Coding Approach for Low-Power 3D Interconnects

Lennart Bamberg
ITEM.ids, University of Bremen
bamberg@item.uni-bremen.de

Robert Schmidt
ITEM.ids, University of Bremen
rschmidt@item.uni-bremen.de

Alberto Garcia-Ortiz
ITEM.ids, University of Bremen
agarcia@item.uni-bremen.de

Abstract

Through-silicon vias (TSVs) in 3D ICs show a significant power consumption, which can be reduced using coding techniques. This work presents an approach which reduces the TSV power consumption by a signal-aware bit assignment which includes inversions to exploit the MOS effect. The approach causes no overhead and results in a guaranteed reduction of the overall power consumption. An analysis of our technique shows a reduction in the TSV power consumption by up to 48 % for real correlated data streams (e.g. image sensor), and 11 % for low-power encoded random data streams.

1 Introduction

3D integration is a promising solution to overcome the challenges that arise with the limit of Moore's law. To connect the dies of a 3D system on chip (3D SoC), through-silicon via (TSV) arrays are typically used as they yield to a short delay and a high reliability [1]. Previous work shows that shifting from 2D to 3D integration, employing TSVs, allows for a significant reduction in the circuit footprint and delay, but often increases the power consumption [2].

The system power consumption is significantly affected by TSVs as they suffer from capacitive coupling which additionally impairs the signal integrity [3]. In TSV arrays, the coupling capacitances are large due to the relatively large TSV dimensions and the conductive substrate [4]. Additionally, the high number of aggressors in 3D further increases the coupling. Thus, coupling is a critical design concern for 3D integrated circuits (3D ICs) and consequently caught the attention of academia and industry (e.g. [4–15]).

Most previous works deal with coupling modeling [4–12] and coupling suppression using manufacturing techniques [9–12]. However, these techniques significantly increase the production cost and further impair the already critical TSV yield [1]. Additionally, most manufacturing techniques aim for signal integrity optimization, while leaving the overall power consumption unaffected [9–11]. Since coupling is a pattern dependent phenomena [16], data encoding approaches have recently been proposed which reduce the coupling peaks, without affecting the manufacturing [13–15]. These techniques again improve the signal integrity but also increase the TSV count, leading to an even increased overall TSV power consumption [3]. Thus, despite its importance, low-power techniques for TSVs have not yet been properly researched.

Low-power coding is very efficient for planar metal-wires [3]. Metal-wires only show a significant coupling with their two adjacent neighbors and the coupling capacitance between each adjacent

metal-wire pair has the same size. In contrast, TSVs have a maximum of eight adjacent neighbors and due to the different distances between direct and diagonal adjacent neighbors, combined with the E-field sharing effect [10], several capacitance values exist in a TSV array [5]. Hence, traditional low-power coding techniques are not directly applicable for TSV arrays. Thus, there is a need for new efficient low-power strategies for the TSVs in 3D integration.

In this work we present the first 3D low-power coding approach. Typical 3D SoCs take advantage of heterogeneous integration [17]: sensor, processor and memory elements are fabricated in individual dies, using the most efficient technology for each die. Afterwards the dies are stacked and connected by TSVs. In such systems, the patterns traversing the TSVs are often temporally correlated and/or normally distributed, resulting in bits with different switching properties [18]. We show that these properties can be exploited to effectively reduce the power consumption by an intelligent bit-to-TSV assignment, since the capacitances of TSV arrays are heterogeneous [5]. An additional fixed inversion of some bits before the transmission, realized by using inverting TSV drivers, can further decrease the power consumption, mainly due to the MOS effect. Our approach only affects the local bit-to-TSV assignments within the individual TSV arrangements, while the global net-to-TSV-bundle assignment remains routing optimal. Thus, the overhead costs are negligible.

A key contribution of this work is a formal method to find the optimal bit-to-TSV assignment (including inversions), that minimizes the TSV power consumption for any given data stream and TSV arrangement. To overcome the need for the exact data properties, systematic bit-to-TSV assignments, generally applicable for DSP signals, are contributed as well. A wide set of analyses, for real and synthetic data, shows that our approach can reduce the power consumption of modern TSVs by over 40 %, despite its negligible costs.

The remainder of this work is structured as follows: after some preliminaries, the method to determine the optimal assignment is derived in Sec. 3. Systematic assignments for DSP signals are presented in Sec. 4, which are compared to an optimal assignment for real data streams in Sec. 5. In Sec. 6, the combination of our approach with traditional low-power codes is briefly discussed. Experimental results are presented in Sec. 7. Finally, a conclusion is drawn.

2 Preliminaries: TSV model

The power consumption of TSVs can be precisely estimated by the power consumption related with its capacitances [6]. Thus, to calculate the TSV power consumption, the capacitance matrices for modern TSV arrays are required. In this work, the capacitance matrices are extracted by means of electromagnetic field simulations for 3D structures of TSV arrays, using Ansys Q3D extractor.

In the analyzed structures, the TSVs are regularly placed in a $M \times N$ array, where M and N are arbitrarily defined. The distance between the centers of two direct neighbored vias is constant and denoted by d. The cylindrical TSVs of length l and radius r are made up of copper. The TSVs traverse through the p-doped silicon substrate, which has a conductivity σ of 10 S/m. For DC insulation, each TSV is surrounded by a SiO_2 dielectric of thickness $r/5$. In the

model, the geometry parameters d and r are varied, in order to analyze different global TSV dimensions predicted by the ITRS for the year 2018. The length of the TSVs is defined by the substrate thickness equal to $50\,\mu m$. A TSV, its dielectric and the substrate form a metal oxide semiconductor (MOS) junction. Thus, a TSV is surrounded by a depletion region, which is modeled in the Q3D extractor as a depleted substrate area ($\sigma = 0$) [19]. The width of a depletion region surrounding TSV i is calculated by solving the exact Poisson's equation for an average TSV voltage of $pr_i{\cdot}V_{dd}$, where pr_i is the probability of a logical 1 on TSV i. $V_{dd} = 1\,V$ is the power supply voltage.

For the final validation of our work, circuit simulations are used. Therefore, full 3π-RLC circuits of the TSV arrays are also extracted.

3 Power-optimal bit-to-TSV assignment

In this section we derive our approach to reduce the TSV power consumption by choosing the optimal (fixed) bit-to-TSV assignment. The approach does not induce a circuit nor a TSV overhead. The only overhead cost is a slight increase in the local metal-wire lengths as the prior global net-to-array assignment remains routing optimal.

To quantify precisely the costs of our approach, we analyze a 3×3 TSV array, including the local routing, for a commercial 40 nm technology and TSVs with a radius of $2\,\mu m$ and a minimum pitch of $8\,\mu m$. In detail, we analyze all possible bit-to-TSV assignments, while considering the array symmetry. The worst-case routing only increases the path parasitics by a maximum of 0.4 %, versus a routing which aims for a local wire length minimization. The overall mean parasitic increase for all assignments is below 0.2 % with a standard deviation below 0.1 %. Thus, the effect of the local routing is marginal as TSV parasitics are dominant. Additionally, due to Keep-out-Zone restrictions, no active components are located nearby TSV arrays. Thus, we do not face a metal-layer-utilization problem. Summarized, the overhead costs for our approach are in fact negligible.

In the following, we derive a formula to calculate the power-optimal bit-to-TSV assignment. Thereby, we do not only consider the possibility of a mere reordering, but also the transmission of negated bits over some interconnects.

First of all, we review the model for the power consumption of (capacitive) interconnect structures, stated in Ref. [6]. Due to the thick oxides, leakage currents can be neglected for interconnects. The mean dynamic power consumption of an N bit interconnect for an initial assignment, mapping bit i (b_i) to interconnect i, is equal to:

$$P = \frac{V_{dd}{}^2 f}{2}\left(\sum_i^N \mathbf{E}\{\Delta b_i^2\}C_{i,i} + \sum_{i,j}^N \mathbf{E}\{\Delta b_i^2 - \Delta b_i \Delta b_j\}C_{i,j}\right). \quad (1)$$

Here, the first term $V_{dd}{}^2 f/2$ depends on the power supply voltage V_{dd} and the clock frequency f, which are not affected by a coding approach. Thus, in the following we use the mean power consumption normalized by this factor: $P_n = 2P/V_{dd}{}^2 f$. In Eq. 1, $C_{i,i}$ is the ground capacitance of interconnect i, and $C_{i,j}$ is the coupling capacitance between the two interconnects i and j. Furthermore, $\mathbf{E}\{\}$ is the expectation operator. Δb_i represents the switching of bit i, which is either 1 (0 to 1 transition), 0 (no transition), or −1 (1 to 0 transition). Thus, $\mathbf{E}\{\Delta b_i^2\}$ is the self switching probability of interconnect i. While the power consumption due to the ground capacitance of an interconnect i is determined only by its self switching, Δb_i, the power consumption associated with a coupling capacitance $C_{i,j}$, is additionally affected by a switching on interconnect j, Δb_j. Compared to the scenario where only interconnect i toggles ($\Delta b_j = 0$): the contribution of $C_{i,j}$ to the power consumption is doubled when

interconnect j toggles in the opposite direction ($\Delta b_i \Delta b_j = -1$) and vanishes if it toggles in the same direction ($\Delta b_i \Delta b_j = 1$).

The normalized power consumption P_n, can also be expressed using Frobenius inner product ($\langle\rangle$) of two matrices $\mathbf{T}$ and $\mathbf{C}$:

$$P_n = \langle \mathbf{T}, \mathbf{C} \rangle. \quad (2)$$

Here, $\mathbf{C}$ is the capacitance matrix, with capacitance $C_{i,j}$ on entry i,j. $\mathbf{T}$ presents the switching probabilities of the bits:

$$\mathbf{T} = \mathbf{T_s} \mathbf{1}_{N \times N} - \mathbf{T_c}, \quad (3)$$

where $\mathbf{T_s}$ is a matrix with the self switching probabilities $\mathbf{E}\{\Delta b_i^2\}$ on the diagonal entries, and zeros on the remaining entries. $\mathbf{T_c}$ represents the coupling probabilities with zeros on the diagonal entries and $\mathbf{E}\{\Delta b_i \Delta b_j\}$ on entry i,j. $\mathbf{1}_{N \times N}$ is a matrix of ones.

Since the capacitances of the $\mathbf{C}$ matrix are heterogeneous [5, 10], the assignment of the logical bits to the TSVs affects the power consumption. Moreover, a fixed inversion of some of the logical bits before the transmission may potentially decrease the $\mathbf{T}$ entries. If some bit pairs of the data stream are negatively correlated ($\mathbf{E}\{\Delta b_i \Delta b_j\} < 0$), initially in between them the likelihood of transitions in the opposite direction ($\Delta b_i \Delta b_j = -1$, causing the highest power consumption) is higher than the likelihood of aligned transitions ($\Delta b_i \Delta b_j = 1$, causing the lowest power consumption). In this scenario, the transmission of one of the two bits i or j negated over an interconnect (e.g. $\overline{b_i} \rightarrow$ interconnect x) results in a positive spatial correlation, since $\mathbf{E}\{\Delta \overline{b_i} \Delta b_j\} = -\mathbf{E}\{\Delta b_i \Delta b_j\} > 0$, and consequently to a reduced power consumption. Additionally, the bit assignment, including inversions, can also affect the TSV capacitances. Due to the MOS effect, an increased 1-bit probability on a TSV enlarges the width of its depletion region, resulting in up to 40 % lower capacitance values [6]. Therefore, for TSVs, transmitting data streams where the bit probabilities are not equally balanced ($\mathbf{E}\{\mathbf{b}\} \neq \mathbf{E}\{\overline{\mathbf{b}}\}$), the capacitance values depend on the assignment, including possible inversions. In the following we model these aspects.

First, we consider the switching matrix $\mathbf{T}$. The effect of a reassignment, including inversions, on $\mathbf{T}$ is mathematically expressed as:

$$\mathbf{T}' = \mathbf{T_s}'\mathbf{1}_{N \times N} - \mathbf{T_c}' = \mathbf{A}_\pi \mathbf{T_s} \mathbf{A}_\pi^{\mathbf{T}} \mathbf{1}_{N \times N} - \mathbf{A}_\pi \mathbf{T_c} \mathbf{A}_\pi^{\mathbf{T}}. \quad (4)$$

Here, $\mathbf{A}_\pi$ is a permutation matrix [20], which also performs inversions. A valid $\mathbf{A}_\pi$ has one 1 or one −1 in each column/row while all other matrix entries are 0. To assign the ith bit of the data stream to line j, $A_{\pi j,i}$ is set to 1. To assign the negated bit to the line, $A_{\pi j,i}$ is set to −1. Thus, for an exemplary 3 bit interconnect structure, to assign bit 3 negated to line 1, bit 1 to line 2 and bit 2 to line 3:

$$\mathbf{A}_\pi = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad (5)$$

Second, we derive a mathematical method to estimate the TSV capacitance matrix $\mathbf{C}$ depending on the bit-to-TSV assignment. The exact bit probability — capacitance relation is very complex, and consequently not suitable to determine the optimal assignment at high levels of abstraction. However, a linear regression to estimate the capacitance values as a function of the bit probabilities has a normalized root mean square error below 2 % [6]. Thus, the following equation can be used to estimate the size of a coupling capacitance for an assignment of the bits i and j to the TSVs i and j:

$$C_{i,j} = C_{0,i,j} + \Delta C_{i,j}(\mathbf{E}\{b_i\} + \mathbf{E}\{b_j\}), \quad (6)$$

where $C_{0,i,j}$ is the capacitance value for all 1-bit probabilities equal to zero and $\Delta C_{i,j}$ is the derivation of the capacitance value with increasing 1-bit probability $\mathbf{E}\{b_i\}$ or $\mathbf{E}\{b_j\}$. Since our requirement is

a formula where an inversion of the bits leads to one simple negation in the formula, we use a shifted form of Eq. 6:

$$C_{i,j} = C_{R,i,j} + \Delta C_{i,j}(\epsilon_i + \epsilon_j). \qquad (7)$$

Here $C_{R,i,j}$ is the capacitance value for all bit probabilities equal to ½ ($C_{R,i,j} = C_{0,i,j} + \Delta C_{i,j}$). ϵ_i is mathematically expressed as:

$$\epsilon_i = \mathbf{E}\{b_i\} - \frac{1}{2}. \qquad (8)$$

Since $\mathbf{E}\{\bar{b}_i\} = 1 - \mathbf{E}\{b_i\}$, an inversion of b_i, negates the ϵ_i value. Thus, the capacitance matrix as a function of $\mathbf{A}_\pi$ is expressed as:

$$\mathbf{C}' = \mathbf{C}_R + \Delta\mathbf{C} \circ (\mathbf{A}_\pi \epsilon \mathbf{1}_{1\times N} + \mathbf{1}_{N\times 1} \epsilon^T \mathbf{A}_\pi{}^T), \qquad (9)$$

where $\mathbf{C}_R$ and $\Delta\mathbf{C}$ are matrices containing the $C_{R,i,j}$ and $\Delta C_{i,j}$ values. ϵ is the vector of ϵ_i values. $\circ$ is the Hadamard operator.

Finally, we can determine the power-optimal bit assignment $\hat{\mathbf{A}}_\pi$:

$$\hat{\mathbf{A}}_\pi = \underset{\mathbf{A}_\pi \in S_N}{\arg\min} \left(\langle \mathbf{T}', \mathbf{C}' \rangle \right), \qquad (10)$$

where Eq. 4 and Eq. 9 are substituted for $\mathbf{T}'$ and $\mathbf{C}'$. S_N is the set of valid permutation matrices including all possible inversions.

In practice $\hat{\mathbf{A}}_\pi$ is determined with any of the several optimization tools available to reduce the computational complexity. Although overall up to several hundreds of TSVs exist in modern 3D ICs, the runtime of an optimization is negligibly low for our problem as it is executed for each TSV bundle individually whose size is relatively small. In this work, we exemplary use simulated annealing [21] to determine the optimal mapping.

4 Systematic TSV assignments for DSP signals

In some scenarios a sample data stream, required to obtain $\mathbf{T}$, may not be known at design time. In this case, the basic characteristics of the data can be used to obtain systematic assignments. As an example, in this section we focus on systematic assignments, applicable for DSP signals as they build an important data type in SoCs.

The bit-level characteristic of DSP signals are well understood [18], and only briefly summarized in the following. In many DSP signals, due to a zero mean normal distribution of the patterns, MSB pairs are strongly correlated ($\mathbf{E}\{\Delta b_i \Delta b_j\} \gg 0$). Additionally, a temporal pattern correlation affects the self switching ($\mathbf{E}\{\Delta b_i^2\}$) of the MSBs. The self switching probability is ½ for no pattern correlation and decreases with an increasing pattern correlation. Generally, the LSBs tend to be uncorrelated ($\mathbf{E}\{\Delta b_i \Delta b_j\} = 0$; $\mathbf{E}\{\Delta b_i^2\} = $ ½). Furthermore, all bit probabilities are equal to ½. Consequently, the capacitance matrix is assignment independent, resulting in:

$$P'_n = \langle \mathbf{T}'_s \mathbf{1}_{N\times N} - \mathbf{T}'_c, \mathbf{C_R} \rangle. \qquad (11)$$

Because of the positive bit correlations, we present systematic assignments without bit inversions. More precisely, we present two systematic approaches: one exploiting a temporal pattern correlation and one exploiting a mean-free normal distribution of the patterns.

First, we analyze temporally correlated, equally distributed patterns. An equal distribution causes no spatial bit correlation, which implies: $\mathbf{E}\{\Delta b_i \Delta b_j\} = 0$ for all $i \neq j$. Thus, for equally distributed signals, all elements of $\mathbf{T}'_c$ are zero. Therefore, Eq. 11 simplifies as:

$$P'_n = \langle \mathbf{T}'_s \mathbf{1}_{N\times N}, \mathbf{C_R} \rangle = \sum_i T'_{s,i,i} C_{T,i}, \qquad (12)$$

where $C_{T,i}$ is the sum of all capacitances connected to interconnect i. $T'_{s,i,i}$ is the ith diagonal entry of $\mathbf{T}'_s$, which is equal to the self switching probability of the bit transmitted over interconnect i.

Therefore, to minimize P', bits with the highest self switching probability $\mathbf{E}\{\Delta b_i^2\}$ have to be transmitted over TSVs with the lowest
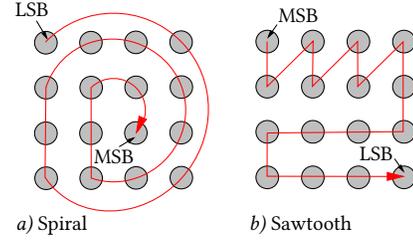
Figure 1. Systematic bit-to-TSV assignments: Spiral for correlated signals and Sawtooth (ST) for normally distributed signals.
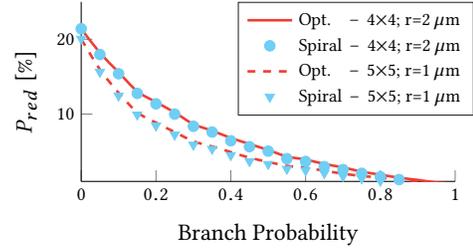
Figure 2. Decrease in power consumption (P_{red}) due to the optimal and the Spiral bit-to-TSV assignment for sequential data streams.

overall capacitance $C_{T,i}$ and vice versa. In TSV arrays, corner TSVs have the lowest overall capacitance, and edge TSVs have a lower overall capacitance than TSVs in the middle of an array [5]. Thus, the optimal assignment maps the bits with the highest self switching to the array corners. The bits with the next highest self switching are mapped to the array edges. The remaining bits are mapped to the array middle. Since the MSBs of correlated patterns show the lowest self switching, our systematic assignment for correlated patterns forms a spiral, as illustrated in Fig. 1.a.

We validate our Spiral mapping for synthetic sequential data streams with varying branch probability, as their patterns are equally distributed and temporally correlated. With the branch probability, the temporal pattern correlation varies. The simulated power consumption reductions, compared to a worst-case random assignment, due to the Spiral and the optimal assignment, are shown in Fig. 2 for two TSV arrays: a 4×4 array with $r = 2\,\mu m$; $d = 8\,\mu m$, and a 5×5 array with $r = 1\,\mu m$; $d = 4.5\,\mu m$. Fig. 2 reveals that the power consumptions for both assignments, optimal and Spiral, are almost equal. This proves the optimality of the systematic approach.

As a second scenario, we investigate a systematic assignment for mean-free normally distributed but temporally uncorrelated patterns. This implies that the self switching probability of each bit is ½. Thus, all diagonal elements of $\mathbf{T}'_s$ are ½ independent of the assignment, which results in a normalized power consumption of:

$$P'_n = \langle \frac{1}{2} \cdot \mathbf{1}_{N\times N} - \mathbf{T}'_c, \mathbf{C_R} \rangle = \frac{1}{2}\sum_i C_{T,i} - \sum_{i,j} T'_{c\,i,j} C_{i,j}, \qquad (13)$$

where $T'_{c\,i,j}$ is the correlation between the two bits transmitted over the interconnects i and j. Therefore, in order to reduce the power consumption, highly correlated bit pairs (large $\mathbf{E}\{\Delta b_i \Delta b_j\}$) have to be assigned to TSV pairs connected by a large coupling capacitance. In TSV arrays, the biggest coupling capacitances are located between corner TSVs and their two direct adjacent edge TSVs, due to the reduced E-field sharing effect [5]. MSBs of normally distributed signals have the highest cross-correlation. Thus, our second systematic assignment has to map the MSB onto a corner and the next lower
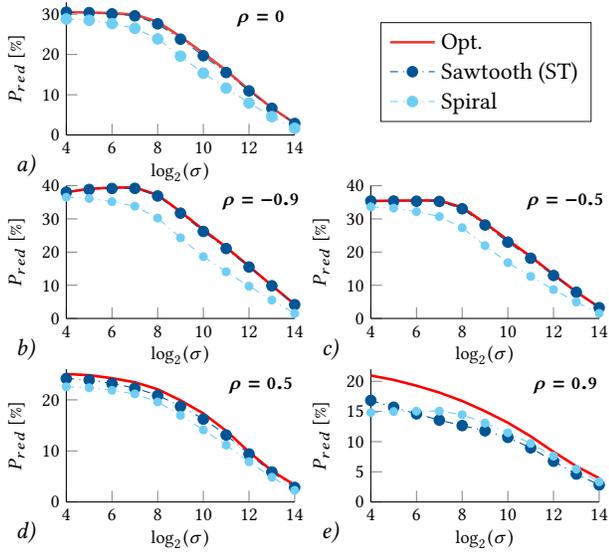
Figure 3. Power consumption reduction (P_{red}) due to our mapping approaches for uncorrelated (3.a) and correlated (3.b-3.e; $\rho \neq 0$), Gaussian distributed data streams with standard deviation σ.

Figure 4. Power consumption reduction (P_{red}) for an optimal/Spiral assignment and image sensor patterns. "+xS" indicates x stable lines.

significant bit onto one of its direct adjacent edge TSVs. The following bits, recursively, have to be mapped by finding the TSV in the array that has the biggest accumulated coupling capacitance with all previously assigned TSVs. Finally, our systematic assignment results in the MSB to LSB mapping illustrated in Fig. 1.b. Over the first two rows the bits, from the MSB downwards, are mapped in a sawtooth manner. From the third row on, a simple row-by-row mapping is used. Fig. 3.a shows the reduction in the power consumption due to the optimal and the Sawtooth (ST) assignment for the transmission of Gaussian distributed 16 b pattern sets, over a 4×4 TSV array ($r = 2\,\mu m$; $d = 8\,\mu m$). The results are plotted over the standard deviation of the patterns, to analyze different normal distributions. The figure underlines the optimal nature of the Sawtooth assignment for normally distributed, temporally uncorrelated patterns.

However, in some real applications, temporally correlated and normally distributed signals occur. For these data streams, the optimal TSV assignment is not trivial and dependent on the correlation quantities. As shown in Fig. 3.b-3.e, for negatively correlated, Gaussian distributed patterns the Sawtooth mapping leads to the lowest power consumption (reduction up to 40 %), while for a positive temporal correlation neither Sawtooth nor Spiral mapping lead to the optimal power consumption. However, compared to a random assignment both approaches still lead to a significant improvement.

Summarized, if it is not possible to determine the optimal assignment by means of Eq. 10, which guarantees the lowest possible power consumption, the proposed Sawtooth mapping should be applied for normally distributed signals and the Spiral mapping for primarily temporally correlated signals.

5 Comparison of systematic and optimal mapping for real DSP signals

Until this point, we investigated the efficiency of our proposed technique for synthetic DSP signals. In the following, we analyze and compare the efficiencies of the systematic and the optimal bit-to-TSV assignments for real DSP signals. Thereby, we focus on an important class of systems: heterogeneous 3D SoCs. Two
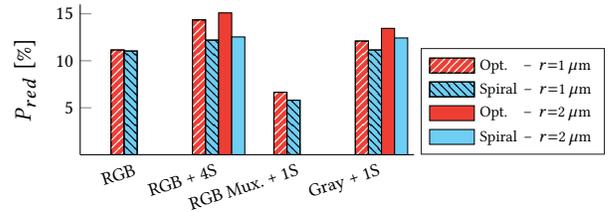
commercially relevant examples are 3D VSoCs [17], including dies for image sensing and dies for digital image processing, and SoCs with an dedicated MEMS sensor die, bonded to a digital die [22].

5.1 Vision system on chip

In contrast to pure CMOS image sensors, VSoCs are used to capture and process the images in a single chip. This overcomes the limitations of traditional systems due to expensive image transmissions between sensor and processor, especially for high frame rates [17].

In a 3D VSoC some dies are dedicated for image sensing and digitalization and some for image processing. In this subsection, we investigate our approach for the transmission of digitalized image pixels from a sensing layer to a processing layer.

The first three analyses are performed for data stemming from a 0–255 RGB image sensor using a Bayer pattern filter [23]. First, we analyze the parallel transmission of all four RGB colors (1 red, 2 green, 1 blue) of each Bayer pattern pixel over one 32 b (4×8) TSV array. For the second analysis, we assume four additional TSVs in the array (resulting in a 6×6 array): one TSV carrying an enable signal, one redundant TSV for yield enhancement and two power/ground TSVs to supply the sensor. In the third analysis the four colors of each pixel are transmitted one after another (RGB Mux.) over a 3×3 array including an additional enable signal. The fourth analysis, is performed for a data stream stemming from a 0–255 grayscale image sensor. Here, the transmission of one pixel per clock cycle over a 3×3 array including an enable signal is investigated. All analyzed data streams are composed of pictures of cars, people and landscapes.

For all analyses, the reduction in the power consumption, against random assignments, is investigated for the optimal and the Spiral assignment, since the strong correlation of adjacent pixels generally results in a temporal pattern correlation. Redundant, enable and power/ground signals are considered as (almost) stable. Redundant and enable signals are assumed as set to logical 0 when unused, which may be exploited by inversions. V_{dd} and GND lines are always on logical 1 and logical 0, respectively, but an inversion for power lines is not possible and consequently forbidden for the assignment. For the simultaneous transmission of a complete RGB pixel, the bits of the four color components are interleaved one-by-one for the Spiral mapping. Since stable lines are perfectly correlated, they are added as MSBs for the Spiral mapping.

For the global TSV dimensions, we choose the minimum ones predicted for the year 2018 ($r = 1\,\mu m$; $d = 4\,\mu m$). To show the effect of varying TSV geometries, the power consumption for the 3×3 and the 6×6 array is also investigated for $r = 2\,\mu m$ and $d = 8\,\mu m$.

The simulated power reductions due to the various assignments are reported in Fig. 4. The results show that the Spiral mapping is almost optimal for the transmission of image sensor patterns without stable lines and always leads to a power reduction of 11–13 %, except for the multiplexed colors, where the reduction is only
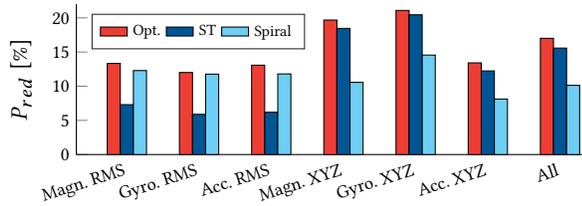
Figure 5. Decrease in the power consumption (P_{red}) for our optimal/systematic approach and MEMS sensor signals.

5 %. Here, due to the multiplexing, the pixel correlation is lost why only the reassignment of the stable line results in a power reduction.

With stable lines in the TSV array, the power reduction due to an optimal assignment is up to 2.5 percentage point higher, as it considers inversions and the coupling properties of stable lines. Therefore, with stable lines, the potential to reduce the power consumption, using our approach, is also higher.

Summarized, the simulations show that both, optimal and systematic, bit-to-TSV assignments can effectively reduce the TSV power consumption in 3D VSoCs. However, in the presence of additional stable lines, the optimal approach has a noticeably higher gain.

5.2 MEMS sensors in a 3D system on chip

In this subsection we analyze the efficiency of our technique for MEMS sensor data, transmitted from a sensing to a processing layer. Therefore, sensor signals from a modern smartphone in various daily use scenarios are used. Analyzed is a magnetometer, an accelerometer and a gyroscope, all sensing on three axes (x, y and z) with a resolution of 16 b. We assume a transmission of one sample per time step over a 4×4 array with $r = 2\,\mu m$ and $d = 8\,\mu m$. We analyze the transmission of the single data streams for two scenarios. In the first one, only the root mean square (RMS) values resulting from the three axis values are transmitted. In the second one, the x-, y- and z-axis values are regularly interleaved/multiplexed (XYZ). For completeness, we also analyze the transmission of all three data streams over one TSV array. Thereby, a regular pattern-by-pattern multiplexing of the three XYZ-interleaved data streams is assumed. Here, we investigate both systematic bit-to-TSV assignments since normally distributed and temporally correlated data streams occur.

The simulated mean power consumption reductions against random assignments are shown in Fig. 5. The figure reveals that, for the interleaved data streams, the proposed Sawtooth mapping is only slightly worse than the proposed optimal assignment which reduces the power consumption by up to 21.1 %. Generally, the single axis values are normally distributed and temporally correlated. However, for interleaved data streams temporal correlation is lost. Thus, these scenarios build examples for temporally uncorrelated, normally distributed signals, since an interleaving does not affect the pattern distribution. The small gain for the optimal bit-to-TSV assignment over the Spiral mapping is caused by the fact that not all sensor signals are perfectly mean-free.

In contrast, for the RMS data streams, the Spiral mapping significantly outperforms the Sawtooth mapping because here the patterns are unsigned (no zero mean) and spatially correlated. However, for the RMS data streams, the maximum possible power reduction due to a reassignment is 13.3 %, which is significantly lower than the maximum power reduction for the interleaved data streams.

In conclusion, for real data streams, the exploitation of a mean-free normal distribution is more efficient than the exploitation of a

temporal pattern correlation. Furthermore, due to non idealities in real signals, the optimal approach has a slightly higher gain than a systematic one. But generally, both assignments, systematic and optimal, lead to a significantly reduced TSV power consumption.

6 Combination with data encoding

Modifying the data properties using encoding techniques is a well established low-power approach [3]. Our proposed technique enables the use of existing low-power coding techniques, designed to reduce the power consumption of metal-wires and/or gates, for 3D integration in the most efficient way by finding the optimal bit-to-TSV assignment. Thus, if an encoding is already applied for other components, no additional overhead is required for the TSV coding. Unencoded, most data streams generally have a balanced number of 0- and 1-bits. However, data encoding techniques often lead to a large fraction of 0-bits [3], which affects the power consumption in 3D negatively. Here, our optimal mapping further boosts the efficiency of the coding approach by transmitting inverted bits. Generally, an inversion is realized by using inverting buffers instead of non-inverting ones (or vice versa) on both sides of the TSV. However, inversions can be also hidden in the coder/decoder.

For example, Gray coding is a popular approach to reduce the power consumption of gates and metal-wires. The nth binary-to-Gray encoders output is equal to the nth input XORed with the $n+1^{th}$ input ($Y[n] = X[n] \oplus X[n + 1]$). Consequently, due to the spatial MSB correlation in normally distributed signals, Gray coding results in bits nearly stable on logical 0 for this kind of data. This reduces the switching activities but also decreases the 1-bit probabilities. Here the required inversions for the optimal bit-to-TSV assignment can be realized inside the Gray encoder and decoder: XOR operations are swapped with XNOR operations to obtain the negated code words which increases, instead of decreases, the 1-bit probabilities, while leaving the switching activities unaffected. Since XOR and XNOR operations have the same costs, this optimization of the data encoding technique is overhead free.

7 Experimental results

In this section, our approach is investigated for real signals and traditional coding approaches by means of Spectre circuit simulations in combination with the results from the Q3D extractor. Here, TSV arrays with $r = 1\,\mu m$ and $d = 4\,\mu m$, including the connection to the metalization are analyzed. For the circuit simulations, 22 nm Predictive Technology Model drivers of strength six are employed. The clock frequency is set to 3 GHz. Considered is the overall power consumption, including leakage and the drivers. To report values independent of the TSV count in the array, and redundant bits in the data stream, the power consumptions reported in Fig. 6 are scaled to values for an effective transmission of 32 b per clock cycle.

The power consumption is investigated for the transmissions of four different data streams, if our optimal approach is applied, and if not. The first data stream contains the MEMS sensor signals from Subsec. 5.2, where, for 3,900 cycles, patterns of a single axis of one sensor are transmitted. Subsequently, patterns stemming from the next sensor are transmitted for 3,900 cycles and so on, until data for all axes and sensors has been transmitted. We refer to this data stream as "Sensor Seq.". For the second data stream "Sensor Mux.", the patterns belonging to the individual axes are interleaved one-by-one. For the first two data streams, a pattern width of 16 b and a 4×4 array is chosen. The results show that multiplexed sensor data leads to a significantly higher power consumption, since the pattern
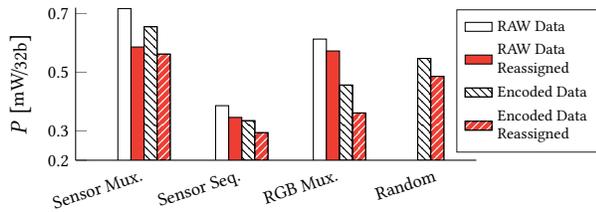
Figure 6. TSV power consumption (including drivers and leakage) if the proposed optimal bit-to-TSV assignment is used and if not.

correlation is lost. Nevertheless, because of limited buffer capabilities in a sensing layer, the more common scenario is the multiplexed transmission. However, temporal correlation can be retained using Gray encoding, which can be realized in the sensors analog-to-digital (AD) converters to avoid a noticeable implementation overhead. Thus, Gray encoding is twice analyzed for the sensor data: once in the traditional way and once in combination with our proposed optimal bit-to-TSV assignment. For the multiplexed binary sensor data, our approach, without Gray encoding, leads to a decrease in the power consumption by 18.3 %. In contrast, the plain Gray encoding even leads to a higher power consumption (reduction 8.6 %). A combination of the Gray encoder with our approach, more than doubles the coding efficiency (power reduction 21.7 %).

The third data stream contains multiplexed RGB Bayer filter colors for different pictures. The 8 b data stream is transmitted together with a redundant line for yield enhancement over a 3×3 array. As known from Subsec. 5.1, the temporal pattern correlation, caused by the pixel correlation, is lost if the RGB colors are multiplexed. Additionally, no normal distribution is present in the patterns. This leads to a dramatic increase in the interconnect power consumption and to no gain for a Gray encoding. To retain spatial correlation for the multiplexed data, a correlator is used [3], which again can be hidden in the AD converters. For a new R, G or B value, the value is first bitwise XORed with the previous value of the same color and subsequently transmitted. Since consecutive R, G or B values are highly correlated, this again leads to MSBs nearly stable on zero. Thus, the correlator produces a spatial and temporal bit correlation, reducing the power consumption and increasing the potential gain of a bit-to-TSV assignment. Thus, the correlated RGB data transmission over a 3×3 array, including one redundant TSV, is also analyzed. Again, our approach swaps XOR with XNOR operations for the correlator and transmits the inverted redundant line to maximize the number of 1-bits. While our approach only leads to a decrease in the power consumption by 6.8 % for the unencoded data, our approach combined with the correlator leads to a dramatic decrease in the power consumption from 0.61 mW to 0.36 mW (−41 %). In contrast, the plain correlator only reduces the power consumption by 25.2 %.

To show the general usability of our approach for all kinds of data, the last analyzed data stream is a random 7 b data stream, encoded to a 8 b data stream using the coupling invert encoding from Ref. [24], transmitted together with a flag with a set probability of 0.01 % over a 3×3 array. The coupling encoding is derived for the physical structure of metal-wires, and thus intrinsically not suitable for TSVs. Here, we assume a 3D network on chip, where the data is mainly transmitted over 2D links and a dedicated encoding for each 3D link is too cost intensive. However, the coding approach leads to a spatial correlation in between some bits and to a temporal bit correlation. This and the set probability of the flag is exploited by our low-power approach. Thus, also for the random 2D coded data,

our approach reduces the TSV power consumption by 11.2 %. This shows the efficiency of our technique for a wide set of applications.

Observe, that in this section the TSV dimensions are equal to the minimum ones predicted by the ITRS for the year 2018. For thicker TSVs and/or wider TSV pitches, which is the common case today, our approach causes an even higher reduction in the TSV power consumption (e.g. up to 48 % for $r = 2\,\mu$m and $d = 8\,\mu$m).

8 Conclusion

This work presents an approach to reduce the TSV power consumption by an intelligent, physical-effect-aware, local bit-to-TSV assignment, which exploits the stochastic bit properties of the transmitted data. Analyses for a large set of real and synthetic data streams underline the importance and efficiency of our low-power approach which is able to reduce the power consumption of modern TSVs by over 40 %, without inducing noticeable overhead costs.

Acknowledgments

References

[1] V. F. Pavlidis *et al.*, *Three-Dimensional Integrated Circuit Design.* Elsevier Science & Technology Books, 2017.

[2] D. H. Kim and S. K. Lim, *Impact of TSV and device scaling on the quality of 3D ICs.* Springer, 2015, pp. 1–22.

[3] A. Garcia *et al.*, "Low-power coding: trends and new challenges," *Journal of Low Power Electron.*, 13(3):356–370, 2017.

[4] C. Xu *et al.*, "Compact AC modeling and performance analysis of through-silicon vias in 3-D ICs," *IEEE Trans. Electron Devices*, 57(12):3405–3417, 2010.

[5] L. Bamberg *et al.*, "Edge effects on the TSV array capacitances and their performance influence," *Integration, the VLSI Journal*, 61:1–10, 2018.

[6] L. Bamberg and A. Garcia, "High-level energy estimation for submicrometric TSV arrays," *IEEE Trans. Very Large Scale Integ. (VLSI) Systems*, 25(10):2856–2866, 2017.

[7] S. Piersanti *et al.*, "Algorithm for extracting parameters of the coupling capacitance hysteresis cycle for TSV transient modeling and robustness analysis," *IEEE Trans. Electromagn. Compat.*, 59(4):1329–1338, 2017.

[8] A. E. Engin and S. R. Narasimhan, "Modeling of crosstalk in through silicon vias," *IEEE Trans. Electromagn. Compat.*, 55(1):149–158, 2013.

[9] C. Qu *et al.*, "Modeling and optimization of multiground TSVs for signals shield in 3-D ICs," *IEEE Trans. Electromagn. Compat.*, 59(2):461–467, 2017.

[10] Y. Peng *et al.*, "Silicon effect-aware full-chip extraction and mitigation of TSV-to-TSV coupling," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, 33(12):1900–1913, 2014.

[11] J. Cho *et al.*, "Modeling and analysis of through-silicon via (TSV) noise coupling and suppression using a guard ring," *IEEE Trans. Compon. Packag. Manuf. Technol.*, 1(2):220–233, 2011.

[12] Z. Xu and J. Q. Lu, "Three-dimensional coaxial through-silicon-via (TSV) design," *IEEE Electron Device Lett.*, 33(10):1441–1443, 2012.

[13] R. Kumar and S. P. Khatri, "Crosstalk avoidance codes for 3D VLSI," in *Design, Automation Test in Europe Conf. Exhibition (DATE)*, 2013, pp. 1673–1678.

[14] Q. Zou *et al.*, "3DLAT: TSV-based 3D ICs crosstalk minimization utilizing less adjacent transition code," in *2014 19th Asia and South Pacific Design Automation Conf. (ASP-DAC)*, 2014, pp. 762–767.

[15] X. Cui *et al.*, "An enhancement of crosstalk avoidance code based on fibonacci numeral system for through silicon vias," *IEEE Trans. Very Large Scale Integ. (VLSI) Systems*, 25(5):1601–1610, 2017.

[16] C. Duan *et al.*, *On and off-chip crosstalk avoidance in VLSI design.* Springer, 2010.

[17] D. Lie *et al.*, "Impact of heterogeneous technology integration on the power, performance, and quality of a 3D image sensor," *IEEE Trans. Multi-Scale Comput. Syst.*, 2(1):61–67, 2016.

[18] P. E. Landman and J. M. Rabaey, "Architectural power analysis: The dual bit type method," *IEEE Trans. Very Large Scale Integ. (VLSI) Systems*, 3(2):173–187, 1995.

[19] T. Bandyopadhyay *et al.*, "Rigorous electrical modeling of through silicon vias (TSVs) with MOS capacitance effects," *IEEE Trans. Compon. Packag. Manuf. Technol.*, 1(6):893–903, 2011.

[20] R. A. Brualdi, *Combinatorial matrix classes.* Cambridge University Press, 2006.

[21] V. Granville *et al.*, "Simulated annealing: a proof of convergence," *IEEE Trans. Pattern Analysis and Machine Intell.*, 16(6):652–656, 1994.

[22] F. Niklaus and A. C. Fischer, "Heterogeneous 3D integration of MOEMS and ICs," in *2016 Int. Conf. on Optical MEMS and Nanophotonics (OMN)*, 2016, pp. 1–2.

[23] S.-Y. Lee and A. Ortega, "A novel approach of image compression in digital cameras with a Bayer color filter array," in *Int. Conf. on Image Processing*, 2001, pp. 482–485.

[24] M. Palesi *et al.*, "Data encoding schemes in networks on chip," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, 30(5):774–786, 2011.