Critical Length Estimation for TSV-Based 3D Sub/Near-Threshold Circuits

A Thesis presented by

Ajay Chandrasekhar

to

The Graduate School

in Partial Fulfilment of the

Requirements

for the Degree of

Master of Science

in

Electrical and Computer Engineering

Stony Brook University

May 2013

Stony Brook University

The Graduate School

Ajay Chandrasekhar

We, the thesis committee for the above candidate for the

Master of Science degree, hereby recommend

acceptance of this thesis.

Dr. Emre Salman - Thesis Advisor

Assistant Professor, Electrical and Computer Engineering Department

Dr. Sangjin Hong - Second Reader

Professor, Electrical and Computer Engineering Department

This thesis is accepted by the Graduate School

Charles Taber

Interim Dean of the Graduate School

Abstract of the thesis

Critical Length Estimation for TSV-Based 3D Sub/Near-Threshold Circuits

by

Ajay Chandrasekhar

Master of Science

in

Electrical and Computer Engineering

Stony Brook University

2013

Comparisons are performed between *RLC* interconnects and *RLC* models of via-first and via-last TSVs in super-threshold, near-threshold, and sub-threshold regions in order to quantify the benefits of TSV-based 3D integration in near/sub-threshold regions. It is observed that the propagation delay of interconnects is highly sensitive to variations in capacitance in near/sub-threshold regions as compared to super-threshold regions. Critical lengths for *RLC* interconnects (optimized for delay) are defined as the lengths at which the interconnects have propagation delays equal to that of the TSVs. Similarly, critical lengths are also defined for root mean square (RMS) power consumption. Critical length for via-first TSVs is determined to be $130\ \mu m$ and that for via-last TSVs is determined to be $1.6\ mm$. The critical lengths for delay and power are found to be approximately the same. In addition, effect of voltage-dependent TSV MOS capacitance on calculation of delay and power is observed. Since the gap between voltage-dependent TSV MOS capacitance and the liner capacitance reduces with decrease in supply voltage, the percentage error in delay and power are lower at sub-threshold voltage than at near-threshold and super-threshold voltages. For via-last TSVs, the error in delay reduces from 57.24 % in super-threshold region to 42.59 % in sub-threshold region. The error in power consumption of via-last TSVs, decreases from 34.47 % in super-threshold region to 0.58 %

in sub-threshold region. For via-first TSVs, error in delay decreases from 41.23 % in super-threshold region to 26.21 % in sub-threshold region, whereas the error in power computation decrease from 21.8 % in super-threshold region to 0.04 % in sub-threshold region.

Contents

1 Introduction **1**

1.1 Motivation for 3D integration . . . 2

1.2 Through silicon via . . . 2

1.3 Challenges involved . . . 4

1.4 Objective . . . 4

1.5 Outline of the thesis . . . 4

2 TSV-Based 3D integration **5**

2.1 Via-first TSV . . . 5

2.2 Via-middle TSV . . . 6

2.3 Via-last TSV . . . 6

3 Determination of critical length **7**

3.1 Sensitivity of transistor delay to capacitance . . . 7

3.2 Sensitivity of lumped *RC* interconnect delay . . . 10

3.2.1 Sensitivity to interconnect capacitance . . . 11

3.2.2 Sensitivity to interconnect resistance . . . 11

3.3 Interconnect model . . . 12

3.4 Repeater insertion . . . 13

3.4.1 *RLC* repeater insertion . . . 14

3.5 TSV model . . . 14

3.5.1 Calculation of TSV parameters . . . 15

3.6 Analysis setup . . . 16

3.7 Estimation of critical length 17
3.7.1 Via-first TSV 17
3.7.2 Via-last TSV 20
3.8 Effect of TSV MOS capacitance 22
3.8.1 Variation of MOS capacitance with bias voltage 22
3.8.2 TSV *RLC* Model 24

4 Conclusion **26**

Bibliography **28**

List of Figures

1.1 Heterogeneous integration using TSVs . . . 3

2.1 TSV Structure . . . 5

3.1 Simulation results of t-V_{dd}-C relationship . . . 9

3.2 Analytical graph showing sensitivity of delay to capacitance for different V_{dd} values . . . 10

3.3 Delay-V_{dd} curve for various interconnect capacitances . . . 11

3.4 Delay-V_{dd} curve for various interconnect resistances . . . 12

3.5 Interconnect model . . . 13

3.6 *RLC* model of a TSV . . . 15

3.7 *RLC* interconnect analysis setup . . . 17

3.8 TSV analysis setup . . . 17

3.9 Ratio of *RLC* interconnect delay to via-first TSV delay . . . 18

3.10 Ratio of *RLC* interconnect power to via-first TSV power . . . 19

3.11 Contour curves at various delay ratios for via-first TSVs . . . 19

3.12 Ratio of *RLC* interconnect delay to via-last TSV delay . . . 20

3.13 Ratio of *RLC* interconnect power to via-last power . . . 21

3.14 Contour curves at various delay ratios for via-last TSVs . . . 21

3.15 Variation of MOS capacitance with bias voltage . . . 22

3.16 Capacitance per unit length vs bias voltage . . . 23

List of Tables

3.1 Via-first TSV dimensions and parameters 16
3.2 Via-last TSV dimensions and parameters 16
3.3 Capacitance values for via-first TSV 23
3.4 Capacitance values for via-last TSV 23
3.5 Delay values for via-first TSV 24
3.6 Delay values for via-last TSV 24
3.7 Power values for via-first TSV 25
3.8 Power values for via-last TSV 25

Chapter 1

Introduction

The semiconductor industry is driven by the requirement of systems that deliver high performance and consume less power. This power-performance challenge is unlikely to disappear anytime in the future. Adding to the list of challenges is the inter-chip communication bottleneck [1]. Since off-chip interconnects are longer than on-chip interconnects, this presents a major performance bottleneck to modern systems. 3D integration is a step towards alleviating this bottleneck, since stacking multiple chips one on top of another increases the number of neighbours that a component has access to, and can be used to minimize critical path delays [2]. Therefore, 3D integration can be viewed as a high-performance optimization of chips.

However, there is still the issue of power consumption. In the past, reduction of supply voltage has resulted in substantial decrease in power consumed by devices (due to quadratic dependence of dynamic power on voltage). However, it is difficult to reduce voltages due to increasing leakage current. Operation of transistors in near-threshold and sub-threshold regions is an attractive method towards ultra-low power operation, but these steps prioritize power consumption over performance [3]. This is because at deep sub-threshold voltages, the increase in delay supersedes the decrease in power [4]. Thus, the potential applications for these modes of operation are confined to sensors and at the most, to personal mobile devices (PMDs). Near-threshold operation is an attempt to widen the application scope of sub-threshold circuits [4]. Operation of conventional interconnects in these regions of operation results in significant increase in delay [1], which has been found to be highly sensitive to the capacitance of the

interconnect. This implies that the longer the interconnect, the more significant the delay. Thus, global interconnects are subject to the largest increase in propagation delay, if they operate in near/sub-threshold regions. Insertion of TSVs that operates in near/sub-threshold region in order to replace global interconnects can help reduce the critical path delay in systems, thereby improving performance as well as achieving low-power operation.

1.1 Motivation for 3D integration

A system generally has multiple components such as processor, memory, I/O and other peripherals. Each of these components are sometimes fabricated on different chips, owing to differences in process flows involved in optimal design of each of these components. These chips are then connected via interconnections. The placement of these functional blocks and their optimal routing are separate design stages. However, even for an optimal design, the global interconnects, which are connections among functional blocks, have significantly higher delay compared to connections within each of the chips. The connections within the chips are called on-chip interconnects and the connections among functional blocks are called off-chip interconnects.

Sub/near-threshold operation has significant potential in low-power applications, but have a major limitation in performance [4]. The main reason for this is the significantly lower currents that slow down the charge/discharge process of capacitors, further raising propagation delay. The sensitivity of this delay to capacitance increases as we proceed to sub-threshold voltages. This worsens the situation for global interconnects in sub-threshold, which results in significant increase in delay. Therefore, 3D integration in sub-threshold can help alleviate this performance bottleneck, while at the same time, further reduces power consumption due to the sub-threshold operation.

1.2 Through silicon via

Through silicon via (TSV) is a method of 3D integration whereby a connection is made between two stacked silicon substrate planes [5]. This differs from a conventional via in the sense

that it passes through silicon substrate to provide connection across multiple planes. This provides an additional dimension for packing multiple functional units within a limited amount of silicon area. However, the disadvantage of TSVs lie in their large diameters [6]. Once a TSV is etched between two planes, the area on the substrates corresponding to the TSVs cannot be used for any other purpose. In other words, TSVs occupy non-trivial portions of silicon area, which could have otherwise been used for additional functional blocks or horizontal interconnections.

It is preferable to have the critical path in any system on the on-chip interconnects rather than off-chip interconnects. Communication among functional blocks is a major performance bottleneck. One solution to this issue is a system-on-chip (SoC), in which all of these components are fabricated on the same die. However, for SoC-based systems, it is difficult to optimize each functional block since all of these blocks are fabricated on the same substrate. Alternatively, we can design each of these functional blocks on separate dies and stack them on top of the other. In this way, we can optimize the process flow of each die depending on the functional block that is implemented on that die. These dies are stacked together and connected using TSVs. Unlike conventional off-chip interconnects, the TSVs are faster, eliminating much of the performance bottleneck of the system. Figure 1.1 shows the idea of heterogeneous integration using TSVs. In short, TSVs can be used to connect multiple dies of different types on a single chip [7].

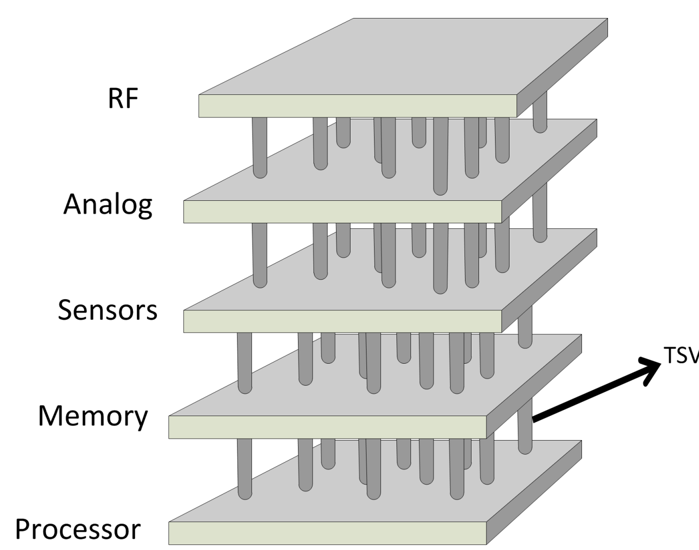

Figure 1.1: Heterogeneous integration using TSVs

1.3 Challenges involved

Despite the benefits, TSV-based 3D integration has drawbacks. One of the limitations is the relatively large size of TSVs [6]. This means that TSVs occupy large portions of silicon area. This translates to lower productivity of chips. Therefore, TSV placement is essentially a trade-off between performance and area. This is because placing a TSV costs silicon area, but there is no guarantee that the desired performance benefits be achieved. One example is if this TSV does not lie in the critical path of the system. This results in unnecessary use of silicon area. Another challenge is the lack of CAD tool support for 3D integration [8].

1.4 Objective

The objective of this thesis is to combine TSV-based 3D integration [5] and near/sub-threshold operation [4] in order to achieve low power without incurring the heavy performance penalties associated with near/sub-threshold circuits. In this thesis, critical length of an interconnect, optimized for delay, is defined as the length of the interconnect that is equivalent to a TSV in terms of propagation delay. Critical length for RMS power consumption is similarly defined. We analyze the variation of the critical lengths for delay and power as the supply voltage changes. We also explore the impact of TSV MOS capacitance [12] on propagation delay and power consumption in near/sub-threshold regions for both via-first and via-last TSVs.

1.5 Outline of the thesis

Chapter 2 provides an insight into TSV-based 3D integration, along with types of through silicon via (TSV) and their characteristics. Chapter 3 aims to quantify the feasibility of TSVs in near/sub-threshold regions of operation, and their comparison to conventional interconnects optimized for delay. Chapter 3 also elucidates the need for considering TSV MOS capacitance when modelling TSVs for near/sub-threshold operations. Chapter 4 concludes the thesis and summarizes the results.

Chapter 2

TSV-Based 3D integration

TSVs are broadly classified based on the stage at which they are fabricated, as via-first, via-middle, and via-last [6]. Figure 2.1 shows the basic structure of a cylindrical TSV, which will be used in all the analyses outlined in this work. It can be seen that the TSV comprises the filling material, which forms the electrical connection and an oxide liner to isolate the filling material from the surrounding substrate. The radius of the TSV is the radius of the cylindrical shape filled using the filling material. The TSV is characterised by the resistance and inductance of the TSV filling material, and the capacitance due to the MOS structure of the TSV and the oxide.

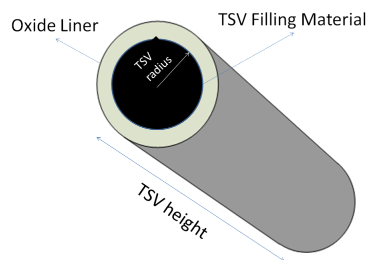

Figure 2.1: TSV Structure

2.1 Via-first TSV

Via-first TSVs are fabricated before the front-end-of-line (FEOL) process [6]. This means that the TSVs are first fabricated on the silicon substrate before any of the steps involved in

transistor fabrication. The FEOL process generally comprises high temperature processes. Therefore, it is necessary to use a material that can withstand high temperature. Thus, copper cannot be used as filling material for the TSV, even though it has a high conductivity. Polysilicon is used in via-first TSVs since it can withstand high temperatures, even though it has low conductivity relative to copper [6].

2.2 Via-middle TSV

Via-middle TSVs are fabricated after the high-temperature FEOL processes but before the back-end-of-line (BEOL) processes which involve metallization and formation of interconnects [6]. After fabrication of the TSVs, the wafer is thinned down to the appropriate depth. Since the TSVs are formed after all the high-temperature processes, there is no longer the requirement for the filling material to be able to withstand such high temperatures. Therefore, tungsten or copper can be used as filling material for via-middle TSVs [6].

2.3 Via-last TSV

Via-last TSVs are fabricated after the BEOL processes, and also after the wafer-thinning process. This has to deal with the challenge of thinning the wafer to the correct depth, since the TSVs have not yet been fabricated at this stage [6]. Copper is mostly used as TSV filling material for this type of TSVs since this does not go through any high-temperature processes, and the fabrication of TSVs occurs after the BEOL process [6].

Chapter 3

Determination of critical length

This chapter outlines the analyses performed and the results obtained as part of the thesis. Section 3.1 quantifies the sensitivity of propagation delay along an interconnect to the capacitance of the interconnect. Section 3.2 shows analysis results that demonstrate the sensitivity of delay to capacitance and resistance as the supply voltage decreases. Section 3.3 explains the interconnect model used, and Section 3.4 uses repeater insertion in order to optimize the interconnect for comparison with TSVs. Section 3.5 explains the TSV model used. Section 3.7 outlines the analyses performed and the critical lengths of interconnects obtained. Section 3.8 discusses the effect of TSV MOS capacitance on delay and power computations.

3.1 Sensitivity of transistor delay to capacitance

Analysis of sensitivity of propagation delay to capacitance is useful for the comparison of interconnects and TSVs since capacitance affects the delay more than other parameters, especially in near/sub-threshold voltages. The following analysis elaborates on the sensitivity of interconnects to the interconnect capacitance. In the sub-threshold region, drain current is given by the equation [3]

$$I_d = I_{d0}\frac{W}{L}e^{\left(\frac{V_{gs}-V_{th}}{nv_T}\right)}\left(1 - e^{\frac{-V_{ds}}{v_T}}\right). \tag{3.1}$$

I_{d0} depends on process parameters, W is channel width, L is channel length, V_{ds} is drain-source voltage, V_{gs} is gate-source voltage, V_{th} is threshold voltage, v_T is thermal voltage (25 mV) and n is sub-threshold factor, which is calculated as $1 + \frac{C_{dep}}{C_{ox}}$, where C_{dep} is depletion capacitance and C_{ox} is oxide capacitance. Drain induced barrier lowering (DIBL) is assumed to be negligible, so threshold voltage is approximately constant with respect to V_{ds}. Taking partial derivative of I_d with respect to V_{dd} gives

$$\frac{\partial I_d}{\partial V_{dd}} \approx \left[I_{d0} \frac{W}{L} e^{\left(\frac{-n.V_{ds} - V_{th}}{nv_T}\right)} \left(\frac{1}{n.v_T}\right) \right] e^{\left(\frac{V_{dd}}{n.v_T}\right)}. \tag{3.2}$$

The term within the square brackets is approximately constant with V_{dd} variation (V_{ds} changes only slightly with V_{dd} since the device is ON). Thus, we have an equation of the form

$$\frac{\partial I_d}{\partial V_{ds}} \approx k.e^{bV_{dd}} \approx \frac{1}{R_{on}}, \tag{3.3}$$

where

$$k = \left[I_{d0} \frac{W}{L} e^{\left(\frac{-n.V_{ds} - V_{th}}{nv_T}\right)} \left(\frac{1}{n.v_T}\right) \right],$$

$$b = \frac{1}{n.v_T}.$$

R_{on} is the on-resistance of the transistor. The *RC* delay for an inverter is determined by using the equation that describes the charging process of a capacitor C through a resistance R

$$V_{out} = V_{dd} \times \left(1 - e^{-\frac{t}{RC}}\right), \tag{3.4}$$

where V_{out} is the output voltage, which is the voltage across the capacitance C, V_{dd} is the supply voltage which constitutes the steady-state voltage of the fully-charged capacitor, t is the time. Re-arranging this equation gives

$$t = RC \ln\left(\frac{V_{dd}}{V_{dd} - V_{out}}\right). \tag{3.5}$$

In analyzing transistor delay, $V_{dd} - V_{out} = V_{ds}$ and $R = R_{on}$. This gives

$$t = R_{on} C \ln \left(\frac{V_{dd}}{V_{ds}} \right). \tag{3.6}$$

Thus,

$$t = \frac{C}{k} e^{-\frac{V_{dd}}{n.v_T}} \ln \frac{V_{dd}}{V_{ds}}. \tag{3.7}$$

Taking partial derivative of delay with respect to capacitance gives

$$\frac{\partial t}{\partial C} = \frac{1}{k} \times e^{(-b.V_{dd})} \ln \frac{V_{dd}}{V_{ds}}.$$

Here, $\frac{\partial t}{\partial C}$ quantifies the sensitivity of delay to capacitance. It is determined by the exponential function and the logarithmic function (k is constant). Thus, as V_{dd} decreases, the exponential term causes the delay to rise, beyond a particular point where the logarithmic function supersedes the exponential term. After this point, the fall in delay due to the logarithmic function dominates. Figure 3.1 shows the shows variation of delay with V_{dd} for various capacitances. Figure 3.2 shows the sensitivity of delay to capacitance for different values of V_{dd} .

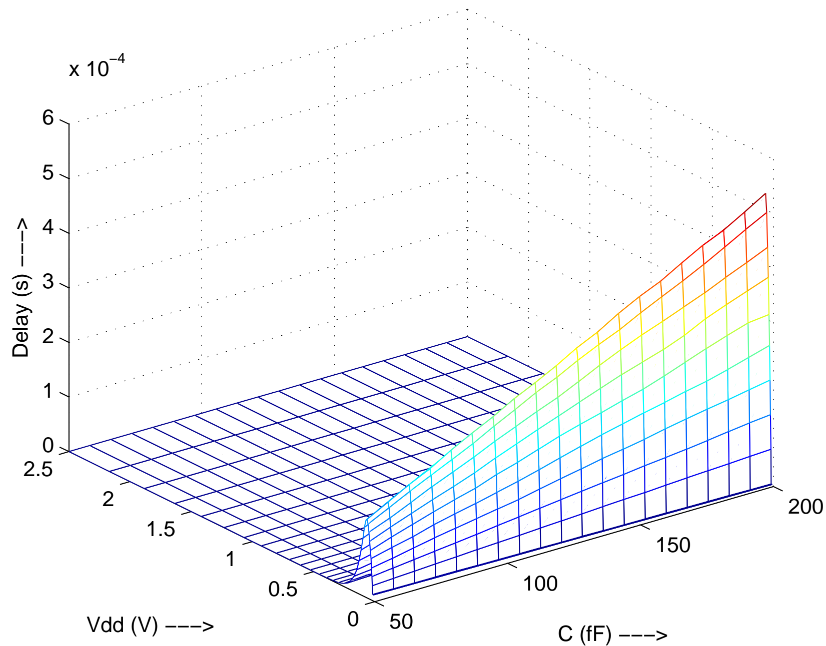

Figure 3.1: Simulation results of t-V_{dd}-C relationship

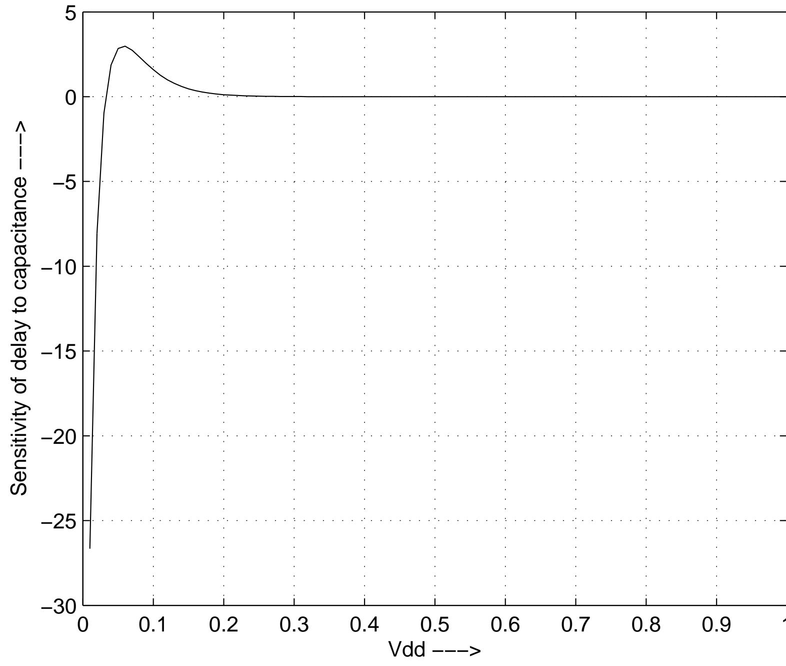

Figure 3.2: Analytical graph showing sensitivity of delay to capacitance for different V_{dd} values

In Figure 3.2, sensitivity of delay to capacitance is given by

$$\frac{\partial t}{\partial C} = e^{-\frac{V_{dd}}{0.0325}} \times \frac{\ln \frac{V_{dd}}{0.0325}}{0.0325}.$$

Figure 3.2 is a quantitative estimate in order to show the nature of the function that defines the variation of delay with supply voltage. It is observed from Figures 3.1 and 3.2 that sensitivity of transistor delay to capacitance increases as supply voltage decreases.

3.2 Sensitivity of lumped *RC* interconnect delay

In this sub-section, the effect of *RC* interconnect parameters on delay is analysed. The behaviour of *RC* interconnect depends on both the resistance and capacitance. The aim is to explain the behaviour of *RC* interconnect in near-threshold and sub-threshold regions. Subsection 3.2.1 deals with the sensitivity of delay of an *RC* interconnect to capacitance. Subsection 3.2.2 deals with sensitivity of delay of the *RC* interconnect to resistance.

3.2.1 Sensitivity to interconnect capacitance

Delay vs V_{dd} curves are plotted for various values of capacitance at constant resistance, as shown in Figure 3.3. Sensitivity to capacitance is higher in sub-threshold region. This can be explained by observing that the driver resistance dominates in sub-threshold region since the driver resistance increases exponentially with decreasing V_{dd}.

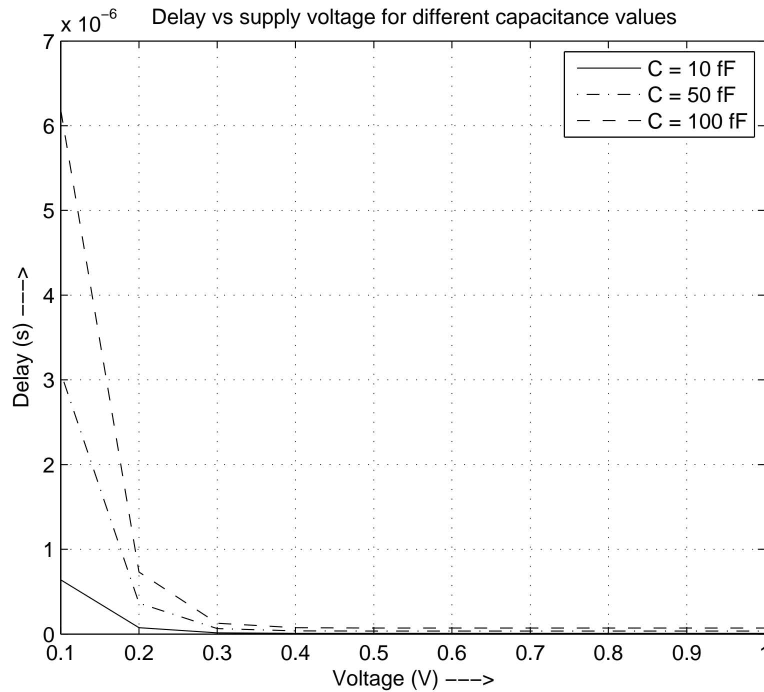

Figure 3.3: Delay-V_{dd} curve for various interconnect capacitances

3.2.2 Sensitivity to interconnect resistance

Delay vs V_{dd} curves are plotted for various values of resistance at constant capacitance in Figure 3.4. Sensitivity to interconnect resistance is low in sub-threshold region. The interconnect resistance remains constant with supply voltage while driver resistance increases exponentially. Thus, the interconnect resistance becomes less significant compared to driver resistance as supply voltage decreases. Therefore, sensitivity to resistance decreases at sub-threshold voltage levels.

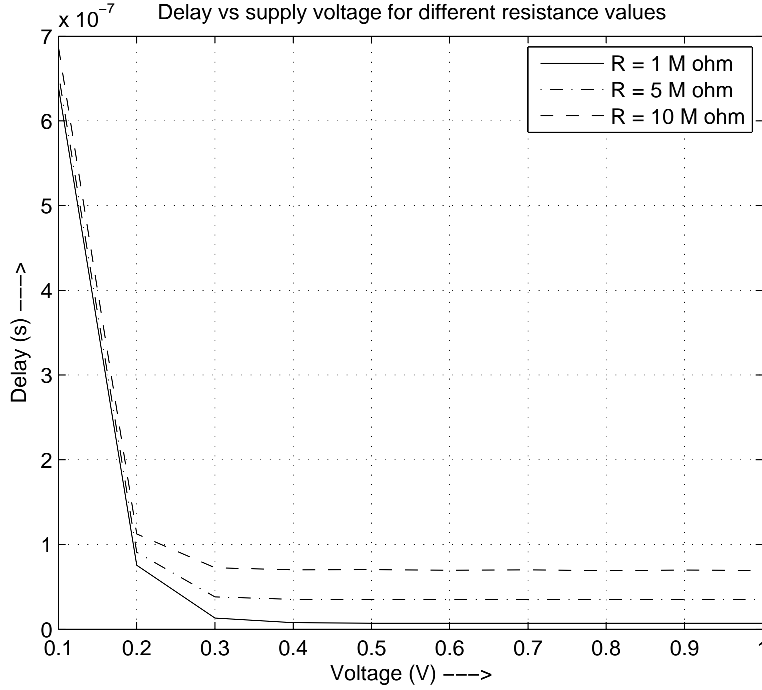

Figure 3.4: Delay-V_{dd} curve for various interconnect resistances

3.3 Interconnect model

The interconnect model is a lumped representation of distributed interconnects. The interconnect is broken down into multiple lumped sections of *R*, *L*, *C* (for *RLC* interconnects). The number of lumped sections n is chosen so that $[16]$

$$n \geq \frac{5l\sqrt{LC}}{t_r}, \tag{3.8}$$

where l is length of the interconnect, L and C are inductance and capacitance per unit length respectively, and t_r is the shortest rise time of the input signal. The maximum frequency of interest is given by $[16]$

$$f_{max} \approx \frac{0.35}{t_r}. \tag{3.9}$$

Equation (3.9) is used to determine the TSV AC resistance which is described in Subsection 3.5.1.

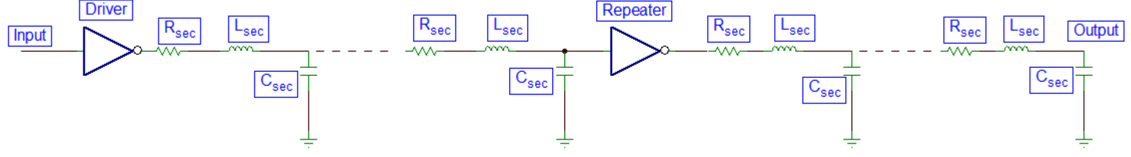

Figure 3.5: Interconnect model

Figure 3.5 shows the interconnect model used for comparison. R_{sec}, C_{sec} and L_{sec} are the resistance, inductance, and capacitance respectively, of each lumped section. Let l be the total length of the interconnect, and suppose n is the number of lumped sections used to model this interconnect. Also, let R_{pl}, L_{pl}, C_{pl} denote the resistance per unit length, inductance per unit length and capacitance per unit length, respectively. Then, $R_{sec}, L_{sec}, C_{sec}$ are calculated as follows

$$R_{sec} = R_{pl} \times \frac{l}{n}, \tag{3.10}$$

$$L_{sec} = L_{pl} \times \frac{l}{n}, \tag{3.11}$$

$$C_{sec} = C_{pl} \times \frac{l}{n}. \tag{3.12}$$

where n is calculated from (3.8).

Critical length is defined as the value of l that makes the propagation delay of a delay-optimized interconnect equal to that of a TSV. Critical length is similarly defined for RMS power consumption. It is the value of l that makes the RMS power consumption of the interconnect equal to that of a TSV whose *RLC* model is shown in Figure 3.6.

3.4 Repeater insertion

In order to make a fair comparison between interconnects and TSVs, interconnects are optimized for delay. This is achieved by repeater insertion. Repeater insertion is a method used to reduce the propagation delay across long interconnects by placing repeaters at regular intervals

along the interconnects [15]. Repeater insertion essentially reduces the quadratic dependence of delay on interconnect length to linear. However, there exists an optimum number of repeaters beyond which the delay overhead of the repeaters offsets any reduction in delay. Subsection 3.4.1 deals with repeater insertion for *RLC* interconnects.

3.4.1 *RLC* repeater insertion

The concept of *RC* repeater insertion is extended to include inductance for repeater insertion along *RLC* interconnects. The optimum size of repeaters h_{RLCopt} is [16]

$$h_{RLCopt} = \sqrt{\frac{R_0 C_{int}}{R_{int} C_0}} \times \left[1 + 0.16 \left(T_{L/R}\right)^3\right]^{-0.24}. \tag{3.13}$$

The optimum number of repeaters k_{RLCopt} is [16]

$$k_{RLCopt} = \sqrt{\frac{R_{int} C_{int}}{2.3 R_0 C_0}} \times \left[1 + 0.18 \left(T_{L/R}\right)^3\right]^{-0.3}, \tag{3.14}$$

where R_{int} is the interconnect resistance, C_{int} is the interconnect capacitance, R_0 is the output resistance of a minimum-sized inverter and C_0 is the input capacitance of a minimum-sized inverter. $T_{L/R}$ is the parameter that characterises the relative importance of interconnect inductance L_{int} [16] ,

$$T_{L/R} = \sqrt{\frac{L_{int}}{R_{int} R_0 C_0}}. \tag{3.15}$$

3.5 TSV model

The TSV model is a lumped *R, L, C* interconnect model (pi model), but with the *R, L, C* replaced by $R_{tsv}, L_{tsv}, C_{tsv}$. Figure 3.6 shows the *RLC* model of a TSV. The C_{tsv} parameter is only the liner capacitance of the TSV (C_{ox}). The effect of TSV MOS capacitance is explored in more detail in Section 3.8.

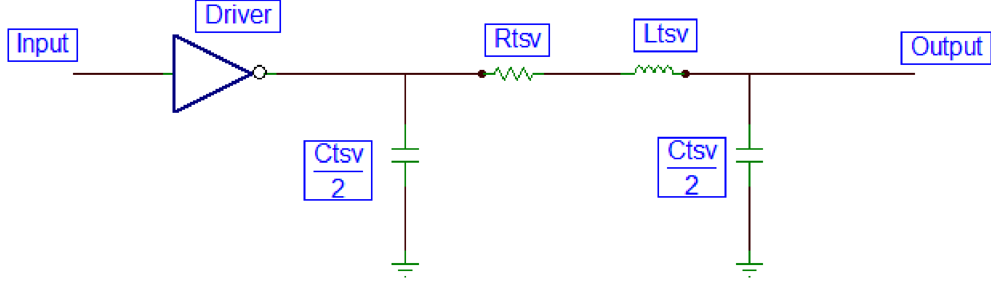

Figure 3.6: *RLC* model of a TSV

3.5.1 Calculation of TSV parameters

This section deals with the calculation of the TSV parameters such as resistance (R_{tsv}), inductance (L_{tsv}) and capacitance (C_{tsv}) as a function of TSV dimensions. In order to compare TSVs with interconnects, two-dimensional *RLC* model of TSV is used. The TSV resistance is given by [5]

$$R_{tsv} = \sqrt{(R_{tsv}^{ac})^2 + (R_{tsv}^{dc})^2}, \tag{3.16}$$

where R_{tsv}^{ac} is the AC resistance and R_{tsv}^{dc} is the DC resistance. DC resistance is given by

$$R_{tsv}^{dc} = \frac{\rho_f l_{tsv}}{\pi (r_{tsv})^2}, \tag{3.17}$$

where ρ_f is the resistivity of the TSV filling material, l_{tsv} is the height of the TSV and r_{tsv} is the radius of the TSV. AC resistance is given by

$$R_{tsv}^{ac} = \frac{\rho_f l_{tsv}}{2\pi r_{tsv} f}, \tag{3.18}$$

where f is the operating frequency of the TSV, as given by (3.9). TSV inductance is given by

$$L_{tsv} = \frac{\mu_0}{4\pi} \times \left[2l_{tsv} \ln \frac{2l_{tsv} + \sqrt{(r_{tsv})^2 + (2l_{tsv})^2}}{r_{tsv}} + r_{tsv} - \sqrt{(r_{tsv})^2 + (2l_{tsv})^2} \right], \tag{3.19}$$

where μ_0 is the permeability of vacuum. TSV capacitance is given by

$$C_{tsv} = \frac{2\pi\epsilon_{ox}l_{tsv}}{\ln\frac{r_{tsv}+t_{ox}}{r_{tsv}}}, \tag{3.20}$$

where ϵ_{ox} is the permittivity of SiO_2 (oxide liner) and t_{ox} is the thickness of the oxide liner [17]. This equation describes the oxide liner capacitance, and does not consider the voltage dependence of the TSV MOS capacitance [12]. Delay and power computations considering TSV MOS capacitance are described in Section 3.8.

3.6 Analysis setup

Analyses are performed in HSPICE using a 65 nm CMOS technology node. Dimensions of via-first and via-last TSVs are determined according to [14], as shown in Tables 3.1 and 3.2. Figure 3.7 shows the analysis setup for the repeater-inserted *RLC* interconnect, driving a load capacitance of $10fF$. Figure 3.8 shows the analysis setup for the TSV driving a load capacitance of $10fF$. The driver is a minimum-sized inverter (aspect ratio of NMOS is 1.85 and aspect ratio of PMOS is twice that of NMOS).

Radius of the TSV	$2\mu m$
Height of the TSV	$10\mu m$
TSV resistance	5.7Ω
TSV inductance	$4.2pH$
TSV capacitance	$23fF$

Table 3.1: Via-first TSV dimensions and parameters

Radius of the TSV	$5\mu m$
Height of the TSV	$60\mu m$
TSV Resistance	$20m\Omega$
TSV Inductance	$35pH$
TSV Capacitance	$283fF$

Table 3.2: Via-last TSV dimensions and parameters

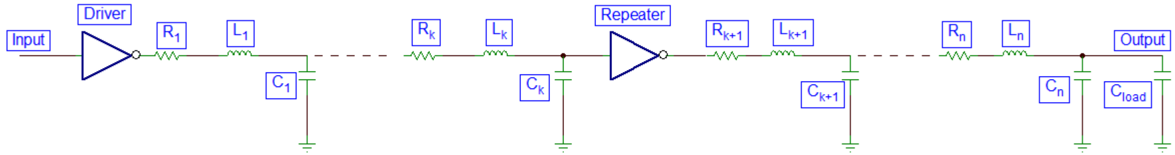

Figure 3.7: *RLC* interconnect analysis setup

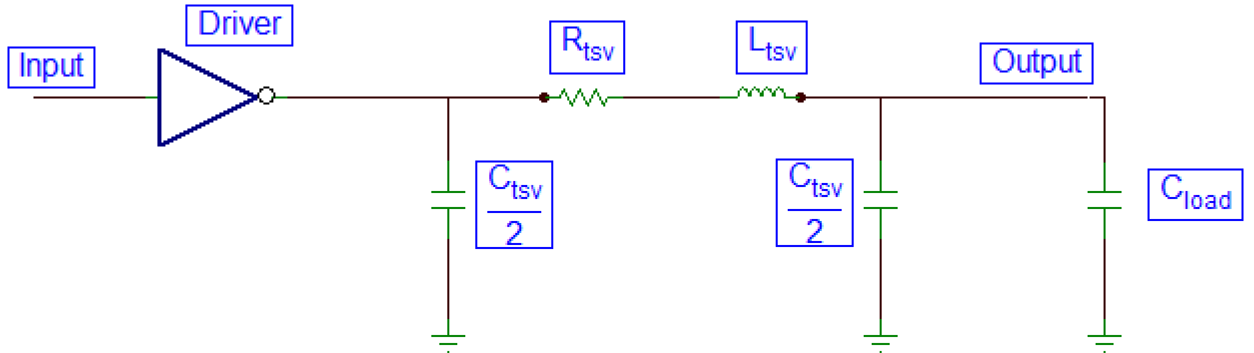

Figure 3.8: TSV analysis setup

3.7 Estimation of critical length

This section compares the delay-optimized interconnects with TSVs. Ratio of the delay of interconnect to the delay of TSV is plotted since this quantifies how much benefit is accrued by TSV-based 3D integration. This analysis is performed for a range of supply voltages that covers the super-threshold region as well as the near-threshold, and sub-threshold regions.

It is found that the via-last TSV is equivalent (in terms of propagation delay) to an *RLC* interconnect of length $1.6mm$ while via-first TSV is equivalent to an *RLC* interconnect of length $130\mu m$. This result is due to the fact that the TSV capacitance of via-last TSV is $283fF$ while that of via-first TSV is $23fF$. The contour plots for the delay ratios are also included. The critical lengths remain the same for RMS power consumption.

3.7.1 Via-first TSV

Figure 3.9 shows the variation of the ratio of *RLC* interconnect delay to via-first TSV delay. The change in ratio is negligible with change in supply voltage. This is because the delay of

the TSV also varies exponentially with decrease in supply voltage, due to the dominant driver resistance in near/sub-threshold regions. Figure 3.10 shows the variation of the ratio of RMS power consumed by *RLC* interconnect to that consumed by via-first TSV. Figure 3.11 shows the contour plot for the ratio of interconnect delay to delay of the via-first TSV. The contour plot shows the negligible change in the interconnect-to-TSV delay ratio with supply voltage.

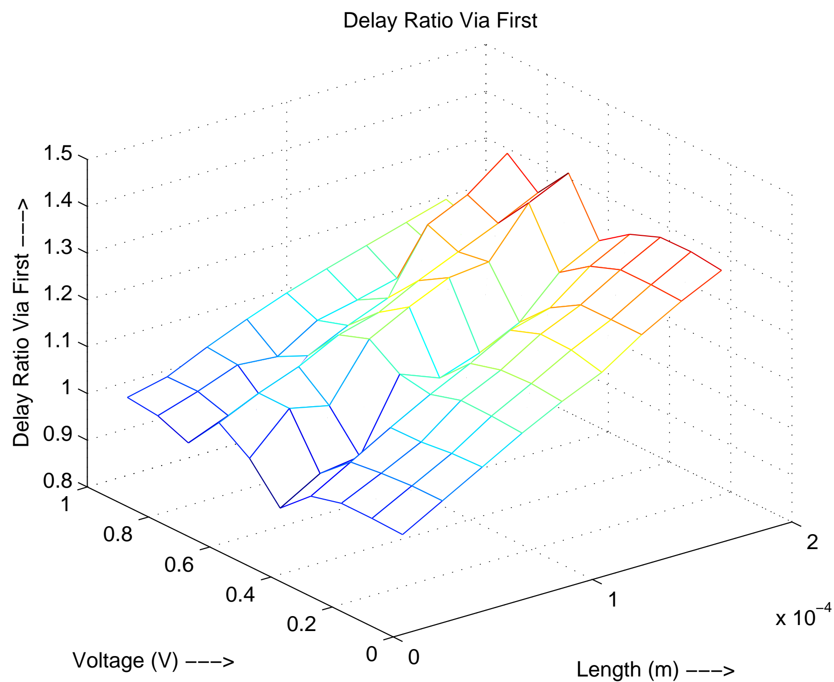

Figure 3.9: Ratio of *RLC* interconnect delay to via-first TSV delay

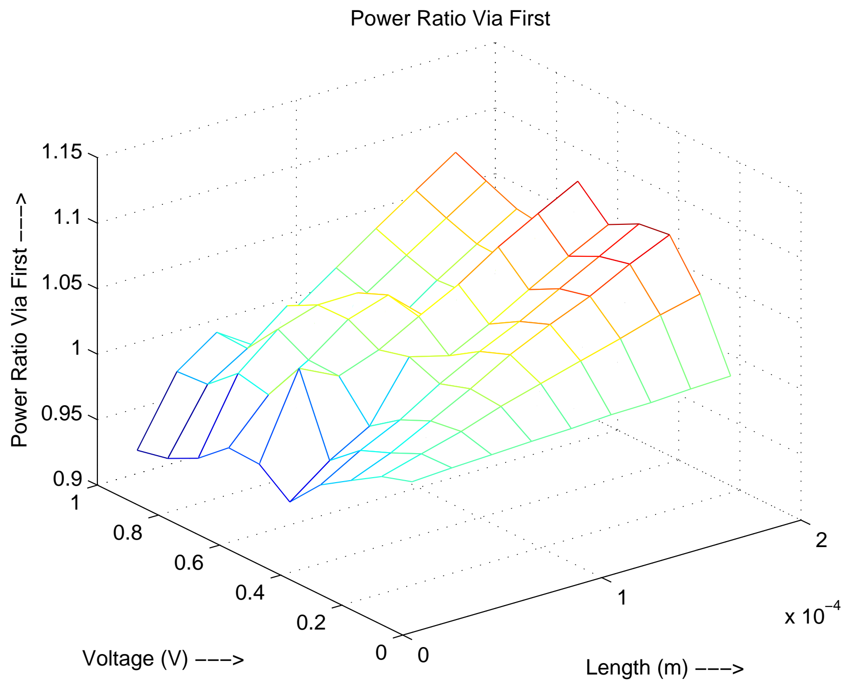

Figure 3.10: Ratio of *RLC* interconnect power to via-first TSV power

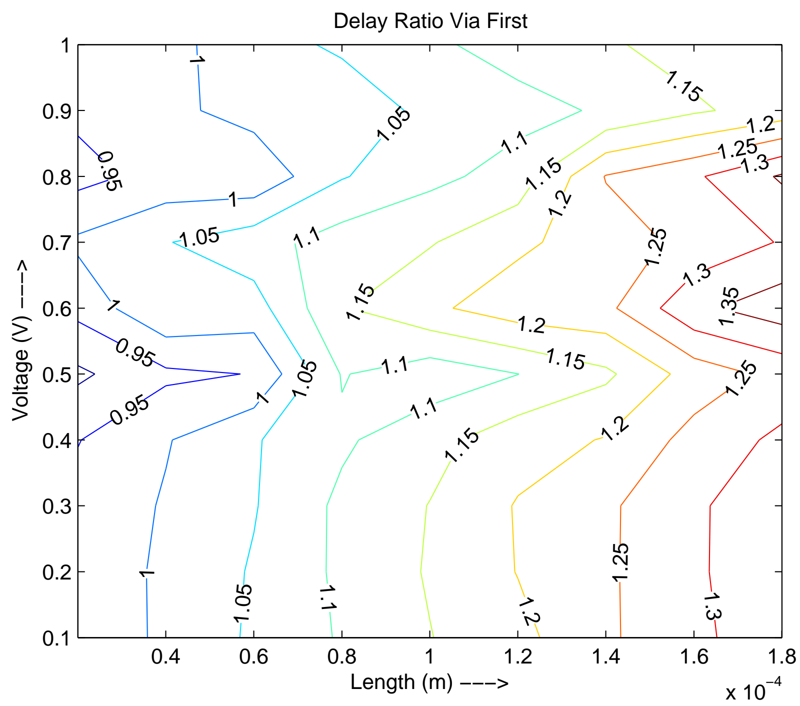

Figure 3.11: Contour curves at various delay ratios for via-first TSVs

3.7.2 Via-last TSV

Figure 3.12 shows the ratio of *RLC* interconnect delay to the delay of via-last TSV. The plot shows that there is negligible change in the ratio with change in supply voltage. That is, whether the region of operation is super-threshold or near/sub-threshold, the ratio remains roughly constant. Since the capacitance of the TSV and the interconnect remain the same, the delay ratio does not change significantly along the V_{dd} axis. However, this does not consider the variation of the TSV MOS capacitance.

Figure 3.13 shows the variation of the ratio of RMS power consumed by the *RLC* interconnect to that of the via-last TSV. A similar trend to that of the delay ratio is found since the *R,L,C* parameters of the interconnect do not change with supply voltage. Figure 3.14 shows the contour plot for the interconnect-to-TSV delay ratio vs supply voltage graphs for via-last TSVs. The contour plot shows negligible change in the interconnect-to-TSV delay ratio with supply voltage.

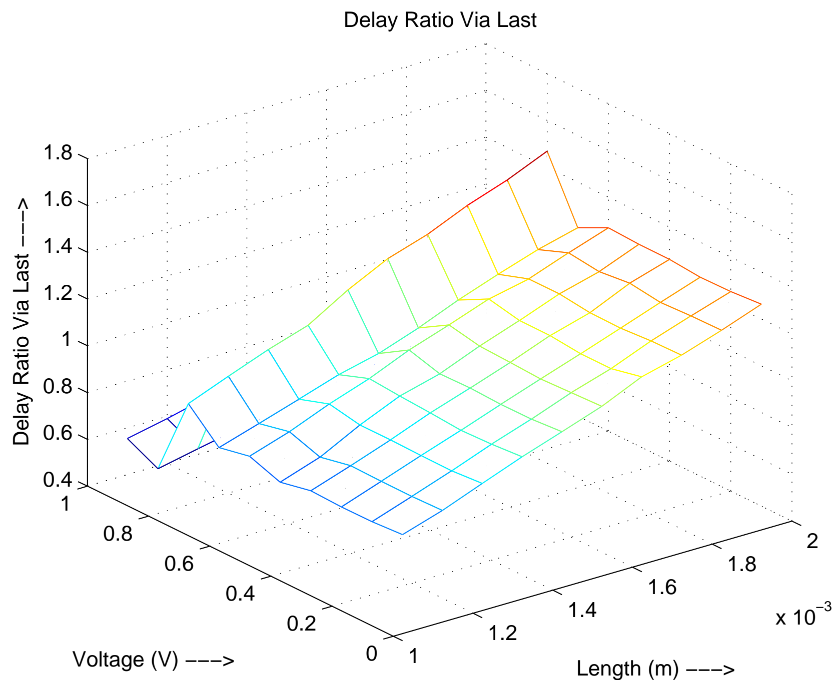

Figure 3.12: Ratio of *RLC* interconnect delay to via-last TSV delay

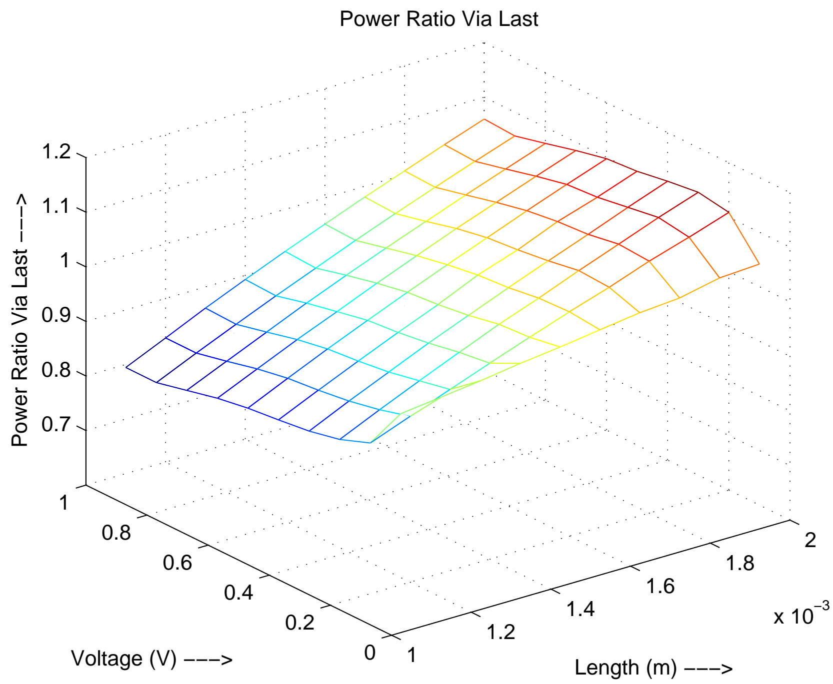

Figure 3.13: Ratio of *RLC* interconnect power to via-last power

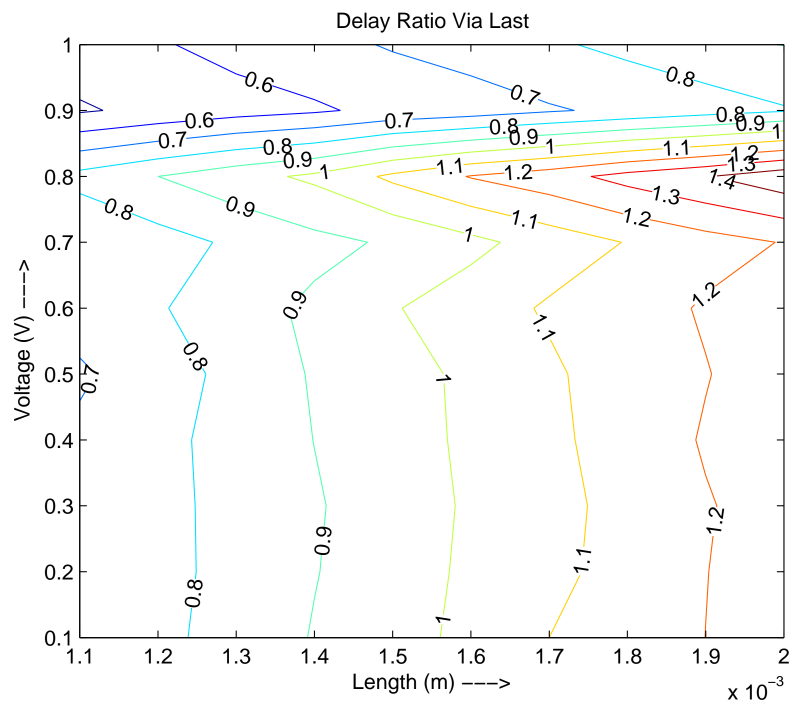

Figure 3.14: Contour curves at various delay ratios for via-last TSVs

3.8 Effect of TSV MOS capacitance

The analyses performed in the previous section are based on a TSV model that calculates the parameters of the TSVs based on their dimensions. These models ignore the voltage-dependence of the TSV MOS capacitance [12].

Subsection 3.8.1 first investigates the variation of MOS capacitance with respect to bias voltage and applies this to TSVs, which have a MOS structure ('M' is the filling material of TSV, 'O' is the oxide liner of the TSV, 'S' is the silicon substrate). Subsection 3.8.2 quantifies the error in delay and power values when the TSV MOS capacitance is ignored.

3.8.1 Variation of MOS capacitance with bias voltage

In order to explore the effect of MOS capacitance on TSVs at near/sub-threshold regions, it is useful to understand the variation of MOS capacitance with bias voltage.

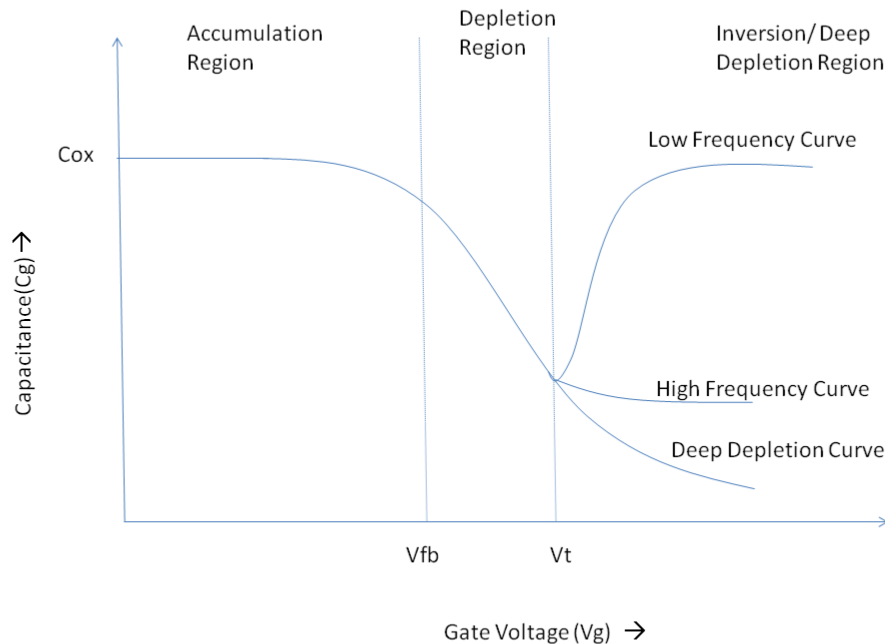

Figure 3.15: Variation of MOS capacitance with bias voltage [12]

Figure 3.15 shows the variation of MOS capacitance of a transistor with bias voltage for three different regions [12] : accumulation, depletion and inversion. The MOS structure of a TSV exhibits similar behaviour, with the exception that the gate voltage along x-axis is replaced

by the bias voltage of the TSV. Bias voltage of the TSV is defined as the potential difference between the TSV filling material and the substrate. It can be observed from Figure 3.15 that in high frequency, the TSV MOS capacitance increases with decrease in bias voltage with the maximum being the oxide liner capacitance C_{ox}. This means that the error in computation of delay and power by ignoring the voltage dependence of TSV MOS capacitance decreases as bias voltage is reduced. Capacitance (per unit length)-voltage graphs for various TSV diameters [12], shown in Figure 3.16, are used to obtain effective capacitance values for via-first and via-last TSVs. Tables 3.3 and 3.4 show the TSV liner capacitance and effective TSV MOS capacitance values at different supply voltages.

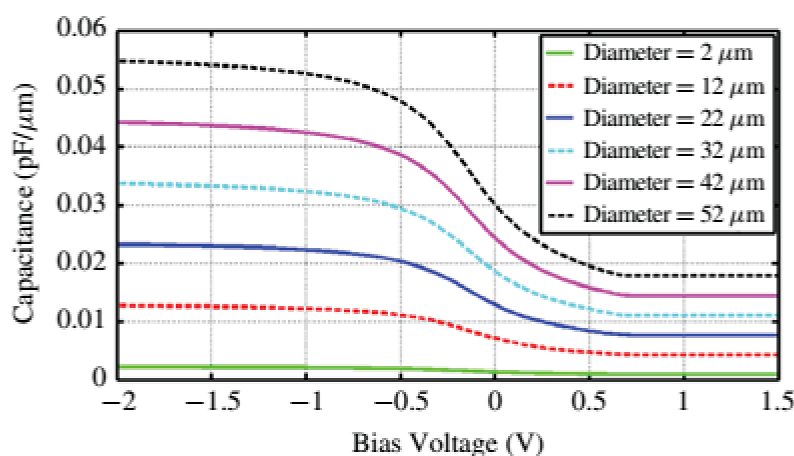

Figure 3.16: Capacitance per unit length vs bias voltage of the TSV [12]

Supply Voltage	Liner Capacitance, C_{ox} (fF)	Effective Capacitance, C_{eff} (fF)
Super-threshold (1 V)	60	30
Near-threshold (0.5 V)	60	35
Sub-threshold (0.1 V)	60	40

Table 3.3: Capacitance values for via-first TSV [12]

Supply Voltage	Liner Capacitance, C_{ox} (fF)	Effective Capacitance, C_{eff} (fF)
Super-threshold (1 V)	720	300
Near-threshold (0.5 V)	720	360
Sub-threshold (0.1 V)	720	400

Table 3.4: Capacitance values for via-last TSV [12]

3.8.2 TSV *RLC* Model

Tables 3.5 and 3.6 show the propagation delays with and without considering voltage dependence of TSV MOS capacitance. It is observed that the error in propagation delay is reduced in sub-threshold region since the TSV MOS capacitance is closer to the liner capacitance than at super-threshold region.

Tables 3.7 and 3.8 show the RMS power consumption with and without TSV MOS capacitance. Similar to delay, the error in power consumption also reduces with supply voltage since the difference between TSV MOS capacitance and liner capacitance is less at sub-threshold region than at super-threshold region. The error in power consumption is approximately zero at sub-threshold region due to the sufficiently low sub-threshold current (compared to super-threshold current). Thus, for an incremental change in capacitance, the difference in power consumption is sufficiently small.

Supply Voltage	Delay for C_{ox} (s)	Delay for C_{eff} (s)	Percentage Error
Super-threshold $(1V)$	$1.401e-11$	$8.233e-12$	41.23
Near-threshold $(0.5V)$	$7.440e-11$	$4.561e-11$	38.7
Sub-threshold $(0.1V)$	$5.212e-08$	$3.846e-08$	26.21

Table 3.5: Delay values for via-first TSV

Supply Voltage	Delay for C_{ox} (s)	Delay for C_{eff} (s)	Percentage Error
Super-threshold $(1V)$	$1.190e-10$	$5.088e-11$	57.24
Near-threshold $(0.5V)$	$6.935e-10$	$3.551e-10$	48.8
Sub-threshold $(0.1V)$	$5.022e-07$	$2.883e-07$	42.59

Table 3.6: Delay values for via-last TSV

Supply Voltage	Power for C_{ox} (W)	Power for C_{eff} (W)	Percentage Error
Super-threshold (1 V)	$1.830e-06$	$1.431e-06$	21.8
Near-threshold (0.5 V)	$2.755e-07$	$2.475e-07$	10.16
Sub-threshold (0.1 V)	$7.633e-09$	$7.630e-09$	0.04

Table 3.7: Power values for via-first TSV

Supply Voltage	Power for C_{ox} (W)	Power for C_{eff} (W)	Percentage Error
Super-threshold (1 V)	$5.286e-06$	$3.464e-06$	34.47
Near-threshold (0.5 V)	$6.125e-07$	$4.607e-07$	24.78
Sub-threshold (0.1 V)	$7.727e-09$	$7.682e-09$	0.58

Table 3.8: Power values for via-last TSV

Chapter 4

Conclusion

In Chapter 1, we have introduced the need for high-performance, low-power systems, and the benefits that 3D integration offers in this regard. The challenges involved in 3D integration have also been outlined. Benefits and limitations of sub/near-threshold operation have been described.

Chapter 2 has summarized TSV-based 3D integration and the types of TSVs based on the CMOS process flow in which the TSVs are fabricated as well as the TSV filling materials used.

Chapter 3 has dealt with the issues involved in interconnects in sub-threshold circuits, and highlighted the advantages of 3D integration for sub-threshold circuits. Increase in sensitivity of delay to capacitance has been quantitatively explained. Critical lengths of interconnects for delay and power remain constant in super-threshold as well as near/sub-threshold regions.

Critical lengths for via-first and via-last TSVs are determined to be $1.6\ mm$ and $130\ \mu m$ respectively. This means that, in terms of propagation delay, a via-last TSV is equivalent to a delay-optimized *RLC* interconnect of length $1.6\ mm$, and a via-first TSV is equivalent to a delay-optimized *RLC* interconnect of length $130\ \mu m$.

The effects of TSV MOS capacitance on delay and power have been explored at super-threshold (1 V), near-threshold (0.5 V) and sub-threshold (0.1 V) regions. The error in delay reduces with decrease in supply voltage. The error in power follows a similar trend. This is because the TSV MOS capacitance is closer to the oxide liner capacitance at sub-threshold

than at super-threshold. For via-last TSVs, the error in delay reduces from 57.24 % in super-threshold region to 42.59 % in sub-threshold region. The error in power consumption of via-last TSVs decreases from 34.47 % in super-threshold region to 0.58 % in sub-threshold region. For via-first TSVs, error in delay decreases from 41.23 % in super-threshold region to 26.21 % in sub-threshold region. The error in power computation for via-first TSVs decreases from 21.8 % in super-threshold region to 0.04 % in sub-threshold region.

Bibliography

[1] S.D. Pable and M. Hasan, "Ultra-low-power signaling challenges for subthreshold global interconnects," *Integration, The VLSI Journal*, Vol. 45, No. 2, pp 186-196, 2012.

[2] V. F. Pavlidis and E. G. Friedman, "Three-Dimensional Integrated Circuit Design", Morgan Kaufmann, 2009.

[3] A. Wang, B. H. Calhoun, and A. Chandrakasan, "Sub-threshold design for ultra low-power systems," Springer, 2005.

[4] D. Markovic, C.C.Wang, L.P. Alarcon, T.T. Liu, and J.M. Rabaey, "Ultralow-Power Design in Near-Threshold Region," *Proceedings of the IEEE*, Vol. 98, No. 2, pp 237-252, February 2010.

[5] G. Katti, M. Stucchi, K. De Meyer, and W. Dehaene, "Electrical modeling and characterization of through silicon via for 3-D ICs," *IEEE Transactions on Electron Devices*, Vol. 57, No. 1, pp 256-262, Jan. 2010.

[6] A.C. Hsieh and T.T. Hwang, "TSV Redundancy: Architecture and Design Issues in 3-D IC," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 20, No. 4, April 2012.

[7] E. Salman, M. H. Asgari, and M. Stanacevic, "Signal integrity analysis of a 2-D and 3-D integrated potentiostat for neurotransmitter sensing," *Proceedings of the IEEE Biomedical Circuits and Systems Conference*, pp 17-20, 2011.

[8] M.K. Hsu, Y.W. Chang, and V. Balabanov, "TSV-aware analytical placement for 3D IC designs," *ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp 664-669, June

2011.

[9] G. Katti, M. Stucchi, K. De Meyer, and W.Dehaene, "Through Silicon Via Capacitance Reduction Technique to Benefit 3-D IC Performance," *IEEE Electron Device Letters*, Vol. 31, No. 6, pp 549-551, June 2010.

[10] I. Savidis and E.G. Friedman, "Closed-Form Expressions of 3D-Via Resistance, Inductance, and Capacitance," *IEEE Transactions on Electron Devices*, Vol. 56, No. 9, pp 1873-1881, September 2009.

[11] K. Salah, A.E. Rouby, H. Ragai, K. Amin, and Y. Ismail, "Compact Lumped Element Model for TSV in 3D-ICs," *IEEE International Symposium on Circuits and Systems*, pp 2321-2324, May 2011.

[12] T. Bandyopadhyay, K.J. Han, D. Chung, R. Chatterjee, M. Swaminathan, and R. Tummala, "Rigorous Electrical Modeling of Through Silicon Vias (TSVs) with MOS Capacitance Effects," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, Vol. 1, No. 6, pp 893-903, June 2011.

[13] G. Chen and E.G. Friedman, "An RLC Interconnect Model Based on Fourier Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 24, No. 2, pp 170-183, February 2005.

[14] S.M. Satheesh and E. Salman, "Power Distribution in TSV-based 3D Processor-Memory Stacks," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS)*, Vol. 2, No. 4, pp. 692-703, December 2012.

[15] H.B. Bakoglu, "Circuits, Interconnections and Packaging for VLSI," Addison Wesley, 1990.

[16] E. Salman and E.G. Friedman, "High-Performance Integrated Circuit Design," McGraw-Hill, 2012.

[17] E. Salman, "Noise Coupling Due to Through Silicon Vias (TSVs) in 3-D Integrated Circuits," *IEEE International Symposium on Circuits and Systems*, pp 1411-1414, May

2011.

[18] A. Raychowdhury, B.C. Paul, S. Bhunia, and K. Roy, “Computing With Subthreshold Leakage: Device/Circuit/Architecture Co-Design for Ultralow-Power Subthreshold Operation,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol.13, No.11, pp 1213-1224, November 2005.

[19] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, “Low-power CMOS digital design,” *IEEE J. Solid-State Circuits*, Vol. 27, No. 4, pp 473-484, April 1992.

[20] H. Soeleman and K. Roy, “Ultra-low power digital subthreshold logic circuits,” *Proceedings of the IEEE International Symposium on Low Power Electronics and Design*, pp 94-96, 1999.

[21] D.H. Kim, K. Athikulwongse, and S. K. Lim, “A Study of Through-Silicon-Via Impact on the 3D Stacked IC Layout,” *Proceedings of the IEEE International Conference on Computer-Aided Design*, pp 674-680, 2009.

[22] R. Weerasekera, M. Grange, D. Pamunuwa, H. Tenhunen, and L.R. Zheng, “Compact Modelling of Through-Silicon Vias (TSVs) in Three-Dimensional (3-D) Integrated Circuits,” *IEEE Conference on 3D System Integration*, pp 1-8, September 2009.

[23] D.E. Khalil,Y. Ismail, M. Khellah, T. Karnik, and V. De, “Analytical Model for the Propagation Delay of Through Silicon Vias,” *IEEE International Symposium on Quality Electronic Design*, pp 553-556, March 2008.