Introduction

This paper reviews several interesting works in CMOS amplifier design, bandwidth extension for CMOS amplifier and two analog building blocks implemented by CMOS amplifiers. It is assumed that the reader is familiar with “amplifier biasing circuits”, “MOS device characteristics” and genetic programming (GP).

Initially, we will review the design of operational amplifiers. They are composed of a differential amplifier, a gain and an output stage. The differential amplifier offers a variety of advantages and is always used as the input to the overall amplifier. Since it provides a common mode rejection, it eliminates noise common on both inputs, while at the same time amplifying any difference between the inputs. It also provides gain. The gain stage is typically a common source or cascade type amplifier. The output stage provides high current driving capability for either driving large capacitive or resistive loads. The output stage typically will have low output impedance and high signal swing characteristics. Figure 1 presents a simple operational amplifier. The biasing circuit is a simple current mirror driver. [1]

CMOS amplifier design using GP:

A large number of systems are developed to fulfill automated analog circuit design. In most systems, there is a circuit generator to synthesize the first draft of a circuit. Then circuit parameters are optimized to meet the target specifications. They are mostly done based on “try and test” method, which provides the practical and optimized circuits only if relevant information exists in the systems. Using GP, circuits can be synthesized without the knowledge of required topologies. Therefore, GP can be used for topology selection, invention and parameter optimization. Using GP to design CMOS amplifiers is demonstrated in [2] in which a circuit is represented by a tree structure. The tree is the construction of evolving functions which evolve a single wire into a complex structure of components (The available components are wire, floating connection, resistor, capacitor, inductor, voltage and current independence source, N-type and P-type transistors.). To design a circuit, the system requires a number of startup configurations from the user, which are circuit specifications, technological parameters, and embryonic circuits. The topological evolution starts from an embryonic...
circuit, which generally contains a wire connecting input to output. The evolving is done using connection modifying functions. They can provide parallel, serial, cross-linked connections. They may swap the connections or even create a gain-stage. Knowledge is provided by the user by adding known basic modules as additional information for the evolution. The basic modules that are used for the evolution of CMOS amplifiers are shown in figure 2.

Then current-flow analysis is done which translates a GP-tree representing an evolving circuit into current-flow lists for each circuit. The lists describe the connection of components based upon the current flowing within the circuit. Correction rules are performed on the current-flow lists to alter the connection of components, remove isolated parts, or remove any parts which do not affect performance. Once the current-flow lists are created, a number of analyses can be done to obtain some characteristics, including the direction of current, floating list, isolated list and transistor region. The remaining circuits are then evaluated for fitness values, which are divided into two groups of hard constraints and soft constraints. The hard constraints are the user specifications for the desired circuit. They are required to meet the target for the evolution to be considered a success. The soft constraints have a lower priority. They provide better optimized circuits.

For the evolution of CMOS amplifiers, the hard constraints include the following: open-loop gain, gain bandwidth product, phase margin, dc-offset, power dissipation, voltage swing, linearity penalty. The soft ones are CMOS transistors violation penalty and complexity. A number of candidates are then selected from evaluated circuits for refining performance. The process changes one of the component parameters at a time. It then evaluates the fitness value and compares the altered circuit with the original and selects the better one. Improved candidates are then inserted back into the generation and sorted. If no circuit meets the target specifications, new generations are then evolved from the current generation. The process runs until all necessary specifications are satisfied.

Bandwidth extension for CMOS amplifiers:

CMOS is a good solution for communication trends that transmit and receive data at high speeds with low error rates, power and cost. Some communication circuits demand the broadband amplifiers but CMOS parasitic limits the performance of broadband amplifiers. Therefore it motivates the use of bandwidth extension technique. Here we review a conventional technique and the improved version which are proposed in [3]. Shunt peaking is a bandwidth extension
technique in which an inductor connected in series with the load resistor shunts the output capacitor (figure 3.a). Treating the transistor as a small-signal dependent current source, the gain is simply the product of \( Z(s) \) and \( g_m \). As \( g_m \) is approximately constant, only \( Z(s) \) is considered hereafter. For the shunt-peaked network:

\[
Z(s) = \frac{V_{out}}{I_{in}} = \left( \frac{1}{sC} \right) \frac{R + sL}{1 + sRC + s^2LC}
\]

The inductor introduces a zero in that increases the impedance with frequency, compensates the decreasing impedance of \( C \), and thus extends the 3 dB bandwidth. An equivalent explanation for increased bandwidth is reduced risetime. That is, the inductor delays current flow to the resistive branch so that more current initially charges which reduces risetime.

Substituting the -3 dB bandwidth of the reference common source amplifier and normalizing to the impedance at DC gives:

\[
Z_N(s) = \frac{1 + s/m\omega_0}{1 + s/\omega_0 + s^2/m\omega_0^2}, \quad \omega_0 = 1/RC, \quad m = R^2C/L
\]

For shunt peaking, \( m=1.41 \) gives the maximum BWER of 1.84. This extension comes with 1.5 dB of peaking. A maximally flat gain is achieved for \( m=1+1.41 \) but BWER is reduced to 1.72. Although the increased impedance of the inductor accounts for the bandwidth improvement, it also leads to peaking in the response. To eliminate peaking with maximum BWER, in shunt with the inductor a capacitor is added that should be large enough to negate peaking but small enough to not significantly alter the gain response. A common-source amplifier incorporating such a bridged-shunt network is shown in figure 3.b. Compared to previous \( Z_{N}(s) \), it introduces another pole and zero. A BWER of 1.83 is achieved with a flat gain response, in contrast to the shunt-peaked design with a nearly identical BWER of 1.84 but 1.5 dB of peaking. The advantage of bridged-shunt peaking over shunt peaking is that the maximum bandwidth is achieved for a larger value of \( m \), which translates to a smaller inductance with smaller die area, higher self-resonant frequency, etc.

A CMOS fully balanced differential difference amplifier:

A versatile analog building block is differential difference amplifier (DDA). The DDA is a five-terminal device, with two differential input ports, \((V_{pp} - V_{pn})\) and \((V_{np} - V_{nn})\). The output of the DDA can be expressed as: \( V_o = A_o[(V_{pp} - V_{pn}) - (V_{np} - V_{nn})] \). Analogous to the traditional op-amp, when a negative feedback is applied, the differential voltages of the two input ports become equal “\( V_{pp} - V_{pn} = V_{np} – V_{nn} \)” as \( A_o \) goes to infinity. Also, like the op-amp, the DDA consists
mainly of two stages: a differential pair with active loads and a second gain stage. In DDA design, there is a tradeoff between achieving high input differential range, low noise, and low distortion. In [4] a CMOS fully balances version of DDA (FBDAA) is presented, which does not suffer from this problem. A fully differential architecture of the DDA can be designed in much the same way as the conventional op-amp. This results in fully balanced outputs: \(V_{op} = -V_{on} = A_0 \left[ (V_{pp} - V_{pn}) - (V_{np} - V_{nn}) \right]\). The proposed FBDDA includes the common-mode feedback circuit to establish the common-mode output voltage, and without it the common-mode voltage output would drift. It determines the output common-mode voltage and controls it such that it is equal to some specified voltage even with the presence of large differential signals. Two similar class AB rail-to-rail output stages are incorporated to achieve well-determined low standby power consumption with good output current driving capability. Both outputs of FBDDA are fed back to two of the inputs. The negative feedbacks may be applied to the same differential pair or the feedbacks to both differential pairs. The first one behaves similar to the DDA since the differential pair which has no feedback applied to it, will not exhibit a virtual short between its inputs, whereas, the second will behave like an op-amp because the feedbacks are applied to each of the two differential pairs, and hence each pair will exhibit virtual short between its inputs. Therefore, both differential pairs will operate near \(V_d = 0\) and \(g_{mn} (W/L)_a\) can be designed as large as desired to achieve low noise, high input range, and low distortion simultaneously. Then several essential op-amp based circuits; the voltage buffer, the non-inverting amplifier, and the state-variable filter are designed using the FBDDA. All of the circuits were tested and found to perform as expected.

**A temperature-stable CMOS variable gain amplifier:**

Variable-gain amplifier is used for controlling the transmission signal power or for adjusting the received signal amplitude. In [5], a CMOS VGA using subthreshold exponential region transistors is proposed. A subthreshold region MOS transistor is suitable because there is no gate current and a simple control circuit is available. Also a master–slave (M/S) control feedback technique is applied. As shown in figure 4.a, the test chip comprises an (intermediate frequency) IF VGA and a quadrature demodulator (QDEM). Each stage has a pair of variable \(g_m\) elements. In the first two stages, transistors biased in the square-law region are used as the variable “\(g_m\)” elements. Transistors biased in the subthreshold exponential region are used in the following two stages and a mixer input stage of the QDEM. The strategy for design of a gain-control circuit is: 1) To achieve a wide linear-in-dB control range, \(M1\) (in figure 4.b) biased in a subthreshold exponential region is used for exponential voltage-to-current conversion in the control and bias block shown in figure 4.a.
2) Temperature dependence of the gain-control characteristic is compensated by the voltage converter which converts the gain-control voltage to the gate voltage of the transistor M1. 3) To achieve a temperature-stable gain at Vref and stable gain-control ratio dB simultaneously, a differential architecture is adopted for the voltage converter, and the common-mode output voltage and differential-mode gain are independently controlled.

![Figure 4.a. Test chip](image1)

![Figure 4.b. Gain control circuit with M/S control](image2)

![Figure 4.c. Basic circuit and amplifier](image3)

Note that the square-root function is transformed to a linear function by the logarithmic conversion, i.e., \( \log \sqrt{x} = 0.5 \log x \).

As shown in figure 4.b. (M/S) control technique is used for gain-control circuit. The master circuit generates bias voltage, which controls the gain of the voltage converter in the slave circuit. So, the gain is the same as the gain of the master converter. Both voltage converters in the master and slave circuits have a common-mode feedback circuit. The output voltage of the master voltage converter is exponentially converted to an output current by the exponential region transistors M2 and M3. If the reference voltage is temperature insensitive, “Icont” will be insensitive to the temperature. It means that the gain-control ratio \( \Delta dB/\Delta Vc \) is insensitive to the temperature when the amplifier gain is proportional to the bias current. Since the exponential amplifiers (in last two stages of figure 4.a) and the square-law amplifiers (in first two stages of figure 4.a) have different temperature dependence, two individual bias and control circuits are required to stabilize the entire gain. Two bias circuits with the same structure but different temperature dependence are employed, as shown in figure 4.c. When M5 and M6 are in exponential region, the bias circuit operates as a \( V_T \) reference circuit. \( V_T \) is proportional to the absolute temperature. In amplifier stage the output current (M8 drain current) is also proportional to the absolute temperature. Therefore, the gain which is defined as: \( G=g_{m8}*R_L=(dI_{D8}/dV_{GS8})*R_L=K_1 R_L/R_S \), will only depend on resistance ratio and independent of temperature, (temperature dependence of \( I_{D8} \) will be compensated by \( V_T \) term in \( V_{GSS} \)). The same argument is true when transistors operate in a square-law region, when the gain is \( G=K_2* R_L/R_S \).

In the implemented circuit the IF amplifier consists of a variable \( g_m \) input amplifier and a transimpedance output buffer. When the input transistors, M9 and M10 are in an exponential region, \( g_m \) is proportional to the bias current \( I_{cont} \). When they are in a square-law region, \( g_m \) is proportional to the square root of \( I_{cont} \). In both cases, \( \log(g_m) \) is proportional to \( \log(I_{cont}) \). The transimpedance output buffer has low input and output impedances, which makes the frequency response
of the IF amplifier wide and flat. The transresistance of this buffer is approximately the same as the feedback resistance. Consequently, the amplifier gain is $g_{\text{m}} R_f$, which is the same as the gain calculated before.

References: