Timing Uncertainty in 3-D Clock Trees Due to Process Variations and Power Supply Noise

Hu Xu, Vasilis F. Pavlidis, Xifan Tang, Wayne Burleson, and Giovanni De Micheli

***Abstract*— Clock distribution networks are affected by different sources of variations. The resulting clock uncertainty significantly affects the frequency of a circuit. To support this analysis, a statistical model of skitter, which consists of clock skew and jitter, for 3-D clock trees is introduced. The effect of skitter on both the setup and hold time slacks is modeled. The variation of skitter is shown to be underestimated up to 36% if process variations and dynamic power supply noise are considered separately, which highlights the importance of this unified treatment. Potential scenarios of supply noise in 3-D integrated circuits (ICs) are investigated. 3-D circuits generated from industrial benchmarks are simulated to show the skitter under these scenarios. The mean and standard deviation of skitter can vary up to 60% and 51%, respectively, due to the different amplitudes and phases of supply noise. The tradeoff between skitter and the power consumed by clock trees is also shown. A set of guidelines are presented to decrease skitter in 3-D ICs. By applying these guidelines to industrial benchmarks, simulations show a decrease in the mean skitter up to 31%.**

***Index Terms*— 3-D ICs, clock jitter, clock skew, clock tree, power supply noise, process variations, skitter.**

I. Introduction

In very deep submicrometer integrated circuits, clock distribution networks are significantly affected by different sources of variations, such as static process variations, dynamic voltage, and thermal variations [1]. The resulting clock uncertainty due to these variations consists of clock skew and jitter. These uncertainties can severely constrain the highest clock frequency of a circuit. In addition, the design of robust clock distribution networks requires a comprehensive analysis and proper mitigation of these variations.

3-D integration emerges as a promising solution to alleviate the increasing interconnect delay and to enhance the density of devices in modern integrated circuits [2]. Multiple planar circuits (tiers) with different technologies can be vertically stacked, which complicates the variation analysis for 3-D circuits. The effect of process variations in 3-D integrated circuits (ICs) has been discussed for both datapaths [3], [4] and clock distribution networks [5]–[8]; the latter is the focus of this paper. Process variations are modeled as die-to-die (D2D) and within-die (WID) variations [9]. Transistors and interconnects within one die are uniformly affected by the same D2D variations. WID variations, nevertheless, affect these devices both randomly and systematically [10]. Since clock paths span more than one tier in 3-D ICs, the process variability of clock distribution networks differs from 2-D ICs [5]–[8]. Dissimilar to 2-D ICs, 3-D clock distribution networks are affected by D2D variations from more than one tier. This situation complicates the analysis of process variations in 3-D clock paths and the distribution of clock paths among tiers significantly affects the statistical skew variations.

The fluctuation of supply voltage, called power supply noise, is another source of variations. This supply noise significantly affects the electrical characteristics of clock buffers [1]. The effect of power supply noise on 2-D clock distribution networks has been investigated in [11]–[13]. The effect of power supply noise on 3-D clock distribution networks, however, has not been adequately explored.

For 3-D ICs, different power distribution networks (PDNs) have been investigated. A 3-D PDN similar to a 2-D network is implemented in [14], while 3-D PDNs with different characteristics among tiers are discussed in [15]. The effect of this different power supply noise on the timing uncertainty of 3-D clock distribution networks, however, remains unclear. The change of the clock uncertainty with both the different characteristics of supply noise and process variations is investigated in this paper.

In most of the previous works, the effect of process variations and power supply noise is discussed separately in terms of skew and jitter. Clock skew, the difference in delay among clock paths, is considered to be significantly affected by process variations and is well modeled for 2-D ICs [10], [16]. A model of process-induced skew in 3-D clock trees is proposed in [5]. The other constituent of clock uncertainty is clock jitter, which is the deviation of the edge of clock signal from the ideal temporal occurrences. Clock jitter can be described in three ways: period jitter, cycle-cycle jitter, and phase jitter (or time interval error) [1]. Period jitter is the difference between the measured clock period and the ideal period, which is the most explicit description of the clock jitter within a circuit. Jitter is produced by the phase-locked loop (PLL) and the clock distribution network. PLL jitter can be mitigated by careful PLL design [17]. The jitter produced in

Manuscript received April 24, 2012; revised October 2, 2012; accepted November 4, 2012. Date of publication January 11, 2013; date of current version October 14, 2013. This work was supported in part by the Swiss National Science Foundation under Grant 260021 126517/1, the European Research Council under Grant 246810 NANOSYS, and Intel Braunschweig Labs, Germany.

H. Xu, X. Tang, and G. De Micheli are with the Integrated Systems Laboratory, EPFL, Lausanne 1015, Switzerland (e-mail: hughesxuh@gmail.com; xifan.tang@epfl.ch; giovanni.demicheli@epfl.ch).

V. F. Pavlidis is with the School of Computer Science, University of Manchester, Manchester M13 9PL, U.K. (e-mail: pavlidis@cs.man.ac.uk).

W. Burleson is with the Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, MA 01003 USA (e-mail: burleson@ecs.umass.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2012.2230035

1063-8210 © 2013 IEEE

clock distribution networks is mainly due to the power supply noise on the clock buffers [18]. The effect of the power supply noise on period jitter in 2-D ICs is analyzed in [11] and [13], while to the best of authors' knowledge, this paper discusses period jitter in 3-D ICs for the first time.

Nevertheless, clock distribution networks are simultaneously affected by process variations and power supply noise. For 2-D ICs, a statistical timing analysis method considering process variations and power supply noise is proposed in [19], where full-chip simulations are required to obtain the distribution of power supply noise. Moreover, the effect of these variations on clock distribution networks is not adequately explored. The combination of skew and jitter, "skitter," is introduced in [20] to model the co-effect of all sources of variations on clock distribution networks, while no closed-form formula is given to model the distribution of skitter. A subcircuit is designed to measure the skitter in [20], which can be utilized to mitigate undesired skitter during operation [21]. If the skitter is high, frequent recovery, and adaptation procedures have to be executed to correctly transfer data. Moreover, these architectural procedures cannot be used for each pair of clock sinks. Consequently, by better understanding the behavior of skitter, this component of clock uncertainty can be mitigated through the proper design of clock distribution networks. In addition, the overhead of the adaptive circuits and architectural procedures can be reduced. A simplified model for skitter in 2-D ICs is proposed in [22], where only the uniform D2D variations and supply noise are considered.

The combined effect of process variations and dynamic power supply noise on 3-D clock distribution networks has not been explored, although clock skew and jitter need to be treated cohesively. Consequently, both theoretical insight and practical design issues on this effect are presented in this paper. The main contributions of this paper are as follows.

1) A statistical model for skitter consisting of skew and jitter is proposed for 3-D clock trees.
2) The skitter is investigated under different scenarios of dynamic power supply noise in 3-D ICs. Simulations on industrial benchmarks show that separately treating process variations and supply noise can significantly underestimate the clock uncertainty.
3) The tradeoffs among skitter and power are presented. The allocation of buffers among tiers is also investigated, where the effect of the third physical dimension in timing uncertainty is described.
4) Based on the observed behavior, a set of design guidelines are presented to mitigate skitter in 3-D clock distribution networks. Case studies are presented to illustrate the efficiency of these guidelines in the design process of 3-D clock distribution networks.

The remainder of this paper is organized as follows. A methodology to obtain the delay variation of a buffer stage is presented in the following section. A statistical model to describe the skitter in 3-D clock trees is presented in Section III. Simulation results and discussions are presented in Section IV. The skitter is analyzed under different scenarios of power supply noise in a 3-D circuit. The necessity of simultaneously considering skew and jitter is demonstrated. The tradeoff between skitter and power consumption follows. In Section V, related design guidelines are proposed to mitigate skitter, lowering the complexity and/or the number of circuits required to adapt clock frequency to prevent from timing failures. Case studies on 3-D ICs generated from industrial benchmarks are also presented. The conclusions are drawn in Section VI.

II. Delay Variation of a Buffer Stage

The distribution of the delay of a buffer stage is modeled in this section. The delay of a buffer stage d consists of the delay of the buffer d_{b} and the interconnect [horizontal wire and/or through silicon via (TSV)] d_{I}. The variation of d is a random variable affected by both process variations and power supply noise.

A. Delay Variation Due to Process Variations

Since the variation of parameters due to process variations is typically within a small range, the delay of a buffer stage considering the parameter variations can be approximated by the first order Taylor expansion [23]

$$\begin{aligned} d(tr, \vec{P}, C_{\mathrm{lw}}) &= d_{\mathrm{b}}(tr, \vec{P}, C_{\mathrm{lb}}) + d_{\mathrm{I}}(\vec{P}, C_{\mathrm{lw}}) \\ &\approx \overline{d} + \sum_{p \in \vec{P}} \left(\left.\frac{\partial d}{\partial p}\right|_0 \Delta p \right). \end{aligned} \quad (1)$$

The input slew of this buffer stage is denoted by tr. The capacitive load seen at the output of the buffer and wire is denoted by C_{lb} and C_{lw}, respectively. The nominal delay is $\overline{d}$ and the subscript "0" denotes the partial derivative with nominal parameters. The set of parameters affected by process variations is denoted by $\vec{P}$. Each parameter is modeled by a random variable. For instance, if the variation of channel length of three buffers is considered, $\vec{P}$ is $\{L_{\mathrm{b},1}, L_{\mathrm{b},2}, L_{\mathrm{b},3}\}$. The variation of a parameter Δp consists of WID and D2D variations

$$\Delta p = \Delta p_{\mathrm{WID}} + \Delta p_{\mathrm{D2D}} \quad (2)$$

where Δp_{D2D} is consistent among buffers (interconnects) within the same die, while Δp_{WID} varies among the components within the same die [9]. The partial derivatives in (1) are determined by

$$\begin{aligned} \frac{\partial d}{\partial p} = &\frac{\partial d_{\mathrm{b}}}{\partial tr}\frac{\partial tr}{\partial p} + \frac{\partial d_{\mathrm{b}}}{\partial C_{\mathrm{lb}}}\frac{\partial C_{\mathrm{lb}}}{\partial p} + \frac{\partial d_{\mathrm{b}}}{\partial p} \\ &+ \frac{\partial d_{\mathrm{I}}}{\partial C_{\mathrm{lw}}}\frac{\partial C_{\mathrm{lw}}}{\partial p} + \frac{\partial d_{\mathrm{I}}}{\partial p}. \end{aligned} \quad (3)$$

The partial derivatives in (3) are determined by the expressions of d_{b} and d_{I}. The expression of $d_{\mathrm{b}}(tr, \vec{P}, C_{\mathrm{lb}})$ can be obtained through analytic formulas [24] or adjoint sensitivity analysis with SPICE-based simulations [23]. To achieve higher accuracy, the latter is used in this paper. For horizontal wires, the expression of $d_{\mathrm{I}}(\vec{P}, C_{\mathrm{lw}})$ is determined by the RLC interconnect model proposed in [25] and [26].

The variations introduced by TSVs have been discussed in [7] and [8], where the TSV stress-induced delay variation of buffers is well modeled. In this paper, the keep-out-zone of

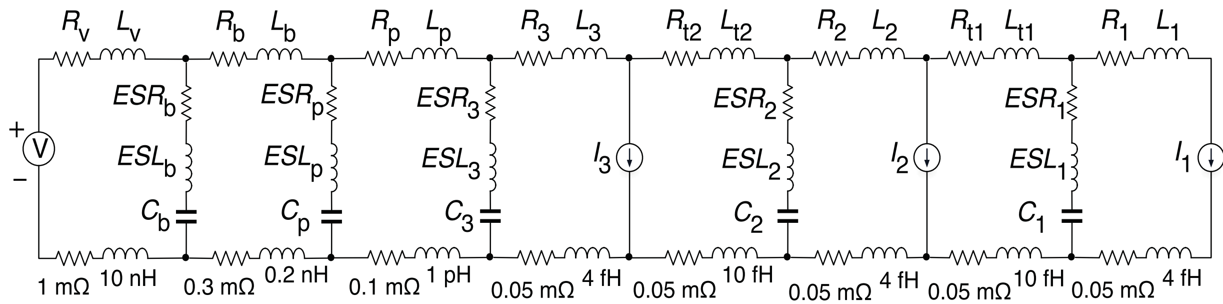


Fig. 1. Simplified circuit used to simulate the global resonant power supply noise.

TSVs is assumed to be large enough ($\leq$10 μm [8], [27]) to mitigate the effect of TSV stress. Consequently, TSVs are modeled as RLC wires with different electrical characteristics from the horizontal interconnects.

B. Power Supply Noise in 3-D ICs

In addition to process variations, the power supply noise also affects the variation of buffer delay. The applied supply voltage $V_{\rm dd}$ is determined by the supply noise υ, $V_{\rm dd} = V_{\rm dd0} + \upsilon$. To efficiently model the power supply noise in a 3-D IC, the following points are considered.

1) High-Frequency Power Supply Noise: The power supply noise with high frequency ($>$ 5 GHz) is due to fast current transients of local devices [28], e.g., caused by clock edges. This high-frequency noise tends to remain highly localized due to the small high-frequency inductive current loops, the frequent return paths provided by power rails, and the fast energy dissipation by the parasitic RLC of the power grid. High-frequency current lasts for only a short time and will be transformed to a low-frequency noise determined by die–package interaction similar to the resonant noise [28].

2) Mid-Frequency Power Supply Noise: Although high-frequency power supply noise is extremely localized and quickly diminishes, the mid-frequency power supply noise (1–2 GHz) caused by mid-low frequency transients affects the performance of neighboring gates [28]. This mid-frequency supply noise will also be transformed to low-frequency resonant noise. Nevertheless, the duration of the mid-frequency noise is longer than the high-frequency noise.

High- and mid-frequency power supply noise can be directly described by random variables with probabilistic formulations, as modeled in [29]. These random variables can be directly included in $\vec{P}$ in (1). To obtain the distribution of these variables, the switching activity of all the cells is required. Afterwards, a full-chip transient simulation of the PDN is performed to determine the temporal and spatial change of the power supply noise. Moreover, both the high- and mid-frequency supply noise can be significantly reduced by RC filters [30] distributed across the chip. Consequently, the focus of this paper is mainly on the low-frequency and long lasting supply noise.

3) Low-Frequency Power Supply Noise: Low-frequency power supply noise (lower than hundreds of MHz) is determined by the die-level resonant supply noise [11], [30]. This noise is typically stimulated by the simultaneous switching of large number of transistors, e.g., a wakeup operation. This resonant supply noise globally affects each tier [13], [15], [28], [30], [31]. A simplified model used to simulate the resonant supply noise in a complete 3-D PDN is illustrated in Fig. 1 [13], [15], [30], [31]. A three-tier circuit is shown in this figure. The power/ground (P/G) signal is supplied from the voltage regulator module (VRM), through the board and package to the circuit.

Resistances and inductances are denoted by R and L, respectively, and are assumed to be the same for both the power and ground paths. The VRM, the board, the package, and the third, the second, and the first tiers are denoted by the subscripts υ, b, p, 3, 2, and 1, respectively. These notations are shown along with the elements of the power path while the corresponding values are depicted along the ground path. The decoupling capacitors are denoted by C, while the equivalent series resistance and inductance of these capacitors are denoted by ESR and ESL, respectively. The transient currents in different tiers are denoted by current sources I. The resistance and inductance of TSVs are denoted by $R_{\rm t}$ and $L_{\rm t}$, respectively.

The resonant supply noise seen by different tiers corresponding to a wakeup operation (i.e., in this context a significant current demand) is illustrated in Fig. 2(a). Due to the decoupling capacitors at different levels, three voltage droops with different resonant frequencies can be seen in each waveform of resonant supply noise [13]. These voltage droops are determined by the supply impedance of the PDN at different frequencies, as illustrated in Fig. 2(b). The supply impedance at different frequencies is determined by the RLC characteristics of the PDN.

1) The first droop of the resonant supply noise: this droop of resonant noise is mainly determined by the LC tank formed between the package inductance and the on-chip capacitance. Due to the limited number of on-chip decoupling capacitors, the first droop is typically the worst resonant supply noise and the major concern for a PDN [11]–[13], [28]. The frequency of the first droop of the supply noise is typically between tens of MHz and 400 MHz [11].

   Note that in some works, the first droop of power supply noise denotes the high- or mid-frequency power supply noise [28]. As previously mentioned, this supply noise affects a circuit locally. To avoid confusion, the first droop of resonant supply noise refers to the global resonant noise determined by the package inductance and on-chip capacitance in the following context.

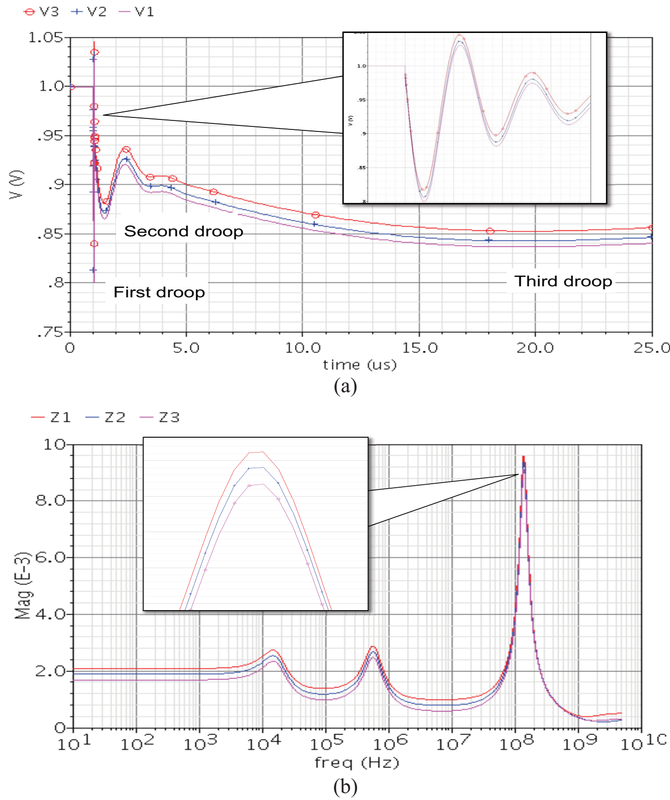


Fig. 2. Resonant power supply noise in 3-D ICs, where (a) and (b) are the resonant supply noise and the impedance of the PDN in different tiers, respectively.

2) The second droop of the resonant supply noise: this droop is determined by the package and board decoupling capacitors. The second droop is much smaller and slower than the first droop due to larger decoupling capacitors in the package.
3) The third droop of the resonant supply noise: this droop is determined by the board decoupling capacitors. Since large capacitors can be used on the board, the third droop can be efficiently mitigated although it lasts for a long time [13].

If there is no decoupling capacitance in the package, two main droops will be seen in the waveform of the power supply noise. Since the first droop of the resonant supply noise is typically the deepest, it is the main focus in this paper. Consequently, "resonant supply noise" refers to the first droop of the resonant supply noise in the following context for simplicity.

The damped sinusoidal waveform can be used to describe the worst resonant noise [11]–[13]. Assuming a clock edge arrives at the source of a clock path at time zero, t_j is the time when this clock edge arrives at buffer j. The supply noise to buffer j at time t_j can be expressed as

$$v(t_j) = V_n e^{-\epsilon t_j} \sin(2\pi f_n t_j + \phi) \quad (4)$$

$$t_j = \sum_{i=1}^{j-1} d_i. \quad (5)$$

The clock frequency is much higher than the resonant noise frequency and the clock path delay is, typically, lower than the period of the resonant noise. Due to the deep voltage drop, the first period of the resonant noise causes the worst clock jitter [11]. Consequently, to investigate the effect of the worst supply noise on clock distribution networks, (4) can be approximated by an undamped sinusoidal waveform [11]

$$v(t_j) \approx V_n \sin(2\pi f_n t_j + \phi). \quad (6)$$

According to (1) and (5), d_i, t_j, and $v(t_j)$ are all random variables. Since Δt_j is low as compared with $\bar{t}_j$, $v(t_j)$ can also be approximated by the first order Taylor expansion

$$v(t_j) \; 7 = \; \bar{v}(t_j) + \Delta v(t_j) \approx \bar{v}(t_j) + \left.\frac{\partial v(t_j)}{\partial t_j}\right|_0 \Delta t_j \quad (7)$$

$$\Delta v(t_j) \approx 2\pi V_n f_n \cos(2\pi f_n \bar{t}_j + \phi) \sum_{i=1}^{j-1} \Delta d_i. \quad (8)$$

The amplitude V_n and frequency f_n are determined by the switch current and the characteristics of the circuits. The initial phase ϕ is the phase of the resonant noise when the investigated clock edge arrives at the source of the clock path.

4) Resonant Noise Versus On-Chip Current: In 3-D ICs, the current dissipated by the tiers can differ due to the different numbers and sizes of devices. The amplitude and the frequency of the resonant supply noise change with the current within different tiers. The resonant noise corresponding to four cases of switching current is simulated for the PDN shown in Fig. 1. These switching currents are listed in Table I. The pulse width and the rise and fall time of the switching current are all 1 ns. The resulting V_n and f_n are reported in Fig. 3.

As illustrated in Fig. 3, different current distribution introduces nonnegligible difference in V_n among tiers (ΔV_n). Both the different IR-drop and resonance impedance among tiers contribute to this ΔV_n. The resonant frequency is similar

TABLE I
FOUR CASES OF SWITCH CURRENT WITHIN A THREE-TIER CIRCUIT

Case	cur1	cur2	cur3	cur4
I_1 [A]	0	10	20	40
I_2 [A]	20	20	20	20
I_3 [A]	40	30	20	0

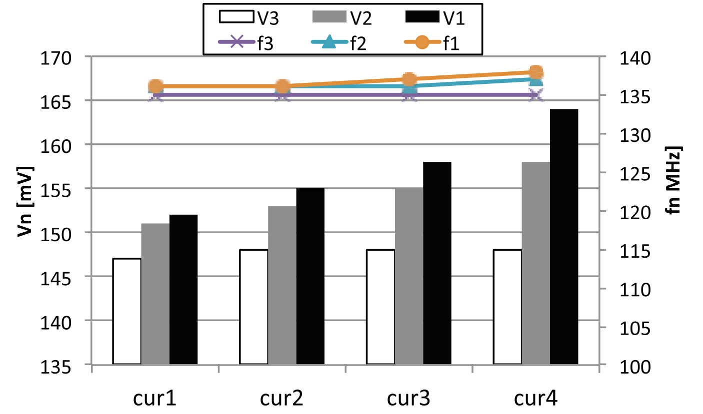


Fig. 3. Amplitude and frequency of the resonant noise versus switching current in different tiers.

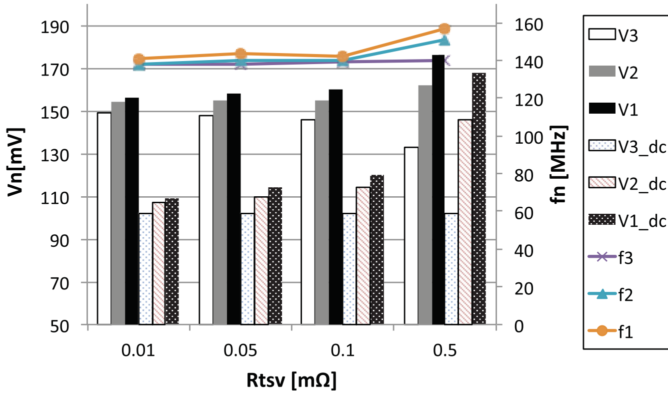


Fig. 4. Resonant supply noise and IR-drop versus total resistance of TSVs.

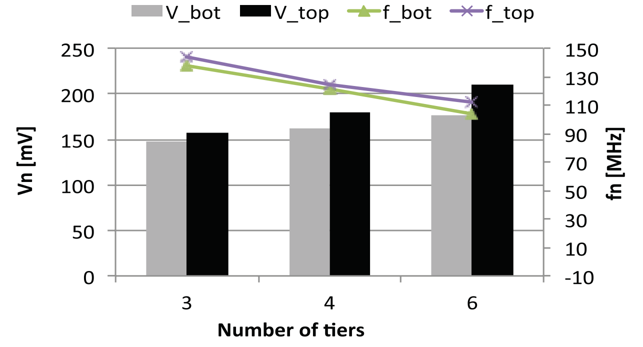


Fig. 5. Resonant noise versus the number of tiers.

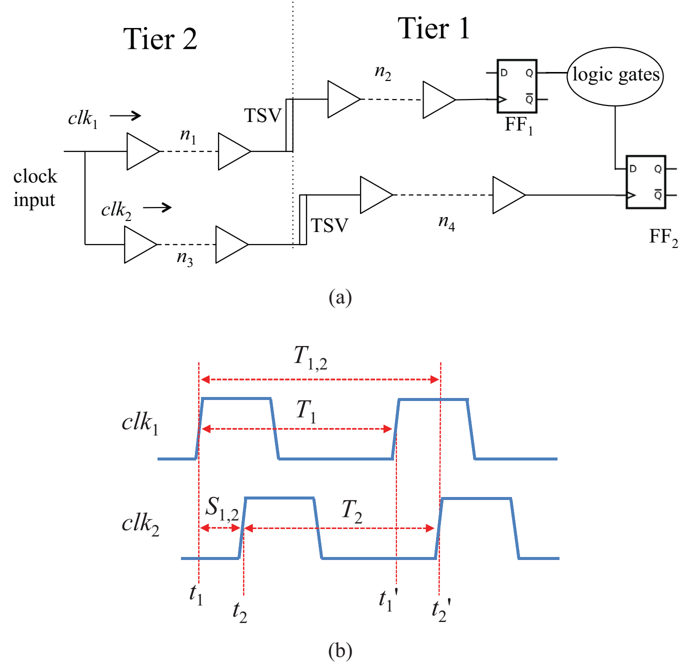


Fig. 6. Clock uncertainty between 3-D clock paths. (a) Two paths and flip-flops. (b) Corresponding clock signals.

among tiers ($\Delta f_n \leq 3$ MHz) and does not change significantly with the current.

5) Resonant Noise Versus Resistance of TSVs: The electrical characteristics of TSVs depend on the manufacturing technology [32], [33]. The change of the power supply noise with the total resistance of TSVs (R_{tsv}) is illustrated in Fig. 4. In this figure, the amplitude and frequency of the overall supply noise are denoted by V3–V1 and f3–f1, respectively. The DC IR-drop is denoted by V3_dc–V1_dc. Since the resonant noise is stimulated by a current pulse, the effect of IR-drop is also included in V3–V1.

Larger R_{tsv} introduces higher IR-drop to the first and second tiers. In the third tier, the DC IR-drop is not affected by R_{tsv}, since this tier is directly connected to the package (see Fig. 1). Nevertheless, higher R_{tsv} decreases the quality factor (Q factor) of the circuit in resonance, which decreases the amplitude of the resonance. Consequently, V_n in the third tier decreases with R_{tsv}.

In the first and second tiers, V_n is determined by both the resonance and the IR-drop. Consequently, V2 and V1 increase with R_{tsv} due to the significantly increased IR-drop. Nevertheless, the increase in V2 and V1 is not as high as the increase in the DC IR-drop due to the lower Q factor.

6) Resonant Noise Versus Number of Tiers: The resonant noise for different number of tiers in a 3-D IC is plotted in Fig. 5. The switch current and on-die capacitance are assumed identical for all tiers. As shown in Fig. 5, ΔV_n between the bottom and top tiers increases with the number of tiers. As more dies are vertically stacked, the difference in resonant noise among tiers increases.

C. Delay Variation Simultaneously Considering Process Variations and Power Supply Noise

According to (1), the delay variation Δd is also affected by the input slew Δtr, which is determined by the previous buffer stage. Considering the effect of Δv and Δtr on Δd, the delay variation of the jth buffer stage can be modeled as

$$\Delta d_j \approx \sum_{p\in\vec{P}_j}\left(\frac{\partial d_j}{\partial p}\bigg|_0 \Delta p\right)+\frac{\partial d_j}{\partial v}\bigg|_0 \Delta v(t_j)+\frac{\partial d_j}{\partial tr}\bigg|_0 \Delta tr_j. \quad (9)$$

The set of statistical parameters of the jth buffer stage is denoted by $\vec{P}_j$, which is a subset of the entire parameter set, $\vec{P}_j \subseteq \vec{P}$. The input slew of the jth buffer stage Δtr_j can be determined similar to (9)

$$\Delta tr_j \approx \sum_{p\in\vec{P}_j}\left(\frac{\partial tr_j}{\partial p}\bigg|_0 \Delta p\right)+\frac{\partial tr_j}{\partial v}\bigg|_0 \Delta v(t_{j-1}) + \frac{\partial tr_j}{\partial tr_{j-1}}\bigg|_0 \Delta tr_{j-1}. \quad (10)$$

Substituting (8) and (10) into (9), Δd_j can be recursively determined considering both process variations and power supply noise. The coefficients in (9) and (10) are obtained through adjoint sensitivity analysis as previously mentioned. The resulting expression (9) is used to determine skitter in the following section.

III. Model of Skitter in 3-D Clock Trees

The definition of the clock skew, period jitter, and skitter in this paper is illustrated in Fig. 6. The clock signal is fed into the 3-D clock tree from the primary clock driver.

Two flip-flops are driven by this clock signal, denoted as FF_1 and FF_2, respectively.

The waveforms clk_1 and clk_2 in Fig. 6(b) correspond to the clock signal driving FF_1 and FF_2, respectively. The time where the first rising edge in Fig. 6(b) arrives at the clock input is defined as the origin. The time when this edge arrives at FF_1 and FF_2 is, respectively, denoted by t_1 and t_2. The arrival time of the next rising edge is t_1' and t_2'. The numbers of buffers from the clock input to FF_1 and FF_2 are denoted by $n_1 + n_2$ and $n_3 + n_4$, respectively. The skew between the first edge of clk_1 and clk_2 is $S_{1,2}$. The measured clock periods after the first edge for FF_1 and FF_2 are T_1 and T_2, respectively. The ideal clock period is T_{clk}. The corresponding period jitters are $J_1 = T_1 - T_{\text{clk}}$ and $J_2 = T_2 - T_{\text{clk}}$.

A. Effect of Skitter on Setup Time Slack

Assuming the data is transferred from FF_1 to FF_2 within one clock cycle, $T_{1,2}$ is the time interval that affects the highest clock frequency of the circuit. The setup time requirement needs to be satisfied for the system to work correctly [1]. The setup time slack $\text{slack}_{\text{setup}}$ is defined as

$$\text{slack}_{\text{setup}} = T_{1,2} - \max(D_{1,2}) - t_{\text{setup}} \tag{11}$$

$$T_{1,2} = (t_2 - t_1) + T_2 = S_{1,2} + J_2 + T_{\text{clk}} \tag{12}$$

where $\max(D_{1,2})$ denotes the longest data transfer time from FF_1 to FF_2. The setup time for FF_2 is t_{setup}, specified in the cell library. Consequently, the variation of $\text{slack}_{\text{setup}}$ is affected by the variation of $T_{1,2}$, called "setup skitter" $J_{1,2}$

$$J_{1,2} = S_{1,2} + J_2 = t_2' - t_1 - T_{\text{clk}}. \tag{13}$$

To avoid setup time violations, $\text{slack}_{\text{setup}} \geq 0$ is required in any operating condition.

According to (5) and (13), skitter $J_{1,2}$ is the linear combination of the delay of buffer stages

$$J_{1,2} = \sum_{k=1}^{n_3+n_4} d_{2,k}' - \sum_{k=1}^{n_1+n_2} d_{1,k} \tag{14}$$

$$\overline{J}_{1,2} = \sum_{k=1}^{n_3+n_4} \overline{d'}_{2,k} - \sum_{k=1}^{n_1+n_2} \overline{d}_{1,k} \tag{15}$$

$$\Delta J_{1,2} = \sum_{k=1}^{n_3+n_4} \Delta d_{2,k}' - \sum_{k=1}^{n_1+n_2} \Delta d_{1,k} \approx \sum_{p \in \vec{P}} \left(\left. \frac{\partial J_{1,2}}{\partial p} \right|_0 \Delta p \right) \tag{16}$$

where $d_{2,k}'$ is the delay of the kth buffer stage along the path to FF_2 for the second clock edge. The mean skitter $\overline{J}_{1,2}$ is determined by the mean delay of all the buffer stages considering the mean voltage supply noise (without process variations). Substituting (9) into (16), the partial derivatives $(\partial J_{1,2}/\partial p)|_0$ are obtained. Consequently, skitter $J_{1,2}$ is approximated by the first order Taylor expansion.

Assuming all the parameters are described by Gaussian distributions, $\Delta J_{1,2}$ can be also approximated by a Gaussian distribution

$$\Delta J_{1,2} \sim \mathcal{N}(0, \sigma_{J_{1,2}}^2) \tag{17}$$

$$\sigma_{J_{1,2}}^2 = \sum_{p \in \vec{P}} \left(\left. \frac{\partial J_{1,2}}{\partial p} \right|_0^2 \sigma_p^2 \right) \tag{18}$$

$$+2 \sum_{p,q \in \vec{P}} \left(\left. \frac{\partial J_{1,2}}{\partial p} \right|_0 \left. \frac{\partial J_{1,2}}{\partial q} \right|_0 \text{cov}(p, q) \right) \tag{19}$$

where $\text{cov}(p, q)$ denotes the covariance between two parameters. Assuming D2D variations are independent from WID variations [3], [23], $\sigma_p^2 = \sigma_{p(\text{D2D})}^2 + \sigma_{p(\text{WID})}^2$. The covariance between two parameters is determined according to the tiers to which these parameters are related and the spatial correlation between these parameters

$$\text{cov}(p, q) = \begin{cases} 0, & \text{if } p, q \text{ are of different type or} \\ & \quad \text{belong to different tiers} \\ \text{cov}(p, q)_{\text{WID}} + \sigma_{p(\text{D2D})}\sigma_{q(\text{D2D})}, & \text{otherwise} \end{cases} \tag{20}$$

where the WID covariance $\text{cov}(p, q)_{\text{WID}}$ is determined by the spatial correlation between parameters p and q within the same tier. Statistically, the devices (wires) close to each other have higher correlation than those far from each other. This spatial correlation can be obtained from fabricated wafers [34] or through a spatial correlation model [10], [23]. Due to the lack of industrial wafer data, the extracted covariance matrix from [10] is used in the simulations of this paper.

As shown in (19) and (20), the variance of setup skitter $\sigma_{J_{1,2}}^2$ highly depends on the covariance between process-induced parameters. In 2-D ICs, the change of $\text{cov}(p, q)$ is mainly determined by $\text{cov}(p, q)_{\text{WID}}$, since the parameters of the same type are affected by the same D2D variations. Therefore, the distribution of clock paths only affects $\sigma_{J_{1,2}}^2$ by changing the WID covariance. In 3-D circuits, however, D2D variations vary among tiers, and WID covariance among tiers is zero. Consequently, the distribution of clock paths will affect the skitter variation in a more complicated way, as discussed in Section V.

B. Effect of Skitter on Hold Time Slack

In addition to the setup time slack, hold time slack also significantly affects the design of ICs. The hold violation can also cause the failure of the entire system [1]. Moreover, this failure cannot be removed by lowering the clock frequency of the system. As illustrated in Fig. 6(b), the hold time slack is modeled as

$$\text{slack}_{\text{hold}} = \min(D_{1,2}) - S_{1,2} - t_{\text{hold}} \tag{21}$$

where the hold time requirement t_{hold} is also specified in the cell library. The "hold skitter" affecting $\text{slack}_{\text{hold}}$ is determined by $S_{1,2}$, which is the skew between clk_1 and clk_2. Note that $S_{1,2}$ is affected by both process variations and power supply noise.

To correctly latch the data in FF_2, $\text{slack}_{\text{hold}} \geq 0$ is required to avoid hold time violations in any operating condition.

TABLE II
VARIATIONS OF DEVICES, HORIZONTAL WIRES, AND TSVs

Parameters	Nominal	3σ (D2D)	3σ (WID)
Channel length [nm]	32	1.5	2.5
Threshold voltage [mV]	242	24.2	24.2
Wire width [nm]	225	22.5	11.3
Wire height [nm]	388	19.4	9.7
ILD thickness [nm]	252	18.9	9.5
TSV resistance [mΩ]	133	39.9	39.9
TSV capacitance [fF]	52	15.6	15.6

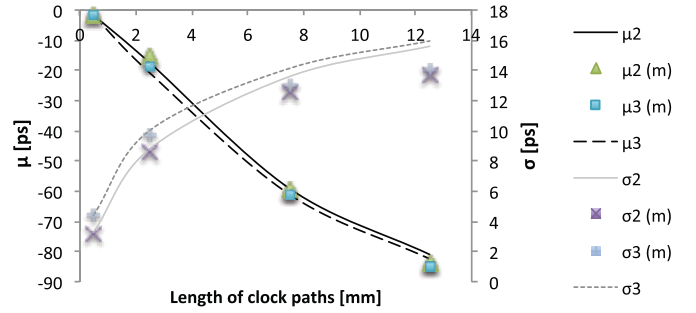


Fig. 7. Skitter versus length of 3-D clock paths.

From Fig. 6(b), $S_{1,2}$ can be determined as

$$S_{1,2} = t_2 - t_1 = \sum_{k=1}^{n_3+n_4} d_{2,k} - \sum_{k=1}^{n_1+n_2} d_{1,k}$$
$$\approx \sum_{k=1}^{n_3+n_4} \overline{d}_{2,k} - \sum_{k=1}^{n_1+n_2} \overline{d}_{1,k} + \sum_{p\in\vec{P}} \left(\frac{\partial S_{1,2}}{\partial p}\bigg|_0 \Delta p \right). \quad (22)$$

Similarly to (17) and (19), the distribution of $\Delta S_{1,2}$ can be modeled as

$$\Delta S_{1,2} \sim \mathcal{N}(0, \sigma^2_{S_{1,2}}) \quad (23)$$
$$\sigma^2_{S_{1,2}} = \sum_{p\in\vec{P}} \left(\frac{\partial S_{1,2}}{\partial p}\bigg|_0^2 \sigma_p^2 \right) + 2 \sum_{p,q\in\vec{P}} \left(\frac{\partial S_{1,2}}{\partial p}\bigg|_0 \frac{\partial S_{1,2}}{\partial q}\bigg|_0 \mathrm{cov}(p,q) \right) \quad (24)$$

where the partial derivatives are obtained similar to the coefficients in (16). As shown through (1) to (24), both the setup and hold time violations are simultaneously affected by the process variations and power supply noise. This effect and the accuracy of the proposed model are discussed in the following section.

IV. SIMULATION RESULTS

The paths of a 3-D clock tree with clock buffers inserted are simulated and discussed in this section. The electrical parameters of the transistors are based on a 32-nm PTM model [35]. The parameters of the interconnects are based on an Intel 32-nm interconnect technology [11]. The parameters of TSVs are based on data from [32]. Both the horizontal wires and TSVs are modeled by π segments in SPICE-based simulations. The proposed model is implemented in MATLAB. All the simulations are performed in a Scientific Linux server (Intel Xeon 2.67 GHz, 24 cores, 24-GB memory).

The variations considered in the simulations are listed in Table II. The D2D and WID ΔL_{b} are extracted based on ITRS data [36]. The wire variations and ΔV_{th} are based on [23]. The variations of TSVs are based on [7]. Note that other sources of variations can also be described by the proposed modeling approach. For example, the TSV stress-induced delay variation in [8] can be included. In this case, the distribution of d_{B} in (1) is adapted based on the distance between the buffer and TSVs and the given formula of stress-induced buffer delay.

In the following sections, the setup skitter $J_{1,2}$ is first compared for clock paths with different lengths. Second, based on various scenarios of the power supply noise (from Section II-B), both the setup skitter $J_{1,2}$ and hold skitter $S_{1,2}$ are compared for two different distributions of clock paths. The concurrent effect of process variations and power supply noise on the setup and hold time slacks is analyzed. The tradeoff between skitter and power is also presented.

A. Setup Skitter Versus Length of Clock Paths

The change of setup skitter with the length of clock paths is investigated in this section. In the simulations, the length ranges from 0.5 to 12.5 mm within 2- and 3-tier circuits. Buffers are inserted to produce a 10% T_{clk} input slew for the next stage. To emphasize the relation between skitter and the length of clock paths, all tiers are assumed to experience similar supply noise ($V_{\mathrm{n}} = 90$ mV, $f_{\mathrm{n}} = 400$ MHz, $\phi = 270°$ [11]). Each pair of paths is averagely distributed across different tiers, as shown in Fig. 6(a). The resulting $\mu_{J_{1,2}}$ and $\sigma_{J_{1,2}}$ are illustrated in Fig. 7, where the suffixes "2" and "3" denote the results for 2- and 3-tier circuits, respectively.

The data from SPICE-based Monte Carlo simulations and the proposed model [labeled with the (m)] are both depicted in Fig. 7. As shown in this figure, both $\mu_{J_{1,2}}$ and $\sigma_{J_{1,2}}$ deteriorate with the length of clock paths. This behavior can be described by the proposed model with a reasonable accuracy. The error of the proposed model is below 11% for $\mu_{J_{1,2}}$ and 12% for $\sigma_{J_{1,2}}$, respectively. Not surprisingly, long clock paths introduce high skitter in 3-D clock trees.

Proposition 1: Both the mean and standard deviation of setup skitter increase with the length of clock paths.

The skitter has been simulated for no TSV variation, 5% TSV variations ($\sigma/\mu = 5\%$), and 15% TSV variations, respectively. The difference in $\sigma_{J_{1,2}}$ among these three cases is around 1 ps for all the clock paths. This situation shows that TSV variations are a second-order effect, consistent with the results reported in [7].

B. Skitter Versus V_n in Different Tiers

3-D PDNs with different amplitudes of power supply noise among tiers are investigated in this section. Due to the different switching current in power supply networks and the vertical resistance of P/G TSVs among tiers, the devices in different tiers can be subjected to different ΔV_{n}, as shown

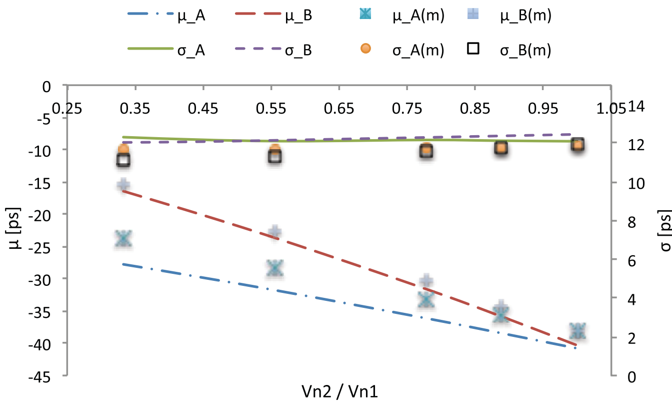


Fig. 8. Skitter for $V_{\mathrm{n1}} = 90$ mV and different V_{n2}.

in Figs. 3 and 4. The tier closer to the P/G pads experiences lower supply noise [15].

The clock paths spanning two tiers with 20 buffers [$n_1 + n_2 = n_3 + n_4 = 20$, see Fig. 6(a)] are taken as an example. The clock source is located in Tier 2. The total length of each path is 5 mm. The initial phase ϕ (270°) and frequency f_{n} (400 MHz) are assumed to be the same for both tiers. Two distributions of clock paths are discussed: (A) $n_1 = n_2 = n_3 = n_4 = 10$ and (B) $n_1 = n_3 = 15, n_2 = n_4 = 5$. Distribution (A) denotes the equally-divided 3-D clock paths. Distribution (B) represents placing the longest segment of clock paths in Tier 2. To depict the accuracy of the model, the simulation results of the setup skitter $J_{1,2}$ for $V_{\mathrm{n1}} = 90$ mV and different V_{n2} are shown in Fig. 8. As shown in this figure, $\mu_{J_{1,2}}$ changes significantly with V_{n2}, while $\sigma_{J_{1,2}}$ does not vary a lot with V_{n2}. This behavior is accurately described by the proposed model. The change of setup and hold skitter with both V_{n2} and V_{n1} is discussed in the following sections.

1) Setup Skitter $J_{1,2}$ Versus V_n: The change of $J_{1,2}$ with (V_{n2}, V_{n1}) is illustrated in Fig. 9. As shown in Fig. 9(a) and (b), for distribution (A), $\mu_{J_{1,2}}$ increases significantly with both V_{n2} and V_{n1}, since higher supply noise introduces greater period jitter. The clock paths of (A) are equally distributed among tiers. As a result, $\mu_{J_{1,2}}$ is affected by V_{n1} and V_{n2} in the same way. For distribution (B), however, the situation is different. As shown in Fig. 9(c) and (d), $\mu_{J_{1,2}}$ is mainly determined by V_{n2}, since the longest segment of the clock paths in (B) is placed in Tier 2.

Proposition 2: For unequally-distributed clock paths, the mean skitter is mainly determined by the tier where the longest part of the clock paths is placed.

As shown in Fig. 9(a) and (b), assuming $V_{\mathrm{n1}} = 0.09$ mV, distribution (A) produces higher $\mu_{J_{1,2}}$ than (B) for different V_{n2}. This difference in $\mu_{J_{1,2}}$ increases with ΔV_{n} ($\Delta V_{\mathrm{n}} = V_{\mathrm{n1}} - V_{\mathrm{n2}}$), from 1 to 42% of μ_{J_A}. The reason is that the majority of buffers in (B) is located in Tier 2, which is more susceptible to V_{n2}. More generally, given $V_{\mathrm{n1}} > V_{\mathrm{n2}}$, the mean skitter of (B) is always lower than (A).

Consequently, the distribution of clock paths in 3-D ICs significantly affects the mean skitter due to the different V_{n} among tiers. However, in 2-D circuits, this mean skitter does not vary significantly with the distribution of clock paths due to the common effect of the global resonant noise at low frequencies [28].

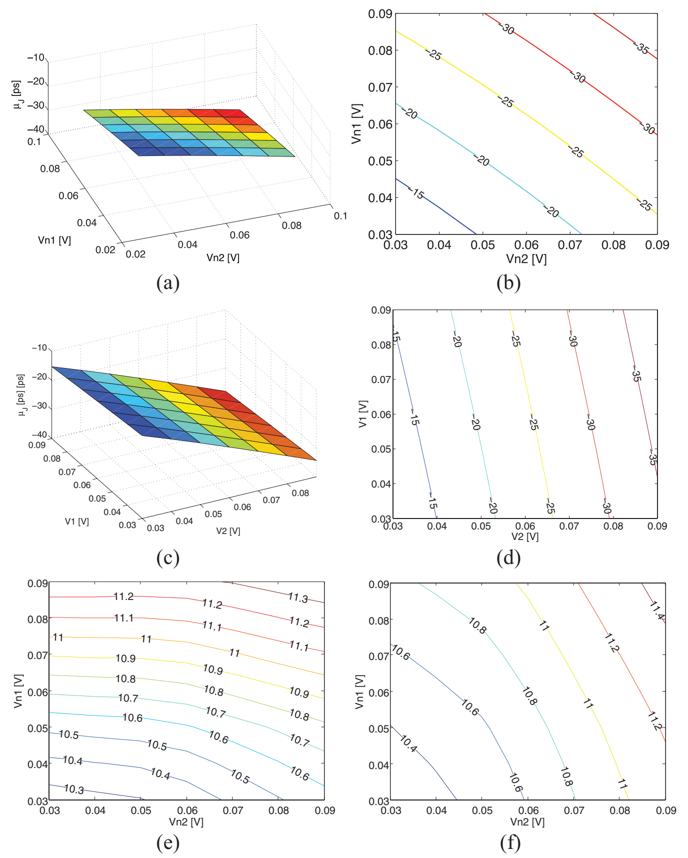


Fig. 9. Setup skitter versus (V_{n2}, V_{n1}), where (a) and (b) are the 3-D plot and contour for $\mu_{J_{\mathrm{A}}}$ in distribution (A), respectively. (c) and (d) are the 3-D plot and contour for $\mu_{J_{\mathrm{B}}}$ in distribution (B), respectively. (e) and (f) are the contours of $\sigma_{J_{\mathrm{A}}}$ and $\sigma_{J_{\mathrm{B}}}$, respectively.

The standard deviation $\sigma_{J_{1,2}}$ of (A) and (B) is illustrated in Fig. 9(e) and (f), respectively. Similar to $\mu_{J_{1,2}}$, $\sigma_{J_{1,2}}$ also increases with V_{n1} and V_{n2}. Nevertheless, $\Delta\sigma_{J_{1,2}}$ is relatively low as compared with $\Delta\mu_{J_{1,2}}$.

2) Hold Skitter $S_{1,2}$ Versus V_n: The mean value of $S_{1,2}$ is relatively low (≤ 0.5 ps), since the two clock paths have the same number, size, and distribution of buffers. Nevertheless, $\sigma_{S_{1,2}}$ is nonnegligible for both distributions (A) and (B), as illustrated in Fig. 10(a) and (b), respectively. Similar to $\sigma_{J_{1,2}}$, $\sigma_{S_{1,2}}$ increases with V_{n1} and V_{n2} but $\Delta\sigma_{S_{1,2}}$ is lower than 1.5 ps.

Proposition 3: The standard deviation of the setup and hold skitter increases with the amplitude of resonant supply noise.

C. Effect of ϕ on Skitter

The skitter under the power supply noise with different ϕ is investigated in this section. As shown in Fig. 2(a), the initial phase ϕ of the supply noise is similar among tiers ($\phi_1 = \phi_2$). The change of $J_{1,2}$ and $S_{1,2}$ with ϕ is illustrated in Fig. 11, where $V_{\mathrm{n1}} = 0.09$ V and $V_{\mathrm{n2}} = 0.07$ V.

1) Setup Skitter $J_{1,2}$ Versus ϕ: As shown in Fig. 11(a) and (b), the difference in ϕ results in significant change not only in $\mu_{J_{1,2}}$, but also in $\sigma_{J_{1,2}}$. For instance, the highest $\sigma_{J_{1,2}}$ is 41% higher than the lowest one for Distribution (A) in

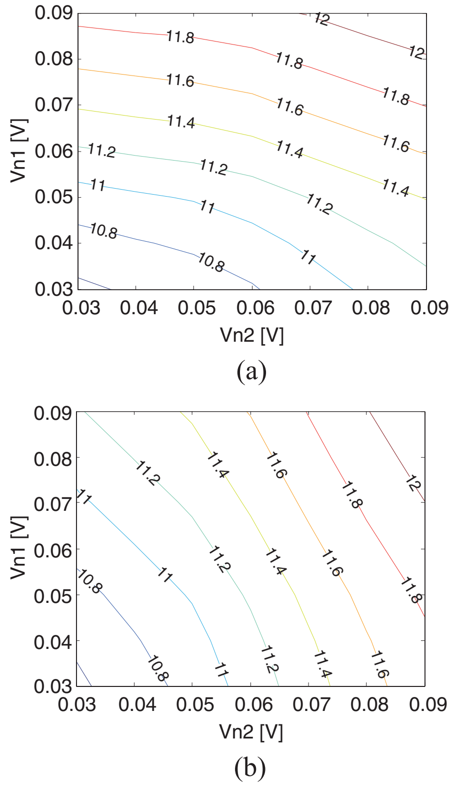


Fig. 10. Hold skitter versus (V_{n2}, V_{n1}), where (a) and (b) are the contours for σ_{S_A} and σ_{S_B}, respectively.

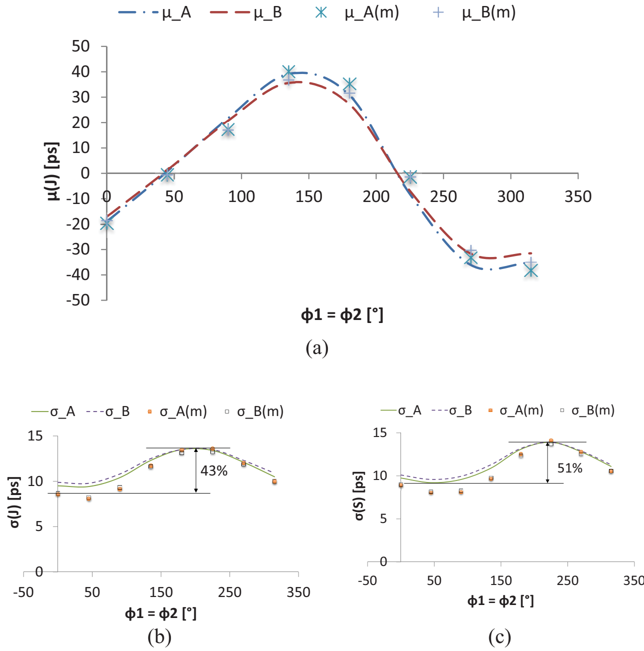


Fig. 11. Skitter versus different ϕ ($\phi_1 = \phi_2$), where (a) is the change of $\mu_{J_{1,2}}$. (b) and (c) are the change of $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$, respectively.

Fig. 11(b). The worst $\mu_{J_{1,2}}$ occurs when ϕ_1 and ϕ_2 are both around 270°, similar to the conclusion for 2-D ICs in [11]. The worst $\sigma_{J_{1,2}}$, however, occurs when $\phi \approx 225°$. Therefore, if the initial phase is not 270°, the skitter can be still high due to the high $\sigma_{J_{1,2}}$. The difference in $\sigma_{J_{1,2}}$ is low between

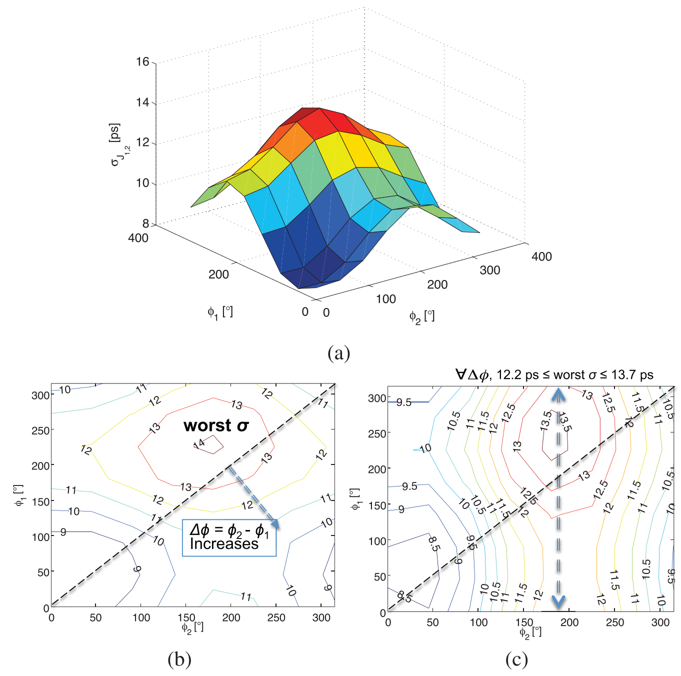


Fig. 12. Skitter $J_{1,2}$ versus shifted ϕ_1 and ϕ_2, where (a) and (b) are the 3-D plot and contour map of $\sigma_{J_{1,2}}$ versus (ϕ_2, ϕ_1) for distribution (A), respectively. (c) is the contour map of $\sigma_{J_{1,2}}$ for distribution (B).

distributions (A) and (B) since in either case, the distribution of the clock paths to FF_1 and FF_2 is identical.

2) Hold Skitter $S_{1,2}$ Versus ϕ_n: The effect of ϕ_1 and ϕ_2 on $S_{1,2}$ is shown in Fig. 11(c). Due to the similarity between the two clock paths, the resulting $\mu_{S_{1,2}}$ is relatively low. The standard deviation, however, is significantly affected by ϕ. As illustrated in Fig. 11(b) and (c), the change of $\sigma_{S_{1,2}}$ is similar to $\sigma_{J_{1,2}}$.

Proposition 4: For the setup and hold skitter, σ changes considerably with the phase of the power supply noise. The highest σ and μ of skitter do not happen at the same initial phase of the supply noise.

Considering the clock paths and waveforms shown in Fig. 6, ϕ is determined by the time when the first clock edge arrives at the input of clock paths. The worst σ can be obtained by traversing all the possible ϕ. Due to the excessive time required by Monte Carlo simulations, the proposed model is highly efficient to determine the worst skitter and the corresponding ϕ for multitier circuits, as compared with Monte Carlo simulations.

3) Effect of Phase-Shifting of the Supply Noise on Skitter: Several techniques, such as RC filtered buffers and "stacked" phase-shifted buffers [12], have been proposed to shift the ϕ seen by the clock paths. In 3-D clock distribution networks, these techniques can be applied to a part of the clock paths in a different tier to increase $\Delta\phi$ among tiers. The change of $\sigma_{J_{1,2}}$ versus the shifted (ϕ_1, ϕ_2) for distribution (A) is shown in Fig. 12(a) and (b). As shown in Fig. 12(b), the dashed line depicts the $\sigma_{J_{1,2}}$ for $\phi_1 = \phi_2$, which denotes the skitter without phase-shifting. As shown by the arrow, the highest $\sigma_{J_{1,2}}$ decreases with $\Delta\phi = \phi_2 - \phi_1$. In this case, since ϕ_2 and ϕ_1 are not simultaneously equal to 270°, the worst $\mu_{J_{1,2}}$ is also decreased.

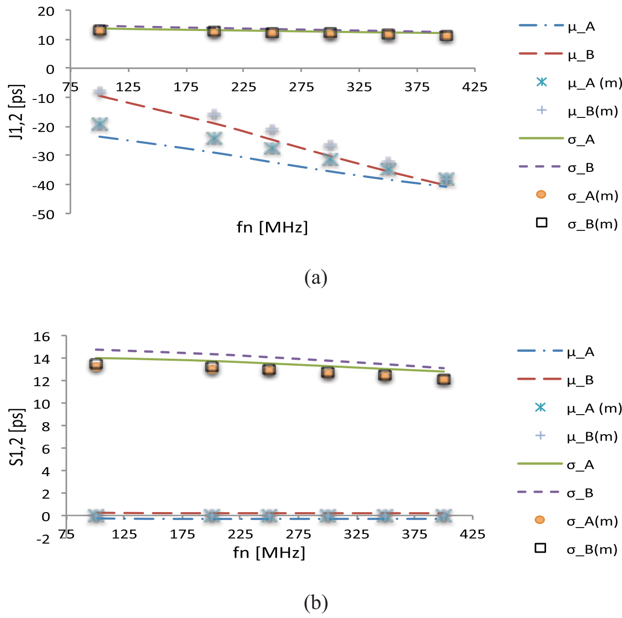
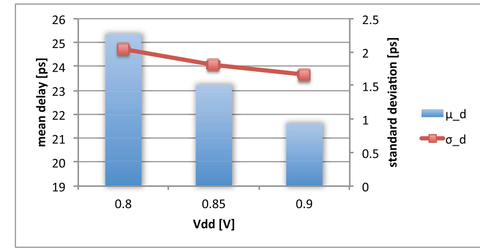
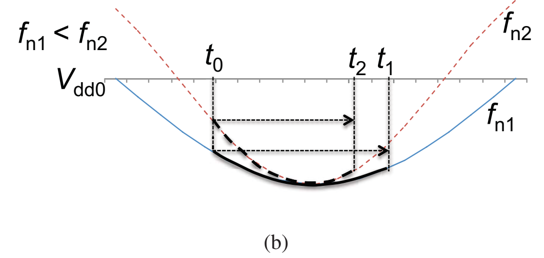


Fig. 13. Skitter versus f_n. The change of $J_{1,2}$ and $S_{1,2}$ is illustrated in (a) and (b), respectively.



(a)



(b)

Fig. 14. Effect of the change of f_n on delay variation, where (a) is the mean and standard deviation of buffer delay versus V_{dd} and (b) is the supply voltage to a clock path during the propagation of a clock edge.

In Fig. 12(c), however, $\sigma_{J_{1,2}}$ of distribution (B) highly depends on ϕ_2. This behavior is due to that $\sigma_{J_{1,2}}$ is dominated by the supply noise in the second tier. In this case, shifting ϕ among tiers provides less than 1.5 ps decrease in $\sigma_{J_{1,2}}$, as shown by the dashed line with arrows.

Proposition 5: For equally-distributed clock paths across 3-D ICs, the worst skitter can be decreased by properly shifting ϕ among tiers with phase-shifted clock distribution.

Note that the proper $\Delta\phi$ should be determined by traversing all the combinations of ϕ in different tiers. The number of combinations increases exponentially with the number of tiers, which implies a large number of simulations. Again, the proposed model provides a highly efficient way to determine a valid shift in ϕ for multitier circuits to decrease skitter.

D. Effect of f_n on Skitter

The effect of the frequency of power supply noise on skitter is investigated in this section. This frequency is usually considered similar among tiers [15], as shown in Figs. 3 and 4. Different f_n are investigated, herein, to demonstrate the change of skitter with the frequency of supply noise. The amplitude V_n and phase ϕ are assumed to be the same among tiers, where $V_{n1} = V_{n2} = 90$ mV and $\phi_1 = \phi_2 = 270°$. The simulation results are illustrated in Fig. 13.

Similar to the effect of V_n, f_n greatly affects $\mu_{J_{1,2}}$. For instance, $\mu_{J_{1,2}}$ increases with f_n up to 70% for distribution (B). The variation of skitter, however, decreases with f_n. The resulting $\Delta\sigma_{J_{1,2}}$ and $\Delta\sigma_{S_{1,2}}$ are up to 15% for both distributions (A) and (B). This behavior is due to the decreased voltage seen by the clock buffers during the clock propagation. The change of μ_d and σ_d for the delay of two inverters (a clock buffer) in series is illustrated in Fig. 14(a). Both μ_d and σ_d decrease with V_{dd}. As shown in Fig. 14(b), assume that the clock edge seeing the worst σ_J arrives at the input of the clock path at t_0. When f_n increases from f_{n1} to f_{n2}, the propagation time of this edge decreases from t_1 to t_2 and the supply voltage within this duration increases. This higher supply voltage introduces lower σ in the buffer delay, which causes lower $\sigma_{J(1,2)}$ and $\sigma_{S(1,2)}$ according to (19) and (24).

Proposition 6: The mean setup skitter increases significantly with the frequency of power supply noise, while both $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$ decrease with this frequency.

As shown in Figs. 7–13, the proposed statistical model for skitter exhibits reasonably high accuracy as compared with SPICE-based simulations. For the worst-case $\mu_{J_{1,2}}$ ($\sigma_{J_{1,2}}$) in Figs. 7–13, the error is −11% (−12%), −7% (−10%), −8% (−4%), and −10% (−9%), respectively. The behavior of skitter under different scenarios of supply noise can be correctly described by the proposed model. Since $\sigma_{J_{1,2}}$ varies with power supply noise, process variations, and power supply noise need to be simultaneously modeled to correctly describe the clock uncertainty.

The difference in mean skitter varies up to 60% due to the different V_n among planes. $\sigma_{J_{1,2}}$ can vary up to 51% due to different ϕ [see Fig. 11(b) and (c)]. Decreasing the variation as well as the mean skitter helps to improve the robustness of 3-D clock distribution networks.

E. Tradeoffs Between Skitter and Power Consumption

The power consumed by clock distribution networks constitutes a significant portion of the total power consumed by a circuit [1]. The power consumption of the clock network under different constraints on skitter is investigated in this section. A pair of clock paths with the length of 5 mm is simulated. These paths are both equally distributed across two tiers, where $V_{n1} = 0.09$ V and $V_{n2} = 0.08$ V. The skitter and power are determined by Monte Carlo simulations in this

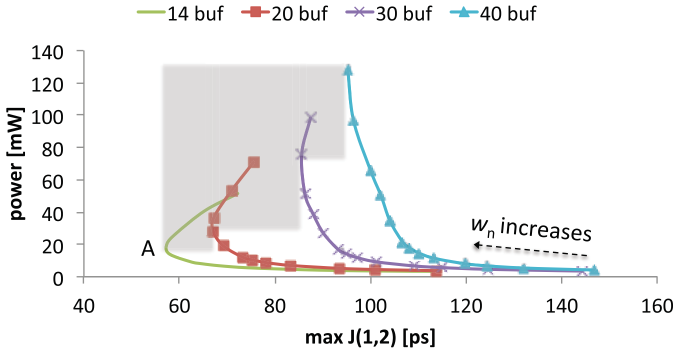


Fig. 15. Tradeoff between power and setup skitter $\max(J_{1,2})$.

TABLE III
3-D ICs BASED ON IBM CLOCK NETWORK BENCHMARKS

	No. of Sinks	No. of Buffers	Area [mm^2]	t_s [h]	t_m [s]	Speedup
r3	862	2128	9.8×9.6	1.8	45	142×
r4	1903	4695	12.7×12.7	1.9	53	129×
r5	3101	7496	14.5×14.3	2.4	56	154×

section. Different numbers (14 to 40) and sizes (W_n) of clock buffers are inserted along the clock paths.

Considering the Gaussian distribution of the setup skitter $J_{1,2}$ in (17), $J_{1,2}$ falls in the range $[\mu_{J_{1,2}} - 3\sigma_{J_{1,2}}, \mu_{J_{1,2}} + 3\sigma_{J_{1,2}}]$ with a probability of 99.7%. Within this range, $\max(J_{1,2})$ is used to indicate the worst (maximum) skitter. For improved readability, the absolute value of $\max(J_{1,2})$ is shown, where $\max(J_{1,2}) = |\mu_{J_{1,2}}| + 3\sigma_{J_{1,2}}$. The total power consumption under different constraints on $\max(J_{1,2})$ for the above clock paths is illustrated in Fig. 15. The shaded area depicts the inferior buffer solutions. Point A denotes the lowest skitter that can be obtained. In the unshaded area, skitter decreases as the buffer size and power increase. For the same constraint in skitter, the clock paths with fewer buffers are more power-efficient.

As shown within the unshaded area, the clock paths with fewer buffers produce lower skitter. For the clock paths with 14 buffers, as the constraint becomes lower than 68 ps, significant power overhead is shown. For example, to decrease the $\max(J_{1,2})$ from 68 to 58 ps (15% improvement), the buffers are sized up from 4 to 10 μm. The resulting power consumption increases from 6.9 to 14.4 mW (109% increase). In conclusion, pursuing extreme constraints on clock skitter results in high overhead in power.

Proposition 7: Skitter decreases as the number of buffers decreases. Skitter can also be decreased by sizing up buffers at the expense of power consumption.

V. Case Study for 3-D Clock Trees and Discussion

Based on Propositions 1 to 7, a set of guidelines is provided to support the design of robust 3-D clock distribution networks. The objective of these guidelines is to decrease skitter in 3-D ICs.

Guideline 1: Given the freedom to choose among tiers for the clock paths in a 3-D circuit, the mean skitter can be decreased by placing most of the clock path length in those tiers that exhibit the lowest supply noise.

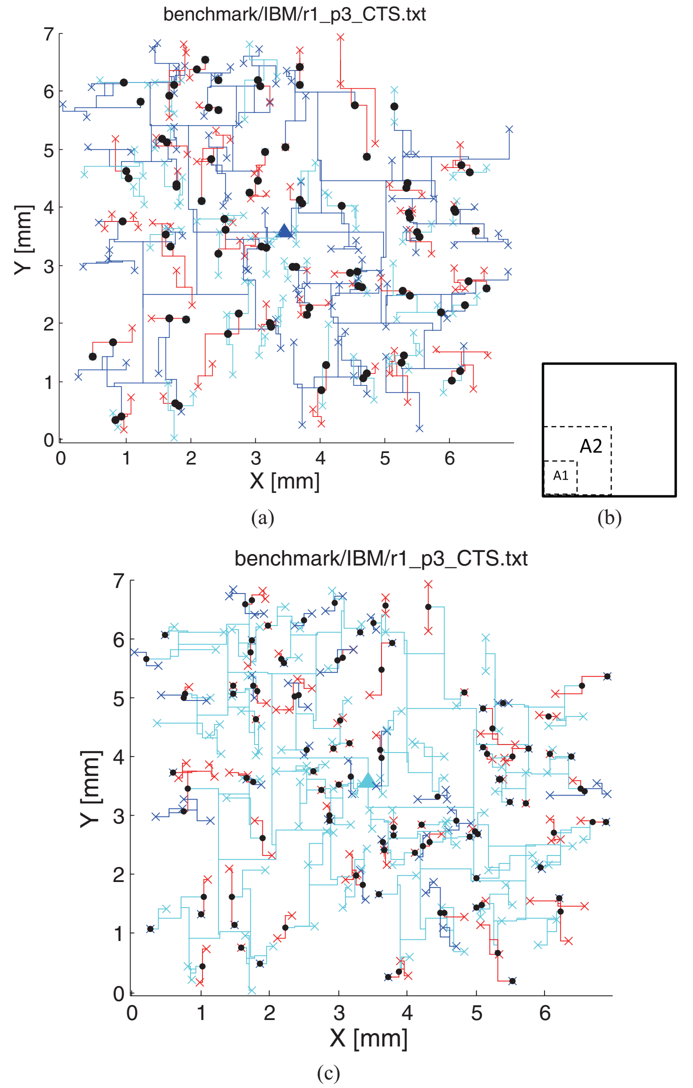


Fig. 16. Synthesized 3-D clock tree with the majority of clock buffers in the (a) first and (c) third tier. (b) Regions where the skitter is measured.

Guideline 2: For 3-D clock paths equally distributed among tiers, the worst-case $\mu_{J_{1,2}}$ and $\sigma_{J_{1,2}}$ can be decreased by shifting ϕ among different tiers.

Guideline 3: By decreasing the frequency of resonant supply noise, $\mu_{J_{1,2}}$ can be decreased by trading off $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$.

Guideline 4: By properly sizing up the clock buffers, a tradeoff between skitter and power consumption can be exploited.

To illustrate the role of these guidelines, several examples of synthesized 3-D clock trees are simulated and analyzed in this section. The 3-D circuits are generated from IBM clock benchmarks [37] by randomly distributing the clock sinks to different tiers [7]. The 3-D clock trees are synthesized with a 3-D MMM+DME algorithm based on [38]. The buffers are inserted with a constraint of 50 fF on the capacitive load. Each clock buffer is formed by an inverter ($W_\text{n} = 4.83$ μm and

TABLE IV
SKITTER IN 3-D ICs GENERATED FROM IBM CLOCK DISTRIBUTION NETWORK BENCHMARKS

Benchmark		A1 C1	A1 C2	A1 C3	A1 C4	A1 Impr1[1]	A1 Impr2	A1 Error[2]	A2 C1	A2 C2	A2 C3	A2 C4	A2 Impr1	A2 Impr2	A2 Error
Setup μ [ps]	r3	−52.6	−52.1	−44.0	−35.9	31%	18%	−5%	−53.7	−53.1	−44.8	−36.4	31%	19%	−7%
	r4	−66.3	−65.0	−58.6	−48.8	25%	17%	−3%	−69.3	−68.6	−62.1	−52.0	24%	16%	−7%
	r5	−64.8	−62.9	−56.8	−47.6	24%	16%	3%	−67.3	−66.5	−59.9	−50.2	25%	16%	−1%
Setup σ [ps]	r3	8.5	11.2	9.6	10.5	7%	−9%	−10%	11.5	15.2	13.9	13.1	14%	6%	−6%
	r4	10.7	16.6	12.0	11.3	32%	6%	−8%	10.8	15.4	16.0	15.6	−2%	2%	−7%
	r5	8.5	12.9	11.6	12.5	2%	−8%	−9%	11.8	16.0	13.9	18.5	−16%	−33%	−8%
Hold σ [ps]	r3	8.5	11.4	10.1	10.3	10%	−1%	−7%	11.5	15.6	14.4	13.2	15%	9%	−7%
	r4	10.7	14.5	13.6	11.5	21%	16%	−7%	10.8	15.1	15.6	15.6	−3%	0%	−9%
	r5	8.5	11.5	11.1	11.5	0%	−4%	−5%	11.8	15.9	15.6	17.6	−10%	−13%	−6%

[1] Impr1 and Impr2 are the improvements of C4 over C2 and C3, respectively.
[2] Error is the maximum error of the proposed model as compared with SPICE-based Monte Carlo simulations.

$W_p = 2.1W_n$). An example of the resulting 3-tier clock trees for "r1" benchmark (267 sinks) is illustrated in Fig. 16(a). The clock source, clock sinks, and TSVs are denoted by ▲, ×, and •, respectively. The clock networks in tiers 1, 2, and 3 are denoted in blue, red, and green, respectively.

The skitter is measured within two different regions, as illustrated in Fig. 16(b). For both regions A_1 and A_2, the skitter is reported between the pair of the farthest sinks. The three largest IBM benchmarks r3, r4, and r5 are simulated. SPICE simulations are performed for the paths of interest with 2000 Monte Carlo simulations. The features of these benchmarks are shown in Table III, where the CPU time is also listed. Note that the simulation time is only for the selected clock paths, not for the entire clock tree. The initial phase and the frequency of the supply noise are assumed to be the same among the three tiers ($f_{n1} = f_{n2} = f_{n3} = 400$ MHz). The amplitudes V_n are assumed to differ among tiers ($V_{n1} = 0.09$ V, $V_{n2} = 0.08$ V, $V_{n3} = 0.065$ V).

The skitter is reported in Table IV. The highest mean skitter is obtained when $\phi_1 = \phi_2 = \phi_3 = 270°$ and the highest σ is reported for $\phi_1 = \phi_2 = \phi_3 = 200°$. Four design practices are compared with each other.

1) Case 1 (C1), the majority of the clock tree is located in Tier 1. The $\mu_{J_{1,2}}$ is obtained by only considering power supply noise. $\sigma_{J_{1,2}}$ is determined by only considering process variations.
2) Case 2 (C2), the majority of the clock tree is also located in Tier 1, but the power supply noise and process variations are simultaneously modeled. The $\mu_{J_{1,2}}$ and $\sigma_{J_{1,2}}$ are determined by considering both variations.
3) Case 3 (C3), the majority of the clock tree is placed in the middle tier (Tier 2).
4) Case 4 (C4), the majority of the tree is placed in Tier 3. The modeling approach in C3 and C4 is the same as in C2.

In C1 and C2, most of the clock buffers are placed in Tier 1, which is adjacent to the heat sink, to constrain the increase in the temperature of the circuit. In C3, the majority of the clock tree is placed in the middle tier to decrease the number of TSVs and power consumption, as suggested in [38]. In C4, based on Guideline 1, the majority of the clock tree is located in Tier 3 (with the lowest V_n), as illustrated in Fig. 16(c).

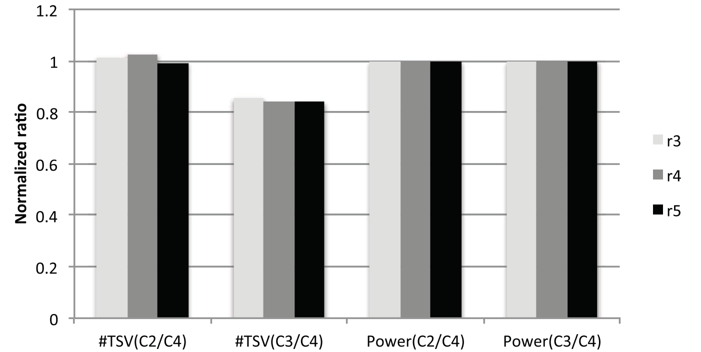


Fig. 17. Normalized number of TSVs and power for Cases 2–4.

As shown in Table IV, $\mu_{J_{1,2}}$ in Case 1 is similar to Case 2. Nevertheless, $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$ are significantly underestimated in Case 1, for both regions A_1 and A_2. As compared to Case 2, the difference in $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$ is up to 36%. This difference shows the necessity of simultaneously modeling process variations and power supply noise.

Proposition 8: Separately modeling process variations and power supply noise significantly underestimates the variation of skitter.

The difference between the proposed model and SPICE-based Monte Carlo simulations is listed in the "Error" column of Table IV. For all $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$, the error of the proposed model is below 10% as compared to Monte Carlo simulations. The error in μ is below 7% for $J_{1,2}$. Considering the greater than 129× speedup in CPU time as reported in Table III, the proposed model provides an efficient way to accurately model skitter.

In Case 2, the majority of the CDN is placed in the tier adjacent to the heat sink. In Case 3, the majority of the CDN is placed in the middle tier to reduce the number of TSVs and power consumption [38]. The number of TSVs and the power consumption of the entire tree for Cases 2 to 4 are illustrated in Fig. 17. The results are normalized over Case 4. As proposed in [38], Case 3 produces the fewest TSVs [see "#TSV(C2/C4)" and "#TSV(C3/C4)"]. The total power

is similar among the three cases due to the similar number of clock buffers, as shown by "Power (C2/C4)" and "Power (C3/C4)." The distribution of this power, however, differs due to the different distribution of buffers among tiers.

Case 4 mitigates the mean skitter trading off the number of TSVs and the power distribution. As illustrated in Figs. 3 and 4, the tier next to the package has the lowest V_{n}. Consequently, $\mu_{J_{1,2}}$ of Case 4 is significantly improved over Cases 2 and 3, as shown by the first three rows of Impr1 and Impr2, respectively. This improvement ranges from 16% up to 31%. This comparison shows the efficiency of Guideline 1 in decreasing mean skitter. For several paths, however, $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$ in Case 4 increase over Cases 2 and 3. This situation is due to the change of the topology of the clock trees. For instance, for the pair of paths in A2 and circuit r5, the number of buffers after the merging point of these paths increases as compared to Case 2. These buffers are located in different tiers. Consequently, $\sigma_{J_{1,2}}$ and $\sigma_{S_{1,2}}$ both increase.

VI. Conclusion

The combined effect of process variations and dynamic power supply noise on clock skew and jitter in 3-D clock trees was investigated for the first time. The combination of skew and jitter was described by clock skitter. A statistical model was proposed to obtain the distribution of skitter in 3-D clock trees. The skitter affecting both the setup and hold time slacks was modeled. Various sources of process variations within each tier and different resonant power supply noise among tiers can be described by this model. The proposed model was verified through SPICE-based Monte Carlo simulations. For the worst-case skitter, the error of the model is below 11% and 12% for the mean and standard deviation of skitter, respectively.

In addition to the analytic model, practical design issues were also addressed. The resonant power supply noise in 3-D ICs was shown to vary among tiers due to different electrical characteristics of PDNs. The skitter in 3-D clock trees varies significantly according to different scenarios of power supply noise and different distributions of clock paths. A set of guidelines were presented to mitigate skitter under different cases of power supply noise. A decrease in the mean skitter up to 31% was obtained in a case study by applying these guidelines.

References

[1] T. Xanthopoulos, *Clocking in Modern VLSI Systems*. New York: Springer-Verlag, 2009.

[2] V. Pavlidis and E. Friedman, *Three-Dimensional Integrated Circuit Design*. San Mateo, CA: Morgan Kaufmann, 2009.

[3] S. Garg and D. Marculescu, "3D-GCP: An analytical model for the impact of process variations on the critical path delay distribution of 3D ICs," in *Proc. Int. Symp. Qual. Electron. Design*, Mar. 2009, pp. 147–155.

[4] S. Reda, A. Si, and R. Bahar, "Reducing the leakage and timing variability of 2D ICs using 3D ICs," in *Proc. IEEE/ACM Int. Symp. Low Power Electron. Design*, Aug. 2009, pp. 283–286.

[5] H. Xu, V. Pavlidis, and G. De Micheli, "Process-induced skew variation for scaled 2-D and 3-D ICs," in *Proc. IEEE/ACM Syst. Level Interconnect Predict. Workshop*, Jul. 2010, pp. 17–24.

[6] H. Xu, V. F. Pavlidis, and G. De Micheli, "Skew variability in 3-D ICs with multiple clock domains," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2011, pp. 2221–2224.

[7] X. Zhao, S. Mukhopadhyay, and S. K. Lim, "Variation-tolerant and low-power clock network design for 3D ICs," in *Proc. Electron. Compon. Technol. Conf.*, Jun. 2011, pp. 2007–2014.

[8] J. Yang, J. Pak, X. Zhao, S. K. Lim, and D. Z. Pan, "Robust clock tree synthesis with timing yield optimization for 3D-ICs," in *Proc. Asia South Pacific Design Autom. Conf.*, Jan. 2011, pp. 621–626.

[9] K. A. Bowman, A. R. Alameldeen, S. T. Srinivasan, and C. B. Wilkerson, "Impact of die-to-die and within-die parameter variations on the clock frequency and throughput of multi-core processors," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 17, no. 12, pp. 1679–1690, Dec. 2009.

[10] A. Agarwal, V. Zolotov, and D. Blaauw, "Statistical clock skew analysis considering intradie-process variations," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 23, no. 8, pp. 1231–1242, Aug. 2004.

[11] J. Jang, O. Franza, and W. Burleson, "Compact expressions for supply noise induced period jitter of global binary clock trees," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 1, pp. 66–79, Dec. 2010.

[12] D. Jiao, J. Gu, and C. Kim, "Circuit design and modeling techniques for enhancing the clock-data compensation effect under resonant supply noise," *IEEE J. Solid-State Circuits*, vol. 45, no. 10, pp. 2130–2141, Oct. 2010.

[13] K. L. Wong, T. Rahal-Arabi, M. Ma, and G. Taylor, "Enhancing microprocessor immunity to power supply noise with clock-data compensation," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp. 749–758, Apr. 2006.

[14] G. H. Loh, Y. Xie, and B. Black, "Processor design in 3D die-stacking technologies," *IEEE Micro*, vol. 27, no. 3, pp. 31–48, May 2007.

[15] P. Jain, P. Zhou, C. H. Kim, and S. S. Sapatnekar, "Thermal and power delivery challenges in 3D ICs," in *Three Dimensional Integrated Circuit Design* (Integrated Circuits and Systems), Y. Xie, J. Cong, and S. Sapatnekar, Eds. Boston, MA: Springer-Verlag, 2010.

[16] D. Harris and S. Naffziger, "Statistical clock skew modeling with data delay variations," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 9, no. 6, pp. 888–898, Dec. 2001.

[17] B. Razavi, *Phase-Locking in High-Performance Systems: From Devices to Architectures*. New York: Wiley, 2003.

[18] M. Saint-Laurent and M. Swaminathan, "Impact of power-supply noise on timing in high-frequency microprocessors," *IEEE Trans. Adv. Packag.*, vol. 27, no. 1, pp. 135–144, Feb. 2004.

[19] T. Enami, S. Ninomiya, and M. Hashimoto, "Statistical timing analysis considering clock jitter and skew due to power supply noise and process variation," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E93-A, no. 12, pp. 2399–2408, Dec. 2010.

[20] R. Franch, "On-chip timing uncertainty measurements on IBM microprocessors," in *Proc. IEEE Int. Test Conf.*, Oct. 2007, pp. 1–7.

[21] M. S. Gupta, J. A. Rivers, P. Bose, G.-Y. Wei, and D. Brooks, "Tribeca: Design for PVT variations with local recovery and fine-grained adaptation," in *Proc. IEEE/ACM Int. Symp. Microarchit.*, Dec. 2009, pp. 435–446.

[22] H. Xu, V. F. Pavlidis, W. Burleson, and G. De Micheli, "The combined effect of process variations and power supply noise on clock skew and jitter," in *Proc. Int. Symp. Qual. Electron. Design*, Mar. 2012, pp. 320–327.

[23] H. Chang and S. Sapatnekar, "Statistical timing analysis under spatial correlations," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 9, pp. 1467–1482, Sep. 2005.

[24] K. Shinkai, M. Hashimoto, A. Kurokawa, and T. Onoye, "A gate delay model focusing on current fluctuation over wide-range of process and environmental variability," in *Proc. IEEE/ACM Int. Conf. Comput. Aided Design*, Nov. 2006, pp. 47–53.

[25] Y. Ismail, E. Friedman, and J. Neves, "Equivalent elmore delay for RLC trees," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 19, no. 1, pp. 83–97, Jan. 2000.

[26] G. Chen and E. Friedman, "Low-power repeaters driving RC and RLC interconnects with delay and bandwidth constraints," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 14, no. 2, pp. 161–172, Feb. 2006.

[27] L. Yu, W.-Y. Chang, K. Zuo, J. Wang, D. Yu, and D. Boning, "Methodology for analysis of TSV stress induced transistor variation and circuit performance," in *Proc. Int. Symp. Qual. Electron. Design*, Mar. 2012, pp. 216–222.

[28] S. Pant and E. Chiprout, "Power grid physics and implications for CAD," in *Proc. IEEE/ACM Design Autom. Conf.*, Sep. 2006, pp. 199–204.

[29] T. Enami, S. Ninomiya, and M. Hashimoto, "Statistical timing analysis considering spatially and temporally correlated dynamic power supply noise," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 28, no. 4, pp. 541–553, Apr. 2009.

[30] R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Köse, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors*, 2nd ed. New York: Springer-Verlag, 2011.

[31] P. Jain, T.-H. Kim, J. Keane, and C. H. Kim, "A multi-story power delivery technique for 3D integrated circuits," in *Proc. ACM Int. Symp. Low Power Electron. Design*, Jul. 2008, pp. 57–62.

[32] G. Katti, M. Stucchi, K. De Meyer, and W. Dehaene, "Electrical modeling and characterization of through silicon via for three-dimensional ICs," *IEEE Trans. Electron Devices*, vol. 57, no. 1, pp. 256–262, Jan. 2010.

[33] I. Savidis and E. G. Friedman, "Closed-form expressions of 3-D via resistance, inductance, and capacitance," *IEEE Trans. Electron Devices*, vol. 56, no. 9, pp. 1873–1881, Sep. 2009.

[34] P. Friedberg, J. Cain, and C. Spanos, "Modeling within-die spatial correlation effects for process-design co-optimization," in *Proc. Int. Symp. Qual. Electron. Design*, 2005, pp. 516–521.

[35] *ASU Predictive Technology Model*. (2008) [Online]. Available: http://www.eas.asu.edu/~ptm/

[36] *International Technology Roadmap for Semiconductors*, (2010). [Online]. Available: http://www.itrs.net

[37] R. S. Tsay. (2000, May). *IBM Clock Benchmarks* [Online]. Available: http://vlsicad.ucsd.edu/GSRC/bookshelf/Slots/BST/#III

[38] X. Zhao, J. Minz, and S. K. Lim, "Low-power and reliable clock network design for through-silicon via (TSV) based 3D ICs," *IEEE Trans. Compon., Packag. Manuf. Technol.*, vol. 1, no. 2, pp. 247–259, Feb. 2011.

Hu Xu received the B.E. degree in automation from Tsinghua University, Beijing, China, the M.Sc. degree in computer architecture from Peking University, Beijing, and the Ph.D. degree in information and computer sciences from the Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, in 2005, 2008, and 2012, respectively.

He is currently a Research Assistant with the Integrated Systems Laboratory, EPFL. His current research interests include modeling and design techniques of 3-D integration under various sources of variations, timing analysis, and other physical design issues for very large scale integration.

Vasilis Pavlidis received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Rochester, Rochester, NY, in 2003 and 2008, respectively.

He was with INTRACOM S.A., Athens, Greece, from 2000 to 2002. He was with Synopsys Inc., Mountain View, CA, in 2007. He was a Post-Doctoral Fellow with the Integrated Systems Laboratory, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, from 2008 to 2012. He is currently an Assistant Professor with the Computer Science Department, University of Manchester, Manchester, U.K., within the Advanced Processor Technologies Group. He has authored the book entitled *Three-Dimensional Integrated Circuit Design*. His current research interests include interconnect modeling and analysis, 3-D integration, and other issues related to VLSI design.

Xifan Tang was born in Shanghai, China, in 1989. He received the B.Sc. degree in microelectronics from Fudan University, Shanghai, China, in 2011. He is currently pursuing the M.Sc. degree in electrical engineering with the Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

His current research interests include concern computer-aided design for very large scale integration.

Wayne Burleson received the B.S.E.E. and M.S.E.E. degrees from the Massachusetts Institute of Technology, Cambridge, and the Ph.D. degree in ECE from the University of Colorado, Boulder.

He is currently a Professor of electrical and computer engineering with the University of Massachusetts Amherst, Amherst, where he has been since 1990. He was a Custom Chip Designer and a Consultant with the semiconductor industries. He was a Visiting Professor with ENST, Paris, France, from 1996 to 1997, LIRM, Montpellier, France, in 2003, and EPFL, Lausanne, Switzerland, from 2010 to 2011. He is involved in research on hardware security, reconfigurable computing, content-adaptive signal processing, RFID, and multimedia instructional technologies. His current research interests include very large scale integration, including circuits and CAD for low-power, long interconnects, clocking, reliability, thermal effects, process variation, and noise mitigation. He has authored or co-authored over 180 refereed papers in journals and conferences.

Giovanni De Micheli received the Nuclear Engineer degree from the Politecnico di Milano, Milan, Italy, in 1979, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of California at Berkeley, Berkeley, in 1980 and 1983, respectively.

He is currently a Professor and the Director of the Institute of Electrical Engineering and of the Integrated Systems Centre, the Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, where he is the Program Leader of the Nano-Tera.ch program. He was Professor of electrical engineering with Stanford University, Stanford, CA. He has authored or co-authored over 500 papers in journals and conferences. He has authored the book entitled *Synthesis and Optimization of Digital Circuits* (McGraw-Hill, 1994), co-authored and/or co-edited eight other books. His citation H-index is 76 according to Google Scholar. His current research interests include several aspects of design technologies for integrated circuits and systems, such as synthesis for emerging technologies, networks on chips, 3-D integration, heterogeneous platform design including electrical components and biosensors, and data processing of biomedical information.

Prof. De Micheli was a recipient of the 2012 IEEE/CAS Mac Van Valkenburg Award for contributions to theory, practice, and experimentation in design methods and tools, the 2003 IEEE Emanuel Piore Award for contributions to computer-aided synthesis of digital systems, the Golden Jubilee Medal for outstanding contributions to the IEEE CAS Society in 2000, the D. Pederson Award for the best paper at the IEEE Transactions on CAD/ICAS in 1987, and several Best Paper Awards, including DAC in 1983 and 1993, DATE in 2005, and Nanoarch in 2010 and 2012. He is a Fellow of the ACM and a member of the Academia Europaea. He is member of the Scientific Advisory Board of IMEC and STMicroelectronics. He was with the IEEE at several capacities, namely the Division 1 Director from 2008 to 2009, a Co-Founder and the President Elect of the IEEE Council on EDA from 2005 to 2007, the President of the IEEE CAS Society in 2003, and the Editor-in-Chief of the IEEE TRANSACTIONS ON CAD/ICAS from 1987 to 2001. He has been the Chair of several conferences, including DATE since 2010, pHealth since 2006, VLSI SOC since 2006, DAC since 2000, and ICCD since 1989.