Chapter 2

Literature Survey

2.1 Introduction

The design of a complex analog macro-block like a phase-locked loop (PLL) or an analog-to-digital converter (ADC) or filters can be typically decomposed into smaller basic cells *e.g.* comparators, operational amplifiers, *etc.*, each of which is an interconnection of basic building blocks such as differential input stage, current sources/sinks, current mirrors, current and voltage references, differential pairs, output buffers, *etc.* Each of these basic building blocks can be designed in various possible architectures. Each of these architectures has its distinct features and its limitations. In this chapter, the existing literature has been explored to identify the gaps and hence the scope for the present work.

2.2 Basic Building Blocks

Some of the important building blocks of analog IC have been discussed in this section.

2.2.1 Differential Input Stage Amplifier

A differential pair is widely used as the input stage of the operational amplifiers. The configuration of a CMOS differential pair is shown in Fig. 2.1. It is made of two transistors with their source in common, fed by a current source. The transistors may either be n-channel, as shown, or p-channel and they are matched

to each other. The main function of the differential stage is to amplify the differential input signal and reject any common-mode component.

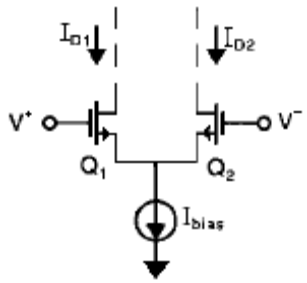

Figure 2.1 A CMOS Differential Pair.

Other design architectures of differential amplifier *e.g.* for current sensing [Bird89], for sensing amplifier in memory [Matsuda79], self-biased configuration for high speed applications [Bazes91], fully differential amplifier for large bandwidth [Ahn02], and optimization methods for achieving high CMRR [Areny91], increasing common-mode input-voltage range [Huijsing85, Bazes90, Wassenaar94] and stability [Brooks93], improved frequency compensation [Nelson93, Ho98], output swing stabilization [Johnson95, Suzuki09].

2.2.2 Current Sources/Sinks

A current sink and current source are two terminal components whose current at any instant of time is independent of the voltage across their terminals. The current of a current sink or source flows from the positive node, through the sink or source to the negative node. The simplest source/sink is a pMOS/nMOS transistor connected between V_{DD}/V_{SS} and the output node. It suffers from two limitations, one output resistance, which ought to be more. Secondly, the limited range of output voltage over which it works properly. Another implementation is the Cascode current source/sink, which has an improved value of output resistance [Allen87].

2.2.3 Current Mirrors

A current mirror is used to generate a replica (if necessary, it may be attenuated or amplified) of a given reference current. It uses the principle that if two identical MOS transistors are operating in their saturation region and their gate-source potentials are equal, then the channel currents should be equal. Electrically, a current mirror is a current controlled current source (CCCS). In real circuits, current mirrors are not able to accomplish the function of a CCCS, exactly. The current gain factor can only be positive while the output impedance, the dynamic range and the speed are finite. Moreover, the current to be copied is not measured ideally as it would be necessary to show a short circuit. Instead, to measure the reference current, a diode connected MOS transistor is normally used. The most commonly used circuits that accomplish the current mirror function are [Maloberti03]:

- Simple current mirror
- Wilson current mirror
- Improved Wilson current mirror [Palumbo93]
- Cascode current mirror
- Modified cascode current mirror
- High compliance current mirror
- Regulated cascode current mirror

All these circuits may deviate from ideal behaviour due to imperfect geometrical matching, technological parameter mismatch and parasitic resistances.

2.2.4 Current and Voltage References

Most of the basic building blocks use a current reference or voltage reference. The current value controls the trans-conductance of transistors and in turn,

influences the static and dynamic properties of the circuit in addition to the power consumption. Voltage reference is required to provide additional intermediate bias voltages. To avoid extra pins, designer prefers to generate them on chip. The characteristics of references are identical to those of the ideal voltage and current source. The references have more precision and stability than ordinarily found in a source. A reference is typically dependent upon the load connected to it, which possibly can be isolated by using a buffer amplifier in between and the high performance of the reference could be maintained. Therefore, an important designer task is to endow the master currents and bias voltages of the system with the required accuracy and independent of supply voltage and temperature. Further, a high performance voltage reference can be used to implement a high performance current reference.

A very crude voltage reference can be made from a voltage divider between the two power supplies. Two possible implementations are — one using nMOS transistors in diode arrangement and second by using both nMOS and pMOS transistors in diode arrangement [Maloberti03]. Unfortunately, the voltage divider also transfers to its output almost the same fraction of the noise and disturbances present on the supply lines. There are more design techniques used to generate an accurate and precise voltage reference. The accurate voltage elements made available by the CMOS technology are:

- The difference between the threshold voltages of an MOS transistor.
- The base-to-emitter voltages (V_{BE}), of a parasitic bipolar transistor.
- The thermal voltage, $V_T = kT/q$.

These three techniques can be used to design three different voltage references. Voltage reference based on threshold difference is temperature-independent and has a remarkable quality. But, it requires a technology that provides transistors of the same type and with different thresholds, which is rare and very expensive. The second voltage reference based on V_{BE} has a negative temperature coefficient

while the third voltage reference based on V_T has a positive temperature coefficient. Another type of voltage reference in the band-gap voltage reference that is able to keep the temperature coefficient to zero or near to zero in the given temperature range. It operates on the principle of compensating the negative temperature coefficient of V_{BE} with the positive temperature coefficient of the thermal voltage, V_T.

2.2.5 Output Buffers

All intermediate stages, which are used in op-amp or comparator before the output buffer, have the common characteristics of large output resistance to get a large voltage gain. A large output resistance can be undesirable when the load consists of a small resistance and/or large capacitance, where as a small load resistance requires a large current in order to provide a large output voltage swing. A large load capacitor requires large output currents to supply charging currents needed to meet transient response requirements. In order to provide a sufficient output current on a steady state or transient basis, it is necessary to use a low-resistance output buffer.

The primary objective of the CMOS output buffer is to function as a current transformer. Most output buffers have a high current gain and a low voltage gain. The specific requirements of an output buffer are:

- To provide sufficient output power in the form of voltage or current.
- To avoid signal distortion.
- To be efficient.
- To provide protection from abnormal conditions (short circuit, over temperature, *etc.*).

The first requirement is to meet the primary objective. The second requirement results from the fact that the signal swings are large and the non-linearities that

normally are not encountered in small-signal amplifiers will become important in output amplifiers. The third requirement is born out of the need to minimize power dissipation in the driver transistors themselves compared with that dissipated in the load. The fourth requirement is normally met with CMOS output stages since MOS devices are, by nature, self-limiting.

There are several approaches to implement the output buffer. The simplest one is an inverter with active load. It gives the best output dynamic range but an asymmetrical output driving capability. Other configurations used are cascode, and cascode with cascode load. The other configurations that give lower output resistance and keep some control on output resistance linearity are source follower and push-pull output buffers.

2.3 Basic Cells

The above-mentioned basic building blocks can be assembled to design basic cells *i.e.* operational amplifiers (Op-Amps) and comparators.

2.3.1 Operational Amplifiers

Operational amplifiers, usually referred to as op-amps, are key elements in analog processing elements. Ideally they perform the function of a voltage controlled current source with an infinite voltage gain. Only those circuit implementations that are specifically used to achieve the op-amp function in CMOS integrated VLSI systems are of relevance here. When used inside an integrated architecture, op-amps are mainly employed to drive capacitive loads, namely gates of transistors, capacitors or arrays of capacitors. This makes the request of having a low output impedance of little importance. Therefore, very often, op-amps are replaced by Operational Trans-conductance Amplifiers (OTAs) whose output impedance is quite high.

A large number of architectures for op-amps and OTAs are available in the literature *e.g.* Miller compensated two-stage op-amp, RC compensated two-stage op-amp that use two stages first the differential stage and second is the simple CMOS inverter. Other two stage op-amps include OTA op-amp, two-stage cascoded op-amp, Folded-Cascode op-amp [Hershenson98, Allen87], Active-cascode op-amp [Gray01], Symmetrical CMOS OTA [Laker94], uncompensated two-stage CMOS op-amp [Gregorian99] which is derived from its bipolar counterpart [Widlar69], two-stage op-amp with cascode differential stage [Gregorian99], improved uncompensated CMOS op-amp [Gregorian79]. Several of single stage op-amps such as simple OTA op-amp [Hershenson98], telescopic cascode, and mirrored cascode have been reported in [Maloberti03]. Other CMOS configurations reported are Cascode CMOS, cascode symmetrical CMOS OTA, symmetrical Miller CMOS OTA with high PSRR [Laker94], CMOS OTA with high PSRR [Steyaert90], folded cascode CMOS OTA [Laker94, Maloberti03]. Fully differential op-amps have been reported by several authors [Johns97, Laker94, Maloberti03, Razavi01, Gray01]. [Sedra87] has reported an Operational Current Amplifier (OCA) for differential current amplification.

2.3.2 Comparators

A comparator is the basic component mainly used in analog-to-digital converters. A comparator is the circuit that compares an analog signal with another analog signal, and outputs a binary signal based upon the comparison. There can be two possible types of input — voltage or current. In the former case, the input voltage is measured with respect to a given reference level. Therefore, the comparator determines whether the amplitude of the input is higher or not than a reference. When current is the input variable, the comparator determines whether the input current is flowing in or out of the input terminal.

One of the simplest practical implementation of CMOS comparator circuit is the current-sink inverter. The gain of this comparator implementation is low, but

could be improved substantially using cascode techniques. The major drawback is not the gain, but the fact that trip point is dependent upon the power supply. In addition, there is a limited range at which the trip point can be placed while still maintaining adequate gain. These problems can be solved by using differential input scheme. The key attribute of this circuit is its ability to amplify the difference between the inverting and non-inverting inputs. As a result, the trip point of a comparator can be made independent of process and supply variations. But, it cannot be controlled perfectly due to process-dependent input-offset voltage, which can be caused by threshold voltage, geometry, and/or temperature mismatch. Therefore, neither of the two proposed circuits performs the comparator function satisfactorily. But, they can have satisfactory performance if they are combined. In this two-stage comparator [Allen87], first stage is the differential stage and second is the inverting stage. With this configuration the limitations of the two are overcome.

Often a comparator is placed in a very noisy environment in which it must detect signal transitions at the threshold point. If the comparator is fast enough (depending upon the frequency of the most prevalent noise) and the amplitude of the noise is great enough, the output will also be noisy. In such a situation, a modification on the transfer characteristics of the comparator is desired. Specifically, hysteresis is needed in the comparator. There are several ways to accomplish it, primarily by involving some form of positive feedback. The most popular method is by introducing it in the input differential stage [Allstot82]. Regenerative Comparator [Gregorian99] has been used to convert a very slowly varying input signal into an output with abrupt changes, or in a noisy environment to detect an input signal crossing a threshold level.

In precision applications, such as high-resolution A/D converters, large input-offset voltages cannot be tolerated. While systematic offset can be nearly eliminated with proper design, random offsets still remain, and are unpredictable. Fortunately, there are offset-cancellation techniques in MOS technology. The

practical implementations of an auto-zeroed comparator are given by [Yee78, Allen87, Maloberti03]. CMOS cascode comparator uses pre-charging operations for biasing and auto-zeroing. Other comparators, such as, differential comparator, op-amp comparator, switched-capacitor comparator, *etc.* have been given in the literature [Gregorian99].

For high-accuracy applications, an effective way for reducing the dc offset voltage due to the feed-through charge is to use a fully differential comparator [Martin83, Razavi92, Razavi95, Maloberti03].

2.4 Design-Synthesis Methods

Circuit synthesis is the inverse operation of circuit analysis, where the sub-block parameters, such as device sizes and bias values, are given and the resulting performance of the overall block is calculated, as is done in SPICE. Synthesis in the analog domain refers to the automatic design of a circuit with a known qualitative behavior, such as amplification, to meet the performance (rather than behavioral) specifications of the problem [Ochotta98]. During synthesis, block performance is specified and values for the sub-block parameters needed to meet these specifications have to be determined. This inverse process is not a one-to-one mapping, but usually is an under-constrained problem with many degrees of freedom.

The different analog circuit synthesis methods that have been explored up till now can essentially be classified into four groups [Hershenson98].

2.4.1 Classical Optimization Methods

Classical optimization methods, such as steepest descent, sequential quadratic programming and Lagrange multiplier methods, have been widely used in analog circuits CAD. The general-purpose optimization codes NPSOL [Gill86] and

MINOS are used in [Maulik92, Chang92]. Other CAD methods based on classical optimization methods, and extensions such as a minimax formulation, include OPASYN [Koh90], OAC [Onodera90], and STAIC [Harvey92]. These classical methods can be used with complicated circuit models, including full SPICE simulations in each iteration, as in DELIGHT.SPICE [Nye88].

The main disadvantage of classical methods is that they only find locally optimal designs. This means that it is possible that some other set of design parameters, far away from the one found, may result in a better design. The same problem arises in determining feasibility — they may fail to find a feasible design, even if one exists. In order to avoid local solutions, the minimization method is carried out from many different initial designs. This increases the likelihood of finding the globally optimal design but it also destroys one of the advantages of classical methods, *i.e.* speed, since the computation effort is multiplied by the number of different initial designs that are tried. It also requires human intervention to give good initial designs, which makes the method less automated. Also, these methods become slow if complex models are used, as in DELIGHT.SPICE [Nye88].

2.4.2 Knowledge-Based Methods

Knowledge-based and expert-systems methods have been widely used in analog circuits in late 1980s. Specific heuristic design knowledge about the circuit topology under design was encoded explicitly in some computer executable form that was then executed during the synthesis run for a given set of input specifications to obtain the design solution. The knowledge was encoded in different ways in different systems.

The IDAC tool [Degrauwe87] used manually derived and prearranged design plans or design scripts to carry out the circuit sizing. The design equations specific for a particular circuit topology had to be derived and the degrees of

freedom in the design had to be solved explicitly during the development of the design plan using simplifications and design heuristics. The big advantage of using design plans is their fast execution speed, which allows for fast performance space explorations [Harjani89]. The big disadvantages are the lack of flexibility and the large time needed to develop a plan for each topology and design target (typically 4 times than needed to actually design the circuit once [Beenker93]), as analog design heuristics are very difficult to formalize in a general and context-independent way.

OASYS [Harjani89] adopted a similar design plan based sizing approach, but explicitly introduced hierarchy in the design of analog circuits and also added a heuristic approach towards topology selection to the system. Hierarchy allowed to reuse design plans of lower-level cells while building up higher-level cell design plans, and therefore also leveraged the number of device-level schematics covered by one top-level topology template. Collecting and ordering all the design knowledge in the design plan however still remained a time-consuming job. The approach was later on adopted in the commercial MIDAS system [Beenker93].

The other ways to encode the knowledge have been explored as well, such as in BLADES [Turky89] which is a rule-based system to size analog circuits and DARWIN [Kruiskamp95] which uses genetic algorithms or evolution system.

2.4.3 Global Optimization Methods

Global optimization methods such as branch-and-bound [Xinghao96] and simulated annealing [Laarhoven87, Wong88] have also been used, *e.g.*, in [Maulik95]. Branch-and-bound unambiguously determines the global optimal design but it is extremely slow, with computation growing exponentially with problem size. Global Simulated Annealing (SA) is another very popular method that can avoid becoming trapped in a locally optimal design. In principle it can compute the globally optimal solution, but in practical implementations there is

no guarantee at all as termination is heuristic. Like classical and knowledge-based methods, SA allows a very wide variety of performance measures and objectives to be handled. Simulated annealing has been used in several tools such as ASTR/OBLX [Ochotta96] and OPTIMAN [Gielen90, Gielen91]. The main disadvantage of SA is that it can be very slow, and in practice it cannot guarantee a global optimal solution.

2.4.4 Convex Optimization and Geometric Programming Methods

In a convex optimization problem we minimize a convex objective function subjected to linear equality constraints, and inequality constraints that are expressed as upper bounds on convex functions. The great practical advantage of convex optimization is beginning to be widely appreciated, mostly due to the development of extremely powerful interior-point methods for general convex optimization problems in last five years [Nesteroy94, Wright97]. These methods can solve large problems, with thousands of variables and tens of thousands of constraints, very efficiently. One great advantage of convex optimization, compared to general-purpose optimization methods, is that the global solution is always found, regardless of the starting point and infeasibility is also unambiguously detected. One of the disadvantages is that the types of problems, performance specifications, and objectives that can be handled are far more restricted than any of the methods described above.

Geometric Programming (GP) is a special type of convex optimization problem, which is recently being used for transistor and wire sizing in digital circuits [Fishburn85, Shyu88, Sapatnekar93, Sapatnekar96] and in analog circuits [Hershenson01, Mandal01]. But the current GP transistor models support the traditional long channel MOS transistors and the GP models for submicron devices are still under development.

2.5 Concept of Figure of Merit

The concept of Figure of Merit to estimate the performance in terms of power, propagation delay and area started in digital circuits and later was adopted by analog circuit designers as well. In analog domain Figures of Merit have been defined for four basic circuits [Brederlow01] Low Noise Amplifiers (LNA) [Yoshida06], Voltage Controlled Oscillators (VCO) [Muer00], Power Amplifiers (PA) and Analog-to-Digital Converters (ADC). The concept of figure of merit has also been applied to evaluate the performance of instrumentation amplifiers [Menolfi99], to study the performance in terms of small signal and large signal behavior of driver/output stage amplifiers [Leung00, Ramos02], for the linearity test of differential amplifiers [Nam03], to study testability of mixed-signal circuits and systems [Soma01], for architecture selection in active filter designs, *etc*. The concept has been widely used in performance evaluation, comparison and selection of ADCs [Walden99, Vogels03], to design of sigma-delta ADC [Rabii97, Marques98], flash ADC [Yoshii84], asynchronous ADC [Allier05] and speed-power tradeoffs of high speed ADC [Uyttenhove02] and many more.

The use of figure of merit is also seen in literature to evaluate the performance of RF CMOS circuits [Woerlee01, Walden99a].

2.6 Gaps Identified

The existing literature in books and journals was explored. So far mostly it is said that for analog circuits, there is no area constrain and only the performance is the requirement. However, there are per pixel processing applications, where area is becoming a major constraint. For example, consider a camera with only a resolution of 4 Mega-pixels. It requires an IC, which captures the image/signal. We require an array of 2K $\times$ 2K preprocessing circuits one for each pixel. If we assume that the size of IC is such that 1 cm^2 area is available for these circuits, we get only 1 cm^2/4Mega = 10^8 $\mu m/4 \times 10^6$ = 25 μm^2 for the processing of a pixel!

Some researchers have indicated sharing of resources between columns or group of 4 pixels, *etc.* [Pain93, Espejo94, Fowler95, Mendis97, Fossum97, Yang99, Kleinfelder01, Bermak02, Lin04, Dudek05, Eltoukhy06].

It was found that there had been no attempt so far to design and optimize analog integrated circuits in the area-constrained environment. And hence there is a scope to work in this direction.

Another application domain is the night vision cameras employed for strategic applications, where signals to be captured are normally weak and noise becomes very significant. And therefore, noise is one of the significant parameters, which can be taken along with the area-constraint.

Hence, this gap sets the exploration path and motivation for the present research work.