Surface versus Performance Trade-offs: A Review of Layout Techniques

Pietro M. Ferreira, Emilie Avignon-Meseldzija, Philippe Bénabès, Francis Trélin
Université Paris-Saclay, CentraleSupélec, CNRS, Lab. de Génie Électrique et Électronique de Paris,91192,Gif-sur-Yvette,France.
Sorbonne Université, CNRS, Lab. de Génie Électrique et Électronique de Paris, 75252, Paris, France.
e-mail: maris@ieee.org

***Abstract*— Selecting the relevant layout techniques is a key point to obtain a high-performance integrated circuit. Most of the common layout techniques, beside allowing the improvement of performance, also leads to an area overhead. Moreover, this area overhead is generally not accurately evaluated. It is proposed in this review to analyze and to evaluate the surface versus performance trade-off in three types of circuits : digital, low-frequency and radiofrequency analog circuits. Each circuit is post-layout simulated using BiCMOS SiGe 55 nm technology from STMicroelectronics. The first analysis evaluates the surface, power consumption and speed trade-off in a digital circuit implementing a 16-bit gray counter, when selecting different combinations of gates from the B55 digital library. The second analysis focuses on the implementation of an accurate capacitor ratio for switched capacitor circuits and quantifies the surface versus accuracy performance. The third analysis evaluate the performance trade-off for six different layout techniques applied on a negative resistor required for a VCO.**

***Index Terms*— surface constraints, layout techniques, performance compromise, icLayoutRender**

I. Introduction

Unarguably, minimal Silicon surface is the main target of designers in integrated circuits (IC). However, minimal sized circuits are not necessarily the best option for optimal performance [1]. In digital circuits, design trade-off involves optimizing speed, surface, and power consumption. Surface is considered minimal for specific digital gates, while the speed and the power consumption trade-off is a design choice. Analog circuits often require additional attention to gain, linearity, and noise to work properly, while radio frequency (RF) circuit urges for input and output impedance matching, stability, and power efficiency.

According to Mezhiba and Friedman, IC design goals have evolved from simply minimum area to speed versus power trade-off [1]. Design criteria have been changing over process technology evolution to enable minimum production costs, autonomous operation, and fast calculation. In nanometer scaled technologies, however, design productivity, testability and reliability have become a major concern. A variety of layout solution strategies have been studied over past two decades [2]. Literature revision pointed out a large panel of design methodologies to achieve the best circuit performance for specific applications. However, Silicon surface is not often in the center of attention in such papers.

In either gate-level or system-level layout, computer-aided design (CAD) tools have proposed a diversity of solutions to cope the distance between digital and analog IC design. At transistor-level, however, not much has been done, as designers keep using rectangular shaped Complementary-Metal-Oxide-Semiconductor (CMOS) transistors. Gimenez and his team are one of a few research groups proposing a variety of transistors geometries to incorporate new effects seldom explored [3]. Some of those geometries are circular annular, pillar surrounding gate, diamond, octo, ellipsoidal, fish, and wave MOS transistors. These layout techniques can help the designers to meet a better surface versus performance trade-off.

This work aims to review common layout techniques in digital library selection, passive device, and transistor implementation. To this end, an extensive literature review of performance trade-offs is proposed in a novel point of view, i.e. circuit surface. When circuit surface is considered as not the exclusive cost in IC design, some interesting conclusions are revealed. Analysis provided in this paper has proved that common sense assumptions to optimize circuit surface can lead to high leakage current, to process variability weakness, to congestion in metal connections, and to low number of vias.

This paper is organized as follows. Section II. presents a brief about the CMOS technology chosen for circuit examples and the layout rendering tool used in illustrations. Section III. reviews how the best suitable gates family can be selected to obtain the best performance compromise during digital synthesis. Section IV. focus on the surface and accuracy trade-off to integrate a capacitance ratio at the detriment of the surface. Section V. studies the performance trade-off when layout techniques like common-centroid, interdigitated fingers, double-gate access, and planar EM propagation are considered in a VCO design. Finally, conclusions are drawn.

II. Background

A. 55 nm BiCMOS Technology

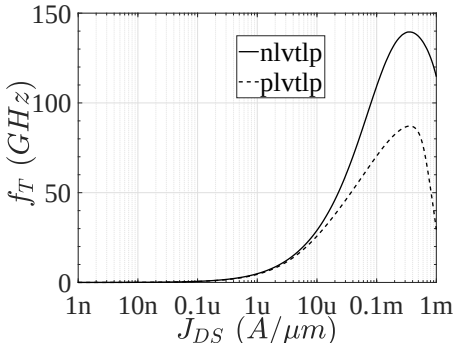
The study is carried out using the BiCMOS SiGe 55 nm technology from ST Microelectronics (B55). This technology includes 8 metal levels, in addition of a top copper metal layer of 3 μm thick. The B55 technology offers low power low threshold ($I_{ON}/I_{OFF} \approx 150k$) MOS transistors dedicated to microwave applications (i.e. self-oscillation frequency $f_{max} > 110$ GHz). Besides, the B55 enables the integration of passive components like inductors, transmission lines and varactors with good quality factors (larger than 10), and digital circuits [4]. Mature BSIM4 and PSP models are provided having process variability and temperature variation from -40 to 175 °C models from silicon measurements [4].

Advanced technologies such as B55 also offers several families of gates with low (lvt), standard (svt) and high (hvt)

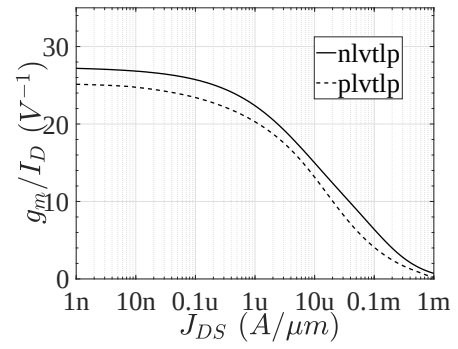
Digital Object Identifier 10.29292/jics.v17i1.589

threshold voltage. Variable oxide thickness leads to low (lp) and great power (gp) options. Combined, six transistor flavors are available, being: lvtlp, lvtgp, svtlp, svtgp, hvtlp, and hvtgp. For digital design, gates libraries are classified as LL, GL, LS, GS, LH, and GH for the respective transistor flavors described. Section III. considers all six libraries separately and two multi-library design solutions.

Section IV. considers analog passive device selection using Metal-Insulator-Metal capacitors using Metal-5 and Metal-6 over a high-permittivity oxide. Section V. considers only n-type (nlvtlp) lvtlp transistors. To this end, a technology parameter extraction is carried out. Figure 1(a) illustrates the transition frequency (f_T) over different transistor inversion levels considering a drain-to-source current density (J_{DS}) variation. Figure 1(b) illustrates the transconductance over drain current ratio (g_m/I_D) for the same J_{DS} variation. One may refer to [5] for more details in g_m/I_D methodology and parameter extraction.



(a)



(b)

Figure 1: B55 characteristics extraction for low-threshold voltage low-power NMOS (nlvtlp) and PMOS (plvtlp) transistors, being (a) the f_T and (b) the g_m/I_D versus J_{DS}.

B. Layout Image Rendering

Most publications illustrate IC layouts in lower picture quality than the simulation illustrations. As a result, the circuit layout becomes hardly readable even for experienced IC designers. This lack of clarity hinds the required physical design solution to address state-of-the-art IC performance, which is the objective of this paper. An open-source tool, name icLayoutRender, is available as a user-friendly transcription of a GDSII file in a high-quality vectorial image [6]. In this review, one may observe the readability improvements when IC Layout Render tool is used. The authors would strongly invite readers to further zoom layout pictures to observe the details in circuit routing and metal connections. Unfortunately, printed versions of this work may lack of layout clarity.

III. Performance Compromise in Gate Library Selection

Electronic design automation (EDA) is the tool to tackle the complexity of digital functions. Hierarchical design flow decomposes register-transfer level (RTL) synthesis in standard cells of a target technology. For the design of digital functions, three main criteria are to be considered: the clock frequency, the area, and the power consumption.

Digital functions are synthesized using standard cells from a library which provides multiple fanout options. Cells with different fanouts are implemented through different transistor sizes, threshold voltage and oxide thickness. Compromises must be studied in the literature to increase the maximum frequency of a digital function to minimize the counterparts in its area and its consumption, even running at the same frequency. Choosing the best suitable gates family helps finding the best performance compromise. This review considers B55 gate libraries. The voltage supply can also vary in a range of 20%, which allows more optimization. To evaluate how these parameters vary, this review considers a small IP requiring a few hundreds of standard cells. Digital function trade-offs are discussed in the following subsections.

A. Digital Synthesis Literature

Lakshminarayana and Jha have discussed power and area optimized solutions for a synthesis framework in [7]. Their algorithm produced circuits whose area and power consumption are comparable to or better than EDA solutions using flattened synthesis. Roy et al. have proposed an efficient algorithm to synthesize prefix graph structures towards optimal speed and area trade-off [8]. Their approach has improved speed by 3% and area by 9%. The obtained RTL graph structure should be considered jointly to EDA tools to consider practical design issues as wire congestion and power consumption.

From all studies development since them, optimal digital synthesis is still a hot research topic even after more than two decades. Cao, Bale, and Trefzer have proposed multi-objective evolutionary optimization flow in [9]. Their work has shown that EDA solutions can be improved when provided a lager panel of cell libraries in a blend of multiple gate fanouts. Those multi-library designs are often lower power and higher speed performance than single-library solutions.

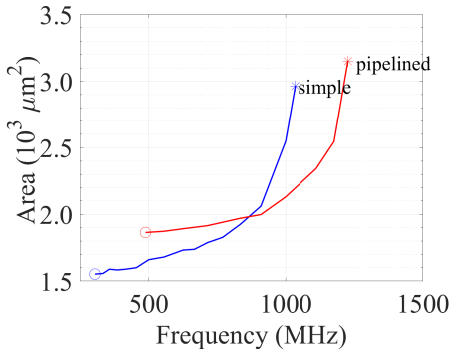
Geng et al. have proposed a heuristic design space exploration using graph neural processes and iterative EDA flow [10]. The golden metric for area power and delay trade-off is found without a cell library choice constraint. Enforcing a golden RTL graph solution in gate synthesis, this work has presented an area of 44.5 x 44.5 μm^2, a power of 6.6 mW, and a delay of 334 ps.

Pipelining is a powerful technique to dramatically increase the performance using storage elements. A variety

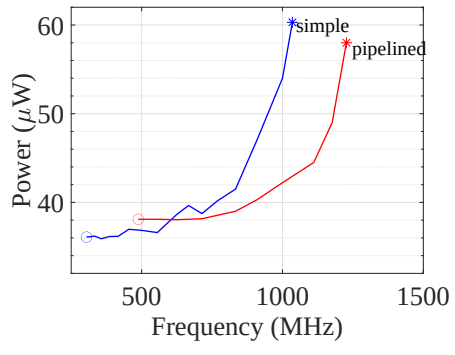
of pipeline architecture has been proposed, considering their advantages and drawbacks. Recently, Ayatollahi et al. revised few pipeline solutions to improve performance in digital circuits using 65 nm CMOS technology [11]. Pipeline scheduler automation is proposed to improve the speed performance from 1% to 128%, while reducing the register stages area from 20% to 74%. Until now, the area overhead required for pipeline stage is the bottleneck for a better compromise.

B. *Design-level optimization*

The digital function designed here is a 16-bit synchronous gray counter followed by a gray-to-binary synchronous decoder. It is implemented with Genus Synthesis Solution from Cadence using B55 in a design for several clock constraints (from few MHz to GHz). Figure 2 illustrates the compromises between area (see Fig. 2(a)) and power consumption (see Fig. 2(b)) versus the maximal frequency without any constraint in synthesis starting point (highlighted by a circle marker). In Fig. 2, the main synthesis parameter is the expected clock period, depicted in blue continuous line for simple 16-bit counter and in red continuous line for a pipelined version. In Fig. 2(b), power consumption estimations are done for a 100 MHz clock. By increasing the clock maximal frequency (i.e. increasing speed constraints), the design is optimized at the expense of gates increase and therefore power consumption. If the frequency is increased up to its maximum, one can get the fastest possible design (highlighted by a start marker).



(a)



(b)

Figure 2: Compromises in 16-bit synchronous gray counter followed by a gray-to-binary synchronous decoder synthesis, being (a) maximum frequency and area, and (b) maximum frequency and power consumption.

It can be deduced that the area and consumption at constant frequency are doubled by going from no-constraint to the maximal clock frequency. One way to release the constraints is to introduce some pipeline (red continuous line in Fig. 2) with flip-flops in the longest combinatorial paths. To understand the way to write a pipeline, let's consider a simple counter described in VHDL as:

```
If rising_edge(clock) then
 Counter <= counter +1;
End if
```

A pipeline can be introduced in the middle of the counter as

```
If rising_edge(clock) then
 counter(7 downto 0) <=
        counter(7 downto 0) +1 ;
If (carry='1') then
 counter(15 downto 8) <=
          counter(15 downto 8) +1
end if ;
If counter(7 downto 0)=''11111110'' then
 carry='1';
else
 carry <= '0';
end if;
```

Notice that the carry is decided one clock step in advance, therefore the comparison is made with 254 and not 255.

Including pipeline in the longest paths of the designed digital function enables the maximal frequency from 1.05 to 1.25 GHz (red curves in Fig. 2). It can be noticed that with low frequency constraints the initial design is more efficient in terms of area and power consumption. However, the pipelined design is more efficient in terms of area when the clock is higher than 900 MHz (see Fig. 2(a)), and more efficient in terms of power after 600 MHz (see Fig. 2(b)). Thus, pipelining the design remains a valuable effort for high-speed designs. All following analyses were done with the pipelined version of the digital function.

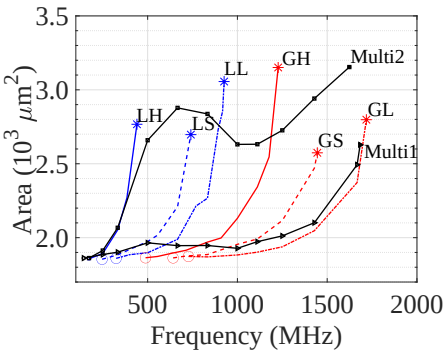


Figure 3: Compromises in the proposed digital function synthesis in terms of maximum frequency and area for LL, GL, LS, GS, LH, and GH single gate library families; for Multi1 (no power constraint) and Multi2 (minimal leakage current).

C. *Gate Library Selection*

The other main criterion in digital synthesis is the choice of the gate library. Considering the libraries LH (continuous blue line), LS (dashed blue line), LL (dash-dotted blue

line), GH (continuous red line), GS (dashed red line), and GL (dash-dotted red line) (see SubSec. A. for details), synthesis has been carried out using single gate families and multi-families. In synthesis version named Multi1 (triangle-marked continuous black line), all gates are available, and no power constraint is given. In synthesis version named Multi2 (square-marked continuous black line), a preference of minimizing leakage current has been given. Figure 3 plots the area as a function of the maximal frequency among six single families and both multi-family's solutions (Multi1 and Multi2).

One can observe that using faster gates (for instance GS or GL libraries) always produces a smaller design. Besides, at low frequencies all available gates can be used. However, when frequency increases, some of the gate libraries are no more suitable as they are too slow. It is noticeable that minimal and maximal area do not depend on the transistor flavors in gate libraries used. When working with multi-family's design, Multi1 (without power constraints) area is always minimized. On the other hand, minimizing the leakage in Multi2 version leads to larger areas which do not increase (varies between 2500 and 3000 μm^2) as counterpart solutions (exponential increase) as the frequency increases. Multi2 solution achieves a maximum frequency as high as Multi1 or GL solutions, but with a 10% to 15% area overhead. At mid frequency range (from 800 MHz to 1 GHz), Multi2 solution overcome LL and GH solutions with a 10% smaller area.

To highlight Multi1 and Multi2 compromises, one should run a more precise analysis in terms of power considering static (P_{stat}) and dynamic (P_{dyn}) power consumption. The P_{stat} shall increase due to leakage current, while P_{dyn} is proportional to the clock frequency. Figure 4 shows P_{stat} and P_{dyn} analyses among six single families, Multi1, and Multi2 solutions.

Figure 4(a) brings out two remarkable observations: (i) the P_{stat} per gate remains almost constant for each family, and (ii) it can vary in a factor of a thousand from the lowest power-hungry library (LH) to the highest one (GL). Multi2 solution presented a lower P_{stat} than Multi1 solution up to 1 GHz, where Multi1 solution overcomes it with a high clock frequency and slightly lower P_{stat}. Besides, Figure 4(b) reveals that P_{dyn} per MHz depends essentially on the number of gates in the design. Indeed, P_{dyn} is very few sensitive to the library selected while P_{stat} is not. Multi2 solution presented the lowest P_{dyn} per MHz, while Multi1 presented a P_{dyn} behavior similar to great-power families (red lines).

If now one considers the total power consumption ($P_{stat} + P_{dyn}$), Fig. 4(c) depicts an increasing power over frequency for Multi1 and Multi2 solutions. Such behavior is very similar to the one observed in Fig. 4(a) highlighting the compromise in power and no-power constraint (Multi1 versus Multi2). This result highlights the predominant portion of P_{stat} over total power and the advantageous design preference of minimizing leakage current in Multi2 solution. Moreover, the lowest power consumption is depicted for the lowest power-hungry library (LH), even if area increases, but at the expense of maximum speed.

D. Supply Voltage Selection

Finally, supply voltage (V_{DD}) selection is a key parameter to reduce power consumption at the expense of clock frequency. Here, supply voltage selection will be considered only for GH library since single gate libraries presented a similar compromise. Besides, the GH solution presents the optimal compromise among area, speed, and power consumption according to Fig. 3 and 4.

Figure 5 presents the GH solution, while V_{DD} options are selected. Available V_{DD} selection is depicted as 0.9 V in blue continuous line, 0.95 V in black continuous line, and 1.05 V in red continuous line. Figure 5(a) illustrates area and clock frequency compromise for V_{DD} options. One may observe that increasing the V_{DD} leads to a frequency increase, but solution at 1.05 V supply presented a 13% area reduction. Figure 5(b) shows the total power consumption as a function of the clock frequency. It is interesting to notice that for high frequency ($\geq$ 1 GHz), it is preferable to increase the supply voltage, in order to reduce the constraints. Such design solution uses fewer gates, and therefore presets a smaller area and a lower power consumption.

Total power consumption shall also be analyzed by its separated components static and dynamic power as depicted in Fig. 6. As expected, P_{stat} and P_{dyn} values are proportional to the square of the supply voltage, i.e. $V_{DD} = 1.05$ V always presents the highest power consumption. Indeed, power reduction is due to an optimal number of gates in $V_{DD} = 1.05$ V in comparison to other two supply voltage choices.

E. Discussions on Gate Library Selection

Herein optimal digital function, some choices have to be done during the design and the synthesis of a digital IP such as gate library family and supply voltage. The constraints evaluated in this work are the area and the power consumption at the expense of clock frequency. If power consumption can be minimized, for example by clock gating or by switching off the supply of a part of the chip, then the leakage power is the major concern for the design. If the digital function runs at constant frequency without possibility to operate in low power, then the total consumption is the only key element in this analysis. Increasing the clock frequency is always done at the expense of an increasing area; such compromise is not linear. If a minimum area is necessary, then the best compromise can be to increase the supply voltage or use faster gates instead of letting the synthesis tool automatically optimize the speed of the digital function. However, area can drastically increase when the design reaches the limit of the technology speed. Regardless, including pipeline in the function design and limiting the longest combinatorial paths are always beneficial at high frequencies.

IV. Passive Device Layout Selection

A. Switched-Capacitor Synthesis Background

From the end of the eighties, switched-capacitor filters are extensively used for analog discrete-time signal processing, due to their compactness and accuracy [12]. They can be found in various applications, from audio signal processing [13], biomedical signal processing [14] to modern communication system [15]. Besides its advantages, switched-capacitor filters also have a critical aspect in which the

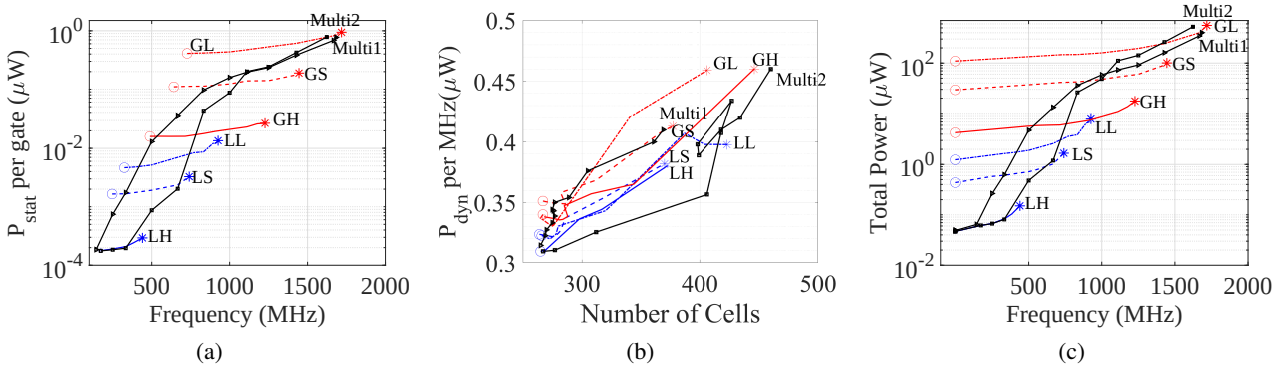


Figure 4: Analysis of static and dynamic power consumption as a function of the synthesis constraints: (a) P_{stat} per gate over clock frequency, (b) P_{dyn} per MHz over the number of gates, and (c) total power consumption over clock frequency.

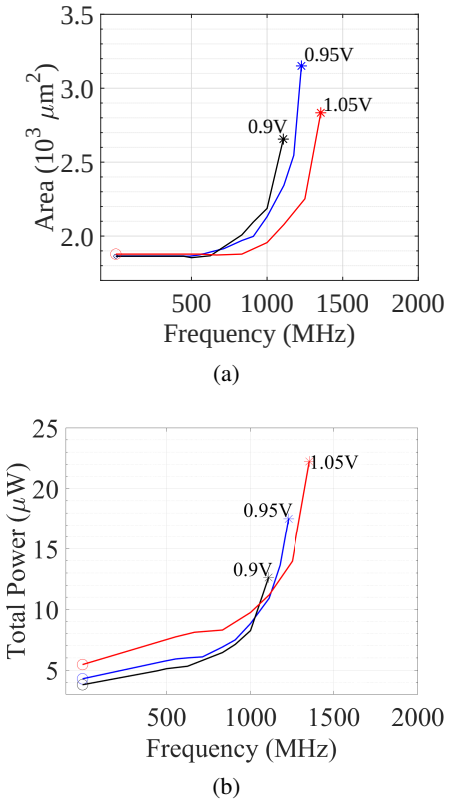


Figure 5: GH library compromises over V_{DD} selection, being (a) maximum frequency and area, and (b) maximum frequency and power consumption.

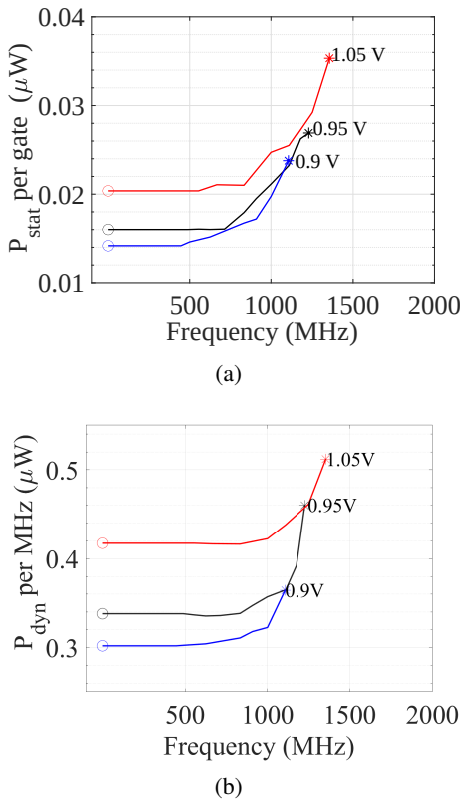


Figure 6: Analysis of static and dynamic power consumption for GH library selection: (a) P_{stat} per gate over clock frequency, and (b) P_{dyn} per MHz over the number of gates.

matching of capacitor ration determines the coefficient of the discrete-time transfer function of the filter.

In [16], McNutt, LeMarquis, and Dunkley have identified five sources of systematic capacitance mismatch and developed a list of layout rules leading to accuracy as good as 0.1% on capacitance ratio. Based on these rules, Khalil et al. have developed for lab use an automatic tool to generate capacitor array to optimize capacitance ratio and evaluated the efficiency of this tool through a measurement of fifty test chip [17]. In [18] and [19], Soares, Mesquita, and Petraglia have proposed a genetic algorithm to obtain the best approximation for capacitance ratio for implementing nonrational filter coefficients with a reasonable number of unit cells. To evaluate and measure capacitance mismatch in CMOS integrated circuits, Soares, Petraglia, and Campos have proposed a quick overview of existing method and two new figures of merit in [20].

Following Subsections focus on the surface and accuracy trade-off to integrate such capacitance ratio. More particularly, it is evaluated the outcome of using strictly unit cell integration for ratio accuracy to the detriment of the surface.

B. Theoretical Consideration on Uncertainty Dimension

Among the five sources of systematic capacitance mismatch identified in [16], mismatch in dimension due to the photolithography inaccuracy is considered the most impacting phenomena. To limit the impact of this phenomenon, the

good practice is to integrate the ratio into unit cells for capacitance ratio sizing. By doing so, it also allows to implement a common centroid structure to compensate the different gradients occurring in the chip. The second source of systematic capacitance mismatch identified in [16] is proximity effects. The solution to mitigate it is to integrate dummies to the capacitors array. However, at what cost in terms of surface and for what gain in terms of capacitance ratio accuracy does these practices (unit cells and dummies) have?

To answer this question, let's consider a capacitance ratio $C_1/C_2 = 1/3$ and the four topologies presented in Fig. 7. In the four illustrated topologies, the representation is scaled with the reference unit dimension W to evaluate the occupied surface in each case; α is the capacitance per surface unit; the required minimum space between each capacitor is considered $0.1W$. H and L are the final dimension of each topology, used to calculate the surface ($S = H \times L$). All capacitors are square for simplicity of the analysis. Each case is studied in terms of gradient compensation, proximity effects and calculate the relative error on the ratio using the dimension uncertainty ΔW, which is related to photolithography device positioning accuracy. Typical positioning inaccuracy for modern photolithography devices is few nanometers.

To evaluate for each topology regarding dimension uncertainty, one should first establish the relative error on the ratio C_1/C_2 as

$$\begin{aligned}\frac{\Delta(C_1/C_2)}{C_1/C_2} &= \frac{C_2}{C_1}\left[\frac{\partial(C_1/C_2)}{\partial C_1}\cdot\Delta C_1 + \frac{\partial(C_1/C_2)}{\partial C_2}\cdot\Delta C_2\right] \\ &= \frac{C_2}{C_1}\left[\frac{1}{C_1}\cdot\Delta C_1 + \frac{C_1}{C_2^2}\cdot\Delta C_2\right].\end{aligned} \tag{1}$$

After obvious simplification, the relative error on the ratio C_1/C_2 is

$$\frac{\Delta(C_1/C_2)}{C_1/C_2} = \frac{1}{C_1}\Delta C_1 - \frac{1}{C_2}\Delta C_2. \tag{2}$$

To calculate the relative error on C_1, due to dimension uncertainty, one shall use the formula

$$\frac{\Delta C_1}{C_1} = \frac{1}{C_1}\frac{\partial C_1}{\partial W_1}\Delta W. \tag{3}$$

By replacing C_1 by αW_1^2, one obtains

$$\frac{\Delta C_1}{C_1} = \frac{2}{W_1}\Delta W. \tag{4}$$

In the same way, the relative error on C_2 regarding dimension uncertainty ΔW is

$$\frac{\Delta C_2}{C_2} = \frac{1}{C_2}\frac{\partial C_2}{\partial W_2}\Delta W, \tag{5}$$

with $C_2 = \alpha W_2^2$, one obtains

$$\frac{\Delta C_2}{C_2} = \frac{2}{W_2}\Delta W. \tag{6}$$

From (2), (4) and (6), one may conclude that the expression of the relative error on the capacitance ratio C_1/C_2 as a function of dimensions W_1 and W_2

$$\frac{\Delta(C_1/C_2)}{C_1/C_2} = \left(\frac{2}{W_1} - \frac{2}{W_2}\right)\Delta W. \tag{7}$$

C. Analysis of Case A

As illustrated in Fig. 7 Case A, C_1 and C_2 are one piece and placed on a different gradient line. Therefore, there is no gradient compensation. Furthermore, each capacitor can be impacted by proximity effects, as they are not surrounded by identical components. The surface of this topology can be deduced from the dimension H and L

$$\begin{aligned}H &= 1.7321\cdot W + 0.1\cdot W + W \\ L &= 1.7321\cdot W,\end{aligned}$$

leading to

$$S = H \times L = 4.9 \times W^2. \tag{8}$$

In this configuration, the dimension of C_1 is $W_1 = W$. Then, the value of C_1 is

$$C_1 = \alpha W_1^2 = \alpha W^2, \tag{9}$$

and the dimension of C_2 is $W_2 = 1.7321W$, so its value is

$$C_2 = \alpha \dot{W}_2^2 = \alpha\,(1.7321W)^2 = 3\alpha W^2 = 3C_1. \tag{10}$$

Thus, the wanted capacitance ratio is $C_1/C_2 = 1/3$. The relative error on the capacitance ratio using (7) in Case A is

$$\frac{\Delta(C_1/C_2)}{C_1/C_2} = 0.845\frac{\Delta W}{W}. \tag{11}$$

D. Analysis of Case B

Case B has C_2 capacitors divided into six capacitors in parallel in which is better in terms of gradient compensation. Indeed, compared to Case A, different parts of C_2 are located on different gradient lines leading to gradient compensation. Regarding proximity effects, Case B is more sensitive as all components are not surrounded by identical components. The surface of this topology can be deduced from the dimension H and L

$$\begin{aligned}H &= 3\cdot 0.7071\cdot W + 2\cdot 0.1\cdot W \\ L &= 2\cdot 0.7071\cdot W + 2\cdot 0.1\cdot W + W,\end{aligned}$$

leading to

$$S = H \times L = 6.07 \times W^2. \tag{12}$$

This is a larger surface than in Case A. In Case B, C_1 is identical to Case A, but C_2 is obtained by six-unit cells in parallel having the value $1/2C_1$ as

$$C_2 = 6\frac{1}{2}C_1 = 6\alpha W_2^2 = 6\alpha(0.7071W)^2 = 3\alpha W^2 = 3C_1. \tag{13}$$

Thus, $W_2 = 0.7071W < W_1$. The relative error on the capacitance ratio calculated with (7) in this Case B is

$$\frac{\Delta(C_1/C_2)}{C_1/C_2} = -0.8285\frac{\Delta W}{W}. \tag{14}$$

It can be noticed that this relative error on the capacitance ratio has slightly decreased compared to Case A. This is due to the fact that W_2 is closer to W_1 compared to the dimension in configuration A.

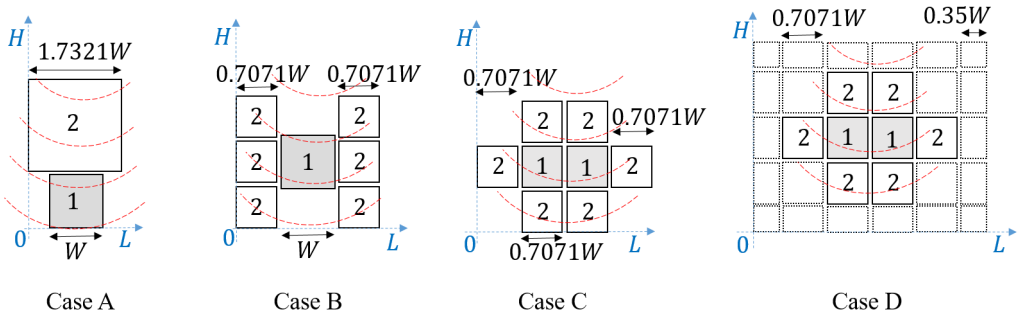


Figure 7 Considered topologies for capacitance ratio layout. Red dotted lines represent gradient lines.

E. *Analysis of Case C*

In Case C, the capacitor C_1 is divided into two capacitors $1/2C_1$ and the capacitors C_2 is divided into six capacitors in parallel of value $1/2C_1$. Case C is satisfactory in terms of gradient compensation but still sensitive to proximity effects as all components are not surrounded by identical components. The surface of this topology, which can be deduced from Fig. 7 is

$$S = H \times L = 7.26 \times W^2 \tag{15}$$

This surface is greater than one presented in Case A and B. The capacitors C_1 and C_2 integrated into unit cells equal to $1/2C_1$ with dimension $W_1 = W_2 = 0.7071W$ can be expressed as:

$$C_2 = 6\frac{1}{2}C_1 = 6\alpha W_2^2 = 6\alpha(0.7071W)^2 = 3\alpha W^2 = 3C_1$$
$$C_1 = 2\frac{1}{2}C_1 = 2\alpha W_1^2 = 2\alpha(0.7071W)^2 = \alpha W^2$$

In this case, due to the dimension $W_1 = W_2 = 0.7071W$, one may derive the condition:

$$\frac{\Delta C_2}{C_2} = \frac{\Delta C_1}{C_1} \tag{16}$$

leading to a zero relative error on the capacitance ratio:

$$\frac{\Delta(C_1/C_2)}{C_1/C_2} = 0 \tag{17}$$

Case C with identical unit cells is the best for dimension uncertainty.

F. *Analysis of Case D*

Case D is the best configuration for all aspects as gradient compensation, dimension uncertainty, and also proximity effects. This is due to the fact that all components are surrounded by identical components. The surface of this topology, which can be deduced from the Fig. 7 is:

$$S = H \times L = 12.97 \times W^2 \tag{18}$$

This is the greatest surface compared to Cases A, B and C.

G. *Surface and Accuracy Compromise*

Based on (11), (14) and (17) established for the relative error on the ratio C_1/C_2, the relative error is plotted as a function of the surface in Fig. 8. For the positioning accuracy of the photolithography device, one may consider $\Delta W = 5$ nm to plot the error for $W = 4$ μm and $W = 6$ μm. From Fig. 8, one can observe that the maximum relative error on C_1/C_2 is obtained for the smallest dimension $W = 4$ μm. This is, since the positioning accuracy of $\Delta W = 5$ nm, the same whatever the size of component. Consequently, the first choice to be made in terms of accuracy on the capacitance ratio is to take large component values in detriment of the surface. For example, the combination C_1=1 pF and C_2=3 pF will give better accuracy on the ratio compared to the combination C_1=100 fF and $C_2 = 300$ fF in detriment of the surface.

Besides, one can also observe, when comparing Case A with B, that as soon as the unit cell is closer to C_1 the absolute relative error on the ratio decreases. In Case C, the relative error is canceled because relative error on C_1 is the same as for C_2. The advantage of Case D in terms of isolation from proximity effects cannot be observed from theoretical analysis, but it illustrates the increased necessary area for integration. Integrating a ratio with small values of capacitances for a ratio is a sensitive task for the trade-off accuracy and surface. In what follows, the four Cases (A, B, C and D) are considered for integrating the ratio C_1/C_2=1/3 with $C_1 = 200$ fF and $C_2 = 600$ fF.

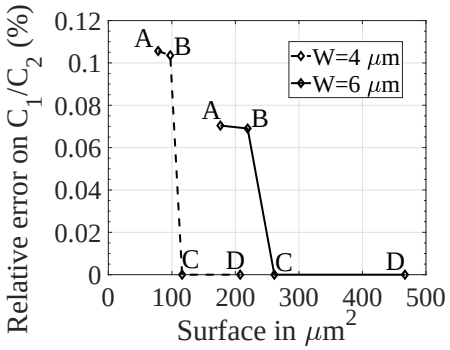


Figure 8 Relative error on C_1/C_2 as a function of surface.

H. *Illustration using B55 technology*

A simulation-based evaluation is proposed to verify the capacitances mismatch for the integrating-ratio C_1/C_2, a test

bench proposal with a switched capacitors integrator is used. The switches and operational amplifier are ideal; only the capacitors are considered in post-layout extracted view as presented in Fig. 9. The frequency of control signal Φ1 and Φ2 is 100 kHz.

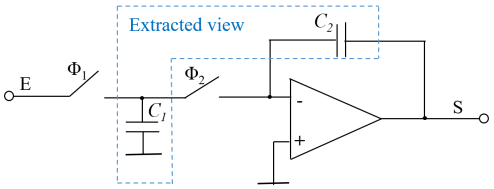


Figure 9: Test bench illustration for a switched capacitor integrator under test.

Simulations are done in the time domain. To deduce the capacitor ratio, test bench imposes an input pulse signal of 100 μs to get the final value at the output, which is ideally supposed to -3.333 V, as presented in Fig. 10. Using a nine-sample Monte Carlo analysis, the final values will change depending on the capacitance ratio as it is illustrated in the inset of Fig. 10. This is the information used here to determine the experimental value of the ratio C_1/C_2 under circuit working conditions.

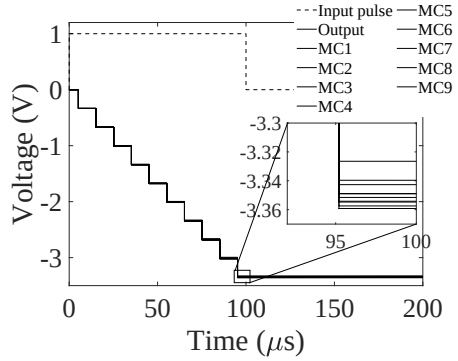


Figure 10: Time domain Monte Carlo Analysis for Input signal (dotted line) and nine-sample output signal (continuous line)

The four layout configurations are presented in Fig. 11(a). Capacitors used are MIM capacitors using Metal-5 over Metal-6 high permittivity oxide. Dimension of each layout configuration is summarized in Table I. Each of these configurations has been simulated for 1000 Monte Carlo samples, considering a similar analysis as depicted in Fig. 10. Results are presented in Fig. 11(b) and the standard deviation of the ratio in each case is presented in Table I. As predicted in theoretical analysis, the two best histograms and standard deviations are obtained for layout C and D. As predicted also, it can be observed that layout B has a slightly better standard deviation because C_1 has dimension closer to unit cells of C_2 compared to layout A. The common centroid configuration in layout B also contribute to this improvement. It can also be observed that the standard deviation obtained for layout D is slightly stronger than the one obtained for layout C. This result can be explained the following two reasons:

- There are no proximity effects in layout C, like heating or strong-signal spurious in the vicinity, to see the real impact of isolation in layout D
- In layout D, capacitors have been spaced and oriented all in the same direction to apply the technique of routing presented in [21]. With this technique of routing, there are more line lengths and less compact layout leading to increased parasitic effects.

Table I. Dimension of for Each Layout Cases

Case	L (μm)	H (μm)	$S = H \cdot L$ (μm^2)	σ (m)
A	17.4	28.3	492.4	0.711
B	34.8	34.5	1200.6	0.687
C	40.3	31.3	1261.4	0.128
D	60.3	54.4	3280.3	0.145

V. Transistor Layout Selection

Transistor-level layout is often laborious and essential to assure performance specifications. Besides Gimenez and few other authors exploring novel CMOS transistors geometries [3], the majority of designers prefers the classic rectangular shape, composed of a number of transistor fingers (N_f) sized with a unitary finger width (W_f), and organized in multipliers (M) of parallel cells. Gate length (L_g) is usually chosen minimal for high speed or from twice to four times the minimal for better linearity. Transistor sizing combined with its bias imposes a transconductance (g_m) as a key parameter for gain and noise performance. Conversely, transistor biasing directly operates over inversion coefficient (IC) and thus power consumption.

Reliability has become an outsider requirement in transistor performance, being mandatory in harsh environments. Maricau and Gielen have given an overview of important CAD tools considering performance reliability in nanometer CMOS [22]. Pan and Graeb have highlighted the area overhead required to improve reliability in [23]. In the literature, simple current mirror and Miller OTA are the design examples where a 50% area overhead is mandatory to achieve a better than 96% yield for ten-year lifetime. Regarding RF circuits, reliability restrain the bias design space in [24, 25] which impacts more the speed and power consumption trade-off rather than the required surface overhead. Recent works put forward reliability constraint in an egalitarian importance as power consumption and speed [26].

A. State-of-the-art Techniques

Layout techniques have been discussed in the literature, and most of the designers know the requirements leading to a preference for common-centroid and interdigitated fingers. That knowledge is often detailed in microelectronics textbooks as [27] chapter 2.4. However, most of IC designs are by CAD automated leading to compact and undistinguished blocks of transistors. If design performance is met, laborious layout is avoided. Eventually, custom layout is mandatory to meet high speed, low noise, high linearity, or matching.

Cathelin et al. have introduced potentialities of digital-enabled CMOS technologies for millimeter wave (mmW) applications in [28]. There, several transistor layout topologies are considered to minimize the extrinsic RF parasitic

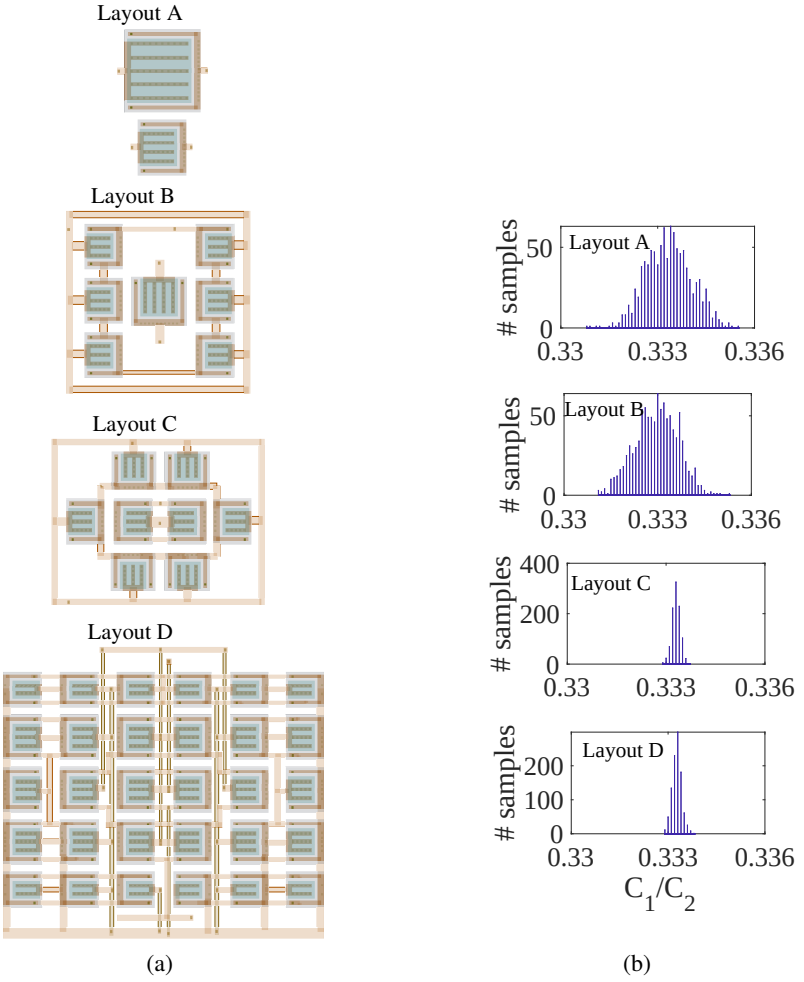


Figure 11 (a) C_1/C_2 Capacitor Layout, and (b) histogram for 1000-point Monte Carlo PLS results for Case A, B, C and D.

elements added in intrinsic transistor core. Their work concludes that exists an optimal transistor size

$$S = M \cdot N_f \cdot W_f \cdot L_g, \tag{19}$$

considering double-gate access and fingers sized in $1\ \mu m \leq W_f \leq 2\ \mu m$. Cathelin's work suggests a N_f and an M part of a compact design choice. For 65 nm ST technology node, they conclude that sized as $W_{tot} = 60\ \mu m\ L_g = 60\ nm$, being $W_f = 1\ \mu m$, $N_f = 12$, $M = 5$ a transistor may achieve a 17% improvement in f_T and 55% improvement in f_{max} if specific layout techniques are respected. Figure 12 illustrates the proposed NMOS transistor layout using double-gate access and planar electromagnetic wave (EM) propagation to maximize mmW performance.

Since then, some authors have been considering high-speed layout technique from [28]. Zhang, Fan, and Sinencio have proposed a linearization technique for high-frequency wide-band LNAs [29]. Using a 130 nm CMOS process, common-source and common-gate LNAs are design-optimized. A high-frequency linearization method is carried out considering impedance matching, power consumption, gain and NF trade-off. Silicon surface is depicted being comparable to the presented layout technique.

A different layout technique for mmW power-efficient transistors is introduced in [30] and confronted to conventional layout as in analog layout books [27]. Liang's and Razavi's proposal considers a unitary transistor cell interdigitated and in a common centroid. However, the double-gate access and EM propagation as in [28] are neglected. In [30], the optimal transconductance should be

$$g_m = \frac{N_f \cdot g_{m,u}}{1 + N_f^2 \cdot \frac{g_{m,u} \cdot Z_u}{3}}, \tag{20}$$

where $g_{m,u}$ is the transconductance of the unitary transistor cell; Z_u is the source degeneration impedance due to parasites components in unitary transistor connections. Moreover, the ideal $g_m \approx N_f \cdot g_{m,u}$ is obtained only if an optimal number of fingers ($N_{f,opt}$) is respected, as

$$N_{f,opt} = \sqrt{\frac{3}{g_{m,u} \cdot Z_u}}. \tag{21}$$

Proposed technique is illustrated in a VCO design having an optimal $W_f = 1.2\ \mu m$, $M = 4$, and $N_f = 20$ for a power efficiency improvement from 4% using conventional layout to 11% with the proposal.

Later, other authors have preferred power-efficient layout from [30]. Using 130 nm CMOS node, Costa et al have proposed two classes of design to explore a linearity and NF trade-off in [31]. High-linear LNA requires main transistors sized $M = 24$, $N_f = 4$, $W_f = 2\ \mu m$, while best-NF LNA requires a $W_f = 1.77\ \mu m$ being a 10% area reduction. Both design examples have similar power consumption, gain

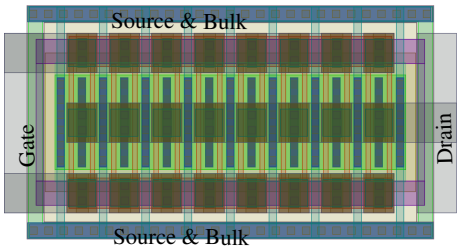


Figure 12: Transistor layout topology considering a minimum of extrinsic RF parasitic elements added in intrinsic transistor core

and impedance matching in the depicted surface and performance trade-off.

Rathore and Darji have compared three layouts for a same sizing of a folded cascode amplifier in [32]. A standard and quasi-automatic layout are first studied, where no analog layout techniques are employed, and the place and route automatically generated by the CAD tool. A second layout consider using multi-finger transistors, which minimize parasitic capacitances between channel and substrate and some parasitic resistances in interconnections. A third version of the layout uses multi-finger transistors and also includes common centroid technique to minimize mismatch in differential pairs due to process variations. It is remarkable the area reduction when specific layout techniques are considered (second and third ones), which leads to an important speed improvement (fewer parasitic capacitors). However, multi-finger layout (second) has shown an important gain reduction, which reinforces the [30] studies in finding an $N_{f,opt}$.

Be differential pairs in amplifiers or be a negative resistance in VCOs, transistor matching is often the reason of performance degradation due to asymmetry in the connections. Lee and Park have proposed a symmetric layout technique to manage the process mismatch in VCOs, by removing asymmetrical crossover metal lines [33]. Proposal improved phase noise and output power at the expense of Silicon surface. The proposal has presented a similar layout technique as introduced in [30], and the benefits of symmetric layout and area increase to a better performance trade-off.

Guitton et al. have explored the circuit design space for a LNA considering IC and W_{tot} for the main amplifier in the classic gain, impedance matching and NF trade-off in [34]. Analytical equations are depicted to aid designers. Surface versus performance trade-off is highlighted. A narrowband LNA presented an optimal IC of 0.75 and 2.4 and a W_{tot} of 20 μm and 30 μm for NMOS and PMOS respectively. However, a wide-band LNA requires IC of 3.2 and 10 and a W_{tot} of 150 μm for NMOS and PMOS respectively.

Tasneem and Mahbub have considered the noise and the power consumption trade-off in a high-gain amplifier in [35]. The better trade-off is proposed for L_g twice minimum length and a low IC (i.e. $g_m/I_D \geq 20$ S/A). Tasneem and Mahbub have highlighted the width and length of the input transistors (differential pairs) should be increased to an optimum noise and power consumption trade-off. One may conclude that minimal area leads to a worse performance.

Low-phase noise is a hard-to-meet performance in VCOs, since it is directly related to the LC tank quality factor and negative transconductance. Moezzi and Bakhtiar have proposed an inductance energy factor optimization for a better phase noise in [36]. The proposal includes an interposed network that follows the LC tank. Thus, phase noise improvement does not have to be constrained to the minimum realizable inductance and the maximum quality factor. However, some of presented interposed network solutions include inductive transformers, which are hardly integrated in a small area. Dumont et al. have proposed a different phase noise improvement in 27 GHz VCO using injection-locked oscillator technique in [37]. Silicon surface required is dominated by inductors, while transistor sizing is $W = 8\ \mu m$, $L = 60\ nm$. Noise improvement is, thus, guaranteed by an external low-noise lower-frequency oscillator. Most design limitation is the increasing power consumption required in both [36, 37] techniques.

B. *VCO Design Optimization*

To illustrate the impact of transistor layout constraints in circuit performance, a voltage-controlled-oscillator (VCO) is chosen as [30, 33]. Figure 13(a) illustrates a passive LC-tank oscillator, the inductor is a differential coil of 557 pH. The capacitor is implemented using a differential varicap of $C_{min} = 1.2$ fF and $C_{max} = 1.9$ fF like the one measured in [38]. Expected oscillation frequency is 19 GHz, which requires a transistor $f_T \geq 20$ GHz. In Fig. 1(a), the required bias is $1\mu A/\mu m \leq J_{DS} \leq 100\mu A/\mu m$ from moderate to strong inversion. Thus, power consumption could be traded for a better performance. To compensate LC-tank losses, a pair of NMOS transistors (nlvtlp) is designed as a negative $g_m \geq 5$ mS, which turns in device sizing of $W_{tot} = 96\ \mu$m and $L = 60$ nm [30].

Figure 13(b) illustrates the layout of the VCO under analysis having an area of 241 x 270 μm^2. For a fair analysis on the impact of transistor layout selection, required current mirror for I_{ref} bias and on-chip balun for output power matching are considered ideal. Thus, only nlvtlp layout selection may differ post-layout and electrical simulations. One may observe that VCO area is dominated by the coil area in Fig. 13(b). Moreover, transmission lines (see V_{op} and V_{om}) are mandatory for a proper 50 Ω-matching at such frequency. Transmission lines are designed from available Pcell using Metal-8 for the RF signal, and from Metal-7 to Metal-1 ground plane. All those layout constraints leave the designer a transistor surface constrained as highlighted in Fig. 13(b) (see the gray shaded rectangle), i.e. transistor layout selected must fit in that area. Nevertheless, such transistor layout selection is not without consequences in VCO performance as is described in the following subsections.

C. *Transistor Layout Selection*

As most designers do, the first layout selected is a classic automatic layout using Pcell constraints. No specific layout design technique is considered except by choosing a $W_f = 2.4\ \mu$m, which leads to a $M \cdot N_f = 40$ organized in only one line. Figure 14(a) illustrates M_1 on the bottom and M_2 on the top of the picture. Metal lines are chosen to route the sources to ground using either Metal1 or Metal 2, thus gates and drains are connected using Metal3 or Metal4. Higher

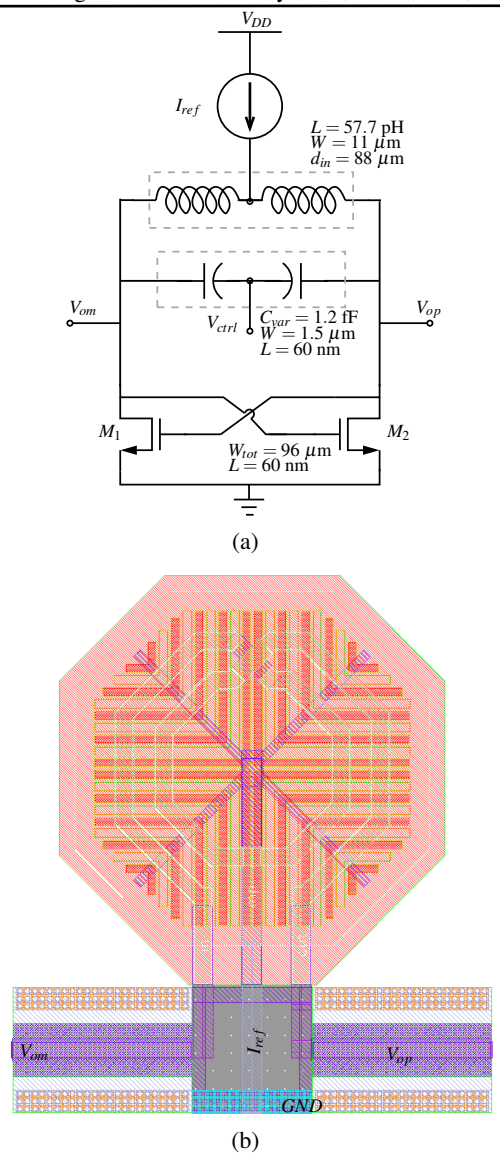


Figure 13: Voltage controlled oscillator under test: (a) the schematic, and (b) the layout (241 x 270 μm^2) illustration.

metal lines are considered only when a connection to the transmission line is required.

Another classic practice among designers is attempting to draw a layout as custom cells without considering state-of-the-art techniques described in this review. No specific layout design technique is considered except by choosing a $W_f = 4$ μm, which leads to a $M \cdot N_f = 24$ organized in two lines. Figure 14(b) illustrates M_1 on the left and M_2 on right. Small metal lines are chosen to route gates and drains in Metal1 or Metal2, while sources are connected directly to the transistor substrate as a ground line. Finally, a larger Metal4 line is used to connect to transmission lines. One may observe that such a designer has drawn a layout using a metal line crossover very similar to what [33] suggests to never do. As presented in the literature, such metal line crossover shall increase the parasitic capacitance between V_{op} and V_{om}.

Experienced designers would rather consider common-centroid layouts including M_1 and M_2 interdigitated gates. If such attempt considers a similar layout as proposed by [30, 33] to guarantee the condition (21), one may find a $W_f = 4$ μm leading to a $M \cdot N_f = 24$, but organized in four lines. Figure 14(c) shows a common centroid interdigitated M_1 and M_2 respecting the layout guidelines of [30], which is expected a considerable output power improvement. Large Metal3 and Metal4 lines are considered in gate and drain connections, while crossovers are avoided as suggested by [33]. Sources are connected to transistor bulk using Metal1 and Metal2. Higher metal lines are considered only when a connection to the transmission line is required.

A different attempt is considered now as similar to layout techniques described in textbooks as [27] chapter 2.4. Then, layout selection turns to different aspect ratio of previous described, it is more vertical instead of being more horizontal (previous one). One may consider a $W_f = 2.4$ μm, which leads to a $M \cdot N_f = 40$ organized in only one line. Figure 14(d) shows a common centroid interdigitated M_1 and M_2 including dummies on the left and right transistor sides. Large Metal2 lines are considered in gate and in drain connections, while crossovers are avoided as V_{op} is routed in the bottom and V_{om} in the top. Sources are connected to transistor bulk using Metal1 and Metal2. Higher metal lines are considered only when a connection to the transmission line is required.

To minimize extrinsic RF parasitic elements added in intrinsic transistor core mmW transistor layout, double-gate connections are considered and (19) met. Considering RF layout techniques as [28] is mandatory for high frequency operation as 19 GHz oscillators. One may consider a $W_f = 1.2$ μm, which leads to a $M \cdot N_f = 80$ organized in four lines. Figure 14(e) illustrates the mmW transistor layout selection including a common centroid interdigitated M_1 and M_2, double-gate connections using four contacts on each, and dummy transistors for further protection. Besides, EM propagation from transmission lines to transistors are considered in a same orientation, i.e. horizontal, using the thicker Metal8 layer as close as possible to the transistor core. Such constraints, however, lead to routing I_{ref} and V_{ctrl} all around the transistors area using Metal2 or Metal3 lines.

In Fig. 14(e), one may observe how hard is V_{op} and V_{om} routing constraints. Such constrained layout leads to a substantial number of crossovers and fewer vias if compared to previous custom layout techniques (see Fig. 14(c) and 14(d). Thus, a second attempt of mmW transistor layout is considered here. Figure 14(f) shows the same RF layout techniques as [28] and considered $W_f = 1.2$ μm and a $M \cdot N_f = 80$ organized in four lines. Relaxed constraints in Metal8 routing enable a Metal5 lines for gate and drain connections, while I_{ref} and V_{ctrl} lines cross transistor core. A substantial number of vias are considered now from Metal5 to Metal8, but still crossovers could not be reduced. One may conclude that compact RF layout technique suffers from such limitation, which is not necessarily a problem for LNAs as in [28].

D. *Post-Layout Simulations*

Different attempts could be done, but the authors decided to limit to only six different layouts. Figure 15 presents a fair comparison between the six attempts using post-layout simulations. Speed performance is asserted by the oscillation frequency (f_{osc}) variation over power consumption compromise. Noise performance is considered by the phase noise

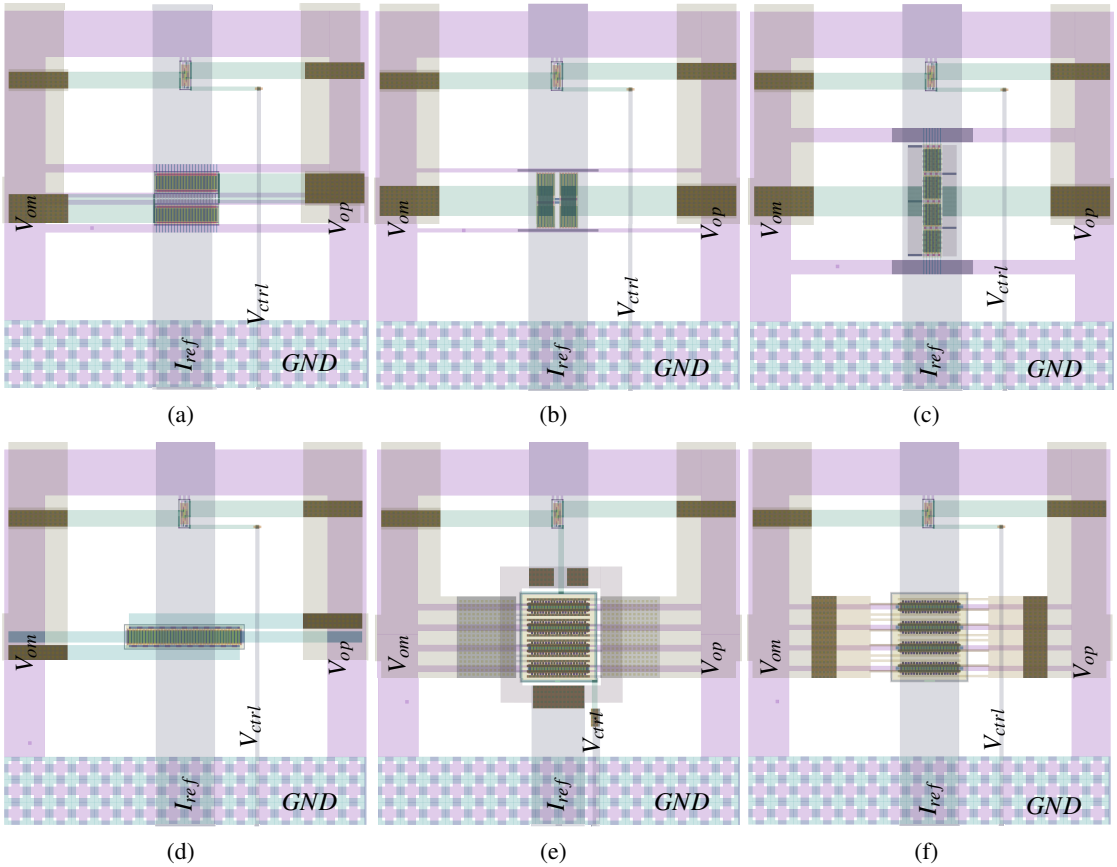


Figure 14: Voltage controlled oscillator layout (59 x 63 μm^2): (a) classic using automatic routing and Pcell constraints, (b) classic using manual routing and custom cells, (c) full custom considering common-centroid interdigitated cells in four lines, (d) full custom considering common-centroid interdigitated cells in one line, (e) full-custom RF technique considering double-gate access, EM propagation common centroid interdigitated cells for ticker Metal8 routing, and (f) full-custom RF technique for a thinner Metal5 routing.

calculated at a 1 MHz shift ($L(1\,MHz)$). Gain performance is verified through transistors transconductance (g_m), which drives a 50 Ω-matched load. Linearity performance is estimated using the total harmonic distortion (THD) figure of merit. Dynamic range is obtained through and ideal matched balun in a similar test bench as [30] to estimate the output power (P_{out}).

Figure 15 compiles the six common characteristics in transistor design, being: the speed, the noise, the gain, the linearity, the output power at full dynamic range, and the g_m/I_D bias point. All these results are obtained from a PSS/PNOISE Spectre post-layout simulation (PLS) for $V_{ctrl} = 0$ V compared to electrical simulations. The gray-continuous line in Fig. 15 represents the electrical simulation of the VCO, which is a reference for ideal performance without layout parasitic devices. In Fig. 15, layouts (a) and (b) are classic layout (no specific technique from literature) using either automatic or custom cells. They are represented in black continuous- and dashed lines respectively. In Fig. 15, layout (c) and (d) consider common-centroid interdigitated transistors and reduced the crossovers between V_{op} and V_{om} as suggested in [27, 30, 33]. They are represented in blue continuous- and dashed lines respectively. In Fig. 15, layouts (e) and (f) consider double-gate access and EM propagation as in [28] in addition to common-centroid interdigitated layout technique. They are represented in red continuous- and dashed lines respectively.

Figure 15(a) illustrates the oscillation frequency (f_{osc}) variation versus the power consumption. As expected, the electrical simulation is the one who achieves the maximum f_{osc}. A common sense among designers would point out parasitic capacitors between V_{op} and V_{om} nodes which are unavoidable. It is remarkable to observe that layout (b), the one using custom cells and no specific technique from literature, is the fastest observed result. Bottom line, using state-of-the-art layout techniques only leads the designer to a compact layout with a very complicated connection grid, where eventually more V_{op} and V_{om} crossings exist. For speed point of view, the layout technique which is the most important is the one described in [33] to limit the crossover. Such issue may not be observed in an amplifier design since.

Figure 15(b) illustrates the phase noise ($L(\Delta f)$) at 1 MHz for different power consumption values. As expected, $L(1\,MHz)$ drops at high power consumption since most of the noise is thermal noise and linear dependent to the g_m. Conversely, the layout (b) is the worst option for noise performance. Moreover, the RF layout techniques applied in layout (e) and (f) have led to a 10 dBc/Hz phase noise reduction being an excellent choice for low noise MOS layout.

Figure 15(c) completes previous observations with the g_m over the power consumption variation. Layout parasitic drops the g_m of up to 5 mS (gray line compared to others).

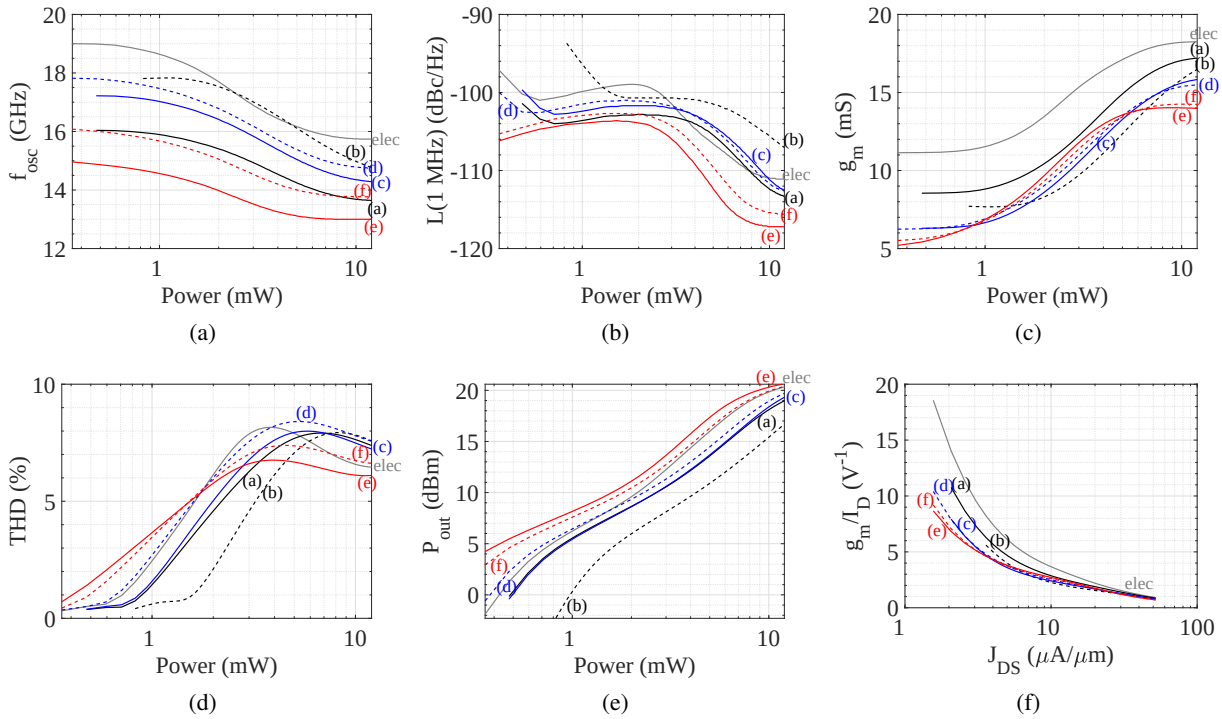


Figure 15: Voltage controlled oscillator post-layout simulation results over power consumption compromise: (a) speed from f_{osc}, (b) noise from $L(1\ MHz)$, (c) gain from g_m, (d) linearity from THD, (e) dynamic range from P_{out}, analysis and (f) g_m/I_D characteristics over layout parasitic.

Such results eventually reveal why the f_{osc} and $L(\Delta f)$ decrease leading to a slower oscillator with a better noise performance. In an amplifier point of view, such g_m drop shall turn in a 6 to a 12 dB gain loss for low- to high-power operation respectively, comparing electrical against PLS of layout (e). To circumvent this problem, a designer must include design margins overdesigning the gain by few dB.

Figure 15(d) illustrates the total harmonic distortion (THD) versus power consumption. Few PLS results are much more linear than electrical simulations where layouts (a) and (b) (without layout techniques) stand out as less than 5 % non-linearity up to 3 mW. It is interesting to observe that layouts (e) and (f) (with RF layout techniques) present a higher non-linearity at low-power ($Power \leq 2$ mW) which turns in a lower non-linearity at high-power ($Power \geq 2$ mW) in comparison to electrical simulations. Layout (d) is the one which best copies the electrical behavior in terms of non-linearity, i.e. more ideal.

Figure 15(e) illustrates the output power (P_{out} in dBm) for an ideal balun and matched load versus the power consumption. It was expected that layout (c) from [30] shall stand as the best option to deliver the maximum output power. However, one may observe that all other layouts proposed using literature techniques presented a better performance. In [30], layout (c) is compared to (b), and similar conclusions could be drawn from the results of this review. However, RF layout techniques, see layout (e) and (f), depicted the best P_{out} performance of about 3dBm greater than layout (c) (comparing red lines with blue-continuous line).

Figure 15(f) depicts a PLS result of Fig. 1(b) in terms of g_m/I_D versus J_{DS} when parasites are considered. One may observe a g_m/I_D preeminent reduction for all layout solutions, in which moderate inverted designs are the most affected by parasites. Indeed, layouts (e) and (f) are the lowest g_m/I_D, which justifies their poor speed and linearity performance while noise is improved. A PLS g_m/I_D versus J_{DS} is an interesting tool to explain the compromise shift.

When literature layout techniques are considered, a designer does them for an accurate VCO reconfiguration and circuit yield. Let's now consider a fixed power consumption of 1 mW and a $-2.5\ V \leq V_{ctrl} \leq 2.5\ V$ variation from a 101-point PSS/PNOISE Spectre PLS. For each case, an average value is calculated for noise, gain, linearity, and output power. The obtained results are as summarized in Tab. II.

Table II.: Average performance for a fixed 1 mW power consumption over $-2.5\ V \leq V_{ctrl} \leq 2.5\ V$ variation .

Layout	f_{osc} (GHz)	$L(1\ MHz)$ (dBc/Hz)	g_m (mS)	THD (%)	P_{out} (dBm)
Fig. 14(a)	15.9	−103.6	8.8	1.35	5.46
Fig. 14(b)	17.8	−96.1	7.7	0.6	0.17
Fig. 14(c)	17.0	−102.4	6.6	1.5	5.34
Fig. 14(d)	17.5	−101.4	6.8	2.6	6.35
Fig. 14(e)	14.6	−103.9	6.8	3.6	8.1
Fig. 14(f)	15.7	−102.9	6.7	3.4	7.5

Nevertheless, a designer is interested in how f_{osc} varies over V_{ctrl} range to better configure the VCO oscillation frequency. Figure 16 illustrates the speed performance reconfiguration over control voltage normalized to the average f_{osc} for each layout case (see Tab. II). One may observe that layouts (d) and (f) are the best candidates to reproduce

the expected reconfigurability depicted in electrical simulations. Besides, layouts (a) and (b) are the worst candidates, as expected, since they do not use common layout techniques (common-centroid and interdigitated cells). As a result, previous analysis in Fig. 15(a) is foreseen by a narrow point of view resulting in mislead conclusions. Indeed, parasitic capacitors are important in for the proposed f_{osc} and must be considered in the design margin of C_{var}, i.e. by a VCO redesign.

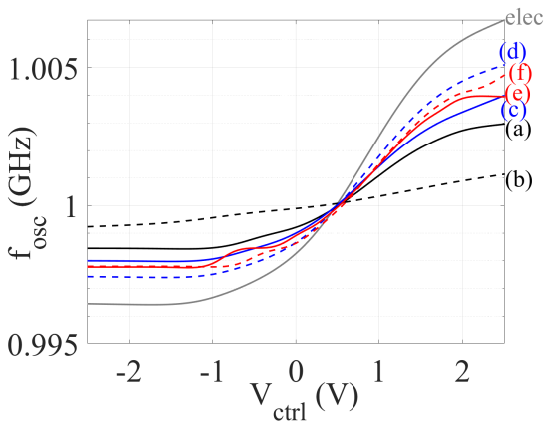


Figure 16: VCO speed performance (f_{osc}) reconfiguration over control voltage (V_{ctrl}) normalized to the average one for each layout case (see Tab. II).

E. Discussions in Transistor Layout Selection

Layout (a) is by far not the best layout, but obtained performance is as good as presented by layout (c) (see Fig. 15). Moreover, layout (a) is the least expensive layout in terms of development cost since it uses automatic placing and routing. Under some circumstances, as low development costs or inexperienced designers, layout (a) is the best option.

Layout (b) highlights that the good intentions of designing custom cells are not enough to assure good performance. Misleading conclusions might be drawn about layout (b) if a designer does not verify circuit performance as a whole. A novice designer might observe Fig. 15(a) and 15(d) to conclude that layout (b) is fast and presents high linearity. However, this is a biased point of view. Indeed, the reason for such observation is a lack of functionality. Layout (b) does not deliver enough power (see Fig. 15(e)); it has a high phase noise (see Fig. 15(b)); it lacks V_{ctrl} reconfigurability (see Fig. 16). Therefore, it does not work at all.

Layouts (c) and (d) reveal that complex common-centroid and interdigitated fingers lead to an important congestion in metal connections. Besides, layout (c) is worse than (d). Layout (d) presents the best compromise in the whole performance analysis depicted in Fig. 15 and 16. Layout (c) is slightly better than layout (d) in terms of P_{out} (see Fig. 15(e)) as claimed in [30]. In the opinion of the author of this review, this is not enough to subdue layout (d) long list of betterments.

Layouts (e) and (f) reveal that ticker metal routing leads to an important congestion in metal connections, and layout (e) is worse than (f). Double-gate access and planar EM propagation are essential only in RF applications. In most of the cases, they impact the measured performance, which is hard to estimate using only PLS results. Layout (f) presents a similar compromise as presented by layout (d) depicted in Fig. 15 except for speed performance. There, parasitic capacitors are important in for the proposed f_{osc} and VCO must be redesigned with appropriate margins in C_{var}. Nevertheless, EM simulations (out of the scope of this review) might lead to totally different conclusions.

VI. Conclusions

First, the minimum area is necessary in digital systems. Moreover, the best compromise can be increasing the supply voltage or using faster gates instead of an automatic blind usage of the synthesis tools. Nevertheless, area can drastically increase when the design speed reaches the limit of the technology. Regardless, including pipeline in the function design and limiting the longest combinatorial paths are always the best design practices at high frequencies.

Second, the necessity to integrate capacitance ratio with unit cells and common centroid placement has been demonstrated mathematically and illustrated on a practical design example. Post-layout simulation results were considered. Unit cells, common centroid, and dummies are unavoidable layout technique capacitance ratio, especially using small capacitance values since they are more sensitive to dimension uncertainty.

Third, the best performance is not often achieved with the state-of-the-art transistor layout techniques. One may observe from this review that there is a compromise between techniques, compactness, and parasites in circuit layout. Common-centroid, interdigitated fingers, double-gate access, and planar EM propagation are common and important layout techniques. However, if the layout constraints are too hard (due to compactness) many metals crossing shall exist. Besides, a limited number of vias might be the only option to respect spacing design rules. Having a high quantity of metal crossings and a low number of vias, a compact layout often led to an increase in parasites. Parasites increase the leakage current, reduce the speed, limit the gain, increase the noise, or decrease the linearity, i.e. worsen performance.

Acknowledgments

The authors would like to acknowledge Francis Trélin for all these years of work in our team. This year is our last one working together since his retirement is due soon. The authors congratulate Francis Trélin for his uncountable efforts during his career. He always sought the best layouts for each of our various circuit tape outs motivated us to write this review.

The authors would like to thank Pedro Sartori Locatelli, master student at UFMG, for his careful reading of our manuscript and their many insightful comments and suggestions.

References

[1] A. V. Mezhiba and E. G. Friedman, "Trade-Offs in CMOS VLSI Circuits," in *Trade-Offs in Analog Circuit Design*, C. Toumazou, G. Moschytz, B. Gilbert, and G. Kathiresan, Eds. Springer, Boston, MA, 2002, pp. 75–114. [Online]. Available: doi.org/10.1007/0-306-47673-8_3

[2] R. A. Rutenbar and J. M. Cohn, "Layout tools for analog ICs and mixed-signal SoCs," in *ACM Proc. Int. Symp. Physical Design*. San

Diego, CA, USA: ACM Press, May 2000, pp. 76–83. [Online]. Available: doi.org/10.1145/332357.332378

[3] S. P. Gimenez, *Layout Techniques for MOSFETs*. Morgan & Claypool, 2016. [Online]. Available: doi.org/10.2200/S00681ED1V01Y201510EET006

[4] P. Chevalier, et al., "A 55 nm triple gate oxide 9 metal layers SiGe BiCMOS technology featuring 320 GHz f_T / 370 GHz f_{MAX} HBT and high-Q millimeter-wave passives," in *Int. Electron Devices Meeting, IEDM*, San Francisco, CA, USA, Dec 2014, pp. 3.9.1–3.9.3. [Online]. Available: doi.org/10.1109/IEDM.2014.7046978

[5] P. Jespers, *The gm/ID Methodology, a sizing tool for low-voltage analog CMOS Circuits*. Boston, MA: Springer US, 2010. [Online]. Available: doi.org/10.1007/978-0-387-47101-3

[6] J. R. R. O. Martins, F. Alves, and P. M. Ferreira, "IC-Layout Render : Image rendering tool for integrated circuit layout in Python," *Open Software*, vol. pp, no. pp, pp. 1–5, 2021. [Online]. Available: doi.org/10.5281/zenodo.5618268

[7] G. Lakshminarayana and N. Jha, "High-level synthesis of power-optimized and area-optimized circuits from hierarchical data-flow intensive behaviors," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst*, vol. 18, no. 3, pp. 265–281, Mar 1999. [Online]. Available: doi.org/10.1109/43.748157

[8] S. Roy, et al., "Towards Optimal Performance-Area Trade-Off in Adders by Synthesis of Parallel Prefix Structures," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst*, vol. 33, no. 10, pp. 1517–1530, Oct 2014. [Online]. Available: doi.org/10.1109/TCAD.2014.2341926

[9] L. Cao, S. J. Bale, and M. A. Trefzer, "Multi-Objective Digital Design Optimization via Improved Drive Granularity Standard Cells," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 11, pp. 4660–4671, Nov 2021. [Online]. Available: doi.org/10.1109/TCSI.2021.3109239

[10] H. Geng, et al., "High-Speed Adder Design Space Exploration via Graph Neural Processes," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst*, vol. pp, no. pp, pp. 1–14, Sep 2021. [Online]. Available: doi.org/10.1109/TCAD.2021.3114262

[11] F. S. Ayatollahi, et al., "AMPS: An Automated Mesochronous Pipeline Scheduler and Design Space Explorer for High Performance Digital Circuits," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. pp, no. pp, pp. 1–12, Jan 2022. [Online]. Available: doi.org/10.1109/TCSI.2021.3138139

[12] C. Solomon, "Switched-capacitor filters: precise, compact, inexpensive," *IEEE Spectrum*, vol. 25, no. 6, pp. 28–32, Jun 1988. [Online]. Available: doi.org/10.1109/6.4561

[13] S. Ma, et al., "A Discrete-Time Audio ΔΣ Modulator Using Dynamic Amplifier With Speed Enhancement and Flicker Noise Reduction Techniques," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 2, pp. 333–343, Feb 2020. [Online]. Available: doi.org/10.1109/JSSC.2019.2941540

[14] J. Xu, et al., "Noise Optimization Techniques for Switched-Capacitor Based Neural Interfaces," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 5, pp. 1024–1035, 2020. [Online]. Available: doi.org/10.1109/TBCAS.2020.3016738

[15] F. Ferreira, P. M. Ferreira, and S. Binsfeld Ferreira, "A Flexible Low-Cost Discrete-Time Wake-up Receiver for LoRaWAN applications," in *2021 19th IEEE International New Circuits and Systems Conference (NEWCAS)*, no. 2. IEEE, Jun 2021, pp. 1–4. [Online]. Available: doi.org/10.1109/NEWCAS50681.2021.9462784

[16] M. McNutt, S. LeMarquis, and J. Dunkley, "Systematic capacitance matching errors and corrective layout procedures," *IEEE Journal of Solid-State Circuits*, vol. 29, no. 5, pp. 611–616, May 1994. [Online]. Available: doi.org/10.1109/4.284714

[17] D. Khalil, et al., "Evaluation of Capacitor Ratios in Automated Accurate Common-Centroid Capacitor Arrays," in *Sixth International Symposium on Quality of Electronic Design (ISQED'05)*. IEEE, 2005, pp. 143–147. [Online]. Available: doi.org/10.1109/ISQED.2005.54/

[18] C. F. Soares and A. Petraglia, "Capacitance ratio approximation in SC filters via genetic algorithm," in *IEEE Int. Symp. Circuits Systems*, 2008, pp. 516–519. [Online]. Available: doi.org/10.1109/ISCAS.2008.4541468

[19] C. F. Soares, A. C. Mesquita Filho, and A. Petraglia, "Optimizing capacitance ratio assignment for low-sensitivity SC filter implementation," *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 3, pp. 375–380, 2010. [Online]. Available: doi.org/10.1109/TEVC.2009.2025033

[20] C. F. Soares, A. Petraglia, and G. S. De Campos, "Methodologies for evaluating and measuring capacitance mismatch in CMOS integrated circuits," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 64, no. 2, pp. 101–105, 2017. [Online]. Available: doi.org/10.1109/TCSII.2016.2554758

[21] C.-W. Lin, et al., "Mismatch-Aware Common-Centroid Placement for Arbitrary-Ratio Capacitor Arrays Considering Dummy Capacitors," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 12, pp. 1789–1802, Dec 2012. [Online]. Available: doi.org/10.1109/TCAD.2012.2204993

[22] E. Maricau and G. Gielen, "Computer-Aided Analog Circuit Design for Reliability in Nanometer CMOS," *IEEE Trans. Emerg Select Topics Circ. Syst.*, vol. 1, no. 1, pp. 50–58, Mar 2011. [Online]. Available: doi.org/10.1109/JETCAS.2011.2135470

[23] X. Pan and H. Graeb, "Microelectronics Reliability Reliability optimization of analog integrated circuits considering the trade-off between lifetime and area," *Microelectron. Reliab.*, vol. 52, no. 8, pp. 1559–1564, 2012. [Online]. Available: doi.org/10.1016/j.microrel.2011.09.026

[24] P. M. Ferreira, H. Petit, and J.-F. Naviner, "A synthesis methodology for AMS/RF circuit reliability: Application to a DCO design," *Microelectron. Rel.*, vol. 51, no. 4, pp. 765–772, Dec 2010. [Online]. Available: doi.org/10.1016/j.microrel.2010.11.002

[25] P. M. Ferreira, H. Petit, and J.-F. Naviner, "A New Synthesis Methodology for Reliable RF front-end Design," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2011, pp. 1–4. [Online]. Available: doi.org/10.1109/ISCAS.2011.5938204

[26] C. M. D. Costa and P. Baltus, "Design Methodology for Industrial Internet-of-Things Wireless Systems," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 5529–5542, Feb 2021. [Online]. Available: doi.org/10.1109/JSEN.2020.3031659

[27] T. C. Carusone, D. A. Johns, and K. W. Martin, *Analog Integrated Circuit Design*, 2nd ed. Danvers, MA: John Wiley & Sons, Inc., 2012.

[28] A. Cathelin, B. Martineau, and N. Seller, "Deep-Submicron Digital CMOS Potentialities for Millimeter-wave Applications," in *IEEE Proc. of RFIC*, no. 1, 2008, pp. 8–11. [Online]. Available: doi.org/10.1109/RFIC.2008.4561384

[29] H. Zhang, X. Fan, and E. S. Sinencio, "A Low-Power, Linearized, Ultra-Wideband LNA Design Technique," *IEEE J. Solid-State Circuits*, vol. 44, no. 2, pp. 320–330, Feb 2009. [Online]. Available: doi.org/10.1109/JSSC.2008.2011033

[30] C. Liang and B. Razavi, "A layout technique for millimeter-wave PA transistors," in *IEEE Radio Frequency Integr. Circuits Symp.* Baltimore, MD, USA: IEEE, Jun 2011, pp. 1–4. [Online]. Available: doi.org/10.1109/rfic.2011.5946225

[31] A. L. T. Costa, H. Klimach, and S. Bampi, "High linearity 24 dB gain wideband inductorless balun low-noise amplifier for IEEE 802.22 band," *Analog Integrated Circuits and Signal Processing*, vol. 83, no. 2, pp. 187–194, Feb 2015. [Online]. Available: doi.org/10.1007/s10470-015-0531-1

[32] S. Rathore and A. D. Darji, "Minimum area layout of high bandwidth folded cascode operational amplifier using SCL 180nm CMOS technology," in *IEEE Int. Conf. on Inf., Commun., Instrum. and Control*, Indore, India, aug 2017, pp. 1–6. [Online]. Available: doi.org/10.1109/ICOMICON.2017.8279035

[33] M. Lee and C. Park, "A 2.4-GHz VCO using a symmetric layout technique to minimize mismatches in differential signals," in *Int. Conf. Inf. Networking*. Chiang Mai, Thailand: IEEE, Jan 2018, pp. 583–585. [Online]. Available: doi.org/10.1109/ICOIN.2018.8343185

[34] G. Guitton, et al., "Design of CMOS LNA with the Inversion Coefficient," in *IEEE New Circuits Syst. Conf.* Montreal, QC, Canada: IEEE, Jun 2018, pp. 182–186. [Online]. Available: doi.org/10.1109/NEWCAS.2018.8585633

[35] N. T. Tasneem and I. Mahbub, "Design of A 52.5 dB Neural Amplifier with Noise-Power Trade-off," in *IEEE Int. Symp. Midwest Circuits Syst.* Dallas, TX, USA: IEEE, Aug 2019, pp. 921–924. [Online]. Available: doi.org/10.1109/MWSCAS.2019.8885224

[36] M. Moezzi and M. Sharif Bakhtiar, "Design of LC Resonator for Low Phase Noise Oscillators," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 2, pp. 169–180, Feb 2016. [Online]. Available: doi.org/10.1109/TCSI.2015.2512704

[37] R. Dumont, et al., "A Low-Noise mm-Wave Injection-Locked Oscillator designed in 65nm Partially Depleted SOI CMOS Technology," in *IEEE New Circuits Syst. Conf.*, Toulon, France, Jun 2021. [Online]. Available: doi.org/10.1109/NEWCAS50681.2021.9462758

[38] P. M. Ferreira, et al., "Fully Integrated Interferometry-Based Reflectometer for High-Impedance Instrumentation," *IEEE Trans. Microw. Theory Techn*, vol. 66, no. 8, pp. 3901–3908, Aug 2018. [Online]. Available: doi.org/10.1109/TMTT.2018.2831699