital circuits are located, and due to the low substrate resistance this causes crosstalk problems. Now a brief explanation of the origin of the substrate noise is given.

In figure 2 schematically the typical power routing of a digital CMOS IC is shown. The digital part of the IC is here simplified to a (huge) inverter. The digital part has its own VDDD and VSSD pins. The CMOS logic is switching and this means current spikes through the VDDD and VSSD pins. The current spikes are due to (dis)charging of capacitances and short circuit current as present in CMOS gates [1]. In the digital standard cells normally substrate contacts are present for latch-up reasons. A single substrate contact has a resistance of several kOhms, however on a large IC with - say 500k - gates, the substrate is

1. This work has been published before at the , ETTCD'97 conference, August 1997. Budapest, Hungary.

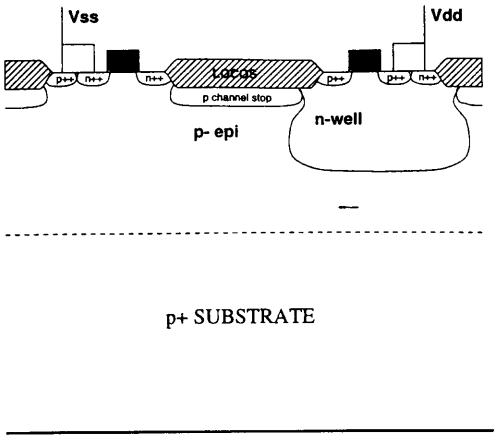


Figure 1: cross section of a p-substrate p-epi CMOS IC

very well connected to VSSD. The problem is now the self inductance of the VDDD and - more serious for analog - the self inductance of the VSSD pins. The self inductance of a single bondwire is typically 5nH and the resulting voltage drop over the VSSD bondwire is $V_{bondwire} = L_{vssd} * d_{Iss}/dt$. In a system the PCB (Printed Circuit Board) ground is usually the reference for analog and digital circuits and thus the substrate noise is equal to $V_{bondwire}$ [2].

If no precautions are taken then substrate noise can be several Volts. However in this case even pure digital circuits will not operate correctly. A practical method to limit substrate noise is to decouple the digital supplies with on-chip capacitors (as denoted with Cdd in figure 2). If these capacitors are large enough then the peak currents needed are drawn from these capacitors. A problem is that these capacitors need area, and are thus expensive. Furthermore the series resistance of these capacitors must be low (the RC time constants must be in the GHz range). Consequently the capacitors must be merged with the logic. Another problem with these capacitances is the LC-resonance effect with the inductance of the bondwire. Therefore a series resistance (as denoted with Rdd in figure 2) must be present to damp the oscillations [2]. Rdd can be a parasitic metal resistance in the VDDD path. Another option to limit substrate noise is to use many VSSD pins. Usually many VDDD pins are also needed for correct operation of the digital

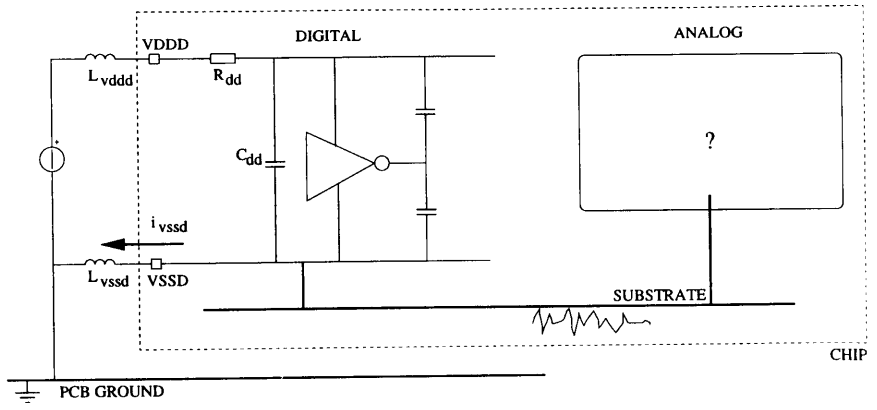


Figure 2: The origin of substrate noise: switching of digital circuits result in dI_VSSD/dt, and thus in a voltage drop over the VSSD bondwire.

part and separate supply pins for the digital I/O are used. Also lowering the digital supply voltage will help to reduce the substrate noise, however at the penalty of speed. It can also be shown that the internal digital supplies are oscillating in antiphase[3]. Disconnecting substrate contacts from the digital Vss and balancing the capacitance of the substrate to both supplies is a possibility to suppress groundbounce [3]. However this requires knowledge about existing capacitance within the circuit in any state. Furtermore digital librariy cells with a separate substrate rail are needed, (to prevent latch-up) which consumes more chip area.

Substrate noise can thus be limited, be it at the cost of money; either area, pins or dedicated libraries. Therefore substrate noise will never be reduced to the millivolt level. Always some $300mV_{pp}$ will be present, dictated by proper functioning of the digital part. Even so, in general the frequency content of the substrate noise is not known, because of multiple clock domains on the same IC.

Guard rings for shielding the analog part from the digital substrate have no effect in CMOS with a low-ohmic substrate [4]. It's the same as building a fence around your house to keep it safe from earthquakes: the substrate noise comes from the bottom of the analog circuit. Substrate noise will simply be there, and the analog circuitry must therefore simply be able to deal with this interference.

4 PROBLEMS IN ANALOG

In this section the problems in analog will be discussed. Since analog circuits share the same substrate these circuits will always be affected by substrate noise. three mechanisms can be distinguished;

1) *direct incoupling*. The - normally - high frequency substrate signal couples directly into the analog circuit. If the frequency content of the substrate signal is outside the signal band in the analog part, this needs not to be harmful as long as the analog circuit behaves like a linear circuit. This behaviour can be investigated in an AC simulation. Unfortunately true linear circuits are rare.

2) *demodulation*. If a substrate signal couples into an analog circuit which contains non-linear elements (which is normally the case) then demodulation can occur. Even if the frequency content of the substrate noise is outside the signal band, demodulation or AM detection can result in noise in the signal band of the analog circuit. An example of this is described in section 6, where substrate noise in a bandgap reference circuit results in a DC shift of the output voltage. Transient analyses are needed to tackle these problems.

3) *sampling*. In a mixed-signal IC normally an AD converter is present, which means a sampling operation on the analog signal. If substrate noise has coupled into the analog circuits preceding the AD converter (Clamp, AGC, filter, buffer, etc.) at a frequency outside the signal frequency band, this sampling operation will fold the substrate noise back into the signal band.

In conclusion it can be mentioned that analog modules and circuits must have a good rejection for substrate signals. The rest of this paper discusses how this can be achieved.

Consider a very simple analog circuit like the current source in figure 3a. The current source is supplied with a clean analog VSS (analog ground) and the gate has a parasitic capacitance Cgate. Due to the na-

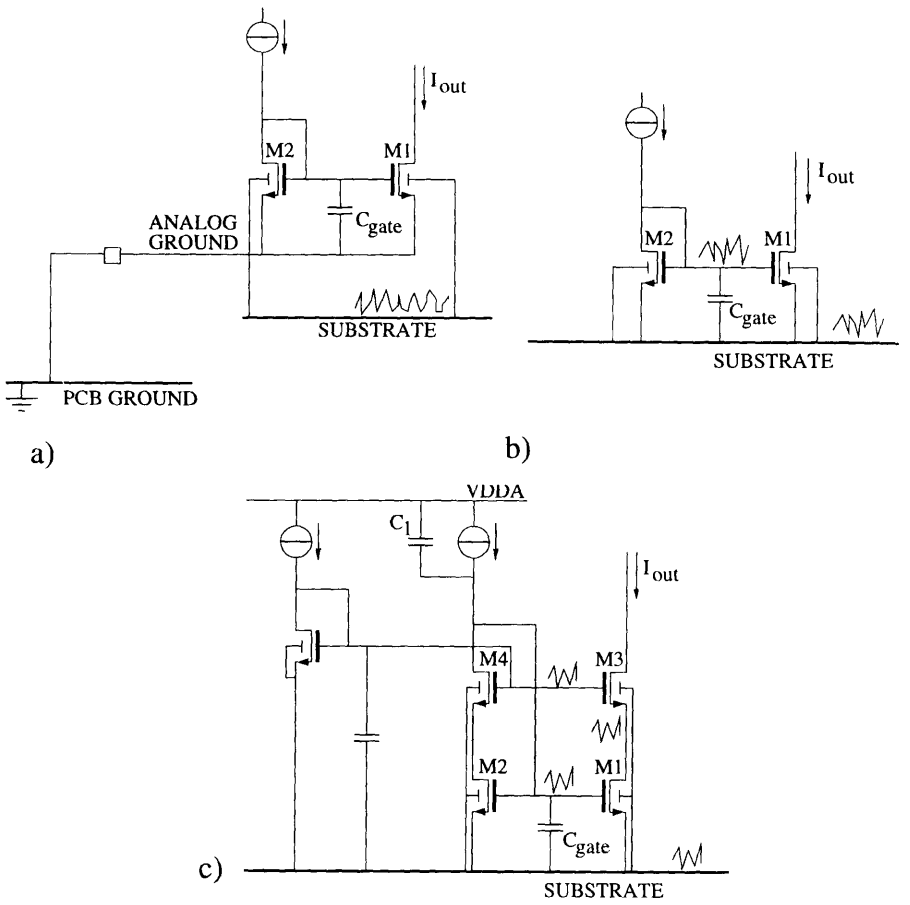


figure 3: example to illustrate the effect of substrate noise on analog circuits a) current source, realized by a current mirror, with a clean VSSA. b) improvement of the current source of 3a) with VSSA connected to the substrate c) improvement of the current source of 3b) by means of small cascode transistors

ture of the CMOS process, the backgate of M1 is coupled to the substrate. For high frequencies the Vgs of the MOST is fixed due to Cgate, and the substrate noise on the backgate directly modulates the drain current of M1. It is therefore better not to use a clean analog ground, but to use a VSS, equal to local substrate voltage. This is shown in figure 3b. The AC content of the substrate noise is now present on the gate, source and backgate of M1. The drain is not coupled to the substrate and the substrate noise couples only via the drain-bulk and drain-source admittances. This is better than coupling via the backgate. For noise and matching reasons the dimensions of M1 and M2 are normally far from the minimum as dictated by the technology, resulting in still a large drain-bulk capacitance of M1 and M2. For this reason the current source is often provided with cascode transistors. Figure 3c illustrates how these cascode transistors should be biased in order to have small substrate sensitivity. Now the drain, gate, source and backgate of M1 and the gate, source and backgate of M3 have the same AC (substrate) signals. The output impedance of the current source is high including the capacitance since M3 can have small dimensions. Note that the capacitance C1 should be kept small for good substrate rejection.

Thus by choosing the right references, in this case the substrate for NMOS transistors, performance can be improved. In the next section a more general approach will be given.

5 STRATEGY FOR ANALOG

In this section several design rules are given to make circuits for analog signal processing less sensitive for substrate noise.

- Use NMOS transistors only as DC current sources. These NMOS transistors should be referred to the substrate and not to a clean VSS.
- Use PMOS transistors for signal handling: i.e. PMOS as differential pairs and signal handling current mirrors. PMOS transistors have an n-well which can be used to shield the transistor from the substrate. Be sure to put enough well contacts to the VDDA, and be aware of series resistances in the well, which are normally not properly modeled.
- Make analog circuits fully differential, with a possibly clean common mode level. A common mode control circuit must suppress interference. Matching and large signal behaviour is still limiting the effect of balancing.
- Use shielding of resistors and wires only for signals not referred to the substrate. Shielding can be done with n-well, poly or one of the lower

metal layers, connected to VDDA. Be careful with series resistances in n-well and high-ohmic poly, since the RC time constant will limit the effect of shielding. Bondpads carrying analog signals can be shielded with n-well.

The remaining problem is now the analog interface. Analog signals are usually single-ended and defined w.r.t. the "clean" PCB ground. Inside the IC large parts of the analog circuits are referred to the substrate (all NMOS) and the signals are differential. This makes interfacing a serious matter.

Figure 4 shows the recommended supply and reference routing for a mixed-signal IC. The digital part has several VDDD and VSSD pins, and the substrate is contacted to VSSD in the digital part as mentioned before. The analog part has a separate VDDA, since VDDD will be polluted by the digital. The VSS of analog (VSSA) should be connected to the substrate with enough substrate contacts, and should contain the same signal as the substrate. VSSA is thus NOT clean w.r.t. PCB ground, and there is actually no difference between VSSA and VSSD. In order to interface with the outside world of the IC, a clean reference signal on chip is needed. This signal is denoted as "analog ground". Analog ground is connected to PCB ground via a bondwire and thus di/dt of this bondwire must be (almost) zero. The pin of analog ground therefore may only carry DC signals or signals with relatively low frequencies, depending on the demands. The analog ground wires on chip, must be shielded from the substrate (with n-well or lower interconnect layers, connected to VDDA) and no unwanted signals ought to couple into analog ground.

The input signal is referred to PCB ground and is fed into a first stage on the IC, which is, in the example of figure 4, a transconductance amplifier. Important is that this first stage has analog ground as a reference for the signal. The output of the first stage is preferably differential, in order to be less sensitive to substrate or supply noise. The common mode level of the differential signals should be clean. After optional analog preprocessing the analog signal can be converted to the digital domain for further signal processing. The signal can be converted into the analog domain via a DA converter and may be followed by postprocessing functions, such as smooth filtering. The differential

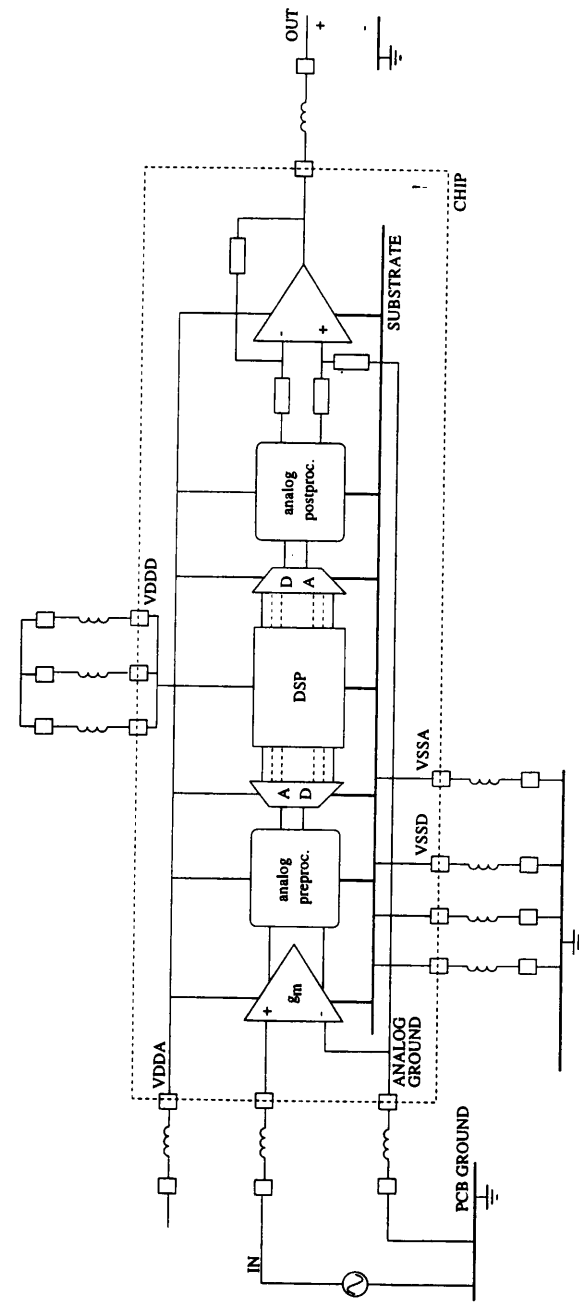


Figure 4. recommended supply and reference routing for a mixed signal IC.

signal must be converted to a single-ended signal referred to analog ground. This is also illustrated in figure 4. It would even be better if the "outside world" signals of the IC are differential, however this is often not feasible for cost reasons.

Summarizing it can be said that: the supply of the circuits are referred to the substrate and the signals are referred to a clean analog ground.

6 EXAMPLES

In this section a practical example of the effect of substrate noise on a bandgap circuit is illustrated. Figure 5a shows a well known bandgap voltage reference [5], generating a reference voltage V_{bg} with respect to the substrate. This means that the output voltage of the bandgap circuit V_{bg} is constant w.r.t. the substrate. (and thus not constant w.r.t. analog ground, if there is substrate noise). The core of the circuit is R1,R2, R3, Q1 and Q2. The folded cascode amplifier (all MOSFETS) is responsible for the proper feedback, needed for correct bandgap operation. This bandgap is denoted here as bandgap "A"

For correct operation of bandgap A, all NMOS transistors are referred to the substrate, as illustrated in the example of a current mirror of figure 3b. The bases of the parasitic vertical pnp transistors are also connected to the substrate. In order to keep the feedback loop stable a capacitor Cc has been added. A further advantage of Cc is that the gate-source voltage of M9 is low-pass filtered, and thus filters more or less incoupling substrate noise.

It appeared to be very important during the evaluations of these types of bandgap circuits how the n-well of the input differential pair is connected. One can connect the n-well to VDDA or to the common source node of the differential pair. Figure 6a shows the simulation results of the two possibilities. The figure shows the bandgap voltage versus time while a 10MHz clock signal of 400mVpp is present on the substrate. If the n-well is connected to the VDDA the bandgap voltage starts drifting away from the nominal value (figure 6a-curve 1). Note that even if the ripple is low-pass filtered, the DC value is not correct and will depend on the substrate noise which is highly unwanted. The explanation is that the base and emitter of Q1 and Q2 follow the substrate noise. So

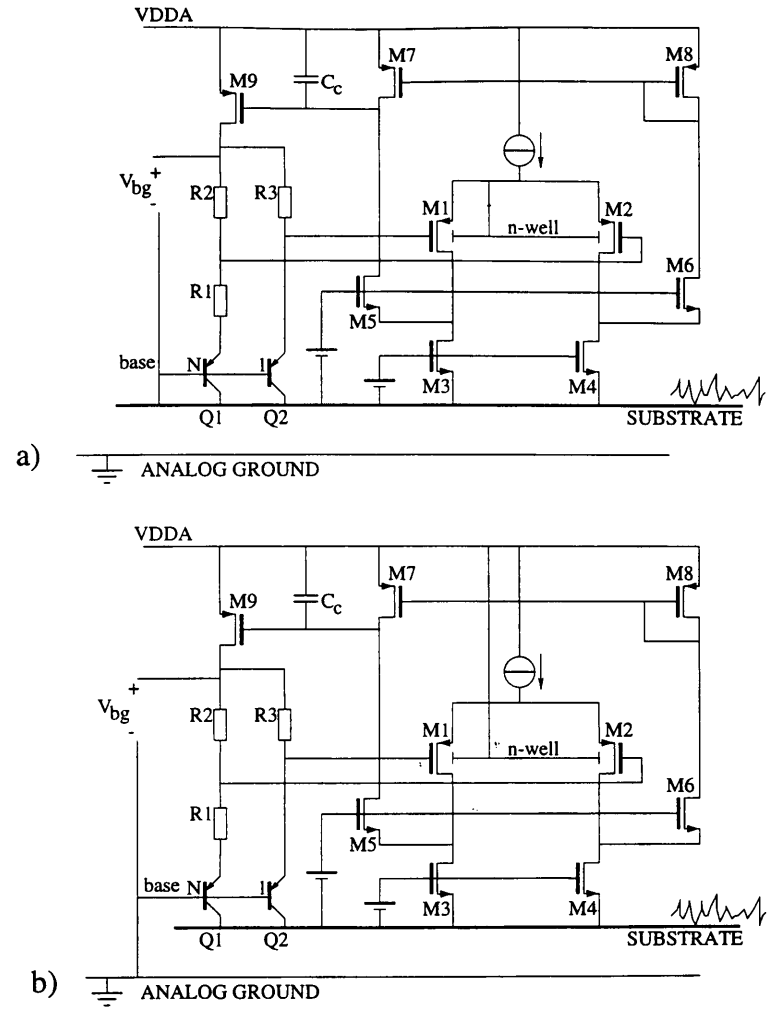


figure 5. a) bandgap type A, generates a reference voltage w.r.t. substrate b) bandgap type B, generates a reference voltage w.r.t. analog ground.

do the gates, sources and drains of M1 and M2. If the n-well is connected to the VDDA, the source-well capacitances of M1 and M2 conduct a current related to the substrate noise and thus modulate the currents of M1 and M2. Due to non-linear effects (see section 4) this results in DC shift of the bandgap voltage. If the n-well is connected to

the common source, as shown in figure 5a, then all terminals (drain , gate, source and well) of M1 and M2 have the same substrate-related signal and the modulation does not occur. Important is that the capacitance from common source to VDDA is held small. The results of the simulations with n-well connected to the common source node is shown in figure 6a, curve 2.

The other bandgap - type B - is shown in figure 5b. The reference voltage is now wanted w.r.t. the "clean" analog ground, and the bases of Q1 and Q2 are now connected to this analog ground. If the n-well of the differential pair M1, M2 would be connected to the common source node, then the bandgap voltage would drift away as shown in figure 6b, curve 1. This is because the n-well to substrate capacitance picks up the substrate noise and pollutes the common source node, which should not follow the substrate noise in this case. The result is again modulation of the currents in M1 and M2. The correct connection of the n-well in bandgap B is to the VDDA. Now the main terminals of M1 and M2 (gate, source, well) are "clean". The noise picked up by the n-well is routed towards VDDA. The resulting simulations are given in figure 6b, curve 2.

The simulations have been verified by measurements on a large mixed-signal IC. Figure 7 shows a measured reference voltage, derived from a bandgap circuit. The reference voltage is plotted as a function of the clock frequency of the IC. For the wrong n-well connection the reference shows a large deviation around 10MHz clock for 3 samples of the IC (curves 1). We modified the connection with a FIB (Focussed Ion Beam) station on the 3 samples and the result is clear! (curves 2)

7 CONCLUSIONS

Substrate noise is one of the key problems in mixed analog/digital ICs. Although measures are known to reduce substrate noise, the noise will never be completely eliminated for cost reasons. Analog circuits on digital ICs simply have to be resistant to substrate noise. A general strategy has been given which can be summarized as: the analog circuits must be referred to the substrate and the analog signals are referred to a clean analog ground. Furthermore several design constraints are given to minimize the effect of substrate noise on analog. Two

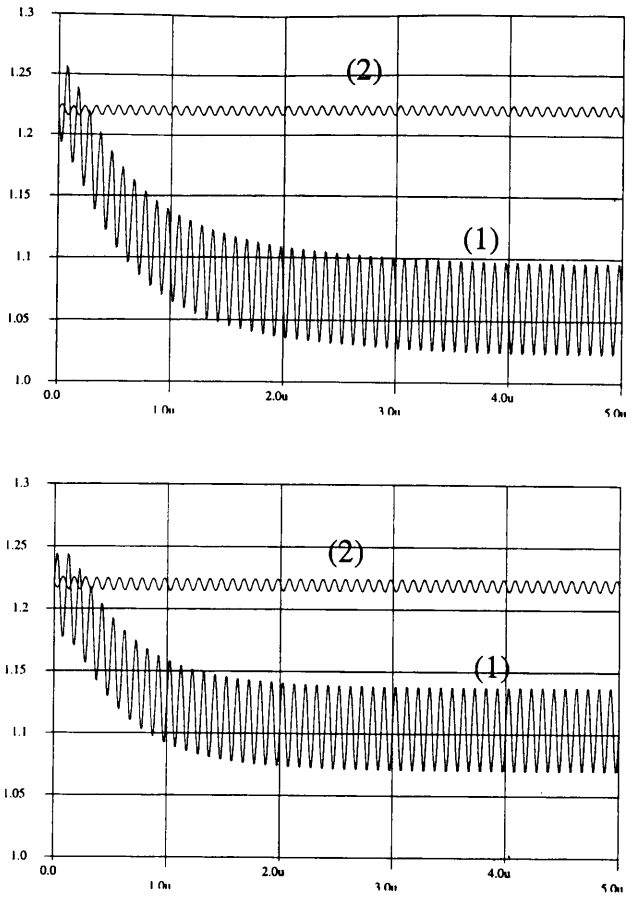


figure 6 Simulation results of the bandgap reference. a) Bandgap circuit A, with output referred to the substrate. Curve 1 with n-well to VDDA (wrong) , curve 2 with n-well to common source node of differential pair (correct). b) Bandgap circuit B, with output referred to the analog ground Curve 1 with n-well to common source node of differential pair (wrong), curve 2 with n-well to VDDA (correct).

bandgap circuits have been discussed and it has been shown that apparently minor design issues, such as the connection of an n-well of a PMOS differential pair, can have large impact on the substrate sensi-

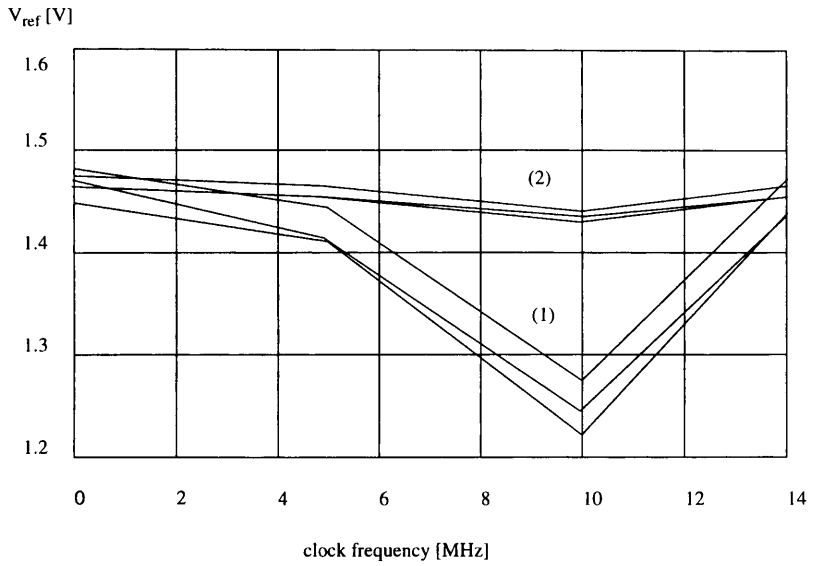


Figure 7.Measured reference voltage, derived from a bandgap circuit as a function of the clock frequency of the IC. curves (1) wrong n-well connection, curve (2) correct n-well connection.

tivity of this circuit. This has been verified with measurements.

Acknowledgment: The techniques and results described in this paper are a result of the contributions of many colleagues within Philips Research and Philips Semiconductors. We wish to thank these colleagues for their contributions. Special thanks are given to W. Relyveld for doing the bandgap simulations.

8 References

[1] H.J.M. Veendrick, "Short-Circuit Dissipation of Static CMOS Circuitry and its Impact on the Design of Buffer Circuits", IEEE Journal of Solid-State Circuits, Vol. SC-19, August 1984, pp 468-473.

[2] D.W.J. Groeneveld, "Groundbounce in CMOS", Digest of technical papers, workshop on Advances in Analog Circuit design, Leuven 1993.

[3] T.J. Schmerbeck, R.A. Richetta and L.D. Smith, "A 27MHz Mixed Analog/Digital Magnetic Recording DSP Using Partial Response Signalling with Maximum Likelihood Detection", Digest of technical papers, ISSCC 1991,, Feb 1991, pp 136-137.

[4] D.K. Su, M.J. Loinaz, S. Masui, B.A. Wooley, "Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal Integrated Circuits", IEEE Journal of Solid-State Circuits, Vol. 28, No. 4, April 1993, pp420-430.

[5] K.E. Kuijk, "A reference voltage source" IEEE Journal of Solid-State Circuits, vol SC-8, pp 222-226, June 1973..

TECHNOLOGY IMPACTS ON SUBSTRATE NOISE

François J. R. Clément
EPFL, Lausanne, Switzerland

ABSTRACT

The electromagnetic substrate behavior of integrated circuits (IC) is reviewed and the significant parasitic phenomenons are presented. The technology impact is examined from three complementary point of view. The respective influence of the lightly-doped and epitaxial wafers is exposed. The fabrication process steps changing the substrate characteristics are addressed for CMOS and bipolar technologies. The die-attachment is considered as a mean to reduce substrate parasitics.

1 INTRODUCTION

With the increasing complexity of mixed digital-analog designs, and with the decreasing feature size of current technologies, taking into account parasitic coupling through the substrate has become a key issue in reducing time-to-market of new circuits [1, 2]. Understanding the wide variety of parasitic effects represent a major concern. Additionally, the increasing impact of IC fabrication technology on parasitic behaviors is especially challenging in terms of circuit design [7-13] and modeling [14-26].

The fabrication of an integrated circuit entails a long and costly process. Current processes use planar technology [3, 5]. In this method, an initial bulk material is altered through successive steps to create electrical devices. The bulk material is either lightly doped — doping concentration around 10^{15} [cm^{-3}] — or epitaxial — a lightly doped layer built on top of an

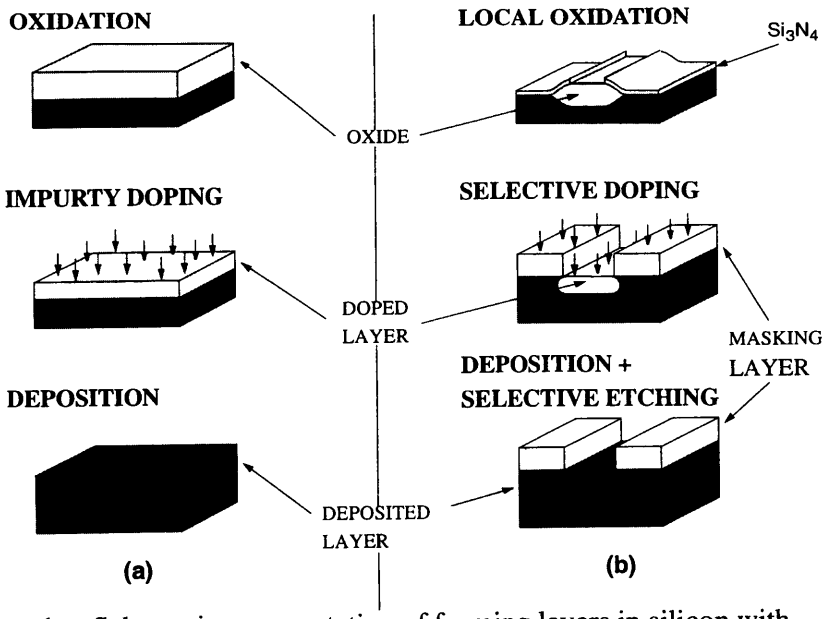


Figure 1: Schematic representation of forming layers in silicon with planar technology: (a) uniform, (b) selective

heavily doped material with a concentration of 10^{19} [cm^{-3}] —, according to the type of wafer introduced in the fabrication facility. The main process steps include growth, oxidation, deposition, implantation and diffusion. A set of masks is used to determine which areas of the circuit are concerned by each process step. Figure 1 illustrates how the original substrate is modified uniformly or selectively using masks.

The masks are computed from graphic geometries known as the «layout». Hence, as shown in figure 2, the vertical structure of an integrated circuit at a particular surface point (x, y) is completely determined by the process and by the layout.

Today's technologies have become very complex to reduce size and to preserve characteristics of integrated devices. There exists many different processes from the most common to the more specific suitable for particular fields of applications like high-speed or high-voltage.

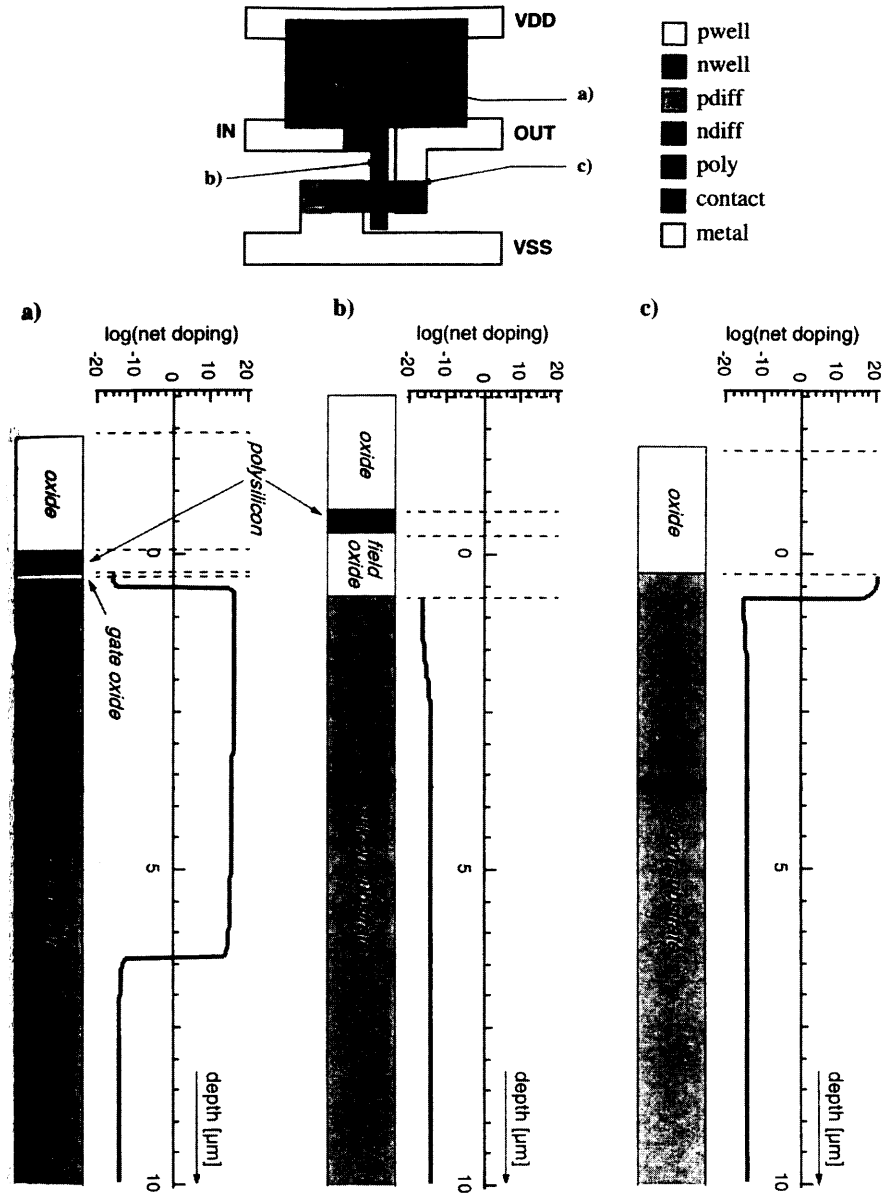


Figure 2: Layout of CMOS inverter and corresponding IC vertical structures for a standard process

The number of fabrication steps varies from 100 for the simplest technologies to 400 for the most complex. The corresponding number of masks varies as well between 10 and 40. Hence, for a specific technology, the vertical doping profile of the substrate is completely determined by the combination of the masks. The substrate characteristics will change significantly with the wafer type, the fabrication process as well as the die-attachment technique.

Parasitic currents will flow differently according to the substrate structure and the die-attachment. The substrate physics addressed in Section 2 determines simplification assumptions and relevant substrate characteristics from semiconductor physics. The significant substrate behaviors and their corresponding models are summarized in Section 3. Section 4 reviews each type of wafer and presents their corresponding models and characteristics. The specifics of each fabrication process is examined in Section 5.

2 Substrate Physics

The following set of assumptions is justified to simplify the substrate characterization:

H2.1 *The well-substrate junctions are reverse biased. The violation of this hypothesis would result in a short-circuit of the power supply.*

H2.2 *Specific semiconductor behavior, like surface inversion, concerns device modeling and will not be considered here.*

H2.3 *No latch-up occurs in the substrate during normal function.*

H2.4 *The maximum wavelength of the magnetic field is much greater than the die size. Therefore, inductive coupling is neglected.*

2.1 Resistive effect

Inside a doped semiconductor, the conductivity σ is given by:

$$\sigma = q(n\mu_n + p\mu_p) \tag{2.1}$$

where q is the electronic charge, μ_n and μ_p represents the mobility of the n- and p-carriers, n and p stand for the n respectively p carrier density. The μ_n and μ_p parameters vary as functions of the total semiconductor doping,

and temperature. Illustration of these dependencies can be found in [3] (pp.138-140) and [2] (p.48). Furthermore, mobilities are limited by the saturation velocity of the carriers. Nevertheless, this limitation is reached for high electric field which should not be reached in the substrate during normal operation.

2.2 Capacitive effect

Silicon has a relative dielectric constant:

$$\varepsilon_{r_{Si}} = 11,7 \tag{2.2}$$

which leads to the absolute dielectric constant:

$$\varepsilon_{Si} = \varepsilon_{r_{Si}}\varepsilon_0 = 1,035 \qquad \left[\frac{pF}{cm}\right] \tag{2.3}$$

Capacitive and resistive effects occur throughout the substrate. However, for substrates doped homogeneously, the capacitance of the substrate can be neglected in the operating frequency range. In the frequency domain the equivalent impedance Y_s for a piece of substrate such as illustrated in figure 3 is given by (2.4).

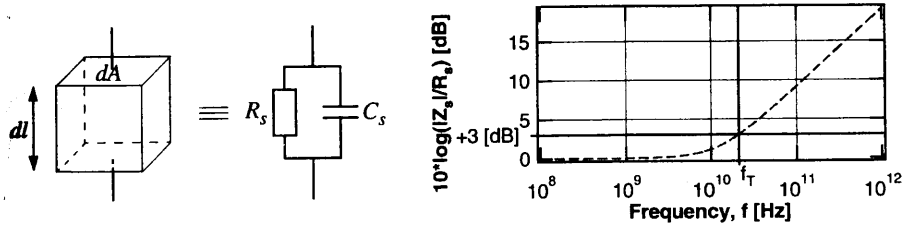


Figure 3: RC model for a piece of homogeneous substrate

$$Y_s = \frac{1 + pR_sC_s}{R_s} = \frac{1 + j\omega T_s}{R_s} \tag{2.4}$$

Therefore introducing (2.1) and (2.3) leads to substrate time constant T_s which no longer relates to the piece dimensions as shown in equation (2.5).

$$T_s = R_sC_s = \frac{\rho_s dl}{dA} \cdot \frac{\varepsilon_s dA}{dl} = \frac{\varepsilon_0\varepsilon_{r_{Si}}}{q(\mu_p p + \mu_n n)} \tag{2.5}$$

For low frequencies, substrate resistance R_s is more important and the associated capacitance C_s can be neglected. As pulsation ω increases, capacitive effect rises to become equal to the resistive effect at the cut-off frequency f_T defined by:

$$\frac{1}{R_s} = \omega_T C_s = 2\pi f_T C_s \Rightarrow f_T = \frac{1}{2\pi T_s} = \frac{q(\mu_p p + \mu_n n)}{2\pi\varepsilon_0 \varepsilon_{r_{Si}}} \tag{2.6}$$

The minimum f_T is achieved for lightly-doped p-type substrate because mobility is lower for holes than for electrons. A normal initial carrier concentration of 10^{15} [cm^{-3}] yields a minimum cut-off frequency:

$$f_{T_{min}} = \frac{q\mu_p p}{2\pi\varepsilon_0 \varepsilon_{r_{Si}}} = \frac{1{,}602\times10^{-19} \cdot 457 \cdot 10^{15}}{2\pi \cdot 1{,}035\times10^{-12}} = 11{,}3\times10^{9} \quad [Hz] \tag{2.7}$$

2.3 Depletion regions

More complex phenomena occur when a junction with two different material types is formed. Those PN junctions are inherent to CMOS processes since wells have to be created in the substrate.

The formation of a PN junction gives rise to a space-charge region due to impurity ionization and majority carrier diffusion. In the substrate, those structures are reverse biased (i.e. there is no current flowing through the junction) and they behave as variable capacitor. The capacitance C_t of such arrangements is given by:

$$C_t = \frac{A\varepsilon_{Si}}{X} \tag{2.8}$$

where A stands for the junction's area and X is the depletion region thickness. The later depends from the doping profile of the junction and from potential difference of both side. One can deduce from [4] a general relation to evaluate the thickness of the space-charge region:

$$X = [\alpha(\Delta\psi_0 + V_{NP})]^{\gamma} \tag{2.9}$$

where the constant $\Delta\psi_0$ represents the junction's built-in voltage which is determined by:

$$|\Delta\psi_0| = \frac{kT}{q}\ln\frac{n_{0N}p_{0P}}{n_i^2} \tag{2.10}$$

and the constants α and γ depend on the doping profile:

	α	γ
asymmetric step junction	$\frac{2\varepsilon(N_A+N_D)}{qN_AN_D}$	$\frac{1}{2}$
linearly graded junction (a = doping gradient)	$\frac{12\varepsilon}{qa}$	$\frac{1}{3}$

Table 1: Junction capacitance constants α and γ for two different doping profiles

In nonequilibrium situations, substrate junctions can momentarily become forward biased at switching time. For instance, this can occur while switching on an n-channel transistor. At that time p-type majority carriers flowing through the substrate resistance to build the channel will increase the potential near the drain or source junction. The junction becomes forward biased and electrons will diffuse in the substrate through the junction.

Moreover, because of the reverse bias of substrate-well junctions, the part of minority carriers reaching any substrate-well junction is pulled through the depletion region by the junction's electric field. Actually, the structure made by n^+ drain/p^- substrate/n well behave as a parasitic npn bipolar transistor, as illustrated in figure 4 below. Similarly, a p^+ drain/n well/p^- substrate forms a parasitic pnp bipolar structure. Despite the poor characteristics of such devices due to the large base dimension, they present a major risk if no precaution is taken to limit the voltage sweep near source or drain junctions.

2.4 Latch-up

All CMOS circuits have a potential problem called latch-up, that is related to parasitic bipolar transistors. Figure 4 shows how, when fabricating CMOS devices, parasitic bipolar transistors are also created.

Latch-up structure is made by two pnp and npn parasitic bipolar where the collector of either transistor is connected to the base of the other, forming a pnpn parasitic SCR (Semiconductor Controlled Rectifier). Under certain conditions such as terminal overvoltage stress, transient displacement currents or ionizing radiation, lateral currents can cause sufficient substrate or well voltage drop. Therefore, emitter-base junctions become forward biased activating both bipolar devices. When current gain product is sufficient, the SCR can be switched to a low impedance state. This condition is defined as latch-up. Latch-up can result in momentary or permanent loss of circuit function, depending on ability of the power supply to source the excess current.

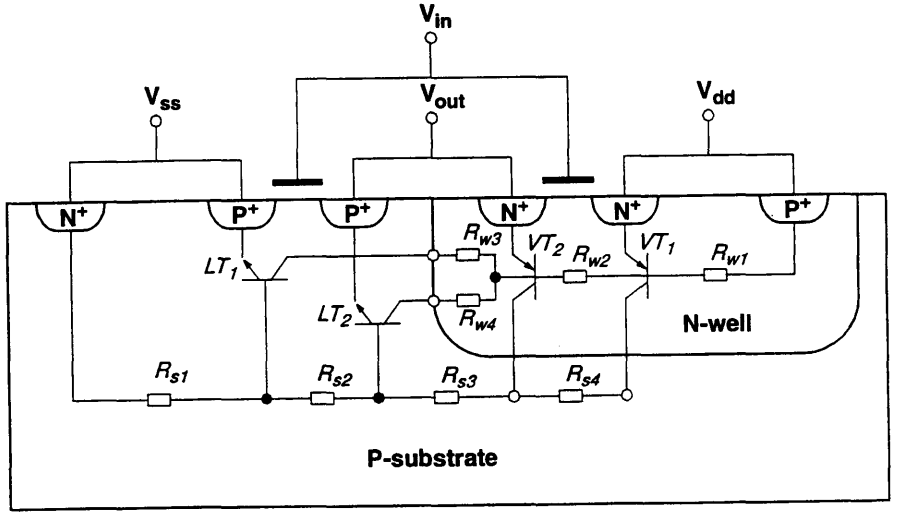


Figure 4: Parasitic bipolar transistors responsible for potential latch-up problems in an inverter fabricated with an N-well CMOS technology

Latch-up as been widely studied and the complex phenomena involved have been described, explained and modeled in depth [6]. Actually, latch-up can be avoided by observing appropriate technological and design rules. Hence, modeling of this parasitic substrate behavior is excluded from the present work.

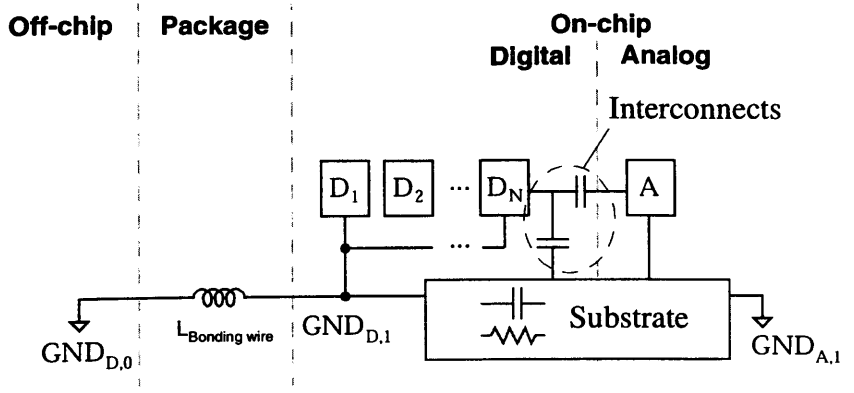


Figure 5: Simplified representation of IC parasitic effects

3 Parasitic Substrate Effects

The increasing influence of parasitic effects occurring in mixed-signal ICs significantly degrade the system performances. IC parasitics have a very complex influence through different medium, as illustrated in Figure 5.

The substrate influences the behavior of a circuit design in a parasitic manner. Current flowing to ground through the bulk creates a voltage drop which affects the device operation. Furthermore, the capacitance of the substrate delays signal transmission to different locations of the device, giving rise to parasitic behavior. Moreover, the substrate is not a perfect isolation between devices, leading to undesirable crosstalk in integrated circuits.

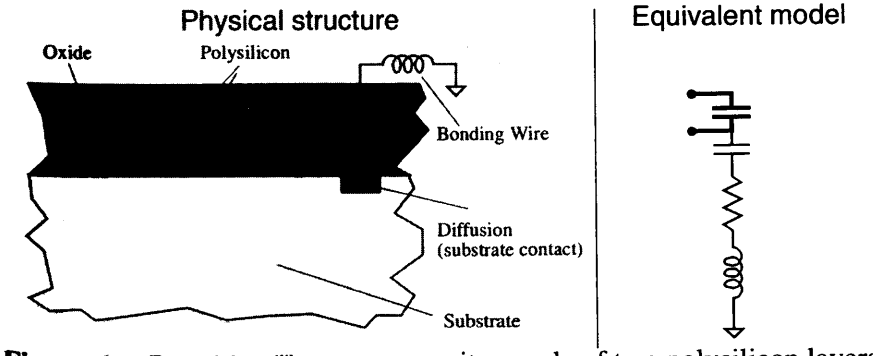


Figure 6: Parasitic effect on a capacitor made of two polysilicon layers

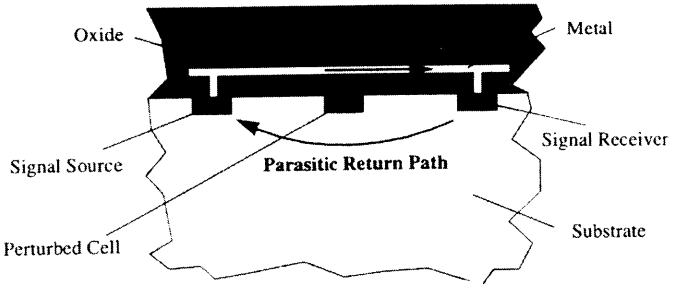


Figure 7: Substrate as a parasitic return path

Firstly, the bulk affects integrated devices by adding parasitic resistances and capacitances. Figure 6 provides an example where the substrate, together with bonding wires, introduces a parasitic RLC structure on a capacitor. For example, a 10 [pF] capacitance together with a typical bonding inductance of 4[nH] has a resonant frequency of 800 [MHz], which is a source of instabilities and oscillations.

Secondly, substrate behaves as a noise vehicle. Perturbations are produced by the digital cells. Therefore, the noise is present on digital interconnects carrying switching signal as well as on digital power supply lines which potential is bouncing due to the existence of bonding wires. Then, noise reaches sensitive analog cells through the network of interconnects and through the substrate. The latter is becoming a very critical piece of high-end IC design as the feature size of the technology decreases. There exist two distinct aspects to substrate noise.

On one hand, as illustrated in figure 7, the substrate is used as a parasitic return path for signals carrying relevant information. Crosstalk happens if a sensitive cell present along this parasitic path is perturbed by the signal. In this context, the parasitic current flow is parallel to the silicon surface.

On the other hand, the bulk might be driving AC noise to the ground. Here, the least resistive path is be followed and the noise flow is determined by the distribution of substrate contacts to AC ground. As depicted in figure 8, the presence of a backside contact gives rise to a vertical current.

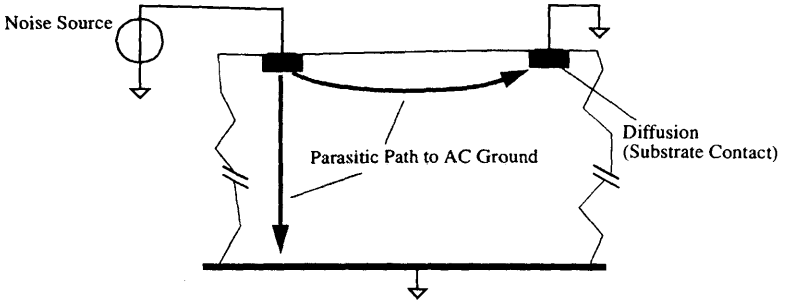


Figure 8: Substrate as a parasitic path to AC ground

The occurrence of one or the other type of parasitic effect, together with the techniques available to reduce the parasitic impact, relates significantly to the kind of wafer and process used for the IC fabrication.

4 Wafer Impact

The substrate characteristics relate significantly on the type of wafer used together with the backside processing and packaging. Understanding the wafer impact as well as the backside influence is critical in choosing the correct noise reduction technique.

There exists two major type of wafer. A lightly doped wafer is made of a silicon material uniformly doped with a typical concentration of 10^{15} [cm^{-3}]. Epitaxial wafer consist of a similar lightly doped material grown by epitaxy on top of an heavily doped material — 10^{19} [cm^{-3}] is a

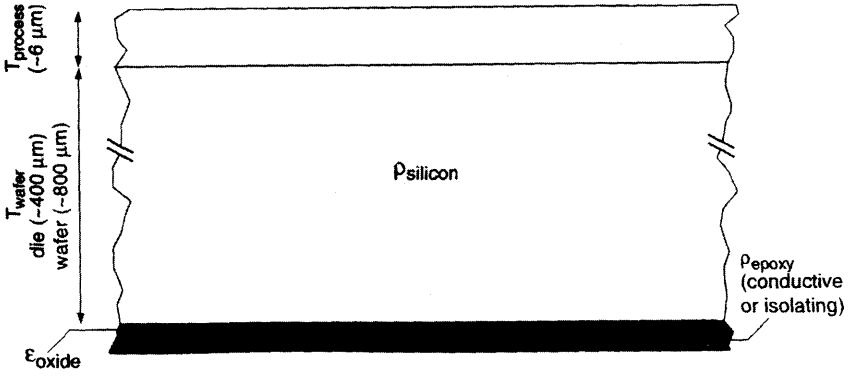


Figure 9: Basic substrate structure

common doping concentration. The thickness of the lightly doped section vary from 5 to 8 [µm]. The difference in resistivity of the two kind of wafers results in a large variation of the electric field distribution. Sub-sections 4.1 and 4.2 present the distribution of the electric field for lightly doped and epitaxial wafers, and present simplified models to help understanding the different of behaviors of the two kind of bulks.

4.1 Lightly Doped Wafer

Lightly doped wafers are characterized by an homogeneous doping concentration around 10^{15} [cm^{-3}]. With a corresponding bulk resistivity approaching 15 [Ω·cm], the current flow distribution through substrates between two surface positions relates significantly to the distance between the source and the receiver.

A MEDICI simulation of the current flow and equipotential lines between two surface contact with respect to the distance is depicted in Figure 10.

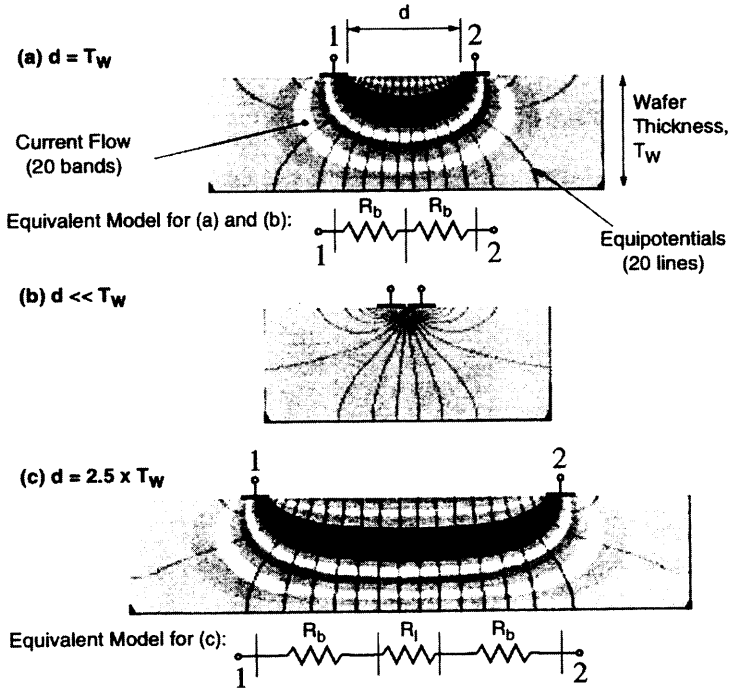


Figure 10: Substrate current flow in a lightly doped wafer with non-conductive epoxy on the backside

Each band holds approximately 5% of the total current. The backside is assumed non-conductive.

For the sample (a), the contact distance has been set equal to the wafer thickness. The proposed modeling with two identical resistances R_b respects the structure symmetry. Therefore, the equivalent resistance between the surface contacts, $R_{12} = 2\ R_b$, varies non-linearly as the distance decreases owing to the important bending of the electric field. As plotted in figure 10(b), the bending becomes greater as the contacts are getting closer. Adversely, when the distance increases beyond the wafer thickness as shown in figure 10(c), the bending does no longer change with the distance. The equivalent resistance R_{12} is then made of two elements taking into account the bending, plus a medium component which value increases linearly with the distance.

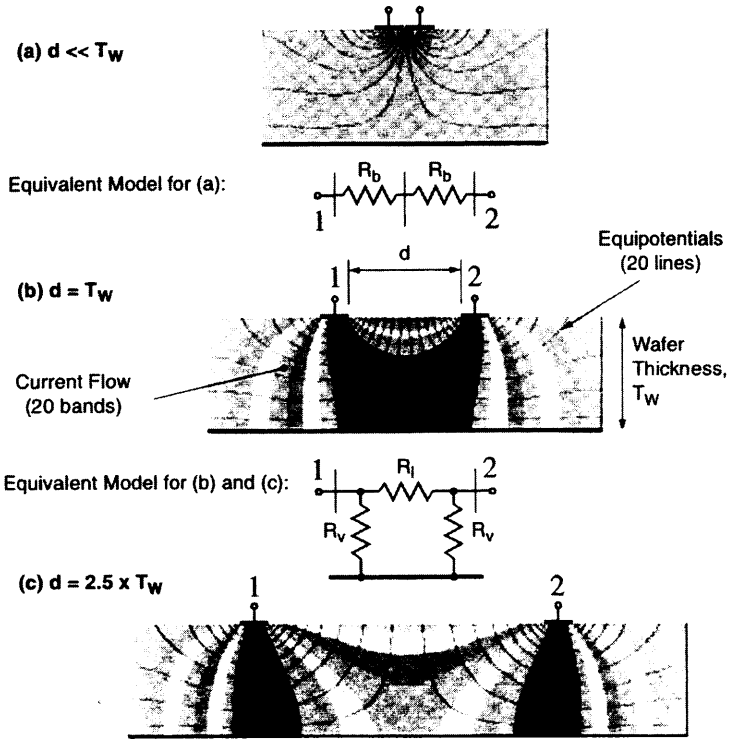


Figure 11: Substrate current flow in a lightly doped wafer with conductive epoxy on the backside

When the backside is made conductive and left floating, the behavior is very similar for small distances such as in figure 11(a). However, as the space between the surface contacts reaches T_W, the low-resistivity backside makes a preferred path for the current. For d close to the wafer thickness — see figure 11(b) — more than 50% of the total current is flowing through the backside contact. Therefore, the model is improved to separate the lateral component, R_l, from the vertical one, R_v. Figure 11(c) shows how, when the distance further increases, the lateral effect becomes negligible causing the overall resistance value between the surface contacts to levels off.

The substrate resistance is plotted in figure 12 as a function of the distance, for a wafer thickness of 400 [μm]. The bulk is made of P-doped silicon with a concentration of $7 \cdot 10^{14}$ [cm^{-3}]. The contacts have a size of 100 [μm] and the resistance values are provided for a structure width of 100 [μm]. For non-conductive epoxy, the linear behavior is clearly apparent beyond the wafer thickness. The use of a conductive die attachment causes a levelling off of the substrate resistance.

Independently of the type of epoxy used to attach the die to the package, the linear relationship applied to compute the resistance of interconnects or diffusion layers — i.e. $R_{12} = \rho_l \cdot A$, where ρ_l is the specific layer resistivity in [Ω/square] and A is the layer area in [square] — is not accurate for the substrate. If the third dimension is considered, the resistance variation

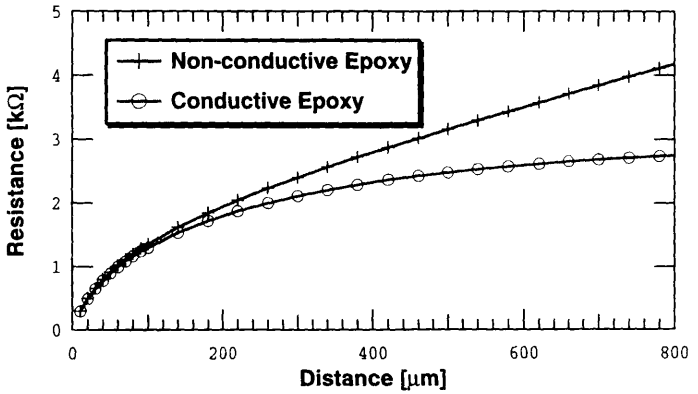


Figure 12: Lightly doped substrate resistance as a function of distance for a wafer thickness of 400 [μm]

becomes more complex as the current bends in this direction as well. On a real chip where many contacts are present, a complex resistive mesh is required to model the interaction through the substrate [9].

4.2 Epitaxial Wafer

For heavily doped bulks with a lightly doped epitaxial layer the substrate current flow can be modeled by the simple structure represented in figure 13. With the very low resistivity of the heavily doped region— typically 0.015 [Ω·cm] — the bulk is considered as one single electrical node [13].

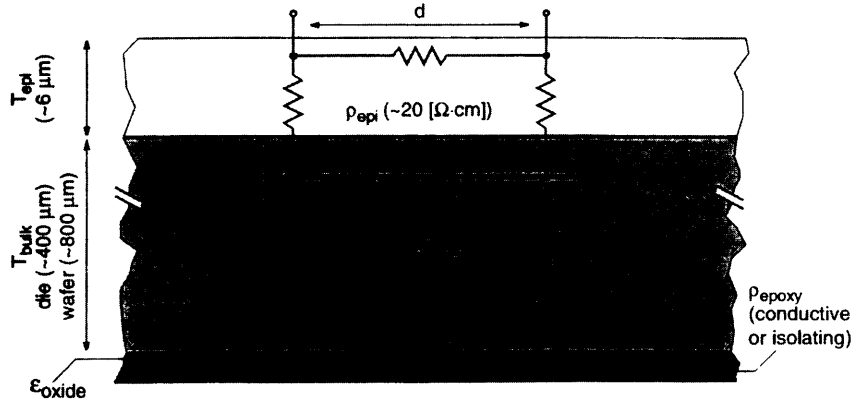


Figure 13: Structure of an epitaxial wafer

The resistive model of an epitaxial wafer is influenced by distance as the one of a lightly doped wafer with conductive epoxy. For two contacts closer than $4 \cdot T_{epi}$, a significant portion of the total current flows in the epitaxial layer. The substrate model of figure 14(a) must be used to account for the lateral element, R_l, as well as the vertical component, R_v. Adversely, when the distance is larger than $4 \cdot T_{epi}$, the current in the epitaxial layer flows vertically and two resistances are sufficient to model the substrate, such as illustrated in figure 14(b). Typical values for R_v are around 2000 [$\Omega/\mu m^2$].

As the resistance levels off beyond $4 \cdot T_{epi}$ — approximately 25 [μm] — distance cannot be used to isolate sensitive and perturbing cells. Therefore, the only efficient method to avoid substrate perturbations is to reduce as much as possible the amount of noise injected in the bulk (c.f. [7], chapter

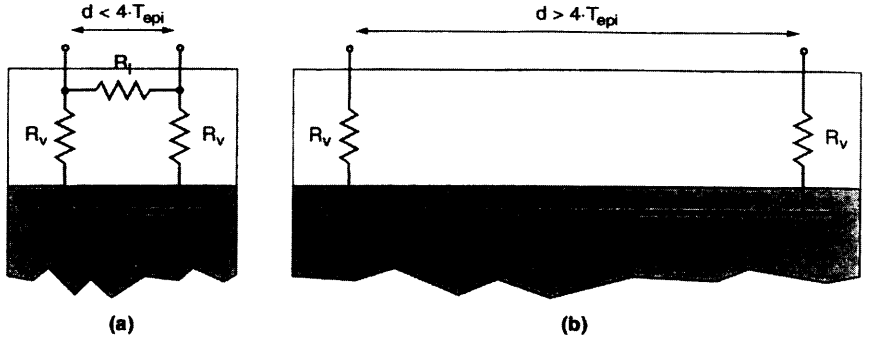


Figure 14: Resistive models for an epitaxial wafer

6). One possible solution consists of using an additional pin to connect the substrate backside to an AC ground separate from the digital power supply. Nevertheless, this technique has a frequency limitation related to skin effect.

As mentioned in section 3, when the bulk is used as a parasitic return path, the substrate current flow is parallel to the surface. Owing to the skin effect, high-frequency signals have a limited penetration depth and the backside is inefficient in collecting such perturbations [1]. The skin depth, T_{skin}, is defined as the distance from the surface beyond which the current density is 1/e of the surface current density. Furthermore, the current density bellow two skin depths is negligible.

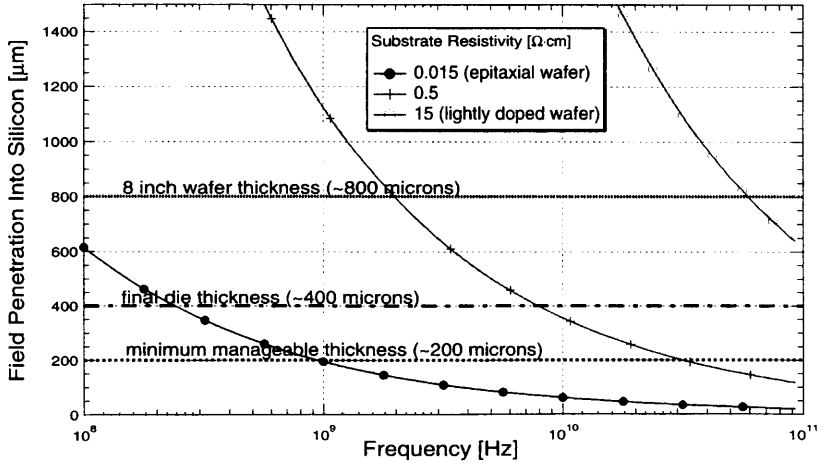


Figure 15: Penetration depth of a lateral current flow for different bulk resistivities

The effect relates to the bulk resistivity, ρ [Ω·cm], to the permeability, μ [H/cm] — equal to μ_0 for silicon —, and to the signal frequency, f [Hz]:

$$T_{skin} = \sqrt{\frac{\rho}{\pi \mu f}} \quad [\text{cm}] \tag{4.1}$$

The penetration depth for different resistivities is plotted in Figure 15. For a bulk resistivity around 0.015 [Ω·cm], high-frequency perturbations above 600 [MHz] flowing at the silicon surface are not collected by the backside ground.

5 Fabrication Processes

Most of the current processes are based either on MOS or bipolar technologies. The first permits to achieve high densities of integration while the second is favored for implementing high speed functions. The process steps that influence the most substrate parasitics are summarized in figure 16.

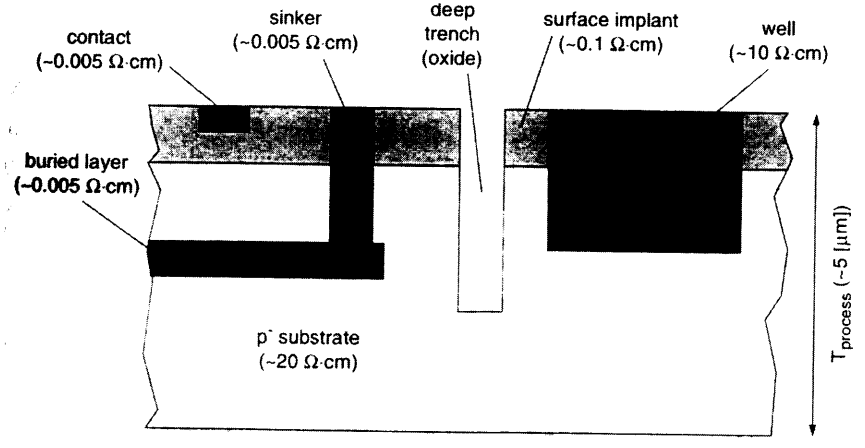


Figure 16: Typical process steps influencing substrate parasitics

The surface implant, also known as p-tub or channel stop, has a usual thickness around 1 [μm] with a resistivity of 0.1 [Ω·cm]. The first purpose of this layer is to avoid parasitic inversion of the silicon surface from the lowest metal layers. Additionally, the low resistivity decreases the risk of latch-up in MOS processes. This is the layer that affects the parasitic sub-

strate behavior the most significantly. The example of surface implant provided in Figure 17 exhibit a surface doping two order of magnitude higher than the bulk concentration. Accordingly, the current density is significantly higher and most of the substrate noise is flowing in this layer.

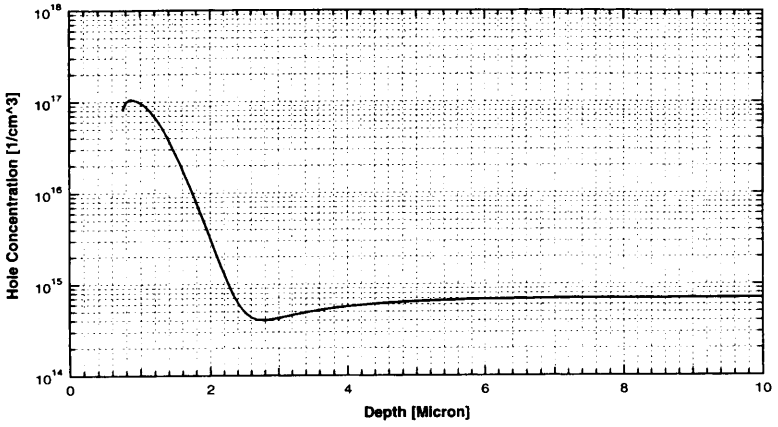


Figure 17: Example of surface implant profile in standard 2mm MOS process

The surface implant can be broken using the wells available in MOS processes, or the deep trench existing in bipolar technologies. As a result, the noise is forced to flow in the substrate depth where the resistivity is significantly higher. The use of a well is further improved by applying a reverse bias potential that increases the depletion width.

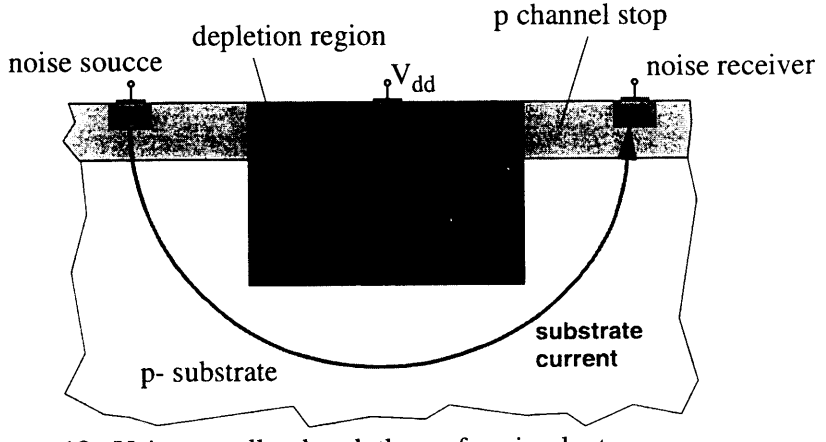


Figure 18: Using a well to break the surface implant

6 References

[1] T. Schmerbeck, "Noise coupling in mixed-signal ASICs," *Low-power HF microelectronics: a unified approach*, pp. 373-430, G. Machado - Editor, IEE, 1996.

[2] N. Verghese, T. Schmerbeck, and D. Allstot, *Simulation Techniques and Solutions for Mixed-Signal Coupling in Integrated Circuits*, Kluwer Academic Publishers, 1995.

[3] R. W. Dutton and Z. Yu, *Technology CAD - Computer Simulation of IC Processes and Devices*. Kluwer Academic Publishers, 1993.

[4] R. M. Warner and B. L. Grung, *Semiconductor-Device Electronics*. Rinehart and Winston, Inc., 1991.

[5] J. Y. Chen, *CMOS Devices and Technology for VLSI*. Prentice-Hall, 1990.

[6] R. R. Troutman, *Latchup on CMOS Technology - The Problem and Its Cure*. Kluwer Academic Publishers, 1986.

[7] T. Blalack, "Switching Noise in Mixed-Signal Integrated Circuits", *Thesis Dissertation*, Stanford University Department of Electrical Engineering, December 1997.

[8] A. Pun, T. Yeung, J. Lau, F. J. R. Clement and D. Su, "Experimental Results and Simulation of Substrate Noise Coupling via Planar Spiral Inductor in RF ICs", *IEEE International Electron Device Meeting*, December 1997.

[9] T. Blalack, J. Lau, F. Clement, and B. Wooley, "Experimental Results and Modeling of Noise Coupling in a Lightly Doped Substrate," *IEDM Technical Digest*, pp. 623-626, December 1996.

[10] K. Makie-Fukuda, T. Kikuchi, T. Matsuura, and M. Hotta, "Measurement of Digital Noise in Mixed-Signal Integrated Circuits," *IEEE J. Solid-State Circuits*, vol. 30, no. 2, pp. 87-92, February 1995.

[11] T. Blalack, B. A. Wooley, "The Effects of Switching Noise on an Oversampling A/D Converter", *IEEE International Solid-State Circuit Conference*, pp. 200-201, February 1995.

[12] R. Merrill, W. Young, and K. Brehmer, "Effect of Substrate Material on Crosstalk in Mixed Analog/Digital Integrated Circuits," *IEEE International Electron Devices Meeting*, pp. 433-436, December 1994.

[13] D. Su, M. Loinaz, S. Masui, and B. Wooley, "Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal Integrated Circuits", *IEEE J. Solid-State Circuits*, vol. 28, no. 4, pp. 420-430, April 1993.

[14] R. Gharpurey, M. C. Chang, U. Erdogan, R. Aggarwal and J. P. Mattia, "R.F. MOSFET Modeling Accounting for Distributed Substrate and Channel Resistances with Emphasis on the BSIM3v3 SPICE Model", *IEEE International Electron Device Meeting*, December 1997.

[15] R. Gharpurey and S. Hosur, "Transform Domain Techniques for Efficient Extraction of Substrate Parasitics", *IEEE International Conference on Computer-Aided Design*, pp. 461-467, December 1997

[16] J. Casalta, X. Aragones, and A. Rubio, "Substrate Coupling Evaluation in BiCMOS Technology," *IEEE J. Solid-State Circuits*, vol. 32, no. 4, pp. 598-603, April 1997.

[17] M. Pfost, H. Rein, and T. Holzwarth, "Modeling Substrate Effects in the Design of High-Speed Si-Bipolar ICs," *IEEE J. Solid-State Circuits*, vol. 31, no. 10, pp. 1493-1501, October 1996.

[18] K. Kwan, I. Wemple, and A. Yang, "Simulation and Analysis of Substrate Coupling in Realistically-Large Mixed-A/D Circuits," *IEEE Symposium on VLSI Circuits*, pp. 184-185, June 1996.

[19] R. Gharpurey and R. G. Meyer, "Modeling and Analysis of Substrate Coupling in Integrated Circuits", *IEEE J. Solid-State Circuits*, vol. 31, no. 3, pp. 344-353, March 1996.

[20] N. K. Verghese, D. J. Allstot and M. A. Wolfe, "Verification Techniques for Substrate Coupling and Their Application to Mixed-Signal IC Design", *EEE J. Solid-State Circuits*, vol. 31, no. 3, pp. 354-365, March 1996.

[21] A. Viviani, J. P. Raskin, D. Flandre, J. P. Colinge and D. Vanhoenacker, "Extended study of crosstalk in SOI-SIMOX substrates," *IEEE International Electron Devices Meeting*, pp. 713-716, December 1995.

[22] J. P. Raskin, D. Vanhoenacker, J. P. Colinge and D. Flandre, "Coupling Effects in High-Resistivity Simox Substrates for VHF and Microwaves Applications", *IEEE International SOI Conference*, pp. 62-63, October 1995.

[23] I. L. Temple and A. T. Yang, "Mixed-Signal Switching Noise Analysis Using Voronoi-Tesselated Substrate Macromodels", *32nd IEEE Design Automation Conference*, pp. 439-444, June 1995.

[24] T. Smedes, N.P. van der Meijs, A.J. van Genderen, P.J.H. Elias and R.R.J. Vanoppen, "Layout Extraction of 3D Models for Interconnect and Substrate Parasitics", *25th European Solid-State Device Research Conference*, pp. 397-400, 1995.

[25] S. Mitra, R. A. Rutenbar, L. R. Carley and D. J. Allstot, "A Methodology for Rapid Estimation of Substrate-Coupled Switching Noise", *IEEE Custom Integrated Circuit Conference*, pp.129-132, May 1995.

[26] K. Joardar, "A Simple Approach to Modeling Cross-Talk in Integrated Circuits," *IEEE J. Solid-State Circuits*, vol. 29, no. 10, pp. 1212-1219, October 1994.

DESIGN TECHNIQUES TO REDUCE SUBSTRATE NOISE

Tallis Blalack
Snaketech, San Jose, California

ABSTRACT

Substrate noise fundamentals are first presented. Next, reducing the noise created by the digital circuitry is discussed. Decreasing the sensitivity of the analog circuitry is covered, and isolation techniques to minimize noise coupling between areas are identified. Finally, a design example is presented to illustrate the difficulty in correctly identifying substrate-noise coupling.

1 INTRODUCTION

In a mixed-signal integrated circuit, performance issues can arise due to the integration of digital and analog circuitry on the same chip. The analog portion of a chip may work well when fabricated separately, but a decrease in performance is observed when the digital circuitry is added. This does not mean that a designer need only reduce the transfer of noise to the analog circuitry to deal with switching noise problems. Attempts to minimize the coupling of switching noise into sensitive analog circuits must take into account three different areas: the amount of noise generated in the digital circuitry, the sensitivity of the analog circuitry to this noise, and the transfer of the noise from the digital portion of the chip into the analog section. Each area is important, and focussing on only one or two may result in a design that does not meet the required specifications. In the digital circuitry it is

essential to identify the noise sources in order to reduce the noise generation. In the analog section, the circuitry must be carefully laid out and the design must be as resistant to noise as possible. Finally, steps to reduce the transfer of noise from the digital circuitry to the analog circuitry are required.

Because each aspect of the noise problem varies from design to design, and from fabrication technology to fabrication technology, there is no universal solution to the minimization of switching noise problems. Each solution will be uniquely defined by the characteristics of the design and the technology. The following CMOS guidelines [1] [2] [3] may not translate directly to other fabrication technologies [4] [5] [6] [7] [8]. As with any design trade-offs, modifying a design to achieve better noise rejection may have an associated cost, such as additional die area, more package pins, or a more expensive package.

2 SUBSTRATE NOISE FUNDAMENTALS

The two primary mechanisms by which noise is introduced into the substrate are supply bounce and capacitive coupling. A brief description of supply bounce is presented along with an example of parasitic capacitances that commonly occur in the layout of an integrated circuit. Once a voltage transient has been induced in the substrate, the type of substrate can make a substantial difference in the manner in which noise is transmitted to other circuits on the chip. Although a detailed knowledge of semiconductor fabrication techniques is not necessary to understand substrate noise in ICs, some understanding of the type of substrate used in the fabrication process is required. Therefore, the two most common types of substrates used in CMOS fabrication will also be discussed.

2.1 Supply Bounce

Figure 1 illustrates a cross-section of a package cavity with bond wires connecting the chip to the package traces. The inductance of the package and bond wires can lead to supply bounce. During a node transition, current is drawn from either the power supply or ground. If output drivers transition

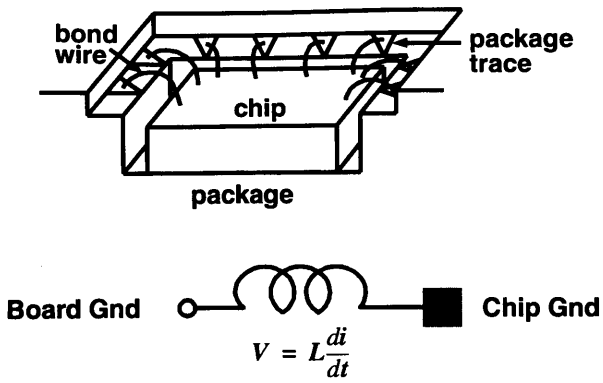


Figure 1: On-chip supply bounce results from the voltage drop across bond-wire and package inductances.

or a large number of nodes change state, a large amount of current is drawn from the supply and induces a voltage drop of $V = L\frac{di}{dt}$ across the inductance of the bond wire and package trace. This voltage drop between the board supply and the chip causes an on-chip supply to "bounce" and results in noise on the supply. As a consequence, the digital power and ground can be very noisy, and a separate power and ground are normally supplied to the analog portion of the chip to isolate the more sensitive analog circuitry from the digital supply noise.

2.2 Capacitive Coupling

Parasitic capacitances exist at every node in a circuit. A parasitic capacitance between two circuit nodes is present, not by design, but due to the physical layout of the circuit. Figure 2 illustrates several of the parasitic coupling capacitances that exist in the layout of an integrated circuit. $C_{L\text{-}L}$ is the line-to-line capacitance that exists between two nearby lines. Line-to-line coupling is not a problem if noisy and sensitive lines are not routed together in the layout. $C_{L\text{-}W}$ and $C_{L\text{-}S}$ are the capacitances between a line and an underlying well or substrate, respectively, and are normally large for clock lines and supply lines. $C_{W\text{-}S}$ is the capacitance between a well and the substrate. If the well contact is connected to a digital supply,

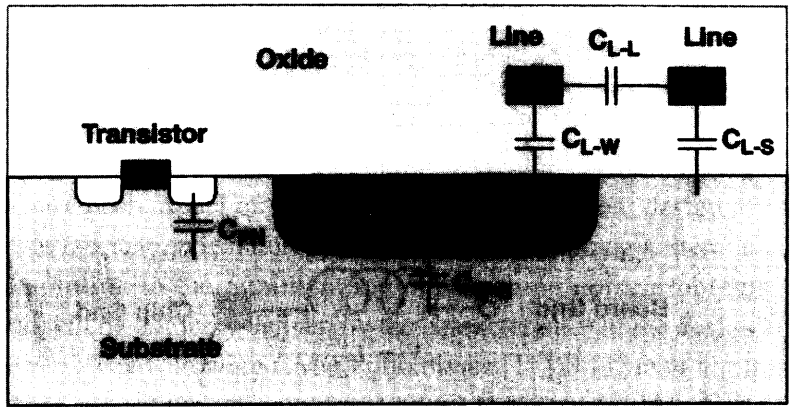


Figure 2: Examples of parasitic coupling capacitances in the layout of an integrated circuit.

noise on the supply may be coupled through the well's junction capacitance to the substrate. Finally, C_{PN} is the capacitance of a drain/source p-n junction to the substrate. If an output driver is switching or a large block of digital circuitry is switching, C_{PN} may be large enough to couple a noticeable voltage transient into the substrate.

Circuit simulation of an entire chip is often impractical due to simulation time and memory constraints. In these cases, the sensitive analog circuitry can be simulated together with a simple circuit that emulates the switching noise produced by a large block of digital circuitry, as shown in Figure 3. An inverter driving a p-n junction is sized to duplicate the transient current injected into the substrate by the logic block [9]. A logic timing simulator can be used to develop an equivalent switching noise source.

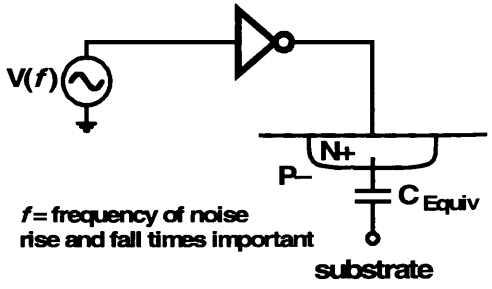


Figure 3: Simple circuit model used to emulate a digital noise source.

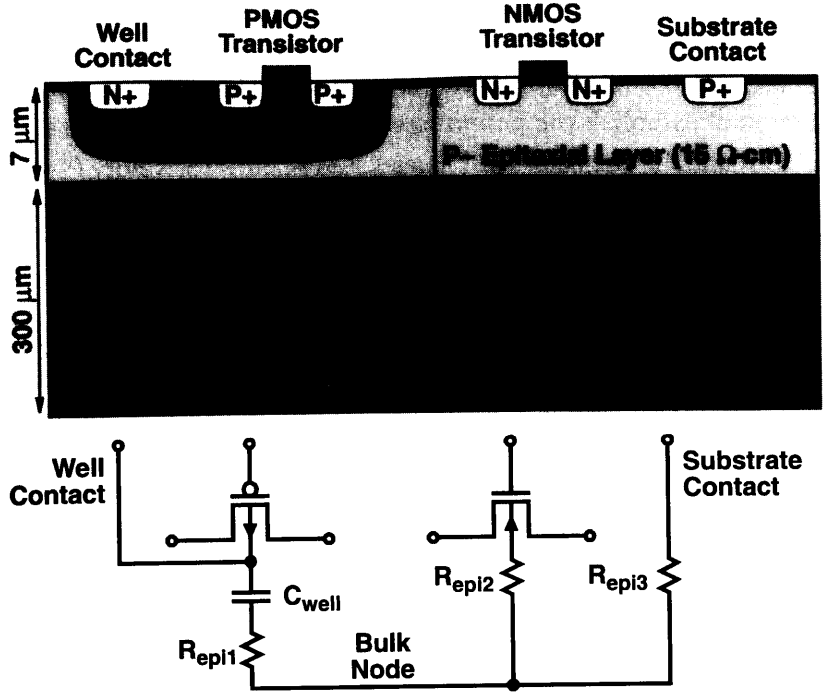


Figure 4: Cross-section of a heavily-doped, p-type substrate with a lightly-doped epitaxial layer and the corresponding circuit model.

2.3 Epitaxial Substrates

Many CMOS fabrication processes today use a lightly-doped epitaxial layer grown on top of a heavily-doped bulk substrate. The lightly-doped epitaxial layer provides a tightly controlled level of doping for device performance, while the low resistivity of the heavily-doped bulk helps to prevent latchup. P-type substrates and an n-well technology are used in the figures and the corresponding descriptions, but the discussion applies equally for n-type substrates with a p-well process.

The top portion of Figure 4 illustrates the cross-section of a 300-µm heavily-doped, p-type wafer with a 7-µm lightly-doped epitaxial layer. The resistivity of the bulk substrate is typically two or more orders of magnitude

less than the resistivity of the epitaxial region. In the n-well technology depicted, the NMOS transistor is formed directly in the epitaxial layer, while the PMOS transistor is placed in an n-well. The substrate contact and well contact establish the potential of the localized well and substrate regions around the contacts. The channel stop is a low resistivity P– layer at the top of the substrate that prevents inversion of the bulk regions outside the transistor channel areas.

The bottom portion of Figure 4 shows the circuit model corresponding to the chip cross-section. Experimental results have shown that once a noise source and sensitive node are separated by more than four times the epitaxial thickness, noise propagates through the heavily-doped bulk substrate and the coupling becomes independent of distance. Substrate noise effects can then be included in circuit simulations by modeling the heavily-doped bulk substrate as a single node [10]. Circuit models are used to represent the transistors with their associated junction capacitances. Resistors R_{epi1}–R_{epi3} represent the spreading resistance through the epitaxial layer, and C_{well} represents the well-to-substrate junction capacitance.

2.4 Lightly-Doped Bulk Substrates

Lightly-doped bulk substrates are lower cost than epitaxial substrates. Latchup, historically a concern in lightly-doped substrates, has become less troublesome as supply voltages have been lowered. The trend to lower supply voltages, motivated by decreasing device dimensions and low power issues, has made it less likely that latchup will occur, or will be sustained if it occurs, and is leading to the wider usage of lightly-doped substrates. Figure 5 illustrates a cross-section of a lightly-doped, p-type substrate.

For a lightly-doped bulk process, a resistive-mesh model of the substrate is necessary for simulation, and noise coupling is very layout dependent. Experimental results have shown that both lightly-doped and epitaxial substrates can be viewed as a purely resistive structure below several GHz [5] [11]. Without an efficient simulation capability utilizing a model of the substrate, it is extremely difficult to predict how digital switching noise will affect circuit performance [12].

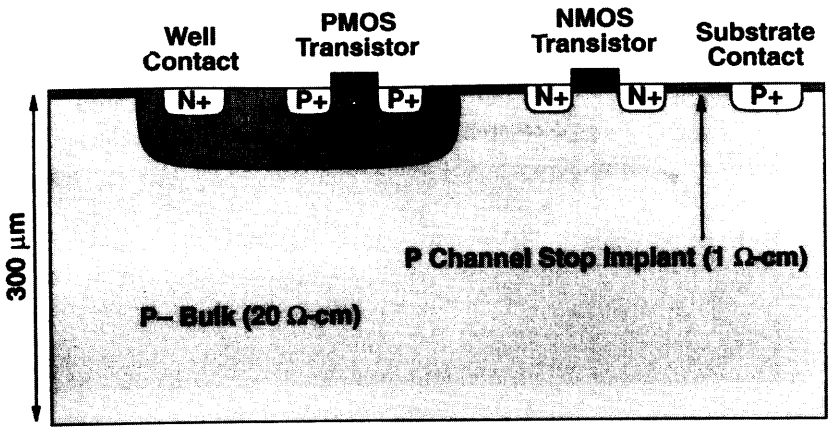


Figure 5: Cross-section of a lightly-doped, p-type substrate.

3 QUIETING THE TALKER

Reducing the noise created by the digital circuitry is a logical place to start. If the noise can be diminished, it may be possible to ease the analog circuitry design constraints and spend less time on isolation issues. This section addresses noise generation issues that are likely to be common across many designs, since individual designs will have different sources of noise. In general, turning off functions not in use and reducing the number of gates switching simultaneously will help minimize the switching noise created.

Certain nodes in a design have a large parasitic capacitance to the substrate. These nodes typically include supply networks, buses, output drivers, clock distribution networks, and other circuits with large fanout. Fast voltage transitions on these nodes couple noise into the substrate. To minimize this noise coupling, the rise and fall times of the circuitry driving these nodes should be as large as the design constraints will allow [13] [14]. Slowing the voltage transitions on large capacitive nodes reduces coupling into the substrate in a two-fold manner. The increased transition times reduce the instantaneous current drawn from the supply. If the current spike

the differential input and can be described using the single-ended circuit. During the sampling period, ϕ_1, switches S1 and S3 are closed and the input voltage is connected to node A. A charge proportional to the difference between the input voltage and the common-mode voltage is stored on the sampling capacitor, C1. After switch S3 opens, node B is left floating and the charge on node B must remain constant. S1 opens on clock phase ϕ_{1D} a few nanoseconds later so that charge injection from the switch cannot change the amount of charge stored on node B. The sampling instant when ϕ_1 falls is the most sensitive timing point of the analog circuitry, and digital switching noise can disturb the sampling process. On a large chip or a chip with a high clock rate for the digital circuitry, it may not be possible to adjust the timing of the digital section to avoid switching during this time.

4 DESENSITIZING THE LISTENER

No matter how much care is taken during the design of the digital section of a chip, the digital circuitry is likely to generate significant amounts of transient noise. Futhermore, if an analog circuit is designed as a building block in a cell library, there is no way to anticipate what circuitry will be placed on-chip with the analog block. Therefore, the analog circuitry should be designed to be as insensitive as possible to the introduction of noise through the substrate, supply lines, and clocks. Although most analog designers would normally take the following suggestions into account during the design phase, the presence of switching noise highlights these considerations.

The use of a fully differential design provides common-mode rejection, which implies that any noise coupled equally into both sides of the design will be attenuated [16]. By designing for a high common-mode rejection ratio (CMRR), many of the switching noise effects will be seen as common-mode and thus rejected. In a similar manner to designing for a good CMRR, a high power-supply rejection ratio (PSRR) is desired. The on-chip analog supplies are often noisy due to current drawn by the analog circuitry and capacitive coupling from the substrate. A high PSRR will minimize supply noise coupling into the signal path.

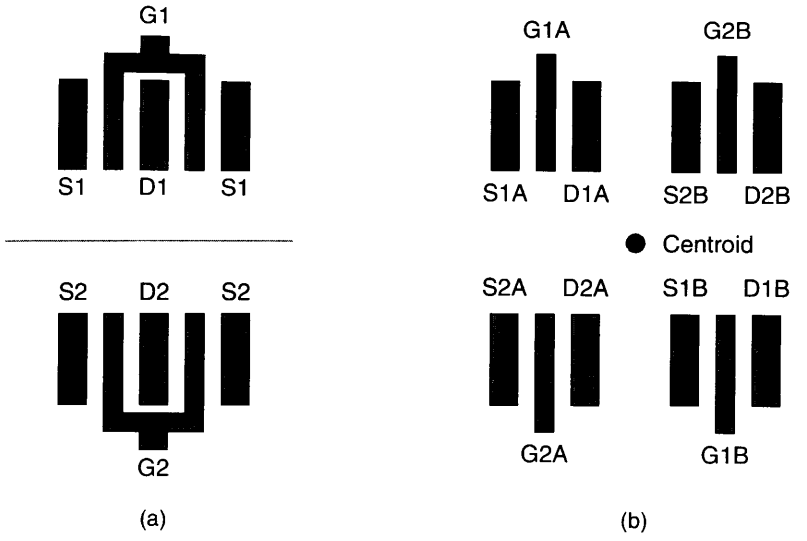


Figure 8: Layout showing (a) mirror symmetry and (b) common-centroid symmetry.

To implement a fully differential circuit with high common-mode rejection, the layout must have the same parasitic coupling to each of the two differential paths. This may require either mirror symmetry or a common-centroid geometry for the layout [17] [18] [19]. Figure 8(a) illustrates mirror symmetry. The differential design has a center line about which one side of the differential path is a mirror image of the other. This ensures that any parasitics to adjacent lines or to the substrate will be identical. In mirror symmetry, the current flow through the mirrored devices should be parallel to the line of symmetry.

Although mirror symmetry may produce identical parasitics for each differential path, location-dependent mismatches may require a common-centroid layout. Figure 8 (b) illustrates a common-centroid geometry in which the transistors of (a) are centered around a common point. Process gradients across a die may change device parameters between two locations, affecting threshold voltage, oxide thickness, resistivity, etc. Also, in the case of a lightly-doped bulk substrate, the noise coupled to a specific node will depend on the location of that node. A common-centroid layout

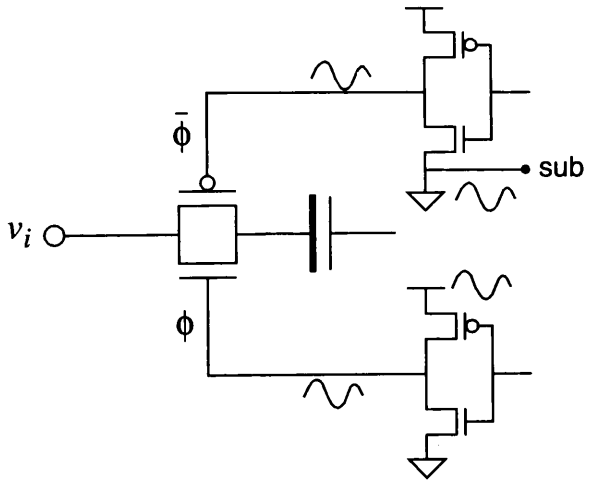


Figure 9: Coupling paths from the supplies to the clock lines.

ensures that process gradients will be identical between the two sides of a differential circuit and that any noise will couple equally to both sides of the differential path. Although a common-centroid arrangement provides better matching, it presents a more complex and time-consuming layout task.

Figure 9 illustrates noise coupling from the supplies onto the clock lines. Noise on the analog clocks can couple into the signal path even when fully differential circuitry is utilized [20]. To achieve acceptable noise levels, it may be necessary to connect the clock drivers that drive the analog switches to the analog power and ground if the digital supplies are noisy. As Figure 9 shows, if a substrate contact is connected to the ground of the clock driver, substrate noise and ground noise will be indistinguishable. In order to keep substrate noise off the clock lines, the substrate contact should not be connected to the supplies of the clock drivers. In an A/D converter based on the use of sigma-delta modulation, or a similar architecture in which the analog circuitry has noise shaping properties that attenuate noise introduced after the front-end, the use of a continuous-time front-end may reduce noise coupling effects. Because the sampling process is very sensitive to noise, moving the sample-and-hold portion of the architecture to later in the signal path will attenuate noise introduced during sampling. However, a continuous-time front end is more sensitive to clock jitter, so the frequency

requirements of a design will determine whether the circuit architecture is selected based on clock jitter or transient noise constraints [21] [22].

5 ISOLATION

Isolating analog circuitry from digital circuitry requires careful floorplanning of a chip. Noisy and sensitive lines should not be run next to one another because of capacitive and inductive coupling between lines. Similarly, digital signals should not be routed over or through the analog portion of the chip. The floorplan should also ensure that the package pin assignments do not route sensitive analog signals near digital I/Os, supplies, or clock signals.

Isolation techniques will vary dramatically depending on the process technology and the design. Switching noise that couples through the substrate in an epitaxial process requires different techniques to combat the problem than noise coupling in a lightly-doped bulk substrate. In both cases distance isolation is difficult to achieve. For an epitaxial process, experimental data has shown that physical separation does not reduce noise coupling, which led to the development of the single-node model for the heavily-doped bulk. In a lightly-doped substrate, experimental data has shown that the substrate contacts necessary to prevent latchup can reduce the effective "electrical" distance between two points [12].

Although additional pins are required, it is generally necessary to provide separate digital and analog supplies for mixed-signal circuits. Even with analog circuitry designed for good power-supply rejection, if the digital and analog supplies are connected together, the large current spikes from the digital circuitry can create supply voltage fluctuations that affect the performance of the analog circuitry. Some designs may require multiple sets of power and ground connections to further divide circuit blocks. If a circuit block requires a large amount of current quickly, such as output drivers driving large capacitive loads at high speed, it may be best to have separate supplies for that circuit block.

Proper connection of the substrate contacts is a critical piece in addressing switching noise problems. Although package and design constraints

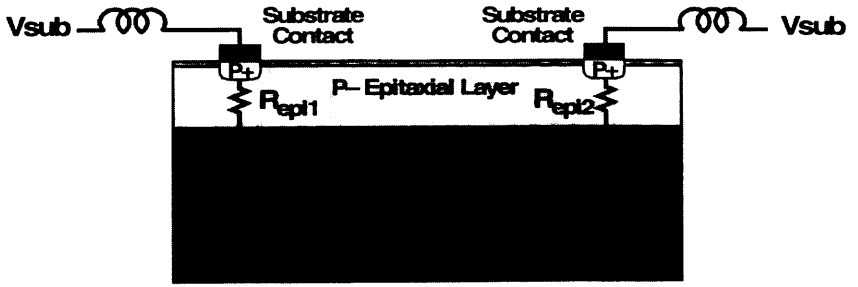


Figure 10: The single-node bulk in an epitaxial process effectively shorts all supplies connected to the substrate.

may limit a designer's choices in dealing with the problem, every effort should be made to connect the substrate contacts near the sensitive analog region to a quiet supply. In an epitaxial process, the substrate should be connected to only one supply. As Figure 10 illustrates, if more than one supply is used to bias the substrate, the low resistivity of the heavily-doped bulk acts as a short between the supplies. The supply chosen to bias the substrate should carry as little current as possible to avoid voltage drops across the bond wire and package inductance. Therefore, for an epitaxial process, the best substrate connection is to a separate substrate supply, or if an extra package pin cannot be obtained, to the analog supply. Some ASIC libraries automatically connect the substrate to the local supply. This, in essence, provides the worst substrate connection because the substrate is then connected to a noisy digital supply.

For a lightly-doped substrate, more than one supply should be used to contact the substrate. The high resistivity of the bulk can be used to provide isolation between areas, and each area should have the substrate connected to a local supply to provide a low resistance return path for substrate current, as shown in Figure 11. Substrate connections may provide a low resistance path to couple noise from one area of a chip to another. To prevent this, digital sections that create significant noise should have their own supplies with the substrate connected to one of the supplies. Contacting the substrate locally and using multiple sets of power and ground will reduce coupling through the substrate contacts and will attenuate a noise transient

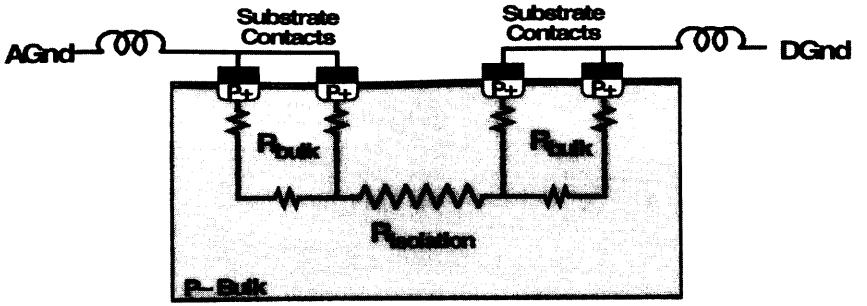
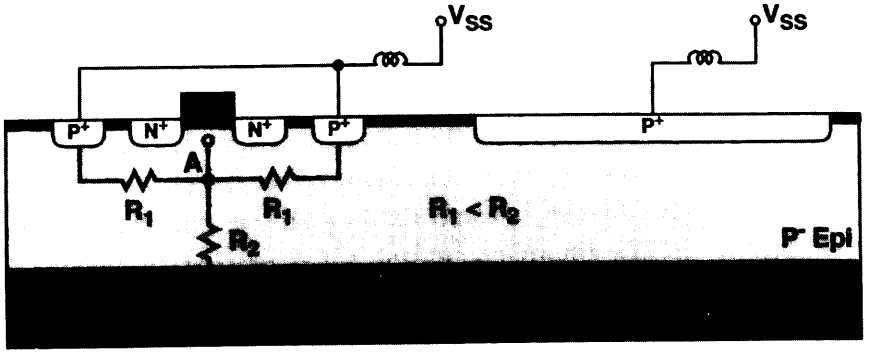


Figure 11: The bulk resistance provides isolation between different sets of substrate contacts in a lightly-doped process.

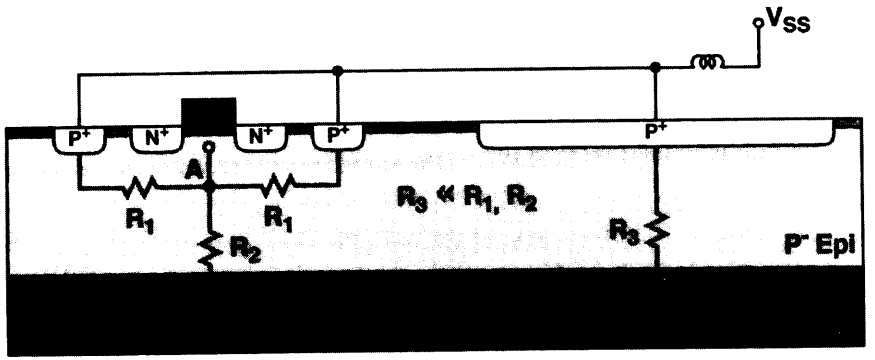
traveling through the substrate as it moves from one set of substrate contacts to another. In some fabrication processes, a metallized ring contacting the substrate is placed around the outside of the chip to seal the edge from alkali ions that may enter the field oxide and affect the yield [23]. If design guidelines allow the edge seal to be broken, the ring should be severed where it provides a coupling path between different sets of substrate contacts.

Reducing the inductance in the substrate bias path will reduce the noise in the substrate. A lower inductance value may be obtained by either changing packages, connecting more package pins to the substrate node, or adding a backside contact. To minimize the noise in the substrate for an epitaxial process, a low inductance backside contact is best. For a lightly-doped bulk process, a backside contact may be beneficial if the circuit includes large on-chip power devices that introduce large currents into the substrate. Contacting the chip on the backside adds additional cost in both processing and packaging and may increase the settling time of the noise.

Devices integrated in a well have the well-to-substrate capacitance in series with any noise coupling into the well from the substrate. This capacitance, along with the well's connection to AC ground through the supply, will provide some degree of isolation. Wells can also be used to break the channel stop implant. In a lightly-doped bulk, the channel stop implant provides a low resistivity region for noise to couple across the die. Placing a



(a)



(b)

Figure 12: (a) P+ guard ring with a dedicated bond pad. (b) P+ guard ring connected to a large substrate contact.

well between the digital and analog areas will force noise deeper into the substrate and result in a larger resistance between the noise source and sensitive areas.

In an epitaxial substrate, guard rings can either help to reduce noise coupling at a sensitive node, or they can couple noise into the sensitive node [10]. In order to reduce switching noise, a guard ring must decouple the epitaxial layer from the noisy bulk in the immediate vicinity of the sensitive node. Resistors R_1, R_2, and R_3 in Figure 12 denote spreading resistances through the lightly doped epitaxial layer. In Figure 12 (a) if R_1 is small with respect to R_2, which can be accomplished by placing the guard ring close to the transistor, the noise at node A will be smaller than the noise in the

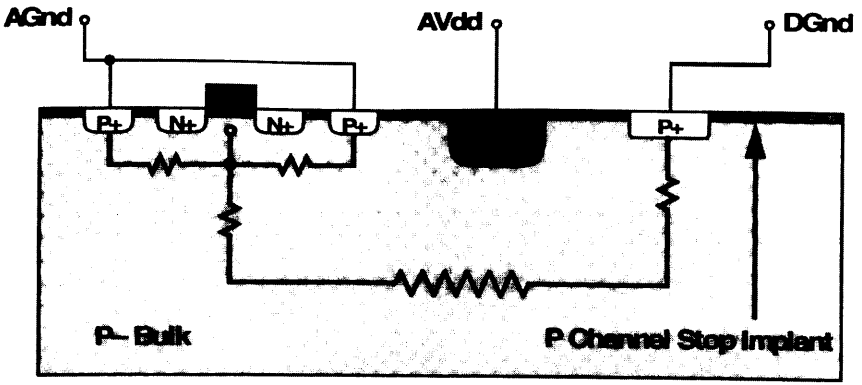


Figure 13: Breaking the channel stop implant layer forces current into the substrate and helps provide isolation between two regions of a chip.

heavily doped bulk. However, if the guard ring is not connected to a dedicated bond wire and is instead connected to a large substrate contact, as in Figure 12 (b), the small resistance R_3 provides a low resistance path to couple noise from the bulk to node A.

In a lightly-doped bulk process, simulations show guard rings to be much more effective [1] [5] [24]. Guard rings around a noise source provide low resistance path to AC ground for the noise and help minimize the amount of noise injected into the substrate. Guard rings around a sensitive ircuit help to decouple noise from the circuit and ensure that noise will couple equally into both sides of a differential design. Finally, a well or guard band between sections will break current flow in the channel stop region, as illustrated in Figure 13.

The low resistance path to AC ground provided by guard rings and other substrate contacts plays an important role in noise coupling. As the circuit model in Figure 14 illustrates, the substrate noise at a receiving node is a function of, not only the resistance between the noise source and the receiving node (R_{SR}), but also the resistance to AC ground from each of these points (R_{SG} and R_{RG}). Figure 15 further illustrates the dependence of the noise coupling on the resistance to substrate ground [12]. The right axis (dSR) is the distance between the source and the receiver. As the noise

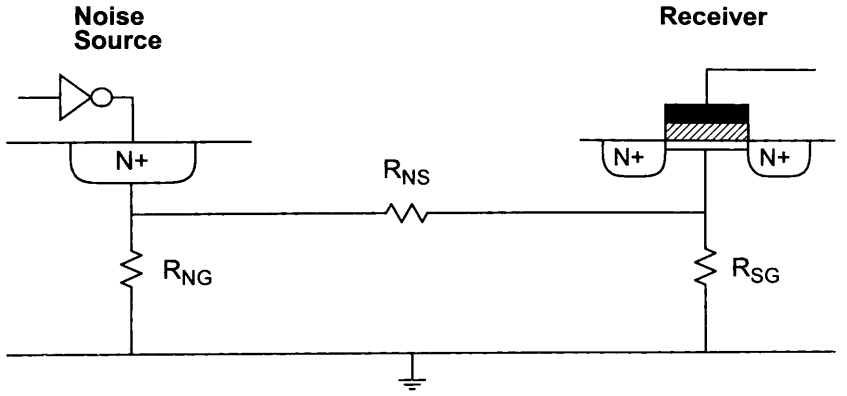


Figure 14: Schematic diagram illustrating the noise dependence on the resistance to ground from both the source and receiver.

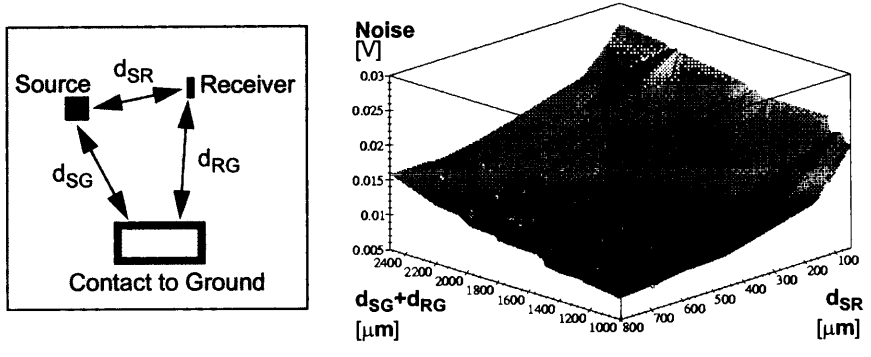


Figure 15: Measurement spread due to the distance from substrate ground.

source and receiving node are placed closer together, the noise increases. The left axis (*dSG+dRG*) is the sum of the distance between the source and ground and the receiver and ground, where ground is the substrate contact. When the distance to the substrate ground is increased, the noise increases. An important point is illustrated in the figure at distances less than 100 μm between the noise source and receiving node. Even at small distances between the noise source and receiving node, an increase in the distance to ground significantly increases the noise coupling.

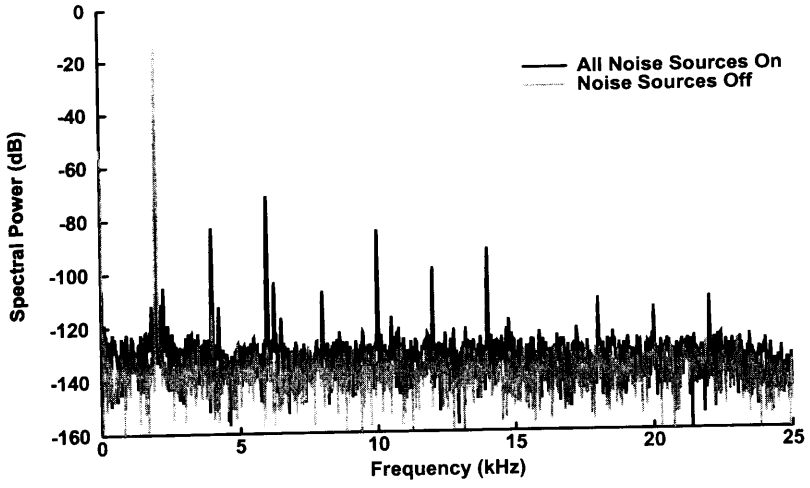


Figure 16: Output spectrum for the oversampling converter with and without noise sources operating.

6 Design Example

This design example is intended to illustrate the difficulty in determining how noise couples into a system. In this experimental chip, noise sources were placed around an oversampling A/D converter. Noise was introduced into the substrate, and the output of the converter was monitored to observe the effects of the noise. As Figure 16 illustrates, when the noise sources were operating, severe harmonic distortion was present in the output spectrum.

The architecture of the data converter and the sensitivity of the harmonic distortion to the timing of the noise and to the input signal level pointed to noise coupling into the sample-and-hold circuitry at the input of the A/D converter. However, the mechanism by which the noise-coupling occurred was unknown. After a long investigation into the problem, two coupling paths were found.

The first noise-coupling mechanism identified was substrate-noise coupling through the p-n junctions of the input transistors into the signal path. Figure 17 depicts a simplified schematic of the differential sample-and-hold circuitry. Switches S1 and S2 have source p-n junctions that connect from

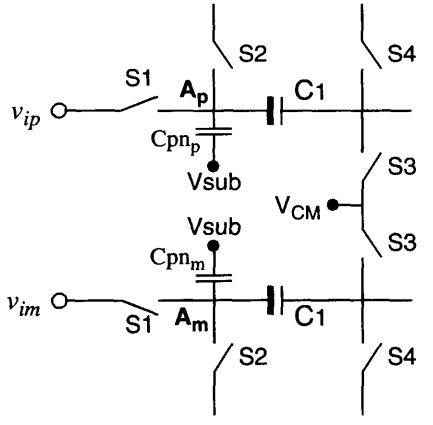


Figure 17: p-n junction capacitive coupling into nodes A_p and A_m.

circuitry. Switches S1 and S2 have source p-n junctions that connect from nodes A_p and A_m to the substrate. Because a different voltage appears across the p-n junction on each side of the differential input, the capacitances from nodes A_p and A_m to the substrate differ in the presence of a nonzero differential input. At larger input amplitudes the capacitive difference between the two sides is greater, which results in a differential noise coupling mechanism that will increase with the input amplitude. Substrate disturbances resulting from noise injected into the substrate will couple through the p-n junctions onto nodes A_p and A_m. A peak-to-peak substrate voltage of 1 V was used for the MIDAS [25] simulation results shown in Figure 18. The level of the third harmonic is –100 dB, which is much lower than the third harmonic observed experimentally and makes the p-n junction a very unlikely candidate as the primary source of the observed harmonic distortion.

The second noise-coupling mechanism was difficult to identifiy due to the fully-differential nature of the design. Noise coupling from the supplies was believed to be common-mode noise, which would undergo common-mode rejection. However, if the gate-to-source capacitance of the NMOS and PMOS input transistors varies as a function of input level, noise on the clock lines will couple differentially through these capacitances onto nodes A_p and A_m and cause a differential disturbance. For small input

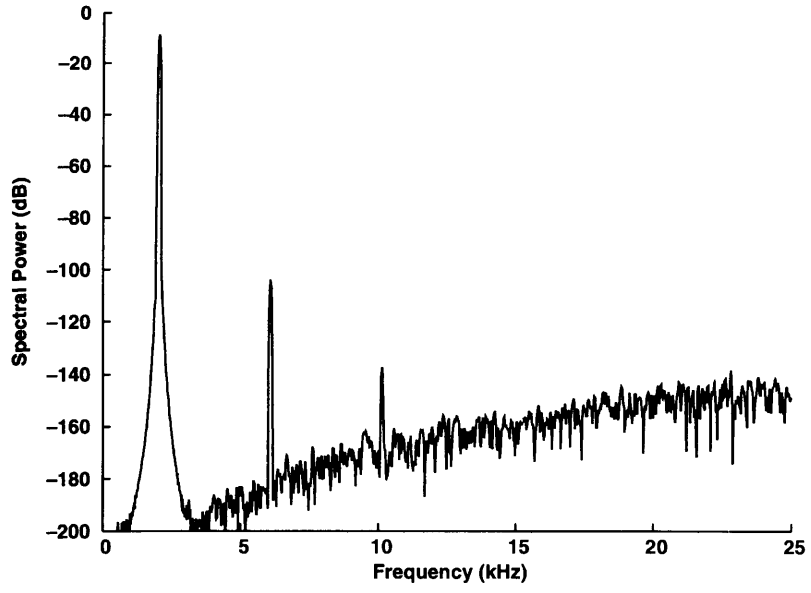


Figure 18: MIDAS output spectrum for the oversampling converter with substrate-noise coupling through the input switch p-n junction capacitances.

amplitudes the input voltage is near the common-mode voltage, and both the PMOS and the NMOS transistors of the transmission gate conduct on each side of the differential input. At large input amplitudes the NMOS sistor conducts on the low voltage side of the differential input, and the PMOS transistor conducts on the high voltage side. One side has an NMOS channel capacitance from the gate-to-source; the other side has a PMOS hannel capacitance from the gate-to-source. This capacitive difference sets up a differential path to couple noise into the input signal, as Figure 19 illustrates.

Significant noise occured on the digital supply in this design due to the large current drawn by the noise sources. Because the clock generation circuitry was connected to this supply, supply noise was coupled onto the clock lines and then into the signal path through the input-transistor channel capacitance. Simulations of the system that included the gate-to-source capacitive difference showed the harmonic tones observed experimentally.

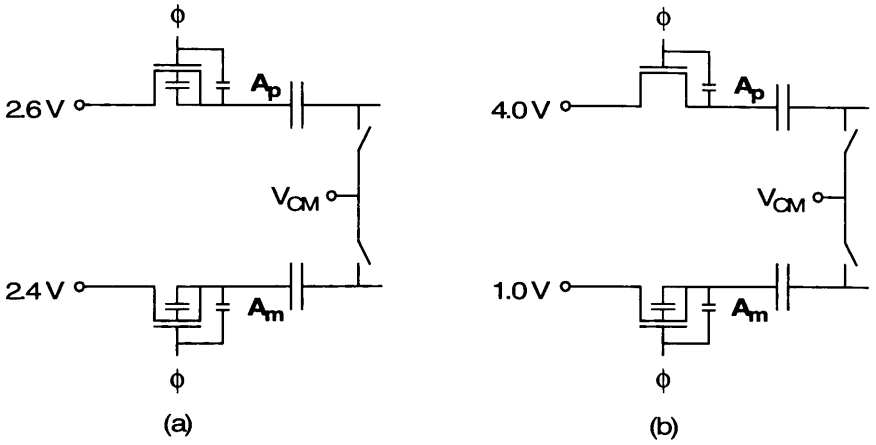


Figure 19: Input circuitry showing a common-mode coupling path (a) and a differential coupling path (b) for clock noise.

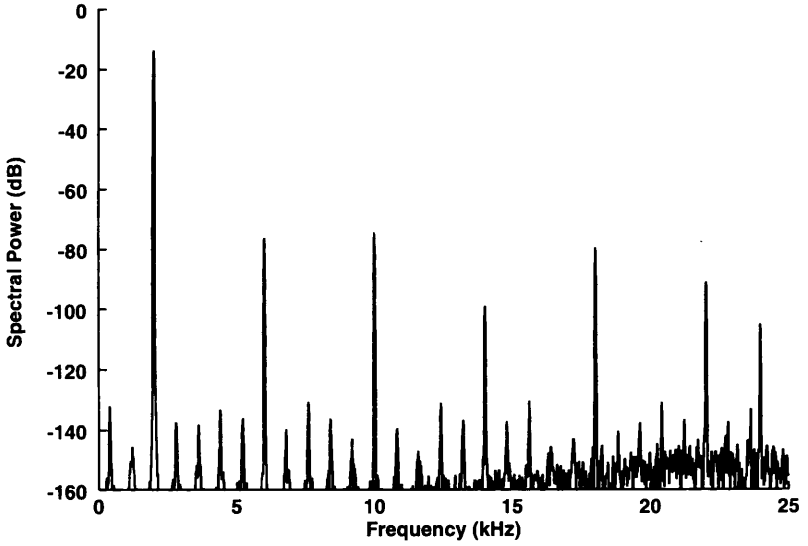


Figure 20: MIDAS output spectrum for the oversampling converter with noise coupling through the input switches gate-to-source capacitances.

The strong correlation between the simulated harmonic tones, shown in Figure 20, and those observed experimentally suggests that the coupling of supply noise through the input switch gate-to-source capacitance is the primary noise coupling mechanism.

This example illustrates the difficulty in evaluating on-chip noise-coupling. Modeling substrate noise coupling is a non-trivial problem, so if a noise problem exists, and the noise-coupling path is not quickly identified, the substrate may be incorrectly assumed to be the conduit for the noise. As substrate-noise modeling programs become more widely available, it will become more feasible to determine if substrate noise is actually the problem.

7 REFERENCES

[1] T. Schmerbeck, "Noise coupling in mixed-signal ASICs," *Low-power HF microelectronics: a unified approach*, pp. 373-430, G. Machado - Editor, IEE, 1996.

[2] N. Verghese, T. Schmerbeck, and D. Allstot, *Simulation Techniques and Solutions for Mixed-Signal Coupling in Integrated Circuits*, Kluwer, 1995.

J. Olmstead and S. Vulih, "Noise Problems in Mixed Analog-Digital Integrated Circuits," *IEEE Custom Integrated Circuits Conference*, pp. 659-662, May 1987.

R. Merrill, W. Young, and K. Brehmer, "Effect of Substrate Material on Crosstalk in Mixed Analog/Digital Integrated Circuits," *IEEE International Electron Devices Meeting*, pp. 433-436, December 1994.

K. Joardar, "A Simple Approach to Modeling Cross-Talk in Integrated Circuits," *IEEE J. Solid-State Circuits*, vol. 29, no. 10, pp. 1212-1219, October 1994.

[6] A. Viviani, J. Raskin, D. Flandre, J. Colinge, and D. Vanhoenacker, "Extended study of crosstalk in SOI-SIMOX substrates," *IEEE International Electron Devices Meeting*, pp. 713-716, December 1995.

[7] M. Pfost, H. Rein, and T. Holzwarth, "Modeling Substrate Effects in the Design of High-Speed Si-Bipolar ICs," *IEEE J. Solid-State Circuits*, vol. 31, no. 10, pp. 1493-1501, October 1996.

[8] J. Casalta, X. Aragones, and A. Rubio, "Substrate Coupling Evaluation in BiCMOS Technology," *IEEE J. Solid-State Circuits*, vol. 32, no. 4, pp. 598-603, April 1997.

[9] M. Nagata and A. Iwata, "A Macroscopic Substrate Noise Model for Full Chip Mixed-Signal Design Verification," *IEEE Symposium on VLSI Circuits*, pp. 37-38, June 1997.

[10] D. Su, M. Loinaz, S. Masui, and B. Wooley, "Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal Integrated Circuits", *IEEE J. Solid-State Circuits*, vol. 28, no. 4, pp. 420-430, April 1993.

[11] F. Clément, *Computer Aided Analysis of Parasitic Substrate Coupling in Mixed Digital-Analog CMOS Integrated Circuits*, Ph.D. Dissertation No. 1449, Swiss Federal Institute of Technology in Lausanne, 1996.

[12] T. Blalack, J. Lau, F. Clement, and B. Wooley, "Experimental Results and Modeling of Noise Coupling in a Lightly Doped Substrate," *IEDM Technical Digest*, pp. 623-626, December 1996.

[13] R. Senthinathan and J. Prince, "Application Specific CMOS Output Driver Circuit Design Techniques to Reduce Simultaneous Switching Noise," *IEEE J. Solid-State Circuits*, vol. 28, no. 12, pp. 1383-1388, December 1993.

[14] D. Cox, D. Guertin, C. Johnson, B. Rudolph, R. Williams, R. Piro, and D. Stout, "VSLI Performance Compensation for Off-Chip Drivers and Clock Generation," *IEEE Custom Integrated Circuits Conference*, pp. 14.3.1-14.3.4, May 1989.

[15] M. Ingels and M. Steyaert, "Design Strategies and Decoupling Techniques for Reducing the Effects of Electrical Interference in Mixed-Mode ICs," *IEEE J. Solid-State Circuits*, vol. 32, no. 7, pp. 1136-1141, July 1997.

[16] K. Makie-Fukuda, T. Kikuchi, T. Matsuura, and M. Hotta, "Measurement of Digital Noise in Mixed-Signal Integrated Circuits," *IEEE J. Solid-State Circuits*, vol. 30, no. 2, pp. 87-92, February 1995.

[17] Y. Tsividis, *Mixed Analog-Digital VLSI Devices and Technology: an Introduction*, pp. 220-236, McGraw-Hill, 1996.

[18] R. Pease, "Comments on 'Analog Layout Using ALAS!'", *IEEE J. Solid-State Circuits*, vol. 31, no. 9, pp. 1364-1365, September 1996.

[19] M. Ismail and T. Fiez, *Analog VLSI: Signal and Information Processing*, pp. 699-726, McGraw-Hill, 1994

T. Blalack and B. Wooley, "The Effects of Switching Noise on an Oversampling A/D Converter," *ISSCC Digest of Technical Papers*, Vol. 38, pp. 200-201, February 1995.

E. van der Zwan and E. Dijkmans, "A 0.2-mW CMOS Sigma-Delta Modulator for Speech Coding with 80-dB Dynamic Range," *IEEE J. Solid-State Circuits*, vol. 31, no. 12, pp. 1873-1880, December 1996.

E. van der Zwan, "A 2.3-mW CMOS Sigma-Delta Modulator for Audio Applications," *IEEE International Solid-State Circuits Conference*, vol. 40, pp. 220-221, February 1997.

S. Wong, Stanford University, private communication, 1997.

K. Kwan, I. Wemple, and A. Yang, "Simulation and Analysis of Substrate Coupling in Realistically-Large Mixed-A/D Circuits," *IEEE Symposium on VLSI Circuits*, pp. 184-185, June 1996.

L. Williams, B. Boser, E. Liu, and B. Wooley, *MIDAS User Manual, Version 2.1*, Integrated Circuits Laboratory, Stanford University, August 1990.

1.2 Gb/s CML TRANSCEIVER WITH 1M CMOS ATM/SDH PROCESSOR IN A BICMOS MONOCHIP

Laurent DUGOUJON

STMicroelectronics,
850, Rue Jean Monnet BP 16 - 38921 Crolles, France

ABSTRACT

A B-ISDN Mixed Signal Design combining very high speed 1.2GHz analog PLLs with high integration level 1Mtr. CMOS Processor is depicted. Architectural choices are presented emphasizing power reduction techniques to achieve 1.5W total consumption. Low noise design style is shown with less than 18ps rms jitter performance. The practical mixed design techniques applied to this chip are described as well as DFT strategy and measurements result.

1. INTRODUCTION

peed and Integration complexity have always been opposite factors. tate-of-the-art mixed signals circuits generally either concentrate on large processing capabilities associated with low speed analog interfaces or very high speed analog functions with small digital control functions.

This work attempts to change this paradigm by using a large spectra BICMOS 0.5µm Technology and specific design techniques to take full benefits of Bipolar and CMOS devices. When complexity and high switching power are needed, the solution is to use CMOS with its superior integration density and associated fast synthesis CAD solutions. Above GHz operation, when low phase noise or jitter are mandatory, the

natural device of choice is then Bipolar.
The philosophy of this work consists in careful partitioning between CMOS and Bipolar and specific actions to maintain minimum power consumption for the whole circuit.

2 . CIRCUIT GLOBAL FUNCTION

The circuit should have two symmetric functions TX and RX allowing simultaneously the conversion-RX/generation-TX of a Serial Data Stream in/from an ATM parallel format. (fig.1). Associated with this Serial/Parallel function, real time complex processing of the transmitted data must be supported [1] [2] [3]. Since SDH and Cell-based protocols are supported, a Framing/Mapping function is used to convert data into SDH frames. The Cell Processor performs Cells delineation [1] with error correction and is associated with OAM for network maintenance specific processing. Finally UTOPIA performs the FIFO interface with asynchronous ATM data-flow.

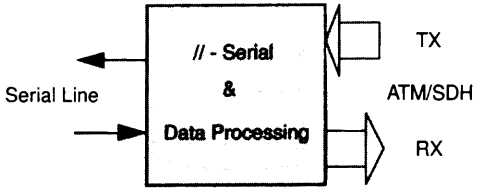


figure 1: Circuit application

The circuit should support these functions for different bit-rates: 155Mbit/s, 622Mbit/s and 1.2Gbit/s for the Transceiver.

3 . PARTITIONING

Partitioning correctly a mixed function system is key to achieving low power and cost optimization. It is valuable at the chip level to determine the best location to implemented a given function.

Figure 2 shows the partitioning into main functions.

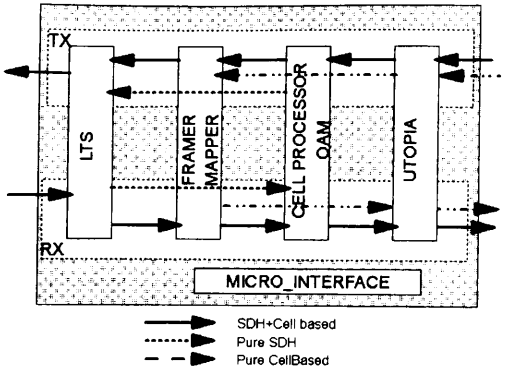


figure 2: Functions partitioning

The LTS block will perform the Serial/Parallel conversion in ECL logic style while the other blocks will be CMOS implemented according to the different protocols supported.
Given this first distribution dictated by interface speed constraints, me functions as bit shifting can be implemented either in the ECL tion or in the CMOS one. An implementation of this shift function the Serial part using a shift register seems to be a good solution. From a deeper analysis of alternative solutions, the possibility of emulating this bit shifting in the parallel section with far smaller area & nsumption penalties was raised. Such partitioning decisions need vileged communications between designers of different specialities ely obtained in conventional specialized team organization.
e final blocks partitioning is illustrated in figure 3.

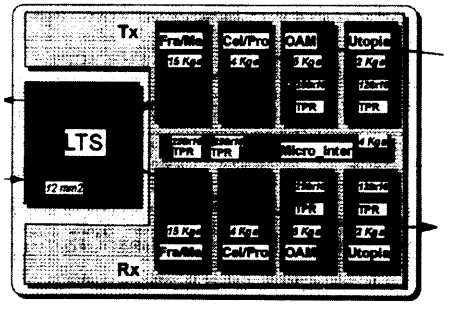


figure 3: Blocks partitioning

The LTS block in ECL logic is finally kept as simple as possible without bit shifting or bit extraction features. Other CMOS blocks are defined in a modular approach allowing shutdown of unused ones and easy re-use in other designs.

4 . ARCHITECTURE CHOICES

We describe first the Parallel CMOS section architecture and then the Serial CML-style part. Our strategy will consist in estimating the relative consumption of the different architectures to determine the "coldest" one.

Parallel section:
In this section, the most important choice is the data-path width. ATM and SDH/SONET protocols are octet-oriented: ATM cells are 53 octet wide.
A natural approach is to use an 8 bit data-path circulating across the different processor stages.
A typical function, Header Error Correction, was exercised with 8, 16 and 32 bit architectures, figure 4. It is clear that the16 bit has the best Power-Area trade-off for this small 320 gate HEC function. We wanted to confirm this result on a more complex block such as the 8.2 Kgate ATM Cell Processor. The 16 bit performance evaluation gave again the best results. Indeed, the area of the control operators based on FSM does not increase with the width of the data-path.
A 16 bit wide data-path is then retained.

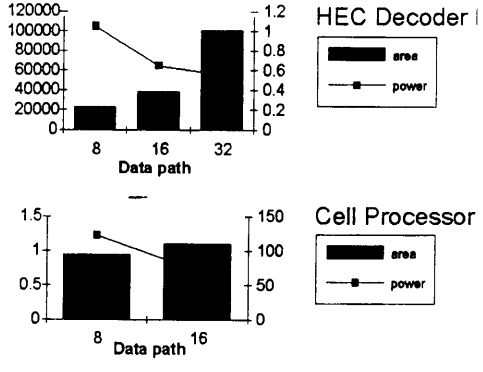


figure 4: Power-Area versus data path

Serial Section:
The goal of this section is to achieve De-serialization in RX or Serialization in TX. The absence of a transmitted High Speed Clock implies generation or recovery of the serial clock for the Serial Data-stream. Different solutions based on PLLs and de/multiplexor were analyzed, the main criteria in this case was to achieve the Speed & Noise performances within a given Power budget of 1W.
The global architecture is given in figure 5.

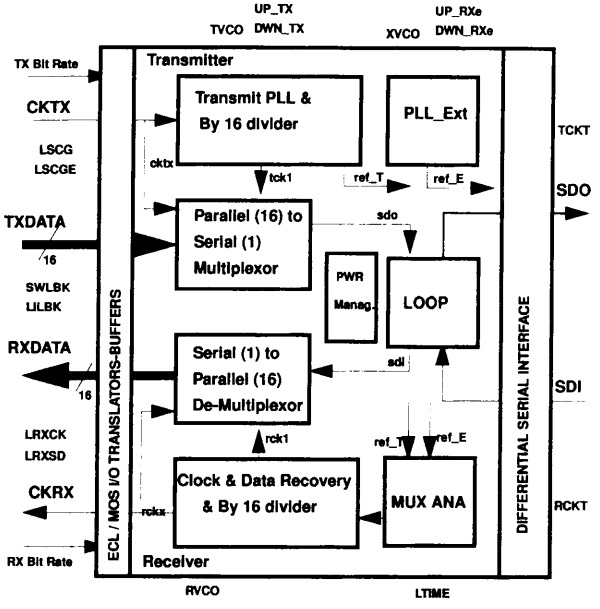


figure 5: LTS architecture

The retained architecture for the PLLs is a classical one [4]. It includes Phase/Frequency comparator, Charge-pump, second order filter, 2.4 GHz VCO and synchronous frequency division stages.
This kind of PLL is used for the Clock&Data Recovery function with specific over-sampling of the incoming data-stream to lock the VCO to the transition positions and then recover the High speed serial clock.
Multiplexor and De-multiplexor architectures were defined to contain the power expenses to a minimum.
Interface blocks provide MOS/ECL level conversion on the processor side and a pre-amplification and line output stage on the Line side.

5 . POWER DECIMATION

CMOS section:

In addition to the already low-power oriented architectural choices previously described, other actions leaded to minimize the total power consumption of the CMOS part.

First, the supply voltage reduction to 3.3V nominal allows the greatest gain, -56% compared to the same implementation with 5V supply voltage. Figure 6 summarizes the techniques used and the associated power cutting results.

Clock gating techniques, both static and dynamic, are applied to put in sleep mode unused blocks or temporarily unused functions. Counter implementation is also optimized by the use of minimum-changes number representation, such as Gray code, as well as Cyclic Polynomial Counter architecture.

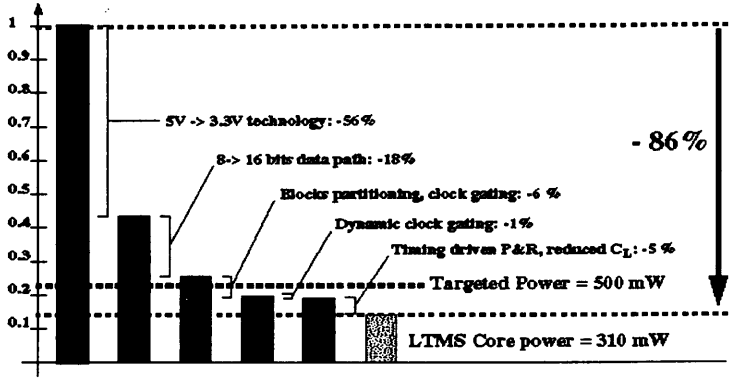


figure 6: Power savings

Finally, capacitive load reduction is achieved through the application of a Timing Driven Place and Route flow that brings some -5% added power savings. All of these techniques permit to ensure the 500mW target and hence minimize the global Noise Power generated by the CMOS digital processor.

BICMOS section:

As low jitter PLLs [5] are integrated in this section, the noise generation from closely associated functions must be carefully managed. In order to minimize switching noise, constant-current implementation

such as ECL or CML types are preferred.

The main limitation of such structures is that power dissipation is constant with respect to operation speed. As maximum 1.2Gb/s operation is supported, leading to 1W consumption, one would have to pay again 1W in the case of 622Mb/s or even 155Mb/s operation. This was not acceptable, so we developed a specific Power-adaptive ECL library able to adjust its consumption to the desired transmission data-rate [4]. Architecture optimization of frequency dividers inside the PLLs and of the multiplexors were performed by mapping the functions in that programmable library. Several parameters are offered to the designer:

Firstly the cell fan-out figure choice which is constant and defined once depending on the circuit topology, secondly, the cell maximum operation speed. This maximum speed is the upper limit of the cell working speed, say Fm, but a 2-bit electrical bus permits to work at Fm/2, Fm/3 or Fm/4 with reduced currents. These two parameters allow the selection of a given pre-defined cell which keeps the minimum power sizing at the different operation speeds.

Figure 7 gives the schematic structure of the cells. The logical swing control is made through PMOS devices in their linear region with small area overhead compared to resistor loads.

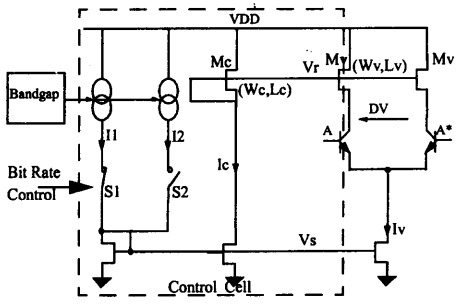


figure 7: power adaptive principle

6 . MIXED DESIGN TECHNIQUES

Noise coupling:

The possible noise sources and applicable techniques to reduce their amplitude or their impact are reviewed:

The first source is related to CMOS output buffers. Indeed these buff-

ers generate large current peaks when charging or discharging external capacitive loads. These current transitions cross the supplies pads, the bonding wires, the package lead-frame and finally the de-coupling capacitor of the supply source. The parasitic L, R, C of this path create instantaneous voltages variation at the die pad and then propagate noise inside the circuit.
A possible solution should consist in dedicating pads to supply the output buffers so that voltage bounce is not dc-forwarded to the other sections of the chip. This is done for the positive VDDBUF supply. For the ground, the addition of a specific local GNDBUF to the existing general GND is more problematic due to substrate shorted through Pwells connections. We used the NISO layer option of our ST BICMOS5 to dc-insulate the noisy grounded Pwells from the clean grounded Psubstrate. See figure 8.

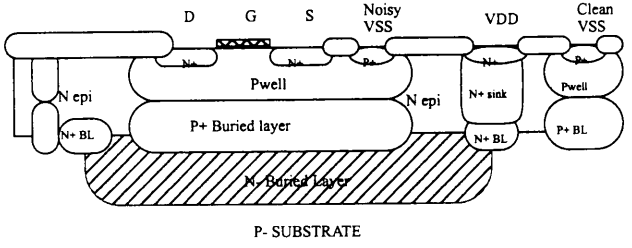


figure 8: Insulated NMOS cross-section

Another noise source is from carrier injection in the substrate when MOS devices are switching on/off through Drain or Source junction capacitance to bulk and, in a reduced manner through the gate to bulk capacitance. Conventional CMOS cells also induce some substrate bounce. So an other large NISO pocket was drawn under all the CMOS blocks. To lower the possible impacts on analog section, the digital supply lines were conservatively sized with many power pads to minimize the digital supply impedance.

Shielding strategy:
Even if precautions are taken to "quiet the talker", the integration of the digital processor with the analog based PLLs of the transceiver

requires us to "insulate the listener" with the use of shielding patterns.

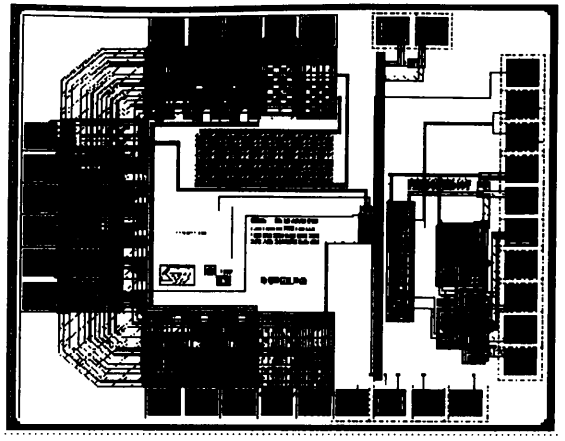


figure 9: Shielding test-chip

We decided to experimentally define the best shielding structure for implementation in our given BiCMOS process and its NISO layer option. Some ten test chip structures were processed using an inverter-based CMOS ring oscillator as a noise generator and a 2.4GHz bipolar VCO as a receptor.
The test chip layout is given figure 9.
The most effective shielding was found to be the combination of NISO layer with N-P-N walls as separation frontier and of an additional P ring around the most sensitive analog cells.

Noise immunity in analog part:
In the ECL section, the switching noise is reduced by the use of bipolar switched constant current digital cells as ECL, or CML whenever possible. Moreover, intra-cell de-coupling capacitors between VCC and analog ground are implemented. Separate power lines reduce the high frequency coupling between PLLs. To further increase the noise immunity of PLLs, control loops are full differential.
Finally, simulations on sensitive blocks were performed with package and bonding parasitic models [4].

Substrate coupling analysis:

Experimental studies through dedicated test chip as reported for shielding choices suffer from fundamental limitations. The main one is the lack of analytic modeling or at least finite element simulations to have better substrate noise insight [6].

Second is the determinant impact of packaging with its associated parasitic effects which limit significantly the test-chip conclusions for different packages.There is therefore an important need to simulate this "substrate channel" to optimize area-noise-package trade-off. For evaluation purpose, a new tool SCA, Substrate Coupling Analysis, was exercised on test-chip structures [7].

Due to very long simulation time needed for VCO jitter visualization, it was not possible to perform comparisons with silicon. However some improvements of SCA and associated DIVA extraction were made thanks to this exercise. Other tools like LAYIN were also considered.

To allow modeling validation, specific structures with Nwell pockets as Emitter and receiver were implemented, see fig. 10.

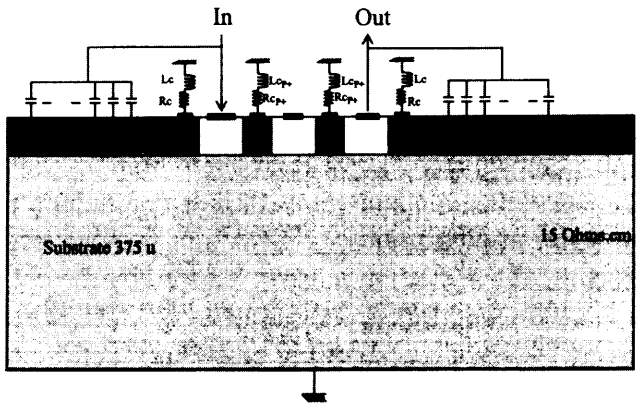


figure 10: Modeling structure cross-section

Inherently to simplifications and coarse decomposition of substrate into equivalent resistors used by these tools, the accuracy they offer on basic structures is lower than manual based ELDO simulation with 2D resistive meshes [8].

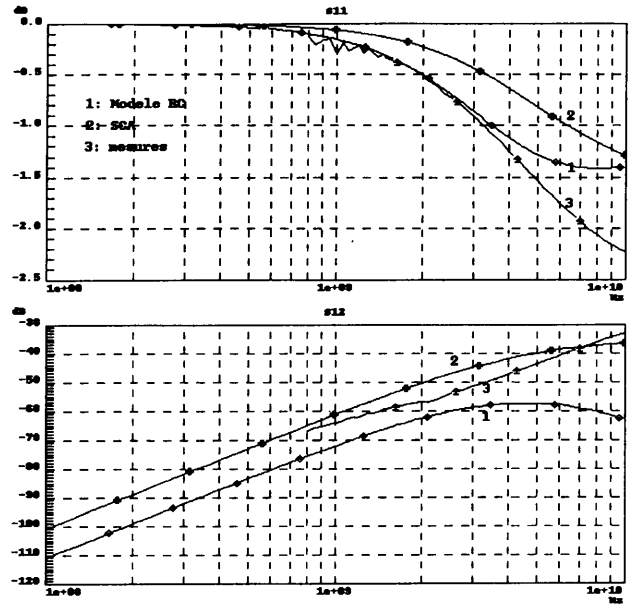


figure 11: Modeling results

Figure 11 shows reflection and transmission coefficient magnitudes comparisons between measurements, SCA and manual resistive mesh decomposition.

Packaging:

Packaging choice is of key importance for noise generation and signal integrity. Deep analysis through concurrent simulation of package and peripheral circuitry such as the 1.2Gb/s ECL output buffer was performed. To allow these simulations with the package seen as an additional "passive component", one needs the package netlist in SPICE format. This demand is increasing as package electrical parameters become pre-dominant in High Speed or RF ICs. From our R&D Cor-

porate Packaging Service, we obtain such package modeling.
The package modeling is achieved with spectral and temporal TDR investigations as well as 3D simulations from geometrical lead-frame description. See fig. 12.

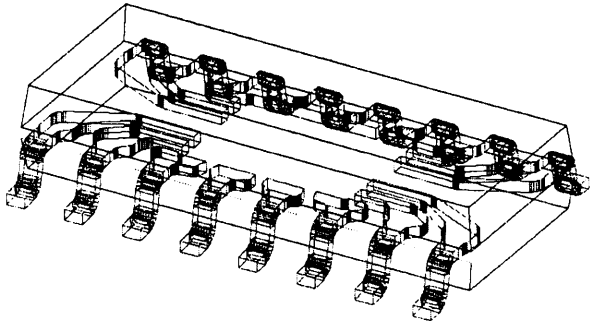


figure 12: SO16 3D view example

The Full matrix of parasitic inductors, resistors and capacitors with mutual coupling is given for all package leads. See fig. 13.

```
L000 1 27 6.91905E-10
L008 9 35 6.91905E-10
R000 18 27 0.000278416
R008 18 35 0.000278416
L001 2 28 4.89217E-10
L009 10 36 4.89217E-10
R001 19 28 0.00022948
R009 19 36 0.00022948
K001_000 L001 L000 0.293539
K009_008 L009 L008 0.293539
L002 3 29 4.89267E-10
L00a 11 37 4.89267E-10
R002 20 29 0.000229558
R00a 20 37 0.000229558
K002_000 L002 L000 0.161384
K00a_008 L00a L008 0.161384
K002_001 L002 L001 0.27386
K00a_009 L00a L009 0.27386
```

figure 13: Package Netlist example (truncated)

The designer should then chose the appropriate pins and import the corresponding netlist inside his circuit simulation.

7 . DESIGN FOR TESTABILITY

Testability issues were addressed with two goals in mind. The first aspect is the functional validation with intensive use of specific test loops, see figure 14. These loops allow the separate stimulation of each function block.
The second point concerns the industrial testing. BISTs operators are

implemented for RAMs, Partial and Full Scan structures enlarge the Fault Coverage above 95%.
JTAG IEEE std 1149-1 is fully supported in the CMOS part I/Os.

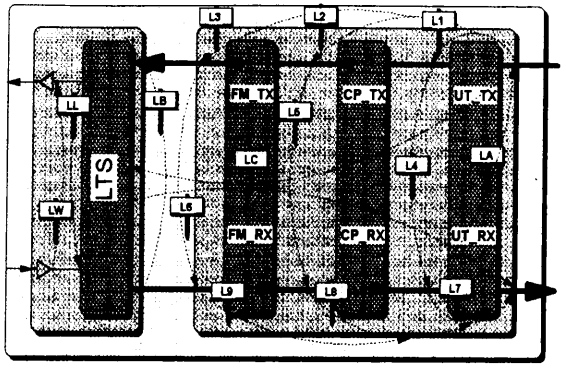


figure 14: Test purposes Loops

8 . OPTIMIZATION FLOW

Design Methodology choices impact the design cycle time and First Time Silicon Success figure.
For the Digital CMOS section, we followed a conventional cell-based design through Synthesis of VHDL netlists. To increase the readability and quality of the RTL code, we evaluated a Graphical Entry Tool named Visual-HDL. That tool was found efficient to develop Finite State Machine codes and brings better material for discussions with the specifications writer, code re-users and... analog Designers.
Back-end of the blocks was achieved by Timing driven Placement & Routing with reduced local iteration loops.

The analog section is characterized by a relative high devices count, 8K devices and a large variety of working frequency disparity among cells, from 2.4GHz down to some 40MHz. Simulating this with a conventional SPICE-like analog simulator would have led to prohibitive CPU time consumption to reach Phase Locked conditions.
To perform faster cycle-time, a top down approach was endorsed [5] using a HDL analog language. The principle of this approach is to propagate the constraints top-down from the HDL global simulation to

the basic cells, then multi-level (HDL + Tr. level) simulations are used to check bottom-up the performances. PLLs and Clock Recovery results fit very well (1%) and the HDLA simulation duration is a fifth of the full transistor simulation time (fig. 15).

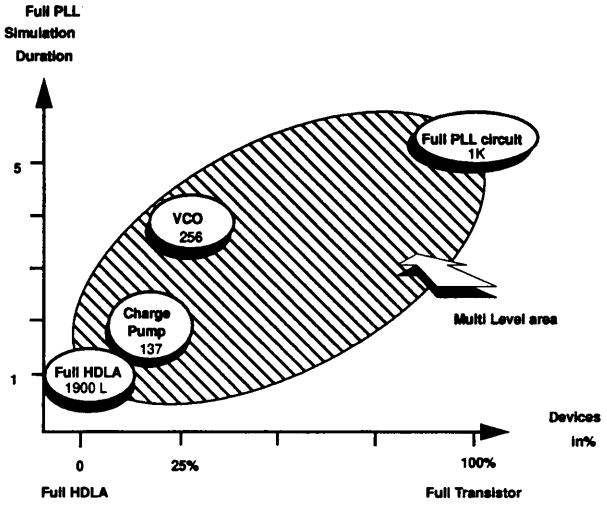


figure 15: Design cycle short-down

Finally, a multi-level simulation of the real cell (tr. level) connected to its environment (HDL cells) is done and validates the Transceiver performance.
The obtained back-annotated description of the full function is easy to transfer to other design groups and needs very little maintenance. It can be re-used to speed-up the development of similar functions in different technologies.

Mixed-mode simulations were not applied due to the nature of the Data stream flow without real-time feed-back between Analog/Digital parts. Back-End was addressed manually with interactive use of Device Level Editor, DLE from Cadence to speed up block assembly.

9 . RESULTS

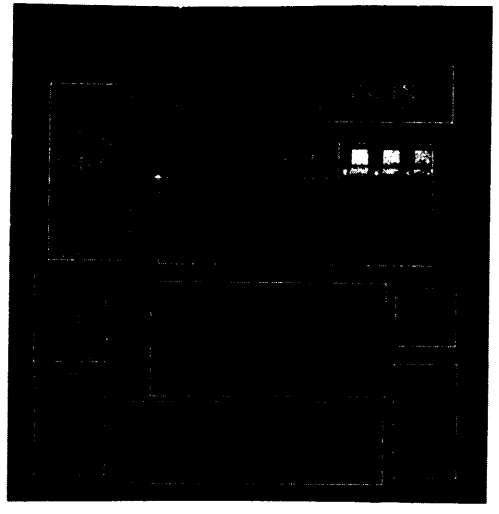


figure 16: Chip micrograph

The die micrograph is displayed in figure 16. Design information is summarized in Table 1.

Table 1:

Technology	0.5 µm BiCMOS 3ML
Power Supply	3.3 Volts
Power Consumption at 622 Mbps	1.5 Watts
Complexity	
ECL	8K devices
CMOS	985K MOS including 46K bit RAM
Die Size	106.2 mm^2
Pin count	208
Package	C, PQFP

Electrical measurements:
The circuit was first validated at the wafer level, All digital CMOS functions were found functional with correct margin above the nominal

38.8MHz. Wafer validation of Serial part was more problematic due to noise problems from the probe test-board. Dissipation characterization was conducted on packaged samples. The main results are displayed in Table 2.

Table 2:

Block	Freq. (MHz)	Bit rate (Mb/s)	Power (mW)
Utopia	50	622	40
LTS in full duplex @3.3 V	155	155	650
	622	622	930
	1244	1244	1010
LTMS	40-622	622	1520

Power consumption measurements exhibit power management efficiency according to the Bit Rate selected.

Serial output measurements:

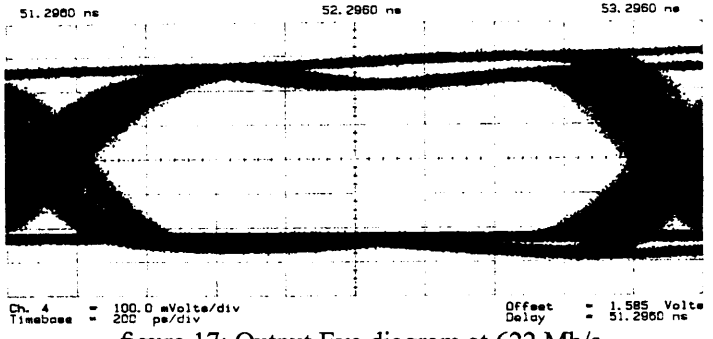


figure 17: Output Eye diagram at 622 Mb/s

Figure 17 shows Eye diagram at 622Mb/s, the most frequently used Bit rate for this kind of circuit. The eye is well opened and typical RMS jitter values of 17.7 ps at 1.244Gbs, 35ps at 622Mbs and 54ps at 155Mbs were observed.

AKNOWLEDGEMENTS

The achievements reported here are the result of a Team-work effort with designers of complementary backgrounds. The author would like to sincerely congratulate and thank Pierre Delerue, Didier Belot, Sebastien Dedieu, Sergio Zocchi, Nicole Bertholet, Christian Couteau and Mario Diaz Nava for their great contributions. Special thanks also to Joseph Borel for his encouragements in the Low Power direction.

10 . CONCLUSIONS

A large mixed-signal circuit was successfully designed in BiCMOS 0.5um process. High speed PLLs were integrated with 1 Mtr. CMOS processor using noise precautions techniques. Special efforts to achieve Low Power were paid in all sections from architecture to layout. Top-down methodology with HDLA macro-modeling was confirmed to speed-up design cycle. BiCMOS process device diversity was deeply exploited to reach minimum power and noise figures.

REFERENCES:

[1] ITU-T recommendation I-432. B-ISDN User-Network Interface Physical Layer Specifications, March 1993.

[2] ITU-T Recommendation G-707(Draft). Network Node Interface for the Synchronous Digital Hierarchy (SDH), November 1995.

[3] 622.08 Mbps Physical Layer Specifications, ATM Forum af-phy-0046.000, January 1996.

[4] A 3.3V Power Adaptive 1244/622/155 Mb/s TRANSCEIVER for ATM, SONET/SDH, Belot et al. ESSCIRC'97.

[5] "SONET Transport Systems: Common Criteria", BellCore GR-253-CORE, December 1995.

[6] Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal ICs, K.Su et al. IEEE JSSC, April 1993.

[7] SCA Preliminary Evaluation, Internal note ST-C.R&D, 5 June 1997.

[8] Modelisation du couplage substrat, B.Salvador, ENSERG June 1997.

MODELING NOISE COUPLING IN MIXED-SIGNAL / RF ICs

Nishath K. Verghese
Apres Technologies Inc.
Cupertino, CA 95014, USA

ABSTRACT

This paper reviews modeling techniques for noise coupling in mixed-signal / RF integrated circuits. Mixed-signal noise coupling is detrimental to integrated circuit performance, particularly in wireless RF circuits as it can swamp out the small received signal prior to amplification or during the mixing process. Modeling noise coupling between on-chip digital and analog components provides a designer with valuable insights into the impact of a noisy environment on IC performance, early in the design phase. Furthermore, the effectiveness of design techniques to mitigate noise coupling can be simulated using these models. Modeling mixed-signal noise coupling requires modeling of the IC environment including the substrate, package and chip interconnects. This paper describes a simple analytical technique to determine a first-order model of the substrate. In conjunction with a simple package model, it provides useful information on the noise coupling paths in a design. Additionally, several techniques for modeling the switching noise generated by digital circuitry are reviewed in this paper. Also, the impact of this noise coupling on the performance of sensitive RF circuitry is modeled using periodic transfer functions. Application of these noise coupling models to real-life problems is illustrated with the help of a design example.

1 INTRODUCTION

The goal of putting mixed-signal "systems on a chip" has been a difficult challenge that is only recently beginning to be met. Some of the more prevalent problems in integrating mixed-signal (analog and digital) circuits are the noise coupling mechanisms on-chip. Switching currents induced by logic circuits work in tandem with the chip/package parasitics to cause ringing in the power-supply rails and in the output driver circuitry. This in turn couples through the common substrate to corrupt sensitive analog signals on the same chip. In fact, substrate coupling is one of the main reasons causing many to put analog and digital on separate ICs. For some sensitive classes of circuits this may be the only alternative; separate ICs mounted in a multi-chip module, separate packages, or COB packaging. In many cases however, the analog can be made to coexist with the switching functions. This may be at the cost of strict partitioning of switching and non-switching functions, extensive special handling, a special semiconductor process, and perhaps even a fully custom design effort. Since a single-chip solution is often the smallest, lowest cost and lowest power implementation, the additional effort is often justified. As the number of logic switching functions increases, the degree of difficulty merging the digital and analog functions will increase [1].

A low-cost, single chip solution is also one of the primary design goals in several wireless applications. A single-chip transceiver is already a reality for certain relatively undemanding applications, such as RF ID, but IC manufacturers are not close to producing a similar device capable of operating in the more demanding application world of cellular telephony. In the meantime, incremental changes are helping provide consumers with handsets that offer major performance and size gains.The biggest obstacle to integrating cellular telephone components is on the receiver channel. As more and more channels are packed into the same range of frequencies, techniques have to be found to eliminate interfering signals. This is particularly difficult in cellular telephony due to the high dynamic range required [2]. For acceptable performance, a

cellular phone has to receive a signal clearly from a basestation miles away, while interfering signals blast away on adjacent channels. Interfering signals on wireless ICs can also arise due to the integration of analog and digital components on the same chip [3]. Compounding the problems of wireless circuit designers dealing with this mixed-signal noise coupling problem is that received signals on-chip are often only in the range of a few tens of microvolts. At signal levels in the microvolts, RF receiver circuits are sensitive to fundamental device noise, let alone mixed-signal substrate-coupled noise.

Several design considerations must be given to the realization of integrated mixed-signal / RF circuits. The potential of computer simulation in aiding these design decisions is becoming increasingly evident. Various techniques have been proposed in literature to determine suitable models of the substrate [4]-[11]. Several methods have also been proposed to macromodel the switching noise generated by on-chip digital circuitry [1], [12]-[14]. Furthermore, recent progress in the development of steady-state methods to simulate periodically time-varying circuits [15],[16] has made it possible to efficiently determine the coupling of noise through RF circuitry. Combining these different models, it is now possible to quickly analyze noise coupling from digital to analog components in mixed-signal / RF circuits [25]. This paper reviews these modeling techniques.

Section 2 discusses some of the problems particular to integrating mixed-signal wireless circuits. Section 3 reviews the noise coupling problem in mixed-signal ICs and the use of substrate, package, and interconnect models for its analysis. A simple analytical technique to determine a first-order substrate model is presented in this section. Section 5 illustrates the use of simplified circuit macromodels for the efficient simulation of noise coupling in mixed-signal/wireless ICs. It includes a review of various macromodeling techniques for switching noise introduced by on-chip digital circuitry. A macromodeling technique for the coupling through RF circuits using periodic transfer functions is described in Section 6. Section 7 illustrates the application of these noise

coupling models to a real-world design problem.

2 MIXED-SIGNAL WIRELESS IC PROBLEMS

During signal propagation and reception in a wireless channel, there exist many sources of noise such as interferers, channel noise and device non-linearity (due to mixers and other front end RF circuitry) that can affect the demodulated signal's integrity. Figure 1 shows some of the possible sources of noise during signal transmission and reception. These interferers and noise mechanisms can be categorized as channel noise sources, front-end receiver nonidealities and mixed-signal noise coupling.

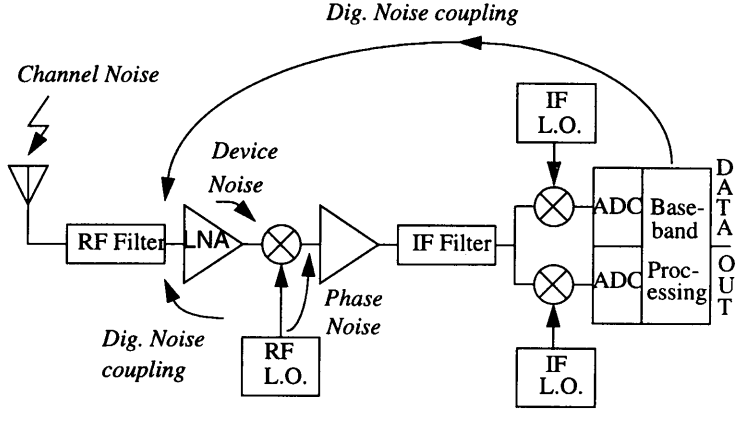


Fig. 1. Typical mixed-signal wireless IC problems.

Channel noise sources include background white Gaussian noise, atmospheric noise, radio propagation losses, Doppler effect, fast fading, multi-path signal losses, and adjacent channel interferers [18],[19]. The wireless channel is constantly changing over time due to the uncertainty

in the channel terrain, signal path propagation, and the mobility of the transceiver unit. Hence, the RF signal received at the antenna is a function of several factors including distance from the base station, signal path geography, obstacles between the transmitter and the receiver, and multi-path propagation of reflective signals. For a typical receiver, the incoming RF signal can range from 20 dBm to -120 dBm-Watts; i.e. with 50–Ω termination, the incoming signal strength can be as low as 0.22 μV modulated to 900MHz to 2GHz [18].

Furthermore, the receiver adds noise to the incoming signal due to RF circuit non-linearity, electronic device noise, and mixed-signal noise coupling. Receiver circuit nonlinearity results in spurious responses or apparent on-channel responses to undesired interfering signals. Intermodulation products during the mixing operation can create spurious responses in the signal frequency range. Other common spurious responses (spurs) include image, half-IF, and Able-Baker spurs [3]. Also, phase noise in the frequency synthesizer can cause undesired interferers to be mixed down, thereby corrupting the IF signal. Phase noise in the frequency synthesizer is caused both by device noise and by mixed-signal substrate-coupled noise mechanisms.

With greater emphasis on integration in modern wireless circuits, mixed-signal coupling is becoming a significant bottleneck in wireless IC development. On-chip integration of sensitive analog RF circuits with the noisy local oscillator (LO), frequency synthesizer, and other high-speed digital circuits can be detrimental to the weak incoming RF signal. The synthesizer consists of a high-speed divider/counter circuit which can produce significant switching noise. Any digital switching noise is of significant importance since cumulatively it can reach several hundred *mVolts* and propagate to the RF section via the substrate and I/O pins. The common substrate can also act as a dc and small-signal feedback path causing variations in dc operating conditions and small signal parameters such as gain and bandwidth [10]. In direct conversion receivers, substrate coupling can cause the LO signal to feedthrough to the LNA (low noise amplifier) input ("LO leakage"), where it gets amplified

and subsequently mixed with itself ("self mixing") resulting in dc offsets. A similar effect can occur if a large interferer leaks from the LNA or mixer input to the LO port and is multiplied by itself [20]. In the rest of this paper the focus will be on the mixed-signal noise coupling problem in integrated circuits.

3 MIXED-SIGNAL NOISE COUPLING

The main source of on-chip noise in mixed-signal ICs is the digital switching noise. In purely digital applications, the CMOS static logic family offers several attractive features including zero static power dissipation, high packing densities, wide noise margins, high operating frequencies, etc. For high-frequency wireless applications however, its major drawback is the generation of a large amount of digital switching noise [21]-[23]. When many static gates change states, a large cumulative current spike flows through parasitic resistances and inductances creating power supply noise voltage spikes known as '*Vdd bounce*' or '*Gnd bounce*'. Some fraction of this noise inevitably propagates to the sensitive analog circuitry through the substrate, power supply lines, bonding wires, package pins, etc., where it often limits the achievable accuracy. Figure 2 illustrates the typical noise coupling problem with a waveform observed on an industrial mixed-signal IC when all its twelve output buffers switched simultaneously [11]. The excessive ground bounce on the digital outputs coupled to the substrate through the junction capacitance of an output buffer device causing input comparators located on the same chip to trigger falsely.

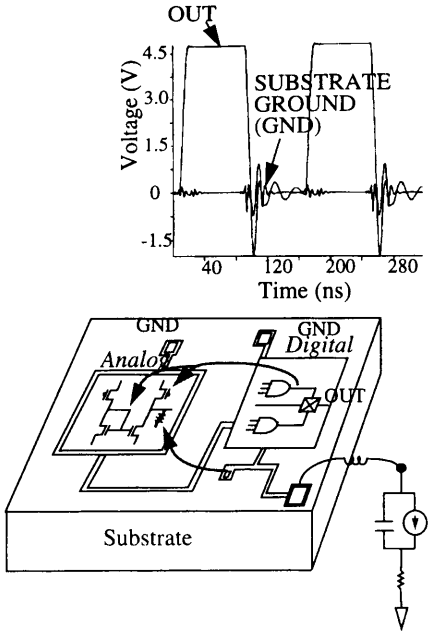


Fig. 2. Noise coupling from digital to analog circuitry via the substrate, package and interconnects in a mixed-signal IC [11].

Typically, the switching noise current can range from 0.1mA to several mA depending on the frequency of operation and device sizes. In a wireless IC, a received RF signal can range from 20 dBm to -120 dBm; i.e. the incoming signal can be as low as 0.22 μV. Considering that the received signal is at 1.8-2GHz, any internal on chip noise can destroy the desired signal. Besides being a coupling path for switching noise the common substrate is also a signal path. Impact ionization in MOS device channels can cause currents to be injected into the substrate even under DC operating conditions, causing substrate biases to vary, which in turn cause variations in MOS threshold voltages, depletion capacitances and other circuit bias and performance quantities. Signal coupling through the substrate can also lead to gain and bandwidth variation in LNAs and

self-mixing problems in direct conversion receivers. Moreover, signal loss in the substrate is a significant concern both in the design of receiver front-end circuitry and in the realization of on-chip high-Q inductors [24].

The overall modeling strategy for noise coupling effects in mixed-signal/RF ICs is to combine simple models of chip parasitics and the circuit for efficient simulation. Since simulating an entire mixed-signal or RF circuit at a transistor level is beyond the scope of traditional circuit simulators, circuit behavior has to be abstracted to a higher level wherever possible.

4 MODELING CHIP PARASITICS

Several chip parasitics must be modeled in an electrical circuit in order to analyze it for mixed-signal noise coupling problems. These include package inductance, device/interconnect capacitance, package/bondwire inductance and substrate resistance (and capacitance).

4.1 Device/Well/Interconnect Capacitance

Every transistor and well on an IC die is coupled capacitively to the substrate through its p-n junction depletion capacitances. Moreover, every interconnect routed on an IC has some capacitance to substrate. This capacitively coupled substrate current is of significant consequence in mixed-signal circuits, due to the presence both of a large number of switching digital nodes that inject current into the substrate and of high impedance analog nodes that are affected by this injected current. Since the amount of injected current is directly proportional to the slew rate of the switching voltage, at higher rates of circuit operation the substrate coupling problem is greatly aggravated. Moreover, with decreasing technology feature sizes, the interconnect capacitances to substrate are becoming increasingly important. To account for capacitively coupled substrate currents, a parasitic capacitance extraction can be performed on the design to determine all significant capacitances to substrate in the circuit. Well resistances in Fig. 3 can be calculated using the methodology described in Section 4.3.

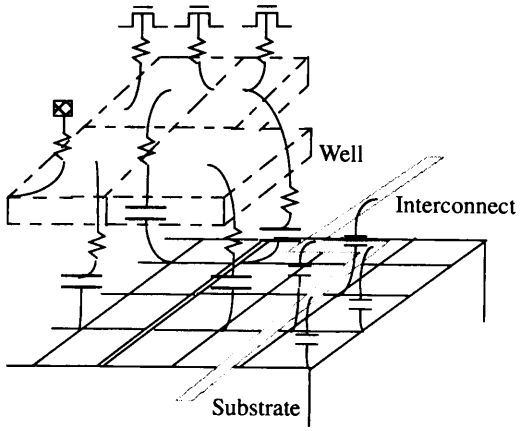


Fig. 3. Wells and interconnects capacitively coupled to substrate.

4.2 Package Inductance

The effect of non-ideal (inductive) power supplies has a tremendous impact on the amount of substrate coupled switching noise in an IC design. Since the bondwires and package pins associated with the substrate supplies have finite and often large inductances, any substrate current picked up by these supplies can cause large glitches in the value of the substrate supply bias. This phenomenon is referred to as inductive or Ldi/dt noise. The presence of parasitic inductances in the substrate supplies (Note: there are typically several separate digital and analog supplies connected to substrate) can severely aggravate the noise coupling problem and much of current mixed-signal IC design methodology focuses on techniques to minimize their effect. For simulation purposes, it is necessary to use suitable package inductance models in the supply leads to accurately analyze substrate-coupled switching noise [1],[5].

A simple chip-package model [1] is illustrated in Fig. 4 where RAPV, LAPV and RACV represent the package resistance, package inductance and on-chip resistance respectively in the analog VDD line while RAPG, LAPG and RACG represent similar parasitics in the analog ground line. CAC represents the chip VDD to GND capacitance while CAV and CAG represent the capacitances from the analog VDD and GND lines to substrate. Similar parasitics are also illustrated for the digital GND and VDD lines.

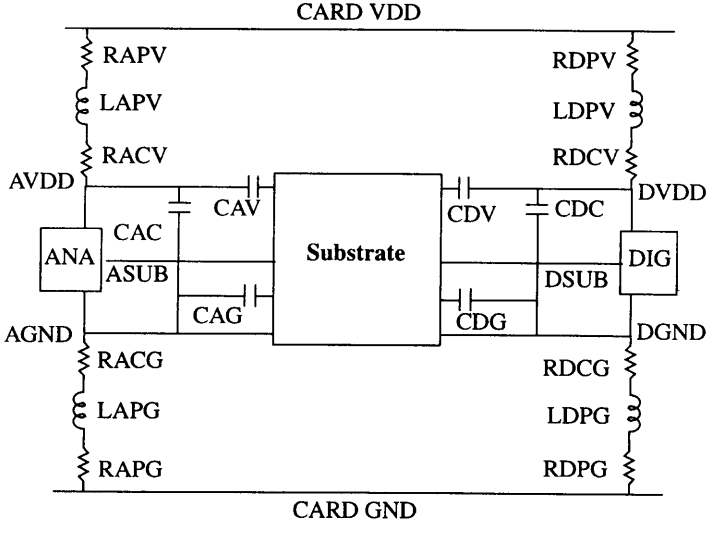


Fig. 4. An illustrative parasitic model of a chip in its package [1].

4.3 Substrate Resistance / Capacitance

Outside of the active areas formed by devices and substrate contacts, the substrate can be treated as consisting of layers of uniformly doped semiconductor material. Neglecting the effects of magnetic-fields on-chip, a simplified form of Maxwell's equations can be applied to the substrate yielding:

$$\frac{1}{\rho}\nabla^2 V(r,t) + \varepsilon\frac{\partial}{\partial t}(\nabla^2 V(r,t)) = -\dot{q}(r,t) \tag{1}$$

where ρ is the resistivity and ε the permittivity of the uniformly-doped semiconductor. $V(r, t)$ is the transient voltage vector and $\dot{q}(r,t)$ is the rate of generation of charge at location $r=(x,y,z)$ on the substrate. Assuming a 3-D semi-infinite substrate that goes to infinity in all but one of the six spatial directions as shown in Fig. 5, the solution to (1) in the Laplace domain for the voltage at any point on the substrate, v_2 due to a current, i_1 injected into the substrate a distance r away, is:

$$v_2(s) = \frac{\rho}{2\pi r}\cdot\frac{i_1(s)}{s(\rho\varepsilon)+1} \quad . \tag{2}$$

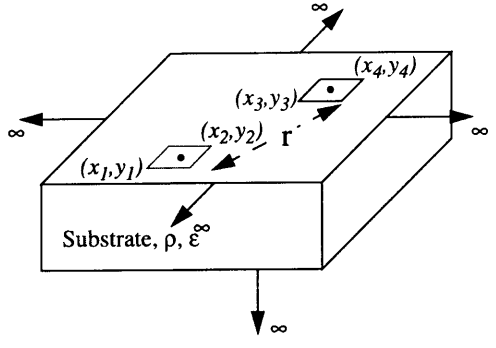


Fig. 5. Two "points" separated by distance, r on semi-infinite substrate.

Consequently, the substrate impedance $z_{21}(s) = v_2(s)/i_1(s)$ has a single pole response with a 3dB frequency given by the reciprocal of the relaxation time constant, $\tau = \rho\varepsilon$ of the substrate. When the two "points" (ports) have finite rectangular size we integrate (2) over their areas and divide the resulting equation by the product of the areas to give the following expression for substrate impedance [17]:

$$z_{12}(s) = \frac{z_{12}}{s(\rho\varepsilon)+1} \tag{3}$$

$$z_{12} = \frac{\rho}{2\pi K A_1 A_2} \sum_{k=1}^{4} \sum_{m=1}^{4} (-1)^{k+m} \left[\frac{b_m^2 - C^2}{2} a_k \ln(a_k + \eta) \right.$$

$$\frac{a_k^2 - C^2}{2} b_m \ln(b_m + \eta) - \frac{1}{6}(b_m^2 - 2C + a_k^2)\eta$$

$$\left. - C a_k b_m \operatorname{atan}\left(\frac{a_k b_m}{\eta C}\right) \right] \tag{4}$$

where

$$\eta = \sqrt{a_k^2 + b_m^2 + C^2} \tag{5}$$

$$a_1 = |x_1 - x_4| \;,\; a_2 = |x_2 - x_4| \;,\; a_3 = |x_2 - x_3| \;,\; a_4 = |x_1 - x_3|$$

$$b_1 = |y_1 - y_4| \;,\; b_2 = |y_2 - y_4| \;,\; b_3 = |y_2 - y_3| \;,\; b_4 = |y_1 - y_3| \tag{6}$$

$$A_1 = (x_2 - x_1)(y_2 - y_1) \tag{7}$$

$$A_2 = (x_4 - x_3)(y_4 - y_3) \tag{8}$$

The variables, $C=0$ and $K=1$ when both ports are on the substrate surface and $((x_1, y_1), (x_2, y_2))$, $((x_3, y_3), (x_4, y_4))$ are the coordinates of the rectangular ports. The self impedance term, $z_{11}(s) = v_1(s)/i_1(s)$ can also be calculated using (4) and has a single pole response with the same 3dB frequency.

Now, the net substrate admittance between the two ports, $y(s)$ can be determined from Fig. 6 as:

$$y(s) = (s(\rho\varepsilon)+1)\left(y_{12} + \frac{y_{10} \cdot y_{20}}{y_{10} + y_{20}}\right) \tag{9}$$

where $y_{10} = y_{11} - y_{12}$, $y_{20} = y_{22} - y_{12}$ and z_{11}, z_{22}, z_{12}, y_{11}, y_{22} and y_{12} are related as in (10).

$$(s(\rho\varepsilon)+1) \cdot \begin{bmatrix} y_{11} & -y_{12} \\ -y_{12} & y_{22} \end{bmatrix} = \frac{1}{s(\rho\varepsilon)+1} \cdot \begin{bmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{bmatrix}^{-1} \tag{10}$$

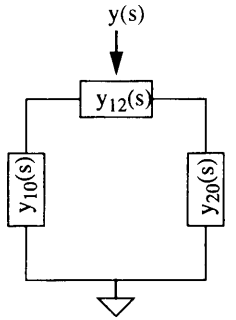


Fig. 6. Computing net substrate admittance between the two ports.

Consequently, the admittance $y(s)$ has a zero in its response at the same 3dB frequency and can be modeled as a parallel combination of conductance and capacitance:

$$y(s) = G + sC \tag{11}$$

$$G = y_{12} + \frac{y_{10} \cdot y_{20}}{y_{10} + y_{20}} \;;\quad C = (\rho\varepsilon) \cdot G \tag{12}$$

Note that the 3dB frequency scales with substrate doping and is nearly 150 GHz for a 1 Ω-cm substrate. Consequently, for most frequencies of interest in mixed-signal ICs, the substrate capacitance can be ignored.

One possible macromodel proposed by Smith and Schmerbeck [1] involves the use of chains of inverters to approximate logic switching activity as shown in Fig. 8. There are several inverters in each string to take into account the fact that the output switching occurs over a number of gate delays after the clock edge. The bottom inverter string is driven by a signal 180 degrees out of phase relative to the top inverter string to emulate simultaneous out of phase switching. Each inverter model includes a multiplication factor to increase the amount of output switching power without modeling large numbers of inverters and each inverter in the string can have a different multiplication factor [1].The multiplication factor for each inverter can be determined by simulating the logic circuit involved using an event-driven simulator to determine switching instants. Since switching activity is not guaranteed to be constant from clock edge to clock edge, the weighting factors may have to change from clock edge to clock edge.

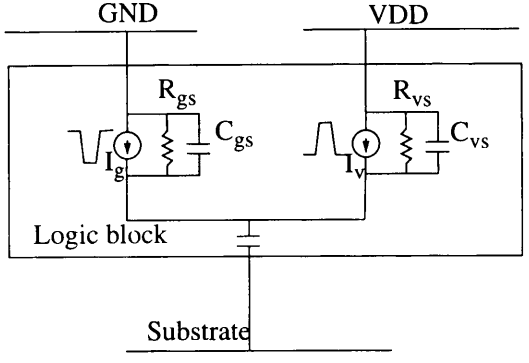


Fig. 9. Simple linear macromodel of logic block for switching noise[12].

A simple linear macromodel as shown in Fig. 9 has been used to quickly evaluate switching noise effects for power distribution synthesis [12]. The current sources I_v and I_g model the current switching on the

digital Vdd and ground rails. In addition, the finite resistance of the sources is modeled along with an equivalent capacitance across the sources. The capacitance to substrate represents the net junction and interconnect capacitances in the logic block.

A modeling methodology for digital switching noise introduced in [13] involves the precharacterization of the substrate current injected by each logic component as a function of its input switching pattern. Each signature noise signal is stored in a library consisting of all such signatures for every logic component of the circuit. For a given input pattern to the logic circuit, an event driven simulator is used to record every transition (event) in the circuit. Utilizing this event information and the precomputed signature noise signals for each gate and assuming spatial independence of all the noise sources, the net noise in the substrate is computed as a convolution of the impulse train of events and the corresponding noise signatures.

Although the assumption that the noise sources are spatially independent may fail in the absence of a heavily-doped bulk substrate, the aforementioned methodology is quite useful in determining simple linear macromodels of logic blocks as in Fig. 9. The current models, I_v and I_g for a given logic block can effectively be determined by convolving its noise signature with the event train associated with that block.

A simpler macromodeling strategy can be arrived at by assuming that coupling from the digital blocks to the substrate occurs solely through the substrate supply line [14]. Then, a macromodel as shown in Fig. 10 is constructed for all logic circuits connected to that supply. Since typically multiple digital supplies are used to bias the substrate, one such macromodel is required for each substrate supply used. The current I_{dvdd} flowing from VDD to ground in the macromodel of Fig. 10 can be computed using a suitable power analysis tool. CDC is the net chip VDD to ground capacitance while the resistors and inductors represent package parasitics. It should be noted that in using such a macromodeling approach, appropriate modifications must be made in designs where a

Kelvin ground is used to bias the substrate (i.e., a quiet ground different from the circuit ground), since in such cases the short-circuit displacement component of I_{dvdd} does not flow directly to the substrate as suggested in Fig. 10.

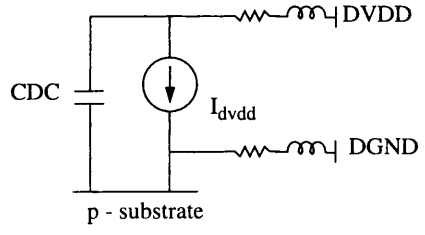


Fig. 10. Macromodel of switching noise in the digital substrate ground line.

6 MODELING RF / ANALOG CIRCUIT COUPLING

A characteristic of RF circuits is that they often process signals that contain closely-spaced large frequencies. Transient simulation of such circuits using a traditional circuit simulator can be extremely expensive because the simulator must take very small time steps to follow the large frequency carrier (LO) signal and simulate over a long period of time in order to allow the circuit to reach steady state. More efficient simulation techniques have emerged recently with the advent of efficient matrix-free iterative methods [15], [16], [26]-[28] for the periodic analyses of RF circuits. Characterizing a given RF circuit with periodic analyses is a two-step process. First, a periodic steady-state (PSS) analysis with only the large periodic stimulus applied is performed to set the periodic operating point of the given circuit. In a mixer for example, only the LO is applied whereas in a switched capacitor filter only the clock is applied. Finding the periodic steady-state solution of a circuit means finding the

initial condition for the circuit's associated system of differential equations such that the solution at the end of the period matches the initial condition. Once the periodic operating point is computed, a periodic small signal analysis is performed with the second tone applied to the circuit. This is done by linearizing the circuit about its periodic operating point and computing the steady state response of the periodically varying linear circuit assuming that it is driven with a small sinusoid of arbitrary frequency. The time required to compute the response of the linearized circuit is about the same as that required to compute the periodic steady-state response to the large stimulus alone, independent of the frequency of the sinusoid. Note that when applying a small sinusoid to a linear periodically time-varying circuit, the circuit responds with sinusoids at many frequencies. Hence, the periodic small signal analysis essentially computes a series of transfer functions, one for each frequency where the input and output frequencies are offset by the harmonics of the LO [27],[28].

Two periodic small signal analyses are possible. A periodic AC (PAC) analysis computes the small signal response from a single stimulus at a single frequency to every node in the circuit at every output frequency (i.e., input frequency and all frequencies offset from it by a harmonic of the LO). A periodic transfer function (PXF) analysis on the other hand, computes transfer functions to a single output at a single frequency from every source in the circuit at every input frequency (i.e., output frequency and all frequencies offset from it by a harmonic of the LO). The latter analysis is particularly useful in determining suitable macromodels of an RF circuit for the simulation of substrate coupling effects. Using a PXF analysis, it is possible to compute transfer functions from the bulk node of every device in the RF circuit to a specific output at a given frequency of interest. Note that for a given output frequency of interest this results in a set of transfer functions for each bulk node, where each transfer function models the frequency conversion effect inherent in the RF circuit. Hence, each transfer function models the gain of the circuit in converting a sideband frequency component of the bulk

node voltage to the specified frequency component of the output voltage.

Once the transfer functions have been computed for the RF circuit, it no longer needs to be represented at the transistor level. A transient simulation can be performed on the logic circuit using the switching noise macromodels and the parasitics coupling the RF circuit and the logic circuits through the substrate and package. The transient noise responses at the bulk nodes of the RF devices can be postprocessed to determine their equivalent noise spectra, and the frequency components of interest can be multiplied with the transfer functions determined earlier to compute the net spurious response at the prespecified output.

A similar macromodeling approach can be employed for analog linear time invariant systems where standard small signal transfer function analysis is used to determine transfer functions from the substrate to the circuit outputs over all frequencies of interest. Once the transfer functions are computed, a behavioral macromodel of the circuit can be formulated, taking into account both the circuit's functional and substrate coupling behavior.

7 A DESIGN EXAMPLE

The methodology described in this paper has been applied to the verification of the transmit section of a portable radio front end IC. Measured results on the fabricated IC indicated an RF spur (undesired or spurious signal) at the output of an up-conversion mixer (modulator) in the transmit section of the circuit. A 310 MHz carrier signal is divided by two and this drives the quadrature modulator. The modulating signals are at 1 MHz, so the expected modulator output is at 154 MHz or 156 MHz depending on whether I leads Q or vice-versa. Also present on-chip is a crystal buffer amplifier that is used to generate a clock signal at the reference frequency of 17 MHz. The measured RF spur at the output of the modulator is at 153 MHz which corresponds to the ninth harmonic of the 17 MHz clock frequency. Since this spur is present at the modulator output even in the absence of the 1MHz modulating signal, it was evident that this signal had coupled through the substrate and package to the output of the modulator.

Block diagrams of the modulator and reference frequency generator are shown in Fig. 11 and Fig. 12 respectively.

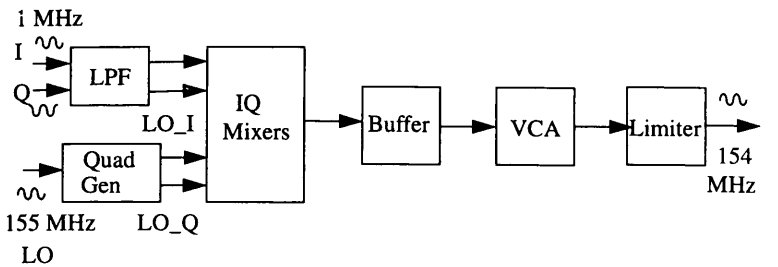


Fig. 11. Block-level schematic of the up-conversion mixer.

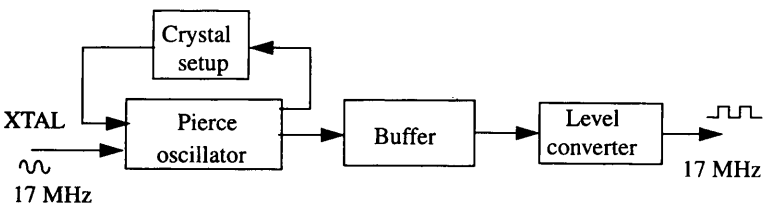


Fig. 12. Block-level schematic of the reference frequency generator.

The modulator has a image-reject architecture where the in-phase and quadrature-phase modulating signals are mixed with the in-phase and quadrature-phase LO signals respectively. The desired frequency component at the outputs of the mixers is the same while the undesired component differs in phase. A simple addition of the two mixer outputs can-

cels the undesired component. Each of the I and Q mixers is a conventional bipolar doubly-balanced commutating mixer as shown in Fig. 13 where the LO signal is sufficiently large to make the input transistors behave as switches. Therefore, each differential pair alternately steers the modulating signal (plus any bias I_Q) into the output.

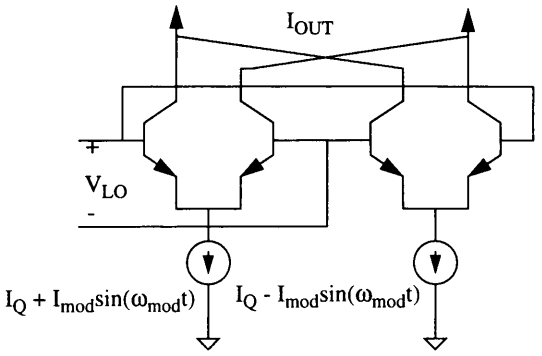


Fig. 13. Active doubly-balanced commutating mixer employed in the modulator.

The modulator and frequency generator blocks are part of a 16mm x 16mm single-chip radio front end fabricated in a BiCMOS technology on a substrate of resistivity,5 Ω-cm and thickness, 25 mils. The backside of the substrate is connected to a flag/heat sink through a non-conductive epoxy. The substrate is grounded via substrate contacts and guard rings on the substrate surface, surrounding the logic and analog circuitry. To isolate the sensitive circuits, different substrate supplies were used to bias the modulator circuit and the reference frequency generator. For further isolation, these two blocks are separated by a large distance on the lightly-doped substrate as shown in Fig. 14 which shows the layout of the frequency generator on the left and the modulator on the right separated by almost 5 mm on the 5 Ω-cm substrate. An extraction of the chip substrate showed that the most significant coupling between the two circuits is approximately 93 Ω between the two substrate ground lines. Together with the package/bond-wire inductance of 8 nH, at 153 MHz the substrate provides an attenuation of approximately 20 dB from the noisy digital ground line to the modulator ground. Note that neither substrate supply is a Kelvin ground. (i.e, each substrate ground is also a circuit ground)

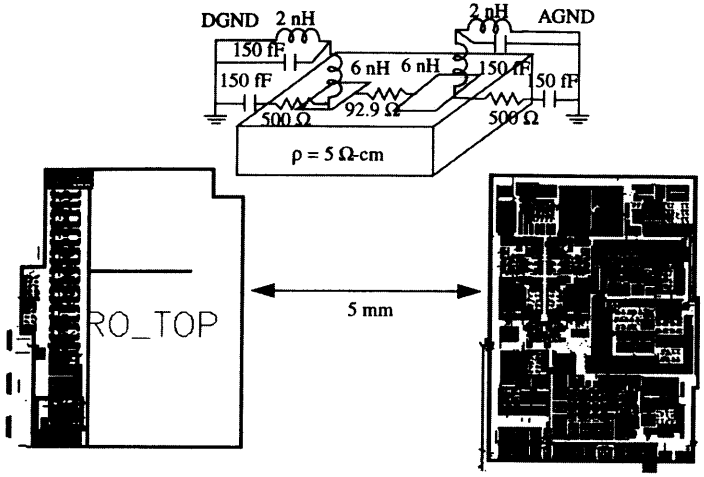


Fig. 14. Physical design of the modulator and crystal buffer amplifier on the chip substrate. Inset shows a simplified model of chip substrate and package.

The periodic operating point of the modulator under that application of just the carrier signal was first determined by performing a periodic steady-state analysis on the modulator at 155 MHz. Each of the differential outputs of the modulator under the large signal excitation of the carrier is plotted in Fig. 15.

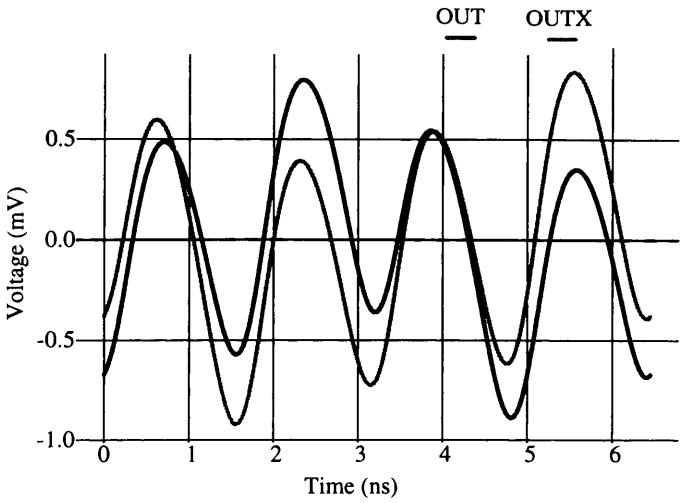


Fig. 15. Periodic steady state analysis output showing carrier feedthrough of modulator under application of the carrier signal only.

To verify operation of the mixer, a periodic ac analysis was performed with a small 1 MHz signal applied to the circuit at the modulating inputs of the circuit linearized about its previously computed periodic operating point. The PAC analysis results in output responses at frequencies, $f_{out} = f_{mod} + kf_{LO}$ where f_{LO} is the LO frequency (155 MHz) , k=..-2, -1, 0, 1, 2,.. is the harmonic number of the LO signal and f_{mod} is the modulating signal frequency (1 MHz). Results of the PAC analysis showing the carrier feedthrough (155 MHz) and undesired upper sideband (156 MHz) relative to the desired lower sideband (154 MHz) are plotted in Fig. 16.

261

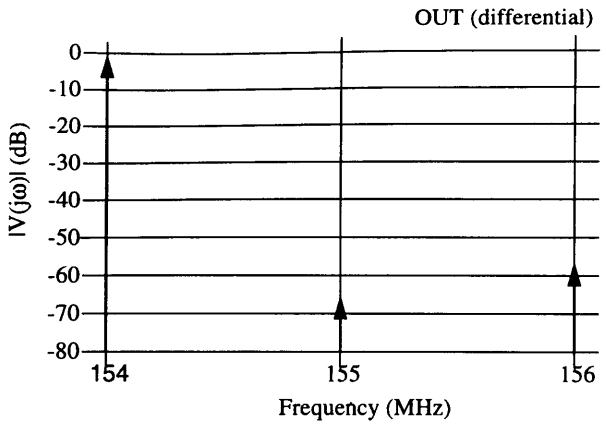


Fig. 16. Periodic ac analysis of up-conversion mixer under application of carrier and 1MHz modulating signal showing carrier feedthrough and undesired upper sideband at output relative to lower sideband.

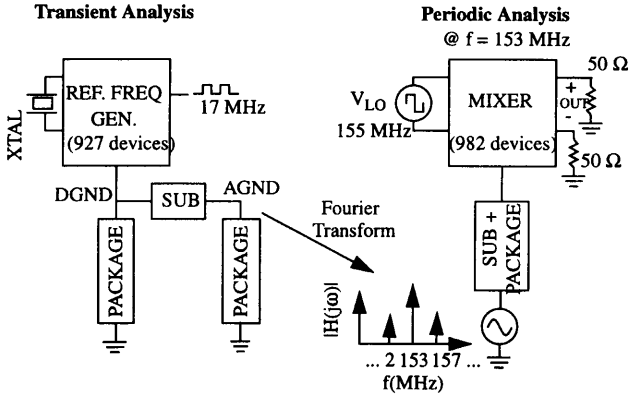


Fig. 17. The mixed-signal noise coupling analysis strategy for the RF IC.

The mixed-signal noise coupling analysis strategy to determine the output spur at 153 MHz for this circuit is illustrated in Fig. 17. First, a macromodel of the RF circuit is determined to characterize the coupling from the modulator substrate to the modulator output. To do so, a periodic small-signal transfer function analysis is performed on the modulator circuit linearized about its previously computed periodic operating point. For a given output frequency component of interest, f_{out} (in this case f_{out}=153 MHz), the PXF analysis results in transfer functions to the modulator output at f_{out} from the substrate at frequencies, $f_{sub} = f_{out} + kf_{LO}$ where k, and f_{LO} are the harmonic number and frequency of the LO signal as before. The computed transfer functions inherently model the conversion of spectral components in the substrate at 153 MHz, 2MHz, 157 MHz etc. to the output at 153 MHz due to mixing with the fundamental, first and second harmonics of the 155 MHz carrier frequency respectively. Next, to determine the actual noise at the modulator substrate, a transistor-level circuit simulation was performed on just the reference frequency generator and the substrate and package parasitics. (Since the reference frequency generator was small enough to be simulated at the transistor-level, no logic circuit macromodeling was used.)

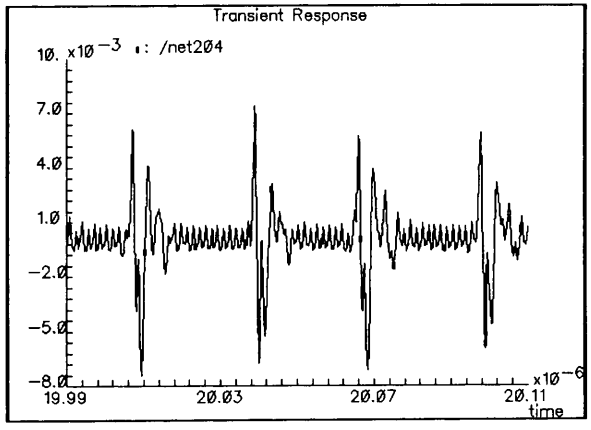


Fig. 18. Simulated transient noise waveform on the modulator substrate ground line.

Figure 18 shows the resulting transient noise waveform at the substrate ground line of the modulator. A Fourier transform of the transient noise waveform at the modulator substrate was then performed and the resulting spectral components at 153 MHz, 2 MHz, 157 MHz, etc. were multiplied with the previously computed transfer functions to the modulator output and summed to determine the net output spur at 153 MHz. Under application of only the LO (carrier) signal, the simulated differential spurious output of the modulator at 153 MHz is -72 dbm into a 50 Ω load which matches fairly well with the measured result of -75 dbm. Note that the spur is both measured and simulated in the absence of the 1 MHz modulating signal to isolate the mixed-signal coupling component at 153 MHz from the third order intermodulation product of the 155 MHz LO with the 1MHz signal.

With roughly 1900 devices, 717 nodes and 3234 equations to solve, a transient analysis of the modulator and reference frequency generator together would have required nearly 2 days of computation to simulate the 20 periods that are required for the modulator to attain steady state. Using periodic analysis instead, a macromodel of the modulator (containing 982 devices, 438 nodes and 1445 equations) was obtained in under an hour of CPU time. Another hour was required to simulate the transient noise coupling from the reference frequency generator.

8 CONCLUSIONS

Modeling techniques to for noise coupling in mixed-signal / RF integrated circuits were reviewed in this paper. Efficient simulation of mixed-signal noise coupling is possible using macromodels of the logic circuits, RF circuits, substrate and package. A simple analytical technique to determine a first-order model of the substrate was described. Several techniques for the macromodeling of logic switching noise were reviewed. Additionally, a technique to determine a macromodel of RF circuit coupling using periodic transfer functions was described. A design example illustrated an application of such models to a mixed-signal

RF IC design.

9 ACKNOWLEDGMENTS

Thanks are due to Dave Allstot, Tim Schmerbeck, Bob Stanisic, Sujoy Mitra, Ken Kundert, Keith Nabors, Dundar Dumlugol, Jacob White, Wen-Kung Chu and Won-Young Jung for several helpful discussions and to Andy Zocher of Motorola Cellular Research Labs for providing the design example

10 REFERENCES

[1]N.K. Verghese, T.J. Schmerbeck, and D.J. Allstot, *Simulation Techniques and Solutions for Mixed-Signal Coupling in Integrated Circuits*, Kluwer Academic Publishers, 1995.

[2]C. Huang, "RF Research chases elusive one-chip radio." *Op-Ed, EE Times*, Issue 959, pp. 45, June 23, 1997.

[3]S. Kiaei, D.J. Allstot, K. Hansen and N.K. Verghese, "Noise Considerations for Mixed-Signal RF IC Transceivers," *ACM Journal on Wireless Networks*, Special Issue on VLSI in Wireless Networks, Mar/Apr 1998.

[4]T.J. Schmerbeck, R. Richetta, and L. Smith, "A 27 MHz Mixed-Signal Magnetic Recording Channel DSP Using Partial Response Signalling with Maximum Likelihood Detection," *Tech. Digest of the IEEE International Solid State Conference*, pp. 136-137, Feb. 1991.

[5]D.K. Su, M.J. Loinaz, S.Masui and B.A. Wooley, "Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal Integrated Circuits," *IEEE Journal of Solid State Circuits*, vol. 28, no.4, pp. 420-430, April 1993.

[6]T. Smedes, N.P. van der Meijs, A.J. van Genderen, "Boundary Element Methods for Capacitance and Substrate Resistance Calculations in a VLSI Layout Verification Package," *Proceedings of ELECTROSOFT'93*, pp. 337-344, July 1993.

[7]B.R. Stanisic, N.K. Verghese, R.A. Rutenbar, L.R. Carley and D.J. Allstot, "Addressing Substrate Coupling in Mixed-Mode ICs: Simulation and Power Distribution Synthesis," *IEEE Journal of Solid State Circuits*, vol. 29, no. 3, pp. 226-238, March 1994.

[8]F. Clement, E. Zysman, M. Kayal and M. Declercq, "LAYIN: Toward a Global Solution for Parasitic Coupling Modeling and Visualization," *Proceedings of the Custom Integrated Circuits Conference*, pp. 537-540, May 1994.

[9] I. L. Wemple and A.T. Yang, "Mixed-Signal Switching Noise Analysis Using Voronoi-Tessellated Substrate Macromodels," *Proceedings of the Design Automation Conference*, pp. 439-444, June 1995.

10]R. Gharpurey and R.G. Meyer, "Modeling and Analysis of Substrate Coupling in Integrated Circuits," *IEEE Journal of Solid State Circuits*, vol. 31, no. 3, pp. 344-353, March 1996.

N.K. Verghese, D.J. Allstot and M.A. Wolfe, "Verification Techniques for Substrate Coupling and Their Application to Mixed-Signal IC Design," *IEEE Journal of Solid State Circuits*, vol. 31, no. 3, pp. 354-365, March 1996.

B.R. Stanisic, R.A. Rutenbar and L.R. Carley, "Addressing Noise Decoupling in Mixed-Signal ICs: Power Distribution Design and Cell Customization," *IEEE Journal of Solid State Circuits*, vol. 30, no. 3, pp. 321-326, March 1995.

13]P. Miliozzi, L. Carloni, E. Charbon, and A.L. Sangiovanni-Vincentelli, "SUBWAVE: a Methodology for Modeling Digital Substrate Noise Injection in Mixed-Signal ICs," *Proceedings of the Custom Integrated Circuits Conference*, pp. 385-388, May 1996.

[14]S. Mitra, R.A. Rutenbar, L.R. Carley and D.J. Allstot, "A Methodology for Rapid Estimation of Substrate-Coupled Switching Noise," *Proceedings of the Custom IC Conference*, pp. 129-132, May 1995.

[15]R. Telichevesky, K. Kundert, I. Elfadel and J. White, "Fast Simula-

tion Algorithms for RF Circuits", *Proceedings of the Custom Integrated Circuits Conference*, pp. 437-444, May 1996.

[16]R. Melville, P. Feldmann, and J. Roychowdhury, "Efficient multitone distortion analysis of analog integrated circuits," *Proceedings of the Custom Integrated Circuits Conference*, pp. 241-244, May 1995.

[17]A. E. Ruehli and P.A. Brennan, "Efficient Capacitance Calculations for Three-Dimensional Multiconductor Systems," *IEEE Transanctions on Microwave Theory and Techniques*, vol. 21, no. 2, pp. 76-82, Feb. 1973.

[18]G. C. Hess, *Land Mobile Radio System Engineering*, Norwood, MA, 1993.

[19]Rappaport T. S., *"Wireless Communications, Principle and Practice,"* IEEE Press, Prentice Hall PTR, 1996

[20]B. Razavi, "Design Considerations for Direct-Conversion Receivers, "*IEEE Transactions on Circuits and Systems - II, Analog and Digital Signal Processing*, vol. 44, no. 6, pp. 428-435, June 1997.

[21]G.H. Warren and C. Jungo, "Noise, crosstalk and distortion in mixed analog/digital integrated circuits," *Proc. of the IEEE Custom Integrated Circuits Conference*, pp. 12.1.1-12.1.4., May 1988.

[22]B.J. Hosticka and W. Brockherde, "The art of analog circuit design in a digital VLSI world," *Proc. of the IEEE International Symposium on Circuits and Systems*, pp. 1347-1350, April 1990.

[23]D.J. Allstot, S. Kiaei, and R.H. Zele, "Analog logic techniques steer around the noise," *IEEE Circuits and Devices Magazine,* vol. 9, pp. 18-21, Sept. 1993.

[24]J.R. Long and M.A. Copeland, "The Modeling, Characterization, and Design of Monolithic Inductors for Silicon RF IC's," *IEEE Journal of Solid State Circuits*, vol. 32, no. 3, pp. 357-369, March 1997.

[25]N. K. Verghese and D.J. Allstot, Verification of RF and Mixed-Signal Integrated Circuits for Substrate Coupling Effects," *Proceedings of the Custom Integrated Circuits Conference*, pp. 363-370, May 1997.

[26]R. Telichevesky, K. Kundert and J. White, "Efficient Steady-State Analysis based on Matrix-Free Krylov-Subspace Methods," *Proceedings of the Design Automation Conference*, pp. 480-484, June 1995.

[27]R. Telichevesky, K. Kundert and J. White, "Efficient AC and Noise Analysis of Two-Tone RF Circuits," *Proceedings of the Design Automation Conference*, pp. 21-3.1 - 21-3.6, June 1996.

[28]R. Telichevesky, K. Kundert and J. White, "Receiver Characterization using Periodic Small-Signal Analysis," *Proceedings of the Custom Integrated Circuits Conference*, pp. 449-452, May 1996.

TOP-DOWN DESIGN OF MIXED-MODE SYSTEMS: CHALLENGES AND SOLUTIONS

Georges G. E. GIELEN*
Katholieke Universiteit Leuven, ESAT-MICAS
Kardinaal Mercierlaan 94, B-3001 Leuven, Belgium
e-mail : gielen@esat.kuleuven.ac.be

*also research associate of the Belgian National Fund of Scientific Research (FWO-V)

ABSTRACT

An overview is presented of the challenges and available solutions in applying top-down design to mixed analog-digital systems. The progress in VLSI technology allows the integration of complex systems on a chip, containing both analog and digital parts. In order to boost the design productivity and guarantee the optimality of such systems while meeting the time to market constraints, a systematic top-down design approach has to be followed with sufficient time and attention paid to system-level architectural design before proceeding to the detailed block design. Also the reuse of macrocells will be unavoidable. System exploration tools will be presented that enable analog-digital co-design and allow to analyze architectural alternatives and trade-offs such as analog-digital partitioning. The crucial elements for analog high-level design (analog behavioral models and power estimators) will then be described in detail. Finally, an overview is given of current techniques for analog circuit synthesis and for mixed-signal layout assembly. Clearly, a top-down approach will be unavoidable for designing future complex systems, but also the remaining challenges to be solved for will be highlighted.

1. INTRODUCTION

Advances in deep submicron VLSI integrated circuit processing technologies offer designers the possibility to integrate more and more functionality on one and the same die, enabling in the near future the integration of complete systems that before occupied one or more printed circuit boards onto a single piece of silicon [1]. An increasing part of these integrated systems contain digital as well as analog circuits, and this in application areas like telecommunications, automotive and multimedia among others.

The drive towards integrated mixed-mode analog/digital ICs is however posing a big problem. The complexity of the systems that can be integrated on a single ASIC can only be mastered by using advanced computer-aided design (CAD) tools. For the digital circuits, commercial simulation and synthesis tools, especially at the logic and layout level, have been around for some years now and a considerable part of the digital design flow has been automated, starting the design from a RT language description. For more complex integrated systems-on-a-chip the use of soft or hard IP (intellectual property) blocks or so-called "virtual components" is expected to provide the productivity boost needed to realize these systems in the required time to market constraints [2]. On the analog side, however, the picture is much more clouded. The growing interest in mixed-signal ICs is exposing the lack of mature analog CAD tools that can boost the efficiency of analog designers. Despite the progress in academia on analog circuit and layout synthesis [3], the commercial offerings are basically limited to some clone of the SPICE simulator and an interactive layout environment. Due to this lack of supporting CAD tools and also due to the more knowledge-intensive nature of analog design, analog design and layout productivity cannot keep up with the time to market constraints of present-day applications. Therefore, although the analog circuits typically occupy only a small part of the area in mixed-signal ICs, they require a disproportionally large part of the overall design time (see Fig. 1) and are often responsible for design failures, necessitating expensive redesign runs. In addition, the specific problems posed by the integration of analog and digital circuits on the same chip (such as for instance substrate noise crosstalk) are not sufficiently handled by existing digital-oriented

tools. As a result, many important tools are today still missing to cover the design flow of mixed-mode ICs.

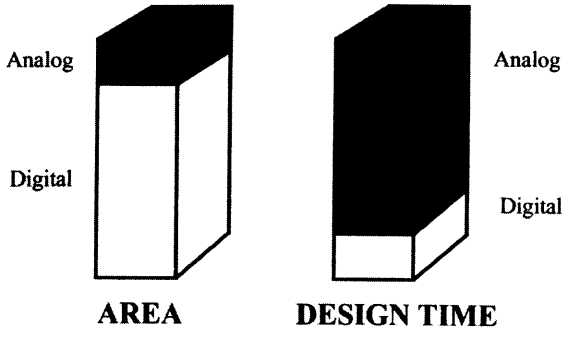


Fig. 1. Problem of the low analog design productivity compared to digital design.

This paper will discuss a top-down design methodology for mixed-signal ASICs and the subtleties to implement this in practice. The methodology will be described in section 2. A technique for analog high-level synthesis will be presented in section 3. Analog behavioral modeling and power/area estimation are important parts of the hierarchical design flow and will be discussed in detail. Analog circuit-level synthesis will be described in section 4, while the layout synthesis will be discussed in section 5. The methodology will be illustrated with several examples. Conclusions will be drawn in section 6.

2. MIXED-MODE TOP-DOWN DESIGN METHODOLOGY

The ideal design flow for a mixed analog/digital ASIC is schematically shown in Fig. 2. It consists of four basic steps [4] :

1) ***system specification*** - First the functionality of the complete ASIC is to be described in some formal way independent of the way of implementation (the format could be a language description or something else). This system specification can then be verified with system-level simulations to check the functionality of the chip within its system context. This specification phase is a very crucial step in the design of the ASIC.

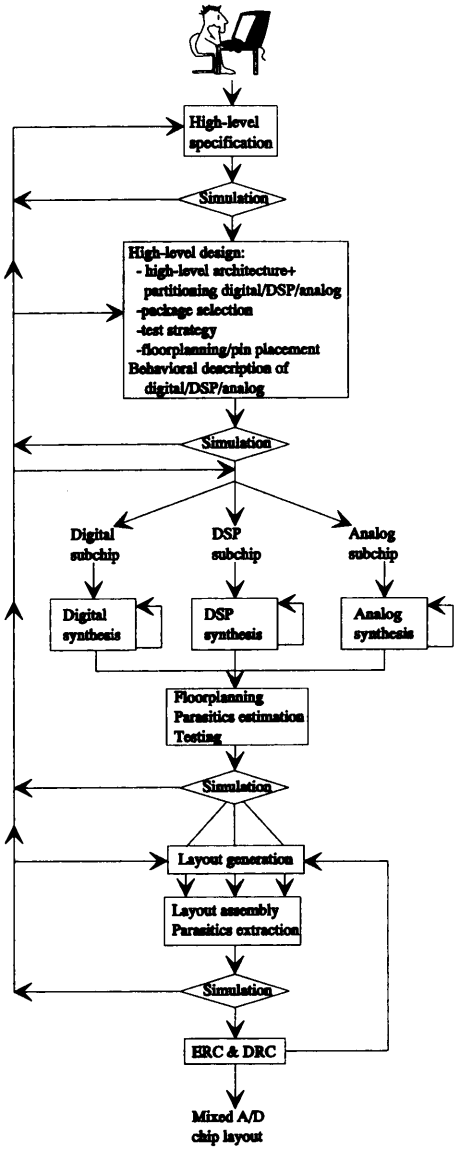


Fig. 2. Outline of the top-down design flow of a mixed-signal IC.

2) ***high-level architectural design*** - The next step is the design of a high-level architecture (a kind of block diagram), including the partitioning in digital, DSP and analog parts, as well as issues like package selection, pin placement and initial floorplanning, test strategy development... Each of the different parts (digital, DSP, analog) is then described separately by high-level (generic) behavioral models. A mixed-mode simulator is used to verify the correctness and consistency of this description with the system specification.

3) ***block synthesis*** - Starting from the behavioral descriptions, the different parts are then synthesized separately, since the different types of circuits (digital, DSP, analog) require different approaches and tools. The analog approach will be further detailed below. Block synthesis also has to be performed when using soft IP. Afterwards, the floorplan can be generated and the wiring parasitic estimates updated. Testability is included here as well. The complete chip can now be simulated to verify the chip performance. When the specifications are not met, redesign iterations have to be carried out, for example by trying a different partitioning.

4) ***layout generation*** - If the chip performance is within the specifications, the layout can be generated. Different tools have to be used to generate the layouts of analog and digital blocks separately according to the constraints of the floorplan. (For hard IP blocks this step is trivial.) The layouts then have to be assembled by mixed-signal block place and route tools and the real interconnect parasitics can be extracted. The system performance then has to be verified again. If the simulated chip performance is still within the specifications, mask production can be started. If not, new design iterations are necessary. The verification of entire integrated systems most likely requires the use of detailed behavioral models extracted from the lower-level implementations in order to make the system simulations feasible.

2.1. Design flow for the analog blocks

For the design of the analog parts in these mixed-mode systems most (academic) analog synthesis approaches today adopt a top-down hierarchical performance-driven design methodology [5,6,7]. This means that the design is performed over different hierarchical levels and is driven by the performance specifications. As opposed to the digital world, however, the hierarchical levels do not correspond to different levels of signal abstraction, but rather to a functional decomposition where large higher-level blocks are decomposed into

smaller lower-level subblocks in order to make the total design tractable. For example, an analog-to-digital converter contains a comparator as a subblock. For the design of the converter many implementation details of the comparator can be made abstraction of as long as it implements the required functionality. This also implies that the subblocks can be represented by behavioral models during the design of the higher-level block. A behavioral model models the input-output behavior of the corresponding block with the block's circuit specifications as parameters of the model.

The design strategy followed in the considered analog design hierarchy can then be summarized as in Fig. 3. The starting point at any level are the specifications for the block under design. In between any two hierarchical levels *i* and *i+1* the following steps are then performed :

a) top-down path :

- **topology selection**, where for the block under design a proper topology is chosen that is capable to meet the specifications. This topology is defined as an interconnection of subblocks.

- **specification translation** where the specifications of the block under design are translated or mapped into individual specifications for each of the subblocks in the selected topology. At the lowest hierarchical level the subblocks are single devices and specification translation reduces to sizing.

- **verification** of the design using simulations. At higher levels these simulations require the use of generic behavioral models for the subblocks, since the subblock implementations are not yet known at this stage. At the lowest hierarchical level these can be SPICE simulations.

These top-down steps are then repeated for each of the subblocks down the hierarchy until a level is reached that allows a physical implementation, at which moment the bottom-up path is started. This level normally is the device level, but could also be a higher level in case analog standard cells or IP blocks are used.

b) bottom-up path :

- **layout generation** where the layout of the block is assembled from the layouts of the subblocks. At the lowest hierarchical level this is device-level layout generation. At higher levels this is block place and route.

- layout extraction followed again by **verification** of the design using detailed simulations (this time including the extracted layout parasitics). At higher levels the blocks become so complex that device-level circuit simulations become untractable, and again behavioral models have to be used. The difference is now that the behavioral models must accurately represent the behavior of the actual implementations, and therefore have to be extracted from the lower-level implementations.

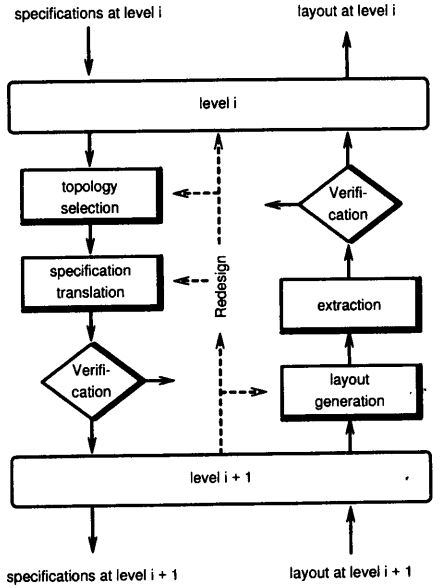


Fig. 3. Hierarchical analog design strategy.

2.2. Advantages and difficulties of the top-down design flow

There are a number of difficulties when applying this methodology to analog blocks. First of all, the definition of the analog hierarchy is not at all clear nor standardized. Splitting constraints over subblocks is most feasible only when the subblocks are not too much interacting with each other; otherwise decomposition is not effective. Secondly,

even at higher levels low-level details (e.g. matching limitations, circuit nonidealities, layout effects...) may be important to determine the feasibility or optimality of a solution. The behavioral models used at higher levels must therefore include such effects as well to the extent possible. Since however not everything can be anticipated or modeled accurately at higher levels, design iterations up and down the hierarchy are very typical in analog designs. Finally, in order to take meaningful decisions at higher levels in the design hierarchy, information from lower-level implementations has to be available at these higher levels. This information can be under the form of feasible performance space and power/area estimators, which indicate the obtainable performances of a block, respectively return an estimate of the power/area needed to implement the block for given specification values. The first guarantees that no unrealistic solutions are being explored at higher levels in the design hierarchy, the second allow to assess the optimality of alternative solutions. Generating all this information from the bottom up however requires a lot of work (to be repeated for every new technology process).

Nevertheless, hierarchical top-down design is unavoidable for complex designs, which require a divide and conquer strategy to be completed successfully. This is for instance already the case for Delta-Sigma A/D converters, which are typically first designed at the level of the converter with more or less ideal integrators and comparators, and only in a subsequent step are the integrators and comparators designed in detail. A similar approach is clearly needed for even more complex systems. The main advantages of the top-down design flow are then :

- ***reduction (or avoidance) of design iterations***: by performing more analysis and simulation at higher levels in the design hierarchy before starting the detailed implementation of the individual blocks, many problems can be anticipated, e.g. timing problems at the analog/digital interface. The use of "executable specifications" and high-level block descriptions ("virtual prototypes") will also result in a clear and better coordination between different design teams, e.g. between the system designers and the circuit designers. All this largely reduces the need for design iterations up and down the hierarchy, speeding up the overall design cycle time.

- ***more optimal designs*** : another advantage is that careful exploration and optimization at higher levels in the design hierarchy results in far bigger savings in power and silicon real estate (e.g. by choosing a proper system architecture) than all detailed power/area grindings at the lowest level. The designers can perform a trade-off analysis for the overall system instead of looking at every block on its own.

The next section will now describe the analog high-level synthesis step in more detail, whereas section 4 later on will describe analog circuit-level synthesis.

3. ANALOG HIGH-LEVEL SYNTHESIS

Analog high-level synthesis is the translation of analog system-level specifications into an architecture of subblocks, in which the specifications of all subblocks have been completely derived. During the top-down design phase the subblocks however are not yet device-level circuit implementations. Hence they have to be represented as behavioral blocks modeled by their functional input-output behavior.

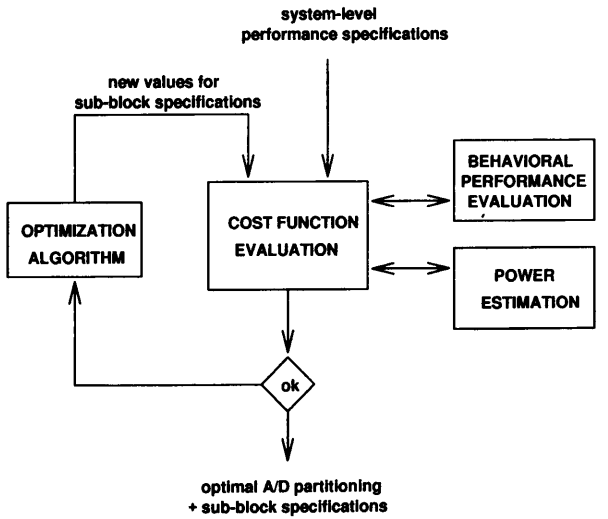


Fig. 4. Flow of analog high-level specification translation and analog-digital partitioning.

The translation of the system-level specifications into individual specifications for the subblocks can then occur in either of two ways: 1) either the whole mapping is encoded in a predeveloped design plan, like was the case in tools such as IDAC [8] or OASYS [5], or 2) optimization techniques are used to find the optimal set of subblock specifications [7,9,10,11]. The latter approach is the most flexible, and requires the least development effort per architecture. The approach is schematically depicted in Fig. 4. The optimization algorithm searches for the optimal set of subblock specifications such that the overall system meets the targeted specifications and at the same time a user-defined design objective (e.g. the overall power consumption) is minimized. This implies that at each iteration of the optimization both the performance of the system and the optimization targets are evaluated. The performance can be evaluated by carrying out a behavioral simulation of the archictuture, and/or by evaluating (if available) a set of equations that describe the architecture performances in terms of the subblock specifications which are the optimization parameters. If equations can be derived, the equation-based approach is in general faster, but the simulation-based approach is more generally applicable. The design targets can be estimated using power and area estimators that describe the power or area required to implement a certain block as a function of the block's specifications. This analog high-level synthesis approach will now be illustrated with some examples. The important basic elements (behavioral models and power/area estimators) will then be described afterwards.

3.1. Examples of analog high-level synthesis

3.1.1. ORCA : RF front-end exploration and optimization

A first example is obtained with the ORCA tool [12] which allows the architectural-level exploration and optimization of analog RF receiver front-ends. The tool contains a library with behavioral models and power estimation functions for the typical blocks encountered in receivers. The behavioral model of each block includes both its nominal behavior (amplification, filtering, mixing) as well as its most important nonidealities (e.g. noise, distortion, aliasing, phase noise, etc). In ORCA the simulations are performed with a dedicated

frequency-domain simulation algorithm that processes stochastic input power spectra typical for the targeted application (e.g. GSM, DECT). The power spectral distributions at every node in the circuit are calculated in a very efficient though slightly approximate way to enable a short response time of ORCA to the designer during system exploration [12].

The tool allows the designer to interactively explore alternative RF receiver topologies and to investigate design trade-offs within each topology at the architectural level, before designing each individual block. As the performance analysis routine has also been integrated within an optimization loop, the tool can also perform an optimal high-level synthesis of a given topology towards a specific application. It then determines the optimal specifications for the individual RF building blocks such that the complete RF front-end meets the required receiver signal quality while the overall power consumption of the entire topology is minimized. The power values are obtained from the power estimation functions of the different blocks.

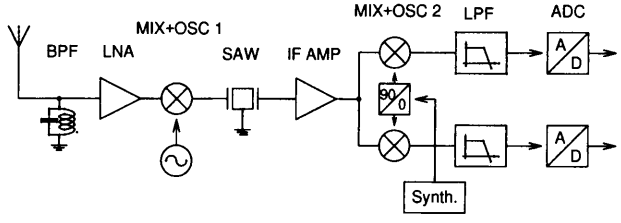


Fig. 5. A combined IF zero-IF receiver topology.

Fig. 5 shows an example of an RF receiver topology, a combination of an IF and a zero-IF receiver. The performance of a receiver is characterized by the ratio between the power of the undemodulated wanted signal at the output (after the A/D converter) and the total power of all the unwanted signals that are located in the same frequency band as the wanted signal. These unwanted signals include noise, distortion, aliasing and phase noise contributions, introduced by the different blocks in the topology. Fig. 6 shows a typical output of a simulation run with ORCA for the topology of Fig. 5. The topology and the building block specifications have been provided by the user

of the tool. The signal band of interest in this example is 100 kHz. The power spectrum of the wanted and the different unwanted signals (noise, aliased signals and distortion) are displayed separately.

The optimal total power consumption of the topology of Fig. 5 for different required signal quality values as resulting from repeated ORCA optimization runs is shown in Fig. 6. The optimum building block specifications for a receiver quality of 40 dB are summarized in Table 1. The reliability of these results of course depends on the accuracy of the behavioral models and the power estimators, and the values therefore only have to be considered as initial starting point for the designers.

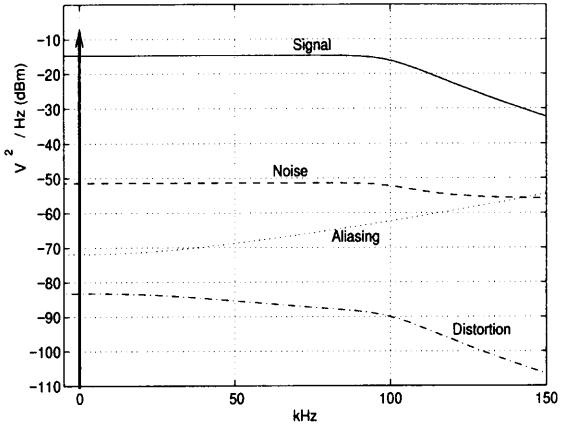


Fig. 6. Typical ORCA simulation result showing the spectra of wanted and unwanted signal components separately.

The optimal total power consumption of the topology of Fig. 5 for different required signal quality values as resulting from repeated ORCA optimization runs is shown in Fig. 7. The optimum building block specifications for a receiver quality of 40 dB are summarized in Table 1. The reliability of these results of course depends on the accuracy of the behavioral models and the power estimators, and these values therefore only have to be considered as initial starting point for the design.

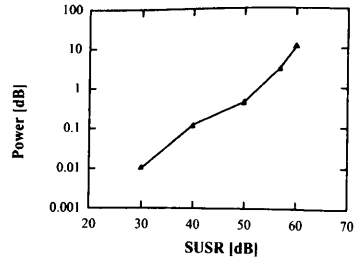


Fig. 7. Optimal power consumption of the complete receiver versus required signal quality.

$block_i$	F_i [MHz]	BW_i [MHz]	A_i [dB]	DR_i [dB]	P_i [mW]
BPF	910.0	54.0	0.0	∞	0.0
LNA	0.0	2680.0	28.0	50.0	30.5
MIX 1	793.0	1738.0	3.4	55.0	14.5
SAW	117.0	55.0	0.0	∞	0.0013
IF AMP	0.0	410.0	14.0	53.0	2.8
MIX 2	117.0	384.0	14.0	45.0	1.1
LPF	0.0	1.23	1.2	74.0	0.126
ADC	0.0	8.61	0.0	65.0	0.216

Table 1. Resulting optimized specifications for the building blocks of an RF front-end receiver.

3.1.2. Analog-digital partitioning

Another important problem in mixed-signal system design is finding the optimal partitioning between analog and digital [13]. Today this partitioning is often performed ad hoc, with some rough calculations, by an experienced system designer. It is however often not feasible to investigate many alternative solutions in the design space. And although there might be good reasons for implementing as many functions in the digital domain as possible, two problems might result. Firstly, the specification constraints on the few remaining analog blocks might become too stringent to realize in the given technology. For instance, the required analog-to-digital converter specifications in terms of resolution and bandwidth might become unrealizable.

Secondly, the digital solution might require more power than the solution with more analog functionality. Therefore, high-level exploration can quantitatively find the optimal analog-digital partitioning in mixed-signal systems. Optimal here means lowest overall power consumption for the complete system. The method is intended to assist system designers with quantitative data in their comparing different alternative system architectures.

The approach is an extension of Fig. 4 where the optimization algorithm now not only optimizes the specifications of the subblocks in the system architecture but also some parameters controlling small architectural variations (in casu the boundary between analog and digital). This is illustrated for the baseband signal processing part of a direct conversion receiver as shown in Fig. 8. The partitioning trade-off is in the position of the analog-to-digital converter and the amount of filtering in the analog versus the digital domain. Hence, the order of the analog low-pass filter and of the digital FIR filter are the architectural optimization variables. Subblock specifications include for instance the resolution and the speed of the analog-to-digital converter. The main system-level specification to be met is a 10-dB signal-to-noise requirement at the output of the digital multirate filter. The system is simulated using the ORCA tool [12], extended with behavioral simulation models for digital (multirate) filters.

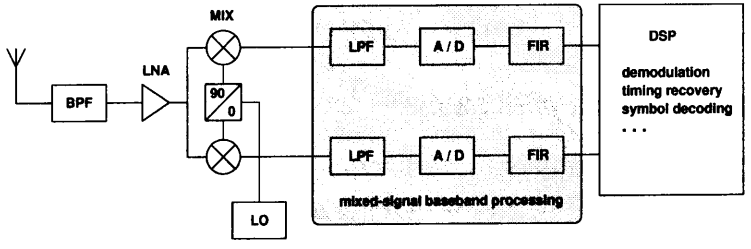


Fig. 8. Baseband signal processing channel in a direct conversion receiver.

The resulting (estimated) power consumption varies as a function of the order of the analog filter (and inversely also of the digital filter), as shown in Fig. 9. The smallest overall power consumption is obtained with a 5th-order analog filter and a 28-tap digital filter. Hence this also

determines the optimal insertion point of the analog-to-digital converter. Fig. 9 clearly shows that there is a lot of power to be gained by carefully making the trade-off between analog and digital filtering in this example (note that the power axis is logarithmic !).

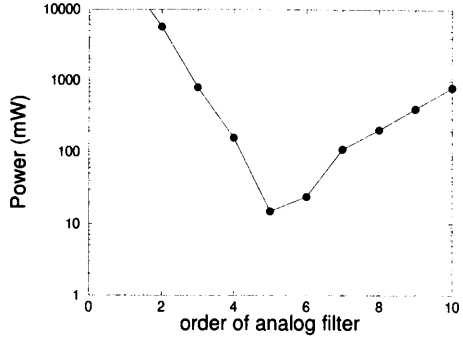


Fig. 9. Power consumption of the baseband signal processing channel as a function of the amount of analog filtering (and inversely the amount of digital filtering).

3.1.3. Particle detector front-end optimization

Another example of analog high-level synthesis is the front-end for a particle detector read-out system [9] as shown in Fig. 10. The most important system-level specifications for this interface are speed (frame time), accuracy, sensitivity and power consumption. Again the general flow of Fig. 4 is followed. The optimization variables are the design parameters of the subblocks used in the analog architecture, e.g. the number of bits and speed of the ADCs or the gain, speed and noise performance of the CSA and PSA. Also some architectural variations are being optimized, such as the number of analog memory cells M and the number of parallel ADCs N. The cost function to be minimized contains two sorts of terms: (1) terms that penalize the solution if the system performance specifications are not satisfied, and (2) the optimization targets such as power or area. In this example the performance evaluation at each iteration of the optimization is performed using a general-purpose analog behavioral simulator such as SABER, in combination with analytic formulas for selected characteristics that are too time-consuming to simulate. An example of

such an equation is the expression for the pile-up error which otherwise would require a full transient simulation.

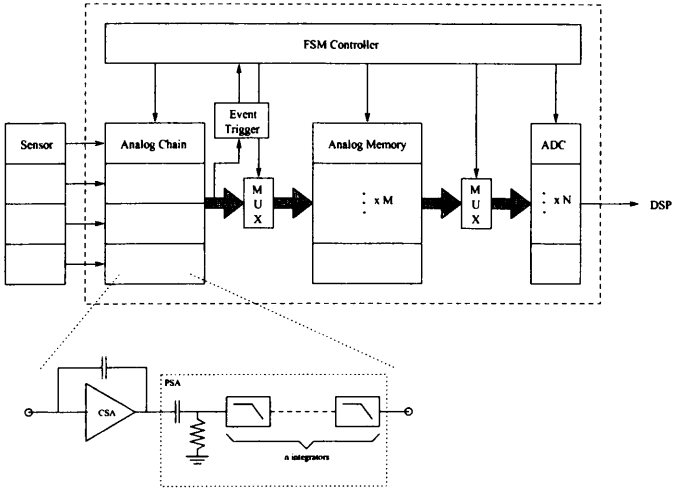


Fig. 10. Block diagram of particle detector front-end.

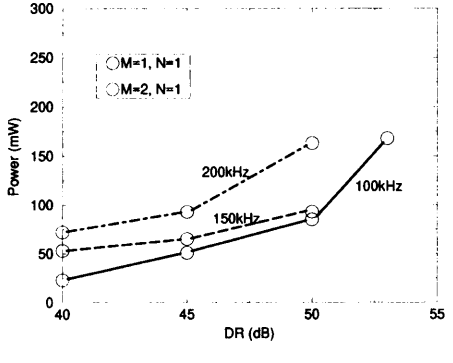


Fig. 11. Result of high-level synthesis of the particle detector front-end.

Fig. 11 shows the results of a number of optimizations with different speed (in kHz) and accuracy (dynamic range) requirements for the radiation detector interface. The mimimum power is indicated together with the architectural variation that has been chosen. Clearly, when the speed requirement increases, more analog memories are needed to

temporarily store input samples. Table 3 shows the specifications for the CSA-PSA subblock as resulting from the optimization with the system-level specifications of Table 2 as input. Similarly the specifications for the other subblocks are derived.

system specification	unit	value
radiation detector capacitance	pF	80
number of input channels	-	4
technology	-	0.7 um CMOS
power supply voltage	V	-2.5/+2.5
minimum energy level	keV	10
maximum energy level	keV	1600
accuracy	keV	5
frame time	us	≤ 5
overall power consumption	mW	minimal

Table 2. System-level specifications for the particle detector front-end.

CSA-PSA specification	unit	value
radiation detector capacitance	pF	80
technology	-	0.7 um CMOS
power supply voltage	V	-2.5/+2.5
maximum peaking time	us	1.1
noise	RMS e	≤ 800
gain CSA	mV/fC	20
gain integrator stage PSA	-	2.7
gain-bandwidth CSA	MHz	50
output voltage range	V	-1/+1
estimated power consumption	mW	21

Table 3. Specifications for the CSA-PSA subblock as derived by analog high-level synthesis.

3.2. Analog behavioral model generation

A crucial element for the hierarchical top-down approach to succeed is the disposition of analog behavioral models that describe the behavior of an analog block without knowing the details of the underlying circuit implementation. These models must describe the desired behavior of the block, while still including the major nonidealities of real implementations with sufficient accuracy. For this purpose a

library of generic behavioral models for frequently used analog blocks can be developed. For example, the ORCA tool [12] contains such a library for RF front-end applications. However, if the nonidealities are not well modeled in the models or the models generate phantom phenomena, then non-optimal or in the worst case erroneous design decisions might be taken at higher levels in the design hierarchy. For example, some of the subblock specifications in the selected architecture (e.g. the phase noise spec of a local oscillator) might turn out not to be realizable in practice, requiring overall design iterations across the hierarchy. On the other hand, including all effects before a circuit implementation is known is not possible either. Therefore, it can never be guaranteed that all effects can be anticipated at a higher level (e.g. layout parasitics). This problem is also true in today's very deep submicron digital design where interconnect heavily impacts the overall delay.

The situation is different during bottom-up system verification. In this case the underlying implementation is known in detail, but the goal of behavioral modeling is to reduce the time required to simulate the block as part of a larger system. In this case peculiarities of the block's implementation should be reflected as much as possible in the model without slowing down the simulation too much. The same is also true when this behavioral model accompanies the block as one of the data needed for a Virtual Component in an IP catalog [2].

The key problem remains how to generate a good behavioral model. For this different approaches have been tried out. We can distinguish :

1) **black box approaches :** these consider the block as a black box and only look at the input-output behavior of the block. These models can be more generic, but it is difficult to guarantee their accuracy over a wide range.

   Analytic approach - One possibility here is to describe the model in analytic equations. For example, the dynamic behavior (settling time and glitch energy) of a current-steering DAC as shown in Fig. 12 can easily be described by superposition of an exponentially damped sine and a shifted hyperbolic tangent [14]:

$$i_{out} = A_{gl}\sin\left(\frac{2\pi}{t_{gl}}(t-t_0)\right)\exp\left(-sign(t-t_0)\frac{2\pi}{t_{gl}}(t-t_0)\right)+ \frac{level_{i+1}-level_i}{2}\tanh\left(\frac{2\pi}{t_{gl}}(t-t_0)\right)+\frac{level_{i+1}+level_i}{2} \quad (1)$$

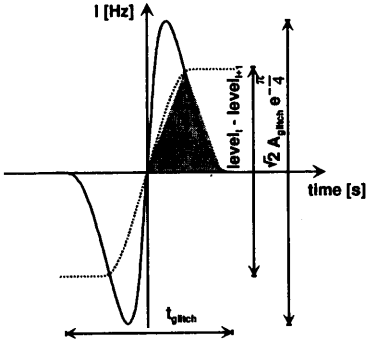


Fig. 12. Dynamic behavior of a current-steering digital-to-analog converter.

Fig. 13 compares the response of the behavioral model with parameter values extracted from SPICE simulations with SPICE simulation results of the original circuit. The speed-up in CPU time is a factor 874 (!!) while the error is below 1%.

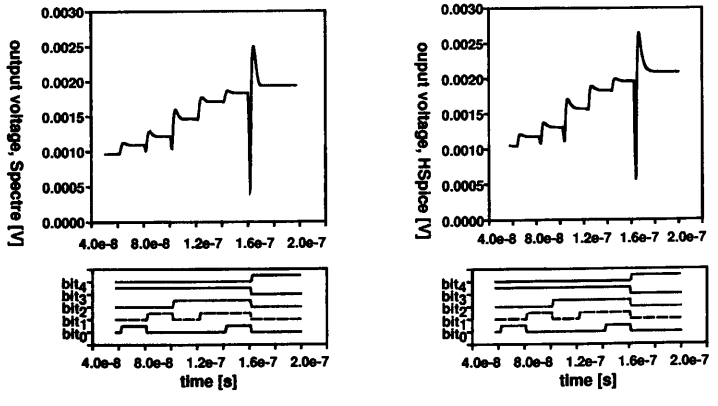


Fig. 13. Comparison between the device-level simulation results (on the right) and the response of the behavioral model (on the left).

Neural network approach - Another possible black box approach is the use of an artificial neural network that is being trained with SPICE simulation results of the real circuit until the response of the network matches closely enough the response of the real circuit [15]. At that moment the network has become an implicit model of the circuit. Difficulties with this approach however are that it is not trivial to decide on a good neural network structure for every circuit, that the training set must exercise all possible operating modes of the circuit and that the resulting model is specific for one particular implementation of the circuit.

2) **white box approaches :** in this case detailed information from the actual circuit implementation is used to build up the behavioral model. The resulting model is specific for the particular circuit but there is a higher guarantee that it tracks the real circuit behavior well in a wide range.

   A typical approach here uses symbolic analysis techniques to first generate the exact set of describing algebraic/differential equations of the circuit, which are then simplified within a given error bound of the exact response using both global and local simplifications [16]. The resulting simplified set of equations then constitutes the behavioral model of the circuit and tracks nicely the behavior of the circuit. The biggest drawback however is that the error estimation is difficult and for nonlinear circuits heavily depends on the targeted response.

3.3. Analog power estimation

The second crucial element at all higher levels in the top-down approach, as well for system-level exploration as for analog high-level synthesis, are accurate and efficient power and area estimators needed to assess the optimality of different alternatives. Such estimators are functions that predict the power or area that is going to be consumed by a circuit implementation of an analog block (e.g. an analog-to-digital converter) with given specification values (e.g. resolution and speed). Such functions can be obtained in two ways.

Functions or procedures - A first possibility is the derivation of analytic functions or procedures that return the power or area estimate. For example, a quite general model for the power consumption P_i of a

block i with specifications for the gain, bandwidth and dynamic range (A_i, BW_i, DR_i) is given by :

$$P_i = FOM_i \cdot kT \cdot DR_i \cdot A_i^2 \cdot BW_i \tag{2}$$

where FOM_i is the figure of merit of block i, which is a parameter used to calibrate the expression to the values of real implementations. Such functions can be derived for many different blocks. Their big advantage is that they normally track power and area well when traversing the design space. The drawback is that their derivation is not straightforward.

Extraction from samples - The second possibility is to extract the estimators from a whole sampling set of available or generated designs through interpolation or fitting. Either an explicit function is fitted to the sampling set, or an implicit model is built up, e.g. under the form of an artificial neural network. The latter approach was prototyped in a tool called EsteMate [17]. First the neural network is trained with a large set of designs that have for instance been generated with an analog synthesis tool. The trained network can then be called to predict the circuit's power consumption as a function of the specifications which are input to the network. The advantage of this approach is that it accurately models the complex interactions between the different specifications. The disadvantages are that it needs a large set of training samples and that the network is not guaranteed to track well outside the training space. Fig. 15 gives an example of the power estimator that is embedded in such a neural network for the class AB opamp of Fig. 14. A network with two hidden layers with 9 and 6 neurons respectively was used. The training took about 15 minutes of CPU time for both the power and the area estimator, and an evaluation of the trained network takes 0.3 msec of CPU time.

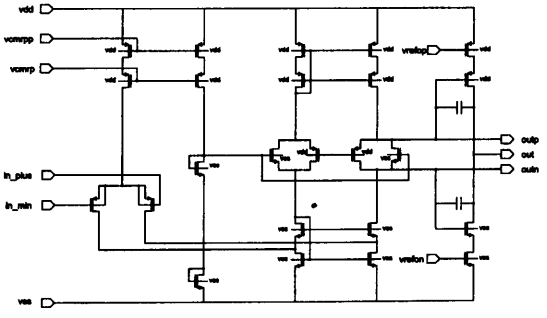


Fig. 14. Example class AB opamp.

Fig. 15 shows an explicit two-dimensional cross-section of the multi-dimensional power estimation function that is implicitly embedded in the neural network. The figure shows the relation of the power as a function of the gain-bandwidth and slew rate requirements, where all other specifications are fixed at a low value, except the offset which has to be smaller than 35 mV. As can be seen the power increases with the required speed (the gain-bandwidth *GBW*). A high slew rate (*SR*) also causes a high power consumption. At low speeds, the power is determined by the offset specification, which causes the leveling off of the power estimator function.

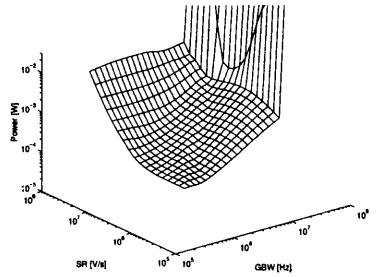


Fig. 15. View of the power estimator of the class AB opamp as embedded in the trained neural network.

Similar to power and area estimators also feasibility functions are needed that limit the high-level synthesis to realizable values of the building block specifications. These can be implemented under the form of functions (e.g. a trained neural network) that return whether a block is feasible or not, or as the geometrically calculated feasible performance space of a circuit (e.g. using polytopes [18] or using radial base functions [19]).

4. ANALOG CIRCUIT-LEVEL SYNTHESIS

Once the specifications for each of the lowest-level analog blocks have been determined, each individual block still has to be implemented as a full transistor circuit with the lowest possible power or area consumption. This can be done by means of an analog synthesis tool like the AMGIE system [7] that generates an optimized

custom design for each specific application. Starting from the given specifications, the system selects the most promising circuit topology out of the system's library, then determines the optimal device sizes as to meet the specifications while minimizing the circuit's power and/or area consumption, and finally automatically generates a full custom layout using the performance-driven analog layout tool LAYLA [20] that has been integrated in the system. After the sizing step and after the layout extraction step the circuit is also fully verified using automatic characterization and datasheet generation. The input to the AMGIE system are therefore the required performance specifications for the circuit to be synthesized, and the output are a fully sized circuit, its layout and its extracted datasheet. The control panel of the system which displays the design hierarchy and the design flow is shown in Fig. 16. User interaction is performed through this control panel. The system is embedded in a commercial EDA environment to guarantee compatibility with existing design frameworks, which are also used for the design of the digital parts.

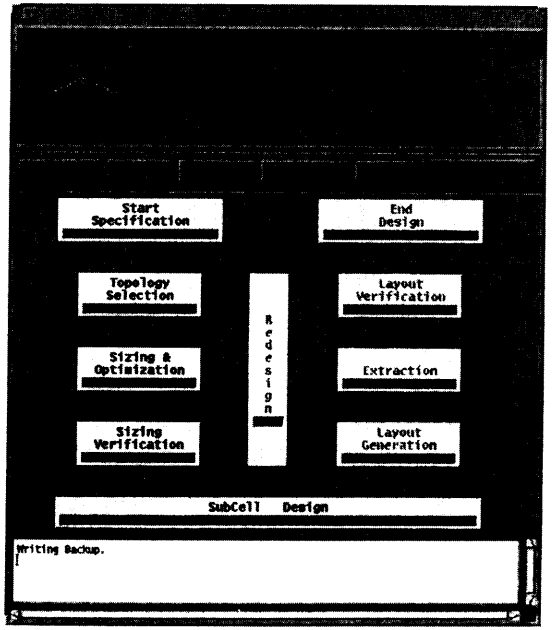


Fig. 16. Control panel of the AMGIE analog synthesis system.

Fig. 17 shows an example of a circuit that has been synthesized with the AMGIE system. It is a charge-sensitive amplifier (CSA) and 4th-order pulse-shaping amplifier (PSA) front-end for particle/radiation detector applications such as in scientific stellar satellites or nuclear physics experiments. The specifications for this CSA-PSA block for a particular scientific space mission were derived using high-level synthesis. The results of circuit-level synthesis are shown in Table 4. The specifications are contained in the second column. Compared to an earlier manual design for the same specifications (third column) the AMGIE system (fourth column) could reduce the power consumption with a factor of almost 6 (from 40 mW to 7 mW) with even a slightly smaller area! The final generated layout of this circuit is shown in Fig. 18. Besides a reduction of the power also the total design time was drastically reduced, hence showing that an analog synthesis tools increases both the optimality of analog designs and the productivity of the analog designers.

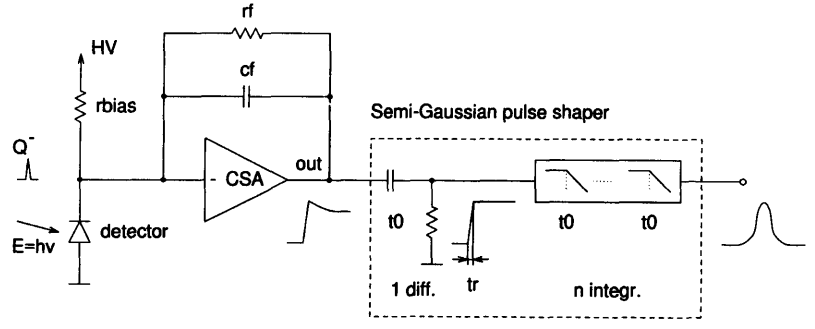


Fig. 17. CSA-PSA circuit.

CSA-PSA specification	spec	manual design	automated
peaking time	< 1.1 µs	1.1 µs	1.1 µs
counting rate	> 200 kHz	200 kHz	294 kHz
noise	< 1000 RMS e⁻	750 RMS e⁻	905 RMS e⁻
gain	20 mV/fC	20 mV/fC	21 mv/fC
output range	> -1..1 V	-1..1 V	-1.5..1.5 V
power	minimal	40 mW	7 mW
area	minimal	0.7 mm²	0.6 mm²

Table 4. Comparison between manual and synthesized results for the CSA-PSA circuit.

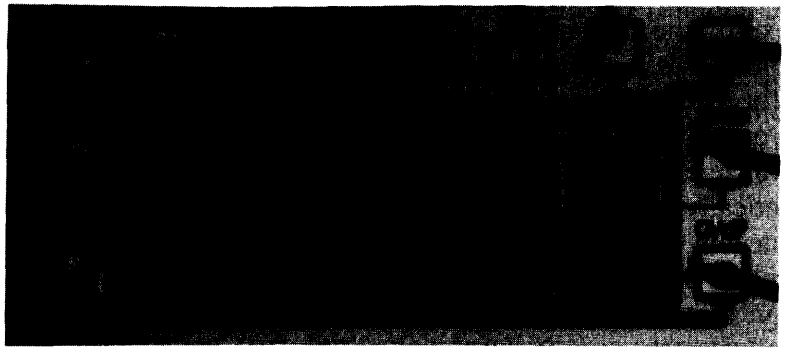


Fig. 18. Synthesized layout of the CSA-PSA circuit.

5. LAYOUT SYNTHESIS OF MIXED-SIGNAL SYSTEMS

The final phase in the design flow is the layout assembly. This comprises different steps: floorplanning, block layout synthesis, block placement, global and detailed routing (including the power grid) and final verification.

5.1. Floorplanning

The first step is floorplanning, which can be performed with estimated information at early stages in the design flow to provide early feedback for the further implementation of the blocks, and with full detail after completing the sizing of each of the blocks. An important aspect for floorplanning mixed-signal systems is that geometrical aspects like area and net length should not be the only criteria to be taken into account, but also electrical constraints like substrate noise or thermal interactions. Commercial tools do not provide such features, but such techniques have been explored at research level. The WRIGHT tool [21] for instance uses simulated annealing to determine the floorplan, and has a fast substrate noise coupling evaluator built in so that a simplified view of substrate noise influences the floorplan.

5.2. Circuit-level layout generation

Once the floorplan has been determined, the layouts of the different blocks can be generated accordingly. For the digital blocks,

commercial standard cell place & route tools can be used, which today fully exploit the many metal layers offered by modern deep submicron VLSI processes. For the analog cells the layout is mostly handcrafted, typically using parameterized procedural device generators to create the layouts of individual devices or of special device structures (like e.g. a differential pair). Nevertheless, academic tools exist that can automatically synthesize analog circuit-level layouts while guaranteeing performance [20,22,23]. The LAYLA tool for instance implements such a performance-driven device place and route layout strategy that minimizes the layout area while enforcing typical constraints such as symmetry and while keeping the performance degradation introduced by the layout parasitics within the margins allowed by the user [20]. This tool was used to generate the layouts of the cells in Fig. 18, and these layouts were then combined hierarchically according to the defined floorplan.

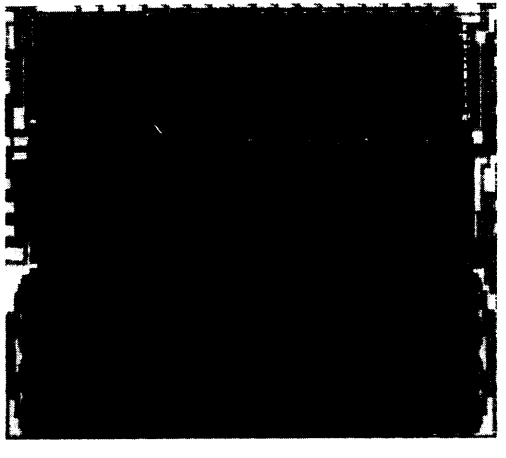


Fig. 19. Generated layout of a 14-bit digital-to-analog converter.

On the other hand, for circuits with a more regular structure, other layout techniques are needed. This is true for the generation of ROMs or RAMs, but also for array-type analog circuits like current-steering digital-to-analog converters, folding/interpolating analog-to-digital converters, etc. The MONDRIAAN tool [24] was developed for this purpose and it translates global layout specifications into a detailed placement and interconnections of all basic cells in the array. The tool

was used to generate the layouts of the switch array (middle) and the current cell array (bottom) of the 14-bit current-steering DAC of Fig. 19. The block at the top is the decoder which was synthesized from VDHL code with a logic synthesis tool and the layout of which was generated with a commercial digital standard cell place & route tool.

5.3. Block place and route

After generating the layout of the individual blocks, the next step in the system design flow is block place and route. For this commercial tools exist but they still don't include all the constraints typically needed in mixed-signal designs, such as the handling of arbitrarily shaped blocks and complex symmetries, as well as the avoidance of signal interactions (crosstalk) and noise couplings. The academic WREN tool [25] comprises both a mixed-signal global router and channel router. The tool uses the notion of SNR-style (signal-to-noise ratio) constraints for incompatible signals, and strives to comply with designer-specified noise rejection limits on critical signals.

Critical in mixed-signal system layout is also the power grid design. In the mixed-signal case not only connectivity, ohmic drops and electromigration effects have to be considered, but also noise constraints (including transient effects like current spikes) and arbitrary (non-tree) grid topologies. The RAIL tool [26] addresses such concerns by casting mixed-signal power grid synthesis as a routing problem that uses fast AWE-based linear system evaluation to electrically model the entire power grid, package and substrate during layout while trying to satisfy dc, ac and transient performance constraints.

5.4. Verification

The final step is the detailed verification of the entire system layout. For mixed-signal systems this is today still a very big problem, both at the layout level and at the electrical level. At the layout level, DRC, ERC and LVS can easily be done for the different blocks. DRC and ERC can also easily be done for the entire chip, but due to the different tools typically used for analog and digital blocks LVS of complete analog-digital systems is not at all trivial.

The situation is even worse at the electrical level. After extraction of

the parasitics from the layout the performance of the individual blocks can be verified using detailed simulations, but due to the complexity no complete device-level simulation of the entire system is feasible. Therefore, the performance of individual blocks has to be abstracted into behavioral models, which are then used to simulate and verify the system performance. There does not (yet?) exist, however, a systematic methodology to extract analog behavioral models, and if these models do not capture the important critical effects (e.g. transients) the system simulations might not detect certain functionality hazards. The system verification today is therefore merely a check of correct connectivity than a true proof of functionality and performance, being even far from any formal verification proof.

Another difficult problem in mixed-signal designs is the verification of all unwanted signal interactions at the system level that can cause parametric malfunctioning of the chip. These interactions can come from capacitive or (at higher frequencies) inductive crosstalk, from supply line or substrate interactions, from thermal interactions, from interactions through the package, etc. Especially the analysis of substrate noise has received much attention in recent years. Efficient techniques have been developed to simulate the impact of digital switching noise injected in the substrate on sensitive analog nodes elsewhere in the same substrate [27,28]. This however has to be combined with an analysis of the digital switching activity and of the impact on the analog circuit performance in order to cover the complete problem. Also the influence of the package is receiving more and more attention in recent years, especially for high-frequency applications such as encountered in RF.

6. CONCLUSIONS

An overview has been given of the challenges and available solutions in applying top-down design to mixed analog-digital systems. In order to establish the design productivity and design optimality needed to design future systems on a chip within the tightening time to market constraints, a systematic top-down design approach has to be followed with sufficient time and attention paid to system-level architectural design before proceeding to the detailed block design. Also the reuse

of available macrocells will be unavoidable. A top-down approach for mixed-signal systems has then been presented and the different advantages and limitations have been discussed. Examples of system-level architectural exploration, including analog-digital partitioning, have been given. Techniques to derive the crucial elements for analog high-level design, i.e. analog behavioral models and power estimators, have then been described in detail. Finally, the progress in techniques and tools for analog circuit and layout synthesis as well as for mixed-signal layout assembly has been presented. Especially the impact of parasitic interactions such as crosstalk and substrate noise couplings are critical limitations to the performance of mixed-signal systems.

It can be concluded that a systematic top-down approach will be unavoidable for designing future complex systems, but that major challenges remain in the accuracy of analog behavioral modeling used for high-level explorations and top-down design refinement and in the detailed system verification (including all parasitic interactions). Also the mapping from system-level specifications (an "algorithm") into an appropriate architecture remains an open problem.

ACKNOWLEDGEMENTS

The author acknowledges Prof. Willy Sansen and all Ph.D. researchers who have contributed to the reported results, including Walter Daems, Geert Debyser, Bart De Smedt, Stéphane Donnay, Koen Lampaert, Erik Lauwers, Francky Leyn, Piet Vanassche, Jan Vandenbussche, Geert Van der Plas, Wim Verhaegen. The author also acknowledges the support of the many funding companies and organisations (ESPRIT, IWT, MEDEA).

REFERENCES

[1] Semiconductor Industry Association (SIA), "The national technology roadmap for semiconductors," 1994.

[2] "Virtual Socket Interface Architecture Document," version 1.0, VSI Alliance, 1997.

[3] L.R. Carley, G. Gielen, R. Rutenbar, W. Sansen, "Synthesis tools for mixed-signal ICs: progress on frontend and backend

strategies," proceedings ACM/IEEE Design Automation Conference (DAC), pp. 298-303, 1996.

[4] S. Donnay et al., "Using top-down CAD tools for mixed analog/digital ASICs: a practical design case," Kluwer International Journal on Analog Integrated Circuits and Signal Processing, Special Issue on Modeling and Simulation of Mixed Analog-Digital Systems, Vol. 10, pp. 101-117, June-July 1996.

[5] R. Harjani, R.Rutenbar, L.R. Carley, "OASYS: a framework for analog circuit synthesis," IEEE Transactions on Computer-Aided Design, Vol. 8, No. 12, pp. 1247-1265, December 1989.

[6] H. Chang et. al., "A top-down, constraint-driven design methodology for analog integrated circuits," Kluwer Academic Publishers, 1997.

[7] G. Gielen et al., "An analog module generator for mixed analog/digital ASIC design," John Wiley International Journal of Circuit Theory and Applications, Vol. 23, pp. 269-283, July-August 1995.

[8] M. Degrauwe et al., "IDAC: an interactive design tool for analog CMOS circuits," IEEE Journal of Solid-State Circuits, Vol. 22, No. 6, pp. 1106-1115, December 1987.

[9] S. Donnay, G. Gielen, W. Sansen, "High-level power minimization of analog sensor interface architectures," Integrated Computer-Aided Engineering, 1998.

[10] E. Ochotta, R. Rutenbar, L.R. Carley, "Synthesis of high-performance analog circuits in ASTRX/OBLX," IEEE Transactions on Computer-Aided Design, Vol. 15, No. 3, pp. 273-294, March 1996.

[11] F. Medeiro, B. Pérez-Verdú, A. Rodríguez-Vázquez, J. Huertas, "A vertically-integrated tool for automated design of ΣΔ modulators," IEEE Journal of Solid-State Circuits, Vol. 30, No. 7, pp. 762-772, July 1995.

[12] J. Crols, S. Donnay, M. Steyaert, G. Gielen, "A high-level design and optimization tool for analog RF receiver front-ends," proceedings International Conference on Computer-Aided Design (ICCAD), pp. 550-553, 1995.

[13] S. Donnay, G. Gielen, W. Sansen, "High-level analog/digital partitioning in low-power signal processing applications,"

proceedings 7th international workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS), pp. 47-56, 1997.

[14] J. Vandenbussche, G. Van der Plas, G. Gielen, M. Steyaert, W. Sansen, "Behavioral model for D/A converters as VSI Virtual Components," proceedings Custom Integrated Circuits Conference, May 1998.

[15] J. Vandenbussche, P. Lemmerling, J. Suykens, G. Gielen, "Behavioral modeling of analog circuits using neural networks," proceedings international workshop on Behavioral Modeling And Simulation (BMAS), pp. 1-8, October 1997.

[16] C. Borchers, L. Hedrich, E. Barke, "Equation-based behavioral model generation for nonlinear analog circuits," proceedings Design Automation Conference (DAC), pp. 236-239, 1996.

[17] G. Van der Plas, J. Vandenbussche, G. Gielen, W. Sansen, "EsteMate: a tool for automated power and area estimation in analog top-down design and synthesis," proceedings Custom Integrated Circuits Conference (CICC), pp. 139-142, 1997.

[18] P. Veselinovic et al., "A flexible topology selection program as part of an analog synthesis system," proceedings European Design & Test Conference (ED&TC), pp. 119-123, 1995.

[19] R. Harjani, J. Shao, "Feasibility and performance region modeling of analog and digital circuits," Kluwer International Journal on Analog Integrated Circuits and Signal Processing, Vol. 10, No. 1, pp. 23-43, January 1996.

[20] Koen Lampaert, Georges Gielen, Willy Sansen, "Analog layout generation for performance and manufacturability," Kluwer Academic Publishers, 1998.

[21] S. Mitra, R. Rutenbar, L. Carley, D. Allstot, "Substrate-aware mixed-signal macrocell placement in WRIGHT," IEEE Journal of Solid-State Circuits, Vol. 30, No. 3, pp. 269-278, March 1995.

[22] E. Malavasi, E. Felt, E. Charbon, A. Sangiovanni-Vincentelli, "Automation of IC layout with analog constraints," IEEE Transactions on Computer-Aided Design, Vol. 15, No. 8, pp. 923-942, August 1996.

[23] J. Cohn, D. Garrod, R. Rutenbar, L. Carley, "Analog device-level layout automation," Kluwer Academic Publishers, 1994.

[24] G. Van der Plas, J. Vandenbussche, G. Gielen, W. Sansen, "Mondriaan: a tool for automated layout synthesis of array-type analog blocks," proceedings Custom Integrated Circuits Conference (CICC), May 1998.

[25] S. Mitra, S. Nag, R. Rutenbar, L. Carley, "System-level routing of mixed-signal ASICs in WREN," International Conference on Computer-Aided Design (ICCAD), November 1992.

[26] B. Stanisic, N. Verghese, R. Rutenbar, L. Carley, D. Allstot, "Addressing substrate coupling in mixed-mode ICs: simulation and power distribution synthesis," IEEE Journal of Solid-State Circuits, Vol. 29, No. 3, March 1994.

[27] N. Verghese, D. Allstot, T. Schmerbeck, "Simulation techniques and solutions for mixed-signal integrated circuits," Kluwer Academic Publishers, 1995.

[28] R.Gharpurey, R. Meyer, "Modeling and analysis of substrate coupling in integrated circuits," IEEE Journal of Solid-State Circuits, Vol. 31, No.3, pp 344-353, March 1996.

LOW-NOISE AND RF POWER AMPLIFIERS FOR TELECOMMUNICATION

Rudy van de Plassche

In telecommunication systems low-noise amplifiers are front-ends of receiver designs. At the transmitter part a high-performance high-efficiency power amplifier is a critical design. Examples of both system parts are described in this section.

In the first paper by Thomas Lee design issues for low-noise CMOS amplifiers are reviewed. Technology scaling and power considerations result in possible noise figures of 1.5 dB at 2 GHz using 0.5-micron CMOS technology.

In the second paper by Janssens et al the design of low-power wide-band CMOS amplifiers in submicron technologies is analyzed. Using a special topology a low-power design is demonstrated in 0.5 micron. Analyzing the input impedance of the circuit Non-Quasi Static RF CMOS modeling is needed to accurately predict the input impedance of the system.

The third paper by Baltus introduces a special bipolar technology with removed substrate to obtain ultra low-power building blocks for telecommunication applications. The removal of the substrate and replacement of this substrate by a glass carrier results in very low parasitic capacitances. As a result system performance at an order of magnitude lower power dissipation is obtained. Cost of the substrate reduction operation is small compared to the total process cost.

The fourth paper by Sevenhans et al introduces silicon germanium technology for high frequency applications. Different examples of circuits like quadrature mixers, amplifiers, VCO's and pre-power amplifiers are described.

The fifth paper by Weber describes the modeling of silicon bipolar power amplifiers for telecom applications. Models for on-chip inductors, interconnect, substrate coupling and packaging are introduced.

The last paper by Taylor describes design considerations for power amplifiers using GaAs MESFET technology. Device, package and circuit approaches are reviewed for 1-3 watt power amplifiers in the frequency range up to 1.9 GHz. A comparison with other technologies is included.

THE DESIGN OF NARROWBAND CMOS RF LOW-NOISE AMPLIFIERS

Thomas H. Lee
Stanford University Center for Integrated Systems
Stanford, California, USA

ABSTRACT

General conditions for minimizing the noise figure of any linear two-port are reviewed before considering the specific case of a MOSFET low-noise amplifier (LNA). It is shown that the minimum noise figure cannot be obtained over an arbitrarily large bandwidth with networks of low order. For narrowband operation, however, one may construct simple amplifiers whose noise figure and power gain are close to the theoretical optima allowed within an explicit power constraint, and which simultaneously present a specified impedance to the driving source. The effects of overlap (drain-gate) capacitance, short-channel carrier heating, substrate resistance ("epi noise"), and gate interconnect resistance are also considered. Amplifier noise figures of 1.5dB or better at 10mW are achievable in the 1-2GHz range with 0.5μm technology, and improve with scaling.

1. INTRODUCTION

That the signals delivered by the antenna in modern wireless systems can be in the submicrovolt range underscores the acute need for low-noise amplification. Furthermore, the transfer characteristics of filters interposed between the antenna and LNA are frequently quite sensitive to the quality of the terminations. Thus, the design problem is often compounded by the additional requirement that the amplifier exhibit a specified (and usually real) input impedance. It should seem intuitively reasonable that it is difficult to satisfy this last condition with an inherently capacitive device such as a MOSFET, if broadband low-noise oper-

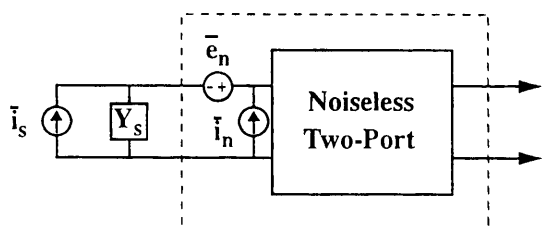
ation is required. This paper will first establish a theoretical foundation for that intuition, then explore a narrowband topology that provides near-optimum gain and noise figure while providing a real input resistance. Addition of an explicit power consumption design constraint to "classical" noise optimization will be seen to lead to a definite optimum device width. The effect of drain-gate overlap capacitance will be considered, as will the degradation in noise figure arising from short-channel effects and the thermal noise of both the substrate and gate interconnect material.

2. CLASSICAL NOISE OPTIMIZATION

What we term "classical" noise optimization begins with the assumption that one is given a linear two-port whose characteristics are fixed. The classical approach then yields the optimum source impedance (from a noise performance viewpoint).

A noisy two-port may be modeled as a noiseless two-port to which a noise voltage and noise current are connected:

FIGURE 1. Noisy two-port (source model shown outside dashed boundary)



The noise factor for this model is then:

$$F = \frac{\overline{i_s^2} + \overline{\left|i_n + Y_s e_n\right|^2}}{\overline{i_s^2}} \tag{1}$$

Because e_n and i_n model the noise due to all of the sources within the original two-port, they may be correlated. Thus, express i_n as the sum of two currents: i_c, which is correlated with e_n, and i_u, which is not:

$$i_n = i_c + i_u \tag{2}$$

Because i_c is fully correlated with e_n, they are proportional:

$$i_c = Y_c e_n \tag{3}$$

The *correlation admittance* Y_c is not necessarily related to a measurable input admittance.

Using the foregoing definitions, the noise factor may be expressed as follows:

$$F = 1 + \frac{G_u + \left|Y_c + Y_s\right|^2 R_n}{G_s} = 1 + \frac{G_u + \left[\left(G_c + G_s\right)^2 + \left(B_c + B_s\right)^2\right] R_n}{G_s} \tag{4}$$

where we have explicitly decomposed each admittance into a sum of a conductance G and a susceptance B, and where the following traditional substitutions have been made:

$$R_n \equiv \frac{\overline{e_n^2}}{4kT\Delta f}, G_u \equiv \frac{\overline{i_u^2}}{4kT\Delta f}, G_s \equiv \frac{\overline{i_s^2}}{4kT\Delta f} \tag{5}$$

The noise factor is a minimum when the following conditions hold:

$$B_s = -B_c = B_{opt} \tag{6}$$

$$G_s = \sqrt{\frac{G_u}{R_n} + G_c^2} = G_{opt} \tag{7}$$

Hence, when the source and correlation susceptances are algebraic inverses, and the source conductance is equal to the value in Eqn. 7, the following minimum noise factor is achieved:

$$F_{min} = 1 + 2R_n\left[G_{opt} + G_c\right] = 1 + 2R_n\left[\sqrt{\frac{G_u}{R_n} + G_c^2} + G_c\right] \tag{8}$$

In general, the noise factor may be expressed as:

$$F = F_{min} + \frac{R_n}{G_s}\left[\left(G_s - G_{opt}\right)^2 + \left(B_s - B_{opt}\right)^2\right] \tag{9}$$

Contours of constant noise factor are circles centered about (G_{opt}, B_{opt}) in the admittance plane (and also on a Smith chart, because the bilinear transformation that maps the two preserves circles).

Although minimizing the noise factor is qualitatively similar to maximizing power transfer, the source admittances leading to these two conditions are almost never the same, as the correlation and input admittances, for example, are rarely equal (except by coincidence). Therefore, one generally cannot enjoy maximum power gain and minimum noise figure simultaneously [1].

3. MOSFET TWO-PORT NOISE PARAMETERS

The MOSFET noise model has two sources arising from thermal fluctuations of channel charge. First, there is a drain noise current source:

$$\overline{i_{nd}^2} = 4kT\gamma g_{d0}\Delta f \tag{10}$$

where g_{d0} is g_{ds} evaluated at zero V_{DS}, and γ is unity in triode and 2/3 in saturation, at least in the long-channel limit. Some older references use g_m, but g_{d0} is better related to channel charge in short-channel devices.

A gate noise current (unmodeled by SPICE) also flows, from channel potential fluctuations coupling capacitively into the gate terminal:

$$\overline{i_{ng}^2} = 4kT\delta g_g \Delta f \tag{11}$$

where δ is twice γ in long devices, and

$$g_g = \frac{\omega^2 C_{gs}^2}{5g_{d0}} \tag{12}$$

The gate and drain noise currents have a correlation coefficient c of $j0.395$ for long-channel devices. The value of c in the short-channel regime is presently unknown.

In what follows, we neglect C_{gd} to simplify the derivation. The primary effect of C_{gd} (in the cascoded designs we will consider) is on the input impedance. To derive the four noise parameters, first short-circuit the input port. Reflect the drain current noise back to the input as a noise voltage and recognize that the ratio of these quantities is simply g_m. Thus,

$$\overline{e_n^2} = \frac{\overline{i_{nd}^2}}{g_m^2} = \frac{4kT\gamma g_{d0}\Delta f}{g_m^2} \tag{13}$$

from which it is apparent that the equivalent input noise voltage is completely correlated, and in phase, with the drain current noise. Thus, we can immediately determine that

$$R_n \equiv \frac{\overline{e_n^2}}{4kT\Delta f} = \frac{\gamma g_{d0}}{g_m^2} \tag{14}$$

The equivalent input noise voltage by itself does not fully account for the drain noise, however, because a noisy drain current also flows even when the input is *open*-circuited (and even if we additionally ignore induced gate current). Under this open-circuit condition, dividing the drain current noise by the transconductance yields an equivalent input voltage which, when multiplied in turn by the input admittance, gives us the value of an equivalent input current noise that completes the modeling of i_{nd}:

$$\overline{i_{n1}^2} = \frac{\overline{i_{nd}^2}\left(j\omega C_{gs}\right)^2}{g_m^2} = \frac{4kT\gamma g_{d0}\Delta f\left(j\omega C_{gs}\right)^2}{g_m^2} = \overline{e_n^2}\left(j\omega C_{gs}\right)^2 \tag{15}$$

where we have assumed that the input admittance of a MOSFET is purely capacitive. This assumption is satisfied well below ω_T and if appropriate high-frequency layout practice is observed to minimize gate resistance. Given this assumption, Eqn. 15 shows that the input noise current i_{n1} is

in quadrature, and therefore completely correlated, with the equivalent input noise voltage e_n.

The total equivalent input current noise is the sum of the reflected drain noise contribution of Eqn. 15 and the induced gate current noise. The induced gate noise current itself consists of two terms, i_{ngc}, and i_{ngu}, which are fully correlated and uncorrelated, respectively, with the drain current noise. Hence, the correlation admittance is:

$$Y_c = j\omega C_{gs} + \frac{i_{ngc}}{e_n} = j\omega C_{gs} + \frac{g_m}{i_{nd}} \cdot i_{ngc} = j\omega C_{gs} + g_m \cdot \frac{i_{ngc}}{i_{nd}} \tag{16}$$

which may ultimately be expressed as:

$$Y_c = j\omega C_{gs} + g_m \cdot c\sqrt{\frac{\delta\omega^2 C_{gs}^2}{5\gamma g_{d0}^2}} = j\omega C_{gs} + \frac{g_m}{g_{d0}} \cdot c\sqrt{\frac{\delta}{5\gamma}} \cdot \omega C_{gs} \tag{17}$$

If c continues to be purely imaginary, even in the short-channel regime, we finally have:

$$Y_c = j\omega C_{gs} + j\omega C_{gs} \frac{g_m}{g_{d0}} \cdot |c|\sqrt{\frac{\delta}{5\gamma}} = j\omega C_{gs}\left(1 + \alpha|c|\sqrt{\frac{\delta}{5\gamma}}\right) \tag{18}$$

where

$$\alpha = \frac{g_m}{g_{d0}} \tag{19}$$

Since α is unity for long-channel devices, and progressively decreases as channel lengths shrink, it is one measure of the departure from the long-channel regime.

Eqn. 18 shows that the correlation admittance is purely imaginary, so that $G_c = 0$. More significant, however, is the fact that Y_c does not equal the admittance of C_{gs}, although it is proportional it. Hence, the MOSFET is typical in that one cannot optimize power transfer and noise figure simultaneously. To explore further the important implications of this observation, we derive the last noise parameter, G_u.

Using the definition of the correlation coefficient, we may express the induced gate noise as follows:

$$\overline{i_{ng}^2} = \overline{\left(i_{ngc} + i_{ngu}\right)^2} = 4kT\Delta f\delta g_g |c|^2 + 4kT\Delta f\delta g_g \left(1 - |c|^2\right) \tag{20}$$

The very last term in Eqn. 20 is the uncorrelated portion of the gate noise current, so that, finally:

$$G_u \equiv \frac{\overline{i_u^2}}{4kT\Delta f} = \frac{4kT\Delta f\delta g_g\left(1 - |c|^2\right)}{4kT\Delta f} = \frac{\delta\omega^2 C_{gs}^2\left(1 - |c|^2\right)}{5g_{d0}} \tag{21}$$

The four noise parameters are summarized in the following table:

TABLE 1. Equivalent two-port noise parameters for MOSFET

Parameter	Expression
G_c	≈ 0
B_c	$\omega C_{gs}\left(1 + \alpha|c|\sqrt{\frac{\delta}{5\gamma}}\right)$
R_n	$\frac{\gamma g_{d0}}{g_m^2} = \frac{\gamma}{\alpha} \cdot \frac{1}{g_m}$
G_u	$\frac{\delta\omega^2 C_{gs}^2\left(1 - |c|^2\right)}{5g_{d0}}$

With these parameters, we can determine the source impedance that minimizes the noise figure:

$$B_{opt} = -B_c = -\omega C_{gs}\left(1 + \alpha|c|\sqrt{\frac{\delta}{5\gamma}}\right) \tag{22}$$

$$G_{opt} = \sqrt{\frac{G_u}{R_n} + G_c^2} = \alpha\omega C_{gs}\sqrt{\frac{\delta}{5\gamma}\left(1-|c|^2\right)} \tag{23}$$

The optimum source susceptance is thus inductive in character, but has a capacitive frequency variation. Furthermore, the optimum source conductance varies linearly with frequency. Synthesizing a network to provide these characteristics over a large frequency range is challenging to say the least, so that achieving a broadband noise match to a MOSFET is fundamentally difficult.

Whenever a noise match is achieved, the corresponding minimum noise factor is:

$$F_{min} = 1 + 2R_n\left[G_{opt} + G_c\right] = 1 + \frac{2}{\sqrt{5}}\frac{\omega}{\omega_T}\sqrt{\gamma\delta\left(1-|c|^2\right)} \tag{24}$$

Note that if there were no gate current noise, or *if the gate and drain current noise sources were fully correlated*, the minimum noise figure would be 0dB. Contrary to widely held beliefs, the mere presence of gate or drain noise by itself does not necessarily impose a fundamental noise figure penalty. MOS amplifiers exhibit nonzero minimum noise figures because of the existence of gate and drain noise currents that are uncorrelated with each other.

Improvements in ω_T that accompany scaling improve the noise figure at any given frequency. To illustrate this point, let us assign numerical values to the parameters in Eqn. 24. Because the behavior of some of these parameters in the short channel regime is unknown, we will make some pessimistic guesses to arrive at conservative estimates of F_{min}. Measurements of γ reveal that it can be 2-3 times larger in short devices than predicted by long-channel theory. Values for δ in the short channel regime have never been reported, unfortunately, so we will assume that it is also augmented by a factor of two to three. Because one mechanism, thermally-driven channel charge fluctuations, gives rise to both gate and drain noise, this last assumption appears reasonable. Finally, assume that $|c|$ remains equal to 0.395. The following table shows F_{min} as a function of frequency (normalized to g_m/C_{gs}) if short channel effects cause a pessimistic tripling of γ and δ:

TABLE 2. Estimated F_{min} ($\gamma = 2$, $\delta = 4$)

$g_m/\omega C_{gs}$	F_{min} (dB)
20	0.5
15	0.6
10	0.9
5	1.6

Clearly, excellent noise figures are possible for MOSFETs, even with increased γ and δ.

4. SECOND-ORDER EFFECTS

There are two additional noise sources that ought to be considered in any detailed LNA design. These are the thermal noise of the substrate (epi noise) and of the gate interconnect. Epi noise can be treated as primarily equivalent to increasing the effective value of γ [2]:

$$\frac{\overline{i_{nd}^2}}{\Delta f} = 4kT\left(\gamma g_{d0} + g_{mb}^2 R_{epi}\right) = 4kTg_{d0}\left(\gamma + \frac{g_{mb}^2 R_{epi}}{g_{d0}}\right) \tag{25}$$

The quantity in the last set of parentheses is the effective value of γ. Using typical values, one finds that the increase in γ is usually below the range of 0.1 to 0.2 (i.e., about 10%), and is thus generally negligible by itself. Furthermore, although epi noise also gives rise to a noisy gate current (whose magnitude can sometimes exceed that of the fundamental induced gate noise), this term is *fully correlated* with the epi-induced drain noise, so substrate resistance has only a minor effect on the fundamental limits of noise performance. The liberal use of nearby substrate taps is always helpful in any case.

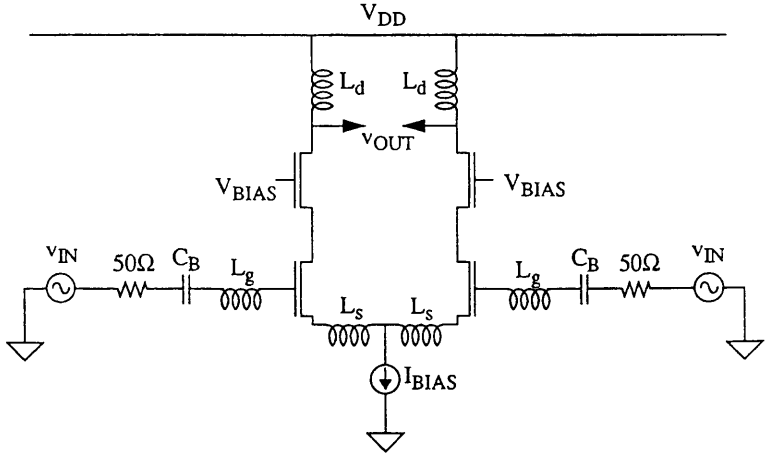
The thermal noise of the resistive gate electrode material is easily mitigated with careful layout that is already consistent with good high-fre-

quency practice. If a gate finger is contacted at only one end, it may be shown that this distributed *RC* structure has an equivalent noise resistance of $R_{SH}W/3L$, where R_{SH} is the sheet resistivity of the gate material. Thus, one must subdivide a device into a sufficient number of segments to make the total noise resistance as small as desired. Excessively short fingers are to be avoided because of the added parasitic capacitance that accompanies the wiring for each finger. Typical finger widths tend to be in the range of 5-20μm, although this is hardly a fixed rule.

5. CIRCUIT IMPLICATIONS AND IMPLEMENTATIONS

Having accepted the difficulty of achieving broadband low-noise operation, we now consider a topology designed specifically for narrowband applications:

FIGURE 2. Differential narrowband LNA (simplified)



In this differential cascoded structure, the source-degenerating inductance L_S interacts with the gate-source capacitance and g_m to produce a real term in the input impedance without paying the noise penalty of an ordinary resistance. Inductance L_G provides the additional degree of freedom required in general to allow for a desired resonant frequency of the input loop.

s discussed previously, the optimum source susceptance has an induc- ve character; this circuit provides that inductance at one frequency. We ave also noted that the correlation susceptance of a MOSFET differs mewhat from that of the gate-source capacitance. However, using typi- l parameter values, the difference is seen to be reasonably small (under %). Hence, the source susceptance that yields optimum noise perfor- nce differs little from the value that produces a resonance at the oper- ng frequency. Thus, this topology permits the near-simultaneous ievement of optimum noise and maximum gain. At the same time, it sents a specified real impedance to the input source.

nay be shown that the input resistance is modified downward by the in-gate capacitance in this cascoded circuit. Although an exact analy- s difficult, a useful approximation is given by:

$$Re\left[Z_{in}\right] \approx \frac{\omega_T L_S}{1+2C_{gd}/C_{gs}} \tag{26}$$

foregoing assumes equal widths for the main and cascoding devices.

hough the cascoding device does contribute noise of its own, the iso- n it provides between source and load is highly desirable. The noise re penalty is generally 0.5dB or less. In situations where that degrada- is unacceptable and the coupling between input and output circuits be tolerated, the cascoding device may be eliminated to yield an roved noise figure.

differential connection consumes twice the power for a given noise re than its single-ended counterpart, but has the valuable attribute of nsitivity to common-mode parasitic reactance in series with the bias- current source. Hence, one need not control or model this reactance. gle-ended designs require infinitely more attention in order to func- as expected.

at is missing from the foregoing is what is missing from classical ise optimization: Guidance on how to choose the width of the transis- s. In the integrated circuit design problem, the device is not fixed, but ssical optimization ignores this valuable degree of freedom. Addition- ly, power consumption is not considered at all in the classical approach.

The necessary modification of the design procedure has been worked out in [3], and it leads to the following approximate expression for the optimum device width:

$$W_{optP} = \frac{3}{2}\frac{1}{\omega L C_{ox} R_s Q_{sP}} \approx \frac{1}{3\omega L C_{ox} R_s} \quad (27)$$

where Q_{sP} is a parameter related to the ideal Q of the input network and has a value in the range of 3-5. For typical process parameters of today, the product of width and operating frequency work out to very roughly 500μm-GHz for a 50Ω system.

It is shown in [3] that deviations from the optimum width do not cause dramatic degradation of noise figure except at low bias currents, where the absolute noise figure is likely to be high anyway. At higher power levels, the optimum conditions are relatively flat, so uncertainties in device models may be comfortably accommodated.

The existence of an optimum width may be understood qualitatively as follows. For a fixed power budget (actually, bias current), a very narrow device has high ω_T, which tends to improve noise figure. However, a narrow device also requires a high-impedance input matching network, and this increases the prominence of gate noise, which tends to degrade noise figure. A very wide device, on the other hand, has fewer problems with gate noise, but the low current density degrades ω_T, increasing the prominence of drain noise. The optimum width balances these two effects to yield the minimum noise figure for a given power budget.

6. EXPERIMENTAL RESULTS

Measurements of noise figure on single-ended, single-device LNAs are in excellent accord with the foregoing expectations. At 1GHz, minimum noise figures slightly under 1dB have been achieved in a 0.5μm technology on bias currents of 2-5mA, with power gains ranging from 13.5dB to 18.5dB. The noise figures remain below 2dB at 500μA, while the gain drops to 8dB [4]. The performance of these LNAs is not necessarily representative of what may be achieved in practice because these LNAs were not designed to provide a specified input resistance. Rather, low-

loss tuners were adjusted to yield these minima. Nevertheless, they provide a valuable check on the entries of Table 2.

As an example of a somewhat more practical design, a differential LNA for use in an integrated GPS receiver has achieved a 2.4dB noise figure at 1.6GHz with a total bias current of under 5mA (2.5mA per side) [5]. To keep noise figure small, on-chip spiral inductors are not used in the gate circuit because their lossiness is too significant. Perhaps more important than the noise figure achieved (only about 1-1.5dB above F_{min}), however, is that the measured performance is close to the predicted noise figure of 2.5dB (assuming a doubling of γ over long-channel values) obtained with a variant of SPICE modified to accommodate gate noise. Since measurements at relatively low noise figure values are much more sensitive to errors than those made at high noise figures, this level of tracking between theory and measurement is reassuring.

7. CONCLUSION

It has been shown that the portion of gate noise current that is uncorrelated with the drain noise current is of fundamental importance in setting a lower bound on the noise figure. Thermal noise from the substrate and gate interconnect material degrades noise performance, but these effects are either of second order, or may be made so through proper layout.

Acknowledgment of the correct noise model, coupled with a modification of classical noise optimization to incorporate an explicit power constraint, leads to a narrowband LNA architecture which simultaneously achieves near optimum noise and power gain while providing a specific input resistance. The optimum performance is achieved when the product of device width and operating frequency is approximately 500μm-GHz in a 50Ω system. Excellent agreement between theoretical predictions and experiment is observed.

8. ACKNOWLEDGMENTS

The author acknowledges with great pleasure having learned much during the almost-completed thesis work of Mr. Derek Shaeffer in the area of LNAs and related topics, and also thanks Mr. Gabriel Brenna for characterizing the noise of a great number of CMOS LNAs as part of his

Diploma Thesis work, carried out at Stanford and Advanced Micro Devices, with the assistance of Dr. Natalino Camilleri.

9. REFERENCES

[1] T. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Chapter 10, Cambridge University Press, 1998.

[2] R. Jindal, "Distributed Substrate Resistance Noise in Fine-Line NMOS Field-Effect Transistors," *IEEE Trans. on Electron Devices*, vol. ED-32, no. 11, Nov. 1985, pp. 2450-2453.

[3] D. Shaeffer and T. Lee, "A 1.5V, 1.5GHz CMOS Low Noise Amplifier," *IEEE J. Solid-State Circuits*, vol. 32, no. 5, May, 1997, pp. 745-759.

[4] G. Brenna, "LNA Research," Diploma Thesis, École Polytechnique Fédérale de Lausanne, Feb. 1998.

[5] D. Shaeffer et al., "A 115mW CMOS GPS Receiver," *ISSCC Digest of Technical Papers*, Feb. 1998.

Design of Broadband Low-Noise Amplifiers in Deep-Submicron CMOS technologies

Johan Janssens and Michiel Steyaert
KULeuven, ESAT-MICAS, Heverlee, Belgium

Abstract

In general it is believed that the implementation of low-noise RF amplifiers in CMOS at a power dissipation competitive with bipolar technologies requires the use of narrow-band techniques. In this paper the design of power-efficient broadband low-noise amplifiers is explored, i.e., without using accurately tuned LC-tanks nor exploiting the 'overdrive' capabilities provided by on-chip inductors.
Based on a topology example of a low-power broadband low-noise amplifier, some aspects of broadband LNA design are illustrated. A prototype design, dimensioned in a 0.5 µm CMOS technology, is described and measured. It is explained how the underlying mechanism of the Non-QuasiStatic effect affects the input impedance even at practical RF frequencies, suggesting the use of NQS models also for these frequencies.

CMOS Receiver Front-Ends

The power dissipation of CMOS receiver front-ends strongly depends on the selected receiver architecture; When the heterodyne architectures are implemented in CMOS, they are significantly more power hungry than their bipolar equivalent. The excess power dissipation originates, apart from CMOS frequency limitations, from the intrinsically lower driving efficiency of MOS transistors,

(gm/I), when compared to bipolar transistors: heterodyne receiver architectures require external filters which need to be driven at a low impedance level. The high-Q filters are necessary since they perform the image rejection and the channel selection. Integrating these filters is also very difficult. Accordingly, the large number of off-chip components makes this architecture not attractive to build highly integrated receivers.

By using alternative receiver architectures, like the low-IF [1] or wideband-IF [2] receiver, the number of external nodes can be minimized, resulting in a significant reduction of the power consumption. The degree of integration can also be much higher as there is in principle no need for external filters; The image rejection is performed using a quadrature demodulation scheme while the channel selection is done by a simple low-pass filter. These functions can to a large extent be carried out in the digital domain.

In practice, of course, some intermediate filtering is needed because of the linearity limitations and the limited swing of practical building blocks: e.g., typically, an HF filter suppresses the out-of-band blocking signals to relax the dynamic range requirements. The use of a bandpass filter between the LNA and the downconversion mixer is however not necessary when a sufficiently linear LNA and a non-switching mixer are used; A bandpass filter removes the second and third order harmonics of the RF signal which are generated within the LNA due to a limited linearity; A switching mixer would mix these harmonics down to the wanted signal frequency. A CMOS technology has the advantage that it allows for the realization of highly linear mixers [3] based on transistors in the linear region. To complete the reception path, a power-efficient CMOS LNA is required.

CMOS Low-Noise Amplifiers

Low-power low-noise amplifier designs frequently employ inductors in a relatively high-Q LC-load which is tuned to the working frequency [2,4,5,6,7,8,9,10]. As the inductor cancels the capacitive part of the impedance on a certain node, the gain-bandwidth limitation is effectively neutralized, resulting in a low

power consumption. A drawback of using narrow-band techniques is the fast roll-off of the gain; This requires an accurate prediction of the resonance frequency and may even involve tuning. Moreover, the performance of integrated spiral inductors in standard CMOS technologies heavily depends on the substrate doping level: the higher the resistivity of the substrate (the lower the impurity concentration), the better the quality of the inductor. However, as typically a low-resistivity, highly doped substrate is used - mainly for its resistance to latch-up in digital CMOS circuits - eddy currents in the substrate significantly degrade the performance of the inductor. The inductor performance can still be improved by removing the silicon substrate underneath, but then we end up in extra processing steps.

There exist some alternative implementations that do not rely on high quality passive inductors. In [11], the inductor is 'simulated' by a transistor scheme (gyrator) [11]. The power consumption of this class of amplifiers is however quite high for a given noise figure because the gyrator generates excess noise. In [12], the spiral inductor is used as a compensation element in the feedback loop of a broadband amplifier. The quality-factor of the inductor does not need to be high as it is shunted with the 50 ohm impedance of the source.

In general it is believed that the implementation of RF low-noise amplifiers in CMOS at a power dissipation competitive with bipolar technologies requires the use of narrow-band techniques. In this paper the design of power-efficient broadband LNAs is explored, i.e., without using accurately tuned, high-Q LC tanks defining the amplification and without exploiting the 'overdrive' capabilities provided by on-chip inductors.

Broadband Low-Noise Amplifier Design

Based on a topology example of a low-power broadband low-noise amplifier, some aspects of broadband LNA design are illustrated. A prototype, dimensioned in a 0.5 μm CMOS technology, is described and measured.

Topology

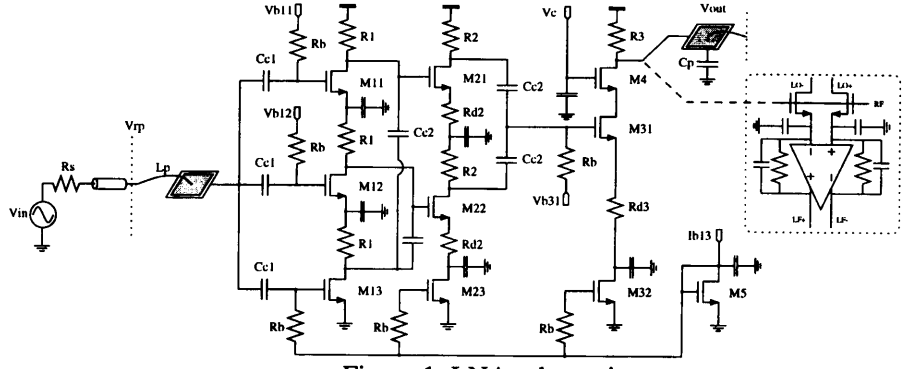


Figure 1: LNA schematic

The schematic of the LNA is shown in figure 1. The amplifier is composed of three cascaded nMOS gain stages: M1x, M2x and M3x. The multi-stage topology guarantees a sufficiently high reverse isolation so that the circuit can be used in a low-IF or zero-IF receiver as precisely these architectures are most suited to build highly integrated CMOS transceivers. In these receiver architectures, the local oscillator frequency and the wanted signal frequency are very close to each other or even the same. Leakage to the antenna of the large-amplitude LO signal that is driving the succeeding mixer stages needs to be prevented at all cost, because this can result in corruption of the wanted signal.

The power consumption of this LNA is reduced by applying a current-reuse technique [6]; Gain stages M1x and M2x consist of three, respectively two, nMOS gain blocks, stacked on top of each other in the same current branch. Application of this technique leads to considerable current savings when generating the transconductance that is needed to create gain and lower the noise figure (NF). In the case of this particular input-stage it results in an almost 3 times higher effective gm for the same DC current. This is not achieved with the current-reuse technique that was presented in [6], because of the limited KP of a pMOS device compared to an nMOS device. An other advantage of the current-reuse technique is that the available voltage headroom is used more effectively; A

single analog power supply of e.g. 3.3 V can be used for the whole front-end without causing a power penalty for the LNA.

Each gain block is loaded with polysilicon resistors Rx such that the gain is inherently broadband. The output nodes of each gain block are combined with capacitors Cc2 and form one effective signal node at radio frequencies. Because a fixed fraction of the coupling capacitors (~1/10) acts as a parasitic to the substrate, a trade-off is necessary between signal transfer and loading of the preceding stage. In practice, the value of the coupling capacitor Cc2 is limited to about 1 or 2 pF.

Large on-chip capacitors are inserted into the schematic to decouple the nodes between the stacked gain blocks, isolating them from each other at high frequencies. As a result, the gain blocks are activated above $gm/2\pi C$, generating a low-frequency –3dB point in the forward gain characteristic: the larger the capacitor value, the sooner the amplifier starts having its nominal gain characteristic. To ensure sufficient linearity, the last two stages are degenerated by polysilicon strips Rdx, lowering the effective voltage excursion at the gates. The capacitors that were originally inserted to provide virtual ground nodes between the different gain blocks can also be employed for this purpose; By reducing the capacitance value, the capacitors are operating as noiseless, frequency-dependent feedback elements. Yet, A drawback of lowering the capacitance value is that the nodes between the stages start moving, eventually degrading the total gain and compromising the reverse isolation. It is important to notice that due to the intrinsically better linearity of a CMOS device compared to bipolar transistors, an open-loop structure can be used for the input stage. In this way extra noise generated by the feedback system is avoided.

The linearity performance of low-noise amplifiers is generally described by the input-referred 3rd order intercept point (IIP3); At high input powers the signal quality mainly degrades due to in-band distortion components that are generated by 3rd order intermodulation. As out-of band signals are in general orders of magnitude larger than the wanted signal, the mixing of out-of-band

signals towards an in-band intermodulation product needs to be avoided by all means.

To understand the main mechanism behind 3rd order distortion in a submicron CMOS transistor [13], we start from the well-known equation for the drain current of a short channel transistor,

$$I_{ds} = \frac{\mu_0 C_{ox}}{2n} \cdot \frac{W}{L} \cdot \frac{(V_{GS} - V_T)^2}{1 + \Theta \cdot (V_{GS} - V_T)} \tag{1}$$

with

$$\Theta = \theta + \frac{\mu_0}{L_{eff} \cdot v_{\max} \cdot n} \tag{2}$$

where θ stands for the mobility degradation due to the transversal electrical field (surface scattering at the oxide-silicon interface) and the $\frac{\mu}{L_{eff} \cdot v_{max} \cdot n}$-term models the degradation caused by the longitudinal electric field (electrons reaching the thermal saturation speed). As the θ-term is small in today's technologies (increasingly better quality of the oxide-silicon interface), it can often be neglected relatively to the longitudinal term. For a typical 0.5 µm CMOS technology, the Θ-parameter equals about 0.9.

It can be seen from (1) that for large values of $V_{GS} - V_T$ the current becomes a linear function of $V_{GS} - V_T$. The transistor is then operating in the velocity saturation region. For smaller vales of $V_{GS} - V_T$, the effect of Θ consists apparently in 'linearizing' the quadratic characteristic...but in reality the effect results in an intermodulation behavior that is worse than in the case of quadratic transistors! Indeed, we will have a slightly lower amount of 2nd order intermodulation, but it comes at the cost of 3rd order intermodulation !

The following equations for the intermodulation ratios *IMx* [13] can be found by calculating the Taylor-expansion of the drain current around a certain $V_{GS} - V_T$ value:

$$IM2 = \frac{v}{V_{GS} - V_T} \cdot \frac{1}{(1+r)\cdot(2+r)} \tag{3}$$

and

$$IM3 = \frac{3}{4} \frac{v}{(V_{GS} - V_T)} \cdot \frac{v}{V_{SV}} \cdot \frac{1}{(1+r)^2 \cdot (2+r)} \tag{4}$$

where

$$V_{SV} = \frac{1}{\Theta} \tag{5}$$

represents the transit voltage between strong inversion and velocity saturation and

$$r = \frac{V_{GS} - V_T}{V_{SV}} \equiv \Theta \cdot (V_{GS} - V_T) \tag{6}$$

denotes the relative amount of velocity saturation.

Based on expression (4), one can directly derive an expression for *IIP3*:

$$IIP3 \equiv 11.25 + 10 \cdot Log_{10}\left((V_{GS} - V_T) \cdot V_{SV} \cdot (1+r)^2 \cdot (2+r)\right) \tag{7}$$

This value is normalized to 0 VdBm, the voltage that corresponds with a power of 0 dBm in a 50 Ohm resistor. For the 0.5 µm technology that was mentioned before and a $V_{GS} - V_T$-value of 0.2 Volt, *IIP3* equals +9.5 VdBm. It is worth noting that for a given L_{eff}, the intrinsic *IIP3*-value of a transistor is only a function of the gate overdrive.

In figure 2, formula (7) is evaluated for a minimum length transistor in three different technologies. As can be seen from the figure, for a given L_{eff} the linearity becomes better with increasing gate-overdrive. For small gate-overdrives, the increase in *IIP3* is proportional to the square root of $V_{GS} - V_T$. At high $V_{GS} - V_T$ values (near velocity saturation), the increase in *IIP3* becomes even more pronounced. However, this region of operation exhibits a very low transconductance efficiency (gm/I_{ds}), particularly for submicron transistors where this parameter is given by

$$\frac{gm}{I_{ds}} = \frac{2}{V_{GS} - V_T} \cdot \frac{1 + \Theta \cdot (V_{GS} - V_T)}{1 + 2\Theta \cdot (V_{GS} - V_T)} \tag{8}$$

The influence of L_{eff} on *IIP3* can also be extracted from figure 2; For practical values of the gate overdrive, the linearity gets worse with decreasing gate-length, because V_{SV} is proportional to L_{eff} (equations (2) and (3)). For large values of the gate overdrive, there is a point where the intermodulation performance of a short channel transistor gets better compared to a large-channel one, because the first already enters the velocity saturation region.

When a certain *IIP3* is required, there are basically two methods to ensure this: using a high enough $V_{GS} - V_T$ or using some kind of feedback mechanism like e.g., source degeneration. It can be shown that for the same equivalent *gm* and the same distortion

performance, the required DC current is lower when local feedback at the source is applied. It comes however at the cost of a larger transistor width, eventually compromising the amplifier bandwidth.

Let us now continue with the example topology. The LNA output buffer consists of a degenerated cascode amplifier. Cascode transistor M4 shields the LNA output from the drain parasitics of M31, at the same time providing a low impedance to these capacitances and ensuring extra reverse isolation. The optimum size of M4 is about half the size of M31.

The DC bias current of each stage is determined by current mirror M5-My3. Part of the transistors in the signal path are DC-biased by means of high-ohmic polysilicon resistors connected to their gates.

In order to combine low-noise operation with a low power consumption, the use of a low $V_{GS}-V_T$-value (~0.2 Volt) for the input transistors is required. This allows a high *gm* at low current levels. However, as low $V_{GS}-V_T$-values mean wide transistors, the input bandwidth decreases, degrading the high-frequency noise figure. In order to counteract the noise figure increase caused by input bandwidth limitations, the series inductance of a bondwire can be employed to generate a low-Q second-order boost at the amplifier's input. The low quality-factor of the input transfer function is ensured by the 50 ohm source impedance, making it

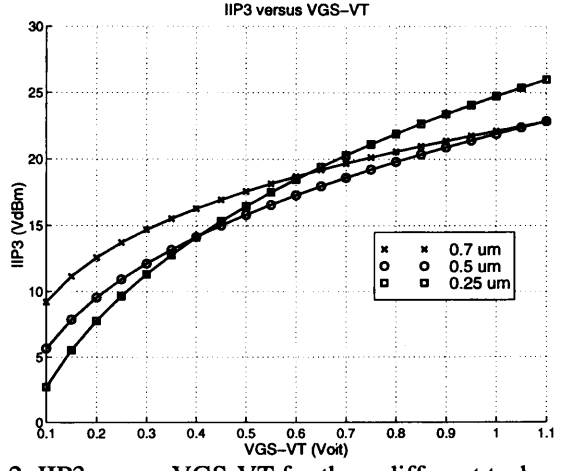


Figure 2: IIP3 versus VGS-VT for three different technologies.

robust with respect to variations in total package inductance. Even though the net quality-factor is low, the second-order section can change an input loss of 3 dB in a gain of 1 dB, extending the frequency band of low-noise operation.

In the classical receiver topologies, the low-noise amplifier has to drive an external 50 ohm filter. This is not an easy job for CMOS transistors. By using alternative receiver architectures (like e.g. low-IF or zero-IF), the problem of the low intrinsic driving capability of CMOS can be avoided; In this example, the LNA drives the capacitive input impedance of two (I&Q-channel) linear downconversion mixers (figure 1) in a low-IF receiver architecture. Because this particular mixer topology can handle the blocking levels (*IIP3* ~ +15 dBm) and because the image rejection is performed in the digital domain by a quadrature demodulation scheme, no intermediate filtering between the LNA and the downconverters is required any longer. Hence it is not necessary to lower the output impedance to 50 ohm because the mixers can be on-chip with their gates directly connected to the LNA output terminal. The impedance level at the output of the LNA is then only constrained by the desired output bandwidth. To evaluate the LNA performance with the correct load, a bonding pad with a capacitance value identical to the mixer input capacitance is connected to the LNA output terminal.

The transistor sizes, element values and biasing information of a prototype dimensioned in a 0.5 µm CMOS technology are summarized in table 1. The design draws only a DC current of 3.4 mA.

Table 1: Prototype dimensions.

	W/L	*Ids*	*Vgs-VT*	*Rx*	*Rdx*
M1x	160/0.5	1.2 mA	~0.22 Volt	350 Ohm	-
M2x	100/0.5	0.7 mA		350 Ohm	20 Ohm
M3x	200/0.5	1.5 mA		300 Ohm	70 Ohm
M4	100/0.5				-

Prototype measurement results

The LNA die is glued onto a thick-film alumina substrate and wire bonded from the pads to 50 ohm lines. The full S-parameter set with respect to the reference planes is measured from 100 MHz to 3 GHz and converted to the forward voltage gain *Vout/Vin*. Figure 3 shows both the forward and the reverse gain of the LNA. The amplifier provides a forward gain of 14.8 dB in a −3dB band ranging from below 100 MHz (near 50 MHz) up to 700 MHz. At 900 MHz, the gain is 9 dB. The reverse isolation is better than 41 dB over the full measuring range, making the amplifier suited for low-IF applications.

The power spectral density of the output noise is measured by connecting a low noise amplifier to the LNA output. By referring the measured noise spectrum to the input, using the S-parameters of the full noise-measurement set-up, noise figure data is extracted continuously from 700 MHz up to 1.1 GHz. A plot of the extracted noise figure is shown in figure 4. The noise figure stays between 2.3 dB and 3.3 dB up to a frequency of 970 MHz.

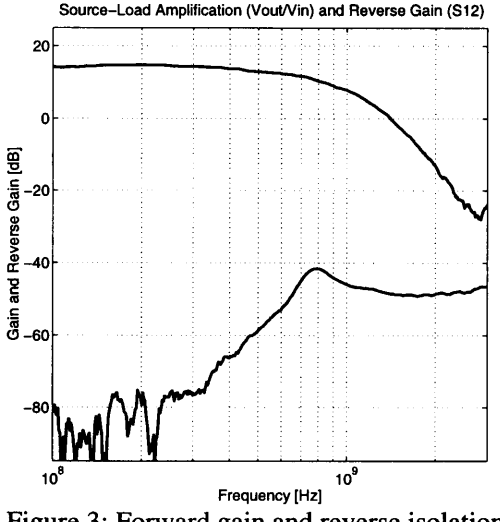


Figure 3: Forward gain and reverse isolation.

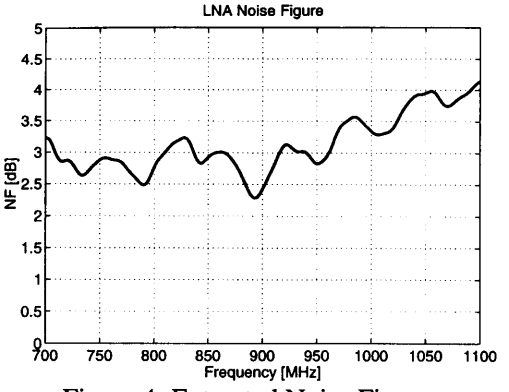


Figure 4: Extracted Noise Figure.

The input-referred 3rd order intercept point (*IIP3*) is extracted by sweeping the power of two closely spaced tones that are applied at the LNA input. Figure 5 shows the fundamental tones and the 3rd order intermodulation products for a two-tone test around 700 *MHz*. The two curves extrapolate to an *IIP3* of –4.7 VdBm

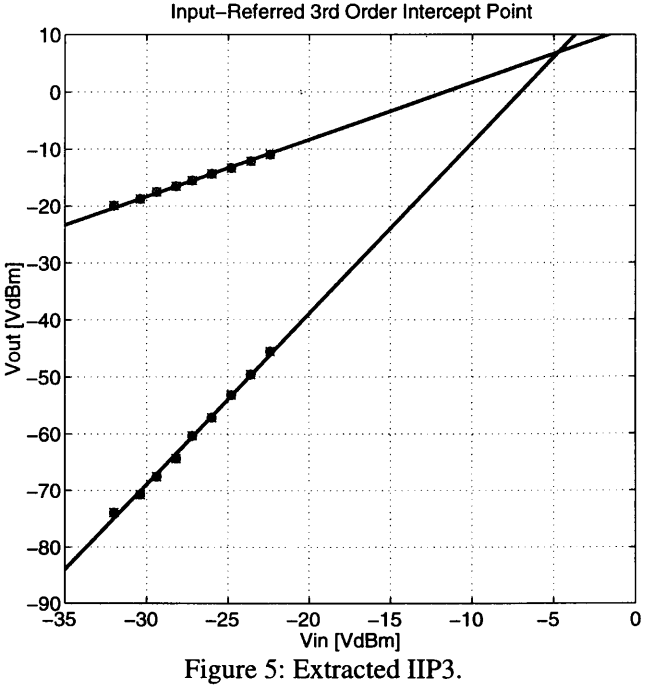


Figure 5: Extracted IIP3.

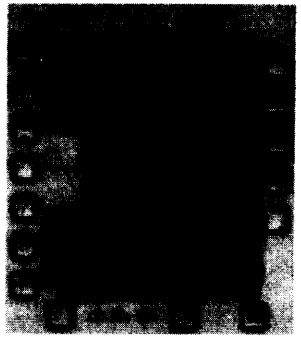


Figure 6: Die Photo.

Table 2: Prototype performance.

	3.0 Volt
	3.4 mA
	50 - 700 MHz
	2.3 - 3.3 dB
	14.8 dB 9 dB
	> 41 dB
	-4.7 VdBm
	1.0 x 1.2 mm²
Technology	0.5 µm CMOS low-ohmic substrate

(VdBm being defined as the voltage that corresponds with a power level of 0 dBm in 50 ohm).

Figure 6 shows a microphotograph of the 0.5 µm prototype. The total die area is 1.2 mm². The active core only occupies an area of 0.12 mm². A large part of the area is occupied by the decoupling capacitors that define the low-frequency –3dB point in the forward gain. By decreasing the dimensions of the decoupling capacitors, the area can be drastically reduced: e.g., by moving f_{-3dB} to 450 MHz, the total integrated capacitance can be divided by a factor of 9, resulting in a total area of only 0.37 mm². The total power consumption is only 10 mW for a power supply of 3.0 Volt. The measured performance of the prototype LNA is summarized in table 2.

LNA input-impedance

The reported performance is achieved without adding any external, tuned impedance transforming networks. When the prototype is used in a practical system, some impedance matching has to be done to correctly terminate the HF filter. If impedance matching is done at the input, a higher working frequency and a higher gain can be achieved at the cost of a higher frequency selectivity [6] and a lower linearity [10]. In this section, some reflections are made on the present input impedance of the prototype and on providing a matched input impedance.

Based on S11-measurements, the input impedance of the LNA can be extracted; Its real and imaginary part is shown in figure 7. At 1.5 GHz, series resonance occurs between the input capacitance and the bondwire inductance. The classical MOS model predicts a perfect series resonance between the bondwire inductor and the gate source capacitor: the real part of the impedance should be zero. However, the measurement yields a real part of 15 Ohm! The reason for this apparent "anomaly" is described in the next paragraphs.

It is well-known that the interaction of the ground inductance with the transconductance and the gate-source capacitance of transistors can generate a real part in the input impedance [10], its expression given by

$$(gm \cdot L_{gnd})/C_{GS} \qquad (9)$$

The technique is frequently employed to ensure a 50 Ohm input match [10]. However, because the chip is connected to the ground plane with six very short bondwires, this effect can not explain the 15 Ohm. The parasitic gate resistance can also not be responsible for this value, because the input transistors are made of very short fingers (for noise reasons)

There is also another effect that has to be taken into account. The gate-source capacitance is in reality not standing between the gate and source terminals (like in all HSpice models), but acts in fact as a distributed capacitance to the inversion layer. Depending on the distance along the channel, every infinitesimal transistor section

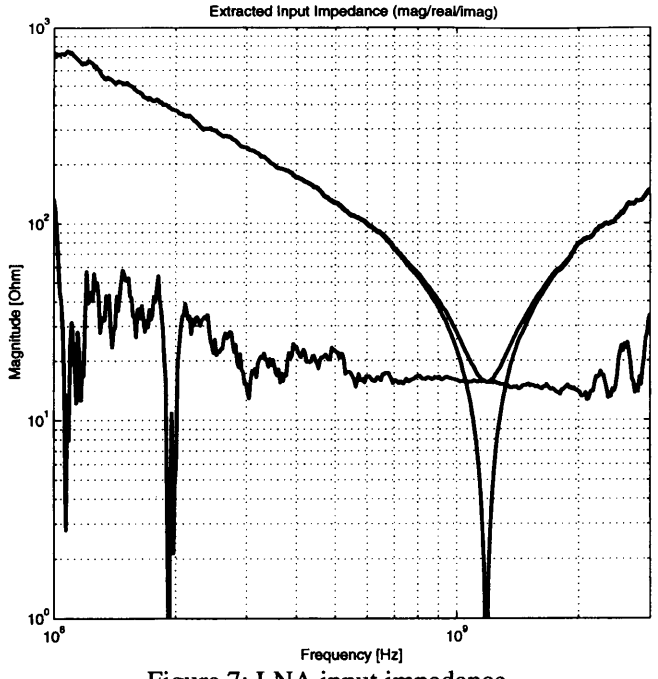


Figure 7: LNA input impedance.

sees an equivalent conductance to the source which depends on the local charge distribution. This phenomenon gives rise to the Non-QuasiStatic effect; When the gate electrode is excited at very high frequencies, the finite channel conductance limits the speed of the inversion layer charge build-up (τ_{NQS}).

In general, it is believed that this "loading"-effect only becomes important (read: "necessary to model") at very high frequencies e.g., at some fraction of the cut-off frequency, f_T. Indeed, the time constant of the phenomenon, τ_{NQS}, is in the order of several f_T's. Transposing this reasoning to the input impedance, this means that at normal frequencies the imaginary part (capacitive part) of the input impedance is much larger than the real part (resistive part).

Nevertheless, when an inductor (e.g., a bondwire) is put in series with the gate, at a certain frequency the capacitive part of the input impedance of the FET is cancelled by a series resonance phenomenon. At this frequency, the total input impedance behaves purely resistive! For all purposes, the oxide capacitance is made

transparent to the signal such that the channel is directly "visible". It is worth to point out that this frequency can be made arbitrarily low, just by changing the value of the series inductor!

Thus, although the working frequency is much lower than f_T or $1/\tau_{NQS}$, a non-quasi static model is needed to "correctly" describe the MOS transistor at RF. The non-quasistatic model in [14] characterizes the real part of the gate-source impedance with a resistor of $1/5gm$. When we evaluate this expression for the input section of the prototype, we get a value of 18.3 Ohm. This is close to the measured value of 15 Ohm, especially when we take into account that the calculated value of 18.3 Ohm is somewhat reduced by the parallel bondpad capacitance.

The classical MOS impedance matching technique [10], was introduced to cope with the 'purely capacitive MOS input impedance' by artificially generating the 50 Ohm impedance using inductive source feedback; In this matching scheme, an input voltage component that is in-phase with the input current is produced by first making a capacitive current component (90 degrees out of phase relatively to the input current) and by rotating it back over 90 degrees by steering it through the source inductor:

$$v_{in,real} = i_{in} \cdot \frac{1}{j\omega C_{GS}} \cdot gm \cdot j\omega L \qquad (10)$$

Although this method gives satisfying results [2,10], there are still some imperfections. First of all, the resistive effect originating from the non-quasistatic mechanism, the bulk transconductance (gm_b) and mainly the parasitic capacitances C_{SB} and C_{GD} are not taken into consideration. Secondly, the current that is steered through the gate electrode is -by construction- at the same impedance level as the source, Rs (e.g. 50 Ohm). This constraint limits the effective transconductance to

$$gm_{eff} = \frac{\omega_T}{\omega} \cdot \frac{1}{Rs} \qquad (11)$$

However, as the gate-source impedance of a MOS transistor already exhibits a real part owing to the NQS effect (equal to $1/5gm$[14]), it is also possible to directly match the source

impedance Rs to the NQS channel resistance.Then, the effective transconductance can be expressed by

$$gm_{eff} = \frac{\omega_T}{\omega} \cdot \frac{1}{\sqrt{Rs \cdot Rg}} \qquad (12)$$

with

$$Rg = R_{parasitic} + \frac{1}{5gm} \qquad (13)$$

It is worth pointing out that this transconductance value can be significantly higher than (11). Moreover, as the source of the transistor is grounded, gm_b and C_{SB} do not play any role. The matching section can be implemented by employing the well-known two-element matching schemes.

Both matching techniques can be combined into one hybrid method, optimizing the trading-off between gain, linearity and noise.

Comparison to other realizations

Table 3 compares the performance of the CMOS LNA prototype to both existing bipolar realizations and previously published CMOS low-noise amplifiers (mostly narrow-band LNAs). As can be seen from the table, the implemented broadband LNA provides an interesting alternative to and even competes with existing narrow-band designs.

Table 3: performance comparison.

Ref.	Freq. [MHz]	NF [dB]	Power [mW]	Gain [dB]	IIP3 [dBm]	Process [µm]	*	**	***
[4]	900	7.5	36 (3.3 V)	11		0.8 CMOS	x		
[5]	900	2.0	15 (3.0 V)	20	-20.0	0.6 CMOS		x	
[6]	900	2.8	20 (2.7 V)	15.6	-3.2	0.5 CMOS	x		
[7]	940	2.4	- (3.0 V)	13.2	-9	Bipolar	x		
[8]	900	2.2	40 (5.0 V)	16	-10.0	BiCMOS			
[9]	1000	3.5	15.6 (3.0 V)	21	12.5	1.0 CMOS	x	x	x
[10]	1500	3.5	30 (1.5 V)	22	-9.3	0.6 CMOS	x	x	
[11]	940	5.3	41 (3.3 V)	20	-8.6	0.9 CMOS			
[12]	900	2.7	35 (2.7 V)	24	-14.2	0.4 CMOS			
[this]	900	3.3	10 (3.0 V)	14.8 (9)	-4.7	0.5 CMOS			

* uses external, tuned resonator(s) for impedance matching
** relies on integrated inductors that need to be accurately tuned
*** substrate underneath inductors is removed to improve Q factor

Conclusions

The design of submicron CMOS broadband LNAs is clarified using a case-study. Special attention is given to the nonlinearities that occur in deep-submicron CMOS transistors. The case study concludes by describing a measured prototype. It is described why and how the same mechanism that gives rise to the non-quasistatic effect is active and can be made visible at frequencies much below the cut-off frequency, suggesting the use of non-quasistatic models at these frequencies. A comparison illustrates that broadband LNAs can provide an interesting alternative to and can even compete with narrow-band designs.

References

[1] J. Crols and M. Steyaert, "Low-IF Topologies for High Performance Analog Front-Ends of Fully Integrated Receivers," accepted for publication in IEEE Trans. on Circuits and Systems II, 1997

[2] J.C.Rudell et al., "A 1.9 GHz Wide-Band IF Double Conversion CMOS Receiver for Cordless Telephone Applications," in IEEE J. Solid-State Circuits, vol.32, no.12, pp.2071-2088, December 1997.

[3] J. Crols and M. Steyaert, "A 1.5 GHz highly linear CMOS downconversion mixer," in IEEE J. Solid-State Circuits, vol.30, no.7, pp.736-742, July 1995.

[4] S. Sheng et al., "A Low-Power CMOS Chipset for Spread-Spectrum Communications," Proc. ISSCC, pp. 346-347, San Francisco, Feb. 1996.

[5] B. Razavi, "A 900-MHz CMOS Direct Conversion Receiver," Digest of Technical Papers, 1997 Symposium on VLSI circuits, pp. 113-114, Kyoto, June 1997.

[6] A. N. Karanicolas, "A 2.7V 900MHz CMOS LNA and Mixer," Proc. ISSCC, pp.50-51, San Francisco, Feb. 1996.

[7] K. Irie et al., "A 2.7 V GSM Transceiver IC," Proc. ISSCC, pp. 302-303, San Francisco, Feb. 1997.

[8] R.G. Meyer and W.D. Mack, "A 1GHz BiCMOS RF Front-End IC," IEEE J. Solid-State Circuits, vol.29, No. 3, pp.350-355, Mar. 1994.

[9] A. Rofougaran et al., "A 1 GHz CMOS RF Front-End IC with Wide Dynamic Range," Proc. ESSCIRC, Lille, pp.250-253, Sept. 1995.

[10] D. Shaeffer and T. Lee, "A 1.5 V, 1.5 GHz CMOS Low Noise Amplifier," Proc. VLSI Circuits Symposium, Honolulu, pp.32-33, June 1996.

[11] Y.J. Shin and K. Bult ,"An Inductorless 900 MHz RF Low-Noise Amplifier in 0.9um CMOS," Proc. CICC, Santa Clara, pp.513-516, May 1997.

[12] J. Janssens, M. Steyaert and H. Miyakawa, "A 2.7 Volt CMOS Broadband Low Noise Amplifier," Digest of Technical Papers, 1997 Symposium on VLSI circuits," pp.87-88, Kyoto, June 1997.

[13] M. Steyaert et al., "RF CMOS Design: Some Untold Pitfalls," Proc. AACD, Lausanne, April 1996.

[14] Y. P. Tsividis, "Operation and modeling of the MOS transistor", Chapter 6, pp, 399, McGraw Hill, 1987.

Put your power into SOA LNAs!

Peter Baltus
Philips Research Laboratories
Eindhoven, The Netherlands

Low noise amplifiers are ultimately limited in performance and power dissipation by parasitics introduced by the IC process. The important parasitics can be determined by investigating the tradeoffs between power, performance and the performance limits of simple building blocks. Most of these parasitics are related to the silicon substrate. Silicon-On-Anything is an IC technology in which the substrate is completely substituted by another material. The technology also includes an NPN device optimized specifically for low power RF. By adapting the LNA design methods, a reduction in power dissipation by one order of magnitude has already been demonstrated. Directions for further improvements are indicated.

1 Introduction

The main function of a Low Noise Amplifier (LNA) is to amplify the RF signal (typically from an antenna) to levels that can be more easily processed in other circuits such as mixers, without adding significant amounts of noise in the LNA itself (fig. 1).

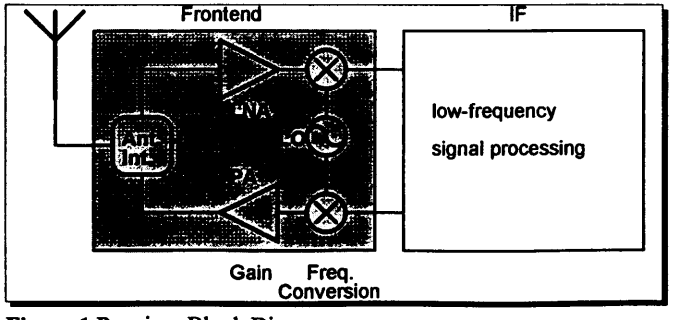


Figure 1 Receiver Block Diagram

This main function needs to be implemented within the limits for the distortion

levels set by the system. In this paper, we will concentrate on integrated LNAs for portable narrowband RF systems in the 1GHz to 5GHz range. In addition to the RF performance, low power dissipation and low cost are important properties for such LNAs.

1.1 Problem description

The signal levels at the input of the LNA can be quite small, often barely above (and sometimes indeed far below) the thermal noise floor. Since many systems require a large (in the order of 10dB) signal to noise ratio at the output of the receiver in order to achieve satisfactory performance, the main concern in an LNA is to amplify the signal without adding a lot of noise. The amplified signal can then be further processed by other circuits, whose noise contributions will be smaller compared to the now much larger signal.

1.2 Example

In a typical cellular communication system, the lowest signal level from the antenna that needs to be received reliably is about -102dBm or 63fW (!). Such a system has some signal loss between the antenna and the LNA input because of filters, PCB interconnect and external matching circuits, that can add up to about 3dB. This leaves only -105dBm or about 32fW of signal at the LNA input, which is about 15dB over the thermal noise floor in a 200kHz bandwidth.
The required SNR for such a system is about 9dB, which leaves in this case only 6dB (=15dB-9dB) for the total receiver noise figure. To meet this receiver noise figure, both a high gain and low noise figure of the LNA are essential.

1.3 Outline of the paper

This paper will argue that significant progress in low power LNAs can only be made by a combination of improvements in both the design methodology and IC technology. First, the factors that limit LNA performance will be investigated (section 2). One of these limits, the IC technology, will then be further discussed in section 3. In this section, we will also introduce the Silicon-on-Anything IC process, which reduces several of these limits. In section 4, existing LNA designs will be presented, and advanced design techniques that take advantage of this IC process will be proposed.

2 Limitations for low power

There are a number of limitations in the implementation of an LNA that force the designer to make a trade-off between the most important specification points of an LNA that were mentioned in the previous sections:

- gain
- noise figure
- distortion
- power dissipation

Moreover, this trade-off has to take into account the boundary conditions defined by the system in which the LNA should work: source impedance, load impedance, input and output matching requirements, power supply voltage and packaging.

Assuming that we are able to find a trade-off by means of circuit and/or process optimizations that allows a significant reduction in power, a number of consequences are unavoidable:

- The base emitter voltage (V_{be}) of bipolar processes has a tendency to increase in successive generations of RF bipolar processes, because the current density in the base emitter junction is increased to achieve smaller time constants. Since it is necessary to have a power supply voltage of at least one V_{be}, and preferably more, the power supply voltage for bipolar circuits is unlikely to drop below 0.9V. Current LNA designs are typically optimized for power supply voltages of 1.8V or 2.7V. Therefore, large reductions in power dissipation (by about an order of magnitude or more) will have to be achieved through reduced supply currents. Lower currents at (almost) constant power supply voltages result in higher impedance levels on chip. External impedance levels are unlikely to increase much above 200Ω because of practical limits in e.g. filter and transmission line impedances. Low power design will therefore require impedance adjustments at inputs and outputs.

- Driving external circuits with relatively low impedance levels tends to require a large amount of power, compared to the on-chip signal processing. Low power designs will therefore tend to migrate towards

high integration levels with as few external connections as possible. Of course, this requires that the IC process provides most necessary components with sufficient quality, such as inductors, varicaps, switches, etc. An added benefit of this migration is that it supports miniaturization and cost reduction.

2.1 Getting signal energy into the LNA

The transfer of available signal energy from the input source into the LNA is determined by the source impedance and the input impedance of the LNA. This input impedance can be chosen from three ideal values:

- optimum termination impedance for the filter that often precedes the LNA
- optimum impedance for achieving the highest gain (power match)
- optimum impedance for achieving the lowest noise figure (noise match)

Ideally, these three impedances should be the same value. The power match and filter termination impedances are typically set in the system definition, and are often the same or at least very close. What is left for the LNA designer is to either make the noise match impedance (almost) identical to the power match impedance, or to provide at least a good compromise between these impedances.

This design problem is often complicated by the influence of the package and bondpad-related parasitics. The package typically provides both a bondwire inductance, a number of parasitic capacitances to ground, and mutual inductive and capacitive coupling between the pins. Fig. 2 shows a simplified electrical model for an IC package.
For a more accurate model, additional inductive and capacitive coupling between non-adjacent pins should be added.

The quality of the parasitic elements in the package is typically very high (> 100), and therefore the main effect of the package is to introduce crosstalk and frequency dependent impedance transforms:

- The dominant crosstalk mechanism depends on the impedance levels at the pins. For low impedance levels (below 100Ω) magnetic coupling typically dominates. For high impedance levels (above 1kΩ) capacitive coupling is likely to be dominant. Low power ICs are

therefore more likely to suffer from crosstalk through capacitive coupling. Attention should be paid to the accuracy of the capacitive elements in the model, since circuits with more traditional impedance levels are less sensitive to inaccuracies of these capacitive elements. Apart from these differences, all the usual tricks for reducing crosstalk (symmetry, shielding through grounded adjacent pins etc.) still apply.

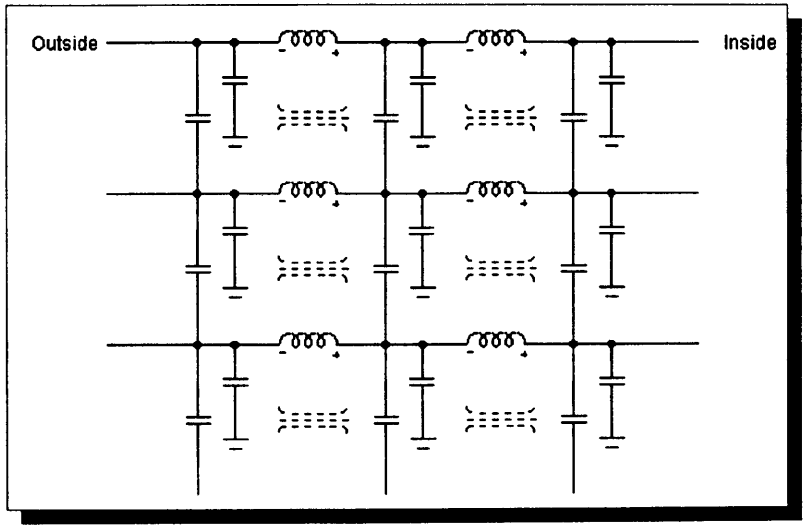


Figure 2 Simplified electrical model for IC package parasitics

- The impedance transform, since it is almost loss-less, can easily be taken into account in narrowband systems by adjusting the on-chip input impedance of the LNA. The narrowband nature of the systems is important since the transform by the package is frequency dependent, and providing a stable input impedance over a larger range of frequencies at the package pin requires different techniques. Of course, this does limit the possibilities of packaging the same circuit in different packages.

The bondpad and ESD diode in most IC processes have parasitic capacitances with significant losses, in the order of 1pF and 100Ω. The losses introduced by these parasitics depend very much on the ratio between the impedance of these parasitics and the input impedance of the LNA. In practice, it will be difficult to achieve good noise performance

for an LNA with input impedance levels much above 200Ω because of these parasitics. Low power LNAs will therefore have to include low loss (and therefore low noise) impedance transforms, at least at the input. Building this impedance transform with discrete components or micro strips outside the IC is sub-optimal, because now the bondpad and ESD device still operate in a high impedance environment. The best solution is to either provide low-parasitic bondpads and ESD diodes, or high quality components that can be used to build on-chip low loss impedance transformers.

2.2 *Getting signal energy out of the LNA*

At the output, we have similar constraints as at the input of the LNA. However, if the LNA is to be used as a part of an integrated frontend in which it is not necessary to have external filtering between LNA and mixers (e.g. image cancelation or zero-IF frontends) then the requirement for providing the optimum source impedance of the external filter is of course no longer necessary. Also, the output impedance level might be jointly optimized with the design of the input of the next stage, allowing for much larger freedom than would be available otherwise. If an external filter is required, similar concerns about packaging, bondpad and ESD parasitics apply as at the input of the LNA, although typically the noise performance is not as critical as at the input because of the gain provided by the LNA itself.

One area that needs special attention is the implementation of the output impedance of the LNA. In a typical design, this impedance is physically present at the output, resulting in a dissipation in the LNA at least equal to the dissipation in the load. A traditional solution is to give up on power matching between stages on an IC [1] to achieve better power efficiency. This approach inherently achieves lower gain than would be theoretically possible for the given dissipation, and is therefore not very attractive. Alternative approaches have been presented [2] that might provide an optimum combination of gain and matching also at RF frequencies.

In many systems, it is often implicitly assumed that the *IP3* not only defines the distortion for very low input signal levels, but also indicates the large signal behaviour of a circuit. Strictly speaking, this assumption is not correct since the *IP3* is based on an extrapolation from the

distortion at very low signal levels. There are many circuits in which the traditional "rule of thumb" (*IP3* is about 10dB above 1dB compression) does not apply. Since most RF circuits are still class A circuits, current consumption is mostly determined by the 1dB compression point (P_{1dB}). Low power LNA designs therefore will use a combination of techniques to improve *IP3* while using non-class A circuits to simultaneously allow high P_{1dB} at low quiescent current levels.

2.3 *Minimizing the DC power energy input into the LNA*

The power dissipation of the LNA is determined by a combination of trade-offs between noise, gain, distortion and power dissipation, and by parasitic transistor behaviour introduced in the fabrication of the active device. We will first investigate the trade-offs, and then the parasitic transistor behaviour.

2.4 *Power vs. Performance trade-offs*

Building Block	Impedance Transform	Parallel	Cascade
Z_{in}	αZ_{in}	$\frac{Z_{in}}{2}$	Z_{in}
Z_{out}	Z_{out}	$\frac{Z_{out}}{2}$	Z_{out}
G_p	G_p	G_p	G_p^2
F	F	F	$F+\frac{F-1}{G_p}$
$IP3$	$IP3$	$2\,IP3$	$\frac{IP3}{G_p+1}$
P_{dc}	P_{dc}	$2\,P_{dc}$	$2\,P_{dc}$

Table 1 Trade-offs between power dissipation and performance

A general discussion of LNA performance limits is complicated by the many circuit structures that can be used to implement an LNA. Therefore, we will first investigate the trade-offs that are possible by building LNAs from imperfect building blocks by using transforms that are independent of the structure of the building block itself. In each of these transforms, we will assume a basic gain block with a power gain G_p, noise factor F, third order intercept point $IP3$, input impedance Z_{in}, output impedance Z_{out} and power dissipation P_{dc}. The parameters of the transformed blocks are compared to the parameters of the original building block. An overview of the transforms that will be discussed is shown in tables 1 and 2.

Building Block	Par/Cascade Hybrid	Attenuation	Feedback
Z_{in}	$\frac{Z_{in}}{G_p+1}$	Z_{in}	Z_{in}^*
Z_{out}	$\frac{Z_{out}}{G_p+1}$	Z_{out}	$\frac{Z_{out}}{1+\alpha G}$
G_p	G_p^2	αG_p	$\frac{G_p}{1+\alpha G_p}$
F	$F+\frac{F-1}{G_p}$	$1+\frac{F-1}{\alpha}$	F
$IP3$	$IP3$	$\frac{IP3}{\alpha}$	$(1+\alpha G)IP3$
P_{dc}	$2(G_p+1)P_{dc}$	P_{dc}	P_{dc}

Table 2 More power dissipation vs. performance trade-offs

To simplify the expressions, we will assume that the input impedance

equals the output impedance. This is not essential for the transforms, since this can always be achieved by putting an impedance transformer at the input or the output of a building block. In all transforms, we will also investigate how the transform scales to a transform with constant power dissipation.

2.4.1 Impedance transform

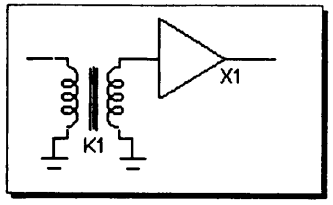


Figure 3 Impedance transform

In narrowband systems, a low loss transformer can often be implemented cheaply at the PCB level, and sometimes (depending on the IC technology) also on chip. When put in front of our building block, an impedance transform with a factor α only changes the input impedance but leaves all other performance parameters unchanged. However, this transform will prove to be very useful as a part of the subsequent transforms.

2.4.2 Parallel building blocks

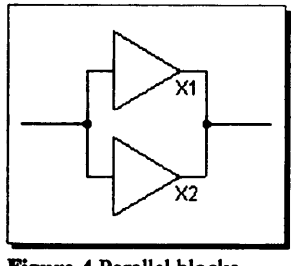


Figure 4 Parallel blocks

The parallel transform results imply that we can always improve the IP3 of a block with at least 3dB at the cost of doubling the current. As an added inconvenience, the input and output impedances are halved, but this can be easily remedied by applying impedance transforms with α=2 at the input and α=0.5 at the output. There might be methods or other circuit structures that provide more than 3dB for a doubling of current, but 3dB is always possible. Please note that the inverse is not always true: reducing the IP3 by 3dB to save 50% current only works if it is possible to actually scale all devices to half their size with all parasitics scaled proportionally. For any given IC technology, there is an end to this type of scaling because of limits in lithography and perimeter parasitics, but a technology which provides scaled parasitics to very low currents and device sizes offers additional flexibility in design.

2.4.3 Cascaded blocks

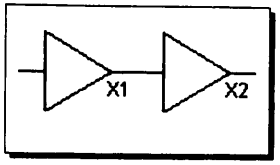


Figure 5 Cascaded blocks

For two cascaded blocks, the gain increases exponentially with power dissipation. However, in contrast to the parallel transform, both the noise figure and the IP3 change (compared to a single block) because of this transform. If the gain is high compared to the noise factor, then the total noise figure will be only slightly higher than the noise figure of a building block. However, the intercept point will decrease with approximately the gain. It would be preferable to have a transform that trades gain vs. power dissipation without affecting the noise and distortion.

2.4.4 Parallel-Cascade Hybrid

By combining the cascade and parallel transforms, we can approximate the transform which increases the gain while keeping noise figure and IP3 constant. This is achieved by putting G_p+1 groups of two cascaded building blocks in parallel. For gains much larger than 1, the power dissipation increases with twice the gain increase. For building blocks with lower gain, this factor is even worse. Obviously, gain is expensive. If the gain is much larger than the noise factor, the noise factor does not change significantly.

It would perhaps have been more obvious to build this hybrid transform by cascading a first stage, consisting of a two parallel building block, and a second stage, consisting of *2* G_p building blocks in parallel. The stages would then be coupled through an impedance transformer with $\alpha = Gp$. This is more consistent with the common practice in receivers to increase the IP3 of subsequent stages about proportionally to the accumulated gain up to that point. It is left as an exercise to the reader to verify that this approach results in the same trade-off for gain vs. power dissipation, but more complex math.

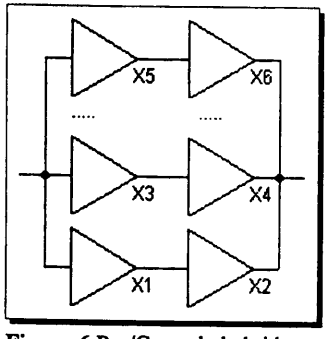


Figure 6 Par/Cascade hybrid

2.4.5 Attenuation

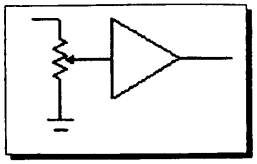


Figure 7 Attenuation

Attenuation is typically easy to implement both on and off chip. It is assumed that the attenuator itself does not introduce significant additional noise. For noise factors much larger than 1, the dynamic range of the total circuit is not significantly affected, but the gain is reduced.

2.4.6 Feedback

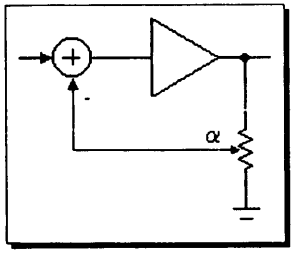


Figure 8 Feedback

Noiseless negative feedback with a factor α increases the dynamic range at the input of the LNA with the same factor as the gain decreases, as opposed to the attenuator, where the gain decreases with constant dynamic range. Since this increased dynamic range is obtained at zero cost in power dissipation, feedback is preferable over attenuation.

The expressions for feedback are valid both for positive and negative values of α. The gain for negative values of α (positive feedback) increases with about the same factor as IP3 decreases, while the noise factor remains unchanged. By combining positive feedback with parallel building blocks, it is possible to increase the gain proportionally to power dissipation without affecting IP3 or noise factor.

These transforms show that for every circuit, the gain or the IP3 can be increased independently, at the cost of a similar increase in power dissipation.

2.5 Technology Limits

Now that we have discussed how we can transform the performance by building combinations of building blocks with feedback and impedance transformers, the performance limits of building blocks themselves needs to be investigated. This discussion will be based on bipolar devices because of their good RF performance (compared to CMOS) and relatively low cost (compared to GaAs). We will start by investigating the properties of a single transistor gain stage. More complex gain stages can be built from this single transistor stage by applying combinations of the

transforms discussed in the previous sections.

2.5.1 Transistor configuration

A bipolar transistor can be used in three basic configurations: common base, common emitter and common collector. The power gain in each of these configurations is different. As a very first approximation, the common collector configuration generates mainly current gain, the common base configuration generates mainly voltage gain, and the common emitter configuration combines approximately the current gain of the common collector with the voltage gain of the common base. This is demonstrated in fig. 9, which shows the power gain for a transistor in the three basic configurations with the same bias current. From a gain vs. power point of view, obviously the common emitter is the most attractive.

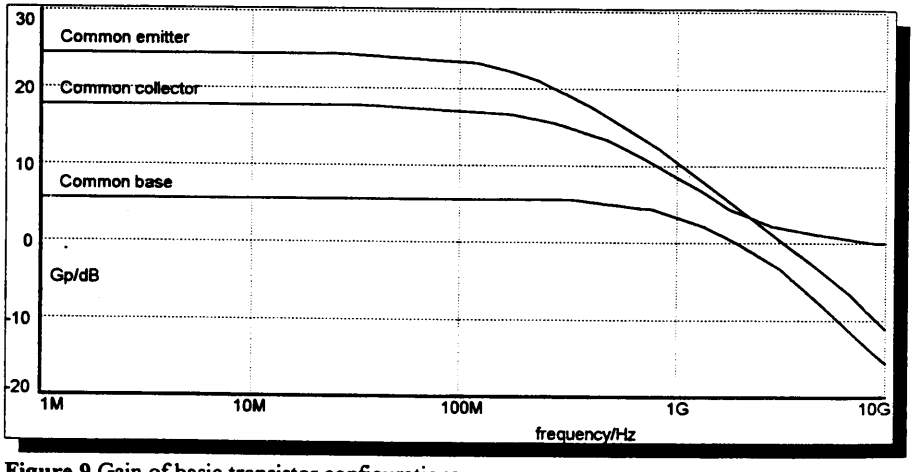


Figure 9 Gain of basic transistor configurations

2.5.2 Transistor figures of merit

The performance of a common emitter transistor can be described through a few figures of merit. The most commonly used figure of merit is f_T. However, f_T is not a very relevant figure of merit for a common emitter stage, since the conditions in which the f_T is defined (shorted output, input driven by current source) are rather unusual for RF circuits. Instead, we use a figure of merit which drives the transistor from a voltage source and loads the transistor for 20dB of voltage gain. This

gives a figure of merit for transistor input bandwidth, f_v, and another for output bandwidth, f_{out} [3]. The input bandwidth is mostly determined by the base resistance and the combined parasitic capacitors at the base. Since the diffusion capacitance varies proportional to the collector current, f_v changes approximately inversely proportional to the collector current (fig. 10).

The output bandwidth is typically dominated by the collector-related parasitic capacitances and external load resistance. Because f_{out} is defined for constant gain, the external load resistance decreases with increasing current. Therefore, f_{out} changes proportional to the collector current. The largest common emitter bandwidth is achieved for the current at which $f_{out} = f_v$. This bandwidth is called f_a [4]. In typical bipolar and BiCMOS processes, this f_a is located at bias levels higher than the biasing for peak f_T. This is the reason why in many RF circuits common emitter circuits that require high bandwidths are biased near peak f_T, even though the f_T itself does not affect the bandwidth significantly. For low power LNAs it would be more useful to have devices where the f_a is located at much lower currents. To achieve this, typically the collector-substrate capacitance has to be reduced significantly.

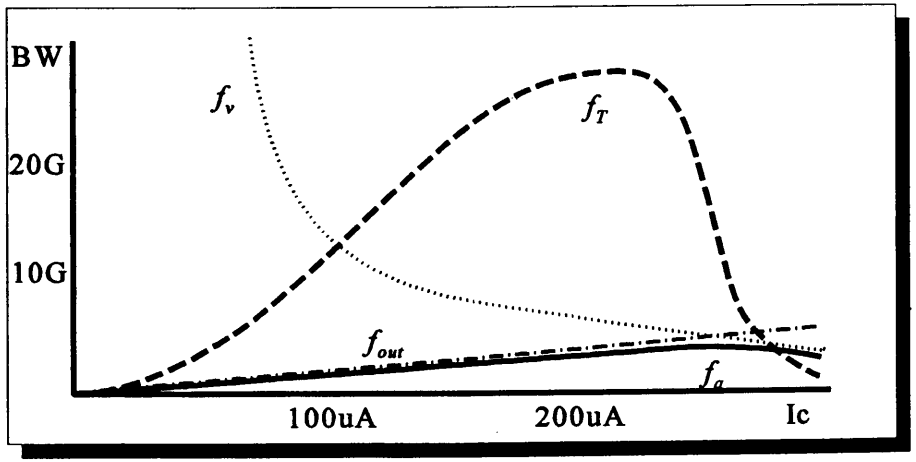


Figure 10 Input bandwidth f_v, output bandwidth f_{out} and f_T versus collector current.

The distortion of a common emitter stage at low frequencies is dominated by the exponential voltage to current conversion in the base-emitter junction. At higher frequencies, the non-linearity of parasitic capacitors

in the device start to contribute significantly to the total distortion. However, once a device is optimized to provide high output bandwidths, the influence of these capacitors becomes relatively small again. This implies that optimized RF devices have similar input *IP3* voltages, which only depend on the biasing of the transistor. This is often used as an argument to "prove" that the power dissipation of an LNA (or any RF circuit, for that matter) only depends on the dynamic range and does not significantly benefit from improvements in device parasitics. However, in the previous sections it has been shown that the important trade-off is not between power dissipation and IP3 at the input, but power dissipation and *IP3* at the output. For that reason, any improvement in gain of a common emitter stage that does not significantly affect the input *IP3* can be turned into an input *IP3* improvement at the original gain through feedback. In a common emitter stage with an optimized RF device, this applies to the power gain that occurs after the base emitter junction non-linearity. This implies that a high RF output impedance contributes significantly to the improvement of the power dissipation vs. *IP3* trade-off.

2.5.3 Interconnect

The output bandwidth of a common emitter stage is in many existing

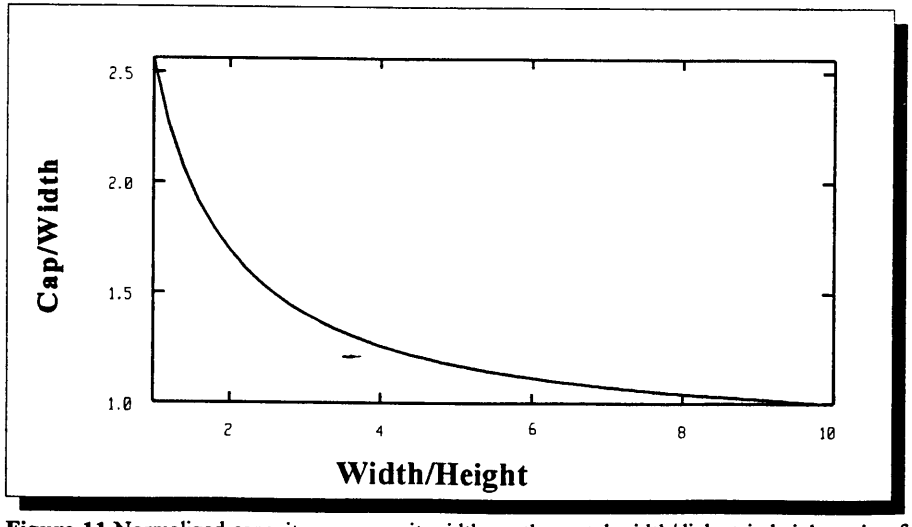


Figure 11 Normalized capacitance per unit width vs. the metal width/dielectric height ratio of interconnect

designs already significantly reduced (in the order of 30%) because of the parasitic capacitance of the interconnect. In low power LNAs, the higher impedance levels make the circuits even more sensitive to the interconnect capacitance. Fringing makes it very difficult to further reduce the capacitance of interconnect by making wires narrower (fig. 11).

2.5.4 Passive devices

For the on-chip impedance transforms, and in order to minimize the number of times the signal needs to go off-chip, high quality passive components are required. In most IC processes, capacitors with sufficiently high Q are available. However, inductors are limited by substrate related parasitics (parasitic capacitance and eddy currents) to Q values around 10 for typical bipolar processes. Even worse, the resonance frequency of these inductors tends to be so low that the maximum Q for inductors with the high values needed for low power LNAs (typically 10nH and up) occurs at frequencies below the operational frequencies of many modern communication systems.

The need for high quality inductors is most obvious [5] but higher integration levels and lower power dissipation could also be facilitated by other passive components such as varicaps, switches and strip lines.

3 Technologies for low power

Obviously, it is necessary to drastically change the IC process in order to reduce LNA power consumption by one order of magnitude or more. The main problems (transistor performance at low currents, interconnect parasitics and high quality inductors) are related to the silicon substrate. There are several processes (SOI, SOS, GaAs) which reduce or eliminate the losses associated with the substrate. In the next section we will introduce SOA, yet another process that eliminates substrate losses. The important differences with the existing solutions are:

- SOA includes an unusual transistor that is specifically optimized for low power RF circuits
- SOA can be processed in a standard bipolar silicon fab, and only relies on SOI starting material and postprocessing to remove the substrate.

3.1 *SOA technology*

Silicon-On-Anything (SOA) is a bipolar IC technology in which the active layer of an SOI wafer is glued onto a substrate of choice *after* processing [6]. Afterwards, the silicon substrate of the SOI wafer is removed completely by selective etching. Glass is typically selected for the substrate because it is cheap and has low losses over a wide frequency range.

From a design point of view, this process offers very significant improvements in virtually all critical parameters that have been identified in the previous sections:

- A lateral NPN transistor with 0.1μm² emitter area using 0.5μm lithography. With this NPN, a common emitter stage provides 2.4GHz of bandwidth at 20dB of gain with only 10μA current [7];
- Interconnect (including bondpads) with 5 to 20 times lower parasitic capacitance to ground;
- Almost perfect isolation between circuit blocks;
- Integrated inductors with Q values up to 35;
- An *individual* trade off between f_T, base resistance, Early voltage and breakdown voltages for *every* transistor in the design by changing the collector drift region (L_{cdr}).

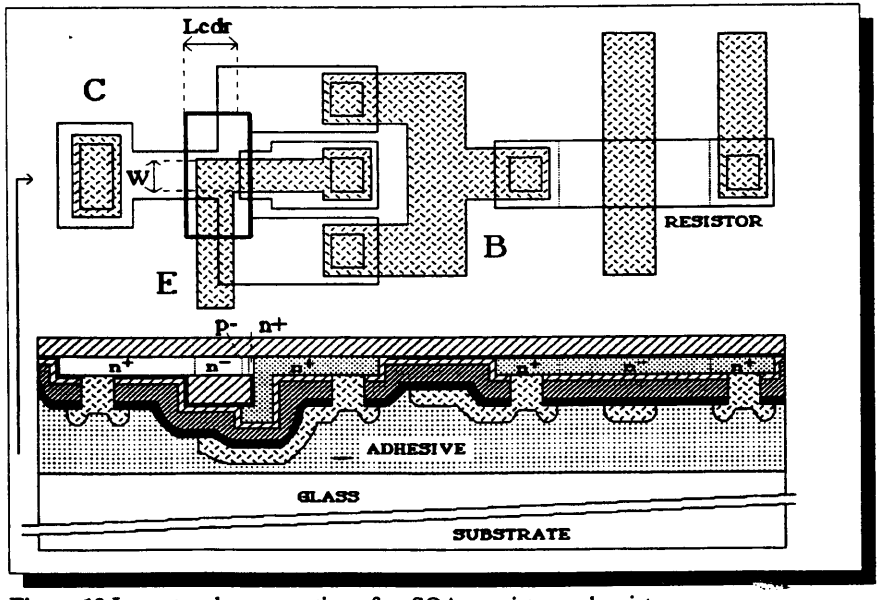


Figure 12 Layout and cross-section of an SOA transistor and resistor

The process also provides 15kΩ/□ poly resistors, 1.5nF/mm² capacitors, PIN diodes, varicaps, PNPs and JFETS, and 2.5 metal layers, all in only 14 mask steps. Fig. 12 shows the layout and cross section of an SOA transistor connected to a polysilicon resistor.

3.2 *SOA device parameters*

Since the SOA NPN device is a lateral device, it is only possible to set the "width" of the emitter by layout. The"length" is set by the thickness of the epi layer, which allows 0.2μ and even 0.1μ emitter widths using standard 0.5μ lithography. The emitter dimension that can be changed is the distance between the two base contacts. For RF circuits, this dimension is virtually always chosen to be minimum.

The lateral device architecture allows the collector drift region (L_{cdr}) to be chosen through layout. This allows a trade-off between f_T, base resistance, and breakdown voltages on a transistor-by-transistor basis. Table 3 summarizes some transistor parameters for two different values of L_{cdr}.

L_{cdr}	1.5μ	1.0μ	unit
f_T	6.6	9.2	GHz
I_{fT}	14	63	μA
f_{max}	10.5	11.2	GHz
f_V	5.1	3.4	GHz
R_b	7.6	12.4	kΩ
C_{je}	354	563	aF
C_{js}	550	550	aF
R_e	295	360	Ω

Table 3 SOA Device parameters

From this table it is obvious that the f_T is low compared to other modern bipolar processes. As has been argued in previous sections, the f_T is not a very relevant parameter for predicting low power RF performance. The f_{max} in table 3 and the gain-bandwidth product of 24GHz at 10μA are more relevant and show the potential of transistors optimized specifically for low power RF.

4 SOA LNA design

When designing RF circuits in SOA using conventional circuit topologies, the specific properties of the active device and the availability of high-Q inductors result in different trade-offs than in more conventional IC processes. However, to get the best performance at the lowest power dissipation in SOA, unconventional circuit topologies will be required.

4.1 *Conventional SOA LNAs*

A conventional SOA LNA will typically have a passive low-loss transformer to transform the external impedance into higher on-chip values. By implementing this transformer as a tapped LC circuit, some band selectivity can be obtained at the same time for no additional cost.

At the external side of the LC circuit, the impedance that can be achieved is limited by the capacitive or inductive current that flows through the part of the LC tank parallel to the external input. Such a current adds an imaginary part to the transformed input impedance that would need additional compensation if it became significant. To keep this current small, the impedance of the capacitor and inductor have to be large compared to the input impedance. In practice, this is limited by the physical size of the inductor. From fig. 14 it is obvious that even for a 200Ω input, the inductors are already responsible for a significant part (about 25%) of the total die area (= 3mm x 3mm).

At the internal side of the LC circuit, the losses of the LC tank have to be negligible compared to the signal power. The quality factor Q has to be much larger than Z_0/Z_{in}, with Z_0 the characteristic impedance of the LC tank ($\sqrt{L}$ / $\sqrt{C}$) and Z_{in} the input impedance of the circuit connected to the output of the LC tank (i.e. the LNA input transistors).

The LC circuit is typically followed by a class A input stage designed mainly for low noise and high gain, similar to a traditional LNA input transistor, but at lower current levels. A common emitter stage provides the highest gain at low current levels. In practice, there will be a trade-off between the noise introduced by the limited Q of the inductors into the high input impedance of common-emitter input transistors, and the lower input impedance but also lower gain of a common base input stage.

4.2 *Examples and results of conventional SOA LNAs*

Combining all considerations from the previous sections into a single practical design has proven to be rather difficult. Experience with traditional RF IC design is only marginally helpful, since the acquired "feel" for typical impedance levels, currents, parasitics etc. points quite often in exactly the wrong direction. Many calculations had to be double checked and triple checked for this reason. Still, this is just a matter of training, and should not pose a significant long-term problem.

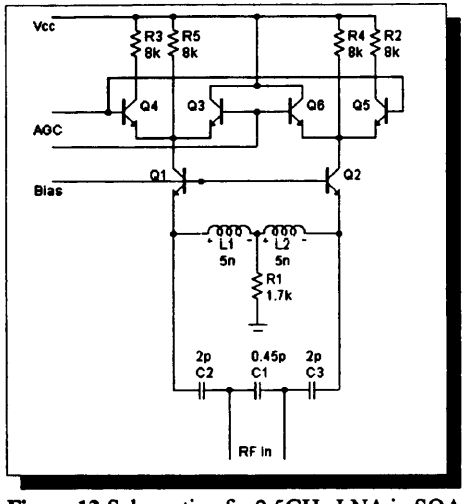


Figure 13 Schematic of a 2.5GHz LNA in SOA

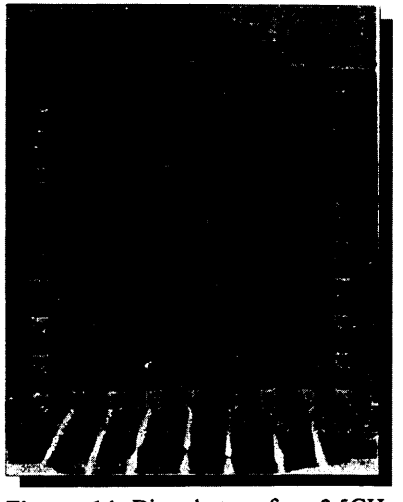


Figure 14 Die photo of a 2.5GHz frontend with LNA in SOA

Fig. 13 shows the schematic of a 2.5GHz LNA in SOA. It is part of a 2.5GHz WLAN frontend shown in fig. 14 [8]. The RF input of the LNA uses a combination of LC circuits and electronic buffers to achieve the required impedance transform. In this case, a limited impedance transform in the LC circuit was chosen to optimize the noise performance. The integrated LC type transformer steps up the external 200Ω impedance to 300Ω. This relatively low output impedance feeds into a differential common base stage (2x7 transistors running at 24μA each) with 300Ω input impedance. The output of this stage is connected to two differential pairs which provide variable gain operation. These feed into the RF inputs of the mixers. In table 4 the measured performance of this LNA is compared to a similar LNA in a traditional 1μm BiCMOS process [9].

The power dissipation has been reduced by one order of magnitude by using conventional high impedance design methods and SOA technology. The internal impedance levels are 8kΩ in SOA and 500Ω in BiCMOS.

	SOA	BiCMOS
Frequency	2.5GHz	1.9GHz
Gain	20dB	20dB
NF	4dB	3.5dB
IP3	-20dBm	-20dBm
Pdc	1mW	10mW

Table 4 Performance of SOA and BiCMOS LNAs

The result of the *IP3* measurement is shown in fig. 15.

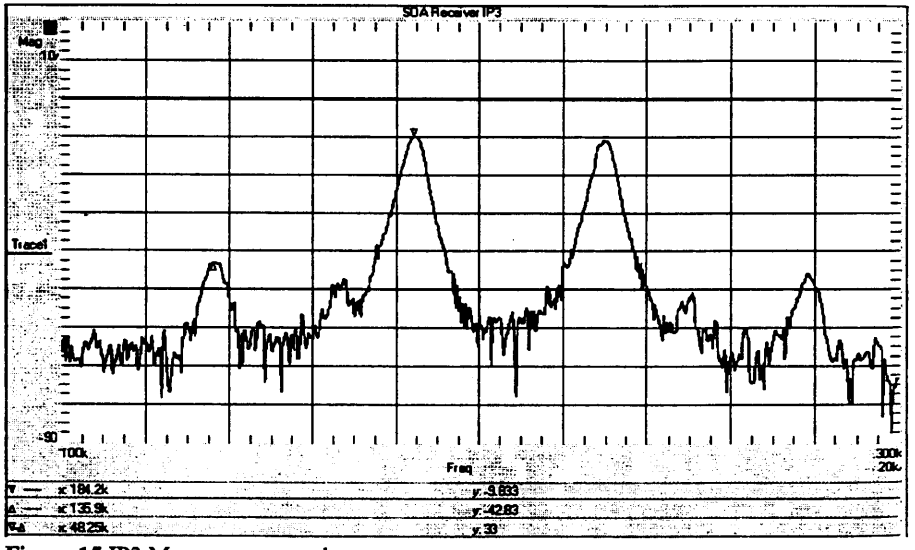


Figure 15 IP3 Measurement result

4.3 Advanced SOA LNA design

To push the performance of high dynamic range circuits even further than is possible with conventional high impedance design and SOA, we have to consider the results of the performance limits and trade-offs discussion earlier in this paper. These results suggest that significant further improvements are possible by using some or all of the following design techniques:

- common emitter input stages with high-impedance (tapped LC) loading together with the LC input transformer from the current LNA
- feedback across the emitter stages to improve the (small signal) *IP3*
- feedforward linearization techniques similar to those used at lower frequencies
- class AB biasing for the LNA to provide efficient output power and high 1dB compression points when needed
- adaptive output matching similar to [2] to improve on output impedance related power loss
- techniques to reduce the collector emitter voltage (V_{ce}) of NPN devices

5 Conclusions

Low power LNA design has been constrained by device and interconnect parasitics in traditional IC processes. SOA solves all important technology limitations for low power LNA design. Using mostly conventional LNA topologies in combination with high internal impedance levels and input transformers, significant improvements in LNA performance have been shown in this paper. Further improvements in low power LNA design using advanced design methods seem to be within reach.

References

[1] D. Pozar, "Microwave Engineering", Addison-Wesley Publishing Company, Reading, MA, U.S.A., 1990, pg. 99

[2] B. Nauta, "Analog Video Line Driver with Adaptive Impedance Matching", proc. ISSCC '98, San Francisco, U.S.A., February 5-7, 1998, pg. 318-319

[3] P. Baltus, "Influence of Process and Device Parameters on the Performace of Portable RF Communication Circuits", Proc. ESSDERC '94, Edinburgh, Scotland, September 11-15, 1994, pg. 3-11

[4] G. Hurkx, "The Relevance of fT and fmax for the Speed of a Bipolar CE Amplifier Stage", proc. Of the BCTM, Mineapolis, U.S.A., September 1996

[5] J. Burghartz, "RF Circuit Design Aspects of Spiral Inductors on Silicon", proc. ISSCC '98, San Francisco, U.S.A., February 5-7, 1998, pg. 246-247

[6] R. Dekker et al., "An Ultra Low-Power RF bipolar technology on glass", Techn. Digest IEDM, 1997

[7] R. Dekker et al., "An Ultra Low Power Lateral Bipolar Polysilicon Emitter Technology on SOI", Techn. Digest IEDM, 1993

[8] T. Wagemans et al., "A 3.5mW 2.5GHz Diversity Receiver and a 1.2mW 3.6GHz VCO in Silicon-on-Anything", digest. ISSCC '98, San Francisco, U.S.A., February 5-7, 1998, pg. 250-251

[9] P. Baltus et al., "DECT Zero IF Receiver Front End", proc. AACD '93, Leuven, Belgium, April 6-8 1993

RADIO TRANSCEIVER CIRCUITS IN SILICON GERMANIUM

Jan Sevenhans, Bart Verstraeten

Alcatel Bell, Antwerpen, Belgium,
e-mail : jsev@sh.bel.alcatel.be ,

ABSTRACT

In this paper some basic aspects of RF-bipolar and Silicium Germanium technology are described to indicate the determining factors that have allowed the succes of Silicon in the GHz range wireless applications. In addition the basic circuits are discussed for the low voltage design of a single chip cellular transceiver. Also preliminary measurement results of a realization in a 50 GHz Silicon Germanium technology are shown for the receiver and the transmitter section of the single chip transceiver first silicon with an integrated VCO. The circuits operate on a 2.7 to 3.6 Volt battery.

1. INTRODUCTION ON SILICON GERMANIUM

In the late eighties the first silicon bipolar RF technologies were maturing to provide about 10Ghz f_T npn transistors. Several achievements in silicon technology were necessary to make this possible. The poly emitter transistor with the As shallow emitter junction into the implanted base contacted by poly was the motor at that time for the introduction of silicon bipolar technology in the emerging cellular phone applications. 10GHz f_T was just enough to serve the basic transceiver functions for cellular applications below 1GHz carrier frequency. The good old Gilbert cell mixer in 5V RF bipolar technology was the key circuit to integrate the RF transceivers in heterodyne and homodyne radiotelephones.

Over the nineties silicon bipolar RF technologies moved further up to 20 ... 30GHz with the help of trench isolation, silicidized poly for emitter and base and the general progress in lithography towards sub 1 micron dimensions. But 20 to 30 GHz fT is about the limit for npn bipolar technologies today, even with selective implanted collectors, because finaly the vertical diffusion depth of the base is determining the base transit time.

So for further progress on the Silicon Bipolar RF performance the mobility of the carriers in the base was the blocking issue. Inspired by the heterobipolar approach in GaAs technologies with superior RF performance, germanium was introduced to realize a heterojunction in the silicon base. This is shown in figure 1.1 : a germanium profile is superimposed on the p-type base implant. Two approaches are generally applied: a graded germanium profile creating an electric field in the base region to accelerate the carriers on their way diffusing from the emitter to the collector junction and the second approach applies a constant germanium doping over the base layer to increase the mobillity of the minority carriers in the base and reducing the base resistance. Both approaches result in a higher f_T for a lower base resistance.

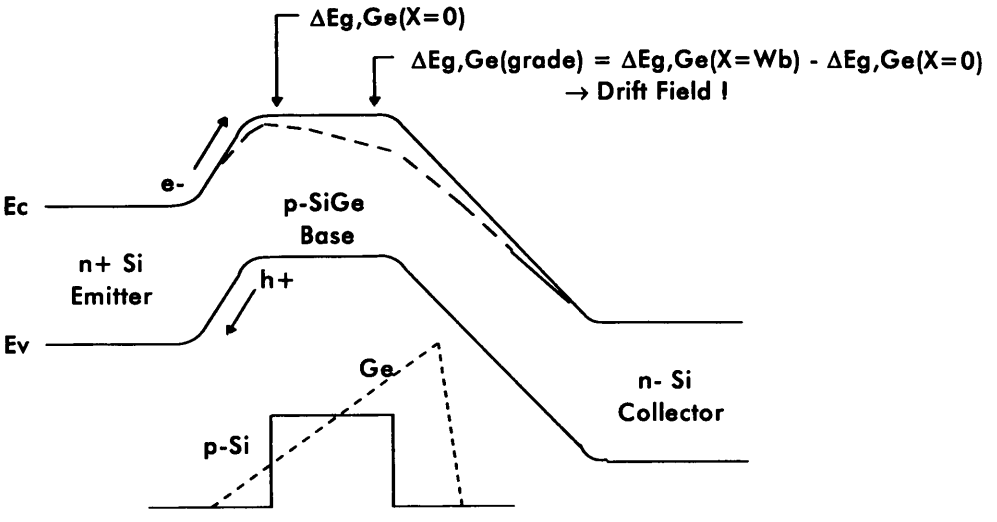


Fig. 1.1 : Electron potential in the heterojunction bipolar SiGe transistor

Some research groups tried to implant the Germanium in the silicon, others have done codeposition of silicon and Germanium with chemical vapour deposition or similar epitaxial techniques. Codeposition of the silicon and the germanium base layer is today the most succesfull technique in the foundries, although the implanted approach has very attractive advantages except for the lattice damage as a result of the big Germanium atoms implated in between the silicon : 10 to 20 Germanium atoms per 100 silicon atoms.

In the late nineties the silicium germanium npn transistor is going to production in several foundries. The growing market for RF wireless integration is convincing the silicon foundries to invest in high performance RF silicon technologies rather than GaAs, because of the compatibility of the silicon technologies with the main stream cmos fabrication equipment. UMTS, wireless LANs and other >2Ghz emerging applications will need the >50GHz f_T performance of Silicon Germanium [SiGe] and the low r_B for transmit noise and receiver sensitivity performance at low battery power. The <2GHz applications will exploite the SiGe performance to complete the full radio single chip integration to reduce the cost by eliminating discretes on the radio pcb.

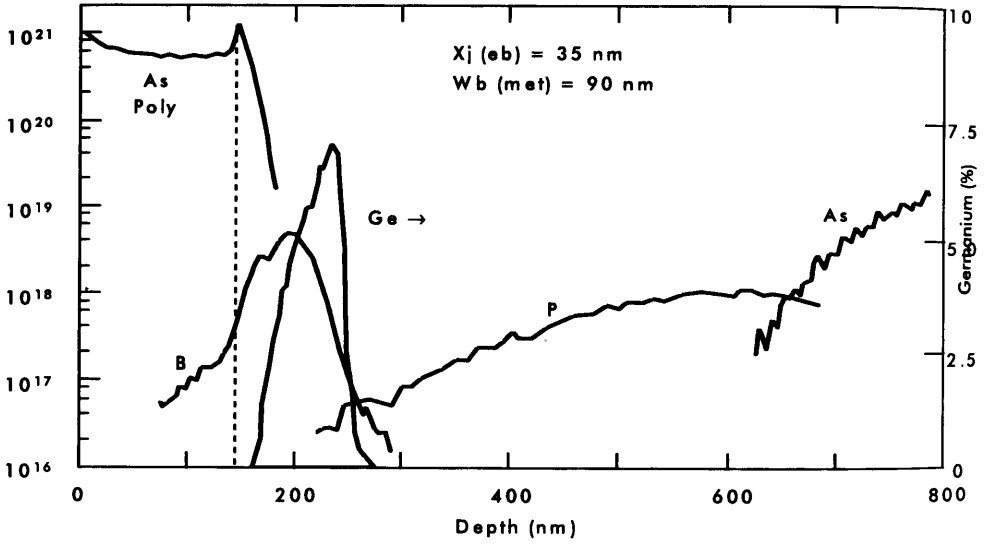


Fig. 1.2 : Doping profile of SiGe heterojunction bipolar transistor.

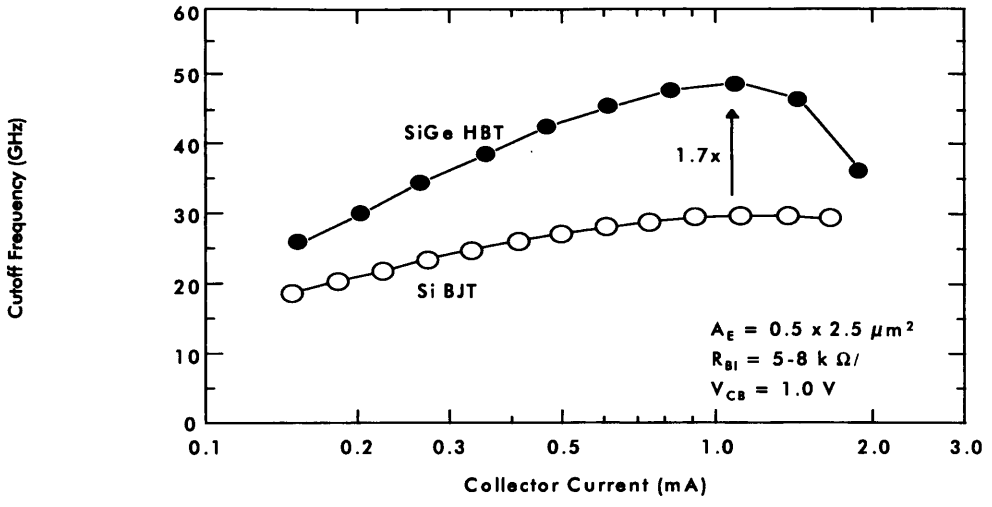


Fig. 1.3 : Transition frequency of a pure bipolar and a heterojunction SiGe transistor with comparable geometry.

2. RF TRANSCEIVER DESIGN

Since about 7 years, RF radio transceiver single chip integration for 5 Volt supply is a fact for cellular applications in production. To reduce the number of batteries and the cost and the size of a handportable cellular phone, new circuits for 3 V operation are today's issue for radio silicon bipolar integration in industrial research. Although preliminary research results were published on building blocks for 2 V operation, the full integration of a complete 2.7 Volt battery transceiver, including prescaler, LNA, receive and transmit mixers, baseband low noise receive post mixer amplifiers (PMA), RF pre–power–amplifier (PPA), and local oscillator VCO and quadrature phase shifter is a design challenge in today's commercial RF bipolar technologies. The 5 V Gilbert cell mixer is marginally applicable as the power supply is reduced to 2.7V, because it has 3 levels of transistors between the supply rails and 1 Volt collector–emitter voltage is low to achieve the peak f_T. In the circuits discussed in this paper, 2 transistors in series between the rails was

applied as design rule. To achieve this, the stages were capacitively coupled. Except for the consequence that low frequency test signals are blocked, the capacitive coupling is the easy solution for optimum bias level condition in each of the stages. DC–coupling the stages makes it difficult to match the bias level of each stage to the input of the next stage. To comply with cheap low frequency (LF) test, LF bypass cells across the capacitors are activated in the LF production test.

THE RECEIVER TOPOLOGY

The receiver is a balanced quadrature demodulator. In an input differential pair with emitter degeneration, the single ended LNA signal is converted to a balanced radio signal to drive the mixers over a capacitive coupling (See fig. 1, 2, 3). On the output collectors of the mixers external capacitors are connected to build a 300 kHz Low Pass RC first order filter suppressing the blocking levels at e.g. 3 MHz carrier distance. The +45° and the -45° Local Oscillator (LO) signals are derived from the LO input signal by an RC/CR on chip quadrature generator (QG). The equivalent input noise of the receiver is 16 dB.

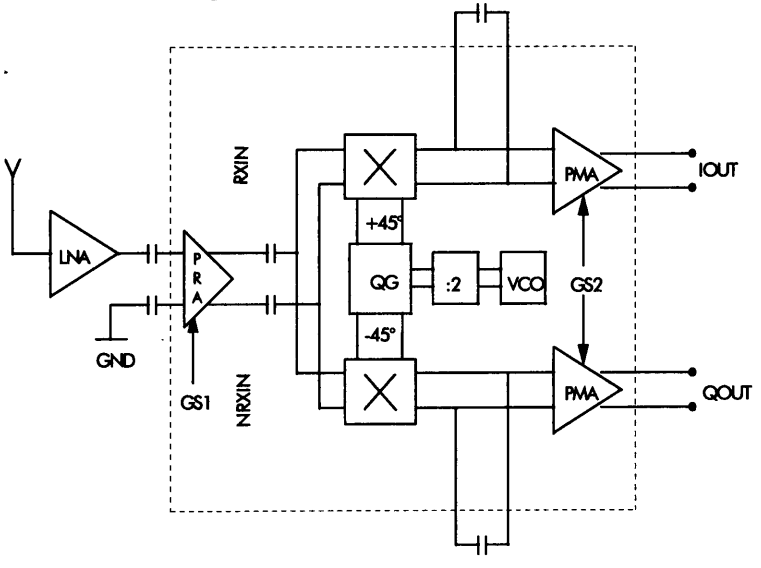


Fig. 2.1 : Receiver Architecture

THE RECEIVE PRE–AMPLIFIER (PRA)

The purpose of the PRA in the receiver is two fold : provide 12 dB gain switching and boost the input signal by 8.5 dB in maximum gain mode. The convenient circuit to use consists of 2 differential pairs connected in parallel to the 2 input pins RXIN and NRXIN. Driven from an external single ended LNA the radio signal hits the PRA on the RXIN input pin. The complementary pin NRXIN is capacitively coupled to ground. Activating the current sources low gain (L.G.) or high gain (H.G.) makes it possible to switch between the 2 gain modes and the emitter followers provide the low source impedance to drive the receive mixer input stage. The current (I) in the pre–amplifier is optimized to meet the linearity requirements for the blocking levels (-23 dBm antenna referred) and the interfering channels.

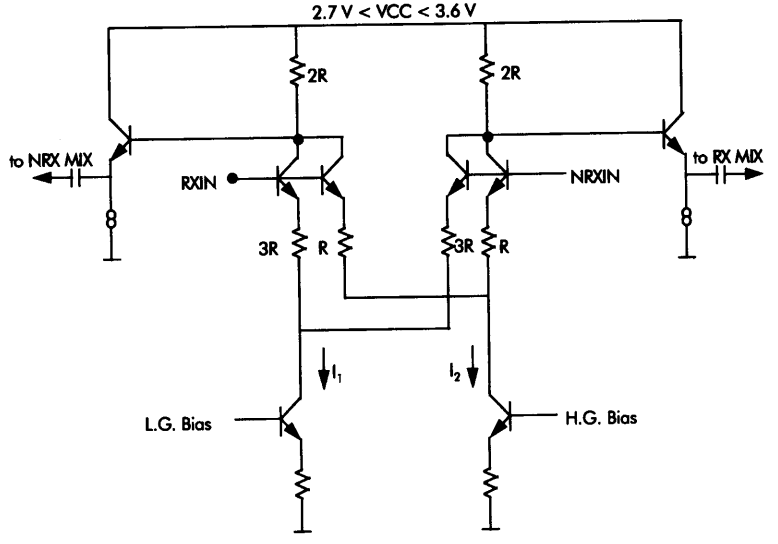


Fig. 2 .2: Receiver Pre–amplifier

THE RECEIVE MIXERS

The receive pre–amplifier (PRA) provides sufficient common mode rejection to have a balanced drive on the RX–mixer (Fig. 3) transistors T_1 and T_2, capacitively coupled to the PRA. The 4 switching transistors (T_3 . . . T_6) act as a balanced mixer on the current mode radio signal in I_1 and I_2. The input bases of the switching mixer are DC coupled through the emitter followers T_7 and T_8 to the quadrature generators DC output level. For good gain matching on the I and Q channel, the mixer transistors (T_3 . . . T_6) are hard switched by the +/-45° LO signal. The mixers conversion gain is 0 dB. An external capacitor (Cext) provides low pass filtering to suppress the blocking signals and the adjacent channels to avoid saturation in the bipolar circuitry. Further filtering is necessary in the base band CMOS AGC circuits and the final channel filtering is implemented in digital circuitry. Conclusive on the mixers, the common mode rejection of the PRA and the capacitive coupling of the balanced RX signal to the RXMIX and NRXMIX (Fig. 3) input make it possible to build a double balanced mixer using only 2 levels of NPN transistors between the supply rails.

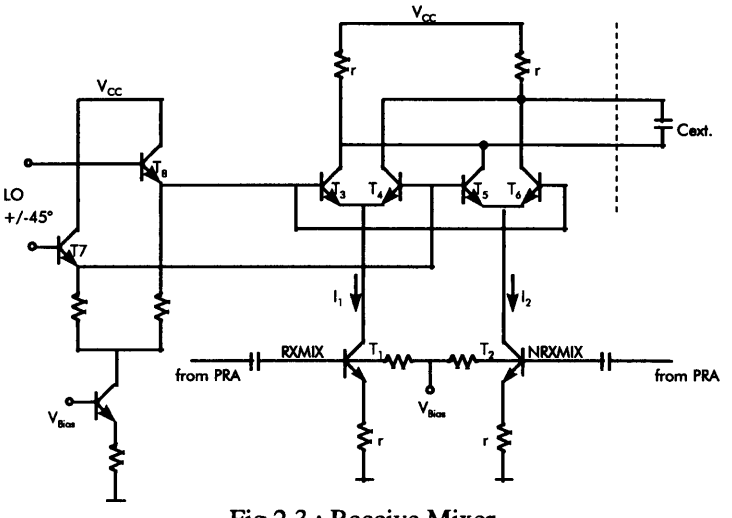


Fig.2.3 : Receive Mixer

THE RECEIVE POST MIXER AMPLIFIER (PMA)

To cope with the noise level of the CMOS analog base band signal processor, 16.5 dB gain in the post mixer amplifier is necessary to overcome the 40 dB noise figure of the CMOS circuitry. This brings the total maximum gain of the integrated receiver at 25 dB. The low Rb in the SiGe transistors allows for a low noise preamp in the RF-front-end. A double asymmetrically degenerated differential pair (ADDP) maximizes the linear range of the PMA. Each of the ADDP's provides a maximum flat gain curve, one for the positive excursions of the RX–signal, and one for the negative excursions. For a smooth cross over, cross coupling resistors between the ADDP's are included.

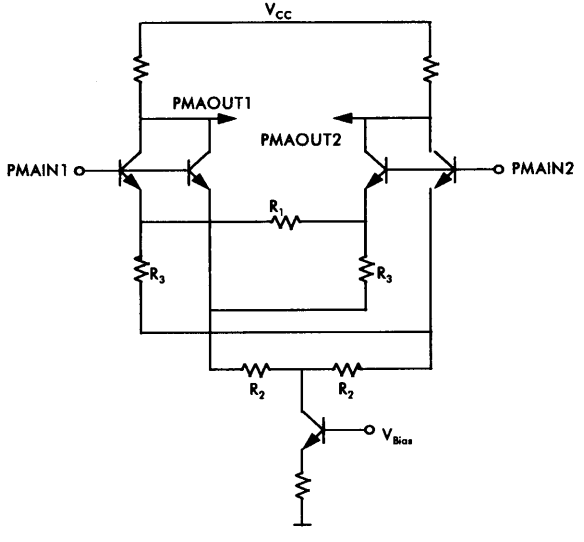


Fig. 2.4 : Receiver Post Mixer Amplifier

THE TRANSMITTER TOPOLOGY

The transmitter is basically a zero IF single side band mixer followed by a pre–power amplifier (PPA).. To use the available 2.7 V supply range efficiently, the base band I and Q balanced inputs were realized in current mode. So the only voltage swing on the mixers inputs are the local oscillator (LO) signals delivered by the quadrature generator (QG) with +45°/-45° phase relation. The modulated collector currents of the I and Q mixer enter a resistive summing node to feed a voltage mode RF signal to the on chip pre–power amplifier (PPA). The PPA delivers 0 dBm output power over a 1:2 transformer into the 50 ohm input of an external Power Amplifier that boosts the signal to the antenna output level.

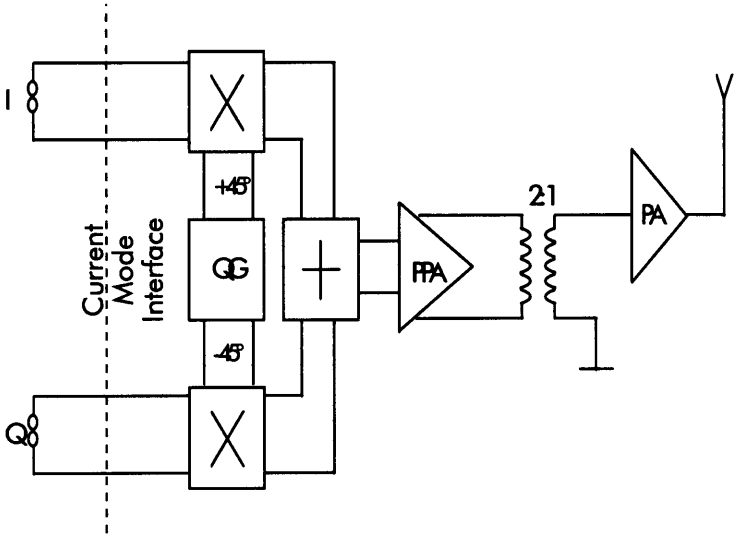


Fig. 2.5 : Transmitter Topology

THE TRANSMIT MIXER

As shown in Fig. 5 the transmitter uses 2 amplitude modulator (AM) cells driven by quadrature LO signals to obtain SSB modulation. The AM mixer cells basically consist of 2 differential pairs and 2 emitter followers to buffer the LO inputs to the mixer. The base band inputs I, NI and Q, NQ are driven by a floating current source in the CMOS base band component. The base band I and Q channel current mode signal is driven to the emitter nodes of the switching differential pairs. To separate the base band circuitry from the RF switching circuits a reverse RF filter is included between the base band I/Q input and the emitters of the switching transistors. The RF filter consists of an external capacitor (Cext.) and an internal resistor R.

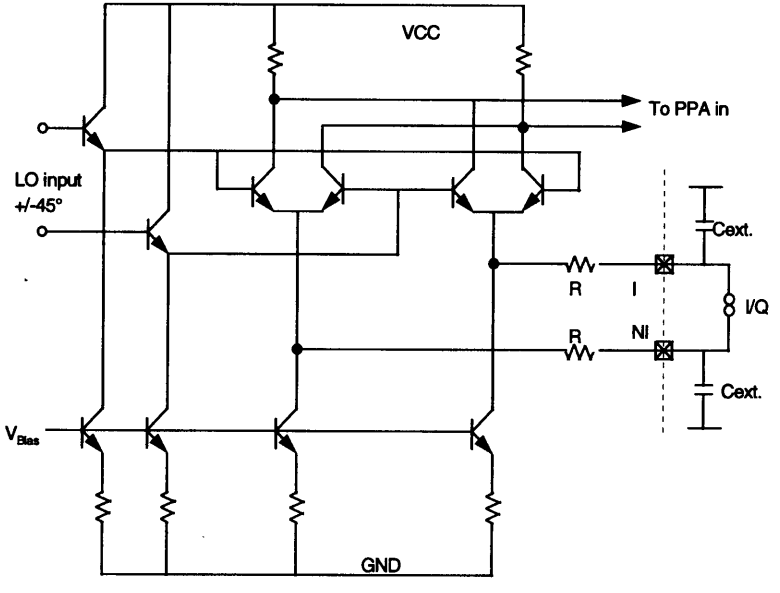


Fig. 2.6 : Transmit RF Mixer

THE TRANSMIT PRE–POWER AMPLIFIER

Except for the additional test circuit T_1, T_2, the pre–power amplifier consists of a basic differential pair and 2 emitter followers. The pre–power amplifier operates in full class A.

To drive 0 dBm into the 50 ohm load through the 2:1 transformer, the swing on the open collectors is over $1.2V_{PP}$ The followers at the PPA-input buffer the mixers. The result is: 3 transistors between the 2.7 V supply rails. But, with the base inputs of the followers biased close to V_{CC} , the open-collectors are biased at a V_{BE} reverse collector to base and the tail current source remains at V_{CC} - 2 V_{BE}. As the test circuit is activated, T_1 and T_2 in Fig. 7 open a low frequency (LF) test path for the PPA and the TX mixers.

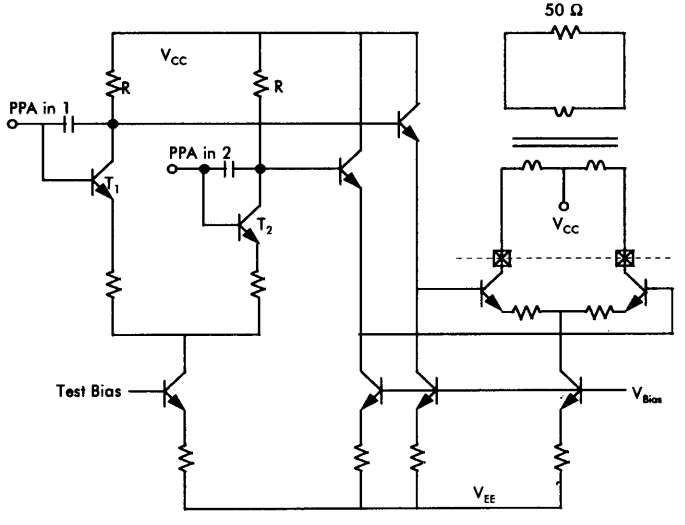


Fig. 2.7 : Pre–power Amplifier

THE QUADRATURE GENERATOR

The quadrature generator is expected to deliver both LO signals with a 90° phase difference accurate to within 1° phase error. This is possible using just an RC and a CR filter to obtain the +45° and -45° phase relation. The phase accuracy is well covered but technology variations are the cause of an important variation of gain. To obtain the necessary gain matching of the I and the Q channel, limiting amplifiers, realized as double differential pairs to obtain a "horizontal cascode" operation, are hard switched by the RC and the CR's output voltages (Fig. 8).

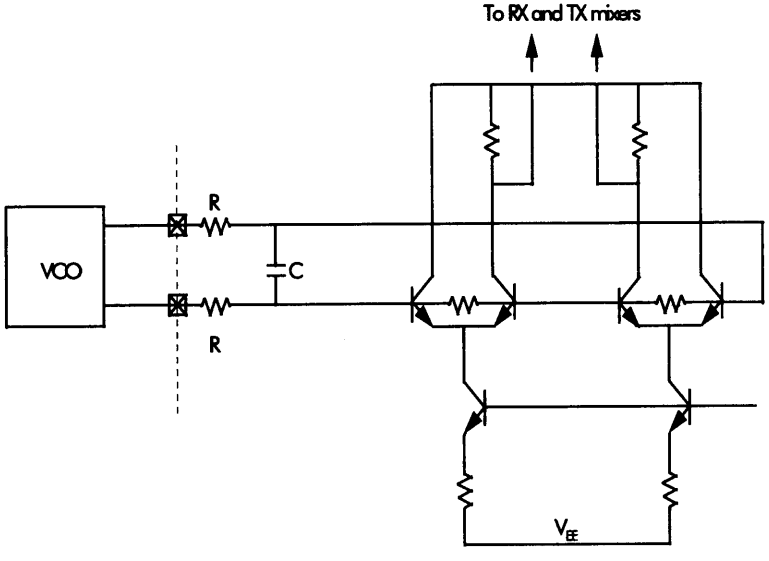


Fig. 2.8 : RC +45° Phase Shifter

THE VCO

Looking at the VCO schematic in figure 9 the transistors T_1 and T_2 are the active elements. The emitters of T_1 and T_2 are AC coupled and the bias current is provided by the 2 current sources T_3 and T_4. The resonator of the VCO consists of the inductors L_1 , L_2 and C_1, C_2 and the base emitter Junctions of transistors T_5 and T_6 act as the varicap, with the control voltage (VCV) on the base. Figure 12 shows the measured phase noise. Switching in C_3 and C_4 makes the VCO jump from the RX to the TX band. The phase noise at 10kHz from the carrier is -80 dBc/Hz and -100dBc/Hz at 100kHz from the carrier, see fig. 12.

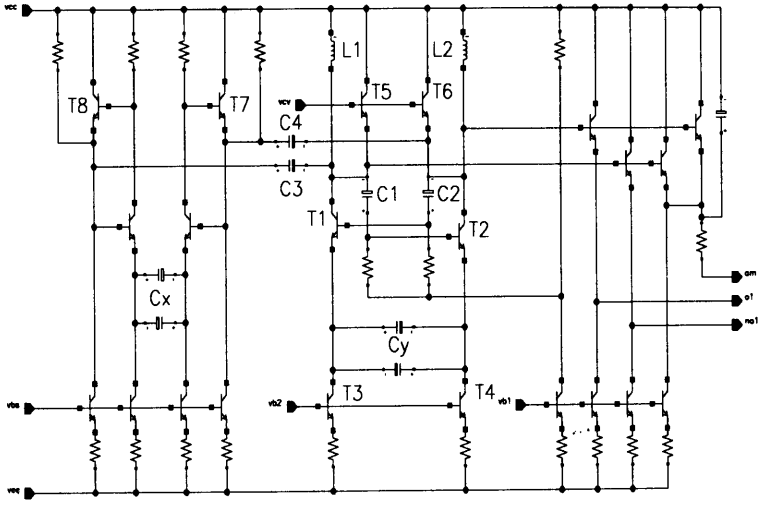


Fig. 2.9 : VCO schematic.

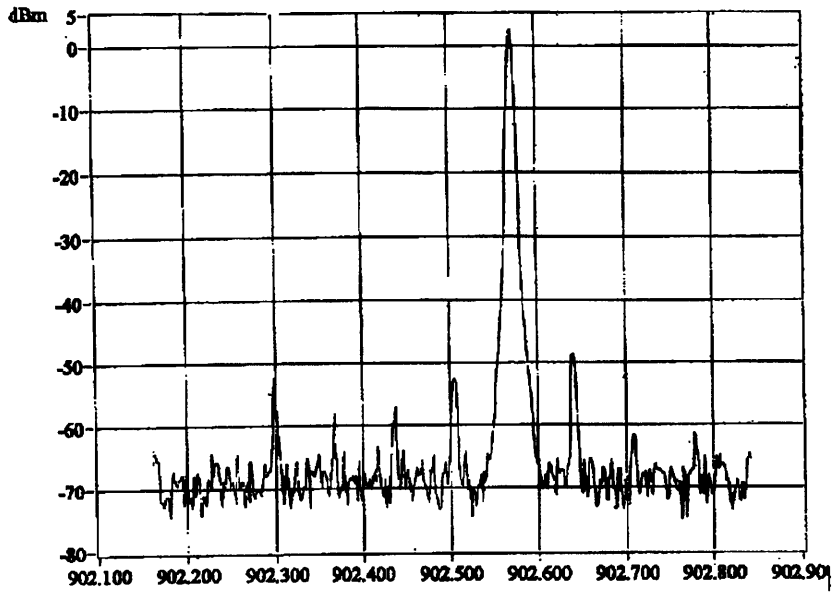


Fig. 2.10 : Measured Transmitter Output Spectrum

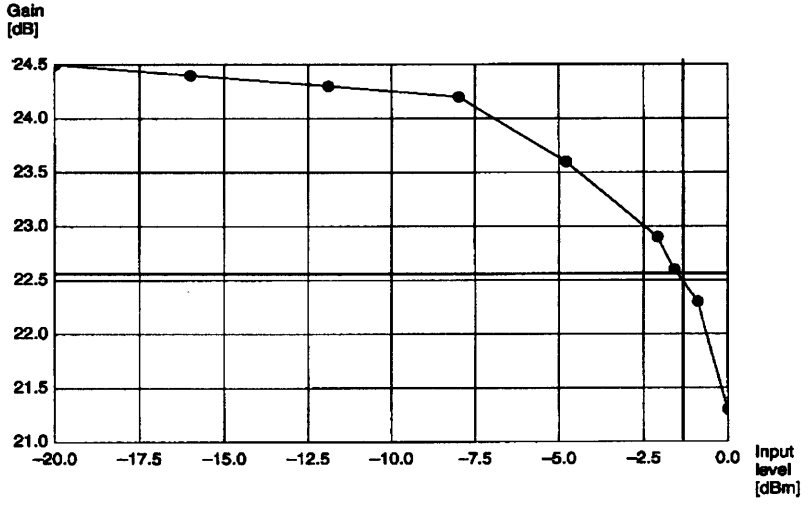


Fig. 2.11 : Measured receiver gain compression in maximum gain mode for a blocking level at 3 MHz from the carrier. The performance copes easily with the required system specifications

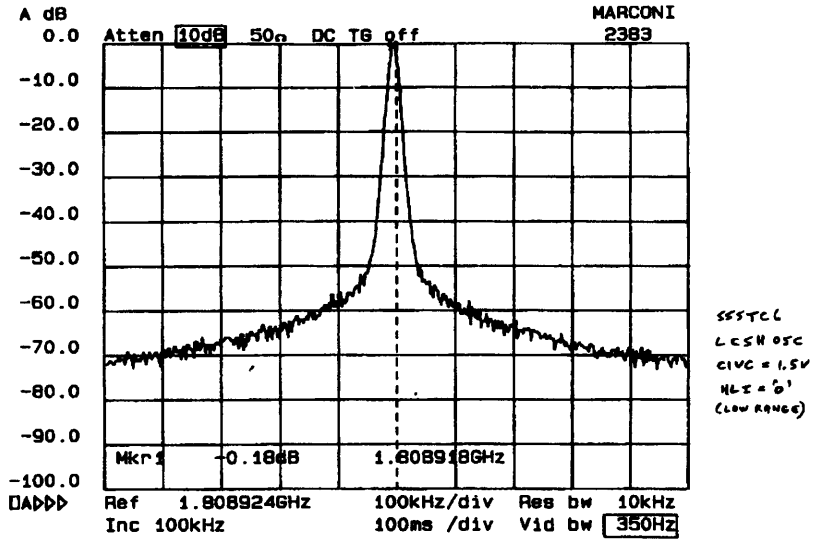


Fig. 2.12 : VCO Phase noise spectrum

CONCLUSION

A 2.7 V radio transceiver for 900/1900 MHz was realized in a 50 GHz f_T silicon germanium technology. The single chip transceiver contains the RX mixers, TX mixers, quadrature generator, amplifiers, quadrature phase shifters, the VCO and the 64/65 prescaler. Most of the circuits were realized in a 20 GHz bipolar technology before. In the 50GHz SiGe technology the noise figure was improved with 2dB and the additional bandwidth allows for true dual mode operation with identical gain curves for both 900 and 1900Mhz operation. The transceiver drives an external power amplifier and needs at least 15 dB external gain of a Low Noise Amplifier in front of RX mixers. An integrated LNA is optional and feasible in the 50GHz SiGe technology with low R_b enabling a 2 dB noise figure.

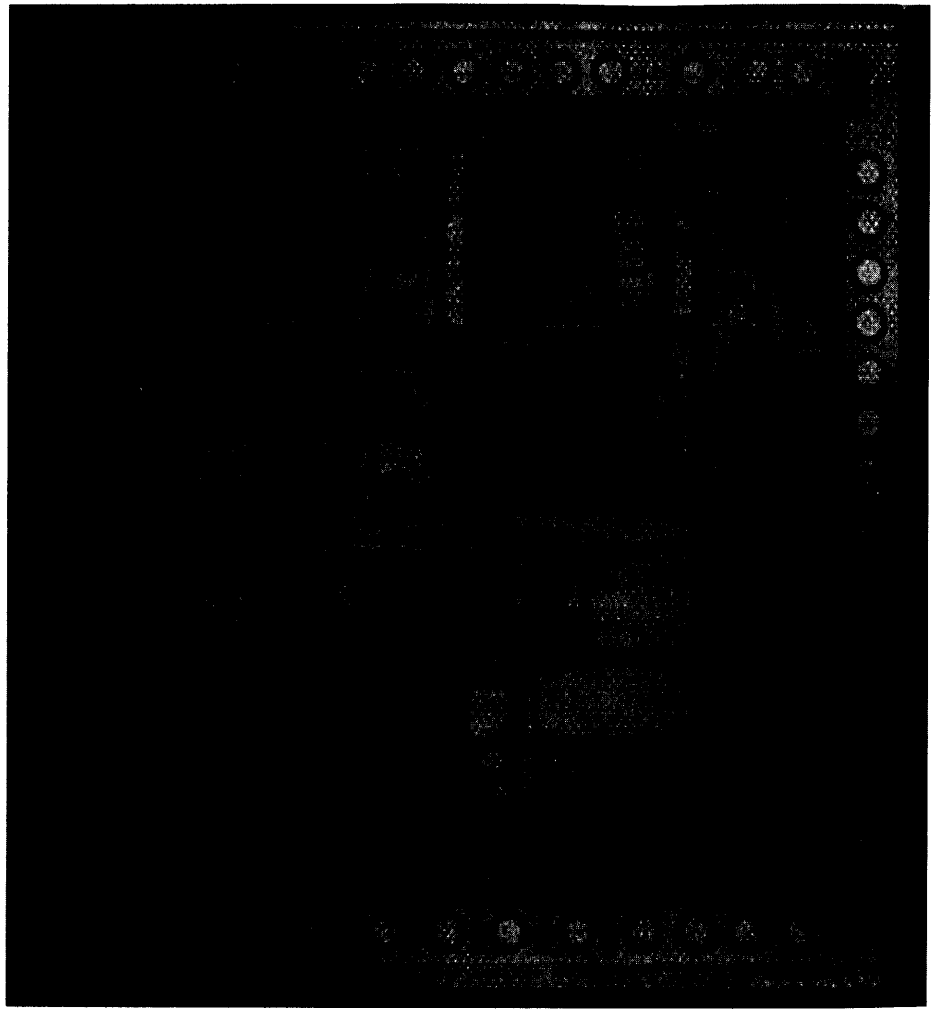


Fig. 2.13 : Photomicrograph of the tranceiver

ACKNOWLEDGEMENT

The authors thank Paul Vanleene, Eric Duvivier, Gianni Puccio, Corine Berland, Celine Dars, Eric Dussauby, Pierre Genest, for their contribution to the project.

REFERENCES

(1)M. M^{c}Donald, " A 2.5 GHz Bicmos Image -Reject Front End," ISSCC Digest of technical papers, pp. 144-145, Feb., 1993.

(2)J.Crols, M.Steyaert , "A Fully integrated 900MHz CMOS Double Quadrature Downconverter," ISSCC Digest of technical papers, pp. 136-137, Feb., 1995.

(3)C.D.Hull, R.R.Chu, J.L.Tham, "A Direct conversion receiver for 900MHz spread spectrum Digital cordless telephone," ISSCC Digest of technical papers, pp. 344-345, Feb., 1996.

(4)J. Craninck, M. Steyaert, "A 1.8 GHz Low-Phase-Noise Spiral-LC CMOS VCO," Symposium on VLSI circuits Digest of technical papers, pp. 30-31, June 1996.

(5) Tsuneo Tsukahara, Masayuki Ishikawa, Masahiro Muraguchi, "A 2V GHz Si-bipolar Direct Conversion Modulator," ISSCC Digest of technical papers, pp. 40-41, Feb., 1994.

(6) Asad Abidi, "Direct Conversion Radio Transceivers for Digital Communications,"ISSCC Digest of technical papers, pp. 186-187,363-364, Feb., 1995.

(7) Jan Craninckx, M. Steyaert, "A 1.8Ghz Low-Phase Noise CMOS VCO Using Optimized Hollow Spiral Inductors," IEEE J. Solid state Circuits, May 1997, pp. 736-744.

(8) Jurgen Arndt, TEMIC, Heilbronn, For sharing some of his experience and material on Silicon Germanium technology.

(9) John D. Cressler, "Silicon Germanium Heterojunction Bipolar Technology: The Next Leap in Silicon?", ISSCC Digest of technical papers, pp. 24-27, Feb. 1994.

Modeling for Si-Bipolar Power Amplifiers

Dr. S. Weber
Siemens AG Dept. HL HF PE PA
Munich, Germany
stephan.weber@hl.siemens.de

Abstract

Designing RF PA's you will propably find a too large difference between simulation and measurements. Therefore accurate modeling of all used parts is needed to make simulation becoming an accurate design tool. The most important parasitic influences are identified and their modeling is shown.

1. INTRODUCTION

The market of RF power amplifiers for mobile communication is presently dominated by GaAs MESFET technology. At the low frequency end (e.g. GSM f_0=900MHz) or at low power levels (e.g. DECT P_{Ant}=250mW) Si BJT or MOSFET technology is becoming more important, due to their lower die costs and the possibility for a higher level of integration. For GaAs there are also further drawbacks e.g. the need for a supply switch and the negative gate voltage. Because there is still a drawback in Silicon compared to GaAs in respect to high frequency performance, especially at high powers (e.g. f_T, f_{max}, breakdown voltages, no chip vias in Si, large substrat losses, no thick Au metalization, etc.) you must get maximum performance from your Si technology.

This performance is not only limited to the intrinsic devices, e.g. described by $f_T(I_C)$ etc., but also by parasitic elements, e.g. in the substrate, die interconnections, package, thermal performance and external components (especially in the matching networks). An accurate modeling for the active desvices and all the parasitics is needed (figure 1).

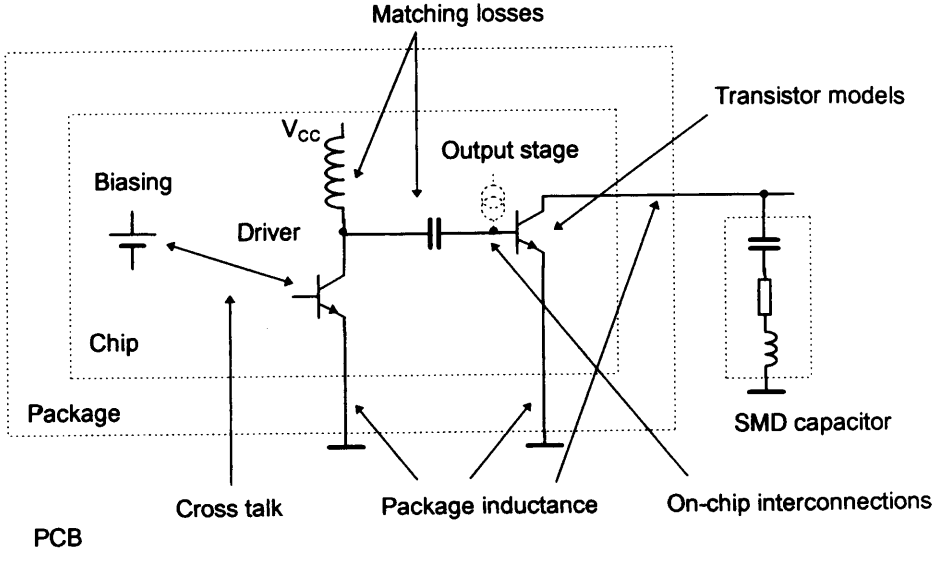


Figure 1 : Some of the main modeling problems in RF PAs

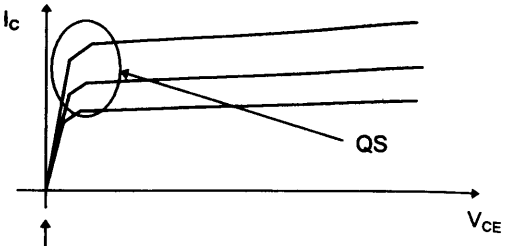
2. MODELING THE BJT

In most standard circuit simulators like SPICE, SPECTRE, SABER, etc. only the Gummel-Poon model is included, maybe with some extensions. In the case of RF PA's first you need good parameters for this basic model. Often even the GP parameters are not fully understand by the users and some of them are only guessed or even at their default values. In real world PA's also the reverse parameters (e.g. the transit time TR) and the series resistances becomes very important. There are also some effects not included in the GP model at all (figure 2). To model the quasi saturation region or the accurate dependance of f_T vs bias conditions and temparature you need extended models. Using high f_T-technologies you may get problems with breakdown voltages [1]. For a typical 20GHz-f_T-transistor the V_{CEO} is only about 4.5V. The V_{CER} or V_{CBO} is much higher (about 15..20V) but in reality the maximum usable collector-emitter voltage is between this limits. As a rule of thumb the supply voltage V_{CC} should not ecxeed V_{CEO}, the RF swing gives then higher collector voltages up to 2 .. 3 times V_{CC} (due to inductice load and the influence of the harmonics) [2]. This margin is needed to get safe operation even in the presence of large VSWR values and production tolerances. Some further margin may be achieved by a low impedance drive of the RF transistors because V_{CER} depends strongly on I_C and the source resistance R itself. Transistor characterization is most important for the output devices. If you cannot extract data directly from the large output structures you need to extrapolate from smaller structures. This should not be done by simply using the area factor because e.g. the ratio between edge and area capacitances changes with the transistor size. If the amplifier works in saturation mode (in most systems like AMPS, GSM, DECT) accurate transistor modeling is important not so much for the simulation of the output power and efficiency but more important to design the driver stage and the interstage matching.

Figure 2 : a) Some draw backs in standard BJT models

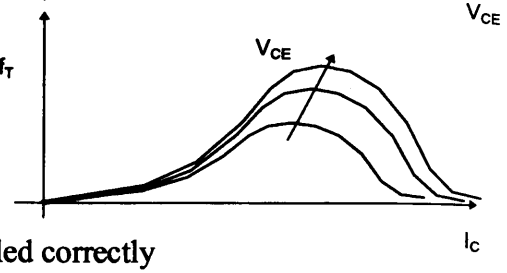
Quasi saturation

New parameters needed
Known models are quite unaccurate
Dynamic parameters also changed



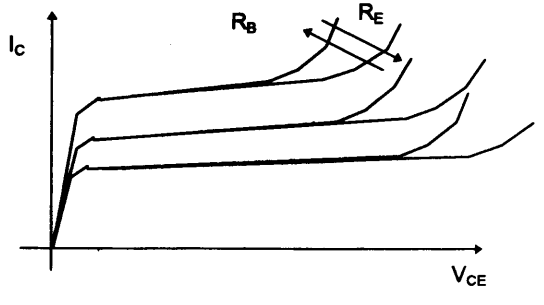
f_T-curve

SPICE model with TF, XTF, ITF, VTF is not sufficient
Temperature dependance and low voltages/high currents are not modeled correctly



Breakdown

You have to look at V_{CEO}, V_{CBO} and most important V_{CER} which depends on current and resistors



b) Influence of transistor performance on an 1.9GHz-PA

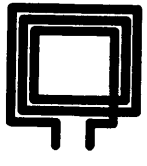
Technology	f_{Tmax}	V_{CEO}	V_{CER}	V_{CCmax}	P_{out}	PAE
#1	≈26GHz	4.0V	16.3V	4.1V	28dBm	39%
#2	≈16GHz	6.9V	18.6V	5.4V	27dBm	36%
#3	≈19GHz	4.1V	20.9V	6.1V	27dBm	35%
#4	≈10GHz	8.7V	21.0V	>7V	25dBm	28%

3. ON-CHIP INDUCTORS AND DIE INTERCONNECTIONS

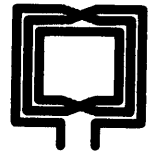
In modern RF technologies with an f_T of 20GHz or more RF amps often looks like AF amplifiers. But at high power levels (e.g. >10mW depending on frequency) you get much better performance in respect to noise, stability, spurious rejection and efficiency using on-chip inductors [3,4] (figure 3) instead of RC coupling. In a standard Si technology with 3 Al layers and a substrate resistance of 10Ωcm you get coils with maximum Q factors of about 5 to 8 (typically at 1 to 3GHz) with inductances of 1 to 8nH and areas of about 100x100µm² to 400x400µm². 2½ or 3-D field simulators, test chips or special programs are useful to understand on-chip coils. A π-circuit model with L_s, R_s, C_{ox} and R_{sub} is almost sufficient for circuit analysis up to 2GHz.

Figure 3 : a) Typical configurations of on-chip inductors

Simple Spiral



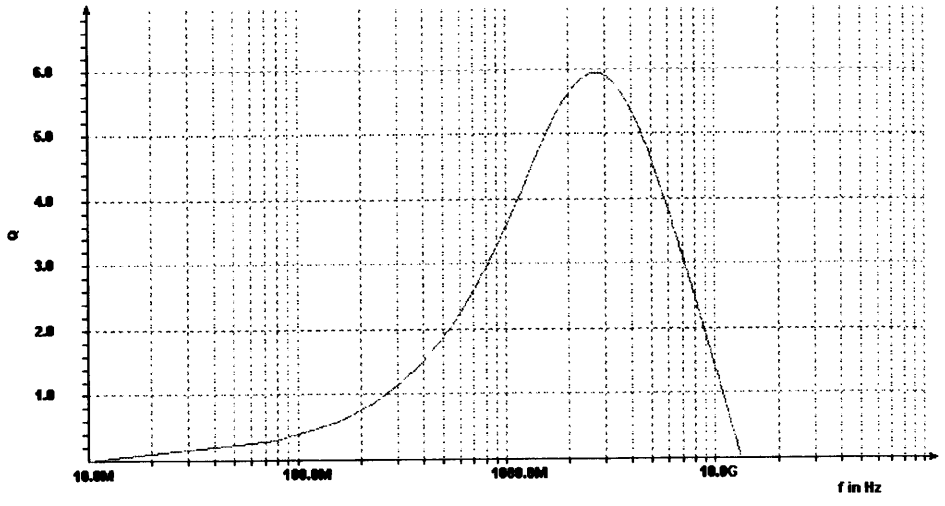
Symmetric Coil



Variations : Tapped coils, center tapped coils, coupled coils

Hints : Use large coils for low frequencies to minimize series resistance and use smaller ones at high frequencies to minimize substrate loss. Do not use the center of the coil, because inner turns have small inductance but some series resistance. Use upper metal layers, not metal 1.

Figure 3 : b) Typical graph Q vs f (L is 4nH) :

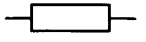


In some CAD systems you may get an extraction of the parasitics in the interconnections [5]. But many of these tools are written for digital or low power circuits, so they are looking only at the series resistances and the parasitic capacitances to the substrate but magnetical coupling is not included (figure 4 to 7). In the case of RF power amplifiers the impedances becoming quite low (few Ohms at the input of the power transistors), therefore the series inductances is in many cases much more important than the shunt capacitances. In Si technology with a substrate height of about 300µm you get about 0.5 to 1.4nH/mm inductance per length (depending on width and distance to ground planes). Using large structures like on-chip inductors, MOS capacitors with some pF or output power transistors the length of interconnections is very different from zero, e.g. a line of 300µm (e.g. from the driver output to the base input of the power stage) may gives an impedance of about j3Ω at 2GHz (figure 8). This is often much more than the ohmic part and in the range of the input impedance at the base of a power transistor.

Accurate modeling of the on-chip coils and the interconnections is very important for the design of the interstage matching and has a great influence on the frequency response of the PA.

Figure 4 : Parasitics in on-chip interconnections

$R = R' \cdot l = R_{sheet} \cdot l/w$

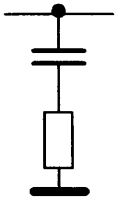


=> Voltage drop, damping, noise

$L = L' \cdot l$

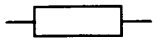
=> Phase shift and impedance transformation

$C = C' \cdot l = \varepsilon_0 \cdot \varepsilon_r \cdot l \cdot w/d_{ox}$ + edge capacitances



=> Damping, cross talk and substrate coupling

Figure 5 : Ohmic series losses

$R = R' \cdot l = R_{sheet} \cdot l/w$ 

Example : l=0.5mm, w=10µm 50Ω-System 5Ω-System

Al1	50mΩ	=> 2.5Ω	=> -0.215dB	-1.94dB
Al2/3	30mΩ	=> 1.5Ω	=> -0.13dB	-1.21dB

(5Ω is approximately the typical input impedance at the base of an output power transistor)

Skin effect : δ = 1.8µm at 1.9GHz => only small influence!

=> Interconnetion is critical

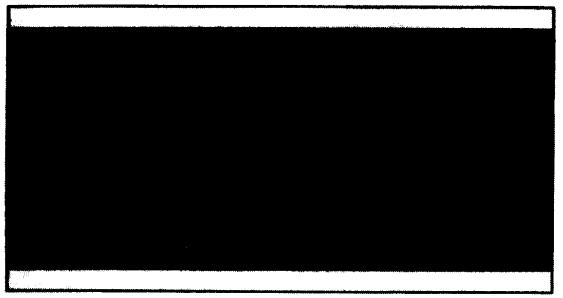
=> Series resistances of integrated MOS capacitors are critical

Figure 6 : Transmission lines

$Z_L^2 = R' + j\omega L' / (G' + j\omega C') \approx L'/C'$ (for low loss lines, but not on Si!)

$v = c_o/\sqrt{\varepsilon_{r(eff)}} \approx 1/\sqrt{L'C'}$

$L = L' \cdot l$ 



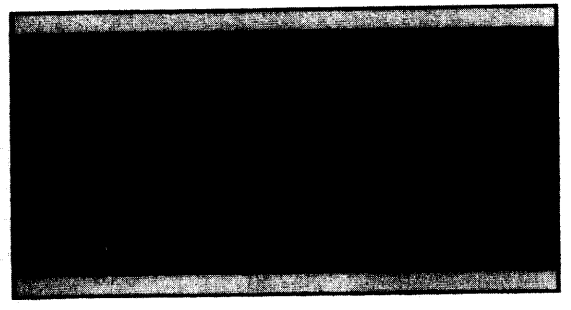
H-field comes deep into the substrate

=> effective distance $\approx d_{sub}$ (typical 300µm) => L' is relative large!

w	L' (calculated as microstrip line)
2µm	1.4nH/mm
5µm	1.23nH/mm
10µm	1.09nH/mm (about 7Ω for 500µm at 2GHz)
25µm	0.9nH/mm

L's is smaller for odd mode , but not so much

$C = C' \cdot l = \varepsilon_o \cdot \varepsilon_r \cdot l \cdot w/d_{ox}$ + edge capacitances



E-field stops at the Si surface (because $\omega\varepsilon E << \sigma E$)

=> effective distance $\approx d_{ox}$ (typical 2µm) => C′ high + series resistance

w	C′ (Alu2, calculated as microstrip line)
2µm	≈65fF/mm
5µm	≈110fF/mm
10µm	≈190fF/mm (about 800Ω for 500µm)
25µm	≈440fF/mm

In low impedance circuits (as PAs) the series losses are much more critical than the capacitances to substrate.

=> Layout should optimized not for minimum C_{sub}, but minimum L_s+R_s

Figure 7 : Transmission line parameters

w	Z_L (at low frequencies)		
2µm	≈146Ω	=>	Z_L has typical values, but it is now
25µm	≈50Ω		complex and frequency dependant

w	v_{ph}	ε_{reff} (at low frequencies)		
25µm	≈50Mm/s	35	=>	ε_{reff} is very high
2µm	≈100Mm/s	8	=>	much phase shift & transformation

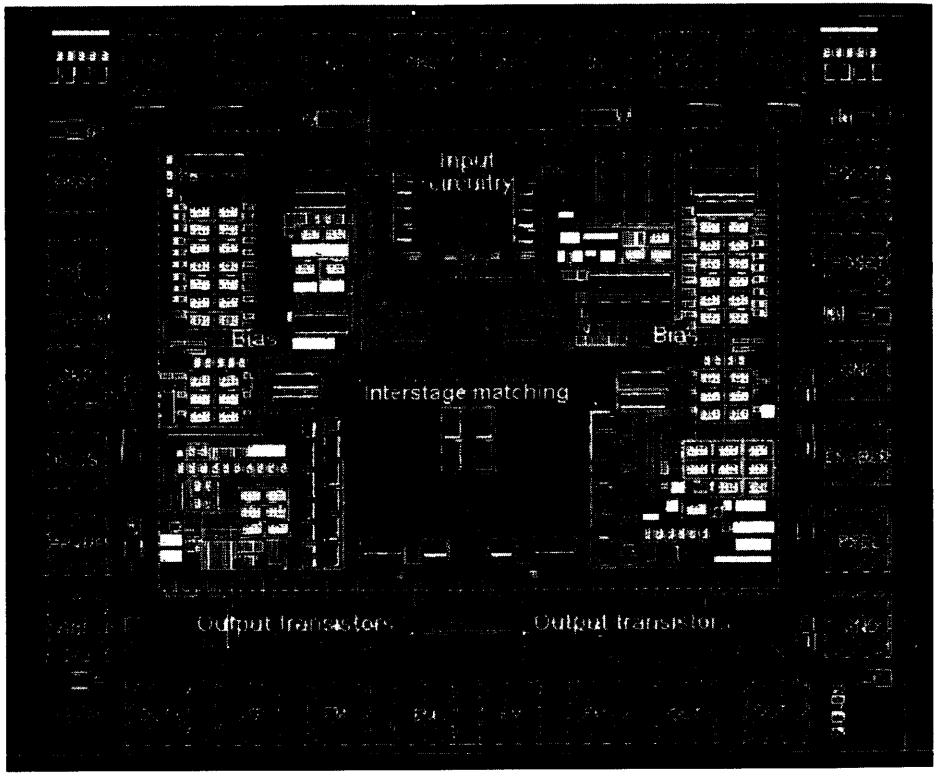
=> **„Slow-Wave-Mode“**

What is more important R_s or L_s?

w	$\omega L/R = Q_s$ at 2GHz
25µm	9.4!!
2µm	2

=> **Series inductance is more critical (due to impedance, not noise) than the ohmic part**

Figure 8 : Layout of a balanced 1.9GHz-DECT-PA



=> There are many parasitics

4. SUBSTRATE AND PACKAGE MODELING

The package has a strong influence of the RF performance. There is much literature available and accurate modeling is one important key for a good RF PA design. This sounds simple but practice has shown that the package influence is in many cases more critical then the simulation shows. Therefore there is much to do and if you get a complete model you have to asks if it is this really correct. What are the limitations, e.g. in frequencies? Is the model based on physics? Is there a comparision between simulation and modeling? The only way seems to be to make a model which bases both on 3D-field simulations and measurements. For PA´s this model has to include also the losses due to ohmic resistances, skin effect and dielectric. The package influence is of cause most important for the output and the ground pins. Good grounding conditions are needed for high RF gain, good stability and isolation. Error voltages on ground plane or in the substrate [6] often have a negative influence on the bias circuit. Using not to low bias currents and guard rings for the isolation between RF and bias circuits is important to minimize these effects. The subtrate losses also cause a drop in RF gain which has to be modeled. Our simulations also shows, that using a single subtrate node (like in small signal IC´s) is far from reality and not sufficient.

5. MATCHING NETWORKS AND EXTERNAL COMPONENTS

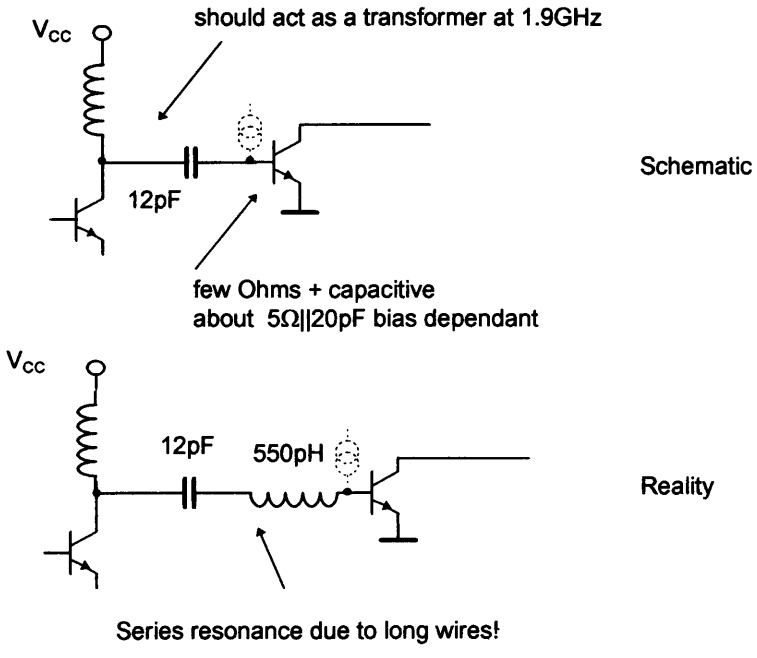
For high output power levels you need an external matching network because you get too much losses with on-chip elements which reduces the efficiency. It also appears that it is not possible to design the linear impedance transformation networks only by linear simulation due to influence of the harmonics (especially for high power amplifiers). Therefore it is better to use a time domain/harmonic balance simulator (e.g. APLAC) to optimize the matching elements in conjunction with the nonlinear output stage. As a starting point it is of cause suitable to use e.g. Smith diagram technics or a linear circuit simulator with an optimizer (like Touchstone, SuperCompact or APLAC).

Designing a balanced amplifier you get a better performance in respect to substrate coupling, ground noise, harmonics etc. But an output balun is not easy to design and has maybe higher losses than a simple matching network. Also the influence of the ground inductance is only effectively reduced in near class A operation. Decoupling is also important for PA's. The supply impedance should be low not only at the frequency of operation f_0 but also from DC to about $3 \cdot f_0$.

Figure 9 : Parasitics in a typical RF PA

Part	**Influence**	**Comments**
Transistor models	May have a large influence, especially on interstage matching !	GP may be sufficient, but not in all cases. High current/low voltage region is critical!
Capacitances to substrate	Often a low influence (not for transistor or MOS-C capacitances)	This is different to low power/high impedance designs.
Series resistors	Medium influence. Look also at the MOS capacitances	Reduces gain
Series inductances	Large influence ! Not only as feedback in BJT emitters stages	Changes frequency response
On-chip coils	Medium influence. A Q of 5..7 is realistic. You have to include the lines to the coil	Modeling is not too difficult, but Q is limited for typical Si technologies
Package model	Strong influence due to series inductances	Difficult to model
Substrate model	Medium influence on bias and RF performance	Difficult to model, important for mixed mode designs

Figure 10 : Influence of interconnection parasitics for interstage matching



=> Collector of the driver directly connected to the low impedance base

=> Only small voltage swing at the collector

=> Driver current must be very high, due to mismatch!

=> Resonance frequency is too low, low efficiency

=> Even the circuit topology may be changed due to parasitic elements not only the parameter values

6. CONCLUSIONS

Accurate simulations are possible for RF power amplifiers, but only if all parasitics are modeled correctly (figure 9). In practice this is often not the case and also hard to reach at all. If you want to build not only a single PA then the additional design effort should be accepted, because parasitic elements not only influence the performance but may also modify your topology (figure 10). If you want to create a new design you get a much better starting point because much of the needed knowledge is in your models.

Literature :

[1] E.O. Johnson, "Physical Limitations on Frequency and Power Parameters of Transistors", RCA Review, Vol. 26, June 1965
[2] Ted Johansson, "Inside the RF Power Transistor, Applied Microwave & Wireless", September/October 1997, p. 34ff
[3] J. Crols, P. Kinget, J. Craninckx, M. Steyart. "An Analytical Model of Planar Inductors on Lowly Doped Silicon Substrates for High Frequency Analog Design up to 3GHz", Symposium on VLSI Circuits, 1996
[4] J.R. Long, M.A. Copeland, "The Modeling, Chracterization, and Design of Monolythic Inductors for Silicon RF IC's", IEEE Journal of Solid-State Circuits, Vol. 32, No. 3, March 1997, p. 357ff
[5] H. Hasegawa, M. Furukawa, H. Yanai, "Properties of Microstrip Line on Si-SiO_2 Systems", IEEE Transactions on Microwave Theory and Techniques, November 1971, p. 869ff
[6] K. Jodar, "A Simple Approach to Modeling Cross-Talk in Integrated Circuits", IEEE Journal Solid-State Circuits, Vol. 29, No. 10, October 1994, p. 1212ff

Design Considerations for GaAs MESFET RF Power Amplifiers

Stewart S. Taylor

TriQuint Semiconductor
2300 NE Brookwood Parkway
Hillsboro, Oregon 97124 USA
(503) 615-9372, staylor@tqs.com

ABSTRACT

Practical aspects of low-voltage, 1-3 watt RF power amplifiers for wireless phones will be presented, particularly as they relate to GaAs MESFET integrated circuits. Device, package, and circuit approaches will be reviewed. Typical performance of GaAs MESFET power amplifiers for 840MHz cellular AMPS & TDMA, 1900MHz PCS TDMA, and DCS 1800 will be presented.

INTRODUCTION

Mobile telephones constitute a fast-growing consumer market for wireless RF circuits. Small size, light weight, low cost, long talk time, and new features are being offered to or demanded by consumers. At the same time, power supply voltages are coming down. While this is generally positive for the digital circuits in the phone and may place only modest challenges for baseband and receiver circuits, it imposes more severe requirements on transmitter circuits. This results from

the knee voltage and on-resistance of the active devices, and the increased losses in the passive matching networks. Power amplifier designers need a combination of suitable device technology, innovative circuit design, low-loss passive networks, and cost-effective packaging to meet these challenges.

GaAs MESFET technology is well-suited to RF power amplifiers (PAs) for battery operated phones. This results from a high f_{max} (high power gain), high low-field mobility (low knee voltage, R_{on}), high-quality passive components (gold metal, semi-insulating substrate), high breakdown voltage, and electrical robustness. To be sure, most high-volume, low-cost GaAs MESFET PAs need support circuitry, adding some additional cost and complexity. While this may appear significant to some RF designers, it is relatively minor and straightforward for IC designers, and is far outweighed by the improvement in efficiency and linearity for the phone manufacturers.

Power Amplifier Requirements

RF power amplifiers for mobile phones are large-signal devices with power gain that drive the antenna, usually through a switch or a filter. They dramatically impact battery life (talk time), particularly at high power levels, and require good linearity in some cases (TDMA, CDMA). Power amplifiers require good thermal design, including IC layout and suitable packaging.

Output power requirements typically vary from 1-3 watts, delivered from power supplies of 3 to 5 volts. PAs need to be low cost, small size, and easy to use. Efficiency at full power, the single most important specification, typically needs to be at least 60% for AMPS, 50% for GSM, 40% for TDMA, and 30% for CDMA. This level of performance is required for the complete power amplifier including the application circuit, and is often in contrast to load-pull data for a wafer-probed single device. Modern PAs need to be dual-mode devices (linear and saturating), which can usually be accommodated with appropriate bias circuits, and must be reliable and robust into

mismatched loads with a 10:1 VSWR. A mismatched load at full power can produce device voltages 3 to 4 times greater than the power supply and dramatically increase device power dissipation. Many new phones are dual band. If dual-band PAs can occupy less board space and cost less than two separate single-band PAs, they will be embraced in the marketplace. Table 1 is a summary of some common PA requirements.

	AMPS	IS-54/136 TDMA	CDMA	GSM (Class 4)	DCS1800
Tx Frequency	824-850 MHz	824-850 MHz	.824-850 MHz	880-915 MHz	1710-1785MHz
Eff. Rad Power	600mW/28dBm	600mW/28dBm	250mW/24dBm	2W/33dBm	1W/30dBm
PA Power	1.4W/31.5dBm	1.3W/31dBm	600mW/28dBm	3W/35dBm	1.8W/32.5dBm
Duty Cyle	100%	33%	100%	12.5%	12.5%
Peak/Ave Power	0dB	4-6dB	8-10dB	0dB	0dB
η @ P_{max}	>60%	>40%	>30%	>50%	>40%
Modulation	FM	$\pi/4$ QPSK	BPSK/QPSK	0.3 GMSK	0.3 GMSK
Envelope	Constant	Varying	Varying	Constant	Constant

Table 1: PA Standards and Summary

Impact of Low Voltage on PA Performance

A lower battery voltage reduces the voltage swing from the active devices. This, in turn, requires a larger current for the same output power. A greater output current increases I^2R losses including that of the output active device and matching network, for the same Q. Employing a larger output device to overcome the series I^2R losses results in greater shunt output resistance losses. A larger output device for a fixed bias current may result in lower gain due to a reduction in f_{max}. Linearity and/or efficiency are degraded for a fixed knee voltage with smaller supplies. Larger devices employed to overcome some of these limitations increase the die size and cost. All of this, while challenging, tends to favor III-V devices, such as GaAs MESFETs, because of their high low-field mobility.

Saturating Power Amplifiers

Saturating power amplifiers are relatively straight-forward to design and implement. Generally, a class B or class AB output stage is power matched (load line/load pull) at both the input and output, while paying proper attention to stability and load mismatch. The device is sized to meet the power requirement and to maximize efficiency. High power at low voltage and high frequency may pose a challenge as the package/interconnect may limit the transformed load impedance presented to the device. Multiple devices with output power combining may improve performance in some cases. The driver stage is usually class AB. Both the driver and output stage are generally overdriven to achieve full power and high efficiency. Technologies with high power gain improve efficiency, requiring less power from the previous stage. Output tuning may include harmonic termination by design or accident to shape the current and voltage waveforms to achieve high efficiency. The selection of the passive elements in the output matching network and their Q's play an important role in the efficiency of the power amplifier.

Linear Power Amplifiers

In contrast to saturating PAs, which are relatively easy to design, linear PAs are more difficult and less intuitive to implement and apply with satisfactory results. Output power, efficiency, and linearity all interact. For a given circuit implementation, improving one specification is usually at the expense of another. Generally, one would like a device with high inherent linearity, and that can be driven gracefully into compression. With sufficient (hopefully minimal) power backoff, acceptable linearity can be achieved. Linearity is generally quantified by adjacent channel power ratio (ACPR) and alternative channel power ratio (AltCPR). This can be related to the AM/AM and AM/PM characteristics of the PA. GaAs MESFETs, with a low knee voltage and a C_{gs} that varies moderately with bias, have somewhat forgiving AM/AM and AM/PM characteristics. Figure 1 shows a typical ACPR and AltCPR plot from a spectrum analyzer.

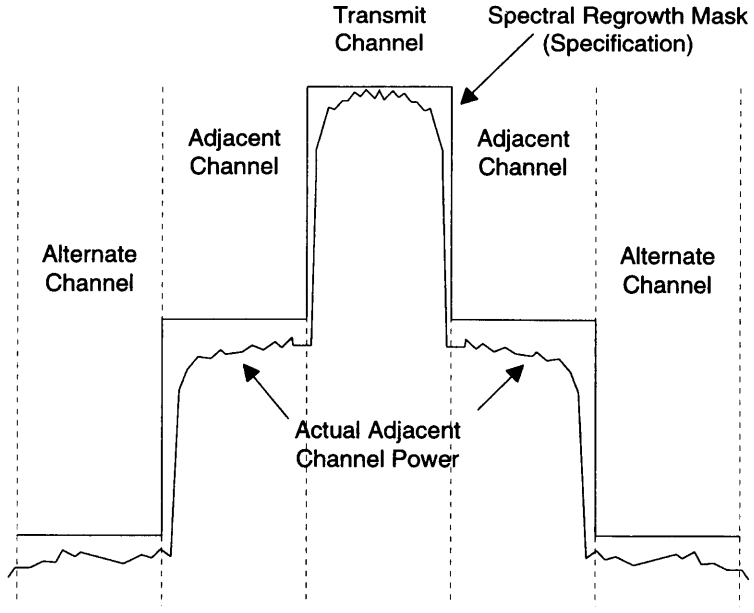


Figure 1: Alternate and Adjacent Channel Power Regrowth

A 1 Watt, 3V Cellular Damps Power Amplifier

Figure 1 is a block diagram of a 1 watt, 3V DAMPS power amplifier. It consists of a two gain stages, bias circuitry, a negative supply generator, and PMOS interface circuit/negative supply sense.

The simplified core PA is shown in Figure 2. The input matching network consists of L1 and C1 while the interstage matching network consists of L2, C2, and L6. R1 and C6 improve the amplifier stability. The output matching network is composed of L4 and C3. L3 is an RFC and C4 is a DC block. The output transistor periphery is 19mm.

Figure 3 shows a simplified current mirror bias circuit that establishes the drain current of dummy transistor Q1 through the use of negative feedback with amplifier A1. I_2, the bias current of the transistor in one of the RF amplifiers, scales with I_1 according to

$$I_2 \cong I_1 \frac{W_2}{W_1} \cong \frac{V_1}{R_1} \frac{W_2}{W_1}$$

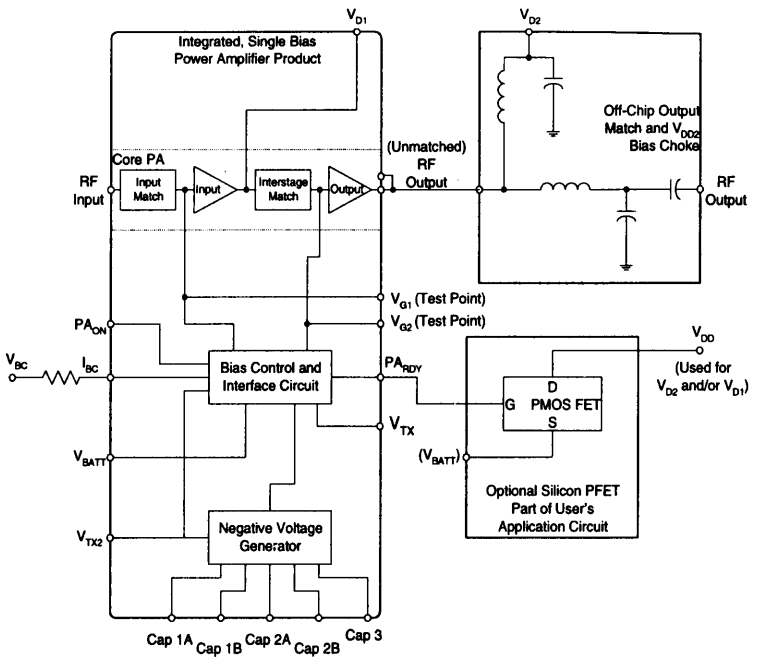


Figure 2: Power Amplifier Block Diagram

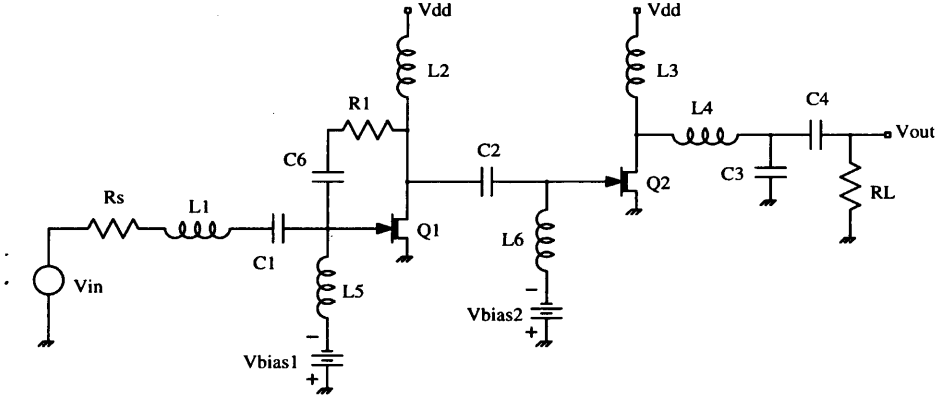


Figure 3: Simplified Power Amplifier

Two of these circuits form Vbias1 and Vbias2 shown in Figure 3. An important feature of the bias amplifier in the current mirror is to operate class AB when sinking current. A simplified circuit diagram of this output stage is shown in Figure 5. Q3 establishes the nominal-

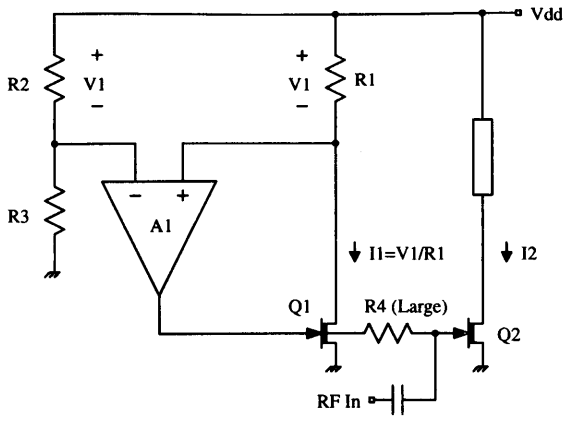


Figure 4: Simplified Biasing

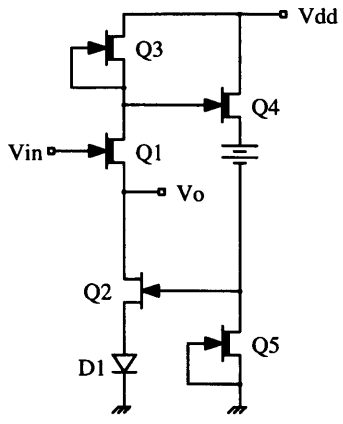


Figure 5: Output Stage

bias current in source follower Q1 and pull down transistor Q2. When Vout goes sufficiently positive, the gate voltage of Q2 is increased by Q3 and Q4, which go positive. This allows Q2 to accommodate up to 8mA of leakage current from the gate of the FET if necessary while standing less than 0.5mA. There is also a series resonant RF trap connected (not shown) from the output of the bias amplifier to ground to suppress the RF signal. This is necessary to avoid overload and a corresponding bias point shift under conditions of large RF drive.

The bias current in the first stage is chosen to be relatively low (10mA). It is found experimentally and with computer simulation that operating the first stage in compression creates an AM/PM curve that partially compensates for that of the output stage. This allows the amplifier to be driven more into compression while meeting adjacent channel power ratio (AdjCPR) and alternate channel power ratio (AltCPR), improving efficiency. If the bias current of the first stage is too low, the AltCPR is too large. If the bias current is too high, the AdjCPR is too large. The low bias current in the input stage allows the output stage to be biased at a lower current, improving efficiency.

Figure 6 is a simplified schematic diagram of the PMOS interface circuit which includes a negative supply voltage sense. This latter feature is essential to protect the amplifier in case the negative supply

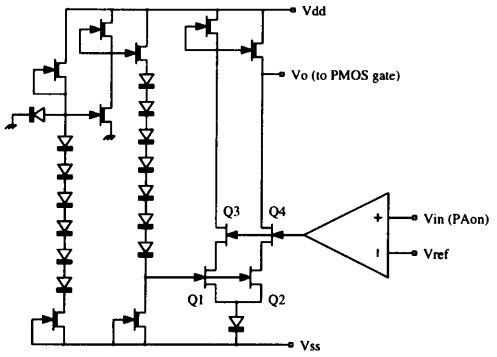


Figure 6: Interface Circuit

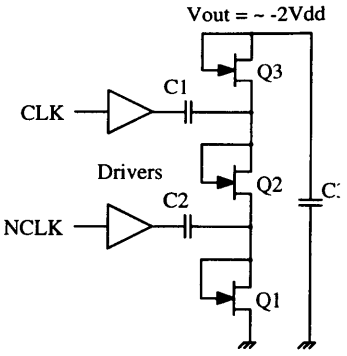


Figure 7: Voltage Doubler

is not present for any reason, including excessive gate leakage in the output transistor. It also delays the turn-on of the PMOS switch until the negative supply is sufficiently large to reduce large DC currents in the amplifier FETs. PAon passes through a noninverting gate whose threshold is set to 1.5V. The output of this buffer drives a NAND gate whose other input is driven by the cascade of two inverters, the first of which has its input connected to the negative supply. The threshold of this negative supply sense is set by several diodes to be approximately -2.5V.

Figure 7 is a simplified schematic diagram of the negative supply generator, which is a voltage doubler. This was required for 2.7V operation with MESFETs with -2.7V worst-case pinchoff voltage since the bias circuitry requires voltage headroom for the transistors to operate in saturation. The voltage doubler consists of two driver circuits, coupling and filter capacitors C1-C3, and rectifying diode-connected EFET transistors Q1-Q3. The diode-connected EFETs rectify with less voltage loss than Schottky diodes, producing a more negative output voltage.

Figures 8 and 9 show simplified circuits for the relaxation oscillator and driver for the voltage doubler. The oscillator frequency is nominally 15MHz, with maximum spurious output levels of -80dBc. The driver circuit is an AC coupled emulation of a standard super-buffered bootstrapped NMOS circuit. The gates of DFETs Q3 and Q5

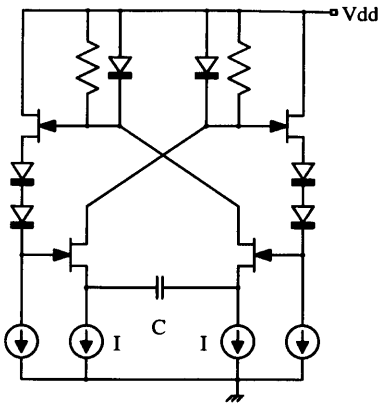


Figure 8: Oscillator

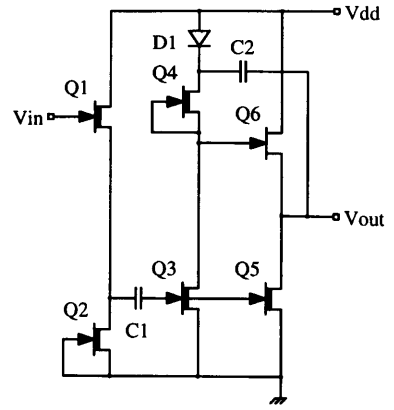


Figure 9: Driver

can be taken below ground to shut them off while the gate of EFET Q6 can be driven above Vdd to make its Vds very small. This results in an efficient circuit that swings within 0.1V of ground and Vdd.

Table 2 is a summary of the performance of the power amplifier, which is assembled in a TSSOP20 package with a down-set paddle that enhances the thermal and electrical performance. Figure 7 shows the adjacent and alternate channel power ratios for 30dBm output power. Figure 16 contains a photograph of the die.

Linearization Techniques

The objective of all linearization techniques is to improve efficiency. By improving linearity, a power amplifier can be driven further into

Po	Vbatt	η @ Po	S11	S21	AdjCPR	AltCPR	Load	Rugg. & Stability
30dBm	3.4V	45%	-12dB @ 10dBm	30dB @ 10dBm	-30dBc	-50dBc	2:1 VSWR, All ∠s	10:1 VSWR in-Band
31dBm	3.4V	50%			-27dBc	-46dBc	50 ohms	10:1 VSWR in-Band

Table 2: Performance Summary

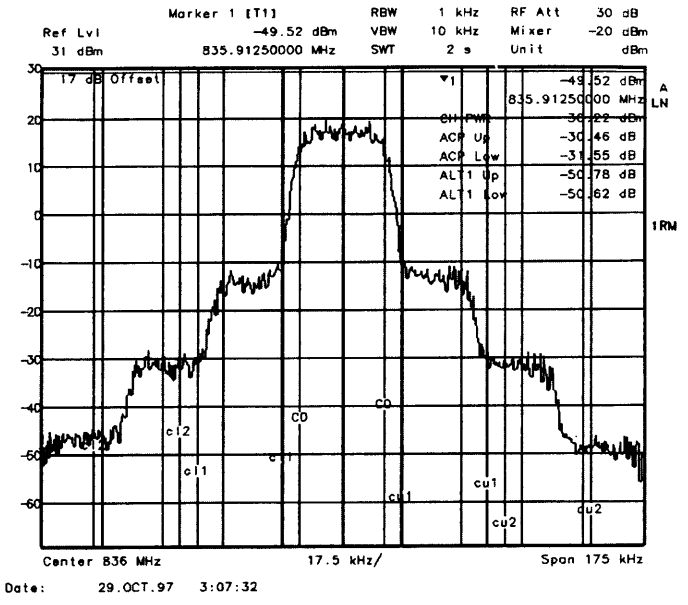


Figure 10: Alternate and Adjacent Channel Power at 30dBm

compression while still meeting the prescribed linearity requirements, thus improving efficiency. Optimum tuning and biasing improve linearity. Beyond this, practical linearization techniques for mobile phones include predistortion, Cartesian feedback, and envelope and phase feedback.

Predistortion is accomplished with a circuit that has approximately complementary AM/AM and AM/PM characteristics of the power amplifier over some range that results in improved linearity, particularly in compression. This is shown in Figure 11. Figure 12 shows two simple circuits that can be employed to predistort the signal driving a typical power amplifier. A shifting bias point with drive signal causes gain expansion and a decreasing AM/PM characteristic. Generally, these types of circuits have insertion loss that must be made up elsewhere in the amplifier.

Figure 13 is a block diagram of Cartesian feedback used to linearize a power amplifier. The RF signal is demodulated to baseband and then

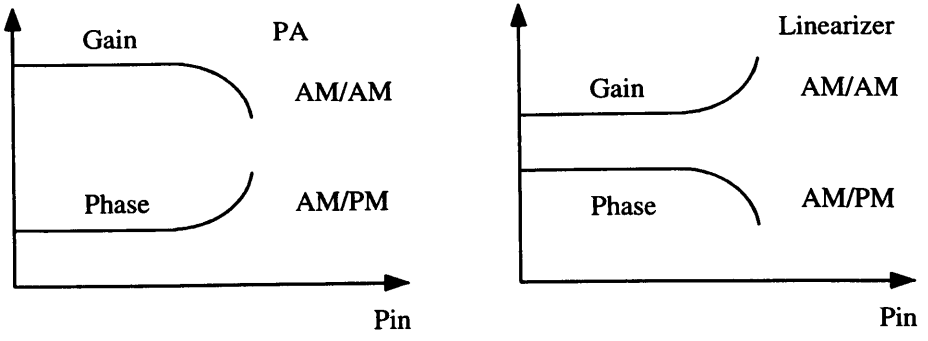


Figure 11: Conceptual Approach to Linearization by Predistortion

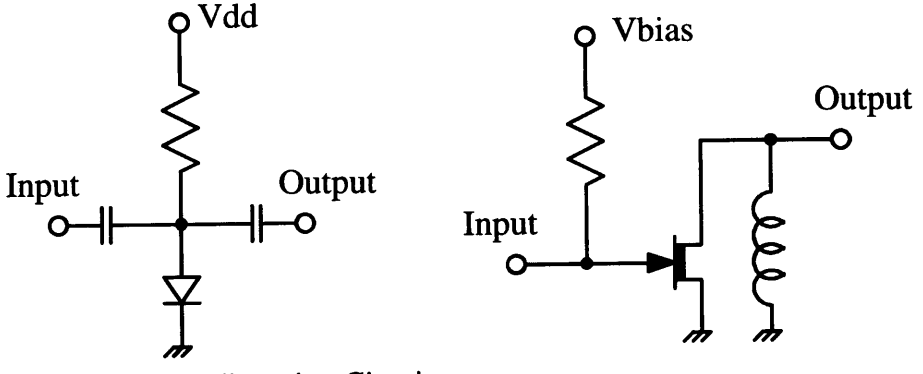


Figure 12: Predistortion Circuits

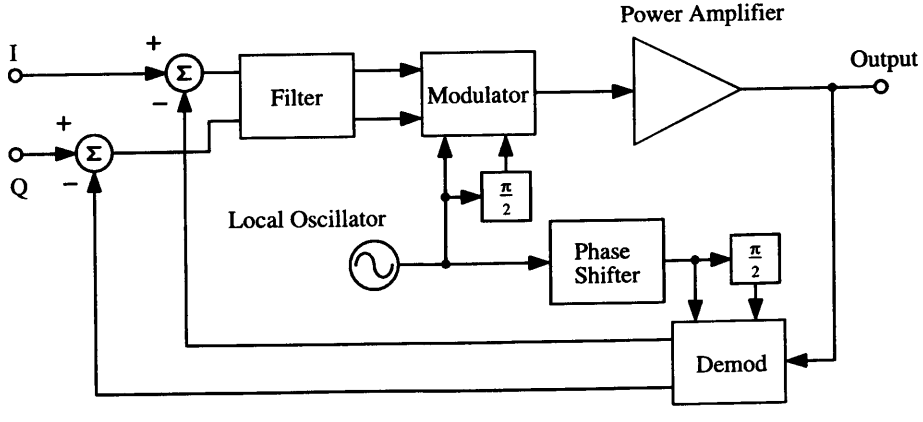


Figure 13: Cartesian Feedback

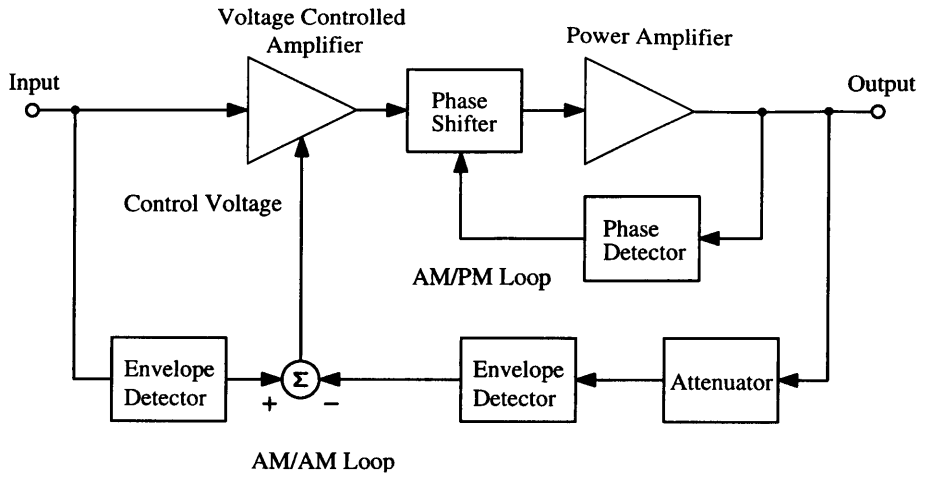


Figure 14: PA with Envelope and Phase Correcting Feedback Loops

nclosed in a feedback loop to linearize the PA. The baseband signal is, in effect, predistorted to compensate for the power amplifier. If the characteristics of the PA change with process or temperature, this approach, unlike the open-loop predistortion circuit shown previously, can track these changes. Cartesian feedback puts the onus on the phone designer rather than the PA designer for the most part. A possible limitation of Cartesian feedback is that if the power amplifier is driven heavily into compression, the resulting AM/PM can cause the feedback loop to become unstable. A modified technique retains most of the same elements with the addition of data converters and memory to measure and store the result of the demodulated and predistorted signal driving the PA. This method becomes an open-loop system that is periodically calibrated to correctly predistort the baseband signal under changing conditions at the expense of additional hardware and power dissipation.

Figure 14 is a block diagram of a linearized power amplifier with envelope and phase correcting feedback. This approach improves the AM/AM and AM/PM characteristics of the power amplifier and is self-contained. It may require significant circuitry, die area, and power.

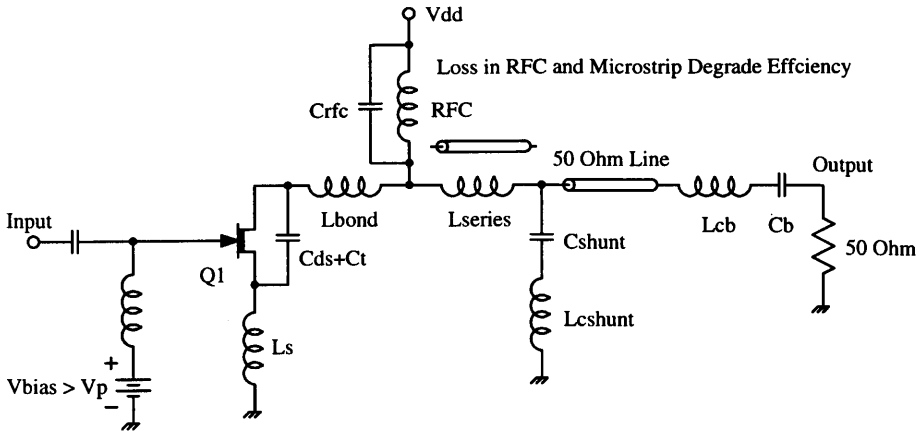


Figure 15: PA Output Stage and Matching Network

@ Pomax	Larger Shunt C	Larger Series L	Larger Vdd	Larger Id (bias)
ACPR	better (lower)	worse (larger)	better (lower)	little change
AltCPR	worse (larger)	better (lower)	little change	better (lower)
Efficiency	worse (lower)	better (larger)	worse (lower)	little change

Table 3. Impact of Tuning, Bias, and Voltage on PA Performance

Output Tuning

The output matching network of a power amplifier usually consists of an RFC (or a small inductor for DC return current), a series inductor, and a shunt capacitor. This relatively simple circuit is usually adequate for transforming the 50 ohm load impedance down to 1 to 4 ohms. The self resonance of the various components sometimes plays a role in the performance of the power amplifier. Sometimes harmonic tuning is used to improve efficiency, or a π network is used to suppress harmonics. Unless this is inside the package or module, the resulting cost, board space, and/or additional losses may be a deterrent to the customer. Figure 15 is an approximate representation of an output stage with a typical matching network. For linear power amplifiers, Table 3 indicates typical changes in performance with tuning, supply voltage, and bias current.

Packaging & Assembly

Cost-effective packaging is essential to the success of power amplifiers. Low thermal and electrical impedances are critical to PA performance. Acceptable junction temperature is needed first for reliability, and then for performance. Output power and efficiency both degrade at higher temperatures.

The equations below show how electrical performance is degraded by the inductance to ground, usually associated with assembly (bond wires). The inductance to ground adds a real component to the input impedance which lowers the effective f_{max} and power gain of the active device. Notice that the low C_{gs} of FETs (because of low g_m) alleviates this situation somewhat compared to bipolar transistors. The small-signal power gain is proportional to f_t / L_s, the square of V_{dd}-V_{knee}, and inversely proportional to frequency squared and output power. Some

$$f_{\max} \cong \frac{f_t}{2}\sqrt{\frac{r_d}{r_i}} \cong \frac{1}{4\pi}\sqrt{\frac{g_m r_d}{C_{gs} L_s}} \cong \frac{1}{4\pi}\sqrt{\frac{1}{\gamma C_{gs} L_s}},\ g_m \cong 2\sqrt{\beta W I_d}\ ,\ f_t \cong \frac{g_m}{2\pi C_{gs}}$$

$$\frac{p_o}{p_{in}} \cong (R_L \| R_o)\left(\frac{f_t}{f}\right)^2 \frac{1}{r_i + 2\pi f_t L_s} = A_I^2 \frac{R_L}{Z_{in}}$$

$$\frac{p_o}{p_{in}} \cong (R_L \| R_o)\frac{f_t}{2\pi L_s f^2} = G_p,\ R_o \cong r_d + 2\pi f L_s (1 + g_m r_d)$$

$$P_o|_{\max} \cong \frac{(V_{batt} - V_{knee})^2}{2R_L}\ ,\ R_L \cong \frac{(V_{batt} - V_{knee})^2}{2P_o}\ ,\ R_L << R_o$$

$$\frac{p_o}{p_{in}} \le \frac{(V_{dd} - V_{knee})^2 f_t}{4\pi L_s f^2 P_o|_{\max}} \text{ (equality assumes } R_L << R_o \text{ and } r_i << 2\pi f_t L_s)$$

technologies, such as LDMOS, make up in lower L_s what they lack in f_t. While ground inductance is manageable at 900MHz, it becomes a challenge at 1900MHz.

Figure 16 is a photograph of a TSSOP20 package and a modern GaAs MESFET PA die.

Power Amplifier Efficiency

The following expression is an approximation for the power-added

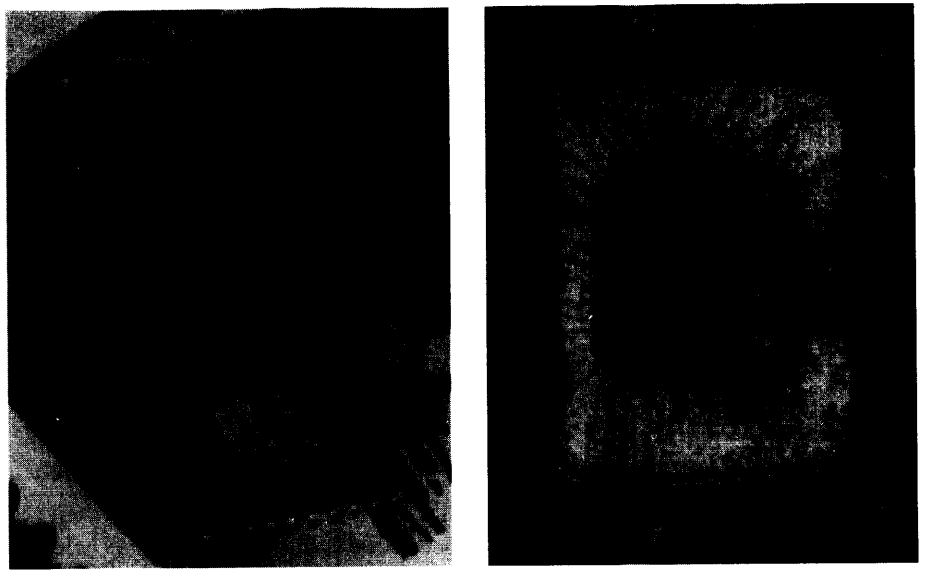


Figure 16: TSSOP20 Package and Bonding Arrangement of a Modern GaAs MESFET Power Amplifier

efficiency of a two-stage power amplifier. It incorporates the information from the above equations as well as the losses associated with the interstage matching network due to on-chip inductor Q (not too important), voltage swing loss due the drop across L_s, and an assumed driver stage efficiency and output resistance. Figure 17 is a graph of PA efficiency using the equation below with some assumed device

$$\eta_{add} = \frac{1}{\frac{1}{\eta_{drain}} + (1.26)\frac{\left(1 + \frac{r_{d1}}{R_p}\right)}{\eta_{driver} G_p}}\ ,\ \text{For a Class B Amplifier 1dB in Compression}$$

$$\eta_{add} \cong \frac{1}{\frac{4}{\pi}\frac{V_{batt}}{(V_{batt} - V_{knee})\sqrt{1 - \left(\frac{2\pi f L_s}{R_L}\right)^2}}\left(1 + \frac{R_L}{R_o}\right) + \left(\frac{1.26}{\eta_{driver}}\frac{1 + \frac{R_L}{R_o}}{R_L}\frac{2\pi L_s f^2}{f_t}\right)\left(1 + \frac{r_{d1}}{Qf_t}\frac{\sqrt{\frac{f_t}{2\pi L_s (r_{d1} - 2\pi f_t L_s)}}}{1 + \frac{2\pi f_t L_s}{r_{d1} - 2\pi f_t L_s}}\right)}$$

$$R_L \cong \frac{(V_{batt} - V_{knee})^2}{2P_o},\ R_o \cong \sqrt{r_d^2 + \left(2\pi f^2 L_s [1 + g_m r_d]\right)^2}$$

$$\eta_{total} \cong \eta_{output\ match}\, \eta_{add}$$

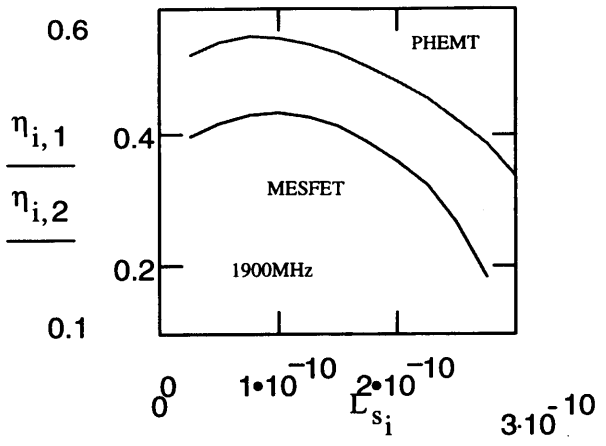
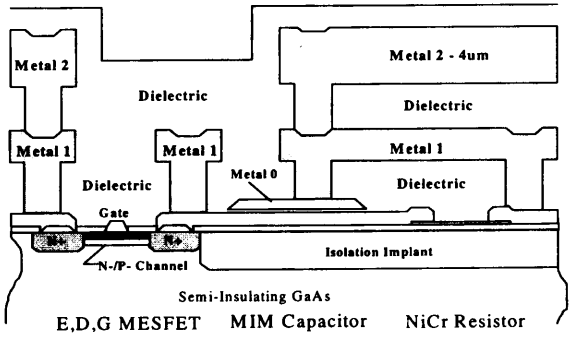


Figure 17: Maximum PA Efficiency vs Ground (Source) Inductance

parameters for a GaAs MESFET and PHEMT at 1900MHz. While it is clear that inductance to ground greater than approximately 100pH rapidly degrades performance, we observe that some inductance is beneficial. This is because it increases the output impedance of the active device more rapidly that it degrades power gain and signal swing. This inductance may also improve linearity to some extent.

A Modern GaAs MESFET Process

Figure 18 is a summary of a modern ion implanted GaAs MESFET process used to make 3V power amplifiers. This technology has three



- RF Transceiver Process
- Vt = 0.15V EFET
- Vt = -0.6V DFET
- Vt = -2.2V GFET
- Three Layer Metal
- Low ε BCB Dielectric
- f_{max} = 50GHz
- C/A = 1200pF/mm^2
- Q = 20 @ 2GHz

Figure 18: Modern Ion Implanted GaAs MESFET Process

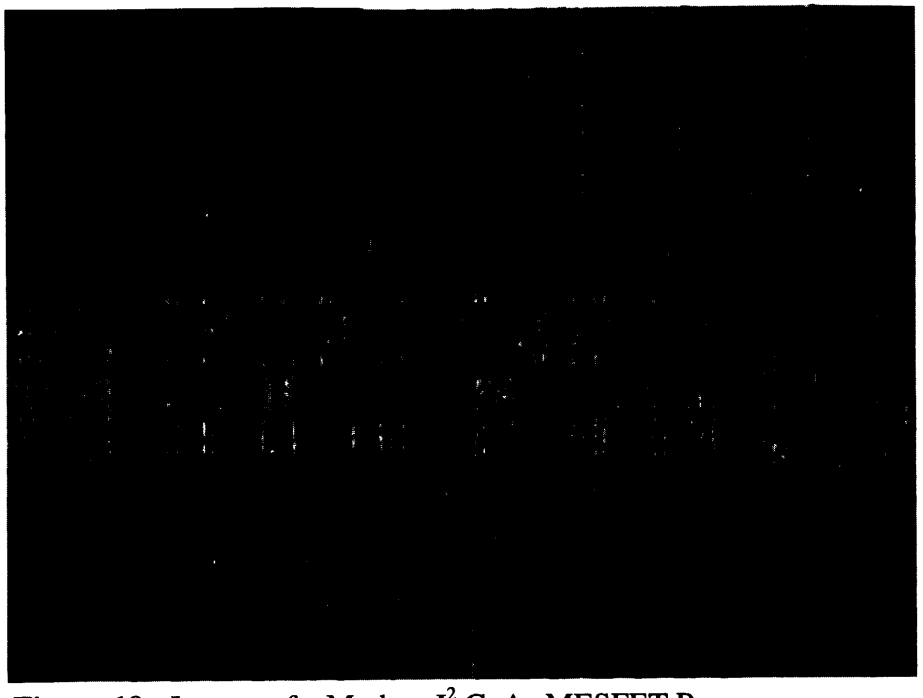


Figure 19: Layout of a Modern I^2 GaAs MESFET Process

layers of gold interconnect, the upper two being 2um and 4um thick. Inductors and transformers with Q's of 20 or more can be implemented by stacking the metal layers. The dielectric is low ε BCB. MIM capacitance is 1200pF/mm^2 while NiCr sheet resistance is 50Ω per square with a very low temperature coefficient. Three FETs are available, one enhancement and two depletion. The threshold voltages are nominally 0.15V, -0.6V, and -2.2V. f_t and f_{max} are 20GHz and 50GHz, respectively. Figure 19 is a photograph of part of the layout of a circuit showing MESFETs, MIM capacitors, inductors, and the three layers of interconnect. Figure 20 is a plot of I_d and g_m versus V_{gs} for the -2.2V power device. It achieves linearity and power handling capability comparable to PHEMT devices.

GaAs MESFET Power Amplifier Performance

Tables 3 and 4 show the performance of commercial 3V GaAs

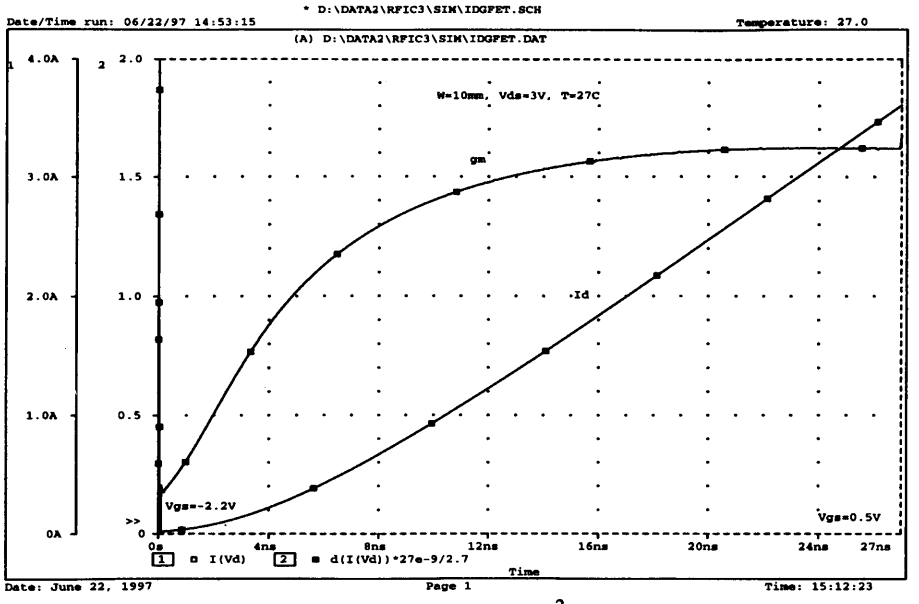


Figure 20. I_d and g_m vs V_{gs} for a Modern I^2 GaAs MESFET Process

MESFET power amplifiers. The efficiencies are generally high while other specifications.

	Rated P_{OUT}	Adj. Channel Power	Alt. Channel Power	Efficiency
TQ7121 Cell Band	30 dBm	-30 dBc Typ	-49 dBc Typ	45 % Typ
TQ7621 PCS Band	30 dBm	-30 dBc Typ	-49 dBc Typ	35 % Typ

Table 4. 3.4V Linear Power Amplifiers for IS-136 (TDMA)

Technology Comparison

Table 5 is a comparison of several technologies used for making power amplifiers. Each can be made to work trading-off cost, performance, and volume (capacity/delivery) according to customer requirements.

	P_{OUT}	Efficiency	Standard
TQ7111 Cell/ISM Band	32 dBm	60 % Typ	AMPS, ETACS, NTACS, ISM
TQ7641 PCS Band	32 dBm	45 % Typ	PCS-1900
TQ7541 DCS Band	32 dBm	45 % Typ	DCS-1800

Table 5. 3.4V Saturated Power Amplifiers for FM and GMSK Mod.

	Si LDMOS	GaAs MESFET	PHEMT	GaAs HBT
Cost	Low	Moderate	High	Highest
Volume/Delivery	High	High	Moderate	Moderate
Efficiency	Good	Very Good	Excellent	Very Good
Low Voltage	Moderate	Good	Very Good	Moderate
Linearity	Good	Very Good	Very Good	Good
Gain	Moderate	Good	Very Good	Good
Single Supply	Yes	Yes (Charge Pump)	Yes	Yes
Vbatt Switch	No	Yes	Maybe	No
On-Chip Passives	Poor	Good	Good	Good
Reliability	Good	Good	Good	?

Table 6. Technology Comparison

Future Trends

MCM packaging technology may allow high-quality passive networks to be integrated with active devices and RF ICs to produce compact, high-performance power amplifiers with module functionality. Linearization techniques will allow amplifiers to be driven further into compression to improve efficiency while meeting ACPR and AltCPR. Single output transistor power amplifiers will emerge as low-cost, area-efficient solutions for dual-band phones. GaAs MESFET technology will continue to play an important role in high-volume, low-cost power amplifiers for mobile phones.