Ground Bounce in Digital VLSI Circuits

Payam Heydari, *Member, IEEE,* Massoud Pedram, *Fellow, IEEE*

Abstract- **This paper is concerned with the analysis and optimization of the ground bounce in digital CMOS circuits. First, an analytical method for calculating 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 propagation delay and the optimum tapering factor of a multistage buffer is discussed and a mathematical relationship for total propagation delay in the presence of the ground bounce is obtained. Effect of an on-chip decoupling capacitor on the ground bounce waveform and circuit speed is analyzed next and a closed form expression for the peak value of the differential-mode component of the ground bounce in terms of the 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.**

Index Terms —— **Signal integrity, Noise, Ground bounce, Tapered buffer, On-chip decoupling capacitor, Skew control.**

I. INTRODUCTION

Signal integrity is a crucial problem in VLSI circuits and is becoming increasingly important as the minimum feature size of devices shrinks to 130 nanometers and below. A major component of the circuit noise is the inductive noise. In fact, faster clock speeds and larger number of devices and I/O drivers as dictated by Moore's Law (and therefore larger value of total circuit current) have resulted in increased amount of this type of noise in the power and ground planes (i.e., the $\frac{dI}{dt}$ noise, also known as the power/ground bounce). It is a critical and challenging design task to control the amount of inductive noise that is inserted into the power planes.

Package pins, bonding wires, and on-chip IC interconnects all have parasitic inductances. When an inductor current experiences time-domain variation, a voltage fluctuation is generated across the inductor. This voltage is proportional to the inductance of the chip-package interface and the rate of change of the current. As a result, when the logic cells in a circuit are switched on and off, the voltage levels at the power distribution lines of the circuit fluctuate. This inductive noise is sometimes referred to as the simultaneous switching noise because it is most pronounced when a large number of I/O drivers switch simultaneously.

P. Heydari is with the Department of Electrical and Computer Engineering, University of California, Irvine, CA 92697-2625 USA.
M. Pedram is with the Department of Electrical Engineering Systems, University of Southern California, Los Angeles, CA 90089 USA.

To quantify the magnitude of the inductive noise and its effect on the circuit performance by way of an example, assume that the effective inductance between the ground of the chip and the ground of the Printed-Wiring Board (PWB) plane is 10nH, and the rate of change of the switching current is 5 mA/nsec per I/O pin (*c.f.* Fig. 1). Assuming that 16 drivers switch at the same time (to provide data for a 16-bit bus), the peak value of the ground bounce is about 0.8V, which is quite large in the context of today's operating voltage levels and can therefore cause harmful effects on the circuit, such as false switching in the logic gates, especially in dynamic logic gates, timing failure, and timing jitter in the on-chip clock generators.
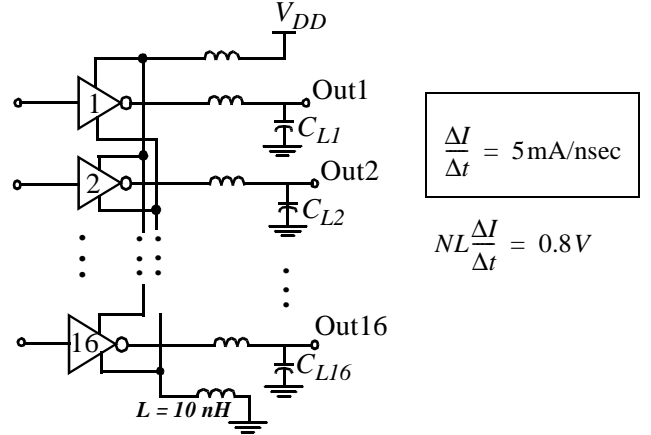

Fig. 1. A simplified circuit schematic of 16 output buffers switching simultaneously.

Due to the large slew rates of the currents flowing through the bond wires and package pins, the ground and supply voltage seen by the output drivers experience bouncing due to the parasitics associated with the package and connections to the chip. Fluctuations on the supply and ground rails are further increased when output drivers switch simultaneously. A number of researchers have studied the power and ground (P/G) bounce problem. The P/G bounce noise is the switching noise on the power-supply and ground lines which consists of the resistive *IR* drop due to bond wire and trace resistances, inductive ΔI-noise due to the chip-package interface inductance including bond wire self-inductance, trace self inductance, trace-to-trace mutual inductance, and capacitive coupling due to the chip-package interface cross-coupling capacitances. While, due to circuit innovations and device scaling, the speed and accuracy of integrated circuits have steadily increased, the performance of packages, especially for low-cost applications, has not significantly improved. This trend follows from the rather poor scalablility of packaging technology and the design environments in which these packages are being employed.

1

In the past a number of approaches have been proposed to analyze the power/ground bounce and its effect on the performance of VLSI circuits. In [1], Senthinathan *et al.* described an accurate technique for estimating the peak ground bounce noise by observing a negative local feedback that is actually present in the current path of the driver. The work, however, suffers from an unrealistic assumption about the time-domain variations of the switching current waveform. More specifically, paper [1] assumes that the switching currents of output drivers have a tri-angular wave-shapes. In [2], Vaidyanath *et al.* relax this assumption by deriving an 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 obtain time domain waveform of the ground bounce and use a 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 MOS 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 [4] considers the more realistic case where the drivers may 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 quantitative 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 utilize the characteristics of the ground bounce waveform, the proposed technique is far from being accurate. In addition, it increases the propagation delay through the output buffers.

This paper is devoted to the analysis and optimization of the ground bounce. We use a simple, yet accurate, circuit to model the chip-package interface parasitics. The ground bounce is addressed with no assumptions about the form of the switching current or noise voltage waveforms. Throughout this paper, the main focus will be on the ground bounce noise. However, the same approach can also be used for power-supply noise analysis. We circumvent the drawbacks of previous approaches by adopting a more accurate chip-package interface model consisting of resistive, and inductive components. Next, the effect of ground bounce on the tapered buffer design is considered, and a mathe-

matical approach is adopted to consider the ground bounce effect on the propagation delay and the optimal tapering factor. We next address the impact of an 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 for the peak value of the differential-mode component of the ground bounce as a function of the decoupling capacitor. We also 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.

II. CIRCUIT MODELING

In this section a simplified circuit model for the chip-package interface parasitics based on the layout schematic of the output pad drivers, the bond wires, and package pins is presented. We also discuss how to analyze the ground bounce in the general case of several off-chip drivers switching as the same time.

A. Modeling the chip-package interface parasitics

Fig. 2 depicts the layout schematic of three identical output pad drivers along with bond wires and package pins.
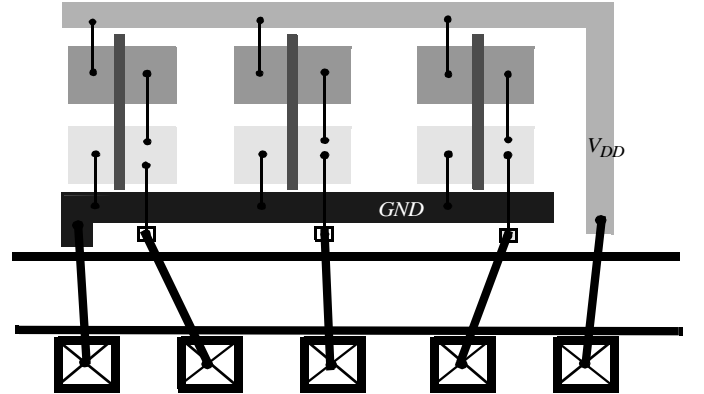

Fig. 2. A layout schematic for the output drivers along with the bonding wires, pads and package pins.
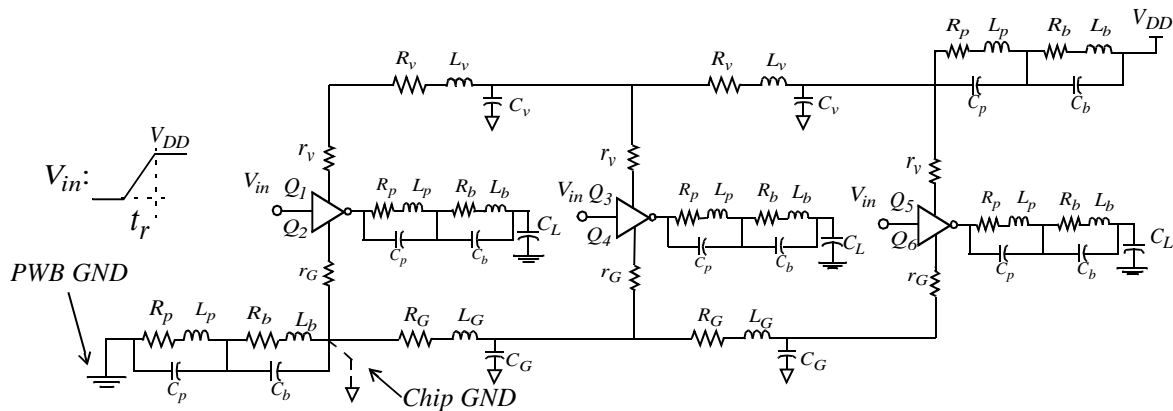

Fig. 3. The circuit representation of Fig. 2

In general, there are two major inductive components which contribute to the total ground bounce: the inductive noise due to the on-chip interconnect, and the inductive noise due to the chip-package interface consisting of bond wire self-inductance, trace and pin self-inductances, and trace-to-trace mutual inductance.

Shown in Fig. 3 is the electrical model of Fig. 2. The on-chip power/ground wires are modeled as a single RLC circuit (R_G, L_G, C_G) for the on-chip ground wire and a single RLC circuit (R_v, L_v, C_v) for the on-chip power-supply wire. The off-chip drivers are normally placed in a close proximity to the pads and bond wires. Therefore the on-chip ground and power lines exhibit small electrical parasitics and thus the two representative RLC circuits for on-chip power and ground wires, (R_v, L_v, C_v) and (R_G, L_G, C_G), are very small compared to the chip-package interface parasitics and can thus be ignored. Bond wires, and package traces and pins are modeled using two separate RLC circuits denoted by (R_b, L_b, C_b) and (R_p, L_p, C_p), respectively, as depicted in Fig. 3.

There are two *RLC* circuits in series that are on the path from the chip ground to the PWB ground. Table I lists the typical chip-package interface parasitics for the CPGA, PPGA, H-PBGA, and TQFP packages [11] [12]. The values listed assume that the package is mounted on the mother-board using a socket, so the pin/land parasitics include the socket effects as well as connecting via parasitics inside the package.

TABLE I
SUMMARY OF PACKAGE I/O LEAD ELECTRICAL PARASITICS
FOR DIFFERENT PACKAGES

Electrical Parameters	Wire-bonding Package Type			
	CPGA	PPGA	H-PBGA	TQFP
Bond wire/die bump R_b (mΩ)	126-165	136-188	114-158	70-150
Bond wire/die bump L_b (nH)	2.3-4.1	2.5-4.6	2.1-4.1	1-4
Bond wire/die bump C_b (pF)	0.2-0.5	0.1-0.3	0.2-0.6	0.1-0.3
Pin/Land R_p (mΩ)	20	20	0	90-97
Pin/Land L_p (nH)	4.5-7.0	4.5-7.0	4.0-6.0	3-5
Pin/Land C_p (pF)	0.1	0.1	0.02	0.1-0.3

Based on practical values specified in Table I, the circuit model can be simplified to a series *RL* circuit because the magnitude response of the capacitive reactance at today's target clock frequencies is more than ten times larger than that of the equivalent impedance of the series *RL* circuit. The ground wiring connection of Fig. 3 is thus simplified to the circuit shown in Fig. 4 where all the ground and power wires reach a single point. The upcoming ground bounce analysis will be performed on the circuit shown in Fig. 4 where N output drivers driving off-chip capacitive loads. According to this figure, R and L represent the ground and power chip-package interface parasitics while R_w and L_w represent the load terminal parasitics. t_r is the rise-time of the input waveform and T is the cycle time.

Another design aspect which should be considered when analyzing the ground bounce is that the output pad driver has a large dimension because it should drive a large capacitor. To estimate the range of transistor sizes used in output buffers, assume that there is a single CMOS driver driving a 50 pF off-chip capacitor. Also assume that the off-chip operating frequency is around 200MHz. A simple calculation reveals that in order to charge up this capacitor to the supply voltage of 2.5V in less than 20% of the total clock period, the required current is 125mA. The device parameters provided by MOSIS for a 0.25μm NMOS device is: $t_{ox} = 58\overset{\circ}{A}$, $\mu_{n,0} = 320$ cm^2/V.sec, and V_{TH0}=0.4V. Using square-law MOS model the *W/L* ratio is approximately 290, which for a minimum channel-length transistor gives rise to a channel width of 73μm. The large current drive requirement for the off-chip CMOS drivers often demands the use of tapered buffer chains.
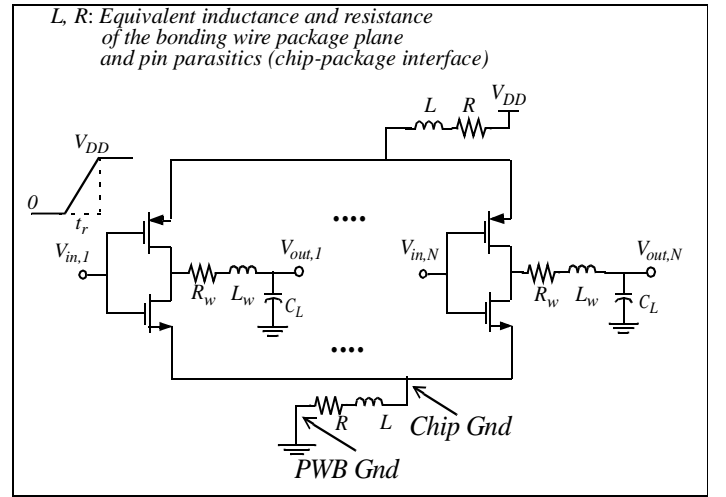

Fig. 4. Circuit schematic of N output pad drivers.

B. Multiple output drivers

Ground bounce can become very large when multiple output drivers switch simultaneously. In this case the ground bounce equation is first calculated for a single driver. To account for the switching effects of multiple output drivers, the NMOS (PMOS) gain parameter, $\beta_{n(p)}$ is modified as the summation of gain parameters of transistors in individual switching drivers. However, 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 circuit representation of the problem is depicted in Fig. 5 (a). To consider the effect of inactive drivers, assume that the gate terminals of quiet drivers are at logic level "HIGH", which causes the NMOS transistors to be in the linear region and the PMOS transistors to be in the cutoff region. The output terminals of the quiet drivers are at logic level "LOW". Unfortunately, the outputs of quiet drivers are exposed to the coupled noise coming from the supply and ground rails. Shown in Fig. 5 (b) is the circuit of Fig. 5 (a) while the NMOS transistors of quiet drivers are modeled approximately by their drain-source resistances. The AC ground nodes of r_{DS} resistors experience the coupled fluctuations from ground and supply lines as also depicted in Fig. 4 (b). As a result, the amount of current flowing through the quiet NMOS transistors

will be negligible, i.e., quiet drivers do not affect the ground bounce analysis. As mentioned earlier, the contribution of the N switching drivers on the ground bounce is taken into account by calculating the ground bounce due to the switching action of a single buffer and then modifying the device gain parameter, $\beta_{n(p)}$.
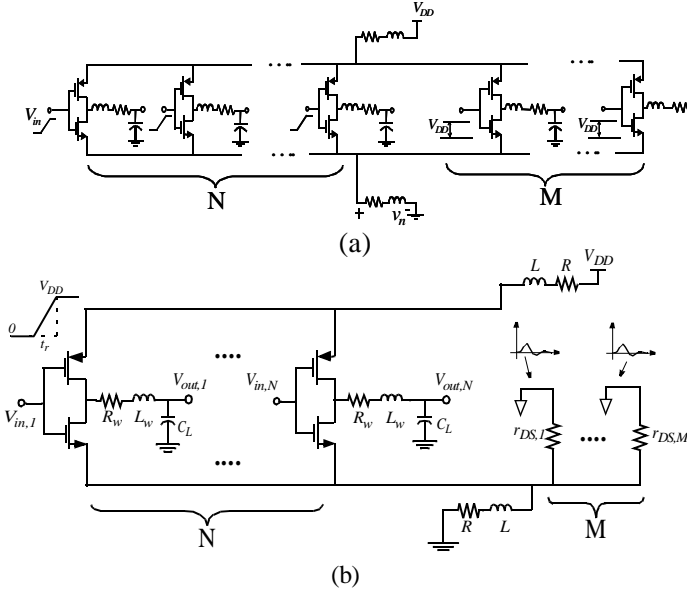

(a)

(b)

Fig. 5. *N* output drivers switch simultaneously and the remaining *M* drivers are quiet. (a) the actual circuit. (b) the simplified circuit.

III. OFF-CHIP GROUND BOUNCE ANALYSIS

Consider an off-chip buffer driving a large capacitive load, as depicted in Fig. 6. The chip-package interface parasitics are modeled using series RL circuits for ground, power and signal paths. Our goal is to obtain the ground bounce through a detailed circuit analysis. This approach is easily extended to include a more general case in which several buffers switch simultaneously.
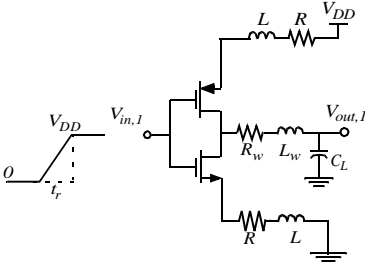

Fig. 6. An output driver driving C_L. The series RL circuits model the chip-package interface parasitics

Since the input waveform is comprised of two different shapes, a ramp voltage and a flat voltage part, and since the NMOS and PMOS devices operate in different regions of operations during the transition from one input state to another, in what follows, the ground bounce for each of these two input shapes is analyzed separately. Notice that a similar analysis may be performed for the power-supply bounce during the low-to-high output transitions. During the ground bounce analysis, only the effect of NMOS current on the ground bounce is considered i.e.,

the effect of PMOS currents is ignored [3], [6]. Fig. 7 demonstrates the input, output, and ground bounce waveform in a circuit consisting of a large off-chip inverter implemented in 0.25μm CMOS process and a 4pF capacitive load.
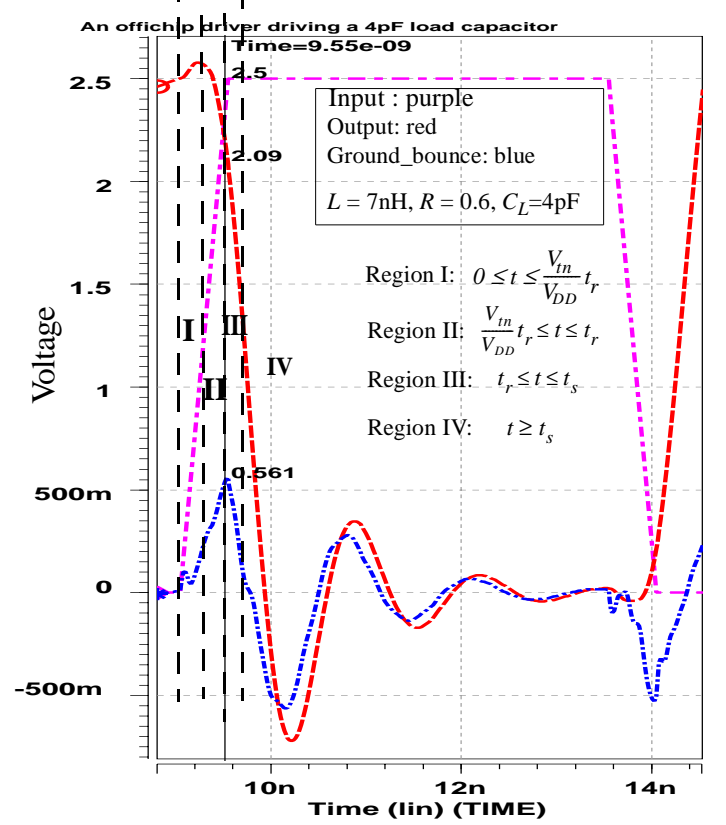

Fig. 7. Ground bounce from HSPICE simulation results of a 0.25μm CMOS process and classification of regions in ground bounce analysis

The chip-package interface parasitics for ground, supply, and signal lines are modeled by series RL circuits whose values are indicated in Fig. 7. The NMOS transistor is off so long as $v_{gs} < V_{tn}$. As v_{gs} exceeds the threshold voltage, the NMOS transistor first enters its saturation region. The transistor stays in saturation region during the entire low-to-high input transition because the off-chip driver drives large capacitive loads, and as a result, the driver output slowly decreases from a logic high to a logic low. As the output voltage decreases, so does the drain-source voltage of the NMOS transistor. Eventually at $t = t_s$ ($t_s > t_r$) the NMOS transistor makes a transition from saturation region to linear region. The NMOS transistor will stay in the linear region for $t > t_s$, until the next edge transition. This particular form of device operating-mode transition in the presence of large capacitive loads allows one to utilize the BSIM3 MOS model with some simplifications as detailed next.

According to the BSIM3 model for the short-channel NMOS transistor [8], the NMOS I-V equations for the saturation and linear regions are as follows:

$$I_D = \begin{cases} W\upsilon_{sat}C_{ox}W(V_{GS}-V_{tn}-V_{DS,sat}) & V_{DS} \geq V_{DS,sat} \\ \mu_n C_{ox}\left(\dfrac{W}{L}\right)\dfrac{1}{1+\dfrac{V_{DS}}{LE_{sat}}}\left(V_{GS}-V_{tn}-\dfrac{V_{DS}}{2}\right)V_{DS} & V_{DS} \leq V_{DS,sat} \end{cases}$$

(1)

where $V_{DS,sat} = \dfrac{LE_{sat}(V_{GS}-V_{tn})}{LE_{sat}+V_{GS}-V_{tn}}$. From the aforementioned discussion, the NMOS transistor will enter the linear region when the input voltage is V_{DD} due to the presence of large off-chip capacitive loads. The overdrive voltage is thus constant at V_{DD}-V_{tn}. As a result, the nonlinear relationship between the saturated drift velocity, $\upsilon_{d,sat}$, and the drain-source saturation voltage, $V_{DS,sat}$, is evaluated at $V_{GS} = V_{DD}$. This leads to the following simplified transistor equation that holds true for the output drivers:

$$I_D = \begin{cases} \beta_n(V_{GS}-V_{tn}) & V_{DS} \geq V_{DS,sat} \\ \dfrac{1}{2}\mu_n C_{ox}\left(\dfrac{W}{L}\right)(V_{DD}-V_{tn})V_{DS} & V_{DS} \leq V_{DS,sat} \end{cases}$$

(2)

where
$$\beta_n = \dfrac{0.5\mu_n C_{ox}(W/L)}{\dfrac{1}{V_{DD}-V_{tn}}+\dfrac{\gamma}{LE_{sat}}}$$

and
$$V_{DS,sat} = \dfrac{LE_{sat}(V_{GS}-V_{tn})}{LE_{sat}+\gamma(V_{DD}-V_{tn})}$$

To account for the voltage variation of the source node, a constant modifying factor, γ, varying between 0.7-0.9 is added to the formulation. In this paper we assume that $\gamma=0.8$.

A similar current-voltage relationship can be derived for the PMOS transistor. Running several experiments and comparing the results with the HSPICE simulation reveals that this simplification causes at most 2% error.

To obtain a better estimate for the ground bounce waveform, we distinguish between four different sub-intervals. Our approach is to derive the closed-form expressions for the ground bounce at each of these sub-intervals by solving the characteristic ordinary differential equation (ODE) coming out of the circuit analysis. We omit the details of how the differential equations are solved and only provide the final expressions.

A. Ramp input

During the low-to-high input transition the NMOS transistor of the output driver experiences multiple region transitions. 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 and obtain the ground bounce waveform for each of these regions.

1) Region I ($0 \leq t \leq \dfrac{V_{tn}}{V_{DD}}t_r$):

During this interval, the NMOS transistor operates in its weak-inversion region. 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 the gate-

to-bulk capacitance C_{gb} of the transistor. Remember that in weak inversion, $C_{gs}=C_{gd}\approx 0$ because the inversion layer contains little charge. However, C_{gs} can be thought as the series combination of the oxide and depletion capacitors [13], Therefore,

$$C_{gs}+C_{gb}+C_{gd} \approx C_{gb} = WL\left(\dfrac{C_{ox}C_{js}}{C_{ox}+C_{js}}\right)$$

where C_{ox} is the parallel plate gate-to-channel capacitor, and C_{js} is the depletion-region capacitor. Writing a KVL for the signal path consisting of the input source, capacitor C_{gb}, and the series RL circuit leads to the following ODE:

$$\dfrac{d^2 v_n}{dt^2}+\left(\dfrac{R}{L}\right)\dfrac{dv_n}{dt}+\dfrac{v_n}{LC_{gb}} = \dfrac{R}{L}\dfrac{V_{DD}}{t_r} \qquad 0 \leq t \leq \dfrac{V_{tn}}{V_{DD}}t_r$$

(3)

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

(4)

where $\alpha = \dfrac{R}{2L}$, $\omega_n^2=\dfrac{1}{LC_{gb}}$, $\omega_d=\sqrt{\omega_n^2-\alpha^2}$, $\theta = \tan^{-1}\left(\dfrac{\omega_d}{\alpha}\right)$.

2) Region II ($\dfrac{V_{tn}}{V_{DD}}t_r \leq t \leq t_r$):

If the overdrive voltage is larger than the NMOS threshold voltage, V_{tn}, the NMOS transistor turns on and operates in its saturation region. The off-chip driver's drain current flowing through the chip-package interface parasitics will generate the ground bounce, v_n. Utilizing the characteristic I-V equation of a saturated NMOS transistor (Eq. (2)) and writing a simple KVL for the drain current path consisting of NMOS transistor and the RL circuit in the circuit shown in Fig. 6, leads us to the following ODE:

$$\dfrac{dv_n}{d\tau}+\dfrac{\left(R+\dfrac{1}{\beta_n}\right)}{L}v_n = \left(\dfrac{V_{DD}}{t_r}-\dfrac{R}{L}V_{tn}\right)+\dfrac{R}{L}\dfrac{V_{DD}}{t_r}\tau \qquad 0 \leq \tau \leq \tau_r$$

(5)

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

Solving the above ODE for v_n yields the following expression for the ground bounce in the interval $(V_{tn}/V_{DD})t_r \leq t \leq t_r$ in terms of t:

$$v_n(t) = \left(v_n(t_0)e^{-p(t-t_0)}+\dfrac{U_s}{p}(1-e^{-p(t-t_0)})\right.$$
$$\left.+\dfrac{U_r}{p^2}[p(t-t_0)-(1-e^{-p(t-t_0)})]\right)u(t-t_0) \qquad t_0 \leq t \leq t_r$$

(6)

where:

$$p=\dfrac{\left(R+\dfrac{1}{\beta_n}\right)}{L}, U_s=\dfrac{V_{DD}}{t_r}-\dfrac{R}{L}V_{tn}, U_r=\dfrac{R}{L}\dfrac{V_{DD}}{t_r}, t_0 = \dfrac{V_{tn}}{V_{DD}}t_r.$$

and with the following initial condition:

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

Note that Eq. (4) is a monotonically increasing function of time. Therefore, the peak value of the voltage $v_n(t)$ occurs at the end of this operating region when $t = t_r$. For $t > t_r$ the input settles at V_{DD} and the ground bounce amplitude decreases. This

means that $v_n(t_r)$ is a global maximum for the ground bounce waveform.

B. Constant input

For $t > t_r$ the input waveform reaches the supply voltage V_{DD}. The rate of change of drain current flowing through the series RL circuit decreases, which causes the ground bounce to decrease in time as well. The ground bounce analysis is performed on two distinct sub-intervals; $t_r \leq t \leq t_s$ when the NMOS transistor is still in the saturation region, and $t \geq t_s$ when the transistor makes a transition to the linear operating region.

1) Region III ($t_r \leq t \leq t_s$):

Over the time interval $t_r \leq t \leq t_s$, the input is flat at V_{DD} and the NMOS transistor operates in the saturation. The corresponding ODE becomes:

$$\frac{dv_n}{dt} + \frac{\left(R + \frac{1}{\beta_n}\right)}{L} v_n = \frac{R}{L}(V_{DD} - V_{tn}) \qquad t_r \leq t \leq t_s \quad (7)$$

with the following initial condition:

$$v_n(t_r^-) = v_n(t_r^+)$$

The ground bounce over this time interval is:

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

$$t_r \leq t \leq t_s \quad (8)$$

At $t = t_s$ the NMOS transistor experiences a transition in its operating mode and enters the linear region. The drain-source voltage of NMOS transistor at $t = t_s$ is equal to the saturated drain-source voltage, $V_{DS,sat}$. The equivalent circuit at $t = t_s^-$ is demonstrated in Fig. 8.
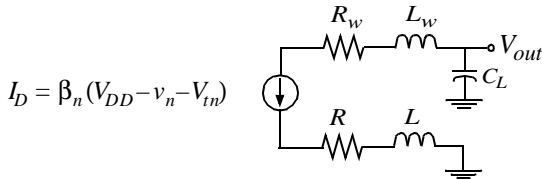

Fig. 8. The equivalent circuit of Fig. 6. at $t = t_s^-$

t_s is determined by solving the following KVL equation around the loop:

$$V_{DD} - \frac{\beta_n}{C_L}(V_{DD} - V_{tn})(t_s - t_r) - \frac{\beta_n}{C_L} \int_{t_r}^{t_s} v_n(t) dt - \left(1 + \left|\frac{R_w + j\omega_c L_w}{R + j\omega_c L}\right|\right) v_n(t_s)$$

$$\approx V_{DS,sat} \quad (9)$$

where ω_c is the clock frequency in rad/sec, and $v_n(t_r)$ is obtained by evaluating (6) at $t = t_r$. $V_{DS,sat}$ is given by Eq. (2) in which $V_{GS} = V_{DD} - v_n(t_s)$.

2) Region IV ($t > t_s$):

For $t > t_s$, the NMOS transistor enters the linear region and is modeled by a voltage dependent finite on-resistance r_{DS}. Shown in Fig. 9, the equivalent circuit consisting of the load capacitor,

C_L, the parasitic resistance and inductance, R_w and L_w, 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.
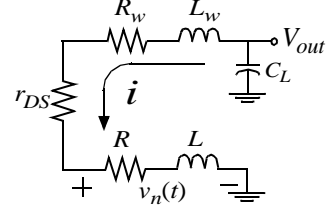

Fig. 9. The equivalent circuit of Fig. 6. at $t = t_s^+$

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. This implies that r_{DS} values of off-chip drivers lie within the range of 20Ω-80Ω . In practice, $(R + R_w + r_{DS}) < 2\sqrt{(L + L_w)/C_L}$ and the ground bounce experiences a decaying oscillatory waveform over $[t_s, T/2]$ as also shown in Fig. 7. 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 current i satisfies the following second-order ODE:

$$\frac{d^2 i}{dt^2} + \left(\frac{R_w + R + r_{DS}}{L}\right)\frac{di}{dt} + \frac{1}{LC_L} i = 0 \qquad t \geq t_s \quad (10)$$

with the following initial conditions:

$$i(t_s) = I = \beta_n(V_{DD} - V_{tn} - v_n(t_s))$$

$$\frac{di}{dt}(t_s) = I' = \frac{(1 + R\beta_n) v_n(t_s) - R\beta_n(V_{DD} - V_{tn})}{L}$$

The solution to Eq. (10) is utilized to derive ground bounce voltage $v_n(t)$ over the time interval $t \geq t_s$:

$$v_n(t) = (V_c \cos\omega_d'(t - t_s) + V_s \sin\omega_d'(t - t_s)) e^{-\alpha'(t - t_s)} \quad t \geq t_s \quad (11)$$

where

$$\alpha' = \frac{R_w + r_{DS} + R}{2(L + L_w)} \quad , \quad \omega_n'^2 = \frac{1}{(L + L_w)C_L} \quad , \quad \omega_d'^2 = \omega_n'^2 - \alpha'^2,$$

$$V_c = RI + LI' \quad , \quad V_s = \left(\frac{R\alpha' - L\omega_n'^2}{\omega_d'}\right)I + \left(\frac{R - L\alpha'}{\omega_d'}\right)I'.$$

Fig. 10 compares our analytical approach with the HSPICE simulation for the three output drivers switching simultaneously

and with the chip-package interface parameter values specified in the figure.
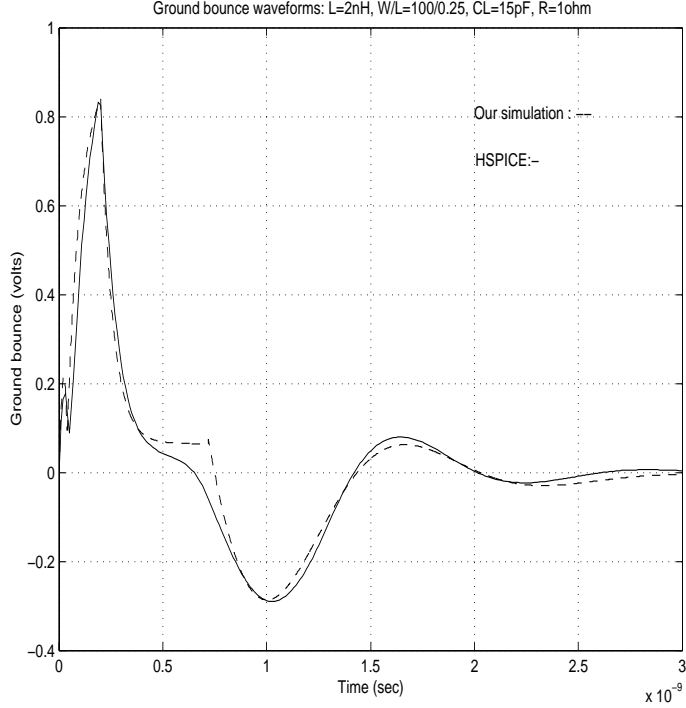

Fig. 10. Ground bounce simulation. (dash): our simulation (plane): HSPICE.

Clearly our analysis can follow the HSPICE simulation in $0 \leq t \leq 0.5$ nsec. The undershoot time is predicted within 1% error. The error in the transition between the exponential and the decaying oscillatory case is due to the error in modeling the time-varying nonlinear voltage-controlled on-resistance of the MOS device.

IV. TAPERED BUFFER DESIGN FOR GROUND BOUNCE OPTIMIZATION

To drive large off-chip capacitances with a minimum of propagation delay, it is necessary to use an output buffer consisting of a number of CMOS inverters with gradually increasing driving capability as also indicated in Fig. 11 [14]. The tapering factor u is the ratio between the gate aspect ratios of two consecutive inverters in the inverter chain.

$$\frac{(W/L)_i}{(W/L)_{i-1}} = u \quad ; \text{ for } i = 1, ..., P$$
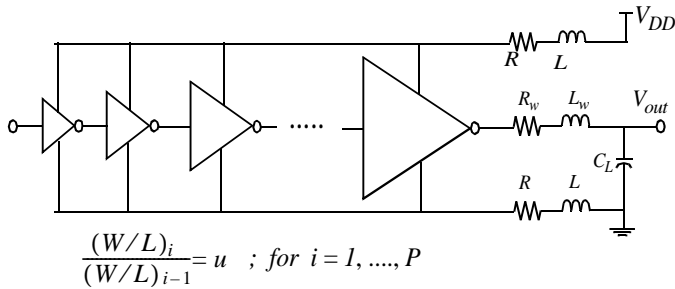
Fig. 11. A CMOS buffer consisting of a series of inverters with gradually increasing driving capability. The chip-package inter-

face parasitics are modeled using series RL circuits.

The ground bounce causes an increase in the propagation delay of the output buffer and thus affects the optimal scaling factor in a series of tapered buffers [6]. As a result, the analytical results for the optimum scaling factor and the optimum number of output drivers proposed in [14] and [15] are no longer valid because they do not address the effects of non-ideal ground and power lines. As a consequence, new formulas that account for the power/ground noise effect on the tapered buffer design is required. From section III.A recall 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 flattened ramp input and under the ideal ground condition ($R, L=0$).

Fig. 12 shows the result of the HSPICE simulation performed on a 0.25μm CMOS inverter. The device parameters are taken from TSMC 0.25μ single-poly, five metal CMOS process provided by MOSIS. The device characteristics are also specified in the figure.
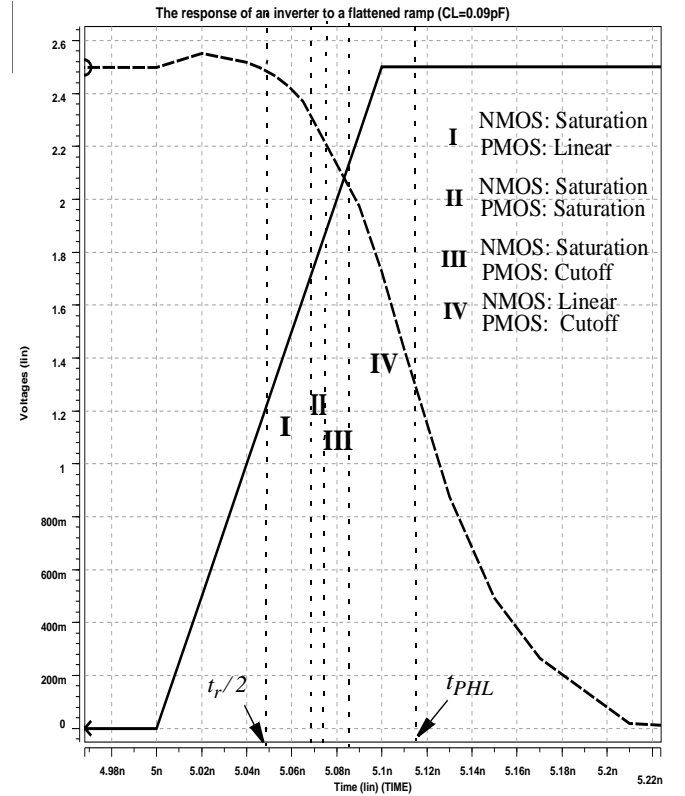

Fig. 12. The input and output waveforms of an inverter simulated by HSPICE. $(W/L)_n$=5/0.25, $(W/L)_p$=10/0.25 (in terms of μm), and C_L=0.09pF.

According to Fig. 12 four different operating regions are distinguished in the time interval $[t_r/2, t_{PHL}]$. The regions of operations are summarized in Table II.

TABLE II
SUMMARY OF CMOS INVERTER OPERATING REGION TRANSITIONS

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

As shown in Fig. 12, the PMOS transistor spends a short amount of time in the saturation region, particularly when the inverter is driving a large capacitive load. Therefore, it is assumed that the PMOS transistor makes a transition from linear region to cutoff region. This assumption means that Region II and Region III can be merged into one single interval since the error introduced by this merging is negligible. To obtain the propagation delay, the time at which the voltage across the load capacitance discharges through the NMOS transistor to $V_{DD}/2$ must be calculated. The propagation delay, which is defined as the time difference between 50% points of the input and output waveforms, is derived through the current-voltage relationship of the load capacitance, C_L:

$$\frac{V_{DD}}{2} = V_{DD} - \frac{1}{C_L}\int_{t_r/2}^{t_{PHL}} i(t)dt \qquad (12)$$

$$t_{PHL} = t_{PHL,0} + t_{r0}\left(\frac{1}{4}\left(\frac{1}{2} - \frac{\beta_p}{\beta_n}\right)\left(1 - \frac{2V_T}{V_{DD}}\right) + \varepsilon\right) \qquad (13)$$

where

$$V_T = V_{tn} = |V_{tp}|$$

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

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

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

where $t_{PHL,0}$ represents 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 delay expression is derived for the low-to-high transition of the output, except that in Eq. (13) β_p and τ_n are replaced by β_n and τ_p, respectively, and vice versa:

$$t_{PLH} = t_{PLH,0} + t_{r0}\left(\frac{1}{4}\left(\frac{1}{2} - \frac{\beta_n}{\beta_p}\right)\left(1 - \frac{2V_T}{V_{DD}}\right) + \varepsilon\right) \qquad (14)$$

where t_{PLH} represents the propagation delay for the low-to-high transition of the output. Finally the total propagation delay is:

$$t_d = \frac{t_{PLH} + t_{PHL}}{2} \qquad (15)$$

Notice that Eq. (15) yields a closed-form delay expression for a CMOS inverter by using Eq. (2), which in turn accounts for the short-channel effects of the transistors. In contrast to other delay estimates published recently in [16] and [17], Eq. (15) provides a simple expression that relates the inverter delay to load impedance as well as the device parameters such as $\beta_{n(p)}$ and device threshold voltage V_T. To verify the accuracy of the delay expression given in (15), we set up an experiment in which a CMOS inverter with $(W/L)_n = 250$ and $(W/L)_p = 500$ is driving a large capacitive load. The capacitance of the load varies between 2pF and 10pF. The input to the inverter is a flattened ramp input with a rise-time varying between 80psec and 500psec. The delay of this inverter is calculated using four different approaches: HSPICE simulation, delay expression proposed in [18], the delay expression proposed in [17], and Eq. (12). Reference [18] uses the square-law model to characterize the MOS transistor. With the on-set of velocity saturation, this model no longer holds true. We modify the delay formula proposed in [18] correspondingly by making the average drain current of transistors I_{av} to be proportional to V_{DD} instead of V_{DD}^2. Results of this comparison are summarized in Table III.

TABLE III
COMPARISON BETWEEN SIMULATED INVERTER DELAY USING HSPICE (LEVEL 49, 0.25μ PROCESS), AND THE MODIFIED DELAY FORMULA PROPOSED IN [18], THE DELAY FORMULA PROPOSED IN [17], AND EQ. (15). DELAYS ARE GIVEN IN psec, AND TRANSISTOR SIZES ARE GIVEN IN μm

t_r (psec)	C_L (pF)	Propagation delay (psec)			
		HSPICE	Paper [18]	Paper [17]	Eq. (15)
80	2	119.5	65.68	112.5	113.2
90	4	235.02	111.51	226.4	230.3
100	3	173.93	92.75	158.6	153.4
150	5	269.68	150.23	245.6	254.3
200	6	330.8	185.5	318.1	325.0
250	8	428.4	242.92	414.2	420.2
300	7	485.57	236.13	477.3	474.3
400	9	506.87	308.2	478.6	497.2
500	6	305.64	294.95	297.3	295.3
500	10	577.13	361.13	548.9	570.1

From Table III one can see that Eq. (15) will provide a better delay estimation for large capacitive loads. The reason for this is that the simplified MOS I-V equations in (2) will become more accurate for larger loads.

Having derived a closed-form expression for the inverter delay, the propagation delay of a tapered buffer can be obtained as explained by the next Lemma.

Lemma 1. Consider a chain of P inverters, each made up 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:

$$t_p = u\left[\frac{[\eta A + 1]^P - 1}{\eta A}\right]t_{p0} + \left[\frac{[\eta A + 1]^P - 1}{\eta}\right]t_{r0} \qquad (16)$$

where $A = \frac{1}{8} - \frac{1}{8}\left(\frac{\beta_n}{\beta_p} + \frac{\beta_p}{\beta_n}\right)$ and 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. (13)).

Proof: According to Eq. (13) for a single inverter the propagation delay can be thought as a linear combination of a term representing the delay for an input excitation with a zero valued rise time and 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 equations (13), (14) and (15).

Now suppose that a chain of P inverters is given. If the gate aspect-ratio of 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. (13) it is observed 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 of the delay expression is scaled up by scaling factor u. As a consequence, the equation for the first inverter is:

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

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

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

Next, one must determine the rise time of the second inverter in terms of the rise time of the first one. According to the assumption made in Lemma 1, we can write:

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

By combining 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 subsequent inverters can be obtained in terms of the propagation delay of previous stages:

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

The total propagation delay is the summation of propagation delays of all 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 in (16).

❑

Lemma 1 is utilized to design the multistage tapered buffer driving large off-chip capacitors while accounting for the non-zero rise and fall times as well as the short-channel effects of the MOS transistors. Our goal is to design a chain of tapered inverters that is capable of driving a large off-chip capacitance C_L with a minimum of propagation delay. To achieve the same delay for the last stage, it is required that $C_L = xC_{in} = u^P C_{in}$. Therefore, the total propagation delay depends on the tapering factor u, the ratio between the external load capacitance C_L of the last stage and C_{in}, x, and the input rise-time t_r. Also in practice, η in Eq. (16) is a real number between 1 and 2. Fig. 13 shows the effect of non-zero input rise time on the optimum tapering factor for various values of x. The optimal tapering factor increases with increasing number of stages. For instance in the case of 10 stage buffers shown in the figure, the optimal tapering factor is the well known $e = 2.7182$ if $t_r = 0$, but it becomes approximately 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 is modeled by a single inductor. The delay of a single inverter in the presence of non-ideal inductive chip-package interface parasitics for the power and ground connections is derived. It turns out that the delay increases by an additional factor due to the presence of ground bounce. This additional term is inversely proportional to the input transition time:

$$t_{p0,GBN} = t_{p0} + \frac{\delta}{t_{r0}} \qquad (17)$$

where
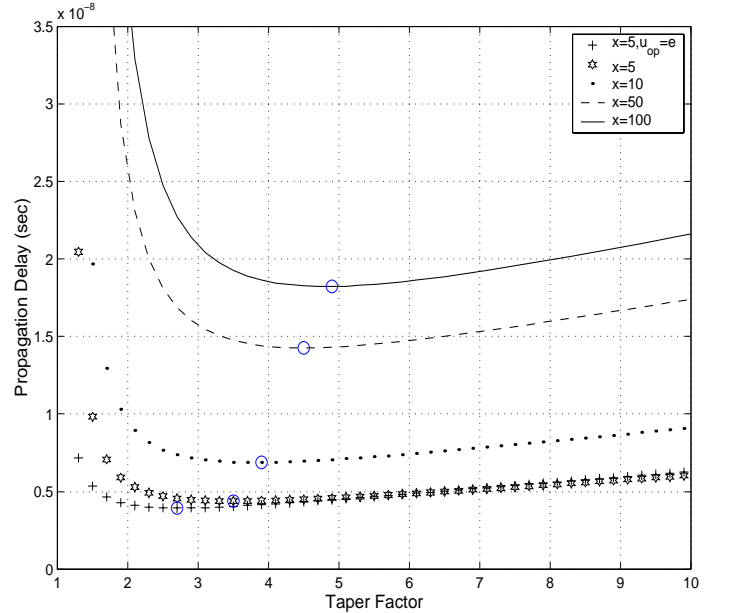$$\delta = L^2\left(\frac{\beta_n^2 + \beta_p^2}{2}\right)\frac{V_{DD}}{V_{DD} - V_T}.$$

Fig. 13. Propagation time vs. tapering factor for various values of x

The analysis is extended to include the effect of ground bounce on the optimum number of stages of a tapered buffer. In a tapered buffer chain, the output transition time of the driving inverter is the input transition time of the driven inverter. The transition time is a function of the buffer's tapering factor which is also seen from Eq. (16). The transition time of propagating signal significantly affects the magnitude of the ground bounce. Reducing the tapering 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 from Eq. (17). 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. Therefore we expect that in the presence of noisy power/ground lines, the number of inverters decreases and the tapering factor increases. Another important observation is that the output transition times of first stages in a tapered buffer are not affected by ground bounce due to relatively small magnitudes of ground bounce during their switching transitions [6]. Using this observation, 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:

$$t_{p, GBN} = t_{p, initial} + \frac{\delta \cdot A}{(\eta A + 1)^{P-2}(ut_{p0, GBN} + t_{r0}) - ut_{p0, GBN}} \tag{18}$$

where $t_{p,initial}$ has the same form as t_p given in Eq. (14) except that t_{po} is replaced by $t_{p0,GBN}$.

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

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

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. (18) the same steps are taken as in the proof of Lemma 1.
❑

Fig. 14 shows a plot of t_p vs. the tapering factor for both cases of the ground bounce being present (non-ideal ground plane) and the ideal ground plane. As we expect the optimum tapering factor increases and therefore the optimum number of buffers decreases accordingly. For instance for $x=100$, the optimal tapering factor increases from 4.8 to 5.7. This discussion confirms that the optimum tapering factor should be increased in the presence of the ground bounce.

According to Fig. 14, for a given x, the propagation delay of the tapered buffer drastically increases as a result of taking the power/ground noise into consideration.
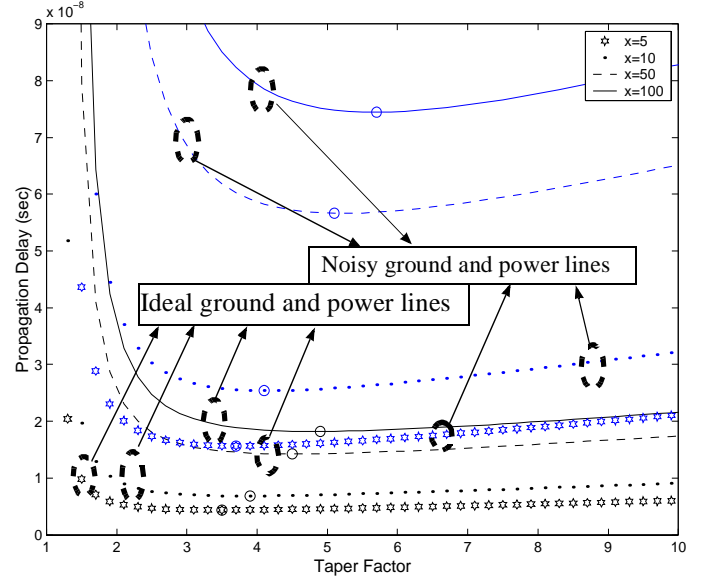

Fig. 14. The effect of ground bounce on the optimal tapering factor.

V. ON-CHIP DECOUPLING CAPACITOR

We need to properly estimate the amount of required on-chip decoupling capacitors. Overestimation is costly from the area point of view whereas underestimation may lead to noise margin problems. The key advantage of a large on-chip decoupling capacitor is that it forces the same fluctuations to appear on both on-chip power and ground planes. Fig. 15 (a) shows the result of HSPICE simulation on a circuit consisting of five identical off-chip drivers in standard 0.25μm CMOS process with $(W/L)_n=40$ and $(W/L)_p=100$, driven by five smaller inverters with $(W/L)_n=16$ and $(W/L)_p=40$. The drivers switch simultaneously while driving five 2pF capacitors. On-chip power and ground wires are connected to the off-chip power and ground traces through bond wires and package pins whose parasitics are modeled by series RL circuits ($R=0.5$, $L=10nH$). The ground bounce is a periodic function whose half-period variation for the rising edge of the input signal to the second driver is predicted using the detailed analysis provided in section III. Interestingly, the power supply noise for a balanced driver is the reverse of a shifted version of the ground bounce by half a period as also indicated in Fig. 15 (a):

$$v_n^{supply}(t) = V_{DD} - v_n\left(t - \frac{T}{2}\right) \tag{19}$$

Fig. 15 (b) demonstrates the result of HSPICE simulation on the same circuit, but in the presence of a 20pF decoupling capacitor. The 20pF decoupling capacitor cancels out the differential component of the power/ground fluctuations and causes the power and ground fluctuations to be nearly identical.
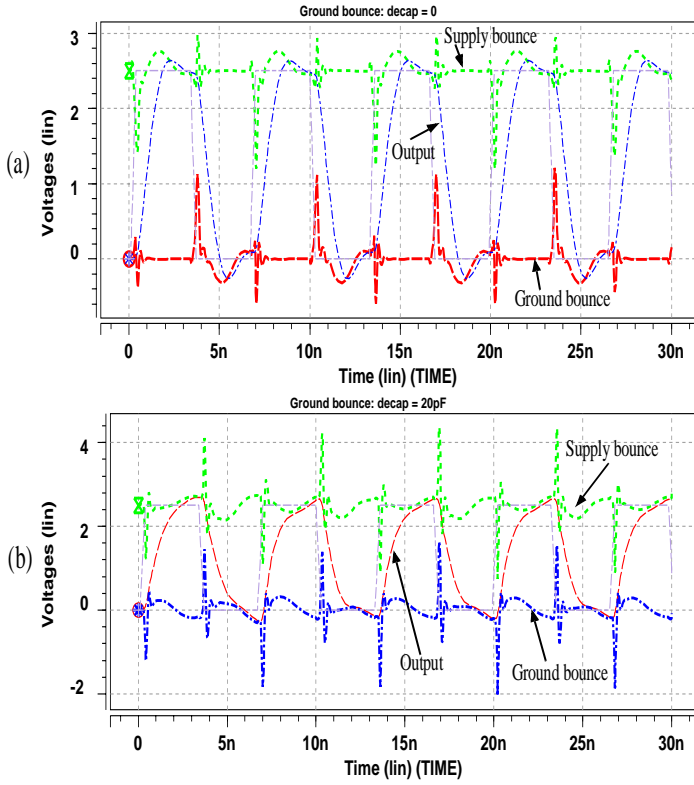
10

Fig. 15. Power/ground noise caused by simultaneous switching of five inverters. (a) power/ground noise without decoupling capacitor. (b) power/ground noise with a 20pF decoupling capacitor.

A very large decoupling capacitor reduces the peak values of the effective fluctuations on power and ground wires which is referred to as the effective P/G noise ($v_n^{P/G}(t) = v_n^{supply}(t) - v_n(t)$), and smooths out the oscillations. Fig. 16 shows the effective P/G noise of the same circuit with different values of decoupling capacitors. A large decoupling capacitor reduces the peak value of the effect P/G noise and causes the effective P/G noise to become a sinusoidal waveform with a period equal to half of the clock cycle time.
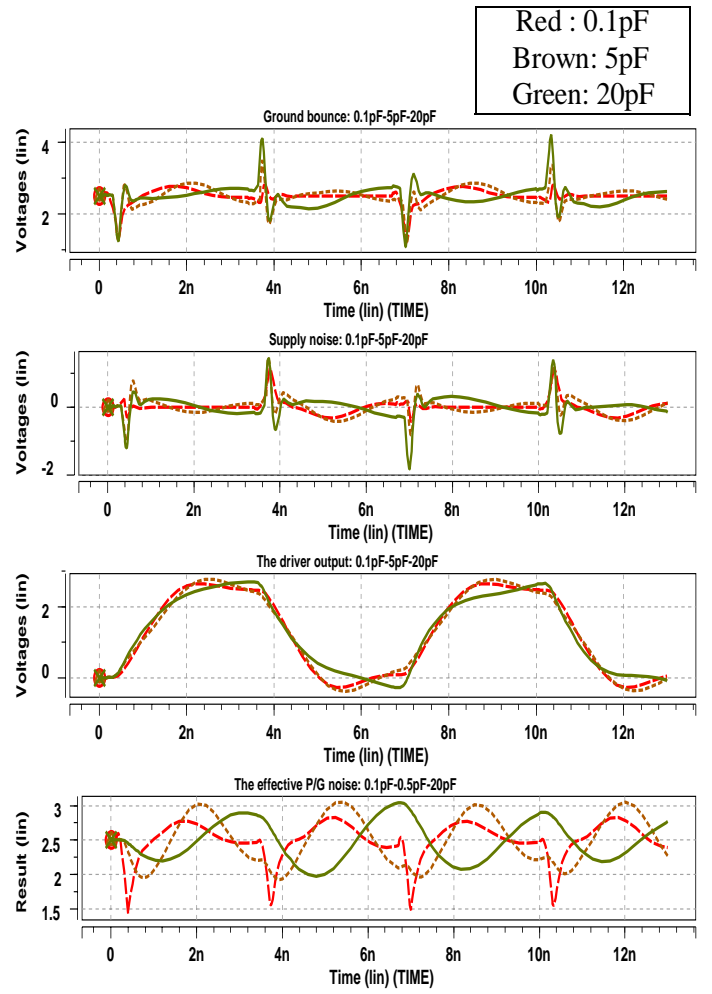

Fig. 16. Power/ground noise with different values of decoupling capacitors. (a) ground bounce. (b) power-supply noise. (c) driver output. (d) the effective P/G noise.

From Fig. 16 one concludes that the effective P/G noise can be decomposed into a common-mode component and a differential-mode component. The output pad buffers consist of tapered inverter chains 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 busses, also shows up on the input line. This implies that the common-mode component of bounces on supply and ground wires due to chip-package parasitics does not affect the circuit performance. In the presence of a large on-chip decoupling capacitor, power and ground fluctuations become in-phase signals and the differential-mode component will be filtered out. Therefore, 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.

11

Fig. 17 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.
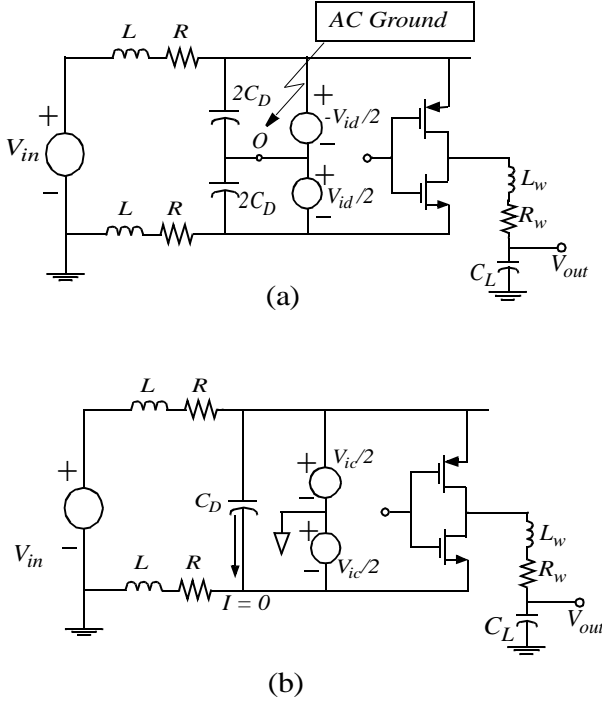

(a)

(b)

Fig. 17. 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.

Similar to the signal analysis of a differential amplifier [13], for the differential-mode circuit, the decoupling capacitor is virtually replaced by two identical capacitors each twice the original decoupling capacitance value. The two virtual voltage sources, $V_{id}/2$ and $-V_{id}/2$, exhibit the differential-mode component of the effective *P/G* noise. Shown in Fig. 17(a), these two voltage sources are 180 degrees out of phase, and node *O* thus becomes an AC ground. The equivalent decoupling capacitor, $2C_D$, 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 supply voltage and electrical parameters of the circuit is:

$$\frac{d^2 v_{nd}}{dt^2} + 2\alpha_c \frac{dv_{nd}}{dt} + \omega_{n,c}^2 v_{nd} = E_1 \frac{t}{t_r} + E_0 \qquad 0 \le t \le t_r \quad (20)$$

where

$$\omega_{n,c}^2 = \frac{1 + R\beta_n}{2LC_D} \quad ; \quad \alpha_c = 0.5\left(\frac{\beta_n}{2C_D} + \frac{R}{L}\right) \quad ;$$

$$E_0 = \frac{\beta_n}{2C_D}\left(\frac{V_{DD}}{t_r} - \frac{R}{L}V_{tn}\right); \quad E_1 = \frac{R\beta_n}{2LC_D}V_{DD}.$$

Eq. (20) is a second-order ODE, which causes the voltage $v_{nd}(t)$ to exhibit two different responses in the time-interval

$[0, t_r]$ depending on the circuit electrical values. If $\alpha_c \ge \omega_{n,c}$, the circuit shows an overdamped response, whereas if $\alpha_c < \omega_{n,c}$, the circuit shows an underdamped response. In both cases $v_{nd}(t)$ will increase with time over the time-interval $[0, t_r]$. We obtain a closed-form relationship between the maximum value of ground bounce and the on-chip decoupling capacitor for each of the two responses. This relationship can help circuit designers choose the correct amount of the decoupling capacitor based on a certain allowable peak value of the differential-mode component of the ground bounce.

A. Overdamped response

With a sufficiently large value, the decoupling capacitor, C_D, will be able to smooth out the ringing. In practice, the chip-package interface parasitic resistance and inductance vary between 0.4Ω-2Ω and 2nH-15nH, respectively. For an overdamped response over the time-interval $[0, t_r]$, the decoupling capacitance needs to be at least 3nF which is excessively large. Therefore, the overdamped response rarely occurs in reality [19][20]. The peak value of the ground bounce is obtained approximately by solving Eq. (20) and setting $t=t_r$ which yields:

$$v_{nd}(t_r) = \frac{E_1}{\omega_{n,c}^2} - \frac{2\alpha_c E_1 t_r}{\omega_{n,c}^4} + \frac{E_1/t_r - E_0 p_1}{p_1^2(p_2 - p_1)}e^{-p_1 t_r}$$
$$+ \frac{E_1/t_r - E_0 p_2}{p_2^2(p_1 - p_2)}e^{-p_2 t_r} \tag{21}$$

where $p_{1,2} = \alpha_c \pm \sqrt{\alpha_c^2 - \omega_{n,c}^2}$

Eq. (21) is utilized to obtain the relationship between the peak value of the differential-mode component of the ground bounce and the decoupling capacitor as also shown in Fig. 18. It is easily verified that $v_{nd}(t_r)$ is a monotonically decreasing function in terms of C_D. In this case the v_{nd} waveform after adding the decoupling capacitor is an exponential-like waveform.

B. Underdamped response

If $\alpha_c < \omega_{n,c}$, the circuit shows a damped oscillatory transitions in the interval $[0, t_r]$. A large decoupling capacitor will be able to reduce the differential-mode of the effective *P/G* noise to negligible values. However, the noise waveform still experiences small ringings because the chip-package interface parasitic resistance is small.

Once again, the peak value of the ground bounce is obtained approximately by solving Eq. (20) and setting $t=t_r$ which yields:

$$v_{nd}(t_r) = \frac{E_1}{\omega_{n,c}^2} - \frac{2\alpha_c E_1 t_r}{\omega_{n,c}^4} + \frac{1}{\omega_{n,c}\omega_{d,c}} \times$$
$$\left[\frac{E_1}{t_r}\frac{1}{\omega_{n,c}}\sin(\omega_{d,c} t_r - 2\Phi_c) - E_0 \sin(\omega_{d,c} t_r + \Phi_c)\right]e^{-\alpha_c t_r} \tag{22}$$

Once again, it is verified that $v_{nd}(t_r)$ is a monotonically decreasing function of C_D because $\frac{\partial v_{nd}(t_r)}{\partial C_D} < 0$. This is depicted in Fig. 18. Therefore, by increasing its value the differential-mode component of the noise can be reduced to arbitrarily small values. Furthermore, $v_{nd}(t_r)$ is also a monotonically decreasing function of C_D and its value approaches zero for sufficiently large values of C_D. In this case the v_{nd} waveform after adding the

decoupling capacitor is a damped oscillatory waveform. Fig. 18 shows the variation of the peak value of the ground bounce in terms of C_D. In Fig. 18, if $\alpha_c \geq \omega_{n,c}$, Eq. (21) is utilized whereas if $\alpha_c < \omega_{n,c}$, Eq. (22) is used.
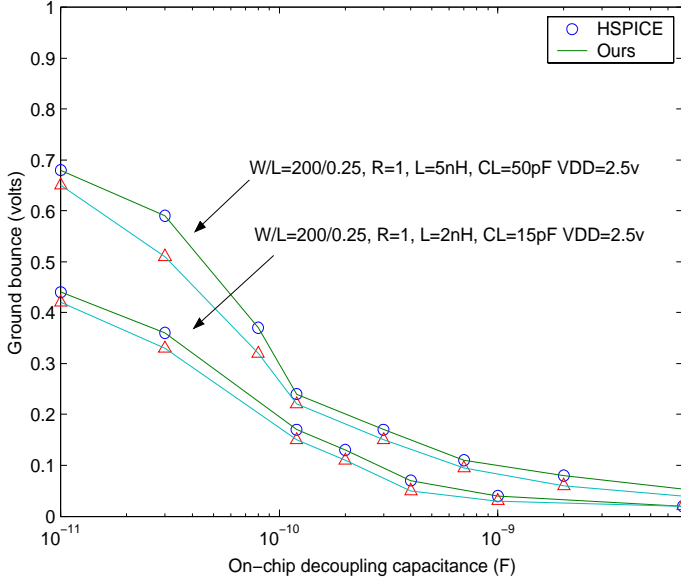

Fig. 18. Variation of the differential-mode part of ground bounce vs. the on-chip decoupling capacitor for three different *W/L* ratios.

VI. 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 prohibit all the buffers from switching simultaneously. This is easily done by inserting a chain of buffers in the signal path to the output drivers. Based on a special property of the ground bouce waveform, one can 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 described below.

As shown in section III.B.2 the ground bounce declines toward zero as a damped oscillatory waveform and therefore it experiences an undershoot. If the switching time of the next driver is tuned to occur at exactly the same time that the ground bounce passes through its undershoot point, then the peak value of the ground bounce will be maximally attenuated. Suppose as before 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.

$$min \quad v_n(t_r) \quad (23)$$

$$\textbf{s.t.} \quad \sum_{i=1}^{k} \tau_i \leq T_c \qquad 1 \leq k \leq N+M$$

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 the number of drivers that 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 must determine the time when this undershoot occurs. Differentiating Eq. (11) with respect to time variable t gives the time τ_d at which the waveform experiences an undershoot.

$$\tau_d = t_s + \Delta t \quad \text{where} \quad \Delta t = \frac{1}{\omega'_d} \tan^{-1} \left(\frac{V_s \tan \Phi' - V_c}{V_s + V_c \tan \Phi'} \right) \quad (24)$$

where $\Phi' = \tan^{-1} \left(\frac{\omega'_d}{\alpha'} \right)$. 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. 19. $\lfloor T_c / \tau_d \rfloor - 1$ drivers are equally triggered by τ_d seconds from each other. The rest of drivers are triggered such that the $\lfloor T_c / \tau_d \rfloor + 1$ *st* driver switches simultaneously with the first driver. The $\lfloor T_c / \tau_d \rfloor + 2nd$ driver switches simultaneously with the second driver and so on. Fig. 19 depicts the skew control of three output drivers under the assumption that the time constraint $T_c = T/2$ half of the clock period.
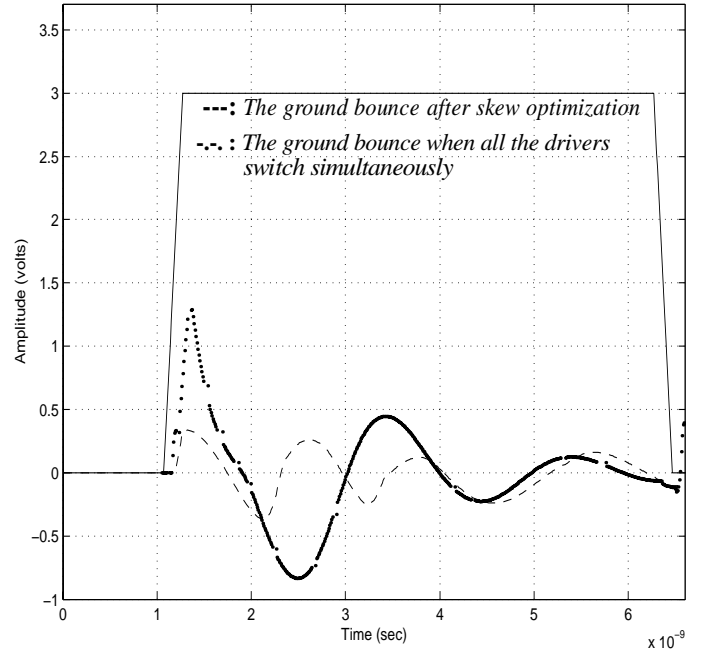

Fig. 19. Ground bounce control by skewing the switching times of three drivers.

In practice, the parasitics of the chip-package interface are widely unknown in the early stages of the circuit design. However, skewing the off-chip drivers to prevent them from switching simultaneously will considerably reduce the ground bounce. As a practical approximation, we can assume τ_d to be roughly equal to three-four times the input rise-time and use the above methodology under this new assumption for the buffer skewing.

VII. CONCLUSION

A detailed analysis and optimization of the off-chip ground bounce using an accurate and simple chip-package interface circuit model was proposed. The effect of ground bounce on the tapered buffer design was studied, and a mathematical analysis was introduced. Next the effect of the 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 proposed. 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] Intel Packaging Databook, http://www.intel.com/design/ packtech/packbook.htm.

[12] AMD Packaging Design, http://www.amd.com/us-en/Processors/ProductInformation/0,,30_118_1850_1860,00.html.

[13] P. R. Gray, P. J. Hurst, S. H. Lewis, R. G. Meyer, Analysis and Design of Analog Integrated Circuits, pp. 66-71, John Wiley and Sons, 2001.

[14] N.Hedenstierna, K. O. Jeppson, "CMOS Circuit Speed and Buffer Optimization," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, No. 2, pp. 270-281, March 1987.

[15] 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.

[16] A. Chatzigeorgiou,S. Nikolaidis, I. Tsoukalas, "Modeling CMOS Gates Driving RC Interconnect Loads," *IEEE Trans. Circuits and System II: Analog and Digital Signal Processing*, vol. 48, No. 4, pp. 413-418, April 2001.

[17] K. T. Tang, E. G. Friedman, "Delay and Power Expressions Characterizing a CMOS Inverter Driving an RLC Load," *IEEE Int'l Symp. on Circuits and Systems*, pp. III-283-III-286, May 2000.

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

[19] T. Gabara, W. Fischer, "Capacitive coupling and quantized feedback applied to conventional CMOS technology," *IEEE Proc. Custom Integrated Circuits Conf.*, pp. 281-284, May 1996.

[20] R. Evans, "Effects of losses on signals in PWB's," *IEEE Trans. Comp., Packag., and Manufact. Technol. - Part B*, vol. 17, No. 2, pp. 217-222, May 1994.

Payam Heydari (S'98-M'00) his B.S. degree in Electronics Engineering and M.S. degree in Electrical Engineering from the Sharif University of Technology, Tehran, Iran in 1992 and 1995, respectively. He received his Ph.D. degree in Electrical Engineering at the University of Southern California in 2001.

During the summer of 1997, he was with Bell-labs, Lucent Technologies where he worked on noise analysis in deep submicron VLSI circuits. He worked at IBM T. J. Watson Research Center on gradient-based optimization and sensitivity analysis of custom integrated circuits during the summer of 1998. Since August 2001, he has been an Assistant Professor of electrical engineering at the University of California, Irvine, where his research interests are design of high-speed analog, RF, and mixed-signal integrated circuits, and analysis of signal integrity and high-frequency effects of on-chip interconnects in high-speed VLSI circuits.

Dr. Heydari has received the Best Paper Award at the 2000 IEEE International Conference on Computer Design (ICCD). He has also received the Technical Excellence Award from the Association of Professors and Scholars of Iranian Heritage in California in 2001. He serves as a member of the Technical Program Committees of the IEEE Design and Test in Europe (DATE), the International Symposium on Physical Design (ISPD), and the International Symposium on Quality Electronic Design (ISQED).

Massoud Pedram received a B.S. degree in Electrical Engineering from the California Institute of Technology in 1986 and M.S. and Ph.D. degrees in Electrical Engineering and Computer Sciences from the University of California, Berkeley in 1989 and 1991, respectively. He then joined the department of Electrical Engineering - Systems at the University of Southern California where he is currently a professor.

Dr. Pedram has served on the technical program committee of a number of conferences, including the Design automation Conference (DAC), Design and Test in Europe Conference (DATE), Asia-Pacific Design automation Conference (ASP-DAC), and International Conference on Computer Aided Design (ICCAD). He served as the Technical Co-chair and General Co-chair of the International Symposium on Low Power Electronics and Design (SLPED) in 1996 and 1997, respectively. He was the Technical Program Chair of the 2002 International Symposium on Physical Design. Dr. Pedram has published four books, 60 jour-

nal papers, and more than 150 conference papers. His research has received a number of awards including two ICCD Best Paper Awards, a Distinguished Paper Citation from ICCAD, a DAC Best Paper Award, and an IEEE Transactions on VLSI Systems Best Paper Award. He is a recipient of the NSF's Young Investigator Award (1994) and the Presidential Faculty Fellows Award (a.k.a. PECASE Award) (1996).

Dr. Pedram is a Fellow of the IEEE, a member of the Board of Governors for the IEEE Circuits and systems Society, an IEEE Solid State Circuits Society Distinguished Lecturer, a board member of the ACM Interest Group on Design Automation, and an associate editor of the IEEE Transactions on Computer Aided Design, the IEEE Transactions on Circuits and Systems, and the ACM Transactions on Design Automation of Electronic Systems. His current work focuses on developing computer aided design methodologies and techniques for low power design, synthesis, and physical design. For more information, please go to URL address: http://atrak.usc.edu/~massoud/.