Data analytics and machine learning for continued semiconductor scaling

Luigi Capodieci

A novel methodology for quantifying design space coverage has great potential for yield optimization of integrated circuits at the 7nm node and beyond.

Although there has been a rapid and greatly publicized growth of data analytics and machine learning methodologies across many applications, and in virtually every industry, these developments seem to have almost completely been missed in the semiconductor integrated circuit (IC) space. With the 14nm process technology node currently in production, and both 10 and 7nm nodes at different stages of development, the IC ‘ecosystem’ is being restructured and consolidated across its four traditional components (i.e., fabless design companies, electronic design automation and intellectual property suppliers, process and metrology tools suppliers, and silicon foundries). There are, however, intrinsic technology factors (e.g., the continual deceleration of geometric scaling and the delayed introduction of key patterning technologies) that are primary sources of disruption to this restructuring. There are also critical hidden gaps and bottlenecks in the design-to-manufacturing data information pipeline. The deployment of carefully selected data analytics techniques (with/without machine learning algorithms) therefore represents a strategic opportunity to enable a 2 year/node (‘more-Moore’) cycle at 10nm and below in the semiconductor industry.

In this work, we present a survey of the state-of-the-art and ongoing developments in data analytics and machine learning. We also offer a perspective on the functional interactions and data information flows for IC design-to-manufacturing, and we discuss the risks and opportunities that arise from the introduction of big-data analytics and machine-learning technologies. Although the terms ‘data analytics’ (or ‘big data’) and ‘machine learning’ comprise a vast constellation of mathematical methodologies, computational platforms, and system-level solutions, we discuss only a few examples in this article (i.e., those that have novel applications for physical design-to-process yield co-optimization).

Design-technology co-optimization (DTCO) is an established approach that involves leveraging collaborations and cross-domain expertise among physical designers, lithographers, process integration, and yield engineering. DTCO can thus be used to simultaneously define fundamental layout components of cell architecture and to generate geometric design and interconnect router rules, while ensuring that the physical design is manufacturable (i.e., within the expected capabilities of

Figure 1. Illustrating the generation of a topological network map for semiconductor integrated circuit (IC) research and development. Left: The physical domain (i.e., a layout) is broken down into abstract element components that are then represented as a set of multidimensional points. Right: The high-dimensional data cloud thus represents the physical domain space in topological graph space. (Courtesy of Motivo Data Analytics.)

Continued on next page
Figure 2. Illustration of the use of machine learning, graph traversal, and search to predict the locations of yield detractors. Both supervised and unsupervised machine learning are applied to predict yield-detractor patterns, via graph traversal and node clustering. A known problematic pattern (in blue on the left) is algorithmically broken down into fundamental components (also in blue), which are then clustered into key components (red). The clustered key components can be extrapolated to identify new problematic nodes (yellow), which are finally mapped back onto specific locations in the physical design layout. (Courtesy of Motivo Data Analytics.)

patterning and other critical process modules). Even in the presence of restricted (or highly restricted) design rules, the main drawback of traditional DTCO is the limited number of layout configurations that can be considered, evaluated, and optimized against process capabilities (both in simulations and in silicon test vehicles). This results in a limited predictability of manufacturing yield for actual IC products, as these are not adequately represented by DTCO layout variables. Indeed, from a combinatorial perspective, the space complexity of layout variants is very large (typically with $10^{20}$ cardinality for single-layer layouts, and up to $10^{80}$ cardinality for multilayer layouts).

However, we have recently proposed a rigorous quantitative methodology for fully characterizing systematic layout variability. Our approach involves the definition of the physical design space coverage (DSC), which is a unifying abstraction for all components of the design-to-manufacturing data flow. The DSC is therefore the necessary mathematical foundation for a practical data analytics computation framework. Despite the apparently intractable space complexity, various algebraic combinatorial techniques can be used to compute all possible layout design variations in less than quadratic time. Moreover, in most practical cases (i.e., for real IC products) the computations can be made in almost linear time. The DSC can be understood intuitively as a measure of all combinatorial variations in a layout, at all length scales (i.e., a direct correlation of its intrinsic geometric entropy). We also note that combinatorial techniques for the quantitative characterization of an abstract space are already commonly adopted in several big data domains (e.g., Web searches, genomics, sentiment text analysis, cybersecurity, and anomaly detection). For IC research and development (and production), the most immediate application of DSC is for quantifying how similar/dissimilar a given layout design is to another layout. This is therefore a measure of how representative a layout is of any other layout, with respect to process manufacturability and yield. With our approach, it is thus possible to implement a new
(analytics-based) DTCO, in which human engineering expertise (i.e., with regard to layout components such as cell architecture and routed interconnections) is greatly augmented by computational tools for determining the extent of a design space that has been evaluated (and optimized). To achieve this we conduct computationally mapping—with the use of well-established (albeit somewhat esoteric) topological data analysis techniques—of the physical domain space (e.g., geometric layout features, edges, polygons, and layers) onto a high-dimensional topological network map. The conceptual construction of such a topological network map is shown in Figure 1, where the physical domain (a layout or a set of layouts) is broken down into abstract elemental components. These components are then represented as a set of multidimensional points (a data cloud). The data cloud is further organized as a topological network (multigraph) by identifying and storing the intrinsic data relationships between the points, together with extrinsic (domain-specific) connections. We then use a class of standard algorithms to compute the topological network properties. These, in turn, are then used to extract (machine-driven) insights back into the physical domain space.

To demonstrate how design-process-yield optimization can be transformed into a graph traversal and search problem, we consider a simple example: the prediction of yield detractors (hotspot layout configurations, under certain process conditions), based on a set of physical measurements (e.g., scanning electron microscope metrology or failure analysis). In the physical space (geometric layout) domain, empirical observations are typically clustered according to their physical attributes. Experimental noise and information loss during dimension reduction, however, can severely limit the proper identification (and prediction) of actual yield-detractor clusters. In contrast, in topological network space, physical data points are represented by subnetworks (subgraphs, with specific sets of nodes and edges), where commonalities and root causes can be identified algorithmically. Specifically, several machine-learning methodologies (supervised and unsupervised learning) are directly applicable to graph traversal and the search for ‘common causal nodes’ (see Figure 2). This is just one simple example of the many applications of data analytics and machine learning to the semiconductor design-to-manufacturing space. The novelty of this methodological approach lies in the use of a topological network space (where the analytics and machine learning are implemented) rather than traditional and noise-prone approaches in which raw physical data is directly analyzed and clustered (i.e., machine learning).

In summary, we have provided a brief overview of how advanced data analytics and machine learning methodologies can, and should, be deployed for the joint optimization of physical design, manufacturing process, and IC product yield in the semiconductor industry. Although there are many opportunities to apply these methodologies in the IC field, only a few semiconductor-specific approaches have so far been developed. For example, our recent introduction of quantitative techniques for the characterization of the complete design space coverage (based on rigorous combinatorics theory) shows great promise. Our DSC approach, when combined with graph search and machine learning, has the potential to open up new research avenues for physical design and yield optimization, e.g., an entire new class of yield-optimization solutions for the 10, 7, and 5nm nodes (and beyond). Our development activities are currently focused on two complementary fronts: triple (and quadruple) layout decomposition and directed self-assembly (DSA). So far, we have successfully applied coverage analytics to minimize (and optimize) the number of shapes needed to decompose 7nm random logic layouts and we are now working on 5nm design rules. Similarly, we have demonstrated that DSC has the potential for operationally extracting practical template ‘alphabets’ for DSA.

Author Information

Luigi Capodieci
KnotPrime
Santa Cruz, CA

Luigi Capodieci has worked on lithographic imaging, pattern and process simulations, resolution enhancement technology, optical proximity correction, and design-for-manufacturing for more than 20 years while at Advanced Micro Devices and GLOBALFOUNDRIES. He is currently the chief technology officer of KnotPrime, a data analytics startup company in the field of anomaly detection and applied algorithmic intelligence.

References


Continued on next page