Received 31 May 2024; revised 17 September 2024 and 30 September 2024; accepted 11 October 2024.
Date of publication 17 October 2024; date of current version 7 November 2024.

Digital Object Identifier 10.1109/JXCDC.2024.3482464

Scaling Logic Area With Multitier Standard Cells

FLORIAN FREYE[1], CHRISTIAN LANIUS[1], HOSSEIN HASHEMI SHADMEHRI[1], DIANA GÖHRINGER[2] (Member, IEEE), and TOBIAS GEMMEKE[1] (Senior Member, IEEE)

[1]Chair of Integrated Digital Systems and Circuit Design, 52074 Aachen, Germany
[2]Chair of Adaptive Dynamic Systems, 01069 Dresden, Germany

CORRESPONDING AUTHOR: F. FREYE (freye@ids.rwth-aachen.de)

This work was supported in part by the Federal Ministry of Education and Research (BMBF) under Grant 16ME0399 (NEUROTEC II) and Grant 03ZU1106CA (Clusters4Future—NeuroSys).

ABSTRACT While the footprint of digital complementary metal–oxide–semiconductor (CMOS) circuits has continued to decrease over the years, physical limitations for further intralayer geometric scaling become apparent. To further increase the logic density, the international roadmap for devices and systems (IRDS) predicts a transition from a single layer of transistors per die to monolithically stacking transistors in multiple tiers starting in 2031. This raises the question of the extent to which these can be exploited in 3-D standard cells to improve logic density. In this work, we investigate the scaling potential of realizing standard cells employing two or three dedicated tiers. For this, specific multitier virtual physical design kits are derived based on the open ASAP7. A typical RISC-V implementation realized in a classic standard cell library is used to identify the subset of the most relevant standard cells. In accordance with the virtual physical design kit (PDK), 3-D derivatives of the single-tier standard cells are crafted and evaluated with respect to achievable logic density considering standard synthesis benchmarks and blocks on the architecture level.

INDEX TERMS Advanced scaling, monolithic 3-D integration, multitier circuits, standard cell library.

I. INTRODUCTION

AFTER decades of continuous technology scaling following Dennard's law, 2-D geometric scaling starts to face limitations with each technology generation taking more and more development time. To tackle this challenge, researchers in academia and industry have started to consider 3-D integration to further increase the density of functionality integrated into a chip. In recent years, die stacking has been the predominantly employed method for 3-D integration. Besides integrating more transistors per area, this has the potential advantage of improved manufacturing yield and offers another degree of freedom, as different system components can be manufactured in different technology nodes. However, while heterogeneous integration improves the von Neumann bottleneck of communication between computation units and memory as is done for example in modern flash memories, it is still limited by comparatively low through-silicon-via (TSV) density.

Monolithic integration is considered a better solution providing lower cost and high integration density. Monolithic integration of multiple active tiers can be done either in the front end of line (FEOL) or the back end of line (BEOL). Any of which has to respect the reduced thermal budget, which prevents a simple adoption of multiple layers of classic complementary metal–oxide–semiconductor (CMOS) devices. Novel devices such as carbon nano-tube field-effect transistors (CNTFETs) and memristive switching devices are processed using far lower temperatures and are, therefore, exceptionally well suited for such integration. Successful monolithic integration of ReRAM, CNTFETs, and classical CMOS was shown in [1], [2], and [3]. However, a high-yielding technology for mass production of logic chips is still out. In the meantime, the International Roadmap for Devices and Systems (IRDS) projects lateral stacking of gate all-around (GAA) CMOS in so-called "tier" in its most recent roadmap in the category "More Moore." Here, beyond CMOS devices are considered mostly for application-specific use, such as neuromorphic computing [4].

In the following, we consider this type of tiered stacking as FEOL 3-D integration. For FEOL integration, a division into two subcategories is proposed in [5]: gate- and transistor-level 3-D integrations. Both options are illustrated in Fig. 1. In gate-level 3-D integration, each individual standard cell is restricted to a single tier, thereby avoiding a redesign. Several

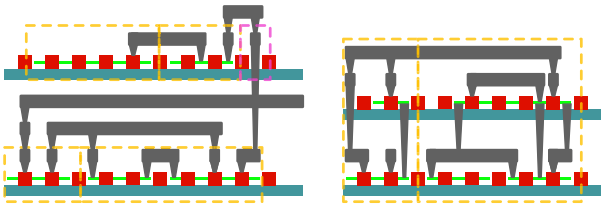

FIGURE 1. Comparison of gate-level integration (left) and transistor-level integration (right). The boundaries of a standard cell are highlighted by the orange boxes. The pink box represents the MIV cell necessary for gate-level integration.

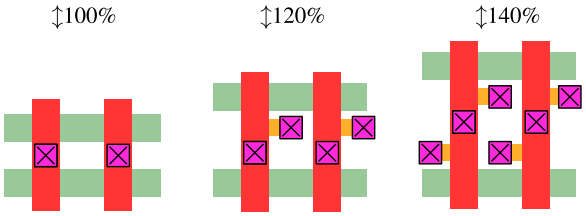

FIGURE 2. Tradeoff between required cell height and gate contacts per CPP.

works have focused on this gate-level stacking for evaluating the advantages of monolithic 3-D integration [6], [7], [8]. However, such gate-level 3-D integration requires considerable wiring layers between the CMOS tiers. Besides the need for novel place and route (P&R) algorithms, it increases the cost and complexity in manufacturing. Furthermore, connecting the tiers requires significant area overhead, as special through-vias cells used for routing have to be inserted. As reported in [7], 18% of all nets travel between tiers for a dual-tier design. Other works such as [6] and [9] ignore this aspect, thereby providing overly optimistic reductions in the achievable footprint.

Several studies have looked into transistor-level 3-D integration with vertical field-effect transistors (VFETs) [10], [11]. However, the actual development of VFETs has been impaired in the past by problems with asymmetric switching behavior. The remaining challenges made laterally stacked transistors appear advantageous [12], which is the focus of this work: transistor-level 3-D integration assuming laterally stacked devices. The assumptions that we make for this virtual multitier technology will be outlined in Section II. In Section III, we will determine a reduced set of standard cells to use and introduce novel techniques to stack them to a second and third tier. Section IV will present the impact of the optimized cells on the area utilization of representative digital designs. We will make use of several standard synthesis benchmarks and typical architectural blocks to evaluate the standard cells of the different virtual technologies. Finally, Section V summarizes our findings.

II. VIRTUAL MULTITIER TECHNOLOGY

Fundamentally, we assume laterally stacked GAA devices in our virtual multitier technology. Thereby, the transistor channel is represented by a single nanosheet layer. Furthermore, the device realization follows the concept presented in [12]. Each conduction type of transistor is arranged along a specific quasi-continuous diffusion. Stacking of different conduction types within one tier as realized in complemenary field-effect transistor (CFET) devices is not considered here. In any case, it appears challenging to go beyond a two-tier design for a CFET technology [12].

In terms of intracell connectivity, a metal layer is required to enable the parallel connection of two transistors within a single tier. This metal layer gets connected to the drain/source regions by means of V0. This amount of intracell connectivity with a single metal layer is considered necessary to realize 3-D standard cells of reasonable compactness. More layers provide only marginal benefits while further increasing manufacturing complexity and cost. Furthermore, tiers get connected by monolithic inter-tier vias (MIVs), which reach directly from the drain/source of a higher tier to the metal layer of the tier below.

Backside power delivery is expected to bring improvements to area, power, and performance [13] and is, thus, assumed here. In [14], Intel demonstrated its PowerVia technology through which the buried power rails are directly contacted by a backside power delivery network. This leaves the front side entirely to the signal network, hence avoiding the contradicting optimization goals considering the requirements of signal routing and power delivery. Hence, power is directly available to the bottom tiers. Upper tiers, however, have to use MIVs to connect to the supply rails. In addition, diffusion breaks are assumed to be single diffusion breaks.

An important bottleneck for 3-D integration is pin density and intracell connectivity across tiers. As gates are stacked on top of each other, simply providing access to lower tier gates is challenging-worse still when considering the necessary within-tier routing. A further challenge imposed by the MIVs is the passing of a signal through a layer. If a MIV passes an active region, it will unavoidably connect to it. Therefore, a direct pass-through of a signal from a bottom tier to a metal on an upper tier is not possible without also connecting the active area of the upper tier, which lies in between. The same is true if a tier is in the middle of the two tiers that have to be connected. Besides the introduction of a diffusion break to allow for such pass-through, a gate contact MIV can be used if not all gates in the cell are connected.

To accommodate for the accessibility of the lower tier, we assume an increase in cell height by one track pitch per added tier. For high-density cells, the IRDS Roadmap projects a cell height of five track pitches, following current height reduction trends [4]. For two- and three-tier designs, we, therefore, assume 6T and 7T cell heights, providing pass-through resources to lower tiers and intratier routing. The resulting change in cell height of 120% for the two-tier variant and 140% for the three-tier variant is depicted in Fig. 2.

A 3-D sketch of an AOI221 gate drawn using these assumptions is shown in Fig. 3(a). To represent the design in a planar fashion, a multitier stick diagram is introduced. The multitier stick diagram consists of as many 2-D stick diagrams as there

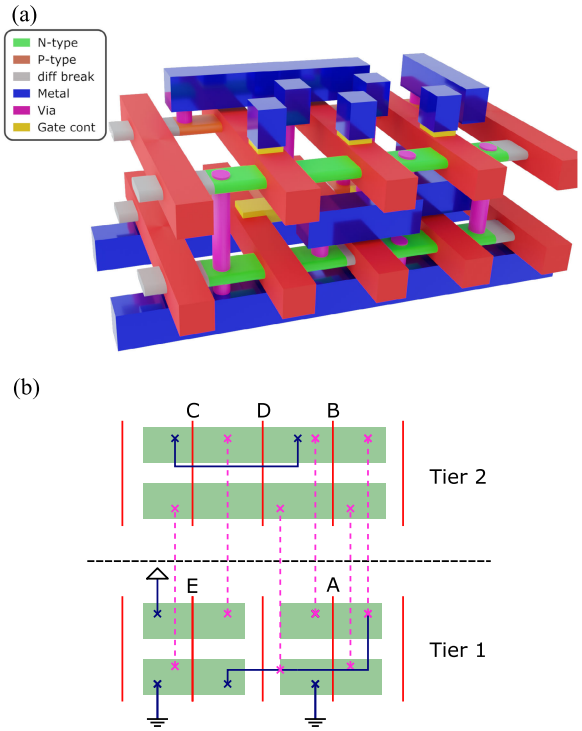

FIGURE 3. **AOI221 gate using given technology assumptions. (a) Three-dimensional visualization. (b) Two-tier stick diagram.**

are tiers. The individual tiers are connected over pink lines representing MIVs. An example for the same AOI221 gate is provided in Fig. 3(b).

III. APPROACH

A. CELL SELECTION

In order to determine a set of standard cells, which is representative of a typical digital design, the synthesized netlist of an RISC-V processor is considered. Fig. 4 shows the top ten cells sorted by contribution in terms of area and count, respectively. Due to the expected high overlap between both sets, their unions contain a total of 11 unique cells.

Because of the duality of the AOI and OAI gates in terms of transistor network, we extend the selection by their respective duals. In addition, the synthesis requires the addition of an integrated clock gating (ICG) cell and buffers to run properly. Finally, the set consists of the following 16 cells: DFF, INV, NAND2, NOR2, OAI21, OAI22, OAI221, OAI222, XNOR2, XOR2, AOI21, AOI22, AOI221, AOI222, ICG, and BUF. The inverter and buffer are both implemented in multiple drive capabilities. Fig. 5 shows a breakdown of the selected cells in the processor design, with the excluded cells lumped in "other." We see that we cover more than 85% of all used cell instances and about 88% of the cell area with the given cell selection.

B. STANDARD CELL AREA ESTIMATION

This set of standard cells is mapped to two- and three-tier technologies following the assumptions, as detailed in Section II. While 2-D standard cells have well-established

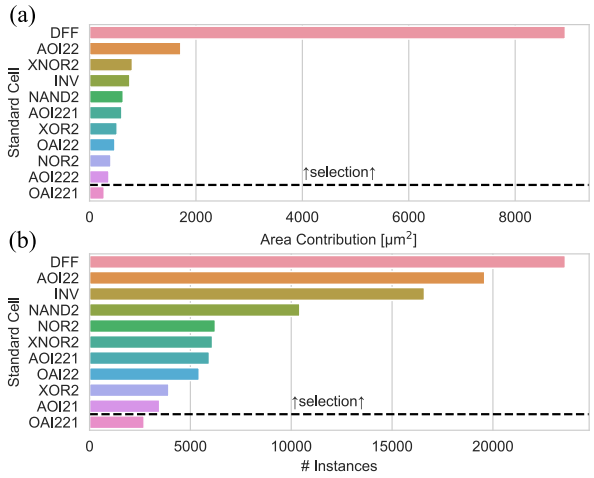

FIGURE 4. **(a) Top ten cells in terms of area use and (b) number of instances for an RISC-V processor. The processor register file explains the high number of D-flipflops.**

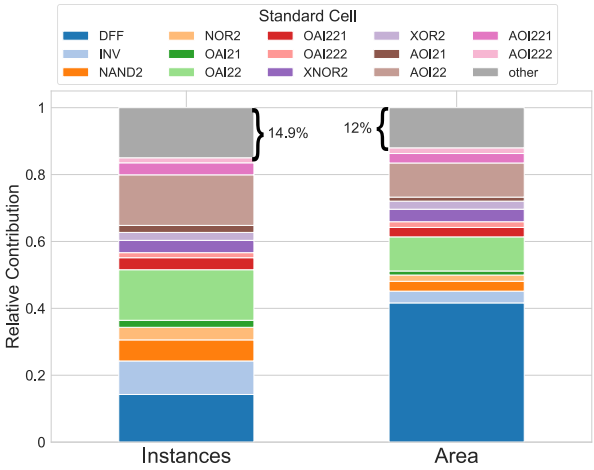

FIGURE 5. **Amount of neglected cells is below 15% for both number of instances and area.**

methodologies that give a systematic approach to finding an optimal solution in terms of area efficiency, similar methodologies do not yet exist for 3-D variants. In order to find acceptable solutions, the classic 2-D approach is followed as a starting point, i.e., Eulerian paths in the cell's stick diagram are identified. While a transition from one path to another results in a diffusion break for the 2-D variant, in the 3-D variant, different paths can be moved to a different tier avoiding the necessity for the diffusion break. In the second step, the Eulerian paths, therefore, get distributed evenly among the available tiers. Other heuristics for determining the layout include placing paths with the most VDD/VSS connections on the bottom layer to facilitate easy access to backside power and flipping paths to allow for maximum signal/power alignment. If there is a significant difference in length between two paths, or if a cell contains only a single Eulerian path, utilizing only a single tier per path would not fully leverage the potential for area savings. In such a case, the Eulerian path can be "wrapped around" into a 3-D shape.

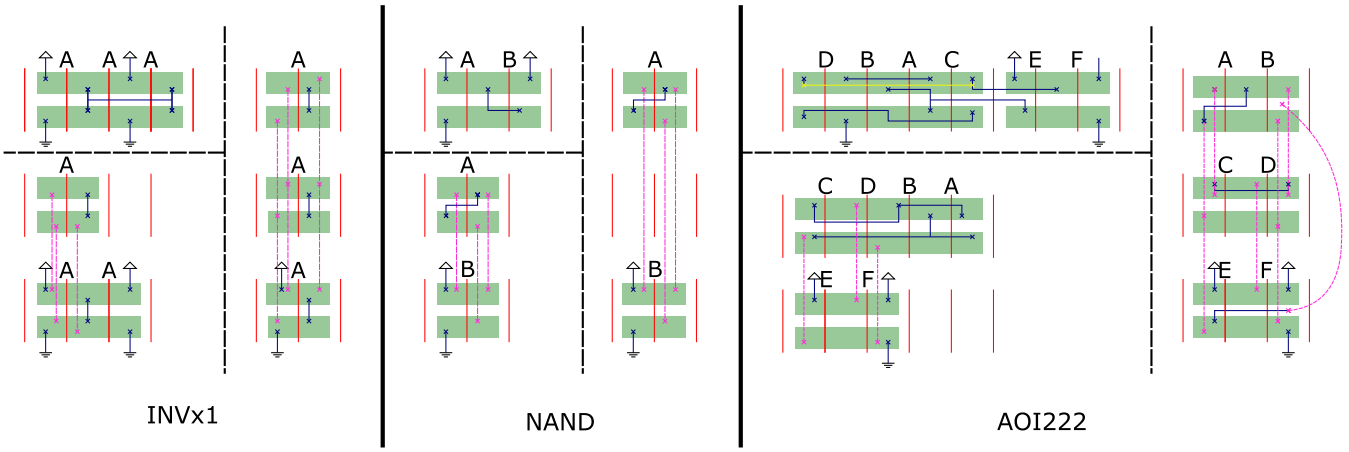

FIGURE 6. Examples of stick diagrams for three different standard cells for one-, two-, and three-tier technologies.

TABLE 1. Area as a function of the tier for selected standard cells.

Standard Cell	1 Tier		2 Tier		3 Tier	
	CPPs	Area	CPPs	Area	CPPs	Area
ASYNC_DFFHx1	24	100%	12	60%	9	54%
INVxp33	2	100%	2	120%	2	145%
INVxp67	3	100%	2	80%	2	97%
INVx1	4	100%	3	90%	2	72%
NAND2xp33	3	100%	2	80%	2	97%
NOR2xp33	3	100%	2	80%	2	97%
OAI21xp33	4	100%	3	90%	2	72%
AOI21xp33	4	100%	3	90%	2	72%
OAI22xp33	5	100%	3	72%	3	87%
AOI22xp33	5	100%	3	72%	3	87%
OAI221xp33	7	100%	4	69%	3	62%
AOI221xp33	7	100%	4	69%	3	62%
OAI222xp33	8	100%	5	75%	3	54%
AOI222xp33	8	100%	5	75%	3	54%
XNOR2xp5	7	100%	4	69%	3	62%
XOR2xp5	7	100%	4	69%	3	62%
ICGx1	18	100%	10	67%	7	56%
BUFx2	8	100%	5	75%	4	72%
BUFx3	11	100%	6	65%	5	66%

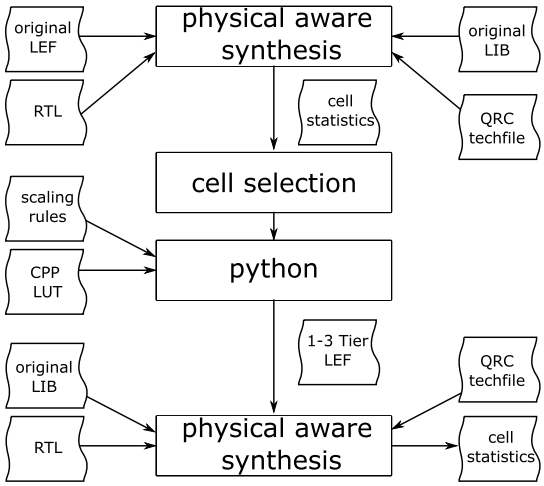

FIGURE 7. EDA methodology.

Fig. 6 shows the three examples of these heuristics applied to standard cell circuits. While the INVx1 has low design complexity, its area is inflated in the single-tier variant to reach the necessary drive strength. Here, the single path can be "wrapped around" to higher tiers such that the area is reduced to 83% in the three-tier implementation. Even though the NAND2 gate provides slightly higher complexity, it can capitalize less on the availability of a third tier, as only two transistors per conduction type are required. As a case study of a more complex cell the AOI222 gate is considered. Here, the Euler paths get distributed among the two tiers for the dual-tier variant. For the three-tier variant, the longer path gets wrapped to reach over two tiers for even tighter integration. The path segment with most VDD/VSS contacts is moved to the bottom tier as advised by the given heuristics. This leads to the output pin ending on the bottom tier, preventing final pin access. However, as six gate contacts are spread over three contacted polypitch (CPP), there are three unused gate contacts available. This used gate contact is indicated by the curved pink MIV.

Table 1 shows the results of the scaling analysis conducted on the given set of standard cells. As the cell height increases according to the necessary gate contacts (see Fig. 2), some cells, such as the INVxp033, do not benefit from moving to higher tiers. On the other hand, more complex cells, such as the AOI222 cells, can significantly benefit from the increased number of tiers.

C. EDA METHODOLOGY

Realizing improvements on the level of standard cells, it is important to also quantify the impact on the level of architectural blocks. A naive approach to this estimation would be replacing the cells in a given design with their scaled variants, obtaining the global scaling factor. However, we expect the synthesis to favor cells that are more suitable for multi-tier implementation. To obtain an understanding of the area improvements for a more complex design, we, therefore, modify the standard cell data and rerun the synthesis with scaled cell sizing. As a baseline, we use the standard cells of the ASAP7 technology. To ensure a fair comparison, we also modify the baseline one-tier variants to benefit from advantageous technology assumptions such as single diffusion breaks. A flowchart of the library generation is depicted in Fig. 7. Here, the synthesis is set to use the LEF for area information and the original timing libraries only for their

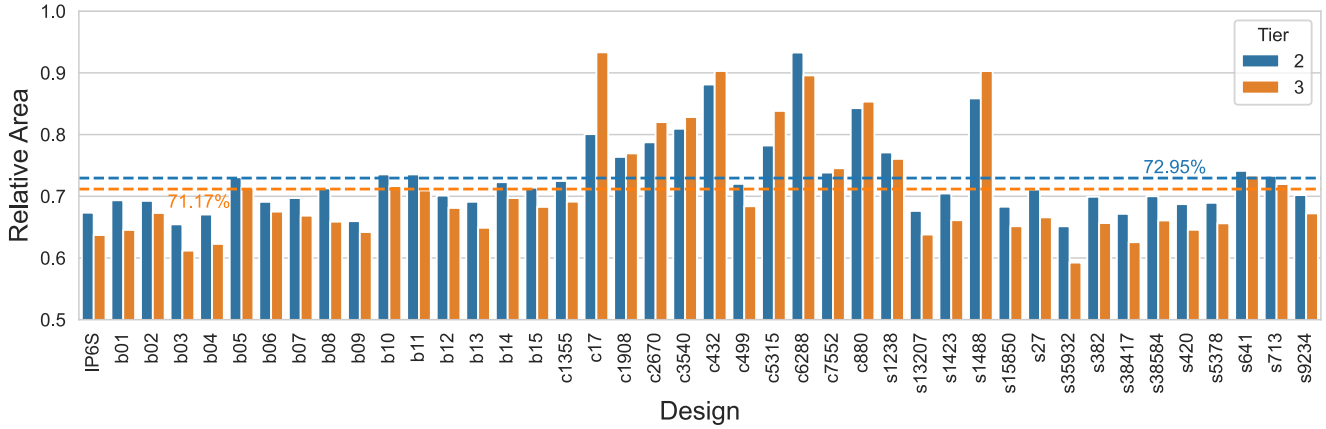

FIGURE 8. Detailed breakdown of area improvement of all 42 benchmark designs. Dashed lines indicate the mean over all tier-one and tier-two results, respectively.

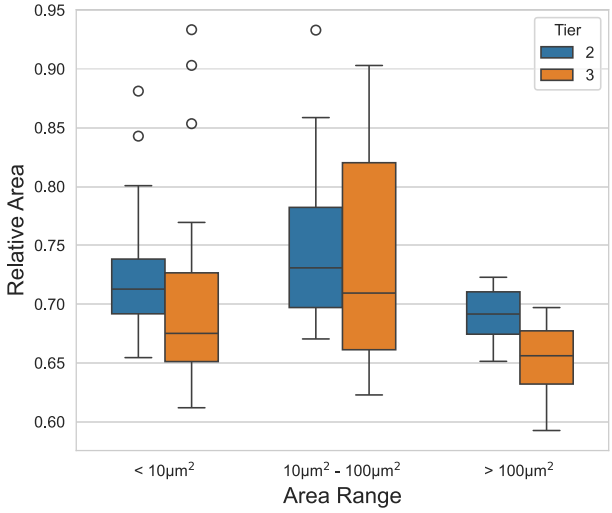

FIGURE 9. Area improvement throughout the complete set of benchmark circuits separated by the initial area.

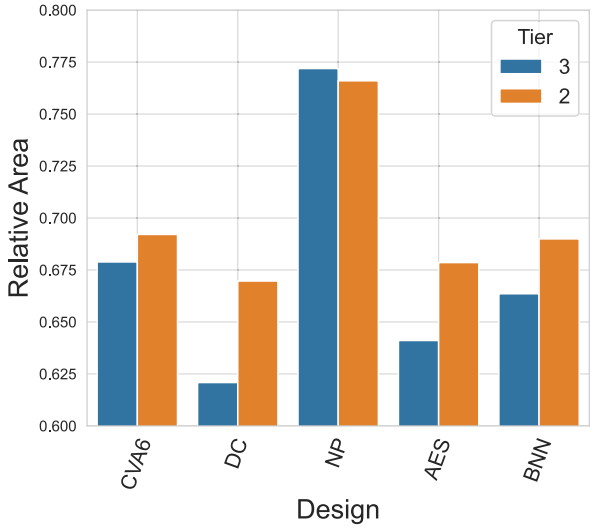

FIGURE 11. Scaling results for multiple complex RTL designs.

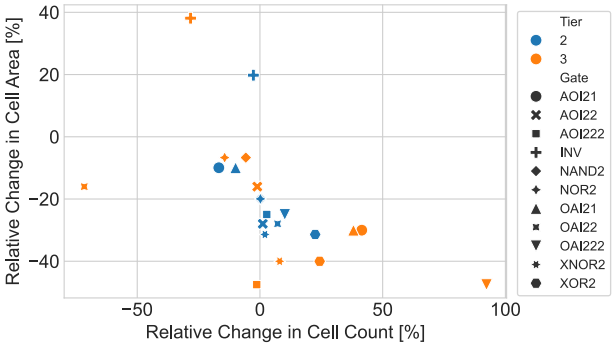

FIGURE 10. Relative change in the number of instances compared to a relative change in the cell area.

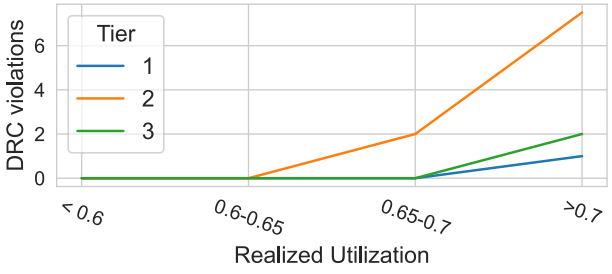

FIGURE 12. Relevant DRC violations over realized utilization.

timing information. The physical aware synthesis estimates wire capacitances based on placement assumptions.

IV. RESULTS

In the first step, we collect the designs from several synthesis benchmarks of digital designs [15], [16], [17]. Then, we filter out instances that are supersets of other benchmark instances in order to avoid a bias in the results of these designs. The change in area is shown in relation to the initial one-tier synthesis results for all 42 considered benchmark designs in Fig. 8.

As seen in Table 1, larger standard cells tend to feature higher area reduction. It appears reasonable to assume that more complex designs or equivalently designs with a larger area feature more opportunities to actually use larger cells.

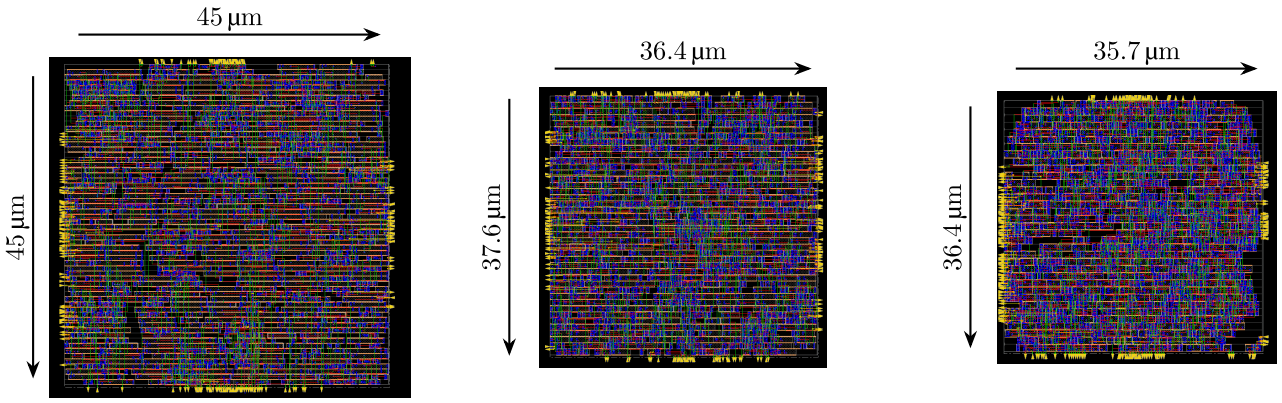

FIGURE 13. Layout view of the three-tier variants for "s13207": one tier (left), two tiers (middle), and three tiers (right).

Hence, sorting the design by their initial tier one area should lead to higher area reductions for larger designs. Splitting the results into three groups, such a trend can be confirmed as visualized in Fig. 9. A stronger and more robust scaling can be observed for larger designs with outliers surpassing a 25% reduction.

Furthermore, it is interesting to validate whether synthesis actually favors cells with better area scaling. In order to verify such a shift, we investigate the cells used in the best-performing design. Fig. 10 shows the relative change in cell count compared to the change in area from the one-tier variant. Here, we filter out DFFs and ICG cells, as their number is bound by the given RTL code and only a single cell with the given functionality is available. Especially the three-tier variant seems to make good use of more complex gates, as the number of OAI222 gates is almost doubled. For both cases, the largest burden for further scaling seems to be the unfavorable scaling of inverter cells, as these do not benefit from tier stacking. We can also observe that simpler standard cells such as NOR2 still are necessary in all given variants, limiting the potential for area reduction.

As much as the set of benchmarks is frequently used to evaluate the quality of the design flow, they may lack specific characteristics only found in actual designs. Furthermore, the benchmark circuits tend to be mostly small in complexity, i.e., they might not reveal the full potential seen for larger designs capable of using more complex cells. Hence, we ran another benchmarking using the following set of actual architectural block designs: the RISC-V CVA6 processor (CVA6) [18], an accelerator for depthwise convolutions (DCs) [19], the nAIxt neuroprocessor for the simulation of biological neural networks (NP) [20], a cryptographic AES accelerator (AES) [21], and a convolution unit for binary neural networks (AES) [22]. The corresponding results are shown in Fig. 11. Here, the largest design is placed on the leftmost side, with the other designs decreasing in size toward the right. We combine the designs larger than 100 μm^2 from Fig. 9 with the results from Fig. 11 to obtain a median scaling potential of 32.14% for the two-tier variant and 33.03% for the three-tier variant.

As much as the relative benefit is design-dependent, the results from Figs. 9 and 11 both indicate only limited benefit of the three-tier solution. Considering the integration complexity and related higher cost, the three-tier solution appears unsuited to provide a benefit under the assumption taken in Section II.

TABLE 2. Comparison table.

	[10]	[6]	[7]	**This work**
Integration Level	Transistor	Gate	Gate	**Transistor**
FET Style	VFET	NA	NA	**lateral GAA**
# Tiers	1-4	1-4	1-5	**1-3**
EDA step	PnR	PnR	PnR	**Synthesis**
MIVs considered	yes	no	yes	**yes**
# Analyzed Designs	3	4	3	**48**
Area reduction	27.9%	73.98%	80%	**33.03%**

A comparison of this work with similar works is found in Table 2. As can be seen, this work is the only work considering lateral GAA-FETs in a transistor-level stacking integration style.

A. PROSPECTIVE PLACE AND ROUTE

A challenge, which can arise from the realized increase in density, is routing congestion. As pins move further together, connecting them becomes challenging. In a first-order assessment, we analyze this challenge by adding a PnR step at the end of the flow given in Fig. 7. Again, we revisit the designs given in Fig. 8. As backside power delivery is assumed here, no power routing was performed. The achieved improvement in the floorplan area falls into a similar order of magnitude as the previously shown results for the gate area. For the two-tier variant, the floorplan area improves by 26.5%, and for the three-tier variant, the improvement increases to 28.9%, on average. As the transition to 5T standard cell height represents a significant increase in cell density for the ASAP7 technology already at the one-tier variant, DRC violations can be observed. To offer a perspective on routability, we, thus, show the median number of DRC violations of the performed study in Fig. 12.

As the DRC violations rise already at lower area utilizations for the two-tier variant, routing these designs seems to be significantly harder. Even though the track height increases for the two-tier variant, the overall area decrease still poses a higher challenge for routing. The further increase

in track height for the three-tier implementation leads to some improvement in routability, as the DRC violations are significantly reduced compared to the two-tier implementation. However, the one-tier variant still offers the best routability for the given utilization range. If future technologies want to capitalize on the presented transistor-level tier stacking, further improvements in the BEOL, therefore, are unavoidable. As the one-tier variant already presents challenges to meet DRC requirements, this is the case also for gate-level stacking if cell height projections of the IRDS are to be met. However, the necessary improvement seems to be lower for the gate-level stacking technique. A layout view after the added PnR step is displayed for the example of "s13207" of the benchmark circuits in Fig. 13.

V. CONCLUSION

In this work, we have investigated the potential of geometric scaling using 3-D standard cells in multitier technology nodes. For the quantitative assessment, virtual physical design kits (PDKs) are developed based on the existing ASAP7 technology amended with reasonable assumptions on such multitier technology. These virtual one-, two-, and three-tier PDKs are used to redesign a selected set of standard cells. Furthermore, typical synthesis benchmarks and actual complex architectural blocks are collected for benchmarking. The designs are then run through a standard digital design flow to determine the resulting area in each case.

We observe that an area improvement of more than 25% can be reached for both dual- and triple-tier solutions with the dual-tier variant achieving a more consistent improvement. There are two factors that prevent the area from shrinking further. First, there is a necessary increase in cell height to properly connect all transistor gates across tiers. Second, diffusion breaks at the sides of the standard cells do not scale in proportion to the number of tiers. Considering that a three-tier technology is more expensive than a two-tier technology, the latter is clearly preferred.

It is important to note that works on gate-level 3-D integration typically neglect the area overhead by MIVs. Accounting for a realistic MIV area and considering the increasing manufacturing cost with the number of tiers, we expect diminishing returns after a certain amount of tiers for this case as well.

Finally, a combination of both methodologies might achieve an overall minimum in cost and maximum in area reduction. To find this sweet spot, further analysis using such a hybrid approach has to be conducted. Such a study should also consider the modeling of parasitics in order to benchmark the design with respect to power and performance as well.

A final analysis of routability indicates that significant improvements in the BEOL are a prerequisite before the projections of the IRDS can be achieved. This was observed for a cell height of five tracks, as used for gate-level stacking and for the higher but denser cells obtained by transistor-level stacking.

REFERENCES

[1] T. Srimani et al., "Heterogeneous integration of BEOL logic and memory in a commercial foundry: Multi-tier complementary carbon nanotube logic and resistive RAM at a 130 nm node," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2020, pp. 1–2.

[2] M. M. Shulaker et al., "Three-dimensional integration of nanotechnologies for computing and data storage on a single chip," *Nature*, vol. 547, no. 7661, pp. 74–78, Jul. 2017.

[3] R. An et al., "A hybrid computing-in-memory architecture by monolithic 3D integration of BEOL CNT/IGZO-based CFET logic and analog RRAM," in *IEDM Tech. Dig.*, Dec. 2022, p. 18.

[4] IRDS More Moore IFT Team, *IRDS 2022 More Moore*, IEEE, New York, NY, USA, 2022.

[5] Y.-J. Lee and S. K. Lim, "Ultrahigh density logic designs using monolithic 3-D integration," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 32, no. 12, pp. 1892–1905, Dec. 2013.

[6] S. D. Lin, P. P. Pande, and D. H. Kim, "Optimization of dynamic power consumption in multi-tier gate-level monolithic 3D ICs," in *Proc. 17th Int. Symp. Quality Electron. Design (ISQED)*, Mar. 2016, pp. 29–34.

[7] A. Tamir, M. Salem, J. Lin, Q. Alasad, and J.-S. Yuan, "Multi-tier 3D IC physical design with analytical quadratic partitioning algorithm using 2D P&R tool," *Electronics*, vol. 10, no. 16, p. 1930, Aug. 2021.

[8] M. Jung et al., "How to reduce power in 3D IC designs: A case study with OpenSPARC T2 core," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2013, pp. 1–4.

[9] W.-T. J. Chan, A. B. Kahng, and J. Li, "Revisiting 3DIC benefit with multiple tiers," in *Proc. 18th Syst. Level Interconnect Predict. Workshop*, B. Taskin and T.-Y. Ho, Eds., New York, NY, USA, 2016, pp. 1–8.

[10] D. Lee, C.-T. Ho, I. Kang, S. Gao, B. Lin, and C.-K. Cheng, "Many-tier vertical gate-all-around nanowire FET standard cell synthesis for advanced technology nodes," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 7, no. 1, pp. 52–60, Jun. 2021.

[11] T. Song, "Many-tier vertical GAAFET (V-FET) for ultra-miniaturized standard cell designs beyond 5 nm," *IEEE Access*, vol. 8, pp. 149984–149998, 2020.

[12] S. Ye, L. Liu, Y. Ma, and Y. Wang, "Stacked lateral gate-all-around metal–oxide–semiconductor field-effect transistors and their three-dimensional integrated circuits," *Silicon*, vol. 15, no. 5, pp. 2467–2478, Apr. 2023.

[13] S. Yang et al., "PPA and scaling potential of backside power options in N2 and A14 nanosheet technology," in *Proc. IEEE Symp. VLSI Technol. Circuits*, Jun. 2023, pp. 1–2.

[14] W. Hafez et al., "Intel PowerVia technology: Backside power delivery for high density and high-performance computing," in *Proc. IEEE Symp. VLSI Technol. Circuits*, Jun. 2023, pp. 1–2.

[15] F. Brglez and H. Fujiwara, "A neutral netlist of 10 combinational benchmark circuits and a targeted translator in FORTRAN," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Jun. 1985.

[16] F. Brglez, D. Bryan, and K. Kozminski, "Combinational profiles of sequential benchmark circuits," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 1989, pp. 1929–1934.

[17] *ITC99 Benchmark*. Accessed: May 31, 2024. [Online]. Available: https://www.cerc.utexas.edu/itc99-benchmarks/bench.html

[18] F. Zaruba and L. Benini, "The cost of application-class processing: Energy and performance analysis of a Linux-ready 1.7-GHz 64-bit RISC-V core in 22-nm FDSOI technology," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 11, pp. 2629–2640, Nov. 2019.

[19] Y. Chen, J. Lou, C. Lanius, F. Freye, J. Loh, and T. Gemmeke, "An energy-efficient and area-efficient depthwise separable convolution accelerator with minimal on-chip memory access," in *Proc. IFIP/IEEE 31st Int. Conf. Very Large Scale Integr. (VLSI-SoC)*, Oct. 2023, pp. 1–6.

[20] K. Kauth, C. Lanius, and T. Gemmeke, "nAIxt: A light-weight processor architecture for efficient computation of neuron models," in *Proc. 37th Int. Conf. Archit. Comput. Syst. (ARCS)*, 2024, pp. 1–15.

[21] S. Zhang, M. Wabnitz, and T. Gemmeke, "A compact 1,257-Gbps/W byte-serial AES accelerator for IoT applications in 22 nm," in *Proc. IEEE Asian Solid-State Circuits Conf. (A-SSCC)*, Nov. 2023, pp. 1–3.

[22] T. Stadtmann, C. Latotzke, and T. Gemmeke, "From quantitative analysis to synthesis of efficient binary neural networks," in *Proc. 19th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2020, pp. 93–100.