1. Introduction

Settling time is a primary design parameter in operational transconductance amplifiers (OTAs) used for high-speed applications such as pipeline A/D converters. For scaled CMOS technologies, an OTA design methodology using $g_m/I_D$ lookup tables is proposed to minimize its power consumption [1]. A major problem in this previous method is that settling time was not included in a target specification, but was converted into crossover frequency $f_c$ with an empirical approach. In this paper, we introduce an iterative optimization sequence to design OTAs, which can achieve the target settling time with the minimum power consumptions.

2. Topology and Specification of OTA

We optimize a two-stage Miller-compensated OTA for a 10-bit 27MS/s pipeline A/D converter applicable to NTSC (National Television System Committee) and related. A circuit topology of the OTA with associated capacitors in the charge-redistribution phase of switched capacitor circuit is shown in Fig. 1, and the target specification is summarized in Table I. The optimization flow derives gate length $L$ and width $W$ for each numbered transistor as well as the minimized bias currents denoted as $I_{DB}$ and $I_{DS}$. We predetermined the capacitors, $C_{g1}$, $C_{f}$, $C_t$, and $C_{e}$, from the specification of dynamic range (noise).

3. OTA Design Optimization

$g_m/I_D$ lookup table methodology and problem overview

$g_m/I_D$ design methodology has been attracted attentions as a transistor sizing tool for low-power designs since $g_m/I_D$, the ratio of transconductance $g_m$ over drain current $I_D$, can specify all device operation regions including strong, moderate, and weak inversion regions [3]. SPICE-based lookup tables, representing transit frequency $f_t$, intrinsic gain $g_m/g_{ds}$, where $g_{ds}$ is the output conductance of a transistor, parasitic gate-to-drain capacitance $C_{gd}$ and drain-to-bulk capacitance $C_{db}$, and current density $I_D/W$ as a function of $g_m/I_D$ for various $L_s$, are employed to optimize the transistor operation region and the design parameters [1].

Settling time $t_s$ is defined by the sum of linear settling time $t_{lin}$ and slewing time $t_{slew}$, which are given by

$$t_{lin} = -\frac{k_t}{2\pi f_c}\ln\left(1 - \frac{B C_f}{C_f + C_L}\right)$$

$$t_{slew} = \frac{C_L}{2Bf_D}\left(V_{id,step} - 2\sqrt{2}\sqrt{\frac{g_m}{I_D}}\right)$$

Here, $k_t$ is the acceleration coefficient calculated from the phase margin and dynamic settling error $\varepsilon_d$, and $B$ is the closed-loop feedback factor equal to $C_f/(C_f+C_t+C_{g1})$ where $C_{g1}$ is the total gate capacitance of input transistor $M_1$; moreover, $C_t$ and $V_{id,step}$ are compensation capacitor and differential input step voltage, respectively [1].

In the typical optimization flow to minimize the power consumption, $C_{g1}$ and $C_{g2}$, the total gate capacitance of second-stage transistor $M_2$, are selected as primary design variables, and their optimum values are obtained by sweeping $C_{g1}$ and $C_{g2}$ to find that the sum of $I_{DB}$ and $I_{DS}$ becomes the minimum value. Since $f_c$ is given by the following equation (3), a given value of $t_{lin}$ can be converted to a requirement on $g_m$, since $C_{f}$ is determined from dynamic range (noise) specification. Consequently, $t_{lin}$ can be included in the optimization flow.

$$f_c = \frac{B g_{m1}}{2\pi C_f}$$

On the other hand, $t_{slew}$ as given in (2), is only obtained after the design of $g_m/I_D$ and $I_{DB}$. As the result, $t_{slew}$ and $t_s$ cannot generate any constraints for the design optimization; furthermore, the dependence of parasitic capacitance on $g_m/I_D$ generates a non-negligible settling time error between the optimization flow and simulation result.

Settling time optimization

From equations (1)-(3), settling time $t_s$ can be given by,
\[ t_s = t_{lin} + t_{slew} = (1 + \alpha) t_{lin} \]  
\[ \alpha = \frac{t_{slew}}{t_{lin}} = \left( \frac{g_{m1}/I_{D1}}{g_{m2}/I_{D2}} \right)^{\frac{1}{2}} \left( \frac{I_{D1}}{I_{D2}} \right)^{\frac{1}{2}} \left( 1 - \frac{2\kappa}{C_f} \right) \]  
\[ \kappa = \ln \left( 1 - \frac{C_f}{C_L} \right) \]

For given values of \( g_{m1}/I_{D1} \) and \( C_{gg} \), \( \alpha \) can be considered as constant; thus, \( t_s \) specification generates the requirements on \( t_{lin} \) and associated \( f_c \). By using this proportional relationship between \( t_s \) and \( t_{lin} \) for given \( g_{m1}/I_{D1} \) and \( C_{gg} \), we can develop an iteration algorithm to optimize the settling time and power consumption in the following sequence.

1. Obtain an initial value of \( t_{lin} \) from \( t_{spec}/(1 + \alpha_{init}) \), where \( t_{spec} \) is the target specification of settling time, and \( \alpha_{init} \) is calculated from initial values of \( g_{m1}/I_{D1} \) (>20 to generate the minimum \( t_{lin} \)) and \( C_{gg1} (=C_f+C_L) \) according to (5).

2. Optimize primary design variables, \( C_{gg1} \) and \( C_{gg2} \), to realize the minimum current condition for the given \( t_{lin} \). This optimization flow is based on the previously published one [1]; however, the parasitic capacitances, \( C_{gg} \) and \( C_{db} \), are recalculated iteratively according to the optimization sequence of \( g_{m1}/I_{D1} \) for all transistors in order to take the dependence of \( C_{gg} \) and \( C_{db} \) on \( g_{m1}/I_{D1} \) into account.

3. Calculate \( t_s \) from the recalculated value of \( \alpha \), which is a function of optimized values of \( g_{m1}/I_{D1} \) and \( C_{gg1} \) as given in (5).

4. If recalculated \( t_s \) satisfy \( t_{spec} \) within a range from \(+0\%\) to \(-1\%\), complete the optimization sequence, and obtain \( W \) from the current density plot. If \( t_{lin} \) is out of this range, repeat from the sequence 2 after the replacement of the original \( t_{lin} \) to \( t_{lin}(t_{spec}/t_s) \).

The entire optimization sequence is illustrated in Fig. 2, where the additional optimization procedures for the settling time and parasitic capacitance are highlighted by the bold lines; while the table lookup sequences are highlighted by the dotted lines.

4. Experimental Results

The proposed optimization flow along with \( g_{m1}/I_{D1} \) lookup tables generated by SPICE simulations are implemented in MATLAB. To verify the versatility of the optimization flow, two 0.18\( \mu \)m CMOS technologies are adopted for the comparison. Fig. 3 plots the converging characteristics of the settling time recalculation for two technologies. Up to six iterations are necessary to optimize the settling time and power consumption simultaneously.

The design results of transistor sizes and bias currents are compared in Table II. The transistor width \( W \) is optimized with the unit size of 2.5\( \mu \)m. The optimized value of \( C_g \) is identically 1pF for both technologies, and \( R_f \) are 224Q2 and 279Q2 for Tech. 1 and Tech. 2, respectively. The difference in \( L_s \) between technologies is caused by the difference in the intrinsic gain, and the difference in the bias current is attributed to the process frequency characteristics. SPICE simulation results for two design results are summarized in Table III. Both simulation results satisfy the target specification without a tweaking of design parameters. The settling time difference between the optimization and simulation is less than 0.6nsec (4.1%), and is resulted from the bias current difference due to the use of unit \( W \) and channel length modulation. The difference in power consumption between the optimization and simulation is less than 4.0%, which presents a superior correspondence.

Acknowledgements

This work is supported by VLSI Design and Education Center (VDEC), The University of Tokyo with the collaboration with Synopsys Corporation.

References