Analysis and Optimization of Ground Bounce in Digital CMOS Circuits

Payam Heydari and Massoud Pedram
Dept. of EE-Systems, University of Southern California
Los Angeles, CA 90089
E-mail: {payam, massoud}@sahand.usc.edu

Abstract-- *This paper is concerned with the analysis and optimization of the ground bounce in digital CMOS circuits. First, an analytical method for calculating of the ground bounce is presented. The proposed method relies on accurate models of the short-channel MOS device and the chip-package interface parasitics. Next the effect of ground bounce on the buffer propagation delay and the optimum taper factor is discussed and a mathematical relationship for total propagation delay in the presence of the ground bounce is obtained. Effect of the on-chip decoupling capacitor on ground bounce waveform and the circuit performance is analyzed next and a closed form expression for the peak value of the differential-mode component of the ground bounce in terms of on-chip decoupling capacitor is provided. Finally a design methodology for controlling the switching times of the output drivers to minimize the ground bounce is presented.*

I. INTRODUCTION

Signal integrity is a crucial problem in many VLSI circuits and is becoming increasingly important as minimum feature size shrinks to sub-quarter micron. A major component of the circuit noise is the power/ground bounce due to simultaneous switching of I/O pins. Faster clock speeds and larger number of devices and I/O drivers have resulted in increased amount of inductive noise in the power and ground plane.

A number of researchers have studied the ground bounce problem. In [1], Senthinathan *et al.* described an accurate technique for estimating the peak ground bounce noise by considering a negative local feedback present in the current path. The work however suffers from the assumption about the triangular form of the switching current waveform. In [2], Vaidyanath *et al.* relax this assumption by deriving the expression for peak value of ground bounce value under the more realistic and milder assumption that the ground bounce is a linear function of time during the output transition of the driver. The authors do not however obtain the time domain waveform of the ground bounce and use a very simplistic model of the pad-pin parasitics (i.e., inductance only).

More recently a number of researchers have tried to consider the short channel effects of the devices on the ground bounce waveform [3][4][5]. While most prior works were concentrated on the case where all the drivers switch simultaneously, paper [5] considers the more realistic case in which the drivers possibly switch at different times. The idea of considering the effects of ground bounce on the tapered buffer has been presented in a paper by Vemuru [6]. The author however does not provide the mathematical analysis required for designing the optimum number of drivers in the tapered buffer chain. In [7], Vittal *et al.* describe an algorithm based on integer linear programming to skew the switching time of the drivers to minimize the ground bounce. However since the ground bounce is analyzed by a high level approach and does not make use of the characteristics of the ground bounce waveform, the proposed technique is not effective. In addition it increases the propagation delay through the output buffers.

In this paper, the ground bounce is addressed with no assumptions about the form of the switching current or noise voltage waveforms. We circumvent the drawbacks of previous approaches by adopting an accurate chip-package interface model consisting of resistive, inductive components. The effect of ground bounce on the tapered buffer design is considered and a mathematical approach is adopted to consider the ground bounce effect on the propagation delay and the optimal taper factor. We next address the impact of on-chip decoupling capacitor on the peak value of the ground bounce. This is an important problem since the decoupling capacitors are widely used to control the ground bounce and to reduce the resonant frequency of the power and ground network [8][9][10]. We thus present a method to find a closed-form expression of peak value of the differential-mode component of the ground bounce as a function of the decoupling capacitor. Finally we propose a technique to skew the output buffers. By this method the peak amplitude of the ground bounce is reduced to at least 65% of its value when all the drivers switch simultaneously. Our technique does not introduce a large delay after skewing the switching times of buffers. Finally the effect of ground bounce on the tapered buffer design is considered and a mathematical method for the optimization of tapered buffers is provided.

The remainder of the paper is organized as follows. In section II a circuit model for the chip package interface of the chip is presented. Next a circuit technique is presented to consider the impact of multiple output drivers on ground bounce. Finally, the ground bounce is analyzed for the resistive and inductive chip-package interface. Section III discusses the tapered buffer design for ground bounce optimization. The effect of decoupling capacitor on ground bounce reduction is analyzed in Section IV. Section V discusses the skew control for ground bounce optimization. Section VI contains our concluding results.

II. OFF-CHIP GROUND BOUNCE ANALYSIS

To analyze the ground bounce, a circuit model based on the layout schematic of the output pad drivers along with the bonding wires and package pins should be used. With continuous scaling of technology and increasing the switching speeds of integrated circuits, the on-chip inductive effect has become more important than before, but it is still overshadowed by the off-chip inductive noise. Hence on-chip interconnect is modeled as a resistive-capacitive lowpass π circuit. As a simplification which is justified by the available data for typical values of *R*, *L*, and *C* in [11], the circuit model can be simplified to a series *RL* circuit. The circuit schematic of N output drivers driving off-chip capacitors is shown in Fig. 1. According to this figure, *R* and *L* represent the ground and power chip-package interface parasitics while R_w and L_w stand for the load terminal parasitics. t_r is the rise-time of the input waveform and *T* is the cycle time.

Another consideration is to use an accurate device model which includes the short channel effects of the MOS device. With rapid decrease of the feature sizes of the MOS devices the short channel effects must be accounted for. Basically as mentioned in [8], these effects including mobility degradation, velocity saturation, hot carrier effects, and output impedance cause a change from a squared current dependence in the saturation region to a pseudo-linear one. We make the following observations.

- The effect of channel-length modulation can be ignored due to the scaling of supply level [10].
- L_{eff} is 0.13μ-0.25μ in current technologies. The supply voltage level is 1.2*V*-3.0*V*. Considering these values, we simplify the i_d-v_{ds} equation in the saturation region as indicated in Eq. (1):

This research is funded in part by SRC under contract no. 98-DJ-606.

- To come up with a closed form delay expression later in section III, the lateral electric field in short-channel transistors is assumed to be a constant in terms of drain-source voltage. This assumption gives more accurate results than the long channel formulations in which the lateral electric field is totally ignored.

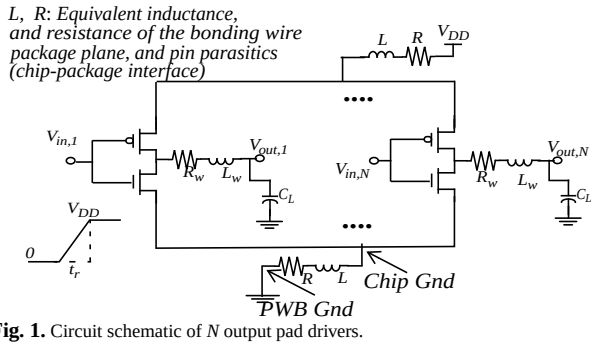

Fig. 1. Circuit schematic of *N* output pad drivers.

$$i_d = \begin{cases} \beta_n(v_{gs} - V_{tn}) & v_{ds} \geq V_{ds,\,sat} \\ 2\beta_n\left((v_{gs} - V_{tn}) - \frac{v_{ds}}{2}\right) & v_{ds} \leq V_{ds,\,sat} \end{cases} ;\ \beta_n = \frac{0.5K_n(W/L)}{\frac{1}{V_{DD} - V_{tn}} + \frac{1}{LE_c}} \quad (1)$$

This simplification has experimentally been proved to cause at most 2% error. Details are omitted due to lack of space.

II.a. Multiple output drivers

The ground bounce can become very large when multiple output drivers switch simultaneously. In this case the ground bounce equation is obtained for a single driver with a modified gain parameter, $\beta_{n,eq}$ which is the summation of the gain parameters of individual drivers. In reality not all the drivers switch exactly at the same time. Similar to [5], we assume that *N* output drivers switch simultaneously while the remaining *M* drivers are quiet. The quiet drivers are in the linear region, therefore their impact is accounted for by an equivalent resistor, R_{eq}, as shown in Fig. 2. The equivalent parasitic circuit seen by the active drivers thus consists of a series *RL* circuit in parallel with the equivalent on-resistances of NMOS devices. Using parallel-to-series transformation [12], and assuming that $R_{eq} >> |\omega L|$,*R* we obtain a simple equivalent *RL* circuit as depicted in Fig. 2.

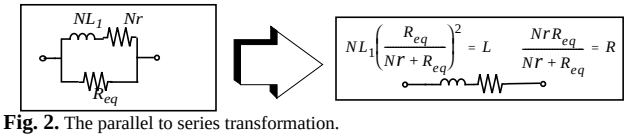

Fig. 2. The parallel to series transformation.

Since the input waveform comprises of two different shapes, a ramp part and a constant input part, in what follows we analyze the ground bounce for each of these two different input parts separately. Our approach is to derive the closed form expressions for the ground bounce by solving the characteristic ordinary differential equation (ODE) coming out of the circuit analysis. Due to the lack of space we omit details of how the differential equations are solved and only provide the final expressions.

II.b. Ramp input

In this case the NMOS transistor is off for $v_{gs}<V_{tn}$. As v_{gs} increases and becomes larger than the threshold voltage, the transistor enters the saturation region. Unlike what is commonly assumed, the ground bounce is not zero for $[0\ ,\ (V_{tn}/V_{DD})t_r]$. Therefore we need to decompose the interval $[0\ ,\ t_r]$ into two subintervals.

II.b.1. $0 \leq t \leq \frac{V_{tn}}{V_{DD}} t_r$

When the transistor is operating in its weak inversion region the amount of drain current flowing through the drain path is very small. Instead there is another current path from input to the ground network provided by C_{gs} of the transistor. C_{gs} is approximately equal to $C_{gs} = (2/3)C_{ox}WL$ [13], where C_{ox} is the parallel plate gate-to-channel capacitor.The ground bounce is:

$$v_n(t) = \frac{V_{DD}}{t_r\omega_d}e^{-\alpha t}\sin\omega_d t + \left(2\alpha\frac{V_{DD}}{t_r\omega_n^2}\right)\left[1 - \frac{\omega_n}{\omega_d}e^{-\alpha t}\sin(\omega_d t + \theta)\right] \quad (2)$$

where $\alpha = \frac{R}{2L}$, $\omega_n^2 = \frac{1}{LC_{gs}}$, $\omega_d = \sqrt{\omega_n^2 - \alpha^2}$, $\theta = \text{atan}\left(\frac{\omega_d}{\alpha}\right)$

II.b.2. $\frac{V_{tn}}{V_{DD}} t_r \leq t \leq t_r$

The ground bounce waveform is given by:

$$v_n(\tau) = \left(v_n(0)e^{-p\tau} + \frac{U_s}{p}(1 - e^{-p\tau}) + \frac{U_r}{p^2}[p\tau - (1 - e^{-p\tau})]\right)u(\tau)\ ; \quad 0 \leq \tau \leq \tau_r \quad (3)$$

where $p = \left(R + \frac{1}{\beta_n}\right)/L$; $U_s = \frac{V_{DD}}{t_r} - \frac{R}{L}V_{tn}$; $U_r = \frac{R}{L}\frac{V_{DD}}{t_r}$

$\tau = t - \frac{V_{tn}}{V_{DD}}t_r$ and $\tau_r = \left(1 - \frac{V_{tn}}{V_{DD}}\right)t_r$.

and with the following initial condition:

$$v_n(0) = \left.\frac{V_{DD}}{t_r\omega_d}e^{-\alpha t}\sin\omega_d t + \left(2\alpha\frac{V_{DD}}{t_r\omega_0^2}\right)\left[1 - \frac{\omega_0}{\omega_d}e^{-\alpha t}\sin(\omega_d t + \theta)\right]\right|_{t\,=\,(V_{tn}/V_{DD})t_r}$$

The peak value of the ground bounce is given by setting $t = t_r$.

II.c. Constant input

The ground bounce in this interval is:

$$v_n(t) = \left[v_n(t_r)e^{-p(t - t_r)} + \frac{R}{R + 1/\beta_n}(V_{DD} - V_{tn})\left(1 - e^{-p(t - t_r)}\right)\right]u(t - t_r) \quad t_r \leq t \leq t_s \quad (4)$$

with the following initial condition:

$v_n(t_r^{\,-}) = v_n(t_r^{\,+})$ = *Evaluate Eq. (4) at* $\tau = \tau_r$

The time t_s is the time when the ground bounce is within 1% of its steady state value ($(R/(R + 1/\beta_n))(V_{DD} - V_{tn})$) and is determined by the following equation:

$$t_s = t_r + \frac{1}{p}\ln\left(\frac{(R + 1/\beta_n)v_n(t_r) - R}{0.01R(V_{DD} - V_{tn})}\right) \quad (5)$$

After time t_s the MOS transistor enters its linear region and is modeled by a voltage dependent finite on-resistance r_{DS}. The circuit consisting of the load capacitor, C_L, the parasitic wire resistance and capacitance, R_w and L_w, the MOS device as a voltage dependant resistance, r_{DS} , and the chip-package interface equivalent parasitics, *R* and *L*, all in series, is solved to obtain the ground bounce voltage. Note that during the design of the output drivers, their *W/L* ratio is assumed to be large enough so that they can provide sufficient current for the off-chip load. On the other hand, having a large r_{DS} causes a large transient power dissipation in the MOS devices of the output drivers which is undesirable. So r_{DS} values are usually in the range of 10Ω-50Ω . Therefore, $(R + R_w + r_{DS}) < 2\sqrt{(L + L_w)/C_L}$ and the ground bounce experiences a decaying oscillatory waveform in $[t_s\ ,\ T/2]$ interval as also shown in Fig. 3. Since in each cycle of the oscillation the electric energy across the load capacitor converts to the electromagnetic energy stored in the electromagnetic field across the inductor and dissipated energy in the resistor, we have a complete fluctuation around the steady-state which is zero volt in this case and the ground bounce passes through a minimum undershoot. The ground bounce waveform is now given by Eq. (6):

$$v_n(t) = L_e\omega_n' I\left(\frac{R}{L_e}\sin\omega_d'(t - t_s) + \cos(\omega_d'(t - t_s) + \Psi)\right)e^{-\alpha'(t - t_s)} \quad t \geq t_s \quad (6)$$

where $\alpha' = \frac{R_e}{2L_e}$; $\omega_n' = \frac{1}{\sqrt{L_eC_L}}$; $\omega_d' = \sqrt{\omega_n'^2 - \alpha'^2}$; $\Psi = \text{atan}\left(\frac{\alpha'}{\omega_d'}\right)$

$R_e = R_w + r_{DS} + R$ and $L_e = L + L_w$.

Fig. 3. compares our simulation with the HSPICE simulation for the three output drivers switching simultaneously and with the chip-package interface parameter values specified in the figure. Clearly our analysis can follow the HSPICE simulation in $0 \le t \le 0.5$ nsec. The undershoot time is predicted within 1% error. The error in the transition between the exponential and the decaying oscillatory case comes from the error in modeling the time-varying nonlinear on-resistance of the MOS device when working in the linear region.

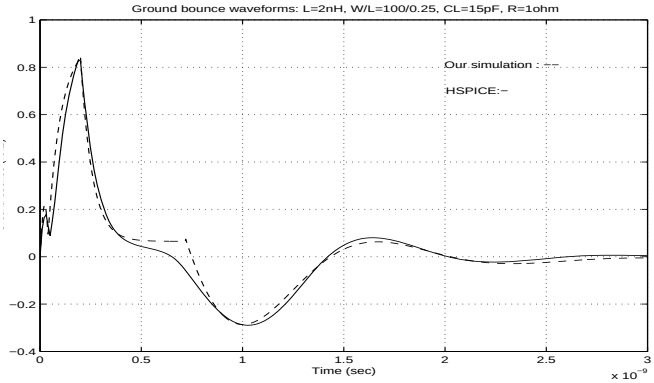

Fig. 3. Ground bounce simulation. (dash): our simulation (plane): HSPICE.

III. TAPERED BUFFER DESIGN FOR GROUND BOUNCE OPTIMIZATION

The ground bounce directly increases the propagation delay of the output buffer and thus affects the optimal scaling factor in a multistage tapered buffer [6]. As a result, the expressions obtained in [6] and [14] for the optimal scaling factor and the optimum number of output drivers are no longer valid and a new analysis is required. From section II we know that the ground bounce is dependent on the nonzero input transition time of the driver. Hence, the first step is to derive the propagation delay of a single driver having short-channel devices controlled by a real ramp input and under the ideal ground condition (*R, L=0*). Fig. 4. shows the result of the HSPICE simulation of an inverter in 0.25μ digital CMOS technology. The device model parameters taken from TSMC 0.25μ single-poly, five metal process technology provided by MOSIS which uses BSIM3v3 MOS model. The device characteristics are also specified in the figure. According to Fig. 4. four different operating regions are distinguished in the time interval $[t_r/2\ ,\ t_{PHL}]$. The regions of operations are also summarized in table 1. The table helps us to determine the region of operation for each NMOS and PMOS transistor throughout the analysis.

As shown in Fig. 4., the PMOS transistor spends a short amount of time in the saturation region. Hence the second and the third intervals can be merged into one single interval since the error introduced by this merging is negligible. To obtain the propagation delay we must obtain the time when the voltage across the load capacitance discharges through the NMOS transistor to $V_{DD}/2$. The propagation delay which is defined as the time difference between 50% points of the input and the output waveforms, is derived via the current-voltage relationship of the load capacitance, C_L:

$$\frac{V_{DD}}{2} = V_{DD} - \frac{1}{C_L}\int_{\frac{t_r}{2}}^{t_{PHL}} i(t)dt \qquad (7)$$

The integral in Eq. (7) is divided into three distinct integrals and the corresponding NMOS and PMOS equations are used for each of these integrals. We assume that the output is decaying exponentially during the transition with a time constant determined by the finite on-resistance of NMOS transistor, and the load and diffusion capacitances. The small overshoot that appears in the output Fig. 4. is due to the feedforward path from the input to the output node through the gate-drain capacitances of NMOS and PMOS transistors. Omitting details of the derivations, the propagation delay for the high-to-low transition is obtained as:

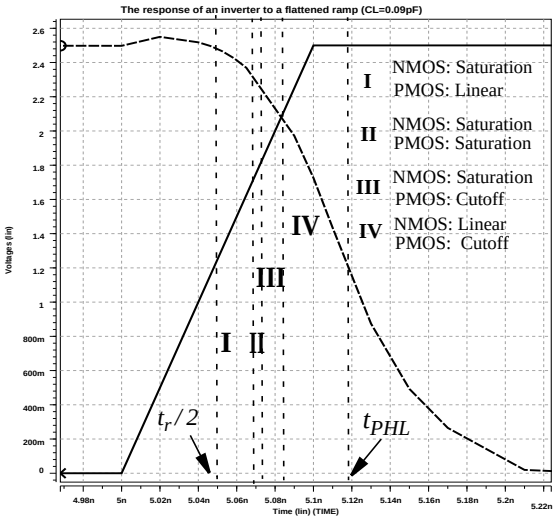

Fig. 4. The input and output waveforms of an inverter simulated by HSPICE. $(W/L)_n$=5/0.25, $(W/L)_p$=10/0.25 (in terms of μm), and C_L=0.09pF.

Table 1: SUMMARY OF CMOS INVERTER OPERATION

	Region I	**Region II**	**Region III**	**Region IV**
NMOS	Saturation	Saturation	Saturation	Linear
PMOS	Linear	Saturation	Cutoff	Cutoff

$$t_{PHL} = t_{PHL,0} + t_{r0}\left(\frac{1}{4}\left(\frac{1}{2} - \frac{\beta_p}{\beta_n}\right)\left(1 - \frac{2V_T}{V_{DD}}\right) + \varepsilon\right) \qquad (8)$$

where

$$\varepsilon = \frac{1}{2}\left[1 + \frac{1 - V_T/V_{DD}}{1 - 2V_T/V_{DD}}\right]$$

$$t_{PHL,0} = \frac{C_L}{2(1 - 2V_T/V_{DD})}\left(\frac{1}{\beta_n}\right) - \left[\frac{V_T/V_{DD}}{1 - 2V_T/V_{DD}}\tau_n \ln\left(\frac{1}{1 - V_T/V_{DD}}\right)\right]$$

$$\tau_n = r_{DS}(C_L + C_{db,n})$$

where $t_{PHL,0}$ is the 50% propagation delay in the ideal case of having an ideal step input and $C_{db,n}$ is the drain-bulk junction capacitance. t_{r0} is the input rise-time of the single driver. A similar expression is obtained for the low to high transition of the output, except that in Eq. (8) β_p and τ_n are replaced by β_n and τ_p, respectively, and vice versa:

$$t_{PLH} = t_{PLH,0} + t_{r0}\left(\frac{1}{4}\left(\frac{1}{2} - \frac{\beta_n}{\beta_p}\right)\left(1 - \frac{2V_T}{V_{DD}}\right) + \varepsilon\right) \qquad (9)$$

and finally the total propagation delay is:

$$t_d = \frac{t_{PLH} + t_{PHL}}{2} \qquad (10)$$

Table 2. compares the simulated values of an inverter delay to the values obtained by Eq. (10) and the values obtained by equation proposed in [15] which widely used by circuit designers. For all the cases depicted in table 2 the load capacitance is 0.09pF.

From table 2 one can see that the error increases when W_p/W_n. decreases. The obvious reason is that the high-to-low transition experiences less delay (in the order of 3) compared to

the low-to-high transition. This asymmetry cannot be accurately predicted by [15] whereas Eq. (10) gives a closer result for the delay.

After obtaining a closed-form expression for delay we try to obtain the propagation delay of a tapered buffer. The following lemma helps us determine the propagation delay of a chain of tapered buffers.

Table 2: COMPARISON BETWEEN SIMULATED INVERTER DELAY VALUES (HSPICE LEVEL 49, 0.25μ PROCESS) AND THE VALUES PROPOSED BY [15] AND THOSE DERIVED BY EQ. (10). DELAYS ARE GIVEN IN nsec, AND TRANSISTOR SIZES ARE GIVEN IN μm

$\frac{t_r}{T/2}$	$\frac{0.05}{0.6}$	$\frac{0.05}{1}$	$\frac{0.09}{1}$	$\frac{0.15}{0.6}$	$\frac{0.4}{2}$	$\frac{0.2}{2}$
$W_p/W_n = 10/5$ Simulation	0.073	0.073	0.084	0.086	0.12	0.094
$W_p/W_n = 10/5$ Paper [15]	0.056	0.056	0.072	0.073	0.98	0.088
$W_p/W_n = 10/5$ Eq. (10)	0.068	0.068	0.08	0.08	0.1	0.091
$W_p/W_n = 10/10$ Simulation	0.04	0.040	0.047	0.052	0.066	0.067
$W_p/W_n = 10/10$ Paper [15]	0.03	0.03	0.04	0.046	0.057	0.057
$W_p/W_n = 10/10$ Eq. (10)	0.037	0.037	0.042	0.05	0.06	0.06
$W_p/W_n = 5/10$ Simulation	0.037	0.037	0.086	0.095	0.114	0.103
$W_p/W_n = 5/10$ Paper [15]	0.03	0.03	0.05	0.059	0.08	0.069
$W_p/W_n = 5/10$ Eq. (10)	0.032	0.032	0.069	0.085	0.096	0.091

Lemma 1. Consider a chain of P inverters, each made of short channel devices. Assume that the gate aspect ratio of each stage is u times larger than that of the previous stage. As an approximation assume that the rise time of any stage is η times larger than the propagation delay of the previous stage plus the rise time of the previous stage (i.e. $t_{r,i} = \eta t_{d,i} + t_{r,i-1}$, for $2 \le i \le P$). Then the total propagation delay is given by:

$$t_p = u\left[\frac{[\eta A + 1]^P - 1}{\eta A}\right] t_{p0} + \left[\frac{[\eta A + 1]^P - 1}{\eta}\right] t_{r0} \qquad (11)$$

where $A = \frac{1}{8} - \frac{1}{8}\left(\frac{\beta_n}{\beta_p} + \frac{\beta_p}{\beta_n}\right)$.

t_{p0} is the propagation delay of a minimum size inverter driving another minimum size inverter when the input rise time is zero (i.e. the first term of Eq. (10)).

Proof: According to Eq. (10) for a single inverter the propagation delay can be thought as a term representing the delay for an input excitation with a zero valued rise time plus a term representing a linearly dependent function of the rise time.

$$t_{d1} = t_{p0} + A t_{r0} \qquad (L1.1)$$

where A is the coefficient of t_{r0} in Eqs (8) and (9).

Now suppose that we have a chain of P inverters. If the gate aspect ratio of the inverters is gradually scaled up with a constant factor of u , then the load capacitor seen by each inverter is scaled up by the same factor. So are the gain factors β_n and β_p of transistors. By revisiting Eq. (10) we see that only the first term of the delay expression is affected by scaling and the second term remains unaffected. Hence for each stage the first term is scaled up by the u factor. The equation for the first inverter is:

$$t_{d1} = u t_{p0} + A t_{r0}$$

The equation for the second inverter has a similar mathematical form:

$$t_{d2} = u t_{p0} + A t_{r1} \qquad (L1.2)$$

Next, we must determine the rise time of the second inverter in terms of the rise time of the first one. According to our approximation we can write:

$$t_{r1} = \eta t_{d1} + t_{r0} \qquad (L1.3)$$

By combining the equations (L1.1), (L1.2) and (L1.3) and eliminating the t_{p0} term from equation (L1.2) we obtain:

$$t_{d2} = (\eta A + 1) t_{d1}$$

Similarly the propagation delays of the subsequent inverters can be obtained in terms of the propagation delay of the previous stages:

$$t_{d,i} = (\eta A + 1) t_{d,i-1} \quad ; \quad 2 \le i \le P \qquad (L1.4)$$

The total propagation delay is the summation of propagation delays of all the individual stages.

$$t_p = t_{d1} \sum_{i=0}^{P-1} (\eta A + 1)^i \qquad (L1.5)$$

The above equation is indeed a geometric series that directly yields the desired expression given by (11).

❑

In practice η is a number between 1 and 2. Fig. 5. shows the effect of nonzero input rise time on the optimum taper factor for various values of x (x is the ratio between the load capacitance C_L and the input capacitance of the tapered buffer C_{in}). The optimal taper factor increases with increasing number of stages. For instance in the case of 10 stage buffers shown in the figure, the optimal taper factor is the well known e=2.7182 if $t_r=0$, but it becomes approximately 3.9, otherwise.

Before considering the effect of ground bounce on the total propagation delay of buffer chains, the impact of the ground bounce on the delay of a single buffer is analyzed. To simplify the derivations, the chip-package-interface parasitics will be modeled by a pure inductor. The delay increases by an additional factor due to the ground bounce effect. Skipping details of the mathematical derivations, this additional term turns out to be inversely proportional to the input transition time:

$$t_{p0,GBN} = t_{p0} + \frac{\delta}{t_{r0}} \qquad \text{where } \delta = L^2\left(\frac{\beta_n^2 + \beta_p^2}{2}\right)\frac{V_{DD}}{V_{DD} - V_T}$$

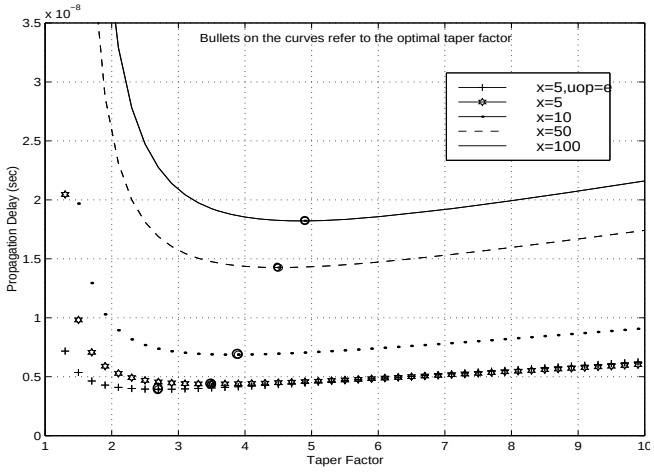

Fig. 5. Propagation time vs. taper factor for various values of x.

Now let's see what the effect of ground bounce is on the optimal number of stages in the tapered buffer. Reducing the taper factor causes the propagation delays of the earlier stages of the multistage buffer to be reduced accordingly. Smaller

propagation delay results in reduced input transition time to the final stages of the tapered buffer. By reducing the input transition time, the ground bounce peak amplitude increases as indicated by the equation for $t_{p0,GBN}$. Larger amplitudes of the ground bounce reduces the current capability of the MOS devices and consequently results in an increase in the propagation delay of the multi-stage buffer. We expect that the optimal taper factor decreases as a result of the noisy ground. The total propagation delay in the presence of the ground bounce is obtained using lemma 2.

Lemma 2. For a multistage tapered buffer with the same specification as in lemma 1 and in the presence of the ground bounce, the total propagation delay is obtained by the following equation:

$$t_{p,GBN} = t_{p,initial} + \frac{\delta \cdot A}{(\eta A+1)^{P-2}(ut_{p0,GBN}+t_{r0}) - ut_{p0,GBN}} \qquad (12)$$

where $t_{p,initial}$ has the same form as t_p given in Eq. (11) except that t_{po} is replaced with $t_{p0,GBN}$.

Proof: The proof for this lemma is similar to the proof of the lemma 1, except that the propagation delay of each stage has an additional term compared to Eq. (L1.4). More precisely:

$$t_{d,i} = (\eta A+1)^{i-1} t_{p0} + \frac{\delta}{t_{r,i-1}} \qquad (L2.1)$$

This additional term is inversely proportional to the rise-time of the previous stage due to the effect of the inductor. To obtain the desired Eq. (12) the same steps can be taken as in the proof of lemma1.

❑

Fig. 6. shows a plot of t_p vs. the taper factor for both cases of the ground bounce being present (nonideal ground plane) and the ideal ground plane. As we expect the optimum taper factor increases and therefore the optimum number of buffers decreases accordingly. For instance for *x = 100*, the optimal taper factor increases from 4.8 to 5.7. This discussion confirms that the optimum taper factor should be increased in the presence of the ground bounce.

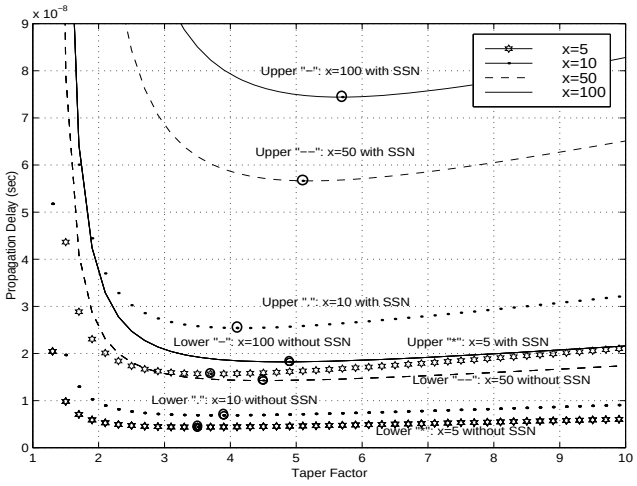

Fig. 6. The effect of ground bounce on the optimal taper factor.

IV. ON-CHIP DECOUPLING CAPACITOR

We need to properly estimate the required amount of the on-chip decoupling capacitors. Overestimation is costly from the area point of view whereas underestimation may lead to noise margin problem. The main effect of on-chip decoupling capacitor is that it forces the same fluctuations to appear on both the on-chip power and ground wires. The output pad buffers consist of tapered inverter chains in order to drive large off-chip capacitors with a short transition time. As a result, the input to the last stage of the output buffer is driven by another predriver stage and the common-mode noise component which appears on the *P/G* bus also appears on the input line. Therefore the bouncing of supply and ground wires due to chip-package parasitics does not affect the circuit performance. The relevant steps that should be taken to correctly compute the value of the decoupling capacitors are:

- Decompose the circuit into two distinct parts, one used for the differential-mode component and the other used for the common-mode component of *P/G* fluctuations.
- Analyze the differential-mode circuit and compute the correct amount of on-chip decoupling capacitor.

Fig. 7. depicts the two circuits corresponding to common-mode and differential-mode fluctuations on *P/G* wires along with the relevant values of voltages and currents shown in this figure. As can be seen, for the differential-mode circuit the decoupling capacitor is virtually replaced by two identical capacitors each twice the original decoupling capacitance. value. Because the two voltages $V_{id}/2$ and $-V_{id}/2$, indicated by the virtual voltmeters V in Fig. 7., are 180 degrees out of phase, node *O* in Fig. 7.a becomes an AC ground. The decoupling capacitor gets duplicated and is placed in parallel with other chip-package interface parasitics. Furthermore, since the input to the buffer is fed from the previous stage, the differential-mode component on the supply line also appears on the input line of the buffer. Considering the above discussions, the differential equation relating the differential-mode component of noise fluctuations, v_{nd}, to drain current of the device is:

$$2LC_D\frac{d^2v_{nd}}{dt^2} + 2RC_D\frac{dv_{nd}}{dt} + v_n = L\frac{di_{ds}}{dt} + Ri_{ds} \qquad 0 \le t \le t_r \quad (13)$$

where C_D stands for the decoupling capacitor which has a large value.

We obtain a closed-form relationship between the maximum value of ground bounce and the on-chip decoupling capacitor. This relationship can help the designer choose the correct amount of the decoupling capacitor based on a certain allowable peak value of the differential-mode component of the ground bounce. Note the value of C_D is usually sufficiently large so that it smooths out the ringing. In other words, adding a large decoupling capacitor results in the overdamped condition. Since the on-chip decoupling capacitor is large, this condition is often satisfied. The peak value of the ground bounce is obtained by solving Eq. (13) and setting $t=t_r$ which yields:

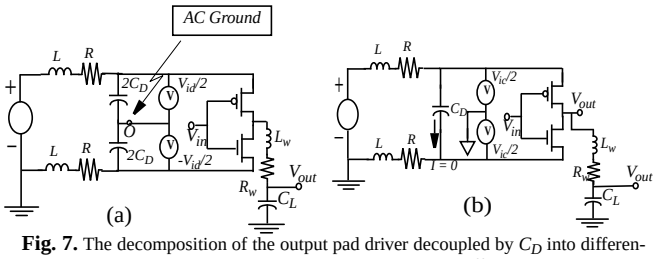

Fig. 7. The decomposition of the output pad driver decoupled by C_D into differential-mode and common-mode equivalent circuits. (a) The differential-mode circuit. (b) The common-mode circuit.

$$v_{nd}(t_r) = \frac{E_0}{\lambda_n^2} - \frac{2\mu E_1}{\lambda_n^4} + \frac{E_1}{\lambda_n^2}t_r + \frac{E_1 - E_0 p_1}{p_1^2(p_2-p_1)}e^{-p_1 t_r} + \frac{E_1 - E_0 p_2}{p_2^2(p_1-p_2)}e^{-p_2 t_r} \qquad (14)$$

where $\lambda_n^2 = \frac{1+R\beta_n}{2LC_D}$; $\mu = 0.5\left(\frac{\beta_n}{2C_D} + \frac{R}{L}\right)$; $E_0 = \frac{\beta_n}{2C_D}\left(\frac{V_{DD}}{t_r} - \frac{R}{L}V_{tn}\right)$

$E_1 = \frac{R\beta_n}{2LC_D}V_{DD}$; $p_{1,2} = -\mu \pm \sqrt{\mu^2 - \lambda_n^2}$.

Eq. (14) at $t=t_r$ can be utilized to obtain the relationship between the peak value of the differential-mode component of the ground bounce and the decoupling capacitor. It is easily veri-

fied that $v_{nd}(t_r)$ is a monotonically decreasing function in terms of C_D as also shown in Fig. 8. Therefore by increasing its value, the differential-mode component of the noise can be arbitrarily reduced. Meanwhile $v_{nd,max}$ is also a monotonically decreasing function in terms of C_D and its value goes to zero for sufficiently large values of C_D. In this case the v_{nd} waveform after adding the decoupling capacitor is an exponential-like waveform and the maximum value of $v_{nd,max}$ is simply obtained by Eq. (14). Fig. 8. shows the variation of the peak value of the ground bounce in terms of C_D .

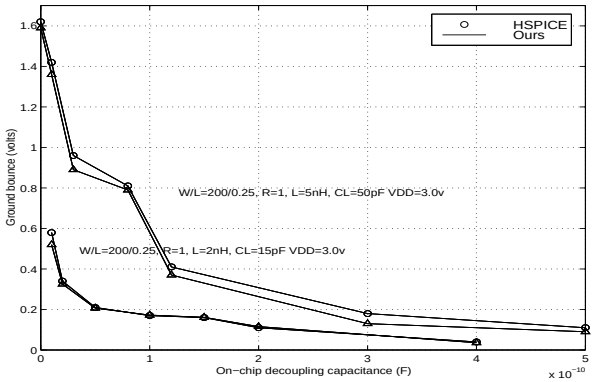

Fig. 8. Variation of the differential-mode part of ground bounce vs. the on-chip decoupling capacitor for three different *W/L* ratios.

V. SKEW CONTROL FOR GROUND BOUNCE OPTIMIZATION

One way to further minimize the peak ground bounce amplitude is to delay the switching time of the output buffers, and thereby to prohibit all the buffers from switching simultaneously. Because of special waveform property we propose an optimum skew time for switching of output buffers under which the ground bounce is attenuated up to 65% of its original value.

As in section II.c the ground bounce declines toward zero as a damped oscillatory waveform and therefore it experiences an undershoot. Now if we tune the switching time of the next driver to occur at exactly the same time that the ground bounce passes through its undershoot point, then its peak value would be maximally attenuated. Let's suppose that there are *N+M* output drivers. The problem can be expressed as minimizing the ground bounce such that the total skew time is less than a delay constraint, T_c.

$$min \quad v_n(t_r) \qquad \textbf{s.t.} \quad \sum_{i=1}^{k} \tau_i \le T_c \qquad 1 \le k \le N+M$$

Since the output drivers have the same physical dimensions, we can equate all the skew times ($\tau_i = \tau_d$ *for all i*). The ratio $\lfloor T_c/\tau_d \rfloor$ gives us the number of drivers which are allowed to be skewed within a certain time constraint T_c. If the total number of output drivers are greater than this ratio, then we have to wrap around and set the switching time of $\lfloor T_c/\tau_d \rfloor + 1$ driver to the switching time of the first driver and so on. As mentioned above $v_n(t)$ experiences an undershoot, we should determine the time when this undershoot occurs.

$$\tau_d = t_s + \Delta t$$

$$\Delta t = \begin{cases} \dfrac{\operatorname{asin}\left(\dfrac{R}{R_e}(1+\cos 2\theta')\right) - \Phi'}{\omega_d'} & if \quad \cos\theta' \le \dfrac{R_e}{2R}\cos(\omega_d'\Delta t + \Phi - \theta') \\ \dfrac{\pi - \operatorname{asin}\left(\dfrac{R}{R_e}(1+\cos 2\theta')\right) - \Phi'}{\omega_d'} & if \quad \cos\theta' > \dfrac{R_e}{2R}\cos(\omega_d'\Delta t + \Phi - \theta') \end{cases}$$

By introducing τ_d seconds delay in switching the second driver, the ground bounce will be reduced by more than 60% as shown in Fig. 9. $\lfloor T_c/\tau_d \rfloor - 1$ drivers are equally triggered by τ_d seconds from each other. The rest of drivers are triggered such that the $\lfloor T_c/\tau_d \rfloor + 1st$ driver switches simultaneously with the first driver. The $\lfloor T_c/\tau_d \rfloor + 2\ nd$ driver switches simultaneously with the second driver and so on. Fig. 9. depicts the skew control of three output drivers under the assumption that the time constraint $T_c = T/2$ half of the clock period.

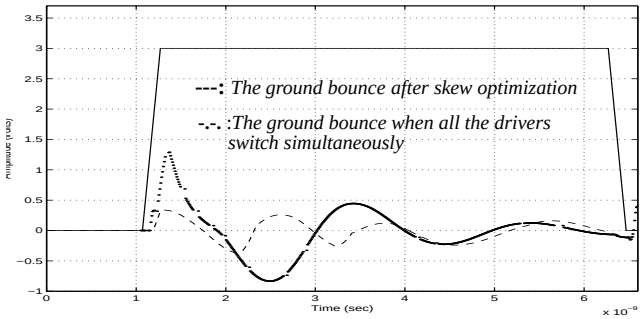

Fig. 9. Ground bounce control by skewing the switching times of three drivers.

IV. CONCLUSION

A detailed analysis and optimization of the off-chip ground bounce using accurate and simple chip-package interface circuit model was presented. The effect of ground bounce on the tapered buffer design was studied and a mathematical analysis was introduced. Next the effect of on-chip decoupling capacitor was analytically investigated and a method to find a closed form expression for the peak value of the differential-mode component of the ground bounce as a function of the decoupling capacitor was presented. Finally a new skew control method for ground bounce optimization was proposed. Experimental results confirmed the effectiveness of this method in reducing the ground bounce.

REFERENCES

[1] R. Senthinathan, J. L. Prince, "Simultaneous Switching Ground Noise Calculation for Packaged CMOS Devices", *IEEE. J. of Solid-State Circuits*, vol. 26, No. 11, pp. 1724-1728, Nov. 1991.

[2] A. Vaidyanath, B. Thoroddsen, and J. L. Prince, "Effects of CMOS Driver Loading Conditions on Simultaneous Switching Noise", *IEEE Trans. Comp., Packag., Manufact. Technol.*, vol. 17, no. 4, Nov. 1994.

[3] S. R. Vemuru, "Accurate Simultaneous Switching Noise Estimation Including Velocity-Saturation Effects", *IEEE Trans. on Comp., Packag., and Manufact. Technol. - Part B*, vol. 19, No. 2, May 1996 .

[4] S. Jou, W. Cheng, Y. Lin, "Simultaneous Switching Noise Analysis and Low Bouncing Design", *IEEE Custom Integrated Circuit Conference*, pp. 25.5.1-25.5.4, May 1998.

[5] H. Cha, O Kwon, "A New Analytic Model of Simultaneous Switching Noise in CMOS Systems", *IEEE Proc. Electronic. Comp. and Technolo Conference*, pp. 615-621, May 1998.

[6] S. R. Vemuru, "Effects of Simultaneous Switching Noise on the Tapered Buffer Design", *IEEE Trans. VLSI Systems*, vol. 5, no. 3, Sept. 1997.

[7] A. Vittal, A. Ha, F. Brewer, M. Marek-Sadowska, "Clock Skew Optimization for Ground Bounce Control", *Proc. IEEE/ACM International Conference on CAD*, pp. 395-399, 1996.

[8] J. M. Rabaey, *Digital Integrated Circuits: A Design Perspective*, pp. 477-482, Prentice-Hall, 1996.

[9] D. Singh, J. M. Rabaey, M. Pedram, F. Cattour, S. Rajkapol, N. Sehgal, T. J. Mozdzen, "Power Conscious CAD tools and Methodologies: A Perspective", *Proc. IEEE*, vol. 83, No. 4, pp. 570-594, April 1995.

[10] H. H. Chen, J. S. Neely, "Interconnect and Circuit Modeling Techniques for Full-chip Power Supply Noise Analysis", *IEEE Trans. on Comp., Packag., and Manufact. Technol. - Part B*, vol. 21, No. 3, August 1998.

[11] C. H. Harper, Electronic Packaging and Interconnection Handbook, pp. 10.1-10.60, Second edition, McGraw-Hill, 1997.

[12]L. O. Chua, C.A. Desoer, E.S. Kuh, *Linear and Nonlinear Circuits*, McGraw Hill, New York, 1987.

[13]D. A. Johns, K. Martin, *Analog Integrated Circuit Design*, Ch. 21, pp. 33-39, John Wiley and Sons, 1997.

[14] N. C. Li, G. L. Haviland, A. A. Tuszynski, "CMOS Tapered Buffer", *IEEE J. of Solid-State Circuits*, vol.30, pp. 1005-1008, August 1990.

[15]D. Hodges, H. Jackson, *Analysis and Design of Digital Integrated Circuits*, McGraw Hill, 1988.