# Through-Silicon-Via Aware Interconnect Prediction and Optimization for 3D Stacked ICs*

Dae Hyun Kim, Saibal Mukhopadhyay, and Sung Kyu Lim
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, Georgia
{daehyun, saibal.mukhopadhyay, limsk} @ ece.gatech.edu

## ABSTRACT

Individual dies in 3D integrated circuits are connected using through-silicon-vias (TSVs). TSVs not only increase manufacturing cost, but also incur silicon area, delay, and power overhead. However, the effects of TSV overheads have not been studied thoroughly in the literature. In this paper, we analyze the impact of TSVs on silicon area and wirelength. We derive a new 3D wirelength distribution model considering TSV size. Based on this new prediction model, we explain the impact of several design parameters newly introduced in 3D ICs. We also present a case study to show how the model can help make early design decisions for 3D ICs.

## Categories and Subject Descriptors

B.7.2 [**Integrated Circuits**]: Design Aids—*Simulation, Verification*; J.6 [**Computer-Aided Engineering**]: Computer-aided design (CAD)

## General Terms

Algorithms, Design, Experimentation, Theory

## Keywords

TSV, Through Silicon Via, Interconnect Prediction, Wirelength Distribution, 3D IC, Rent's Rule

## 1. INTRODUCTION

Technology advances have pushed functional integration to such a high level that the interconnect and package represent real barriers to further progress. While significant research effort has been expended on several different technology fronts, three-dimensional (3D) integration is now emerging as a leading contender in the challenge of meeting performance, power, cost, and size demands through this decade and beyond. The 3D integrated circuit is an emergent technology that vertically stacks multiple dies with a die-to-die interconnect so called through-silicon-via (TSV). TSV provides the possibility of arranging digital functional unit blocks across multiple dies at a very fine level of granularity. This results in a decrease in the overall wire length, which translates into less wire delay and less power. Advances in 3D integration and packaging are undoubtedly gaining momentum and have become of critical interest to the semiconductor community.

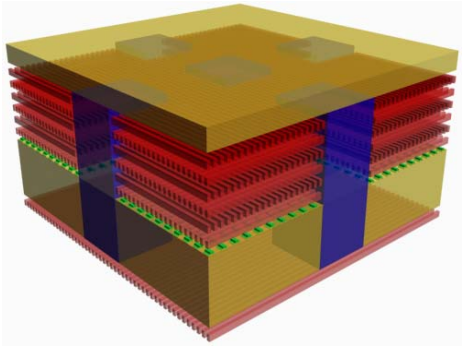

**Figure 1: Via-last through-silicon-via (TSV), where each TSV is surrounded by neighboring TSVs and wires. A typical size of TSV is much larger than that of global wires.**

The advantage of shorter wirelength mainly originates from the usage of TSVs. However, TSVs do have its negative impact. For example, TSVs consume silicon area as shown in Figure 1, and the additional area for TSVs increases total chip area. Moreover, TSVs act as obstacles during placement and routing. Depending on the types of TSVs, via-first TSVs occupy the device (= placement) layer, via-last TSVs occupy both the device and metal (= routing) layers. Therefore, one may need to increase chip area to address the placement and routing congestion caused by TSV insertion. These factors are primarily due to the non-negligible size of the TSVs (1 to 10$\mu m$ diameter typically). Therefore, their impact on area, power, and delay are indeed significant. However, most existing works related to TSV, especially in the field of interconnect prediction tend to ignore the TSV size impact on the overall silicon footprint area of 3D ICs.

In this paper we present a detail analysis of the effect of TSVs on critical properties of 3D ICs. This paper makes following contributions:

- We study the impact of signal and P/G (Power/Ground) TSVs on silicon area, footprint area, and wirelength.
- We build a new 3D wirelength distribution model incorporat-

*This material is based upon the work supported by the National Science Foundation under CAREER Grant No. CCF-0546382, the Center for Circuit and System Solutions (C2S2), and the Interconnect Focus Center (IFC).


*SLIP'09,* July 26–27, 2009, San Francisco, California, USA.

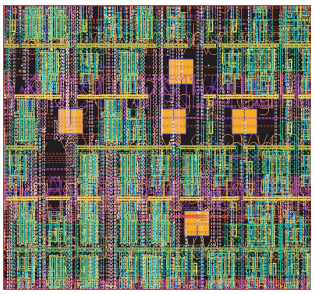

**Figure 2: TSV placed (orange square, shown in Cadence Virtuoso)**

ing various characteristics of 3D ICs such as TSV manufacturing types, whitespace, routing congestion, and so on.

- We study the impact of various 3D-related design parameters on silicon area, footprint area, and wirelength distribution.
- We show a case study of how the proposed models can be used for early design decisions on whether, for a particular circuit, 3D is a better option or not.

The rest of the paper is organized as follows. Section 2 reviews the prior works on TSV-aware 3D wirelength modeling; Section 3 explains the preliminaries such as types of TSVs, bonding styles, and stacking options; Section 4 presents the TSV-aware chip area model including signal and P/G TSVs; Section 5 presents the proposed 3D wirelength distribution model; Section 6 presents the proposed model; Section 7 shows the impacts of various parameters on silicon area, footprint area and wirelength; Section 8 draws conclusions.

## 2. RELATED WORKS

After the successful prediction of Davis' 2D wirelength distribution (WLD) model [1], a few works have extended it to 3D WLD model [2, 3, 4]. While other works assume one vertical pitch (= die height) is same as one gate pitch, the authors of [2, 4] introduced a new parameter $r$, which is the strata-to-gate-pitch ratio. The strata pitch (= die height) varies in a wide range depending on manufacturing technology such as die thinning, TSV materials, microfluidic channels for cooling, and so on. Since the gate placers should use less number of TSVs as the strata pitch goes up, the inclusion of $r$ has a significant importance. However, [2, 4] do not provide closed-form formulas, so the computation time is very long.

[5] extended Davis' 2D WLD model by introducing a new parameter $p_{gates}$, which is the percentage of die area occupied by logic gates. This is to explain the impact of whitespace existing in the placement layer. The authors show that the impact of $p_{gates}$ on wirelength could be as large as $10\%$ of the total wirelength.

While many studies have been done on 3D WLD, the impact of TSV size has not been mentioned in any of them. As a simple example, a $10\mu m \times 10\mu m$ signal TSV is comparable to about $50$ gates in terms of area in $45nm$ technology. If we use one million TSVs of this size, the TSVs occupy area of $50$ million gates, which is prohibitive. Thus, TSV size and count should be considered in 3D ICs.

## 3. PRELIMINARIES

The two most popular ways to fabricate TSVs are "via-first"and "via-last" processes, depending on when the via is implemented

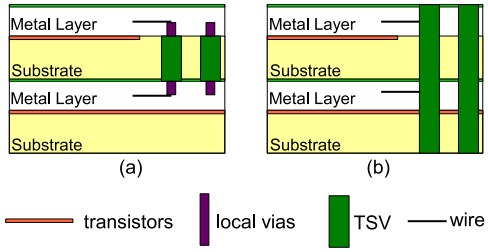


**Figure 3: Two types of TSVs. (a) via-first (b) via-last**

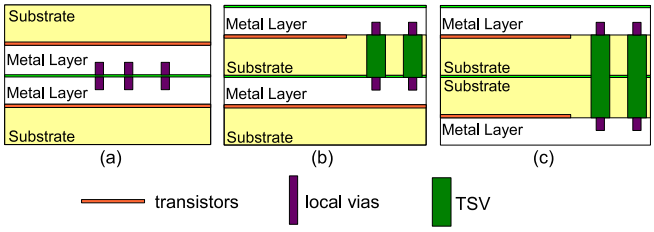


**Figure 4: Three types of bonding. (a) F2F (Face-to-Face) (b) F2B (Face-to-Back) (c) B2B (Back-to-Back)**

in the production process [6] (see Figure 3). Via-last TSVs are realized once the CMOS devices are completed and after the grinding and thinning process for wafer thinning. Via-last TSVs occupy all three layers: bulk, device, and metal layers, thereby becoming serious layout obstacles (see Figure 2). Via-first TSVs are implemented on the wafer prior to any production process, even before CMOS device fabrication. Via-first, however, is technically more challenging. Via-first TSVs are smaller compared to via-last, and TSVs occupy only two layers: bulk and device. This causes less interference with other layout objects.

Figure 4 shows three different bonding styles. Face-to-face (F2F) uses local vias or relatively small TSVs to connect to other dies. On the other hand, face-to-back (F2B) and back-to-back (B2B) bonding do need TSVs to maintain TSV aspect ratio [7]. As TSV size becomes bigger, more silicon area is used for the same number of TSVs.

3D stacking can be done at three different levels of granularity: gate-level, block-level, and chip-level. The gate-level stacking allows individual gates to be placed in any die in the 3D stack, whereas the block-level stacking requires that all gates in the same block stay together in the same die. However, each block can be placed in any die in the stack. The last chip-level stacking simply stacks entire 2D dies without any inter-die optimization. In terms of the number of TSVs required, the gate-level stacking contains the highest TSV count, whereas the chip-level stacking requires the lowest TSV count.

## 4. TSV-AWARE CHIP AREA MODEL

From this section, we assume F2B (Face-to-Back) bonding is applied to all dies. Other bonding styles, however, can be modeled in a similar way.

We introduce *TSV cell* to calculate area easily. A TSV is inside a TSV cell which has some whitespace around the TSV depending on design rules (see Figure 2). It comes from the minimum distance between two adjacent TSVs, between a TSV and an adjacent metal wire, or between a TSV and a transistor.

Table 1 shows the notations used in our modeling. Assuming that TSVs are evenly distributed between any two dies, the following

**Table 1: Notations**

| | |
|---|---|
| $N_{TSV,S}$ | average number of signal TSVs between two dies |
| $N_{TSV,PG}$ | average number of P/G TSVs between two dies |
| $A_{TSVcell,S}$ | area of a signal TSV cell |
| $A_{TSVcell,PG}$ | area of a P/G TSV cell |
| $A_{2D}$ | silicon area of a 2D chip |
| $A_{3D}$ | total silicon area of a 3D chip |
| $N_{DIE}$ | # dies |
| $A_{2DFP}$ | footprint area of a 2D chip |
| $A_{3DFP}$ | footprint area of a 3D chip |
| $N_{gates}$ | total # gates |
| $N_S$ | average # gates in a die |
| $p_{gates}$ | the percentage of die area occupied by logic gates [5] |
| $r$ | die-to-gate-pitch ratio in Figure 4 [2] |
| $M_S[l]$ | # gate pairs separated by $l$ gate pitches in a die |
| $M_z[v]$ | # die pairs separated by $v$ vertical pitches |
| $M_{tz}[l,v]$ | # gate pairs separated by $v$ vertical pitches and total $l$ gate pitches |
| $M_t[l]$ | # gate pairs separated by $l$ gate pitches in 3D |
| $I_{exp}[l]$ | # interconnects between two gate pairs separated by $l$ gate pitches |
| $\Gamma$ | normalization constant |
| $i(l)$ | normalized WLD without TSV size |
| $i^*(l)$ | normalized WLD with TSV size |
| $N_h(v)$ | the number of wires whose vertical length is $v$ |

equations hold.

$$A_{TSV,S} = N_{TSV,S} \cdot A_{TSVcell,S}$$

$$A_{TSV,PG} = N_{TSV,PG} \cdot A_{TSVcell,PG}$$

$$A_{3D} = A_{2D} + (N_{DIE} - 1) \cdot (A_{TSV,S} + A_{TSV,PG}) \quad (1)$$

$$A_{3DFP} = \frac{A_{3D}}{N_{DIE}}$$

$$\frac{A_{3D}}{A_{2D}} = 1 + (N_{DIE} - 1) \cdot (\frac{A_{TSV,S} + A_{TSV,PG}}{A_{2D}}) \quad (2)$$

$$\frac{A_{3DFP}}{A_{2DFP}} = \frac{1}{N_{DIE}} + (1 - \frac{1}{N_{DIE}}) \cdot (\frac{A_{TSV,S} + A_{TSV,PG}}{A_{2D}}) \quad (3)$$

Equation (1) shows the total silicon area of the 3D chip. Since we assume F2B bonding, the bottommost die does not have TSVs. The additional silicon area, therefore, is the TSV area between two dies multiplied not by $N_{DIE}$ but by $N_{DIE} - 1$.

$A_{3D}/A_{2D}$ is 1, and $A_{3DFP}/A_{2DFP}$ is $1/N_{DIE}$ if TSV size is zero. In this case, the additional silicon area becomes zero because $A_{TSVcell,S}$ and $A_{TSVcell,PG}$ are zero. However, the silicon area and footprint area ratios of 3D to 2D are strongly related to the occupancy rate of signal and P/G TSVs as shown in equation (2) and (3). Therefore, TSV size impact should be considered during wirelength prediction.

# 5. INTERCONNECT PREDICTION MODEL

Although several papers such as [8] reported estimated wirelengths after placement, they did not include effects of increased area. Since the footprint area ratios are bigger than the ideal values ($1/N_{DIE}$) as shown in equation (3), the real wirelength becomes longer than expected. In this section, we derive a new 3D wirelength distribution which explains impact of TSV size.

## 5.1 New 3D Design Parameters

Before deriving the new wirelength distribution model, we introduce a few parameters which explain various phenomena caused by TSV insertion in 3D.

- $P_{TSV,place}$ : There are usually whitespaces in ICs [9] thus we can use those for TSV insertion. Some whitespaces, however, cannot be used because they are used for decap and minimization of congestion, the whitespaces are too small for TSV insertion, or they are far away from appropriate TSV locations. In these cases, we have to increase silicon area to insert TSVs. The increased area is formulated as

$$\Delta A_{TSV,place} = P_{TSV,place} \cdot A_{TSV} \quad (4)$$

$$P_{TSV,place} \geq 0 \quad (5)$$

where $A_{TSV}$ is the total area of inserted TSVs. If $P_{TSV,place}$ is 0, we can insert TSVs into the existing whitespace of the chip so that no additional silicon area is needed. If $P_{TSV,place}$ is 1, on the other hand, we have to increase the chip area whenever we insert TSVs because the existing whitespace cannot be used for TSV insertion. $P_{TSV,place}$ can be greater than 1 because insertion of a TSV cell may need a row creation if the design is based on standard cell libraries.

- $P_{TSV,route}$ : For via-first fabrication (Figure 3 (a)), routing congestion is mainly caused by connections between metal wires and TSVs. For via-last fabrication (Figure 3 (b)), routing congestion is mainly caused by inserted TSVs which make wires bypass the TSVs. We add this parameter to explain the different degree of routing congestion caused by various types of TSVs and bonding styles, and circuit characteristics such as # nets, # gates, and so on. The increased area is formulated as

$$\Delta A_{TSV,route} = P_{TSV,route} \cdot A_{TSV} \quad (6)$$

$$P_{TSV,route} \geq 0 \quad (7)$$

where $A_{TSV}$ is the total area of inserted TSVs. If $P_{TSV,route}$ is 0, no routing congestion is caused by TSV insertion, which happens when there are already enough space for connection between metal wires and TSVs. If $P_{TSV,route}$ is greather than 0, on the other hand, some wires bypassing TSVs cause congestion around the TSVs. In this case, we have to insert whitespace to resolve the congestion.

Then, the total silicon area in equation (1) becomes

$$A_{3D} = A_{2D} + \Delta A_{TSV,place} + \Delta A_{TSV,route} \quad (8)$$

- $B$ (Granularity parameter) : Placement can be done at gate-level or block-level. We can also choose a specific granularity of block size for block-level placement. This parameter explains how big blocks we use for 3D placement and is defined as the average number of blocks in a die. Therefore, gate-level placement is done when $B$ is equal to $N_{gates}/N_{DIE}$ and the most coarse block-level placement is done when $B$ is 1.

$$1 \leq B \leq \frac{N_{gates}}{N_{DIE}} \quad (9)$$

## 5.2 Gate-level 3D Wirelength Distribution

Now we show the new wirelength distribution considering TSV size at gate-level (see Table 1 for the notations). The normalized wirelength distribution without consideration of TSV size in [2, 4] is as follows.

$$M_{tz}[l,v] = M_z[v]M_S[l - vr] + N_S(N_{DIE} - v)\delta[l - vr] \quad (10)$$

$$M_t[l] = \sum_{v=0}^{N_{DIE}-1} M_{tz}[l,v] \quad (11)$$

$$i(l) = \Gamma \cdot M_t[l] \cdot I_{exp}[l] \quad (12)$$

The first term in the right-hand side of equation (10) becomes zero for wires whose horizontal length is zero (call these *PV* wires). On the other hand, the second term becomes zero for wires whose horizontal length is nonzero (call these *NPV* wires).

*PV* wires are not affected by TSV insertion because their horizontal wirelength is zero while the horizontal wirelength of *NPV* wires are affected by TSV insertion. Therefore we rewrite equation (12) as follows.

$$i(l) = \Gamma \cdot M_t[l] \cdot I_{exp}[l] = i_h(l) + i_v(l) \quad (13)$$

$$i_h(l) = \Gamma \cdot I_{exp}[l] \cdot \sum_{v=0}^{N_{DIE}-1} M_z[v] M_S[l - vr] \quad (14)$$

$$i_v(l) = \Gamma \cdot I_{exp}[l] \cdot \sum_{v=0}^{N_{DIE}-1} N_S(N_{DIE} - v)\delta[l - vr] \quad (15)$$

where $i_h(l)$ consists of *NPV* wires and $i_v(l)$ consists of *PV* wires. We modify only $i_h(l)$ because TSV size affects only the horizontal wirelength.

$i_h(l)$ in equation (14) can be rewritten as follows.

$$i_h(l) = \sum_{v=0}^{N_{DIE}-1} i_h(v, l) \quad (16)$$

$$i_h(v, l) = \Gamma \cdot I_{exp}[l] \cdot M_z[v] M_S[l - vr] \quad (17)$$

$$N_h(v) = \sum_{l=1}^{2\sqrt{N_S}} i_h(v, l) \quad (18)$$

where $i_h(v, l)$ is the wirelength distribution of wires whose total length is $l$ gate pitches and vertical length is $v$ vertical pitches. $N_h(v)$ is the number of *NPV* wires whose vertical length is $v$ vertical pitches. This number should be conserved for each $v$.

Now we derive the new WLD by re-normalization as follows.

$$i_h^*(v, l) = \Gamma(v)^* \cdot I_{exp}^*[l] \cdot M_z[v] M_S^*[l - vr] \quad (19)$$

$$\Gamma(v)^* = \frac{N_h(v)}{\sum_{l=1}^{2\sqrt{N_S^*}} I_{exp}^*[l] \cdot M_z[v] M_S^*[l - vr]} \quad (20)$$

$$N_S^* = A_{3D}/N_{DIE} \quad (21)$$

where $\Gamma(v)^*$ is the re-normalization constant for *NPV* wires whose vertical length is $v$ vertical pitches, $I_{exp}^*[l]$ is the modified expected number of interconnects connecting two gate socket pairs at a distance of $l$, and $M_S^*[l]$ is the modified total number of gate socket pairs at a distance of $l$. As seen in the above equations, $i_h^*(v, l)$ was re-normalized separately.

Then the new distribution becomes as follows.

$$i_h^*(l) = \sum_{v=0}^{N_{DIE}-1} i_h^*(v, l) \quad (22)$$

$$i^*(l) = i_h^*(l) + i_v(l) \quad (23)$$

The modified $M_S^*[l]$ is as follows.

$$M_S^*[l] = M_S[l] - OVR[l] \quad (24)$$

where $OVR[l]$ is calculated by computing the number of gate pairs which are $l$ gate pitches away. One of the gates in the pair should be inside a TSV cell.

1: Compute average intra-block wirelength
2: Compute average inter-block wirelength
3: Compute # TSVs
4: Compute average intra-block wirelength with TSV size
5: Compute average inter-block wirelength with TSV size

**Figure 5: Derivation of block-level 3D wirelength distribution**

$I_{exp}^*(l)$ is computed as follows.

$$N_A^* = 1$$

$$N_B^*[l] = N_B[l] \cdot (1 - R_{tg}) \cdot p_{gates}$$

$$N_C^*[l] = N_C[l] \cdot (1 - R_{tg}) \cdot p_{gates}$$

$$I_{exp}^*[l] = \frac{p_{gates} \cdot \alpha k}{N_A^* N_C^*[l]} \begin{bmatrix} (N_A^* + N_B^*[l])^p + (N_B^*[l] + N_C^*[l])^p \\ -(N_B^*[l])^p - (N_A^* + N_B^*[l] + N_C^*[l])^p \end{bmatrix}$$

$$R_{tg} = \frac{L_{TSV}}{\sqrt{A_{3D}/N_{TSV}}}$$

where $L_{TSV}$ is the width of a TSV cell, and $N_{TSV}$ is the total number of TSVs. In our model, we fix the gate pitch, which is same as $L_{gate}$. Therefore, $p_{gates}$ is defined as follows.

$$p_{gates} = \frac{N_{gates}}{A_{2D}/{L_{gate}}^2} \quad (25)$$

## 5.3 Block-level 3D Wirelength Distribution

The modeling of block-level 3D wirelength distribution is done hierarchically. The derivation flow is shown in Figure 5.

### 5.3.1 Intra-block wirelength without TSV size

Intra-block wirelength without TSV size is computed by 2D wirelength distribution. In order to compute this, we use equations in [5]. The total number of interconnects in a block is calculated as follows.

$$I_{B,intra} = \alpha \cdot k \cdot N_{Bg} \cdot (1 - {N_{Bg}}^{p-1}) \quad (26)$$

where $\alpha$, $k$ and $p$ are Rent's constants [10], and $N_{Bg}$ is the number of gates in a block.

### 5.3.2 Inter-block wirelength without TSV size

Inter-block wirelength without TSV size is computed by 3D wirelength distribution. In this case, however, we treat a block as a gate and apply equation (12). The total number of inter-block connections is calculated as follows.

$$I_{total} = \alpha \cdot k \cdot N_{gates} \cdot (1 - {N_{gates}}^{p-1}) \quad (27)$$

$$I_{B,inter} = I_{total} - I_{B,intra} \cdot N_{DIE} \cdot B \quad (28)$$

where $I_{total}$ is the total number of interconnects in the circuit and $B$ is the granularity parameter defined in Section 5.1.

### 5.3.3 Computation of # TSVs

We assume that an inter-block connection exists between two gates separated by distance $D_B$. $D_B$ is defined as follows.

$$D_B = n_H \cdot L_H + n_V \cdot r \quad (29)$$

where $L_H$ is the average distance between two adjacent blocks in a die, and $n_H$ and $n_V$ are integers greater than or equal to zero. Then the total number of TSVs is computed in a similar way shown in [2].

**Table 2: Validation of our prediction on TSV count for block-level placement.**

| circuit | [8] Folding-2 | ours ($B$=1) | Dif. (%) | [8] Folding-4 | ours ($B$=4) | Dif. (%) |
|---|---|---|---|---|---|---|
| ibm01 | 1671 | 1595 | 4.55 | 2476 | 3852 | -55.57 |
| ibm03 | 4125 | 2487 | 39.71 | 5909 | 6006 | -1.64 |
| ibm04 | 2940 | 2850 | 3.06 | 6388 | 6883 | -7.75 |
| ibm06 | 4116 | 3285 | 20.19 | 9077 | 7933 | 12.60 |
| ibm07 | 5932 | 4233 | 28.64 | 8755 | 10222 | 16.76 |
| ibm08 | 5801 | 4638 | 20.05 | 10181 | 11199 | -10.00 |
| ibm09 | 4540 | 4690 | -3.30 | 8257 | 11326 | -37.17 |
| ibm13 | 7696 | 6594 | 14.32 | 13071 | 15923 | -21.82 |
| ibm15 | 15128 | 10845 | 28.31 | 23662 | 26187 | -10.67 |
| ibm18 | 12077 | 13425 | -11.16 | 28287 | 32415 | -14.59 |
| | Absolute Dif. | | 17.33 | Absolute Dif. | | 18.86 |

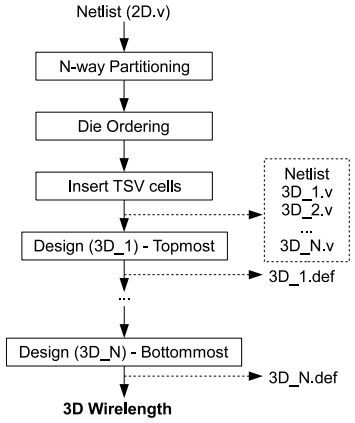


**Figure 6: 3D Circuit Design Scheme**

### 5.3.4 *Intra-block wirelength with TSV size*

The number of TSVs in a block is computed as follows.

$$N_{B,inter,TSV} = \frac{N_{B,TSV}}{(N_{DIE}-1)\cdot B} \tag{30}$$

where $N_{B,TSV}$ is the total number of TSVs obtained in the previous section. TSVs inserted into a block increase the block area. This area affects the wirelength of intra-block wires. The intra-block wirelength with TSV size is computed in a similar way shown in equation (19).

### 5.3.5 *Inter-block wirelength with TSV size*

The increased block area also affects the wirelength of inter-block *NPV* wires because the distance between two gates from two different blocks is increased. On the other hand, we assume that *PV* wires in block-level distribution are not affected by the increased block area for simplification.

The computation of inter-block wirelength with TSV size is done by equation (19).

# 6. VALIDATION OF OUR 3D MODEL

## 6.1 Validation of TSV Count

The authors of [8] designed 3D chips by folding 2D designs for various benchmarks. This design scheme is similar as block-level placement thus we compare the TSV counts in [8] with our block-level prediction.

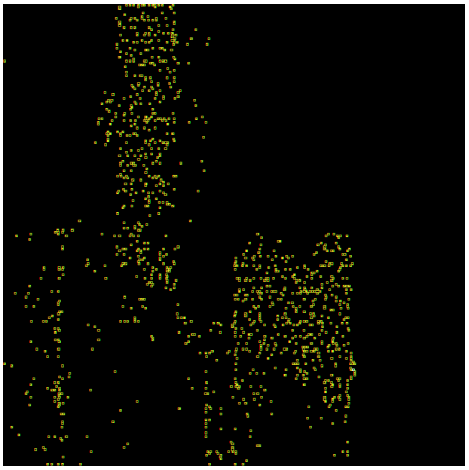

**Figure 7: Snapshot of TSVs inserted in the topmost die of the circuit "Ind 3" (see Table 3) in Cadence SoC Encounter. There are 1186 TSVs (yellow) and 80483 standard cells (black). Die area is $592\mu m \times 592\mu m$.**

**Table 4: Impact of TSV size consideration on wirelength. $N_{DIE}$ : 4, $p_{gates}$ : 1, $L_{gate}$ : $1.37\mu m$ and $L_{TSV}$ : $1.37\mu m$. $B = N_{gates}/N_{DIE}$ (gate-level).**

| $r$ | $N_{gates}$ | # TSVs | 2D WL | TSV size consideration | 3D WL | $\Delta$ WL |
|---|---|---|---|---|---|---|
| 5 | 1M | 0.66M | 17.23 | no | 11.23 | -34.82% |
| | | | | yes | 14.35 | -20.07% |
| | 100M | 75.2M | 53.96 | no | 29.82 | -44.74% |
| | | | | yes | 40.00 | -25.87% |
| 30 | 1M | 0.17M | 17.23 | no | 13.37 | -22.40% |
| | | | | yes | 14.82 | -13.99% |
| | 100M | 24.3M | 53.96 | no | 30.37 | -43.72% |
| | | | | yes | 34.51 | -36.05% |

Table 2 shows the comparison of TSV count for all the circuits reported in [8]. We note that our predictions match well with the reported numbers in most cases, although absolute difference in some cases is up to 30%.

## 6.2 3D Circuit Design Scheme

Figure 6 shows our 3D circuit design scheme for validation of 3D wirelength distribution. We first synthesize HDL source files with Synopsys Design Compiler [11]. Then we do N-way partitioning where N is the number of dies (=$N_{DIE}$). The area balancing factor we used is 0.05 (5%). Since TSV cells will be inserted to all dies except the bottommost die (notice that we are assuming F2B bonding) and inserting TSV cells increases die area, dies are sorted in the order of die area before TSV insertion so that the largest die is laid at the bottommost location. After inserting TSV cells, we do placement using Cadence SoC Encounter [12] for the topmost die (3D_1 in Figure 6) and extract TSV cell locations. Since the locations of these cells affect the placement of the next die (3D_2), we feed the locations into SoC Encounter during the placement of 3D_2. Placement for the remaining dies is done in a similar way.

Then routing (both global and detailed) is done for each placement result (3D_#.def). We used Cadence SoC Encounter for routing.

Figure 7 shows the snapshot of TSV cells in Cadence SoC Encounter. Yellow points are TSV cells and black area is actually

**Table 3: Validation of our prediction on wirelength. $N_{DIE} = 3$, $P_{TSV,place} = 1$ and $P_{TSV,route} = 0$.**

| | | | | Wirelength ($\mu m$) | | | | |
|---|---|---|---|---|---|---|---|---|
| circuit | # gates | # nets | # TSVs | 3D Design | Prediction (w/o TSV size) | Dif. | Prediction (with TSV size) | Dif. |
| Ind 1 | 11295 | 11839 | 688 | 315792 | 276465 | −12.45% | 300540 | −4.83% |
| Ind 2 | 29706 | 29979 | 1217 | 569482 | 452222 | −20.59% | 488465 | −14.22% |
| wb_conmax | 62028 | 63158 | 719 | 1108165 | 953519 | −13.96% | 994221 | −10.28% |
| Ind 3 | 260579 | 262357 | 1799 | 6991301 | 6320040 | −9.60% | 6504680 | −6.96% |
| netcard | 651674 | 653155 | 2261 | 23303968 | 21168400 | −9.16% | 21560700 | −7.48% |

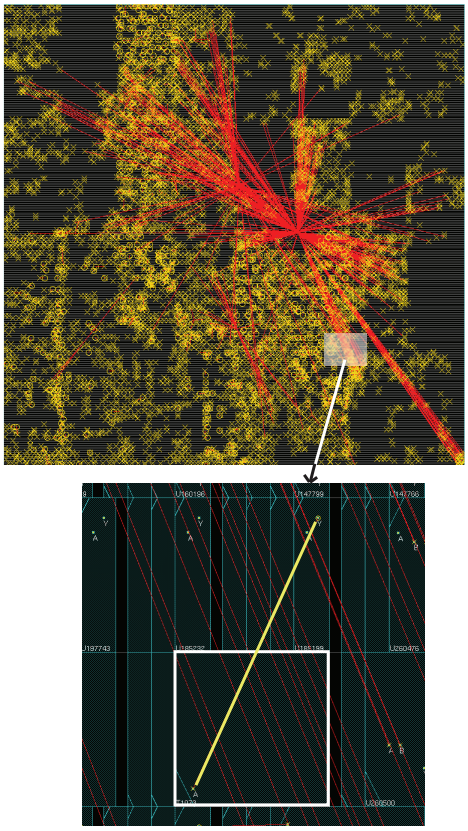

**Figure 8: Snapshot of connections to TSVs in the topmost die of the circuit "Ind 3" in Cadence SoC Encounter. There are 1186 connections to TSVs among 82167 nets. Die area is $592\mu m \times 592\mu m$. The white square is a TSV cell and thick yellow line shows the connection between the TSV cell and an inverter cell.**

filled with standard cells which are not shown. Figure 8 shows all the connections to TSV cells.

## 6.3 Validation of Wirelength Prediction

We used three industrial circuits (Ind 1, Ind 2 and Ind 3) and two IWLS'05 benchmark circuits (wb_conmax and netcard) [13] for validation of our wirelength prediction.

Table 3 shows the comparison of total wirelength for each circuit. As seen in the table, both predictions underestimate for the three circuits but prediction considering TSV size is more accurate.

Since this 3D design method does not optimize gate placement globally, we believe that the wirelength of circuits globally optimized in 3D with the same number of TSVs will be shorter than the wirelength in Table 3. Then the prediction will match the wirelength more closely.

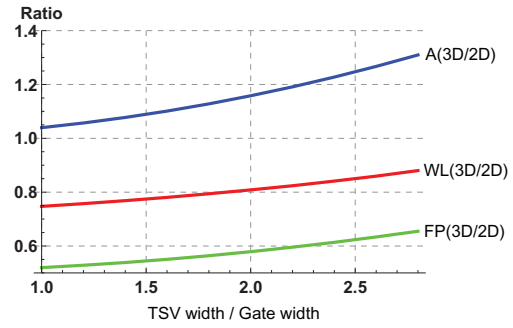


**Figure 9: Impact of TSV size on silicon area (A), footprint area (FP) and wirelength (WL). $r$ : 100**

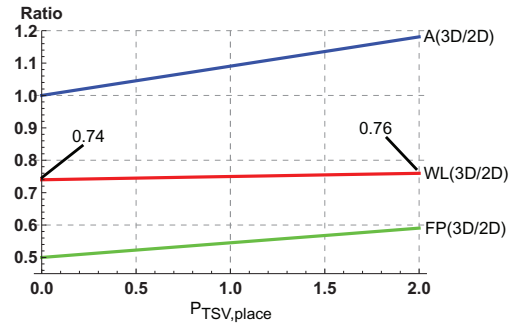


**Figure 10: Impact of $P_{TSV,place}$ on silicon area (A), footprint area (FP) and wirelength (WL).**

In order to make the prediction even more accurate, we need to determine the parameters such as $P_{TSV,place}$ and $P_{TSV,route}$ more carefully. In Table 3, we used fixed parameter values regardless of circuit characteristics to show how our prediction behaves.

# 7. IMPACT STUDY OF DESIGN PARAMETERS

## 7.1 Impact of TSV Size

Table 4 shows the impact of TSV size on wirelength. We note that the wirelength differences are about 10% to 20%. If $r$ is small, 3D placers tend to use more TSVs and the difference becomes greater. This is because the TSVs occupy larger space so that silicon area increase by TSV insertion affects the wirelength significantly. The difference also becomes more noticeable if the TSV size is relatively bigger than the gate size.

## 7.2 Impact Study

The Rent's constants in our experiments are $\alpha = 0.75$, $k = 4$ and $p = 0.75$. The parameters used for this study are as follows

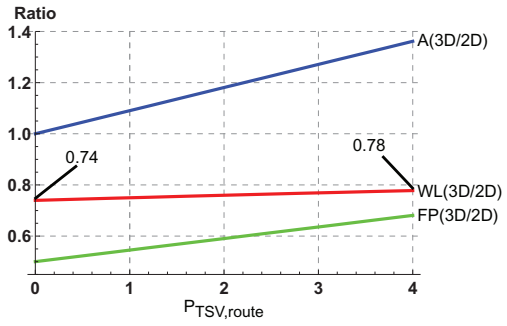


**Figure 11: Impact of $P_{TSV,route}$ on silicon area (A), footprint area (FP) and wirelength (WL).**

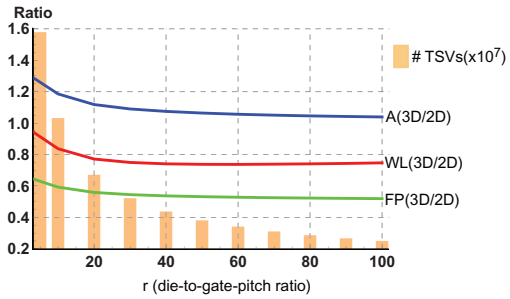


**Figure 12: Impact of $r$ on silicon area (A), footprint area (FP) and wirelength (WL).**

(see Table 1). $N_{gates} = 40M$. $L_{gate}$, which is the physical gate width, is $1.37\mu m$. The variable parameters are as follows (if not specified in each case): 2D die size = $100mm^2$, $L_{TSV} = 1.37\mu m$, $P_{TSV,place} = 1$, $P_{TSV,route} = 0$, $r = 30$ and $N_{DIE} = 2$. Lastly, we used gate-level stacking option, where $B = N_{gates}/N_{DIE}$.

- TSV size (Figure 9) : As TSV size increases, silicon area and footprint area increase, and so does wirelength. 3D WL becomes bigger than 2D WL if TSV size continues to go up, which means we cannot benefit from 3D with respect to WL. **In short, silicon area, footprint area and WL increase as TSV size increases.**

- $P_{TSV,place}$ (Figure 10) : 3D silicon area increases as $P_{TSV,place}$ goes up. Therefore, gates are spreaded out so that WL slightly increases. Since the silicon area increase strongly depends on TSV size as well as TSV count, the three ratios will become bigger if TSV size or TSV count increases. Moreover, even though the WL increase in this figure is small, this parameter should be kept as small as possible to save the cost for silicon area, i.e., better placement tool is needed. **In short, silicon area, footprint area and WL increase as $P_{TSV,place}$ increases.**

- $P_{TSV,route}$ (Figure 11) : This parameter affects the three metrics in the same way as $P_{TSV,place}$. However, the range of $P_{TSV,route}$ is larger than $P_{TSV,place}$ because of via-last fabrication. If a circuit is seriously congested, via-last TSVs will cause many wires to overlap with each other, thereby requiring $P_{TSV,route}$ to be more than 1 or 2. Figure 11 confirms a similar impact trend as $P_{TSV,place}$. **In short, silicon area, footprint area and WL increase as $P_{TSV,route}$ increases.**

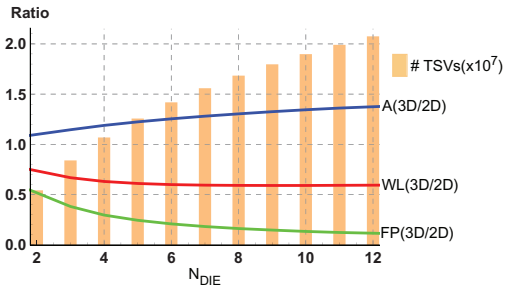


**Figure 13: Impact of $N_{DIE}$ on silicon area (A), footprint area (FP) and wirelength (WL).**

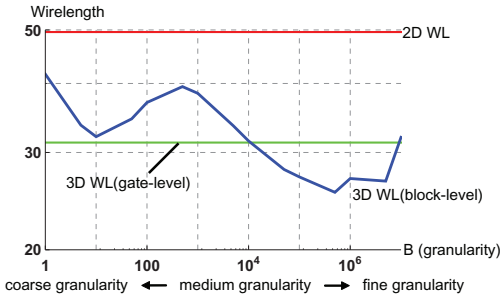


**Figure 14: Impact of $B$ on wirelength (WL). $N_{DIE}$ : 4.**

- $r$ (Figure 12) : Bigger $r$, which means taller die, means placers use less number of vertical connections. As the figure shows, the number of TSVs decrease as $r$ increases. The silicon area, footprint area and WL also decrease because less number of TSVs are used. However, note that the number of TSVs cannot be decreased below the min-cut size in real circuits. **In short, silicon area, footprint area and WL decrease as $r$ increases.**

- $N_{DIE}$ (Figure 13) : We use more and more TSVs when the number of dies increases. This increases silicon area, but WL decreases because of more vertical connections. The WL decrease saturates at some point. **In short, silicon area increases but footprint area and WL decrease as $N_{DIE}$ increases.**

- $B$ (Figure 14, 15 and 16) : As expected, when $B$ is 1, we have one big block in each die (coarse granularity). So, the silicon area increase is small, but WL decrease is also small. As $B$ goes up, silicon area and footprint area generally increase, but WL fluctuates. Wirelength reaches the minimum usually at fine granularity at which one block has about 20 to 100 gates. Area ratio reaches the minimum when $B$ is 1. **In short, silicon area and footprint area ratios increase but saturate at some point. On the other hand, WL fluctuates and reaches the minimum at fine granularity.**

## 7.3 Case Study

We show a case study in this section to demonstrate how to use our model for early decision making for 3D ICs. The technology parameters used are as follows: number of gates is 10M under $90nm$. TSV is via-last and its size is $2 \times 2\mu m$. Die height ratio $r$ is 20, and the 2D die area is $78mm^2$. Lastly, we set $P_{TSV,place} = 1.0$ and $P_{TSV,route} = 0.5$.

These are the steps we follow in our decision making: (1) We choose a circuit to be fabricated in 3D and calculate the number

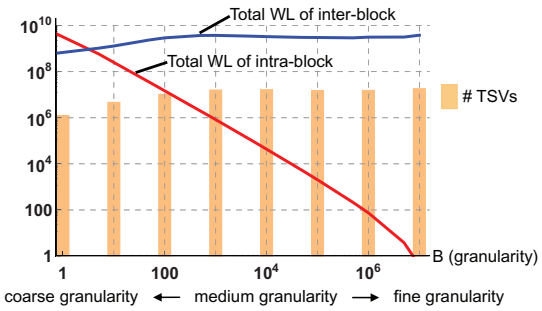


**Figure 15: Impact of $B$ on total wirelength of intra-block and inter-block, and TSV count. $N_{DIE}$ : 4.**

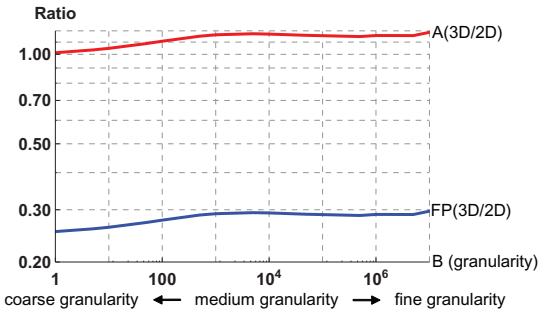


**Figure 16: Impact of $B$ on silicon area (A) and footprint area (FP). $N_{DIE}$ : 4.**

of gates. (2) We select fabrication technologies for the circuit and TSVs including bonding and TSV types. (3) We simulate the circuit in 2D with existing tools to estimate how much decaps we need, how serious the congestion is, how much the power consumption is, and so on. (4) Then we estimate two additional parameters $P_{TSV,place}$ and $P_{TSV,route}$. (5) Then we change $N_{DIE}$ and $B$ to estimate how many TSVs are used, how large the additional silicon area is needed and how much the wirelength is decreased. Table 5 shows the simulated values.

Based on this result, we notice that 3D placement in "large" block-level stacking (coarse granularity) uses less number of TSVs, thereby achieving smaller silicon area increase while decreasing sufficient amount of wirelength. On the other hand, 3D placement in "small" block-level stacking (finer granularity) uses many TSVs but decreases wirelength a lot. Our choice depends on what the most important factor is. If the yield of TSV fabrication is low so that TSV cost is high, coarse granularity 3D placement is the best option. If the TSV cost is low but die bonding cost is high, 2-die or 3-die stacking with fine granularity is the best choice. If TSV and die bonding costs are low and silicon area is not a concern, 6-die stacking with fine granularity is the best. This case study shows that medium granularity is worse than coarse granularity stacking with respect to TSV count, silicon area ratio, and wirelength. However, the trends vary depending on technologies, circuit size, and so on, as we can see in Figure 14.

**Table 5: Case study for early design exploration**

| $N_{DIE}$ | $B$ (Granularity) | # TSVs (mil) | $\frac{A_{3D}}{A_{2D}}$ | $\Delta$ WL (%) |
|---|---|---|---|---|
| 2 | coarse | 0.394 | **1.030** | -19.63 |
| | medium | 1.164 | 1.089 | -0.91 |
| | fine | 1.300 | 1.100 | **-21.52** |
| | gate-level | 1.453 | 1.112 | -14.13 |
| 4 | coarse | 0.456 | **1.035** | -23.37 |
| | medium | 2.362 | 1.181 | -12.81 |
| | fine | 2.626 | 1.202 | **-32.84** |
| | gate-level | 3.005 | 1.231 | -24.44 |
| 6 | coarse | 0.479 | **1.037** | -24.10 |
| | medium | 3.073 | 1.236 | -18.98 |
| | fine | 3.463 | 1.266 | **-35.60** |
| | gate-level | 4.017 | 1.308 | -26.29 |

# 8. CONCLUSION

In this paper, we first studied TSV-aware chip area model. As TSV size and TSV count have a significant impact on chip area, we derived a new wirelength distribution model considering TSV size and placement level. We included a few parameters which explain characteristics of 3D circuits during the derivation. The experimental results show that the wirelength increased by TSVs is not negligible so that we have to put the TSV count under our control during 3D placement. We also showed early design exploration which helps decision making for going from 2D to 3D ICs.

# 9. REFERENCES

[1] J. A. Davis, V. K. De, and J. D. Meindl, "A Stochastic Wire-Length Distribution for Gigascale Integration (GSI)-Part I: Derivation and Validation," in *IEEE Trans. on Electron Devices*, 1998.

[2] J. W. Joyner, P. Zarkesh-Ha, J. A. Davis, and J. D. Meindl, "A Three-Dimensional Stochastic Wire-Length Distribution for Variable Separation of Strata," in *Proc. IEEE Int. Interconnect Technology Conference*, 2000.

[3] R. Zhang, K. Roy, C.-K. Koh, and D. B. Janes, "Stochastic Wire-Length and Delay Distributions of 3-Dimensional Circuits," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, 2000.

[4] J. W. Joyner, "Opportunities and Limitations of Three-dimensional Integration for Interconnect Design," in *doctoral thesis*, 2003.

[5] D. C. Sekar, A. Naeemi, R. Sarvari, J. A. Davis, and J. D. Meindl, "IntSim: A CAD tool for Optimization of Multileve Interconnect Networks," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, 2007.

[6] J. U. Knickerbocker and et al, "Development of next-generation system-on-package (SOP) technology based on silicon carriers with fine-pitch chip interconnection," in *IBM J. Res. Dev. 49(4/5)*, 2005.

[7] ——, "Three-dimensional silicon integration," in *IBM J. Res. Dev. 52(6)*, 2008.

[8] J. Cong, G. Luo, J. Wei, and Y. Zhang, "Thermal-Aware 3D IC Placement Via Transformation," in *Proc. Asia and South Pacific Design Automation Conf.*, 2008.

[9] M. Y. vonne Lanzerotti, G. Fiorenza, and R. A. Rand, "Interpretation of Rent's Rule for Ultralarge-Scale Integrated Circuit Designs, With an Application to Wirelength Distribution Models," in *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 2004.

[10] B. S. Landman and R. L. Russo, "On a Pin Versus Block Relationship For Partitions of Logic Graphs," in *IEEE Trans. on Computers*, 1971.

[11] Synopsys, "Design Compiler," http://www.synopsys.com.

[12] Cadence, "Soc Encounter," http://www.cadence.com.

[13] International Workshop on Logic and Synthesis, "IWLS 2005 Benchmarks," http://www.iwls.org/iwls2005/benchmarks.html.