Model to Hardware Matching

For nano-meter Scale Technologies

Sani R. Nassif,
IBM Research – Austin,
nassif@us.ibm.com

Abstract— **Our ability to reliably predict the outcome of a semiconductor manufacturing process has been steadily deteriorating. This is happening because of two important factors. First, the overall CMOS technology slowdown has led to rapidly increasing complexity in the process and in its interaction with design. This has in turn caused an increase in the number and magnitude of systematic sources of mismatch between simulation models (both at the Technology-CAD and at the circuit simulation levels) and hardware measurements. Second, manufacturing variability resulting from random as well as systematic phenomena -long a source of concern only for analog design- is becoming important for digital design as well and thus its prediction is now a first order priority. Process complexity and the challenges of accurately modeling variability have conspired to increase the error in performance predictions, leading to a gap in model to hardware matching.**

In this paper, we will review these issues and show examples of potential solutions to this problem some of which are currently being developed in IBM, and some which are longer term and would benefit greatly from the attention of the academic community.

Keywords-component; model to hardware correlation; manufacturing variations; modeling.

I. WHAT IS A MOSFET

Until relatively recently (say in the 1990s) the definition of the layout features which determine the performance of a fabricated MOSFET were relatively simple. The intersection between the diffusion and polysilicon mask shapes determined the dimensions of the MOSFET, and those dimensions sufficed to characterize the behavior of the transistor. We do not mean to imply that the scaling of these dimensions did not make transistor modeling more difficult, but rather to point out that the behavior of the transistor was determined by *local* geometry and that its extraction from the layout was a straightforward task.

As the industry migrated to sub-wavelength lithography, the difference between the drawn mask shapes and actual printed (wafer) regions began to increase. This resulted in the introduction of resolution enhancement techniques (RET) and optical pattern correction (OPC) to improve the fidelity of manufactured devices. In spite of advances bordering in many cases on the miraculous, it remains a fact that there is an ever growing gap between device layout as viewed by a designer, and final manufactured shapes as rendered in Silicon. This gap exhibits itself in two ways: (a) the precision with which the dimensions (length and width) of a device can be set are limited by the residual error that remains after RET is applied; and (b) the interaction between shapes on the mask due to interference, flare, and other lithography and illumination related issues means that the actual dimensions of the device are determined by *all* the shapes in the neighborhood of the device.

At the 65nm node and below, assuming the current lithography roadmap remains as it is currently defined, the *radius of influence* that defines the neighborhood of shapes that play a part in determining the characteristics of a MOSFET is expected to increase from the current "nearest feature" to include one-removed features, auxiliary features, corners, vias, and other second order phenomena. This increase in the radius of influence will impact a number of areas:

- It will make the modeling of device behavior more difficult since it will be harder to define a *canonical* or *typical* device from which to perform model characterization. We will come back to this subject later in this paper.
- It will make the circuit extraction phase of design verification (where the layout is converted into a simulatable netlist) more complex since a larger number of geometries will need to be processed.
- It will severely reduce design *composability*, defined as the ability to compose a complex circuit function out of individual simple functions, e.g. building a multi-bit adder out of single-bit adders, which are in turn composed of individual NAND, NOR and similar logic gates.

The trends outlined thus far point to an increasing need to model and comprehend the interaction between design and manufacturing at the device level. A related and important trend is the increase in manufacturing variability, which we will discuss next.

II. VARIABILITY AND UNCERTAINTY

The performance of an integrated circuit is determined by the electrical characteristics of the individual linear and non-linear devices which constitute the circuit. Variations in these device characteristics cause the performance of the circuit to deviate from its intended range and can cause performance

degradation and erroneous behavior. We will refer to this type of variability as *"physical variability"*.

In addition, circuit performance is determined by the environment in which the circuit operates. This includes factors such as temperature, power supply voltage, and noise. Variations in operating environment can have a similar impact on circuit behavior as device variations. We will refer to this type of variability as *"environmental variability"*.

It is tempting to think of physical variability as simply the result of systematic and random manufacturing fluctuations. But a variety of time-dependent wear-out mechanisms such as metal electro-migration and negative-bias threshold instability (NBTI) cause device characteristics to change over time –albeit with a time constant on the scale of months and years. In contrast, environmental variability is very much a function of time but at the same time-scale as that of the operation of the circuit, e.g. in the Nano-second range for a typical GHz design.

Whether physical or environmental, we can classify components of variability in various ways. For example, we can examine their temporal behavior (as we alluded to above). We can also examine their spatial distribution across chip, reticle, wafer, and lot. Most important, however, is the notion of whether we fully understand the interaction between the specific component of variability and the characteristics of the relevant design.

Consider the case where a physical or environmental component of variability is known to be a function of specific design characteristics. For example, it is well known that variation in the channel length of MOSFET devices is related to the orientation of the devices. With a suitable quantitative model relating the variation to design practice, a designer can make the appropriate engineering tradeoff and margin his design so as to minimize the impact of the variations. Such phenomena are often systematic in nature, and we will refer to them simply as *"variability"*.

Now consider the case where a physical or environmental phenomenon is not well understood, such that available information is limited to the magnitude of the variation, but without insight into its quantitative dependence on design. In such a case, the designer has no choice but to perform worst-case analysis, i.e. creating a large enough design margin to correct for the worst possible condition that may occur. Such large margins are usually wasteful in design resources and end up impacting the overall cost and performance. We will refer to these types of phenomena as *"uncertainty"*.

Variability can be designed around and will typically cause small increases in design cost, while uncertainty needs to be margined against and will typically cause large increased in design cost. A key point to remember, however, is that the difference between the two is determined by our ability to understand and model the mechanisms as play, and that an investment in modeling and analysis can sometimes turn a source of uncertainty into a source of variability, thereby reducing design cost and/or improving design performance.

With increasing manufacturing process complexity, more and more phenomena are competing for limited modeling resources. This trend, unchecked, endangers our industry's ability to deliver future design improvements consistent with historical trends. Furthermore, many phenomena are increasing in magnitude as scaling continues [1], requiring more modeling and analysis resources.

The concrete understanding of variability, uncertainty, and their interaction with design depends on our ability to perform detailed design-space-oriented manufacturing process characterization. We will examine this topic next.

III. Characterization vs. Modeling

The prior two sections identified two trends: (1) an increase in the diversity of device implementation and interaction, and (2) an increase in random and systematic variability resulting from the manufacturing process as well as circuit operation. Putting aside the environmental sources of variability, which require a set of techniques for analysis, modeling, and optimization that are beyond the scope of this paper, we will focus on how our understanding of the manufacturing process is created. We will further focus on how this understanding of the process is used to generate *simulatable models* of circuits, acknowledging that Spice-level circuit simulation is usually the basis on which all modern design is built.

In earlier technologies, it was sufficient to create simple scribe-line test structures including a handful of MOSFETs with varying dimensions. The measured characteristics of the devices was used to extract device model parameters (e.g. for a BSIM [2] model). The resultant parameters were considered to be a complete encapsulation of the behavior of the technology. As technology scaled and device models became more complex (e.g. to handle phenomena such as short channel effects) the selection and number of devices that are included in the test structure increased somewhat.

In contrast, the impact of the local layout environment on device behavior in current technologies dramatically increases the number of factors that need to be considered, so the number of potential device layout variations that would need to be studied is much larger than the handful supported by current test structures and modeling strategies. Furthermore, since variability has become an important limiter of design performance, an assessment of within-die fluctuations is desirable. Such an assessment would require sufficient replication of identical structures to allow statistical characterization. Such a characterization may, for example, allow for the comparison of the spread in device behavior for different layout practices –an important degree of freedom for a designer attempting to reduce the impact of variations on a circuit.

Without models of the interaction between design implementation (layout) and manufacturing variability, we run the risk of excessive worst-casing as more and more of these phenomena emerge and challenge current design practices. Thus there is a need to change the manner in which we define the interaction between design and technology. When the characteristics of devices were substantially independent of the manner in which they were laid out, it was sufficient to merely *characterize* the manufacturing process. In current technologies, the focus needs to shift to *modeling* of the manufacturing process, i.e. creating quantitative relationships

between design implementation and design performance. It is only through such models that we can convert uncertainty to variability, and reduce excessive margins and pessimism.

All of these facts point to the need for significantly more complex process modeling test structures. Some of desirable features of such structures are:

- Similarity to design. Structures should have layout features that are similar to realistic designs in order to ensure that the resulting data is meaningful.
- Spatial breadth. Since within-die and within-wafer variability often has strong spatial components, careful attention must be paid to the placement and size of structures in order to gather sufficient data.
- Statistical breadth. As pointed out above, we need a sufficient number of replicas to confidently measure distributions, as well as changes in distributions over time or for different design practices.
- Layout breadth. The structures should explore a sufficiently wide range of layouts to cover all or most of anticipated design practice.
- Ease of measurement. Crucial for both the timely delivery of initial models, as well as the continued monitoring of model to hardware matching as technology learning progresses.

No one structure can accommodate all the needs above, but a family of related structures can. Such a family might include large dense structures that are run infrequently for detailed modeling and characterization, and small scribe-line structures that are run often (perhaps on all wafers) and serve to monitor a smaller number of key variables.

At IBM we are making significant research investments in this area to produce a family of structures that are easy or even trivial to design via the use of test structure compiler technology. Careful attention is paid to density to allow for the maximum possible number of variations and replications and also to test efficiency to ensure that the structures are tested often and consistently. We believe this area will grow in importance and relevance with time, and will be a key enabler for future design/technology interaction.

IV. Modeling And Virtual Fabrication

The previous section alluded to the need for improved models of technology performance and variability, since such models are crucial to the management of pessimism and design margin. But how are such models to be created, characterized and delivered? And what is the role of Computer-Aided Design (CAD) as well as Technology-CAD in the creation of such models?

The MOSFET circuit simulation device model (e.g. BSIM) has been the core definition of the design/technology interface. Such models are standardized, well understood, and are supported by both simulators as well as parameter extractors. The historical trend in this field has been for increasing model complexity (and corresponding number of model parameters) to capture an ever broader set of phenomena. This trend focuses on nominal accuracy of the model with respect to a *"typical"* or *"golden"* device, and can as easily be driven by measurements as by device characteristics generated from a traditional device simulator. Also, the complexity of these models makes for costly characterization cycles that require substantial experience to insure the delivery of robust, reliable and accurate models.

The trend to increased (a) layout dependence (radius of influence) and (b) variability is an opportunity to reexamine current modeling trends. In a regime where the within-die tolerance on key parameters is expected to reach a third or more of the mean value, we need to carefully understand the interplay between *nominal accuracy* (predicting the mean) vs. *statistical accuracy* (predicting the spread). This interaction is central to the understanding of the model to hardware matching issue, and needs to be driven by an understanding of the economic significance of misprediction on the potential profit of a design.

We strongly believe that there is a strong and immediate need for *statistically accurate* MOSFET models which would serve to allow the prediction and reduction of the impact of technology variations on design. Such models are not new, and have been suggested and implemented in the past. What is different now is the need for those models to be enhanced to include the interaction between device layout and device behavior. This is an area of research that can result in immediate benefit.

A *statistical characterization methodology* will also be required in order to insure that such models are easily created, validated, and updated as technology learning occurs. Such a methodology is not as well developed, but will be a corner stone of future design/technology interaction.

Generating early models of technology for exploratory design is crucial, especially due to the ever lengthening design cycles of large complex chips. Such models need to be created with little or no hardware available, and are often updated as more information becomes available. Such models currently focus on capturing the nominal behavior of technology, but need to be enhanced to capture variability. This presents a significant challenge to current methods, but also represents an opportunity for new algorithms and tools.

The concept of a *"virtual factory"* has existed for some time in the Technology-CAD community [3]. Perhaps it is time to revisit the concept of virtual fabrication and enhance it to allow the early modeling of variability as well as the interaction between layout and device performance. Such enhancement should be done such that the virtual factory meshes directly with advanced test structures such that it can readily adapt to new data, new physical mechanisms, and even new layout styles.

V. Model to Hardware Matching

The issues and efforts outlined thus far culminate in the problem of model to hardware matching. We use this term to

denote our ability to predict, via modeling, the behavior of hardware after fabrication. The confidence we have in such predictions is key to the economic success of the semiconductor business since performance and yield play a large part in determining the profitability of a design.

The increasing interaction between layout and device performance is one of many sources of additional potential disparity between models and reality. In fact, a large number of mechanisms, when not modeled with appropriate accuracy, can lead to deviation between the predicted and observed performance. A few examples include:

1. Voltage and Temperature variability across the die due to differing power density [4]. Such variations, when not accounted for, can cause significant error in timing and static power (leakage) estimates.
2. Metal thickness variability due to Chemical-Mechanical Polishing [5]. While physical models exist for such variability, it is still not widely applied in commercial CAD flows.

Considering that the increment in performance between technology generations has been reducing as CMOS nears maturity, and that design margin is directly related to our ability to accurately predict and bound the nominal and statistical performance of our designs, it is clear that the improvement of our modeling breadth and accuracy must have the highest priority.

ACKNOWLEDGMENTS

The author is indebted to the large design/technology interaction community within IBM, which is far too numerous to name, but especially to Anne Gattiker for her insights and collaboration.

REFERENCES

[1] S. Nassif, "Delay Variability: Sources, Impacts and Trends", Proceedings of 2000 IEEE ISSCC Conference.

[2] B. Sheu, D. Scharfetter, P-K. Ko, and M-C. Jeng, "BSIM: Berkeley short-channel IGFET model for MOS transistors", IEEE JSSC, Aug 1987.

[3] P. Lloyd, *et. al.* "Technology CAD at AT&T", Microelectronics Journal, March 1995.

[4] H. Su, F. Liu, A. Devgan, E. Acar and S. Nassif, "Full Chip Leakage Estimation Considering Power Supply and Temperature Variations", ISLPED 2003.

[5] V. Mehrotra, S. Nassif and D. Boning, "Modeling the Effects of Manufacturing Variation on High-Speed Microprocessor Interconnect Performance", 1998 IEEE IEDM Conference.