A Computationally Efficient Technique for the Optimization of Two Stage CMOS Operational Amplifiers

Paolo Bruschi* , Dino Navarrini*, Giovanni Tarroboiro* and Giuseppe Raffa*

Abstract — **In this paper a mixed numerical-analytical method for optimizing the design of two stage operational amplifiers is presented. The main difference with respect to similar approaches is the elimination of as much as possible degrees of freedom of the problem by means of analytical and practical considerations. A built-in MATLAB function has been used to carry out the numerical optimization. The method has been applied to a simple two stage CMOS operational amplifier topology and the results have been verified by means of electrical simulations.**

1 INTRODUCTION

The demand of systems-on-chip or systems-on-package equipped with interfaces to the sensor world has significantly increased in the last years. This has shifted back the interest of researchers and industries to the development of analog circuits.

In order to reduce the long time-to-market involved in designing analog circuits down to the transistor level, new CAD platforms devoted to re-use of currently available cells are being developed.

However, the technology evolution and the continuous scaling down of the supply voltage and current often dictates to completely re-design even the basic building blocks. Generally, as in the case of the operational amplifier, the topology that more likely will suits the requirements can be rapidly selected from a small number of options. The really time consuming phase is then transistor sizing. Several approaches to this problem have been proposed, resulting in some cases in commercial products. Besides convergence problems and the risk of finding locally optimum designs, a disadvantage of many promising methods is slowness. It should be considered that very often the starting requirements are not feasible and the designer has to repeat the optimization procedure several times, gradually relaxing the critical constraints.

Computational efficiency is therefore a key feature to make an optimization tool really useful.

A class of programs that meet this requirement is based on classical constrained optimization, applied to an analytical description of the circuit behaviour. Most circuit performance indicators (e.g. gain bandwidth product) are expressed as a function of a selected set of degrees of freedom (DOF) (typically transistor sizes and bias currents).

One of these indicators is then chosen as a target function to be maximized (or minimized). The other indicators can be used as constraints.

Among the various practical implementations, geometrical programming [1] offers interesting characteristics of convergence and immunity to local minima (or maxima). As a counterpart, it requires all the equations of the circuit to be written in the same "posynomial" form. This lack of generality can be overcome introducing *ad hoc* approximations.

The alternative approach proposed in this work exploits the efficiency of the sequential quadratic programming (SQP) algorithms which, differently from geometrical programming, can handle equations of arbitrary form. The SQP method have been applied to a basic two stage operational amplifier, using KCL and KVL and device equation to eliminate as much as possible DOFs of the system. We demonstrated that the reduction of even a single DOF (the compensation capacitor) results in up to a 100% improvement of the computational efficiency. All the parameters involved in MOSFET modeling can be easily obtained by fitting the simulated characteristics. The result is a very robust and simple tool capable of producing usable designs for a very wide range of specifications.

2 TWO STAGE AMPLIFIER MODELLING AND PERFORMANCE CALCULATION

A simplified small signal equivalent circuit, which can be used to model most two stage operational amplifiers, is shown in figure 1.

G_{m1} and G_{m2} are the equivalent transconductances of the input and output stage, respectively. R_1 and and R_2 are the output resistances of the two stages. The capacitor C_C and the zero nulling resistor R_C form the compensation network. The capacitances C_1 and C_2 can be expressed as:

$$C_1 = C_{o1} + C_{i2} \; ; \quad C_2 = C_{o2} + C_L \tag{1}$$

where C_{o1} is the output capacitance of the first stage, C_L is the load capacitance, C_{i2} and C_{o2} are the input and output capacitances of the second stage.

The input capacitance of the first (differential) stage has not been included in this simplified

* Dipartimento di Ingegneria dell'Informazione, via Diotisalvi, 2, I-56126 Pisa, Italy.
E-mail: p.bruschi@iet.unipi.it

schematization.

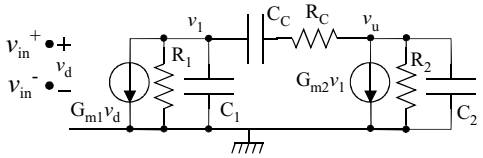

Figure 1: Simplified op-amp model.

In most practical cases the two poles of the circuit are some decades apart. We also suppose that R_C *is* made equal to $1/G_{m2}$, so the zero of the transfer function is perfectly canceled. With these assumptions, which are quite general, the unity gain angular frequency ω_0 and non dominant pole angular frequency ω_2 are given by:

$$\omega_0 = \frac{G_{m1}}{C_C}; \quad \omega_2 = \frac{G_{m2}}{C_1 + C_2}\left(1 + \frac{C_S}{C_C}\right)^{-1} \tag{2}$$

where C_S is the capacitance of the series of C_1 and C_2. In order to achieve a given phase margin we can require that:

$$\omega_2 = \sigma\omega_0 \tag{3}$$

For example, with σ=3 we get a phase margin of 71.5º. Combining equations (2) and (3) and solving for C_C we obtain:

$$C_C = \frac{\sigma}{2}\frac{G_{m1}}{G_{m2}}(C_1 + C_2)\left[1 + \sqrt{1 + \frac{4}{\sigma}\frac{G_{m2}}{G_{m1}}\frac{C_S}{C_1 + C_2}}\right] \tag{4}$$

The gain bandwidth product, given by $\omega_0/2\pi$, is than derived from C_C through Eq. (2). Note that Eq.(4) allows C_C to be removed from the degrees of freedom (DOF) of the optimization problem.

The method has been applied to the simple two stage topology of figure 2.

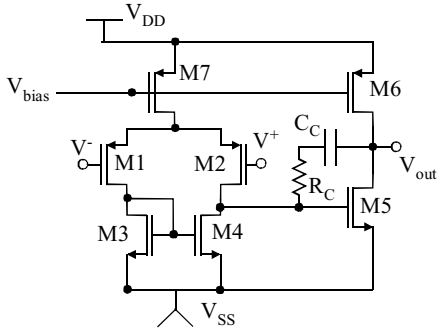

Figure 2. Schematic of the two stage amplifier.

With the DOF simplification given by Eq. (4) and the condition of exact cancellation of the zero, fifteen DOFs remain (the widths and lengths of the seven transistors and the bias voltage). Four DOFs can be eliminated considering that M1=M2 and M3=M4 for obvious symmetry requirements.

In order to achieve a good matching it is advisable to make:

$$L_7 = L_6 \quad \text{and} \quad L_3 = L_5 \tag{5}$$

A symmetrical output swing is also often required, so we get the relationship:

$$(V_{GS} - V_{tn})_5 = \left|(V_{GS} - V_{tp})_6\right| \tag{6}$$

Finally, in order to obtain a null systematic offset we impose that, in quiescent conditions (null differential input voltage), M5 and M6 are in saturation. This, together with KCL at the output node, yields:

$$K_p \frac{W_6}{L_6}\frac{(V_{GS6} - V_{tp})^2}{2} = K_n \frac{W_5}{L_5}\frac{(V_{GS5} - V_{tn})^2}{2} \tag{7}$$

where K_n and K_p are the transconductance factors ($\mu_n C_{ox}$ and $\mu_p C_{ox}$) of the *n*-MOS and *p*-MOS transistors. Eqs. (5-7) eliminate other 4 DOFs. We choose the following set of 7 DOFs to identify the amplifier:

$$W_1,\ L_1\ W_5,\ L_5, L_6, (V_{GS} - V_{tn})_5,\ (V_{GS} - V_{tp})_1$$

It can be easily shown that the remaining transistor sizes and voltages of the circuit can be explicitly derived from the selected DOF set.

The parameters of the small signal equivalent circuit of figure 1 have been expressed as a function of the transistor sizes and operating point. Clearly:

$$G_{m1} = g_{m1}, \quad G_{m2} = g_{m5} \tag{8}$$

The transconductance of the MOSFETs has been calculated using the well known formula:

$$g_m = \mu C_{ox}\frac{W}{L}(V_{gs} - V_t) \tag{9}$$

where μ is the electron (or hole) mobility and the dependence of the factor μC_{ox} on V_{GS}-V_t is fitted with a 4th degree polynomial.

The capacitors C_1 and C_2 are calculated by the following formulas that refer to the charge controlled model of the MOSFET [2]:

$$\begin{aligned} C_1 &= C_{dd2} + C_{dd4} + C_{gg5} \\ C_2 &= C_{dd5} + C_{dd6} + C_L \end{aligned} \tag{10}$$

where C_L is the load capacitance and the other terms are defined as:

$$C_{dd} = \frac{\partial Q_d}{\partial V_d}, \quad C_{gg} = \frac{\partial Q_g}{\partial V_g} \cdot \tag{11}$$

The C_{dd} and C_{gg} parameters are functions of the transistor sizes. We have adopted the simple linear approximations given by:

$$\begin{aligned} C_{dd} &= k_{dw} W \\ C_{gg} &= k_{gw} W + k_{gg} WL \end{aligned} \tag{12}$$

The constants k_{dw}, k_{gw}, k_{gg} are clearly process dependent and can be obtained by means of simulations or inspection of the transistor model file.

Finally, the resistances $R_1 = (r_{d2}^{-1} + r_{d4}^{-1})^{-1}$ and $R_2 = (r_{d5}^{-1} + r_{d6}^{-1})^{-1}$ are obtained estimating the MOSFET output resistances r_d using the classical expression:

$$r_d = \frac{\lambda^{-1}}{I_D} = 2\frac{\lambda^{-1}}{g_m(V_{gs} - V_t)} \tag{13}$$

The parameter λ is a function of L, V_{DS} and $(V_{GS}-V_t)$, fitted by the formula:

$$\lambda^{-1} = (V_{DS} - V_{DS0})^{\frac{L}{L_0}}\left(k_{\lambda 0} + k_{\lambda 1} L + k_{\lambda 2} L^2\right) \tag{14}$$

where V_{DS0} and L_0 are constant parameters and the coefficients $k_{\lambda i}$ are 2nd order polynomial functions of $(V_{GS}-V_t)$ The eleven constants, required to calculate λ, are obtained by fitting the simulated MOSFET characteristics.

The procedure of extracting the various constant used by the program can be automated by scripts that launch the required simulations, collect the data and perform the fitting procedures.

The performance of the circuit is represented by the following indicators: gain bandwidth product (f_0), area occupancy (*area*), DC gain (A_{V0}), input thermal (S_{VT}) and flicker (S_{VF}) noise density, slew rate (s_r), input offset voltage (v_{io}), quiescent dissipated power (P_d). These quantities are calculated by means of analytical functions. The formulas used for f_0, A_{V0}, s_r, *area* and P_d are given by the following equations, where all the quantities can be expressed as a function of the selected set of seven DOFs:

$$\begin{aligned} & f_0 = \frac{G_{m1}}{2\pi C_C}; \quad s_r = 2\pi f_0 (V_{GS} - V_t)_1 \\ & area = \sum W_i L_i + 2L_C \sum W_i + \frac{C_C}{k_C} \\ & A_0 = G_{m1} R_1 G_{m2} R_2; \quad P_d = (I_{D7} + I_{D6})(V_{DD} - V_{SS}) \end{aligned} \tag{15}$$

Here L_C is the minimum drain/source length required to contact them and k_C is the capacitance per unit area of the compensation capacitor.

The flicker and thermal noise density and the random offset are calculated considering only the contribution of the first stage. Classical expressions are used to calculate the contribution of transistors M1-4 to the input voltage noise. Each MOSFET is modeled as an ideal device with an additional noise voltage source in series with the gate. The corresponding spectral density is given by:

$$\text{flicker: } S_{VF}(f) = \frac{N_f}{WL}\frac{1}{f} \tag{16}$$

$$\text{thermal: } S_{VT}(f) = \frac{8}{3} k_B T \frac{1}{g_m} \tag{17}$$

where k_B is the Boltzmann constant and N_f is a process parameter. The offset voltage (standard deviation) is a simple function of the standard deviations of the threshold voltage and of the factor $\beta = \mu C_{ox} W/L$ of the input stage transistors. The latter depends on the transistor sizes and on two process constant C_{vt} and C_β through the expressions:

$$\sigma_{vt} = \frac{C_{vt}}{\sqrt{WL}}; \quad \frac{\sigma_\beta}{\beta} = \frac{C_\beta}{\sqrt{WL}} \tag{18}$$

3 EXAMPLES OF DESIGN OPTIMIZATION

The aim of the program is to optimize a single circuit performance indicator using all the other indicators as inequality constraints. At this stage it is possible to maximize the gain bandwidth product and to minimize one of the following parameters: area, flicker noise and thermal noise. The core of the optimization procedure is the MATLAB function "fmincon", based on the sequential quadratic programming algorithm[3]. This function is particularly suitable for small and medium complexity problems: convergence and computation time are greatly improved by the DOF reduction described in section 2. A series of inequalities are required to define the search range for each DOF. The program starts iterating from an initial point that should be within the search range but that does not require to satisfy the constraints.

Several tests have been performed to check the validity of the method. Convergence was achieved in most cases, for all the four possible target performance indicators, without changing the initial point. Computation times were always of the order of few seconds on a Pentium II 333 MHz platform.

Since an exhaustive evaluation of the program is still in progress, we report here two examples of automatic design, showing the gain bandwidth product and area occupancy optimization functionality of the method. The technological process used for these tests was the BCD6 of STMicroelectronics. The amplifiers resulting from the optimization procedure have been analyzed by means of the electrical simulator ELDO. The

MOSFET model used by the simulator is the Philips level 9.

The specifications imposed for the gain bandwidth and area optimization are shown in table I. In both cases, the requirement of high dc gain (considering the simple non-cascoded two stage topology), low noise and low offset voltage are in contrast with the performance indicator to be optimized and with the low dissipated power. A manual design would be difficult and time consuming.

Optimization	GBW	*area*
P_d (mW)	3	4
s_r (V/µs)	10	10
area (µm^2)	8000	min
v_{io} (mV)	3	5
GBW (MHz)	max	30
S_v thermal V^2/Hz	1.0×10^{-16}	5×10^{-17}
S_v flicker (1Hz) V^2/Hz	4×10^{-12}	5×10^{-12}
A_0 (dB)	106	109
Phase margin	70º	70º

Table I. Specifications used for the tests.

The transistor sizes and V_{GS}-V_t proposed by the program are shown in table II and table III. With these values, two ELDO netlists have been created and simulations have been run to extract the amplifier performances in both cases. The results are shown in table IV. Note that the area occupancy and offset voltage could not be simulated and have been calculated as described in the previous section.

	W (µm)	L (µm)	$\lvert V_{GS}-V_t\rvert$ (V)
M1	186.75	0.95	0.1
M3	11.45	1.45	0.28
M5	160.90	1.45	0.28
M6	619.20	1.15	0.28
M7	87.95	1.15	0.28

Table II. Results of the GBW optimization test.

The simulated performances meet the design requirements of table I except for the noise densities. This is possibly due to the fact of having neglected the noise contribution of the second stage

	W (µm)	L (µm)	$\lvert V_{GS}-V_t\rvert$ (V)
M1	140.50	0.90	0.1
M3	17.50	1.95	0.23
M5	215.10	1.95	0.23
M6	476.16	0.87	0.23
M7	77.42	0.87	0.23

Table III. Results of the *area* optimization test.

Optimization	GBW	*area*
P_d (mW)	3.12	2.28
s_r (V/µs)	26.5	19
area (µm^2)	8000	7627
v_{io} (mV)	3	3.68
GBW (MHz)	42.5	31.1
S_v thermal V^2/Hz	1.1×10^{-16}	1.25×10^{-16}
S_v flicker V^2/Hz	6.3×10^{-12}	3.9×10^{-12}
A_0 (dB)	105	107
Phase margin	59º	60º

Table IV. Amplifier performances resulting from the simulation.

4 CONCLUSIONS

The designs produced by the program have been tested with electrical simulations based on state-of-the-art device models and simulator. In most cases the performances obtained met the initial requirements, showing that, for the examined simple topology, the program is capable of proposing usable designs rather than first estimation circuits. An exception is represented by the noise density which was generally higher than the maximum value used as a constraint. This problem can be simply overcome by re-running the program with stricter specifications. The very short computational times make this operation much more convenient than manually modifying the circuit.

Acknowledgments

The authors would like to thank the R & D group of the STMicroelectronics of Cornaredo (MI, Italy) for providing the device models of the BCD6 process.

References

[1] P.Mandal, V. Visvanathan, “CMOS Op-Amp Sizing Using a Geometric Programming Formulation”, IEEE Trans. on CAD Design of Integrated Circuits and System, vol. 20, no. 1, pp. 22-38, 2001.

[2] D. E. Ward, R. W. Dutton, “A Charge-Oriented Model for MOS Transistors Capacitances”, IEEE Journal of Solid State Circuits, vol.SC-13,, no. 5, pp. 703−708, 1978.

[3] M.J.D. Powell, “Variable Metric Methods for Constrained Optimization”, Mathematical Programming: The state of the Art, (A. Backem, M. Grotschel, B. Korte) Springer Verlag, pp. 288-311, 1983.